Eigenproblems are widely used in linear algebra and routinely encountered in various applications in engineering and computational sciences including quantum Chemistry/Physics (the discretization of highly non-linear governing equations leads to challenging large-scale eigenproblems), Markov Chains based modeling as studied in assignment 2 (upon which the famous “PageRank” algorithm used by Google is based), frequency analysis in structural dynamics, or dimensionality reduction (such as the Principal Component Analysis), to name a few.

Contents

1 Introduction 3
  1.1 A motivating example .......................................................... 3
  1.2 Definition of the eigenproblem .............................................. 3
  1.3 About the characteristic polynomial ....................................... 4

2 About Hermitian matrices 6
  2.1 Hermitian inner product ....................................................... 6
  2.2 Main results ........................................................................... 7

3 Diagonalization 7
  3.1 Diagonalization theorem .......................................................... 7
  3.2 Examples of diagonalization ..................................................... 8
    3.2.1 Example 1 ........................................................................ 8
    3.2.2 Example 2 ........................................................................ 9
    3.2.3 Example 3 ........................................................................ 10
  3.3 Determinants and traces ............................................................ 10

4 Some applications of the diagonalization 12
  4.1 Computing powers of matrices .................................................. 12
  4.2 Linear recurrence sequences .................................................... 12
  4.3 Resolution of systems of ODEs .................................................. 13
    4.3.1 General case .................................................................... 13
    4.3.2 Example .......................................................................... 14

5 Diagonalizability and invertibility of a matrix 15
  5.1 Conditions of diagonalizability ............................................... 15
  5.2 Invertibility of matrices ........................................................... 16
  5.3 Diagonalizability and invertibility ............................................. 16

6 Eigenanalysis for linear transformations 16
  6.1 Main results ............................................................................ 16
  6.2 Examples of linear transformations ........................................... 19
6.3 Geometric interpretation of diagonalization ........................................ 19

7 Multiplicities of eigenvalues .................................................................. 20
1 Introduction

1.1 A motivating example

Consider the following system of ODEs\(^1\)
\[
\begin{aligned}
\dot{x}(t) &= Ax(t), \quad t > 0 \\
x(0) &= x_0,
\end{aligned}
\]
(1)
where \( A \in \mathbb{R}^n \) and \( x = (x_1, \ldots, x_n)^T \). Taking the analytical expression \( x(t) = x_0 e^{at} \) in the one-dimensional case as a guide, consider a solution of the form \( x(t) = p e^{\lambda t} \). Then
\[
\dot{x}(t) = \lambda p e^{\lambda t} = Ax(t) = Ap e^{\lambda t}
\]
which yields
\[
Ap = \lambda p
\]
(2)
since \( e^{\lambda t} > 0 \), where \( p \in \mathbb{R}^n \) and \( \lambda \in \mathbb{R} \) have to be determined. Note that \( p \) has to be a nonzero vector, otherwise the initial conditions are not necessarily satisfied. Equations of the form (2) are referred to as an \textit{eigenproblem} and will be studied in depth in this chapter. In particular we will show that the solution of (1) is unique and defined as
\[
x(t) = \sum_{i=1}^{n} c_i p_i e^{\lambda_i t},
\]
(3)
where \( \{p_i, \lambda_i\} \) are solutions of (2) and \( c_i \) are coefficients to be determined.

1.2 Definition of the eigenproblem

Let \( A \in \mathbb{C}^n \). An eigenproblem is stated as

Find \( \lambda \in \mathbb{C} \) and \( p \in \mathbb{C}^n \) nonzero such that
\[
Ap = \lambda p.
\]

\textbf{Definition 1.} \( p \neq 0 \) is an eigenvector of \( A \) and \( \lambda \) the corresponding eigenvalue. The spectrum of \( A \) is the set of all the eigenvalues of \( A \) and is classically denoted by \( \sigma(A) \subset \mathbb{C} \).

Since we are looking for a nonzero vector \( p \) such that \( (A - \lambda I)p = 0 \), it implies that \( A - \lambda I \) is singular \( \iff \) det \( n(A - \lambda I) = 0 \).

\textbf{Definition 2.} The characteristic polynomial (or eigenpolynomial) is defined as
\[
p_A(\lambda) := \text{det}_n(A - \lambda I).
\]

A real matrix can have complex eigenvalues. For example, consider \( A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \). Then
\[
\text{det}_2(A - \lambda I) = \left| \begin{array}{cc}
1 - \lambda & 1 \\
-1 & 1 - \lambda
\end{array} \right| = 0 \iff \lambda = 1 \pm i \in \mathbb{C}.
\]
The eigenvectors associated with the eigenvalues \( \lambda = 1 + i \) and \( \lambda = 1 - i \) are \((1, i)^T\) and \((1, -i)^T\). The eigenvectors associated with the eigenvalues \( \lambda = 1 + i \) and \( \lambda = 1 - i \) are \((1, i)^T\) and \((1, -i)^T\), respectively.

\textbf{Definition 3.} The eigenspace corresponding to a given an eigenvalue \( \lambda \in \sigma(A) \) is defined as
\[
E_\lambda := \left\{ p \in \mathbb{C}^n \text{ such that } Ap = \lambda p \right\} = \text{null}(A - \lambda I)
\]

\(^1\)For example, think about chemical reactions where \( x \) represents concentrations of species and \( A \) depends on rate constants.
$E_\lambda$ consists of all the (nonzero) eigenvectors of $A$ associated with $\lambda$, in addition to the zero vector.

**Proposition 1.** Let $A \in \mathbb{C}^{n \times n}$ and $\lambda, \mu \in \sigma(A)$ with $\lambda \neq \mu$. Then

$$E_\lambda \cap E_\mu = \{0\}.$$  

**Proof.** The zero vector clearly belongs to $E_\lambda \cap E_\mu$. For the converse inclusion, let $p \in E_\lambda \cap E_\mu$.

$$Ap = \lambda p = \mu p \Rightarrow (\lambda - \mu)p = 0$$

Since $\lambda \neq \mu$ it implies that $p = 0$. \hfill $\square$

Note that for any $\lambda \in \sigma(A)$, $E_\lambda$ is invariant under $A$, namely

$$AE_\lambda \subset E_\lambda.$$  

Indeed, let $p \in E_\lambda$ and consider $q = Ap \in AE_\lambda$. Then $Aq = A(Ap) = A(\lambda p) = \lambda Ap = \lambda q$, that is, $q \in E_\lambda$.

### 1.3 About the characteristic polynomial

**Proposition 2.** Let $A \in \mathbb{C}^{n \times n}$ with $n \geq 2$. Its characteristic polynomial writes

$$p_A(\lambda) = (-1)^n (\lambda^n - tr(A)\lambda^{n-1} + \cdots + (-1)^n det_n(A)),$$

where the trace of $A$ is defined as $tr(A) = \sum_{i=1}^n a_{ii}$.

