Toeplitz Matrices are Unitarily Similar to Symmetric Matrices

by

Jianzhen Liu

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 5, 2017

Keywords: Symmetric matrix, Toeplitz matrix, Numerical range

Approved by

Tin-Yau Tam, Chair, Professor of Mathematics
Ziqin Feng, Assistant Professor of Mathematics
Randall R. Holmes, Professor of Mathematics
Ming Liao, Professor of Mathematics

Abstract

We prove that Toeplitz matrices are unitarily similar to complex symmetric matrices. Moreover, two $n \times n$ unitary matrices that uniformly turn all $n \times n$ Toeplitz matrices via similarity to complex symmetric matrices are explicitly given, respectively. When $n \leq 3$, we prove that each complex symmetric matrix is unitarily similar to some Toeplitz matrix, but the statement is false when $n > 3$.

Acknowledgments

I would first like to acknowledge my advisor Dr. Tin-Yau Tam. He is more than an academic advisor. Indeed he is a mentor of my life. With his endless patience, he has been helping me to adjust to life in a new environment. He encourages all his students to be proactive to attend conferences and give talks and is really helpful to us. Dr. Tam and his wife, Kitty, always provide us a home-like environment.

I also want to thank my family. My wife, Xiang Li, has done her very best to take care of me and our son Aiden. Without her strong support, it would not be possible for me to arrive at this stage. At the beginning of our marriage, we talked about bringing our kid(s) to the commencement and sharing every special moment. Now that dream comes true. Thanks are also given to my brother and sister, who are living in China with my parents. They took good care of my mother when she had a stroke and I was away from home.

I would of course like to thank the Advisory Committee since all committee members are always supportive. The experience in the Department of Mathematics and Statistics is always amazing. The faculty and staff are very friendly. I was never disappointed when seeking help from them. The department provides enough financial support for the graduate students so they are able to attend conferences, in which they can communicate with their peers. Fellow graduate students are also willing to share their experiences in research as well as in teaching to help each other.

I thank Auburn University for providing me with such a friendly environment. As an international student, I have made friends with some local families and built up good relationships. The experience in Auburn has been invaluable.

Thanks should also be given to my coauthors, Professors Mao-Ting Chien, Hiroshi Nakazato, and Tin-Yau Tam. Without their help, it would have been impossible for me to finish my dissertation. I learnt a lot in working on the results. Their enthusiasm and encouragement on doing research have benefited me in my future academic career.

Table of Contents

Chapter 1

Introduction

Denote by $\mathbb{C}_{n\times n}$ the space of all $n \times n$ complex matrices and let $\mathbb{U}(n)$ be the group of $n \times n$ unitary matrices. It is well-known that every $A \in \mathbb{C}_{n\times n}$ is similar to a complex symmetric matrix [32, Theorem 4.4.24]. It means that given an appropriate basis, every linear transformation $L : \mathbb{C}_n \to \mathbb{C}_n$ can be represented by a complex symmetric matrix. When we refine the condition similarity with unitary similarity, one may ask whether every matrix is unitarily similar to a symmetric matrix. This is true when $n = 2$ [44]. However, it is not true [19, Example 7] when $n \geq 3$.

It is natural for us to ask, what kind of matrices are unitarily similar to complex symmetric matrices? There are some characterizations of the unitary similarity of a matrix with a complex matrix. Vermeer [60] gave some equivalent conditions. Liu, Nguelifack and Tam [40] extended the result of Vermeer in the context of an orthogonal symmetric Lie algebra of the compact type. Tener [55] gave some equivalent conditions in terms of the Hermitian decomposition $A = \frac{1}{2}(A+A^*)+\frac{1}{2}(A-A^*)$ of A. Garcia [20] obtained some criteria based on the singular value decomposition under the assumption that the singular values are distinct. In [20], some necessary and sufficient conditions for a 3×3 matrix to be unitarily similar to a complex symmetric matrix are presented, as well as an algorithm whereby an arbitrary 3×3 matrix can be tested. This test is generalized to a necessary and sufficient condition that applies to almost every $n \times n$ matrix. The test is constructive and explicitly exhibits the unitary equivalence to a complex symmetric matrix. Then it is possible to explicitly construct a conjugation C for which A is C-symmetric.

Toeplitz matrices, in which the elements on any line parallel to the principal diagonal are all equal, arise in many areas of systems theory. A basic reference is given by Grenander and Szegö [25], which includes applications to probability theory and statistics. Hartwig and Fisher [28] described some applications in chemical physics and stochastic processes. The survey on filtering theory is given by Kailath [35] in 1974 contains many useful references to Toeplitz and block Toeplitz matrix applications; in particular, canonical matrix fractions are studied by Dickinson [16] in 1974, the concept of innovations for stationary processes is given by Kailath [36], and relationships to the so-called Levinson-and-Chandrasekhar-type equations arising in estimation theory is given in [18] by Friedlander. Toeplitz matrices are also generalized as Toeplitz operators acting on the vector Hardy space $\mathcal{H}$ [7]. Toepltiz matrices are among the most well-studied structured matrices and a wide literature exists concerning Toeplitz matrices and the related matrices. It covers problems like the analysis of asymptotic spectral properties, where tools from functional analysis and operator theory are used; the study of related matrix algebras and fast discrete transforms; the analysis of preconditioners for the iterative solution of Toeplitz systems; the analysis of displacement operators, which enable one to represent the inverse of a Toeplitz matrix in a nice form; the analysis of fast and superfast algorithms for solving Toeplitz and Toeplitz-like linear systems with their interplay with Cauchy-like matrices.

A symmetric matrix is a square matrix that is equal to its transpose. It appears in many problems such as Grunsky inequality [32], moment problems [32], and in data fitting and quadrature applications [42, 1]. Because of their additional structure, symmetric matrices in many ways are simpler to deal with than the general matrices. It is well known that a real symmetric matrix can be diagonalized by an orthogonal transformation, so all the eigenvalues are real. For a general matrix, even it has real entries, the existence of non-real eigenvalue is possible.

Recently, Chien and Nakazato [13] proved that every Toeplitz matrix is unitarily similar to a complex symmetric matrix. Though the symmetric matrix referred in the result can be described, the unitary matrix that turns the given Toeplitz matrix to a complex symmetric matrix is not provided. We wonder that if the unitary matrix U that turns Toeplitz matrices into complex matrices can be given explicitly and in a simple form.

Computing the eigenvalues of a random $n \times n$ matrix is one of the most important tasks in linear algebra. The commonly used methods are turning a random matrix to some structured matrix before using the iterations. We hope to turn a matrix to a matrix with certain form via unitary similarity. Especially, we hope the unitary matrix is not too expensive in terms of complexity. Our main results given in Chapter 4 inspire us to develop more relating results which would be useful in the eigenproblems.

The remaining of this dissertation is organized as follow:

(i) In Chapter 2, we introduce the unitary similarity of two complex matrices. A classic criterion for two matrices being unitarily similar, known as Specht-Pearcy criterion, will be given. We also give Kippenhahn polynomials, which is a useful tool to study the unitary similarity between two matrices.

(ii) In Chapter 3, along with the introduction of Toeplitz matrix and some related forms, we will briefly discuss some important properties of these matrices. We then introduce the conditions for a complex matrix to be unitarily similar to a complex symmetric matrix.

(iii) We will give two unitary matrices that uniformly turn all $n \times n$ Toeplitz matrices into symmetric matrices via similarity and they will be given explicitly in Chapter 4. Thus, this constructive proof is an improvement of one of the main results in [13].

(iv) In Chapter 5, we present a standard form of complex symmetric matrices. We also study the problem of whether every symmetric matrix is unitarily similar to a Toeplitz matrix. The problem can be viewed as the (weak) converse of Theorem 4.1.3 and Theorem 4.1.1. When $n \leq 3$, it is true (see Theorem 4.2.2). However, the answer is negative when $n > 3$.

(v) Proofs to the claim that “not every complex symmetric matrix is unitarily similar to a Toeplitz matrix” are given for the case $n = 4$ and the case $n = 5$ in Chapter 6.

(vi) In Chapter 7, we will introduce some classic algorithms in finding eigenvalues of a matrix. Then we will discuss some possible uses of our results.

(vii) Our work can lead to further research in the some directions, which will be discussed in Chapter 8.

The main results in this dissertation can be found in [12].

Chapter 2

Unitary similarity and Kippenhahn polynomials

In this chapter, we will review some relating definitions concerning unitary similarity between two complex matrices and present some classic conditions of unitary similarity.

We denote by $\mathbb{C}_{n\times n}$ the set of complex matrix and the conjugate transpose of A by A^*.

Definition 2.0.1. A matrix $U \in \mathbb{C}_{n\times n}$ is said to be unitary if $U^*U = I$.

It can be easily seen that $U^{-1} = U^*$ if and only if U is unitary. Note that the set of $n \times n$ unitary matrices is a group, denoted as $\mathbb{U}(n)$. The group $\mathbb{U}(n)$ is a subgroup of general linear group $\mathrm{GL}(n, \mathbb{C})$, which is a Lie group. It plays a very important role in matrix theory. One of the commonly seen applications is unitary similarity. There are some basic equivalent conditions for a matrix U to be unitary:

Theorem 2.0.2. [32, Theorem 2.1.4] If $U \in \mathbb{C}_{n\times n}$, the following conditions are equivalent:

(i) U is unitary.

(ii) U is nonsingular and $U^{-1} = U^*$.

(iii) $U^*U = I$.

(iv) U^* is unitary.

(v) The columns of U form an orthonormal set.

(vi) For every $x \in \mathbb{C}_n$ and $y = Ux$, we have $y^*y = x^*x$.

Definition 2.0.3. Let $A \in \mathbb{C}_{n\times n}$. We define

(i) A complex scalar α is an eigenvalue of A if there is a nonzero vector x such that $Ax = \alpha x$, in which case we say that x is an eigenvector of A.

(ii) If $A^*A = AA^*$, then we call A normal. Note that a matrix A is normal if it is unitarily diagonalizable, that is, $U^*AU = \Lambda$, where U is unitary and Λ is diagonal.

(iii) If $A^* = A$, then we call A is Hermitian. Note that a Hermitian matrix is normal and its eigenvalues are real, that is, $\alpha^* = \alpha$, where α^* denotes the complex conjugate of α.

2.1 Unitary similarity

Two $n \times n$ matrices $A, B \in \mathbb{C}_{n\times n}$ are said to be similar if $B = P^{-1}AP$ for some invertible $n \times n$ matrix P. Similarity is an equivalence relation. Each equivalence class under similarity represent the same linear transformation under different bases. Two similar matrices share some essential properties such as trace and spectrum. Two unitarily similar matrices share even more properties including norm, numerical range, and so on. Because of this, for a given matrix A, we are interested in finding a simple "standard form" B which is similar to A, the study of A then reduces to the study of the simpler matrix B. For example, if A is similar to a diagonal matrix, we call that A is diagonalizable. By observing the diagonal matrix, we can easily see the eigenvalues, singularity and some other properties of A. However, not all matrices are diagonalizable. Over the complex field, every matrix is similar to a matrix in Jordan form, which is unique up to permutation of the Jordan blocks. Being able to turn a matrix to a matrix with a special structure via similarity has always been an important question in linear algebra.

Definition 2.1.1. Two matrices $A, B \in \mathbb{C}_{n\times n}$ are unitarily similar if there exists a unitary matrix U such that $B = U^*AU$, where $U^* = \bar{U}^T$.

Since finding U^* is much easier than finding the inverse of a general invertible matrix, unitary similarity is conceptually simpler than similarity. The equivalence classes under unitary similarity are finer than the equivalence classes under similarity.

One of the fundamentally useful facts of elementary matrix theory is a theorem given by I. Schur.

Theorem 2.1.2. [32, Theorem 2.3.1] Let $A \in \mathbb{C}_{n\times n}$ have eigenvalues $\lambda_1, \cdots, \lambda_n$ in any prescribed order and let $x \in \mathbb{C}_n$ be a unit vector such that $Ax = \lambda_1 x$.

