Absolute value circulant preconditioners for nonsymmetric Toeplitz-related systems

Sean Hon
joint work with Andy Wathen

Mathematical Institute
University of Oxford

NASCA 2018, 3rd July 2018

Outline

Introduction

- We are interested in solving the systems defined by analytic functions $h(z)$ of Toeplitz matrices T_n using Krylov subspace methods, i.e.

$$h(T_n)x = b.$$

- Examples: $h(T_n) = e^{T_n}$, $\sin T_n$, $\cos T_n$, etc.

Motivations:

- 1. Such a problem has been recently of interest for example in [Jin, Zhao, and Tam, 2014] and [Bai, Jin, and Yao, 2015].

  - Applications: e^{T_n} arises in option pricing, etc.

- 2. The idea of converting a nonsymmetric problem to a symmetric one + absolute value preconditioning.

Toeplitz matrices

$$T_n = \begin{bmatrix} a_0 & a_{-1} & \cdots & a_{-n+2} & a_{-n+1} \\ a_1 & \ddots & \ddots & \ddots & a_{-n+2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{n-2} & \ddots & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{bmatrix}$$

- For all $f \in L^1[-\pi, \pi]$, we let

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-\boldsymbol{i}kx} \, dx, \quad k = 0, \pm 1, \pm 2, \ldots,$$

be the Fourier coefficients of f. The function f is called the generating function/spectral symbol of T_n.

Preconditioning for A_n

- Let $A_n \in \mathbb{C}^{n\times n}$.

$$A_n x = b$$

- (Left) Preconditioning:

$$P_n^{-1} A_n x = P_n^{-1} b$$

- To speed up a typical Krylov subspace method, a good preconditioner P_n should satisfy:
  - 1. The product $P_n^{-1} d$ for any vector d needs to be efficiently computed.
  - 2. The spectrum of $P_n^{-1} A_n$ should be clustered.

Circulant matrices

$$C_n = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & \ddots & \ddots & \ddots & c_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_{n-2} & \ddots & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{bmatrix}$$

- (Eigendecomposition [Davis, 1979]) For any $C_n \in \mathbb{C}^{n\times n}$,

$$C_n = F_n^* \Lambda_n F_n$$

where F_n is the Fourier matrix.

Example: Optimal circulant preconditioners for T_n

- We let

$$\mathcal{M}_{F_n} = \{F_n^* \Lambda_n F_n \mid \Lambda_n \text{ is any } n\text{-by-}n \text{ diagonal matrix}\}$$

be the set of all circulant matrices.

- The optimal preconditioner $c(T_n) \in \mathbb{C}^{n \times n}$ [T. Chan, 1988] for T_n is defined to be the minimiser of

$$\|T_n - C_n\|_F$$

over all $C_n \in \mathcal{M}_{F_n}$, where $\|\cdot\|_F$ is the Frobenius norm.

1. Symmetrising real T_n

- Instead of directly dealing with a real Toeplitz system $T_n x = b$, [Pestana and Wathen, 2015] suggested that one can premultiply it by the anti-identity matrix Y_n, defined as

$$Y_n = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{n \times n},$$

to obtain the symmetric system $Y_n T_n x = Y_n b$.

1. Symmetrising real T_n

$$Y_n T_n = \begin{bmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ a_{n-2} & \iddots & \iddots & \iddots & a_{-1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ a_1 & \iddots & \iddots & \iddots & a_{-n+2} \\ a_0 & a_{-1} & \cdots & a_{-n+2} & a_{-n+1} \end{bmatrix}$$

- Since $Y_n T_n$ is symmetric yet (possibly) indefinite, we need a symmetric positive definite preconditioner to use MINRES.

2. Absolute value circulant preconditioner $|C_n|$

- The authors also proposed the absolute value circulant preconditioner
$$|C_n| = F_n^* |\Lambda_n| F_n.$$
- $|C_n|$ is symmetric positive definite provided that C_n is nonsingular.
- They further showed that the eigenvalues of the preconditioned matrix $|C_n|^{-1} Y_n T_n$ are clustered around ± 1.
- One can then use for example MINRES for $Y_n T_n$ with guaranteed convergence that depends only on its spectrum.

