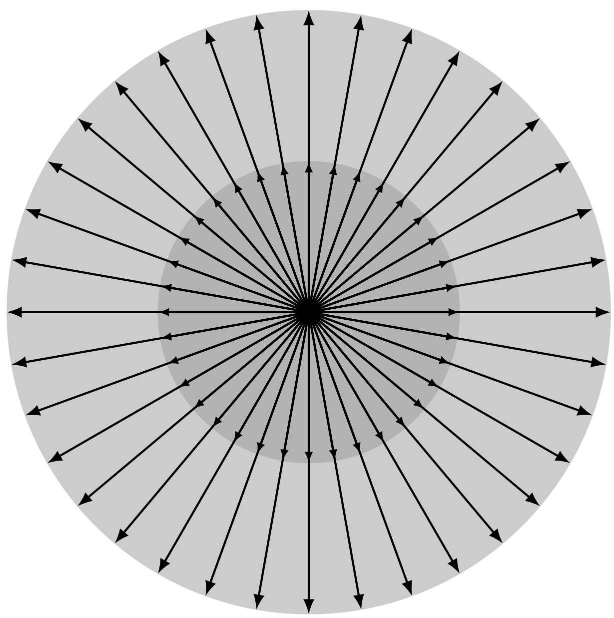
Linear Algebra II

Konstantinos Bizanos



Comprehensive Notes and Exercises

CONTENTS

Preface **7**

1 Similar Matrices **9**

1.1 Definition and properties 9

1.2 Relation to linear maps 12

1.3 Chapter 1 Exercises. 14

2 Polynomials **17**

2.1 Divisibility 17

2.2 Irreducible Polynomials 21

2.3 Polynomials and matrices 24

2.4 Polynomials and linear maps 24

2.5 Chapter 2 Exercises. 26

3 Eigenvalues and Eigenvectors **31**

3.1 Eigenvalues, eigenvectors and eigenspaces of a matrix 31
3.1.1 Eigenvalues and eigenvectors of a matrix 31
3.1.2 Eigenspaces of a matrix 35
3.1.3 Properties of eigenspaces 36
3.2 Eigenvalues, eigenvectors and linear maps 39
3.3 Characteristic polynomial 44
3.4 Chapter 3 Exercises. 50

4 Diagonalizability **59**
4.1 Diagonalizable Matrices 59
4.2 The Major Diagonalizability Criterion 64
4.3 Diagonalizable Linear Maps 68
4.4 Applications of Diagonalization 70
4.5 Exercises of Chapter 4. 74

5 Triangularizable Matrices **83**
5.1 Triangularizable Matrices 83
5.2 Triangularizable Linear Maps 89
5.3 Cayley-Hamilton Theorem 91
5.4 Exercises of Chapter 5. 94

6 Minimal Polynomial **101**
6.1 Minimal Polynomial 101
6.2 Diagonalizability Criterion 107

6.3 Minimal Polynomial of a Linear Map . . . 108
6.4 Simultaneous Diagonalization . . . 109
6.4.1 Invariant Subspaces . . . 109
6.4.2 Simultaneous Diagonalization . . . 111
6.5 Exercises of Chapter 6. . . . 113

7 The Standard Inner Product on $\mathbb{R}^n$ and $\mathbb{C}^n$ **121**
7.1 The Standard Inner Product . . . 121
7.2 Orthonormal Bases . . . 124
7.2.1 Orthonormal bases and the Gram–Schmidt method . . . 124
7.2.2 Orthogonal complement of a subspace of $\mathbb{F}^n$ . . . 128
7.3 Hermitian and Unitary Matrices . . . 131
7.3.1 The matrix A^*, the standard inner product, and matrix multiplication . . . 131
7.3.2 Hermitian Matrices . . . 132
7.3.3 Unitary Matrices . . . 134
7.4 Exercises of Chapter 7. . . . 138

8 Normal Matrices **143**
8.1 Schur's Lemma . . . 143
8.2 Spectral Theorem . . . 144
8.3 Normal Matrices . . . 145
8.4 Exercises of Chapter 8. . . . 151

PREFACE

These notes concern the undergraduate course Linear Algebra II. The content of the notes has been based on the material taught in the undergraduate curriculum of the Department of Mathematics of the National and Kapodistrian University of Athens. The notes constitute an aid for examined (and non-examined) students; however, it must be stated that under no circumstances can these notes replace any corresponding textbook of this subject area. At the end of each chapter there are also practice exercises for the students, with which it is recommended to engage for the optimal understanding of the course material.

After the completion of the theory, the reader can also find solved examination problems from the examinations of the Department of Mathematics of the National and Kapodistrian University of Athens, in order to become familiar with the type of exercises that are usually asked in the examinations.

Finally, it is clear that the notes will contain typographical (and not only) errors, so if you notice mistakes you may point them out at the e-mail: kbfg6@umsystem.edu

Konstantinos Bizanos,
Missouri,
December 26, 2025

CHAPTER 1

SIMILAR MATRICES

1.1 Definition and properties

Definition 1.1.1. Let $A, B \in \mathbb{F}^{n\times n}$. We will say that A and B are **similar** if there exists an invertible matrix P such that $B = P^{-1}AP$.

Observation 1.1.1. Matrix similarity is an equivalence relation.

Proof. Let $A, B, C \in \mathbb{F}^{n\times n}$.

i. A is similar to itself, since: $A = \mathbb{I}_n^{-1}A\mathbb{I}_n$.

ii. If A and B are similar, then B and A are also similar. Indeed, there exists an invertible $P \in \mathbb{F}^{n\times n}$ such that

$$B = P^{-1}AP \Leftrightarrow PB = P(P^{-1}AP) \Leftrightarrow A = PBP^{-1}.$$

iii. If A and B are similar, and also B and C are similar, then A and C are similar. Indeed, there exist invertible matrices P, Q such that

$$B = P^{-1}AP$$

and

$$C = Q^{-1}BQ.$$

Then:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (Q^{-1}P^{-1})A(PQ) = (PQ)^{-1}A(PQ),$$

hence A and C are similar.

□

Example 1.1.1. Let the matrices

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}.$$

We observe that $B = P^{-1}AP$, where

$$P = \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix},$$

therefore the matrices A and B are similar.

Example 1.1.2. Let the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

The matrices A and B are not similar. Indeed, if there existed an invertible matrix P such that:

$$B = P^{-1}AP = P^{-1}\mathbb{I}_2P = P^{-1}P = \mathbb{I}_2,$$

we arrive at a contradiction.

Observation 1.1.2. If a matrix B is similar to $\mathbb{I}_n$, then $B = \mathbb{I}_n$.

Proof. The conclusion follows immediately from Example 1.1.2. □

Reminder 1.1.1. Let $A \in \mathbb{F}^{n\times n}$. The following three integers coincide (that is, they define the *rank* of the matrix A, i.e. $\text{rank}(A)$):

(i) The maximum number of linearly independent columns.

(ii) The maximum number of linearly independent rows.

(iii) The $\dim(\text{Im}(\mathcal{L}_A))$, where $\mathcal{L}_A : \mathbb{F}^{n\times 1} \to \mathbb{F}^{n\times 1}$, $X \mapsto AX$.

Proposition 1.1.1. Let A and B be similar. Then the following hold:

i. $\det(A) = \det(B)$,

ii. $\operatorname{tr} A = \operatorname{tr} A$

iii. $\operatorname{rank} A = \operatorname{rank} B$.

Proof. i. Since A and B are similar, there exists an invertible matrix P such that $B = P^{-1}AP$. Then:

$$\begin{aligned} \det B &= \det(P^{-1}AP) \\ &= \det(P^{-1}) \cdot \det(A) \cdot \det(P) \\ &= \frac{1}{\det(P)} \cdot \det(A) \cdot \det(P) \\ &= \det A \end{aligned}$$

ii. Since A is similar to B, there exists an invertible P such that

$$B = P^{-1}AP$$

Then, from the property

$$\operatorname{tr} AB = \operatorname{tr} BA$$

we have

$$\operatorname{tr} B = \operatorname{tr}\left(P^{-1}AP\right) = \operatorname{tr}\left(APP^{-1}\right) = \operatorname{tr} A.$$

iii. The matrices A and B are similar, hence they are also equivalent, so $\operatorname{rank}(A) = \operatorname{rank}(B)$.

□

Attention! The converse does not hold in general.

Example 1.1.3. For example, the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

are not similar, even though they have equal determinant and equal rank: $\det(A) = \det(B)$ and $\operatorname{rank}(A) = \operatorname{rank}(B)$.

1.2 Relation to linear maps

Question. How do similar matrices appear in nature?

An informal answer is that similar matrices arise from changes of bases of linear maps. The precise answer is given via the following theorem. But first let us recall some useful tools.

Reminder 1.2.1. i. Let $\hat{a}$ be a basis of V, and let $P \in \mathbb{F}^{n\times n}$ be invertible. Then there exists a basis $\hat{b}$ of V with

$$P = \left(1_V : \hat{b}, \hat{a}\right).$$

ii. Let $\hat{a}, \hat{b}$ be bases of V. Then we have

$$\left(1_V : \hat{b}, \hat{a}\right)^{-1} = \left(1_V : \hat{a}, \hat{b}\right).$$

iii. Let $f : U \to V$, $g : V \to W$ be linear maps and let $\hat{a}, \hat{b}, \hat{c}$ be ordered bases of U, V, W respectively. Then:

$$(g \circ f : \hat{a}, \hat{c}) = (g : \hat{b}, \hat{c}) \cdot (f : \hat{a}, \hat{b}).$$

Proof. Indicatively we will prove (i). Since

$$\hat{a} = (a_1, \ldots, a_n)$$

is a basis of V, there exists a linear map $f : V \to V$ with

$$P = (f : \hat{a}, \hat{a}).$$

Since P is invertible, f is an isomorphism, hence

$$f(a_1), \ldots, f(a_n)$$

form a basis of V. Set

$$\hat{b} = (f(a_1), \ldots, f(a_n))$$

and from the definition of the matrix of a linear map we have

$$P = \left(1_V : \hat{b}, \hat{a}\right).$$

□

Theorem 1.2.1. Let $f : V \to V$ be a linear map, let $\hat{v}$ be an ordered basis of V and let $A \in \mathbb{F}^{n\times n}$ with

$$A = (f : \hat{v}, \hat{v}).$$

Let $B \in \mathbb{F}^{n\times n}$. Then the following are equivalent:

i. The matrices A and B are similar.

ii. There exists an ordered basis $\hat{w}$ of V such that $B = (f : \hat{w}, \hat{w})$.

Proof.

- i. $\Rightarrow$ ii.: Let P be invertible with

$$B = P^{-1}AP.$$

Then, by Reminder 1.2, there exists an ordered basis $\hat{w}$ such that

$$P = (1_V : \hat{w}, \hat{v}).$$

Then:

$$B = (1_V : \hat{v}, \hat{w}) \cdot (f : \hat{v}, \hat{v}) \cdot (1_V : \hat{w}, \hat{v}) = (1_V \circ f \circ 1_V : \hat{w}, \hat{w}) = (f : \hat{w}, \hat{w}).$$

- ii. $\Rightarrow$ i.: Let

$$B = (f : \hat{w}, \hat{w})$$

for some ordered basis $\hat{w}$ of V. We will show that A and B are similar. Set $P = (1_V : \hat{w}, \hat{v})$. Then:

$$B = (f : \hat{w}, \hat{w}) = (1_V : \hat{v}, \hat{w}) \cdot (f : \hat{v}, \hat{v}) \cdot (1_V : \hat{w}, \hat{v}) = P^{-1}AP,$$

hence the matrices A and B are similar.

□

1.3 Chapter 1 Exercises.

Group A : 1,2,4,5,6,7 **Group B** : 3

Exercise 1.1. Let $\lambda \in \mathbb{F}$ and $A \in \mathbb{F}^{n\times n}$. Show that if A is similar to $\lambda\mathbb{I}_n$, then $A = \lambda\mathbb{I}_n$.

Exercise 1.2. Let $A, B \in \mathbb{F}^{n\times n}$.

a. If the matrices $A + \lambda\mathbb{I}_n, B + \lambda\mathbb{I}_n$ are similar for some $\lambda \in \mathbb{F}$, show that A, B are similar.

b. Is it true that if A^2, B^2 are similar, then A, B are similar;

Exercise 1.3. Let $A, B \in \mathbb{F}^{n\times n}$ be similar matrices. Prove the following equalities.

a. $\det A = \det B$.

b. $\operatorname{rank} A = \operatorname{rank} B$.

c. $\operatorname{Tr} A = \operatorname{Tr} B$.

Exercise 1.4. Show that for every $a \in \mathbb{R}$,

a. the matrices $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & -a \\ a & -1 \end{pmatrix}$ are not similar.

b. the matrices $\begin{pmatrix} 1 & -a \\ a & -1 \end{pmatrix}, -\begin{pmatrix} 1 & -a \\ a & -1 \end{pmatrix}$ are similar.

Exercise 1.5. A linear map is given $f : \mathbb{R}^2 \to \mathbb{R}^2, f(x, y) = (x + 2y, 2x + y)$.

a. Compute the matrices $(f : \hat{e}, \hat{e})$ and $(f : \hat{a}, \hat{a})$, where $\hat{a}$ is the ordered basis (a_1, a_2), with $a_1 = (1, -1), a_2 = (1, 1)$.

b. Find an invertible P with $(f : \hat{a}, \hat{a}) = P^{-1}(f : \hat{e}, \hat{e})P$ and an invertible Q with

$$(f : \hat{e}, \hat{e}) = Q^{-1} (f : \hat{a}, \hat{a}) \, Q.$$

Exercise 1.6. A linear map is given $f : \mathbb{R}^3 \to \mathbb{R}^3$ with formula

$$f(x, y, z) = (x + y + 2z, 2x + 2y + 4z, 3x + 3y + 6z).$$

a. Show that the set $\{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ is a basis of $\mathbb{R}^3$.

b. Compute the matrix $(f : \hat{a}, \hat{a})$, where $a = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ and P invertible with $(f : \hat{a}, \hat{a}) = P^{-1}(f : \hat{e}, \hat{e})P$.

c. Is it true that there exists an ordered basis $\hat{b}$ of $\mathbb{R}^3$ such that $(f : \hat{b}, \hat{b}) = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 2 & 3 & 1 \end{pmatrix}$?

Exercise 1.7. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$, $f(X) = AX - XA$.

a. Show that the map f is linear.

b. After computing the matrix $B = \left(f : \hat{E}, \hat{E}\right)$, where $\hat{E} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ is the usual ordered basis of $\mathbb{R}^{2\times 2}$, show that $\dim \ker f = \dim \operatorname{Im} f = 2$ and $B^3 = 0$.

c. Is it true that there exists an ordered basis $\hat{b}$ of $\mathbb{R}^3$, such that $(f : \hat{b}, \hat{b}) = \operatorname{diag}(1, -1, 0, 0)$?

Exercise 1.8. Examine which of the following statements are true. In each case justify your answer with a proof or counterexample. Let $A, B \in \mathbb{F}^{n\times n}$ be similar matrices.

a. If $A = \mathbb{I}_n$, then $B = \mathbb{I}_n$.

b. If $B = -A \in \mathbb{F}^{3\times 3}$, then A and B are not invertible.

c. The matrices $\begin{pmatrix} A & \\ & A \end{pmatrix}, \begin{pmatrix} B & \\ & B \end{pmatrix} \in \mathbb{F}^{2n\times 2n}$ are similar.

d. The matrices $\begin{pmatrix} A & \\ & C \end{pmatrix}, \begin{pmatrix} B & \\ & C \end{pmatrix} \in \mathbb{F}^{(n+m)\times(n+m)}$ are similar, for every $C \in \mathbb{F}^{m\times m}$.

CHAPTER 2

POLYNOMIALS

2.1 Divisibility

First, by $\mathbb{F}[x]$ we denote the set of polynomials with coefficients from $\mathbb{F}$.

Every $a(x) \in \mathbb{F}[x]$ with $a(x) \neq 0$ is written uniquely in the form

$$a(x) = a_n x^n + \cdots + a_1 x + a_0, \quad a_n \neq 0.$$

With the previous notation, $n = \deg a(x)$ is called the **degree** of $a(x)$, while a_n is called the **leading coefficient** of $a(x)$.

Observation 2.1.1. i. $\deg\left(a(x) + b(x)\right) \leq \max\left\{\deg a(x), \deg b(x)\right\}$,

ii. $\deg\left(a^m(x)\right) = m \cdot \deg a(x)$.

We consider the following operations on $\mathbb{F}[x]$:

$$a(x) + b(x), \quad a(x) \cdot b(x), \quad \lambda \cdot a(x)$$

with $a(x), b(x) \in \mathbb{F}[x]$ and $\lambda \in \mathbb{F}$.
Thus, $\mathbb{F}[x]$ becomes an $\mathbb{F}$-vector space with respect to addition and scalar multiplication.

Example 2.1.1. Let the polynomials

$$a(x) = a_m x^m + \cdots + a_1 x + a_0$$

and

$$b(x) = b_n x^n + \cdots + b_1 x + b_0$$

in $\mathbb{F}[x]$. Then, for the polynomial $c(x) = a(x) \cdot b(x)$ we have:

$$c_j = \sum_{0 \leq t} a_{j-t} \cdot b_t.$$

Definition 2.1.1. Let $a(x), b(x) \in \mathbb{F}[x]$. We will say that $a(x)$ **divides** $b(x)$ in $\mathbb{F}[x]$ if there exists $c(x) \in \mathbb{F}[x]$ such that

$$b(x) = a(x) \cdot c(x)$$

and we denote it by

$$a(x) \mid b(x).$$

Example 2.1.2.

i. We have that $x^2 + x + 1 \mid x^3 - 1$, since

$$x^3 - 1 = (x^2 + x + 1)(x - 1).$$

ii. For every polynomial $a(x) \in \mathbb{F}[x]$ we have $a(x) \mid 0$.

iii. In general, $0 \mid a(x)$ if and only if $a(x) = 0$.

Observation 2.1.2. If $a(x) \in \mathbb{F}[x]$ divides two polynomials $b(x), c(x) \in \mathbb{F}[x]$, then it also divides every polynomial of the form

$$f(x)b(x) + g(x)c(x)$$

for every $f(x), g(x) \in \mathbb{F}[x]$.

Proof. We have that $a(x) \mid b(x)$, that is, there exists a polynomial $q_1(x) \in \mathbb{F}[x]$ such that:

$$b(x) = a(x)q_1(x).$$

Similarly, there exists a polynomial $q_2(x) \in \mathbb{F}[x]$ such that:

$$c(x) = a(x)q_2(x).$$

Then we obtain:

$$\begin{aligned} f(x)b(x) + g(x)c(x) &= f(x)a(x)q_1(x) + g(x)a(x)q_2(x) \\ &= a(x)\left[f(x)q_1(x) + g(x)q_2(x)\right]. \end{aligned}$$

Thus, we conclude that

$$a(x) \mid f(x)b(x) + g(x)c(x),$$

for every $f(x), g(x) \in \mathbb{F}[x]$. □

Theorem 2.1.1 (Euclidean Division)**.** Let $a(x), b(x) \in \mathbb{F}[x]$ with $a(x) \neq 0$. Then there exist unique $q(x), r(x) \in \mathbb{F}[x]$ such that

$$b(x) = a(x)q(x) + r(x)$$

with $r(x) = 0$ or $\deg r(x) < \deg a(x)$.

Example 2.1.3. We consider the polynomials $a(x) = x^2 + 1$ and $b(x) = x^3 - 2x + 1$. Then we have the following:

$$b(x) = x \cdot a(x) + (-3x + 1).$$

Application 2.1.1. Let $f(x) \in \mathbb{F}[x]$ and $c \in \mathbb{F}$. Then

$$c \in \mathbb{F} \quad \text{is a root of} \quad f(x) \qquad \Longleftrightarrow \qquad x - c \big| \; f(x).$$

Proof. Suppose that $x - c \,|\, f(x)$, then there exists $g(x) \in \mathbb{F}[x]$ such that

$$f(x) = (x - c) \cdot g(x) \Rightarrow f(c) = (c - c)g(c) = 0.$$

Conversely, by Theorem 2.1.1 there exist $q(x), r(x) \in \mathbb{F}[x]$ such that:

$$f(x) = (x - c)q(x) + r(x), \; r(x) = 0 \quad \text{or} \quad \deg r(x) < \deg\,(x - c).$$

Hence $\deg r(x) = 0$ with

$$f(c) = (c - c)q(c) + r(c) = 0 \Leftrightarrow r(c) = 0.$$

That is, $r(x) = 0$ and

$$f(x) = (x - c)q(x).$$

□

Definition 2.1.2. Let $f(x), g(x) \in \mathbb{F}[x]$, not both zero. A $d(x) \in \mathbb{F}[x]$ is called the **greatest common divisor** of $f(x), g(x)$ if the following hold:

i. $d(x)$ is monic (the leading coefficient of $d(x)$ equals 1).

ii. $d(x)$ divides both $f(x)$ and $g(x)$.

iii. If there is another common divisor $d'(x) \in \mathbb{F}[x]$ with $d'(x) \mid f(x)$ and $d'(x) \mid g(x)$, then $d'(x) \mid d(x)$.

Theorem 2.1.2. Let $f(x), g(x) \in \mathbb{F}[x]$, not both zero.

i. There exists a unique greatest common divisor of $f(x), g(x)$.

ii. Let $d(x) = \gcd(f(x), g(x))$.

Then there exist $a(x), b(x) \in \mathbb{F}[x]$ such that:

$$d(x) = f(x) \cdot a(x) + g(x) \cdot b(x).$$

Definition 2.1.3. The polynomials $f(x), g(x) \in \mathbb{F}[x]$ are called **relatively prime** if:

$$\gcd(f(x), g(x)) = 1.$$

Example 2.1.4. For $\gcd(x-a, x-b)$ we have:

$$\gcd(x-a, x-b) = \begin{cases} 1, & \text{if } a \neq b \\ \\ x-a, & \text{if } a = b \end{cases}$$

More generally, if $p(x) \in \mathbb{F}[x]$ is an irreducible polynomial and $f(x) \in \mathbb{F}[x]$, then

$$\gcd(f(x), p(x)) = \begin{cases} p(x), & p(x) | f(x) \\ \\ 1, & p(x) \not| f(x) \end{cases}.$$

Application 2.1.2. Let $a(x), b(x) \in \mathbb{F}[x]$ be relatively prime. Then:

i. If $a(x) \mid b(x) \cdot c(x)$ with $c(x) \in \mathbb{F}[x]$, then $a(x) \mid c(x)$.

ii. If $a(x) \mid c(x)$ and $b(x) \mid c(x)$, then $a(x) \cdot b(x) \mid c(x)$.

Proof. i. Assume that $\gcd(a(x), b(x)) = 1$. By Theorem 2.1.2, there exist $a'(x), b'(x) \in \mathbb{F}[x]$ such that:

$$1 = a'(x)a(x) + b'(x)b(x) \tag{2.1}$$

$$\Rightarrow \quad c(x) = c(x)a'(x)a(x) + c(x)b'(x)b(x).$$

By assumption, $a(x) \mid b(x)c(x)$, that is, there exists $q(x) \in \mathbb{F}[x]$ such that

$$b(x)c(x) = q(x)a(x).$$

Hence:

$$c(x) = c(x)a'(x)a(x) + q(x)a(x)b'(x) = a(x)[a'(x)c(x) + q(x)b'(x)]$$

so $a(x) \mid c(x)$.

ii. From relation (2.1) and since $a(x) \mid c(x)$ and $b(x) \mid c(x)$, it follows that $a(x)b(x) \mid c(x)$.

□

2.2 Irreducible Polynomials

Definition 2.2.1. A polynomial $p(x) \in \mathbb{F}[x]$ of positive degree is called **irreducible** in $\mathbb{F}[x]$ if there do **not** exist $a(x), b(x) \in \mathbb{F}[x]$ such that:

$$a(x)b(x) = p(x) \quad \text{and} \quad \deg a(x) < \deg p(x), \quad \deg b(x) < \deg p(x).$$

Example 2.2.1.

1. Every $p(x) \in \mathbb{F}[x]$ with $\deg p(x) = 1$ is irreducible.
2. The polynomial $x^2 + 1 \in \mathbb{R}[x]$ is irreducible, while $x^2 + 1 \in \mathbb{C}[x]$ is not, since

$$x^2 + 1 = (x - i)(x + i).$$

Observation 2.2.1. Let $p(x) \in \mathbb{F}[x]$ be irreducible and monic. Then:

$$\gcd(p(x), q(x)) = \begin{cases} 1, & \text{if } p(x) \nmid q(x), \\ \\ p(x), & \text{if } p(x) \mid q(x). \end{cases}$$

Proposition 2.2.1. Let $f(x) \in \mathbb{C}[x]$. If $z \in \mathbb{C}$ is a root of $f(x)$, then its conjugate $\bar{z}$ is also a root of $f(x)$. [1]

[1] In $\mathbb{C}$: For every $z \in \mathbb{C}$ there exist unique $a, b \in \mathbb{R}$ such that $z = a + bi$. That is, $\mathbb{C}$ is a vector space over $\mathbb{R}$ with basis $\{1, i\}$ and $\dim_{\mathbb{R}} \mathbb{C} = 2$.

Question 2.2.1. Which are the irreducible polynomials in $\mathbb{R}[x]$ and in $\mathbb{C}[x]$?

i. The irreducible polynomials of $\mathbb{C}[x]$ are the linear polynomials. [2]

ii. The irreducible polynomials in $\mathbb{R}[x]$ are the linear ones or the quadratic ones with $\Delta < 0$.

Theorem 2.2.1. Every polynomial $f(x)$ of positive degree can be written uniquely as follows:

$$f(x) = c \cdot p_1^{n_1}(x) \cdots p_s^{n_s}(x),$$

where $c \in \mathbb{C}$ and the $p_i(x)$ are monic, pairwise distinct, irreducible polynomials.

Example 2.2.2. 1. Let $f(x) = x^3 - 1 \in \mathbb{F}[x]$. Depending on the field $\mathbb{F}$:

- If $\mathbb{F} = \mathbb{R}$, then $x^3 - 1 = (x-1)(x^2+x+1)$.
- If $\mathbb{F} = \mathbb{C}$, then:

$$x^3 - 1 = (x-1)\left(x - \frac{-1+\sqrt{3}i}{2}\right)\left(x - \frac{-1-\sqrt{3}i}{2}\right).$$

2. For $g(x) = x^4 + 1$, we observe that:

$$g(x) = \left(x^2 - \sqrt{2}x + 1\right)\left(x^2 + \sqrt{2}x + 1\right).$$

Proposition 2.2.2. Let

$$f(x) = c_1 p_1(x)^{m_1} \cdots p_s(x)^{m_s}, \quad g(x) = c_2 p_1(x)^{n_1} \cdots p_s(x)^{n_s},$$

where the $p_i(x)$ are monic, irreducible, pairwise distinct, and $0 \le m_i, n_i$. Define:

$$d_i = \min\{m_i, n_i\}, \quad d(x) = p_1(x)^{d_1} \cdots p_s(x)^{d_s}.$$

Then:

$$d(x) = \gcd\{f(x), g(x)\}.$$

Example 2.2.3. In $\mathbb{R}[x]$, consider:

$$f(x) = 3(x-5)^{10}(x^2+x+1)^6, \quad g(x) = -(x-5)^4(x-7)^4(x^2+x+1)^{10}.$$

Then:

$$\gcd(f(x), g(x)) = (x-5)^4(x^2+x+1)^6.$$

[2]This is equivalent to the Fundamental Theorem of Algebra: Every polynomial of positive degree with complex coefficients has a root in $\mathbb{C}$.

Definition 2.2.2. Let $a \in \mathbb{F}$ be a root of $f(x) \in \mathbb{F}[x]$. The greatest integer m such that $(x-a)^m \mid f(x)$ is called the **multiplicity** of the root a in $f(x)$. It is denoted by $m = \tau(a)$.

Example 2.2.4. For $\mathbb{F} = \mathbb{R}$ and

$$f(x) = (x-2)^5(x-3)(x^2+x+1),$$

we have:

$$\tau(2) = 5, \quad \tau(3) = 1.$$

Definition 2.2.3. A root a of $f(x)$ is called **simple** if its multiplicity equals 1. Otherwise, the root is called **multiple**.

Proposition 2.2.3. Let $a \in \mathbb{F}$ be a root of $f(x) \in \mathbb{F}[x]$. Then a is a multiple root of $f(x)$ if and only if it is also a root of the derivative $f'(x)$.

Proof. If a is a multiple root, then $(x-a)^2 \mid f(x)$, that is, there exists $g(x) \in \mathbb{F}[x]$ such that:

$$f(x) = (x-a)^2 g(x) \Rightarrow f'(x) = 2(x-a)g(x) + (x-a)^2 g'(x),$$

so $f'(a) = 0$.
Conversely, let $f(a) = f'(a) = 0$. Since $f(a) = 0$, there exists $g(x) \in \mathbb{F}[x]$ with $f(x) = (x-a)g(x)$. Then:

$$f'(x) = (x-a)g'(x) + g(x) \Rightarrow f'(a) = g(a) = 0 \Rightarrow (x-a) \mid g(x),$$

hence $f(x) = (x-a)^2 h(x)$ for some $h(x)$, i.e. a is a multiple root. □

Corollary 2.2.1. Let $f(x) \in \mathbb{F}[x]$. If $\gcd(f(x), f'(x)) = 1$, then every root of $f(x)$ is simple.

Proof. The proof is immediate via Proposition 2.2.3. □

2.3 Polynomials and matrices

Definition 2.3.1. Let $f(x) = f_m x^m + \cdots + f_1 x + f_0 \in \mathbb{F}[x]$ and $A \in \mathbb{F}^{n\times n}$. By $f(A)$ we denote the matrix:

$$f(A) = f_m A^m + \cdots + f_1 A + f_0 I_n.$$

Example 2.3.1. If $f(x) = -3x + 5 \in \mathbb{R}[x]$, then:

$$f(A) = -3A + 5I_n, \quad \text{for every } A \in \mathbb{F}^{n\times n}.$$

Observation 2.3.1. Let $f(x), g(x) \in \mathbb{F}[x]$, and let $h(x) = f(x) + g(x)$, $\kappa(x) = f(x)g(x)$. Then, for every $A \in \mathbb{F}^{n\times n}$ we have:

$$h(A) = f(A) + g(A), \quad \kappa(A) = f(A)g(A).$$

Example 2.3.2.

1. Consider the polynomials $f(x) = x^2 - x$ and $g(x) = x + 1$ in $\mathbb{R}[x]$. Then $k(x) = f(x) \cdot g(x) = x^3 - x$, hence:

$$k(A) = f(A) \cdot g(A) = A^3 - A = A(A - I_n)(A + I_n).$$

2. If $b(x) = q(x) \cdot a(x) + r(x)$, then $b(A) = q(A) \cdot a(A) + r(A)$.

2.4 Polynomials and linear maps

Definition 2.4.1. Let $f : V \to V$ be a linear map and let $a(x) = a_m x^m + \cdots + a_1 x + a_0 \in \mathbb{F}[x]$. We define:

$$a(f) : V \to V, \quad a(f) = a_m f^m + \cdots + a_1 f + a_0 1_V.$$

Observation 2.4.1. Let $a(x), b(x) \in \mathbb{F}[x]$, with $c(x) = a(x) + b(x)$, $d(x) = a(x) \cdot b(x)$ and $f : V \to V$ a linear map. Then:

$$c(f) = a(f) + b(f), \quad d(f) = a(f) \circ b(f).$$

Example 2.4.1. If $P(x) = x^2 - 1 \in \mathbb{F}[x]$, then:

$$P(f) = f^2 - 1_V = (f - 1_V) \circ (f + 1_V).$$

Proposition 2.4.1. Let $f : V \to V$ be a linear map, let $\hat{v}$ be an ordered basis of V and let $A = (f : \hat{v}, \hat{v})$. Then, for every $\varphi(x) \in \mathbb{F}[x]$, we have:

$$(\varphi(f) : \hat{v}, \hat{v}) = \varphi(A).$$

Proof.

$$\begin{aligned}(\varphi(f) : \hat{v}, \hat{v}) &= (\varphi_m f^m + \cdots + \varphi_1 f + \varphi_0 1_V : \hat{v}, \hat{v}) \\ &= \varphi_m (f^m : \hat{v}, \hat{v}) + \cdots + \varphi_1 (f : \hat{v}, \hat{v}) + \varphi_0 (1_V : \hat{v}, \hat{v}) \\ &= \varphi_m A^m + \cdots + \varphi_1 A + \varphi_0 I_n = \varphi(A).\end{aligned}$$

□

Example 2.4.2. i. Consider the linear map $f : \mathbb{F}^3 \to \mathbb{F}^3$ with matrix:

$$A = (f : \hat{v}, \hat{v}) = \begin{pmatrix} -2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

Then if $g = f^2 + 3f + 1_{\mathbb{F}^3}$, we have:

$$(g : \hat{v}, \hat{v}) = A^2 + 3A + I_n = \begin{pmatrix} 11 & 1 & 4 \\ -6 & 5 & 0 \\ -1 & 3 & -1 \end{pmatrix}.$$

ii. Let $f : V \to V$ be a linear map and let $\varphi(x) \in \mathbb{F}[x]$ such that $\varphi(f) = 0$. If the constant term $\varphi_0 \neq 0$, then f is an isomorphism.