(a) There is a unitary $U = [x \quad u_2 \quad \cdots \quad u_n] \in \mathbb{C}_n$ such that $U^*AU = T = [t_{ij}]$ is upper triangular with diagonal entries $t_{ii} = \lambda_i, i = 1, ..., n$.

(b) If $A \in \mathbb{R}_{n\times n}$ has only real eigenvalues, then x may be chosen to be real and there is a real orthogonal $Q = [x \quad q_2 \quad ... \quad q_n] \in \mathbb{R}_{n\times n}$ such that $Q^TAQ = T = [t_{ij}]$ is upper triangular with diagonal entries $t_{ii} = \lambda_i, i = 1, ..., n$.

This theorem is also called Schur Triangularization Theorem. Theorem 2.1.2 can be extended to a commuting family of complex matrices, that is, all matrices in the family can be reduced simultaneously to an upper triangular form by a unitary similarity.

Definition 2.1.3. A Hilbert space is a vector space $\mathcal{H}$ with an inner product $\langle \cdot, \cdot \rangle$ such that the norm defined by

$$|f| = \sqrt{\langle f, f \rangle}$$

turns $\mathcal{H}$ into a complete metric space.

A classical problem of operator theory is:

Let A and B be operators acting on a complex Hilbert space $\mathcal{H}$. How can one determine whether A and B are unitarily similar, that is, $B = U^*AU$ for some unitary operator U, where U^* is the Hermitian adjoint operator of U?

More precisely, the problem is to find a set of invariants that completely determine an operator up to unitary similarity. For a finite-dimensional Hilbert space $\mathcal{H}$, there are some criteria developed in the literature to determine if two matrices are unitarily similar.

Theorem 2.1.4. (Specht, [51]) Let $A,\, B \in \mathbb{C}_{n\times n}$. A is unitarily similar to B if and only if for all $r > 0$ and integers $\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_r \geq 0$,

$$\operatorname{tr}(A^{\alpha_1} A^{*\beta_1} \cdots A^{\alpha_r} A^{*\beta_r}) = \operatorname{tr}(B^{\alpha_1} B^{*\beta_1} \cdots B^{\alpha_r} B^{*\beta_r}).$$

This result actually is not practical in terms of checking the unitary similarity between two matrices. Pearcy [48] provided a refinement of Specht's theorem but it is still not practical.

Theorem 2.1.5. (Pearcy, [48]) Let $A,\, B \in \mathbb{C}_{n\times n}$. A is unitarily similar to B if and only if

$$\operatorname{tr}(A^{\alpha_1} A^{*\beta_1} \cdots A^{\alpha_r} A^{*\beta_r}) = \operatorname{tr}(B^{\alpha_1} B^{*\beta_1} \cdots B^{\alpha_r} B^{*\beta_r}).$$

holds for all monomials for which $\sum_i \alpha_i + \sum_i \beta_i \leq 2n^2$, where $\alpha_1, \cdots, \alpha_r, \beta_1, \cdots, \beta_r \geq 0$ are integers.

Note that Theorem 2.1.5 is a finite criterion, in theory computer can carry out the process. The number of traces required to determine the unitary similarity for word of degree at most 2^{n^2} arising from Pearcy's result in low-dimensional cases. It can be easily seen that when $n = 2$, three traces are enough to determine the unitary class; for the case $n = 3$, seven traces suffice [48, 50].

2.2 Kippenhahn polynomials

To study the unitary similarity between two matrices $A, B \in \mathbb{C}_{n\times n}$, we introduce a tool called Kippenhahn polynomial. First we introduce some invariants of the unitary similarity.

Definition 2.2.1. [15] The k-th numerical range of $A \in \mathbb{C}_{n\times n}$ is the set

$$\Lambda_k(A) := \{z \in \mathbb{C} : PAP = zP \text{ for some } k\text{-dimensional orthogonal projection } P\},$$

$1 \leq k \leq n$. When $k = 1$, $\Lambda_k(A)$ is reduced to the classical numerical range defined as

$$W(A) := \{\xi^* A \xi : \xi \in \mathbb{C}^n, \xi^* \xi = 1\},$$

which has been well studied in the literature [17, 27, 33, 37]. One of the most important properties about the numerical range is convexity.

Theorem 2.2.2. (Toeplitz-Hausdorff Theorem) Let $A \in \mathbb{C}_{n \times n}$. $W(A)$ is a compact convex set.

We remark that compactness is trivial as $W(A)$ is the image of the unit sphere under the continuous map $x \to x^* A x$.

If A is a normal matrix, that is, $A^* A = A A^*$, with eigenvalues $\lambda_1, \cdots, \lambda_n$, $W(A)$ is the convex hull of the eigenvalues of A. The numerical range of a Hermitian matrix H is a closed interval on the real axis, whose endpoints are the minimum and maximum eigenvalues of H. There are many important consequences of the convexity of numerical range.

Theorem 2.2.3. [33, Theorem 1.3.4] For each $A \in \mathbb{C}_{n \times n}$ there exists a unitary matrix $U \in \mathbb{U}(n)$ such that all the diagonal entries of $U^* A U$ equal to $\frac{\operatorname{tr} A}{n}$.

It is well-known that every complex matrix $A \in \mathbb{C}_{n \times n}$ can be uniquely split into two components so that $A = \Re(A) + i\Im(A)$, where $\Re(A) = (A + A^*)/2$ and $\Im(A) = (A - A^*)/(2i)$. Note that $\Re(A)$ and $\Im(A)$ are Hermitian matrices, $i\Im(A)$ is skew-Hermitian. The k-th numerical range $\Lambda_k(A)$ is completely determined by the following ternary form:

$$F_A(x, y, z) = \det(x\Re(A) + y\Im(A) + zI_n).$$

Kippenhahn [37] proved this result when $k = 1$. More precisely, $W(A)$ is the convex hull of the real affine part of the dual curve of $F_A(x, y, z) = 0$. In [21], it was proved that the equations $\Lambda_k(A) = \Lambda_k(B)$ $(1 \leq k \leq n)$ for $n \times n$ matrices A, B hold only if $F_A = F_B$. A matrix A and its transpose A^T have the common ternary form $F_A = F_{A^T}$.

It is known that if matrices A and B are unitary similar, then their numerical ranges are the same. However, the converse is in general not true. For example, $W(A) = W(B)$ is the closed unit disk centered at the origin, where

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

They are not similar since they have different spectra.

Helton and Spitovsky [31] showed that for every $A \in \mathbb{C}_{n\times n}$, there exists a complex symmetric $B \in \mathbb{C}_{n\times n}$ satisfying $F_B(x,y,z) = F_A(x,y,z)$, hence $\Lambda_k(A) = \Lambda_k(B)$. Their result depends on a result in [30], which answers affirmatively to the conjectures raised in [17] and [39], namely, for a hyperbolic ternary form $F(x,y,z)$, there exist real symmetric matrices H and K such that $F(x,y,z) = F_{H+iK}(x,y,z)$. The result of [30] provides us motivation to study the class of matrices which are unitarily similar to symmetric matrices. In [14], a method to construct symmetric matrices H, K starting from a hyperbolic form $F(x,y,z)$ is explicitly given when the curve $F(x,y,z) = 0$ has genus 0 or 1.

Chapter 3

Toeplitz matrices and symmetric matrices

Toeplitz matrix and symmetric matrices are two kinds of matrices that are well-studied and they play important roles across many areas. In this chapter, we will first introduce Toeplitz matrix and some of its relatives, such as Hankel matrix and Toeplitz band matrix. Some classical results about their properties will be discussed. We then introduce symmetric matrix and explore some algorithms about Toeplitz matrix and complex symmetric matrix.

3.1 Toeplitz matrices and related forms

Typical problems modelled by Toeplitz matrices include the numerical solution of certain differential equation, and certain integral equation, the computation of spline functions, time series analysis, signal and image processing, Markov chains and queueing theory, polynomial, and power series computations.

3.1.1 Toeplitz matrix

An $n \times n$ matrix $T = (a_{ij})$ is called a Toeplitz matrix if $a_{ij} = a_{k\ell}$ for every pairs (i, j), (k, ℓ) satisfying $i-j = k-\ell$. In this case, a_{ij} is denoted by a_{i-j} for some $a_0, a_{\pm 1}, a_{\pm 2}, \ldots, a_{\pm(n-1)}$. Explicitly,

$$T = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \ddots & & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \cdots & \cdots & a_2 & a_1 & a_0 \end{pmatrix}.$$

The entries of A are constant down the diagonals parallel to the main diagonal. The Toeplitz matrices

$$B = \begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{pmatrix} \text{ and } F = \begin{pmatrix} 0 & & & \mathbf{0} \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ \mathbf{0} & & 1 & 0 \end{pmatrix}$$

are called backward shift and forward shift because of their effect on the elements of the standard basis $\{e_1, \cdots, e_n\}$. Moreover, $F = B^T$ and $B = F^T$. A Toeplitz matrix T can be written as a linear combination of B, F and their powers, that is, $T = \sum_{k=0}^{n-1} a_{-k} B^k + \sum_{k=1}^{n-1} a_k F^k$.

An upper triangular Toeplitz matrix T can be presented as $T = \sum_{k=0}^{n} a_{-k} B^k$. We let $B^0 = I$. T is nonsingular if and only if $a_0 \neq 0$. Moreover,

$$T^{-1} = \sum_{k=0}^{n-1} b_{-k} B^k$$

is also an upper triangular Toeplitz matrix, where $b_0 = a_0^{-1}$ and $b_{-k} = a_0^{-1}(\sum_{m=0}^{k-1} a_{m-k} b_{-m})$ for $k = 1, 2, \cdots, n-1$. It can be clearly seen that the upper triangular Toeplitz matrices form a commutative algebra. Upper triangular $n \times n$ Toeplitz matrices along with some other special Toeplitz matrices and Toeplitz-like matrices, play important roles in some algorithms.

3.1.2 Hankel matrix

The term Hankel matrix, is coined by Gantmacher (1960) and is also called orthosymmetric matrix. It has constant skew diagonal. More specifically, a Hankel matrix is of the form:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_2 & a_3 & a_4 & \cdots & a_{n+1} \\ a_3 & a_4 & a_5 & \cdots & a_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n+1} & a_{n+2} & \cdots & a_{2n-1} \end{pmatrix}.$$

Complex Hankel matrices are special complex symmetric matrices. Hankel matrices and Toeplitz matrices can be related by making use of the square matrix

$$J = \begin{pmatrix} 0 & & 0 & 1 \\ & \iddots & \iddots & 0 \\ 0 & \iddots & \iddots & \\ 1 & 0 & & 0 \end{pmatrix}$$

which has units along the secondary diagonal and zeros elsewhere, noting that $J^2 = I$, the unit matrix. The effect of premultiplying any matrix by J reverses the order of its rows, and postmultiplying reverses the order of its columns. If A is a Hankel matrix, then JA and AJ are Toeplitz and $(JA)^T = AJ$. Although Toeplitz matrices and Hankel matrices are closely related to each other, their eigenvalues could be very different. There is extensive literature on inverting Hankel matrices or solving such linear systems. However, efficient eigenvalue algorithms for structured matrices are still under development. Taking advantage of two properties, namely that a complex Hankel matrix is symmetric and that a permuted Hankel matrix can be embedded in a circulant matrix, F. T. Luk [43] developed an $O(n^2 \log n)$ algorithm that can find all the eigenvalues of an $n \times n$ Hankel matrix.

