Ioan Radu PETER

Szilárd Csaba LÁSZLÓ

Adrian VIOREL

ELEMENTS OF LINEAR ALGEBRA

U.T. PRESS

Cluj-Napoca, 2014

ISBN 978-973-662-935-8

Ioan Radu PETER
Szilárd Csaba LÁSZLÓ
Adrian VIOREL

ELEMENTS OF LINEAR ALGEBRA

U.T. PRESS
Cluj-Napoca, 2014

Editura U.T.PRESS
Str.Observatorului nr. 34
C.P.42, O.P. 2, 400775 Cluj-Napoca
Tel.:0264-401.999 / Fax: 0264 - 430.408
e-mail: utpress@biblio.utcluj.ro
www.utcluj.ro/editura

Director: Prof.dr.ing. Daniela Manea
Consilier editorial: Ing. Călin D. Câmpean

Multiplicarea executată la editura U.T.PRESS.

ISBN 978-973-662-935-8
Bun de tipar: 17.01.2014
Tiraj: 100 exemplare

Contents

Preface

The aim of this textbook is to give an introduction to Linear Algebra, and at the same time to provide insight into concepts that are useful in various applications.

Since the intended reader is considered to be at undergraduate level, most possibly with little or no experience with abstract algebra, we try to build an approach that is self contained and straightforward. We achieve this by including simple yet detailed proofs of almost all results. In addition, fully solved problems and examples accompany the presentation of new concepts and results along with a section containing proposed problems at the end of each chapter.

The structure as such is based on seven chapters, starting with the recollection of the nuts and bolts of matrices (Chapter 1) before entering the core of the book (Chapters 2, 3 and 4) which covers: Vector paces, Linear maps between vector spaces and Eigenvalue problems. Two further chapters deal with the case of vector spaces that are endowed with more geometric structure, namely we present Inner product spaces (Chapter 5) and Operators on inner product spaces (Chapter 6). The final chapter is a briefing to the analytic geometry of quadratic curves and surfaces.

Last but not least, the authors gratefully acknowledge the support of Prof. Ioan Raşa and Prof. Dorian Popa who have carefully read the manuscript at different stages suggesting valuable improvements.

1

Matrices

1.1 Basic definitions, operations and properties.

Definition 1.1. *A matrix of dimension $m \times n$ with elements in a field $\mathbb{F}$, (where usually $\mathbb{F} = \mathbb{R}$, or $\mathbb{F} = \mathbb{C}$), is a function $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to \mathbb{F}$,*

$$A(i, j) = a_{ij} \in \mathbb{F},\ \forall i \in \{1, 2, \ldots, m\},\ j \in \{1, 2,\ \ldots, n\}.$$

Usually an $m \times n$ matrix is represented as a table with m lines and n columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}.$$

Hence, the elements of a matrix A are denoted by a_{ij}, where a_{ij} stands for the number that appears in the i^{th} row and the j^{th} column of A (this is called the (i, j) entry of A) and the matrix is represented as $A = (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$.

We will denote the set of all $m \times n$ matrices with entries in $\mathbb{F}$ by $\mathcal{M}_{m,n}(\mathbb{F})$ respectively, when $m = n$ by $\mathcal{M}_n(\mathbb{F})$. It is worth mentioning that the elements of

$\mathcal{M}_n(\mathbb{F})$ are called square matrices. In what follows, we provide some examples.

Example 1.2. Consider the matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \text{ respectively } B = \begin{pmatrix} i & 2+i & 0 \\ -3 & \sqrt{2} & -1+3i \end{pmatrix},$$

where i is the imaginary unit. Then $A \in \mathcal{M}_3(\mathbb{R})$, or in other words, A is a real valued square matrix, meanwhile $B \in \mathcal{M}_{2,3}(\mathbb{C})$, or in other words, B is a complex valued matrix with two rows and three columns.

In what follows we present some special matrices.

Example 1.3. Consider the matrix $I_n = (a_{ij})_{i,j=\overline{1,n}} \in \mathcal{M}_n(\mathbb{F}), a_{ij} = 1$, if $i = j$ and $a_{ij} = 0$ otherwise. Here $1 \in \mathbb{F}$, respectively $0 \in \mathbb{F}$ are the multiplicative identity respectively the zero element of the field $\mathbb{F}$.

Then

$$I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

and is called the *identity matrix* (or *unit matrix*) of order n.

Remark 1.4. Sometimes we denote the identity matrix simply by I.

Example 1.5. Consider the matrices $O = (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F})$ having all entries the zero element of the field $\mathbb{F}$. Then

$$O = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

and is called the *null matrix* of order $m \times n$.

Example 1.6. Consider the matrices $A = (a_{ij})_{i,j=\overline{1,m}} \in \mathcal{M}_n(\mathbb{F})$ given by $a_{ij} = 0$ whenever $i > j$, respectively $a_{ij} = 0$ whenever $i < j$. Then

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}, \text{ respectively } A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

is called *upper triangular*, respectively *lower triangular* matrix.

If all entries outside the main diagonal are zero, A is called a *diagonal* matrix. In this case we have

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

Addition of Matrices.

If A and B are $m \times n$ matrices, the *sum* of A and B is defined to be the $m \times n$ matrix $A + B$ obtained by adding corresponding entries. Hence, the addition operation is a function

$$+ : \mathcal{M}_{m,n}(\mathbb{F}) \times \mathcal{M}_{m,n}(\mathbb{F}) \to \mathcal{M}_{m,n}(\mathbb{F}),$$

$$(a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} + (b_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} = (a_{ij} + b_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}, \forall\, (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}, (b_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F}).$$

In other words, for $A, B \in \mathcal{M}_{m,n}(\mathbb{F})$ their sum is defined as

$$C = A + B = (c_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}}$$

where $c_{ij} = a_{ij} + b_{ij}$ for all $i \in \{1, 2, \dots, m\}$, $j \in \{1, 2, \dots, n\}$.

Properties of Matrix Addition.

Let $O \in \mathcal{M}_{m,n}(\mathbb{F})$ the null matrix of size $m \times n$. For a given matrix $X = (x_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F})$ we denote by $-X$ its *additive inverse* (*opposite*), that is, $-X = (-x_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F})$. For every $A, B, C \in \mathcal{M}_{m,n}(\mathbb{F})$ the following properties hold:

1. $A + B$ is again an $m \times n$ matrix (closure property).
2. $(A + B) + C = A + (B + C)$ (associative property).
3. $A + B = B + A$ (commutative property).
4. $A + O = O + A = A$ (additive identity).
5. $A + (-A) = (-A) + A = O$ (the additive inverse).

It turns out that $(\mathcal{M}_{m,n}(\mathbb{F}), +)$ is an Abelian group.

Scalar Multiplication.

For $A \in \mathcal{M}_{m,n}(\mathbb{F})$ and $\alpha \in \mathbb{F}$ define $\alpha A = (\alpha a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$. Hence, the *scalar multiplication* operation is a function

$$\cdot : \mathbb{F} \times \mathcal{M}_{m,n}(\mathbb{F}) \to \mathcal{M}_{m,n}(\mathbb{F}),$$

$$\alpha \cdot (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}} = (\alpha \cdot a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}, \forall\, \alpha \in \mathbb{F}, (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}, \in \mathcal{M}_{m,n}(\mathbb{F}).$$

Properties of Scalar Multiplication.

Obviously, for every $A, B \in \mathcal{M}_{m,n}(\mathbb{F})$ and $\alpha, \beta \in \mathbb{F}$ the following properties hold:

1. αA is again an $m \times n$ matrix (closure property).
2. $(\alpha\beta)A = \alpha(\beta A)$ (associative property).

3. $\alpha(A+B) = \alpha A + \alpha B$ (distributive property).

4. $(\alpha+\beta)A = \alpha A + \beta A$ (distributive property).

5. $1A = A$, where 1 is the multiplicative identity of $\mathbb{F}$ (identity property).

Of course that we listed here only the left multiplication of matrices by scalars. By defining $\alpha A = A\alpha$ we obtain the right multiplication of matrices by scalars.

Example 1.7. If $A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ -2 & 2 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & 1 \\ 0 & -1 & 2 \end{pmatrix}$, then $2A - B = \begin{pmatrix} 3 & -2 & 0 \\ -1 & 5 & -3 \\ -4 & 5 & -2 \end{pmatrix}$ and $2A + B = \begin{pmatrix} 1 & -2 & 4 \\ 1 & 3 & -1 \\ -4 & 3 & 2 \end{pmatrix}$.

Transpose.

The *transpose* of a matrix $A \in \mathcal{M}_{m,n}(\mathbb{F})$ is defined to be a matrix $A^\top \in \mathcal{M}_{n,m}(\mathbb{F})$ obtained by interchanging rows and columns of A. Locally, if $A = (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$, then $A^\top = (a_{ji})_{\substack{j=\overline{1,n} \\ i=\overline{1,m}}}$.

It is clear that $(A^\top)^\top = A$. A matrix, that has many columns, but only one row, is called a *row matrix*. Thus, a row matrix A with n columns is an $1 \times n$ matrix, i.e.

$$A = (a_1 \;\; a_2 \;\; a_3 \;\; \dots \;\; a_n).$$

A matrix, that has many rows, but only one column, is called a *column matrix*. Thus, a column matrix A with m rows is an $m \times 1$ matrix, i.e.

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

Obviously, the transpose of a row matrix is a column matrix and viceversa, hence, in inline text a column matrix A is represented as

$$A = (a_1 \ \ a_2 \ \ \ldots \ \ a_m)^\top.$$

Conjugate Transpose.

Let $A \in \mathcal{M}_{m,n}(\mathbb{C})$. Define the *conjugate transpose* of $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{C})$ by $A^\star = (\overline{a}_{ji})_{\substack{j=\overline{1,n}\\ i=\overline{1,m}}}$, where $\overline{z}$ denotes the complex conjugate of the number $z \in \mathbb{C}$. We have that $(A^\star)^\star = A$ and $A^\top = A^\star$ whenever A contains only real entries.

Properties of the Transpose.

For every $A, B \in \mathcal{M}_{m,n}(\mathbb{F})$ and $\alpha \in K$ hold:

1. $(A + B)^\top = A^\top + B^\top$.

2. $(A + B)^\star = A^\star + B^\star$.

3. $(\alpha A)^\top = \alpha A^\top$ and $(\alpha A)^\star = \overline{\alpha} A^\star$.

Symmetries.

Let $A = (a_{ij})_{\substack{i=\overline{1,n}\\ j=\overline{1,n}}} \in \mathcal{M}_n(\mathbb{F})$ be a square matrix. We recall that

- A is said to be a *symmetric matrix* whenever $A = A^\top$ (locally $a_{ij} = a_{ji}$ for all $i, j \in \{1, 2, \ldots n\}$).

- A is said to be a *skew-symmetric matrix* whenever $A = -A^\top$ (locally $a_{ij} = -a_{ji}$ for all $i, j \in \{1, 2, \ldots n\}$).

- A is said to be a *hermitian matrix* whenever $A = A^\star$ (locally $a_{ij} = \overline{a}_{ji}$ for all $i, j \in \{1, 2, \ldots n\}$).

- A is said to be a *skew-hermitian matrix* whenever $A = -A^\star$ (locally $a_{ij} = -\overline{a}_{ji}$ for all $i, j \in \{1, 2, \ldots n\}$).

It can be easily observed that every symmetric real matrix is hermitian, respectively, every skew-symmetric real matrix is skew-hermitian.

Example 1.8. The matrix $A = \begin{pmatrix} 1 & -2 & 4 \\ -2 & 0 & 3 \\ 4 & 3 & 2 \end{pmatrix}$ is a symmetric matrix, meanwhile the matrix $B = \begin{pmatrix} 0 & 1 & -3 \\ -1 & 0 & 3 \\ 3 & -3 & 0 \end{pmatrix}$ is a skew-symmetric matrix. The matrix $C = \begin{pmatrix} 1 & 1+i & i \\ 1-i & 3 & 3-2i \\ -i & 3+2i & 2 \end{pmatrix}$ is a hermitian matrix, meanwhile the matrix $D = \begin{pmatrix} -i & 2-i & -3i \\ -2-i & i & 2+3i \\ -3i & -2+3i & 0 \end{pmatrix}$ is a skew-hermitian matrix.

Matrix multiplication.

For a matrix $X = (x_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F})$ we denote by X_{i*} its i^{th} row, i.e. the row matrix

$$X_{i\star} = (x_{i1} \;\; x_{i2} \;\; \ldots \;\; x_{in}).$$

Similarly, the j^{th} column of X is the column matrix

$$X_{\star j} = (x_{1j} \;\; x_{2j} \;\; \ldots \;\; x_{mj})^\top.$$

It is obvious that

$$(X^\top)_{i\star} = (X_{\star i})^\top,$$

respectively

$$(X^\top)_{\star j} = (X_{j\star})^\top.$$

We say that the matrices A and B are *conformable* for multiplication in the order AB, whenever A has exactly as many columns as B has rows, that is $A \in \mathcal{M}_{m,p}(\mathbb{F})$ and $B \in \mathcal{M}_{p,n}(\mathbb{F})$.

For conformable matrices $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,p}}}$ and $B = (b_{jk})_{\substack{j=\overline{1,p}\\ k=\overline{1,n}}}$ the *matrix product* AB is defined to be the $m \times n$ matrix $C = (c_{ik})_{\substack{i=\overline{1,m}\\ k=\overline{1,n}}}$ with

$$c_{ik} = A_{i\star}B_{\star k} = \sum_{j=1}^{p} a_{ij}b_{jk}.$$

In the case that A and B failed to be conformable, the product AB is not defined.

Remark 1.9. Note, the product is not commutative, that is, in general, $AB \neq BA$ even if both products exists and have the same shape.

Example 1.10. Let $A = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}$ and $BA = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

Rows and columns of a product.

Suppose that $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,p}}} \in \mathcal{M}_{m,p}(\mathbb{F})$ and $B = (b_{ij})_{\substack{i=\overline{1,p}\\ j=\overline{1,n}}} \in \mathcal{M}_{p,n}(\mathbb{F})$.

There are various ways to express the individual rows and columns of a matrix product. For example the i^{th} row of AB is

$$\begin{aligned} C_{i\star} &= [AB]_{i\star} = [A_{i\star}B_{\star 1} \;\; A_{i\star}B_{\star 2} \;\; \ldots \;\; A_{i\star}B_{\star n}] = A_{i\star}B \\ &= \begin{pmatrix} a_{i1} & a_{i2} & \ldots & a_{ip} \end{pmatrix} \begin{pmatrix} B_{1\star} \\ B_{2\star} \\ \vdots \\ B_{p\star} \end{pmatrix} \end{aligned}$$

There are some similar representations for individual columns, i.e. the j^{th} column is

$$\begin{aligned} C_{\star j} &= [AB]_{\star j} = [A_{1\star}B_{\star j} \;\; A_{2\star}B_{\star j} \;\; \ldots \;\; A_{m\star}B_{\star j}]^\top = AB_{\star j} \\ &= \begin{pmatrix} A_{\star 1} & A_{\star 2} & \ldots & A_{\star p} \end{pmatrix} \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{pj} \end{pmatrix} \end{aligned}$$

Consequently, we have:

1. $[AB]_{i\star} = A_{i\star}B \quad (i^{th} \quad \text{row of} \quad AB)$.

2. $[AB]_{\star j} = AB_{\star j} \quad (j^{th} \quad \text{column of} \quad AB)$.

3. $[AB]_{i\star} = a_{i1}B_{1\star} + a_{i2}B_{2\star} + \cdots + a_ipB_{p\star} = \sum_{k=1}^{p} a_{ik}B_{k\star}$.

4. $[AB]_{\star j} = A_{\star 1}b_{1j} + A_{\star 2}b_{2j} + \cdots + A_{\star p}b_{pj} = \sum_{k=1}^{p} A_{\star k}b_{kj}$.

The last two equations have both theoretical and practical importance. They show that the rows of AB are combinations of rows of B, while the columns of AB are combinations of columns of A. So it is waisted time to compute the entire product when only one row or column is needed.

Properties of matrix multiplication.

Distributive and associative laws.

For conformable matrices one has:

1. $A(B+C) = AB + AC$ (left-hand distributive law).
2. $(B+C)A = BA + CA$ (right-hand distributive law).
3. $A(BC) = (AB)C$ (associative law).

For a matrix $A \in \mathcal{M}_n(\mathbb{F})$, one has

$$AI_n = A \quad \text{and} \quad I_n A = A\ ,$$

where $I_n \in \mathcal{M}_n(\mathbb{F})$ is the identity matrix of order n.

Proposition 1.11. *For conformable matrices $A \in \mathcal{M}_{m,p}(\mathbb{F})$ and $B \in \mathcal{M}_{p,n}(\mathbb{F})$, on has*

$$(AB)^\top = B^\top A^\top\ .$$

The case of conjugate transposition is similar:

$$(AB)^\star = B^\star A^\star\ .$$

Proof. Let $C = (c_{ij})_{\substack{i=\overline{1,n}\\ j=\overline{1,m}}} = (AB)^\top$. Then for every $i \in \{1, 2, \ldots, n\}$, $j \in \{1, 2, \ldots, m\}$ one has $c_{ij} = [AB]_{ji} = A_{j\star}B_{\star i}$. Let us consider now the (i,j) entry of $B^\top A^\top$.

$$\begin{aligned}
[B^\top A^\top]_{ij} &= (B^\top)_{i\star}(A^\top)_{\star j} = (B_{\star i})^\top (A_{j\star})^\top = \sum_{k=1}^{p} [B^\top]_{ik}[A^\top]_{kj} \\
&= \sum_{k=1}^{p} b_{ki}a_{jk} = \sum_{k=1}^{p} a_{jk}b_{ki} \\
&= A_{j\star}B_{\star i}
\end{aligned}$$

$\square$

Exercise. Prove that for every matrix $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F})$ the matrices $AA^\top$ and $A^\top A$ are symmetric matrices.

For a matrix $A \in \mathcal{M}_n(\mathbb{F})$, one can introduce its m^{th} power by

$$A^0 = I_n,\ A^1 = A,\ A^m = A^{m-1}A.$$

Example 1.12. If $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ then $A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$. Hence $A^m = A^{m(\mod)4}$.

Trace of a product. Let A be a square matrix of order n. The trace of A is the sum of the elements of the main diagonal, that is

$$\operatorname{trace} A = \sum_{i=1}^{n} a_{ii}\ .$$

Proposition 1.13. *For $A \in \mathcal{M}_{m,n}(\mathbb{C})$ and $B \in \mathcal{M}_{n,m}(\mathbb{C})$ one has* $\operatorname{trace} AB = \operatorname{trace} BA$.

Proof. We have

$$\operatorname{trace} AB = \sum_{i=1}^{m}[AB]_{ii} = \sum_{i=1}^{m}(A)_{i\star}(B)_{\star i} = \sum_{i=1}^{m}\sum_{k=1}^{n} a_{ik}b_{ki} =$$

$$\sum_{i=1}^{m}\sum_{k=1}^{n} b_{ki}a_{ik} = \sum_{k=1}^{n}\sum_{i=1}^{m} b_{ki}a_{ik} = \sum_{k=1}^{n}[BA]_{kk} = \operatorname{trace} BA.$$

□

Block Matrix Multiplication.

Suppose that A and B are partitioned into submatrices-referred to as blocks- as indicated below:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1r} \\ A_{21} & A_{22} & \dots & A_{2r} \\ \vdots & \vdots & \dots & \vdots \\ A_{s1} & A_{s2} & \dots & A_{sr} \end{pmatrix} \text{ and } B = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1t} \\ B_{21} & B_{22} & \dots & B_{2r} \\ \vdots & \vdots & \dots & \vdots \\ B_{r1} & B_{r2} & \dots & B_{rt} \end{pmatrix}$$

We say that the partitioned matrices are conformable partitioned if the pairs (A_{ik}, B_{kj}) are conformable matrices, for every indices i, j, k. In this case the product AB is formed by combining blocks exactly the same way as the scalars are combined in ordinary matrix multiplication. That is, the (i, j) block in the product AB is

$$A_{i1}B_{1j} + A_{i2}B_{2j} + \dots A_{ir}B_{rj} \ .$$

Matrix Inversion.

For a square matrix $A \in \mathcal{M}_n(\mathbb{F})$, the matrix $B \in \mathcal{M}_n(\mathbb{F})$ that satisfies

$$AB = I_n \text{ and } BA = I_n$$

(if exists) is called the *inverse* of A and is denoted by $B = A^{-1}$. Not all square matrices admits an inverse (are invertible). An invertible square matrix is called *nonsingular* and a square matrix with no inverse is called *singular matrix*. Although not all matrices are invertible, *when an inverse exists, it is unique*. Indeed, suppose that X_1 and X_2 are both inverses for a nonsingular matrix A. Then

$$X_1 = X_1 I_n = X_1(AX_2) = (X_1A)X_2 = I_n X_2 = X_2$$

which implies that only one inverse is possible.

Properties of Matrix Inversion. For nonsingular matrices $A, B \in \mathcal{M}_n(\mathbb{F})$, the following statements hold.

1. $(A^{-1})^{-1} = A$

2. The product AB is nonsingular.

3. $(AB)^{-1} = B^{-1}A^{-1}$.

4. $(A^{-1})^\top = (A^\top)^{-1}$ and $(A^{-1})^\star = (A^\star)^{-1}$.

One can easily prove the following statements.

Products of nonsingular matrices are nonsingular.

If $A \in \mathcal{M}_n(\mathbb{F})$ is nonsingular, then there is a unique solution $X \in \mathcal{M}_{n,p}(\mathbb{F})$ for the equation

$$AX = B, \text{ where } B \in \mathcal{M}_{n,p}(\mathbb{F}),$$

and the solution is $X = A^{-1}B$.

A system of n linear equations in n unknowns can be written in the form $Ax = b$, with $x, b \in \mathcal{M}_{n,1}(\mathbb{F})$, so it follows when A is nonsingular, that the system has a unique solution $x = A^{-1}b$.

1.2 Determinants and systems of linear equations

Determinants.

For every square matrix $A = (a_{ij})_{\substack{i=\overline{1,n}\\ j=\overline{1,n}}} \in \mathcal{M}_n(\mathbb{F})$ one can assign a scalar denoted $det(A)$ called the determinant of A. In extended form we write

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

In order to define the determinant of a square matrix, we need the following notations and notions. Recall that by a permutation of the integers $\{1, 2, ..., n\}$ we mean an arrangement of these integers in some definite order. In other words, a permutation is a bijection $\sigma : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$. It can easily be observed that the number of permutations of the integers $\{1, 2, ..., n\}$ equals $n! = 1 \cdot 2 \cdot \ldots \cdot n$. Let us denote by S_n the set of all permutations of the integers $\{1, 2, ..., n\}$. A pair (i, j) is called an inversion of a permutation $\sigma \in S_n$ is $i < j$ and $\sigma(i) > \sigma(j)$. A permutation $\sigma \in S_n$ is called even or odd according to whether the number of inversions of σ is even or odd respectively. The sign of a permutation $\sigma \in S_n$, denoted by sgn (σ), is $+1$ if the permutation is even and -1 if the permutation is odd.

Definition 1.14. *Let $A \in \mathcal{M}_n(\mathbb{F})$. The determinant of A is the scalar defined by the equation*

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn } (\sigma)\, a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}.$$

It can easily be computed, that for $A = (a_{ij})_{\substack{i=\overline{1,2} \\ j=\overline{1,2}}} \in \mathcal{M}_2(\mathbb{F})$, one has

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Similarly, if $A = (a_{ij})_{\substack{i=\overline{1,3} \\ j=\overline{1,3}}} \in \mathcal{M}_3(\mathbb{F})$, then its determinant can be calculated by the rule

$$\det(A) =$$

$$a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Example 1.15. If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ then

$\det(A) = 1 \cdot 5 \cdot 9 + 3 \cdot 4 \cdot 8 + 2 \cdot 6 \cdot 7 - 3 \cdot 5 \cdot 7 - 1 \cdot 6 \cdot 8 - 2 \cdot 4 \cdot 9 = 0.$

Laplace's theorem.

Let $A \in \mathcal{M}_n(\mathbb{F})$ and let k be an integer, $1 \leq k \leq n$. Consider the rows $i_1 \ldots i_k$ and the columns $j_1 \ldots j_k$ of A. By deleting the other rows and columns we obtain a submatrix of A of order k, whose determinant is called a **minor** of A and is denoted by $M_{i_1 \ldots i_k}^{j_1 \ldots j_k}$. Now let us delete the rows $i_1 \ldots i_k$ and the columns $j_1 \ldots j_k$ of A.. We obtain a submatrix of A of order $n-k$. Its determinant is called the complementary minor of $M_{i_1 \ldots i_k}^{j_1 \ldots j_k}$ and it is denoted by $\widetilde{M}_{i_1, \ldots i_k}^{j_1 \ldots j_k}$. Finally let us denote (the so called cofactor)

$$A_{i_1 \ldots i_k}^{j_1 \ldots j_k} = (-1)^{i_1 + \cdots + i_k + j_1 + \cdots + j_k} \widetilde{M}_{i_1 \ldots i_k}^{j_1 \ldots j_k}.$$

The adjugate of A is the matrix $\operatorname{adj}(A) = \left((A_i^j)_{\substack{i=\overline{1,n} \\ j=\overline{1,n}}} \right)^\top$, that is

$$\operatorname{adj}(A) = \begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & \cdots & A_n^2 \\ \vdots & \vdots & \cdots & \vdots \\ A_1^n & A_2^n & \cdots & A_n^n \end{pmatrix}$$

The next result provides a computation method of the inverse of a nonsingular matrix.