Proof. Let $\varphi(x) = \varphi_m x^m + \cdots + \varphi_1 x + \varphi_0$. Then:

$$\varphi(f) = \varphi_m f^m + \cdots + \varphi_1 f + \varphi_0 1_V = 0.$$

Move the constant term:

$$1_V = -\frac{1}{\varphi_0} \cdot (\varphi_m f^m + \cdots + \varphi_1 f) = \left[-\frac{1}{\varphi_0} \cdot (\varphi_m f^{m-1} + \cdots + \varphi_1) \right] \circ f.$$

Hence f is left- and right-invertible, i.e. an isomorphism. □

2.5 Chapter 2 Exercises.

Group A : 1,2,3,5,6,7,8,9,11,15 **Group B :** 4,10,12,13,14,16

Exercise 2.1. Let $f(x), p(x) \in \mathbb{F}[x]$ where $p(x)$ is monic and irreducible. Show that

$$\gcd(f(x), p(x)) = 1 \quad \text{or} \quad \gcd(f(x), p(x)) = p(x).$$

Exercise 2.2. Find $\gcd\left(x^2+1, x^{2018}+1\right)$ and $\gcd\left(x^2+1, x^{2018}-1\right)$.

Exercise 2.3. a. Let $f(x) \in \mathbb{F}[x]$ and $a, b \in \mathbb{F}$ with $a \neq b$. Find the remainder of the division of $f(x)$ by $(x-a)(x-b)$.

b. Find all values of $c, d \in \mathbb{R}$ such that $(x-1)(x-2)\big|x^{100}+cx^5+dx+1$.

c. Find all values of $c, d \in \mathbb{R}$ such that $(x-1)^2\big|x^{100}+cx^5+dx+1$.

Exercise 2.4. The polynomials $f(x) = 2x^3 - 3x^2 + 6x + 5$ and $g(x) = x^3 + ax^2 + x + 1$ are given, where $a \in \mathbb{R}$.

a. Find the roots in $\mathbb{C}$ of $f(x)$.

b. For which values of a do $f(x), g(x)$ have a common real root?

c. Find the factorization of $g(x)$ into a product of monic irreducibles in $\mathbb{R}[x]$ if one of its roots in $\mathbb{C}$ is i.

Exercise 2.5. Let $f(x), g(x) \in \mathbb{R}[x]$, where $f(x) = x^5 - x^4 - x^2 + x,\ g(x) = x^2 + x - 6$. Find their gcd and lcm.

Exercise 2.6. Let $f(x), g(x) \in \mathbb{R}[x]$, where $f(x) = x^3 - x^2 + x - 1,\ g(x) = x^2 + x - 2$. Find the matrices $A \in \mathbb{R}^{n\times n}$ such that $f(A) = g(A) = 0$.

Exercise 2.7. Let $f(x), g(x) \in \mathbb{F}[x]$ with $\gcd(f(x), g(x)) = 1$.

a. Show that there is no $A \in \mathbb{F}^{n\times n}$ with $f(A) = g(A) = 0$.

b. Is it true that for every $h(x) \in \mathbb{F}[x]$ there exist $a(x), b(x) \in \mathbb{F}[x]$ such that $h(x) = a(x)f(x) + b(x)g(x)$?

Exercise 2.8. Show that every root in $\mathbb{C}$ of $f(x)$ is simple in the cases

a. $f(x) = x^n - 1$,

b. $f(x) = x^n + x + 1$.

Exercise 2.9. Let $f\colon \mathbb{F}^3 \to \mathbb{F}^3$ be the linear map with $(f : \hat{a}, \hat{a}) = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$, where $\hat{a}$ is an ordered basis of $\mathbb{F}^3$ and $\varphi(x) = x^2 + 3x + 1 \in \mathbb{F}[x]$. Find the matrix $(\varphi(f)\colon \hat{a}, \hat{a})$.

Exercise 2.10. Let $A \in \mathbb{F}^{n\times n}$ and $\varphi(x) \in \mathbb{F}[x]$. Show the following.

a. If A is diagonal, $A = \text{diag}(a_1, \cdots, a_n)$, then $\varphi(A) = \text{diag}(\varphi(a_1), \cdots, \varphi(a_n))$.

b. If A is of the form $A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$, where $A_i \in \mathbb{F}^{n_i \times n_i}$ and $n_1 + \cdots n_k = n$ ('block diagonal'), then $\varphi(A) = \begin{pmatrix} \varphi(A_1) & & \\ & \ddots & \\ & & \varphi(A_k) \end{pmatrix}$. (**Note**. We mean that the invisible entries are 0.)

c. If A is upper triangular, $A = \begin{pmatrix} a_1 & & * \\ & \ddots & \\ & & a_n \end{pmatrix}$, then $\varphi(A) = \begin{pmatrix} \varphi(a_1) & & * \\ & \ddots & \\ & & \varphi(a_n) \end{pmatrix}$.

d. If A is of the form $A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ & & A_k \end{pmatrix}$, where $A_i \in \mathbb{F}^{n_i \times n_i}$ and $n_1 + \cdots n_k = n$ ('block upper triangular'), then $\varphi(A) = \begin{pmatrix} \varphi(A_1) & & * \\ & \ddots & \\ & & \varphi(A_k) \end{pmatrix}$.

Exercise 2.11. Let $A \in \mathbb{F}^{n\times n}$.

a. Let $\varphi(x) \in \mathbb{F}[x]$ with nonzero constant term and $\varphi(A) = 0$. Show that A is invertible.

b. Let $A^5 = 0$. Show that the matrix $\varphi(A)$ is invertible, where $\varphi(x) = x^4 - x^3 + x^2 - x + 1$.

Exercise 2.12. The conclusion in question b. is called **Lagrange's Theorem**. For $\mathbb{F} = \mathbb{R}$, it says that through n distinct points of the plane there passes a unique polynomial curve of degree at most $n-1$, analogous to the fact that through two distinct points of the plane there passes a unique line. Let $\lambda_1, \cdots, \lambda_n \in \mathbb{F}$ be distinct. Consider the vector space $\mathbb{F}_{n-1}[x]$ of all polynomials of degree at most $n-1$ and the map

$$f\colon \mathbb{F}_{n-1}[x] \to \mathbb{F}^n, f(\varphi(x)) = (\varphi(\lambda_1), \cdots, \varphi(\lambda_n)).$$

a. Show that the map f is linear, one-to-one, and onto.

b. Show that for every $a_1, \cdots, a_n \in \mathbb{F}$ there exists a unique $\varphi(x) \in \mathbb{F}_{n-1}[x]$ such that $\varphi(\lambda_1) = a_1, \cdots, \phi(\lambda_n) = a_n$.

c. Find a polynomial $\varphi(x)$ such that $\varphi(1) = 2, \varphi(2) = 1, \varphi(-1) = 1$.

d. Show that the $\varphi(x)$ of subquestion b. is given by the relation $\varphi = \sum_{j=1}^{n} a_j \varphi_j(x)$, where

$$\varphi_j(x) = \prod_{k=1, k\neq j}^{n} \frac{x - \lambda_k}{\lambda_j - \lambda_k}.$$

Exercise 2.13. Let $A \in \mathbb{R}^{n\times n}$ and $B = \begin{pmatrix} A & I_n \\ 0 & A \end{pmatrix} \in \mathbb{R}^{2n\times 2n}$.

a. Show that $f(B) = \begin{pmatrix} f(A) & f'(A) \\ 0 & f(A) \end{pmatrix}$ for every $f(x) \in \mathbb{R}[x]$, where $f'(x)$ is the derivative of $f(x)$.

b. Show that if $(A - I_n)^{2013}(A - 2I_n)^{2014} = 0$, then $(B - I_{2n})^{2014}(B - 2I_{2n})^{2015} = 0$.

Exercise 2.14. Show that for every $A \in \mathbb{F}^{n\times n}$ there exists a nonzero $\varphi(x) \in \mathbb{F}[x]$ of degree at most n^2 such that $\varphi(A) = 0$.

Exercise 2.15. Let $A, B, P \in \mathbb{F}^{n\times n}$ such that $B = P^{-1}AP$. Show that $\varphi(B) = P^{-1}\varphi(A)P$ for every $\varphi(x) \in \mathbb{F}[x]$.

Exercise 2.16. Let $a_1, \cdots, a_n \in \mathbb{R}$. Set

$$e_i = \sum_{1\leq t_1<\cdots<t_i\leq n} a_{t_1}a_{t_2}\cdots a_{t_i},\ i = 1, \cdots, n.$$

For example, if $n = 3$, then

$$e_1 = a_1 + a_2 + a_3,\ e_2 = a_1a_2 + a_1a_3 + a_2a_3,\ e_3 = a_1a_2a_3.$$

Show that if $e_i > 0$ for each $i = 1, \cdots, n$, then $a_i > 0$ for each $i = 1, \cdots n$.

Exercise 2.17. Examine which of the following statements are true. In each case justify your answer with a proof or counterexample. Let $f(x), g(x), h(x) \in \mathbb{F}[x]$.

a. If $f(x)\big|g(x)h(x)$, then $f(x)\big|g(x)$ or $f(x)\big|h(x)$.
b. Let $f(x)$ be irreducible. If $f(x)\big|g(x)h(x)$, then $f(x)\big|g(x)$ or $f(x)\big|h(x)$.
c. If $f(x)\big|h(x)$ and $g(x)\big|h(x)$, then $f(x)g(x)\big|h(x)$.
d. If $f(x)\big|h(x)$, $g(x)\big|h(x)$ and $\gcd(f(x), g(x)) = 1$, then $f(x)g(x)\big|h(x)$.
e. Let $A \in \mathbb{F}^{n\times n}$. If $f(A) = g(A) = 0$, then $f(x), g(x)$ are not relatively prime.

CHAPTER 3

EIGENVALUES AND EIGENVECTORS

3.1 Eigenvalues, eigenvectors and eigenspaces of a matrix

3.1.1 Eigenvalues and eigenvectors of a matrix

Definition 3.1.1. Let $A \in \mathbb{F}^{n\times n},\ \lambda \in \mathbb{F}$ and $X \in \mathbb{F}^{n\times 1},\ X \neq 0$. If the relation

$$AX = \lambda X, \tag{3.1}$$

holds, we say that λ is an **eigenvalue** of A and X is a corresponding **eigenvector** of A associated with the eigenvalue λ.

Example 3.1.1. Let

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix}, \quad X = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

i. We have $AX = X$, hence 1 is an eigenvalue of A and X is a corresponding eigenvector.

ii. We have $AY = -Y$, hence -1 is an eigenvalue of A and Y is an eigenvector.

iii. We have $AZ = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}$. There is no $\lambda \in \mathbb{R}$ such that $\lambda Z = AZ$, hence Z is not an eigenvector of A.

Example 3.1.2. Let the matrix
$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$
We will find the eigenvalues and eigenvectors of A.

Proof. Let $X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2\times 1}$ and $\lambda \in \mathbb{R}$. From relation (3.1) we have:

$$AX = \lambda X \Leftrightarrow \begin{pmatrix} x+3y \\ 4x+2y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} \Leftrightarrow \begin{cases} x + 3y = \lambda x \\ \\ 4x + 2y = \lambda y \end{cases} \Leftrightarrow \begin{cases} (1-\lambda)x + 3y = 0 \\ \\ 4x + (2-\lambda)y = 0 \end{cases}$$

The system has a nontrivial solution if and only if:
$$\det \begin{pmatrix} 1-\lambda & 3 \\ 4 & 2-\lambda \end{pmatrix} = 0 \Rightarrow (1-\lambda)(2-\lambda) - 12 = \lambda^2 - 3\lambda + 10 = 0.$$
Hence the eigenvalues of A are:
$$\lambda = 5 \quad \text{and} \quad \lambda = -2.$$

Now we find the corresponding eigenvectors:

i. For $\lambda = -2$: From the system we get $3x + 3y = 0 \Rightarrow y = -x$, hence:
$$V(-2) = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} : x \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle.$$

ii. For $\lambda = 5$: From $-4x + 3y = 0 \Rightarrow y = \frac{4}{3}x$, hence:
$$V(5) = \left\{ \begin{pmatrix} x \\ \frac{4}{3}x \end{pmatrix} : x \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 \\ \frac{4}{3} \end{pmatrix} \right\rangle.$$

□

Example 3.1.3. We consider the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{F}^{2\times 2}.$$

We distinguish the following cases:

a. Assume that $A \in \mathbb{R}^{2\times 2}$ and let $\lambda \in \mathbb{R}$ and $X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2\times 1}$. Then:

$$AX = \lambda X \Leftrightarrow \begin{pmatrix} y \\ -x \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} \Rightarrow \begin{cases} \lambda x - y = 0 \\ x + \lambda y = 0 \end{cases}$$

The determinant of the system is:

$$\det \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} = \lambda^2 + 1 \neq 0, \quad \text{for every } \lambda \in \mathbb{R}.$$

Hence, there are no eigenvalues and eigenvectors of A over $\mathbb{R}$.

b. If $A \in \mathbb{C}^{2\times 2}$, then for $\lambda \in \mathbb{C}$ and $X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^{2\times 1}$, we have:

$$AX = \lambda X \Leftrightarrow \begin{pmatrix} y \\ -x \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} \Rightarrow \begin{cases} \lambda x - y = 0 \\ x + \lambda y = 0 \end{cases}$$

The system has a nonzero solution if and only if:

$$\det \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} = \lambda^2 + 1 = 0 \Rightarrow \lambda = i \text{ or } \lambda = -i.$$

i. For the eigenvalue $\lambda = i$, from $ix - y = 0 \Rightarrow y = ix$. Hence:

$$V(i) = \left\{X \in \mathbb{C}^{2\times 1} \mid y = ix\right\} = \left\{\begin{pmatrix} x \\ ix \end{pmatrix}\right\} = \left\langle \begin{pmatrix} 1 \\ i \end{pmatrix} \right\rangle$$

is the set of eigenvectors of A corresponding to the eigenvalue $\lambda = i$.

ii. For the eigenvalue $\lambda = -i$, from $-ix - y = 0 \Rightarrow y = -ix$. Hence:

$$V(-i) = \left\{X \in \mathbb{C}^{2\times 1} \mid y = -ix\right\} = \left\{\begin{pmatrix} x \\ -ix \end{pmatrix}\right\} = \left\langle \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\rangle$$

is the set of eigenvectors of A corresponding to the eigenvalue $\lambda = -i$.

Properties 3.1.1. Let $A \in \mathbb{F}^{n \times n},\ \lambda \in \mathbb{F}$. The following are equivalent:

i. λ is an eigenvalue of A.

ii. There exists $X \in \mathbb{F}^{n \times 1}$ with $X \neq 0$ such that $(A - \lambda \mathbb{I}_n)X = 0$.

iii. $\det(A - \lambda \mathbb{I}_n) = 0$.

Proof.

- i. $\Rightarrow$ ii.: By definition, there exists $X \neq 0$ such that

$$AX = \lambda X \iff (A - \lambda \mathbb{I}_n)X = 0.$$

- ii. $\Rightarrow$ iii.: The implication follows from the well-known proposition:

If $B \in \mathbb{F}^{n \times n}$ and $BX = 0$ has a nonzero solution, then $\det B = 0$.

- iii. $\Rightarrow$ i.: If $\det(A - \lambda \mathbb{I}_n) = 0$, then there exists $X \neq 0$ with

$$(A - \lambda \mathbb{I}_n)X = 0 \Rightarrow AX = \lambda X,$$

hence λ is an eigenvalue.

□

Corollary 3.1.1.

i. A matrix $A \in \mathbb{F}^{n \times n}$ is invertible if and only if 0 is not an eigenvalue of A.

ii. If A is upper or lower triangular with diagonal entries $a_1, \ldots, a_n$, then:

$$\det(A - \lambda \mathbb{I}_n) = 0 \Leftrightarrow \prod_{i=1}^{n} (a_i - \lambda) = 0.$$

That is, λ is an eigenvalue of A if and only if $\lambda = a_i$ for some i.

iii. λ is an eigenvalue of A if and only if it is an eigenvalue of A^t.

Proof.

i. The proof is left as an exercise to the reader.

ii. The proof is left as an exercise to the reader.

iii. From:

$$\det(A - \lambda \mathbb{I}_n) = \det\left((A - \lambda \mathbb{I}_n)^t\right) = \det(A^t - \lambda \mathbb{I}_n),$$

we conclude that λ is an eigenvalue of A^t.

□

Example 3.1.4. We observe that 2 is an eigenvalue of the matrix:

$$A = \begin{pmatrix} 0 & 2 & 3 \\ 2 & 0 & 3 \\ 1 & 0 & 2 \end{pmatrix},$$

since:

$$\det(A - 2I_3) = \det \begin{pmatrix} -2 & 2 & 3 \\ 2 & -2 & 3 \\ 1 & 0 & 0 \end{pmatrix} = 0.$$

Also, 0 is **not** an eigenvalue of A because:

$$\det A = 6 \neq 0.$$

3.1.2 Eigenspaces of a matrix

Definition 3.1.2. Let $A \in \mathbb{F}^{n\times n}$ and let λ be an eigenvalue of A. The **eigenspace** of A corresponding to λ is the set:

$$V_A(\lambda) = \left\{ X \in \mathbb{F}^{n\times 1} \mid AX = \lambda X \right\}.$$

Observation 3.1.1.

i. $V_A(\lambda)$ is the set of eigenvectors of A corresponding to λ, together with the zero vector.

ii. $V_A(\lambda)$ is a subspace of $\mathbb{F}^{n\times 1}$, since it is the solution set of the homogeneous system

$$(A - \lambda \mathbb{I}_n)X = 0.$$

Theorem 3.1.1. Let $A \in \mathbb{F}^{n\times n}$ and let $\lambda \in \mathbb{F}$ be an eigenvalue of A. Then:

$$\dim V_A(\lambda) = n - \text{rank}(A - \lambda \mathbb{I}_n).$$

Proof. Consider the linear map

$$\mathcal{L}_B : \mathbb{F}^{n\times 1} \to \mathbb{F}^{n\times 1}, \quad \mathcal{L}_B(X) = BX,$$

where

$$B = A - \lambda \mathbb{I}_n.$$

Then:

i. $\dim \ker \mathcal{L}_B + \dim \text{Im}\mathcal{L}_B = n$ (Dimension Theorem),

ii. $\dim \text{Im}\mathcal{L}_B = \text{rank}B$.

Since $\ker \mathcal{L}_B = V_A(\lambda)$, we obtain:

$$\dim V_A(\lambda) = \dim \ker \mathcal{L}_B = n - \text{rank}(A - \lambda \mathbb{I}_n).$$

□

3.1.3 Properties of eigenspaces

Proposition 3.1.1. Let $\lambda_1, \dots, \lambda_s$ be distinct eigenvalues of the matrix A and $X_1, \dots, X_s \in \mathbb{F}^{n\times 1}$ with $X_i \in V_A(\lambda_i)$ for each $i \in \{1, \dots, s\}$. If

$$X_1 + \cdots + X_s = 0,$$

then

$$X_1 = X_2 = \cdots = X_s = 0.$$

Proof. We will use induction on s.

- Base case: For $s = 1$, clearly $X_1 = 0$.

- Inductive step: Assume the statement holds for $s-1$. Suppose $X_1 + \cdots + X_s = 0$ with each $X_i \in V_A(\lambda_i)$. Then:

$$AX_1 + \cdots + AX_s = A(X_1 + \cdots + X_s) = A(0) = 0,$$

hence:

$$\lambda_1 X_1 + \cdots + \lambda_s X_s = 0.$$

Subtract the relation $\lambda_1(X_1 + \cdots + X_s) = 0$:

$$(\lambda_1 X_1 + \cdots + \lambda_s X_s) - \lambda_1(X_1 + \cdots + X_s) = 0$$

$$\iff \quad (\lambda_2 - \lambda_1)X_2 + \cdots + (\lambda_s - \lambda_1)X_s = 0.$$

For $i \geq 2$, we have

$$(\lambda_i - \lambda_1)X_i \in V_A(\lambda_i)$$

and since the eigenvalues are distinct,

$$\lambda_i \neq \lambda_1 \Rightarrow \lambda_i - \lambda_1 \neq 0.$$

Therefore, by the inductive hypothesis we get

$$X_2 = \cdots = X_s = 0$$

and from the original assumption it follows that $X_1 = 0$. □

Corollary 3.1.2. Eigenvectors of a matrix A corresponding to distinct eigenvalues are linearly independent.

Proof. Let $X_1, \ldots, X_s$ be eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_s$, respectively. If

$$m_1 X_1 + \cdots + m_s X_s = 0$$

for some $m_i \in \mathbb{F}$, then by Proposition 3.1.1 we have

$$m_i X_i = 0$$

that is $m_i = 0$, since $X_i \neq 0$. Hence the vectors are linearly independent. □

Application 3.1.1. Let X, Y be eigenvectors of the matrix A corresponding to different eigenvalues. Show that $aX + bY$ is not an eigenvector of A, if $ab \neq 0$.

Proof. Assume $AX = rX$ and $AY = mY$ with $r \neq m$, and that $aX + bY$ is an eigenvector for some eigenvalue λ. Then:

$$\begin{aligned} & A(aX + bY) = \lambda(aX + bY) \\ \Rightarrow \quad & arX + bmY = a\lambda X + b\lambda Y \\ \Rightarrow \quad & (ar - a\lambda)X + (bm - b\lambda)Y = 0. \end{aligned}$$

Since X, Y are eigenvectors with distinct eigenvalues, they are linearly independent (by Corollary 3.1.2), hence the corresponding coefficients must be zero:

$$\begin{cases} ar - a\lambda = 0 \Rightarrow \lambda = r \\ bm - b\lambda = 0 \Rightarrow \lambda = m \Rightarrow r = m \end{cases}$$

a contradiction. Therefore $aX + bY$ is not an eigenvector if $ab \neq 0$. □

Proposition 3.1.2. Let $A \in \mathbb{F}^{n\times n}$, $\varphi(x) \in \mathbb{F}[x]$, let λ be an eigenvalue of A and X a corresponding eigenvector. Then $\varphi(\lambda)$ is an eigenvalue of the matrix $\varphi(A)$ and X is a corresponding eigenvector. That is:

$$V_A(\lambda) \subseteq V_{\varphi(A)}(\varphi(\lambda)).$$

Proof. By assumption we have $AX = \lambda X$. Note:

$$A^2X = A(AX) = A(\lambda X) = \lambda(AX) = \lambda^2 X.$$

We generalize by induction on the exponent.

Claim: For every $m \in \mathbb{N}$ we have: $A^m X = \lambda^m X$.

Proof of Claim:

- **Base case:** For $m = 1$, it holds immediately: $AX = \lambda X$.
- **Step:** Assume $A^m X = \lambda^m X$. Then:

$$A^{m+1}X = A(A^m X) = A(\lambda^m X) = \lambda^m(AX) = \lambda^m \cdot \lambda X = \lambda^{m+1}X.$$

  Hence the claim holds for every $m \in \mathbb{N}$.

Now let $\varphi(x) = \varphi_m x^m + \cdots + \varphi_0$. Then:

$$\varphi(A)X = (\varphi_m A^m + \cdots + \varphi_0 \mathbb{I}_n)\, X = (\varphi_m \lambda^m + \cdots + \varphi_0)\, X = \varphi(\lambda)X.$$

Since $X \neq 0$, it is an eigenvector of $\varphi(A)$ for the eigenvalue $\varphi(\lambda)$. □

Example 3.1.5. Attention! The inclusion is not equality in general. If

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

then:

$$V_A(1) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle.$$

But if $\varphi(x) = x^2$, then
$$\varphi(A) = A^2 = I_2,$$
and:
$$V_{\varphi(A)}(1) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^{2\times 1}.$$

Hence
$$V_A(1) \subsetneqq V_{\varphi(A)}(1).$$

Example 3.1.6. a. If $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector of a matrix A for eigenvalue -1, then it is also an eigenvector of the matrix $A^{1821} + 10A$, with eigenvalue:
$$(-1)^{1821} + 10(-1) = -1 + (-10) = -11.$$

b. If $A \in \mathbb{F}^{2\times 2}$ has eigenvalues 2 and -3, then for $\varphi(x) = x^2 + 5$, the eigenvalues of $\varphi(A)$ are:
$$\varphi(2) = 2^2 + 5 = 9, \quad \varphi(-3) = 9 + 5 = 14.$$

c. If B is a matrix with eigenvalues -1 and 1, then the matrix B^2 has the unique eigenvalue 1, since:
$$(-1)^2 = 1, \quad 1^2 = 1.$$

3.2 Eigenvalues, eigenvectors and linear maps

Definition 3.2.1. Let $f\colon V \to V$ be a linear map, $\lambda \in \mathbb{F}$ and $v \in V$ with $v \neq 0$. If
$$f(v) = \lambda v,$$
then λ is called an **eigenvalue** of f and v is an **eigenvector** of f corresponding to the eigenvalue λ.

If λ is an eigenvalue of f, then the corresponding **eigenspace** of f is the set:
$$V_f(\lambda) = \{v \in V \mid f(v) = \lambda v\}.$$

Observation 3.2.1. If $f\colon V \to V$ is a linear map and λ is an eigenvalue of it, then:

$$V_f(\lambda) = \ker(f - \lambda \cdot 1_V) \leq V.$$

Proposition 3.2.1. Let $f\colon V \to V$ be a linear map. If $\hat{v}$ is any ordered basis of V and $A = (f\colon \hat{v}, \hat{v})$, then:

$$\dim V_f(\lambda) = \dim V - \dim \mathrm{Im}(f - \lambda 1_V) = \dim V - \mathrm{rank}(A - \lambda \mathbb{I}_n).$$

Example 3.2.1. Consider the linear map $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, $f(x, y, z) = (0, 0, x + y)$. Find a basis of each eigenspace.

Proof. We look for $(x, y, z) \in \mathbb{R}^3$ such that:

$$f(x, y, z) = \lambda(x, y, z), \quad \lambda \in \mathbb{R}.$$

From:

$$(0, 0, x + y) = (\lambda x, \lambda y, \lambda z)$$

we get the system:

$$\begin{cases} \lambda x = 0 \\ \lambda y = 0 \\ \lambda z = x + y \end{cases}.$$

The system has a nonzero solution if and only if:

$$\det \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 1 & 1 & -\lambda \end{pmatrix} = 0 \Leftrightarrow \lambda = 0.$$

Hence, from the first two equations we have $x = -y$. Therefore:

$$V_f(0) = \left\{(x, -x, z) \in \mathbb{R}^3 \mid x, z \in \mathbb{R}\right\} = \langle (1, -1, 0), (0, 0, 1) \rangle .$$

The set $\{(1, -1, 0), (0, 0, 1)\}$ is linearly independent, hence it is a basis of the eigenspace $V_f(0)$. $\square$

Attention to this specific example!

Example 3.2.2. Let $V = \mathbb{R}_2[x]$ and $f\colon \mathbb{R}_2[x] \to \mathbb{R}_2[x]$, $f(\varphi(x)) = \varphi(x) + \varphi'(x)$ be a linear map. Find a basis for each eigenspace of f.

Proof. Let $\varphi(x) = ax^2 + bx + c \in \mathbb{R}_2[x]$. Then:

$$f(\varphi(x)) = \varphi(x) + \varphi'(x) = ax^2 + bx + c + 2ax + b = ax^2 + (2a + b)x + (b + c).$$

If $f(\varphi(x)) = \lambda \cdot \varphi(x)$, then:

$$ax^2 + (2a + b)x + (b + c) = a\lambda x^2 + b\lambda x + c\lambda.$$

Thus we get the system:

$$\begin{cases} a(1-\lambda) = 0 \\ 2a - b(1-\lambda) = 0 \\ b + c(1-\lambda) = 0 \end{cases}$$

The system has a nonzero solution if and only if:

$$\det \begin{pmatrix} 1-\lambda & 0 & 0 \\ 2 & 1-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{pmatrix} = 0 \Leftrightarrow \lambda = 1.$$

For the eigenvalue $\lambda = 1$ we have $a = b = 0$, hence $V_f(1) = \langle 1 \rangle$ and the basis of the eigenspace is $\{1\}$. □

Proposition 3.2.2. Let $f\colon V \to V$ be a linear map and $A = (f : \hat{a}, \hat{a})$ with respect to some basis $\hat{a}$ of V, and $\lambda \in \mathbb{F}$. Then:

1. i. λ is an eigenvalue of f if and only if it is an eigenvalue of A.
2. ii. If $v \in V$, then $v \in V_f(\lambda)$ if and only if $[v]_{\hat{a}} \in V_A(\lambda)$.
3. iii. The set $\{v_1, \ldots, v_m\}$ is a basis of $V_f(\lambda)$ if and only if

$$\{[v_1]_{\hat{a}}, \ldots, [v_m]_{\hat{a}}\}$$

[1] is a basis of $V_A(\lambda)$.

Proof. i. Let λ be an eigenvalue of f. Then there exists $v \neq 0$ such that

$$f(v) = \lambda v \iff (f - \lambda 1_V)(v) = 0,$$

that is, $f - \lambda 1_V$ is not invertible, hence $A - \lambda \mathbb{I}_n$ is not invertible, i.e.

$$\det(A - \lambda \mathbb{I}_n) = 0,$$

so λ is an eigenvalue of A.

ii. We have

$$v \in V_f(\lambda) \Leftrightarrow f(v) = \lambda v \Leftrightarrow A[v]_{\hat{a}} = \lambda [v]_{\hat{a}} \Leftrightarrow [v]_{\hat{a}} \in V_A(\lambda).$$

iii. The map $v \mapsto [v]_{\hat{a}}$ is an isomorphism $g\colon V \to \mathbb{F}^n$, and by (ii) it restricts to an isomorphism $V_f(\lambda) \to V_A(\lambda)$.

□

Example 3.2.3. Let $V = \mathbb{R}_2[x]$, $\hat{a} = (a_1, a_2, a_3)$ with

$$a_1 = 1, \quad a_2 = x + 1, \quad a_3 = x^2 + 1$$

and let $f\colon \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ such that:

$$A = (f\colon \hat{a}, \hat{a}) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

i. Show that $\hat{a}$ is a basis of V.

ii. Is it true that $a_1 - a_3$ is an eigenvector?

iii. Find a basis for each eigenspace of f.

iv. Find a basis for each eigenspace of A.

[1] For example, if $\hat{a} = (a_1, a_2, a_3)$ and $v = a_1 + 2a_2 + 5a_3$, then

$$[v]_{\hat{a}} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}.$$

Proof. i. The polynomials a_1, a_2, a_3 span (show why) V and their number equals $\dim \mathbb{R}_2[x] = 3$, hence they are linearly independent and therefore a basis of $\mathbb{R}_2[x]$.

ii. Note that $f(a_1) = f(a_2) = f(a_3) = a_2$. Take $\lambda \in \mathbb{R}$ such that

$$f(a_1-a_3) = \lambda(a_1-a_3) \Leftrightarrow f(a_1)-f(a_3) = \lambda(a_1-a_3) \Leftrightarrow \lambda(a_1-a_3) = 0 \Leftrightarrow a_1 - a_3 \neq 0\lambda = 0.$$

Hence $a_1 - a_3$ is an eigenvector of f with corresponding eigenvalue 0.

iii. Let $v \in V$ and $\lambda \in \mathbb{R}$ such that

$$\begin{aligned} f(v) = \lambda v \quad &\iff \quad f(r_1a_1 + r_2a_2 + r_3a_3) = \lambda(r_1a_1 + r_2a_2 + r_3a_3) \\ &\iff \quad r_1f(a_1) + r_2f(a_2) + r_3f(a_3) = \lambda r_1a_1 + \lambda r_2a_2 + \lambda r_3a_3 \\ &\iff \quad \lambda r_1a_1 + a_2(\lambda r_2 - r_1 - r_2 - r_3) + \lambda r_3a_3 = 0 \\ &\iff \quad \begin{cases} \lambda r_1 = 0 \\ r_1 + r_2(1-\lambda) + r_3 = 0 \\ \lambda r_3 = 0 \end{cases} \end{aligned}$$

Thus the system has a nonzero solution if and only if

$$\det \begin{pmatrix} \lambda & 0 & 0 \\ 1 & 1-\lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = 0 \iff \lambda = 0 \quad \text{or} \quad \lambda = 1.$$

Hence:

1. $$V_f(0) = \{r_1a_1 + r_2a_2 + r_3a_3 \mid r_1 + r_2 + r_3 = 0\} = \langle a_2 - a_1, a_3 - a_1 \rangle$$
   linearly independent (show why), so $\{a_2 - a_1, a_3 - a_1\}$ is a basis of $V_f(0)$.
2. $$V_f(1) = \{r_1a_1 + r_2a_2 + r_3a_3 \mid r_1 = r_3 = 0\} = \{r_2a_2\} = \langle a_2 \rangle .$$
   Hence $\{a_2\}$ is a basis of $V_f(1)$.

iv. By Proposition 3.2.2 we have that

$$\{[a_3 - a_1]_{\hat{a}}, [a_2 - a_1]_{\hat{a}}\}$$

is a basis of $V_A(0)$ and

$$\{[a_2]_{\hat{a}}\}$$

a basis of $V_A(1)$.