3.1.3 Toeplitz band matrix

For a Toeplitz matrix $T_n = (a_i)$, if there are integers p, q, $0 \le p, q \le n-1$, such that

$$a_p \neq 0, a_{-q} \neq 0, a_r = 0, \text{ if } r > p \text{ or } r < -q,$$

we call T_n is a Toeplitz band matrix. Toeplitz band matrix was first introduced by Trench [57] in 1974. A simple case of Toeplitz band matrix has the form

$$T_n = \begin{pmatrix} a_0 & a_{-1} & \cdots & a_{-m} & 0 & \cdots & 0 \\ a_1 & a_0 & a_{-1} & & & \ddots & \vdots \\ \vdots & a_1 & a_0 & \ddots & & & 0 \\ a_m & & \ddots & \ddots & & & a_{-m} \\ 0 & \ddots & & & & & \vdots \\ \vdots & \ddots & \ddots & & \ddots & a_0 & a_{-1} \\ 0 & \cdots & 0 & a_m & \cdots & a_1 & a_0 \end{pmatrix}.$$

We call that T_n is a banded Toeplitz matrix with width m. If $a_i = a_{-i}$ for all $i = 1, \cdots, m$, T_n is a $n \times n$ banded symmetric matrix. For example,

$$T_5 = \begin{pmatrix} a_0 & a_1 & a_2 & 0 & 0 \\ a_1 & a_0 & a_1 & a_2 & 0 \\ a_2 & a_1 & a_0 & a_1 & a_2 \\ 0 & a_2 & a_1 & a_0 & a_1 \\ 0 & 0 & a_2 & a_1 & a_0 \end{pmatrix}$$

is a 5×5 banded symmetric Toeplitz matrix with bandwidth 2.

Trench [58] in 1964 derived an algorithm for inverting a Hermitian Toeplitz matrix with $O(n^2)$ operations (rather than $O(n^3)$, as required by standard matrix inversion methods) and stated a similar algorithm for the non-Hermitian case. In 1974, Trench [57] simplified the more general algorithm given in [58] for the case where of Toeplitz band matrix. Mentz [45] has developed an algorithm for calculating the inverse of an $n \times n$ symmetric Toeplitz band matrix with a band-width of $2m + 1$. The inverse is given in terms of the roots of a polynomial equation, but is only valid if n is large in relation to m, Hoskins, Ponzo [34] and Rehnqvist [49] gave inversion methods for special cases of symmetric band matrices. These algorithms require that all principal minors be nonzero.

3.1.4 Circulant matrix

A circulant matrix C is a special Toeplitz matrix having the form

$$C = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & \cdots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \cdots & \cdots & c_{n-3} \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ c_2 & c_3 & c_4 & & \ddots & c_1 \\ c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_0 \end{pmatrix}.$$

Each row is the previous row cycled forward one step and the entries in each row are a cyclic permutation of those in the first. Circulant matrices arise and are prevalent in areas like applications involving the discrete Fourier transform and the study of cyclic codes for error correction. There is a very beautiful result on the eigenvectors of circulant matrices:

Theorem 3.1.1. [24, Theorem 3.1] Every circulant matrix C has eigenvectors

$$y^{(m)} = \frac{1}{\sqrt{n}}(1, e^{\frac{2\pi i m}{n}}, \cdots, e^{\frac{2\pi i m(n-1)}{n}})^T, \quad m = 1, \cdots, n-1$$

with eigenvalues

$$\lambda_m = \sum_{k=0}^{n-1} c_k e^{\frac{e2imk}{n}}, \quad m = 1, \cdots, n-1,$$

and can be expressed in the form $C = U\Lambda U^*$, where U has the eigenvectors as columns in order and Λ is $\operatorname{diag}(\lambda_k)$. In particular all circulant matrices share the same eigenvectors, the same matrix U works for all circulant matrices, and any matrix of the form $C = U\Lambda U^*$ is circulant.

3.2 Some behaviors of Toeplitz matrices

Due to great importance, a wide range of literature focuses on Toeplitz matrices and the behavior of Toeplitz matrices. The most complete references were given by Grenander and Szegö [25] and Widom [61]. A more recent text devoted to the subject is Böttcher and Silbermann [9] in 1999. However, some of the behavior are particularly critical to the study of Toeplitz matrices.

3.2.1 Szegö's Theorem

Szegö's Theorem is probably the most famous and most important result on Toeplitz matrices. Given a sequence of Toeplitz matrices $\{T_n\}$, Szegö's Theorem deals with the asymptotic behavior of the eigenvalues of T_n as n goes to infinity. Though we will not explicitly review Szegö's Theorem here, one can see more details in [24].

3.2.2 Toeplitz (or Hankel) decomposition

One of the top ten algorithms of the twentieth century is the "decompositional approach to matrix computation". Several decompositions of a matrix can be seen in almost every linear algebra book. Their applications are extensively seen and being cornerstones of modern computations.

QR decomposition (also called a QR factorization) of $A \in \mathbb{C}_{n\times n}$ is a decomposition of a matrix into the product $A = QR$ of an orthogonal matrix Q and an upper triangular matrix R. QR decomposition is often used to solve the linear least squares problem, and is the basis for a particular eigenvalue algorithm, the QR *algorithm*. QR decomposition is the matrix version of the Gram-Schmidt Orthogonalization Process.

LU decomposition (also called LU factorization) factors a square matrix as the product of a lower triangular matrix L and an upper triangular matrix U. The LU decomposition can be viewed as the matrix form of Gaussian elimination if no row permutation is involved in the elimination. Computers usually solve square systems of linear equations using the LU decomposition, and it is also a key step when inverting a matrix, or computing the determinant of a matrix.

The polar decomposition of $A \in \mathbb{C}_{n\times n}$ is of the form $A = UP$, where U is a unitary matrix and P is a positive-semidefinite matrix. Intuitively, the polar decomposition separates A into a component that stretches the space along a set of orthogonal axes, represented by P, and a rotation (with possible reflection) represented by U.

The singular value decomposition (SVD) is a factorization of $A = V\Sigma U^*$ for $A \in \mathbb{C}_{n\times n}$, where $\Sigma = \operatorname{diag}(\sigma_1, \cdots, \sigma_n)$ and $\sigma_1 \geq \cdots \geq \sigma_n$ are the singular values of A. Notice that $U, V \in \mathbb{U}(n)$ are not uniquely determined. However, if the singular values are distinct, then U and V are respectively unique up to the post-multiplication of a diagonal unitary matrix. Clearly $A^*AU = U^2$ and $AA^*V = V^2$. The SVD can be obtained from the polar decomposition and vice versa. It has many useful applications in signal processing and statistics.

Matrix decompositions provide a platform on which a variety of scientific and engineering problems can be solved. Once computed, they may be reused repeatedly to solve new problems involving the original matrix and may often be updated or downdated with respect to small changes in the original matrix. Furthermore, they permit reasonably simple rounding-error analysis and afford high-quality software implementations.

Very recently, Ye and Lim [63] developed a new kind of decomposition that utilizing Toeplitz matrices and Hankel matrices.

Theorem 3.2.1. (Ye and Lim, [63]) Every $n \times n$ matrix is a product of finite Toeplitz (Hankel) matrices.

A Toeplitz or a Hankel decomposition of a given matrix A may not be as easily computable as LU or QR, but once computed, these decompositions can be reused ad infinitum for any problem involving A. If A has a known Toeplitz decomposition with r factors, one can solve linear systems in A within $O(rn \log^2 n)$ time via any of the superfast algorithms in [2, 8, 11, 53, 59]. However, there is no efficient algorithm so far to explicitly decompose a matrix to Toeplitz matrices or Hankel matrices. The situation makes this decomposition not as applicable as the previous classic decompositions.

3.3 Complex symmetric matrices

An $n \times n$ matrix S is called a symmetric matrix if $S^T = S$. The entries are symmetric with respect to the main diagonal. So if the entries are written as $S = (a_{ij})$, then $a_{ij} = a_{ji}$, for all indices i and j. Complex symmetry is a purely algebraic property, and it has no effect on the spectrum of the matrix. Indeed, for any given set of n numbers,

$$\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C},$$

there exists a complex symmetric $n \times n$ matrix S whose eigenvalues are just the prescribed numbers.

Theorem 3.3.1. [32, Theorem 4.4.24] Each $A \in \mathbb{C}_{n\times n}$ is similar to a complex symmetric matrix.

The result is derived directly from the fact that each Jordan block is similar to a symmetric matrix and each $A \in \mathbb{C}_{n\times n}$ is similar to a direct sum of Jordan blocks.

A complex symmetric matrix may not even be diagonalizable. For example, consider the complex symmetric matrix

$$S = \begin{pmatrix} 2i & 1 \\ 1 & 0 \end{pmatrix},$$

where $i = \sqrt{-1}$. The only eigenvalue of this matrix is $\lambda = i$, with algebraic multiplicity 2 but geometric multiplicity 1. In fact, the Jordan normal form of S is

$$Z^{-1}SZ = \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix},$$

where $Z = \begin{pmatrix} i & 1 \\ 1 & 0 \end{pmatrix}$. Thus, S is not diagonalizable.

However, a complex symmetric matrix can be diagonalized using a unitary matrix. Thus if A is a complex symmetric matrix, there is a unitary matrix U such that UAU^T is a real diagonal matrix. This result is referred to as the Autonne-Takagi factorization. Autonne-Takagi factorization is also called symmetric singular value decomposition (SSVD). It was originally proved by Autonne [3] in 1915 and Takagi [54] in 1925 and rediscovered with different proofs by several other mathematicians. When A is a complex symmetric matrix, Autonne-Takagi factorization takes the form:

$$A = U\Sigma U^T,$$

where U is unitary and Σ is diagonal. The columns of U are called the Takagi vectors.

A necessary and sufficient condition for a complex symmetric to be diagonalizable is:

Theorem 3.3.2. [32, Theorem 4.4.27] Let $S \in \mathbb{C}_{n\times n}$ be symmetric. Then S is diagonalizable if and only if it is complex orthogonally diagonalizable, that is, $S = O\Lambda O^T$, where O is a complex orthogonal matrix and Λ is diagonal.

3.4 Characterizations of unitary similarity of a square complex matrix to a symmetric matrix

Complex square matrix being unitarily similar to a complex symmetric matrix arises from many subjects and one of them is eigenproblem. In numerical linear algebra, the preprocessing step of a typical algorithm for computing eigenvalues and singular values of a matrix is turning the matrix to an intermediate matrix admitting low cost iterations in the second step. For some subclasses of normal matrices, e.g., Hermitian, skew-Hermitian, and unitary matrices, the intermediate matrix shapes admit a low storage cost $O(n)$ and, as such, permit the design of QR algorithms with linear complexity steps. Unfortunately, for the generic normal matrix class, the intermediate matrices are of Hessenberg form, requiring $O(n^2)$ storage and resulting in a quadratic cost for each QR-step. An upper Hessenberg matrix contains zeros below the first subdiagonal. If the matrix is symmetric or Hermitian, then the form is tridiagonal. This matrix has the same eigenvalues as the original, but less computation is needed to reveal them. An alternative intermediate condensed form might thus result in significant computational savings. To achieve this goal, the use of intermediate complex symmetric matrices that can be constructed using unitary similarities is useful. The problem of determining whether a square complex matrix is unitarily similar to a complex symmetric one has been intensively studied. In general, it is difficult to determine if a complex matrix is unitarily similar to a complex symmetric matrix. The first general approach was given by Vermeer, who obtained the following characterizations:

Theorem 3.4.1. [60, Theorem 3] Let $A \in \mathbb{C}_{n\times n}$. The following statements are equivalent.

(i) A is unitarily similar to a complex symmetric matrix.

(ii) There is a symmetric unitary matrix U such that UAU^* is symmetric.

(iii) There is a symmetric unitary matrix U and a symmetric matrix S such that $A = SU$.

(iv) There is a symmetric unitary matrix V such that $VAV^* = AT$.

Definition 3.4.2. A function $C : \mathbb{C}_n \to \mathbb{C}_n$ is called a *conjugation* if it satisfies:

(i) $C(\alpha x + \beta y) = \bar{\alpha} Cx + \bar{\beta} Cy$ for all $x, y \in \mathbb{C}_n$ and all $\alpha, \beta \in \mathbb{C}$ (conjugate-linear).