**Proof.** For any $\lambda \in \mathbb{C}$ we have

$$p_A(\lambda) = \det_n(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - \lambda \\ \end{vmatrix} = (-1)^n (\lambda^n + c_1 \lambda^{n-1} + c_2 \lambda^{n-2} + \cdots + c_{n-1} \lambda + c_n).$$  

(6)

The coefficient $(-1)^n$ in factor comes from the fact that the only source of the $\lambda^n$ term is given by the product of the terms in the diagonal. We immediately deduce

$$p_A(0) = \det_n(A) = (-1)^n c_n \Rightarrow c_n = (-1)^n \det_n(A).$$

Let us show now that $c_1 = -tr(A)$. Using Laplace formula successively, it can be shown that $p_A$ can also be expanded as

$$p_A(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) + \frac{c_2}{2} \lambda^{n-2} + c_3 \lambda^{n-3} + \cdots + c_n.$$  

(7)

The term $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$ is the product of the diagonal terms. The remaining terms $c_2 \lambda^{n-2} + \cdots + c_n$ correspond to cofactors $(-1)^{i+j} \det(M_{ij})$ where the $i$-th row and $j$-th column are left out with $i \neq j$. But when $i \neq j$ one factor in $\lambda$ is left out in the row $i$ and in the column $j$, meaning that the maximal power we can get in $\lambda$ is $n-2$. In addition, it can be shown (by induction) that

$$(a_{11} - \lambda) \cdots (a_{nn} - \lambda) = (-1)^n \lambda^n + (-\lambda)^{n-1} (a_{11} + a_{22} + \cdots + a_{nn}) + \cdots$$

Comparing terms in $\lambda^{n-1}$ in (6) and (7), we obtain

$$(-1)^n c_1 = (-1)^{n-1} tr(A) \Rightarrow c_1 = (-1)^{2n-1} tr(A) = -tr(A)$$

which concludes the proof. Note this result could also be proved by induction on $n$. \hfill $\square$

**Proposition 3.** $A$ has at most $n$ distinct eigenvalues in $\mathbb{C}$.
Proof. Since \( p_A \) is a polynomial of degree \( n \) it admits at most \( n \) distinct complex roots.

**Reminder (complex numbers):** For any complex number \( z = a + ib \in \mathbb{C} \), where \( a, b \in \mathbb{R} \) and \( i^2 = -1 \), the complex conjugate of \( z \) is defined as \( \overline{z} = a - ib \). The modulus of \( z \) is given by \( |z| = \sqrt{a^2 + b^2} \) and is such that \( |z|^2 = z \overline{z} \). A complex number can also be represented by its polar form \( z = r(\cos(\theta) + i\sin(\theta)) = re^{i\theta} \). Using this representation the modulus of \( z \) is given by \( |z| = r \) and its complex conjugate writes \( \overline{z} = r(\cos(\theta) - i\sin(\theta)) = re^{-i\theta} \).

**Proposition 4.** Let \( A^* = A^T \) the transpose conjugate of \( A \in n\mathbb{C}^n \). The eigenvalues of \( A^* \) are the complex conjugate of the eigenvalues of \( A \).

Proof. We first need to prove that \( \det_n(A^*) = \overline{\det_n(A)} \). For this we will use the same type of arguments for proving \( \det_n(A^T) = \det_n(A) \) (see chapter 9). From \( \tilde{A} = EA \) we have \( \tilde{A} = A^*E^* \) and then

\[
\det_n(\tilde{A}) = \det_n(E)\det_n(A), \quad (8)
\]

\[
\det_n(\tilde{A}^*) = \det_n(E^*)\det_n(A^*). \quad (9)
\]

Since \( \tilde{A} \) is upper triangular with diagonal entries \( \tilde{a}_{ii} \) and \( \tilde{A}^* \) is lower triangular with diagonal entries \( \overline{a}_{ii} \), we deduce that

\[
\det_n(\tilde{A}^*) = \prod_{i=1}^n \overline{a}_{ii} = \overline{\det_n(A)}. \quad (10)
\]

Next, it can be shown that

\[
\det_n(E^*) = \overline{\det_n(E)} \quad (11)
\]

for any elementary matrix \( E \). Indeed, from

\[
E^* = \begin{cases} 
E(i,j) & \text{type I} \\
E(\overline{i};i) & \text{type II} \\
E(\overline{i};j,i) & \text{type III}
\end{cases}
\]

we get

\[
\det_n(E^*) = \begin{cases} 
-1 = \det_n(E) = \overline{\det_n(E)} & \text{type I} \\
\overline{\lambda} = \overline{\det_n(E)} & \text{type II} \\
1 = \det_n(E) = \overline{\det_n(E)} & \text{type III}
\end{cases}
\]

From (8), (9) and (10) it follows

\[
\det_n(\tilde{A}^*) = \det_n(E^*)\det_n(A^*) = \overline{\det_n(\tilde{A})} = \overline{\det_n(E)\det_n(A)}. \]

We finally deduce that \( \det_n(A^*) = \overline{\det_n(A)} \) using (11). Thus it holds

\[
p_{A^*}(\overline{\lambda}) = \det_n(A^* - \overline{\lambda}1) = \det_n((A - \lambda1)^*) = \overline{\det_n(A - \lambda1)} = \overline{p_A(\lambda)} \quad (12)
\]

meaning that the roots of \( p_{A^*} \) are \( \overline{\lambda} \) with \( \lambda \in \sigma(A) \) since

\[
\lambda \in \sigma(A) \iff p_A(\lambda) = 0 \iff p_{A^*}(\overline{\lambda}) = 0 \iff \overline{\lambda} \in \sigma(A^*). \]

We conclude this section with the following results:

- Let \( A \in n\mathbb{C}^n \). Then the transposed matrix \( A^T \) has the same eigenvalues as \( A \) since they have the same characteristic polynomial:

\[
p_{A^T}(\lambda) = \det_n((A^T - \lambda1)) = \det_n((A - \lambda1)) = p_A(\lambda).
\]
Let \( A \in \mathbb{R}^n \) be a real matrix. If \( A \) has complex eigenvalues they occur in complex conjugate pairs. Indeed, since the characteristic polynomial of \( A \) has real coefficients its holds
\[
p_A(\lambda) = 0 \Rightarrow p_A(\lambda) = (-1)^n \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n \lambda + c_n = p_A(\lambda) = 0,
\]
namely \( \lambda \) and \( \overline{\lambda} \) are eigenvalues of \( A \). This property does not hold for matrices which have some entries in \( \mathbb{C} \setminus \mathbb{R} \). For example, the eigenvalues of \( A = \begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix} \) are 1 and \( i \), but \( -i \) is not an eigenvalue of \( A \).

## 2 About Hermitian matrices

**Definition 4.** \( A \in \mathbb{C}^n \) is an Hermitian matrix if \( A^* = A \). If \( A \) is a real matrix we retrieve the definition of a symmetric matrix, \( A^T = A \).

### 2.1 Hermitian inner product

Let \( V \) be a vector space over \( \mathbb{C} \). An application \( \langle \cdot, \cdot \rangle : V \times V \to \mathbb{C} \) is an Hermitian inner product if
\[
\begin{align*}
(i) \quad & \langle u, u \rangle \geq 0, \quad \forall u \in V \quad \text{(nonnegative)} \\
(ii) \quad & \langle u, u \rangle = 0 \quad \text{if} \quad u = 0 \quad \text{(positive definite)} \\
(iii) \quad & \langle u + v + w, u + v + w \rangle = \langle u, u \rangle + \langle v, v \rangle + \langle w, w \rangle, \quad \forall u, v, w \in V \\
(iv) \quad & \langle u, \alpha v \rangle = \alpha \langle u, v \rangle, \quad \forall u, v \in V, \quad \forall \alpha \in \mathbb{C} \\
(v) \quad & \langle v, u \rangle = \overline{\langle u, v \rangle}, \quad \forall u, v \in V \quad \text{(Hermitian symmetric)}
\end{align*}
\]