□

3.3 Characteristic polynomial

Definition 3.3.1. Let $A \in \mathbb{F}^{n\times n}$ with $A = (a_{ij})$. The **characteristic polynomial** of A is:

$$\chi_A(x) = \det(A - x\mathbb{I}_n) = \det \begin{pmatrix} a_{11} - x & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - x \end{pmatrix}.$$

Example 3.3.1.

1. If $A = (a)$, then $\chi_A(x) = \det(a - x) = a - x$.

2. If $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}$, then:

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} 1-x & 3 \\ 4 & 2-x \end{pmatrix} \\ &= (1-x)(2-x) - 12 \\ &= x^2 - 3x - 10 \\ &= (x+2)(x-5). \end{aligned}$$

3. If $A = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 1 & 3 \\ 0 & 4 & 2 \end{pmatrix}$, then:

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} 2-x & 3 & 4 \\ 0 & 1-x & 3 \\ 0 & 4 & 2-x \end{pmatrix} \\ &= (2-x)[(1-x)(2-x)-12] \\ &= (2-x)(x+2)(x-5). \end{aligned}$$

Properties 3.3.1. Let $A \in \mathbb{C}^{n\times n}$. Then:

i. $\chi_A(x) = \chi_{A^t}(x)$

ii. If A is upper or lower triangular with diagonal entries a_{ii}, then:

$$\chi_A(x) = (-1)^n (x-a_{11})\cdots(x-a_{nn}).$$

iii. If $A_i \in \mathbb{F}^{n_i\times n_i}$ for $i = 1, \dots, s$ and $A \in \mathbb{F}^{n\times n}$ with

$$n = n_1 + \cdots + n_s$$

of the form:

$$A = \begin{pmatrix} A_1 & * & * & \cdots & * \\ 0 & A_2 & * & \cdots & * \\ \vdots & \vdots & \ddots & \cdots & * \\ 0 & 0 & 0 & \cdots & A_s \end{pmatrix},$$

then:

$$\chi_A(x) = \chi_{A_1}(x)\cdots\chi_{A_s}(x).$$

Proof. i. We have:

$$\chi_{A^t}(x) = \det(A^t - x\mathbb{I}_n) = \det\left((A - x\mathbb{I}_n)^t\right) = \det(A - x\mathbb{I}_n) = \chi_A(x).$$

ii. Let A be upper triangular, that is:

$$A = \begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Then:

$$\chi_A(x) = \det(A - x\mathbb{I}_n) = \det \begin{pmatrix} a_{11} - x & * & \cdots & * \\ 0 & a_{22} - x & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} - x \end{pmatrix} = \prod_{i=1}^{n} (a_{ii} - x).$$

iii. We use the following lemma:

Lemma 3.3.1. (a) If $B_1 \in \mathbb{F}^{n_1 \times n_1}$, $B_2 \in \mathbb{F}^{n_2 \times n_2}$ and $B = \begin{pmatrix} B_1 & * \\ 0 & B_2 \end{pmatrix}$, then:

$$\det B = \det B_1 \cdot \det B_2.$$

(b) If $B = \begin{pmatrix} B_1 & * & \cdots & * \\ 0 & B_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_s \end{pmatrix}$ with $B_i \in \mathbb{F}^{n_i \times n_i}$ and $n = n_1 + \cdots + n_s$, then:

$$\det B = \prod_{i=1}^{s} \det B_i.$$

Proof of the lemma.

(a) If $B_1 = \mathbb{I}_{n_1}$, then the property follows immediately by expanding along the first column of B. Similarly for $B_2 = \mathbb{I}_{n_2}$ by expanding along the last row. In the general case:

$$\begin{pmatrix} B_1 & * \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} \mathbb{I}_{n_1} & 0 \\ 0 & B_2 \end{pmatrix} \begin{pmatrix} B_1 & * \\ 0 & \mathbb{I}_{n_2} \end{pmatrix},$$

hence:

$$\det B = \det B_1 \cdot \det B_2.$$

(b) The general result follows by induction, applying (1) successively.

Back to the proof.

Since A has the form:

$$A = \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_s \end{pmatrix},$$

we have:

$$\chi_A(x) = \det(A - x\mathbb{I}_n) = \det \begin{pmatrix} A_1 - xI_{n_1} & * & \cdots & * \\ 0 & A_2 - xI_{n_2} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_s - xI_{n_s} \end{pmatrix} = \prod_{i=1}^{s} \chi_{A_i}(x).$$

□

Proposition 3.3.1. Let $A \in \mathbb{F}^{n\times n}$. Then:

- The leading coefficient of the characteristic polynomial $\chi_A(x)$ is $(-1)^n$.
- If $A = (a_{ij})$, then:

$$\chi_A(x) = \prod_{i=1}^{n}(a_{ii} - x) + \psi(x),$$

  where $\deg \psi(x) \leq n - 2$.

Proof. From the definition of the characteristic polynomial:

$$\chi_A(x) = \det(A - xI_n).$$

Applying Lemma 3.3.1, we know that the determinant expansion includes all terms of the form:

$$(-1)^\kappa \cdot (a_{1j_1} - x\delta_{1j_1}) \cdots (a_{nj_n} - x\delta_{nj_n}),$$

where $(j_1, \ldots, j_n)$ is a permutation of $\{1, \ldots, n\}$.
The only term that contains x^n is when $(j_1, \ldots, j_n) = (1, \ldots, n)$, i.e. the product:

$$(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x).$$

This yields a degree n polynomial with leading term $(-x)^n = (-1)^n x^n$.
All other terms contain at most $n - 1$ occurrences of x, hence they contribute to a polynomial of degree at most $n - 1$.
Moreover, by a direct observation:

$$\prod_{i=1}^{n}(a_{ii} - x) = (-1)^n x^n + \cdots + (-1)^1 \operatorname{tr} Ax + \det A,$$

so the difference:

$$\chi_A(x) - \prod_{i=1}^{n}(a_{ii} - x)$$

is a polynomial of degree at most $n-2$, i.e.:

$$\chi_A(x) = \prod_{i=1}^{n}(a_{ii} - x) + \psi(x), \quad \deg \psi(x) \leq n-2.$$

□

Proof. Consider the matrix

$$B = A - x\mathbb{I}_n = \begin{pmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{pmatrix}.$$

Consider the term

$$(-1)^\kappa b_{1j_1} \cdots b_{nj_n}$$

with

$$(j_1, \dots, j_n) \neq (1, \dots, n).$$

Hence there exists t with $j_t \neq t$, so

$$b_{tj_t} \neq a_{tt} - x.$$

Then $b_{tj_t} = a_{ts}$ for some $t \neq s$. Since a_{ts} lies in column s, in the original product there is no other element from column s. Hence the element $a_{ss} - x$ does not appear and therefore $\deg \psi(x) \leq n-2$.

It remains to show that in the sum of Lemma 3.3.1 the term $(a_{11} - x) \cdots (a_{nn} - x)$ appears with coefficient $+1$. Indeed: (i) This term appears in the sum (by induction on n and expansion along the first row), and (ii) it does not cancel with another one, due to its uniqueness. □

Example 3.3.2. For the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{F}^{2\times 2}$ we have:

$$\chi_A(x) = (a-x)(d-x) - bc.$$

Corollary 3.3.1. Let $A \in \mathbb{F}^{n\times n}$ and $\chi_A(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$. Then:

1. $\det A = a_0$ and $\operatorname{Tr}(A) = (-1)^{n-1} a_{n-1}$,
2. if $\chi_A(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$, then $\det A = \prod_{i=1}^{n} \lambda_i$ and $\operatorname{Tr}(A) = \sum_{i=1}^{n} \lambda_i$.

Proof. 1. It is clear that $\det A = \chi_A(0) = a_0$. From Proposition 3.3.1 we have:

$$\chi_A(x) = \prod_{i=1}^{n}(a_{ii} - x) + \psi(x), \quad \deg \psi(x) \leq n-2, \quad A = (a_{ij}).$$

Hence, by matching coefficients of x^{n-1} we obtain:

$$a_{n-1} = (-1)^n \mathrm{Tr}(A) \quad \Rightarrow \quad \mathrm{Tr}(A) = (-1)^{n-1} a_{n-1}.$$

2. From the characteristic polynomial we have $\chi_A(0) = \lambda_1 \cdots \lambda_n = \det A$. Also, from the above:

$$\mathrm{Tr}(A) = (-1)^n a_{n-1} = (-1)^n(-1)^n(\lambda_1 + \cdots + \lambda_n) = \lambda_1 + \cdots + \lambda_n.$$

□

Example 3.3.3. Let the matrix $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. From Corollary 3.3.1 (i) we have:

$$\chi_A(x) = x^2 - \mathrm{Tr}(A)x + \det A = x^2 - 3x - 10.$$

Moreover:

$$\chi_A(x) = (5-x)(-2-x), \text{ i.e. the eigenvalues are } \lambda_1 = 5, \lambda_2 = -2.$$

Hence, by (ii) of the corollary:

$$\mathrm{Tr}(A) = \lambda_1 + \lambda_2 = 3, \quad \det A = \lambda_1 \cdot \lambda_2 = -10.$$

Proposition 3.3.2. If $A, B \in \mathbb{F}^{n\times n}$ are similar, then $\chi_A(x) = \chi_B(x)$.

Proof. Since A is similar to B, there exists an invertible $P \in \mathbb{F}^{n\times n}$ such that $B = P^{-1}AP$. It is left as an exercise to the reader to show that for every $\varphi(x) \in \mathbb{F}[x]$ we have:

$$\varphi(B) = P^{-1}\varphi(A)P.$$

Therefore:

$$\chi_B(x) = \det(B - x\mathbb{I}_n) = \det(P^{-1}AP - x\mathbb{I}_n) = \det(P^{-1}(A - x\mathbb{I}_n)P) = \chi_A(x).$$

□

Definition 3.3.2. Let $f\colon V \to V$ be a linear map and $A = (f\colon \hat{a}, \hat{a})$ for some ordered basis $\hat{a}$ of V. The characteristic polynomial of f is defined by:

$$\chi_f(x) = \chi_A(x).$$

3.4 Chapter 3 Exercises.

Group A : $1, 2, 3, 4, 5, 6, 9, 10, 11, 13, 14, 18, 19, 20, 21, 32, 34, 35$,
Group B: $7, 8, 12, 15, 16, 17, 22, 23, 24, 25, 26, 29, 30, 32, 36$,
Group C : $27, 28$

Exercise 3.1. a. Let

$$A = \begin{pmatrix} 1 & 1 & 1 & -5 \\ 1 & 1 & -5 & 1 \\ 1 & -5 & 1 & 1 \\ -5 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{C}^{4\times 4} \quad \text{and} \quad X = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{C}^{4\times 1}.$$

Is X an eigenvector of A? Is 6 an eigenvalue of A?

b. Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 1 & -2 & 2 \\ 0 & -3 & 4 \\ 0 & -2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

Exercise 3.2. Let $A \in \mathbb{F}^{n\times n}$ and $\varphi(x) \in \mathbb{F}[x]$.

a. Show that if $\lambda \in \mathbb{F}$ is an eigenvalue of A with corresponding eigenvector X, then $\varphi(\lambda)$ is an eigenvalue of $\varphi(A)$ with corresponding eigenvector X.

b. Let $A = \begin{pmatrix} 2 & 0 & 3 \\ -3 & -2 & 4 \\ 1 & 0 & 3 \end{pmatrix}$. Find (without computations) an eigenvalue and a corresponding eigenvector of $B = A^{1821} + I_3$.

c. * Let $\mathbb{F} = \mathbb{C}$. Show that for every eigenvalue λ of $\varphi(A)$ there exists an eigenvalue λ_i of A such that $\lambda = \varphi(\lambda_i)$.

Exercise 3.3. Let $A = \begin{pmatrix} 5 & 3 & 3 \\ -3 & -1 & -3 \\ -3 & -3 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$ and $X = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \in \mathbb{R}^{3\times 1}$.

a. Is it true that X is an eigenvector of A? If yes, find two different bases of the eigenspace $V_A(\lambda)$, where λ is the eigenvalue corresponding to the above eigenvector.

b. Is it true that X is an eigenvector of $A^{1821} + I_3$?

c. Find a matrix $B \in \mathbb{R}^{3\times 3}$ with $X \in V_B(3)$.

Exercise 3.4. Find the eigenvectors of $A = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} \in \mathbb{F}^{2\times 2}$ in the cases

a. $\mathbb{F} = \mathbb{R}$

b. $\mathbb{F} = \mathbb{C}$.

Exercise 3.5. Find a basis for each eigenspace of the matrices

a. $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

b. $B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

Exercise 3.6. Compute, for the various values of a, the dimensions of the eigenspaces of $A = \begin{pmatrix} 1 & a & 4 \\ 0 & 1 & 0 \\ 0 & 2 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

Exercise 3.7. Let $A = (a_{ij}) \in \mathbb{F}^{n\times n}$ such that for every $j = 1, \cdots, n$, we have $\sum_{i=1}^{n} a_{ij} = 1$. Show the following.

a. There exists a nonzero $X \in \mathbb{F}^{n\times 1}$ such that $AX = X$.

b. If A is invertible and $A^{-1} = (b_{ij})$, then for every $j = 1, \cdots, n$, we have $\sum_{i=1}^{n} b_{ij} = 1$.

Exercise 3.8. Let $\lambda \neq \mu$ be two eigenvalues of a matrix $A \in \mathbb{F}^{n\times n}$ with corresponding eigenvectors u, v. Then

a. u, v are linearly independent and

b. for every $a, b \in \mathbb{F} - \{0\}$, $au + bv$ is not an eigenvector of A.

Exercise 3.9. a. Is it true that 2 is an eigenvalue of the linear map

$$f : \mathbb{R}^4 \to \mathbb{R}^4,\ f(x, y, z, w) = (x + w, 2y + z, 3z + w, x + w)\ ;$$

Is it true that $\big(1, 0, -1, 2\big)$ is an eigenvector of f?

b. Find the eigenvalues and eigenvectors of the linear map

$$f : \mathbb{R}^3 \to \mathbb{R}^3,\ f(x, y, z) = (x - y.2x + 3y + 2z, x + y + 2z).$$

c. Let $f : \mathbb{F}^2 \to \mathbb{F}^2$ be the linear map defined by $f(e_1) = -e_2$ and $f(e_2) = e_1$, where $\hat{e} = \{e_1, e_2)$ is the standard basis of $\mathbb{F}^2$. Compute the eigenvalues and eigenvectors of f when i. $\mathbb{F} = \mathbb{R}$ and ii. $\mathbb{F} = \mathbb{C}$. Give a geometric interpretation of the result in i. .

Exercise 3.10. a. Find the possible eigenvalues of the linear map $f : V \to V$ in each of the cases

i. $f^2 = 1_V$,
ii. $f^2 = f$.

b. Then prove the following statement. If $\varphi(f) = 0$ for some $\varphi(x) \in \mathbb{F}[x]$, then every eigenvalue of the $\mathbb{F}$-linear map $f\colon V \to V$ is a root of $\varphi(x)$.

c. Prove the following statement. If $\varphi(A) = 0$ for some $\varphi(x) \in \mathbb{F}[x]$ and $A \in \mathbb{F}^{n\times n}$, then every eigenvalue of A is a root of $\varphi(x)$.

Exercise 3.11. a. For which $a \in \mathbb{R}$ is $(1, 1)$ an eigenvector of the linear map

$$f : \mathbb{R}^2 \to \mathbb{R}^2,\ f(x, y) = (x + ay, 2x + y)\ ;$$

b. Find the eigenvalues and eigenvectors of the linear maps

1. $f : \mathbb{R}^3 \to \mathbb{R}^3,\ f(x, y, z) = (4x, 2y - 5z, y - 2z)$,
2. $g : \mathbb{C}^3 \to \mathbb{C}^3,\ f(x, y, z) = (4x, 2y - 5z, y - 2z)$.

Exercise 3.12. A linear map $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ is given, with $f(x^2 + x) = 2x^2 + 2x$, $f(x + 1) = 2x + 3$ and $f(1) = x + 3$.

a. Find the eigenvectors of f and a basis for each eigenspace of f.

b. Is it true that f is an isomorphism?

c. Is it true that $f^4 - 6f - 4 \cdot 1_{\mathbb{R}_2[x]}$ is an isomorphism?

d. Find two linearly independent eigenvectors of $f^4 - 6f - 4 \cdot 1_{\mathbb{R}_2[x]}$.

Exercise 3.13. Find the eigenvalues and eigenvectors of the linear maps

a. $g : \mathbb{R}_2[x] \to \mathbb{R}_2[x],\ g(\phi(x)) = \phi(1)x$

b. $h : \mathbb{R}_2[x] \to \mathbb{R}_2[x],\ h(\phi(x)) = \phi'(x)$, where $\phi'(x)$ is the derivative of $\phi(x)$.

Exercise 3.14. Let $A \in \mathbb{C}^{3\times 3}$ with $\chi_A(x) = -x^3 + 3x^2 - 2x$.

a. Is A invertible?

b. Is $(A - 3I_3)(A - 4I_3)$ invertible?

c. Compute the determinant of $A^2 - 2A - 15I_3$.

d. Is it true that there exists an ordered basis $\hat{a}$ such that for the linear map

$$f : \mathbb{C}^3 \to \mathbb{C}^3,\ f(x, y, z) = (x + 2y + 3z, 2y + 3z, 3z)\ ,$$

we have $(f : \hat{a}, \hat{a}) = A$?

e. Find $\chi_{A^2}(x)$.

f. Is it true that there exists $B \in \mathbb{C}^{3\times 3}$ such that $AB - BA = A^k$ for some positive integer k?

g. Is it true that there exists an integer $k > 1$ with $A^k = A^t$, where A^t is the transpose of A?

Exercise 3.15. Let $A, B \in \mathbb{F}^{n\times n}$, where A is invertible. Show that $\chi_{AB}(x) = \chi_{BA}(x)$. (Note. The conclusion also holds without the assumption that A is invertible, see exercise 27.)

Exercise 3.16. Let $A \in \mathbb{F}^{n\times n}$ be invertible and $x_A(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, where $a_0 \neq 0$. Show that

$$\chi_{A^{-1}}(x) = (-1)^n \left[x^n + \frac{a_{n-1}}{a_0} x^{n-1} + \cdots + \frac{a_1}{a_0} + \frac{(-1)^n}{a_0} \right].$$

Exercise 3.17. Let

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in \mathbb{F}^{n \times n}.$$

a. Show that the characteristic polynomial of A is $(-1)^n(x^n + a_{n-1}x^{n-1} + \cdots + a_0)$.

b. Show that if λ is an eigenvalue of A, then

$$\begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}$$

is an eigenvector of A^t.

Exercise 3.18. Find the characteristic polynomial of

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{F}^{n \times n}.$$

Exercise 3.19. Find the eigenvalues of the matrix

$$C = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix} \in \mathbb{C}^{5 \times 5}.$$

Exercise 3.20. Let $A \in \mathbb{C}^{n \times n}$ be invertible.

a. Show that λ is an eigenvalue of A if and only if $\dfrac{1}{\lambda}$ is an eigenvalue of A^{-1}.

b. Suppose A is similar to A^{-1} and n is odd. Show that 1 or -1 is an eigenvalue of A.

Exercise 3.21. Let $A \in \mathbb{C}^{4\times 4}$ such that $\chi_A(x) \in \mathbb{R}[x]$, $\det A = -13$, $\operatorname{Tr}(A) = 4$ and one eigenvalue of A is $2-3i$. Find the eigenvalues of A.

Exercise 3.22. Let $A \in \mathbb{C}^{n\times n}$ be invertible. Show that if A is similar to $-A$, then n is even, $n = 2m$, and the characteristic polynomial of A is of the form $(x^2-\rho_1)\cdots(x^2-\rho_m)$, where $n \geq 2$.

Exercise 3.23. Find the eigenspaces of the linear map $f\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},\ A \mapsto A^t$, where $n \geq 2$.

Exercise 3.24. Consider the diagonal matrices

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},\ B = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} \in \mathbb{F}^{n\times n}.$$

Show that the following statements are equivalent.

a. A, B are similar.

b. There exists a permutation $\sigma \in S_n$ such that $b_i = a_{\sigma(i)}$ for each $i = 1, \cdots, n$.

c. $\chi_A(x) = \chi_B(x)$.

Exercise 3.25. Let $a, b \in \mathbb{F}$. Find the characteristic polynomial, the eigenvalues, and the eigenvectors of

$$A = \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ b & b & b & \cdots & a \end{pmatrix} \in \mathbb{F}^{n\times n}.$$

Exercise 3.26. Let $a, b \in \mathbb{F}$ with $a \neq b$. Show that the characteristic polynomial of

$$A_n = \begin{pmatrix} 0 & a & a & \cdots & a \\ b & 0 & a & \cdots & a \\ b & b & 0 & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ b & b & b & \cdots & 0 \end{pmatrix} \in \mathbb{F}^{n\times n}$$

is $\dfrac{(-1)^n}{a-b}\left[a(x+b)^n - b(x+a)^n\right]$.

Exercise 3.27. Let $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times m}$. Show that $(-1)^n x^n \chi_{AB}(x) = (-1)^m x^m \chi_{BA}(x)$. (Consequently, if $m = n$, then $\chi_{AB}(x) = \chi_{BA}(x)$.)

Exercise 3.28. Let $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{F}$ and $C = (a_i bj) \in \mathbb{F}^{n\times n}$. Using the previous exercise (or otherwise) find $\chi_C(x)$ and the eigenvalues of C.

Exercise 3.29. Let $n \geq 1$ and

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix} \in \mathbb{F}^{2n\times 2n}.$$

Find the characteristic polynomial, the eigenvalues, and the eigenvectors of A. Find the dimension of each eigenspace of A.

Exercise 3.30. Let $A, B \in \mathbb{C}^{n\times n}$, $C = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$ and $D = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$. Then

a. $\chi_C(x) = \chi_{A+B}(x) \cdot \chi_{A-B}(x)$.

b. $\chi_D(x) = \chi_{A+iB}(x) \cdot \chi_{A-iB}(x)$.

c. If the eigenvalues of A are $\lambda_1, \cdots, \lambda_n$, then the eigenvalues of the matrix $\begin{pmatrix} A & A \\ A & A \end{pmatrix}$ are

$$2\lambda_1, \cdots, 2\lambda_n, \underbrace{0, \cdots 0}_{n}.$$

Exercise 3.31. Let $a, b \in \mathbb{R}$. It is given that the matrices $A, B \in \mathbb{R}^{3\times 3}$ are similar, where

$$A = \begin{pmatrix} 1 & a & 1 \\ a & 1 & b \\ 1 & b & 1 \end{pmatrix}, \; B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Find a, b.

Exercise 3.32. Find the characteristic polynomial of the linear map $f^2 = f \circ f$, where

$$f : \mathbb{R}^3 \to \mathbb{R}^3,\ f(x, y, z) = (0, x, y).$$

Exercise 3.33. Find the characteristic polynomial of the linear map $f^2 = f \circ f$, where

$$f : \mathbb{R}^3 \to \mathbb{R}^3,\ f(x, y, z) = (0, x, y).$$

Exercise 3.34. An ordered basis $\hat{u} = (u_1, u_2, u_3)$ of $\mathbb{R}^3$ is given and the linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ with corresponding matrix $A = (f : \hat{u}, \hat{u}) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

a. Find $\chi_f(x)$ and $\chi_{f^2}(x)$.

b. Is it true that $u_1 + u_2 + 2u_3$ is an eigenvector of f? Same question for u_1.

c. Find a basis for each eigenspace of A.

d. Find a basis for each eigenspace of f.

e. We know that $V_f(0) \subseteq V_{f^2}(0)$. Is it true that we have equality?

f. Is it true that there exists a linear map $g : \mathbb{R}^3 \to \mathbb{R}^3$ such that $f(g(v)) = v$ for every $v \in \mathbb{R}^3$?

Exercise 3.35. Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ and $f : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2},\ f(X) = AX - XA$. After showing that f is linear, find a basis for each eigenspace of f.

Exercise 3.36. Consider the vector space $F(\mathbb{R}, \mathbb{R})$ of functions $\mathbb{R} \to \mathbb{R}$ and the subspace V spanned by the functions $\sin x, \cos x$. Find a basis of each eigenspace of the linear maps

a. $f : V \to V,\ f(\phi(x)) = \phi'(x)$ (derivative),

b. $g : V \to V,\ g(\phi(x)) = \phi''(x)$ (second derivative).

Exercise 3.37. Show that for each

a. $A \in \mathbb{C}^{2\times 2},\ \chi_A(x) = x^2 - Tr(A)x + \det_A$,

b. $A \in \mathbb{C}^{3\times 3},\ \chi_A(x) = -x^3 + Tr(A)x^2 - Tr(adj(A))x + \det A$.

Exercise 3.38. Examine which of the following statements are true. In each case give a proof or a counterexample.

a. If λ is an eigenvalue of $A \in \mathbb{F}^{n\times n}$ and μ an eigenvalue of $B \in \mathbb{F}^{n\times n}$, then $\lambda + \mu$ is an eigenvalue of $A + B$.

b. If λ is an eigenvalue of $A \in \mathbb{F}^{n\times n}$ and μ an eigenvalue of $B \in \mathbb{F}^{n\times n}$, then $\lambda\mu$ is an eigenvalue of AB.

c. Every $A \in \mathbb{R}^{2\times 2}$ has at least one real eigenvalue.

d. Every $A \in \mathbb{R}^{3\times 3}$ has at least one real eigenvalue.

e. If 2 is an eigenvalue of A^2, where $A \in \mathbb{R}^{n\times n}$, then $\sqrt{2}$ is an eigenvalue of A.

f. If $\chi_A(x) = \chi_B$, where $A, B \in \mathbb{F}^{n\times n}$, then A, B are similar.

g. Let $A, B \in \mathbb{F}^{n\times n}$. Then $\phi(A), \phi(B)$ are similar for every $\phi(x) \in \mathbb{F}[x]$.

h. There exists $A \in \mathbb{F}^{3\times 3}$ with eigenvalues $0, 1, 2, 3$.

i. If v is an eigenvector of the linear map $f : V \to V$ and $v \in \ker f$, then 0 is an eigenvalue of f.

j. Let $A \in \mathbb{R}^{3\times 3}$ with $\chi_A(x) = -(x^2 - 1)(x - 5)$. Then there exists a linear map $f : \mathbb{R}^3 \to \mathbb{R}^3$ and an ordered basis $\hat{a}$ of $\mathbb{R}^3$ with $f(1, 0, 0) = 3 \cdot (1, 0, 0)$ and $(f : \hat{a}, \hat{a}) = A$.

k. Let $A \in \mathbb{F}^{n\times n}$. If -1 is an eigenvalue of A, then there exists a nonzero $X \in \mathbb{F}^{n\times 1}$ with $A^2X = X$.

CHAPTER 4

DIAGONALIZABILITY

4.1 Diagonalizable Matrices

Definition 4.1.1. Let $A \in \mathbb{F}^{n\times n}$. We say that A is **diagonalizable** if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that $P^{-1}AP = \Delta$, where Δ is a diagonal matrix.

Observation 4.1.1. Let $\Delta = P^{-1}AP$ with $\Delta = \text{diag}(a_1, a_2, \ldots, a_n)$. Since similar matrices have the same characteristic polynomial, we have:

$$\chi_A(x) = \chi_\Delta(x) = (a_1 - x)(a_2 - x)\cdots(a_n - x).$$

Here $a_1, a_2, \ldots, a_n$ are the eigenvalues of the matrix A.

Example 4.1.1. 1. Let $A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ and $P = \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix}$. Then we observe that:

$$P^{-1}AP = \text{diag}(-2, 5).$$

Hence the matrix A is diagonalizable.

2. Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. We will show that A is not diagonalizable. Indeed, if there existed an invertible matrix P with:

$$P^{-1}AP = \text{diag}(a_1, a_2),$$

then, by Remark 4.1.1 and since the eigenvalues of A are $a_1 = a_2 = 1$, we would have to have:

$$P^{-1}AP = \mathbb{I}_2 \quad \Leftrightarrow \quad A = \mathbb{I}_2,$$

which is a contradiction.

3. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$. The matrix A is not diagonalizable, because:

$$\chi_A(x) = x^2 + 1$$

has no roots in $\mathbb{R}$. Hence A has no eigenvalues in $\mathbb{R}$ and, according to Remark 4.1.1, it is not diagonalizable.

4. Let $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2}$. The matrix A is diagonalizable, since for the invertible matrix:

$$P = \begin{pmatrix} -1 & 1 \\ i & i \end{pmatrix}$$

we have:

$$P^{-1}AP = \text{diag}(i, -i).$$

Question 4.1.1. The examples raise some basic questions. Let $A \in \mathbb{F}^{n\times n}$.

1. When is A diagonalizable?
2. If A is diagonalizable, how do we find matrices P and Δ such that $P^{-1}AP = \Delta$?

If $A \in \mathbb{F}^{n\times n}$, we denote by $A^{(i)}$ the ***i*-th column** of A. For example, if

$$A = \begin{pmatrix} 1 & -3 \\ 4 & 2 \end{pmatrix},$$

then:

$$A^{(1)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \quad A^{(2)} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}.$$

With this notation we have:

$$A = \left(A^{(1)}, A^{(2)}, \cdots, A^{(n)}\right).$$

i. If $A, B \in \mathbb{F}^{n\times n}$, then $(AB)^{(i)} = AB^{(i)}$.

ii. Let $\hat{E} = (E_1, E_2, \cdots, E_n)$ be the standard ordered basis of $\mathbb{F}^{n\times 1}$. Then $I_n^{(i)} = E_i$ and in fact $A^{(i)} = AE_i$.

Proof. i. If $B = (B^{(1)}, B^{(2)}, \cdots, B^{(n)})$, then:

$$AB = (AB^{(1)}, AB^{(2)}, \cdots, AB^{(n)}),$$

and from the definition of matrix multiplication we have:

$$(AB)^{(i)} = AB^{(i)}.$$

ii. We observe that:

$$AE_i = A\mathbb{I}_n^{(i)} = (A\mathbb{I}_n)^{(i)} = A^{(i)}.$$

□

Observation 4.1.2. Let $A, P, \Delta \in \mathbb{F}^{n\times n}$ with P invertible such that $P^{-1}AP = \Delta$, where Δ is not necessarily diagonal. The following are equivalent:

i. The i-th column of P is an eigenvector of A with eigenvalue λ.

ii. The i-th column of Δ equals λE_i.

Proof. Assume that $P^{-1}AP = \Delta$. Then:

$$AP = P\Delta.$$

i. If $AP^{(i)} = \lambda P^{(i)}$, then:

$$\Delta^{(i)} = (P^{-1}AP)^{(i)} = P^{-1}AP^{(i)} = \lambda P^{-1}P^{(i)} = \lambda \mathbb{I}_n^{(i)} = \lambda E_i.$$

ii. If $\Delta^{(i)} = \lambda E_i$, then:

$$P\Delta^{(i)} = \lambda P E_i = \lambda P^{(i)},$$

so:

$$AP^{(i)} = \lambda P^{(i)}.$$

Since P is invertible, $P^{(i)} \neq 0$. □

Theorem 4.1.1. 1st Diagonalizability Criterion Let $A \in \mathbb{F}^{n\times n}$. The following are equivalent:

i. A is diagonalizable.

ii. There exists a basis of $\mathbb{F}^{n\times 1}$ consisting of eigenvectors of A.