(ii) $C(C(x)) = x$ for all $x \in \mathbb{C}_n$ (involutive).

(iii) $\|Cx\| = \|x\|$ for all $x \in \mathbb{C}_n$ (isometric).

An example of conjugation C is the so-called *flip conjugation*, that is, $C(x) = (\bar{x}_n, \cdots, \bar{x}_1)^T$ for any $x = (x_1, \cdots, x_n)^T \in \mathbb{C}_n$.

The matrices that are unitarily equivalent to complex symmetric matrices can be characterized in terms of C-symmetry.

Theorem 3.4.3. (Garcia and Putinar, [19]) A matrix A is unitarily similar to a complex symmetric matrix if and only if there exists a conjugation C such that $A = CA^*C$.

Note that such A is called *C-symmetric*.

Garcia [20] obtained a criteria based upon the diagonalization of A^*A and AA^* under the assumption that the singular values are distinct.

Theorem 3.4.4. [20, Theorem 1] Suppose that $A \in \mathbb{C}_{n\times n}$ has distinct singular values. If

(i) $u_1, u_2, \cdots, u_n$ are unit eigenvectors of A^*A corresponding to the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, respectively,

(ii) $v_1, v_2, \cdots, v_n$ are unit eigenvectors of AA^* corresponding to the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, respectively,

then A is unitary similar to a complex symmetric matrix if and only if

$$|\langle u_i, v_j\rangle| = |\langle u_j, v_i\rangle|,$$

$$\langle u_i, v_j\rangle\langle u_j, v_k\rangle\langle u_k, v_i\rangle = \langle u_i, v_k\rangle\langle u_k, v_j\rangle\langle u_j, v_i\rangle,$$

holds for

$$1 \leq i \leq j \leq k \leq n.$$

The proof of Theorem 3.4.4 given in [20] is a constructive one. The proof implies that if A satisfies the two necessary and sufficient conditions, then there exist unimodular constants $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that $Cu_i = \alpha_i v_i$ for $i = 1, 2, \cdots, n$, where C is a conjugate-linear operator.

The condition of distinct singular values is not necessary. Basing upon singular value decomposition, a different proof to Theorem 3.4.4 and some related results are given by Liu, Nguelifack and Tam.

Theorem 3.4.5. [40, Theorem 2.2] Let $A \in \mathbb{C}_{n\times n}$ and let $\Sigma = \sigma_1 I_{n_1} \bigoplus \cdots \bigoplus \sigma_m I_{n_m}$, where $\sigma_1 > \sigma_2 > \cdots > \sigma_m$ are the distinct singular values of A with multiplicities $n_1, \cdots, n_m$, respectively. The following statements are equivalent.

(i) A is unitarily similar to a complex symmetric matrix.

(ii) There is a SVD $A = Y\Sigma X^*$ such that X^*Y is symmetric.

(iii) For any SVD $A = V\Sigma U^*$, there are block diagonal unitary matrices $Q := Q_1 \bigoplus \cdots \bigoplus Q_m$ and $Q' := Q_1 \bigoplus \cdots \bigoplus Q_{m-1} \bigoplus Q'_m$, conformal to Σ, such that $(UQ)^*VQ'$ is symmetric. If $\sigma_m > 0$, then $Q = Q'$.

Note that Vermeer's theorem [60] can be deduced by Theorem 3.4.5.

Chapter 4

Every Toeplitz matrix is unitarily similar to a complex symmetric matrix

In this chapter, we are going to prove that every Toeplitz matrix T is unitarily similar to a complex symmetric matrix S. We will explicitly give two unitary matrices U such that $S = U^*TU$. Thus the corresponding complex symmetric matrices can also be given. We then consider the inverse of the claim when $n \leq 3$.

4.1 Two unitary matrices that turn Toeplitz matrices to symmetric matrices via similarity

The $n \times n$ Jordan block $J_n(0)$ corresponding to the zero eigenvalue is a Toeplitz matrix and [32, p. 208] $J_n(0)$ is unitarily similar to the symmetric matrix

$$\frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} + \frac{i}{2}\begin{pmatrix} 0 & \cdots & 0 & -1 & 0 \\ \vdots & \iddots & \iddots & \iddots & 1 \\ 0 & \iddots & \iddots & \iddots & 0 \\ -1 & \iddots & \iddots & \iddots & \vdots \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$$

via $U = \frac{1}{\sqrt{2}}(I + iJ) \in \mathbb{U}(n)$, where J is the $n \times n$ backward identity.

The property that $J_n(0)$ is unitarily similar to a symmetric matrix is also true for arbitrary Toeplitz matrix $T \in \mathbb{C}_{n\times n}$ and it was proved by Chien and Nakazato [13]. We are going to give a different proof. Indeed, we explicitly give unitary matrices U that *uniformly* transform all Toeplitz matrices in $\mathbb{C}_{n\times n}$ to symmetric matrices.

A Toeplitz matrix can be viewed as a linear combination of Jordan block $J_n(0)$, the transpose of $J_n(0)$, and their powers. It is not difficult to see that the unitary matrix $U = \frac{1}{\sqrt{2}}(I + iJ) \in \mathbb{U}(n)$ can also turn all Toeplitz matrices to symmetric matrices.

Theorem 4.1.1. Every Toeplitz matrix T is unitarily similar to a symmetric matrix $B = (b_{ij})$ via the unitary matrix $U = \frac{1}{\sqrt{2}}(I + iJ) \in \mathbb{U}(n)$. More specifically,

$$b_{ij} = \frac{1}{2}(a_{i-j} + a_{j-i}) + \frac{i}{2}(a_{i+j-n-1} - a_{n+1-i-j}).$$

Proof. Since $U^* = \frac{1}{\sqrt{2}}(I - iJ)$ and $JTJ = T^T$, where $J = J_n(0)$, we have

$$\begin{aligned} U^*TU &= \frac{1}{2}(I - iJ)T(I + iJ) \\ &= \frac{1}{2}(T + iTJ - iJT + T^T) \\ &= \frac{1}{2}(T + T^T) + \frac{i}{2}(TJ - JT) \end{aligned}$$

Note that $T + T^T$ is symmetric, TJ and JT are Hankel matrices (see [32, 0.9.8]), which are symmetric. Hence, U^*TU is symmetric. $\square$

Example 4.1.2. When $n = 4$, $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 1 & i & 0 \\ 0 & i & 1 & 0 \\ i & 0 & 0 & 1 \end{pmatrix}$ and

$$U^*TU = \frac{1}{2} \begin{pmatrix} 2a_0 + i(a_{-3} - a_3) & a_1 + a_{-1} + i(a_{-2} - a_2) & a_2 + a_{-2} + i(a_{-1} - a_1) & a_3 + a_{-3} \\ a_1 + a_{-1} + i(a_{-2} - a_2) & 2a_0 + i(a_{-1} - a_1) & a_1 + a_{-1} & a_2 + a_{-2} + i(a_1 - a_{-1}) \\ a_2 + a_{-2} + i(a_{-1} - a_1) & a_1 + a_{-1} & 2a_0 + i(a_1 - a_{-1}) & a_1 + a_{-1} + i(a_2 - a_{-2}) \\ a_3 + a_{-3} & a_2 + a_{-2} + i(a_1 - a_{-1}) & a_1 + a_{-1} + i(a_2 - a_{-2}) & 2a_0 + i(a_3 - a_{-3}) \end{pmatrix}$$

We remark that in Theorem 4.1.1 the unitary U that uniformly turns all Toeplitz matrices to symmetric matrices via similarity is not unique.

The results given in [19] state that for a conjugation C over $\mathbb{C}_n$, there exists an orthonormal basis $\{e_1, \cdots, e_n\}$ such that $Ce_i = e_i$ for all $i = 1, \cdots, n$. Note that this orthonormal basis $\{e_1, \cdots, e_n\}$ is called C-*real*. If there exists a conjugation C that makes A C-symmetric, the unitary matrix $U = (e_1 \quad e_2 \quad \cdots \quad e_n)$ can turn A to a complex symmetric matrix U^*AU. Direct computation yields the result that every Teoplitz matrix T is C-symmetric under the flip conjugation C. Then it leaves us to find a C-real orthonormal basis under the flip conjugation C. We have the following theorem.

Theorem 4.1.3. Every Toeplitz matrix $T \in \mathbb{C}_{n\times n}$ is unitarily similar to a symmetric matrix. Moreover, the following $U \in \mathbb{U}(n)$ uniformly turns all Toeplitz matrices in $\mathbb{C}_{n\times n}$ into symmetric matrices via similarity:

1. When $n = 2m$, with $m \geq 1$,

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & & & & & i \\ & \ddots & & & \iddots & \\ & & 1 & i & & \\ & & 1 & -i & & \\ & \iddots & & & \ddots & \\ 1 & & & & & -i \end{pmatrix}.$$

2. When $n = 2m + 1$, with $m \geq 1$,

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & & & & & & i \\ & \ddots & & & & \iddots & \\ & & 1 & 0 & i & & \\ & & 0 & \sqrt{2} & 0 & & \\ & & 1 & 0 & -i & & \\ & \iddots & & & & \ddots & \\ 1 & & & & & & -i \end{pmatrix}.$$

Proof. Clearly U is unitary and we write $U = (u_1 \cdots u_n)$ in column form and let $B := U^*TU = (b_{st})$. Our goal is to show that B is symmetric. Note that $b_{st} = u_s^* T u_t$.

(1) When $n = 2m$,

$$u_k = \begin{cases} (e_k + e_{2m-k+1})/\sqrt{2} & k \leq m \\ (e_{2m-k+1} - e_k)i/\sqrt{2} & k > m. \end{cases}$$

By straightforward computation, we have

$$b_{st} = b_{ts} = \begin{cases} \frac{1}{2}(a_{t-s} + a_{s-t} + a_{s+t-2m-1} + a_{2m+1-s-t}), & s \leq m, t \leq m \\ \frac{1}{2}i(a_{t-s} - a_{s-t} + a_{s+t-2m-1} - a_{2m+1-s-t}), & s \leq m, t > m \\ \frac{1}{2}(a_{t-s} + a_{s-t} - a_{s+t-2m-1} - a_{2m+1-s-t}), & s > m, t > m. \end{cases}$$

(2) When $n = 2m+1$,

$$u_k = \begin{cases} (e_k + e_{2m-k+2})/\sqrt{2} & k \le m \\ e_{m+1} & k = m+1 \\ (e_{2m-k+2} - e_k)i/\sqrt{2} & k > m+1. \end{cases}$$

Straightforward computation yields

$$b_{st} = b_{ts} = \begin{cases} \frac{1}{2}(a_{t-s} + a_{s-t} + a_{s+t-2m-2} + a_{2m+2-s-t}) & s \le m, t \le m \\ \frac{\sqrt{2}}{2}(a_{m+1-s} + a_{s-m-1}) & s \le m, t = m+1 \\ \frac{1}{2}i(a_{t-s} - a_{s-t} + a_{s+t-2m-2} - a_{2m+2-s-t}) & s \le m, t > m+1 \\ a_0 = \frac{1}{2}2a_0 & s = m+1, t = m+1 \\ \frac{\sqrt{2}}{2}i(a_{t-m-1} - a_{m+1-t}) & s = m+1, t > m+1 \\ \frac{1}{2}(a_{t-s} + a_{s-t} - a_{s+t-2m-2} - a_{2m+2-s-t}) & s > m+1, t > m+1. \end{cases}$$

It follows that $B = U^*TU$ is symmetric. $\square$

We remark that the corresponding symmetric matrices $B = U^*TU$ given by Theorem 4.1.1 and Theorem 4.1.3 are different.