Theorem 1.16. *A square matrix $A \in \mathcal{M}_n(\mathbb{F})$ is invertible if and only if* $\det(A) \neq 0$. *In this case its inverse can be obtained by the formula*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A).$$

Corollary 1.17. *A linear system $Ax = 0$ with n equations in n unknowns has a non-trivial solution if and only if* $\det(A) = 0$.

We state, without proof, the Laplace expansion theorem:

Theorem 1.18.

$$\det(A) = \sum M_{i_1\ldots i_k}^{j_1\ldots j_k} A_{i_1\ldots i_k}^{j_1\ldots j_k} \text{ , where}$$

- *The indices $i_1 \ldots i_k$ are fixed*
- *The indices $j_1 \ldots j_k$ runs over all the possible values, such that $1 \le j_1 < \cdots < j_k \le n$.*

As immediate consequences we obtain the following methods of calculating determinants, called row expansion and column expansion.

Corollary 1.19. *Let $A \in \mathcal{M}_n(\mathbb{F})$. Then*

(i) $\det(A) = \sum_{k=1}^{n} a_{ik} A_i^k$, *(expansion by row i)*

(ii) $\det(A) = \sum_{k=1}^{n} a_{kj} A_k^j$, *(expansion by column j).*

Properties of the determinant.
Let $A, B \in \mathcal{M}_n(\mathbb{F})$ and let $a \in \mathbb{F}$. Then

(1) $\det(A^\top) = \det(A)$.

(2) A permutation of the rows, (respectively columns) of A multiplies the determinant by the sign of the permutation.

(3) A determinant with two equal rows (or two equal columns) is zero.

(4) The determinant of A is not changed if a multiple of one row (or column) is added to another row (or column).

(5) $\det(A^{-1}) = \frac{1}{\det(A)}$.

(6) $\det(AB) = \det(A)\det(B)$.

(7) $\det(aA) = a^n \det(A)$.

(8) If A is a triangular matrix, i.e. $a_{ij} = 0$ whenever $i > j$ ($a_{ij} = 0$ whenever $i < j$), then its determinant equals the product of the diagonal entries, that is $\det(A) = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn} = \prod_{i=1}^{n} a_{ii}$.

Rank. Elementary transformations.

A natural number r is called the rank of the matrix $A \in \mathcal{M}_{m,n}(\mathbb{F})$ if

1. There exists a square submatrix $M \in \mathcal{M}_r(\mathbb{F})$ of A which is nonsingular (that is $\det(M) \neq 0$).

2. If $p > r$, for every submatrix $N \in \mathcal{M}_p(\mathbb{F})$ of A one has $\det(N) = 0$.

We denote $\operatorname{rank}(A) = r$.

It can be proved that for $A \in \mathcal{M}_{m,n}(\mathbb{F})$ and $B \in \mathcal{M}_{n,p}(\mathbb{F})$, then

$$\operatorname{rank}(A) + \operatorname{rank}(B) - m \leq \operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\} .$$

Theorem 1.20. *Let $A, B \in \mathcal{M}_n(\mathbb{F})$ with $\det(A) \neq 0$. Then rank$(AB) =$ rank(B).*

Proof. Since $\det(A) \neq 0$, we have $\operatorname{rank}(A) = n$. By using the above notation with $m = p = n$ we obtain $\operatorname{rank}(B) \leq \operatorname{rank}(AB) \leq \operatorname{rank}(B)$. Hence $\operatorname{rank}(AB) = \operatorname{rank}(B)$. $\square$

Definition 1.21. *The following operations are called elementary row transformations on the matrix $A \in \mathcal{M}_{m,n}(\mathbb{F})$:*

1. *Interchanging of any two rows.*

2. *Multiplication of a row by any non-zero number.*

3. *The addition of one row to another.*

Similarly one can define the elementary column transformations.

Consider an arbitrary determinant. If it is nonzero it will be nonzero after performing elementary transformations. If it is zero it will remain zero. One can conclude that the rank of a matrix does not change if we perform any elementary transformation on the matrix. So we can use elementary transformation in order to compute the rank.

Namely, given a matrix $A \in \mathcal{M}_{m,n}(\mathbb{F})$ we transform it by an appropriate succession of elementary transformations- into a matrix B such that

- the diagonal entries of B are either 0 or 1, all the 1's preceding all the 0's on the diagonal.
- all the other entries of B are 0.

Since the rank is invariant under elementary transformations, we have $\text{rank}\,(A) = \text{rank}\,(B)$, but it is clear that the rank of B is equal to the number of 1's on the diagonal.

The next theorem offers a procedure to compute the inverse of a matrix:

Theorem 1.22. *If a square matrix is reduced to the identity matrix by a sequence of elementary row operations, the same sequence of elementary row transformations performed on the identity matrix produces the inverse of the given matrix.*

Example 1.23. Compute the inverse of the matrix $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ by using elementary row operations.

We write $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{(-\frac{1}{3}A_{3\star}+A_{2\star})}{\simeq}$

$$\left(\begin{array}{ccc} 1 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{array}\right) \Bigg| \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 1 \end{array}\right) \overset{(-A_{2\star}+A_{1\star})}{\simeq}$$
$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{array}\right) \Bigg| \left(\begin{array}{ccc} 1 & -1 & \frac{1}{3} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 1 \end{array}\right) \overset{(\frac{1}{2}A_{2\star}, \frac{1}{3}A_{3\star})}{\simeq} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \Bigg| \left(\begin{array}{ccc} 1 & -1 & \frac{1}{3} \\ 0 & \frac{1}{2} & -\frac{1}{6} \\ 0 & 0 & \frac{1}{3} \end{array}\right).$$

Hence $A^{-1} = \left(\begin{array}{ccc} 1 & -1 & \frac{1}{3} \\ 0 & \frac{1}{2} & -\frac{1}{6} \\ 0 & 0 & \frac{1}{3} \end{array}\right)$.

Recall that a matrix is in row echelon form if

(1) All nonzero rows are above any rows of all zeroes.

(2) The first nonzero element (leading coefficient) of a nonzero row is always strictly to the right of the first nonzero element of the row above it.

If supplementary the condition
(3) Every leading coefficient is 1 and is the only nonzero entry in its column,
is also satisfied, we say that the matrix is in reduced row echelon form.
An arbitrary matrix can be put in reduced row echelon form by applying a finite sequence of elementary row operations. This procedure is called the *Gauss-Jordan* elimination procedure.

Existence of an inverse. For a square matrix $A \in \mathcal{M}_n(\mathbb{F})$ the following statements are equivalent.

1. A^{-1} exists (A is nonsingular).

2. $\operatorname{rank}(A) = n$.

3. A is transformed by Gauss-Jordan in I_n.

4. $Ax = 0$ implies that $x = 0$.

Systems of linear equations.

Recall that a system of m linear equations in n unknowns can be written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = b_m. \end{cases}$$

Here $x_1, x_2, \ldots, x_n$ are the unknowns, $a_{11}, a_{12}, \ldots, a_{mn}$ are the coefficients of the system, and $b_1, b_2, \ldots, b_m$ are the constant terms. Observe that a systems of linear equations may be written as $Ax = b$, with $A = (a_{ij})_{\substack{i=\overline{1,m}\\ j=\overline{1,n}}} \in \mathcal{M}_{m,n}(\mathbb{F})$, $x \in \mathcal{M}_{n,1}(\mathbb{F})$ and $b \in \mathcal{M}_{m,1}(\mathbb{F})$. The matrix A is called the coefficient matrix, while the matrix $[A|b] \in \mathcal{M}_{m,n+1}(\mathbb{F})$,

$$[A|b]_{ij} = \begin{cases} a_{ij} \text{ if } j \neq n+1 \\ b_i \text{ if } j = n+1 \end{cases}$$

is called the augmented matrix of the system.

We say that $x_1, x_2, ..., x_n$ is a solution of a linear system if $x_1, x_2, ..., x_n$ satisfy each equations of the system. A linear system is consistent if it has a solution, and inconsistent otherwise. According to the Rouché-Capelli theorem, a system of linear equations is inconsistent if the rank of the augmented matrix is greater than the rank of the coefficient matrix. If, on the other hand, the ranks of these two matrices are equal, the system must have at least one solution. The solution is unique if and only if the rank equals the number of variables. Otherwise the general solution has k free parameters where k is the difference between the number of variables and the rank. Two linear systems are equivalent if and only if they have the same solution set.

In row reduction, the linear system is represented as an augmented matrix $[A|b]$. This matrix is then modified using elementary row operations until it reaches reduced row echelon form. Because these operations are reversible, the augmented matrix produced always represents a linear system that is equivalent to the original. In this way one can easily read the solutions.

Example 1.24. By using Gauss-Jordan elimination procedure solve the following systems of linear equations.

$$\begin{cases} x_1 - x_2 + 2x_4 = -2 \\ 2x_1 + x_2 - x_3 = 4 \\ x_1 - x_2 - 2x_3 + x_4 = 1 \\ x_2 + x_3 + x_4 = -1. \end{cases}$$

We have $[A|b] = \left(\begin{array}{cccc|c} 1 & -1 & 0 & 2 & -2 \\ 2 & 1 & -1 & 0 & 4 \\ 1 & -1 & -2 & 1 & 1 \\ 0 & 1 & 1 & 1 & -1 \end{array}\right) \overset{(-2A_{1\star}+A_{2\star},-A_{1\star}+A_{3\star})}{\simeq}$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 2 & -2 \\ 0 & 3 & -1 & -4 & 8 \\ 0 & 0 & -2 & -1 & 3 \\ 0 & 1 & 1 & 1 & -1 \end{array}\right) \overset{(A_{2\star}\leftrightarrow A_{4\star})}{\simeq}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 2 & -2 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & -2 & -1 & 3 \\ 0 & 3 & -1 & -4 & 8 \end{array}\right) \overset{(A_{2\star}+A_{1\star},-3A_{2\star}+A_{4\star})}{\simeq}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 3 & -3 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & -2 & -1 & 3 \\ 0 & 0 & -4 & -7 & 11 \end{array}\right) \underset{\simeq}{\scriptstyle (\frac{1}{2}A_{3\star}+A_{1\star}, \frac{1}{2}A_{3\star}+A_{.21\star}, -2A_{3\star}+A_{4\star})}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{5}{2} & -\frac{3}{2} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & -2 & -1 & 3 \\ 0 & 0 & 0 & -5 & 5 \end{array}\right) \underset{\simeq}{\scriptstyle (\frac{1}{2}A_{4\star}+A_{1\star}, \frac{1}{10}A_{4\star}+A_{2\star}, -\frac{1}{5}A_{4\star}+A_{3\star})}$$

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & -5 & 5 \end{array}\right) \underset{\simeq}{\scriptstyle (-\frac{1}{2}A_{3\star}, -\frac{1}{5}A_{4\star})} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right).$$

One can easily read the solution $x_1 = 1$, $x_2 = 1$, $x_3 = -1$, $x_4 = -1$.

Recall that a system of linear equations is called homogeneous if $b = (0\,0\,\cdots 0)^\top$ that is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = 0. \end{cases}$$

A homogeneous system is equivalent to a matrix equation of the form

$$Ax = O.$$

Obviously a homogeneous system is consistent, having the trivial solution $x_1 = x_2 = \cdots = x_n = 0$.

It can be easily realized that a homogeneous linear system has a non-trivial solution if and only if the number of leading coefficients in echelon form is less than the number of unknowns, in other words, the coefficient matrix is singular.

1.3 Problems

Problem 1.3.1. By using Laplace's theorem compute the following determinants.

$$D_1 = \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 2 & 3 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 \end{vmatrix}, \ D_2 = \begin{vmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{vmatrix}.$$

Problem 1.3.2. Compute the following determinants.

a) $\begin{vmatrix} 1 & \omega & \omega^2 & \omega^3 \\ \omega & \omega^2 & \omega^3 & 1 \\ \omega^2 & \omega^3 & 1 & \omega \\ \omega^3 & 1 & \omega & \omega^2 \end{vmatrix}$, where $\omega \in \mathbb{C}$ such that the relation $\omega^2 + \omega + 1 = 0$ holds.

b) $\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \epsilon & \epsilon^2 & \dots & \epsilon^{n-1} \\ 1 & \epsilon^2 & \epsilon^4 & \dots & \epsilon^{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \epsilon^{n-1} & \epsilon^{2(n-1)} & \dots & \epsilon^{(n-1)^2} \end{vmatrix}$, where $\epsilon = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$.

Problem 1.3.3. Let $A = (a_{ij})_{\substack{i=\overline{1,n}\\ j=\overline{1,n}}} \in \mathcal{M}_n(\mathbb{C})$ and let us denote
$\overline{A} = (\overline{a}_{ij})_{\substack{i=\overline{1,n}\\ j=\overline{1,n}}} \in \mathcal{M}_n(\mathbb{C})$. Show that

a) $\det(\overline{A}) = \overline{\det(A)}$.

b) If $\overline{a}_{ij} = a_{ji}$, $i, j \in \{1, 2, \dots, n\}$ then $\det(A) \in \mathbb{R}$.

Problem 1.3.4. Let $a_1, a_2, \dots a_n \in \mathbb{C}$. Compute the following determinants.

a) $\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \end{vmatrix}$.

b) $\begin{vmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_n & a_1 & a_2 & \dots & a_{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_2 & a_3 & a_4 & \dots & a_1 \end{vmatrix}$.

Problem 1.3.5. Compute A^n, $n \geq 1$ for the following matrices.

a) $A = \begin{pmatrix} 7 & 4 \\ -9 & -5 \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$, $a, b \in \mathbb{R}$.

b) $A = \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, $A = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}$, $a, b \in \mathbb{R}$.

Problem 1.3.6. Compute the rank of the following matrices by using the Gauss-Jordan elimination method.

a) $\begin{pmatrix} 0 & 1 & -2 & -3 & -5 \\ 6 & -1 & 1 & 2 & 3 \\ -2 & 4 & 3 & 2 & 1 \\ -3 & 0 & 2 & 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & -2 & 3 & -2 \\ 3 & -1 & 1 & -3 & 4 \\ -2 & 1 & 0 & 1 & -1 \\ 2 & 0 & 0 & -1 & 0 \end{pmatrix}$.

b) $$\begin{pmatrix} 1 & -2 & 3 & 5 & -3 & 6 \\ 0 & 1 & 2 & 3 & 4 & 7 \\ 2 & 1 & 3 & 3 & -2 & 5 \\ 5 & 0 & 9 & 11 & -7 & 16 \\ 2 & 4 & 9 & 12 & 10 & 26 \end{pmatrix}.$$

Problem 1.3.7. Find the inverses of the following matrices by using the Gauss-Jordan elimination method.

a) $A = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & 3 \\ 3 & 1 & -1 \end{pmatrix}$.

b) $A = (a_{ij})_{\substack{i=\overline{1,n} \\ j=\overline{1,n}}} \in \mathcal{M}_n(\mathbb{R})$, where $a_{ij} = \begin{cases} 1 \text{ if } i \neq j \\ 0 \text{ otherwise.} \end{cases}$

Problem 1.3.8. Prove that if A and B are square matrices of the same size, both invertible, then:

a) $A(I + A)^{-1} = (I + A^{-1})^{-1}$,

b) $(A + BB^{\top})^{-1}B = A^{-1}B(I + B^{\top}A^{-1}B)^{-1}$,

c) $(A^{-1} + B^{-1})^{-1} = A(A + B)^{-1}B$,

d) $A - A(A + B)^{-1}A = B - B(A + B)^{-1}B$,

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

f) $(I + AB)^{-1} = I - A(I + BA)^{-1}B$,

g) $(I + AB)^{-1}A = A(I + BA)^{-1}$.

Problem 1.3.9. For every matrix $A \in \mathcal{M}_{m,n}(\mathbb{C})$ prove that the products $A^\star A$ and $AA^\star$ are hermitian matrices.

Problem 1.3.10. For a quadratic matrix A of order n explain why the equation

$$AX - XA = I$$

has no solution.

Problem 1.3.11. Solve the following systems of linear equations by using Gauss-Jordan elimination procedure.

a)

$$\begin{cases} 2x_1 - 3x_2 + x_3 + 4x_4 = 13 \\ 3x_1 + x_2 - x_3 + 8x_4 = 2 \\ 5x_1 + 3x_2 - 4x_3 + 2x_4 = -12 \\ x_1 + 4x_2 - 2x_3 + 2x_4 = -12. \end{cases}$$

b)

$$\begin{cases} x_1 - x_2 + x_3 - x_4 + x_5 - x_6 = 1 \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1 \\ 2x_1 + x_3 - x_5 = 1 \\ x_2 - 3x_3 + 4x_4 = -4 \\ -x_1 + 3x_2 + 5x_3 - x_6 = -1 \\ x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6 = 2 \end{cases}$$

Problem 1.3.12. Find $m, n, p \in \mathbb{R}$ such that the following systems be consistent, and then solve the systems.

a)

$$\begin{cases} 2x - y - z = 0 \\ x + 2y - 3z = 0 \\ 2x + 3y + mz = 0 \\ nx + y + z = 0 \\ x + py + 6z = 0 \\ 2e^x = y + z + 2. \end{cases}$$

b)

$$\begin{cases} 2x - y + z = 0 \\ -x + 2y + z = 0 \\ mx - y + 2z = 0 \\ x + ny - 2z = 0 \\ 3x + y + pz = 0 \\ x^2 + y^2 + x^2 = 3. \end{cases}$$

2

Vector Spaces

2.1 Definition and basic properties of a Vector Space

Definition 2.1. *A vector space V over a field $\mathbb{F}$ (or $\mathbb{F}$ vector space) is a set V with an addition $+$ (internal composition law) such that $(V, +)$ is an abelian group and a scalar multiplication $\cdot : \mathbb{F} \times V \to V, (\alpha, v) \to \alpha \cdot v = \alpha v$, satisfying the following properties:*

1. $\alpha(v + w) = \alpha v + \alpha w$, $\forall \alpha \in \mathbb{F}, \forall v, w \in \mathbb{F}$
2. $(\alpha + \beta)v = \alpha v + \beta v, \forall \alpha, \beta \in \mathbb{F}, \forall v \in V$
3. $\alpha(\beta v) = (\alpha\beta)v$
4. $1 \cdot v = v, \forall v \in V$

The elements of V are called vectors and the elements of $\mathbb{F}$ are called scalars. The scalar multiplication depends upon $\mathbb{F}$. For this reason when we need to be exact we will say that V is a vector space over $\mathbb{F}$, instead of simply saying that V is a vector space. Usually a vector space over $\mathbb{R}$ is called a real vector space and a vector space over $\mathbb{C}$ is called a complex vector space.

Remark. From the definition of a vector space V over $\mathbb{F}$ the following rules for calculus are easily deduced:

- $\alpha \cdot 0_V = 0$
- $0_{\mathbb{F}} \cdot v = 0_V$
- $\alpha \cdot v = 0_V \Rightarrow \alpha = 0_{\mathbb{F}}$ or $v = 0_V$.

Examples. We will list a number of simple examples, which appear frequently in practice.

- $V = \mathbb{C}^n$ has a structure of $\mathbb{R}$ vector space, but it also has a structure of $\mathbb{C}$ vector space.
- $V = \mathbb{F}[X]$, the set of all polynomials with coefficients in $\mathbb{F}$ with the usual addition and scalar multiplication is ann $\mathbb{F}$ vector space.
- $M_{m,n}(\mathbb{F})$ with the usual addition and scalar multiplication is a $\mathbb{F}$ vector space.
- $C_{[a,b]}$, the set of all continuous real valued functions defined on the interval $[a, b]$, with the usual addition and scalar multiplication is an $\mathbb{R}$ vector space.

2.2 Subspaces of a vector space

It is natural to ask about subsets of a vector space V which are conveniently closed with respect to the operations in the vector space. For this reason we give the following:

Definition 2.2. *Let V be a vector space over $\mathbb{F}$. A subset $U \subset V$ is called subspace of V over $\mathbb{F}$ if it is stable with respect to the composition laws, that is,*
$v + u \in U, \forall v, u \in U$, and $\alpha v \in U \forall \alpha \in \mathbb{F}, v \in U$, and the induced operations verify the properties from the definition of a vector space over $\mathbb{F}$.

It is easy to prove the following propositions:

Proposition 2.3. *Let V be a $\mathbb{F}$ vector space and $U \subset V$ a nonempty subset. U is a vector subspace of V over $\mathbb{F}$ iff the following conditions are met:*

- $v - u \in U, \forall v, u \in U$
- $\alpha v \in U, \forall \alpha \in \mathbb{F}, \forall v \in U$

Proof. Obviously, the properties of multiplication with scalars, respectively the associativity and commutativity of addition operation are inherited from V. Hence, it remains to prove that $0 \in U$ and for all $u \in U$ one has $-u \in U$. Since $\alpha u \in U$ for every $u \in U$ and $\alpha \in \mathbb{F}$ it follows that $0u = 0 \in U$ and $0 - u = -u \in U$. □

Proposition 2.4. *Let V be a $\mathbb{F}$ vector space and $U \subset V$ a nonempty subset. U is a vector subspace of V over $\mathbb{F}$ iff*

$$\alpha v + \beta u \in U, \forall \alpha, \beta \in \mathbb{F}, \forall v, u \in V.$$

Proof. Let $u, v \in U$. For $\alpha = 1$, $\beta = -1$ we have $v - u \in U$. For $\beta = 0$ and $\alpha \in \mathbb{F}$ we obtain $\alpha v \in U$. The conclusion follows from the previous proposition. □

Example 2.5. Let $S = \{(x, y, z) \in \mathbb{R}^3 | x + y + z = 0\}$. Show that S is a subspace of $\mathbb{R}^3$.

To see that S is a subspace we check that for all $\alpha, \beta \in \mathbb{R}$ and all
$v_1 = (x_1, y_1, z_1), v_2 = (x_2, y_2, z_2) \in S$

$$\alpha v_1 + \beta v_2 \in S.$$

Indeed, since $v_1, v_2 \in S$ we have

$$\begin{aligned} x_1 + y_1 + z_1 &= 0 \\ x_2 + y_2 + z_2 &= 0, \end{aligned}$$

and by multiplying the equations with α and β respectively, and adding the resulting equations we obtain

$$(\alpha x_1 + \beta x_2) + (\alpha y_1 + \beta y_2) + (\alpha z_1 + \beta z_2) = 0.$$

But this is nothing else than the fact that
$\alpha v_1 + \beta v_2 = (\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2, \alpha z_1 + \beta z_2)$ satisfies the equation that defines S. The next propositions show how one can operate with vector subspaces (to obtain a new vector subspace) and how one can obtain a subspace from a family of vectors.

Proposition 2.6. *Let V be a vector space and $U, W \subset V$ two vector subspaces. The sets*

$$U \cap W \text{ and } U + W = \{u + w | u \in U, w \in W\}$$

are subspaces of V.

Proof. We prove the statements by making use of the Proposition 2.4. Let $\alpha, \beta \in \mathbb{F}$ and let $u, v \in U \cap W$. Then $u, v \in U$ and $u, v \in W$. Since U and W are vector spaces it follows that $\alpha v + \beta u \in U$, respectively $\alpha v + \beta u \in W$. Hence $\alpha v + \beta u \in U \cap W$. Now consider $\alpha, \beta \in \mathbb{F}$ and let $x, y \in U + W$. Then $x = u_1 + w_1$, $y = u_2 + w_2$ for some vectors $u_1, u_2 \in U$, $w_1, w_2 \in W$. But then

$$\alpha x + \beta y = (\alpha u_1 + \beta u_2) + (\alpha w_1 + \beta w_2) \in U + W.$$

□

The subspace $U \cap W$ is called the intersection vector subspace, while the subspace $U + W$ is called the sum vector subspace. Of course that these definitions can be also given for finite intersections (respectively finite sums) of subspaces.

Proposition 2.7. *Let V be a vector space over $\mathbb{F}$ and $S \subset V$ nonempty. The set* $\langle S \rangle = \left\{ \sum_{i=1}^{n} \alpha_i v_i : \ \alpha_i \in \mathbb{F} \text{ and } v_i \in S, \text{ for all } i = \overline{1,n},\, n \in \mathbb{N} \right\}$ *is a vector subspace over $\mathbb{F}$ of V.*

Proof. The proof is straightforward in virtue of Proposition 2.4. □

The above vector space is called the vector space generated by S, or the linear hull of the set S and is often denoted by $\mathrm{span}(S)$. It is the smallest subspace of V which contains S, in the sense that for every U subspace of V with $S \subset U$ it follows that $\langle S \rangle \subset U$.
Now we specialize the notion of sum of subspaces, to direct sum of subspaces.

Definition 2.8. *Let V be a vector space and $U_i \subset V$ subspaces, $i = \overline{1,n}$. The sum $U_1 + \cdots + U_n$ is called direct sum if for every $v \in U_1 + \cdots + U_n$, from $v = u_1 + \cdots + u_n = w_1 + \cdots + w_n$ with $u_i, w_i \in U_i, i = \overline{1,n}$ it follows that $u_i = w_i$, for every $i = \overline{1,n}$.*

The direct sum of the subspaces $U_i, i = \overline{1,n}$ will be denoted by $U_1 \oplus \cdots \oplus U_n$. The previous definition can be reformulated as follows. Every $u \in U_1 + \cdots + U_n$ can be written uniquely as $u = u_1 + u_2 + \ldots + u_n$ where $u_i \in U_i$, $i = \overline{1,n}$.
The next proposition characterizes the direct sum of two subspaces.