Example

- For example, we consider the $n \times n$ Grcar matrix T_n, i.e.

$$T_n = \begin{bmatrix} 1 & 1 & 1 & 1 & & \\ -1 & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & 1 \\ & & \ddots & \ddots & \ddots & 1 \\ & & & \ddots & \ddots & 1 \\ & & & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Example

- We have the corresponding symmetrised (Hankel) matrix

$$Y_n T_n = \begin{bmatrix} & & & & -1 & 1 \\ & & & \iddots & \iddots & 1 \\ & & \iddots & \iddots & \iddots & 1 \\ & \iddots & \iddots & \iddots & \iddots & 1 \\ -1 & \iddots & \iddots & \iddots & \iddots & \\ 1 & 1 & 1 & 1 & & \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Example

- For example, $|c(T_4)|$ derived from T_4 is given by

$$|c(T_4)| = \begin{bmatrix} 1.4231 & 0.1250 & 0.0769 & 0.1250 \\ 0.1250 & 1.4231 & 0.1250 & 0.0769 \\ 0.0769 & 0.1250 & 1.4231 & 0.1250 \\ 0.1250 & 0.0769 & 0.1250 & 1.4231 \end{bmatrix}.$$

Problem setup

- Under certain assumptions, we show that $|h(c(T_n))|$ is an effective preconditioner for

$$h(T_n)x = b.$$

  - T_n is the Toeplitz matrix generated by a function $f \in \mathcal{C}[-\pi, \pi]$, the Banach space of continuous complex-valued functions defined on $[-\pi, \pi]$ with the supremum norm $\|\cdot\|_\infty$.
  - $c(T_n)$: the optimal circulant preconditioner for T_n.

- Note that $h(T_n)$ is not Toeplitz in general.

1. Symmetrising real $h(T_n)$

- Given a $h(T_n) \in \mathrm{R}^{n \times n}$, we can premultiply it by Y_n to obtain a symmetric matrix $Y_n h(T_n)$.

- As $Y_n h(T_n)$ is (possibly) indefinite, MINRES with a symmetric positive definite preconditioner should be used.

2. $|h(C_n)|$ as a preconditioner for nonsymmetric $h(T_n)$

- $|h(C_n)|$ is a circulant matrix as

$$|h(C_n)| = F_n^*|h(\Lambda_n)|F_n.$$

- Hence, the product $|h(C_n)|^{-1}d$ for any vector d can be efficiently computed by a few Fast Fourier Transforms (FFTs) in $\mathcal{O}(n \log n)$ operations.

2. $|h(C_n)|$ as a preconditioner for nonsymmetric $h(T_n)$

Theorem (Hon, 2018 & Hon and Wathen 2018)

Suppose $h(z)$ with the radius of convergence r is analytic on $|z| < r$. Let $f \in \mathcal{C}[-\pi, \pi]$ with real Fourier coefficients such that $2\|f\|_\infty < r$. Let $T_n \in \mathbb{R}^{n\times n}$ be the Toeplitz matrix generated by f, let $c(T_n) \in \mathbb{R}^{n\times n}$ be the optimal circulant preconditioner for T_n, and let $Y_n \in \mathbb{R}^{n\times n}$ be the anti-identity matrix. If $\|h(c(T_n))^{-1}\|_2$ is uniformly bounded w.r.t. n, then for all $\epsilon > 0$ there exist positive integers N_ϵ and M_ϵ such that for all $n > N_\epsilon$

$$|h(c(T_n))|^{-1} Y_n h(T_n) = Q_n + R_n + E_n,$$

where Q_n is symmetric and orthogonal,

$$rank(R_n) \leq M_\epsilon,$$

and

$$\|E_n\|_2 \leq \epsilon.$$

Example. Nonsymmetric sinh T_n

- Consider sinh T_n, where T_n is the Grcar matrix.
- For sufficiently large n, we have

$$|\sinh c(T_n)|^{-1} Y_n \sinh T_n = Q_n + \widetilde{R}_n + \widetilde{E}_n.$$

- Since $Y_n \sinh T_n$, $|\sinh c(T_n)|$, and Q_n are all symmetric, the eigenvalues of

$$|\sinh c(T_n)|^{-1} Y_n \sinh T_n$$

are mostly close to ± 1.

Example

- For example, when $n = 4$,

$$\sinh T_4 = \begin{bmatrix} 0.8351 & 0.9110 & 0.7142 & 2.1095 \\ -1.1614 & 0.5847 & 0.2134 & 0.7142 \\ 0.4740 & -0.9378 & 0.5847 & 0.9110 \\ -0.2236 & 0.4740 & -1.1614 & 0.8351 \end{bmatrix} \in \mathbb{R}^{4\times 4}.$$

Example

- We have the symmetrised matrix

$$Y_4 \sinh T_4 = \begin{bmatrix} -0.2236 & 0.4740 & -1.1614 & 0.8351 \\ 0.4740 & -0.9378 & 0.5847 & 0.9110 \\ -1.1614 & 0.5847 & 0.2134 & 0.7142 \\ 0.8351 & 0.9110 & 0.7142 & 2.1095 \end{bmatrix}.$$

Example

- The corresponding $|\sinh c(T_4)|$ derived from T_4 is

$$|\sinh c(T_4)| = \begin{bmatrix} 1.6394 & 0.2971 & 0.5568 & 0.2971 \\ 0.2971 & 1.6394 & 0.2971 & 0.5568 \\ 0.5568 & 0.2971 & 1.6394 & 0.2971 \\ 0.2971 & 0.5568 & 0.2971 & 1.6394 \end{bmatrix}.$$

Example. Nonsymmetric sinh T_n

- sinh T_n, where T_n is a Grcar matrix.