Moreover, if $\{x_1, x_2, \dots, x_n\}$ is a basis of $\mathbb{F}^{n\times 1}$ of eigenvectors of A with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, then setting

$$P = (x_1, x_2, \dots, x_n) \in \mathbb{F}^{n\times n},$$

the matrix P is invertible and:

$$P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Proof. i. $\rightarrow$ ii. Assume there exists an invertible $P \in \mathbb{F}^{n\times n}$ such that $P^{-1}AP = \Delta$, where Δ is diagonal.
Since the i-th column of Δ has the form $\lambda_i E_i$, by Remark 4.1.2 it follows that the i-th column of P, i.e. $P^{(i)}$, is an eigenvector of A corresponding to eigenvalue λ_i.
Since the columns of P form a basis (because P is invertible), we obtain a basis of $\mathbb{F}^{n\times 1}$ consisting of eigenvectors of A.

ii. $\rightarrow$ i. Now assume that $\{x_1, x_2, \dots, x_n\}$ is a basis of $\mathbb{F}^{n\times 1}$ consisting of eigenvectors of A. Thus:

$$Ax_i = \lambda_i x_i, \quad \text{for each } i = 1, \dots, n.$$

Set $P = (x_1, x_2, \dots, x_n)$, so that $P^{(i)} = x_i$. Since the x_i form a basis, P is invertible. By Remark 4.1.2, it follows that:

$$(P^{-1}AP)^{(i)} = \lambda_i E_i,$$

hence $P^{-1}AP$ is diagonal. □

Example 4.1.2. a. Let

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}.$$

Then:

$$\chi_A(x) = (x+2)(x-5), \quad V_A(-2) = \left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\rangle, \quad V_A(5) = \left\langle \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\rangle.$$

The vectors
$$\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$
are linearly independent, since
$$\det \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix} \neq 0,$$
so they form a basis of $\mathbb{R}^{2\times 1}$ and A is diagonalizable. Setting:
$$P = \begin{pmatrix} -1 & 3 \\ 1 & 4 \end{pmatrix},$$
we obtain that P is invertible and $P^{-1}AP = \operatorname{diag}(-2, 5)$.

b. Let
$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 2 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$
We have:
$$\chi_A(x) = (2-x)^2(3-x), \quad V_A(2) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle, \quad V_A(3) = \left\langle \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\rangle.$$
We only have two linearly independent eigenvectors, while we need **three** for diagonalization, because we cannot produce a basis of $\mathbb{R}^{3\times 1}$ from eigenvectors of A. Therefore, A is **not** diagonalizable.

c. Let
$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$
We compute:
$$\chi_A(x) = (2+x)^2(4-x), \quad V_A(-2) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle, \quad V_A(4) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right\rangle.$$
The three eigenvectors are linearly independent, hence they form a basis of $\mathbb{R}^{3\times 1}$ and A is diagonalizable. Setting:
$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix},$$
we have that P is invertible and:
$$P^{-1}AP = \operatorname{diag}(-2, -2, 4).$$

Proposition 4.1.1. i. Let $A \in \mathbb{F}^{n\times n}$ be diagonalizable and let $\varphi(x) \in \mathbb{F}[x]$ be a polynomial. Then the matrix $\varphi(A)$ is also diagonalizable.

ii. If A is diagonalizable and invertible, then the matrix $\varphi(A^{-1})$ is also diagonalizable.

Proof. i. Since A is diagonalizable, there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that:

$$P^{-1}AP = \Delta = \text{diag}(\lambda_1, \dots, \lambda_n).$$

Since φ is a polynomial, we have:

$$\varphi(P^{-1}AP) = \varphi(\Delta), \quad \text{and} \quad \varphi(P^{-1}AP) = P^{-1}\varphi(A)P.$$

Thus:

$$P^{-1}\varphi(A)P = \varphi(\Delta).$$

The matrix $\varphi(\Delta)$ is also diagonal, since:

$$\varphi\left(\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\right) = \begin{pmatrix} \varphi(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \varphi(\lambda_n) \end{pmatrix}.$$

Hence $\varphi(A)$ is diagonalizable.

ii. Since A is invertible and diagonalizable, there exists an invertible P such that:

$$P^{-1}AP = \text{diag}(\lambda_1, \dots, \lambda_n), \quad \text{with } \lambda_i \neq 0 \text{ for each } i.$$

Then:

$$P^{-1}A^{-1}P = \left(P^{-1}AP\right)^{-1} = \text{diag}\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right).$$

So A^{-1} is also diagonalizable. Since φ is a polynomial, from part (i) it follows that $\varphi(A^{-1})$ is also diagonalizable.

□

4.2 The Major Diagonalizability Criterion

Lemma 4.2.1. If $X_i \in V(\lambda_i)$ for $i = 1, \dots, t$ and

$$X_1 + \cdots + X_t = 0,$$

then

$$X_1 = X_2 = \cdots = x_t = 0.$$

Proof. This result has been proved in Proposition 3.1.1. □

Corollary 4.2.1. 4.2.2 Eigenvectors corresponding to distinct eigenvalues are linearly independent.

Proof. This result has been proved in Corollary 3.1.2. □

Corollary 4.2.2. Let $A \in \mathbb{F}^{n\times n}$ and let λ_1, λ_2 be eigenvalues with $\lambda_1 \neq \lambda_2$. Then:

$$V_A(\lambda_1) \cap V_A(\lambda_2) = \{0\}.$$

Proof. The result follows immediately from Lemma 4.2.1. □

Lemma 4.2.2. Let $A \in \mathbb{F}^{n\times n}$ and let $\lambda_1, \dots, \lambda_t$ be distinct eigenvalues. Then:

i. If B_i is a basis of $V_A(\lambda_i)$ for each $i = 1, \dots, t$, then the set $B_1 \cup \cdots \cup B_t$ is a basis of

$$V_A(\lambda_1) + \cdots + V_A(\lambda_t).$$

ii. We have:

$$\dim\left(\sum_{i=1}^{t} V_A(\lambda_i)\right) = \sum_{i=1}^{t} \dim V_A(\lambda_i).$$

Proof. i. Let B_i be a basis of $V_A(\lambda_i)$ with

$$B_i = \{b_{i_1}, \dots, b_{i_{m_i}}\}.$$

By the definition of the sum of subspaces we have:

$$\sum_{i=1}^{t} V_A(\lambda_i) = \left\langle \bigcup_{i=1}^{t} B_i \right\rangle.$$

It suffices to show that $\bigcup_{i=1}^{t} B_i$ is linearly independent. Let:

$$\sum_{i=1}^{t}\sum_{j=1}^{m_i} a_{ij} b_{ij} = 0$$

with $a_{ij} \in \mathbb{F}$. By Corollary 4.2.2 and since eigenvectors belong to eigenspaces with different eigenvalues, we have

$$\sum_{j=1}^{m_i} a_{ij} b_{ij} = 0, \quad \text{for each} \quad i = 1, \ldots, t.$$

Since

$$B_i = \{b_{i_1}, \ldots, b_{i_{m_i}}\}$$

is a basis, all coefficients a_{ij} are zero; hence the union is linearly independent.

ii. From (i) we have that the union of bases of the eigenspaces is a basis of the sum of the eigenspaces. Therefore:

$$\dim \left(\sum_{i=1}^{t} V_A(\lambda_i) \right) = \left| \bigcup_{i=1}^{t} B_i \right| = \sum_{i=1}^{t} \dim V_A(\lambda_i).$$

□

Theorem 4.2.1 (Major Diagonalizability Criterion)**.** Let $A \in \mathbb{F}^{n\times n}$ and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of A. The following are equivalent:

i. A is diagonalizable.

ii. There exists a basis of $\mathbb{F}^{n\times 1}$ consisting of eigenvectors of A.

iii. $V_A(\lambda_1) + \cdots + V_A(\lambda_k) = \mathbb{F}^{n\times 1}$.

iv. $\dim V_A(\lambda_1) + \cdots + \dim V_A(\lambda_k) = n$.

v. The characteristic polynomial can be written as:

$$\chi_A(x) = (-1)^n (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}, \quad \text{with } \dim V_A(\lambda_i) = n_i.$$

Proof.

- Statements i. and ii. are equivalent by Theorem 4.1.1.
- iii. $\iff$ iv. By Lemma 4.2.2, we have:

$$\sum_{i=1}^{k} V_A(\lambda_i) = \mathbb{F}^{n\times 1} \Leftrightarrow \dim \left(\sum_{i=1}^{k} V_A(\lambda_i) \right) = n.$$

- ii $\Rightarrow$ iii. If there exists a basis of eigenvectors, then the eigenspaces cover the whole space.

- *iii.* $\Rightarrow$ *ii.* By Lemma 4.2.2, the union of bases of the eigenspaces is linearly independent, and since it covers the whole space, it is a basis of $\mathbb{F}^{n\times 1}$.

- i. $\Rightarrow$ v. Since A is diagonalizable, it is similar to a diagonal matrix diag$(a_1, \dots, a_n)$, hence:

$$\chi_A(x) = (a_1 - x)\cdots(a_n - x) = (-1)^n \prod_{i=1}^{k}(x - \lambda_i)^{n_i}.$$

  And for each i:

$$\dim V_A(\lambda_i) = n - \text{rank}(A - \lambda_i I_n) = n - \text{rank}(\Delta - \lambda_i I_n) = n_i.$$

- v. $\Rightarrow$ iv. From the assumption $\dim V_A(\lambda_i) = n_i$ for each i, we obtain:

$$\sum_{i=1}^{k} \dim V_A(\lambda_i) = \sum_{i=1}^{k} n_i = n.$$

□

Corollary 4.2.3. If $A \in \mathbb{F}^{n\times n}$ has n distinct eigenvalues, then it is diagonalizable.

Proof. The claim follows immediately from (iv) of Theorem 4.2.1 □

Example 4.2.1. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Then, by Corollary 4.2.3, the matrix A is diagonalizable, since its eigenvalues are $1, 4, 6$, which are distinct.

Definition 4.2.1. Let λ be an eigenvalue of the matrix A. If $(x - \lambda)^{m(\lambda)}$ is the largest power of $x - \lambda$ dividing the characteristic polynomial $\chi_A(x)$, then:

- the number τ (λ) is called the **algebraic multiplicity** of λ,

- while **d** $(\lambda) := \dim V_A(\lambda)$ is called the **geometric multiplicity** of λ.

Theorem 4.2.2. Let $A \in \mathbb{F}^{n\times n}$ and let λ be an eigenvalue of A. Then:

$$\mathbf{d}(\lambda) \leq \tau(\lambda),$$

i.e. the geometric multiplicity of λ is less than or equal to its algebraic multiplicity.

Proof. Let $\tau(\lambda)$ be the algebraic multiplicity of λ, and let $\{v_1, \ldots, v_t\}$ be a basis of the eigenspace $V_A(\lambda)$. By the Basis Extension Theorem, there exists a basis of $\mathbb{F}^{n\times 1}$ of the form:

$$\hat{v} = (v_1, \ldots, v_t, v_{t+1}, \ldots, v_n).$$

Consider the linear map:

$$\mathcal{L}_A\colon \mathbb{F}^{n\times 1} \to \mathbb{F}^{n\times 1}, \quad X \mapsto AX.$$

Then the matrix of $\mathcal{L}_A$ with respect to the basis $\hat{v}$ is:

$$B = (L_A\colon \hat{v}, \hat{v}) = \begin{pmatrix} \lambda\mathbb{I}_t & * \\ 0 & * \end{pmatrix}.$$

This is in block upper-triangular form. From the properties of the characteristic polynomial (e.g. Proposition 3.3.1), we have:

$$\chi_B(x) = \chi_{\lambda I_t}(x) \cdot \chi_*(x).$$

Hence,

$$(x-\lambda)^t \mid \chi_B(x) = \chi_A(x)$$

since A and B are similar. [1] Therefore,

$$t \leq \tau(\lambda).$$

□

4.3 Diagonalizable Linear Maps

Definition 4.3.1. A linear map $f\colon V \to V$ is called **diagonalizable** if there exists an ordered basis $\hat{a}$ of V such that the matrix $(f\colon \hat{a}, \hat{a})$ is diagonal.

Observation 4.3.1. i. The map f is diagonalizable if and only if there exists a basis of V consisting of eigenvectors of f.

ii. Let $f\colon V \to V$ be a linear map and let $\hat{b}$ be an ordered basis of V. Then f is diagonalizable if and only if the matrix $\left(f\colon \hat{b}, \hat{b}\right)$ is diagonalizable.

[1] This holds because the matrix of $\mathcal{L}_A$ with respect to the standard basis is A.

Proof. i. If f is diagonalizable, then there exists an ordered basis $\hat{a} = (a_1, \dots, a_n)$ of V with $(f\colon \hat{a}, \hat{a}) = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, i.e. $f(a_i) = \lambda_i a_i$. Hence there exists a basis of eigenvectors.

Conversely, if $\hat{a}$ is a basis of eigenvectors, then the matrix $A = (f\colon \hat{a}, \hat{a})$ has a basis of $\mathbb{F}^{n\times 1}$ consisting of eigenvectors and is diagonalizable (Theorem 4.1.1). Hence f is diagonalizable.

ii. If f is diagonalizable, then there exists a basis $\hat{a}$ such that $A = (f : \hat{a}, \hat{a})$ is diagonal. But the matrices $(f : \hat{a}, \hat{a})$ and $(f : \hat{b}, \hat{b})$ are similar, hence the latter is also diagonalizable.

Conversely, if $(f : \hat{b}, \hat{b})$ is diagonalizable, then it is similar to a diagonal matrix, i.e. there exists a basis $\hat{a}$ of V such that $(f : \hat{a}, \hat{a})$ is diagonal. □

Example 4.3.1. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map given by:

$$f(x, y) = (x + 3y,\ 4x + 2y).$$

Let the basis $\hat{a} = (a_1, a_2)$, where $a_1 = (1, -1)$ and $a_2 = (3, 4)$. Then:

$$f(a_1) = -2a_1, \quad f(a_2) = 5a_2.$$

That is, both are eigenvectors. Hence the matrix $(f : \hat{a}, \hat{a})$ is:

$$\begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix},$$

which is diagonal. Therefore, f is diagonalizable.

Reminder 4.3.1. Let $f\colon V \to V$ be a linear map, let $\hat{a}$ be an ordered basis of V, and let $A = (f : \hat{a}, \hat{a})$. If λ is an eigenvalue of f, then the map:

$$\varphi\colon V \to \mathbb{F}^{n\times 1}, \quad v \mapsto [v]_{\hat{a}}$$

is a vector space isomorphism and moreover:

$$\varphi(V_f(\lambda)) = V_A(\lambda).$$

Theorem 4.3.1 (Major Diagonalizability Criterion for linear maps)**.** Let $f\colon V \to V$ be a linear map and let $\lambda_1, \dots, \lambda_k$ be its distinct eigenvalues. The following are equivalent:

i. f is diagonalizable.

ii. There exists a basis of V consisting of eigenvectors of f.

iii. $V_f(\lambda_1) + V_f(\lambda_2) + \cdots + V_f(\lambda_k) = V$.

iv. $\dim V_f(\lambda_1) + \cdots + \dim V_f(\lambda_k) = \dim V$.

v. $\chi_f(x) = (-1)^n (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k}$ with $\dim V_f(\lambda_i) = n_i$ for each $i = 1, \ldots, k$.

Proof.

- i. $\Leftrightarrow$ ii. By Remark 4.3.1.
- i. $\Rightarrow$ iii. If f is diagonalizable, then for a suitable basis $\hat{a}$ the matrix $A = (f : \hat{a}, \hat{a})$ is diagonalizable, hence by Theorem 4.2.1 we have $\sum V_A(\lambda_i) = \mathbb{F}^{n\times 1}$. By Reminder 4.3.1, it follows that $V = \sum V_f(\lambda_i)$.
- iii. $\Rightarrow$ i. (Exercise)
- iii. $\Rightarrow$ iv. From the above, using Reminder 4.3.1 and Theorem 4.2.1, we obtain equality of dimensions.
- iv. $\Rightarrow$ iii. (Exercise)
- iv. $\Rightarrow$ v. (Exercise)

□

4.4 Applications of Diagonalization

Diagonalization has wide applications: in computing powers of matrices, in recurrence relations of sequences, in matrix roots, in systems of differential equations, and in many other problems.

Observation 4.4.1. If $P^{-1}AP = \Delta = \text{diag}(\lambda_1, \ldots, \lambda_n)$, then for every $m \geq 1$:

$$(P^{-1}AP)^m = \Delta^m \quad \Leftrightarrow \quad P^{-1}A^m P = \text{diag}(\lambda_1^m, \ldots, \lambda_n^m).$$

Application 4.4.1 (Matrix Powers)**.** Let

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix}.$$

Compute the matrix A^m for every $m \in \mathbb{N}$.

Proof. With straightforward computations we find:

$$V_A(-1) = \left\langle \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\rangle, \quad V_A(1) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle, \quad V_A(2) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle.$$

The matrix A has three distinct eigenvalues, hence it is diagonalizable.

Set:

$$P = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, \quad \text{with } P^{-1}AP = \text{diag}(-1, 1, 2).$$

Then for every $m \in \mathbb{N}$ we have:

$$A^m = P \cdot \text{diag}((-1)^m, 1, 2^m) \cdot P^{-1}.$$

Finally, we obtain:

$$A^m = \begin{pmatrix} 2^m & 1 - 2^m & 1 - 2^m \\ 0 & (-1)^m & 0 \\ 0 & 1 - (-1)^m & 1 \end{pmatrix}.$$

□

Let us now consider the Fibonacci-type recurrence:

$$F_1 = 1, \quad F_2 = 2, \quad F_{n+1} = F_n + F_{n-1}.$$

This sequence often appears in counting problems. For example:

How many binary sequences of length n do not contain two consecutive 1's?

This problem translates into a recurrence, which can be analyzed via diagonalization of a suitable matrix.

Application 4.4.2 (Recurrences – Fibonacci sequence)**.** We have $F_1 = 1$, $F_2 = 2$, $F_{n+1} = F_n + F_{n-1}$ for every $n \geq 2$. We observe that:

$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix}.$$

By induction one proves that:

$$\begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = A^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{where } A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

We compute the characteristic polynomial:

$$\chi_A(x) = x^2 - x - 1,$$

with roots (eigenvalues):

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

The eigenspaces of A are:

$$V_A(\lambda_1) = \left\langle \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \right\rangle, \quad V_A(\lambda_2) = \left\langle \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right\rangle.$$

Set:

$$P = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}, \quad \Delta = \mathrm{diag}(\lambda_1, \lambda_2).$$

Then:

$$A = P\Delta P^{-1} \Rightarrow A^m = P\Delta^m P^{-1}, \quad \text{for every } m \in \mathbb{N}.$$

Therefore, for every $n \geq 3$:

$$\begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = A^{n-2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^n - \lambda_2^n \\ \lambda_1^{n+1} - \lambda_2^{n+1} \end{pmatrix}.$$

Hence:

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right].$$

Application 4.4.3 (Matrix Roots)**.** Let

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix}.$$

Find a matrix $B \in \mathbb{F}^{3\times 3}$ such that $B^3 = A$.

Proof. From Application 4.4.1 we have

$$A = P \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} P^{-1},$$

where

$$P = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Then, setting

$$B = P\text{diag}\left(-1, 1, \sqrt[3]{2}\right) P^{-1}$$

we obtain $B^3 = A$, as required. $\square$

4.5 Exercises of Chapter 4.

Group A: 1, 3, 4, 5, 8, 9, 10, 14, 15, 16, 17, 18, 25, 32, 36
Group B: 2, 6, 7, 12, 13, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 33, 34, 35

Exercise 4.1. Examine which of the following matrices are diagonalizable. If some $A_i \in \mathbb{F}^{n\times n}$ is diagonalizable, find a basis of $\mathbb{F}^{n\times 1}$ consisting of eigenvectors of A_i, an invertible $P_i \in \mathbb{F}^{n\times n}$ such that $P_i^{-1}A_iP_i$ is diagonal, and the matrix $P_i^{-1}A_iP_i$.

a. $A_1 = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$,

b. $A_2 = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$,

c. $A_3 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$,

d. $A_4 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

Exercise 4.2. Let $A \in \mathbb{F}^{n\times n}$ be a diagonalizable matrix.

a. Show that for every positive integer k the matrix A^k is diagonalizable, and more generally that for every $\phi(x) \in \mathbb{F}[x]$ the matrix $\phi(A)$ is diagonalizable.

b. Show that if $A^k = 0$ for some positive integer k, then $A = 0$.

c. Show that if A is invertible, then $\phi(A^{-1})$ is diagonalizable for every $\phi(x) \in \mathbb{F}[x]$.

d. If $\chi_A(x) = (x-3)^{10}$, find A.

e. Let $X \in \mathbb{F}^{n\times 1}$ with $A^kX = 0$ for some positive integer k. Show that $AX = 0$.

f. Suppose that A is invertible and $\mathbb{F} = \mathbb{R}$. Is it possible that $A+A^{-1}$ is similar to $diag(1, 3, 3, \cdots, 3)$?

Exercise 4.3. Let

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \in \mathbb{F}^{3\times 3}.$$

a. Find the eigenvalues of A, a basis for each eigenspace of A, and the dimension of the vector space generated by the eigenvectors of A.

b. Determine whether A is diagonalizable, and if it is diagonalizable, find an invertible matrix P such that $P^{-1}AP$ is diagonal.

Exercise 4.4. Let $A = \begin{pmatrix} 4 & a \\ 3 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$.

a. Prove that the matrix A is diagonalizable if and only if $a > -1/12$.

b. Let $a = 2$. Find invertible matrices $P, Q \in \mathbb{R}^{2\times 2}$ such that $P^{-1}AP$ and $Q^{-1}AQ$ are distinct diagonal matrices.

Exercise 4.5. a. Let $A \in \mathbb{R}^{n\times n}$ be a diagonalizable matrix whose eigenvalues are nonnegative. Show that there exists $B \in \mathbb{R}^{n\times n}$ such that $B^2 = A$.

b. Show that $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ is not diagonalizable and that there is no $B \in \mathbb{R}^{2\times 2}$ such that $B^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Exercise 4.6. Let $A, P, \Delta \in \mathbb{F}^{n\times n}$ such that $AP = P\Delta$ and Δ is diagonal, $\Delta = \text{diag}(\lambda_1, \cdots, \lambda_n)$.

a. Show that for each $k = 1, \cdots, n$ we have $AP^{(k)} = \lambda_k P^{(k)}$, where $P^{(k)}$ is the k-th column of P.

b. Let $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{F}$. Find a matrix $A \in \mathbb{F}^{3\times 3}$ with eigenvalues $\lambda_1, \lambda_2, \lambda_3$ and corresponding eigenvectors

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Is A unique?

Exercise 4.7. Let $A = \begin{pmatrix} * & 0 & * & 0 \\ * & 3 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 4 \end{pmatrix} \in \mathbb{C}^{4\times 4}$ with $\det A = \text{Tr}A = 0$. Show that A is diagonalizable.

Exercise 4.8. Let $A \in \mathbb{F}^{n\times n}$ be an upper triangular matrix of the form

$$A = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix},$$

that is, A is upper triangular and every diagonal entry equals λ. Show that A is diagonalizable if and only if it is diagonal.

Exercise 4.9. Determine whether

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in \mathbb{C}^{n\times n}$$

is diagonalizable.

Exercise 4.10. Find the values of $a, b, c \in \mathbb{R}$ such that

$$A = \begin{pmatrix} 3 & 0 & 0 \\ a & 3 & 0 \\ b & c & -2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

is diagonalizable.

Exercise 4.11. Find the values of $a \in \mathbb{R}$ such that the dimension of the vector space generated by the eigenvectors of

$$A = \begin{pmatrix} 0 & 1 & 0 \\ a & 0 & a \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

is equal to 3.

Exercise 4.12. Let $A, B \in \mathbb{F}^{n\times n}$ such that $AB = BA$. Prove that if A has n distinct eigenvalues, then B is diagonalizable.

Exercise 4.13. Let $A, B \in \mathbb{F}^{n\times n}$ be two diagonalizable matrices. Show that A, B are similar if and only if $\chi_A(x) = \chi_B(x)$.

Exercise 4.14. Find all $a \in \mathbb{F}$ such that the linear map $f : \mathbb{F}^3 \to \mathbb{F}^3$ is diagonalizable in the following cases:

a. $f(x, y, z) = (x + az, 2y, ay + 2z)$,

b. $f(x, y, z) = (ax + y + z, x + ay + z, x + y + az)$.

Exercise 4.15. Determine which of the following linear maps are diagonalizable:

a. $f : \mathbb{F}^3 \to \mathbb{F}^3,\ f(x, y, z) = (x + y, y - z, 2y + 4z)$,

b. $g : \mathbb{F}^3 \to \mathbb{F}^3,\ g(x, y, z) = (2x + y, y - z, 2y + 4z)$,

c. $h : \mathbb{F}_2[x] \to \mathbb{F}_2[x],\ h(\phi(x)) = \phi(1)x$.

Exercise 4.16. Let $f : V \to V$ be a diagonalizable linear map such that $\lambda \in \{-1, 1\}$ for every eigenvalue λ of f. Show that $f^2 = 1_V$.

Exercise 4.17. Let $f : V \to V$ be an isomorphism. Prove the following.

a. If $\lambda \in \mathbb{F}$ is an eigenvalue of f, then $\lambda \neq 0$.

b. $\lambda \in \mathbb{F}$ is an eigenvalue of $f \Leftrightarrow \lambda^{-1}$ is an eigenvalue of f^{-1}.

c. For every $\lambda \in \mathbb{F} - \{0\},\ V_f(\lambda) = V_{f^{-1}}(\lambda^{-1})$.

d. f is diagonalizable $\Leftrightarrow f^{-1}$ is diagonalizable.

Exercise 4.18. Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear map such that there exists an ordered basis $\hat{a} = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ with

$$(f : \hat{a}, \hat{a}) = \begin{pmatrix} 0 & 0 & \lambda_1 \\ 0 & \lambda_2 & 0 \\ \lambda_3 & 0 & 0 \end{pmatrix} .$$

a. Show that f^2 is diagonalizable.

b. Is it true that f is diagonalizable?

c. Suppose that $\lambda_1, \lambda_3 > 0$. Show that $\sqrt{\lambda_1} v_1 + \sqrt{\lambda_3} v_3$ is an eigenvector of f.

Exercise 4.19. Let $f : V \to V$ be a diagonalizable linear map. Show that $\ker f = \ker f^m$ and $\mathrm{Im} f = \mathrm{Im} f^m$ for some positive integer m.

Exercise 4.20. For every positive integer k compute A^k, where

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 0 & -5 & 6 \\ 0 & -3 & 4 \end{pmatrix}.$$

Exercise 4.21. Let

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 2 & 1 \end{pmatrix}.$$

a. Compute the power A^k, $k \geq 1$.

b. Find a matrix $B \in \mathbb{R}^{3\times 3}$ such that $B^3 = A$.

c. How many matrices $B \in \mathbb{C}^{3\times 3}$ can you find such that $B^3 = A$?

Exercise 4.22. Consider the sequence $(a_n),\ n = 1, 2, \cdots$, defined by the terms $a_1 = 1,\ a_2 = 4$ and the recurrence relation $a_n = 2a_{n-1} + 3a_{n-2},\ n = 3, 4, \cdots$. Find the general term a_n in terms of a_1, a_2 and n.

Exercise 4.23.

a. Let $A \in \mathbb{R}^{n\times n}$ be diagonalizable such that $|\lambda| \geq 2$ for every eigenvalue of A. Show that there exists an invertible $B \in \mathbb{R}^{n\times n}$ such that $B + B^{-1} = A$.

b. Show that there is no invertible $B \in \mathbb{R}^{3\times 3}$ such that $B + B^{-1} = I_3$.

Exercise 4.24. Assume that $n \geq 2$.

a. Show that $\mathbb{R}^{n\times n} = U \oplus V$, where $U = \{A \in \mathbb{R}^{n\times n} : A = A^t\},\ V = \{A \in \mathbb{R}^{n\times n} : A = -A^t\}$. Also show that $\dim U = \dfrac{n(n+1)}{2}$, $\dim V = \dfrac{n(n-1)}{2}$.

b. Using the above, prove that the linear map

$$f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n},\ A \mapsto A^t\ ,$$

is diagonalizable and find its characteristic polynomial.

Exercise 4.25. Let $f, g : V \to V$ be two linear maps such that f is diagonalizable and every eigenvector of f is an eigenvector of g. Show that $f \circ g = g \circ f$.

Exercise 4.26. Let $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{F}$ such that the matrix

$$A = \begin{pmatrix} a_1b_1 & \cdots & a_1b_n \\ \vdots & & \vdots \\ a_nb_1 & \cdots & a_nb_n \end{pmatrix} \in \mathbb{F}^{n\times n}$$

is nonzero.

a. Show that $rankA = 1$.

b. Show that A is diagonalizable if and only if $TrA \neq 0$.

Exercise 4.27. Show that the matrix

$$A = \begin{pmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & & \vdots \\ b & b & b & \cdots & a \end{pmatrix} \in \mathbb{F}^{n\times n}.$$

is diagonalizable.

Exercise 4.28. Let $a \in \mathbb{F}$ and let $\hat{\beta} = (v_1, v_2, v_3)$ be an ordered basis of $\mathbb{F}^3$. Consider the linear map $f : \mathbb{F}^3 \to \mathbb{F}^3$ defined by

$$f(v_1) = v_1,\ f(v_2) = 2v_1 - av_2 - v_3,\ f(v_3) = a^2v_2 + av_3.$$

a. Show that f is not diagonalizable.

b. Show that f^n is diagonalizable for every $n \geq 2$.

Exercise 4.29. Let $n \geq 2$. Let $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{F}$ such that not all of them are equal to 0 and $\sum\limits_{i=1}^{n-1} a_ib_i = 0$. Compute the characteristic polynomial of the matrix

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & a_1 \\ 0 & 0 & \cdots & 0 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n-1} \\ b_1 & b_2 & \cdots & b_{n-1} & 0 \end{pmatrix} \in \mathbb{F}^{n\times n}$$

and show that it is not diagonalizable.

Exercise 4.30. Determine which of the following statements are true or false. Justify your answer.

a. There exists a diagonalizable linear map $f : \mathbb{F}^4 \to \mathbb{F}^4$ such that $\chi_f(x) = x^2(x-3)^2$ and $\dim \mathrm{Im} f = 3$.

b. For every $a, b \in \mathbb{R}$, the matrices $\begin{pmatrix} 4 & a \\ 0 & 5 \end{pmatrix}, \begin{pmatrix} 5 & 0 \\ b & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ are similar.

c. Let $f : V \to V$ be a linear map. If $\lambda \neq \mu$ are two eigenvalues of f, then the linear map

$$g : V(\lambda) \oplus V(\mu) \to V(\lambda) \oplus V(\mu),\ g(u+v) = f(u+v)\ ,$$

is diagonalizable.

Exercise 4.31. Let $A \in \mathbb{F}^{n\times n}$ with rank$A = r$. Prove that the characteristic polynomial of A has the form

$$(-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_{n-r}x^{n-r}.$$

Exercise 4.32. Let $A \in \mathbb{C}^{2\times 2}$ and let λ, μ be the eigenvalues of A. Show that if $\lambda \neq \mu$, then for every positive integer k,

$$A^k = \frac{\lambda^k}{\lambda - \mu}(A - \mu\mathbb{I}_2) + \frac{\mu^k}{\mu - \lambda}(A - \lambda\mathbb{I}_2).$$

Exercise 4.33. Let $A \in \mathbb{F}^{n\times n}$ with rank$A = 1$ and $n \geq 2$. Prove the following statements.

a. A is similar to a matrix of the form

$$\begin{pmatrix} 0 & \cdots & 0 & a_1 \\ 0 & \cdots & 0 & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & a_n \end{pmatrix}.$$

b. $TrA \neq 0 \ \Leftrightarrow A$ is diagonalizable.

Exercise 4.34. Consider the linear map $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ defined by $f(x^2+1) = x+1,\ f(x+1) = x+1\ f(1) = x+1$. Set $g = f^{1821} + 2 \cdot 1_V,\ V = \mathbb{R}_2[x]$.

a. Find a basis for each eigenspace of f and each eigenspace of g.

b. Determine whether f, g are diagonalizable.

c. Determine whether f, g are isomorphisms.

