Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Is every matrix similar to a Toeplitz matrix?

Autora: Maria Ferrer Pascual

Directora: Dra. Maria Eulàlia Montoro Lopez
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, January 21, 2022

Abstract

The theory of matrices is a very important field in the world of mathematics, which plays a central role in the study of a large variety of areas in pure and applied mathematics. Is is one of the first topics studied in the degree. I remember when I was studying the first course that I got fascinated by the fact that you can transform a matrix into an easier one through a similarity and these two matrices would preserve basic properties that give a big amount of information of the matrix.

In Algebra courses, we have learned all the geometrical concepts that a matrix can represent and are endless. Besides, there are structured matrices, such as Toeplitz matrices which are useful in many branchs. Their study has been an active field of research since the beginning of the last century, and it remains until today. Many journal papers have been devoted to these matrices and the large interest has a reason: the number of applications that arise from it. They contribute to the discretization of differential and integral equations, in the theory of orthogonal polynomials, trigonometric moments, time series analysis, graph theory and signal processing applications, as it is studied at [2], [9], [14] and [18].

The aim of this work is to bring together these two areas. We are interested in finding which Jordan canonical forms can be realized by some Toeplitz matrix. In other words, if every matrix is similar to a Toeplitz matrix. To do so, we have deeply studied the two only articles published related to our topic ([4] and [5]). Before starting with this research, we will present all concepts that will be needed related to Linear Algebra and Toeplitz matrices. Once we have set the knowledge base required, we will focus on our final purpose.

Unless it is otherwise stated, we suppose we are in the field of complex numbers. With a view to answer our question, we have divided the study in different parts, mainly in two blocks. The first block is dedicated to proving those matrices which are similar to a Toeplitz matrix, which are diagonalizable matrices, nonderogatory matrices and the general case for $n \leq 4$. Regarding the second part, is devoted to matrices of general dimension and we will need to do some strong assumptions like having just one eigenvalue. In this case, we distinguish between matrices of odd and even order. All these investigations will lead us to answer this question.

2010 Mathematics Subject Classification. 15A09, 15A20, 15A29, 15B05

Acknowledgements

I would like to give special thanks to my supervisor Dr. Eulàlia Montoro for her support and encouragement. Without her help, this project could not have been possible. Eulàlia, thanks for all the meetings, for always working side by side with me and all your helpful suggestions. I am greatly indebted to you for your active interest in this paper, for proposing this topic and the countless hours you committed.

This work represents the end of a cycle and, at the same time, it is the beginning of the next one. As such, I am eternally grateful to all people that have been by my side every step of my journey. Specially, I want to thank my parents, Ramon and Meritxell, for all their unconditional support and for always believing in me. Mum and dad, thanks for walking always by my side and helping me jump every stone in the path. I could never thank you enough for all you have done for me. 'All for one and one for all'.

Also, I want to do a special mention to the light of my life, yaeta, thank you for being my grandmother and my best friend since I was born. I am so grateful for all what you have taught me. All I am today is thanks to you. RIP.

Finally, I want to mention my friends from the three cities my heart belongs to: Barcelona, Valencia and Gandesa. Love you all.

Contents

Chapter 1

Introduction

Toeplitz matrices are characterized by the property of following a concrete structure, depending uniquely on the indices of a given sequence numbers. Later we will define it rigorously. Before starting with the mathematical study, we will contextualize these structured matrices.

History and contextualization

Toeplitz matrices are named after Otto Toeplitz (01/08/1981 in Breslau, Germany - 15/02/1940 in Jerusalem). Toeplitz had a strong mathematical influence from his father and his grandfather, Emili Toeplitz and Julius Toeplitz, that were teachers in mathematics and published several mathematical papers. He followed in the footsteps of them and studied mathematics at University of Breslau, where he specialized in the field of algebraic geometry. He received his P.h.D. in 1905.

Figure 1.0.1: Otto Toeplitz

The next year, he moved to Göttingen. At that time, Hilbert was completing his theory of integral equations and Toeplitz was greatly influenced by him and started to study on known classical theories of forms on n-dimensional (finite) spaces for infinite dimensional spaces. He collaborated with the circle of Hilbert's brilliant students. Moreover, he published papers related to Hilbert's spectral theory and he discovered basic ideas of what now are called Toeplitz operators.

With the latter, in 1913, he became a professor at University of Kiel. Years after he did an extensive project of fundamental articles on integral equations. In 1928, he was promoted and he was offered a chain at University of Bonn, he accepted it but five years later, when Hitler raised power in 1933, Toeplitz was dismissed from the office by the National Socialist regime. Until 1938, he involved in political activism against Hitler's Jewish communities. Furthermore, Toeplitz founded a private school in Bonn for Jewish children excluded from the German educational System.

After the famous night of Broken Glass (November 9-10, 1938), he was settled in Palestine and he participated in the renovation of the University of Jerusalem, where he started to work as an administrative adviser while he did private seminars and, two years later, he passed away because of a tuberculosis.

Despite all the complicated political situation in Germany, he never stopped studying mathematics. It is considered to be one of the creators of the spectral theory of linear operators. His chief interest was the theory of infinite linear, bilinear and quadratic forms. He studied the matrices related to these forms in a functional analysis frame. Around 1930, his mathematical research was based on a more general point of view and he developed a general theory of infinite dimensional spaces and he criticized Banach's work for being too abstract.

The work

Once having contextualized our topic, we will focus on the technical part of our work. As we have noticed before, this paper studies if every matrix is similar to a Toeplitz matrix.

Before presenting the body of our work, we must point out that the *inverse eigenvalue problem* concerns the construction of a matrix given its spectrum (the reverse problem is what we will study in this paper). In the case of Toeplitz matrices, this problem is always solvable. That means for a given system of n complex numbers λ_k, there is always a Toeplitz matrix $T \in M_n(\mathbb{C})$ such that the λ_k are eigenvalues of T, counting multiplicities. In the case that λ_k and $dim(Ker(A - \lambda_k I_n)$ are known, then we have the common inverse eigenvalue problem (see [2] and [17]).

In the study of finding the possible Jordan forms that can be realized by some Toeplitz matrix, the first matrices we focus on are diagonalizable matrices (*Chapter 3*), which we prove that are similar to a Toeplitz by taking into account a related Toeplitz form: circulant matrices. Then, we study another type of matrices that are called nonderogatory (*Chapter 4*). In this case, Hessenberg ones are crucial: we first check that a nonderogatory matrix is similar to an upper Hessenberg matrix and, secondly, we see that the latter matrices are similar to a unit upper Hessenberg Toeplitz matrix. Hence, because of its uniqueness, we can talk about the Toeplitz canonical form in the nonderogatory case. Afterwards, we ask ourselves what would happen if we extended this case to a general field $\mathbb{F}$ and we show that the proof constructed before can be adapted to fields of characteristic zero or greater than n (where n is the size of the matrix).

The next chapters discuss the problem if the hypothesis of nondiagonalizable and nonderogatorcy are dropped. We start by studying the general case of a complex matrix of dimension $n \leq 4$ (*Chapter 5*) and, through a constructive proof, we arrive to the conclusion that are similar to a Toeplitz matrix. However, we will prove that the statement is not true if we consider the matrix in the field $\mathbb{R}$.

Subsequently, we study the general case for $n > 5$ (*Chapter 6*). We cannot extend the above-mentioned proof for $n \leq 4$, as checking all the possible Jordan forms and finding the corresponding similarities is lengthy and rough so it makes no sense studying higher dimensions with this strategy. In such case, we reduce ourselves to the case of having just one eigenvalue, concretely $\lambda = 0$, as it plays an are instrumental role in the structure of the kernel of a Toeplitz matrix. Is in this chapter where we find that *not* every matrix is similar to a Toeplitz by finding concrete counterexamples for matrices of odd and even order. In this part we will define a general tuple that defines the possible Jordan forms of a matrix following a concrete structure and check whether if it is Toeplitz admissible or not, just about checking if these Jordan forms can be similar to a Toeplitz matrix. The procedure for the even order is a bit more complicated and we will consider three sub-cases: $n = 4$, $n = 6$ and $n \geq 8$. Distinctly, for $n \geq 8$, we will compact this last problem to find common zeroes of a class of polynomials defined recursively and we will also find counterexamples.

Chapter 2

Preliminaries

The essential prerequisites of this work are a knowledge of matrix theory. However, before starting we need to recall some concepts of previous courses of the degree in *Mathematics* at *University of Barcelona*, because of that, we will skip some of the proofs. In addition, we will introduce some new related concepts that are needed to go deep on our study.

2.1 Notation

First, let us set up and discuss some notations and terminology that will be used along the paper.

- Unless otherwise stated we assume that all the matrices are defined in the field $\mathbb{F} = \mathbb{C}$.
- $M_{n,m}(\mathbb{C})$ stands for $n \times m$ complex matrices, if $n = m$ we simply write $M_n(\mathbb{C})$.
- $A \in M_{n,m}(\mathbb{C})$ with elements a_{ij} is denoted by $A = (a_{ij})_{1 \le i \le n, 1 \le j \le m}$. If $n = m$, its *determinant* is expressed by $det(A)$ or $|A|$ and its *trace* by $tr(A)$. Besides, $vec(A) \in C^{mn}$ denotes the *vectorization* of a matrix, that is obtained by stacking the columns of the matrix A on top of one another.
- We mean by A^t the *transposed* matrix, by $\bar{A}$ the *conjugate* matrix and by A^* the *conjugate transpose* matrix. It is said that $A \in M_n(\mathbb{C})$ is *Hermitian* if and only if $A^* = A$ and it is *Normal* if and only if $A^* A = AA^*$.
- We denote by I_n the *identity matrix.*
- The *permutation matrix* is obtained when we permute the standard basis via $\sigma(n)$, is denoted by $P_\sigma \in M_n(\mathbb{C})$. Particularly, if $\sigma = (1, 2, \ldots, n)$, $Z \in M_n(\mathbb{C})$ denotes the following permutation matrix

$$Z = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \in M_n(\mathbb{C})$$

- The *exchange* or *counteridentity* matrix is a matrix with 1's in the anti-diagonal and it is a special case of a permutation. Therefore we have, $\sigma = \{(1,n)(2,n-1)\dots\}$ and
$$J = \begin{pmatrix} 0 & \cdots & \cdots & 0 & 1 \\ \vdots & & \iddots & \iddots & 0 \\ \vdots & \iddots & \iddots & & \vdots \\ 0 & \iddots & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \in M_n(\mathbb{C})$$

- Let E and F be vector subspaces of finite dimension over $\mathbb{C}$ and $f : E \to F$ be a linear map. We denote the *kernel* of f by
$$Ker(f) = \{x \in E \mid f(x) = 0\} \subset E$$
and the *image* of f by
$$Im(f) = \{y \in F \mid \exists x \in E, y = f(x)\} \subset F$$
Note that every linear application is identified with a matrix ($f \leftrightarrow A$). The *rank* of f is the dimension of the image subspace, that is
$$Rank(f) = dim(Im(f)) = Rank(A)$$

- The *characteristic polynomial* of $A \in M_n(\mathbb{C})$ is denoted by $p_A(x) = det(A - \lambda I)$ and the *minimal polynomial* by $\mu_A(x)$.

- Regarding the study of eigenvalues and eigenvectors, we denote by $Spec(A)$ the *spectrum* of A and its *cardinality* is denoted by $|Spec(A)|$.

- The relation of *similarity* between two given matrices $A, B \in M_n(\mathbb{C})$ is denoted by $\sim$, this relation is given when there exists an invertible matrix $S \in M_n(\mathbb{C})$ such that $B = S^{-1}AS$.

- $\lfloor x \rfloor$ denotes the *integer part* of x.

2.2 Sylvester equations

We now present some concepts related to matrix analysis that will be useful for our work.

Definition 2.2.1. A *Sylvester equation* is a matrix equation of the form
$$AX + XB = C \tag{2.2.1}$$
where $A \in M_m(\mathbb{C})$, $B \in M_n(\mathbb{C})$ and $X, C \in M_{n,m}(\mathbb{C})$.

Definition 2.2.2. Given $A \in M_m(\mathbb{C})$ and $B \in M_n(\mathbb{C})$, $Syl(A, B)$ stands for the *linear Sylvester operator* defined by
$$Syl(A,B) : M_m(\mathbb{C}) \to M_n(\mathbb{C})$$
$$X \mapsto AX - XB$$

Definition 2.2.3. Let $A = (a_{ij})_{1\le i\le m, 1\le j\le n} \in M_{m,n}(\mathbb{C})$ and $B = (b_{ij})_{1\le i\le p, 1\le j\le q} \in M_{p,q}(\mathbb{C})$. The *Kronecker product* of A and B, $A \otimes B$, is defined as

$$A \otimes B \equiv \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} \in M_{mp,nq}(\mathbb{C}) \tag{2.2.2}$$

Observation 2.2.4. The equation $AX + XB = C$ can be easily rewritten via Kronecker product as

$$[(I_n \otimes A) + (B^t \otimes I_m)]\, vec(X) \ = \ vec(C) \tag{2.2.3}$$

The point of the two following theorems is that it allows one to prove Lemma 2.2.7 that will be useful for our object of study.

Theorem 2.2.5. ([13]) Let $A \in M_m(\mathbb{C})$, $B \in M_n(\mathbb{C})$. If $\lambda \in Spec(A)$ and $x \in \mathbb{C}^m$ is a corresponding eigenvector of A and if $\mu \in Spec(B)$ and $y \in \mathbb{C}^n$ is a corresponding eigenvector of B, then $\lambda + \mu$ is an eigenvalue of the Kronecker sum $[(I_n \otimes A) + (B^t \otimes I_m)]$ and $y \otimes x \in \mathbb{C}^{nm}$ is a corresponding eigenvector.

Theorem 2.2.6. ([13]) Let $A \in M_m(\mathbb{C})$ and $B \in M_n(\mathbb{C})$. The equation $AX+XB = C$ has a unique solution $X \in M_{n,m}(\mathbb{C})$ for each $C \in M_{n,m}(\mathbb{C})$ if and only if $Spec(A) \cap Spec(-B) = \emptyset$.

Proof. Applying Observation 2.2.4, we obtain an equivalent equation

$$[(I_n \otimes A) + (B^t \otimes I_m)]vec(X) = vec(C)$$

that it has a unique solution if and only if the rank of $(I_n \otimes A) + (B^t \otimes I_m)$ is maximum and this happens if and only if $0 \notin Spec((I_n \otimes A) + (B^t \otimes I_m))$. Applying Theorem 2.2.5 we have that $\lambda + \mu \neq 0$ with $\lambda \in Spec(A)$ and $\mu \in Spec(B)$, and equivalently, $Spec(A) \cap Spec(-B) = \emptyset$. □

Lemma 2.2.7. ([4]) Let $A \in M_m(\mathbb{C})$, $B \in M_n(\mathbb{C})$ be such that $Spec(A) \cap Spec(B) = \emptyset$. Then for any $C \in M_{n,m}(\mathbb{C})$, we have

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \sim \begin{pmatrix} A & 0 \\ C & B \end{pmatrix} \tag{2.2.4}$$

Proof. The above similarity is obtained by

$$\begin{pmatrix} Id & 0 \\ X & Id \end{pmatrix}^{-1} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} Id & 0 \\ X & Id \end{pmatrix} = \begin{pmatrix} Id & 0 \\ -X & Id \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} Id & 0 \\ X & Id \end{pmatrix} = \begin{pmatrix} A & 0 \\ X(-A) + XB & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$$

And regarding, Theorem 2.2.5 and 2.2.6, X does exist and it is unique. □

2.3 Toeplitz matrices

In this section we give a brief exposition of Toeplitz matrices. We present the definition, some basic properties and examples. We will touch only a few aspects of the theory as it is not our goal of study. For a fuller treatment of the topic, we refer the reader to [3], [9] and [18].