Proposition 2.9. *Let V be a vector space and $U, W \subset V$ be subspaces. The sum $U + W$ is a direct sum iff $U \cap W = \{0_V\}$.*

Proof. Assume that $U + W$ is a direct sum and there exists $s \in U \cap W$, $s \neq 0_V$. But then every $x \in U + W$, $x = u + w$ can be written as

$x = (u - s) + (w + s) \in U + W$. From the definition of the direct sum we have $u = u - s$, $w = w + s$ hence $s = 0_V$, contradiction.
Conversely, assume that $U \cap W = \{0_V\}$ and $U + W$ is not a direct sum. Hence, there exists $x \in U + W$ such that $x = u_1 + w_1 = u_2 + w_2 \in U + W$ and $u_1 \neq u_2$ or $w_1 \neq w_2$. But then $u_1 - u_2 = w_1 - w_2$, hence $u_1 - u_2, w_1 - w_2 \in U \cap W$. It follows that $u_1 = u_2$ and $w_1 = w_2$, contradiction. $\square$

Let V be a vector space over $\mathbb{F}$ and U be a subspace. On V one can define the following binary relation $\mathfrak{R}_U$: let $u, v \in V$, $u \,\mathfrak{R}_U\, v$ iff $u - v \in U$.
It can easily be verified that the relation $\mathfrak{R}_U$ is an equivalence relation, that is

(r) $v \,\mathfrak{R}_U\, v$, for all $v \in V$. (reflexivity)

(t) $u \,\mathfrak{R}_U\, v$ and $v \,\mathfrak{R}_U\, w \Longrightarrow u \,\mathfrak{R}_U\, w$, for all $u, v, w \in V$. (transitivity)

(s) $u \,\mathfrak{R}_U\, v \Longrightarrow v \,\mathfrak{R}_U\, u$, for all $u, v \in V$. (symmetry)

The equivalence class of a vector $v \in V$ is defined as

$$\mathfrak{R}_U[v] = \{u \in V : v \,\mathfrak{R}_U\, u\} = v + U.$$

The quotient set (or factor set) $V/\mathfrak{R}_U$ is denoted by V/U and consists of the set of all equivalence classes, that is

$$V/U = \{\mathfrak{R}_U[v] : v \in V\}.$$

Theorem 2.10. *On the factor set V/U there is a natural structure of a vector space over $\mathbb{F}$.*

Proof. Indeed, let us define the sum of two equivalence class $\mathfrak{R}_U[v]$ and $\mathfrak{R}_U[w]$ by

$$\mathfrak{R}_U[v] + \mathfrak{R}_U[v] = \mathfrak{R}_U[v + w]$$

and the multiplication with scalars by

$$\alpha\mathfrak{R}_U[v] = \mathfrak{R}_U[\alpha v].$$

Then, is an easy verification that with these operations V/U becomes an $\mathbb{F}$ space. □

The vector space from the previous theorem is called the factor vector space, or the quotient vector space.

2.3 Basis. Dimension.

Up to now we have tried to explain some properties of vector spaces "in the large". Namely we have talked about vector spaces, subspaces, direct sums, factor space. The Proposition 2.7 naturally raises some questions related to the structure of a vector space V. Is there a set S which generates V (that is $\langle S\rangle = V$)? If the answer is yes, how big should it be? Namely how big should a "minimal" one (minimal in the sense of cardinal numbers) be? Is there a finite set which generates V? We will shed some light on these questions in the next part of this chapter. Why are the answers to such questions important? The reason is quite simple. If we control (in some way) a minimal system of generators, we control the whole space.

Definition 2.11. *Let V be an $\mathbb{F}$ vector space. A nonempty set $S \subset V$ is called system of generators for V if for every $v \in V$ there exists a finite subset $\{v_1, \dots, v_n\} \subset V$ and the scalars $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ (it is also said that V is a linear combination of $v_1, \dots, v_n$ with scalars in $\mathbb{F}$). V is called dimensionally finite, or finitely generated, if it has a finite system of generators.*

A nonempty set $L \subset V$ is called a linearly independent system of vectors if for every finite subset $\{v_1, \ldots, v_n\} \subset L$ of it $\alpha_1 v_1 + \ldots \alpha_n v_n = 0$ implies that $a_i = 0$ for all $i = \overline{1,n}$.
A nonempty set of vectors which is not linearly independent is called linearly dependent.
A subset $\mathfrak{B} \subset V$ is called basis of V if it is both a system of generators and linearly independent. In this case every vector $v \in V$ can be uniquely written as a linear combination of vectors from $\mathfrak{B}$.

Example 2.12. Check whether the vectors $(0,1,2), (1,2,0), (2,0,1)$ are linearly independent in $\mathbb{R}^3$.

By definition, the three vectors are linearly independent if the implication

$$\alpha_1 (0,1,2) + \alpha_2 (1,2,0) + \alpha_3 (2,0,1) = 0_{\mathbb{R}^3} \Rightarrow \alpha_1 = \alpha_2 = \alpha_3 = 0$$

holds.
Checking the above implication actually amounts (after computing the right hand side) to investigating whether the linear system

$$\begin{cases} & & \alpha_2 & + & 2\alpha_2 & = & 0 \\ \alpha_1 & + & 2\alpha_2 & & & = & 0 \\ 2\alpha_1 & & & + & \alpha_2 & = & 0 \end{cases}$$

has **only** the trivial solution $(\alpha_1, \alpha_2, \alpha_3) = (0,0,0)$ or not. But we can easily compute the rank of the matrix, which is 3 due to

$$\begin{vmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{vmatrix} = -9 \neq 0,$$

to see that, indeed, the system has only the trivial solution, and hence the three vectors are linearly independent.
We have the following theorem.

Theorem 2.13. *(Existence of basis) Every vector space $V \neq 0$ has a basis.*

We will not prove this general theorem here, instead we will restrict to finite dimensional vector spaces.

Theorem 2.14. *Let $V \neq \{0\}$ be a finitely generated vector space over $\mathbb{F}$. From every finite system of generators one can extract a basis.*

Proof. Let $S = \{v_1, \ldots, v_r\}$ be a finite generators system. It is clear that there are nonzero vectors in S (otherwise $V = \{0\}$). Let $0 \neq v_1 \in S$. The set $\{v_1\}$ is linearly independent (because $\alpha v_1 = 0 \Rightarrow \alpha = 0$ from $v_1 \neq 0$). That means that S contains linearly independent subsets. Now $P(S)$ is finite (S being finite), and in a finite number of steps we can extract a maximal linearly independent system, let say $\mathfrak{B} = \{v_1, \ldots, v_n\},\ 1 \leq n \leq r$ in the following way:

$$v_2 \in S \setminus \langle v_1 \rangle,$$

$$v_3 \in S \setminus \langle \{v_1, v_2\} \rangle$$

$$\vdots$$

$$v_n \in S \setminus \langle \{v_1, v_2, \ldots, v_{n-1}\} \rangle.$$

We prove that $\mathfrak{B}$ is a basis for V. It is enough to show that $\mathfrak{B}$ generates V, because $\mathfrak{B}$ is linearly independent by the choice of it. Let $v \in V$. S being a system of generators it follows that it is enough to show that every $v_k \in S,\ n \leq k \leq r$ is a linear combination of vectors from $\mathfrak{B}$. Suppose, by contrary, that v_k is not a linear combination of vectors from $\mathfrak{B}$. It follows that the set $\mathfrak{B} \cup \{v_k\}$ is linearly independent, contradiction with the maximality of $\mathfrak{B}$. $\square$

Corollary 2.15. *Let V be an $\mathbb{F}$ vector space and S a system of generators for V. Every linearly independent set $L \subset S$ can be completed to a basis of V.*

Proof. Let $L \subset S$ be a linearly independent set in S. If L is maximal by the previous Theorem it follows that L is a basis. If L is not maximal, there exists a linearly independent set L_1 with $L \subset L_1 \subset S$. If L_1 is maximal it follows that L_1 is a basis. If it is not maximal, we repeat the previous step. Because S is a finite set, after a finite number of steps we obtain a system of linearly independent vectors $\mathfrak{B}$ which is maximal, $L \subset \mathfrak{B} \subset S$, so $\mathfrak{B}$ is a basis for V, again by the previous Theorem. □

Theorem 2.16. *Let V be a finitely generated vector space over $\mathbb{F}$. Every linearly independent system of vectors L can be completed to a basis of V.*

Proof. Let S be a finite system of generators. The intersection $L \cap S$ is again a system of generators and $L \subset L \cap S$. We apply the previous corollary and we obtain that L can be completed to a basis of V. □

Theorem 2.17. *(The cardinal of a basis). Let V be a finitely generated $\mathbb{F}$ vector space. Every basis of V is finite and has the same number of elements.*

Proof. Let $\mathfrak{B} = \{e_1, \ldots . e_n\}$ be a basis of V, and let $\mathfrak{B}'\{e'_1, \ldots, e'_m\}$ a system of vectors with $m > n$. We show that $\mathfrak{B}'$ can not be a basis for V.
Because $\mathfrak{B}$ is a basis the vectors e'_i can be uniquely written as $e'_i = \sum_{j=1}^n a_{ij} e_j$, $1 \leq i \leq m$. If $\mathfrak{B}'$ is linearly independent, then it follows that $\sum_{i=1}^m \lambda_i e'_i = 0$ implies $\lambda_i = 0$, $i = \overline{1, m}$, or, in other words, the system $\sum_{i=1}^m a_{ij} \lambda_i = 0$, $j = \overline{1, n}$ has only the trivial solution, impossible. □

Definition 2.18. *Let $V \neq \{0\}$ be an $\mathbb{F}$ vector space finitely generated. The number of elements in a basis of V is called the dimension of V (it does not depend on the* *choice of the basis, and it is denoted by* $dim_{\mathbb{F}}V$*). The vector space* V *is said to be of finite dimension. For* $V = \{0\}$ *,* $dim_{\mathbb{F}}V = 0$*.*

Remark 2.19. *According to the proof of Theorem 2.17, if* $dim_{\mathbb{F}}V = n$ *then any set of* $m > n$ *vectors is linear dependent.*

Corollary 2.20. *Let* V *be a vector space over* $\mathbb{F}$ *of finite dimension,* $dim_{\mathbb{F}}V = n$*.*

1. *Any linearly independent system of* n *vectors is a basis. Any system of* m *vectors,* $m > n$ *is linearly dependent.*

2. *Any system of generators of* V *which consists of* n *vectors is a basis. Any system of* m *vectors,* $m < n$ *is not a system of generators*

Proof. a) Consider $L = \{v_1, \ldots, v_n\}$ a linearly independent system of n vectors. From the completion theorem (Theorem 2.16) it follows that L can be completed to a basis of V. It follows from the cardinal basis theorem (Theorem 2.17) that there is no need to complete L, so L is a basis.
Let L' be a system of m vectors, $m > n$. If L' is linearly independent it follows that L' can be completed to a basis (Theorem 2.16), so $\dim_{\mathbb{F}}V \geq m > n$, contradiction.
b) Let $S = \{v_1, \ldots, v_n\}$ be a system of generators which consists of n vectors. From the Theorem 2.14 it follows that a basis can be extracted from its n vectors. Again from the basis Theorem 2.17 it follows that there is no need to extract any vector, so S is a basis.
Let S' be a generators system which consists of m vectors, $m < n$. From the Theorem 2.14 it follows that from S' one can extract a basis, so $\dim_{\mathbb{F}}V \leq m < n$, contradiction. □

Remark 2.21. The dimension of a finite dimensional vector space is equal to any of the following:

- The number of the vectors in a basis.
- The minimal number of vectors in a system of generators.
- The maximal number of vectors in a linearly independent system.

Example 2.22. Let $S = \{(x, y, z) \in \mathbb{R}^3 | x + y + z = 0\}$. Give an example of a basis of S.

In example 2.5 we have shown that S is a subspace of $\mathbb{R}^3$. One can see that, from a geometric point of view, S is a plane passing through the origin, so $\dim S = 2$. This follows also from rewriting S as follows

$$\begin{aligned} S &= \left\{(x, y, z) \in \mathbb{R}^4 | x + y + z = 0\right\} \\ &= \{(x, y, -x - y) \,|x, y \in \mathbb{R}\} \\ &= \{x\,(1, 0, -1) + y\,(0, 1, -1) \,|x, y \in \mathbb{R}\} \\ &= span\,\{(1, 0, -1)\,, (0, 1, -1)\}\,. \end{aligned}$$

The vectors $(1, 0, -1)$ and $(0, 1, -1)$ are linearly independent so they form a basis of S.

Theorem 2.23. *Every linearly independent list of vectors in a finite dimensional vector space can be extended to a basis of the vector space.*

Proof. Suppose that V is finite dimensional and $\{v_1, \ldots, v_m\}$ is linearly independent. We want to extend this set to a basis of V. V being finite dimensional, there exists a finite set $\{w_1, \ldots, w_n\}$, a list of vectors which spans V.

- If w_1 is in the span of $\{v_1, \ldots, v_m\}$, let $B = \{v_1, \ldots, v_m\}$. If not, let $B = \{v_1, \ldots, v_m, w_1\}$.

- If w_j is in the span of B, let B unchanged. If w_j is not in the span of B, extend B by jointing w_j to it.

After each step B is still linearly independent. After n steps at most, the span of B includes all the w's. Thus B also spans V, and being linearly independent, it follows that it is a basis. $\square$

As an application we show that every subspace of a finite dimensional vector space can be paired with another subspace to form a direct sum which is the whole space.

Theorem 2.24. *Let V be a finite dimensional vector space and U a subspace of V. There exists a subspace W of V such that $V = U \oplus W$.*

Proof. Because V is finite dimensional, so is U. Choose $\{u_1, \dots, u_m\}$ a basis of U. This basis of U a linearly independent list of vectors, so it can be extended to a basis $\{u_1, \dots, u_m, w_1, \dots, w_n\}$ of V. Let $W = \langle w_1, \dots, w_n \rangle$.
We prove that $V = U \oplus W$. For this we will show that

$$V = U + W, \text{ and } U \cap W = \{0\}$$

Let $v \in V$, there exists $(a_1, \dots, a_m, b_1, \dots, b_n)$ such that

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_m,$$

because $\{u_1, \dots, u_m, w_1, \dots, w_n\}$ generates V. By denoting
$a_1 u_1 + \cdots + a_m u_m = u \in U$ and $b_1 w_1 + \cdots + b_n w_m = w \in W$ we have just proven that $V = U + W$.
Suppose now that $U \cap W \neq \{0\}$, so let $0 \neq v \in U \cap W$. Then there exist scalars $a_1, \dots, a_m \in \mathbb{F}$ and $b_1, \dots, b_n \in \mathbb{F}$ not all zero, with

$$v = a_1 u_1 + \cdots + a_m u_m = b_1 w_1 + \cdots + b_n w_m,$$

so

$$a_1u_1 + \cdots + a_mu_m - b_1w_1 - \cdots - b_nw_m = 0.$$

But this is a contradiction with the fact that $\{u_1, \ldots, u_m, w_1, \ldots, w_n\}$ is a basis of V, so we obtain the contradiction, i.e. $U \cap W = \{0\}$. □

The next theorem relates the dimension of the sum and the intersection of two subspaces with the dimension of the given subspaces:

Theorem 2.25. *If U and W are two subspaces of a finite dimensional vector space V, then*

$$dim\,(U + W) = dim\,U + dim\,W - dim\,(U \cap W)\ .$$

Proof. Let $\{u_1, \ldots, u_m\}$ be a basis of $U \cap W$, so $\dim U \cap W = m$. This is a linearly independent set of vectors in U and W respectively, so it can be extended to a basis $\{u_1, \ldots, u_m, v_1 \ldots v_i\}$ of U and a basis $\{u_1, \ldots, u_m, w_1, \ldots w_j\}$ of W, so $\dim U = m + i$ and $\dim W = m + j$. The proof will be complete if we show that $\{u_1, \ldots, u_m, v_1 \ldots, v_i, w_1, \ldots, w_j\}$ is a basis for $U + W$, because in this case

$$\begin{aligned}\dim\,(U + W) &= m + i + j \\ &= (m + i) + (m + j) - m \\ &= \dim U + \dim W - dim(U \cap W)\end{aligned}$$

The set $\mathrm{span}\{u_1, \ldots, u_m, v_1 \ldots, v_i, w_1, \ldots, w_j\}$ contains U and W, so it contains $U + W$. That means that to show that it is a basis for $U + W$ it is only needed to show that it is linearly independent. Suppose that

$$a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_iv_i + c_1w_1 + \cdots + c_jw_j = 0\ .$$

We have

$$c_1w_1 + \cdots + c_jw_j = -a_1u_1 - \cdots - a_mu_m - b_1v_1 - \cdots - b_iv_i$$

which shows that $w = c_1w_1 + \cdots + c_jw_j \in U$. But this is also in W, so it lies in $U \cap W$. Because $u_1, \ldots, u_m$ is a basis in $U \cap W$ it follows that there exist the scalars $d_1, \ldots, d_m \in \mathbb{F}$, not all zero, such that

$$c_1w_1 + \cdots + c_jw_j = -(d_1u_1 + \cdots + d_mu_m) \ .$$

But $\{u_1, \ldots, u_m, w_1, \ldots, w_j\}$ is a basis in W, so it is linearly independent, that is all c_i's are zero.
The relation involving a's, b's and c's becomes

$$a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_iv_i = 0 \ ,$$

so a's and b's are zero because the vectors $\{u_1, \ldots, u_m, v_1 \ldots, v_i\}$ form a basis in U. So all the a's, b's and c's are zero, that means that
$\{u_1, \ldots, u_m, v_1, \ldots, v_i, w_1, \ldots, w_j\}$ are linearly independent, and because that generates $U + W$, they form a basis of $U + W$. □

The previous theorem shows that the dimension fits well with the direct sum of spaces. That is, if $U \cap W = \{0\}$, the sum is the direct sum and we have

$$\dim (U \oplus W) = \dim U + \dim W \ .$$

This is true for the direct sum of any finite number of spaces as it is shown in the next theorem:

Theorem 2.26. *Let V be a finite dimensional space, U_i subspaces of V, $i = \overline{1, n}$, such that*

$$V = U_1 + \cdots + U_n \ ,$$

and

$$dim V = dim U_1 + \cdots + dim U_n \ .$$

Then

$$V = U_1 \oplus \cdots \oplus U_n \ .$$

Proof. One can choose a basis for each U_i. By putting all these bases in one list, we obtain a list of vectors which spans V (by the first property in the theorem), and it is also a basis, because by the second property, the number of vectors in this list is $\dim V$.

Suppose that we have $u_i \in U_i, \quad i = \overline{1,n}$, such that

$$0 = u_1 + \cdots + u_n \ .$$

Every u_i is represented as the sum of the vectors of basis of U_i, and because all these bases form a basis of V, it follows that we have a linear combination of the vectors of a base of V which is zero. So all the scalars are zero, that is all u_i are zero, so the sum is direct. $\square$

We end the section with two important observations. Let V be a vector space over $\mathbb{F}$ (not necessary finite dimensional). Consider a basis $\mathfrak{B} = (e_i)_{i \in I}$ of V.

We have the first representation theorem:

Theorem 2.27. *Let V be a vector space over $\mathbb{F}$ (not necessary finite dimensional). Let us consider a basis $\mathfrak{B} = (e_i)_{i \in I}$. For every $v \in V, v \neq 0$ there exist a unique subset $\mathfrak{B}' \subseteq \mathfrak{B}$, $\mathfrak{B}' = \{e_{i_1}, \ldots, e_{i_k}\}$ and the nonzero scalars $a_{i_1}, \ldots, a_{i_k} \in \mathbb{F}^*$, such that*

$$v = \sum_{j=1}^{k} a_{i_j} e_{i_j} = a_{i_1} e_{i_1} + \cdots + a_{i_k} e_{i_k}.$$

Proof. Obviously, by the definition of basis v is a finite linear combination of the elements of the basis. We must show the uniqueness. Assume the contrary, that

$$v = \sum_{i=1}^{n} \alpha_{j_i} e_{j_i} = \sum_{i=1}^{m} \alpha_{k_i} e_{k_i},\ \alpha_{j_i} \neq 0,\ i = \overline{1,n},\ \alpha_{k_i} \neq 0,\ i = \overline{1,m}.$$

Assume that there exists $e_{k_s} \notin \{e_{j_1}, \ldots, e_{j_n}\}$. Then, since $\sum_{i=1}^{n} \alpha_{j_i} e_{j_i} - \sum_{i=1}^{m} \alpha_{k_i} e_{k_i} = 0$ we obtain that $\alpha_{k_s} = 0$, contradiction. Similarly, $e_{j_s} \in \{e_{k_1}, \ldots, e_m\}$, for all $s = \overline{1,n}$. Hence, $m = n$ and one may assume that

$$v = \sum_{i=1}^{n} \alpha_{j_i} e_{j_i} = \sum_{i=1}^{n} \alpha_{k_i} e_{k_i},\ \alpha_{j_i} \neq 0,\ i = \overline{1,n},\ \alpha_{k_i} \neq 0,\ i = \overline{1,n}.$$

Using the relation $\sum_{i=1}^{n} \alpha_{j_i} e_{j_i} - \sum_{i=1}^{n} \alpha_{k_i} e_{k_i} = 0$ again we obtain that $\alpha_{j_i} = \alpha_{k_i}$, $i \in \{1, \ldots, n\}$, contradiction. $\square$

Example 2.28. Show that $\mathcal{B} = \{(1,1), (1,-1)\}$ is a basis of $\mathbb{R}^2$, and find the representation of the vector $v = (3,-1)$ with respect to $\mathcal{B}$ of the vector $v = (3,-1)$.

Our aim is to find the representation of $v = (3,-1)$ with respect to B, that is, to find two scalars $x, y \in \mathbb{R}$ such that

$$v = x\,(1,1) + y\,(1,-1)\,.$$

Expressing the above equality component wise gives a system with two unknowns, x and y

$$\begin{cases} x & + & y & = & 3 \\ x & - & y & = & -1. \end{cases}$$

Its unique solution, and the answer to our problem, is $x = 1, y = 2$.

2.4 Local computations

In this section we deal with some computations related to finite dimensional vector spaces.

Let V be an $\mathbb{F}$ finite dimensional vector space, with a basis $\mathfrak{B} = \{e_1, \ldots, e_n\}$. Any vector $v \in V$ can be uniquely represented as

$$v = \sum_{i=1}^{n} a_i e_i = a_1 e_1 + \cdots + a_n e_n.$$

The scalars $(a_1, \dots, a_n)$ are called the coordinates of the vector v in the basis $\mathfrak{B}$. It is obvious that if we have another basis $\mathfrak{B}^{'}$, the coordinates of the same vector in the new basis change. How we can measure this change? Let us start with a situation that is a bit more general.

Theorem 2.29. *Let V be a finite dimensional vector space over $\mathbb{F}$ with a basis $\mathfrak{B} = \{e_1, \dots, e_n\}$. Consider the vectors $S = \{e_1^{'}, \dots, e_m^{'}\} \subseteq V$:*

$$\begin{aligned} e_1^{'} &= a_{11}e_1 + \cdots + a_{1n}e_n \\ &\dots \\ e_m^{'} &= a_{m1}e_1 + \cdots + a_{mn}e_n \end{aligned}$$

Denote by $A = (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$ the matrix formed by the coefficients in the above equations. The dimension of the subspace $\langle S \rangle$ is equal to the rank of the matrix A, i.e. $\dim\langle S \rangle = \mathrm{rank} A$.

Proof. Let us denote by $X_i = (a_{i1}, \dots, a_{in}) \in \mathbb{F}^n$, $i = \overline{1,m}$ the coordinates of $e_i^{'}$, $i = \overline{1,m}$ in $\mathfrak{B}$. Then, the linear combination $\sum_{i=1}^{m} \lambda_i e_i^{'}$ has its coordinates $\sum_{i=1}^{m} \lambda_i X_i$ in $\mathfrak{B}$. Hence the set of all coordinate vectors of elements of $\langle S \rangle$ equals the subspace of $\mathbb{F}^n$ generated by $\{X_1, ..., X_m\}$. Moreover $e_1^{'}, \dots, e_m^{'}$ will be linearly independent if and only if $X_1, \dots, X_m$ are. Obviously, the dimension of the subspace $\langle X_1, ..., X_m \rangle$ of $\mathbb{F}^n$ is equal to the rank of the matrix

$$\begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix} = A.$$

$\square$

Consider now the case of $m = n$ in the above discussion. The set $S = \{e_1^{'}, \ldots, e_n^{'}\}$ is a basis iff rank$A = n$. We have now

$$\begin{aligned} e_1^{'} &= a_{11}e_1 + \cdots + a_{1n}e_n \\ e_2^{'} &= a_{21}e_1 + \cdots + a_{2n}e_n \\ &\ldots \\ e_n^{'} &= a_{n1}e_1 + \cdots + a_{nn}e_n, \end{aligned}$$

representing the relations that change from the basis $\mathfrak{B}$ to the new basis $\mathfrak{B}^{'} = S$. The matrix $A^\top$ is denoted by

$$P^{(e,e^{'})} = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix}.$$

The columns of this matrix are given by the coordinates of the vectors of the new basis $e^{'}$ with respect to the old basis e!