Exercise 4.35. Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ and let $f : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2},\ f(X) = AX - XA$. Determine whether f is diagonalizable.

Exercise 4.36. If $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the eigenvalues of an invertible $A \in \mathbb{C}^{4\times 4}$, then the eigenvalues of $adj A$ are $\lambda_1\lambda_2\lambda_3,\ \lambda_1\lambda_2\lambda_4,\ \lambda_1\lambda_3\lambda_4,\ \lambda_2\lambda_3\lambda_4$.

Exercise 4.37. Determine which of the following statements are true. In each case, give a proof or a counterexample.

a. Every matrix that is similar to a diagonalizable matrix is diagonalizable.

b. If $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = x(x+1)(x^2+1)$, then A is diagonalizable.

c. If $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = x(x+1)(x^2+1)$, then A is diagonalizable.

d. Let $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = x^2(x-1)(x-2)$. Then A is diagonalizable if and only if $\dim V_A(0) > 1$.

e. If $A, B \in \mathbb{F}^{n\times n}$ are diagonalizable, then $A + B$ is diagonalizable.

f. If $A, B \in \mathbb{F}^{n\times n}$ are diagonalizable, then AB is diagonalizable.

g. Every invertible matrix is diagonalizable.

h. The dimension of the subspace generated by the eigenvectors of $A = \begin{pmatrix} * & 0 & * & 0 \\ * & 3 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 4 \end{pmatrix}$ is at least 2.

CHAPTER 5

TRIANGULARIZABLE MATRICES

5.1 Triangularizable Matrices

Definition 5.1.1. A matrix $A \in \mathbb{F}^{n\times n}$ is called **triangularizable** if there exists an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that $P^{-1}AP = T$ is upper triangular.

Example 5.1.1. The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, which is not diagonalizable (show why), is clearly triangularizable, since it is upper triangular.

Observation 5.1.1.

i. If $A \in \mathbb{F}^{n\times n}$ is triangularizable, then $\chi_A(x)$ is a product of linear factors in $\mathbb{F}[x]$.

ii. If A is diagonalizable, then it is also triangularizable.

Proof.

i. A is triangularizable, i.e., there exists an invertible $P \in \mathbb{F}^{n\times n}$ satisfying

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix},$$

hence we conclude that

$$\chi_A(x) = \chi_{P^{-1}AP}(x) = (\lambda_1 - x)\cdots(\lambda_n - x).$$

ii. This follows immediately from the definition. Attention! **The converse is not true.**

□

Reminder 5.1.1. Let $A, B, P \in \mathbb{F}^{n\times n}$ be matrices such that $P^{-1}AP = B$. The following are equivalent:

i. $P^{(i)}$ is an eigenvector of A with eigenvalue λ.

ii. $B^{(i)} = \lambda E_i$, where $\{E_1, \cdots, E_n\}$ is the standard basis of $\mathbb{F}^{n\times 1}$.

Theorem 5.1.1. A matrix $A \in \mathbb{F}^{n\times n}$ is triangularizable if and only if $\chi_A(x)$ is a product of linear factors.

Proof. If A is triangularizable, then $\chi_A(x)$ is a product of linear factors by Notation 5.1.1.

Conversely, assume that

$$\chi_A(x) = (\lambda_1 - x)\cdots(\lambda_n - x).$$

We use induction on n.

- **Base case.** For $n = 1$, the claim is immediate.
- **Inductive step.** Assume the claim holds for matrices $B \in \mathbb{F}^{(n-1)\times(n-1)}$ whose characteristic polynomial is a product of linear factors.

  Let $u_1 \in \mathbb{F}^{n\times 1}$ be an eigenvector corresponding to the eigenvalue λ_1. Then there exists a basis of $\mathbb{F}^{n\times 1}$ of the form
  $$u = \{u_1, \ldots, u_n\}.$$
  Define the matrix P_1 whose columns are the vectors u_i, i.e.,
  $$P_1^{(i)} = u_i.$$
  Then P_1 is invertible.

  By Reminder 5.1.1, we obtain:
  $$P_1^{-1}AP_1 = \begin{pmatrix} \lambda_1 & * \\ 0 & B_1 \end{pmatrix}, \quad B_1 \in \mathbb{F}^{(n-1)\times(n-1)}.$$

  Moreover,
  $$\chi_A(x) = (\lambda_1 - x)\cdot \chi_{B_1}(x),$$

so $\chi_{B_1}(x)$ is also a product of linear factors. Hence, by the inductive hypothesis, there exists an invertible matrix P_2 such that

$$P_2^{-1} B_1 P_2 = T,$$

with T upper triangular.

Set

$$P = P_1 \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_2 \end{array}\right),$$

which is invertible. Then:

$$\begin{aligned} P^{-1}AP &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_2 \end{array}\right)^{-1} \cdot P_1^{-1} A P_1 \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_2 \end{array}\right) \\ &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_2 \end{array}\right)^{-1} \cdot \begin{pmatrix} \lambda_1 & * \\ 0 & B_1 \end{pmatrix} \cdot \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & P_2 \end{array}\right) \\ &= \begin{pmatrix} \lambda_1 & * \\ 0 & T \end{pmatrix}, \end{aligned}$$

which is upper triangular.

□

Example 5.1.2. Consider the matrix

$$A = \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix}.$$

Then $\chi_A(x) = x^2$, i.e. A is triangularizable. By straightforward computations we find

$$V_A(0) = \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\rangle.$$

Take any invertible P with

$$P^{(1)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in V_A(0),$$

for example

$$P = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Then we obtain

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

Example 5.1.3. Consider the matrix

$$B = \begin{pmatrix} 3 & 4 & 5 \\ 0 & -2 & 1 \\ 0 & -4 & 2 \end{pmatrix}.$$

Using the previous example, where

$$P^{-1}AP = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}, \ A = \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix}, \ P = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$$

and the idea of the proof of Theorem 5.1.1, we set

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$

The matrix Q is invertible, and in fact we know that

$$Q^{-1}BQ = \begin{pmatrix} 3 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}.$$

Example 5.1.4. Consider the matrix

$$A = \begin{pmatrix} 2 & 2 & 0 \\ -1 & 5 & 1 \\ 1 & -1 & 5 \end{pmatrix}.$$

By straightforward computations we find:

$$\chi_A(x) = -(x-4)^3$$

and

$$V_A(4) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

Consider a basis of $\mathbb{R}^{3\times 1}$ that contains

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

for example

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Set

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is invertible, and

$$P^{-1}AP = \begin{pmatrix} 4 & * \\ 0 & B_1 \end{pmatrix},$$

where

$$B_1 = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}.$$

Continue similarly with B_1, where $\chi_{B_1}(x) = (x-4)^2$ and

$$V_{B_1}(4) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle,$$

by setting

$$P_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then P_2 is invertible and

$$P_2^{-1}B_1P_2 = \begin{pmatrix} 4 & * \\ 0 & 4 \end{pmatrix}.$$

Finally, if

$$P = P_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

then P is invertible and moreover

$$P^{-1}AP = \begin{pmatrix} 4 & * & * \\ 0 & 4 & * \\ 0 & 0 & 4 \end{pmatrix}.$$

Example 5.1.5. Consider the matrix

$$A = \begin{pmatrix} 0 & 0 & -3 \\ -1 & 3 & 1 \\ 1 & 0 & 4 \end{pmatrix}$$

with $\chi_A(x) = (1-x)(3-x)^2$.

Method A

We compute that

$$V_A(1) = \left\langle \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \right\rangle, \quad V_A(3) = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

Thus we consider the matrix

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

which is obtained by extending the set $\left\{ \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ to a basis of $\mathbb{R}^{3\times 1}$. Then we know that

$$P^{-1}AP = \begin{pmatrix} 1 & * & * \\ 0 & 3 & * \\ 0 & 0 & 3 \end{pmatrix}.$$

Method B

Observe that $A^{(2)} = 3E_2$. Hence E_2 is an eigenvector of A with eigenvalue 3. Then we follow the idea of the proof as before. In summary, set

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$P_1^{-1}AP_1 = \begin{pmatrix} 3 & * & * \\ 0 & 0 & -3 \\ 0 & 1 & 4 \end{pmatrix}.$$

Setting

$$B_1 = \begin{pmatrix} 0 & -3 \\ 1 & 4 \end{pmatrix},$$

we have $\chi_{B_1}(x) = (x-1)(x-3)$, hence B_1 is diagonalizable. Moreover,

$$V_{B_1}(1) = \left\langle \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\rangle, \quad V_{B_1}(3) = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle.$$

For

$$P_2 = \begin{pmatrix} 3 & 1 \\ -1 & -1 \end{pmatrix}$$

we have

$$P_2^{-1} B_1 P_2 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

Finally, setting

$$P = P_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & -1 & -1 \end{pmatrix},$$

we obtain

$$P^{-1} A P = \begin{pmatrix} 3 & * & * \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

5.2 Triangularizable Linear Maps

Definition 5.2.1. A linear map $f : V \to V$ is called **triangularizable** if there exists an ordered basis $\hat{v}$ of V such that the matrix $(f : \hat{v}, \hat{v})$ is upper triangular.

Observation 5.2.1. Let $f : V \to V$, let $\hat{a}$ be an ordered basis of V, and let $A = (f : \hat{a}, \hat{a})$. The following are equivalent:

i. f is triangularizable.

ii. A is triangularizable.

iii. $\chi_f(x)$ is a product of linear factors.

Proof.

- i. $\leftrightarrow$ ii. This follows immediately from Theorem 1.2
- ii. $\leftrightarrow$ iii. This follows from Theorem 5.1.1 and the fact that $\chi_f(x) = \chi_A(x)$.

□

Example 5.2.1. Let

$$f \colon \mathbb{R}^3 \to \mathbb{R}^3, \quad f(x, y, z) = (2x, x + y + 2z, ay + z).$$

Show that f is triangularizable if and only if $a \geq 0$.

Proof. We observe that

$$A = (f : \hat{e}, \hat{e}) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & a & 1 \end{pmatrix},$$

hence:

$$\chi_f(x) = \chi_A(x) = \det \begin{pmatrix} 2-x & 0 & 0 \\ 1 & 1-x & 2 \\ 0 & a & 1-x \end{pmatrix}.$$

Therefore:

$$\chi_f(x) = (2-x) \cdot \left[(1-x)^2 - 2a\right] = (2-x)(x^2 - 2x + 1 - 2a).$$

The map f is triangularizable if and only if $\chi_f(x)$ factors into linear factors over $\mathbb{R}$, i.e. if the discriminant of the quadratic factor is nonnegative:

$$\Delta = 4 - 4(1-2a) = 8a \geq 0 \Leftrightarrow a \geq 0.$$

□

Reminder 5.2.1. Let T be upper triangular, i.e. of the form

$$T = \begin{pmatrix} t_1 & & * \\ & \ddots & \\ 0 & & t_n \end{pmatrix}.$$

By induction one has

$$T^k = \begin{pmatrix} t_1^k & & * \\ & \ddots & \\ 0 & & t_n^k \end{pmatrix}, \quad \text{for every } k \geq 1.$$

More generally, for every $\varphi(x) \in \mathbb{F}[x]$,

$$\varphi(T) = \begin{pmatrix} \varphi(t_1) & & * \\ & \ddots & \\ 0 & & \varphi(t_n) \end{pmatrix},$$

i.e. it is also upper triangular.

Theorem 5.2.1 (Spectral Mapping)**.** Let $A \in \mathbb{F}^{n \times n}$ with

$$\chi_A(x) = (\lambda_1 - x) \cdots (\lambda_n - x).$$

Then for every $\varphi(x) \in \mathbb{F}[x]$ we have:

$$\chi_{\varphi(A)}(x) = (\varphi(\lambda_1) - x) \cdots (\varphi(\lambda_n) - x).$$

Proof. Let $\varphi(x) \in \mathbb{F}[x]$. Since $\chi_A(x) = (\lambda_1 - x)\cdots(\lambda_n - x)$, the matrix A is triangularizable, i.e. there exists $P \in \mathbb{F}^{n\times n}$ such that

$$P^{-1}AP = T = \begin{pmatrix} t_1 & & * \\ & \ddots & \\ 0 & & t_n \end{pmatrix}.$$

Also, $\varphi(T) = P^{-1}\varphi(A)P$. From this relation and Reminder 5.2.1 we get:

$$\chi_{\varphi(T)}(x) = \chi_{\varphi(A)}(x) = (\varphi(t_1) - x)\cdots(\varphi(t_n) - x),$$

where each t_i is an eigenvalue of A, for $i = 1, 2, \ldots, n$. □

5.3 Cayley-Hamilton Theorem

Motivation. Let $A \in \mathbb{F}^{n\times n}$. We know that $\dim \mathbb{F}^{n\times n} = n^2$ and the number of matrices $I_n, A, A^2, \ldots, A^{n^2}$ is $n^2 + 1$. Hence they are linearly dependent, i.e. there exist $a_0, a_1, \ldots, a_{n^2} \in \mathbb{F}$, not all zero, such that

$$a_{n^2}A^{n^2} + \cdots + a_0\mathbb{I}_n = 0.$$

Setting $\varphi(x) = a_{n^2}x^{n^2} + \cdots + a_0$, we obtain a nonzero polynomial with $\varphi(A) = 0$.

Observation 5.3.1. Let

$$A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix} \in \mathbb{F}^{n\times n}$$

with $A_i \in \mathbb{F}^{n_i\times n_i}$ and $n_1 + n_2 = n$. Then, by induction,

$$A^m = \begin{pmatrix} A_1^m & * \\ 0 & A_2^m \end{pmatrix}, \quad \text{for every } m \geq 1.$$

Consequently, for every $\varphi(x) \in \mathbb{F}[x]$,

$$\varphi(A) = \begin{pmatrix} \varphi(A_1) & * \\ 0 & \varphi(A_2) \end{pmatrix}.$$

Theorem 5.3.1 (Cayley-Hamilton for matrices)**.** Let $A \in \mathbb{F}^{n\times n}$ with

$$\chi_A(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_0.$$

Then:

$$\chi_A(A) = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_0 I_n = 0.$$

Proof. The proof is split into two steps:

Step A. If $A \in \mathbb{C}^{n\times n}$, then it is similar to an upper triangular matrix T. That is, there exists an invertible P such that:

$$A = P^{-1}TP.$$

Hence:

$$\chi_A(A) = \chi_T(A) = \chi_T(P^{-1}TP) = P^{-1}\chi_T(T)P.$$

It suffices to show that

$$\chi_T(T) = 0,$$

i.e. to prove the claim in the special case where the matrix is upper triangular.

Step B. We use induction on n.

• **Base case.** For $n = 1$ the theorem is obvious.

• **Inductive step.** Assume it holds for every upper triangular matrix of size $(n-1)\times(n-1)$.

Let

$$T = \begin{pmatrix} \lambda_1 & * \\ 0 & T_1 \end{pmatrix}$$

where T_1 is upper triangular of size $(n-1)\times(n-1)$ and

$$\chi_T(x) = (\lambda_1 - x)\cdot \chi_{T_1}(x).$$

Thus:

$$\chi_T(T) = (\lambda_1\mathbb{I}_n - T)\cdot \chi_{T_1}(T).$$

We have:

$$\chi_T(T) = \begin{pmatrix} 0 & * \\ 0 & \lambda_1\mathbb{I}_{n-1} - T_1 \end{pmatrix}\begin{pmatrix} \chi_{T_1}(\lambda_1) & * \\ 0 & \chi_{T_1}(T_1) \end{pmatrix}.$$

By the inductive hypothesis, $\chi_{T_1}(T_1) = 0$, hence:

$$\chi_T(T) = \begin{pmatrix} 0 & * \\ 0 & \lambda_1\mathbb{I}_{n-1} - T_1 \end{pmatrix}\begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} = 0.$$

□

Observation 5.3.2. If $\varphi \in \mathbb{F}[x]$, $A \in \mathbb{F}^{n\times n}$ with $\varphi(A) = 0$ and λ is an eigenvalue of A, then $\varphi(\lambda)$ is a root of the polynomial $\varphi(x)$. Consequently, if $A^k = 0$, then every eigenvalue λ of A (in $\mathbb{C}$) is 0.

Proposition 5.3.1. Let $A \in \mathbb{F}^{n\times n}$. The following statements are equivalent:

i. $A^n = 0$,

ii. $A^k = 0$ for some $k \geq 1$,

iii. Every eigenvalue of A in $\mathbb{C}$ is 0.

Proof.

- i. $\Rightarrow$ ii. Obvious, since $A^n = 0$.
- ii. $\Rightarrow$ iii. By Notation 5.3.2, every eigenvalue of A is a root of the annihilating polynomial, hence it must be 0.
- iii. $\Rightarrow$ i. If all eigenvalues of A are 0, then $\chi_A(x) = (-1)^n x^n$. By Theorem 5.3.1 we get

$$\chi_A(A) = (-1)^n A^n = 0,$$

thus $A^n = 0$.

□

Theorem 5.3.2 (Cayley-Hamilton for linear maps)**.** Let $f\colon V \to V$ be a linear map and let $\chi_f(x) = (-1)^n x^n + \cdots + a_0$ be its characteristic polynomial. Then:

$$\chi_f(f) = (-1)^n f^n + \cdots + a_0 \cdot 1_V = 0.$$

Proof. Let $\hat{v}$ be an ordered basis of V and $A = (f : \hat{v}, \hat{v})$. We know that for every $\varphi(x) \in \mathbb{F}[x]$:

$$(\varphi(f) : \hat{v}, \hat{v}) = \varphi(A).$$

For $\varphi(x) = \chi_f(x) = \chi_A(x)$, by Theorem 5.3.1 we have $\chi_A(A) = 0$, hence:

$$(\chi_f(f) : \hat{v}, \hat{v}) = 0 \Rightarrow \chi_f(f) = 0.$$

□

5.4 Exercises of Chapter 5.

Group A: 1,2,3,4,5,6,7,11,14,23,28,34
Group B: 8,9,12,13,15,16,17,18,19,20,21,22,24,25,31,33,35
Group C: 10,26,27,29,30

Exercise 5.1. Prove that if $A \in \mathbb{R}^{2\times 2}$ has at least one real eigenvalue, then A is triangularizable.

Exercise 5.2.

a. Let $A = \begin{pmatrix} -2 & 1 \\ -4 & 2 \end{pmatrix} \in \mathbb{C}^{2\times 2}$. After showing that A is triangularizable, find an invertible $U \in \mathbb{C}^{2\times 2}$ such that $U^{-1}AU$ is triangular.

b. Let $A = \begin{pmatrix} 3 & 4 & 5 \\ 0 & -2 & 1 \\ 0 & -4 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. After showing that A is triangularizable, find an invertible $U \in \mathbb{R}^{3\times 3}$ such that $U^{-1}AU$ is triangular.

c. Let $A = \begin{pmatrix} 2 & 2 & 0 \\ -1 & -2 & 1 \\ 0 & 5 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$. After showing that A is triangularizable, find an invertible $U \in \mathbb{R}^{3\times 3}$ such that $U^{-1}AU$ is triangular.

Exercise 5.3. Find the values of a for which the matrix

$$\begin{pmatrix} 4 & a \\ 3 & 3 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

is triangularizable but not diagonalizable.

Exercise 5.4. Let

$$A = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

a. Find the characteristic polynomial and the dimensions of the eigenspaces of A.

b. Is A diagonalizable?

c. Is A triangularizable? If yes, find an invertible U such that $U^{-1}AU$ is triangular.

Exercise 5.5. Let $\{v_1, v_2, v_3\}$ be a basis of $\mathbb{R}^3$, let $a \in \mathbb{R}$, and let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map such that $f(v_1) = 2v_1,\ f(v_2) = v_1 + v_2 + 2v_3,\ f(v_3) = av_2 + v_3$. Show that f is triangularizable if and only if $a \geq 0$.

Exercise 5.6. Show that there exist infinitely many matrices $A \in \mathbb{R}^{2\times 2}$ such that $A^2 - 5A + 6I_2 = 0$.

Exercise 5.7. Let $A \in \mathbb{R}^{3\times 3}$ with $\chi_A(x) = -x^3 + x$. Show that for every positive integer k:

a. A^k is diagonalizable, and

b. $A^{2k} = A^2$ and $A^{2k+1} = A$.

Exercise 5.8. a. Let $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \cdots, \lambda_n$. Then for every $k \geq 1$, $\mathrm{Tr}(A^k) = \lambda_1^k + \cdots + \lambda_n^k$.

b. Let $A \in \mathbb{R}^{n\times n}$ be a triangularizable matrix such that $\mathrm{Tr}(A^2) = 0$. Show that $A^n = 0$.

c. Let $A \in \mathbb{C}^{n\times n}$ such that $\mathrm{Tr}(A) = \mathrm{Tr}(A^2) = \cdots = \mathrm{Tr}(A^{n-1}) = 0$. Show that if $\mathrm{Tr}(A^n) \neq 0$, then A is

- diagonalizable and
- invertible.

Exercise 5.9. Let $A \in \mathbb{F}^{n\times n}$. Show that the following are equivalent:

a. Every eigenvalue of A in $\mathbb{C}$ equals 0.

b. $A^k = 0$ for some positive integer k.

c. $A^n = 0$.

d. $\mathrm{Tr}(A) = \mathrm{Tr}(A^2) = \cdots = \mathrm{Tr}(A^n) = 0$.

Exercise 5.10. Let $A, B \in \mathbb{F}^{n\times n}$ such that $AB - BA = A$. Prove that $A^n = 0$.

Exercise 5.11. Let $A \in \mathbb{C}^{n\times n}$ be invertible. Show that if $\chi_A = (\lambda_1 - x)\cdots(\lambda_n - x)$, $\lambda_i \in \mathbb{C}$, then

$$\chi_{A^{-1}}(x) = \left(\frac{1}{\lambda_1} - x\right)\cdots\left(\frac{1}{\lambda_n} - x\right).$$

Exercise 5.12. Let $\dim V = n$ and let $f : V \to V$ be a linear map.

a. Show that f is triangularizable if and only if for each $i = 1, \cdots, n$ there exists a subspace $W_i \leq V$ with $\dim W_i = i$, $W_1 \subseteq W_2 \subseteq \cdots \subseteq W_n$, and $f(W_i) \subseteq W_i$.

b. Is it true that f is triangularizable if for each $i = 1, \cdots, n$ there exists a subspace $W_i \leq V$ with $\dim W_i = i$ and $f(W_i) \subseteq W_i$?

Exercise 5.13. Let $A \in \mathbb{F}^{n \times n}$.

a. Show that if A is not invertible, then there exists $f(x) \in \mathbb{F}[x]$ of degree $n-1$ such that $Af(A) = 0$.

b. Show that if A is invertible, then there exists $f(x) \in \mathbb{F}[x]$ of degree $n-1$ such that $A^{-1} = f(A)$.

Exercise 5.14. Let

$$A = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -2 & -1 \\ 0 & 3 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

a. Express A^{-1} as a linear combination of I_3, A, A^2.

b. Prove that $A^{2n} - 2A^{2n-1} = A^2 - 2A$ for every positive integer n.

c. Find a polynomial $\varphi(x) \in \mathbb{R}[x]$ of degree at most 2 such that $A^5 - 2A^4 + 2a + 3I_3 = \phi(A)$.

Exercise 5.15. Let $A \in \mathbb{R}^{n \times n}$ such that $\chi_A(x) = (-1)^n(x^n - x^m - x^{n-m} + 1)$, where $0 < m < n$. Show that there exists a positive integer ν such that A^n is triangularizable.

Exercise 5.16. Let $A \in \mathbb{C}^{n \times n}$ be a non-diagonalizable matrix. Then A is similar to a matrix of the form $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Exercise 5.17. Let $A, B \in \mathbb{F}^{n \times n}$ satisfy $AB = BA = 0$. Show that $\chi_A(A + B) = \chi_A(B) - \det(A) \cdot I_n$.

Exercise 5.18. If $A = (a_{ij}) \in \mathbb{F}^{n\times n}$, define $h(A) = \sum\limits_{i,j} a_{ij}a_{ji}$.

a. Show that if A, B are similar, then $h(A) = h(B)$.

b. Let $\mathbb{F} = \mathbb{C}$. Show that $h(A) = \lambda_1^2 + \cdots + \lambda_n^2$, where $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$ are the eigenvalues of A.

Exercise 5.19. Show that every upper triangular matrix $A \in \mathbb{F}^{n\times n}$ is similar to a lower triangular matrix. Then show that every matrix $B \in \mathbb{C}^{n\times n}$ is similar to a lower triangular matrix.

Exercise 5.20. Let $A \in \mathbb{R}^{n\times n}$ such that $A^n = I_n$. Show that $-n \leq \text{Tr}A \leq n$.

Exercise 5.21. Let V be a $\mathbb{C}$-vector space and let $f, g : V \to V$ be two linear maps such that $f \circ g = g \circ f$. Prove the following.

a. If λ is an eigenvalue of f, then $g(V_f(\lambda)) \subseteq V_f(\lambda)$.

b. The maps f, g have a common eigenvector.

c. There exists an ordered basis of V such that the corresponding matrices of f, g are upper triangular.

d. For every eigenvalue λ of $f - g$ there exist an eigenvalue λ_f of f and an eigenvalue λ_g of g such that $\lambda = \lambda_f - \lambda_g$.

Exercise 5.22. Let $A, B \in \mathbb{C}^{n\times n}$. Consider the linear maps

$$L_A : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n},\ L_A(X) = AX,$$

$$R_B : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n},\ R_B(X) = XB.$$

a. Show that $L_A \circ R_B = R_B \circ L_A$.

b. Show that L_A has the same eigenvalues as the matrix A and that R_B has the same eigenvalues as the matrix B.

c. Assume that A, B have no common eigenvalue. Show that for every $C \in \mathbb{C}^{n\times n}$ there exists a unique $D \in \mathbb{C}^{n\times n}$ such that $AD - DB = C$.

Exercise 5.23. Let $A \in \mathbb{F}^{n\times n}$ and let W_A be the subspace of $\mathbb{F}^{n\times n}$ generated by $I_n, A, A^2, \cdots$. Show that for every $k \geq 0$, $A^{n+k} \in \langle I_n, A, A^2, \cdots, A^{n-1}\rangle$ and hence $\dim W_A \leq n$.

Exercise 5.24. Determine which of the following statements are true. Justify your answers.

a. Let $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = (x^2+1)(x+1)^2$. Then the matrix A^n is triangularizable if and only if n is even.

b. For every $A \in \mathbb{F}^{n\times n}$ there exists a polynomial $\varphi(x) \in \mathbb{F}[x]$ of positive degree such that $\varphi(A) = I_n$.

Exercise 5.25. Let $A \in \mathbb{F}^{n\times n}$ with $\mathrm{rank} A = 1$. Prove the following statements.

a. $A^2 = \mathrm{Tr}(A)\cdot A$.

b. $A^n = 0 \Leftrightarrow \mathrm{Tr}(A) = 0$.

c. A is triangularizable.

d. $\mathrm{Tr}(A) \neq 0 \Leftrightarrow A$ is diagonalizable. (see Exercise 3.26).

Exercise 5.26. Let $A, B, C, D \in \mathbb{F}^{n\times n}$ such that $A^iC = B^iD$ for every $i \geq 1$. Prove that if A, B are invertible, then $C = D$.

Exercise 5.27. Let $A \in \mathbb{C}^{n\times n}$ and let $f_A : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ be the linear map defined by $f_A(B) = AB - BA$. Show that if every eigenvalue of A equals 0, then every eigenvalue of f_A equals 0.

Exercise 5.28. Let V be a real vector space of dimension 3, let $\hat{a} = \{v_1, v_2, v_3\}$ be an ordered basis of V, and let $c \in \mathbb{R}$. Consider the linear map $f : V \to V$ defined by $f(v_1) = 2v_2,\ f(v_2) = -v_1 + 3v_2,\ f(v_3) = cv_1 + v_2 + v_3$.

a. Find all values of c for which f is triangularizable.

b. Find all values of c for which f is diagonalizable.

c. For $c = 0$ find a basis for each eigenspace of f and a basis of V generated by eigenvectors of f.

Exercise 5.29. If $A \in \mathbb{R}^{n\times n}$ is triangularizable and $\operatorname{Tr}(A^2) = Tr(A^3) = \operatorname{Tr}(A^4) = c$, then $c \in \mathbb{Z}_{\geq 0}$ and $\operatorname{Tr}(A^k) = c$ for every positive integer k.

Exercise 5.30. Let $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ have no common eigenvalue. Show that there is no nonzero $X \in \mathbb{F}^{n\times m}$ with $AX = XB$.

Exercise 5.31. Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}$. Show that for every $m \geq 3$ there is no $B \in \mathbb{C}^{3\times 3}$ with $B^m = A$.

Exercise 5.32. Let $A, B \in \mathbb{C}^{n\times n}$ with $(AB)^n = 0,\ n \geq 1$. Then $(BA)^n = 0$.

Exercise 5.33. If $A \in \mathbb{C}^{n\times n}$ has at most one nonzero eigenvalue, then $\det(I_n + A) = 1 + Tr(A)$.

Exercise 5.34. Let $A, B \in \mathbb{C}^{n\times n}$. Show that the matrix $\chi_B(A)$ is invertible if and only if A, B have no common eigenvalue.

Exercise 5.35. Determine which of the following statements are true. In each case give a proof or a counterexample.

a. Let A be an invertible matrix. Then A is triangularizable if and only if A^{-1} is triangularizable.

b. If $A \in \mathbb{F}^{n\times n}$ is triangularizable, then $\varphi(A)$ is triangularizable for every $\varphi(x) \in \mathbb{F}[x]$.

c. Let $A \in \mathbb{R}^{n\times n}$. If A^2 is triangularizable, then A is triangularizable.

d. If $A \in \mathbb{R}^{3\times 3}$, then there exists an invertible $U \in \mathbb{R}^{3\times 3}$ with $U^{-1}AU$ upper triangular.

e. If $A \in \mathbb{R}^{3\times 3}$, then there exists an invertible $U \in \mathbb{R}^{3\times 3}$ with

$$U^{-1}AU = \begin{pmatrix} \lambda & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

f. If $A \in \mathbb{R}^{3\times 3}$ is of the form

$$A = \begin{pmatrix} * & 0 & * \\ * & -5 & * \\ * & 0 & * \end{pmatrix}$$

then there exists an invertible $U \in \mathbb{R}^{3\times 3}$ with

$$U^{-1}AU = \begin{pmatrix} -5 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

g. Let $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = (x-1)^2(x-2)(x-3)$. Then A is triangularizable and not diagonalizable if and only if $\dim V(1) = 1$.

h. Let $f : V \to V$ be a triangularizable linear map and let $U \leq V$ be a subspace such that $f(U) \subseteq U$. Then the restriction of f to U is a triangularizable map.

CHAPTER 6

MINIMAL POLYNOMIAL

6.1 Minimal Polynomial

Motivation. If $A \in \mathbb{F}^{n\times n}$, then there exists a polynomial $\varphi(x) \in \mathbb{F}[x]$, with $\varphi(x) \neq 0$, such that $\varphi(A) = 0$. For example, by the Cayley–Hamilton Theorem we have $\chi_A(A) = 0$. Our goal is to find the **monic** polynomial of *smallest degree* that annihilates A.

Definition 6.1.1. Let $A \in \mathbb{F}^{n\times n}$. The **minimal polynomial** of A, denoted by $m_A(x)$, satisfies:

i. $m_A(x)$ is monic,

ii. $m_A(A) = 0$,

iii. $m_A(x)$ has minimal degree among polynomials satisfying (i) and (ii).

Observation 6.1.1. For every $A \in \mathbb{F}^{n\times n}$ there exists a polynomial satisfying the properties of Definition 6.1.1, and moreover it is **unique**.

Proof.

- **Existence.** Consider the set

$$S = \{\varphi(x) \in \mathbb{F}[x] \mid \varphi(x) \neq 0 \text{ and } \varphi(A) = 0\}.$$

The set S is nonempty, since by Theorem 5.3.1 we have $\chi_A(A) = 0$. Choose a polynomial $\varphi(x) \in S$ of minimal degree. If r is the leading coefficient of $\varphi(x)$, then $r^{-1}\varphi(x)$ also lies in S, is monic, and has minimal degree.