Toeplitz matrices are studied in many fields of mathematics. In pure math: algebra, analysis, combinatorics, graph theory, differential and integral equations, topology, as well as in applied mathematics: numerical integration, image processing, time series analysis, mechanics and among other areas. Again, it is not the aim of this work to do an extensive study of these applications: for a deeper discussion, we refer the reader to [6], [15] and [18].

Definition 2.3.1. A *Toeplitz matrix* $T \in M_n(\mathbb{C})$ (T-matrix, for short) is a matrix which each descending diagonal from left to right is constant. In other words, is a matrix characterized by having constant diagonal entries, that is, a matrix of the form

$$T = \begin{pmatrix} t_0 & t_1 & t_2 & \cdots & \cdots & t_{(n-1)} \\ t_{-1} & t_0 & t_1 & t_2 & \ddots & \vdots \\ t_{-2} & t_{-1} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & t_1 & t_2 \\ \vdots & & \ddots & t_{-1} & t_0 & t_1 \\ t_{-(n-1)} & \cdots & \cdots & t_{-2} & t_{-1} & t_0 \end{pmatrix} \in M_n(\mathbb{C}) \tag{2.3.1}$$

A Toeplitz matrix does not have necessarily to be square, but we will only be considering square matrices in the major part of our work.

2.3.1 Basic properties

T-matrices have a long history and have given rise to important recent applications. To date, many results related to the algebra of Toeplitz matrices and the corresponding applications has been accumulated in journal literature and these results combined already form a well-structured theory. This work is devoted to construct the necessary theory to check if every matrix is similar to a Toeplitz matrix, as we have already mentioned in the preceding pages. Accordingly, we will just enumerate some primary results of Toeplitz matrices to have an elementary knowledge of these matrices that will let us go in dept with our main subject.

1. Regarding the structure of a Toeplitz matrix, it is closed under addition and scalar multiplication. Furthermore, the product of two T-matrices, in general, is not a Toeplitz matrix and the same happens with the computation of its inverse.

2. Given a sequence $t_{-(n+1)}, \ldots, t_0, \ldots, t_{n+1} \in \mathbb{C}$, we write the general term of a Toeplitz matrix as $t_{ij} = t_{j-i}$.

3. Toeplitz matrices $T \in M_{n+2}(\mathbb{C})$ can be represented by a $2n+3$ vector

$$v = (t_{-(n+1)}, \ldots, t_0, \ldots, t_{n+1}) \in \mathbb{C}^{2n+3}$$

In the case of symmetric Toeplitz matrices, they are just determined by the first column or row. Hence, these matrices are represented by

$$v = (t_{-(n+1)}, \dots, t_0) \in \mathbb{C}^{n+2} \text{ or } v = (t_0, \dots, t_{n+1}) \in \mathbb{C}^{n+2}$$

4. The following matrices are Toeplitz

$$B = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in M_{n+1}(\mathbb{C}) \text{ and } F = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \in M_{n+1}(\mathbb{C}) \qquad (2.3.2)$$

   and are called *backward shift* and *forward shift*, respectively, because of their effect on the elements of the standard basis $\{e_1, \dots, e_{n+1}\}$. Note that $B^t = F$ and $F^t = B$.

5. Concerning the anterior point, we have a characterization of T-matrices:

$$T = \sum_{k=1}^{n} t_{-k} F^k + \sum_{k=0}^{n} t_k B^k \in M_{n+1}(\mathbb{C}) \quad \Leftrightarrow \quad T \in M_{n+1}(\mathbb{C}) \text{ is a Toeplitz matrix}$$

6. In a Toeplitz linear system, $Tx = b$ with $T \in M_n(\mathbb{C})$ and $x, b \in \mathbb{C}^n$, we have $2n - 1$ degrees of freedom rather than n^2.

7. Toeplitz matrices are *persymmetric*, in other words, a square matrix which is symmetric with respect to the northeast-to-southwest diagonal. Thus,

$$TJ = JT^t \quad \forall T, J \in M_n(\mathbb{C})$$

   where J is the counteridentity matrix.
   In particular, symmetric Toeplitz matrices are also *centrosymmetric* (symmetric respect to its center), so follow the equality

$$TJ = JT \quad \forall T, J \in M_n(\mathbb{C})$$

   and *bisymmetric* (symmetric with respect to the two main diagonals). Hence,

$$TJ = JT \text{ and } T = T^t \quad \forall T, J \in M_n(\mathbb{C})$$

8. Toeplitz matrices are nearby linked to Fourier series, because the multiplication operator by a trigonometric polynomial compressed to a finite-dimensional space, can be represented by such a matrix. Similarly, one can represent linear convolution as multiplication by a Toeplitz matrix.

2.3.2 Related Toeplitz forms

In this part we will introduce three matrices that are closely related to Toeplitz matrices.

Circulant matrices

Circulant matrices have a lot of interest when studying some properties related to Toeplitz matrices. Furthermore, they will be a key concept when we prove that all diagonalizable matrices are similar to T-matrices (see *Chapter 3*).

Definition 2.3.2. ([10]) A *circulant matrix* $C \in M_n(\mathbb{C})$ is a matrix with the following form

$$C = \begin{pmatrix} c_1 & c_2 & \cdots & c_{n-1} & c_n \\ c_n & c_1 & \ddots & & c_{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_3 & & \ddots & \ddots & c_2 \\ c_2 & c_3 & \cdots & c_n & c_1 \end{pmatrix} \tag{2.3.3}$$

It is easily seen that $C = (c_{jk}) = c_{(k-j+1)}\ (mod\, n)$. We emphasize that the matrix is determined by the first row as each row k with $k = 1, \dots, n$ is identical to the first row ($k = 1$) but moved $k - 1$ positions to the right. That is, in the rows and columns we have cyclic permutations of the elements. Moreover, circulant matrices form a linear subspace of the set of all matrices of dimension n.

Hankel matrices

The matrix now defined is not a Toeplitz matrix despite it is closely related and they always go together.

Definition 2.3.3. ([6]) A *Hankel matrix* $H \in M_n(\mathbb{C})$ has constant skew diagonal, that is, it is constant across the anti-diagonals. One can write more explicitly:

$$H = \begin{pmatrix} h_1 & h_2 & h_3 & \cdots & h_{n-1} & h_n \\ h_2 & h_3 & \cdots & \cdots & h_n & h_{n+1} \\ h_3 & \iddots & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots \\ h_{n-1} & \iddots & \iddots & \iddots & \iddots & \vdots \\ h_n & h_{n+1} & \cdots & \cdots & h_{2n-2} & h_{2n-1} \end{pmatrix} \in M_n(\mathbb{C})$$

It can be easily see that any Hankel matrix is symmetric.

Observation 2.3.4. Let $T \in M_n(\mathbb{C})$ be a Toeplitz matrix and $H \in M_n(\mathbb{C})$ be a Hankel matrix. Then $T = J\,H$, where J is the counteridentity matrix. In other words, Toeplitz matrices turn into Hankel matrices by reversing the order of the rows.

Toeplitz-plus-Hankel matrices

Definition 2.3.5. ([6]) A *Toeplitz-plus-Hankel matrix* is a matrix that the entries are constant along each diagonal plus constant along each anti-diagonal. That is to say that it is Toeplitz and Hankel at the same time.

Example 2.3.6. As the notation of the general matrix is harsh, we will provide an example of dimension 5 to visualize the form of a Toeplitz-plus-Hankel matrix:

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 2 & 1 \\ 2 & 1 & 2 & 1 & 2 \\ 1 & 2 & 1 & 2 & 1 \end{pmatrix} \in M_5(\mathbb{C})$$

Chapter 3

Diagonalizable case

We begin by considering those matrices that are diagonalizable. The following lemma provides a natural and intrinsic characterization of circulant matrices related with the permutation cyclic matrix, which will be essential the proof of Theorem 3.2, the crucial part of this chapter, that proves that every diagonalizable matrix is similar to a Toeplitz.

Lemma 3.1. ([16]) Let $A = (a_{ij})_{1\leq i,j\leq n} \in M_n(\mathbb{C})$ and $Z \in M_n(\mathbb{C})$ the cyclic permutation matrix (see Notation 2.1). Then

$$A \text{ is circulant} \quad \Leftrightarrow \quad AZ = ZA$$

That is, A is invariant under the similarity $A \mapsto Z^{-1}AZ$.

Proof. We suppose $AZ = ZA$, then

$$AZ = \begin{pmatrix} a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ a_{41} & a_{42} & a_{43} & \cdots & a_{4n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \end{pmatrix} = ZA = \begin{pmatrix} a_{1n} & a_{11} & a_{12} & \cdots & a_{1(n-1)} \\ a_{2n} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{3n} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{nn} & a_{n2} & a_{n3} & \cdots & a_{n(n-1)} \end{pmatrix}$$

and by equaling term-wise A must be circulant.
Conversely, if A is a circulant matrix

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

by direct computation $AZ = ZA$ is obtained. □

Theorem 3.2. ([4]) Every diagonalizable matrix is similar to a Toeplitz matrix.

Proof. Let $D = diag(d_0, \ldots, d_{n-1}) \in M_n(\mathbb{C})$ and Z the cyclic permutation matrix

$$Z = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 1 & & & & 0 \end{pmatrix} \in M_n(\mathbb{C}) \tag{3.0.1}$$

We have that $Spec(Z) = \{1, \xi, \xi^2, \dots, \xi^{n-1}\}$, where ξ is a primitive nth root of the unity. Therefore, $Z \sim D' = \text{diag}(1, \xi, \xi^2, \dots, \xi^{n-1})$ and there exists an invertible matrix $Q \in M_n(\mathbb{C})$ such that $Z = Q^{-1}D'Q$.

The task is now to prove that $D \sim p(Z)$ for some $p(x) \in \mathbb{C}[x]$. Let $p(x) = a_0 + \dots + a_n x^n \in \mathbb{C}[x]$, then

$$p(Z) = p(QD'Q^{-1}) = \sum_{i=1}^{n-1} a_i (QD'Q^{-1})^i = \sum_{i=1}^{n-1} Q a_i (D')^i Q^{-1} = Qp(D')Q^{-1} \tag{3.0.2}$$

If we take $p(x)$ as the polynomial obtained by the Lagrange interpolation such that $p(\xi^i) = d_i$, $0 \le i \le n-1$, we obtain that

$$\begin{pmatrix} p(1) & & \\ & \ddots & \\ & & p(\xi^{n-1}) \end{pmatrix} = \begin{pmatrix} d_0 & & \\ & \ddots & \\ & & d_{n-1} \end{pmatrix} \tag{3.0.3}$$

that is, $p(Z) \sim D$. Furthermore, $p(Z)$ is a Toeplitz matrix because $p(Z)$ commutes with Z (see Lemma 3.1). □

In light of having a more intuitive view of the above proof, we will do the construction of it for $n = 2$ and $n = 3$ step by step.

Example 3.3.

- <u>Case $n = 2$</u>:
  Let $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ where $\lambda_1 \neq \lambda_2 \in \mathbb{C}$ (if $\lambda_1 = \lambda_2$, D is a Toeplitz matrix).
  In this case, $Z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $Spec(Z) = \{1, -1\}$. The matrix Q given in the proof is: $Q = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$.
  We define $p(x) \in \mathbb{C}_2[x]$ such as $p(1) = \lambda_1$ and $p(-1) = \lambda_2$. According to the proof,

  $$p(Z) = p\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} p\left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right) \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix} =$$

  $$= \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} \frac{\lambda_1+\lambda_2}{2} & \frac{\lambda_1-\lambda_2}{2} \\ \frac{\lambda_1-\lambda_2}{2} & \frac{\lambda_1+\lambda_2}{2} \end{pmatrix}$$

  where the last matrix is a Toeplitz matrix similar to D.

- <u>Case $n = 3$</u>:
  Let $D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$ where λ_1, λ_2, $\lambda_3 \in \mathbb{C}$. We need to consider three cases:

  - If $\lambda_1 = \lambda_2 = \lambda_3$, it is immediate that we have that D is a Toeplitz matrix.

  - $\lambda_1 = \lambda_2 \neq \lambda_3$, $Spec(Z) = \{\, 1, \frac{-\sqrt{3}i-1}{2}, \frac{\sqrt{3}i-1}{2} \,\}$. In this case

$$Z=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}=\begin{pmatrix}1&\frac{-\sqrt{3}i-1}{2}&\frac{\sqrt{3}i-1}{2}\\1&\frac{\sqrt{3}i-1}{2}&\frac{-\sqrt{3}i-1}{2}\\1&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&\frac{-\sqrt{3}i-1}{2}&0\\0&0&\frac{\sqrt{3}i-1}{2}\end{pmatrix}\begin{pmatrix}1/3&1/3&1/3\\\frac{\sqrt{3}i-1}{6}&\frac{-\sqrt{3}i-1}{6}&1/3\\\frac{-\sqrt{3}i-1}{6}&\frac{\sqrt{3}i-1}{6}&1/3\end{pmatrix}$$

We define $p(x)\in\mathbb{C}_3[x]$ such as $p(1)=\lambda_1$, $p(\frac{-\sqrt{3}i-1}{2})=\lambda_2$ and $p(\frac{\sqrt{3}i-1}{2})=\lambda_3$ and we have

$$p(Z)=\begin{pmatrix}1&\frac{-\sqrt{3}i-1}{2}&\frac{\sqrt{3}i-1}{2}\\1&\frac{\sqrt{3}i-1}{2}&\frac{-\sqrt{3}i-1}{2}\\1&1&1\end{pmatrix}p\left(\begin{pmatrix}1&0&0\\0&\frac{-\sqrt{3}i-1}{2}&0\\0&0&\frac{\sqrt{3}i-1}{2}\end{pmatrix}\right)\begin{pmatrix}1/3&1/3&1/3\\\frac{\sqrt{3}i-1}{6}&\frac{-\sqrt{3}i-1}{6}&1/3\\\frac{-\sqrt{3}i-1}{6}&\frac{\sqrt{3}i-1}{6}&1/3\end{pmatrix}=$$

$$=\begin{pmatrix}\frac{\lambda_3+2\lambda_1}{3}&\frac{-\sqrt{3}\lambda_3 i+\sqrt{3}\lambda_1 i+\lambda_1}{6}&\frac{\sqrt{3}\lambda_3 i-\lambda_3-\sqrt{3}\lambda_1 i+\lambda_1}{6}\\\frac{-\sqrt{3}\lambda_3 i+\sqrt{3}\lambda_1 i+\lambda_1}{6}&\frac{\lambda_3+2\lambda_1}{3}&\frac{-\sqrt{3}\lambda_3 i-\lambda_3+\sqrt{3}\lambda_1 i+\lambda_1}{6}\\\frac{-\sqrt{3}\lambda_3 i-\lambda_3+\sqrt{3}\lambda_1 i+\lambda_1}{6}&\frac{-\sqrt{3}\lambda_3 i+\sqrt{3}\lambda_1 i+\lambda_1}{6}&\frac{\lambda_3+2\lambda_1}{3}\end{pmatrix}$$

- $\lambda_1\neq\lambda_2\neq\lambda_3\neq\lambda_1$, we can proceed analogously to the last case.

Observe that $p(Z)$ is not unique as it is obtained from a polynomial interpolation.