Properties (iii) and (iv) means the application \( \langle \cdot, \cdot \rangle \) is linear in the second argument. Note that properties (iii), (iv) and (v) imply the conjugate linearity in the first argument:
\[
\begin{align*}
(vi) \quad & \langle \alpha u, v \rangle = \overline{\langle v, u \rangle} = \overline{\alpha \langle u, v \rangle} = \overline{\alpha} \langle u, v \rangle \\
(vii) \quad & \langle v + w, u \rangle = \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle = \langle v, u \rangle + \langle w, u \rangle \\
\end{align*}
\]

Now consider \( V = \mathbb{C}^n \). The application \( \langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \) defined as
\[
\langle u, v \rangle := u^T v = \sum_{i=1}^n u_i v_i \tag{13}
\]
is an Hermitian inner product (i.e. satisfies properties (i)–(vii)).\(^2\) In addition it holds
\[
\begin{align*}
\langle A^* u, v \rangle &= \overline{\langle A v, u \rangle} = (\overline{A^T} u)^T v = \overline{\pi^T A^T} v = \overline{\pi^T} A v = \langle u, A v \rangle, \tag{14} \\
\langle A u, v \rangle &= \langle (A^*)^* u, v \rangle = \langle A^* v \rangle.
\end{align*}
\]

Two vectors \( u, v \in \mathbb{C}^n \) are said to be orthogonal if
\[
\langle u, v \rangle = 0.
\]

\(^2\)The dot product \( \langle u, v \rangle = u^T v \) is an inner product over \( \mathbb{R}^n \) but not over \( \mathbb{C}^n \). Indeed, \( \langle u, u \rangle = 0 \not\Rightarrow u = 0 \) for \( u \in \mathbb{C}^n \). It is necessary to consider the Hermitian inner product (13) since for \( u \in \mathbb{C}^n \), \( \langle u, u \rangle = \sum_{i=1}^n |u_i|^2 = 0 \not\Rightarrow u_i = 0 \quad \forall i \not\Rightarrow u = 0.\)
2.2 Main results

**Proposition 5.** Let \( A \in \mathbb{C}^n \) be an Hermitian matrix. Then its eigenvalues are real.

**Proof.** Let \( \lambda \in \sigma(A) \). There exists \( u \neq 0 \) such that \( Au = \lambda u \). Then

\[
\lambda \langle u, u \rangle = \langle u, \lambda u \rangle = \langle u, Au \rangle = \langle A^* u, u \rangle \quad \text{from (iv)}
\]

\[
= \langle Au, u \rangle = \langle \lambda u, u \rangle = \overline{\lambda} \langle u, u \rangle \quad \text{from (vi)}
\]

Since \( u \neq 0 \) then \( \langle u, u \rangle > 0 \) from (ii). This implies \( \lambda = \overline{\lambda} \), that is, \( \lambda \in \mathbb{R} \).

**Proposition 6.** Let \( A \in \mathbb{C}^n \) be an Hermitian matrix. Then eigenvectors corresponding to distinct eigenvalues are orthogonal.

**Proof.** Let \( \lambda, \mu \in \sigma(A) \) with \( \lambda \neq \mu \). From Proposition 5 we know that \( \lambda, \mu \in \mathbb{R} \). We have \( Av = \lambda v \) and \( Aw = \mu w \), with \( v \neq 0 \), \( w \neq 0 \). Hence we have

\[
\mu \langle v, w \rangle = \langle v, \mu w \rangle = \langle v, Aw \rangle = \langle A^* v, w \rangle \quad \text{from (iv)}
\]

\[
= \langle Av, w \rangle = \langle \lambda v, w \rangle = \overline{\lambda} \langle v, w \rangle \quad \text{from (15)}
\]

since \( \lambda \in \mathbb{R} \). We deduce that

\[
(\lambda - \mu) \langle v, w \rangle = 0 \Rightarrow \langle v, w \rangle = 0
\]

since \( \lambda \neq \mu \).

3 Diagonalization

3.1 Diagonalization theorem

Let \( A \in \mathbb{C}^n \). Since the characteristic polynomial of \( A \) is of degree \( n \) it admits \( n \) roots so \( A \) has \( n \) eigenvalues (at most \( n \) different eigenvalues from Proposition 3). Consider the eigenvalues \( \lambda_i \in \mathbb{C} \) and corresponding eigenvectors \( p_i \in \mathbb{C}^n \)

\[
Ap_i = \lambda_i p_i, \quad i = 1, \ldots, n.
\]

The preceding equations can be recast in a matrix form as

\[
AP = P\Lambda
\]

with

\[
P = [p_1 | p_2 | \ldots | p_n] \in \mathbb{C}^n
\]

and

\[
\Lambda = \begin{pmatrix}
\lambda_1 & & \\
& \lambda_2 & \\
& & \ddots & \\
& & & \lambda_n
\end{pmatrix}
\]

is a diagonal matrix containing the corresponding eigenvalues.

**Definition 5.** The matrix \( A \in \mathbb{C}^n \) is diagonalizable if the matrix \( P \in \mathbb{C}^n \) is invertible. In this case it holds

\[
P^{-1}AP = \Lambda,
\]

where \( \Lambda \) is the diagonal matrix containing the eigenvalues of \( A \). We also say that \( P \) diagonalizes \( A \), or that \( A \) is similar to the diagonal matrix \( \Lambda \).
Theorem 1 (Diagonalization theorem). The matrix $A \in \mathbb{C}^n$ is diagonalizable $\iff$ there is a basis for the $\mathbb{C}$-vector space $\mathbb{C}^n$ consisting of eigenvectors of $A$.

Proof. Let’s prove each implication.

$\Rightarrow$ Assume that $A$ is diagonalizable. Hence

$P$ is invertible $\iff$ $\{p_1, \ldots, p_n\} \subset \mathbb{C}^n$ are linearly independent $\iff$ $\{p_1, \ldots, p_n\}$ is a basis for $\mathbb{C}^n$ which is a vector space of dimension $n$ over $\mathbb{C}$

$\Leftarrow$ Assume that $\{p_1, \ldots, p_n\}$ form a basis for $\mathbb{C}^n$. Then the eigenvectors are linearly independent $\iff$ $P$ is invertible from full rank Theorem. If $P$ is invertible then $A$ is diagonalizable.

A more general form of this diagonalization theorem will be given for linear transformations, see Theorem 7.

Corollary 1. Let $A \in \mathbb{R}^n$ be a real-valued matrix. Then $A$ is diagonalizable using a real-valued diagonalization matrix if and only if there is a basis for the $\mathbb{R}$-vector space $\mathbb{R}^n$ consisting of eigenvectors of $A$.

3.2 Examples of diagonalization

3.2.1 Example 1

Let us determine if

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

is diagonalizable. We first find the eigenvalue of $A$ by solving

$$0 = \det_3(A - \lambda I) = \det_3 \begin{pmatrix} 1 - \lambda & 1 & 0 \\ 0 & 2 - \lambda & 2 \\ 0 & 0 & 3 - \lambda \end{pmatrix} = (1 - \lambda)(2 - \lambda)(3 - \lambda)$$

since $A - \lambda I$ is an upper triangular matrix. The eigenvalues of $A$ are $\lambda_1 = 1$, $\lambda_2 = 2$ and $\lambda_3 = 3$. Let’s now determine the corresponding eigenspaces $E_{\lambda_i} = \text{null}(A - \lambda_i I)$. For each eigenspace we will use the RREF of $A - \lambda_i I$.