- **Uniqueness.** Let $m_A(x)$ and $m'_A(x)$ be two monic polynomials that annihilate A and have minimal degree. If $m_A(x) \neq m'_A(x)$, then the difference

$$m_A(x) - m'_A(x)$$

is a nonzero polynomial of degree strictly smaller than $\deg(m_A)$ and it also annihilates A. Then

$$r^{-1}(m_A - m'_A)$$

(where r is the leading coefficient) is monic, annihilates A, and has smaller degree—a contradiction. Hence

$$m_A(x) = m'_A(x).$$

□

Properties 6.1.1. Let $A \in \mathbb{F}^{n\times n}$.

i. If $\varphi(A) = 0$ for some $\varphi(x) \in \mathbb{F}[x]$, then $m_A(x) \mid \varphi(x)$. In particular, $m_A(x) \mid \chi_A(x)$.

ii. Every eigenvalue of A is a root of $m_A(x)$. Every root of $m_A(x)$ is an eigenvalue of A. That is, $m_A(x)$ and $\chi_A(x)$ have the same roots (ignoring multiplicities).[1]

Proof. i. By Euclidean division, there exist polynomials $q(x), r(x) \in \mathbb{F}[x]$ such that

$$\varphi(x) = q(x)\, m_A(x) + r(x), \quad \text{where either } \deg r(x) < \deg m_A(x) \text{ or } r(x) = 0.$$

Then

$$\varphi(A) = q(A)\, m_A(A) + r(A) = r(A).$$

Since $\varphi(A) = 0$, we get $r(A) = 0$. If $r(x) \neq 0$, then $c^{-1}r(x)$, where c is the leading coefficient of $r(x)$, is monic, nonzero, satisfies $r(A) = 0$, and has degree smaller than $\deg(m_A)$—a contradiction. Hence $r(x) = 0$, i.e. $m_A(x) \mid \varphi(x)$.

In particular, for $\varphi(x) = \chi_A(x)$, by Theorem 5.3.1 we have $\chi_A(A) = 0$, hence $m_A(x) \mid \chi_A(x)$.

ii. Let $\lambda \in \mathbb{F}$ and $X \neq 0$ such that $AX = \lambda X$. We know that for every $\varphi(x) \in \mathbb{F}[x]$,

$$\varphi(A)X = \varphi(\lambda)X.$$

In particular, for $\varphi(x) = m_A(x)$ we obtain

$$m_A(A)X = m_A(\lambda)X = 0.$$

[1]For example, it could happen that $\chi_A(x) = -(x-1)^2(x-2)$ and $m_A(x) = (x-1)(x-2)$. It cannot happen that $m_A(x) = (x-1)^2$ or $m_A(x) = (x-1)^2(x-2)(x-3)$.

Since $X \neq 0$, it follows that $m_A(\lambda) = 0$. Thus every eigenvalue of A is a root of $m_A(x)$. The converse follows immediately from (i), since $m_A(x) \mid \chi_A(x)$.

□

Example 6.1.1. Consider the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}.$$

We observe that

$$\chi_A(x) = (x-1)(x-2) = \chi_B(x).$$

Then

$$m_A(x) = (x-1)(x-2)$$

because

$$m_A(x) \mid \chi_A(x) = (x-1)(x-2)$$

and it has the same roots. Similarly,

$$m_B(x) = (x-1)(x-2).$$

Example 6.1.2. Let the matrices be

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

We observe that $\chi_A(x) = -(x-2)^3$, hence

$$m_A(x) \in \{x-2,\ (x-2)^2,\ (x-2)^3\}.$$

We have:

$$A - 2\mathbb{I}_3 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0, \quad (A - 2\mathbb{I}_3)^2 = 0 \Rightarrow m_A(x) = (x-2)^2.$$

For the matrix B:

$$B - 2\mathbb{I}_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \neq 0, \quad (B - 2\mathbb{I}_3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0.$$

Therefore

$$m_B(x) = (x-2)^3.$$

Pay special attention to this example!

Example 6.1.3. Let

$$A = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

a. Find $\varphi(x) \in \mathbb{R}[x]$ with $\deg \varphi(x) \leq 1$ such that $A^{-1} = \varphi(A)$.

b. Find $\psi(x) \in \mathbb{R}[x]$ with $\deg \psi(x) \leq 1$ such that $A^4 + A - 2\mathbb{I}_3 = \psi(A)$.

Proof. First we find $m_A(x)$. We have $\chi_A(x) = -(x-2)^2(x-3)$, hence:

$$m_A(x) = (x-2)^2(x-3) \quad \text{or} \quad m_A(x) = (x-2)(x-3).$$

We compute:

$$(A - 2\mathbb{I}_3)(A - 3\mathbb{I}_3) = 0 \Rightarrow m_A(x) = (x-2)(x-3) = x^2 - 5x + 6.$$

a. Since 0 is not an eigenvalue of A, it follows that A is invertible. From $m_A(A) = 0$, i.e. $A^2 - 5A + 6\mathbb{I}_3 = 0$, solving gives

$$A^{-1} = -\frac{1}{6}(A - 5\mathbb{I}_3) \Rightarrow \varphi(x) = -\frac{1}{6}(x-5).$$

b. By Euclidean division of $x^4 + x - 2$ by $m_A(x) = x^2 - 5x + 6$, we have

$$x^4 + x - 2 = (x^2 + 5x + 19)(x^2 - 5x + 6) + 66x - 166.$$

Hence

$$A^4 + A - 2\mathbb{I}_3 = 66A - 166\mathbb{I}_3 \Rightarrow \psi(x) = 66x - 166,$$

since $m_A(A) = 0$. □

Proposition 6.1.1. Similar matrices have the same minimal polynomial.

Proof. Let $A, B \in \mathbb{F}^{n\times n}$ with $B = P^{-1}AP$, where $P \in \mathbb{F}^{n\times n}$ is invertible. For every $\varphi(x) \in \mathbb{F}[x]$ we have

$$\varphi(B) = P^{-1}\varphi(A)P.$$

Hence $\varphi(A) = 0$ if and only if $\varphi(B) = 0$. Therefore $m_A(x) = m_B(x)$. □

Proposition 6.1.2. If $A \in \mathbb{F}^{n\times n}$ with

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}, \quad \text{where } B \in \mathbb{F}^{n_1\times n_1},\ C \in \mathbb{F}^{n_2\times n_2},\ n = n_1 + n_2,$$

then

$$m_A(x) = \text{lcm}(m_B(x), m_C(x)).$$

Proof. Since

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix},$$

we know that for every $\varphi(x) \in \mathbb{F}[x]$,

$$\varphi(A) = \begin{pmatrix} \varphi(B) & 0 \\ 0 & \varphi(C) \end{pmatrix}.$$

Setting $\varphi(x) = m_A(x)$ we get

$$m_A(A) = \begin{pmatrix} m_A(B) & 0 \\ 0 & m_A(C) \end{pmatrix} = 0,$$

hence $m_A(B) = 0$ and $m_A(C) = 0$. Therefore

$$m_B(x) \mid m_A(x) \quad \text{and} \quad m_C(x) \mid m_A(x) \Rightarrow \text{lcm}(m_B(x), m_C(x)) \mid m_A(x).$$

Define $\psi(x) = \text{lcm}(m_B(x), m_C(x))$. Then

$$\psi(A) = \begin{pmatrix} \psi(B) & 0 \\ 0 & \psi(C) \end{pmatrix} = 0,$$

so $m_A(x) \mid \psi(x)$. Since both are monic, we conclude

$$m_A(x) = \psi(x).$$

□

Observation 6.1.2. If

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_k \end{pmatrix} \in \mathbb{F}^{n\times n}, \quad \text{where } A_i \in \mathbb{F}^{n_i\times n_i},\ n_1 + \cdots + n_k = n,$$

then

$$m_A(x) = \text{lcm}(m_{A_1}(x), \ldots, m_{A_k}(x)).$$

Proof. This follows immediately from Proposition 6.1.2 and induction on k. □

Corollary 6.1.1. Let $A = \mathrm{diag}(\lambda_1, \cdots, \lambda_1, \cdots, \lambda_k, \cdots, \lambda_k)$ with $\lambda_i \neq \lambda_j$ for all $i \neq j$. Then

$$m_A(x) = (x - \lambda_1) \cdots (x - \lambda_k).$$

Proof. This is immediate by applying Notation 6.1.2 to $A_i = \lambda_i I_{n_i}$. □

Observation 6.1.3. Let $A \in \mathbb{F}^{n \times n}$ be diagonalizable, i.e. similar to a diagonal matrix. By Proposition 6.1.1 and Corollary 6.1.1, the polynomial $m_A(x)$ is a product of distinct linear factors.

Example 6.1.4. Consider the matrix

$$A = \left(\begin{array}{cc|cc} 4 & 1 & \multicolumn{2}{c}{\multirow{2}{*}{O}} \\ 2 & 3 & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{O}} & 2 & 0 \\ & & 3 & 0 \end{array}\right) \in \mathbb{R}^{4 \times 4}.$$

Equivalently, write $A = \left(\begin{array}{c|c} B & O \\ O & C \end{array}\right)$ with

$$B = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 0 \\ 3 & 0 \end{pmatrix}.$$

By computation we find

$$m_B(x) = (x-2)(x-5), \quad m_C(x) = x(x-2).$$

By Proposition 6.1.2,

$$m_A(x) = \mathrm{lcm}(m_B(x), m_C(x)) = x(x-2)(x-5).$$

Example 6.1.5. Let $A = \left(\begin{array}{c|c} B & * \\ O & C \end{array}\right)$. Then in general it is not true (why?) that $m_A(x) = \mathrm{lcm}(m_B(x), m_C(x))$. What is always true is that $\mathrm{lcm}(m_B(x), m_C(x)) \mid m_A(x)$.

6.2 Diagonalizability Criterion

Motivation. Let A be diagonalizable. Then A is similar to a diagonal matrix

$$D = \text{diag}(\lambda_1, \cdots, \lambda_1, \cdots, \lambda_k, \cdots, \lambda_k)$$

with $\lambda_i \neq \lambda_j$ for $i \neq j$. By Proposition 6.1.1 and Corollary 6.1.1 we have

$$m_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k),$$

i.e. a product of distinct monic linear factors in $\mathbb{F}[x]$. We will show that the converse also holds.

Theorem 6.2.1 (Diagonalizability criterion via $m_A(x)$)**.** Let $A \in \mathbb{F}^{n \times n}$. Then A is diagonalizable if and only if $m_A(x)$ is a product of distinct monic linear factors in $\mathbb{F}[x]$.

Proof. If A is diagonalizable, the claim follows immediately from Corollary 6.1.1.

Conversely, assume

$$m_A(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

with $\lambda_i \neq \lambda_j$ for $i \neq j$. We know the distinct eigenvalues of A are precisely the λ_i, $i = 1, \ldots, k$. We will show that

$$\mathbb{F}^{n \times 1} = V_A(\lambda_1) + V_A(\lambda_2) + \cdots + V_A(\lambda_k).$$

Define the following polynomials:[2]

$$a_i(x) = \prod_{j \neq i} (x - \lambda_j).$$

Then

$$\gcd(a_1(x), \ldots, a_k(x)) = 1.$$

Hence there exist $b_i(x) \in \mathbb{F}[x]$ such that

$$1 = \sum_{i=1}^{k} a_i(x) b_i(x).$$

Therefore

$$\mathbb{I}_n = \sum_{i=1}^{k} a_i(A) b_i(A),$$

[2] For example, for $k = 3$: $a_1(x) = (x - \lambda_2)(x - \lambda_3)$, $a_2(x) = (x - \lambda_1)(x - \lambda_3)$, $a_3(x) = (x - \lambda_1)(x - \lambda_2)$.

i.e. for every $X \in \mathbb{F}^{n\times 1}$ we have

$$X = \sum_{i=1}^{k} a_i(A)b_i(A)X.$$

Claim: For each $i = 1, \dots, k$, we have $a_i(A)b_i(A)X \in V_A(\lambda_i)$.

Indeed, note that

$$(A - \lambda_i \mathbb{I}_n)a_i(A)b_i(A)X = m_A(A)b_i(A)X = 0,$$

since $(x - \lambda_i)a_i(x) = m_A(x)$. Hence each term $a_i(A)b_i(A)X$ lies in the corresponding eigenspace $V_A(\lambda_i)$.

Thus

$$\mathbb{F}^{n\times 1} \subseteq \sum_{i=1}^{k} V_A(\lambda_i) \Rightarrow \mathbb{F}^{n\times 1} = \sum_{i=1}^{k} V_A(\lambda_i),$$

and by Theorem 4.2.1, the matrix A is diagonalizable. □

6.3 Minimal Polynomial of a Linear Map

Definition 6.3.1. Let $f\colon V \to V$ be a linear map and let $\hat{a}$ be an ordered basis of V. Define $m_f(x) = m_A(x)$, where $A = (f : \hat{a}, \hat{a})$. The polynomial $m_f(x)$ is called the **minimal polynomial** of f.

Observation 6.3.1. Since similar matrices have the same minimal polynomial, the definition of $m_f(x)$ does not depend on the choice of $\hat{a}$.

Properties 6.3.1 (Minimal polynomial of a linear map)**.** Let $f\colon V \to V$ be a linear map. The following hold:

i. $m_f(x)$ is monic, $m_f(f) = 0$, and among polynomials with these properties it has minimal degree.

ii. If $\varphi(f) = 0$ with $\varphi(x) \in \mathbb{F}[x]$, then $m_f(x) \mid \varphi(x)$. In particular, $m_f(x) \mid \chi_f(x)$.

iii. Every eigenvalue of f is a root of $m_f(x)$. The polynomials $m_f(x)$ and $\chi_f(x)$ have the same roots.

iv. The map f is diagonalizable if and only if $m_f(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$ with $\lambda_i \neq \lambda_j$ for $i \neq j$.

Proof. These properties follow immediately from the corresponding results for matrices. As an illustration we prove property (i).

Let $A = (f : \hat{a}, \hat{a})$. We know that for every $\varphi(x) \in \mathbb{F}[x]$,

$$\varphi(A) = (\varphi(f) : \hat{a}, \hat{a})$$

by Proposition 2.4.1. Hence

$$(m_f(f) : \hat{a}, \hat{a}) = m_f(A) = m_A(A) = 0,$$

so $m_f(f) = 0$. Since $m_A(x)$ is monic and has minimal degree, the same is true for $m_f(x)$ by definition. $\square$

6.4 Simultaneous Diagonalization

Question 6.4.1. Let $A, B \in \mathbb{F}^{n\times n}$ be diagonalizable. Then there exist invertible $P_A, P_B \in \mathbb{F}^{n\times n}$ such that $P_A^{-1}AP_A$ and $P_B^{-1}BP_B$ are diagonal. When does there exist a common invertible P such that $P^{-1}AP$ and $P^{-1}BP$ are diagonal?

Observation 6.4.1. Assume such a P exists, i.e. P is invertible with $P^{-1}AP = \Delta_A$ diagonal and $P^{-1}BP = \Delta_B$ diagonal. Then $A = P\Delta_A P^{-1}$ and $B = P\Delta_B P^{-1}$, and we observe that

$$AB = P\Delta_A\Delta_B P^{-1} \quad \text{and} \quad BA = P\Delta_B\Delta_A P^{-1}.$$

Since Δ_A, Δ_B are diagonal, we have $\Delta_A\Delta_B = \Delta_B\Delta_A$, i.e. $AB = BA$. We will show that the converse also holds. We study the corresponding problem for linear maps.

6.4.1 Invariant Subspaces

Definition 6.4.1. Let $f\colon V \to V$ be a linear map. A subspace $U \leq V$ is called f**-invariant** if $f(U) \subseteq U$, i.e. for every $u \in U$ we have $f(u) \in U$.

Example 6.4.1. Let $f\colon V \to V$ be a linear map.

i. The subspaces $\{0\}$ and V are f-invariant.

ii. The subspaces $\ker f$ and $\operatorname{Im} f$ are f-invariant. Indeed:
   - If $u \in \ker f$, then $f(u) = 0 \in \ker f$.
   - If $v \in \operatorname{Im} f$, then there exists $u \in V$ such that $f(u) = v$, hence $f(v) = f(f(u)) \in \operatorname{Im} f$.

iii. Every eigenspace $V_f(\lambda)$ is f-invariant. Indeed, if $u \in V_f(\lambda)$ then $f(u) = \lambda u \in V_f(\lambda)$.

iv. If $U_1, U_2 \leq V$ are f-invariant, then $U_1 + U_2$ is also f-invariant.

v. The sum of f-invariant eigenspaces is f-invariant.

vi. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ with $f(x, y) = (x, x + y)$.
   - $f(e_2) = e_2$ and $f(e_1) = e_1 + e_2$, where e_1, e_2 is the standard basis of $\mathbb{R}^2$.
   - Hence $\langle e_2 \rangle$ is f-invariant, while $\langle e_1 \rangle$ is not (show why).

vii. If $U \leq V$ with $\dim U = 1$, then U is f-invariant if and only if $U = \langle u \rangle$ for some eigenvector u of f (show why).

Observation 6.4.2. Let U be f-invariant. Then $f(U) \subseteq U$, so the restriction of f to U, denoted by f_U, is the linear map

$$f_U\colon U \to U, \quad f_U(u) = f(u), \quad \text{for every } u \in U.$$

Proposition 6.4.1. Let $f\colon V \to V$ be a linear map and let U be an f-invariant subspace of V. If f is diagonalizable, then the restriction $f_U\colon U \to U$ is also diagonalizable.

Proof. By Property 6.3.1(iv), if f is diagonalizable then

$$m_f(x) = (x - \lambda_1) \cdots (x - \lambda_k), \quad \text{with } \lambda_i \neq \lambda_j \text{ for } i \neq j.$$

We claim that $m_{f_U}(x) \mid m_f(x)$. If so, then $m_{f_U}(x)$ is also a product of distinct linear factors, hence f_U is diagonalizable by the same property.

Indeed, for every $u \in U$ we have

$$m_f(f_U)(u) = m_f(f)(u) = 0.$$

Hence $m_{f_U}(x) \mid m_f(x)$. □

6.4.2 Simultaneous Diagonalization

Theorem 6.4.1. Let $f, g\colon V \to V$ be linear maps. The following are equivalent:

i. There exists a basis of V whose elements are eigenvectors of both f and g.

ii. The maps f and g are diagonalizable and, moreover, $f \circ g = g \circ f$.

Proof.
- i. $\Rightarrow$ ii. Assume there exists a basis $\{v_1, \dots, v_n\}$ of V with $f(v_i) = \lambda_i v_i$ and $g(v_i) = \mu_i v_i$ for each i, where $\lambda_i, \mu_i \in \mathbb{F}$.

  We claim that $f \circ g = g \circ f$. Indeed,

$$\begin{aligned}(f \circ g - g \circ f)(v_i) &= f(g(v_i)) - g(f(v_i)) = f(\mu_i v_i) - g(\lambda_i v_i) \\ &= \mu_i f(v_i) - \lambda_i g(v_i) \\ &= \mu_i \lambda_i v_i - \lambda_i \mu_i v_i \\ &= 0.\end{aligned}$$

  Thus $f \circ g = g \circ f$. Finally, by Theorem 4.3.1(ii), it follows that f and g are diagonalizable.

- ii. $\Rightarrow$ i. Since f is diagonalizable, we have

$$V = V_f(\lambda_1) + \cdots + V_f(\lambda_k),$$

  where λ_j are the distinct eigenvalues of f.

  We claim that if $f \circ g = g \circ f$, then each $V_f(\lambda_i)$ is g-invariant. Indeed, let $v \in V_f(\lambda_i)$, i.e. $f(v) = \lambda_i v$. Then

$$g(f(v)) = g(\lambda_i v) = \lambda_i g(v) \Rightarrow f(g(v)) = \lambda_i g(v) \Rightarrow g(v) \in V_f(\lambda_i).$$

  Hence $V_f(\lambda_i)$ is g-invariant.

  Let g_i be the restriction of g to $V_f(\lambda_i)$. Since g is diagonalizable and $V_f(\lambda_i)$ is g-invariant, Proposition 6.4.1 implies that g_i is diagonalizable. Therefore there exists a basis B_i of $V_f(\lambda_i)$ consisting of eigenvectors of g. Since $B_i \subseteq V_f(\lambda_i)$, each element of B_i is also an eigenvector of f.

  Letting $B = \bigcup_{i=1}^k B_i$, we obtain the desired basis of V, whose elements are eigenvectors of both f and g. [3]

□

[3] The point is that $\dim \sum_{i=1}^k V_f(\lambda_i) = \sum_{i=1}^k \dim V_f(\lambda_i)$, hence $\dim V = \sum_i |B_i|$.

Let us see why the corresponding statement holds for matrices. Consider the linear maps

$$L_A, L_B \colon \mathbb{F}^{n\times 1} \to \mathbb{F}^{n\times 1}, \quad L_A(X) = AX, \quad L_B(X) = BX.$$

If Theorem 6.4.1(i) holds, then there exists a basis B of $\mathbb{F}^{n\times 1}$ whose elements are eigenvectors of L_A and L_B.

Hence L_A and L_B are diagonalizable and $L_A \circ L_B = L_B \circ L_A$. This is equivalent to saying that the matrices A and B are diagonalizable and $AB = BA$, since

$$\begin{aligned} BA &= (L_B : \hat{E}, \hat{E}) \cdot (L_A : \hat{E}, \hat{E}) \\ &= (L_B \circ L_A : \hat{E}, \hat{E}) \\ &= (L_A \circ L_B : \hat{E}, \hat{E}) \\ &= (L_A : \hat{E}, \hat{E}) \cdot (L_B : \hat{E}, \hat{E}) = AB. \end{aligned}$$

6.5 Exercises of Chapter 6.

Group A: 1, 2, 3, 4, 5, 7, 10, 12, 15, 16, 17, 18, 19, 21, 30, 32, 36
Group B: 6, 8, 9, 11, 13, 14, 20, 22, 23, 24, 26, 28, 29, 31, 33, 34, 35, 37
Group C: 25, 27, 38, 39, 40

Exercise 6.1. Let $A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}$.

a. Find the minimal polynomial of A.

b. Determine whether A is diagonalizable.

c. Show that A is invertible and find $\varphi(x) \in \mathbb{R}[x]$ of degree at most 1 such that $A^{-1} = \varphi(A)$.

d. Find $\psi(x) \in \mathbb{R}[x]$ of degree at most 1 such that $A^4 = \psi(A)$.

Exercise 6.2. Compute the characteristic and minimal polynomials of

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{F}^{3\times 3}$$

and determine whether A and B are similar.

Exercise 6.3. Let $\hat{v} = (v_1, v_2, v_3)$ be an ordered basis of $\mathbb{R}^3$ and

$$f : \mathbb{R}^3 \to \mathbb{R}^3, \quad f(xv_1 + yv_2 + zv_3) = (3x + y)v_1 + (2y + z)v_2 + (-x - y + z)v_3.$$

Find the minimal polynomial of f and determine whether there exists an ordered basis $\hat{u}$ of $\mathbb{R}^3$ such that $(f : \hat{u}, \hat{u}) = A$, where A is the matrix from the previous exercise.

Exercise 6.4. Consider the linear map $f : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ defined by $f(\varphi(x)) = \varphi'(x) - 2\varphi(x)$.

a. Find the minimal polynomial of f and determine whether f is diagonalizable.

b. Find the dimension of each eigenspace of f.

Exercise 6.5. Let $A \in \mathbb{C}^{n\times n}$ such that $(A + 3_n)(A - 4\mathbb{I}_n)(A + 7\mathbb{I}_n) = 0$. Determine whether A is

a. diagonalizable,

b. invertible.

Exercise 6.6. Determine all $A \in \mathbb{R}^{3\times 3}$ such that $A^3 - 3A^2 + 2A = 0$ and $\mathrm{Tr}(A) = 6$.

Exercise 6.7. Find the characteristic polynomial and the minimal polynomial of

$$A = \begin{pmatrix} 2 & 5 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 2 & 0 \\ 0 & 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix} \in \mathbb{R}^{5\times 5}.$$

Determine whether A is diagonalizable.

Exercise 6.8. Let $n > 1$. Find the minimal polynomial of the linear map

$$f : \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}, \quad f(A) = A^t$$

and determine whether it is diagonalizable.

Exercise 6.9. If $f : V \to V$ is a linear map such that $f^3 = f$, then every $v \in V$ can be written uniquely as $v = v_{-1} + v_0 + v_1$, where $v_\lambda \in \ker(f - \lambda \mathbb{I}_n)$ for $\lambda = -1, 0, 1$.

Exercise 6.10. Show that $m_A(x) = m_{A^t}(x)$ for every $A \in \mathbb{F}^{n\times n}$.

Exercise 6.11. Let $\mathbb{F}^{n\times n}$ and let W_A be the subspace of $\mathbb{F}^{n\times n}$ generated by the matrices A^n with $n \geq 0$. Show that $\dim W_A = \deg m_A(x)$.

Exercise 6.12. Let $A, B, C \in \mathbb{F}^{n\times n}$ and let

$$D = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in \mathbb{F}^{2n\times 2n}.$$

a. Show that if D is diagonalizable, then A and C are diagonalizable.

b. Does the converse of (a) hold?

Exercise 6.13. Find the minimal polynomial of

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \in \mathbb{F}^{n\times n}.$$

Exercise 6.14. Show the following.

a. If $\deg m_A(x) = \deg \chi_A(x)$, then $m_A(x) = (-1)^n \chi_A(x)$.

b. We have $m_A(x) = x^n - 1$ and $m_B(x) = (x-1)^n$, where

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{F}^{n\times n}.$$

Exercise 6.15. Let

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & d & 2 & 0 \\ c & e & f & 2 \end{pmatrix} \in \mathbb{R}^{4\times 4}.$$

Prove that A is diagonalizable if and only if $a = f = 0$.

Exercise 6.16. Let

$$A = \begin{pmatrix} 1 & k & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

a. Find the values of k such that $\deg m_A(x) \leq 2$.

b. For the value of k found above, compute A^{-1} using $m_A(x)$.

c. Show that A^m is not diagonalizable for every positive integer m.

Exercise 6.17. Find the values of $c \in \mathbb{R}$ such that the polynomial $(x-3)^{1821}(x^{1821}-5x+c)$ is annihilated by the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & -1 \\ -6 & 0 & 3 \end{pmatrix}.$$

Exercise 6.18. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map with $m_f(x) = x(x-1)^2$. Find all $a, b, c \in \mathbb{R}$ such that $f^{1821} + af^2 + bf + c \cdot 1_{\mathbb{R}^n} = 0$.

Exercise 6.19. Let

$$A = \begin{pmatrix} a & 0 & a \\ 0 & 1 & 0 \\ 0 & a & -1 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

For each of the following cases find all values of $a \in \mathbb{R}$ (if any) for which the stated property holds.

a. There exists an invertible matrix $P \in \mathbb{R}^{3\times 3}$ such that $P^{-1}AP$ is upper triangular.

b. There exists an invertible matrix $P \in \mathbb{R}^{3\times 3}$ such that $P^{-1}AP$ is diagonal.

c. The matrix A is annihilated by the polynomial $(x-1)(x-2)\cdots(x-2010)$.

Exercise 6.20. Let $A \in \mathbb{C}^{n\times n}$ such that $A^m = \mathbb{I}_n$ for some positive integer m and $\text{Tr}(A) = n$. Prove that $A = \mathbb{I}_n$.

Exercise 6.21. Let $A, B \in \mathbb{R}^{2\times 2}$ with $A \neq \mathbb{I}_2$, $B \neq -\mathbb{I}_2$, and $A^3 - A^2 + A - \mathbb{I}_2 = 0$, $B^3 + B^2 + B + \mathbb{I}_2 = 0$.

a. Show that A and B have the same minimal polynomial.

b. Do they have the same characteristic polynomial?

c. Determine whether A and B are triangularizable.

Exercise 6.22. Let $A \in \mathbb{R}^{3\times 3}$ with $A^2 - 9A + 20\mathbb{I}_3 = 0$. Show that exactly one of the following cases holds:

$A = 4\mathbb{I}_3$ or $A = 5\mathbb{I}_3$ or A is similar to $\text{diag}(4,4,5)$ or A is similar to $\text{diag}(4,5,5)$.

Exercise 6.23. Let $A_i \in \mathbb{R}^{3\times 3}$, $i = 1, \cdots, 5$, with $A_i^2 - 9A_i + 20\mathbb{I}_3 = 0$. Show that two of the matrices A_i are similar.

Exercise 6.24. Let $f, g : V \to V$ be linear maps such that

$$\gcd(m_f(x), m_g(x)) = 1.$$

a. Show that the linear map $m_g(f) : V \to V$ is an isomorphism.

b. Show that if $\ker f \neq \{0_V\}$, then $\ker g = \{0_V\}$.

Exercise 6.25. Let

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 0 & \cdots & 0 & -a_3 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in \mathbb{F}^{n\times n}.$$

In Exercise 3.17 we saw that $\chi_A(x) = (-1)^n(x^n + a_{n-1}x^{n-1} + \cdots + a_0)$. Show that $m_A(x) = (-1)^n \chi_A(x)$.

Exercise 6.26. Let $A \in \mathbb{F}^{n\times n}$ and let $\varphi(x) \in \mathbb{F}[x]$. Show that

$$\varphi(A) \text{ is invertible if and only if } \gcd(\varphi(x), m_A(x)) = 1.$$

Exercise 6.27. Let $A \in \mathbb{R}^{n\times n}$ be an invertible, triangularizable matrix such that $m_A(x) = m_{A^2}(x)$. Show that $(A - \mathbb{I}_n)^n = 0$.

Exercise 6.28. a. Let $A, B \in \mathbb{C}^{2\times 2}$ such that $m_A(x) = m_B(x)$. Show that A and B are similar.

b. Let

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{4\times 4}.$$

Show that $\chi_C(x) = \chi_D(x)$ and $m_C(x) = m_D(x)$, but the matrices C, D are not similar.

Exercise 6.29. Let $A \in \mathbb{F}^{n\times n}$. Consider the linear map $R_A : \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n}$ defined by $R_A(B) = BA$. Show the following.

a. If $\sigma(x) \in \mathbb{F}[x]$, then $\sigma(R_A)(B) = B\sigma(A)$ for every $B \in \mathbb{F}^{n\times n}$, and

b. Is it true that $\chi_{R_A}(x) = \chi_A(x)$?

Exercise 6.30. Let $A, B \in \mathbb{F}^{n\times n}$. We know that $\chi_{AB}(x) = \chi_{BA}(x)$ (see Exercise 3.27). Is it true that $m_{AB}(x) = m_{BA}(x)$?

Exercise 6.31. Determine which of the following statements are true. In each case give a proof or a counterexample.

a. There exists $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = (x-1)(x+1)^3$ and $m_A(x) = (x-1)^2(x+1)$.

b. Let $A \in \mathbb{C}^{n\times n}$ such that $A^5 + 5A + \mathbb{I}_n = 0$. Then A is diagonalizable.

c. There exists $A \in \mathbb{R}^{3\times 3}$ with $m_A(x) = (x-1)(x-3)$ and A similar to a matrix of the form $\begin{pmatrix} * & 0 & * \\ * & 2 & * \\ * & 0 & * \end{pmatrix}$?

Exercise 6.32. If $f : V \to V$ is a diagonalizable linear map and U is an f-invariant subspace of V, then the restriction of f to U is diagonalizable.

Exercise 6.33. Let $A \in \mathbb{F}^{n\times n}$ with $\det A = 0$. Show that there exists a nonzero $B \in \mathbb{F}^{n\times n}$ with $AB = BA = 0$.

Exercise 6.34. Let $A, B \in \mathbb{F}^{n\times n}$ such that $A^3 = A$ and $B^3 = B$. Show the following.

a. A is diagonalizable and $\text{rank}(A) = \text{Tr}(A^2)$.

b. The matrices A, B are similar if and only if $\text{rank}(A) = \text{rank}(B)$ and $\text{Tr}(A) = \text{Tr}(B)$.

Exercise 6.35. Let $A, B \in \mathbb{F}^{n\times n}$ with $A^2 - 3A = B^2 - 3B = 0$. Show the following:

a. If $\text{Tr}(A) = \text{Tr}(B)$, then A and B are similar.

b. If $AB = BA$, then $\varphi(A + B)$ is diagonalizable for every $\varphi(x) \in \mathbb{F}[x]$.

Exercise 6.36. Let $A \in \mathbb{R}^{4\times 4}$ with $\chi_A(x) = x^2(x-1)(x-2)$. Show that if $AX = AY = 0$, where $X = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $Y = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$, then $A^3 = 3A^2 - 2A$.

Exercise 6.37. Let $A, B \in \mathbb{C}^{6\times 6}$ with $m_A(x) = (x-1)(x-2)$ and $\chi_B(x) = (x-3)^4(x-4)^2$. Show that if $V_A(1) \subseteq V_B(3)$ and $V_A(2) \subseteq V_B(4)$, then $AB = BA$.