Remark 3.4. ([9]) It is well known that any symmetric matrix $S\in M_n(\mathbb{R})$ is diagonalizable. Consequently, every symmetric matrix with real coefficients is similar to a Toeplitz matrix. Additionally, if we consider $N\in M_n(\mathbb{C})$ a normal matrix (N is diagonalizable by a unitary matrix), then N is similar to a Toeplitz matrix.

Chapter 4

Nonderogatory case

In this chapter we will relate nonderogatory and Toeplitz matrices. The strategy here is first constructing a similarity between a nonderogatory matrix and an upper Hessenberg matrix. Once having done this, we will introduce some lemmas with the final objective of showing that an upper Hessenberg matrix is similar to a unit upper Hessenberg Toeplitz matrix. Moreover, we will prove that the latter matrix is unique and therefore we will be able to talk about the Toeplitz canonical form.

4.1 Nonderogatory matrices

Definition 4.1.1. $A \in M_n(\mathbb{C})$ is *nonderogatory* if every eigenvalue of A has geometric multiplicity 1. Analogously, if in the Jordan form of A every eigenvalue of A appears in exactly one Jordan block or equivalently if the minimal and the characteristic polynomials of A coincide.

Another characterization of nonderogatory matrices is given in the following lemma:

Lemma 4.1.2. ([12]) $A \in M_n(\mathbb{C})$ is nonderogatory if and only if the only matrices that commute with A are polynomials in A.

4.2 Upper Hessenberg matrices

Definition 4.2.1. ([12]) $H = (h_{ij}) \in M_n(\mathbb{C})$ is *upper Hessenberg* if all its entries below the first subdiagonal are zero, videlicet, $h_{ij} = 0$ whenever $i > j + 1$.
Upper Hessenberg matrices are called *unreduced* when all entries on the first subdiagonal are non-zero, i.e., $h_{i+1,i} \neq 0$. Particularly, H is a *unit* upper Hessenberg matrix if all its subdiagonal entries are equal to one.

Observation 4.2.2. These matrices play an important role in matrix factorization as it simplifies the steps. For instance, in QR factorization that, what is more, assures a unique factorization (see [1]).

Example 4.2.3. An example of an unreduced and a unit Hessenberg matrix are respectively:

$$H_1 = \begin{pmatrix} 2 & 2 & 7 & 2 \\ 2 & 8 & 3 & 2 \\ 0 & 7 & 5 & 6 \\ 0 & 0 & 5 & 9 \end{pmatrix} \in M_4(\mathbb{C}) \text{ and } H_2 = \begin{pmatrix} 2 & 2 & 7 & 2 \\ 1 & 8 & 3 & 2 \\ 0 & 1 & 5 & 6 \\ 0 & 0 & 1 & 9 \end{pmatrix} \in M_4(\mathbb{C})$$

4.3 Nonderogatory matrices and upper Hessenberg matrices

The following proposition deals with the relation of similarity between nonderogatory matrices and unit upper Hessenberg matrices.

Proposition 4.3.1. ([4]) Let $A \in M_n(\mathbb{C})$. The following statements are equivalent:

(a) A is nonderogatory.

(b) A is similar to an upper Hessenberg matrix $H = (h_{ij})$ with $h_{ij} \neq 0$ for $i = j + 1$.

(c) A is similar to a unit upper Hessenberg matrix.

Proof. We will prove the following implications: $(a) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a)$.

<u>(a) ⇒ (c)</u> : Since A is nonderogatory, every eigenvalue has only one Jordan block, and thus their Jordan form is almost unit upper Hessenberg. We may now use Lemma 2.2.7 repeatedly to fill in the missing 1's on the first subdiagonal.

<u>(c) ⇒ (b)</u> : This result is trivial.

<u>(b) ⇒ (a)</u> : Let $H \in M_n(\mathbb{C})$ be a Hessenberg matrix such that $H \sim A$. For all $\lambda \in \mathbb{C}$, since $h_{ij} \neq 0$ for $i = j + 1$, we have that $Rank(H - \lambda I) \geq n - 1$ and using the dimension formula from Linear Algebra we obtain $dim(Ker(H - \lambda I)) \leq 1$. Consequently, the eigenspaces have dimension 1 and this completes the proof.

□

From now on, we will provide an example of every fundamental proof we use as a way of helping the reader to understand easily the vital parts of every proof and give a basic notion to those readers who are not so familiar with mathematical notation.

Example 4.3.2. We will give an example of <u>*(a) ⇒ (c)*</u>: Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in M_3(\mathbb{C})$ be a nonderogatory matrix. It is a simple matter by using the proof of Lemma 2.2.7 where $C = \begin{pmatrix} 0 & 1 \end{pmatrix}$ by taking the change of basis $S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ we obtain an upper Hessenberg matrix as it follows

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}$$

4.4 Nonderogatory matrices and Toeplitz matrices

This part is committed to prove that every nonderogatory matrix is similar to a unit upper Hessenberg Toeplitz matrix, $H \in M_n(\mathbb{C})$. In the above section we have proved that nonderogatory matrices are similar to unit upper Hessenberg matrices. Hence, it remains to prove that a unit upper Hessenberg matrix is similar to a unit upper Hessenberg Toeplitz matrix. The way to prove this is based on fixing any selected diagonal of H and making all the entries along this diagonal the same while keeping the lower diagonals unchanged. Doing this with every diagonal from lower left to upper right, we can obtain the desired matrix.

Since Toeplitz matrices are constant along diagonals, it will be natural to consider matrices from a diagonal perspective. Following standard convention, the northwest-southeast diagonals of a matrix will be numbered $-(n-1), \ldots, 0, \ldots, (n-1)$ starting from the lower left corner.

Definition 4.4.1. For each integer i, let $\triangle_i \subset M_n(\mathbb{C})$ be the subspace of matrices whose non-zero entries are restricted to the ith diagonal. This can be rewritten as

$$\triangle_i = \{A \in M_n(\mathbb{C}) \mid a_{rs} = 0 \quad \text{if} \;\; s - r \neq i\} \tag{4.4.1}$$

Observe that if $|i| > n-1$, then $\triangle_i$ is just the zero matrix.

Example 4.4.2. Following Example 4.3.2 of matrix A we have

$$\triangle_{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; \triangle_0 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \text{ and } \triangle_{-2} = \triangle_1 = \triangle_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Lemma 4.4.3. Let $\triangle_i \subset M_n(\mathbb{C})$.

$$A^{(i)} \in \triangle_i \text{ and } B^{(j)} \in \triangle_j \;\; \Rightarrow \;\; A^{(i)}B^{(j)} \in \triangle_{i+j}.$$

Proof. It is straightforward. □

Remark 4.4.4. Regarding general matrices, it will be useful for our study to have compact notation for a general matrix $A \in M_n(\mathbb{C})$ that consists of sorting all entries except those on the ith diagonal: $A^{(i)} \in \triangle_i$ is defined by

$$a_{rs}^{(i)} = \begin{cases} a_{rs} & \text{if } s - r = i \\ 0 & \text{otherwise} \end{cases}$$

Precisely,

$$A = \sum_i A^{(i)}$$

We need some previous results that will prompt us to our final result.

Lemma 4.4.5. ([4])

(a) Let $P \in M_n(\mathbb{C})$ be any unit upper triangular matrix with non-zero entries on at most two diagonals, i.e., $P = I_n + P^{(k)}$ for some $k \geq 1$. Then P^{-1} can have non-zero entries only on the ith diagonals where $i = 0, k, 2k, 3k, \ldots < n$.

(b) Let $A \in M_n(\mathbb{C})$ be an upper Hessenberg matrix and $B = P^{-1}AP$, where $P = I_n + P^{(k+1)}$ is unit upper triangular with $k \geq 0$. Then

(b.1) B is upper Hessenberg, particularly, $B^{(l)} = A^{(l)}$ for $l \leq k-1$.

(b.2) The kth diagonal of B (the first that could differ from A) is given by

$$B^{(k)} = A^{(k)} + A^{(-1)}P^{(k+1)} - A^{(-1)} + P^{(k+1)}A^{(-1)} \tag{4.4.2}$$

Proof.

(a) It is clear that $P^{(k)}$ is a nilpotent matrix and so $P^{(k)} + I_n$ is an invertible matrix which its inverse is

$$P^{-1} = (P^{(k)} + I_n)^{-1} = \sum_{j=0}^{n-1} (-P^{(k)})^j. \tag{4.4.3}$$

According to Lemma 4.4.3, we have $(P^{(k)})^j \in \triangle_{kj}$. So, P^{-1} is a linear combination of elements that the only entry which it is not zero is diagonal kj with $j \geq 0$.

(b)

(b.1) As a consequence of (a)

$$B = P^{-1}AP = \left(I_n - P^{(k+1)} + \sum_{j=2}^{n-1} (-P^{(k+1)})^j\right) A \left(I_n + P^{(k+1)}\right) = \tag{4.4.4}$$

$$= A + AP^{(k+1)} - P^{(k+1)}A - P^{(k+1)}AP^{(k+1)} + \sum_{j=2}^{n-1} (-1)^j (P^{(k+1)^j}) A \left(I_n + P^{(k+1)}\right)$$

By hypothesis, A is an Upper Hessenberg matrix, hence we can write

$$A = A^{(-1)} + A^{(0)} + \ldots + A^{(n-1)}$$

Substituting in Equation (4.4.4) and using part (a) of this lemma and Lemma 4.4.3, we can observe all the summands in the Equation (4.4.4) belong to the set $\triangle_i$ with $i \in \{-1, \ldots, n-1\}$ as is easy to check by analyzing every summand

$$AP^{(k+1)} = (A^{(-1)} + \ldots + A^{(n-1)}P^{(k+1)} \in \triangle_k + \ldots + \triangle_{k+n}, \tag{4.4.5}$$

$$P^{(k+1)}A = P^{(k+1)}(A^{(-1)} + \ldots + A^{(n-1)})P^{(k+1)} \in \triangle_{2k+1} + \ldots + \triangle_{2k+n+1}$$

$$(P^{(k+1)})^j (A^{(-1)} + \ldots + A^{(n-1)})(I_n + P^{(k+1)}) \in \triangle_{(k+1)j-1} + \ldots + \triangle_{(k+1)j+n-1} +$$

$$+ \triangle_{(k+1)j+k} + \ldots + \triangle_{(k+1)j+k+n}$$

Thus, we see that the non-zero contributions to $B^{(\ell)}$ for $\ell \leq k-1$ come only from A, the first term in (4.4.4). Consequently, we have that B is an upper Hessenberg matrix and $B^{\ell} = A^{\ell}$ for $\ell \leq k-1$.

(b.2) From Equations (4.4.4) and (4.4.5) we have

$$B^{(k)} = A^{(k)} + (AP^{(k+1)})^{(k)} - \left(P^{(k+1)}A\right)^{(k)} - \left(P^{(k+1)}AP^{(k+1)}\right)^{(k)} +$$

$$+\left(\sum_{j=2}^{n-1}(-1)^j(P^{(k+1)j})A\left(I_n+P^{(k+1)}\right)\right)^{(k)}=$$

$$=A^{(k)}+(A^{(-1)}P^{(k+1)})-(P^{(k+1)}A^{(-1)})-(P^{(k+1)}A^{(-k-2)}P^{(k+1)})+$$

$$+\left(\sum_{j=2}^{n-1}(-1)^j((P^{(k+1)j}A)\left(I_n+P^{(k+1)}\right)\right)^{(k)}=A^{(k)}+(AP^{(k+1)})^{(k)}-(P^{(k+1)}A)^{(k)}$$

Therefore,

$$B^{(k)}=A^{(k)}+A^{(-1)}P^{(k+1)}-P^{(k+1)}A^{(-1)}$$

□

The following theorem states the relation of similarity between a nonderogatory matrix and a Toeplitz matrix. Note that the proof strongly depends on the restrictive assumption of having a nonderogatory matrix.

Theorem 4.4.6. ([4]) Every nonderogatory matrix in $M_n(\mathbb{C})$ is similar to a Toeplitz matrix, in particular, to a unit upper Hessenberg Toeplitz matrix.

Proof. As a result of Proposition 4.3.1 we can suppose A is a unit upper Hessenberg matrix. Let $P\in M_n(\mathbb{C})$ be a unit upper triangular matrix of the form $P=I_n+P^{(k+1)}$ with $0\le k\le n-2$ and we consider as in Lemma 4.4.5, $B=P^{-1}AP$. We will prove that we can take P such that all the entries of the kth diagonal of B are equal. By Lemma 4.4.5, we can choose $P^{(k+1)}$ such that

$$B^{(k)}=A^{(k)}+A^{(-1)}P^{(k+1)}-P^{(k+1)}A^{(-1)} \tag{4.4.6}$$

where $A^{(-1)}=\begin{pmatrix}0&&&&\\1&0&&&\\&1&\ddots&&\\&&\ddots&\ddots&\\&&&1&0\end{pmatrix}=F$. (Forward shift Toeplitz matrix (Example 2.3.2)).

Recalling Definition 2.2.2, the Sylvester operator

$$\varphi:\triangle_{k+1}\to\triangle_k \tag{4.4.7}$$

$$X\mapsto FX-XF$$

allows rewriting Equation (4.4.6) as

$$B^{(k)}=A^{(k)}+\varphi(P^{(k+1)})$$

which is well defined as if $P^{(k+1)}\in\triangle_{k+1}$ then $\varphi(P^{(k+1)})\in\triangle_k$.

Our next concern will be the behavior of $Im(\varphi)$ because, in this way, we will be able to study the kth diagonals of $B^{(k)}$.

Let us consider the standard product on $M_n(\mathbb{C})$

$$\langle A,B\rangle=tr(AB^*) \tag{4.4.8}$$

As $\triangle_k$ and $\triangle_{k+1}$ are linear subspaces of $M_n(\mathbb{C})$ we can restrict the inner product to these two subspaces.

If φ^* is the adjoint map of φ we have that

$$\langle \varphi(X), Y \rangle = \langle X, \varphi^*(Y) \rangle \tag{4.4.9}$$

$$\langle \varphi(X), Y \rangle = \langle FX - XF, Y \rangle = tr((FX - XF)Y^t) = tr(FXY^t - XFY^t) = \tag{4.4.10}$$

$$= tr(FXY^t) - tr(XFY^t) = tr(XY^tF) - tr(XFY^t) = tr(XY^tF - XFY^t) = \langle X, F^tY - YF^t \rangle$$

Combining Equations (4.4.9) and (4.4.10) we obtain that

$$F^tY - YF^t = \varphi^*(Y)$$

where the the adjoint map φ^* is

$$\varphi^* : \triangle_k \to \triangle_{k+1} \tag{4.4.11}$$

$$Y \mapsto F^tY - YF^t$$

From $Im(\varphi) = (Ker(\varphi^*))^\perp$, we can express $\triangle_k$ as the following direct sum

$$\triangle_k = Ker(\varphi^*) \oplus Im(\varphi) \tag{4.4.12}$$

Hence, we can compute $Im(\varphi)$ by finding $(Ker(\varphi^*))^\perp$.

$$Ker(\varphi^*) = \{Y \in M_n(\mathbb{C}) \mid \varphi^*(Y) = 0\} = \{Y \in M_n(\mathbb{C}) \mid F^tY - YF^t = 0\}$$

Invoking Proposition 4.1.2 we know that the matrices that commute with a nonderogatory matrix are triangular Toeplitz matrices. Thus,

$$Ker(\varphi^*) = \{T \in \triangle_k \mid T \text{ is an upper triangular Toeplitz matrix}\}$$

Now,

$$Im(\varphi) = (Ker(\varphi^*))^\perp = \{B \mid \langle B, T \rangle = 0 \ \ \forall T \in Ker(\varphi^*)\} = \{B \mid tr(BT^t) = 0 \ \ \forall T \in Ker(\varphi^*)\}$$

Let $T = (t_{ij})_{1 \le i,j \le n} \in M_n(\mathbb{C})$ be an upper triangular Toeplitz matrix, and $B = (b_{ij})_{1 \le i,j \le n} \in M_n(\mathbb{C})$, then

$$tr(BT^t) = \sum_{i=1}^{n} (BT^t)_{ii} = \sum_{j=1}^{n} \sum_{i=1}^{n} (B)_{ij} (T^t)_{ji} = \sum_{i=1}^{n-1} \sum_{j=i}^{n} t_{j-i+1} b_{ij}$$

Hence,

$$Im(\varphi) = \{B = (b_{ij})_{1 \le i,j \le n} \in \triangle_k \mid \sum_{i=1}^{n} \sum_{j=0}^{n} b_{ij} = 0\}$$

We are left with the task of constructing $P^{(k+1)}$. From Equation (4.4.12) we know that $A^{(k)}$ has the unique following decomposition

$$A^{(k)} = T + R$$

where $T \in Ker(\varphi^*)$ and $R \in Im(\varphi)$.