Remarks

- In the matrix notations we have

$$\begin{pmatrix} e_1^{'} \\ e_2^{'} \\ \ldots \\ e_n^{'} \end{pmatrix} = A \begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix} \text{ or } \ (e^{'})_{1,n} = (P^{(e,e^{'})})^\top (e)_{1,n}$$

- Consider the change of the basis from $\mathfrak{B}$ to $\mathfrak{B}^{'}$ with the matrix $P^{(e,e^{'})}$ and the change of the basis from $\mathfrak{B}^{'}$ to $\mathfrak{B}^{''}$ with the matrix $P^{(e^{'},e^{''})}$. We can think at the "composition" of these two changes, i.e. the change of the basis from $\mathfrak{B}$ to $\mathfrak{B}''$ with the matrix $P^{(e,e'')}$. It is easy to see that one has

$$P^{(e,e')}P^{(e',e'')} = P^{(e,e'')}.$$

- If in the above discussion we consider $\mathfrak{B}'' = \mathfrak{B}$ one has

$$P^{(e,e')}P^{(e',e)} = I_n \ ,$$

that is

$$(P^{(e',e)})^{-1} = P^{(e,e')}.$$

At this step we try to answer the next question, which is important in applications. If we have two bases, a vector can be represented in both of them. What is the relation between the coordinates in the two bases?
Let us fix the setting first. Consider the vector space V, with two bases $\mathfrak{B} = \{e_1, \dots, e_n\}$ and $\mathfrak{B}' = \{e_1', \dots, e_n'\}$ and $P^{(e,e')}$ the matrix of the change of basis.
Let $v \in V$. We have

$$v = a_1e_1 + \cdots + a_ne_n = b_1e_1' + \cdots + b_ne_n',$$

where $(a_1, \dots a_n)$ and $(b_1, \dots b_n)$ are the coordinates of the same vector in the two bases. We can write

$$(v) = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \cdot \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} \cdot \begin{pmatrix} e_1' \\ e_2' \\ \dots \\ e_n' \end{pmatrix}.$$

Denote

$$(v)_e = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix}$$

and

$$(v)_{e'} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix}$$

the matrices of the coordinates of v in the two bases.

Denote further the basis columns

$$(e)_{1n} = \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}$$

the column matrix of the basis $\mathfrak{B}$ and

$$(e^{'})_{1n} = \begin{pmatrix} e^{'}_1 \\ e^{'}_2 \\ \dots \\ e^{'}_n \end{pmatrix}$$

the matrix column of the basis $\mathfrak{B}^{'}$, we have

$$v = (v)_e^{\top}(e)_{1n} = (v)_{e'}^{\top}(e^{'})_{1n} = (v)_{e'}^{\top}(P^{(e,e^{'})})^{\top}(e)_{1n}$$

Because v is uniquely represented in a basis it follows

$$(v)_{e'}^{\top}(P^{(e,e^{'})})^{\top} = (v)_e^{\top} \ ,$$

or

$$(v)_{e'} = (P^{(e,e')})^{-1}(v)_e = P^{(e',e)}(v)_e \ .$$

Hence,

$$(v)_e = (P^{(e,e')})(v)_{e'} \ .$$

2.5 Problems

Problem 2.5.1. Show that for $\operatorname{span}(v_1, \ldots, v_n) = V$ one has $\operatorname{span}(v_1 - v_2, v_2 - v_3, \ldots, v_{n-1} - v_n, v_n) = V$.

Problem 2.5.2. Find a basis for the subspace generated by the given vectors in $\mathcal{M}_3(\mathbb{R})$.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 2 \\ 2 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ -2 & 2 & -1 \\ -1 & 2 & 1 \end{pmatrix}.$$

Problem 2.5.3. Let V be a finite dimensional vector space $\dim V = n$. Show that there exist one dimensional subspaces $U_1, \ldots, U_n$, such that

$$V = U_1 \oplus \cdots \oplus U_n.$$

Problem 2.5.4. Find three distinct subspaces U, V, W of $\mathbb{R}^2$ such that

$$\mathbb{R}^2 = U \oplus V = V \oplus W = W \oplus U.$$

Problem 2.5.5. Let U, W be subspaces of $\mathbb{R}^8$, with $\dim U = 3$, $\dim W = 5$ and $\dim U + W = 8$. Show that $U \cap W = \{0\}$.

Problem 2.5.6. Let U, W be subspaces of $\mathbb{R}^9$ with $\dim U = \dim W = 5$. Show that $U \cap W \neq \{0\}$.

Problem 2.5.7. Let U and W be subspaces of a vector space V and suppose that each vector $v \in V$ has a unique expression of the form $v = u + w$ where u belongs to U and w to W. Prove that

$$V = U \oplus W.$$

Problem 2.5.8. In $C[a, b]$ find the dimension of the subspaces generated by the following sets of vectors:

a) $\{1, \cos 2x, \cos^2 x\}$,

b) $\{e^{a_1 x}, \dots, e^{a_n x}\}$, where $a_i \neq a_j$ for $i \neq j$

Problem 2.5.9. Find the dimension and a basis in the intersection and sum of the following subspaces:

- $U = \text{span}\{(2, 3, -1), (1, 2, 2,), (1, 1, -3)\}$,
  $V = \text{span}\{(1, 2, 1), (1, 1, -1), (1, 3, 3)\}$.
- $U = \text{span}\{(1, 1, 2, -1), (0, -1, -1, 2), (-1, 2, 1, -3\}$,
  $V = \text{span}\{(2, 1, 0, 1), (-2, -1, -1, -1), (3, 0, 2, 3)\}$.

Problem 2.5.10. Let U, V, W be subspaces of some vector space and suppose that $U \subseteq W$. Prove that

$$(U + V) \cap W = U + (V \cap W).$$

Problem 2.5.11. In $\mathbb{R}^4$ we consider the following subspace $V = \text{span}\{(2, 1, 0, 1), (-2, -1, -1, -1), (3, 0, 2, 3)\}$. Find a subspace W of $\mathbb{R}^4$ such that $\mathbb{R}^4 = V \oplus W$.

Problem 2.5.12. Let V, W be two vector spaces over the same field $\mathbb{F}$. Find the dimension and a basis of $V \times W$.

Problem 2.5.13. Find a basis in the space of symmetric, respectively skew-symmetric matrices of dimension n.

Problem 2.5.14. Let $V = \{(x_1, \ldots, x_n) \in \mathbb{R}^n |, x_1 + x_2 + \ldots + x_n = 0,\ x_1 + x_n = 0\}$. Find a basis in V.

Problem 2.5.15. Let $\mathcal{M}_n(\mathbb{R})$ be the set of the real square matrices of order n, and $\mathcal{A}_n$, respectively $\mathcal{S}_n$ the set of symmetric, respectively skew-symmetric matrices of order n. Show that $\mathcal{M}_n(\mathbb{R}) = \mathcal{A}_n \oplus \mathcal{S}_n$.

Problem 2.5.16. Let us denote by $\mathbb{R}_n[X]$ the set of all polynomials having degree at most n with real coefficients. Obviously $\mathbb{R}_n[X]$ is a subspace of $\mathbb{R}[X]$ with the induced operations. Find the dimension of the quotient space $\mathbb{R}_n[X]/U$ where U is the subspace of all real constant polynomials.

Problem 2.5.17. Let V be a finite-dimensional vector space and let U and W be two subspaces of V. Prove that

$$\dim((U + W)/W) = \dim(U/(U \cap W)).$$

Problem 2.5.18. Let us consider the matrix

$$M = \begin{pmatrix} 1 & 3 & 5 & -3 & 6 \\ 1 & 2 & 3 & 4 & 7 \\ 1 & 3 & 3 & -2 & 5 \\ 0 & 9 & 11 & -7 & 16 \\ 4 & 9 & 12 & 10 & 26 \end{pmatrix}.$$

Let U and W be the subspaces of $\mathbb{R}^5$ generated by rows 1, 2 and 5 of M, and by rows 3 and 4 of M respectively. Find the dimensions of $U + W$ and $U \cap W$.

Problem 2.5.19. Find bases for the sum and intersection of the subspaces U and W of $\mathbb{R}_4[X]$ generated by the respective sets of polynomials
$\{1 + 2x + x^3, 1 - x - x^2\}$ and $\{x + x^2 - 3x^3, 2 + 2x - 2x^3\}$.

3

Linear maps between vector spaces

Up to now we met with vector spaces. It is natural to ask about maps between them, which are compatible with the linear structure of a vector space. These are called linear maps, special maps which also transport the linear structure. They are also called morphisms of vector spaces or linear transformations.

Definition 3.1. *Let V and W be two vector spaces over the same field $\mathbb{F}$. A linear map from V to W is a map $f : V \to W$ which has the property that $f(\alpha v + \beta u) = \alpha f(v) + \beta f(u)$ for all $v, u \in V$ and $\alpha, \beta \in \mathbb{F}$.*

The class of linear maps between V and W will be denoted by $L_{\mathbb{F}}(V, W)$ or $\mathrm{Hom}_{\mathbb{F}}(V, W)$.

From the definition it follows that $f(0_V) = 0_W$ and

$$f(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i f(v_i),\ \forall\ \alpha_i \in \mathbb{F},\ \forall v_i \in V, i = \overline{1, n}.$$

We shall define now two important notions related to a linear map, the kernel and the image.

Consider the sets:

$$\ker f = f^{-1}(0_W) = \{v \in V | f(v) = 0_w\},\ \text{and}$$

$$\mathrm{im} f = f(V) = \{w \in W | \exists\ v \in V, f(v) = w\}.$$

Definition 3.2. *The sets* $\ker f$ *and* $f(V)$ *are called the kernel (or the null space), respectively the image of* f*.*

An easy exercise will prove the following:

Proposition 3.3. *The kernel and the image of a linear map* $f : V \to W$ *are subspaces of* V *and* W *respectively.*

Example 3.4. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be given by $(x, y) \mapsto (x + y, x + y)$. Find $\ker T$ and $T\left(\mathbb{R}^2\right)$.

By definition

$$\begin{aligned} \ker T &= \left\{(x, y) \in \mathbb{R}^2 | T(x, y) = (0, 0)\right\} \\ &= \left\{(x, y) \in \mathbb{R}^2 | \left(x + y, x + y\right) = (0, 0)\right\} \\ &= \left\{(x, y) \in \mathbb{R}^2 | x + y = 0\right\}. \end{aligned}$$

Geometrically, this is the straight line with equation $y = -x$. Clearly
$\ker T = \mathrm{span}\,\{(1, -1)\}$ and $\dim \ker T = 1$.
From the way T is defined we see that all vectors in the image $T\left(\mathbb{R}^2\right)$ of T, have both components equal to each other, so

$$\begin{aligned} T\left(\mathbb{R}^2\right) &= \{(\alpha, \alpha) | \alpha \in \mathbb{R}\} \\ &= \mathrm{span}\,\{(1, 1)\}. \end{aligned}$$

For the finite dimensional case the dimension of ker and im of a linear map between vector spaces are related by the following:

Theorem 3.5. *Let* $f : V \to W$ *be a linear map between vector spaces* V *and* W *over the field* $\mathbb{F}$*,* V *being finite dimensional.*

$$dim V = dim \ker f + dim \, \mathrm{im} f.$$

Proof. Let n and m be the dimensions of V and $\ker f$, $m \leq n$. Consider a basis $\{e_1, \ldots, e_m\}$ for $\ker f$. The independent system of vectors $e_1, \ldots, e_m$ can be completed to a basis $\{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$ of V.
Our aim is to prove that the vectors $f(e_{m+1}), \ldots, f(e_n)$ form a basis for $f(V)$. It is sufficient to prove that the elements $f(e_{m+1}), \ldots, f(e_n)$ are linearly independent since they generate $f(V)$.
Suppose the contrary, that $f(e_{m+1}), \ldots, f(e_n)$ are not linearly independent. There exist $\alpha_{m+1}, \ldots, \alpha_n \in \mathbb{F}$ such that

$$\sum_{k=m+1}^{n} \alpha_k f(e_k) = 0_W,$$

and by the linearity of f,

$$f(\sum_{k=m+1}^{n} \alpha_k e_k) = 0_W.$$

Hence

$$v' = \sum_{k=m+1}^{n} \alpha_k e_k \in \ker f$$

and v' can be written in terms of $e_1, \ldots, e_m$. This is only compatible with the fact that $e_1, \ldots, e_n$ form a basis of V if $\alpha_{m+1} = \cdots = \alpha_n = 0$, which implies the linear independence of the vectors $f(e_{m+1}), \ldots, f(e_n)$. $\square$

Theorem 3.6. *Let $f : V \to W$ be a linear mapping between vector spaces V and W, and $dim V = dim W < \infty$. Then, $f(V) = W$ iff $\ker f = \{0_V\}$. In particular f is onto iff f is one to one.*

Proof. Suppose that $\ker f = \{0_V\}$. Since $f(V)$ is a subspace of W it follows that $\dim V = \dim f(V) \leq \dim W$, which forces $\dim f(V) = \dim W$, and this implies that $f(V) = W$.

The fact that $f(V) = W$ implies that $\ker f = \{0_V\}$ follows by reversing the arguments. □

Proposition 3.7. *Let $f : V \to W$ be a linear map between vector spaces V, W over $\mathbb{F}$. If f is a bijection, it follows that its inverse $f^{-1} : W \to V$ is a linear map.*

Proof. Because f is a bijection $\forall w_1, w_2 \in W$, $\exists!\ v_1, v_2 \in V$, such that $f(v_i) = w_i, i = 1, 2$. Because f is linear, it follows that

$$\alpha_1 w_1 + \alpha_2 w_2 = \alpha_1 f(v_1) + \alpha_2 f(v_2) = f(\alpha_1 v_1 + \alpha_2 v_2).$$

It follows that $\alpha_1 v_1 + \alpha_2 v_2 = f^{-1}(\alpha_1 w_1 + \alpha_2 w_2)$, so

$$f^{-1}(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 f^{-1}(w_1 +) + \alpha_2 f^{-1}(w_2).$$

□

Definition 3.8. *A linear bijective map $f : V \to W$ between vector spaces V, W over $\mathbb{F}$ is called an isomorphism of the vector space V over W , or isomorphism between the vector spaces V and W.*
A vector space V is called isomorphic to a vector space W if there exists an isomorphism $f : V \to W$. The fact that the vector spaces V and W are isomorphic will be denoted by $V \simeq W$.

Example 3.9. Let V be an $\mathbb{F}$ vector space and V_1, V_2 two supplementary spaces, that is $V = V_1 \oplus V_2$. It follows that $\forall v \in V$ we have the unique decomposition $v = v_1 + v_2$, with $v_1 \in V_1$ and $v_2 \in V_2$. The map

$$p : V \to V_1,\ p(v) = v_1,\ \forall v \in V$$

is called the projection of V on V_1, parallel to V_2.

The map $s : V \to V,\ s(v) = v_1 - v_2,\ \forall v \in V$ is called the symmetry of V with respect to V_1, parallel with V_2.
It is easy to see that for $v \in V_1$, $v_2 = 0$, so $p(v) = v$ and $s(v) = v$, and for $v \in V_2$, $v_1 = 0$, so $p(v) = 0$ and $s(v) = -v$.

3.1 Properties of $L(V, W)$

In this section we will prove some properties of linear maps and of $L(V, W)$.

Proposition 3.10. *Let $f : V \to W$ be a linear map between the linear spaces V, W over $\mathbb{F}$.*

1. *If $V_1 \subseteq V$ is a subspace of V, then $f(V_1)$ is a subspace of W.*
2. *If $W_1 \subseteq W$ is a subspace of W, then $f^{-1}(W_1)$ is a subspace of V.*

Proof. 1. Let w_1, w_2 be in $f(V_1)$. It follows that there exist $v_1, v_2 \in V_1$ such that $f(v_i) = w_i, i = 1, 2$. Then, for every $\alpha, \beta \in \mathbb{F}$ we have

$$\alpha w_1 + \beta w_2 = \alpha f(v_1) + \beta f(v_2) = f(\alpha v_1 + \beta v_2) \in f(V_1).$$

2. For $v_1, v_2 \in f^{-1}(W_1)$ we have that $f(v_1), f(v_2) \in W_1$, so
$\forall\ \alpha, \beta \in \mathbb{F},\ \alpha f(v_1) + \beta f(v_2) \in W_1$. Because f is linear
$\alpha f(v_1) + \beta f(v_2) = f(\alpha v_1 + \beta v_2) \Rightarrow \alpha v_1 + \beta v_2 \in f^{-1}(W_1)$.

□

The next proposition shows that the kernel and the image of a linear map characterize the injectivity and surjectivity properties of the map.

Proposition 3.11. *Let $f : V \to W$ be a linear map between the linear spaces V, W.*

1. *f is one to one (injective)* $\Longleftrightarrow$ $\ker f = \{0\}$.

2. *f is onto (surjective)* $\Longleftrightarrow$ $f(V) = W$.

3. *f is bijective* $\Longleftrightarrow$ $\ker f = \{0\}$ *and* $f(V) = W$.

Proof. 1 Suppose that f is one to one. Because $f(0_V) = 0_W$ it follows that $\ker f = \{0_V\} \subset V$. For the converse, suppose that $\ker f = \{0_V\}$. Let $v_1, v_2 \in V$ with $f(v_1) = f(v_2)$. It follows that $f(v_1 - v_2) = 0$ and because $\ker f = \{0\}$ we have that $v_1 = v_2$. The claims 2. and 3. can be proved in the same manner. $\square$

Next we shall study how special maps act on special systems of vectors.

Proposition 3.12. *Let $f : V \to W$ be a linear map between the linear spaces V, W and $S = \{v_i | i \in I\}$ a system of vectors in V.*

1. *If f is one to one and S is linearly independent, then $f(S)$ is linearly independent.*

2. *If f is onto and S is a system of generators, then $f(S)$ is s system of generators.*

3. *If f is bijective and S is a basis for V, then $f(S)$ is a basis for W.*

Proof. 1. Let $\{w_1, \ldots, w_n\}$ be a finite subsystem from $f(S)$, and $\alpha_i \in \mathbb{F}$ with $\sum_{i=1}^n \alpha_i w_i = 0$. There exist the vectors $v_i \in V$ such that $f(v_i) = w_i$, for all $i \in \{1, \ldots, n\}$. Then $\sum_{i=1}^n \alpha_i w_i = \sum_{i=1}^n \alpha_i f(v_i) = f(\sum_{i=1}^n \alpha_i v_i) = 0$, so $\sum_{i=1}^n \alpha_i v_i = 0$. Because S is linearly independent it follows that $\alpha_i = 0$ for all $i = \overline{1,n}$, so $f(S)$ is linearly independent.

2. Let $w \in W$. There exists $v \in V$ with $f(v) = w$. Because S is a system of generators, there exists a finite family of vectors in S, v_i, and the scalars

$\alpha_i \in \mathbb{F}, i = \overline{1,n}$ such that $\sum_{i=1}^n \alpha_i v_i = v$. It follows that

$$w = f(v) = f(\sum_{i=1}^n \alpha_i v_i) = \sum_{i=1}^n \alpha_i f(v_i).$$

3. Because f is bijective and S is a basis for V, it follows that both 1. and 2. hold, that is, $f(S)$ is a basis for W. $\square$

Definition 3.13. *Let $f, g : V \to W$ be linear maps between the linear spaces V and W over $\mathbb{F}$, and $\alpha \in \mathbb{F}$. We define*

1. *$f + g : V \to W$ by $(f+g)(v) = f(v) + g(v),\ \forall\ v \in V$, the sum of the linear maps, and*

2. *$\alpha f : V \to W$ by $(\alpha f)(v) = \alpha f(v),\ \forall\ v \in V, \forall\ \alpha \in \mathbb{F}$, the scalar multiplication of a linear map.*

Proposition 3.14. *With the operations defined above $L(V,W)$ becomes a vector space over $\mathbb{F}$.*

The proof of this statement is an easy verification.
In the next part we specialize in the study of the linear maps, namely we consider the case $V = W$.

Definition 3.15. *The set of endomorphisms of a linear space V is:*

$$End(L) = \{f : V \to V |\ f \text{ linear } \}.$$

By the results from the previous section, $End(V)$ is an $\mathbb{F}$ linear space.
Let W, U be two other linear spaces over the same field $\mathbb{F}$, $f \in L(V,W)$ and $g \in L(W,U)$. We define the product (composition) of f and g by
$h = g \circ f : V \to U$,

$$h(v) = g(f(v)),\ \forall\ v \in V.$$

Proposition 3.16. *The product of two linear maps is a linear map.*
Moreover, if f and g as above are isomorphisms, then the product $h = g \circ f$ is an isomorphism.

Proof. We check that for all $v_1, v_2 \in V$ and all $\alpha, \beta \in \mathbb{F}$

$$\begin{aligned} h(\alpha v_1 + \beta v_2) &= g(f(\alpha v_1 + \beta v_2)) \\ &= g(\alpha f(v_1) + \beta f(v_2)) \\ &= g(\alpha f(v_1)) + g(\beta f(v_2)) \\ &= \alpha h(v_1) + \beta h(v_2). \end{aligned}$$

The last statement follows from the fact that h is a linear bijection. $\square$

It can be shown that the composition is distributive with respect to the sum of linear maps, so $End(V)$ becomes an unitary ring.
It can easily be realized that:

Proposition 3.17. *The isomorphism between two linear spaces is an equivalence relation.*

Definition 3.18. *Let V be an $\mathbb{F}$ linear space. The set*

$$Aut(V) = \{f \in End(V) |\ f \text{ isomorphism } \}$$

is called the set of automorphisms of the vector space V.

Proposition 3.19. *$Aut(V)$ is a group with respect to the composition of linear maps.*

Proof. It is only needed to list the properties.

1. the identity map I_V is the unit element.

2. $g \circ f$ is an automorphism for f and g automorphisms.

3. the inverse of an automorphism is an automorphism.

$\square$

The group of automorphisms of a linear space is called the general linear group and is denoted by $GL(V)$.

Example 3.20. • **Projectors endomorphisms.** An endomorphism $p : V \to V$ is called projector of the linear space V iff

$$p^2 = p,$$

where $p^2 = p \circ p$. If p is a projector, then:

1. $\ker p \oplus p(V) = V$

2. the endomorphism $q = I_V - p$ is again a projector.

Denote $v_1 = p(v)$ and $v_2 = v - v_1$, it follows that $p(v_2) = p(v) - p(v_1) = p(v) - p^2(v) = 0_V$, so $v_2 \in \ker f$. Hence

$$v = v_1 + v_2, \ \forall\, v \in V,$$

where $v_1, v_2 \in f(V)$ and, moreover, the decomposition is unique, so we have the direct sum decomposition $\ker p \oplus p(V) = V$. For the last assertion simply compute $q^2 = (I_V - p) \circ (I_V - p) = I_V - p - p + p^2 = I_V - p = q$, because p is a projector. It can be seen that $q(V) = \ker p$ and $\ker q = q(V)$. Denote by $V_1 = p(V)$ and $V_2 = \ker p$. It follows that p is the projection of V on V_1, parallel with V_2, and q is the projection of V on V_2 parallel with V_1.

• **Involutive automorphisms.** An operator $s : V \to V$ is called involutive iff $s^2 = I_V$. From the definition and the previous example one has:

1. an involutive operator is an automorphism
2. for every involutive automorphism, the linear operators:

$$\begin{aligned} p_s &: V \to V, p_s(v) = \frac{1}{2}(v + s(v)) \\ q_s &: V \to V, q_s(v) = \frac{1}{2}(v - s(v)) \end{aligned}$$

   are projectors and satisfy the relation $p_s + q_s = 1_V$.

3. reciprocally, for a projector $p : V \to V$, the operator $s_p : V \to V$, given by $s_p(v) = 2p(v) - v$ is an involutive automorphism.

From the previous facts it follows that $p_s \circ s = s \circ p_s = p, s_p \circ p = p \circ s_p = p$. An involutive automorphism s is a symmetry of V with respect to the subspace $p_s(V)$, parallel with the subspace $\ker p_s$.

Example 3.21. Let V be a vector space and $f : V \to V$ a linear map such that $\ker f = \mathrm{im} f$. Determine the set $\mathrm{im} f^2$, where f^2 denotes the composition of f with itself, $f^2 = f \circ f$.