Exercise 6.38. Let $A, B \in \mathbb{C}^{n\times n}$ such that $AB = BA$, $A^{1821} = B^{1821} = \mathbb{I}_n$. Then $A + B + \mathbb{I}_n$ is diagonalizable and invertible.

Exercise 6.39. Show that a matrix $A \in \mathbb{F}^{n\times n}$ is diagonalizable if and only if there exist $a_i \in \mathbb{F}$ and $P_i \in \mathbb{F}^{n\times n}$ such that $A = a_1P_1 + \cdots + a_kP_k$, $P_i^2 = P_i$, and $P_iP_j = P_jP_i$ for all i, j.

Exercise 6.40. Let $A, B \in \mathbb{F}^{n\times n}$. Show that either $m_{AB}(x) = m_{BA}(x)$ or $m_{AB}(x) = x\, m_{BA}(x)$ or $m_{BA}(x) = x\, m_{AB}(x)$.

Exercise 6.41 (Review exercise)**.** Determine which of the following statements are true. In each case give a proof or a counterexample. Let $A \in \mathbb{F}^{n\times n}$.

a. $A^m = 0$ for some positive integer $m \iff A^n = 0$.

b. A is invertible $\iff m_A(0) \neq 0$.

c. If $A^2 = 4A$, then A is diagonalizable.

d. If $B \in \mathbb{F}^{2n\times 2n}$ and $B = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}$, then $m_B(x) = m_A(x)$.

e. If A is invertible, then $m_{AB}(x) = m_{BA}(x)$ for every $B \in \mathbb{F}^{n\times n}$.

CHAPTER 7

THE STANDARD INNER PRODUCT ON $\mathbb{R}^n$ AND $\mathbb{C}^n$

7.1 The Standard Inner Product

Definition 7.1.1. The map

$$\langle \cdot, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad \langle u, v \rangle = u_1 v_1 + \cdots + u_n v_n,$$

where $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, is called the **standard inner product** on $\mathbb{R}^n$.

- If $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$, the **length** (norm) of u is

$$|u| = \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{\langle u, u \rangle}.$$

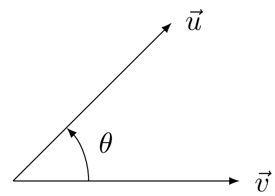
- The vectors u, v are said to be **orthogonal** (perpendicular) if $\langle u, v \rangle = 0$.

Observation 7.1.1. For $n = 2$, one can prove that

$$\cos \vartheta = \frac{\langle u, v \rangle}{|u| \cdot |v|},$$

where ϑ is the angle between the vectors u and v. Hence, $\cos \vartheta = 0$ if and only if

$$\langle u, v \rangle = 0.$$



Example 7.1.1. Let $u, v \in \mathbb{R}^3$ with $u = (1, 1, 2)$ and $v = (-1, -1, 1)$. Then

$$\langle u, v \rangle = (-1) \cdot 1 + (-1) \cdot 1 + 1 \cdot 2 = 0,$$

so u and v are orthogonal.

Properties 7.1.1. For any vectors $u, v, w \in \mathbb{R}^n$ and any $a \in \mathbb{R}$, the following properties of the inner product hold:

i. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,

ii. $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$,

iii. $\langle au, v \rangle = a\langle u, v \rangle$,

iv. $\langle u, av \rangle = a\langle u, v \rangle$,

v. $\langle u, v \rangle = \langle v, u \rangle$,

vi. $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0 \Leftrightarrow u = 0$.

Proof. The proof of the above properties is a simple exercise and is left to the reader (yes, you). □

Example 7.1.2. Let $u, v \in \mathbb{R}^n$ be two orthogonal vectors. Prove:

i. $|u + v|^2 = |u|^2 + |v|^2$,

ii. If $|u| = |v|$, then the vectors $u + v$ and $u - v$ are orthogonal.

Proof. i. Compute:

$$|u + v|^2 = \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle.$$

Since $u \perp v$ we have $\langle u, v \rangle = \langle v, u \rangle = 0$, hence

$$|u + v|^2 = |u|^2 + |v|^2.$$

ii. Compute:

$$\langle u+v, u-v\rangle = \langle u,u\rangle - \langle u,v\rangle + \langle v,u\rangle - \langle v,v\rangle.$$

Since $u \perp v$ and $|u| = |v|$, it follows that

$$\langle u+v, u-v\rangle = |u|^2 - |v|^2 = 0.$$

Thus $u+v$ and $u-v$ are orthogonal.

□

Reminder 7.1.1. The set of **complex numbers** is defined by

$$\mathbb{C} = \left\{ a+bi \mid a,b \in \mathbb{R},\ i^2 = -1 \right\}.$$

- Every $z \in \mathbb{C}$ can be written uniquely as $z = a+bi$, with $a,b \in \mathbb{R}$.
- The **complex conjugate** of $z = a+bi$ is $\overline{z} = a - bi$.
- The **modulus** of z is $|z| = \sqrt{a^2+b^2}$.
- Additional properties:

$$|z|^2 = z\,\overline{z}, \quad \overline{z_1+z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}.$$

Definition 7.1.2. The **standard inner product** on $\mathbb{C}^n$ is the map

$$\langle\cdot,\cdot\rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \quad \langle u,v\rangle = u_1\overline{v_1} + u_2\overline{v_2} + \cdots + u_n\overline{v_n},$$

where $u = (u_1,\ldots,u_n)$ and $v = (v_1,\ldots,v_n)$.

- The **length** of u is

$$|u| = \sqrt{|u_1|^2 + \cdots + |u_n|^2} = \sqrt{\langle u,u\rangle}.$$

- The vectors u, v are **orthogonal** if $\langle u,v\rangle = 0$.

Example 7.1.3. If $u = (1,i)$ and $v = (-1,i)$, then

$$\langle u,u\rangle = 1\cdot\overline{1} + i\cdot\overline{i} = 1+1 = 2, \quad \langle u,v\rangle = 1\cdot(-1) + i\cdot\overline{i} = -1+1 = 0.$$

Hence u and v are orthogonal.

Properties 7.1.2. For any vectors $u, v, w \in \mathbb{C}^n$ and any $a \in \mathbb{C}$, the following hold:

i. $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$

ii. $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$

iii. $\langle au, v\rangle = a\langle u, v\rangle$

iv. $\langle u, av\rangle = \overline{a}\langle u, v\rangle$

v. $\langle u, v\rangle = \overline{\langle v, u\rangle}$

vi. $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0 \Leftrightarrow u = 0$

Proof. As an example, we prove property (iv).

Let $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ be vectors in $\mathbb{C}^n$. Then

$$av = (av_1, \ldots, av_n),$$

and therefore

$$\langle u, av\rangle = u_1\overline{av_1} + \cdots + u_n\overline{av_n} = \overline{a}\,(u_1\overline{v_1} + \cdots + u_n\overline{v_n}) = \overline{a}\,\langle u, v\rangle.$$

□

7.2 Orthonormal Bases

7.2.1 Orthonormal bases and the Gram–Schmidt method

Definition 7.2.1. Let $V \subseteq \mathbb{F}^n$. A basis $\{v_1, \ldots, v_m\}$ of V is called **orthonormal** if:

i. $|v_i| = 1$ for each $i = 1, \ldots, m$,

ii. $\langle v_i, v_j\rangle = 0$ for every $i \neq j$.

Example 7.2.1. For $V = \mathbb{R}^2$, the standard basis $\hat{e} = \{e_1 = (1,0),\, e_2 = (0,1)\}$ is orthonormal. Likewise, the basis

$$\left\{ \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \right\}$$

is an orthonormal basis of $\mathbb{R}^2$ (show why).

These examples can be generalized as follows:

For every angle φ, the vectors $u = (\cos\varphi, \sin\varphi)$ and $v = (-\sin\varphi, \cos\varphi)$ form an orthonormal basis of $\mathbb{R}^2$.

Indeed,

$$|u| = \sqrt{\cos^2\varphi + \sin^2\varphi} = 1, \quad |v| = \sqrt{\sin^2\varphi + \cos^2\varphi} = 1,$$

and

$$\langle u, v \rangle = \cos\varphi(-\sin\varphi) + \sin\varphi\cos\varphi = 0.$$

Observation 7.2.1. Let $\{v_1, v_2, \ldots, v_m\}$ be an orthonormal basis of V and let $v \in V$. Then there exist unique scalars $a_i \in \mathbb{F}$ such that

$$v = a_1 v_1 + \cdots + a_m v_m.$$

Moreover,

$$a_i = \langle v, v_i \rangle.$$

Indeed,

$$\langle v, v_i \rangle = \sum_{j=1}^{m} a_j \langle v_j, v_i \rangle = a_i.$$

Proposition 7.2.1. Every nonzero subspace of $\mathbb{R}^2$ has an orthonormal basis.

Proof. Let $V \leq \mathbb{R}^2$. By the basis existence theorem, V has a basis $\{u_1, u_2\}$. Define

$$v_1 = u_1$$

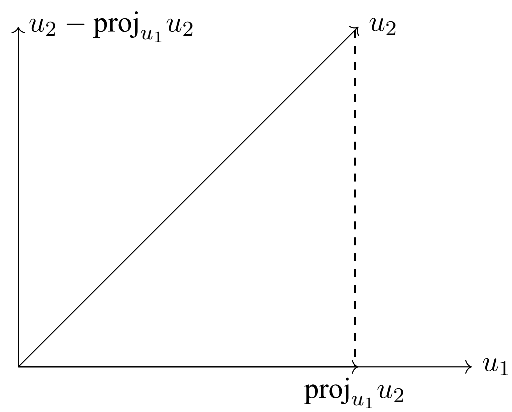
and

$$v_2 = u_2 - \operatorname{proj}_{u_1} u_2,$$

where $\text{proj}_{u_1} u_2$ is the projection of u_2 onto u_1. Then v_1 and v_2 are orthogonal, and

$$\left\{ \frac{v_1}{|v_1|}, \frac{v_2}{|v_2|} \right\}$$

is an orthonormal basis of V. □



Next we study the Gram–Schmidt method and see that the above result generalizes to every subspace of $\mathbb{R}^n$.

Gram–Schmidt Orthonormalization Method

We first describe the method for $n = 3$. Let $V \leq \mathbb{F}^3$ and let $\{u_1, u_2, u_3\}$ be a basis of V. Define the following vectors in V:

i. $v_1 = u_1$

ii. $v_2 = u_2 - \dfrac{\langle u_2, v_1 \rangle}{|v_1|^2} v_1$, where $\text{proj}_{v_1} u_2 = \frac{\langle u_2, v_1 \rangle}{|v_1|^2} v_1$

iii. $v_3 = u_3 - \dfrac{\langle u_3, v_2 \rangle}{|v_2|^2} v_2 - \dfrac{\langle u_3, v_1 \rangle}{|v_1|^2} v_1$, where the terms are the corresponding projections

Then

$$\left\{ \frac{v_1}{|v_1|}, \frac{v_2}{|v_2|}, \frac{v_3}{|v_3|} \right\}$$

is an **orthonormal** basis of V (show why).

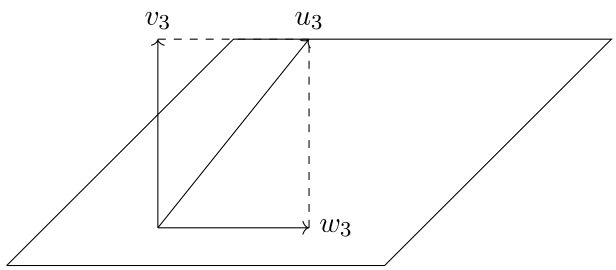


Figure for $n = 3$.

General case. Let $\{u_1, \ldots, u_m\}$ be a basis of V and define recursively $v_1, \ldots, v_m$ by

$$v_1 = u_1, \quad v_i = u_i - \sum_{j=1}^{i-1} \frac{\langle u_i, v_j \rangle}{|v_j|^2} v_j, \quad \text{for } i \geq 2.$$

Then

$$\left\{ \frac{v_1}{|v_1|}, \ldots, \frac{v_m}{|v_m|} \right\}$$

is an **orthonormal** basis of V.

Theorem 7.2.1. Every nonzero subspace of $\mathbb{F}^n$ has an orthonormal basis.

Proof. Let $V \leq \mathbb{F}^n$. By the basis existence theorem, V has a basis $\{u_1, \ldots, u_m\}$. Applying Gram–Schmidt to this basis yields the desired result. □

Example 7.2.2. Find an orthonormal basis of

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}.$$

Proof. First find a basis of V:

$$\begin{aligned} V &= \{(x, y, -x - y) \mid x, y \in \mathbb{R}\} \\ &= \{x(1, 0, -1) + y(0, 1, -1) \mid x, y \in \mathbb{R}\} \\ &= \langle (1, 0, -1),\ (0, 1, -1) \rangle. \end{aligned}$$

Thus $\{u_1, u_2\}$ is a basis, with $u_1 = (1, 0, -1)$ and $u_2 = (0, 1, -1)$. Apply Gram–Schmidt:

$$v_1 = u_1 = (1, 0, -1), \qquad v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{|v_1|^2} v_1 = \left(-\frac{1}{2}, 1, -\frac{1}{2}\right).$$

Therefore an orthonormal basis is

$$\left\{\frac{v_1}{|v_1|}, \frac{v_2}{|v_2|}\right\}, \quad \text{where} \quad \frac{v_1}{|v_1|} = \frac{1}{\sqrt{2}}(1, 0, -1), \quad \frac{v_2}{|v_2|} = \frac{1}{\sqrt{6}}(-1, 2, -1).$$

□

Example 7.2.3. Find an orthonormal basis of $V = \langle a, b, c \rangle$, where

$$a = (1, 1, 1, 1), \quad b = (1, 1, 1, -1), \quad c = (3, 3, 3, -1).$$

Proof. To find a basis of V we row-reduce the matrix with rows a, b, c:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 3 & 3 & 3 & -1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence $\{u_1, u_2\}$ is a basis of V, where $u_1 = (1, 1, 1, 0)$ and $u_2 = (0, 0, 0, 1)$. Here $\langle u_1, u_2 \rangle = 0$, so Gram–Schmidt is not needed. Set $v_1 = u_1$ and $v_2 = u_2$, and obtain

$$\frac{v_1}{|v_1|} = \frac{1}{\sqrt{3}}(1, 1, 1, 0), \quad \frac{v_2}{|v_2|} = (0, 0, 0, 1),$$

which form an orthonormal basis of V. □

7.2.2 Orthogonal complement of a subspace of $\mathbb{F}^n$

Definition 7.2.2. Let $V \leq \mathbb{F}^n$. The **orthogonal complement** of V is the set

$$V^{\perp} = \{u \in \mathbb{F}^n \mid \langle u, v \rangle = 0 \text{ for every } v \in V\}.$$

Proposition 7.2.2. $V^{\perp}$ is a subspace of $\mathbb{F}^n$.

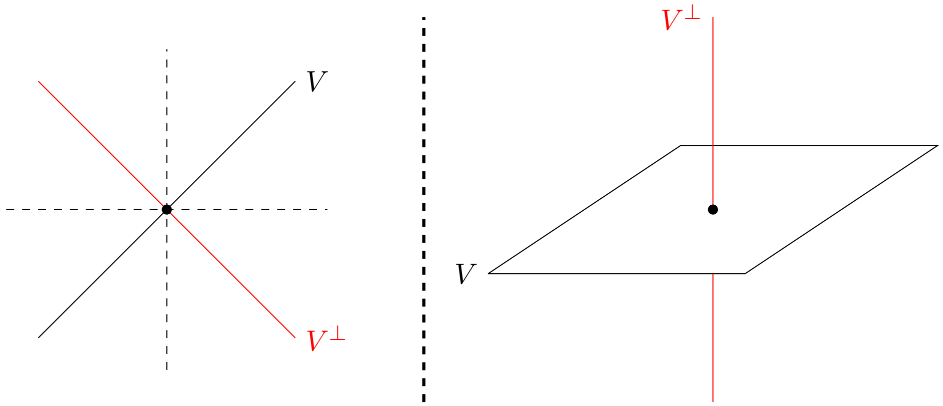
Proof. First, $V^\perp \neq \emptyset$ since $0_{\mathbb{F}^n} \in V^\perp$. Let $u_1, u_2 \in V^\perp$, i.e. $\langle u_1, v\rangle = \langle u_2, v\rangle = 0$ for all $v \in V$. Then

$$\langle u_1 - u_2, v\rangle = \langle u_1, v\rangle - \langle u_2, v\rangle = 0,$$

so $u_1 - u_2 \in V^\perp$. Also, if $u \in V^\perp$ and $\lambda \in \mathbb{F}$, then

$$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle = 0,$$

so $\lambda u \in V^\perp$. □



Example 7.2.4.

Proposition 7.2.3. Let $V \leq \mathbb{F}^n$. Then:

i. $\mathbb{F}^n = V \oplus V^\perp$,

ii. $\dim V^\perp = n - \dim V$,

iii. If $V \leq U$, then $U^\perp \leq V^\perp$,

iv. $(V^\perp)^\perp = V$.

Proof. i. We show that $\mathbb{F}^n = V + V^\perp$ and that $V \cap V^\perp = \{0\}$. By Theorem 7.2.1, V has an orthonormal basis $\{v_1, \ldots, v_m\}$ (if $V = \{0\}$ the claim is immediate).

Let $w \in \mathbb{F}^n$. Write $w = v + (w - v)$, where

$$v = \sum_{i=1}^{m} \langle w, v_i\rangle\, v_i \in V.$$

Then $w - v \in V^{\perp}$, since for each i,

$$\langle w - v, v_i \rangle = \langle w, v_i \rangle - \langle v, v_i \rangle = \langle w, v_i \rangle - \sum_t \langle w, v_t \rangle \langle v_t, v_i \rangle = \langle w, v_i \rangle - \langle w, v_i \rangle = 0.$$

Thus $w \in V + V^{\perp}$. Finally, if $v \in V \cap V^{\perp}$, then $\langle v, v \rangle = 0 \Leftrightarrow v = 0$, hence $V \cap V^{\perp} = \{0\}$.

ii. Immediate from (i), since

$$n = \dim(V \oplus V^{\perp}) = \dim V + \dim V^{\perp}.$$

iii. Immediate from the definition: if $\langle w, u \rangle = 0$ for all $u \in U$, then in particular $\langle w, v \rangle = 0$ for all $v \in V$, so $w \in V^{\perp}$.

iv. Since every $v \in V$ satisfies $\langle v, u \rangle = 0$ for all $u \in V^{\perp}$, we have $V \subseteq (V^{\perp})^{\perp}$. Also,

$$\dim(V^{\perp})^{\perp} = n - \dim V^{\perp} = n - (n - \dim V) = \dim V,$$

so $(V^{\perp})^{\perp} = V$.

□

How to find $V^{\perp}$

Proposition 7.2.4 (Extending an orthonormal basis)**.** Let $V \leq \mathbb{F}^n$. By Theorem 7.2.1, there exists an orthonormal basis of V, say $\{v_1, \dots, v_m\}$. Then there exists an orthonormal basis of $\mathbb{F}^n$ of the form

$$\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}.$$

Proof. Let $\{v_1, \dots, v_m\}$ be an orthonormal basis of V. By the basis extension theorem,

$$\{v_1, \dots, v_m, u_1, \dots, u_k\}$$

is a basis of $\mathbb{F}^n$. Apply Gram–Schmidt to this basis to obtain an orthonormal basis

$$\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$$

of $\mathbb{F}^n$. □

With this notation, $\{v_1, \dots, v_m\}$ is an orthonormal basis of V and $\{v_{m+1}, \dots, v_n\}$ is an orthonormal basis of $V^{\perp}$.

In summary, one method to find an orthonormal basis of $V^{\perp}$ is:

i. Start with a basis $B = \{u_1, \dots, u_m\}$ of V.

ii. Apply Gram–Schmidt to B to obtain an orthonormal basis

$$B' = \{v_1, \dots, v_m\}$$

of V.

iii. Extend B' to a basis of $\mathbb{F}^n$:

$$B'' = \{v_1, \dots, v_m, w_1, \dots, w_k\}.$$

iv. Apply Gram–Schmidt to B'' to obtain an orthonormal basis $\{v_1, \dots, v_n\}$ of $\mathbb{F}^n$. Then

$$W = \{v_{m+1}, \dots, v_n\}$$

is an orthonormal basis of $V^\perp$.

7.3 Hermitian and Unitary Matrices

7.3.1 The matrix A^*, the standard inner product, and matrix multiplication

Definition 7.3.1. Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$. The matrix

$$\overline{A} = \overline{(a_{ij})}$$

is called the **conjugate** of A, and

$$A^* = (\overline{A})^t$$

is the **conjugate transpose** (adjoint) of A.

Example 7.3.1. Consider

$$A = \begin{pmatrix} 2-i & 4 \\ 5+2i & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2}.$$

Then

$$\overline{A} = \begin{pmatrix} 2+i & 4 \\ 5-2i & 0 \end{pmatrix}, \quad A^* = \begin{pmatrix} 2+i & 5-2i \\ 4 & 0 \end{pmatrix}.$$

We will often regard the standard inner product as defined on column vectors, i.e. $\langle \cdot, \cdot \rangle : \mathbb{C}^{n\times 1} \times \mathbb{C}^{n\times 1} \to \mathbb{C}$. If

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^{n\times 1},$$

then

$$\langle X, Y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}.$$

Note that $\langle X, Y \rangle = X^t \overline{Y}$.

Lemma 7.3.1. Let $A \in \mathbb{C}^{n\times n}$. Then:

i. For all $X, Y \in \mathbb{C}^{n\times 1}$, $\langle AX, Y \rangle = \langle X, A^* Y \rangle$.

ii. If $\langle AX, Y \rangle = 0$ for all $X, Y \in \mathbb{C}^{n\times 1}$, then $A = 0$.

Proof. i. We have

$$\langle AX, Y \rangle = (AX)^t \overline{Y}.$$

Also,

$$\langle X, A^* Y \rangle = X^t \overline{A^* Y} = X^t A^t \overline{Y} = (AX)^t \overline{Y}.$$

Hence they are equal.

ii. Take $X = E_i$ and $Y = E_j$, where E_k is the kth standard basis column. Then

$$\langle AX, Y \rangle = \langle AE_i, E_j \rangle = a_{ji}.$$

Since this is 0 for all i, j, all entries of A are zero, hence $A = 0$.

□

7.3.2 Hermitian Matrices

Definition 7.3.2. A matrix $A \in \mathbb{C}^{n\times n}$ is called **Hermitian** if $A^* = A$.

Observation 7.3.1. i. A matrix $A \in \mathbb{R}^{n\times n}$ is Hermitian if and only if it is symmetric.

ii. A matrix $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ is Hermitian if and only if $\overline{a_{ii}} = a_{ii}$ and $\overline{a_{ij}} = a_{ji}$ for all $i \neq j$.

Example 7.3.2. i. The matrix
$$A = \begin{pmatrix} 2 & 4-5i \\ 4+5i & 6 \end{pmatrix}$$
is Hermitian since
$$A^* = (\overline{A})^t = A.$$

ii. The matrix
$$B = \begin{pmatrix} 2 & 4-3i \\ 4+5i & 6 \end{pmatrix}$$
is not Hermitian since
$$B^* = \begin{pmatrix} 2 & 4-5i \\ 4+3i & 6 \end{pmatrix} \neq B.$$

Properties 7.3.1. Let $A \in \mathbb{C}^{n\times n}$ be Hermitian. Then:

i. Every eigenvalue of A is real.

ii. Eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof. i. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A with eigenvector $X \neq 0$, so $AX = \lambda X$. Then
$$\langle AX, X\rangle = \langle \lambda X, X\rangle = \lambda\langle X, X\rangle.$$
By Lemma 7.3.1,
$$\langle AX, X\rangle = \langle X, A^*X\rangle = \langle X, AX\rangle = \langle X, \lambda X\rangle = \overline{\lambda}\,\langle X, X\rangle.$$
Hence
$$(\lambda - \overline{\lambda})\langle X, X\rangle = 0.$$
Since $X \neq 0$, we have $\langle X, X\rangle \neq 0$, therefore $\lambda = \overline{\lambda}$, i.e. $\lambda \in \mathbb{R}$.

ii. Let $\lambda \neq \mu$ be eigenvalues with eigenvectors X, Y:
$$AX = \lambda X, \qquad AY = \mu Y.$$
Then
$$\langle AX, Y\rangle = \langle \lambda X, Y\rangle = \lambda\langle X, Y\rangle,$$
and also
$$\langle AX, Y\rangle = \langle X, A^*Y\rangle = \langle X, AY\rangle = \langle X, \mu Y\rangle = \overline{\mu}\,\langle X, Y\rangle.$$

By (i), $\mu \in \mathbb{R}$, so $\overline{\mu} = \mu$. Thus

$$(\lambda - \mu)\langle X, Y\rangle = 0.$$

Since $\lambda \neq \mu$, it follows that $\langle X, Y\rangle = 0$, i.e. $X \perp Y$.

□

The reader is invited to compare Property 7.3.1(ii) with the familiar fact for arbitrary matrices, namely that distinct eigenvalues correspond to linearly independent eigenvectors.

7.3.3 Unitary Matrices

Definition 7.3.3. A matrix $A \in \mathbb{C}^{n\times n}$ is called **unitary** if

$$AA^* = A^*A = I_n.$$

Observation 7.3.2. Let $A \in \mathbb{C}^{n\times n}$.

i. A is unitary if and only if it is invertible and $A^{-1} = A^*$.

ii. A is unitary if and only if $AA^* = I_n \Leftrightarrow A^*A = I_n$.

iii. If $A \in \mathbb{R}^{n\times n}$, then A is unitary (orthogonal) if and only if it is invertible and $A^{-1} = A^t$.

Example 7.3.3. i. The identity matrix I_n is unitary.

ii. The matrix

$$A_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

is unitary (orthogonal), since $A_\varphi A_\varphi^t = I_2$. Geometrically, A_φ represents a rotation of the plane by angle φ.

iii. The matrix

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

is unitary (orthogonal), since $A = A^t$ and $AA^t = I_3$.

iv. The matrix
$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$
is not unitary, since $AA^t = 5I_2$. However, the matrix $\frac{1}{\sqrt{5}}A$ is unitary.

Proposition 7.3.1. Let $A, B \in \mathbb{C}^{n\times n}$ be unitary. Then:

i. $|\det A| = 1$,

ii. the matrices AB and AB^{-1} are unitary,

iii. the matrix $\left(\begin{array}{c|c} 1 & O \\ \hline O & A \end{array}\right) \in \mathbb{C}^{(n+1)\times(n+1)}$ is unitary.

Proof. i. Since A is unitary, $AA^* = I_n$. Taking determinants,
$$\det A \cdot \det A^* = \det A \cdot \det \overline{A} = \det A \cdot \overline{\det A} = |\det A|^2 = 1,$$
so $|\det A| = 1$.

ii. Since A, B are unitary,
$$(AB)(AB)^* = ABB^*A^* = AA^* = I_n,$$
and similarly $(AB)^*(AB) = I_n$, so AB is unitary. Also $B^*B = I_n$ implies $B^{-1}(B^{-1})^* = I_n$, hence B^{-1} is unitary, so AB^{-1} is unitary.

iii. Immediate, since
$$\left(\begin{array}{c|c} 1 & O \\ \hline O & A \end{array}\right)^* = \left(\begin{array}{c|c} 1 & O \\ \hline O & A^* \end{array}\right).$$

□

Theorem 7.3.1 (Characterizations of unitary matrices)**.** Let $A \in \mathbb{C}^{n\times n}$. The following are equivalent:

i. A is unitary.

ii. $\langle AX, AY\rangle = \langle X, Y\rangle$ for all $X, Y \in \mathbb{C}^{n\times 1}$.

iii. The columns of A form an orthonormal basis of $\mathbb{C}^{n\times 1}$.

iv. The rows of A form an orthonormal basis of $\mathbb{C}^{n\times 1}$.

v. $|AX| = |X|$ for all $X \in \mathbb{C}^{n\times 1}$.

Proof.

- i. $\Rightarrow$ ii. If A is unitary, then $A^*A = I_n$. Hence

$$\langle AX, AY\rangle = \langle X, A^*AY\rangle = \langle X, Y\rangle.$$

- ii. $\Rightarrow$ iii. Take $X = E_i$ and $Y = E_j$. Then $\langle AE_i, AE_j\rangle = \langle E_i, E_j\rangle$, so

$$\langle A^{(i)}, A^{(j)}\rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

  i.e. the columns of A are orthonormal.

- iii. $\Rightarrow$ iv. From $\langle A^{(i)}, A^{(j)}\rangle = \delta_{ij}$ we get $A^t\overline{A} = I_n$, i.e. $A^*A = I_n$, hence the rows are orthonormal.

- iv. $\Rightarrow$ i. If the rows are orthonormal, then $AA^* = I_n$, so A is unitary.

- ii. $\Rightarrow$ v. Put $X = Y$.

- v. $\Rightarrow$ ii. Assume $|AX| = |X|$ for all X. For arbitrary X, Y,

$$|A(X+Y)| = |X+Y|.$$

  Squaring and expanding gives

$$\langle AX, AY\rangle + \langle AY, AX\rangle = \langle X, Y\rangle + \langle Y, X\rangle \quad (\alpha).$$

  Replace Y by iY to get

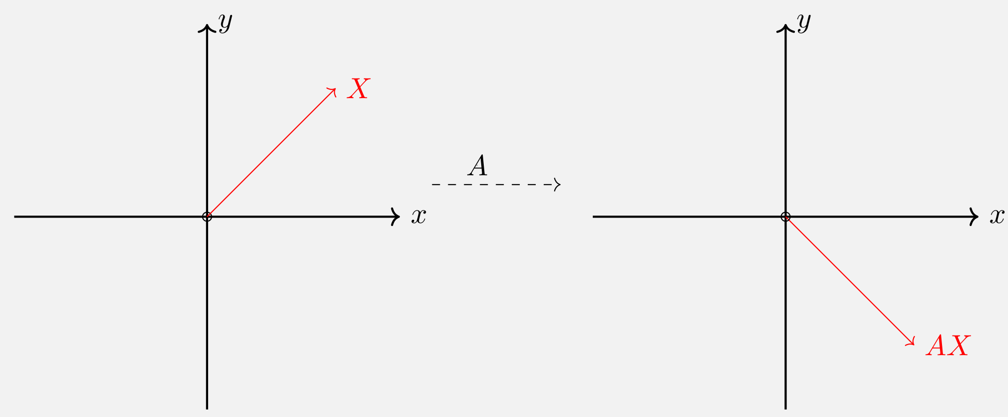
$$-i\langle AX, AY\rangle + i\langle AY, AX\rangle = -i\langle X, Y\rangle + i\langle Y, X\rangle \quad (\beta).$$

  Adding (α) and $i \cdot (\beta)$ yields $\langle AX, AY\rangle = \langle X, Y\rangle$.

□

After proving the theorem, let us add two remarks.

i. Property (v) says that a unitary matrix preserves lengths.

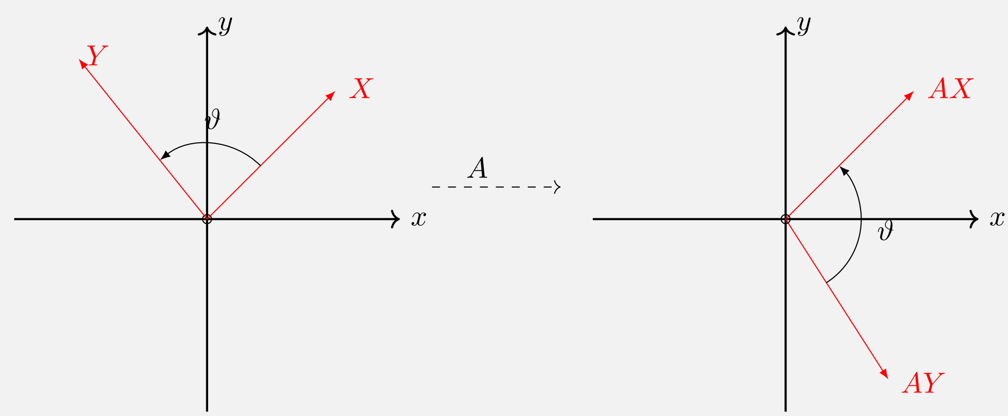


Preservation of vector length.

ii. Property (ii) shows that angles are preserved: if

$$\cos\vartheta = \frac{\langle X, Y\rangle}{|X|\cdot|Y|},$$

then $\cos\vartheta$ does not change under a unitary transformation.



Preservation of angles.

Corollary 7.3.1. Every eigenvalue of a unitary matrix has modulus 1.