- $A - \lambda_1 I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow \widetilde{A - \lambda_1 I} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Then

$$\text{rank}(A - \lambda_1 I) = 2 \Rightarrow \dim(E_{\lambda_1}) = \dim(\text{null}(A - \lambda_1 I)) = 1$$

using the dimension formula. Next, solving

$$(\widetilde{A - \lambda_1 I})p = 0$$

yields $p = (x_1, 0, 0)^T$, that is, $E_{\lambda_1} = \text{null}(A - \lambda_1 I) \subset \text{span}\{(1, 0, 0)^T\}$. Since $\dim(E_{\lambda_1}) = 1$ it follows that $E_{\lambda_1} = \text{span}\{p_1\}$, $p_1 = (1, 0, 0)^T$.

- $A - \lambda_2 I = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \widetilde{A - \lambda_2 I} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Then

$$\dim(E_{\lambda_2}) = \dim(\text{null}(A - \lambda_2 I)) = 3 - 2 = 1.$$
Solving \((A - \lambda I)p = 0\) yields \(p = (x_1, x_1, 0)^T\), that is, \(E_{\lambda_1} = \text{null}(A - \lambda_1 I) \subset \text{span}\{(1, 1, 0)^T\}\). Since \(\dim(E_{\lambda_1}) = 1\) it follows that \(E_{\lambda_1} = \text{span}\{p_2\}, p_2 = (1, 1, 0)^T\).

- \(A - \lambda_3 I = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow A - \lambda_3 I = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}\). Then

\[
\dim(E_{\lambda_3}) = \dim(\text{null}(A - \lambda_3 I)) = 3 - 2 = 1.
\]

Solving \((A - \lambda_3 I)p = 0\) yields \(p = (x_3, 2x_3, x_3)^T\), that is, \(E_{\lambda_3} = \text{null}(A - \lambda_3 I) \subset \text{span}\{(1, 2, 1)^T\}\). Since \(\dim(E_{\lambda_3}) = 1\) it follows that \(E_{\lambda_3} = \text{span}\{p_3\}, p_3 = (1, 2, 1)^T\).

Consider \(P = [p_1, p_2, p_3] = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}\). It holds

\[
\det(P) = 1 \Leftrightarrow P \text{ is invertible} \Leftrightarrow \{p_1, p_2, p_3\} \text{ are linearly independent} \Leftrightarrow \{p_1, p_2, p_3\} \subset \mathbb{R}^3 \text{ is a basis for the } \mathbb{R}-\text{vector space } \mathbb{R}^3 \Leftrightarrow A = P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},
\]

namely \(A\) is diagonalizable. Note that \(\{p_1, p_2, p_3\}\) is also a basis for the \(\mathbb{C}\)-vector space \(\mathbb{C}^3\).

### 3.2.2 Example 2

Let us determine if

\[
A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}
\]

is diagonalizable. Solving

\[
0 = \det_2(A - \lambda I) = (1 - \lambda)^2 + 1 = 0
\]

leads to \(\lambda_1 = 1 + i\) and \(\lambda_2 = 1 - i\), where \(i^2 = -1\) (note that \(\lambda_1\) and \(\lambda_2\) are complex conjugates, see end of section 1.3).

- \(A - \lambda_1 I = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \Rightarrow A - \lambda_1 I = \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}\). Then

\[
\dim(E_{\lambda_1}) = \dim(\text{null}(A - \lambda_1 I)) = 2 - 1 = 1.
\]

Solving \((A - \lambda_1 I)p = 0\) yields \(p = (-ix_2, x_2)^T\), that is, \(E_{\lambda_1} = \text{null}(A - \lambda_1 I) \subset \text{span}\{(1, i)^T\}\). Since \(\dim(E_{\lambda_1}) = 1\) it follows that \(E_{\lambda_1} = \text{span}\{p_1\}, p_1 = (1, i)^T\).

- \(A - \lambda_2 I = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \Rightarrow A - \lambda_2 I = \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}\). Then

\[
\dim(E_{\lambda_2}) = \dim(\text{null}(A - \lambda_2 I)) = 2 - 1 = 1.
\]

Solving \((A - \lambda_2 I)p = 0\) yields \(p = (ix_2, x_2)^T\), that is, \(E_{\lambda_2} = \text{null}(A - \lambda_2 I) \subset \text{span}\{(1, -i)^T\}\). Since \(\dim(E_{\lambda_2}) = 1\) it follows that \(E_{\lambda_2} = \text{span}\{p_2\}, p_2 = (1, -i)^T\).
Consider \( P = [p_1, p_2] = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \). It holds
\[
\det(P) = -2i \neq 0 \iff P \text{ is invertible} \iff \{p_1, p_2\} \text{ are linearly independent} \iff \{p_1, p_2\} \subset \mathbb{C}^2 \text{ is a basis for the } \mathbb{C}-\text{vector space } \mathbb{C}^2 \\
\iff \Lambda = P^{-1}AP = \begin{pmatrix} 1 + i & 0 \\ 0 & 1 - i \end{pmatrix},
\]

namely \( A \) is diagonalizable.

### 3.2.3 Example 3

Let us determine if \( A = \begin{pmatrix} 8 & 1 \\ -9 & 2 \end{pmatrix} \) is diagonalizable. Solving \( 0 = \det_2(A - \lambda I) = \lambda^2 - 10\lambda + 25 = 0 \) leads to \( \lambda_1 = \lambda_2 = 5 \).

- \( A - \lambda_1 I = \begin{pmatrix} 3 & 1 \\ -9 & -3 \end{pmatrix} \Rightarrow \widetilde{A - \lambda_1 I} = \begin{pmatrix} 1 & 1/3 \\ 0 & 0 \end{pmatrix} \). Then
  \( \dim(E_{\lambda_1}) = \dim(\text{null}(A - \lambda_1 I)) = 2 - 1 = 1 \).

Solving \( (A - \widetilde{\lambda_1 I})p = 0 \)

yields \( p = (-x_2/3, x_2)^T \), that is, \( E_{\lambda_1} = \text{null}(A - \lambda_1 I) \subset \text{span}\{-1, 3\} \). Since \( \dim(E_{\lambda_1}) = 1 \) it follows that \( E_{\lambda_1} = \text{span}\{p_1\}, p_1 = (-1, 3)^T \).

- Since \( \lambda_2 = \lambda_1 \) we get the same eigenspace \( E_{\lambda_2} = \text{span}\{p_2\}, p_2 = (-1, 3)^T \).

Consider \( P = [p_1, p_2] = \begin{pmatrix} -1 & -1 \\ 3 & 3 \end{pmatrix} \). Here
\[
\det(P) = 0 \iff P \text{ is not invertible} \iff \{p_1, p_2\} \text{ are linearly dependent} \iff A \text{ is not diagonalizable}
\]

### 3.3 Determinants and traces

We provide here some expressions for determinants and traces of matrices in terms of their eigenvalues. We first start by proving that similar matrices have the same eigenvalues.

**Proposition 7.** Let \( A \in \mathbb{C}^{n \times n} \) and let \( T \in \mathbb{C}^{n \times n} \) be invertible. Then \( A \) and \( B = T^{-1}AT \) (\( A \) and \( B \) are said to be similar matrices) have the same characteristic polynomial, meaning they have the same eigenvalues.