We start by writing down explicitly

$$\begin{aligned} \mathrm{im} f^2 &= \mathrm{im} f \circ f \\ &= f \circ f\,(V) \\ &= f\,(f\,(V))\,. \end{aligned}$$

But, $f\,(V) = \mathrm{im} f = \ker f$ is the set of all vectors which are mapped by f to zero, so

$$\begin{aligned} \mathrm{im} f^2 &= f(\ker f) \\ &= 0. \end{aligned}$$

3.2 Local form of a linear map

Let V and W be two vector spaces over the same filed $\mathbb{F}$, $\dim V = m$, $\dim W = n$, and $e = \{e_1, \ldots, e_m\}$ and $f = \{f_1, \ldots, f_n\}$ be bases in V and W respectively. A linear map $T \in L(V, W)$ is uniquely determined by the values ond the basis e.
We have

$$\begin{aligned} T(e_1) &= a_{11} f_1 + \cdots + a_{1n} f_n, \\ T(e_2) &= a_{21} f_1 + \cdots + a_{2n} f_n, \\ &\vdots \\ T(e_m) &= a_{m1} f_1 + \cdots + a_{mn} f_n, \end{aligned}$$

or, in the matrix notation

$$\begin{pmatrix} T(e_1) \\ T(e_2) \\ \vdots \\ T(e_m) \end{pmatrix} = A \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \text{ where } A = (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}.$$

The transposed of A is denoted by $M_T^{(f,e)}$ and is called the matrix of the linear map T is the bases e and f.
From the definition of the matrix of a linear map it follows that:

Theorem 3.22.
- *For $T_1, T_2 \in L(V, W)$ and $a_1, a_2 \in \mathbb{F}$*

$$M_{a_1 T_1 + a_2 T_2} = a_1 M_{T_1} + a_2 M_{T_2}$$

- *The vector space $L(V, W)$ is isomporphic to $\mathcal{M}_{m,n}(\mathbb{F})$ by the map $T \in L(V, W) \mapsto M_T(\mathbb{F}) \in \mathcal{M}_{m,n}(\mathbb{F})$.*

- *Particularly $End(V)$ is isomorphic to $\mathcal{M}_n(\mathbb{F})$.*

Now we want to see how the image of a vector by a linear map can we expressed. Let $v \in V$, $v = \sum_{i=1}^{m} v_i e_i$, or in the matrix notation $(v)_e^\top (e)_{1m}$, where, as usual

$$(v)_e = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

and

$$(e)_{1m} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{pmatrix}.$$

Now denote $T(v) = w = \sum_{j=1}^{n} w_j e_j \in W$, we have

$$T(v) = (w)_f^\top (f)_{1n}.$$

T being linear, we have $T(v) = \sum_{i=1}^{m} v_i T(e_i)$, or, again in matrix notation:

$$T(v) = (v)_e^\top (T(e))_{1m}.$$

From the definition of $M_T^{(f,e)}$ it follows that

$$(T(e))_{1m} = (M_T^{(f,e)})^\top (f)_{1n}.$$

So finally we have

$$(w)_f^\top (f)_{1n} = (v)_e^\top (M_T^{(f,e)})^\top (f)_{1n}.$$

By the uniqueness of the coordinates of a vector in a basis it follows that

$$(w)_f^\top = (v)_e^\top (M_T^{(f,e)})^\top.$$

Taking the transposed of the above relation we get

$$(w)_f = (M_T^{(f,e)})(v)_e.$$

Example 3.23. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T = \begin{pmatrix} -3 & 0 & 2 \\ 1 & 1 & 0 \\ -2 & 1 & 2 \end{pmatrix}$. Find a basis in $\ker T$ and find the dimension of $T\left(\mathbb{R}^3\right)$.

Observe that the kernel of T,

$$\ker T = \left\{ (x, y, z) \in |T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

is the set of solutions of the linear homogeneous system

$$\left\{ \begin{array}{rcrcrcl} -3x & & & + & 2z & = & 0 \\ x & + & y & & & = & 0 \\ -2x & + & y & + & 2z & = & 0, \end{array} \right. \tag{3.1}$$

the matrix of the system being exactly T. To solve this system we need to compute the rank of the matrix T. We get that

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

and that $\operatorname{rank} A = 2$. To solve the system we chose $x = \alpha$ as a parameter and express y and z in terms of x from the first two equations to get

$$x = \alpha, \quad y = -x, \quad z = \frac{3}{2}x.$$

The set of solutions is

$$\left\{ \left(\alpha, -\alpha, \frac{3}{2}\alpha \right) |\alpha \in \mathbb{R} \right\} = \operatorname{span} \left\{ \left(1, -1, \frac{3}{2} \right) \right\}$$

and hence, a basis in $\ker T$ consists only of $\left(1,-1,\frac{3}{2}\right)$, $\dim \ker T = 1$.
Based on the dimension formula

$$\dim \ker T + \dim T\left(\mathbb{R}^3\right) = \dim \mathbb{R}^3$$

we infer that $\dim T\left(\mathbb{R}^3\right) = 2$.

Proposition 3.24. *Let V, W, U be vector spaces over $\mathbb{F}$, of dimensions m, n, p, and $T \in L(V,W)$, $S \in L(W,U)$, with matrices M_T and M_S, in some basis. Consider the composition map $S \circ T : V \to U$ with the matrix $M_{S\circ T}$. Then*

$$M_{S\circ T} = M_s M_T.$$

Proof. Indeed, one can easily see that for $v \in V$ we have $(T(v)) = M_T(v)$ where $(T(v))$, respectively (v) stand for the coordinate of $T(v)$, respectively v in the appropriate bases. Similarly, for $w \in W$ one has $(S(w)) = M_S(w)$.
Hence, $(S \circ T(v)) = (S(T(v))) = M_S(T(v)) = M_S M_T(v)$, or, equivalently

$$M_{S\circ T} = M_S M_T.$$

□

Let V and W be vector spaces and $T \in L(V,W)$ be a linear map. In V and W we consider the bases $e = \{e_1, \ldots, e_m\}$ and $f = \{f_1, \ldots, f_n\}$, with respect to these bases the linear map has the matrix $M_T^{(f,e)}$. If we consider two other bases $e' = \{e'_1, \ldots, e'_m\}$ and $f' = \{f'_1, \ldots, f'_n\}$ the matrix of T with respect to these bases will be $M_T^{(f',e')}$. What relation do we have between the matrices of the same linear map in these two bases?

Theorem 3.25. *In the above conditions $M_T^{(f',e')} = P^{(f',f)} M_T^{(f,e)} P^{(e,e')}$.*

Proof. Let us consider $v \in V$ and let $w = T(v)$. We have

$$(w)_{f'} = M_T^{(f',e')}(v)_{e'} = M_T^{(f',e')} P^{(e',e)}(v)_e.$$

On the other hand

$$(w)_{f'} = P^{(f',f)}(w)_f = P^{(f',f)}(T(v))_f = P^{(f',f)} M_T^{(f,e)}(v)_e.$$

Taking into account that $(P^{(e',e)})^{-1} = P^{(e,e')}$ we get

$$M_T^{(f',e')} = P^{(f',f)} M_T^{(f,e)} (P^{(e',e)})^{-1} = P^{(f',f)} M_T^{(f,e)} P^{(e,e')}.$$

$\square$

Corollary 3.26. *Let e and e' be two bases of a finite-dimensional vector space V and let $T : V \to V$ be a linear mapping. If T is represented by matrices $A = M_T^{(e,e)}$ and $A' = M_T^{(e',e')}$ with respect to e and e' respectively, then $A' = PAP^{-1}$ where P is the matrix representing the change of basis e to e'.*

3.3 Problems

Problem 3.3.1. Consider the following mappings $T : \mathbb{R}^3 \to \mathbb{R}^3$. Study which one of them is a linear mapping.

a) $T(x_1, x_2, x_3) = (x_1^2, x_2, x_3^2)$.

b) $T(x_1, x_2, x_3) = (x_3, x_1, x_2)$.

c) $T(x_1, x_2, x_3) = (x_1 - 1, x_2, x_3)$.

d) $T(x_1, x_2, x_3) = (x_1 + x_2, x_2 - x_3, x_1 + x_2 + x_3)$.

e) $T(x_1, x_2, x_3) = (x_3, 0, 0)$.

f) $T(x_1, x_2, x_3) = (x_1, 2x_2, 3x_3)$.

Problem 3.3.2. Let $T \in End(V)$ and let $\{e_i : i = \overline{1, n}\}$ be a basis in V. Prove that the following statements are equivalent.

1. The matrix of T, with respect to the basis $\{e_i : i = \overline{1, n}\}$ is upper triangular.

2. $T(e_k) \in \text{span}\{e_1, \dots, e_k\}$ for all $k = \overline{1, n}$.

3. $T(\text{span}\{e_1, \dots, e_k\}) = \text{span}\{e_1, \dots, e_k\}$ for all $k = \overline{1, n}$.

Problem 3.3.3. Let $T_1, T_2 : \mathbb{R}^3 \to \mathbb{R}^3$ having the matrices

$$M_{T_1} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 3 \end{pmatrix},$$

respectively

$$M_{T_2} = \begin{pmatrix} -1 & 4 & 2 \\ 0 & 4 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

in the canonical basis of $\mathbb{R}^3$.

a) Find the image of $(0, 1, -1)$ through $T_1, T_1^{-1}, T_2, T_2^{-1}$.

b) Find the image of $(1, 3, -2)$ through $T_1 + T_2, (T_1 + T_2)^{-1}$.

c) Find the image of $(1, 2, 0)$ through $T_1 \circ T_2, T_2 \circ T_1$.

Problem 3.3.4. Let V be a complex vector space and let $T \in End(V)$. Show that there exists a basis in V such that the matrix of T relative to this basis is upper triangular.

Problem 3.3.5. Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be a linear mapping represented by the matrix

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

Find a basis in $\ker T, \operatorname{im} T$ and the dimension of the spaces $V, W, \ker T$ and $\operatorname{im} T$.

Problem 3.3.6. Show that a linear transformation $T : V \to W$ is injective if and only if it has the property of mapping linearly independent subsets of V to linearly independent subsets of W.

Problem 3.3.7. Show that a linear transformation $T : V \to W$ is surjective if and only if it has the property of mapping any set of generators of V to a set of generators of W.

Problem 3.3.8. Let $T : V \to W$ be a linear mapping represented by the matrix

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

Compute $\dim V, \dim W$ and find a basis in $\operatorname{im} T$ and $\ker T$.

Problem 3.3.9. Find all the linear mappings $T : \mathbb{R} \to \mathbb{R}$ with the property $\operatorname{im} T = \ker T$.

Find all $n \in \mathbb{N}$ such that there exists a linear mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ with the property $\operatorname{im} T = \ker T$.

Problem 3.3.10. Let V, respectively V_i, $i = \overline{1,n}$ be vector spaces over $\mathbb{C}$. Show that, if $T : V_1 \times V_2 \times \cdots \times V_n \to V$ is a linear mapping then there exist and they are unique the linear mappings $T_i : V_i \to V$, $i = \overline{1,n}$ such that

$$T(v_1, \ldots, v_n) = T_1(v_1) + T_2(v_2) + \cdots + T_n(v_n).$$

Problem 3.3.11. (The first isomorphism theorem). If $T : V \to W$ is a linear transformation between vector spaces V and W, then

$$V/\ker T \simeq \operatorname{im} T.$$

[Hint: show that the mapping $S : V/\ker T \to \operatorname{im} T$, $S(v + \ker T) = T(v)$ is a bijective linear mapping.]

Problem 3.3.12. (The second isomorphism theorem). If U and W are subspaces of a vector space V, then

$$(U + W)/W \simeq U/(U \cap W).$$

[Hint: define the mapping $T : U \to (U + W)/W$ by the rule $T(u) = u + W$, show that T is a linear mapping and use the previous problem.]

Problem 3.3.13. (The third isomorphism theorem). Let U and W be subspaces of a vector space V such that $W \subseteq U$. Prove that U/W is a subspace of V/W and that $(V/W)/(U/W) \simeq V/U$.
[Hint: define a mapping $T : V/W \to V/U$ by the rule $T(v + W) = v + U$, show that T is a linear mapping and use the firs isomorphism theorem.]

Problem 3.3.14. Show that every subspace U of a finite-dimensional vector space V is the kernel and the image of suitable linear operators on V.

Problem 3.3.15. Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ having the matrix

$$M_T = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & 0 & -1 & 2 \\ 2 & 5 & 3 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix}$$

in the canonical basis $\{e_1, e_2, e_3, e_4\}$ of $\mathbb{R}^4$.
Find the matrix of T with respect to the following basis.

a) $\{e_1, e_3, e_2, e_4\}$.

b) $\{e_1, e_1 + e_2, e_1 + e_2 + e_3, e_1 + e_2 + e_3 + e_4\}$.

c) $\{e_4 - e_1, e_3 + e_4, e_2 - e_4, e_4\}$.

Problem 3.3.16. A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $T(x_, x_2, x_3) = (x_1 - x_2 - x_3, -x_1 + x_3)$. Let $e = \{(2, 0, 0), (-1, 2, 0), (1, 1, 1)\}$ and $f = \{(0, -1), (1, 2)\}$ be bases in $\mathbb{R}^3$ and $\mathbb{R}^2$ respectively. Find the matrix that represents T with respect to these bases.

4

Proper vectors and the Jordan canonical form

4.1 Invariant subspaces. Proper vectors and values

In this part we shall further develop the theory of linear maps. Namely we are interested in the structure of an operator.

Let us begin with a short description of what we expect to obtain.

Suppose that we have a vector space V over a field $\mathbb{F}$ and a linear operator $T \in End(V)$. Suppose further that we have the direct sum decomposition:

$$V = \bigoplus_{i=1}^{m} U_i,$$

where each U_i is a direct subspace of V. To understand the behavior of T it is only needed to understand the behavior of each restriction $T|_{U_j}$. Studying $T|_{U_j}$ should be easier than dealing with T because U_j is a "smaller" vector space than V.

However we have a problem: if we want to apply tools which are commonly used in the theory of linear maps (such as taking powers for example) the problem is that generally T may not map U_j into itself, in other words $T|_{U_j}$ may not be an operator on U_j. For this reason it is natural to consider only that kind of decomposition for which T maps every U_j into itself.

Definition 4.1. *Let V be an operator on the vector space V over $\mathbb{F}$ and U a subspace of V. The subspace U is called invariant under T if $T(U) \subset U$, in other words $T|_U$ is an operator on U.*

Of course that another natural question arises when dealing with invariant subspaces. How does an operator behave on an invariant subspace of dimension one? Every one dimensional subspace is of the form $U = \{\lambda u | \lambda \in \mathbb{F}\}$. If U is invariant by T it follows that $T(u)$ should be in U, and hence there should exist a scalar $\lambda \in \mathbb{F}$ such that $T(u) = \lambda u$. Conversely if a nonzero vector u exists in V such that $T(u) = \lambda u$, for some $\lambda \in \mathbb{F}$, then the subspace U spanned by u is invariant under T and for every vector v in U one has $T(v) = \lambda v$. It seems reasonable to give the following definition:

Definition 4.2. *Let $T \in End(V)$ be an operator on a vector space over the field $\mathbb{F}$. A scalar $\lambda \in \mathbb{F}$ is called eigenvalue (or proper value) for T if there exists a nonzero vector $v \in V$ such that $T(v) = \lambda v$. A corresponding vector satisfying the above equality is called eigenvector (or proper vector) associated to the eigenvalue λ.*

The set of eigenvectors of T corresponding to an eigenvalue λ forms a vector space, denoted by $E(\lambda)$, the proper subspace corresponding to the proper value λ. It is clear that $E(\lambda) = \ker(T - \lambda I_V)$

For the finite dimensional case let M_T be the matrix of T in some basis. The equality $T(v) = \lambda v$ is equivalent to $M_T v = \lambda v$, or $(M_T - \lambda I_n)v = 0$, which is a linear system. Obviously this homogeneous system of linear equations has a nontrivial solution if and only if

$$\det(M_T - \lambda I_n) = 0.$$

Observe that $\det(M_T - \lambda I_n)$ is a polynomial of degree n in λ, where $n = \dim V$. This polynomial is called the characteristic polynomial of the operator T. Hence, the eigenvalues of T are the roots of its characteristic polynomial.
Notice that the characteristic polynomial does not depend on the choice of the basis B that is used when computing the matrix M_T of the transformation T. Indeed, let B' be another basis and M'_T the matrix of T with respect to this new basis. Further, let P be transition matrix from B to B'. So $M'_T = P^{-1}M_TP$ and $\det(P) \neq 0$. We have

$$\begin{aligned}
\det(P^{-1}M_TP - \lambda I) &= \det(P^{-1}M_TP - P^{-1}(\lambda I)P) \\
&= \det(P^{-1}(M_T - \lambda I)P) \\
&= \frac{1}{\det(P)} \det(M_T - \lambda I)\det(P) \\
&= \det(M_T - \lambda I),
\end{aligned}$$

which proves our claim.

Theorem 4.3. *Let $T \in End(V)$. Suppose that $\lambda_i, i = \overline{1,m}$ are distinct eigenvalues of T, and $v_i, i = \overline{1,m}$ are the corresponding proper vectors. The set $\{v_1, \ldots, v_m\}$ is linearly independent.*

Proof. Suppose, by contrary, that the set $\{v_1, \ldots, v_m\}$ is linearly dependent. It follows that a smallest index k exists such that

$$v_k \in \operatorname{span}\{v_1, \ldots, v_{k-1}\}.$$

Thus the scalars $a_1, \ldots a_{k-1}$ exist such that

$$v_k = a_1 v_1 + \ldots a_{k-1} v_{k-1}.$$

Applying T to the above equality, we get

$$\lambda_k v_k = a_1 \lambda_1 v_1 + \ldots a_{k-1} \lambda_{k-1} v_{k-1}.$$

It follows that

$$0 = a_1(\lambda_k - \lambda_1)v_1 + \cdots + a_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1}.$$

Because we choose k to be the smallest index such that $v_k = a_1 v_1 + \cdots + a_{k-1} v_{k-1}$, it follows that the set $\{v_1, \cdots v_{k-1}\}$ is linearly independent. It follows that all the a's are zero. □

Corollary 4.4. *An operator T on a finite dimensional vector space V has at most $dim V$ distinct eigenvalues.*

Proof. This is an obvious consequence of the fact that in a finite dimensional vector space we have at most $\dim V$ linearly independent vectors. □

The linear maps which have exactly $n = \dim V$ linearly independent eigenvectors have very nice and simple properties. This is the happiest case we can meet with in the class of linear maps.

Definition 4.5. *A linear map $T : V \to V$ is said to be diagonalizable if there exists a basis of V consisting of n independent eigenvectors, $n = dim V$.*

Recall that matrices A and B are similar if there is an invertible matrix P such that $B = PAP^{-1}$. Hence, a matrix A is diagonalizable if it is similar to a diagonal matrix D.

4.2 The minimal polynomial of an operator

The main reason for which there exists a richer theory of operators than for linear maps is that operators can be raised to powers (we can consider the composition of an operator with itself).

Let V be an n-dimensional vector space over a field $\mathbb{F}$ and $T : V \to V$ be a linear operator.

Now, $L(V, V) = End(V)$ is an n^2 dimensional vector space. We can consider $T^2 = T \circ T$ and of course we obtain $T^n = T^{n-1} \circ T$ inductively. We define T^0 as being the identity operator $I = I_V$ on V. If T is invertible (bijective), then there exists T^{-1}, so we define $T^{-m} = (T^{-1})^m$. Of course that

$$T^m T^n = T^{m+n}, \quad \text{for} \quad m, n \in \mathbb{Z}.$$

For $T \in End(V)$ and $p \in \mathbb{F}[X]$ a polynomial given by

$$p(z) = a_0 + a_1 z + \dots a_m z^m \ , \quad z \in \mathbb{F}$$

we define the operator $p(T)$ given by

$$p(T) = a_0 I + a_1 T + \dots a_m T^m.$$

This is a new use of the same symbol p, because we are applying it to operators not only to elements in $\mathbb{F}$. If we fix the operator T we obtain a function defined on $\mathbb{F}[X]$ with values in $End(V)$, given by $p \to p(T)$ which is linear. For $p, q \in \mathbb{F}[X]$ we define the operator pq given by $(pq)(T) = p(T)q(T)$.

Now we begin the study of the existence of eigenvalues and of their properties.

Theorem 4.6. *Every operator over a finite dimensional, nonzero, complex vector space has an eigenvalue.*

Proof. Suppose V is a finite dimensional complex vector space and $T \in End(V)$. Choose $v \in V$, $v \neq 0$. Consider the set

$$(v, T(v), T^2(v), \ldots T^n(v)).$$

This set is a linearly dependent system of vectors (they are $n+1$) vectors and $\dim V = n$. Then there exist complex numbers, $a_0, \ldots a_n$, not all 0, such that

$$0 = a_0 v + a_1 T(v) + \cdots + a_n T^n(v) \ .$$

Let m be the largest index such that $a_m \neq 0$. Then we have the decomposition

$$a_0 + a_1 z + \cdots + a_m z^m = a_0 (z - \lambda_1) \ldots (z - \lambda_m) \ .$$

It follows that

$$\begin{aligned} 0 &= a_0 v + a_1 T(v) + \ldots a_n T^n(v) \\ &= (a_0 I + a_1 T + \ldots a_n T^n)(v) \\ &= a_0 (T - \lambda_1 I) \ldots (T - \lambda_m I)(v) \ . \end{aligned}$$

which means that $T - \lambda_j I$ is not injective for at least one j, or equivalently T has an eigenvalue. □

Remark 4.7. The analogous statement is not true for real vector spaces. But on real vector spaces there are always invariant subspaces of dimension 1 or 2.

Example 4.8. Let $T : \mathbb{F}^2 \to \mathbb{F}^2$ given by $T(x,y) = (-y, x)$. It has no eigenvalues and eigenvectors if $\mathbb{F} = \mathbb{R}$. Find them for $\mathbb{F} = \mathbb{C}$.

Obviously, $T(x,y) = \lambda(x,y)$ leads to $(-y, x) = \lambda(x,y)$, or equivalently

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

The previous system is equivalent to $x = \lambda y$, $(\lambda^2 + 1)y = 0$.
If $\lambda \in \mathbb{R}$ then the solution is $x = y = 0$, but note that $(0, 0)$ is excluded from eigenvectors by definition.
If $\lambda \in \mathbb{C}$ we obtain the eigenvalues $\lambda_1 = i$, $\lambda_2 = -i$ and the corresponding eigenvectors $(i, 1) \in \mathbb{C}^2$, respectively $(-i, 1) \in \mathbb{C}^2$.

Theorem 4.9. *Every operator on an odd dimensional real vector space has an eigenvalue.*

Proof. Let $T \in End(V)$ and $n = \dim V$ odd. The eigenvalues of T are the roots of the characteristic polynomial that is $\det(M_T - \lambda I_n)$. This polynomial is a polynomial of degree n in λ, hence, since n is odd, the equation $\det(M_T - \lambda I_n) = 0$ has at least one real solution. $\square$

A central goal of linear algebra is to show that a given operator $T \in End(V)$ has a reasonably simple matrix in a given basis. It is natural to think that reasonably simple means that the matrix has as many $0's$ as possible.
Recall that for a basis $\{e_k, k = \overline{1, n}\}$,

$$T(e_k) = \sum_{i=1}^{n} a_{ik} e_i \ ,$$

where $M_T = (a_{ij})_{\substack{i=\overline{1,m} \\ j=\overline{1,n}}}$ is the matrix of the operator.

Theorem 4.10. *Suppose $T \in End(V)$ and $\{e_i, i = \overline{1, n}\}$ is a basis of V. Then the following statements are equivalent:*

1. *The matrix of T with respect to the basis $\{e_i, i = \overline{1, n}\}$ is upper triangular.*
2. *$T(e_k) \in span\{e_1, \dots, e_k\}$ for $k = \overline{1, n}$.*
3. *$span\{e_1, \dots, e_k\}$ is invariant under T for each $k = \overline{1, n}$.*

Proof. 1⇔2 obviously follows from a moment's tought and the definition. Again 3⇒2. It remains only to prove that 2⇒3.
So, suppose 2 holds. Fix $k \in \{1, \dots, n\}$. From 2 we have

$$\begin{array}{lrr} T(e_1) \in & \mathrm{span}\{e_1\} \subseteq & \mathrm{span}\{e_1, \dots, e_k\} \\ T(e_2) \in & \mathrm{span}\{e_1, e_2\} \subseteq & \mathrm{span}\{e_1, \dots, e_k\} \\ \vdots & & \\ T(e_k) \in & \mathrm{span}\{e_1, \dots, e_k\} \subseteq & \mathrm{span}\{e_1, \dots, e_k\}. \end{array}$$

So, for v a linear combination of $\{e_1, \dots, e_k\}$ one has that

$$T(v) \in \mathrm{span}\{e_1, \dots, e_k\},$$

consequently 3. holds. □

Theorem 4.11. *Suppose that V is a complex vector space and $T \in End(V)$. Then there exists a basis of V such that T is an upper-triangular matrix with respect to this basis.*

Proof. Induction on the $\dim V$. Clearly this holds for $\dim V = 1$.
Suppose that $\dim V > 1$ and the result holds for all complex vector spaces of dimension smaller then the dimension of V. Let λ be an eigenvalue of T (it exists) and

$$U = \mathrm{im}(T - \lambda I).$$

Because $T - \lambda I$ is not surjective, $\dim U < \dim V$. Furthermore U is invariant under T, since for $u \in U$ there exists $v \in V$ such that $u = T(v) - \lambda v$, hence $T(u) = T(T(v)) - \lambda T(v) = (T - \lambda I)(w) \in U$ where $w = T(v)$.

So, $T|_U$ is an operator on U. By the induction hypothesis there is a basis $\{u_1, \ldots, u_m\}$ of U with respect to which $T|_U$ has an upper-triangular matrix. So, for each $j \in \{1, \ldots, m\}$ we have

$$T(u_j) = T|_U(u_j) \in \operatorname{span}\{u_1, \ldots, u_m\}.$$

Extend the basis $\{u_1, \ldots, u_m\}$ of U to a basis $\{u_1, \ldots u_m, v_1, \ldots v_n\}$ of V. For each $k = \overline{1, n}$

$$T(v_k) = (T - \lambda I)(v_k) + \lambda v_k.$$

By the definition of U, $(T - \lambda I)(v_k) \in U = \operatorname{span}\{u_1, \ldots, u_m\}$. Thus the equation above shows that

$$T(v_k) \in \operatorname{span}\{u_1, \ldots, u_m, v_k\}.$$

From this, in virtue of the previous theorem, it follows that T has an upper-triangular matrix with respect to this basis. $\square$

One of the good points of this theorem is that, if we have this kind of basis, we can decide if the operator is invertible by analysing the matrix of the operator.