Let $\alpha \in \mathbb{C}$ be the average of all the entries in the kth diagonal of $A^{(k)}$, for instance,

$$\alpha = \frac{1}{n-k} \sum_{i=k}^{n-k} a_{i,i+1}$$

We take $T \in \triangle_k$ as the matrix with all α's on the kth diagonal and $R = A^{(k)} - T \in Im(\varphi)$. Let $P^{(k+1)} \in \triangle_{k+1}$ be the unique matrix such that $\varphi(P^{(k+1)}) = -R$. We solve the system of equations by back-substitution, and we obtain

$$B^{(k)} = A^{(k)} + \varphi(P^{(k+1)}) = T$$

being $B^{(k)}$ is a Toeplitz matrix. □

Example 4.4.7. Following with the Example 4.3.2 we will make the construction of the last proof. Let us consider $A = H = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix}$. We can choose the form of P such that all the 0th diagonal of A is constant. Observe that $A^{(-1)}$ is constant. So, we only need to change $A^{(0)}$. We want to find $P \in M_3(\mathbb{C})$ such that

$$P^{-1}AP = T = \begin{pmatrix} * & 0 & 0 \\ 1 & * & 0 \\ 0 & 1 & * \end{pmatrix}$$

and applying Lemma 4.4.5 we have that $P = I_3 + P^{(1)}$. What is more, we have that

$$B^{(k)} = A^{(k)} + \varphi(P^{(k+1)})$$

where

$$\begin{aligned} \varphi : \triangle_1 &\to \triangle_0 \\ \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} &\mapsto FX - XF = \begin{pmatrix} -x & 0 & 0 \\ 0 & x-y & -y+z \\ 0 & 0 & y \end{pmatrix} \end{aligned}$$

and

$$\triangle_1 = \{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\}, \quad \triangle_0 = \{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\}$$

Note that $Im(\varphi) = \{\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\}$. We also consider the adjoint map of φ

$$\begin{aligned} \varphi^* : \triangle_0 &\to \triangle_1 \\ \begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix} &\mapsto F^tY - YF^t = \begin{pmatrix} 0 & -x+y & 0 \\ 0 & 0 & -y+z \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

as we have that

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & y \end{pmatrix}\begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

It is clear that $Ker(\varphi^*) = \{Id\}$. We also have that the unique decomposition of $A^{(0)}$ is

$$A^{(0)} = T + R \text{ with } T \in Ker(\varphi^*),\ R \in Im(\varphi)$$

We take

$$T = \begin{pmatrix} \frac{7}{3} & 0 & 0 \\ 0 & \frac{7}{3} & 0 \\ 0 & 0 & \frac{7}{3} \end{pmatrix}$$

where $\alpha = \frac{7}{3}$ is the average of the elements in the 0th diagonal. Then,

$$R = A^{(0)} - T = \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} = \varphi(P^{(1)})$$

Hence, we obtain that

$$B^{(0)} = A^{(0)} + \varphi(P^{(1)}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & -2/3 \end{pmatrix} = \begin{pmatrix} 7/3 & 0 & 0 \\ 0 & 7/3 & 0 \\ 0 & 0 & 7/3 \end{pmatrix}$$

as

$$\varphi(P^{(1)}) = \begin{pmatrix} -x & 0 & 0 \\ 0 & x-y & 0 \\ 0 & 0 & y \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{2}{3} \end{pmatrix}$$

and therefore $x = \frac{1}{3}$ and $y = \frac{2}{3}$, which implies that $P = \begin{pmatrix} 1 & \frac{1}{3} & 0 \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{pmatrix}$.

Finally, we obtain the desired result

$$A \sim B = \begin{pmatrix} 7/3 & 0 & 0 \\ 1 & 7/3 & 0 \\ 0 & 1 & 7/3 \end{pmatrix}$$

4.5 The Toeplitz canonical form for a nonderogatory matrix

The objective of this section is to prove the uniqueness of the upper Hessenberg Toeplitz matrix defined above. Afterwards, we will be able to refer to it as the *Toeplitz canonical form.*

First, we need to give some related results. The principal significance of the next two lemmas is setting up a recurrence relation for characteristic polynomials of unit upper Hessenberg Toeplitz matrices. On the one hand, Lemma 4.5.1 studies the characteristic polynomial of a unit upper Hessenberg matrix and, on the other hand, Lemma 4.5.2 studies the characteristic polynomial of a Hessenberg matrix. After being set up this, we will give the formal definition of the mentioned canonical form.

Lemma 4.5.1. Let $T_n = \begin{pmatrix} a_1 & a_2 & \cdots & & a_n \\ 1 & a_1 & \ddots & & \vdots \\ 0 & 1 & \ddots & \ddots & \\ \vdots & & & & a_2 \\ 0 & & \cdots & 1 & a_1 \end{pmatrix}$ be a unit upper Hessenberg Toeplitz matrix. We define $p_0(x) = 1$ and $p_n(x) = det(xId - T_n)$. Then,

$$p_n(x) = xp_{n-1}(x) - \sum_{i=1}^{n} a_i p_{n-i}(x) \qquad \text{for } n = 1, 2, \ldots$$

Proof. We will use induction to prove this result.
Base case, $n = 1$:

$$p_1(x) = det(-T_1 + xI) = det(-a_1 + x) = -a_1 + x = xp_0(x) - a_1 p_0(x) \tag{4.5.1}$$

Inductive step: Show that for any $n \geq 1$, if $p_{n-1}(x) = xp_{n-1}(x) - \sum_{i=0}^{n} a_i p_{n-i}(x)$ holds, then $p_n(x)$ holds.
By definition:

$$p_n(x) = det(-T_n + xI) = \begin{vmatrix} -a_1 + x & -a_2 & \cdots & -a_{n-1} & a_n \\ -1 & -a_1 + x & -a_2 & \cdots & -a_{n-1} \\ 0 & -1 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -a_2 \\ 0 & & & -1 & -a_1 + x \end{vmatrix}$$

We apply the Laplace expansion along the first column in order to compute the determinant. Note that the first minor of dimension $(n-1) \times (n-1)$ is exactly $p_{n-1}(x)$. It is immediate that if we remove row i and column 1 for $i = 1, \ldots, n-1$ we get a minor where its last row is $0, \ldots, 0, -1$. And, applying again the Laplace expansion we will clearly obtain that the determinant of the minor is $(-1)p_{n-i}(x)$. Hence,

$$p_n(x) = (-a_1 + x)p_{n-1}(x) + \sum_{i=1}^{n-1} -a_i p_{n-i}(x) = xp_{n-1}(x) - \sum_{i=1}^{n} a_i p_{n-i}(x)$$

□

The above lemma gives the relationship between the coefficients of the characteristic polynomial $p_n(x) \in \mathbb{C}[x]$ in terms of the entries of the Toeplitz matrix $T_n \in M_n(\mathbb{C})$. If we write

$$p_n(x) = x^n + c_{n,1}x^{n-1} + c_{n,2}x^{n-2} + \ldots + c_{n,n-1}x + c_{n,n}$$

then each coefficient $c_{n,i}$ is a polynomial involving all the n complex variables $a_1, a_2, \ldots, a_n$.

The following lemma shows that $c_{n,i}$ involves just the first i variables $a_1, a_2, \ldots, a_n$ and depends linearly on a_i.

Lemma 4.5.2. If $p_n(x) = x^n + c_{n,1}x^{n-1} + c_{n,2}x^{n-2} + \ldots + c_{n,n-1}x + c_{n,n}$ is the characteristic polynomial of $T_n = \begin{pmatrix} a_1 & a_2 & \cdots & & a_n \\ 1 & a_1 & \ddots & & \vdots \\ 0 & 1 & \ddots & \ddots & \\ \vdots & & & & a_2 \\ 0 & & \cdots & 1 & a_1 \end{pmatrix}$, then we have that

$$c_{n,i} = \begin{cases} -na_1 & \text{if } i = 1 \\ -(n-i+1)a_i + d_{ni}(a_1, a_2, \ldots, a_{i-1}) & \text{if } 2 \leq i \leq n \end{cases} \tag{4.5.2}$$

where $d_{ni}(a_1, a_2, \ldots, a_{i-1})$ is a polynomial in $a_1, a_2, \ldots, a_{i-1}$.

Proof. The proof is by induction on n.
Base case, $n = 1$:

$$p_1(x) = det(-T_1 + xI) = det(-a_1 + x) = -a_1 + x \tag{4.5.3}$$

Note that $c_{1,1} = -a_1$.
Base case, $n = 2$:

$$p_2(x) = det(-T_2 + xI) = x^2 - 2a_1 x + (-a_2 + a_1^2) \tag{4.5.4}$$

We see that $c_{2,1} = -2a_1$ and $c_{2,2} = -a_2 + a_1^2$ with $d_{22}(a_1) = a_1^2$.

Inductive step: We assume that c_{ki} follows (4.5.2) for each $k \le n-1$ and $1 \le i \le k$. Now we want to prove for $k = n$:

On the one hand, we know that the characteristic polynomial follows this structure:

$$p_i(x) = x^i + c_{i,1}x^{i-1} + c_{i,2}x^{i-2} + \ldots + c_{i,i-1}x + c_{i,i}$$

On the other hand, by (4.5.1), we have that

$$p_n(x) = x\,p_{n-1}(x) - (a_1 p_{n-1}(x) + \ldots + a_{n-1}p_1(x) + a_n)$$

Substituting (4.5.2) into (4.5.2) we obtain that

$$p_n(x) = x(x^{n-1} + c_{n-1,1}x^{n-2} + \ldots + c_{n-1,n-2}x + c_{n-1,n-1})$$
$$-\big(a_1(x^{n-1} + c_{n-1,1}x^{n-2} + \ldots + c_{n-1,n-2}x + c_{n-1,n-1}) + \ldots + a_{n-1}(x - c_{1,1}) + a_n\big)$$

Now, we consider Equation (4.5.2) for $i = n$ and we will match term to term both equations. Regarding term x^{n-1}, we have that $c_{n,1} = c_{n-1,1} - a_1 = -(n-1)a_1 - a_1 = -na_1$. Where in the penultimate equality we have applied induction hypothesis.
Now, we focus on study what happens with term x^{n-i}, for $i = 2, \ldots, n$.

$$c_{n,i} = c_{n-1,i} - (a_1 c_{n-1,i-1} + a_2 c_{n-2,i-1} + \ldots + a_{i-1}c_{n-i+1,1} + a_i)$$
$$= (c_{n-1,i} - a_i) - (a_1 c_{n-1,i-1} + a_2 c_{n-2,i-2} + \ldots + a_{i-1}c_{n-i+1,1}) = (A) + (B)$$

We apply induction hypothesis to $c_{n-1,i-1}, c_{n-2,i-2} + \ldots + c_{n-i+1,1}$ and we have that (B) is a polynomial in the variables $a_1, \ldots, a_{i-1}$. Now we study A, we apply induction hypothesis to (A) gives: $c_{n-1,i} - a_i = -(n-i+1)a_i$. Putting both facts together we obtain the desired result. ☐

Once introduced the former lemmas, the following theorem gives the uniqueness of the unit upper Hessenberg Toeplitz matrix and it will be reasonable to define the *Toeplitz canonical form.*

Theorem 4.5.3. ([4]) Every nonderogatory matrix $A \in M_n(\mathbb{C})$ is similar to a unique unit upper Hessenberg Toeplitz matrix.

Proof. In Theorem 4.5.3 we have proved that every nonderogatory matrix in $M_n(\mathbb{C})$ is similar to a unit upper Hessenberg Toeplitz matrix. For this reason, the proof is completed by showing that if two upper Hessenberg Toeplitz matrices are similar, then they are equal. Set

$$H_1 = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & a_1 & \ddots & \ddots & a_{n-1} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_2 \\ 0 & \cdots & 0 & 1 & a_1 \end{pmatrix} \quad \text{and} \quad H_2 = \begin{pmatrix} b_1 & b_2 & \cdots & b_{n-1} & b_n \\ 1 & b_1 & \ddots & \ddots & b_{n-1} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & b_2 \\ 0 & \cdots & 0 & 1 & b_1 \end{pmatrix}$$

We suppose that H_1 and H_2 are similar matrices. Consequently, their characteristic polynomials must be the same: $p_{H_1}(x) = p_{H_2}(x)$. By Lemma 4.5.2 we have that

$$\begin{cases} -na_1 = -nb_1 & \text{if } i = 1 \\ -(n-i+1)a_i + d_{ni}(a_1, a_2 \dots, a_{i-1}) = -(n-i+1)b_i + d_{ni}(b_1, b_2 \dots, b_{i-1}) & \text{if } 2 \le i \le n \end{cases}$$

It is clear that $a_k = b_k$ for all $1 \le k \le n$ by using the induction hypothesis that $d_{nk}(a_1, a_2, \dots, a_{k-1}) = d_{nk}(b_1, b_2, \dots, b_{k-1})$ and taking into account that $n - k + 1 \ne 0$ as $k \le n$. Finally, we have proved the uniqueness. □

Definition 4.5.4. A *Toeplitz canonical form* for a nonderogatory matrix $A \in M_n(\mathbb{C})$ is the unit upper Hessenberg Toeplitz matrix $T \in M_n(\mathbb{C})$ similar to A.

4.6 Extension of other fields of the nonderogatory case

In earlier sections, we have been working under the supposition that $\mathbb{F} = \mathbb{C}$ and we have proved that every nonderogatory matrix with entries in the complex field is similar to an upper Hessenberg Toeplitz matrix. One may ask whether the result above is still true if we consider nonderogatory matrices with entries in any field $\mathbb{F}$. To study the general case, we consider the following three cases:

- $Char(\mathbb{F}) = 0$
  In this case, all the construction of the similarity follows as all the operations defined above will be well defined.

- $Char(\mathbb{F}) = p > n$
  The same conclusion of the first case can be drawn for $p > n$ as all the operations will be well defined.

- $Char(\mathbb{F}) = p \le n$.
  Loosely speaking, we have that when we apply the procedure explained above, we can arrive at a point that we divide by zero because of the characteristic of the field and then the proof will be not well defined.
  Consequently, there exists a nonderogatory matrix $A \in M_n(\mathbb{F})$ that is *not* similar to any upper Hessenberg Toeplitz matrix in $M_n(\mathbb{F})$. We will give a counterexample to prove the assertion.