Table: Numbers of iterations with MINRES for $Y_n \sinh T_n$

n	with no preconditioner	with $\lvert \sinh c(T_n) \rvert$
32	39	21
64	106	24
128	172	22
256	391	20

Example. Nonsymmetric sinh T_n

- sinh T_n, where T_n is a Grcar matrix.

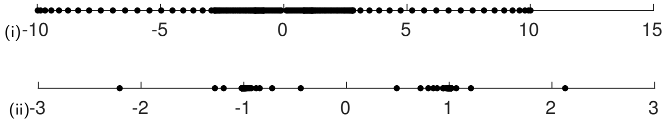

Figure: Spectrum of $Y_n \sinh T_n$ at $n = 256$ (i) without a preconditioner or (ii) with the preconditioner $|\sinh c(T_n)|$.

Conclusions

- We have provided the followings:

  1. One way (Y_n) to convert a nonsymmetric problem to a symmetric one + the absolute value preconditioner $|h(c(T_n))|$, and
  2. a theorem that accounts for the effectiveness of $|h(c(T_n))|$.

- An extension of this work to the multilevel Toeplitz case is straightforward.

Future work

- A direction for future work is to design effective preconditioners for ill-conditioned symmetric matrix $Y_n T_n$.

- Another is to investigate the asymptotic spectral distribution of symmetric matrix $Y_n T_n$.

  - Knowledge about the spectrum of $Y_n T_n$ could help designing good preconditioners.

Asymptotic spectral distribution of symmetrised Toeplitz matrices

- The following simple example illustrates the point: consider

$$T_n = \begin{bmatrix} 2 & & & \\ 1 & 2 & & \\ & \ddots & \ddots & \\ & & 1 & 2 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

- Its spectral symbol is $f(x) = 2 + e^{ix}$ and $n = 512$.

Asymptotic spectral distribution of symmetrised Toeplitz matrices

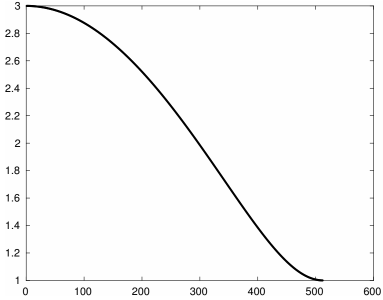

Figure: Singular value distribution of $Y_n T_n$ with T_n generated by $f(x) = 2 + e^{\mathrm{i}x}$ and $n = 512$.

- The singular values of $Y_n T_n$ are distributed as $|f(x)| = \sqrt{5 + 4\cos x}$.

Asymptotic spectral distribution of symmetrised Toeplitz matrices

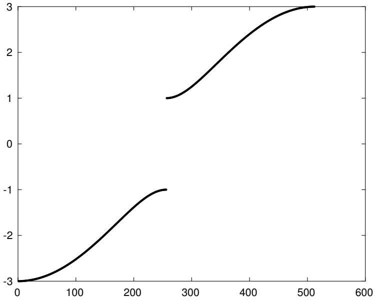

Figure: Spectral distribution of $Y_n T_n$ with T_n generated by $f(x) = 2 + e^{ix}$ and $n = 512$.

- The eigenvalues of $Y_n T_n$ seem to be distributed as $\pm|f|$. In fact, roughly half of them are negative/positive. This is joint work with Stefano Serra-Capizzano.

References

Z. Bai, X. Jin, and T. Yao
Superoptimal preconditioners for functions of matrices.
Numerical Mathematics: Theory, Methods and Applications, 8(4):515–529, 11 2015.

T. Chan
An optimal circulant preconditioner for Toeplitz systems.
SIAM Journal on Scientific and Statistical Computing, 9(4):766–771, 1988.

P. Davis
Circulant matrices
John Wiley & Sons, New York-Chichester-Brisbane, 1979.

S. Hon
Optimal preconditioners for systems defined by functions of Toeplitz matrices.
Linear Algebra and its Applications, 548:148 – 171, 2018.

S. Hon and A. Wathen
Circulant preconditioners for analytic functions of Toeplitz matrices.
Numerical Algorithms, Feb 2018.

X. Jin and Z. Zhao and S. Tam
Optimal preconditioners for functions of matrices
Linear Algebra and its Applications 457, 224 - 243, 2014.

J. Pestana and A. Wathen
A preconditioned MINRES method for nonsymmetric Toeplitz matrices.
SIAM Journal on Matrix Analysis and Applications, 36(1):273–288, 2015.