Theorem 4.12. *Suppose $T \in End(V)$ has an upper triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal are non zero.*

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of V with respect to which T has the matrix

$$M_T = \begin{pmatrix} \lambda_1 & \ldots & & * \\ 0 & \lambda_2 & \ldots & \\ 0 & 0 & \ldots & \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

We will prove that T is not invertible iff one of the λ_k's equals zero. If $\lambda_1 = 0$, then $T(v_1) = 0$, so T is not invertible, as desired.

Suppose $\lambda_k = 0$, $1 < k \leq n$. The operator T maps the vectors $e_1, \ldots, e_{k-1}$ in $\text{span}\{e_1, \ldots, e_{k-1}\}$ and, because $\lambda_k = 0$, $T(e_k) \in \{e_1, \ldots, e_{k-1}\}$. So, the vectors $T(e_1), \ldots, T(e_k)$ are linearly dependent (they are k vectors in a $k-1$ dimensional vector space, $\text{span}\{e_1, \ldots, e_{k-1}\}$. Consequently T is not injective, and not invertible.

Suppose that T is not invertible. Then $\ker T \neq \{0\}$, so $v \in V$, $v \neq 0$ exists such that $T(v) = 0$. Let

$$v = a_1 e_1 + \cdots + a_n e_n$$

and let k be the largest integer with $a_k \neq 0$. Then

$$v = a_1 e_1 + \cdots + a_k e_k,$$

and

$$0 = T(v),$$

$$0 = T(a_1 e_1 + \cdots + a_k e_k),$$

$$0 = (a_1 T(e_1) + \cdots + a_{k-1} T(e_{k-1})) + a_k T(e_k).$$

The term $(a_1 T(e_1) + \cdots + a_{k-1} T(e_{k-1}))$ is in $\text{span}\{e_1, \ldots, e_{k-1}\}$, because of the form of M_T. Finally $T(e_k) \in \text{span}\{e_1 \ldots, e_{k-1}\}$. Thus when $T(e_k)$ is written as a linear combination of the basis $\{e_1, \ldots, e_n\}$, the coefficient of e_k will be zero. In other words, $\lambda_k = 0$. □

Theorem 4.13. *Suppose that $T \in End(V)$ has an upper triangular matrix with respect to some basis of V. Then the eigenvalues of T are exactly of the entries on the diagonal of the upper triangular matrix.*

Proof. Suppose that we have a basis $\{e_1, \ldots, e_n\}$ such that the matrix of T is upper triangular in this basis. Let $\lambda \in \mathbb{F}$, and consider the operator $T - \lambda I$. It has the same matrix, except that on the diagonal the entries are $\lambda_i - \lambda$ if those in the matrix of T are λ_j. It follows that $T - \lambda I$ is not invertible iff λ is equal with some λ_j. So λ is a proper value as desired. □

4.3 Diagonal matrices

A diagonal matrix is a matrix which is zero except possibly the diagonal.

Proposition 4.14. *If $T \in End(V)$ has $\dim V$ distinct eigenvalues, then T has a diagonal matrix*

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

with respect to some basis.

Proof. Suppose that T has $\dim V$ distinct eigenvalues, $\lambda_1, \ldots, \lambda_n$, where $n = \dim V$. Choose corresponding eigenvectors $e_1, \ldots, e_n$. Because nonzero vectors corresponding to distinct eigenvalues are linearly independent, we obtain a set of vectors with the cardinal equal to $\dim V$, that is a basis, and in this basis the matrix of T is diagonal. □

The next proposition imposes several conditions on an operator that are equivalent to having a diagonal matrix.

Proposition 4.15. *Suppose $T \in End(V)$. Denote $\lambda_1, \ldots, \lambda_n$ the distinct eigenvalues of T. The following conditions are equivalent.*

1. *T has a diagonal matrix with respect to some basis of V.*

2. *V has a basis consisting of proper vectors.*

3. *There exists one dimensional subspaces $U_1, \ldots, U_m$ of V, each invariant under T such that*
$$V = U_1 \oplus \cdots \oplus U_m.$$

4. $V = \ker(T - \lambda_1 I) \oplus \cdots \oplus \ker(T - \lambda_n I)$.

5. *$dim\, V = dim$* $\ker(T - \lambda_1 I) + \cdots +$ *dim* $\ker(T - \lambda_n I)$.

Proof. We saw that $1 \Leftrightarrow 2$. Suppose 2 holds. Choose $\{e_1, \ldots, e_m\}$ a basis consisting of proper vectors, and $U_i = \text{span}\{e_i\}$, for $i = \overline{1, m}$. Hence $2 \Rightarrow 3$. Suppose 3 holds. Choose a basis $e_j \in U_j$, $j = \overline{1, m}$. It follows that e_j, $j = \overline{1, m}$ is a proper vector, so they are linearly independent, and because they are m vectors, they form a basis. Thus 3 implies 2.

Now we know that $1, 2, 3$ are equivalent. Next we will prove the following chain of implications
$$2 \Rightarrow 4 \Rightarrow 5 \Rightarrow 2$$

Suppose 2 holds, then V has a basis consisting of eigenvectors. Then every vector in V is a linear combination of eigenvectors of T, that is
$$V = \ker(T - \lambda_1 I) + \cdots + \ker(T - \lambda_n I).$$

We show that it is a direct sum. Suppose that
$$0 = u_1 + \cdots + u_n \ ,$$

with $u_j \in \ker(T - \lambda_j I)$, $j = \overline{1, n}$. They are linearly independent, so all are 0.

Finally $4 \Rightarrow 5$ is clear because in 4 we have a direct sum.

$5 \Rightarrow 2$. $\dim V = \dim\, \ker(T - \lambda_1 I) + \cdots + \dim\, \ker(T - \lambda_n I)$. According to a precedent result, distinct eigenvalues give rise to linear independent eigenvectors.

Let $\{e_1^1, \ldots, e_{i_1}^1\}, \ldots, \{e_1^n, \ldots, e_{i_n}^n\}$ bases in $\ker(T - \lambda_1 I), \ldots, \ker(T - \lambda_n I)$. Then $\dim V = i_1 + \cdots + i_n$, and $\{e_1^1, \ldots, e_{i_1}^1, \ldots, e_1^n, \ldots, e_{i_n}^n\}$ are linearly independent. Hence $V = \operatorname{span}\{e_1^1, \ldots, e_{i_1}^1, \ldots, e_1^n, \ldots, e_{i_n}^n\}$ which shows that 2 holds. $\square$

Example 4.16. Consider the matrix $A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 0 \end{pmatrix}$. Show that A is diagonalizable and find the diagonal matrix similar to A.

The characteristic polynomial of A is

$$\det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - \lambda - 6 = -(\lambda + 1)(\lambda - 2)(\lambda - 3).$$

Hence, the eigenvalues of A are $\lambda_1 = -1, \lambda_2 = 2$ and $\lambda_3 = 3$. To find the corresponding eigenvectors, we have to solve the three linear systems $(A + I)v = 0$, $(A - 2I)v = 0$ and $(A - 3I)v = 0$. On solving these systems, we find that the solution spaces are

$$\{(\alpha, \alpha, 2\alpha) : \alpha \in \mathbb{R}\},$$

$$\{(\alpha, \alpha, -\alpha) : \alpha \in \mathbb{R}\},$$

respectively

$$\{(\alpha, -\alpha, 0) : \alpha \in \mathbb{R}\}.$$

Hence, the corresponding eigenvectors associated to λ_1, λ_2 and λ_3 respectively, are $v_1 = (1, 1, 2)$, $v_2 = (1, 1, -1)$ and $v_3 = (1, -1, 0)$ respectively. There exists 3 linear independent eigenvectors, thus A is diagonalizable.

Our transition matrix is $P = [v_1|v_2|v_3] = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 2 & -1 & 0 \end{pmatrix}$.

We have $P^{-1} = \frac{1}{6}\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & -2 \\ 3 & -3 & 0 \end{pmatrix}$.
Hence, the diagonal matrix similar to A is

$$D = P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Obviously one may directly compute D, by knowing that D is the diagonal matrix having the eigenvalues of A on its main diagonal.

Proposition 4.17. *If λ is a proper value for an operator (endomorphism) T, and $v \neq 0$, $v \in V$ is a proper vector then one has:*

1. *$\forall k \in \mathbb{N}$, λ^k is a proper value for $T^k = T \circ \cdots \circ T$ (k times) and v is a proper vector of T^k.*
2. *If $p \in \mathbb{F}[X]$ is a polynomial with coefficients in $\mathbb{F}$, then $p(\lambda)$ is an eigenvalue for $p(T)$ and v is a proper vector of $p(T)$.*
3. *For T automorphism (bijective endomorphism), λ^{-1} is a proper value for T^{-1} and v is an eigenvector for T^{-1}.*

Proof. 1. We have $T(v) = \lambda v$, hence $T \circ T(v) = T(\lambda v) = \lambda T(v) = \lambda^2 v$. Assume, that $T^{k-1}(v) = \lambda^{k-1}v$. Then $T^k(v) =$

$$T \circ T^{k-1}(v) = T(T^{k-1}(v)) = T(\lambda^{k-1}v) = \lambda^{k-1}T(v) = \lambda^{k-1}\lambda v = \lambda^k v.$$

2. Let $p = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{F}[X]$. Then $p(T)(v) =$

$$a_0I(v) + a_1T(v) + \cdots + a_nT^n(v) = a_0v + a_1(\lambda v) + \cdots + a_n(\lambda^n v) = p(\lambda)v.$$

3. $T^{-1}(v) = u$ such that $T(u) = v$. But $v = \lambda^{-1}T(v) = T(\lambda^{-1}v)$, hence $T(u) = T(\lambda^{-1}v)$. Since T is injective we have $u = \lambda^{-1}v$, or equivalently $T^{-1}(v) = \lambda^{-1}v$. $\square$

Example 4.18. Let $T : V \to V$ be a linear map. Prove that if -1 is an eigenvalue of $T^2 + T$ then 1 is an eigenvalue of T^3. Here I is the identity map and $T^2 = T \circ T$, etc.

From the fact that -1 is an eigenvalue of $T^2 + T$ there exists $v \neq 0$ such that

$$\left(T^2 + T\right) v = -v,$$

or, equivalently

$$\left(T^2 + T + I\right) v = 0.$$

Now, we apply the linear map $T - I$ (recall that the linear maps form a vector space, so the sum or difference of two linear maps is still linear) to the above relation to get

$$(T - I)\left(T^2 + T + I\right) v = 0.$$

Here we have used that, by linearity, $(T - I)\,0 = 0$.
Finally, simple algebra yields $(T - I)\,(T^2 + T + I) = T^3 - I$, so the above equation shows that

$$T^3 v = v,$$

as desired.

4.4 The Jordan canonical form

In a previous section we have seen the endomorphisms which are diagonalizable.

Let V be a vector space of finite dimension n over a field $\mathbb{F}$. Let $T : V \to V$ and let λ_0 be an eigenvalue of T. Consider the matrix form of the endomorphism in a given basis, $T(v) = M_T v$. The eigenvalues are the roots of the characteristic polynomial $\det(M_t - \lambda I_n) = 0$. It can be proved that this polynomial does not depend on the basis and of the matrix M_T. So, it will be called the characteristic polynomial of the endomorphism T, and it will be denoted by $P(\lambda)$, and of course $\deg P = n$. Sometimes it is called the characteristic polynomial of the matrix, but we understand that is the matrix associated to an operator.
Denote by $m(\lambda_0)$ the multiplicity of λ_0 as a root of this polynomial. Associated to the proper value λ_0 we consider the proper subspace corresponding to λ_0:

$$E(\lambda_0) = \{v \in V | T(v) = \lambda_0 v\}.$$

Consider a basis of V and let M_T be the matrix of T with respect to this basis. We havev that:

Theorem 4.19. *With the above notations, the following holds*

$$dim\, E(\lambda_0) = n - rank(M_T - \lambda_0 I) \leq m(\lambda_0).$$

Proof. Obviously is enough to prove the claim in $V = \mathbb{R}^n$. Let $x_1, x_2, \ldots, x_r$ be linearly independent eigenvectors associated to λ_0, so that $\dim E(\lambda_0) = r$. Complete this set with $x_{r+1}, \ldots x_n$ to a basis of $\mathbb{R}^n$. Let P be the matrix whose columns are x_i, $i = \overline{1, n}$. We have $M_T P = [\lambda_0 x_1 | \ldots | \lambda_0 x_r | \ldots]$. We get that the first r columns of $P^{-1} M_T P$ are diagonal with λ_0 on the diagonal, but that the rest of the columns are indeterminable. We prove next that $P^{-1} M_T P$ has the same characteristic polynomial as M_T. Indeed

$$\det(P^{-1} M_T P - \lambda I) = \det(P^{-1} M_T P - P^{-1}(\lambda I) P) =$$

$$\det(P^{-1}(M_T - \lambda I)P) = \frac{1}{\det(P)} \det(M_T - \lambda I)\det(P) = \det(M_T - \lambda I).$$

But since the first few columns of $P^{-1}M_TP$ are diagonal with λ_0 on the diagonal we have that the characteristic polynomial of $P^{-1}M_TP$ has a factor of at least $(\lambda_0 - \lambda)^r$, so the algebraic multiplicity of λ_0 is at least r.

$\square$

The value $\dim E(\lambda_0)$ is called the geometric multiplicity of the eigenvalue λ_0.
Let $T \in End(V)$, and suppose that the roots of the characteristic polynomial are in $\mathbb{F}$. Let λ be a root of the characteristic polynomial, i.e. an eigenvalue of T. Consider m the algebraic multiplicity of λ and $q = \dim E(\lambda)$, the geometric multiplicity of λ.
It is possible to find q eigenvectors and $m - q$ principal vectors (also called generalized eigenvectors), all of them linearly independent, and an eigenvector v and the corresponding principal vectors $u_1, \ldots, u_r$ satisfy

$$T(v) = \lambda v, T(u_1) = \lambda u_1 + v, \ldots, T(u_r) = \lambda u_r + u_{r-1}$$

The precedent definition can equivalently be stated as
A nonzero vector u is called a generalized eigenvector of rank r associated with the eigenvalue λ if and only if $(T - \lambda I)^r(u) = 0$ and $(T - \lambda I)^{r-1}(u) \neq 0$. We note that a generalized eigenvector of rank 1 is an ordinary eigenvector. The previously defined principal vectors $u_1, \ldots, u_r$ are generalized eigenvectors of rank $2, \ldots, r+1$.
It is known that if λ is an eigenvalue of algebraic multiplicity m, then there are m linearly independent generalized eigenvectors associated with λ.
These eigenvectors and principal vectors associated to T by considering all the eigenvalues of T form a basis of V, called the Jordan basis with respect to T. The

matrix of T relative to a Jordan basis is called a Jordan matrix, and it has the form

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix}$$

The J's are matrices, called Jordan cells. Each cell represents the contribution of an eigenvector v, and the corresponding principal vectors, $u_1, \dots u_r$, and it has the form

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \in \mathcal{M}_{r+1}(\mathbb{F})$$

It is easy to see that the Jordan matrix is a diagonal matrix iff there are no principal vectors iff $m(\lambda) = \dim E(\lambda)$ for each eigenvalue λ.

Let M_T be the matrix of T with respect to a given basis B, and J be the Jordan matrix with respect to a Jordan basis B'. Late P be the transition matrix from B to B', hence it have columns consisting of either eigenvectors or generalized eigenvectors. Then $J = P^{-1} M_T P$, hence $M_T = PJP^{-1}$.

Example 4.20. (algebraic multiplicity 3, geometric multiplicity 2) Consider the operator with the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$. Find the Jordan matrix and the transition matrix of A.

The characteristic polynomial of A is $\det(A-\lambda I)=(2-\lambda)^3$, hence $\lambda=2$ is an eigenvalue with algebraic multiplicity 3. By solving the homogenous system $(A-2I)v=0$ we obtain the solution space
$E(2)=\ker(A-2I)=\{(\alpha,2\alpha,\beta):\alpha,\beta\in\mathbb{R}\}$. Hence the dimension of $E(2)$ is 2, consequently the eigenvalue $\lambda=2$ has geometric multiplicity 2. Therefore we can take the linear independent eigenvectors $v_1=(1,2,1)$ respectively $v_2=(0,0,1)$. Note that we need a generalized eigenvector, which can be obtained as a solution of the system

$$(A-2I)u=v_1.$$

The solutions of this system lie in the set $\{(\alpha,2\alpha+1,\beta):\alpha,\beta\in\mathbb{R}\}$, hence a generalized eigenvector, is $u_1=(1,3,0)$.
Note that v_1,u_1,v_2 are linear independent, hence we take the transition matrix
$P=[v_1|u_1|v_2]=\begin{pmatrix}1&1&0\\2&3&0\\1&0&1\end{pmatrix}$. Then $P^{-1}=\begin{pmatrix}3&-1&0\\-2&1&0\\-3&1&1\end{pmatrix}$, hence

$$J=P^{-1}AP=\begin{pmatrix}2&1&0\\0&2&0\\0&0&2\end{pmatrix}.$$

Example 4.21. (algebraic multiplicity 3, geometric multiplicity 1) Consider the operator with the matrix $A=\begin{pmatrix}-1&-18&-7\\1&-13&-4\\-1&25&8\end{pmatrix}$. Find the Jordan matrix and the transition matrix of A.

The characteristic polynomial of A is $\det(A-\lambda I)=-(\lambda+2)^3$, hence $\lambda=-2$ is an eigenvalue with algebraic multiplicity 3. By solving the homogenous system $(A+2I)v=0$ we obtain the solution space

$E(2) = \ker(A+2I) = \{(5\alpha, 3\alpha, -7\alpha) : \alpha \in \mathbb{R}\}$. Hence the dimension of $E(2)$ is 1, consequently the eigenvalue $\lambda = 2$ has geometric multiplicity 1. Therefore we can take the linear independent eigenvector $v = (5, 3, -7)$. Note that we need two generalized eigenvectors, which can be obtained as a solution of the system

$$(A+2I)u_1 = v,$$

respectively

$$(A+2I)u_2 = u_1.$$

The solutions of the first system lie in the set $\{(-\frac{1+5\alpha}{7}, -\frac{2+3\alpha}{7}, \alpha) : \alpha \in \mathbb{R}\}$, hence a generalized eigenvector, for $\alpha = 4$ is $u_1 = (-3, -2, 4)$.
The solutions of the system $(A+2I)u_2 = u_1$ with $u_1 = (-3, -2, 4)$ lie in the set $\{(-\frac{3+5\alpha}{7}, \frac{1-3\alpha}{7}, \alpha) : \alpha \in \mathbb{R}\}$, hence a generalized eigenvector, for $\alpha = 5$ is $u_1 = (-4, -2, 5)$. Note that v, u_1, u_2 are linear independent, hence we take the transition matrix $P = [v_1|u_1|u_2] = \begin{pmatrix} 5 & -3 & -4 \\ 3 & -2 & -2 \\ -7 & 4 & 5 \end{pmatrix}$. Then

$P^{-1} = \begin{pmatrix} -2 & -1 & -2 \\ -1 & -3 & -2 \\ -2 & 1 & -1 \end{pmatrix}$, hence

$$J = P^{-1}AP = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}.$$

4.5 Problems

Problem 4.5.1. Find the eigenvalues and eigenvectors of the operator
$T : C^\infty(1, b) \to C^\infty(1, b),\ T(f)(x) = \frac{f'(x)}{xe^{x^2}}$.

Problem 4.5.2. Find matrices which diagonalize the following: *a)* $\begin{pmatrix} 1 & 5 \\ 3 & 3 \end{pmatrix}$.

b) $\begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix}$.

Problem 4.5.3. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 2 & 1 & -1 \\ 3 & -2 & 3 \\ 2 & -2 & 3 \end{pmatrix}.$$

Problem 4.5.4. Prove that a square matrix and its transpose have the same eigenvalues.

Problem 4.5.5. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 6 & 6 & -15 \\ 1 & 5 & -5 \\ 1 & 2 & -2 \end{pmatrix}.$$

Problem 4.5.6. Find the eigenvalues and eigenvectors of the operator $T : C[-\pi, \pi] \to C[-\pi, \pi]$,

$$T(f)(x) = \int_{-\pi}^{\pi} (x \cos y + \sin x \sin y) f(y) dy.$$

Problem 4.5.7. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 4 & 1 & 1 \\ -2 & 2 & -2 \\ 1 & 1 & 4 \end{pmatrix}.$$

Problem 4.5.8. Find the eigenvalues and eigenvectors of the operator $T : C[-\pi, \pi] \to C[-\pi, \pi]$,

$$T(f)(x) = \int_{-\pi}^{\pi} (\cos^3(x-y) + 1) f(y) dy.$$

Problem 4.5.9. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}.$$

Problem 4.5.10. Find the eigenvalues and eigenvectors of the operator $T : C^\infty(1,2) \to C^\infty(1,2),\ T(f)(x) = \frac{f'(x)}{\sin^2 x}$.

Problem 4.5.11. Triangularize the matrix $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$.

Problem 4.5.12. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}.$$

Problem 4.5.13. Find the eigenvalues and eigenvectors of the operator $T : C^\infty(1,b) \to C^\infty(1,b)$, $T(f)(x) = \frac{f'(x)}{\tan^2 x}$.

Problem 4.5.14. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ -4 & -2 & 1 \\ 4 & 1 & -2 \end{pmatrix}.$$

Problem 4.5.15. Prove that a complex 2×2 matrix is not diagonalizable if and only if it is similar to a matrix of the form $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$, where $b \neq 0$.

Problem 4.5.16. Find the Jordan canonical form and the transition matrix for the matrices

$$\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}, \quad \begin{pmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & -4 \end{pmatrix}.$$

Problem 4.5.17. Prove that if A and B are $n \times n$ matrices, then AB and BA have the same eigenvalues.

Problem 4.5.18. Find the Jordan canonical form and the transition matrix for the matrix

$$\begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}.$$

5

Inner product spaces

5.1 Basic definitions and results

Up to now we have studied vector spaces, linear maps, special linear maps.
We can measure if two vectors are equal, but we do not have something like "length", so we cannot compare two vectors. Moreover, we cannot say anything about the position of two vectors.
In a vector space one can define the norm of a vector and the inner product of two vectors. The notion of the norm permits us to measure the length of the vectors, and compare two vectors. The inner product of two vectors, on one hand induces a norm, so the length can be measured, and on the other hand (at least in the case of real vector spaces), lets us measure the angle between two vectors, so a full geometry can be constructed there. Nevertheless in the case of complex vector spaces, the angle of two vectors is not clearly defined, but the orthogonality is.

Definition 5.1. *An inner product on a vector space V over the field $\mathbb{F}$ is a function (bilinear form) $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ with the properties:*

- *(positivity and definiteness) $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0$ iff $v = 0$.*

- *(additivity in the first slot)* $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$*, for all* $u, v, w \in V$.
- *(homogeneity in the first slot)* $\langle \alpha v, w\rangle = \alpha\langle v, w\rangle$ *for all* $\alpha \in \mathbb{F}$ *and* $v, w \in V$.
- *(conjugate symmetry)* $\langle v, w\rangle = \overline{\langle w, v\rangle}$ *for all* $v, w \in V$.

An inner product space is a pair $(V, \langle\cdot,\cdot\rangle)$*, where* V *is vector space and* $\langle\cdot,\cdot\rangle$ *is an inner product on* V.

The most important example of an inner product space is $\mathbb{F}^n$. Let $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots w_n)$ and define the inner product by

$$\langle v, w\rangle = v_1\overline{w}_1 + \cdots + v_n\overline{w}_n.$$

This is the typical example of an inner product, called the Euclidean inner product, and when $\mathbb{F}^n$ is referred to as an inner product space, one should assume that the inner product is the Euclidean one, unless explicitly stated otherwise.

Example 5.2. Let $A \in \mathcal{M}_2(\mathbb{R})$, $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ be a positive definite matrix, that is $a > 0$, $\det(A) > 0$. Then for every $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R}^2$ we define $\langle u, v\rangle = (v_1\, v_2)A \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$.
It can easily be verified that $\langle\cdot,\cdot\rangle$ is an inner product on the real linear space $\mathbb{R}^2$. If $A = I_2$ we obtain the usual inner product $\langle u, v\rangle = u_1v_1 + u_2v_2$.

From the definition one can easily deduce the following properties of an inner product:

$$\langle v, 0\rangle = \langle 0, v\rangle = 0,$$
$$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle,$$
$$\langle u, \alpha v\rangle = \overline{\alpha}\langle u, v\rangle,$$

for all $u, v, w \in V$ and $\alpha \in \mathbb{F}$

Definition 5.3. *Let V be a vector space over $\mathbb{F}$. A function*

$$\|\cdot\| : V \to \mathbb{R}$$

is called a norm on V if:

- *(positivity)* $\|v\| \geq 0$, $v \in V$, $\|v\| = 0 \Leftrightarrow v = 0$ *;*
- *(homogeneity)* $\|\alpha v\| = |\alpha| \cdot \|v\|, \forall \alpha \in \mathbb{F}, \forall v \in V$;
- *(triangle inequality)* $\|u + v\| \leq \|u\| + \|v\|, \forall u, v \in V$.