**Proof.** We shall use the fact that
\[
\det_n(CD) = \det_n(DC) \tag{17}
\]
for any square matrices \( C, D \). The characteristic polynomial of \( B \) can be written as
\[
p_B(\lambda) = \det_n(T^{-1}AT - \lambda I) \\
= \det_n(T^{-1}(A - \lambda I)T) \\
= \det_n(T(T^{-1}(A - \lambda I))) \\
= \det_n(A - \lambda I) \\
= p_A(\lambda).
\]

\( \square \)
Theorem 2. Let \( A = P \Lambda P^{-1} \in \mathbb{C}^n \) be a diagonalizable matrix, where \( P \) containing eigenvectors of \( A \) is invertible, and \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) is a diagonal matrix containing the eigenvalues of \( A \). Then

(a) \( p_A(\lambda) = p_A(\lambda) = \prod_{i=1}^n (\lambda_i - \lambda) \)
(b) \( \det_n(A) = \det_n(\Lambda) = \prod_{i=1}^n \lambda_i \)
(c) \( \text{tr}(A) = \text{tr}(\Lambda) = \sum_{i=1}^n \lambda_i \)

Proof. (a) Applying Proposition 7 yields

\[
p_A(\lambda) = p_{P^{-1}AP}(\lambda) = p_A(\lambda)
\]

and then

\[
p_A(\lambda) = \det_n(\Lambda - \lambda I) = \begin{vmatrix}
\lambda_1 - \lambda & & \\
& \ddots & \\
& & \lambda_n - \lambda
\end{vmatrix} = \prod_{i=1}^n (\lambda_i - \lambda).
\]

(b) We have

\[
\det_n(A) = \det_n(PP^{-1}A) = \det_n(P^{-1}AP) = \det_n(\Lambda) = \prod_{i=1}^n \lambda_i
\]

(c) Here we shall use the following property for traces\(^3\)

\[
\text{tr}(CD) = \text{tr}(DC)
\]

We have

\[
\text{tr}(A) = \text{tr}(PP^{-1}A) = \text{tr}(P^{-1}AP) = \sum_{i=1}^n \lambda_i.
\]

Remark 1. It is worth mentioning that the formulas \( \det_n(A) = \prod_{i=1}^n \lambda_i \) and \( \text{tr}(A) = \sum_{i=1}^n \lambda_i \) are valid even if \( A \) is not diagonalizable. Indeed, when writing the characteristic polynomial as

\[
p_A(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)
\]

it can be seen that the constant term is \( \prod_{i=1}^n \lambda_i \) and the coefficient in \( \lambda^{n-1} \) is \( (-1)^{n+1} \sum_{i=1}^n \lambda_i \). The result follows by comparing these expressions with the terms given in Proposition 2 (see (5)).

\(^3\)For \( C = [c_{ij}] \) and \( D = [d_{ij}] \) it holds \( \text{tr}(CD) = \sum_{i=1}^n (CD)_{ii} = \sum_{i=1}^n \sum_{j=1}^n c_{ij}d_{ji} = \sum_{j=1}^n \sum_{i=1}^n d_{ji}c_{ij} = \sum_{j=1}^n (DC)_{jj} = \text{tr}(DC) \).
4 Some applications of the diagonalization

4.1 Computing powers of matrices

If \( A \in \mathbb{C}^{n \times n} \) is a diagonalizable matrix then powers of \( A \) can be computed conveniently as

\[
A^k = P \Lambda^k P^{-1}
\]

where \( \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \vdots \\ 0 & & \lambda_n \end{pmatrix} \)

Example. Compute \( A = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix}^{10000} \). Since \( A \) is diagonalizable (see section 6.3) it holds

\[
A^{10000} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2^{10000} + (1/2)^{10000} \\ 2^{10000} - (1/2)^{10000} \end{pmatrix} = \begin{pmatrix} 2^{10000} + (1/2)^{10000} \\ 2^{10000} - (1/2)^{10000} \end{pmatrix}
\]

4.2 Linear recurrence sequences

Let \( \mathbb{R}^\infty \) be the vector space of all infinite sequences \((a_1, a_2, a_3, \ldots)\) of real numbers. Consider the Fibonacci linear recurrence sequence given by

\[
a_{n+1} = a_n + a_{n-1}, \quad n \geq 2
\]

with \( a_1 = 0, a_2 = 1 \). In order to determine the \( n \)-th term of the Fibonacci sequence we shall recast the recurrence relationship (19) in a matrix form

\[
\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}
\]

which can be expressed in terms of the two first terms as

\[
\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = A \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A^2 \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} = \ldots = A^{n-1} \begin{pmatrix} a_2 \\ a_1 \end{pmatrix}.
\]

Let us diagonalize the matrix \( A \). We have

\[
\det_2(A - \lambda I) = 0 \iff \lambda^2 - \lambda - 1 = 0 \\
\iff \lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{or} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}
\]

Note that \( \lambda_1 \) is known as the golden ratio (\( \lambda_1 = 1 + 1/\lambda_1 \)). Next

\[
A - \lambda_1 I = \begin{pmatrix} 1 - \lambda_1 & 1 \\ 1 & -\lambda_1 \end{pmatrix} \Rightarrow A - \lambda_1 I = \begin{pmatrix} 1 & -\lambda_1 \\ 0 & 0 \end{pmatrix}
\]

meaning that

\[
\dim(E_{\lambda_1}) = \dim(\text{null}(A - \lambda_1 I)) = 2 - 1 = 1.
\]

Solving

\[
(A - \lambda_1 I)p = 0
\]
yields \( p = (\lambda_i x_2, x_2)^T \), that is, \( E_{\lambda_i} = \text{null}(A - \lambda_i I) \subset \text{span}\{(\lambda_i, 1)^T\} \). Since \( \dim(E_{\lambda_i}) = 1 \) it follows that \( E_{\lambda_i} = \text{span}\{p_i\} \), \( p_i = (\lambda_i, 1)^T \). Considering \( P = [p_1, p_2] = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \) it can be seen that 

\[
\det(P) = \lambda_1 - \lambda_2 = \sqrt{5} \neq 0 \iff P \text{ is invertible} \iff \{p_1, p_2\} \subset \mathbb{R}^2 \text{ is a basis for the } \mathbb{R}\text{-vector space } \mathbb{R}^2 
\]

\[
P^{-1}A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.
\]

Substituting \( A = P\Lambda P^{-1} \) into (20) it holds

\[
\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} a_2 \\ a_1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^{n-1} & 0 \\ 0 & \lambda_2^{n-1} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & \lambda_1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\]

from which we finally deduce the expression of the \( n \)-th term

\[
a_n = \frac{1}{\sqrt{5}} (\lambda_1^{n-1} - \lambda_2^{n-1}).
\]

### 4.3 Resolution of systems of ODEs

#### 4.3.1 General case

Consider the following system of ODEs

\[
\begin{align*}
\dot{x}(t) &= Ax(t), \quad t > 0 \\
x(0) &= x_0 
\end{align*}
\]

with \( A = [a_{ij}] \in \mathbb{C}^n \) is a matrix independent of time, \( x(t) = (x_1(t), \ldots, x_n(t))^T \in \mathbb{C}^n \) and \( x_0 \in \mathbb{C}^n \).

It is worth mentioning that (21) is a system of coupled ODEs since each component \( x_i \) depends on all the other ones. If \( A \) is diagonalizable then \( \dot{x}(t) = P\Lambda P^{-1}x(t) \), where \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) and \( P \) contains eigenvectors \( p_i \) of \( A \). Multiplying by \( P^{-1} \) on both sides we obtain

\[
P^{-1}\dot{x}(t) = \Lambda P^{-1}x(t),
\]

which motivates the change of variables \( y(t) = P^{-1}x(t) \) to finally get the following auxiliary system of ODEs

\[
\dot{y}(t) = \Lambda y(t).
\]

Since \( \Lambda \) is a diagonal matrix, (22) is a system of decoupled ODEs which writes component-wisely as

\[
\dot{y}_i(t) = \lambda_i y_i(t).
\]

Hence, each component of \( y \) is given by \( y_i(t) = y_i(0)e^{\lambda_i t} \) which in matrix form writes

\[
y(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ & \ddots & 0 \\ & 0 & \ddots \end{pmatrix} y(0)
\]

with initial conditions \( y(0) = P^{-1}x(0) = P^{-1}x_0 \). Coming back to the original variables, the exact solution to (21) can be compactly written as

\[
x(t) = P \begin{pmatrix} e^{\lambda_1 t} & 0 \\ & \ddots & 0 \\ & 0 & \ddots \end{pmatrix} P^{-1}x_0.
\]
The solution can also be expanded as

\[ x(t) = Py(t) = \sum_{i=1}^{n} p_i y_i(t) = \sum_{i=1}^{n} p_i y_i(0)e^{\lambda_i t}, \]  

(24)

where the coefficients \( y_i(0) \) are solution of the linear system \( Py(0) = x_0 \). Since \( P \) is invertible the solution is unique, meaning that \( x(t) \) is uniquely defined in (24).