  Counterexample:
  We suppose $n = 4$ and $Char(\mathbb{F}) = p = 2$. Let

  $$B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in M_n(\mathbb{F})$$

  The characteristic polynomial of B is $p_B(x) = x^4 - x^3$, and therefore $c_{4,1} = -1, c_{4,2} = c_{4,3} = c_{4,4} = 0$. If $T = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ 1 & a_0 & a_1 & a_2 \\ 0 & 1 & a_0 & a_1 \\ 0 & 0 & 1 & a_0 \end{pmatrix} \sim B$, they must have the same characteristic polynomial. From Lemma 4.5.2, we have one of the equations of the system is

$c_{4,1} = -4a_0 = -1$ and as we are in a field of $Char(\mathbb{F}) = 2$, we obtain $0 = -1$, which is a contradiction. In consequence, we have that there is no unit upper Hessenberg Toeplitz matrix defined in $M_4(\mathbb{F})$ can have this characteristic polynomial, which is our assertion.

Chapter 5

General case for n ≤ 4

In this chapter we will restrict ourselves to matrices of low dimension dropping the assumption of being a diagonalizable or a nonderogatory matrix. It is worth pointing out that these conditions cannot be relaxed in general, as all the construction of the results proved closely depends on them. We will show that all 4×4 (or smaller) complex matrices are similar to a Toeplitz matrix.

Theorem 5.1. ([4]) Every $A \in M_n(\mathbb{C})$ with $n \leq 4$ is similar to a Toeplitz matrix.

Proof. The proof falls naturally into three parts:

- n = 2
  In this case, the possible Jordan form of a 2×2 matrix are:

$$J_1 = \begin{pmatrix} \lambda_1 & 0 \\ 1 & \lambda_1 \end{pmatrix}, \; J_2 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix} \text{ and } J_3 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{with } \lambda_1 \neq \lambda_2$$

  Observe J_1 is nonderogatory and J_2 and J_3 are diagonalizable. Hence, the three matrices are similar to a Toeplitz matrix (Theorem 3.2 and 4.5.3).

- n = 3
  We will consider three sub-cases depending on the cardinality of $Spec(A)$, denoted as $|Spec(A)|$.

  - If $|Spec(A)| = 3$, then $A = diag(\lambda_1, \lambda_2, \lambda_3)$ which is similar to a Toeplitz matrix.
  - If $|Spec(A)| = 2$, we have two possible forms:

$$J_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \text{ and } J_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \quad \text{with } \lambda_1 \neq \lambda_2$$

    J_1 is diagonalizable and J_2 is nonderogatory, and thus both are similar to a Toeplitz matrix.

  - If $|Spec(A)| = 1$, we have three possible forms:

$$J_1 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}, \; J_2 = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix} \text{ and } J_3 = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

J_1 and J_2 are Toeplitz so we are done. Observe that $J_3 = \lambda_1 I + N_3$ with

$$N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and if we consider

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

we obtain that

$$S^{-1} N_3 S = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and thus

$$S^{-1} J_3 S = \begin{pmatrix} \lambda_1 & 0 & 1 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix}$$

and J_3 is similar to a Toeplitz matrix.

- <u>n = 4</u>
  The proof is similar in spirit to the case $n = 3$.
  - If $|Spec(A)| = 4$, we are in the diagonalizable case.
  - If $|Spec(A)| = 3$, we have two possible matrices:

$$J_1 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix} \text{ and } J_2 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

    J_1 is diagonalizable and J_2 is nonderogatory, so we are done.
  - If $|Spec(A)| = 2$, we have six possibilities:

$$D_1 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \quad O_1 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \quad ND_1 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 1 & \lambda_2 \end{pmatrix}$$

$$D_2 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \quad O_2 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix} \quad ND_2 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 1 & \lambda_1 & 0 & 0 \\ 0 & 1 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

    Observe we named every matrix according to its type. D_i (diagonalizable) and ND_i (nonderogatory) with $i = 1, 2$ are clearly similar to a Toeplitz matrix.
    It remains to study O_i cases. Our aim is to find a similarity that results with a Toeplitz matrix. The computations are lengthily and muddled, so we have used computer programs (such as Maple and Mathematica) to compute it and the proof is constructive as it follows:

The first thing we do is assuming $\lambda_1 = 0$ and we construct a similarity to $\bar{T}_2$. If we take

$$\bar{T}_2 = \begin{pmatrix} 4+2i & 4+4i & 3+6i & 1+8i \\ 4 & 4+2i & 4+4i & 3+6i \\ 4 & 4 & 4+2i & 4+4i \\ 0 & 4 & 4 & 4+2i \end{pmatrix} \tag{5.0.1}$$

and computing its Jordan form we have

$$\bar{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 16+8i \end{pmatrix} \tag{5.0.2}$$

Now we scale and shift $\bar{J}_2$ and we obtain

$$\frac{\lambda_2 - \lambda_1}{16+8i}\bar{J}_2 + \lambda_1 Id = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & \frac{\lambda_2-\lambda_1}{16+8i} & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

Now we look for a similarity that makes the element $\frac{\lambda_2-\lambda_1}{16+8i}$ in the position (2,3) be a 1. This is $D = diag(1, 16+8i, \lambda_2 - \lambda_1, 1)$. Finally, we have that

$$O_2 = D(\frac{\lambda_2 - \lambda_1}{16+8i}\bar{J}_2 + \lambda_1 Id)D^{-1} = DS_2(\frac{\lambda_2 - \lambda_1}{16+8i})S_2^{-1}D^{-1}$$

This let us conclude that O_2 is similar to the Toeplitz matrix $T_2 = \frac{\lambda_2-\lambda_1}{16+8i}\bar{J}_2 + \lambda_1 Id$.

– $|Spec(A)| = 1$. There are five different possibilities:

$$J_1 = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}, \quad J_2 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

$$J_3 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}, \quad J_4 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \text{ and } J_5 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

Clearly, J_1 and J_2 are already Toeplitz matrices. As well, J_3 and J_4 can be split as: $J_i = \lambda_1 Id + N_i$ for $i = 3, 4$ with

$$N_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } N_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and by applying a permutation we obtain the Toeplitz matrix of the forms:

$$T_3 = \begin{pmatrix} \lambda_1 & 0 & 0 & 1 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix} \text{ and } T_4 = \begin{pmatrix} \lambda_1 & 0 & 1 & 0 \\ 0 & \lambda_1 & 0 & 1 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

To end our proof, we must construct a similarity for J_5. With the same reasoning as before we have that taking

$$T_5 = \begin{pmatrix} 0 & -2i & 2 & -1+2i \\ 4 & 0 & -2i & 2 \\ 8 & 4 & 0 & -2i \\ 16+8i & 8 & 4 & 0 \end{pmatrix} \text{ and } S_5 = \begin{pmatrix} 0 & 2-2i & -1+i & 0 \\ 0 & -4+4i & -4-4i & 6-2i \\ 64 & 32-32i & 32 & -16+16i \\ -64 & -32-96i & -160 & 80-16i \end{pmatrix}$$

and this concludes the proof.

□

Remark 5.2. In *Chapter 5* (concretely, in Theorem 6.2.1) we will see that the condition of being in the field $\mathbb{F} = \mathbb{C}$ is essential to the proof as we will see an example that shows that in $\mathbb{F} = \mathbb{R}$ the theorem is not true.

Chapter 6

General case: Beyond n $\geq$ 5

This chapter deals with the cases of higher dimension. We will show how to dispense with the hypothesis of having general matrices of dimension higher than 5 as we cannot proceed as in the preceding chapter (general case for $n \leq 4$) as the calculus are computationally expensive and very long. Moreover, we will not be able to do constructive proofs or give a general result as we did in the former chapters (diagonalizable and nonderogatory cases).

We will focus on some special cases in order to answer the main question of this paper: if every matrix is similar to a Toeplitz matrix and we will see the answer to the question is negative. To do so we will need to make some specific assumptions.

In what follows, we will present some new notation:

Notation 6.1. Let $A \in M_n(\mathbb{C})$ and $\lambda_0 \in Spec(A)$. We define the integers

$$d_k(A,\lambda_0) = dim(Ker(A-\lambda_0 I_n)^{k+1}) - dim(Ker(A-\lambda_0 I_n)^k)$$

for $k = 0, 1, \ldots$.
Observe that $d_0(A,\lambda_0)$ is the *geometric multiplicity* of λ_0 and $d_k(A,\lambda_0)$ is the number of Jordan blocks corresponding to λ_0 with a size greater than k.

In the general inverse Jordan structure problem, λ_k and d_{kj} with $j = 0, \ldots, q_k$ are given and we look for a matrix A with eigenvalues λ_k and $d_j(A,\lambda_k) = d_{kj}$. In the case of $n \times n$ matrices, the problem has a solution if $\sum_k \sum_{j=0}^{q_k} d_{kj} \leq n$ (see [2] and [7]). If A is Toeplitz, the problem is very complex and it is hard to set if the problem has a solution. So, we will make the following assumption:

Assumption 6.2. We will suppose we have just one eigenvalue $\lambda_0 \in \mathbb{C}$ and, without loss of generality, we can assume $\lambda_0 = 0$ since the identity matrix is Toeplitz. According to the above notation, we write $d_k(A,0) = d_k(A)$.

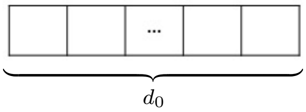
Notation 6.3. From now on

$$S_n = \begin{pmatrix} 0 & \cdots & & 0 \\ 1 & & & \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{pmatrix} \in M_n(\mathbb{C})$$

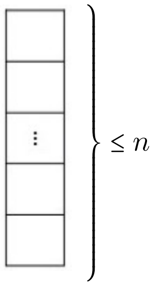
Definition 6.4. ([5]) The $(q+2)$*-tuple* $(n, d_0, \ldots, d_q)$ with $0 \le d_q \le \ldots \le d_0$ and $\sum_{j=0}^{q} d_j \le n$ is *Toeplitz admissible* if there exists a Toeplitz matrix $T \in M_n(\mathbb{C})$ with $d_k(T) = d_k$ for $k = 0, \ldots, q$.

Examples 6.5. In the following examples we will study basic cases of Toeplitz admissible matrices. What is more, we will sketch them using Young diagram that will be useful to visualize and make more understable further results.

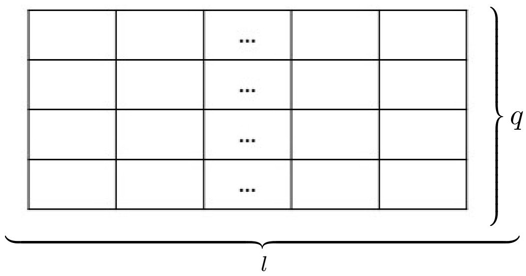
- (n, d_0) with $d_0 \le n$ is Toeplitz admissible since we can take any matrix $T \in M_n(\mathbb{C})$ with $Rank(T) \le n - d_0$.



- $(n, \underbrace{1, \ldots, 1}_{\le n})$ corresponds to a nonderogatory matrix and so it is Toeplitz admissible (Theorem 4.4.6).



- $(n, \underbrace{l, \ldots, l}_{q})$ is Toeplitz admissible for $ql \le n$. According to Notation 6.3 we have that $d_k(S_n^l) = l$ for $kl \le n$.



In general, the study if a tuple is Toeplitz admissible is a difficult problem that we will have to study it carefully.

In the next lemma we will study the structure of the intersection of the kernel and the range of a Toeplitz matrix as it has the same structure as the kernel of a square Toeplitz matrix. Our approach is based on the following result:

Lemma 6.6. ([5]) Given $A \in M_n(\mathbb{C})$, then

$$d_1(A) = dim(Ker(A) \cap Im(A)) \tag{6.0.1}$$

Proof. Let f be a linear map defined as

$$f : Ker(A^2) \to Ker(A) \cap Im(A)$$
$$X \mapsto AX$$

Given $Y \in Ker(A) \cap Im(A)$, we have that $AY = 0$ and $Y = AX = f(X)$ and therefore f is exhaustive.
As well,

$$Ker(f) = \{X \in Ker(A^2) \mid f(X) = 0\} = Ker(A)$$

Applying the first isomorphism theorem, we get that

$$Ker(A^2)/Ker(f) = Ker(A^2)/Ker(A) \cong Im(f) = Ker(A) \cap Im(A)$$

And the result follows. □

The prior lemma points us to study the structure of $Ker(T) \cap Im(T)$ in order to go further in our investigations. At first, we will focus on the structure of $Ker(T)$ and afterwards the intersection specified.

Notation 6.7. Given a vector $u = (u_k)_{k=0}^m \in \mathbb{C}^{m+1}$ then

- $\hat{u} = (u_k)_{k=m}^0 \in \mathbb{C}^{m+1}$ is the *reverse vector.*
- $u(\lambda)$ will denote the polynomial $u(\lambda) = \sum_{k=0}^m u_k \lambda^k$.
- $T_d(u)$ is a Toeplitz matrix of the form

$$T_d(u) = \begin{pmatrix} u_0 & & & 0 \\ \vdots & u_0 & & \\ u_m & \vdots & \ddots & \\ & u_m & & u_0 \\ & & \ddots & \vdots \\ 0 & & & u_m \end{pmatrix} \in M_{m+d,d}(\mathbb{C})$$

Note that $T_d(u)$ is the matrix of the multiplication operator $x(\lambda) \to u(\lambda)x(\lambda)$ where $x(\lambda)$ is a polynomial of degree $d-1$ with respect to the basis λ^i.

Lemma 6.8. ([5]) Let $T = (a_{i-j}) \in M_n(\mathbb{C})$ be a Toeplitz matrix and let $T_k = (a_{i-j})_{i=1}^{2n-k;k}$ for $k = 1, \dots, 2n-1$. Then

$$u \in Ker(T_k) \iff v = (u^t\ 0)^t \in Ker(T_{k+1}) \text{ and } \omega = (0\ , u^t)^t \in Ker(T_{k+1})$$

Proof. Observe that $T_{k-1} \in M_{2n-k+1,k-1}(\mathbb{C})$ is obtained from $T_k \in M_{2n-k,k}(\mathbb{C})$ by deleting the last column and adding a top row in such way that the structure of a Toeplitz matrix is preserved.
If

$$u \in Ker(T_k) \iff T_k u = (\underbrace{0, \dots, 0}_{2n-k}) \iff \sum_{i=2}^{2n-k} \sum_{j=1}^{k} T_{ij} u_j = 0 \iff \sum_{i=1}^{2n-k-1} \sum_{j=1}^{k+1} T_{ij} v_j = 0 \iff$$

$$\iff \sum_{i=1}^{2n-k-1} \sum_{j=1}^{k+1} T_{ij} \omega_j = 0$$

The converse implication follows likewise. □

Example 6.9. As the notation of the last proof is complicated, we will give an example. Let $T = \begin{pmatrix} \frac{-2}{3} & \frac{1}{9} & \frac{4}{27} \\ 1 & \frac{-2}{3} & \frac{1}{9} \\ 0 & 1 & \frac{-2}{3} \end{pmatrix}$, then

$$T_1 = \begin{pmatrix} \frac{4}{27} \\ \frac{1}{9} \\ \frac{-2}{3} \\ 1 \\ 0 \end{pmatrix}, \; T_2 = \begin{pmatrix} \frac{1}{9} & \frac{4}{27} \\ \frac{-2}{3} & \frac{1}{9} \\ 1 & \frac{-2}{3} \\ 0 & 1 \end{pmatrix}, \; T_3 = T, \; T_4 = \begin{pmatrix} 1 & \frac{-2}{3} & \frac{1}{9} & \frac{4}{27} \\ 0 & 1 & \frac{-2}{3} & \frac{1}{9} \end{pmatrix} \text{ and } T_5 = \begin{pmatrix} 0 & 1 & \frac{-2}{3} & \frac{1}{9} & \frac{4}{27} \end{pmatrix}$$

Let us denote by u_k the kernel generating vector of each T_k. By a trivial computation we have that

$$Ker(T_1) = \langle (0) \rangle, \; Ker(T_2) = \langle \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rangle, \; Ker(T_3) = \langle \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ 1 \end{pmatrix} \rangle,$$

$$\langle \begin{pmatrix} 1/3 \\ 2/3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/3 \\ 2/3 \\ 1 \end{pmatrix} \rangle \in Ker(T_4) \text{ and } \langle \begin{pmatrix} 0 \\ 1/3 \\ 2/3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1/3 \\ 2/3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/3 \\ 2/3 \\ 1 \\ 0 \\ 0 \end{pmatrix} \rangle \in Ker(T_5)$$

It is clear that the lemma follows as if we start with $k = 1$ it is trivial that the zero vector will belong to all the subspaces $Ker(T_k)$ by taking the linear combinations with the coefficients equal to 0. If we start with $k = 3$, we have that $u_3 = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ 1 \end{pmatrix}$ and $(u_3, 0)^t$ belongs to T_4 and T_5. The way back works under the same idea.