Example 4.1.4. When $n = 4$, $U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & i \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -i \end{pmatrix}$ and

$$U^*TU = \frac{1}{2}\begin{pmatrix} 2a_0 + a_3 + a_{-3} & a_1 + a_{-1} + a_2 + a_{-2} & i(a_2 - a_{-2} + a_{-1} - a_1) & i(a_3 - a_{-3}) \\ a_1 + a_{-1} + a_2 + a_{-2} & 2a_0 + a_1 + a_{-1} & i(a_1 - a_{-1}) & i(a_2 - a_{-2} + a_1 - a_{-1}) \\ i(a_2 - a_{-2} + a_{-1} - a_1) & i(a_1 - a_{-1}) & 2a_0 - a_1 - a_{-1} & a_1 + a_{-1} - a_2 - a_{-2} \\ i(a_3 - a_{-3}) & i(a_2 - a_{-2} + a_1 - a_{-1}) & a_1 + a_{-1} - a_2 - a_{-2} & 2a_0 - a_3 - a_{-3} \end{pmatrix}$$

4.2 Every complex symmetric matrix is unitarily similar to a Toeplitz matrix when $n \le 3$

Denote by S_n the subspace of complex symmetric matrices in $\mathbb{C}_{n\times n}$ and $\mathbb{T}_n$ the set of Toeplitz matrices in $\mathbb{C}_{n\times n}$. Theorem 4.1.1 and Theorem 4.1.3 imply that $\{U^*TU : U \in \mathbb{U}(n)\}$ and S_n have nonempty intersection for all $T \in \mathbb{T}_n$.

Given $S \in S_n$, can we find a unitary matrix U such that USU^* is Toeplitz? If the answer is affirmative, then it can be viewed as a (weak) converse to Theorem 4.1.1 and Theorem 4.1.3, that is, every symmetric $S \in \mathbb{C}_{n\times n}$ is unitarily similar to a Toeplitz matrix.

It is not hard to see that the claim is true when $n = 2$ since each $A \in \mathbb{C}_{n\times n}$ is unitarily similar to a matrix of equal diagonal entries [33, p.18]. How about the 3×3 case? The answer is affirmative and we are going to prove it. We first note that for any 3×3 complex matrix

$$S = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{12} & s_{22} & s_{23} \\ s_{13} & s_{23} & s_{33} \end{pmatrix},$$

we have

$$USU^* = \frac{1}{2}\begin{pmatrix} s_{11}+s_{33} & \sqrt{2}(s_{12}+is_{23}) & s_{11}+2is_{13}-s_{33} \\ \sqrt{2}(s_{12}-is_{23}) & 2s_{22} & \sqrt{2}(s_{12}+is_{23}) \\ s_{11}-2is_{13}-s_{33} & \sqrt{2}(s_{12}-is_{23}) & s_{11}+s_{33} \end{pmatrix},$$

where U is the unitary matrix given in Theorem 4.1.3. So, if $s_{11}+s_{33}=2s_{22}$, then the matrix USU^* is Toeplitz. Let us return to the case that S is symmetric. If we can find a rotation matrix W such that $B=(b_{ij})=WSW^T$ satisfies

$$b_{11}+b_{33}=2b_{22}, \tag{4.2.1}$$

then we have the desired result by applying the unitary similarity via U to B for the 3×3 case. We will show that such a rotation matrix exists.

Denote by $\mathrm{SO}(n)$ the $n\times n$ proper orthogonal group. Let $S\in S_3$ and $\tilde{S}=S-\frac{1}{3}(\operatorname{tr}S)I_3$. Then $\operatorname{tr}\tilde{S}=0$. If we can show that $\tilde{S}$ is unitarily similar to some Toeplitz matrix T, then S is unitarily similar to the Toeplitz matrix $T+\frac{1}{3}(\operatorname{tr}S)I_3$. Thus we may assume $\operatorname{tr}S=0$.

Lemma 4.2.1. Suppose that $S\in S_3$ satisfying $\operatorname{tr}S=0$. Then there is $W\in\mathrm{SO}(3)$ such that the (2, 2)-entry of WSW^T is 0, and hence

$$WSW^T = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{pmatrix}$$

for some $b_{11},b_{12},b_{13},b_{23},b_{22},b_{33}\in\mathbb{C}$ with $b_{22}=0$, $b_{33}=-b_{11}$. It follows that $b_{11}+b_{33}-2b_{22}=0$.

Proof. Note that $\operatorname{tr} S = 0$. By a result of Brickman [10], (also see [4, 41]), the range

$$W(S) := \{(WSW^T)_{22} : W \in \mathrm{SO}(3)\}$$

is convex. Since s_{11}, s_{22}, $s_{33} \in W(S)$,

$$0 = \frac{1}{3}\operatorname{tr} S = \frac{1}{3}(s_{11} + s_{22} + s_{33}) \in W(S).$$

So there is $W \in \mathrm{SO}(3)$ such that $(WSW^T)_{22} = 0$. $\square$

Theorem 4.2.2. Any 3×3 complex symmetric matrix is unitarily similar to some 3×3 Toeplitz matrix.

Proof. Let $S \in S_3$. Let $\tilde{S} = S - \frac{1}{3}(\operatorname{tr} S)I_3$. By Lemma 4.2.1, there exists $W \in \mathrm{SO}(3)$ such that $\tilde{B} := (\tilde{b}_{ij}) = W\tilde{S}W^T$ and $\tilde{b}_{22} = 0$; thus $\tilde{b}_{11} + \tilde{b}_{33} = 2\tilde{b}_{22}$. Now $B := (b_{ij}) = WSW^T = W\tilde{S}W^T + \frac{1}{3}(\operatorname{tr} S)I_3$ is symmetric and $b_{11} + b_{33} = 2b_{22}$ and $b_{22} = \frac{1}{3}\operatorname{tr} S$, that is, (4.2.1) is satisfied. By the previous discussion

$$UBU^* = UWSW^TU^* = UWS(UW)^*$$

is Toeplitz, where U is the 3×3 unitary matrix given in Theorem 4.1.3. $\square$

Chapter 5

A standard form of complex symmetric matrices

To consider the class of complex symmetric matrices which are unitarily similar to Toeplitz matrices, we introduce a new standard form for complex symmetric matrices.

Lemma 4.2.1 can be extended in the following theorem.

Theorem 5.0.3. Let $S \in S_n$. There exists $W \in \mathrm{SO}(n)$ for which the diagonal entries $(d_1, d_2, \ldots, d_n)$ of the complex symmetric matrix $W^T SW$ satisfy

$$d_j = \frac{2}{n}\operatorname{tr} S - d_{n+1-j}$$

for $j = 1, 2, \ldots, n$. In particular, $d_j = -d_{n+1-j}$ when $\operatorname{tr} S = 0$ for $j = 1, 2, \ldots, n$.

Proof. Without loss of generality, we may assume that $\operatorname{tr} S = 0$. It suffices to prove that there exists $W \in \mathrm{SO}(n)$ for which the diagonal entries $(d_1, d_2, \ldots, d_n)$ of $W^T SW$ satisfy $d_j = -d_{n+1-j}$ for $j = 1, 2, \ldots, n$.

It is trivial when $n = 2$ and Lemma 4.2.1 handles the $n = 3$ case.

Now we let $n \geq 4$. When $n = 2m$, let $C = \operatorname{diag}(1, 0, \ldots, 0, 1)$. By a result of Au-Yeung and Tsing [4] (also see [41, Theorem 11.7]), the range

$$W_C(S) = \{(WSW^T)_{11} + (WSW^T)_{nn} : W \in \mathrm{SO}(n)\}$$

is convex. If $0 \notin W_C(S)$, we can separate $W_C(S)$ from 0 by the line $x = a$ for some $a > 0$ by rotating the range. We may assume $W_C(S) \subset \{z \in \mathbb{C} : \Re(z) \geq a\}$. This relation implies that

$$\begin{aligned}
\Re((WSW^T)_{11} + (WSW^T)_{nn}) &\geq a, \\
\Re((WSW^T)_{22} + (WSW^T)_{(n-1)(n-1)}) &\geq a, \\
&\cdots \\
\Re((W^T)_{mm} + (WSW^T)_{(m+1)(m+1)}) &\geq a,
\end{aligned}$$

and hence $\Re(\operatorname{tr}(WSW^T)) = \Re(\operatorname{tr} S) \geq ma > 0$, which contradicts $\operatorname{tr} S = 0$. So we have $0 \in W_C(S)$, and thus there exists $W \in \mathrm{SO}(n)$ such that the diagonal entries $(d_1, d_2, \ldots, d_n)$ of $W^T SW$ satisfy $d_n = -d_1$. The argument can be used to prove $d_j = -d_{n-j+1}$ by taking $C = \operatorname{diag}(0_{j-1}, 1, 0_{n-2j}, 1, 0_{j-1})$. Hence, we have $d_j = -d_{n+1-j}$, where $j = 1, \cdots, m$.

When $n = 2m+1$, let $C = \operatorname{diag}(0_m, 1, 0_m)$. Using the idea in Lemma 4.2.1 we can prove that there exists $W \in \mathrm{SO}(n)$ such that the diagonal entries $(d_1, \ldots, d_{m+1}, \ldots, d_n)$ of $W^T SW$ satisfy $d_{m+1} = 0$. Then we consider the $(2m) \times (2m)$ matrix $\tilde{S}$ obtained by deleting the $(m+1)$-st row and column from S. We then apply the inductive hypothesis to $\tilde{S}$ and complete the proof of our assertion. □

We consider the class of 4×4 standard form of complex symmetric matrices S with $\operatorname{tr} S = 0$:

$$S = \begin{pmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{12} & s_{22} & s_{23} & s_{24} \\ s_{13} & s_{23} & -s_{22} & s_{34} \\ s_{14} & s_{24} & s_{34} & -s_{11} \end{pmatrix},$$

which is parametrized by the 8 complex numbers $s_{11}, s_{22}, s_{12}, \ldots, s_{34}$. In order to have $S = U^* TU$ for some $T \in \mathbb{T}_4$ and the unitary matrix U given in Section 3, the following two equations are necessary and sufficient conditions:

$$s_{12} + s_{34} - 2s_{22} = 0, s_{13} - s_{24} + 2s_{23} = 0.$$

A general form of $\tilde{T} = USU^*$ for the above S is given by

$$\tilde{T} = \begin{pmatrix} 0 & t_{12} & t_{13} & t_{14} \\ t_{21} & 0 & t_{23} & t_{13} \\ t_{31} & t_{32} & 0 & t_{12} \\ t_{41} & t_{31} & t_{21} & 0 \end{pmatrix},$$

which is parametrized by 8 complex numbers $t_{12}, t_{13}, t_{14}, t_{23}$ and t_{21}, t_{31}, t_{41}, t_{32}. We denote by $\tilde{\mathbb{T}}_4$ the complex vector space of matrices of the form

$$\tilde{T} = \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{11} & t_{23} & t_{13} \\ t_{31} & t_{32} & t_{11} & t_{12} \\ t_{41} & t_{31} & t_{21} & t_{11} \end{pmatrix}.$$

The matrix $\tilde{T}$ is Toeplitz if $t_{23} = t_{12}$, $t_{32} = t_{21}$.

Chapter 6

Not every symmetric symmetric matrix is unitarily similar to a Toeplitz matrix

Although every complex symmetric matrix is unitarily similar to a symmetric matrix when $n \le 3$, there exist symmetric matrices that are not unitarily similar to any Toeplitz when $n \ge 4$. people may understand this intuatively by counting the variables in a Toeplitz matrix and symmetric matrix, respectively. More specifically, an $n \times n$ Toeplitz contains $2n-1$ variables and a complex symmetric matrix of the same size has $\frac{n(n+1)}{2}$ variables, which are different when $n = 4$. However, we actually compare the unitary orbits of the two kinds of matrices. In the following sections, we will compare the dimensions of the two orbits for cases $n = 4$ and $n = 5$, respectively.