Furthermore, expressing the solution under the form (24) reveals what a basis could be for the set of solutions to \( \dot{x} = Ax \), namely \( DA \). We know that \( DA = \text{span}\{p_1 e^{\lambda_1 t}, \ldots, p_n e^{\lambda_n t}\} \). Let us prove these functions are linearly independent. In this aim, consider the linear combination

\[ \sum_{i=1}^{n} c_i p_i e^{\lambda_i t} = 0 \in \mathbb{C}^n, \forall t \geq 0. \]

In particular, for \( t = 0 \) it holds \( \sum_{i=1}^{n} c_i p_i = 0 \). Since \( P \) is invertible, \( \{p_1, \ldots, p_n\} \) are linearly independent \( \Rightarrow c_i = 0, i = 1, \ldots, n \). Hence \( \{p_1 e^{\lambda_1 t}, \ldots, p_n e^{\lambda_n t}\} \) are linearly independent and then form a basis for \( DA \) which is of dimension \( n \).

### 4.3.2 Example

Consider the ODE system

\[ \begin{cases} \dot{x}_1 = \frac{5}{4} x_1 - \frac{3}{4} x_2 \\ \dot{x}_2 = -\frac{3}{4} x_1 + \frac{5}{4} x_2 \end{cases} \]  

(25)

with initial conditions

\[ x_1(0) = x_0 = 1, \quad x_2(0) = y_0 = 2. \]  

(26)

This ODE system can be recast in the matrix form (21) with

\[ A = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix}, \]

\[ x(t) = (x_1(t), x_2(t))^T \text{ and } x_0 = (1, 2)^T. \]

The eigenvalues of \( A \) are \( \lambda_1 = 2, \lambda_2 = 1/2 \) with corresponding eigenvectors \( p_1 = (1, -1)^T \) and \( p_2 = (1, 1)^T \). Consider \( P = [p_1, p_2] = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \). It holds

\[ \det(P) = 2 \Leftrightarrow \{p_1, p_2\} \text{ are linearly independent} \]

\[ \Leftrightarrow \{p_1, p_2\} \subset \mathbb{R}^2 \text{ is a basis for the } \mathbb{R}-\text{vector space } \mathbb{R}^2 \]

meaning that \( A \) is diagonalizable. We can then apply the derivation from section 4.3.1 (see (23)) to express the solution of the ODE system (25-26) as

\[ x(t) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} e^{2t} \\ 0 \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}e^{2t}(x_0 - y_0) + \frac{1}{2}e^{t/2}(x_0 + y_0) \\ -\frac{1}{2}e^{2t}(x_0 - y_0) + \frac{1}{2}e^{t/2}(x_0 + y_0) \end{pmatrix} \]

that is

\[ \begin{cases} x_1(t) = -\frac{1}{2}e^{2t} + \frac{3}{2}e^{t/2} \\ x_2(t) = \frac{1}{2}e^{2t} + \frac{3}{2}e^{t/2} \end{cases} \]
5 Diagonalizability and invertibility of a matrix

5.1 Conditions of diagonalizability

Theorem 3. Let \( A \in \mathbb{C}^n \) with distinct eigenvalues \( \lambda_1, \ldots, \lambda_m \in \mathbb{C} \) with \( m \leq n \). Then the corresponding (nonzero) eigenvectors \( \{p_1, \ldots, p_m\} \) are linearly independent.

Proof. This result directly comes from Theorem 5 where the linear transformation is given by

\[
T : \mathbb{C}^n \rightarrow \mathbb{C}^n \\
x \mapsto Ax
\]

Corollary 2. Let \( A \in \mathbb{C}^n \). If \( A \) has \( n \) distinct eigenvalues then \( A \) is diagonalizable.

Proof.

\[ A \text{ has } n \text{ distinct eigenvalues } \implies \{p_1, \ldots, p_n\} \text{ are linearly independent} \quad \text{Theorem 3} \]
\[ \iff \{p_1, \ldots, p_n\} \text{ is a basis for } \mathbb{C}^n \]
\[ \iff A \text{ is diagonalizable} \quad \text{Theorem 1} \]

Remark 2. The converse is wrong. Take for example the identity matrix \( 1 \) which is diagonalizable (1 is already diagonal) but only has the same eigenvalue 1 repeated \( n \) times.

We now provide a result of diagonalizability which holds for a class of commonly used matrices.

Theorem 4. Let \( A \in \mathbb{C}^n \) be an Hermitian matrix. Then \( A \) is diagonalizable and there exists an orthonormal basis for \( \mathbb{C}^n \) which consists of eigenvectors of \( A \) (or, equivalently, the matrix \( P \) which diagonalizes \( A \) is orthogonal, namely \( P^*P = I \)).

Proof. Theorem 4 holds for any Hermitian matrix even with repeated eigenvalues. However, for simplicity of exposition we will assume the eigenvalues of \( A \) are all distinct. Let \( p_i \) denote eigenvectors of \( A \) and consider \( P = [p_1 | \ldots | p_n] \in \mathbb{C}^n \). From Proposition 6 we know that

\[ \langle p_i, p_j \rangle = P^T p_j = 0, \quad i \neq j, \]

namely the \( p_i \)'s are mutually orthogonal. Let’s prove they are linearly independent. Consider a linear combination

\[ \alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_n p_n = 0, \quad \alpha_i \in \mathbb{C}. \]

For any \( j = 1, \ldots, n \), it holds

\[ 0 = \langle \alpha_1 p_1 + \alpha_2 p_2 + \ldots + \alpha_n p_n, p_j \rangle = \alpha_j \langle p_j, p_j \rangle \]

by orthogonality. Since \( p_j \) is a nonzero vector (this is an eigenvector) it implies that \( \alpha_j = 0 \) for any \( j = 1, \ldots, n \). Hence

\[ \{p_1, \ldots, p_n\} \text{ is linearly independent } \iff \{p_1, \ldots, p_n\} \text{ is an (orthogonal) basis for } \mathbb{C}^n \]
\[ \iff A \text{ is diagonalizable} \]

In addition, if each eigenvector is rescaled by a factor \( \frac{1}{\langle p_i, p_i \rangle^{1/2}} \) then eigenvectors are orthonormal

\[ \langle p_i, p_j \rangle = \delta_{ij} \]
which in matrix form writes as
\[ P^*P = 1, \]  
meaning that \( P \) is an orthogonal matrix.\(^4\)

The preceding result is referred to as an *orthogonal diagonalization* since \( A \) is diagonalized by an orthogonal matrix \( P \) (see (27)). In such a case the diagonal matrix containing the eigenvalues writes \( \Lambda = P^{-1}AP = P^*AP \).