The interest of the following theorem is in the assertion that if we have $T \in M_n(\mathbb{C})$, we can compute $u \in Ker(T)$ such that the columns of $T_d(u)$ form a basis of T (see Notation 6.7). Conversely, given $u \in \mathbb{C}^{m+1}$ and $d = n - m$ we can obtain $T \in M_n(\mathbb{C})$ such that $Ker(T) = T_d(u)$.

Theorem 6.10. ([5]) Let $T \in M_n(\mathbb{C})$ be a singular Toeplitz matrix and $d = dim(Ker(T))$. Then

(a) There exists a vector $u = (u_0, \dots, u_m)$ with $m = n - d$ such that the columns of $T_d(u) \in M_{m+d,d}(\mathbb{C})$ form a basis of $Ker(T)$ and the columns of $T_d(\widehat{u})$ form a basis of $Ker(T^t)$.

(b) The above vector u is unique up to a constant non-zero factor.

(c) Reciprocally, for a given non-zero vector $u \in \mathbb{C}^{m+1}$ and any $n > m$ there exists a Toeplitz matrix $T \in M_n(\mathbb{C})$ such that the columns of $T_d(u)$, $d = n - m$ form a basis of $Ker(T)$.

Proof.

(a) Observe that for $T = 0$, the result is trivial, so we can assume $T \neq 0$. Besides the matrix T, we consider the $(2n-k)\times k$ Toeplitz matrices T_k for $k = 1,\dots,2n-1$ defined in Lemma 6.8. Let m be the largest integer such that $Ker(T_m) = \{0\}$ and $Ker(T_{m+1}) \neq \{0\}$ (since $Ker(T_1) = \{0\}$, m is well defined). Observe that by Lemma 6.8 we have that

$$u \in Ker(T_{m+1}) \Leftrightarrow (u^t, 0)^t \in Ker(T_{m+2}) \Leftrightarrow \dots \Leftrightarrow$$

$$\Leftrightarrow (u, \underbrace{0,\dots,0}_{n-m}) \in Ker(T_{m+n-m}) = Ker(T_n) = Ker(T)$$

and we can assert that the columns of $T_{n-m}(u) \in Ker(T)$.
Let $J \in M_n(\mathbb{C})$ be the backward identity, clearly, $T^t = JTJ$ from this relation, we have that the columns of $T_{n-m}(\widehat{u})$ belong to $Ker(T^t)$.

(b) The uniqueness of u is consequence of $dim(Ker(T_{m+1})) = 1$.

(c) We will give the main ideas of the proof of this part. For simplicity, we assume that $u_m \neq 0$ and $u(\lambda)$ has only simple roots λ_i for $i = 1,\dots,m$. For $t \in \mathbb{C}$ we define the vector $\ell(t) = (t^{k-1})_{k=1}^{2n-1}$ and consider $a = (a_i)_{i=n-1}^{1-n}$ to be the columns of T. Notice that $T_{m+1}u = 0 \Leftrightarrow T_{2n-m-1}(u)^t a = 0$.
If we solve the second system and using the fact that $u(\lambda_i) = 0$, we have that the solutions can be expressed as

$$a = \sum_{i=1}^{m} c_i \ell(\lambda_i) \text{ with } c_i \neq 0$$

What is more, if we study the kernel of $T = (a_{i-j}) \in M_n(\mathbb{C})$, we have that

$$Tx = 0 \Leftrightarrow x(\lambda_i) = 0 \text{ for } i = 1,\dots,m$$

As a result, $x(\lambda)$ must be a multiple of $u(\lambda)$ as the zeros of $u(\lambda)$ are the zeros of $x(\lambda)$. Specifically, $Rank(T) = m$.

In conclusion, T is the desired Toeplitz matrix with kernel generating vector u and $dim(Ker(T)) = d = n - m$.

□

Definition 6.11. The vector u defined as in Theorem 6.10 is a *kernel generating vector* for $T \in M_n(\mathbb{C})$.

Example 6.12. Now we will provide an example of having a kernel generating vector and obtaining a Toeplitz matrix and vice versa in order to clarify the main concepts of the above proof.

- We will check a case where given a Toeplitz matrix of low dimension, we can directly obtain the kernel generating vector:
  Let $T = \begin{pmatrix} 3 & 4 & 5 \\ 2 & 3 & 4 \\ 1 & 2 & 3 \end{pmatrix} \in M_3(\mathbb{C})$ be the Toeplitz matrix which is singular. We have that

$d = dim(Ker(T)) = 1$ and $m = 2$ and therefore $T_1(u) = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}$. We have that

$$T_1 = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \; T_2 = \begin{pmatrix} 4 & 5 \\ 3 & 4 \\ 2 & 3 \\ 1 & 2 \end{pmatrix}, \; T_3 = T \; , \; T_4 = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 \end{pmatrix} \text{ and } T_5 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

In our example, $m = 2$ is the largest integer such that $Ker(T_2) = 0$ and $Ker(T_3) \neq 0$. We have that $Ker(T_3) = \langle(1, -2, 1)\rangle$. If $u \in Ker(T_3)$, then $(u, 0) \in Ker(T_4)$ and $(u, 0, 0) \in Ker(T_5)$.

- Reciprocally, we take a kernel generating vector and we obtain the Toeplitz matrix T:
  If we take $u = (1, 1, -2)$, $u(\lambda) = -2 + \lambda + \lambda^2$, $\lambda_1 = 1$, $\lambda_2 = -2$, $m = 2$, $n = 3$ and $d = 1$. We have to find a Toeplitz matrix $T \in M_3(\mathbb{C})$ such that the kernel are the columns of $T_1(u)$. Equivalently, we need to solve the system $T_3(u)^t a = 0$

$$\begin{pmatrix} 1 & 1 & -2 & 0 & 0 \\ 0 & 1 & 1 & -2 & 0 \\ 0 & 0 & 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} a_{-2} \\ a_{-1} \\ a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} a_{-2} + a_{-1} - 2a_0 \\ a_{-1} + a_0 - 2a_1 \\ a_1 + a_2 - 2a_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} a_{-2} \\ a_{-1} \\ a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -5a_1 + 6a_2 \\ 3a_1 - 2a_2 \\ -a_1 + 2a_2 \\ a_1 \\ a_2 \end{pmatrix}$$

  If we take $a_1 = 2$ and $a_2 = 1$, we get $T = \begin{pmatrix} 0 & 2 & 1 \\ 4 & 0 & 2 \\ -4 & 4 & 0 \end{pmatrix}$

In the next proposition we will present a Toeplitz matrix that will be essential for our concern, the study of $Ker(T) \cap Im(T)$.

Proposition 6.13. ([5]) For any $u \in \mathbb{C}^{m+1}$, the matrix defined by

$$R_d(u) = T_d(\widehat{u})^t T_d(u) \in M_d(\mathbb{C})$$

is a Toeplitz matrix.

Proof. It is straightforward to see that

$$T_d(\widehat{u})^t T_d(u) = \begin{pmatrix} u_m & \cdots & u_0 & 0 & \cdots & 0 \\ 0 & u_m & \cdots & u_0 & & \\ & & \ddots & & \ddots & \\ 0 & & & u_m & \cdots & u_0 \end{pmatrix} \begin{pmatrix} u_0 & & 0 \\ \vdots & u_0 & \\ u_m & \vdots & \ddots \\ & u_m & & u_0 \\ & & \ddots & \vdots \\ 0 & & & u_m \end{pmatrix} = \begin{pmatrix} t_0 & t_1 & t_2 & \cdots & t_m \\ t_{-1} & t_0 & t_1 & \ddots & \vdots \\ t_{-2} & \ddots & \ddots & \ddots & t_2 \\ \vdots & \ddots & \ddots & \ddots & t_1 \\ t_{-m} & \cdots & t_{-2} & t_{-1} & t_0 \end{pmatrix}$$

which is a Toeplitz with

$$t_j = \sum_{k=0}^{m-j} u_{j+k} u_{m-k} \quad \text{and} \quad t_{-j} = \sum_{k=0}^{m-j} u_k u_{m-j-k} \tag{6.0.2}$$

for $j = 0, \ldots, m$ and the remaining entries equal to zero. □

The next theorem will be our tool of investigation and our subsequent considerations rely on the following:

Theorem 6.14. ([5]) Let $T \in M_n(\mathbb{C})$ be a singular Toeplitz matrix, $u \in \mathbb{C}^{m+1}$ a kernel generating vector for T and $d = n - m = dim(Ker(T))$. Then

$$d_1 = dim(Ker(T) \cap Im(T)) = dim(Ker(R_d(u))) \tag{6.0.3}$$

Further to this, if $d_1 = dim(Ker(R_d(u))$ and v is a kernel generating vector for $R_d(u) \in M_d(\mathbb{C})$, then the columns of $T_{d_1}(w) \in M_{m+d_1,d_1}(\mathbb{C})$, with $w(\lambda) = u(\lambda)v(\lambda)$ form a basis of $Ker(T) \cap Im(T)$.

Proof. We know by Theorem 6.10 that $Ker(T) = Im(T_d(u))$. Let $x = T_d(u)\xi \in Ker(T)$, then

$$T_d(u)\xi \in Im(T) \;\Leftrightarrow\; T_d(u)\xi = Ty \;\Leftrightarrow\; z^t T_d(u)\xi = z^t Ty = 0 \;\; \forall z \in Ker(T^t) \;\Leftrightarrow$$

$$\Leftrightarrow T_d(\hat{u})T_d(u)\xi = R_d(u)\xi = 0$$

The last *if and only if* comes from the fact that the columns of $T_d(\hat{u})$ form a basis of $Ker(T^t)$. This gives the following bijection

$$Ker(R_d(u)) \to Ker(T) \cap Im(T)$$

$$\xi \mapsto T_d(u)\xi$$

that implies that $dim(Ker(T) \cap Im(T)) = dim(Ker(R_d(u)))$.
Regarding the second part of the theorem, we have that $Ker(R_d(u)) = Im(T_{d_1}(v))$ as $R_d(u)$ is a Toeplitz matrix (see Proposition 6.13), then we have that

$$\xi \in Ker(R_d(u)) \;\Rightarrow\; \xi = T_{d_1}(v)\xi_2 \;\Rightarrow\; T_d(u)\xi = T_d(u)T_{d_1}(v)\xi_2 \in Ker(T) \cap Im(T)$$

that follows directly that $T_{d_1}(w) = T_d(u)T_{d_1}(v)$ and, hence, the columns of $T_{d_1}(w)$ form a basis of $Ker(T) \cap Im(T)$. □

Corollary 6.15. The triple (n, d, d_1) is Toeplitz admissible if and only if there exists $u \in \mathbb{C}^{n-d+1}$ with $u \neq 0$ such that $d_1 = dim(Ker(R_d(u)))$.

Let

$$\mathscr{M}(n, d_0, d_1) = \{\, T \in M_n(\mathbb{C}) \,\text{Toeplitz} \mid d_0(T) = d_0,\; d_1(T) = d_1 \,\}$$

Notice that

$$\mathscr{M}(n, d_0, d_1) \neq \varnothing \;\Leftrightarrow\; (n, d_0, d_1) \text{ is Toeplitz admissible.}$$

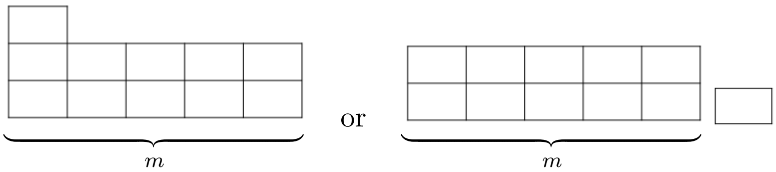
6.1 Matrices of odd order (n = 2m + 1)

In this section we will study the set $\mathscr{M}(2m+1, d_0, d_1)$. First, we will present Theorem 6.1.1, which asserts that all Toeplitz matrices that belong to $\mathscr{M}(2m+1, m, m)$ are similar to $diag(\underbrace{S_2, \ldots, S_2}_{m-1}, S_3) \in M_{2m+1}(\mathbb{C})$. Considering the theorem we will find the first counterexample that leads us to state that not every matrix is similar to a Toeplitz matrix.

Theorem 6.1.1. ([5]) For $m \geq 2$, any $T \in \mathscr{M}(2m+1, m, m)$ is similar to

$$diag(\underbrace{S_2, \dots, S_2}_{m-1}, S_3) \in M_{2m+1}(\mathbb{C}) \text{ with } S_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } S_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Proof. First, let us see $\mathscr{M}(2m+1, m, m)$ in a schematic way:



Let $T \in \mathscr{M}(2m+1, m, m)$ and u a kernel generating vector for T. Note that $d = dim(Ker(T)) = m$ and, considering Theorem 6.14, $d_1 = dim(Ker(R_m(u))) = m$ then, $dim(Im(R_m(u)) = 0$ and therefore $R_m(u) = 0$. By imposing it, we obtain the following system of nonlinear equations

$$\left.\begin{aligned} t_{1-m} &= 2u_0u_2 + u_1^2 = 0 \\ t_{2-m} &= 2u_0u_3 + 2u_1u_2 = 0 \\ t_{3-m} &= 2u_0u_4 + 2u_1u_3 + u_2^2 = 0 \\ &\vdots \\ t_0 &= 2u_0u_{m-1} + 2u_1u_m + \cdots = 0 \\ &\vdots \\ t_{m-2} &= 2u_{m-2}u_{m+1} + 2u_{m-1}u_m = 0 \\ t_{m-1} &= 2u_{m-1}u_{m+1} + u_m^2 = 0 \end{aligned}\right\} \tag{6.1.1}$$

On the one hand, we suppose $u_0 \neq 0$ and we may assume $u_0 = 1$ and we will show that $u_1 = 0$. Suppose that $u_1 \neq 0$ and define

$$x_0 = 1\,,\; x_k = \frac{u_k}{u_1^k} \quad \text{for } k = 2, \dots, m$$

From (6.1.1) we get the following recursion

$$x_0 = x_1 = 1 \;\;,\;\; x_2 = x_2 \text{ and } x_j = -\frac{1}{2}\sum_{k=1}^{j-1} x_k x_{j-k} \text{ for } j > 2 \tag{6.1.2}$$

Note that the system will exclusively depend on x_2. Studying the recursion, we can prove by induction that x_j are real numbers and that the following property is fulfilled

$$\left.\begin{aligned} x_j > 0 &\text{ if } j \text{ is odd} \\ x_j < 0 &\text{ if } j \text{ is even} \end{aligned}\right\} \tag{6.1.3}$$

Base case:
We have assumed that $x_1 = 1$. Let us now check the sign of x_2. We know that $x_2 = \frac{u_2}{u_1^2} = -\frac{1}{2} < 0$ where the last equality follows from $2u_2 + u_1^2 = 0$ (see Equation (6.1.1)).