A normed space is a pair $(V, \|\cdot\|)$, where V is a vector space and $\|\cdot\|$ is a norm on V.

Example 5.4. On the real linear space $\mathbb{R}^n$ one can define a norm in several ways. Indeed, for any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ define its norm as
$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. One can easily verify that the axioms in the definition of norm are satisfied. This norm is called the *Euclidian norm.*
More generally, for any $p \in \mathbb{R}$, $p \geq 1$ we can define
$\|x\| = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$, the so called $p-$*norm* on $\mathbb{R}^n$.
Another way to define a norm on $\mathbb{R}^n$ is $\|x\| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$. This is the so called *maximum norm.*

Definition 5.5. *Let X be a nonempty set. A function $d : X \times X \to \mathbb{R}$ satisfying the following properties:*

- *(positivity)* $d(x, y) \geq 0, \forall x, y \in X$ *and* $d(x, y) = 0 \Leftrightarrow x = y$;
- *(symmetry)* $d(x, y) = d(y, x), \forall x, y \in X$;
- *(triangle inequality)* $d(x, y) \leq d(x, z) + d(z, y), \forall x, y, z \in X$;

is called a metric or distance on X. A set X with a metric defined on it is called a metric space.

Example 5.6. Let X be an arbitrary set. One can define a distance on X by

$$d(x,y) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{otherwise.} \end{cases}$$

This metric is called the *discrete metric* on X. On $\mathbb{R}^n$ the *Chebyshev distance* is defined as

$$d(x,y) = \max_{1 \le i \le n} |x_i - y_i|,\ x = (x_1, x_2, \dots, x_n),\ y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n.$$

In this course we are mainly interested in the inner product spaces. But we should point out that an inner product on V defines a norm, by $\|v\| = \sqrt{\langle v, v\rangle}$ for $v \in V$, and a norm on V defines a metric by $d(v,w) = \|w - v\|$, for $v, w \in V$.

On the other hand, from their generality point of view the metrics are the most general ones (can be defined on any set), followed by norms (which assumes the linearity of the space where is defined) and on the last position is the inner product. It should be pointed that every inner product generates a norm, but not every norm comes from an inner product, as is the case for the max norm defined above.

For an inner product space $(V, \langle \cdot, \cdot \rangle)$ the following identity is true:

$$\left\langle \sum_{i=1}^{m} \alpha_i v_i, \sum_{j=1}^{n} \beta_j w_j \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \overline{\beta}_j \langle v_i, w_j \rangle.$$

Definition 5.7. *Two vectors $u, v \in V$ are said to be orthogonal ($u \perp v$) if $\langle u, v \rangle = 0$.*

In a real inner product space we can define *the angle of two vectors* as

$$\widehat{(v,w)} = \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}$$

We have

$$v \perp w \Leftrightarrow \langle v, w\rangle = 0 \Leftrightarrow \widehat{(v, w)} = \frac{\pi}{2}.$$

Theorem 5.8. ***(Parallelogram law)*** *Let V be an inner product space and $u, v \in V$. Then*

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Proof.

$$\begin{aligned} \|u+v\|^2 + \|u-v\|^2 &= \langle u+v, u+v\rangle + \langle u-v, u-v\rangle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \langle u, u\rangle + \langle +\langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle \\ &= 2(\|u\|^2 + \|v\|^2). \end{aligned}$$

□

Theorem 5.9. ***(Pythagorean Theorem)*** *Let V be an inner product space, and $u, v \in V$ orthogonal vectors. Then*

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof.

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \|u\|^2 + \|v\|^2. \end{aligned}$$

□

Now we are going to prove one of the most important inequalities in mathematics, namely the Cauchy-Schwarz inequality. There are several methods of proof for this, we will give one related to our aims.

Consider $u, v \in V$. We want to write u as a sum between a vector collinear to v and a vector orthogonal to v. Let $\alpha \in \mathbb{F}$ and write u as $u = \alpha v + (u - \alpha v)$. Imposing now the condition that v is orthogonal to $(u - \alpha v)$, one obtains

$$0 = \langle u - \alpha v, v \rangle = \langle u, v \rangle - \alpha \|v\|^2,$$

so one has to choose $\alpha = \frac{\langle u,v \rangle}{\|v\|^2}$, and the decomposition is

$$u = \frac{\langle u, v \rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v \rangle}{\|v\|^2} v \right).$$

Theorem 5.10. ***Cauchy-Schwarz Inequality*** *Let V be an inner product space and $u, v \in V$. Then*

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|.$$

The equality holds iff one of u, v is a scalar multiple of the other (u and v are collinear).

Proof. Let $u, v \in V$. If $v = 0$ both sides of the inequality are 0 and the desired result holds. Suppose that $v \neq 0$. Write $u = \frac{\langle u,v \rangle}{\|v\|^2} v + \left(u - \frac{\langle u,v \rangle}{\|v\|^2} v \right)$. Taking into account that the vectors $\frac{\langle u,v \rangle}{\|v\|^2} v$ and $u - \frac{\langle u,v \rangle}{\|v\|^2} v$ are orthogonal, by the Pythagorean theorem we obtain

$$\begin{aligned} \|u\|^2 &= \left\| \frac{\langle u, v \rangle}{\|v\|^2} v \right\|^2 + \left\| u - \frac{\langle u, v \rangle}{\|v\|^2} v \right\|^2 \\ &= \frac{|\langle u, v \rangle|^2}{\|v\|^2} + \left\| u - \frac{\langle u, v \rangle}{\|v\|^2} v \right\|^2 \\ &\geq \frac{|\langle u, v \rangle|^2}{\|v\|^2}, \end{aligned}$$

inequality equivalent with the one in the theorem.

We have equality iff $u - \frac{\langle u,v \rangle}{\|v\|^2} v = 0$, that is iff u is a scalar multiple of v.

$\square$

5.2 Orthonormal Bases

Definition 5.11. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let I be an arbitrary index set. A family of vectors $A = \{e_i \in V | i \in I\}$ is called an* orthogonal *family, if $\langle e_i, e_j \rangle = 0$ for every $i, j \in I$, $i \neq j$. The family A is called* orthonormal *if it is orthogonal and $\|e_i\| = 1$ for every $i \in I$.*

One of the reason that one studies orthonormal families is that in such special bases the computations are much more simple.

Proposition 5.12. *If $(e_1, e_2, \ldots, e_m)$ is an orthonormal family of vectors in V, then*

$$\|\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_m e_m\|^2 = |\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_m|^2$$

for all $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{F}$.

Proof. Apply Pythagorean Theorem, that is

$$\|\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_m e_m\|^2 = |\alpha_1|^2 \|e_1\|^2 + |\alpha_2|^2 \|e_2\|^2 + \cdots + |\alpha_m|^2 \|e_n\|^2.$$

The conclusion follows taking into account that $\|e_i\| = 1$, $i = \overline{1, n}$. □

Corollary 5.13. *Every orthonormal list of vectors is linearly independent.*

Proof. Let $(e_1, e_2, \ldots, e_m)$ be an orthonormal list of vectors in V and $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{F}$ with

$$\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_m e_m = 0.$$

It follows that $\|\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_m e_m\|^2 = |\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_m|^2 = 0$, that is $\alpha_j = 0$, $j = \overline{1, m}$. □

An orthonornal basis of an inner product vector space V is a basis of V which is also an orthonormal list of V. It is clear that every orthonormal list of vectors of length $\dim V$ is an orthonormal basis (because it is linearly independent, being orthonormal).

Theorem 5.14. *Let $(e_1, e_2, \ldots, e_n)$ be an orthonormal basis of an inner product space V. If $v = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \in V$, then*

- $\alpha_i = \langle v, e_i \rangle$, *for all* $i \in \{1, 2, \ldots, n\}$ *and*
- $\|v\|^2 = \sum_{i=1}^{n} |\langle v, e_i \rangle|^2$

Proof. Since $v = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n$, by taking the inner product in both sides with e_i we have

$\langle v, e_i \rangle = \alpha_1 \langle e_1, e_i \rangle + \alpha_2 \langle e_2, e_i \rangle + \cdots + \alpha_i \langle e_i, e_i \rangle + \cdots + \alpha_n \langle e_n, e_i \rangle = \alpha_i.$

The second assertion comes from applying the previous proposition. Indeed,

$$\|v\|^2 = \|\alpha_1 e_1 + \cdots + \alpha_n e_n\|^2 = |\alpha_1|^2 + \cdots + |\alpha_n|^2 = \sum_{i=1}^{n} |\langle v, e_i \rangle|^2.$$

□

Up to now we have an image about the usefulness of orthonormal basis. The advantage is that in an orthonormal basis the computations are easy, as in the Euclidean two or three dimensional spaces. But how does one go to find them? The next result gives an answer to the question. The following result is a well known algorithm in linear algebra, called the Gram-Schmidt procedure. The procedure is pointed here, giving a method for turning a linearly independent list into an orthonormal one, with the same span as the original one.

Theorem 5.15. ***Gram-Schmidt*** *If $(v_1, v_2, \ldots, v_m)$ is a linearly independent set of vectors in V, then there exists an orthonormal set of vectors $(e_1, \ldots e_m)$ in V,* *such that*

$$span(v_1, v_2, \dots, v_k) = span(e_1, e_2 \dots, e_k)$$

for every $k \in \{1, 2, \dots, m\}$.

Proof. Let $(v_1, v_2, \dots, v_m)$ be a linearly independent set of vectors. The family of orthonormal vectors $(e_1, e_2 \dots, e_m)$ will be constructed inductively. Start with $e_1 = \frac{v_1}{\|v_1\|}$. Suppose now that $j > 1$ and an orthonormal family $(e_1, e_2, \dots, e_{j-1})$ has been constructed such that

$$\text{span}(v_1, v_2, \dots, v_{j-1}) = \text{span}(e_1, e_2, \dots, e_{j-1})$$

Consider

$$e_j = \frac{v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|}$$

Since the list $(v_1, v_2, \dots, v_m)$ is linearly independent, it follows that v_j is not in span$(v_1, v_2, \dots, v_{j-1})$, and thus is not in span$(e_1, e_2, \dots, e_{j-1})$. Hence e_j is well defined, and $\|e_j\| = 1$. By direct computations it follows that for $1 < k < j$ one has

$$\begin{aligned}
\langle e_j, e_k\rangle &= \left\langle \frac{v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|}, e_k \right\rangle \\
&= \frac{\langle v_j, e_k\rangle - \langle v_j, e_k\rangle}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|} \\
&= 0,
\end{aligned}$$

thus $(e_1, e_2, \dots e_k)$ is an orthonormal family. By the definition of e_j one can see that $v_j \in \text{span}(e_1, e_2, \dots, e_j)$, which gives (together with our hypothesis of induction), that

$$\text{span}(v_1, v_2, \dots, v_j) \subset \text{span}(e_1, e_2, \dots, e_j)$$

Both lists being linearly independent (the first one by hypothesis and the second one by orthonormality), it follows that the generated subspaces above have the same dimension j, so they are equal. □

Remark 5.16. If in the Gram-Schmidt process we do not normalize the vectors we obtain an orthogonal basis instead of an orthonormal one.

Example 5.17. Orthonormalize the following list of vectors in $\mathbb{R}^4$:

$$\{v_1 = (0,1,1,0), v_2 = (0,4,0,1), v_3 = (1,-1,1,0), v_4 = (1,3,0,1)\}.$$

First we will orthogonalize by using the Gram-Schmidt procedure.
Let $u_1 = v_1 = (0,1,1,0)$.

$$u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle}u_1 = (0,4,0,1) - \frac{4}{2}(0,1,1,0) = (0,2,-2,1).$$

$$u_3 = v_3 - \frac{\langle v_3, u_1\rangle}{\langle u_1, u_1\rangle}u_1 - \frac{\langle v_3, u_2\rangle}{\langle u_2, u_2\rangle}u_2 = \left(1, -\frac{1}{9}, \frac{1}{9}, \frac{4}{9}\right).$$

$$u_4 = v_4 - \frac{\langle v_4, u_1\rangle}{\langle u_1, u_1\rangle}u_1 - \frac{\langle v_4, u_2\rangle}{\langle u_2, u_2\rangle}u_2 - \frac{\langle v_4, u_3\rangle}{\langle u_3, u_3\rangle}u_3 = \left(\frac{1}{11}, \frac{1}{22}, -\frac{1}{22}, -\frac{2}{11}\right).$$

It can easily be verified that the list $\{u_1, u_2, u_3, u_4\}$ is orthogonal. Take now $w_i = \frac{u_i}{\|u_i\|}$, $i = \overline{1,4}$. We obtain

$$w_1 = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right),$$

$$w_2 = \left(0, \frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right),$$

$$w_3 = \left(\frac{3}{\sqrt{11}}, -\frac{1}{3\sqrt{11}}, \frac{1}{3\sqrt{11}}, \frac{4}{3\sqrt{11}}\right),$$

$$w_4 = \left(\frac{\sqrt{22}}{11}, \frac{\sqrt{22}}{22}, -\frac{\sqrt{22}}{22}, -\frac{2\sqrt{22}}{11}\right).$$

Obviously the list $\{w_1, w_2, w_3, w_4\}$ is orthonormal.

Now we can state the main results in this section.

Corollary 5.18. *Every finitely dimensional inner product space has an orhtonormal basis.*

Proof. Choose a basis of V, apply the Gram-Schmidt procedure to it and obtain an orthonormal list of length equal to $\dim V$. It follows that the list is a basis, being linearly independent. □

The next proposition shows that any orthonormal list can be extended to an orthonormal basis.

Proposition 5.19. *Every orhtonormal family of vectors can be extended to an orthonormal basis of V.*

Proof. Suppose $(e_1, e_2, \ldots, e_m)$ is an orthonormal family of vectors.. Being linearly independent, it can be extended to a basis, $(e_1, e_2, \ldots, e_m, v_{m+1}, \ldots, v_n)$. Applying now the Gram-Schmidt procedure to $(e_1, e_2, \ldots, e_m, v_{m+1}, \ldots, v_n)$, we obtain the list $(e_1, e_2, \ldots, e_m, f_{m+1}, \ldots, f_n)$, (note that the Gram Schmidt procedure leaves the first m entries unchanged, being already orthonormal). Hence we have an extension to an orthonormal basis. □

Corollary 5.20. *Suppose that $T \in End(V)$. If T has an upper triangular form with respect to some basis of V, then T has an upper triangular form with respect to some orthonormal basis of V.*

Corollary 5.21. *Suppose that V is a complex vector space and $T \in End(V)$. Then T has an upper triangular form with respect to some orthonormal basis of V.*

5.3 Orthogonal complement

Let $U \subseteq V$ be a subset of an inner product space V. The *orthogonal complement* of U, denoted by $U^\perp$ is the set of all vectors in V which are orthogonal to every vector in U i.e.:

$$U^{\perp} = \{v \in V | \langle v, u \rangle = 0, \ \forall u \in U\}.$$

It can easily be verified that $U^{\perp}$ is a subspace of V, $V^{\perp} = \{0\}$ and $\{0\}^{\perp} = V$, as well that $U_1 \subseteq U_2 \Rightarrow U_2^{\perp} \subseteq U_1^{\perp}$.

Theorem 5.22. *If U is a subspace of V, then*

$$V = U \oplus U^{\perp}$$

Proof. Suppose that U is a subspace of V. We will show that

$$V = U + U^{\perp}$$

Let $\{e_1, \dots, e_m\}$ be an orthonormal basis of U and $v \in V$. We have

$$v = (\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m) + (v - \langle v, e_1 \rangle e_1 - \cdots - \langle v, e_m \rangle e_m)$$

Denote the first vector by u and the second by w. Clearly $u \in U$. For each $j \in \{1, 2, \dots, m\}$ one has

$$\begin{aligned} \langle w, e_j \rangle &= \langle v, e_j \rangle - \langle v, e_j \rangle \\ &= 0 \end{aligned}$$

Thus w is orthogonal to every vector in the basis of U, that is $w \in U^{\perp}$, consequently

$$V = U + U^{\perp}.$$

We will show now that $U \cap U^{\perp} = \{0\}$. Suppose that $v \in U \cap U^{\perp}$. Then v is orthogonal to every vector in U, hence $\langle v, v \rangle = 0$, that is $v = 0$. The relations $V = U + U^{\perp}$ and $U \cap U^{\perp} = \{0\}$ imply the conclusion of the theorem. □

Proposition 5.23. *If U_1, U_2 are subspaces of V then* *a)* $U_1 = (U_1^{\perp})^{\perp}$.

b) $(U_1 + U_2)^{\perp} = U_1^{\perp} \cap U_2^{\perp}$.

c) $(U_1 \cap U_2)^{\perp} = U_1^{\perp} + U_2^{\perp}$.

Proof. a) We show first that $U_1 \subseteq (U_1^{\perp})^{\perp}$. Let $u_1 \in U_1$. Then for all $v \in U_1^{\perp}$ one has $v \perp u_1$. In other words $\langle u_1, v\rangle = 0$ for all $v \in U_1^{\perp}$. Hence $u_1 \in (U_1^{\perp})^{\perp}$.
Assume now that $(U_1^{\perp})^{\perp} \not\subseteq U_1$. Hence, there exists $u_2 \in (U_1^{\perp})^{\perp} \setminus U_1$. Since $V = U_1 \oplus U_1^{\perp}$ we obtain that there exists $u_1 \in U_1$ such that $u_2 - u_1 \in U_1^{\perp}$ $(*)$.
On the other hand, according to the first part of proof $u_1 \in (U_1^{\perp})^{\perp}$ and $(U_1^{\perp})^{\perp}$ is a linear subspace, hence $u_2 - u_1 \in (U_1^{\perp})^{\perp}$. Hence, for all $v \in U_1^{\perp}$ we have $(u_2 - u_1) \perp v$ $(**)$.
$(*)$ and $(**)$ implies that $(u_2 - u_1) \perp (u_2 - u_1)$ that is $\langle u_2 - u_1, u_2 - u_1\rangle = 0$, which leads to $u_1 = u_2$ contradiction.
b) For $v \in (U_1 + U_2)^{\perp}$ one has $\langle v, u_1 + u_2\rangle = 0$ for all $u_1 + u_2 \in U_1 + U_2$. By taking $u_2 = 0$ we obtain that $v \in U_1^{\perp}$ and by taking $u_1 = 0$ we obtain that $v \in U_2^{\perp}$. Hence $(U_1 + U_2)^{\perp} \subseteq U_1^{\perp} \cap U_2^{\perp}$.
Conversely, let $v \in U_1^{\perp} \cap U_2^{\perp}$. Then $\langle v, u_1\rangle = 0$ for all $u_1 \in U_1$ and $\langle v, u_2\rangle = 0$ for all $u_2 \in U_2$. Hence $\langle v, u_1 + u_2\rangle = 0$ for all $u_1 \in U_1$ and $u_2 \in U_2$, that is $v \in (U_1 + U_2)^{\perp}$.
c) According to a) $((U_1 \cap U_2)^{\perp})^{\perp} = U_1 \cap U_2$.
According to b) and a) $(U_1^{\perp} + U_2^{\perp})^{\perp} = (U_1^{\perp})^{\perp} \cap (U_2^{\perp})^{\perp} = U_1 \cap U_2$.
Hence, $((U_1 \cap U_2)^{\perp})^{\perp} = (U_1^{\perp} + U_2^{\perp})^{\perp}$ which leads to $(U_1 \cap U_2)^{\perp} = U_1^{\perp} + U_2^{\perp}$. $\square$

Example 5.24. Let $U = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 | x_1 - x_2 + x_3 - x_4 = 0\}$. Knowing that U is a subspace of $\mathbb{R}^4$, compute $\dim U$ and $U^{\perp}$.

It is easy to see that

$$\begin{aligned} U &= \left\{(x_1,x_2,x_3,x_4)\in\mathbb{R}^4 | x_1-x_2+x_3-x_4=0\right\} \\ &= \left\{(x_1,x_2,x_3,x_1-x_2+x_3)\,| x_1,x_2,x_3\in\mathbb{R}\right\} \\ &= \left\{x_1\,(1,0,0,1)+x_2\,(0,1,0,-1)+x_3\,(0,0,1,1)\,| x_1,x_2,x_3\in\mathbb{R}\right\} \\ &= \operatorname{span}\left\{(1,0,0,1)\,,(0,1,0,-1)\,,(0,0,1,1)\right\}. \end{aligned}$$

The three vectors $(1,0,0,1)\,,(0,1,0,-1)\,,(0,0,1,1)$ are linearly independent (the rank of the matrix they form is 3), so they form a basis of U and $\dim U = 3$.
The dimension formula

$$\dim U + \dim U^{\perp} = \dim \mathbb{R}^4$$

tells us that $\dim U^{\perp} = 1$, so $U^{\perp}$ is generated by a single vector. A vector that generates $U^{\perp}$ is $(1,-1,1,-1)$, the vector formed by the coefficients that appear in the linear equation that defines U. This is true because the right hand side of the equation is exactly the scalar product between $u^{\perp} = (1,-1,1,-1)$ and a vector $v = (x_1,x_2,x_3,x_4)\in U$.

5.4 Linear manifolds

Let V be a vector space over the field $\mathbb{F}$.

Definition 5.25. *A set $L = v_0 + V_L = \{v_0 + v | v \in V_L\}$, where $v_0 \in V$ is a vector and $V_L \subset V$ is a subspace of V is called a linear manifold (or linear variety). The subspace V_L is called the director subspace of the linear variety.*

Remark 5.26. One can easily verify the following.

- A linear manifold is a translated subspace, that is $L = f(V_L)$ where $f : V \to V,\ f(v) = v_0 + v$.

- if $v_0 \in V_L$ then $L = V_L$.
- $v_0 \in L$ because $v_0 = v_0 + 0 \in v_0 + V_L$.
- for $v_1, v_2 \in L$ we have $v_1 - v_2 \in V_L$.
- for every $v_1 \in L$ we have $L = v_1 + V_L$.
- $L_1 = L_2$, where $L_1 = v_0 + V_{L_1}$ and $L_2 = v_0' + V_{L_2}$ iff $V_{L_1} = V_{L_2}$ and $v_0 - v_0' \in V_{L_1}$.

Definition 5.27. *We would like to emphasize that:*

1. *The dimension of a linear manifold is the dimension of its director subspace.*
2. *Two linear manifolds L_1 and L_2 are called orthogonal if $V_{L_1} \perp V_{L_2}$.*
3. *Two linear manifolds L_1 and L_2 are called parallel if $V_{L_1} \subset V_{L_2}$ or $V_{L_2} \subset V_{L_1}$.*

Let $L = v_0 + V_L$ be a linear manifold in a finitely dimensional vector space V. For $\dim L = k \leq n = \dim V$ one can choose in the director subspace V_L a basis of finite dimension $\{v_1, \dots, v_k\}$. We have

$$L = \{v = v_0 + \alpha_1 v_1 + \cdots + \alpha_k v_k | \alpha_i \in \mathbb{F}, i = \overline{1,k}\}$$

We can consider an arbitrary basis (fixed) in V, let's say $E = \{e_1, \dots, e_n\}$ and if we use the column vectors for the coordinates in this basis, i.e.
$v_{[E]} = (x_1, \dots, x_n)^\top$, $v_{0_{[E]}} = (x_1^0, \dots, x_n^0)^\top$, $v_{j_{[E]}} = (x_{1j}, \dots, x_{nj})^\top$, $j = \overline{1,k}$, one has the *parametric equations of the linear manifold*

$$\begin{cases} x_1 = & x_1^0 + \alpha_1 x_{11} + \cdots + \alpha_k x_{1k} \\ \vdots & \\ x_n = & x_n^0 + \alpha_1 x_{n1} + \cdots + \alpha_k x_{nk} \end{cases}$$

The rank of the matrix $(x_{ij})_{\substack{i=\overline{1,n} \\ j=\overline{1,k}}}$ is k because the vectors $v_1, \ldots, v_k$ are linearly independent.

It is worthwhile to mention that:

1. a linear manifold of dimension one is called line.
2. a linear manifold of dimension two is called plane.
3. a linear manifold of dimension k is called k plane.
4. a linear manifold of dimension $n-1$ in an n dimensional vector space is called hyperplane.

Theorem 5.28. *Let us consider V an n-dimensional vector space over the field $\mathbb{F}$. Then any subspace of V is the kernel of a surjective linear map.*

Proof. Suppose V_L is a subspace of V of dimension k. Choose a basis $\{e_1, \ldots, e_k\}$ in V_L and complete it to a basis $\{e_1, \ldots, e_k, e_{k+1}, \ldots, e_n\}$ of V. Consider $U = \operatorname{span}\{e_{k+1}, \ldots, e_n\}$. Let $T : V \to U$ given by

$$T(e_1) = 0, \ldots T(e_k) = 0, T(e_{k+1}) = e_{k+1}, \ldots, T(e_n) = e_n.$$

Obviously, $T(\alpha_1 e_1 + \cdots + \alpha_n e_n) = \alpha_1 T(e_1) + \cdots + \alpha_n T(e_n) = \alpha_{k+1} e_{k+1} + \cdots + \alpha_n e_n$ defines a linear map. It is also clear that $\ker T = V_L$ as well that T is surjective, i.e. $\operatorname{im} T = U$. $\square$

Remark 5.29. In fact the map constructed in the previous theorem is nothing but the projection on U parallel to the space $\text{span}\{e_1, \dots, e_k\}$.