6.1 4×4 matrices

In this section, we shall prove that the following inclusion is proper:

$$\{UTU^* : T \in \mathbb{T}_4, U \in \mathbb{U}(4)\} \subset \{USU^* : S \in S_4, U \in \mathbb{U}(4)\}.$$

We first establish that the following theorem.

Theorem 6.1.1. The set

$$\{UTU^* : T \in \mathbb{T}_4, \operatorname{tr} T = 0, U \in \mathbb{U}(4)\}, \tag{6.1.1}$$

is parametrized by real 26-variables. The following set

$$\{USU^* : S \in \tilde{\mathbb{T}}_4, \operatorname{tr} S = 0, U \in \mathbb{U}(4)\} \tag{6.1.2}$$

contains the above set (6.1.1) and its dimension is 27. Hence there is a set $S \in \tilde{\mathbb{T}}_4$ with $\operatorname{tr} S = 0$ for which USU^* does not belong to (6.1.1) for any $U \in \mathbb{U}(4)$.

Proof. We first examine the set (6.1.1). Note that the dimension of the real vector space

$$\{T \in \mathbb{T}_4, \operatorname{tr} T = 0\}$$

is 12. The special unitary group SU(4) is a 15-dimensional real analytic manifold with the real tangent space at the identity composed of the 4×4 skew-Hermitian matrices of zero trace:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix}$$

with $x_{11} = ip_1$, $x_{22} = ip_2$, $x_{33} = ip_3$, $x_{44} = -i(p_1 + p_2 + p_3)$, $x_{12} = r_1 + is_1$, $x_{21} = -r_1 + is_1$, $x_{13} = r_2 + is_2$, $x_{31} = -r_2 + is_2$, $x_{14} = r_3 + is_3$, $x_{41} = -r_3 + is_3$, $x_{23} = r_4 + is_4$, $x_{32} = -r_4 + is_4$, $x_{24} = r_5 + is_5$, $x_{42} = -r_5 + is_5$, $x_{34} = r_6 + is_6$, $x_{43} = -r_6 + is_6$, where p_1, p_2, p_3, $r_1, \ldots, r_6$, $s_1, \ldots, s_6$ are 15 real parameters. For a general point $U_g T_0 U_g^*$ of the set

$$\{UTU^* : T \in \mathbb{T}_4, \operatorname{tr} T = 0, U \in \mathrm{SU}(4)\},$$

we shall estimate the dimension of its tangent space. By using the operation

$$Z \mapsto U_g^* Z U_g,$$

we may assume that $U_g = I$ and restrict ourselves to consider the dimension of the tangent space at $T_0 \in \mathbb{T}_4$ with $\operatorname{tr} T_0 = 0$. By the Taylor expansion of an element of a neighborhood of T_0, we obtain

$$\exp(tX)(T_0 + tT_1)\exp(-tX) = T_0 + tT_1 + tXT_0 - tTX_0 + O(t^2)$$

for $X \in \mathbb{C}_{4\times 4}, X^* = -X, \operatorname{tr} X = 0$ and $T_1 \in \mathbb{T}_4, \operatorname{tr} T_1 = 0$. These T_1's form a 12 dimensional real vector space. We compute the dimension of the derivation range

$$\{T_0 X - XT_0 : X \in \mathbb{C}_{4\times 4}, X^* = -X, \operatorname{tr} X = 0\}$$

modulo the vector space

$$\{\tilde{T} \in \mathbb{T}_4 : \operatorname{tr} \tilde{T} = 0\}.$$

We denote by W_{ij} the (i,j)-entry of $W = TX - XT$. Let

$$W_1 = W_{1,1}, W_2 = W_{2,2}, W_3 = W_{3,3}, W_4 = W_{1,2} - W_{3,4}, W_5 = W_{2,1} - W_{4,3},$$

$$W_6 = W_{1,3} - W_{2,4}, W_7 = W_{3,1} - W_{4,2}, W_8 = W_{2,3} - W_{3,4}, W_9 = W_{3,2} - W_{4,3}.$$

Each W_j can be expressed as

$$W_j = t_{j,1}p_1 + t_{j,2}p_2 + t_{j,3}p_3 + \sum_{k=4}^{9} t_{j,k-3}r_k + \sum_{10}^{15} t_{j,k-9}s_k$$

for some complex coefficients $t_{j,k}$. The coefficients $t_{j,k}$ satisfy

$$t_{j,k} = 0,$$

for $j, k = 1, 2, 3$. We are going to give the coefficient vectors

$$P_k = (t_{4,k}, t_{5,k}, t_{6,k}, t_{7,k}, t_{8,k}, t_{9,k}), \quad k = 1, 2, 3.$$

They are

$$\begin{aligned} P_1 &= (0,0,0,0,ia_1-b_1,-ic_1+d_1), \\ P_2 &= 2(ia_1-b_1,-ic_1+d_1,ia_2-b_2,-ic_2+d_2,0,0), \\ P_3 &= (2ia_1-2b_1,-2ic_1+2d_1,2ia_2-2b_2,-2ic_2+2d_2,3ia_1-3b_1,-3ic_1+3d_1), \end{aligned}$$

and hence the vectors $P_j's$ satisfy the linear equation

$$P_3 - P_2 - 3P_1 = 0.$$

Hence, the rank of the 18×15 matrix $(\Re(t_{j,k}), \Im(t_{j,k}))^T$ is necessarily less than or equal to 14. By taking a rather general coefficient a_j, b_j, c_j, d_j, the rank of such a matrix is just 14. Thus, we conclude that the dimension of the set (6.1.1) is 26.

We then examine the set (6.1.2) and show that it has dimension $27 = 16 + 11$. We take a generic matrix $\tilde{T}$ in $\tilde{\mathbb{T}}_4$ with $\operatorname{tr} \tilde{T} = 0$ as follows:

$$\tilde{T} = \begin{pmatrix} 0 & 2-2i & 3+7i & 7+8i \\ 1+4i & 0 & 23+3i & 3+7i \\ 23+7i & 13-3i & 0 & 2-2i \\ -11+11i & 23+7i & 1+4i & 0 \end{pmatrix},$$

at which we consider the tangent space of the set (6.1.2). By the Taylor expansion of an element of a neighborhood of $\tilde{T}$, we obtain

$$\exp(tX)(\tilde{T} + tT_1)\exp(-tX) = \tilde{T} + tT_1 + X\tilde{T} - \tilde{T}X + O(t^2)$$

for $X \in \mathbb{C}_{4\times 4}, X^* = -X, \operatorname{tr} X = 0$ and $T_1 \in \tilde{\mathbb{T}}_4, \operatorname{tr} T_1 = 0$. These T_1's form a 16 dimensional real vector space. We compute the dimension of the derivation range

$$\{\tilde{T}X - X\tilde{T} : X \in \mathbb{C}_{4\times 4}, X^* = -X, \operatorname{tr} X = 0\}$$

modulo the vector space

$$\{T_1 \in \tilde{\mathbb{T}}_4 : \operatorname{tr} T_1 = 0\}.$$

We denote by W_{ij} the (i,j)-entry of $W = TX - XT$. Let

$$W_1 = W_{1,1}, W_2 = W_{2,2}, W_3 = W_{3,3}, W_4 = W_{1,2} - W_{3,4},$$

$$W_5 = W_{2,1} - W_{4,3}, W_6 = W_{1,3} - W_{2,4}, W_7 = W_{3,1} - W_{4,2}.$$

Let

$$\Re(W_j) = c_{2j-1,1}p_1 + c_{2j-1,2}p_2 + c_{2j-1,3}p_3 + c_{2j-1,4}r_1 + \cdots + c_{2j-1,9}r_6 + c_{2j-1,10}s_1$$

$$+ \ldots + c_{2j-1,15}s_6,$$

$$\Im(W_j) = c_{2j,1}p_1 + c_{2j,2}p_2 + c_{2j,3}p_3 + c_{2j,4}r_1 + \cdots + c_{2j,9}r_6 + c_{2j,10}s_1 + \ldots + c_{2j,15}s_6,$$

$j = 1, \ldots, 7$. We consider the 14×15 matrix $\tilde{T} = (c_{ij}) = [A|B]$:

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & -26 & 4 & 0 & 0 \\ 0 & 0 & 0 & -2 & -14 & -19 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & -36 & -26 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & -14 \\ 0 & 0 & 0 & 0 & 26 & 0 & 36 & 0 \\ 0 & 0 & 0 & 0 & 14 & 0 & 0 & 0 \\ 0 & 4 & 4 & 0 & -20 & -46 & -6 & -20 \\ 0 & 4 & 4 & 0 & -5 & -14 & -14 & -5 \\ 0 & 8 & 8 & 0 & -12 & -6 & -46 & -12 \\ 0 & -2 & -2 & 0 & -14 & -14 & -14 & -14 \\ 0 & -14 & -14 & -30 & 0 & -2 & 4 & 0 \\ 0 & 6 & 6 & -11 & 0 & -8 & -4 & 0 \\ 0 & 14 & 14 & -2 & 0 & -4 & 2 & 0 \\ 0 & -46 & -46 & -8 & 0 & 4 & 8 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 0 & 6 & 0 & 3 & 0 & 0 & 0 \\ 0 & 1 & -20 & 18 & 0 & 0 & 0 \\ 0 & -6 & 0 & 0 & -6 & 0 & 0 \\ 0 & -1 & 0 & 0 & 10 & -20 & -20 \\ -3 & 0 & 0 & 0 & 6 & 0 & 0 \\ 0 & -2 & 20 & 0 & -10 & 0 & 0 \\ 0 & 0 & -11 & 14 & -14 & -11 & -11 \\ 0 & 0 & -6 & -46 & 6 & -6 & -6 \\ 0 & 0 & 8 & -14 & 14 & 8 & 8 \\ 0 & 0 & 34 & 6 & -46 & 34 & 34 \\ -30 & -5 & 0 & 8 & 4 & 0 & 0 \\ -11 & -16 & 0 & -2 & 4 & 0 & 0 \\ -2 & 14 & 0 & 4 & 8 & 0 & 0 \\ -8 & 24 & 0 & 4 & -2 & 0 & 0 \end{pmatrix}.$$

We are going to show that the rank of $\tilde{T}$ is at least 11. For this purpose we delete the first two columns and the last two columns from $\tilde{T}$ without increasing the rank. These columns correspond to the variables p_1, p_2, s_5, s_6. We also delete the first two rows and the last row from $\tilde{T}$ without increasing the rank. These two rows correspond to the $(1,1)$-entry of $W = TX - XT$ and the imaginary part of $W_7 = W_{3,1} - W_{4,2}$. Then we obtain an 11×11 real invertible matrix. This shows that the set (6.1.2) has a tangent space of dimension at least 27 at the point $\tilde{T}$. In fact, with some additional computations we can show that the rank is just 11 but it is unnecessary for the proof of the assertion in the theorem, so the set (6.1.1) cannot cover the set (6.1.2). Thus, we just proved that there exists a 4×4 complex symmetric matrix which is not unitarily similar to any 4×4 Toeplitz matrix. $\square$

6.2 5×5 matrices

We shall prove the following theorem.

Theorem 6.2.1. There exists a 5×5 complex symmetric matrix A with $\operatorname{tr} A = 0$ for which F_A is not realized by F_T for any $T \in \mathbb{T}_5$ with $\operatorname{tr} T = 0$.