Inductive step:
We assume the result for $i \leq n = 2m+1$. Let us check the sign for x_{n+1}.

$$x_{n+1} = -\frac{1}{2}\sum_{k=1}^{2m} x_k x_{2m+1-k}$$

At this point, we have to analyze what happens within the summation. If k is even, we have that $x_k x_{n+1-k} < 0$ and the same if k is odd, and so, $x_{n+1} > 0$.
Consequently, we have that

$$2x_{m-1}x_{m+1} + x_m^2 > 0$$

and this is a contradiction as the last equation of (6.1.1) will not be fulfilled. Thus, $u_1 = 0$ and using the recursion (6.1.2) we obtain that $u_j = 0$ for all $j = 2, \ldots, m+1$. According to our notation, this gives us that T is an upper triangular matrix and the first m columns are equal to 0.

$$T = \left(\begin{array}{c|c} \underbrace{0 \ \ldots \ 0}_{m} & T_U \end{array} \right) \in M_{2m+1}(\mathbb{C})$$

where T_U denotes an upper triangular Toeplitz matrix of dimension $(2m+1) \times (m+1)$. Note that T is nilpotent and similar to $diag(\underbrace{S_2, \ldots, S_2}_{m-1}, S_3)$.

If $u_0 = 0$, then by (6.2.11), $u_j = 0$ for all $j = 1, \ldots, m$ and $u_{m+1} \neq 0$. In this case we have that T is lower triangular and the m last columns of T are zero. In the same way as before, such matrix is nilpotent and similar to $diag(\underbrace{S_2, \ldots, S_2}_{m-1}, S_3)$.

$$T = \left(\begin{array}{c|c} T_L & \underbrace{0 \ \ldots \ 0}_{m} \end{array} \right) \in M_{2m+1}(\mathbb{C})$$

where T_L denotes an upper triangular Toeplitz matrix of dimension $(2m+1) \times (m+1)$. □

Example 6.1.2. Let us give an example of the above theorem for the case $n = 5$ and $m = 2$. Hence, we have that $T \in \mathscr{M}(5, 2, 2)$. We know that T is similar to

$$diag(S_2, S_3) = \begin{pmatrix} \boxed{\begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}} & \mathbf{0} \\ \mathbf{0} & \boxed{\begin{matrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{matrix}} \end{pmatrix} \in M_5(\mathbb{C})$$

and let us find T in an explicit form. Solving the system of equations (6.1.1) we obtain that $t_i = 0$ for $i = -4, \ldots, 0$ and $i = 4$. Finally, we have that

$$T = \begin{pmatrix} 0 & 0 & a & b & 0 \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{with } a \neq 0$$

Corollary 6.1.3. For $m \geq 2$, there is no Toeplitz matrix similar to

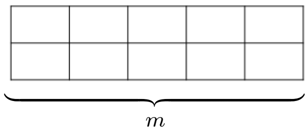
$$diag(\underbrace{S_2, \ldots, S_2}_{m}, c) \quad \text{if } c \neq 0$$

Proof. The tuple $(2m+1, m, m, 0)$, corresponding to the second figure of Theorem 6.1.1, is not Toeplitz admissible as a consequence the same theorem. □

Corollary 6.1.4. For any odd integer $n > 4$ there exists an $n \times n$ matrix that is not similar to a Toeplitz matrix.

6.2 Matrices of even order (n = 2m)

Once we have studied matrices of odd order, we study the even order case. For even n the problem is more complicated. In Theorem 6.1.1, we have studied the set $\mathscr{M}(2m+1,m,m)$. Following the proof of the mentioned theorem, we can sketch $\mathscr{M}(2m,m,m)$ as it follows:



and solving the system (6.1.1) for our case we obtain that any $T \in \mathscr{M}(2m+1,m,m)$ is similar to $diag(\underbrace{S_2,\dots,S_2}_{m})$. Consequently, $\mathscr{M}(2m,m,m)$ is Toeplitz admissible.

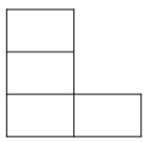
Our objective is now finding tuples that are not Toeplitz admissible with the following forms: $\mathscr{M}(2m,m,m-1)$ or $\mathscr{M}(2m,m-1,m-1)$. For this reason, we will divide our section into three parts: order 4, 6 and greater than 8. The last part consists of defining a recursion of polynomials and finding a counterexample based on them.

6.2.1 Case n = 4

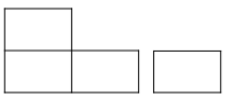
Notice that we have earlier studied this case in *Chapter 5* by studying the possible Jordan forms. Taking into consideration Theorem 5.1, we know that there exists $T \in M_4(\mathbb{C})$ Toeplitz matrices that are similar to $diag(S_2,0,c)$ with $c \neq 0$ or $diag(S_3,0)$. So, we have the statement is true for complex matrices but in this section we will see that the assertion is not true for real matrices.

Theorem 6.2.1. ([5]) The class $\mathscr{M}(4,2,1)$ does not contain real matrices. Furthermore, there is no Toeplitz matrix $T \in M_4(\mathbb{R})$ that is similar to $diag(S_2,0,c)$ with $c \neq 0$ or similar to $diag(S_3,0)$.

Proof. Before starting the rigorous proof, we sketch the Jordan blocks that are:



or



We suppose that $T \in \mathscr{M}(4,2,1)$ and $u = (u_0,u_1,u_2)^t$ is a kernel generating vector for T. Let $R_2(u)$ be the matrix defined in (6.13) as

$$R_2(u) = \begin{pmatrix} 2u_1u_2 & 2u_0u_2 + u_1^2 \\ 2u_0u_2 + u_1^2 & 2u_0u_1 \end{pmatrix} \in M_2(\mathbb{C})$$

Since $d_1 = dim(Ker(R_2(u))) = 1$, we have that

$$det(R_2(u)) = 4u_0^2u_2^2 + u_1^4 = 0 \tag{6.2.1}$$

If $u_i \in \mathbb{R}$, then combining this fact with (6.2.1) we have two possible options:

$$u_0 \neq 0, u_1 = 0, u_2 = 0 \;\; \Rightarrow \;\; u(\lambda) = u_0 \neq 0$$

or

$$u_0 = 0, u_1 = 0, u_2 \neq 0 \;\; \Rightarrow \;\; u(\lambda) = u_2\lambda^2 \neq 0$$

Following the construction of Proposition 6.13 and using that $d_0 = 2$ we have that T is a triangular matrix of the form

$$\begin{pmatrix} 0 & & & \\ 0 & 0 & & \\ a & 0 & 0 & \\ * & a & 0 & 0 \end{pmatrix} \text{ or } \begin{pmatrix} 0 & 0 & a & * \\ & 0 & 0 & a \\ & & 0 & 0 \\ & & & 0 \end{pmatrix} \quad \text{with } \; a \neq 0$$

If we compute the Jordan form of these matrices, we have that they are similar to $diag(S_2, S_2)$ and not to $diag(S_2, 0, c)$ with $c \neq 0$ or $diag(S_3, 0)$. □

6.2.2 Case n = 6

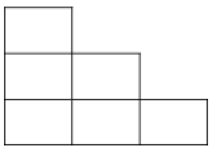
In this section, we will find a triple $(6, d_0, d_1)$ for which the class $\mathscr{M}(6, d_0, d_1) = \varnothing$. Additionally, we find a nilpotent matrix that is not similar to a Toeplitz matrix.

Theorem 6.2.2. ([5]) $\mathscr{M}(6,3,2) = \varnothing$, in other words, there is no $T \in M_6(\mathbb{C})$ Toeplitz matrix that is similar to

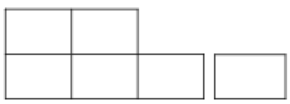
$$diag(S_3, S_2, 0) \qquad or \qquad diag(S_2, S_2, 0, c)$$

with $c \neq 0$.

Proof. We have the following possible Young diagrams:



or



We suppose that $T \in \mathscr{M}(6,3,2)$, then

$$R_3(u) = \begin{pmatrix} 2u_1u_2 + 2u_0u_3 & u_1^2 + 2u_0u_2 & 2u_0u_1 \\ u_2^2 + 2u_1u_3 & 2u_1u_2 + 2u_0u_3 & u_1^2 + 2u_0u_2 \\ 2u_2u_3 & u_2^2 + 2u_1u_3 & 2u_1u_2 + 2u_0u_3 \end{pmatrix} = \begin{pmatrix} t_0 & t_1 & t_2 \\ t_{-1} & t_0 & t_1 \\ t_{-2} & t_{-1} & t_0 \end{pmatrix}$$

and $n = 6$, $m = 3$ and $d = 3$. Observe that in this case we have that $d_1 = dim(Ker(R_3(u))) = 2$ and $dim(Im(R_3(u))) = 1$. Now, we distinguish different cases:

- Case $t_0 \neq 0$ ($\Leftrightarrow \;\; 2u_1u_2 + 2u_0u_3 \neq 0$)
  Considering Theorem 6.10, we know that u and $R_3(u)$ are defined up to non-zero constant factor, then we can assume $t_0 = 2$. In addition, as $d_1 = dim(Ker(R_3(u))) =$

2, we have that $dim(Im(R_3(u))) = 1$ so all the minors of dimension 2×2 are 0. In particular, $t_0^2 - t_2t_{-2} = 0$, and therefore we can take

$$u_0u_1u_2u_3 = 1 \text{ and } u_0u_3 + u_1u_2 = 1 \tag{6.2.2}$$

At first glance, we see that $u_k \neq 0$ for $k = 0,1,2,3$. We do a variable change $u_0u_3 = \gamma$, and $u_1u_2 = \gamma^{-1}$. Then, we have that

$$\gamma\gamma^{-1} = 1\ ,\ \ \gamma + \gamma^{-1} = 1 \Rightarrow \gamma = \frac{1\pm\sqrt{3}i}{2}$$

We continue using the relationship given by the minors 2×2. We now use two relations $t_1^2 = t_0t_2$ and $t_{-1}^2 = t_0t_{-2}$ that drive to

$$4u_0u_2u_3^2 = 4u_1^2u_3^2 + u_2^4 \text{ and } 4u_0^2u_1u_3 = 4u_0^2u_2^2 + u_1^4 \tag{6.2.3}$$

substituting we have

$$4\gamma^3u_2^3u_3 = 4u_3^2 + \gamma^2u_2^6 \text{ and } 4\gamma^3u_2^3u_3 = 4\gamma^4u_2^6 + \gamma^{-2}u_3^2$$

Then by substituting the first expression into the second we have

$$u_3^2 = \frac{\gamma^4(4-\gamma^{-2})}{4-\gamma^{-2}}u_2^6 = \gamma^4u_2^6$$

Hence,

$$u_3 = \pm\gamma^2u_2^3 \tag{6.2.4}$$

Inserting (6.2.4) into the first equality of (6.2.3) we get that

$$\pm 4\gamma^3 = \mp 4 = 4\gamma^2 + 1$$

and this equality is clearly false, so we have a contradiction.

- Case $t_0 = 0$ ($\Leftrightarrow\ 2u_1u_2 + 2u_0u_3 \neq 0$)
  In this case

$$R_3(u) = \begin{pmatrix} 0 & u_1^2 + 2u_0u_2 & 2u_0u_1 \\ u_2^2 + 2u_1u_3 & 0 & u_1^2 + 2u_0u_2 \\ 2u_2u_3 & u_2^2 + 2u_1u_3 & 0 \end{pmatrix} = \begin{pmatrix} t_0 & t_1 & t_2 \\ t_{-1} & t_0 & t_1 \\ t_{-2} & t_{-1} & t_0 \end{pmatrix}$$

  We proceed as before using the fact that $dim(Im(R_3(u)) = 1$. We only have two possible situations:

  - $t_2 \neq 0 \Rightarrow t_1 = t_{-1} = t_{-2} = 0 \Rightarrow u_1^2 = -2u_0u_1,\ u_2^2 = -2u_1u_3$ and $2u_2u_3 = 0$.
    We can have two possible situations: $u_2 = 0$ or $u_3 = 0$.
    * $u_2 = 0 \Rightarrow u_1 = 0 \Rightarrow R_d(u) = 0$ (!)
    * $u_3 = 0 \Rightarrow u_2 = 0 \Rightarrow u_1 = 0 \Rightarrow R_3(u) = 0$ (!)
  - $t_{-2} \neq 0 \Rightarrow t_1 = t_{-1} = t_2 = 0 \Rightarrow \Rightarrow u_1^2 = -2u_0u_1,\ u_2^2 = -2u_1u_3$ and $2u_0u_1 = 0$.
    Hence, we can have $u_0 = 0$ or $u_1 = 0$.
    * $u_0 = 0 \Rightarrow u_1 = 0 \Rightarrow u_2 = 0 \Rightarrow R_d(u) = 0$ (!)
    * $u_1 = 1 \Rightarrow u_2 = 0 \Rightarrow u_0 = 0 \Rightarrow R_d(u) = 0$ (!)

In conclusion, we have proved that $dim(Im(R_3(u))) \neq 1$, that is $d_1 \neq 2$ which proves the theorem. □

The case $n = 6$ arises the following question: is $\mathscr{M}(2m, m, m-1) = \varnothing$ for all $m \geq 3$? We will be able to answer the problem in the next section.

6.2.3 Case n ≥ 8

For higher dimensions, the problem is reduced to the class of polynomials that can be checked for special cases, defined recursively as it follows:

$$p_0(t) = p_1(t) = 1\ ,\ \ p_2(t) = t \ \text{ and } \ p_j(t) = \frac{-1}{2}\sum_{k=1}^{j-1} p_k(t)p_{j-k}(t) \qquad (j > 2) \tag{6.2.5}$$

We will study this recursion as will be the fundamental point of this section to check if there is always a similarity to a Toeplitz matrix $T \in M_n(\mathbb{C})$ for $n \geq 8$. The study of this polynomial recursion will be based on three lemmas and one main theorem.