### 5.2 Invertibility of matrices

**Proposition 8.** Let \( A \in \mathbb{C}^n \). \( A \) is invertible if its eigenvalues \( \lambda_i \) are all nonzero.

**Proof.** We know that \( A \) is invertible if \( \det_n(A) \neq 0 \) (see chapter 9). From Theorem 2 we have \( \det_n(A) = \prod_{i=1}^n \lambda_i \), meaning that \( A \) is invertible if \( \lambda_i \neq 0, i = 1, \ldots, n \).

### 5.3 Diagonalizability and invertibility

**Warning:** There is no direct link between diagonalizability and invertibility for matrices.

- (a) \( A \) is diagonalizable \( \not\Rightarrow \) \( A \) is invertible. Consider any diagonalizable matrix which has a zero eigenvalue, for example \( A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \).

- (b) \( A \) is invertible \( \not\Rightarrow \) \( A \) is diagonalizable. For example \( A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) is invertible (\( \det_2(A) = 1 \)) but only has one eigenvalue \( \lambda = 1 \) with eigenvector \( p = (1, 0)^T \). Since \( \dim(\text{null}(A - \lambda I)) = 1 \) there is no basis for \( \mathbb{R}^2 \) made of eigenvectors of \( A \), so \( A \) is not diagonalizable.

### 6 Eigenanalysis for linear transformations

In this section we consider linear transformations \( T \in \mathcal{L}(V, V) \), where \( V \) is a given vector space over \( \mathbb{F} \). Then \( \lambda \in \mathbb{F} \) is an eigenvalue of \( T \) if there exists a nonzero vector \( p \in V \) such that
\[ Tp = \lambda p. \]

The vector \( p \) is an eigenvector of \( T \) associated to the eigenvalue \( \lambda \).

#### 6.1 Main results

**Theorem 5.** Let \( T \in \mathcal{L}(V, V) \). Let \( \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{F} \) be distinct eigenvalues of \( T \) with corresponding (nonzero) eigenvectors \( p_1, p_2, \ldots, p_m \in V \). Then \( \{p_1, p_2, \ldots, p_m\} \) are linearly independent.

**Proof.** This result is proved by induction.

- **Base case.** Since \( p_1 \) is a nonzero vector, \( \{p_1\} \) is linearly independent.

- **Induction step.** Assume that \( \{p_1, \ldots, p_{m-1}\} \) are linearly independent, and check this is true for \( \{p_1, \ldots, p_m\} \). Consider a linear combination
\[ a_1p_1 + a_2p_2 + \ldots + a_mp_m = 0, \quad a_i \in \mathbb{F}. \]  
Applying \( T \) to the preceding equation and using the linearity of \( T \) it follows
\[ a_1\lambda_1p_1 + a_2\lambda_2p_2 + \ldots + a_m\lambda_mp_m = 0. \]  
\( ^4\)Other version of the proof. Consider the \( p_i \)'s which are mutually orthogonal. After rescaling each eigenvector we get \( P^*P = 1 \). Hence \( 1 = \det_n(P^*P) = \det_n(P^*)\det_n(P) = \det_n(P)\det_n(P) = \det_n(P)^2 \) which means that \( P \) is invertible \( \Leftrightarrow A \) is diagonalizable.
Substracting $\lambda_m$ times (28) to the preceding equation yields

$$a_1(\lambda_1 - \lambda_m)p_1 + a_2(\lambda_2 - \lambda_m)p_2 + \ldots + a_m(\lambda_m - \lambda_m)p_m = 0.$$

By induction assumption $\{p_1, \ldots, p_m\}$ are linearly independent, which implies $a_1 = a_2 = \cdots = a_m = 0$ since $\lambda_i - \lambda_m \neq 0$ for $i = 1, 2, \ldots, m$. The only term remaining term in (28) is

$$\lambda_m p_m = 0$$

and since $p_m$ is a nonzero vector it follows that $\lambda_m = 0$. Hence $\{p_1, \ldots, p_m\}$ are linearly independent, which proves the result for any $m$.

\begin{corollary}
Let $T \in \mathcal{L}(V, V)$ with $\dim(V) = n$. Then $T$ has at most $n$ distinct eigenvalues.
\end{corollary}

\begin{proof}
Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of $T$ with corresponding eigenvectors $\{p_1, \ldots, p_m\}$. From Theorem 5, $\{p_1, \ldots, p_m\}$ are linearly independent. Since $\dim(V) = n$ it follows that $m \leq n$ applying the Fundamental Theorem.
\end{proof}

\begin{remark}
Proposition 3 can be seen as a particular case of Corollary 3 with

$$T : \mathbb{C}^n \rightarrow \mathbb{C}^n,$$

$$x \mapsto Ax.$$

We now provide a result according to which any linear transformation admits at least one eigenvalue.

\begin{theorem}
Let $V \neq \{0\}$ be a vector space over $\mathbb{C}$ such that $\dim(V) = n$. Then any $T \in \mathcal{L}(V, V)$ admits at least one eigenvalue.
\end{theorem}

\begin{proof}
Let $v \in V$, $v \neq 0$. Consider the set of vectors

$$\{v, Tv, T^2v, \ldots, T^nv\}.$$

Since $\dim(V) = n$, this set of $n + 1$ vectors is linearly independent applying the Fundamental Theorem of linear algebra. There exists a collection of scalars $a_0, a_1, \ldots, a_n \in \mathbb{C}$ not all zero such that

$$a_0v + a_1Tv + \cdots + a_nT^nv = 0.$$

Let $m \leq n$ be the largest integer such that $a_m \neq 0$. Necessarily $m \geq 1$ (otherwise $a_0v = 0$ would imply $v = 0$, which is in contradiction with the original assumption). Consider now the polynomial

$$p(z) = a_0 + a_1z + a_2z^2 + \cdots + a_mz^m.$$

Since any polynomial with complex coefficients admits at least one root (Fundamental Theorem of algebra), $p(z) = (z - \lambda_1)\tilde{p}_1(z)$, where $\tilde{p}_1$ is of degree $m - 1$. Repeating the process we end up with the factorization

$$p(z) = c(z - \lambda_1)(z - \lambda_2)\ldots(z - \lambda_m).$$

It then holds

$$0 = a_0v + a_1Tv + \cdots + a_mT^nv$$

$$= (a_0T^0 + a_1T + a_2T^2 + \cdots + a_mT^m)(v)$$

$$= p(T)(v)$$

$$= c(T - \lambda_1Id)(T - \lambda_2Id)\ldots(T - \lambda_mId)(v)$$

meaning that there exists at least one $\lambda_i \in \mathbb{C}$ such that $T - \lambda_iId$ is noninjective, that is $\lambda_i$ is an eigenvalue of $T$.
\end{proof}
Theorem 7 (Diagonalization theorem II). Let \( T \in \mathcal{L}(V, V) \) with \( \dim(V) = n \). Let \( \{p_1, \ldots, p_n\} \) denote the eigenvectors of \( T \). The representation matrix \( [T]^F_F \in \mathbb{F}^{n \times n} \) is diagonal \( \Leftrightarrow \mathcal{F} = \{p_1, \ldots, p_n\} \) is a basis for \( V \).

Proof. Consider two bases \( E \) and \( F \) for the vector space \( V \). From chapter 8 we know that the representation matrices of \( T \) with respect to \( E \) and \( F \) satisfy the relationship

\[
[T]_F^F = P_{F \to E}^{-1} [T]_E^E P_{F \to E},
\]

where \( P_{E \to F} \) is the change of basis matrix from \( E \) to \( F \). Now, consider the eigenvectors \( p_i \) of \( T \) with corresponding eigenvalues \( \lambda_i \), that is, \( Tp_i = \lambda_i p_i \). Then

\( F = \{p_1, \ldots, p_n\} \) is a basis for \( V \) \( \Leftrightarrow \) \( [T]_F^F = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \)

by definition of \( [T]_F^F \).