Proof. The real vector space of all 5×5 Toeplitz matrices T with $\operatorname{tr} T = 0$ is parametrized as

$$\begin{pmatrix} 0 & p_1 + iq_1 & p_2 + iq_2 & p_3 + iq_3 & p_4 + iq_4 \\ p_5 + iq_5 & 0 & p_1 + iq_1 & p_2 + iq_2 & p_3 + iq_3 \\ p_6 + iq_6 & p_5 + iq_5 & 0 & p_1 + iq_1 & p_2 + iq_2 \\ p_7 + iq_7 & p_6 + iq_6 & p_5 + iq_5 & 0 & p_1 + iq_1 \\ p_8 + iq_8 & p_7 + iq_7 & p_6 + iq_6 & p_5 + iq_5 & 0 \end{pmatrix}$$

by 16 real parameters $p_1, \ldots, p_8$, $q_1, \ldots, q_8$. The Kippenhahn polynomial of a general 5×5 unitarily symmetrizable matrix $\tilde{S}$ with $\operatorname{tr} \tilde{S} = 0$ is expressed by

$$F_{\tilde{S}}(t, x, y) = t^5 + c_1 t^3 x^2 + c_2 t^3 xy + c_3 t^3 y^2 + c_4 t^2 x^3 + c_5 t^2 x^2 y + c_6 t^2 xy^2$$

$$+ c_7 t^2 y^3 + c_8 tx^4 + c_9 tx^3 y + c_{10} tx^2 y^2 + c_{11} txy^3 + c_{12} ty^4 + c_{13} x^5 + c_{14} x^4 y$$

$$+ c_{15} x^3 y^2 + c_{16} x^2 y^3 + c_{17} xy^4 + c_{18} y^5,$$

where $c_1, \ldots, c_{18}$ are real coefficients. We will identify $F_{\tilde{S}}(t, x, y)$ by its coefficient vector $\bar{F}_{\tilde{S}} := (c_1, c_2, \ldots, c_{18})$. We consider the subspace

$$\{\bar{F}_T = (c_1, c_2, \ldots, c_{18}) : T \in \mathbb{T}_5, \operatorname{tr} T = 0\}.$$

The map

$$(p_1, q_1, \ldots, p_8, q_8) \mapsto (c_1, \ldots, c_{18})$$

is a polynomial map and hence it is infinitely differentiable. So the set of points $(c_1, \ldots, c_{18})$ for 5×5 Toeplitz matrices T with $\operatorname{tr} T = 0$ has dimension ≤ 16. We show that there exists a 5×5 complex symmetric matrix S for which the linear perturbation $\tilde{S} = 2S + H + iK$ of $2S$ by the matrices H, K:

$$H = \begin{pmatrix} a_1 & a_5 & a_9 & a_{12} & a_{14} \\ a_5 & a_2 & a_6 & a_{10} & a_{13} \\ a_9 & a_6 & a_3 & a_7 & a_{11} \\ a_{12} & a_{10} & a_7 & a_4 & a_8 \\ a_{14} & a_{13} & a_{11} & a_8 & a_{12} \end{pmatrix}, \quad K = \begin{pmatrix} b_1 & 0 & 0 & 0 & 0 \\ 0 & b_2 & 0 & 0 & 0 \\ 0 & 0 & b_3 & 0 & 0 \\ 0 & 0 & 0 & b_4 & 0 \\ 0 & 0 & 0 & 0 & b_5 \end{pmatrix},$$

with $a_{12} = -(a_1+a_2+a_3+a_4)$, $b_5 = -(b_1+b_2+b_3+b_4)$ for real coefficients $b_1, \ldots, b_4, a_1, \ldots, a_{14}$. Then the matrix $\tilde{S} = 2S + H + iK$ has non-vanishing Jacobian

$$\frac{\partial(c_1, \ldots, c_4, c_5, \ldots, c_{18})}{\partial(b_1, \ldots, b_4, a_1, \ldots, a_{14})}$$

at $(b_1, \ldots, b_4, a_1, \ldots, a_{18}) = (0, \ldots, 0, 0, \ldots, 0)$ for some symmetric matrix S. In fact, let

$$S = \begin{pmatrix} 0 & 2+3i & 1-3i & 4+2i & -3+5i \\ 2+3i & 4+2i & 3-2i & 3+5i & 4-6i \\ 1-3i & 3-2i & 2+i & 2+2i & 1-i \\ 4+2i & 3+5i & 2+2i & -2-i & 1+i \\ -3+5i & 4-6i & 1-i & 1+i & -4-2i \end{pmatrix}.$$

Then we compute the above Jacobian by using computer software. This value does not vanish and is about 3.81737×10^{57}. Hence, there exists a 5×5 symmetric matrix A with $\operatorname{tr} A = 0$ for which F_A is not realized by F_T for any $T \in \mathbb{T}_5$ with $\operatorname{tr} T = 0$. $\square$

By comparing the dimensions of the unitary of orbits of the space of Toeplitz matrices of zero trace and the space of complex symmetric matrices of standard form in Theorem 5.0.3

of zero trace, we proved that not every complex symmetric is unitarily similar to a Toepltiz matrix when $n = 4$. Theoretically, this method can be used to prove the result when $n \geq 5$. However, it is costly in terms of complexity. We used Kippenhahn polynomial proved the result when $n = 5$.

Chapter 7

Algorithms for computing eigenvalues of Toeplitz matrices and symmetric matrices

Algorithms for computing the eigenvalues and singular values of a matrix are amongst the most important ones in numerical linear algebra. Finding the eigenvalues of a square matrix A is equivalent to finding the roots of characteristic polynomial $\det(\lambda I - A)$. According to the well-known Abel-Ruffini theorem which states that there is no general formula for the roots of polynomials of degree greater than 4, there is no direct formula solver to the eigenvalues of matrix A except some special cases.

In the eigenproblem of a matrix, Hessenberg matrices play significant roles. The major reason is so-called stability. If we are able to convert a matrix to a triangular matrix in finite steps while preserving the eigenvalues, for example, similarity, we can then see all the eigenvalues on the main diagonal of the triangular matrix. Unfortunately, there is no general method like Gaussian elimination. But it is possible to reach some matrices closed to triangular. An upper Hessenberg matrix is "almost" triangular except the entries on the subdiagonal. Hessenberg matrices and tridiagonal matrices are the commonly used intermediate matrices used in many eigenvalue algorithms because the zero entries reduce the complexity of the problem.

Various methods such as iterative (e.g. Lanczos, Arnoldi) as well as the so-called direct methods (e.g. divide-and-conquer algorithms, GR-methods) exist. Many of the procedures for computing eigenvalues and/or singular values are based on the QR-method. In general, the ideal to find the eigenvalues and eigenvector of a random matrix $A \in \mathbb{C}_{n\times n}$ is to produce better approximate solutions with each iteration. To reduce the complexity in the computations, some matrices that have big advantages are used in iterative algorithms. So, most of the eigenvalue methods are based on a two-step approach. In the first phase, the matrix is transformed to a suitable condensed matrix format, sharing the eigenvalues, and in the second stage the eigenvalues of this condensed matrix are computed. This step usually costs $O(n^3)$ operations. The main purpose of this intermediate matrix is saving valuable computing time. Important subclasses of normal matrices, such as the Hermitian, skew-Hermitian, and unitary matrices admit a condensed matrix represented by only $O(n)$ parameters, allowing subsequent low-cost algorithms to compute their eigenvalues.

Let us give a brief review for some classic algorithms used in the eigenproblems.

7.1 QR-algorithm

Many of the procedures for computing eigenvalues or singular values are based on the QR-algorithm. The QR-algorithm computes a Schur decomposition of a matrix. It is certainly one of the most important algorithm in eigenvalue computations [46].

For a matrix $A \in \mathbb{C}_{n\times n}$ the standard QR-algorithm consists of two steps:

(i) The first step is to transform the matrix A in finite steps by similarity transformation to a suitable shape which will make the iterations in the second step cost as low as possible. Usually, the term "suitable" means Hessenberg form. The first step is essential since generically it reduces the global computational complexity of the next step by one order (e.g. from $O(n^4)$ to $O(n^3)$ for an arbitrary matrix A). Only for specific subclasses of the normal matrix class is an efficient preprocessing step developed, resulting in an $O(n)$ parameter representation of the transformed matrix. In general it is not possible to obtain.

(ii) The second step consists of repeatedly applying QR-steps on the matrix until the eigenvalues are revealed. The overall complexity (number of floating points) of the algorithm is $O(n^3)$, which we will see is not entirely trivial to obtain.

The QR-algorithm is a very powerful algorithm to stably compute the eigenvalues and the corresponding eigenvectors or Schur vectors. All steps of the algorithm cost $O(n^3)$

floating point operations. The one exception is the case where only eigenvalues are desired of a symmetric tridiagonal matrix.

7.2 Lanczos based iterative algorithms for complex symmetric matrices

Note that

(i) Not every complex symmetric matrix is diagonalizable, although it is the case for Hermitian matrices.

(ii) The eigenvalues for an ordinary complex symmetric matrix do not have special properties.

(iii) The straight reduction of a dense complex symmetric matrix to a tridiagonal form is not always stable.

(iv) so there is no robust theory of the complex symmetric tridiagonal eigenproblem.

So there is no effective software to solve the symmetric eigenproblem. While the complex symmetry of A has no effect on the eigenvalues of A. This particular structure can be exploited to halve the work and storage requirements of the general non-Hermitian Lanczos method. The Lanczos algorithm is a direct algorithm devised by Lanczos [38] that is an adaptation of power methods to find the most useful eigenvalues and eigenvectors of an n^{th} order linear system with a limited number, which is much smaller than n, of operations. Although iterative methods are very effective for solving large and sparse eigenproblems, they are not very practical for the dense problems, especially when all the eigenvalues of the matrix are required. Bar-On and Ryaboy [6] in 1997 gave a new fast directed algorithm, which is similar in concept to the standard complex Hermitian eigensolver. The main stages of the algorithm can be summarized as follows:

(i) Tridiagonal reduction. Reduce the complex symmetric matrix $S \in \mathbb{C}_{n\times n}$ into a tridiagonal complex symmetric matrix $T \in \mathbb{C}_{n\times n}$ by a sequence of complex orthogonal transformations.

(ii) Complex orthogonal QR. Compute the complete set of eigenvalues of the tridiagonal matrix, using the complex orthogonal QR algorithm, and extract the eigenvalues whose eigenvectors are of further interest.

(iii) Inverse iteration. Compute the eigenvectors of the tridiagonal matrix, corresponding to the subset of eigenvalues required, by inverse iteration.

(iv) Back transformation. Compute the corresponding eigenvectors of the original dense matrix by back transformations.

This method has been shown in [5] that considerably outperform the general eigensolver when computing the eigenvalues of complex symmetric matrices. Tables are also given in [5] to the advantage of the method in terms of complexity.

7.3 Superfast divide-and-conquer method for Toeplitz matrices

The set of Toeplitz matrices is a very typical and useful class of structured matrices. However, it is nothing special for a Toeplitz matrix when it comes to eigenvalues. However, we expect to have a much fast method to solve the eigenproblem of Toeplitz matrix due to the special structure of a Toeplitz matrix.

A method that is widely used to solve Toeplitz system is using displacement equation methods. With displacement structures, Toeplitz matrices can be transformed into Cauchy-like matrices using the FFT (fast Fourier transform) or other trigonometric transformations. Methods introduce in [26], [29] and [22] have complexity $O(N^2)$. Algorithms for solving Toeplitz systems are usually classified in fast algorithms if their complexity is $O(n^2)$ and superfast algorithms if their complexity is $O(n \log^2 n)$.

For the eigenproblem of the Teoplitz matrices, people usually use QR-algorithm to deal with. The method does not completely utilize the structure of Toeplitz matrix. In [62], a superfast divide-and-conquer method is presented for finding all the eigenvalues as well as all the eigenvectors (in a structured form) of a class of symmetric matrices with off-diagonal ranks or numerical ranks bounded by r, as well as the approximation accuracy of the eigenvalues due to off-diagonal compression. More specifically, the complexity is $O(r^2 n \log n) + O(rn \log^2 n)$, where n is the order of the matrix. Such matrices are often encountered in practical computations with banded matrices, Toeplitz matrices (in Fourier space), and certain discretized problems. In the same paper, the application of the algorithm on Toeplitz matrix is explained. The nearly linear complexity is proven and is verified with applications such as Toeplitz and discretized matrices.

Chapter 8

Future works

Our work can lead to further research in the following directions.

