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

May 24-28, 2010, Novi Sad
Applied Linear Algebra (ALA) 2010
Chebyshev polynomials of a compact set

- Chebyshev polynomials on the interval $[-1; 1]$ [Chebyshev 1859].
- Generalized by [Georg Faber 1920] to the idea of the Chebyshev polynomials of $\Omega$, where $\Omega$ 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

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

$\Omega$ is an interval, a set of discrete points, the unit circle, etc.
Let $A \in \mathbb{C}^{n \times n}$ be normal, i.e.,

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

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

$$\min_{p \in \mathcal{M}_m} \| p(A) \|.$$
Let $A \in \mathbb{C}^{n \times n}$ be normal, i.e.,

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

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

$$\min_{p \in M_m} \| p(A) \|.$$ 

Then

$$\min_{p \in M_m} \| p(A) \| = \min_{p \in M_m} \| p(\Lambda) \| = \min_{p \in M_m(\Omega)} \| p(z) \|_{\infty}$$

where $\Omega = \{\lambda_1, \ldots, \lambda_n\}$.
Let $A \in \mathbb{C}^{n \times n}$ be normal, i.e.,

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

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

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

Then

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

where $\Omega = \{\lambda_1, \ldots, \lambda_n\}$.

The problem for $A$ is solved by the Chebyshev polynomial of $\Omega$. 
Let $A \in \mathbb{C}^{n \times n}$ be a general matrix. We consider the problem

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

- Introduced in [Greenbaum, Trefethen 1994].
- Unique solution $T^A_m(z) \in \mathcal{M}_m$ exists if $m < d(A)$, [Greenbaum, Trefethen 1994; Liesen, T. 2009].
- $T^A_m(z)$ 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].
[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]

\[
\min_{p \in \pi_m} \| p(A)b \| \quad \text{(GMRES)}, \quad \min_{q \in \mathcal{M}_m} \| q(A)b \| \quad \text{(Arnoldi)},
\]

\[
b \approx \{ Ab, \ldots, A^m b \}, \quad A^m b \approx \{ b, \ldots, A^{m-1} b \}.
\]
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]

\[
\min_{p \in \pi_m} \| p(A)b \| \quad \text{(GMRES)},
\]

\[
\min_{q \in \mathcal{M}_m} \| q(A)b \| \quad \text{(Arnoldi)},
\]

\[b \approx \{ Ab, \ldots, A^m b \},\]

\[A^m b \approx \{ b, \ldots, A^{m-1} b \}.\]

One may remove \( b \) from the discussion and pose the following “ideal” approximation problems:

\[
\min_{p \in \pi_m} \| p(A) \| \quad \text{(Ideal GMRES)},
\]

\[
\min_{q \in \mathcal{M}_m} \| q(A) \| \quad \text{(Ideal Arnoldi)},
\]

\[I \approx \{ A, \ldots, A^m \},\]

\[A^m \approx \{ I, \ldots, A^{m-1} \}\]

(Chebyshev polynomial of \( A \))
Motivation Example

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

\[
J_\lambda = \begin{bmatrix}
\lambda & 1 \\
& \ddots & \ddots \\
& & \ddots & 1 \\
& & & \lambda
\end{bmatrix} \in \mathbb{C}^{n \times n}.
\]

**Question:** How do the ideal GMRES and Chebyshev polynomials of \( J_\lambda \) look like?
Let $\lambda \in \mathbb{C}$. Consider an $n$ by $n$ Jordan block

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

**Question:** How do the ideal GMRES and Chebyshev polynomials of $J_\lambda$ look like?

- Ideal GMRES polynomial of $J_\lambda$ - a very difficult problem
  
  [T., Liesen, Faber 2007].

- Chebyshev polynomial of $J_\lambda$ [Liesen, T. 2009]:

$$T_{m}^{J_\lambda}(z) = (z - \lambda)^m.$$
Outline

1. General results

2. Matrices and sets in the complex plane
Outline

1. General results

2. Matrices and sets in the complex plane
Shifts and scaling

**Theorem**  
[Faber, Liesen, T. 2010]

For \( A \in \mathbb{C}^{n \times n} \) and \( \alpha \in \mathbb{C} \) the following hold:

\[
\min_{p \in \mathcal{M}_m} \|p( A + \alpha I)\| = \min_{p \in \mathcal{M}_m} \|p(A)\|, \\
\min_{p \in \mathcal{M}_m} \|p(\alpha A)\| = |\alpha|^m \min_{p \in \mathcal{M}_m} \|p(A)\|.
\]

- **Shift invariance:** Not surprising, because the polynomials are normalized at infinity.