Lemma 6.2.3. ([3]) The generating function of the family of polynomials $\{p_j(t)\}$ which is

$$p(z,t) = \sum_{j=0}^{\infty} p_j(t)z^j$$

is given by

$$p(z,t) = \sqrt{1 + 2z + z^2(1 + 2t)} \tag{6.2.6}$$

Proof. According to the definition of $p_j(t)$ (in (6.2.5)), we have

$$\sum_{i+k=j} p_i(t)p_k(t) = 0 \ \text{ for } \ j > 2$$

Moreover,

$$p(z,t)^2 = \sum_{j=0}^{\infty}\sum_{i+k=j}(p_i(t)p_k(t))z^j = \sum_{j=0}^{2}(\sum_{i+k=j} p_i(t))z^j = A(t) + B(t)z + C(t)z^2$$

It is straightforward by the definition of $p_j(t)$ for $j = 0, 1, 2$ that

$$A(t) = p_0(t) = 1$$

$$B(t) = p_0(t)p_1(t) + p_1(t)p_0(t) = 2$$

$$C(t) = p_0(t)p_2(t) + p_1(t)p_1(t) + p_2(t)p_0(t) = t + 1 + 1 = 2t + 1$$

We isolate $p(z,t)$ and we obtain the desired result. □

Lemma 6.2.4. ([3]) Let $p_j(t)$ as in (6.2.5), then

$$p_j(t) = \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} 2^{j-2k}\binom{1/2}{j-k}\binom{j-k}{k}(2t+1)^k$$

Proof. The basic idea of the proof is using the binomial expansion twice.

$$p(z,t) = \sqrt{1 + 2z + z^2(1 + 2t)} = (1 + 2z + z^2(1 + 2t))^{1/2} = \sum_{j=0}^{\infty}\binom{\frac{1}{2}}{j}(2z + z^2(1 + 2t))^j \tag{6.2.7}$$

and

$$(2z + z^2(1 + 2t))^j = \sum_{k=0}^{j}\binom{j}{k}(2z)^{j-k}(z^2(1 + 2t))^k \tag{6.2.8}$$

Thus, combining (6.2.7) and (6.2.8) we have that

$$p(z,t) = \sum_{j=0}^{\infty}\sum_{k=0}^{j}\binom{1/2}{j}\binom{j}{k}2^{j-k}z^{j-k}z^k(1+2t)^k = \sum_{j=0}^{\infty}\sum_{k=0}^{j}\binom{1/2}{j}\binom{j}{k}2^{j-k}(1+2t)^k z^j = \sum_{j=0}^{\infty}p_j(t)z^j$$

□

Lemma 6.2.5. ([3]) The polynomials $p_j(t)$ satisfy the 3-term recursion

$$(j+2)p_{j+2}(t) + (2j+1)p_{j+1}(t) + (j-1)(2t+1)p_j(t) = 0\ , \quad j \geq 0 \tag{6.2.9}$$

Proof. Let $h(z,t)$ be the generating function of the polynomial family $\{p_j(t)\}$ defined in (6.2.9). Let h_z denote the partial derivative of $h(z,t)$ by z and $h(z,t) = h$. Then, we show that $h(z,t) = p(z,t)$.

$$\sum_{j=0}^{\infty}(j+2)p_{j+2}z^{j+1} = h_z - 1$$

$$\sum_{j=0}^{\infty}(2j+1)p_{j+1}z^{j+1} = 2zh_z - h + 1$$

$$\sum_{j=0}^{\infty}(j-1)p_j z^{j+1} = z^2h_z - zh$$

Summing the three equations we obtain the ordinary differential equation

$$(1+2z+(2t+1)z^2)h_z - (1+(2t+1)z)h = 0 \tag{6.2.10}$$

It is straightforward that $p(z,t)$ satisfies (6.2.10). Since $p(0,t) = h(0,t)$, we conclude that $p(z,t) = h(z,t)$. □

This theorem is essential for the further part of this work as all the next considerations strongly depend on it.

Theorem 6.2.6. ([3]) For $m > 1$ $p_{m+1}(t) = p_m(t) = 0$ has only a trivial solution $t = 0$.

Proof. We will use induction to prove this result.
Base case, $m = 2$: By computing $p_2(t) = t$ and $p_3(t) = -t$, it is evident that the only solution of $p_2(t) = p_3(t)$ is $t = 0$.

Inductive step: We assume the theorem is valid for $m > 1$, then we prove that it is true for $m+1$.
We will prove it by contraposition. We suppose $p_{m+1}(\tau) = p_{m+2}(\tau) = 0$ for some $\tau \neq 0$. From Lemma 6.2.5 we have $(m-1)(2\tau+1)p_m(\tau) = 0$ then $\tau = -\frac{1}{2}$ and applying again the lemma we have

$$(m+1)p_{m+1}(\tau) + (2m-1)p_m(\tau) + (m-2)(2\tau+1)p_{m-1}(\tau) = 0 \Rightarrow p_m(-\frac{1}{2}) = 0 = p_{m+1}(-\frac{1}{2})$$

and we have a contradiction. □

Corollary 6.2.7. ([5]) For $m > 3$, the system of $m - 2$ equations

$$p_{m+2}(t) = p_{m+3}(t) = \ldots = p_{2m-1}(t) = 0$$

has only the trivial solution $t = 0$.

As soon as defined these polynomials, we are now ready for the study of the even order case for $n \geq 8$.

Theorem 6.2.8. ([5]) For $m \geq 4$, $\mathscr{M}(2m, m, m-1) = \varnothing$, precisely, there is no Toeplitz matrix that is similar to

$$diag(\underbrace{S_2, \ldots, S_2}_{m-1}, 0, c) \quad \text{or} \quad diag(S_3, \underbrace{S_2, \ldots, S_2}_{m-2}, 0)$$

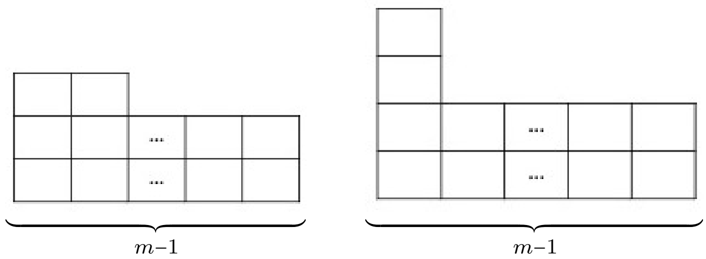
if $c \neq 0$.

The proof of this theorem is based on the following lemma that describes the class of $\mathscr{M}(2m, m-1, m-1)$.

Lemma 6.2.9. ([5]) Let $T \in \mathscr{M}(2m, m-1, m-1)$ and $u = (u_j)_{j=0}^{m+1}$ be a kernel generating vector for T. Then

$$u = (u_0, u_1, 0, \ldots, 0) \;\text{ or }\; u = (0, \ldots, 0, u_m, u_{m+1})$$

Proof. Firstly, we sketch part of the possible Jordan forms (in particular, the nilpotent ones):



Considering Theorem 6.10, we have that $d_1 = dim(Ker(R_{m-1}(u)) = m-1 \Rightarrow R_{m-1}(u) = 0$ by Theorem 6.14. So, we have a system of $2m - 3$ nonlinear equations

$$\left.\begin{aligned} t_{2-m} &= 2(u_0u_3 + u_1u_2) = 0 \\ &\vdots \\ t_0 &= 2u_0u_{m-1} + 2u_1u_m + \cdots = 0 \\ &\vdots \\ t_{m-2} &= 2u_{m-2}u_{m+1} + 2u_{m-1}u_m = 0 \end{aligned}\right\} \tag{6.2.11}$$

We shall consider different cases:

- $u_0 = u_1 = 0 \;\Rightarrow\; u_2 = \cdots = u_{m-1} = 0$, which is our assertion.
- $u_0 = 0$ and $u_1 \neq 0 \;\Rightarrow\; u_2 = \cdots = u_{m+1} = 0$.

- If $u_0 \neq 0$, we can suppose $u_0 = 1$. The first $m-1$ equations of (6.2.11) can be rewritten as a recursion

$$u_j = -\frac{1}{2}\sum_{k=1}^{j-1} u_k u_{j-k} \quad (j = 3, \dots m+1) \tag{6.2.12}$$

Note that we can see u_j as a polynomial in u_2. If we define $u_{m+2}, \dots, u_{2m-1}$ according to the recursion, then the last $m-2$ equations of (6.2.11) can be written as

$$u_{m+1} = \dots = u_{2m-1} = 0 \tag{6.2.13}$$

Now, we can distinguish two sub-cases:

  - $\underline{u_1 = 0}$. In this case, we obtain that

$$u_j = \begin{cases} 0 & \text{if } j \text{ is odd} \\ \alpha_j u_2^{j/2} & \text{if } j \text{ is even, where } (-1)^j \alpha_j > 0 \end{cases}$$

    Hence, the recursion defined in (6.2.13) is only true if $u_2 = 0$, which means that $u_3 = \dots = u_{m+1} = 0$.
  - $\underline{u_1 \neq 0}$. We can define

$$\left.\begin{aligned} x_k &= \frac{u_k}{u_1^k} \\ x_2 &= t = u_2 \end{aligned}\right\} \tag{6.2.14}$$

    Combining Equations (6.2.12) and (6.2.14) we have the following recursion

$$x_j = -\frac{1}{2}\sum_{k=1}^{j-1} x_k x_{j-k} \;\;,\;\; x_0 = 1 \;\;,\;\; x_1 = 1 \;\;,\;\; x_2 = t$$

    If we consider Corollary (6.2.7) and Equation (6.2.13), we obtain that $t = 0 = u_2$, which implies that $u_3 = \dots = u_{m+1} = 0$ and this proves the lemma.

□

Proof of Theorem 6.2.8. Let $T \in \mathscr{M}(2m, m, m-1)$ and $u = (u_j)_{j=0}^m$ a kernel generating vector for T. We have that $dim(Ker(R_m(u))) = m-1$. According to Theorem 6.10, $R_m(u)$ has a kernel generating polynomial of degree 1, this is $v(\lambda) = x_0 + x_1\lambda$. Hence, $R_m(u)T_{m-1}(v) = 0$, from which we can see that

$$0 = T_{m-1}(\hat{v})^t R_m(u) T_{m-1}(v) = T_{m-1}(\hat{v})^t T_m(\hat{u})^t T_m(u) T_{m-1}(v) =$$
$$= T_{m-1}(\hat{w})^t T_{m-1}(w) = R_{m-1}(w)$$

The last equality comes from the fact that $T_m(u)T_{m-1}(v) = T_{m-1}(w)$. Finally, since $R_{m-1}(w) = 0$ and because of Lemma 6.2.9, we have that

$$w = (w_0, w_1, 0, \dots, 0) \text{ or } w = (0, \dots, 0, w_m, w_{m+1})$$

Let us assume we have the first case (then, the second case is analogous): $w(\lambda) = u(\lambda)v(\lambda)$ and $deg(v(\lambda)) = 1$, therefore

$$deg(w(\lambda)) = 1 \Rightarrow deg(u(\lambda)) = 0, \; u = (u_0, 0, \dots, 0)$$

We have that $u(\lambda)$ must be constant and we obtain that $T \sim diag(S_2, \dots, S_2)$. That means that $T \in \mathscr{M}(2m, m, m)$ and $T \notin \mathscr{M}(2m, m, m-1)$, and this concludes the proof. □

Bearing in mind the theorem, we can conclude the following fact:

Corollary 6.2.10.

$$\mathscr{M}(n, d_0, d_1) \neq \varnothing \;\; \Rightarrow \;\; \mathscr{M}(n, d_1, d_1) \neq \varnothing$$

Proof. Let $T \in \mathscr{M}(n, d_0, d_1)$, u be a kernel generating vector for T, v be a kernel generating vector for $R_{d_0}(u)$ and $w(\lambda) = u(\lambda)v(\lambda)$. Then

$$R_{d_1}(w) = T_d(\hat{v})^t R_{d_0}(u) T_d(v) = 0$$

Therefore, any Toeplitz matrix $T' \in M_n(\mathbb{C})$ with $dim(Ker(T')) = d_1$ and a kernel generating vector w, belongs to $\mathscr{M}(n, d_1, d_1)$. □

Corollary 6.2.11. For any $n > 4$ there exists an $n \times n$ matrix that is not similar to a Toeplitz matrix.

Chapter 7

Conclusions

In regard of the objectives of this project that were set at the beginning, we can state that they have been reasonably satisfied and fulfilled.

From my personal point of view, I can ensure that working with this paper has exceeded my expectations. Not only I have worked on one of my favorite subjects of the degree, but I have had the opportunity to introduce myself to read advanced essays of many topics related to the work and realizing that I have the sufficient skills to understand high-level mathematical reports.

Furthermore, the fact of having to rewrite articles [4] and [5] has forced me to learn how to make formal definitions, how to write organized and rigorous proofs, understand deeply every detail written in the articles and find the way to explain it more understandable for the reader. Further to this, at first, I was terrified by the fact of having to write the work in TeX but I got used to it very fast and now it feels more familiar than Word.

Concerning the technical part of the project, the objectives have been met as we have studied in depth the similarities to a Toeplitz in the case of diagonalizable and nonderogatory matrices. Moreover, we have drawn an strategy to find counterexamples to reject our initial question by essentially studying the structure of a Kernel of a Toeplitz matrix and the Toeplitz admissibility of the tuples.

Notwithstanding that we have answered the question, there are many areas of this topic that remain open and due to the short amount of time we have had, we have not been able to study them. I will list the most important, in my point of view, unsolved topic: the study in an arbitrary field $\mathbb{F}$. For example, those that do not have a Jordan form.

Finally, for further reading, I recommend a book that two of the writers of article [4], D. Steven Mackey and Niloufer Mackey, sent me which recaps all the studies related to Toeplitz matrices of the mathematician that has studied most these matrices: Georg Heinig, Dario Andrea Bini, Fabio Di Benedetto, Eugene Tyrtyshnikov, Marc Van Barel, *Structured Matrices in Numerical Linear Algebra.* Volume 30. Springer, 2019.

Bibliography

[1] Aubanell, A.; Benseny, A. ; Delshams, A. *Eines bàsiques de càlcul numèric: amb 87 problemes resolts.* Bellaterra : Publicacions de la Universitat Autònoma de Barcelona, 1991.

[2] Chu M T and Golub G H, *Inverse Eigenvalue Problems: Theory, Algorithms, and Applications.* Oxford University Press, 2005.

[3] Dario Andrea Bini, Fabio Di Benedetto, Eugene Tyrtyshnikov, Marc Van Barel, *Structured Matrices in Numerical Linear Algebra.* Volume 30. Springer, 2019.

[4] D. Steven Mackey, Niloufer Mackey, and Srdjan Petrovic, Is every matrix similar to a Toeplitz matrix? Linear Algebra Appl. 297:87–105, 1999.

[5] G. Heinig, *Not every matrix is similar to a Toeplitz matrix.* Linear Algebra Appl., 332–334:519–531, 2001.

[6] G. Heinig, K. Rost, *Algebraic methods for Toeplitz-like matrices and operators, in: Operator Theory: Advances and Applications*, vol. 13, Birkhauser, Basel, 1984.

[7] G.H. Golub, C. Van Loan *Matrix Computations*, 2nd edition, Baltimore, 1989.

[8] Nikolskii, S. M. *Toeplitz matrices and operators* Cambridge University Press, 2020.

[9] M. Gray *Toeplitz and Circulant matrices: A review.* Standford University.

[10] P.J. Davis, *Circulant matrices*, Wiley, New York, 1979.

[11] P. Delsarte, Y.Genin, *Spectral properties of finite Toeplitz matrices*, vol. 58, Springer, New York, 1984.

[12] R.A. Horn, C.R. Johnson *Matrix Analysis*, Cambridge University Press, New York, 1990.

[13] R.A. Horn, C.R. Johnson *Topics in matrix analysis*, Cambridge University Press, New York, 1991.

[14] S. Nicoloso and U. Pietropaoli, *Coloring Toeplitz graphs.* Manuscript, 2010.

[15] Sheng Bau, *A generalization of the concept of Toeplitz graphs* Center for Discrete Mathematics, Fuzhou University, Fuzhou, China.

[16] Teodoro Lara, *Matrices circulantes.* Divulgaciones Matemáticas Vol. 9 No. 1(2001), pp. 85–10. Venezuela.

[17] T H Jones, N B Willms, *Inverse eigenvalue problems for checkerboard Toeplitz matrices*, IOP Conf. Series: Journal of Physics: Conf. Series 1047, 2018.

[18] U. Grenander and G. Szegö, *Toeplitz forms and their applications.* University of California, Berckley and Los Angeles.