(i) The main results, Theorem 4.1.3 and Theorem 4.1.1 can be deduced to multilevel Toeplitz matrix. Very recently, Koyuncu et al. have shown that any multilevel Toeplitz matrix is unitarily similar to a complex matrix basing on Theorem 4.1.3 and Theorem 4.1.1. The result is not yet published.

(ii) The main results, Theorem 4.1.3 and Theorem 4.1.1, state that every Toeplitz matrix is unitarily similar to a symmetric matrix. Note that our matrices are all finite. It inspires us to ask: whether every Toeplitz operator is unitarily similar to a complex symmetric operator on Hardy space $\mathcal{H}$. There is no answer to this question so far.

(iii) The algorithm given in [62] has advantage on the eigenproblems of some structured matrices such as banded matrices, Toeplitz matrices, and certain discredited problems. However, it is not performing well on the eigenproblem of complex symmetric matrices. Recall that in Chapter 4, we have

$$USU^* = \frac{1}{2}\begin{pmatrix} s_{11}+s_{33} & \sqrt{2}(s_{12}+is_{23}) & s_{11}+2is_{13}-s_{33} \\ \sqrt{2}(s_{12}-is_{23}) & 2s_{22} & \sqrt{2}(s_{12}+is_{23}) \\ s_{11}-2is_{13}-s_{33} & \sqrt{2}(s_{12}-is_{23}) & s_{11}+s_{33} \end{pmatrix}$$

when $n = 3$, where U is the unitary matrix given in Theorem 4.1.3. We use the unitary matrix to turn a symmetric matrix to an almost Toeplitz except the main diagonal. In general, applying unitary matrices in Theorem 4.1.3 and Theorem 4.1.1 may not give us immediate good results. However, it may be helpful to develop a better algorithm by applying the unitary matrices in Theorem 4.1.3 and Theorem 4.1.1 or some others, on complex symmetric matrices or other matrices, to turn the matrices to suitable forms for some superfast algorithms. By transforming a matrix to some structured matrices, we may be able to apply some existing powerful algorithms to solve the eigenproblem faster.

Bibliography

[1] G.S. Ammar, D. Calvetti and L. Reichel, Computation of Gauss-Kronrod quadrature rules with non-positive weights, Electr. Trans. Numer. Anal., 9 (1999), 26–38.

[2] G.S. Ammar and W. B. Gragg, Superfast solution of real positive definite Toeplitz systems, SIAM J. Matrix Anal. Appl., 9 (1988), 61–76.

[3] L. Autonne, Sur les matrices hypohermitiennes et sur les matrices unitaires, Ann. Univ. Lyon, Nouvelle Série I, Fasc., 38 (1915), 1–77.

[4] Y.H. Au-Yeung and N.K. Tsing, Some theorems on the numerical range, Linear and Multilinear Algebra, 15 (1984), 215–218.

[5] I. Bar-On, M. Paprzycki, High performance solution of the complex symmetric eigenproblem, Numerical Algorithms, 18 (1998), 195–208.

[6] I. Bar-On and V. Ryaboy, Fast diagonalization of large and dense complex symmetric matrices, with applications to quantum reaction dynamics, SIAM J. Sci. Compu., 18 (1997), 1412–1435.

[7] N. Bebiano and I.M. Sitkovsky, Numerical ranges of Toeplitz operators with matrix symbols, Linear Algebra Appl., 436 (2012), 1721–1726.

[8] R. B. Bitmead and B. D. O. Anderson, Asymptotically fast solution of Toeplitz and related systems of linear equations, Linear Algebra Appl., 34 (1980), 103–116.

[9] A. Böttcher and B. Silbermann, Introduction to Large Truncated Toeplitz Matrices, Springer, New York, 1999.

[10] L. Brickman, On the field of values of a matrix, Proc. Amer. Math. Soc., 12 (1961), 61–66.

[11] S. Chandrasekaran, M. Gu, X. Sun, J. Xia, and J. Zhu, A superfast algorithm for Toeplitz systems of linear equations, SIAM J. Matrix Anal. Appl., 29 (2007), 1247–1266.

[12] M.T. Chien, J. Liu, H. Nakazato and T.Y. Tam, Toeplitz matrices are unitarily similar to symmetric matrices, Linear and Multilinear Algebra, 14 pages URL: http://dx.doi.org/10.1080/03081087.2017.1330865.

[13] M.T. Chien and H. Nakazato, Matrices unitarily similar to symmetric matrices, Proceedings of the 20th International Conference on Applied Mathematics, Budapest, Hungary (2015), 86–91.

[14] M.T. Chien and H. Nakazato, Computing the determinantal representations of hyperbolic forms, Czech. Math. J., 66 (2016), 633–651.

[15] M.D. Choi, D. W. Kribs and K. Zyczkowski, Higher-rank numerical ranges and compression problems, Linear Algebra Appl., 418 (2006), 828–839.

[16] B. Dickinson, M. Morf and T. Kailath, Canonical matrix fraction and state space descriptions for deterministic and stochastic linear systems, IEEE Trans. On Auto. Control, 19 (1974), 656–667.

[17] M. Fiedler, Geometry of the numerical range of matrices, Linear Algebra Appl., 37 (1981), 81–96.

[18] B. Friedlander, M. Morf, T. Kailath and L. Ljung, Levinson-and-Chandrasekhar-type equations for a general discrete-time linear estimation problem, Proc. IEEE Conference on Decision and Control, Florida (1976), 910–915.

[19] S.R. Garcia and M. Putinar, Complex symmetric operators and applications, Trans. Amer. Math. Soc., 358 (2006), 1285–1315.

[20] S.R. Garcia, D. E. Poore, M.K. Wyse, Unitary equivalence to a complex symmetric matrix: a modulus criterion, Oper. Matrices, 5 (2011), 273–287.

[21] H.L. Gau and P. Y. Wu, Higher-rank numerical ranges and Kippenhahn polynomials, Linear Algebra Appl., 438 (2013), 3054–3061.

[22] I. Gohberg, T. Kailath and V. Olshevsky, Fast Gaussian elimination with partial pivoting for matrices with displacement structure, Math. Comp., 64 (1995), 1557–1576.

[23] G.H. Goluba, H. A. van der Vorst, Eigenvalue computation in the 20th century, J. Comput. Appl. Math., 123 (2000), 35–65.

[24] R.M. Gray, Toeplitz and Circulant matrices: a review, Found. Trends Commun. Inf. Theory, 2 (2006), 155–239.

[25] U. Grenander and G. Szegö, Toeplitz Forms and Their Applications, University of Calif. Press, Berkeley and Los Angeles, 1958.

[26] M. Gu, Stable and efficient algorithms for structured systems of linear equations, SIAM J. Matrix Anal. Appl., 19 (1998), 279–306.

[27] K.E. Gustafson and D.K.M. Rao, Numerical Range: The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.

[28] R.E. Hartwig and M.E. Fisher, Asymptotic behavior of Toeplitz matrices and determinants, Arch. Rational Mech. Anal., 32 (1969), 190–225.

[29] G. Heinig, Inversion of generalized Cauchy matrices and other classes of structured matrices, Linear Algebra for Signal Processing, IMA Vol. Math. Appl., 69 (1994), 95–114.

[30] J.W. Helton and V. Vinnikov, Linear matrix inequality representations of sets, Comm. Pure Appl. Math., 60 (2007), 654–674.

[31] J.W. Helton and I.M. Spitkovsky, The possible shape of the numerical ranges, Oper. Matrices, 6 (2012), 607–611.

[32] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 2013.

[33] R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[34] W.D. Hoskins, P.J. Ponzo, Some properties of a class of band matrices, Math. Comp., 26 (1972), 393–400.

[35] T. Kailath, A view of three decades of linear filtering theory, IEEE Trans. Inform. Theory, 20 (1974), 145–181.

[36] T. Kailath, Some new results and insights in linear least-squares estimation theory, In Proc. First Joint IEEE-USSR Workshop on Info. Theory, Moscow (1975), 97–104.

[37] R. Kippenhahn, Über den wertevorrat einer Matrix, Math. Nachr., 6 (1951), 193–228.

[38] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat'l Bur. Std., 45 (1950), 255–282.

[39] P.D. Lax, Differential equations, difference equations and matrix theory, Comm. Pure Appl. Math., 6 (1958), 175–194.

[40] X. Liu, B.M. Nguelifack and T.Y. Tam, Unitary similarity to a complex symmetric matrix and its extension to orthogonal symmetric Lie algebras, Linear Algebra Appl., 438 (2013), 3789–3796.

[41] C.K. Li and T.Y. Tam, Numerical ranges arising from simple Lie algebras, J. Can. Math. Soc., 52 (2000), 141–171.

[42] F.T. Luk and S. Qiao, Decaying exponentials and signal decomposition, Mathematics in Signal Processing, Oxford University Press, 1998.

[43] F.T. Luk and S. Qiao, A fast eigenvalue algorithm for Hankel matrices, Linear Algebra Appl., 316 (2000), 171–182.

[44] A. McIntosh, The Toeplitz-Hausdorff theorem and ellipticity conditions, Amer. Math. Monthly, 85 (1978), 475–477.

[45] R.P. Mentz, On the inverse of some covariance matrices of Toeplitz type, SIAM J. Appl. Math., 31 (1976), 426–437.

[46] B. N. Parlett, The QR algorithm, Computing Sci. Eng., 2 (2000), 38–42.

[47] C. Pearcy, A complete set of unitary invariants for 3×3 complex matrices, Trans. Amer. Math. Soc., 104 (1962), 425–429.

[48] C. Pearcy, A complete set of unitary invariants for operators generating finite W^*-algebras of type I, Pacific J. Math., 12 (1962), 1405–1414.

[49] L. Rehnqvist, Inversion of certain symmetric band matrices, BIT, 12 (1972), 90–98.

[50] K. S. Sibirskii, A minimum polynomial basis of unitary invariants of a square matrix of order 3, Mat. Zametki, 3 (1968), 291–295.

[51] E. H. Spanier, Algebraic Topology, McGraw-Hill, 1966.

[52] W. Specht, Zur theorie der matrizen II, Jahresber. D. M. V., 50 (1940), 19–23.

[53] M. Stewart, A superfast Toeplitz solver with improved numerical stability, SIAM J. Matrix Anal. Appl., 25 (2003), 669–693.

[54] T. Takagi, On an algebraic problem related to an analytic theorem of Caratheodory and Fejer and on an allied theorem of Landau, Japan. J. Math., 1 (1925), 83–93.

[55] J.E. Tener, Unitary equivalence to a complex symmetric matrix: an algorithm, J. Math. Anal. Appl., 341 (2008), 640–648.

[56] Y. Teranishi, The ring of invariants of matrices, Nagoya Math. J., 104 (1986), 149–161.

[57] W.F. Trench, Inversion of Toeplitz band matrices, Math. of Comp., 28 (1974), 1089–1095.

[58] W.F. Trench, An algorithm for the inversion of finite Toeplitz matrices, J. Soc. Indust. Appl. Math., 12 (1964), 515–522.

[59] M. Van Barel, G. Heinig, and P. Kravanja, A stabilized superfast solver for nonsymmetric Toeplitz systems, SIAM J. Matrix Anal. Appl., 23 (2001), 494–510.

[60] J. Vermeer, Orthogonal similarity of a real matrix and its transpose, Linear Algebra Appl., 428 (2008), 382–392.

[61] H. Widom, Toeplitz matrices, I.I. Hirschman Jr. (Ed.), Studies in Real and Complex Analysis, Math. Assoc. Amer (1965), 179–209.

[62] J. Vogel, J. Xia, S. Cauley and V. Balakrishnan, Superfast divide-and-conquer method and perturbation analysis for structured eigenvalue solutions, SIAM J. Sci. Comput., 38 (2016), A1358–A1382.

[63] K. Ye, L.-H. Lim, Every matrix is a product of Toeplitz matrices, Found. Comput. Math., 16 (2016), 577–598.