- **Paper contains explicit relations between the coefficients of**
  \( T_m^A(\bar{z}) \), \( T_m^{A+\alpha I}(\bar{z}) \), and \( T_m^{\alpha A}(\bar{z}) \).
Example - shift of a matrix

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

\[
A \equiv \begin{bmatrix} J_a & 0 \\ 0 & J_b \end{bmatrix}.
\]

Define

\[
\alpha \equiv \frac{a + b}{2}.
\]

Then

\[
A - \alpha I = \begin{bmatrix} J_\lambda & 0 \\ 0 & J_{-\lambda} \end{bmatrix} \quad \text{where} \quad \lambda \equiv \frac{a - b}{2},
\]

and the previous theorem implies

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

\[ \text{Let } A \in \mathbb{C}^{n \times n} \text{ and a positive integer } m < d(A) \text{ be given. If there exists a unitary matrix } P \text{ such that either } \]

\[ P^*AP = -A \text{ or } P^*AP = -A^T, \]

\[ \text{then } \]

\[ T_m^A(z) = (-1)^m T_m^A(-z). \]

[Faber, Liesen, T. 2010]
\[ A = \begin{bmatrix} J_\lambda & J_{-\lambda} \end{bmatrix}, \quad P = \begin{bmatrix} I^\pm & I^\pm \end{bmatrix}, \]

where \( I^\pm = \text{diag}(1, -1, 1, \ldots, (-1)^{n-1}) \). Then

\[ J_{-\lambda} = -I^\pm J_\lambda I^\pm \Rightarrow P^*AP = -A. \]
Example

\[
A = \begin{bmatrix}
    J_\lambda & J_{-\lambda}
\end{bmatrix}, \quad P = \begin{bmatrix}
    I^\pm & I^\pm
\end{bmatrix},
\]

where \( I^\pm = \text{diag}(1, -1, 1, \ldots, (-1)^{n-1}) \). Then

\[
J_{-\lambda} = -I^\pm J_\lambda I^\pm \Rightarrow P^* A P = -A.
\]

Moreover

\[
T_m^A(A) = \begin{bmatrix}
    T_m^A(J_\lambda) & T_m^A(J_{-\lambda})
\end{bmatrix},
\]

and

\[
\| T_m^A(J_{-\lambda}) \| = \| I^\pm T_m^A(-J_\lambda) I^\pm \| = \| T_m^A(J_\lambda) \|,
\]

i.e., the Chebyshev polynomial of \( A \) 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 for Matrices

Consider a block-diagonal matrix $A = \text{diag}(A_1, \ldots, A_h)$ where $d(A_j) \leq k$, $j = 1, \ldots, h$. Then the matrix

$$T^A_{k, \ell}(A) = \text{diag}(B_1, \ldots, B_h)$$

for $\ell = 1, 2, \ldots$, has at least $\ell + 1$ diagonal blocks $B_j$ such that

$$\| B_j \| = \| T^A_{k, \ell}(A) \|.$$
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 for Matrices** [Faber, Liesen, T. 2010]

Consider a block-diagonal matrix $\mathbf{A} = \text{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_h)$ where $d(\mathbf{A}_j) \leq k$, $j = 1, \ldots, h$. Then the matrix

$$T_{k \cdot \ell}^\mathbf{A}(\mathbf{A}) = \text{diag}(\mathbf{B}_1, \ldots, \mathbf{B}_h) \quad \ell = 1, 2, \ldots,$$

has at least $\ell + 1$ diagonal blocks $\mathbf{B}_j$ such that

$$\| \mathbf{B}_j \| = \| T_{k \cdot \ell}^\mathbf{A}(\mathbf{A}) \|.$$

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

\[ A = \text{diag}(A_1, A_2, A_3, A_4) \]

where each \( A_j = J_{\lambda_j} \) is a 3 \times 3 Jordan block. The four eigenvalues are \(-3, -0.5, 0.5, 0.75\), and \( k = d(A_j) = 3 \).
\[ A = \text{diag}(A_1, A_2, A_3, A_4) \]

where each \( A_j = J_{\lambda_j} \) is a \( 3 \times 3 \) Jordan block. The four eigenvalues are \(-3, -0.5, 0.5, 0.75\), and \( k = d(A_j) = 3 \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( | T_m^A(A_1) | )</th>
<th>( | T_m^A(A_2) | )</th>
<th>( | T_m^A(A_3) | )</th>
<th>( | T_m^A(A_4) | )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.6396</td>
<td>1.4620</td>
<td>2.3970</td>
<td>2.6396</td>
</tr>
<tr>
<td>2</td>
<td>4.1555</td>
<td>4.1555</td>
<td>3.6828</td>
<td>4.1555</td>
</tr>
<tr>
<td>3</td>
<td>9.0629</td>
<td>5.6303</td>
<td>7.6858</td>
<td>9.0629</td>
</tr>
<tr>
<td>5</td>
<td>22.3872</td>
<td>20.7801</td>
<td>17.6382</td>
<td>22.3872</td>
</tr>
<tr>
<td>6</td>
<td>22.6857</td>
<td>22.6857</td>
<td>20.3948</td>
<td>22.6857</td>
</tr>
</tbody>
</table>
Outline

1. General results

2. Matrices and sets in the complex plane
Chebyshev polynomials $T^\Omega_m(z)$ 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 $\Omega$ and $\Psi$ [Kamo, Borodin 1994]

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

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

be the pre-image of $\Omega$ under the polynomial map $p$. Then

$$T^{\Psi}_k(z) = T^\Omega_k(p(z)) .$$
Given are $\lambda_1, \ldots, \lambda_\ell \in \mathbb{C}$ and $n \geq 1$. Consider

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

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

[Reichel, Trefethen 1992] related the pseudospectra of $A$ to their symbol $f_A(z) = D + zE$. A
Chebyshev polynomials for lemniscates

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

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

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

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

Moreover,

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

- 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

\[
\min_{p \in \mathcal{M}_m} \| p(A) \|
\]

into the problem

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

where \( \Omega \) is a set in the complex plane associated with \( A \)?
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,

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,

More details can be found at

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

Thank you for your attention!