In other words, taking a basis made of the eigenvectors of \( T \) diagonalizes its representation matrix, and the diagonalization matrix is the change of basis matrix from the eigenvectors basis to any other basis of \( V \). We will see later how this can be interpreted from a geometrical point of view (see examples in section 6.3).

Links between diagonalization theorems 1 and 7. Consider the linear transformation given by

\( T : \mathbb{C}^n \to \mathbb{C}^n \)

\( x \mapsto Ax \)

and take two bases for \( V = \mathbb{C}^n \), the standard basis \( E = \{e_1, \ldots, e_n\} \) and \( F = \{p_1, \ldots, p_n\} \) which consists in the eigenvectors of \( A \) (with associated eigenvalues \( \lambda_i \)). Then

\( [T]_E^E = A \)

(30)

since \( Te_i = Ae_i = i\)-th column of \( A \), and

\( [T]_F^F = \text{diag}(\lambda_1, \ldots, \lambda_n) = \Lambda \)

(31)

since \(Tp_i = Ap_i = \lambda_i p_i \). Also, the change of basis matrix \( P_{F \to E} \) contains the coordinates of the eigenvectors \( p_i \) expressed in the standard basis, meaning that

\( P_{F \to E} = [p_1 | \ldots | p_n] = P \)

(32)

Hence, substituting (30), (31) and (32) into (29), the statement of Theorem 7 becomes

\( \Lambda = P^{-1}AP \) is diagonal \( \Leftrightarrow \{p_1, \ldots, p_n\} \) is a basis for \( \mathbb{C}^n \)

which is nothing but the statement of Theorem 1.

We finally conclude this section by providing conditions under which the eigenvectors of \( T \) form a basis for the vector space \( V \) (i.e., diagonalizability conditions).

Theorem 8. Let \( T \in \mathcal{L}(V, V) \) with \( \dim(V) = n \). If \( T \) has \( n \) distinct eigenvalues then its eigenvectors \( \{p_1, \ldots, p_n\} \) form a basis for \( V \).

Proof. If \( T \) has \( n \) distinct eigenvalues then \( \{p_1, \ldots, p_n\} \) are linearly independent from Theorem 5. Since \( \dim(V) = n \) then \( \{p_1, \ldots, p_n\} \) is a basis for \( V \).
6.2 Examples of linear transformations

We consider here several examples of linear transformations in the plane

\[ T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \]

\[ x \mapsto Ax \]

Let \( v \neq 0 \) and \( w \perp v \). Consider the basis \( B = \{v, w\} \) for \( \mathbb{R}^2 \) and define the line \( L = \text{span}\{v\} \).

- **Orthogonal projection onto \( L \).**
  \[ T(v) = v \text{ and } T(w) = 0 \Rightarrow A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \]
  Here \( T \) has two eigenvectors (\( v \) and \( w \)) with two distinct eigenvalues (1 and 0).

- **Reflection through \( L \).**
  \[ T(v) = v \text{ and } T(w) = -w \Rightarrow A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \]
  Here \( T \) has two eigenvectors (\( v \) and \( w \)) with two distinct eigenvalues (1 and -1).

- **Shear in the direction of \( L \).**
  \[ T(v) = v \text{ and } T(w) = kv + w \Rightarrow A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \]
  Here \( T \) only has one eigenvector (\( v \)) associated with \( \lambda_1 = 1 \). Since \( \dim(E_{\lambda_1}) = 1 \) the matrix \( A \) (or \( T \)) is not diagonalizable.

- **Rotation through an angle \( \theta \in [0, \pi] \).**
  \[ T(v) = \cos(\theta)v + \sin(\theta)w \text{ and } T(w) = -\sin(\theta)v + \cos(\theta)w \Rightarrow A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \]
  The only values of \( \theta \) for which the eigenvalues and eigenvectors are real are \( \theta = 0 \) and \( \theta = \pi \). If \( \theta = 0 \) then \( A = 1 \) has two real eigenvalues \( \lambda_1 = \lambda_2 = 1 \) with eigenvectors \( \{(1, 0)^T, (0, 1)^T\} \), while when \( \theta = \pi \) then \( A = -1 \) has two real eigenvalues \( \lambda_1 = \lambda_2 = -1 \) with the same eigenvectors. Otherwise when \( \theta \in (0, \pi) \) then \( A \) has two complex eigenvalues \( e^{k\alpha} \) with complex eigenvectors \( \{(i, 1)^T, (-i, 1)^T\} \).

6.3 Geometric interpretation of diagonalization

We come back here to the example discussed at the end of chapter 8, where \( T \) is defined by

\[ T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \]

\[ x \mapsto Ax \]

with \( A = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix} \). We first look for the eigenvalues of \( A \) by solving

\[ \det_2(A - \lambda 1) = 0 = (5/4 - \lambda)^2 - (3/4)^2 \Leftrightarrow \lambda_1 = 2 \text{ or } \lambda_2 = 1/2. \]

- For \( \lambda_1 = 2 \) we have \( A - \lambda_1 1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \Rightarrow \dim(E_{\lambda_1}) = 1. \text{ Then} \]
  \[ (A - \lambda_1 1)p = 0 \Rightarrow x + y = 0 \Rightarrow E_{\lambda_1} = \text{span}\{p_1\}, \text{ } p_1 = (1, -1)^T. \]

\[^5\text{These examples could be left as exercises.}\]
• For $\lambda_2 = 1/2$ we have $\widetilde{A} - \lambda_2 I = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$, so $\dim(E_{\lambda_2}) = 1$. Then

$$(A - \lambda_2 I)p = 0 \Rightarrow x - y = 0 \Rightarrow E_{\lambda_2} = \text{span}\{p_2\}, \; p_2 = (1, 1)^T.$$ 

Since $P = [p_1 | p_2] = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ is invertible ($\det_2(P) = 2$) the matrix $A$ is diagonalizable. Consider the two bases $E = \{e_1, e_2\}$ and $F = \{p_1, p_2\}$. Following the discussion after Theorem 7 in section 6.1), we have

$$[T]_E^F = A, \; [T]_F^E = \Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}, \; P_{E \rightarrow F} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \; P_{F \rightarrow E} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

and

$$A = P_{F \rightarrow E} A P_{E \rightarrow F} \quad (34)$$

We now focus on the geometric interpretation of (34). In this aim we first represent $y = Ax$ using the known expression for $A$ and the standard basis $E$ (see Figure 2). The diagonalization (34) allows to better understand how the linear transformation acts on vectors in the plane (see Figure 3 for an example of visualization):

(i) Change of coordinates from $E$ to $\{p_1, p_2\}$ through $(x', y')^T = P_{E \rightarrow F}(x, y)^T$

(ii) The new coordinates are stretched along the directions $\{p_1, p_2\}$ through $\Lambda(x', y')^T = (2x', y'/2)^T$

(iii) The vector which is obtained is then converted back to the old coordinates through $P_{F \rightarrow E}(2x', y'/2)^T$

7 Multiplicities of eigenvalues
Figure 2: Image (in blue) of the black square through the linear transformation (33).

Figure 3: Visualization of $A(e_1)$ using (34). Each coordinate in the basis $\{p_1, p_2\}$ is stretched by a factor $\lambda_1, \lambda_2$. 