Proof. Let $A \in \mathbb{C}^{n\times n}$ be unitary. The claim follows immediately from Theorem 7.3.1(v) applied to an eigenvector $X \neq 0$ of A. $\square$

7.4 Exercises of Chapter 7.

Group A: 1, 2, 3, 4, 5, 7, 8, 10, 11, 13
Group B: 6, 9, 12, 14, 16, 17, 18, 19
Group C: 15

Exercise 7.1. Let $u, v \in \mathbb{R}^n$.

a. If $\langle u, v \rangle = 0$, then $|u+v|^2 = |u|^2 + |v|^2$. When $n = 2$, this is the Pythagorean theorem.

b. If $|u| = |v|$, then $u + v$ and $u - v$ are orthogonal. When $n = 2$, this says that the diagonals of a rhombus are perpendicular.

c. $|u+v|^2 + |u-v|^2 = 2|u|^2 + 2|v|^2$. Give a geometric interpretation when $n = 2$.

Exercise 7.2. a. Find an orthonormal basis of $\mathbb{R}^3$ that contains the vector $u = \frac{1}{\sqrt{2}}(1, 0, 1)$.

b. Find an orthonormal basis of $\mathbb{R}^3$ that contains the vectors $u_1 = \frac{1}{\sqrt{5}}(1, 0, 2)$ and $u_2 = (0, 1, 0)$.

Exercise 7.3. Let $\{u_1, \ldots, u_n\}$ be an orthonormal basis of $\mathbb{C}^n$ and let $V = \langle u_1, \ldots, u_k \rangle$ with $1 \leq k < n$. Show that an orthonormal basis of $V^\perp$ is $\{u_{k+1}, \ldots, u_n\}$.

Exercise 7.4. Let V be the subspace of $\mathbb{R}^4$ generated by

$$v_1 = (1, 1, -1, -1), \quad v_2 = (1, 2, 3, -1), \quad v_3 = (4, 7, 8, -4).$$

After finding a basis of V, find an orthonormal basis of V and an orthonormal basis of $V^\perp$.

Exercise 7.5. Let

$$V = \{(x, y, z) \in \mathbb{R}^3 : x - 2y + z = 0\}, \qquad W = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}.$$

Find an orthonormal basis for each of the subspaces V, $V^\perp$, $V \cap W$, and $(V \cap W)^\perp$.

Exercise 7.6. Let $W_1, W_2 \leq \mathbb{F}^n$. Show that $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ and $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

Exercise 7.7. Let $A, B \in \mathbb{C}^{n \times n}$. Prove:

a. $(\overline{A})^t = \overline{A^t}$.

b. $\det(A^*) = \det(\overline{A}) = \overline{\det(A)}$.

c. $(A^*)^* = A$.

d. $(\lambda A)^* = \overline{\lambda} A^*$ for every $\lambda \in \mathbb{C}$.

e. $(A + B)^* = A^* + B^*$.

f. $(AB)^* = B^* A^*$.

g. If A is invertible, then $(A^*)^{-1} = (A^{-1})^*$.

Exercise 7.8. Let $A \in \mathbb{C}^{n\times n}$. If $\phi(x) \in \mathbb{C}[x]$, $\phi(x) = a_n x^n + \cdots + a_1 x + a_0$, define

$$\overline{\phi}(x) = \overline{a_n} x^n + \cdots + \overline{a_1} x + a_0.$$

Show:

a. $\chi_{A^*}(x) = \overline{\chi_A}(x)$.

b. $m_{A^*}(x) = \overline{m_A}(x)$.

c. λ is an eigenvalue of A iff $\overline{\lambda}$ is an eigenvalue of A^*.

Exercise 7.9. Let $A \in \mathbb{C}^{n\times n}$ with $A^* A = -A$. Show that A is similar to a diagonal matrix of the form

$$\text{diag}(0, \ldots, 0, -1, \ldots, -1),$$

and that $\text{rank}(A) + \text{rank}(A + I_n) = n$.

Exercise 7.10. Let $A, B \in \mathbb{C}^{n\times n}$ be unitary. Show:

a. $\overline{A}$, A^t, and A^{-1} are unitary.

b. If λ is an eigenvalue of A, then $|\lambda| = 1$ and $\frac{1}{\lambda}$ is an eigenvalue of A^*.

c. $|\det A| = 1$.

d. AB and AB^{-1} are unitary.

Exercise 7.11. Find a unitary matrix whose first row is $\begin{pmatrix} \frac{1}{\sqrt{10}} & 0 & \frac{3}{\sqrt{10}} \end{pmatrix}$.

Exercise 7.12. Let $U \in \mathbb{C}^{n\times n}$ be unitary with $\det(U - I_n) \neq 0$. Then the matrix $H \in \mathbb{C}^{n\times n}$ defined by

$$iH = (U + I_n)(U - I_n)^{-1}$$

is Hermitian.

Exercise 7.13. Let $A \in \mathbb{C}^{n\times n}$. Show that if any two of the following hold, then the third holds as well:

a. A is Hermitian.

b. A is unitary.

c. $A^2 = I_n$.

Exercise 7.14. Let $A \in \mathbb{R}^{n\times n}$ be unitary (orthogonal). Show:

a. If $\det A = 1$ and n is odd, then 1 is an eigenvalue of A.

b. If $\det A = -1$ and n is even, then 1 is an eigenvalue of A.

c. If $\det A = -1$, then -1 is an eigenvalue of A.

Exercise 7.15. Let $A, B \in \mathbb{R}^{n\times n}$ be unitary (orthogonal) with $\det A = -\det B$. Then

$$\det(A + B) = 0.$$

Exercise 7.16. Let $A \in \mathbb{C}^{n\times n}$ with $A^* = -A$. Show:

a. Every eigenvalue of A is of the form $i\mu$ with $\mu \in \mathbb{R}$.

b. The matrix $A + I_n$ is invertible and $\det(A + I_n) > 1$.

c. The matrix $(I_n - A)(I_n + A)^{-1}$ is unitary.

Exercise 7.17. Let $A \in \mathbb{C}^{n\times n}$. Show that if $|AX| = |X|$ for every $X \in \mathbb{C}^{n\times 1}$, then A is unitary.

Exercise 7.18. Let $A \in \mathbb{C}^{n\times n}$ such that $\langle AX, X\rangle = 0$ for every $X \in \mathbb{C}^{n\times 1}$. Show that $A = 0$. Does the same conclusion hold if $A \in \mathbb{R}^{n\times n}$ and $\langle AX, X\rangle = 0$ for every $X \in \mathbb{R}^{n\times 1}$?

Exercise 7.19. Prove Exercise 7.17 using Exercise 7.18.

Exercise 7.20 (Review exercise)**.** Decide which of the following statements are true. In each case give a proof or a counterexample.

a. If $A, B \in \mathbb{C}^{n\times n}$ are Hermitian, then $A + B$ is Hermitian.

b. If $A, B \in \mathbb{C}^{n\times n}$ are Hermitian, then AB is Hermitian.

c. If $A, B \in \mathbb{C}^{n\times n}$ are Hermitian and $AB = BA$, then AB is Hermitian.

d. The matrix
$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \\ 0 & 0 & -\sin\phi & \cos\phi \end{pmatrix}$$
is unitary.

e. If $A, B \in \mathbb{C}^{n\times n}$ are unitary, then every eigenvalue of AB has modulus 1.

f. There is no unitary $A \in \mathbb{C}^{n\times n}$ such that $(A - 2I_n)(A - 3I_n)(A - 4I_n) = 0$.

CHAPTER 8

NORMAL MATRICES

8.1 Schur's Lemma

Lemma 8.1.1 (Schur**).** i. (**Complex version**) For every $A \in \mathbb{C}^{n\times n}$ there exists a unitary matrix $U_A \in \mathbb{C}^{n\times n}$ such that $U_A^{-1}AU_A$ is upper triangular.

ii. (**Real version**) For every triangularizable matrix $A \in \mathbb{R}^{n\times n}$ there exists an orthogonal matrix $U_A \in \mathbb{R}^{n\times n}$ such that $U_A^{-1}AU_A$ is upper triangular.

Proof. i. The proof is similar to the proof of Theorem 5.1.1, with the difference that we seek a unitary matrix U_A.

We proceed by induction on the size n of A.

- **Base case.** For $n = 1$ the result is obvious.
- **Inductive step.** Assume the statement holds for every $(n-1)\times(n-1)$ matrix.
  Let $A \in \mathbb{C}^{n\times n}$. Since A is triangularizable, the characteristic polynomial $\chi_A(x)$ splits into linear factors. Let u be an eigenvector of A with eigenvalue λ. Set

  $$v_1 = \frac{u}{|u|}.$$

  Then there exists an orthonormal basis $\{v_1, \dots, v_n\}$ of $\mathbb{C}^{n\times 1}$. Hence there exists $U_1 \in \mathbb{C}^{n\times n}$ whose columns are

  $$U_1^{(i)} = v_i, \qquad i = 1, \dots, n.$$

By Theorem 7.3.1, the matrix U_1 is unitary, and

$$U_1^{-1}AU_1 = \left(\begin{array}{c|c} \lambda & O \\ \hline O & B \end{array}\right), \qquad B \in \mathbb{C}^{(n-1)\times(n-1)}.$$

By the induction hypothesis, there exists a unitary $U_2 \in \mathbb{C}^{(n-1)\times(n-1)}$ such that $U_2^{-1}BU_2$ is upper triangular. Define

$$U_A = U_1 \left(\begin{array}{c|c} 1 & O \\ \hline O & U_2 \end{array}\right).$$

By Proposition 7.3.1, U_A is unitary, and $U_A^{-1}AU_A$ is upper triangular.

ii. The proof is exactly the same as in (i).

□

8.2 Spectral Theorem

Theorem 8.2.1 (Spectral Theorem)**.** i. (**Complex version**) For every Hermitian matrix $A \in \mathbb{C}^{n\times n}$ there exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU$ is diagonal.

ii. (**Real version**) For every symmetric matrix $A \in \mathbb{R}^{n\times n}$ there exists an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that $U^{-1}AU$ is diagonal.

Proof. i. Let $A \in \mathbb{C}^{n\times n}$ be Hermitian. By Schur's Lemma there exists a unitary $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU = T$, where T is upper triangular. Since $A = A^*$, we have

$$UTU^{-1} = (UTU^{-1})^* = (U^{-1})^*T^*U^* = UT^*U^{-1} \;\Rightarrow\; T = T^*.$$

Since T is upper triangular and equal to its conjugate transpose, it must be diagonal. Moreover, since $T = \overline{T}$, it is real (its diagonal entries are real).

ii. The proof is analogous to (i), noting that a real symmetric $A \in \mathbb{R}^{n\times n}$ is triangularizable over $\mathbb{R}$, since it is triangularizable over $\mathbb{C}$ and by Properties 7.3.1 all its eigenvalues are real.

□

Example 8.2.1. Consider the matrix

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Find an orthogonal matrix P such that $P^{-1}AP$ is diagonal.

Proof. In the usual way, we find that

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\} \text{ is a basis of } V_A(4), \qquad \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \text{ is a basis of } V_A(1).$$

By Properties 7.3.1, every vector in $V_A(4)$ is orthogonal to every vector in $V_A(1)$, since A is real symmetric. Applying Gram–Schmidt separately to each eigenspace, we obtain the following orthonormal bases:

- For $V_A(4)$:
$$\left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$
- For $V_A(1)$:
$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

Thus, setting
$$P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{pmatrix},$$
the matrix P is orthogonal and
$$P^{-1}AP = \text{diag}(4, 1, 1).$$

□

8.3 Normal Matrices

Question 8.3.1. For which matrices $A \in \mathbb{C}^{n\times n}$ does there exist a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU$ is diagonal? We saw in the previous subsection that Hermitian matrices have this property. Are there others?

Observation 8.3.1. Suppose $A \in \mathbb{C}^{n\times n}$ has the above property, i.e. there exists $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU = \Delta$ is diagonal. Then $A = U\Delta U^{-1}$. We compute:

$$AA^* = U\Delta U^{-1} \cdot (U\Delta U^{-1})^* = U\Delta U^{-1}(U^{-1})^*\Delta^* U^* = U\Delta U^{-1} U\Delta^* U^{-1} = U\Delta\Delta^* U^{-1}.$$

Similarly,

$$A^*A = U\Delta^*\Delta U^{-1}.$$

But Δ is diagonal, hence $\Delta\Delta^* = \Delta^*\Delta$. Therefore $AA^* = A^*A$. Later we will see that the converse also holds.

Lemma 8.3.1. If $T \in \mathbb{C}^{n\times n}$ is upper triangular and $TT^* = T^*T$, then T is diagonal.

Proof. We prove the claim by induction on n.

- **Base case.** For $n = 1$ the statement is trivial.
- **Inductive step.** Let

$$T = \left(\begin{array}{c|ccc} t_{11} & t_{12} & \cdots & t_{1n} \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right) \in \mathbb{C}^{n\times n}, \qquad T_1 \in \mathbb{C}^{(n-1)\times(n-1)}$$

with T_1 upper triangular. The condition $TT^* = T^*T$ is equivalent to

$$\left(\begin{array}{c|ccc} t_{11} & t_{12} & \cdots & t_{1n} \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right)\left(\begin{array}{c|ccc} \overline{t_{11}} & 0 & \cdots & 0 \\ \hline \overline{t_{12}} & & & \\ \vdots & & T_1^* & \\ \overline{t_{1n}} & & & \end{array}\right) = \left(\begin{array}{c|ccc} \overline{t_{11}} & 0 & \cdots & 0 \\ \hline \overline{t_{12}} & & & \\ \vdots & & T_1^* & \\ \overline{t_{1n}} & & & \end{array}\right)\left(\begin{array}{c|ccc} t_{11} & t_{12} & \cdots & t_{1n} \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right).$$

Comparing the $(1,1)$-entry yields

$$|t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2 = |t_{11}|^2,$$

hence $t_{12} = \cdots = t_{1n} = 0$. It follows that $T_1T_1^* = T_1^*T_1$, so by the induction hypothesis T_1 is diagonal. Since the first row off-diagonal entries are 0 and T_1 is diagonal, T is diagonal.

□

Theorem 8.3.1. Let $A \in \mathbb{C}^{n\times n}$. The following are equivalent:

i. There exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU$ is diagonal.

ii. $AA^* = A^*A$.

Proof.
- i. $\Rightarrow$ ii. This was shown in Notation 8.3.1.
- ii. $\Rightarrow$ i. Let $A \in \mathbb{C}^{n\times n}$ satisfy $AA^* = A^*A$. By Schur's Lemma there exists a unitary U such that $A = UTU^{-1}$ with T upper triangular. From $AA^* = A^*A$ we obtain $TT^* = T^*T$, and since T is upper triangular, Lemma 8.3.1 implies that T is diagonal.

□

Definition 8.3.1. A matrix $A \in \mathbb{C}^{n\times n}$ is called **normal** if $AA^* = A^*A$.

Example 8.3.1.

i. Every diagonal matrix is normal.

ii. Every Hermitian matrix is normal.

iii. Every unitary matrix is normal.

iv. The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ is not normal, since

$$AA^* = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix} = A^*A.$$

Theorem 8.3.2. Let $A \in \mathbb{C}^{n\times n}$. The following are equivalent:

i. A is normal.

ii. There exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU = \Delta$ is diagonal.

iii. There exists an orthonormal basis of $\mathbb{C}^{n\times 1}$ consisting of eigenvectors of A.

Proof.
- i. $\Leftrightarrow$ ii. This was proved in Theorem 8.3.1.
- ii. $\Rightarrow$ iii. Each column of U is an eigenvector of A. These columns form an orthonormal basis of $\mathbb{C}^{n\times 1}$ since U is unitary.
- iii. $\Rightarrow$ ii. The proof is left as an exercise to the reader.

□

Lemma 8.3.2. Let $B \in \mathbb{C}^{n\times n}$ satisfy $\langle BX, X\rangle = 0$ for every $X \in \mathbb{C}^{n\times 1}$. Then $B = 0$.

Proof. Let $X, Y \in \mathbb{C}^{n\times 1}$. Then

$$\langle B(X+Y), X+Y\rangle = 0$$

is equivalent to

$$\langle BX, X\rangle + \langle BX, Y\rangle + \langle BY, X\rangle + \langle BY, Y\rangle = 0.$$

Hence

$$\langle BX, Y\rangle + \langle BY, X\rangle = 0. \tag{8.1}$$

Replacing Y by iY gives

$$i\langle BX, Y\rangle + i\langle BY, X\rangle = 0. \tag{8.2}$$

Combining (8.1) and (8.2) yields $\langle BX, Y\rangle = 0$ for all X, Y. By Lemma 7.3.1(ii), it follows that $B = 0$. □

Warning! Lemma 8.3.2 does not hold in general if $B \in \mathbb{R}^{n\times n}$ and $\langle BX, X\rangle = 0$ for all $X \in \mathbb{R}^{n\times 1}$. For example, the 90° rotation matrix

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

satisfies the above property, but $B \neq 0$.

Lemma 8.3.3. Let $T \in \mathbb{C}^{n\times n}$ be upper triangular such that every eigenvector of T is also an eigenvector of T^*. Then T is diagonal.

Proof. We prove the claim by induction on n.

- For $n = 1$ the statement is trivial.
- **Inductive step.** Let T be upper triangular and assume every eigenvector of T is also an eigenvector of T^*. Write

$$T = \left(\begin{array}{c|ccc} t_{11} & t_{12} & \cdots & t_{1n} \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right), \qquad T^* = \left(\begin{array}{c|ccc} \overline{t_{11}} & 0 & \cdots & 0 \\ \hline \overline{t_{12}} & & & \\ \vdots & & T_1^* & \\ \overline{t_{1n}} & & & \end{array}\right).$$

  Since E_1 is an eigenvector of T, by hypothesis it is also an eigenvector of T^*. Therefore $(T^*)^{(1)} = \lambda E_1$, hence

$$\overline{t_{12}} = \cdots = \overline{t_{1n}} = 0.$$

Now let

$$X' = \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix}$$

be an eigenvector of T_1, and define

$$X = \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Then X is an eigenvector of T, hence also of T^*. It follows that X' is an eigenvector of T_1^*, and by the induction hypothesis T_1 is diagonal.

□

Theorem 8.3.3. Let $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$. The following are equivalent:

i. A is normal.

ii. There exists a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $U^{-1}AU$ is diagonal.

iii. There exists an orthonormal basis of $\mathbb{C}^{n\times 1}$ consisting of eigenvectors of A.

iv. $|AX| = |A^*X|$ for every $X \in \mathbb{C}^{n\times 1}$.

v. $V_A(\lambda) = V_{A^*}(\overline{\lambda})$ for every eigenvalue λ of A.

vi. Every eigenvector of A is also an eigenvector of A^*.

vii. $\sum_{i,j} |a_{ij}|^2 = \sum_i |\lambda_i|^2$.

Proof.

- The equivalences i., ii., iii. were proved in Theorem 8.3.2.
- i. $\Rightarrow$ iv. If A is normal, then $AA^* = A^*A$. For every $X \in \mathbb{C}^{n\times 1}$:

$$\begin{aligned} & \langle (A^*A - AA^*)X, X\rangle = 0 \\ \iff \quad & \langle AX, AX\rangle = \langle A^*X, A^*X\rangle \\ \iff \quad & |AX| = |A^*X|. \end{aligned}$$

- iv. $\Rightarrow$ i. If $|AX| = |A^*X|$ for every X, then
$$\langle (A^*A - AA^*)X, X\rangle = 0 \quad \text{for all } X.$$
By Lemma 8.3.2, we get $A^*A = AA^*$, hence A is normal.
- i. $\Rightarrow$ v. If A is normal, then $B = A - \lambda I$ is also normal. Using (iv), we have
$$|BX| = 0 \;\Leftrightarrow\; |B^*X| = 0 \;\Rightarrow\; V_A(\lambda) = V_{A^*}(\overline{\lambda}).$$
- v. $\Rightarrow$ i. By Schur's Lemma there exists a unitary U such that $T = U^{-1}AU$ is upper triangular. If UX is an eigenvector of A, then X is an eigenvector of T, and by hypothesis also of T^*. By Lemma 8.3.3, T is diagonal, hence A is normal.
- i. $\Rightarrow$ vi. From i. $\Rightarrow$ v., and the inclusion of eigenspaces, the implication v. $\Rightarrow$ vi. is immediate.
- vi. $\Rightarrow$ i. Trivial from vi. $\Rightarrow$ v. and v. $\Rightarrow$ i.
- i. $\Rightarrow$ vii. If A is normal, then
$$\text{Tr}(AA^*) = \sum_{i,j} |a_{ij}|^2.$$
Also, since A is unitarily diagonalizable, $U^{-1}AU = \Delta$, we get
$$\text{Tr}(AA^*) = \text{Tr}(\Delta\Delta^*) = \sum_i |\lambda_i|^2.$$
- vii. $\Rightarrow$ i. If $\sum_{i,j} |a_{ij}|^2 = \sum_i |\lambda_i|^2$ and $U^{-1}AU = T$ is upper triangular, then
$$\sum_{i,j} |a_{ij}|^2 = \sum_{i,j} |t_{ij}|^2 = \sum_i |\lambda_i|^2 + \sum_{i\neq j} |t_{ij}|^2 \Rightarrow \sum_{i\neq j} |t_{ij}|^2 = 0,$$
so T is diagonal and thus A is normal.

□

After completing the proof, let us make a couple of remarks about Theorem 8.3.3

i. Property (iv) says that if A is normal, then $|AX| = |A^*X|$. For $X = E_i$ we get $|A^{(i)}| = |A_i|$, i.e. the ith column of A has the same length as the ith row of A, for each $i = 1, \dots, n$.

ii. In general, for every $A \in \mathbb{C}^{n\times n}$, if
$$\chi_A(x) = (x - \lambda_1)\cdots(x - \lambda_n),$$
then
$$\chi_{A^*}(x) = (x - \overline{\lambda_1})\cdots(x - \overline{\lambda_n}).$$
However, when A is normal, we also have $V_A(\lambda) = V_{A^*}(\lambda)$ for every eigenvalue λ of A.

8.4 Exercises of Chapter 8.

Group A: 1, 2, 3, 4, 5, 6, 7, 15, 17, 19, 20, 21, 21, 23, 26
Group B: 8, 9, 10, 13, 14, 16, 18, 24, 25, 27
Group C: 11, 12

Exercise 8.1. Determine whether there exists a real orthogonal matrix P such that $P^{-1}AP$ is upper triangular, where

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}.$$

If such a P exists, find one.

Exercise 8.2. Let

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Find an orthogonal $P \in \mathbb{R}^{3\times 3}$ such that $P^{-1}AP$ is diagonal.

Exercise 8.3. Let

$$A = \begin{pmatrix} 4 & 3 & 0 \\ 3 & 12 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

with eigenvalues $1, 3, 13$.

a. Find an orthogonal $U \in \mathbb{R}^{3\times 3}$ such that $U^{-1}AU$ is diagonal.

b. Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be linear with $(f : \hat{a}, \hat{a}) = A$, where $\hat{a}$ is an ordered basis of $\mathbb{R}^3$. Show that

$$f^{40} - 5f^9 + 3f^6 + 1_{\mathbb{R}^3} \neq 0.$$

Exercise 8.4. Let $A \in \mathbb{C}^{n\times n}$, $H = \frac{1}{2}(A + A^*)$, and $S = \frac{1}{2}(A - A^*)$.

a. Show that H is Hermitian and that $S^* = -S$.

b. Show that if every eigenvector of H is also an eigenvector of S, then A is normal.

Exercise 8.5. Show that there exists an orthonormal basis of $\mathbb{C}^{2\times 1}$ consisting of eigenvectors of

$$\begin{pmatrix} 1 & i \\ a & 1 \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

if and only if $|a| = 1$.

Exercise 8.6. Show that if A is normal, then the ith row of A has the same length as the ith column of A, for every i.

Exercise 8.7. Find all normal matrices $A \in \mathbb{C}^{n\times n}$ such that $A^m = 0$ for some m.

Exercise 8.8. Let $A \in \mathbb{C}^{n\times n}$ be normal. Prove:

a. A is Hermitian $\Leftrightarrow$ every eigenvalue of A is real.

b. A is unitary $\Leftrightarrow$ every eigenvalue of A has modulus 1.

Exercise 8.9. a. If $A \in \mathbb{R}^{n\times n}$ is symmetric and $A^k = I_n$, then $A^2 = I_n$.

b. Find all symmetric $A \in \mathbb{R}^{n\times n}$ such that $A^{1821} = I_n$.

c. If $A \in \mathbb{C}^{n\times n}$ is Hermitian and unitary with $\mathrm{Tr}(A) = 0$, then n is even.

d. If $A \in \mathbb{C}^{n\times n}$ is Hermitian and unitary and has at least two distinct eigenvalues, find the minimal polynomial of A.

Exercise 8.10. a. For every $A \in \mathbb{C}^{n\times n}$, the matrix $A + A^* - iI_n$ is invertible.

b. If $A, B \in \mathbb{R}^{n\times n}$ are symmetric and $AB = BA$, then $AB + iI_n$ is invertible.

c. Let $A \in \mathbb{C}^{n\times n}$.

   i. Every eigenvalue of AA^* is real and nonnegative.

   ii. $\det(AA^* + I_n)$ is a real and positive number.

Exercise 8.11. If $A \in \mathbb{C}^{n\times n}$ is normal, then $A^* = f(A)$ for some $f(x) \in \mathbb{C}[x]$.

Exercise 8.12. Let $A \in \mathbb{C}^{n\times n}$ and $B = AA^* - A^*A$. Show that if $AB = BA$, then A is normal.

Exercise 8.13. Find a symmetric $A \in \mathbb{R}^{3\times 3}$ with eigenvalues $1, 1, -1$ such that the eigenspace $V_A(1)$ is spanned by

$$\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad \begin{pmatrix}2\\2\\1\end{pmatrix}.$$

Is A unique?

Exercise 8.14. a. Let $A \in \mathbb{R}^{4\times 4}$ with $\dim V_A(2) = \dim V_A(3) = 2$ and $\langle u, v\rangle = 0$ for every $u \in V_A(2)$ and $v \in V_A(3)$. Show that A is symmetric.

b. Let $A \in \mathbb{R}^{n\times n}$ satisfy $AA^t = A^tA$ and assume $\chi_A(x)$ splits into linear factors in $\mathbb{R}[x]$. Show that A is symmetric.

Exercise 8.15. Let $A \in \mathbb{R}^{n\times n}$ be symmetric and not of the form cI_n, $c \in \mathbb{R}$. Find $m_A(x)$ if

$$(A - 2I_n)^3(A - 3I_n)^4 = 0.$$

Exercise 8.16. If $A \in \mathbb{C}^{n\times n}$ is normal and λ_1, λ_2 are distinct eigenvalues of A, then

$$V_A(\lambda_1) = V_A(\lambda_2)^\perp.$$

Exercise 8.17. Let $A, B \in \mathbb{C}^{4\times 4}$ be normal matrices with

$$\chi_A(x) = (x-1)^2(x-2)^2, \qquad \chi_B(x) = (x-3)^2(x-4)^2.$$

If $V_A(1) = V_B(3)$, show that $AB = BA$.

Exercise 8.18. Let $A \in \mathbb{C}^{n\times n}$. Show that A is Hermitian if and only if $\langle AX, X\rangle \in \mathbb{R}$ for every $X \in \mathbb{C}^{n\times 1}$.

Exercise 8.19. Give an example of $A \in \mathbb{C}^{3\times 3}$ such that there exists a basis of $\mathbb{C}^{3\times 1}$ consisting of eigenvectors of A, but there is no orthonormal basis of $\mathbb{C}^{3\times 1}$ consisting of eigenvectors of A.

Exercise 8.20. Let $B \in \mathbb{C}^{n\times n}$ satisfy

$$(B - \tfrac{1}{2}I_n)^3(B - iI_n)^4 = 0.$$

a. Show that if B is Hermitian, then $B = \frac{1}{2}I_n$.

b. Show that if B is unitary, then $B = iI_n$.

Exercise 8.21. Let $u \in \mathbb{R}^{n\times 1}$ with $|u| = 1$, and set $S = I_n - uu^t \in \mathbb{R}^{n\times n}$.

a. Show that there exists an orthonormal basis of $\mathbb{R}^{n\times 1}$ consisting of eigenvectors of S.

b. Show that $Su = 0$ and $Sv = v$ for every $v \in \mathbb{R}^{n\times 1}$ such that $\langle v, u\rangle = 0$. Then find the dimension of each eigenspace of S.

c. Give a geometric interpretation of S for $n = 2, 3$.

Exercise 8.22. Let $a \in \mathbb{C}$ and

$$A = \begin{pmatrix} 0 & 1 & a \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

a. Is it true that for every a there exists a unitary $U \in \mathbb{C}^{3\times 3}$ such that $U^{-1}AU$ is upper triangular?

b. Is it true that for $a = 1$ there exists a unitary $Q \in \mathbb{C}^{3\times 3}$ such that $Q^{-1}AQ$ is upper triangular?

c. Find all values of a such that there exists an orthonormal basis of $\mathbb{C}^{3\times 1}$ consisting of eigenvectors of A.

d. Let $a = 0$. Find a unitary $U \in \mathbb{C}^{3\times 3}$ such that $U^{-1}AU = \text{diag}(1, -1, 1)$.

Exercise 8.23. Let

$$A = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 4 & 1 \\ 0 & -2 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

a. Find a basis for each eigenspace of A and determine whether A is diagonalizable.

b. Find two linearly independent eigenvectors of $B = A^{12} - 8A^7 + 5A^5 + 4I_3$.

c. Determine whether there exists an ordered basis $\hat{a} = (a_1, a_2, a_3)$ of $\mathbb{R}^3$ such that $(f : \hat{a}, \hat{a}) = A$, where $f : \mathbb{R}^3 \to \mathbb{R}^3$ is the linear map defined by

$$f(a_1) = 3a_1 - 6a_2, \quad f(a_2) = 3a_1 - 8a_2 + 6a_3, \quad f(a_3) = 5a_3.$$

d. Find (if it exists) an invertible $P \in \mathbb{R}^{3\times 3}$ such that $P^{-1}AP$ is upper triangular.

e. Find (if it exists) an orthogonal $U \in \mathbb{R}^{3\times 3}$ such that $U^{-1}AU$ is upper triangular.

Exercise 8.24. Let

$$A = \begin{pmatrix} 0 & 0 & -3 \\ -1 & 3 & 1 \\ 1 & 0 & 4 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

a. Determine whether there exists an orthogonal $P \in \mathbb{R}^{3\times 3}$ such that $P^{-1}AP$ is upper triangular.

b. Let $B = A^{1821} - A^3 + I_3$. Find an invertible $Q \in \mathbb{R}^{3\times 3}$ such that $Q^{-1}AQ$ is upper triangular.

c. If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is linear with $(f : \hat{a}, \hat{a}) = A$, determine whether $f^3 - 3f - 18 \cdot 1_{\mathbb{R}^3}$ is an isomorphism.

Exercise 8.25. Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Consider the linear map

$$L_A : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}, \qquad L_A(X) = AX.$$

Show that

$$\ker(L_A) = (\mathrm{Im}(L_A))^{\perp}.$$

Exercise 8.26. Let $A \in \mathbb{C}^{3\times 3}$ satisfy $A^*A = 4A$.

a. Show that A is Hermitian.

b. Determine whether there exists an orthonormal basis of $\mathbb{C}^{3\times 1}$ consisting of eigenvectors of A.

c. Show that if $\mathrm{rank}(A) = 1$, then there exists a unitary $U \in \mathbb{C}^{3\times 3}$ such that $U^{-1}AU = \mathrm{diag}(4, 0, 0)$.

d. Determine whether $\langle AX, AY\rangle = \langle X, Y\rangle$ for all $X, Y \in \mathbb{C}^{3\times 1}$.

Exercise 8.27. If $T \in \mathbb{C}^{n\times n}$ is upper triangular and every eigenvector of T is also an eigenvector of T^t, then T is diagonal.

Exercise 8.28 (Review exercise)**.** Decide which of the following statements are true. In each case give a proof or a counterexample. Let $A \in \mathbb{C}^{n \times n}$ be Hermitian.

a. If A is unitary and every eigenvalue of A is positive, then $A = I_n$.

b. $\varphi(A)$ is diagonalizable for every $\varphi(x) \in \mathbb{C}[x]$.

c. If $A^m = 0$ for some m, then $A = 0$.

d. If every eigenvalue of A is nonnegative, then there exists a Hermitian $B \in \mathbb{C}^{n \times n}$ with $B^2 = A$.