Theorem 5.30. *Let V, U two linear spaces over the same field $\mathbb{F}$. If $T : V \to U$ is a surjective linear map, then for every $u_0 \in U$, the set $L = \{v \in V | T(v) = u_0\}$ is a linear manifold.*

Proof. T being surjective, there exists $v_0 \in V$ with $T(v_0) = u_0$. We will show that $\{v - v_0 | v \in L\} = \ker T$.
Let $v \in L$. We have $T(v - v_0) = T(v) - T(v_0) = 0$, so $\{v - v_0 | v \in L\} \subseteq \ker T$.
Let $v_1 \in \ker T$, i.e. $T(v_1) = 0$. Write $v_1 = (v_1 + v_0) - v_0$. $T(v_1 + v_0) = u_0$, so $(v_1 + v_0) \in L$. Hence, $v_1 \in \{v - v_0 | v \in L\}$ or, in other words $\ker T \subseteq \{v - v_0 | v \in L\}$.
Consequently $L = v_0 + \ker T$, which shows that L is a linear manifold. □

The previous theorems give rise to the next:

Theorem 5.31. *Let V a linear space of dimension n. Then, for every linear manifold $L \subset V$ of dimension $\dim L = k < n$, there exists an $n - k$-dimensional vector space U, a surjective linear map $T : V \to U$ and a vector $u \in U$ such that*

$$L = \{v \in V | T(v) = u\}.$$

Proof. Indeed, consider $L = v_0 + V_L$, where the dimension of the director subspace $V_L = k$. Choose a basis $\{e_1, \dots, e_k\}$ in V_L and complete it to a basis $\{e_1, \dots, e_k, e_{k+1}, \dots, e_n\}$ of V. Consider $U = \text{span}\{e_{k+1}, \dots, e_n\}$. Obviously $\dim U = n - k$. According to a previous theorem the linear map
$T : V \to U$, $T(\alpha_1 e_1 + \cdots + \alpha_k e_k + \alpha_{k+1} e_{k+1} + \cdots + \alpha_n e_n) = \alpha_{k+1} e_{k+1} + \cdots + \alpha_n e_n$ is surjective and $\ker T = V_L$. Let $T(v_0) = u$. Then, according to the proof of the previous theorem $L = \{v \in V | T(v) = u\}$. □

Remark 5.32. If we choose in V and U two bases and we write the linear map by matrix notation $M_T v = u$ we have the implicit equations of the linear manifold L,

$$\begin{cases} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n = & u_1 \\ \vdots & \\ a_{p1}v_1 + a_{p2}v_2 + \cdots + a_{pn}v_n = & u_p \end{cases}$$

where $p = n - k = \dim U = \operatorname{rank}\,(a_{ij})_{\substack{i=\overline{1,p} \\ j=\overline{1,n}}}$.

A hyperplane has only one equation

$$a_1v_1 + \cdots + a_nv_n = u_0$$

The director subspace can be seen as

$$V_L = \{v = v_1e_1 + \cdots + v_ne_n | f(v) = 0\} = \ker f,$$

where f is the linear map (linear functional) $f : V \to \mathbb{R}$ with
$f(e_1) = a_1, \ldots, f(e_n) = a_n$.
If we think of the hyperplane as a linear manifold in the Euclidean space $\mathbb{R}^n$, the equation can be written as

$$\langle v, a\rangle = u_0, \text{where } \ a = a_1e_1 + \cdots + a_ne_n,\, u_0 \in \mathbb{R}.$$

The vector a is called *the normal vector* to the hyperplane.
Generally in a Euclidean space the equations of a linear manifold are

$$\begin{cases} \langle v, v_1\rangle = & u_1 \\ \vdots & \\ \langle v, v_p\rangle = & u_p \end{cases}$$

where the vectors $v_1, \dots v_p$ are linearly independent. The director subspace is given by

$$\begin{cases} \langle v, v_1 \rangle = & 0 \\ \vdots & \\ \langle v, v_p \rangle = & 0 \end{cases}$$

so, the vectors $v_1, \dots, v_p$ are orthogonal to the director subspace V_L.

5.5 The Gram determinant. Distances.

In this section we will explain how we can measure the distance between some "linear sets", which are linear manifolds.

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and consider the vectors $v_i \in V\ ,\ i = \overline{1, k}$.

The determinant

$$G(v_1, \dots, v_k) = \begin{vmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \dots & \langle v_1, v_k \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \dots & \langle v_2, v_k \rangle \\ \dots\dots & \dots & \dots & \\ \langle v_k, v_1 \rangle & \langle v_k, v_2 \rangle & \dots & \langle v_k, v_k \rangle \end{vmatrix}$$

is called the *Gram determinant* of the vectors $v_1 \dots v_k$.

Proposition 5.33. *In an inner product space the vectors $v_1, \dots, v_k$ are linearly independent iff $G(v_1, \dots, v_k) \neq 0$.*

Proof. Let us consider the homogenous system

$$G \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This system can be written as

$$\begin{cases} \langle v_1, v\rangle = 0 \\ \vdots \\ \langle v_k, v\rangle = 0 \end{cases} \quad \text{where } v = x_1v_1 + \dots x_kv_k.$$

The following statements are equivalent.

The vectors $v_1, \dots, v_k$ are linearly dependent. $\Longleftrightarrow$ There exist $x_1, \dots, x_k \in \mathbb{F}$, not all zero such that $v = 0$. $\Longleftrightarrow$ The homogenous system has a nontrivial solution. $\Longleftrightarrow \det G = 0$. $\square$

Proposition 5.34. *If* $\{e_1, \dots, e_n\}$ *are linearly independent vectors and* $\{f_1, \dots, f_n\}$ *are vectors obtained by Gram Schmidt orthogonalization process, one has:*

$$G(e_1, \dots, e_n) = G(f_1, \dots, f_n) = \|f_1\|^2 \cdot \dots \cdot \|f_n\|^2$$

Proof. In $G(f_1, \dots, f_n)$ replace f_n by $e_n - a_1f_1 - \dots - a_{n-1}f_{n-1}$ and we obtain

$$G(f_1, \dots, f_n) = G(f_1, \dots, f_{n-1}, e_n).$$

By an inductive process the relation in the theorem follows. Obviously $G(f_1, \dots, f_n) = \|f_1\|^2 \cdot \dots \cdot \|f_n\|^2$ because in the determinant we have only on the diagonal $\langle f_1, f_1\rangle, \dots, \langle f_n, f_n\rangle$. $\square$

Remark 5.35. Observe that:

- $\|f_k\| = \sqrt{\dfrac{G(e_1, \dots e_k)}{G(e_1, \dots, e_{k-1})}}$
- $f_k = e_k - a_1f_1 - \dots a_{k-1}f_{k-1} = e_k - v_k$ one obtains $e_k = f_k + v_k$, $v_k \in \operatorname{span}\{e_1, \dots, e_{k-1}\}$ and $f_k \in \operatorname{span}\{e_1, \dots, e_{k-1}\}^\perp$, so f_k is the orthogonal complement of e_k with respect to the space generated by $\{e_1 \dots, e_{k-1}\}$.

The distance between a vector and a subspace

Let U be a subspace of the inner product space V. The distance between a vector v and the subspace U is

$$d(v,U) = \inf_{w \in U} d(v,w) = \inf_{w \in U} \|v - w\|.$$

Remark 5.36. The linear structure implies a very simple but useful fact:

$$d(v,U) = d(v+w, w+U)$$

for every $v, w \in V$ and $U \subseteq V$, that is the linear structure implies that the distance is invariant by translations.

We are interested in the special case when U is a subspace.

Proposition 5.37. *The distance between a vector $v \in V$ and a subspace U is given by*

$$d(v,U) = \|v^{\perp}\| = \sqrt{\frac{G(e_1,\ldots,e_k,v)}{G(e_1,\ldots,e_k)}},$$

where $v = v_1 + v^{\perp},\ v_1 \in U,\ v^{\perp} \in U^{\perp}$ and $e_1,\ldots,e_k$ is a basis in U.

Proof. First we prove that $\|v^{\perp}\| = \|v - v_1\| \leq \|v - u\|, \quad \forall u \in U$. We have

$$\begin{aligned} \|v^{\perp}\| &\leq \|v - u\| \Leftrightarrow \\ \langle v^{\perp}, v^{\perp}\rangle &\leq \langle v^{\perp} + v_1 - u, v^{\perp} + v_1 - u\rangle \Leftrightarrow \\ \langle v^{\perp}, v^{\perp}\rangle &\leq \langle v^{\perp}, v^{\perp}\rangle + \langle v_1 - u, v_1 - u\rangle. \end{aligned}$$

The second part of the equality, i.e. $\|v^{\perp}\| = \sqrt{\frac{G(e_1,\ldots,e_k,v)}{G(e_1,\ldots,e_k)}}$, follows from the previous remark. □

Definition 5.38. *If $e_1,\ldots,e_k$ are vectors in V the volume of the k- parallelepiped constructed on the vectors $e_1,\ldots,e_k$ is defined by $\mathcal{V}_k(e_1,\ldots,e_k) = \sqrt{G(e_1,\ldots,e_k)}$.*

We have the following inductive relation

$$\mathcal{V}_{k+1}(e_1,\dots,e_k,e_{k+1}) = \mathcal{V}_k(e_1,\dots,e_k)d(e_{k+1},\operatorname{span}\{e_1,\dots,e_k\}).$$

The distance between a vector and a linear manifold

Let $L = v_0 + V_L$ be a linear manifold, and let v be a vector in a finitely dimensional inner product space V. The distance induced by the norm is invariant by translations, that is, for all $v_1, v_2 \in V$ one has

$$d(v_1,v_2) = d(v_1+v_0, v_1+v_0) \Leftrightarrow \|v_1 - v_2\| = \|v_1 + v_0 - (v_2+v_0)\|$$

That means that we have

$$\begin{aligned} d(v,L) &= \inf_{w\in L} d(v,w) = \inf_{v_L\in V_L} d(v, v_0 + v_L) \\ &= \inf_{v_L\in V_L} d(v - v_0, v_L) \\ &= d(v - v_0, V_L). \end{aligned}$$

Finally,

$$d(v,L) = d(v-v_0,V_L) = \sqrt{\frac{G(e_1,\dots,e_k,v-v_0)}{G(e_1,\dots,e_k)}},$$

where $e_1,\dots,e_k$ is a basis in V_L.

Example 5.39. Consider the linear manifolds
$L = \{(x,y,z,t)\in\mathbb{R}^4 | x+y+t=2, x-2y+z+t=3\}$,
$K = \{(x,y,z,t)\in\mathbb{R}^4 | x+y+z-t=1, x+y+z+t=3\}$. Find the director subspaces V_L, V_K and a basis in $V_L\cap V_K$. Find the distance of $v=(1,0,2,2)$ from L, respectively K, and show that the distance between L and K is 0.

Since $L = v_0 + V_L$ and $K = u_0 + V_K$ it follows that $V_L = L - v_0$ and $V_K = K - u_0$ for some $v_0 \in L$, $u_0 \in K$. By taking $x = y = 0$ in the equations that describe L we obtain $t = 2$, $z = 1$, hence $v_0 = (0, 0, 1, 2) \in L$. Analogously $u_0 = (0, 0, 2, 1) \in K$. Hence the director subspaces are

$$V_L = \{(x, y, z-1, t-2) \in \mathbb{R}^4 | x + y + t = 2, x - 2y + z + t = 3\} =$$

$$\{(x, y, z, t) \in \mathbb{R}^4 | x + y + t = 0, x - 2y + z + t = 0\},$$

respectively

$$V_K = \{(x, y, z-2, t-1) \in \mathbb{R}^4 | x + y + z - t = 1, x + y + z + t = 3\} =$$

$$\{(x, y, z, t) \in \mathbb{R}^4 | x + y + z - t = 0, x + y + z + t = 0\}.$$

By solving the homogenous systems $\begin{cases} x + y + t = 0 \\ x - 2y + z + t = 0 \end{cases}$, respectively $\begin{cases} x + y + z - t = 0 \\ x + y + z + t = 0 \end{cases}$ we obtain that

$$V_L = \mathrm{span}\{e_1 = (-1, 1, 3, 0), e_2 = (-1, 0, 0, 1)\},$$

respectively

$$V_K = \mathrm{span}\{e_3 = (-1, 1, 0, 0), e_4 = (-1, 0, 1, 0)\}.$$

Since $\det[e_1|e_2|e_3|e_4] = 3 \neq 0$ the vectors e_1, e_2, e_3, e_4 are linearly independent, hence $V_L \cap V_K = \{0\}$. The distance of v from L is

$$d(v, L) = d(v - v_0, V_L) = \sqrt{\frac{G(e_1, e_2, v - v_0)}{G(e_1, e_2)}} = \sqrt{\frac{19}{21}},$$

meanwhile

$$d(v, K) = d(v - v_0, V_K) = \sqrt{\frac{G(e_3, e_4, v - v_0)}{G(e_3, e_4)}} = \sqrt{\frac{4}{3}}.$$

It is obvious that $K \cap L \neq \emptyset$, since the system $\begin{cases} x+y+t=2 \\ x-2y+z+t=3 \\ x+y+z-t=1 \\ x+y+z+t=3 \end{cases}$ is consistent, having solution $(1,0,1,1)$, hence we must have

$$d(L,K)=0.$$

Let us consider now the hyperplane H of equation

$$\langle v-v_0, n\rangle = 0 \ .$$

The director subspace is $V_H = \langle v, n\rangle = 0$ and the distance

$$d(v,H) = d(v-v_0, V_H).$$

One can decompose $v - v_0 = \alpha n + v_H$, where v_H is the orthogonal projection of $v - v_0$ on V_H and αn is the normal component of $v - v_0$ with respect to V_H. It means that

$$d(v,H) = \|\alpha n\|$$

Let us compute a little now, taking into account the previous observations about the tangential and normal part:

$$\begin{aligned} \langle v-v_0, n\rangle &= \langle \alpha n + v_H, n\rangle \\ &= \alpha\langle n,n\rangle + \langle v_H, n\rangle \\ &= \alpha\|n\|^2 + 0 \end{aligned}$$

So, we obtained

$$\frac{|\langle v-v_0,n\rangle|}{\|n\|} = |\alpha|\|n\| = \|\alpha n\|$$

that is

$$d(v, H) = \frac{|\langle v - v_0, n\rangle|}{\|n\|}$$

In the case that we have an orthonormal basis at hand, the equation of the hyperplane H is

$$a_1x_1 + \cdots + a_kx_k + b = 0\ ,$$

so the relation is now

$$d(v, H) = \frac{|a_1v_1 + \cdots + a_kv_k + b|}{\sqrt{a_1^2 + \cdots + a_k^2}}\ .$$

The distance between two linear manifolds

For A and B sets in a metric space, the distance between them is defined as

$$d(A, B) = \inf\{d(a, b)|a \in A\ ,\ \ b \in B\}.$$

For two linear manifolds $L_1 = v_1 + V_1$ and $L_2 = v_2 + V_2$ it easily follows:

$$\begin{aligned} d(L_1, L_2) &= d(v_1 + V_1, v_2 + V_2) = d(v_1 - v_2, V_1 - V_2) && (5.1)\\ &= d(v_1 - v_2, V_1 + V_2). && (5.2)\end{aligned}$$

This gives us the next proposition.

Proposition 5.40. *The distance between the linear manifolds $L_1 = v_1 + V_1$ and $L_2 = v_2 + V_2$ is equal to the distance between the vector $v_1 - v_2$ and the sum space $V_1 + V_2$.*

If we choose a basis in $V_1 + V_2$, let's say $e_1, \ldots, e_k$, then this formula follows:

$$d(L_1, L_2) = \sqrt{\frac{G(e_1, \ldots, e_k, v_1 - v_2)}{G(e_1, \ldots, e_k)}}.$$

Some analytic geometry

In this section we are going to apply distance problems in Euclidean spaces. Consider the vector space $\mathbb{R}^n$ with the canonical inner product, that is: for $\overline{x} = (x_1, \ldots, x_n)$, $\overline{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ the inner product is given by

$$\langle \overline{x}, \overline{y} \rangle = \sum_{i=1}^{n} x_k y_k.$$

Consider D_1 , D_2 two lines (one dimensional linear manifolds), M a point (zero dimensional linear manifold, we assimilate with the vector $\overline{x}_M = \overline{0M}$), P a two dimensional linear manifold (a plane), and H an $n-1$ dimensional linear manifold (hyperplane). The equations of these linear manifolds are:

$$D_1 \; : \; \overline{x} = \overline{x}_1 + s\overline{d}_1,$$

$$D_2 \; : \; \overline{x} = \overline{x}_2 + t\overline{d}_2,$$

$$M \; : \; \overline{x} = \overline{x}_M,$$

$$P \; : \; \overline{x} = \overline{x}_P + \alpha\overline{v}_1 + \beta\overline{v}_2,$$

respectively

$$H \; : \; \langle \overline{x}, \overline{n} \rangle + b = 0,$$

where $s, t, \alpha, \beta, b \in \mathbb{R}$. Recall that two linear manifolds are parallel if the director space of one of them is included in the director space of the other.

Now we can write down several formulas for distances between linear manifolds.

$$\begin{aligned}
d(M, D_1) &= \sqrt{\frac{G(\overline{x}_M - \overline{x}_1, \overline{d}_1)}{G(\overline{d}_1)}}; \\
d(M, P) &= \sqrt{\frac{G(\overline{x}_M - \overline{x}_P, \overline{v}_1, \overline{v}_2)}{G(\overline{v}_1, \overline{v}_2)}}; \\
d(D_1, D_2) &= \sqrt{\frac{G(\overline{x}_1 - \overline{x}_2, \overline{d}_1, \overline{d}_2)}{G(\overline{d}_1, \overline{d}_2)}} \quad \text{if} \quad D_1 \nparallel D_2 \\
d(D_1, D_2) &= \sqrt{\frac{G(\overline{x}_1 - \overline{x}_2, \overline{d}_1)}{G(\overline{d}_1,)}} \quad \text{if} \quad D_1 \parallel D_2 \\
d(M, H) &= \frac{|\langle \overline{x}_M, \overline{n} \rangle + b|}{\|\overline{n}\|} \\
d(D_1, P) &= \sqrt{\frac{G(\overline{x}_1 - \overline{x}_P, \overline{d}_1, \overline{v}_1, \overline{v}_2)}{G(\overline{d}_1, \overline{v}_1, \overline{v}_2)}} \quad \text{if} \quad D_1 \nparallel P
\end{aligned}$$

Example 5.41. Find the distance between the hyperplane $H = \{(x, y, z, t) \in \mathbb{R}^4 : x + y + z + t = 1\}$ and the line $D = \{(x, y, z, t) \in \mathbb{R}^4 : x + y + z + t = 3, x - y - 3z - t = -1, 2x - 2y + 3z + t = 1\}$.

Since $v_0 = (0, 0, 0, 1) \in H$ its director subspace is $V_H = \{(x, y, z, t) \in \mathbb{R}^4 : x + y + z + t = 0\} = \text{spam}\{e_1 = (1, 0, 0, -1), e_2 = (0, 1, 0, -1), e_3 = (0, 0, 1, -1)\}$. Since $u_0 = (1, 1, 0, 1) \in D$ its director subspace is $V_D = \{(x, y, z, t) \in \mathbb{R}^4 : x + y + z + t = 0, x - y - 3z - t = 0, 2x - 2y + 3z + t = 0\} = \text{spam}\{e_4 = (1, 1, 1, -3)\}$.

We have $e_4 = e_1 + e_2 + e_3$ hence $V_D \subset V_H$ that is D and H are parallel. Obviously one can compute their distance by the formula

$$d(D, H) = \sqrt{\frac{G(e_1, e_2, e_3, v_0 - u_0)}{G(e_1, e_2, e_3)}}.$$

But, observe that the distance between these manifolds is actually the distance between a point $M \in D$ and H, hence is more simple to compute from the formula

$d(M,H) = \frac{|\langle \overline{x_M}, \overline{n} \rangle + b|}{\|\overline{n}\|}$, with $\overline{x_M} = u_0$. Indeed the equation of H is $x+y+z+t=1$, thus $\overline{n} = (1,1,1,1)$ and $b = -1$, hence

$$d(D,H) = \frac{|\langle (1,1,0,1), (1,1,1,1) \rangle - 1|}{\|(1,1,1,1)\|} = \frac{2}{2} = 1.$$

5.6 Problems

Problem 5.6.1. Prove that for the nonzero vectors $x, y \in \mathbb{R}^2$, it holds

$$\langle x, y \rangle = \|x\| \|y\| \cos \theta,$$

where θ is the angle between x and y.

Problem 5.6.2. Find the angle between the vectors $(-2,4,3)$ and $(1,-2,3)$.

Problem 5.6.3. Find the two unit vectors which are orthogonal to both of the vectors $(-2,3,-1)$ and $(1,1,1)$.

Problem 5.6.4. Let $u, v \in V$, V an inner product space. Show that

$$\|u\| \leq \|u + av\|, \forall a \in \mathbb{F} \Leftrightarrow \langle u, v \rangle = 0.$$

Problem 5.6.5. Prove that

$$(\sum_{i=1}^{n} a_i b_i)^2 \leq (\sum_{i=1}^{n} i a_i^2)(\sum_{i=1}^{n} \frac{1}{i} b_i^2),$$

for all $a_i, b_i \in \mathbb{R}$, $, i = \overline{1, n}$.

Problem 5.6.6. Let S be the subspace of the inner product space $\mathbb{R}_3[X]$, the space of polynomials of degree at most 3, generated by the polynomials $1 - x^2$ and $2 - x + x^2$, where $\langle f, g \rangle = \int_0^1 f(x) g(x) dx$. Find a basis for the orthogonal complement of S.

Problem 5.6.7. Let $u, v \in V$, V an inner product space. If

$$\|u\| = 3\ ,\ \|u+v\| = 4\ ,\ \|u-v\| = 6,$$

find $\|v\|$.

Problem 5.6.8. Prove or infirm the following statement: There exists an inner product on $\mathbb{R}^2$ such that the norm induced by this scalar product satisfies

$$\|(x_1, x_2)\| = |x_1| + |x_2|,$$

for all $(x_1, x_2) \in \mathbb{R}^2$.

Problem 5.6.9. Show that the planes $P:\ x - 3y + 4z = 12$ and $P_2:\ 2x - 6y + 8z = 6$ are parallel and then find the distance between them.

Problem 5.6.10. Let V be an inner product space. Then it holds:

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}\ ,\ \forall u, v \in V.$$

Problem 5.6.11. If V is a complex vector space with an inner product on it, show that

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2}{4}\ ,\ \forall\ u, v \in V.$$

Problem 5.6.12. Prove that the following set

$$\left\{\frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}\right\}$$

is orthonormal in $C[-\pi, \pi]$, endowed with the scalar product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx.$$

Problem 5.6.13. Show that the set of all vectors in $\mathbb{R}^n$ which are orthogonal to a given vector $v \in \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$. What will its dimension be?

Problem 5.6.14. If S is a subspace of a finite dimensional real inner product space V, prove that $S^\perp \simeq V/S$.

Problem 5.6.15. Let V be an inner product space and let $\{v_1, \ldots, v_m\}$ a list of linearly independent vectors from V. How many orthonormal families $\{e_1, \ldots, e_m\})$ can be constructed by using the Gram-Schmidt procedure, such that

$$\text{span}\{v_1, \ldots, v_i\} = \text{span}\{e_1, \ldots, e_i\}, \ \forall \ \ i = \overline{1, m}.$$

Problem 5.6.16. Orthonormalize the following list of vectors in $\mathbb{R}^4$ $\{(1, 11, 0, 1), (1, -2, 1, 1), (1, 1, 1, 0), (1, 1, 1, 1)\}$.

Problem 5.6.17. Let V be an inner product space and let $U \subseteq V$ subspace. Show that

$$\dim U^\perp = \dim V - \dim U.$$

Problem 5.6.18. Let $\{e_1, \ldots, e_m\}$ be an orthonormal list in the inner product space V. Show that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$$

if and only if $v \in span\{e_1, \ldots, e_m\}$.

Problem 5.6.19. Let V be a finite-dimensional real inner product space with a basis $\{e_1, \ldots, e_n\}$. Show that for any $u, w \in V$ it holds $\langle u, w\rangle = [u]^\top G(e_1, \ldots, e_n)[w]$ where $[u]$ is the coordinate vector (represented as a column matrix) of u with respect to the given basis and $G(e_1, \ldots, e_n)$ is the matrix having the same entries as the Gram determinant of $\{e_1, \ldots, e_n\}$..

Problem 5.6.20. Find the distance between the following linear manifolds.

a) $L = \{(x, y, z, t) \in \mathbb{R}^4 | x + y + t = 1, x - 2y + z = -1\}, K = \{(x, y, z, t) \in \mathbb{R}^4 | y + 2z - t = 1, x + y + z + t = 2, x - y - 2z = -4\}$.

b) $L = \{(x, y, z, t) \in \mathbb{R}^4 | x + y + t = 2, x - 2y + z = 3\}, K = \{(x, y, z, t) \in \mathbb{R}^4 | y + z - t = 1, 2x - y + z + t = 3\}$.

c) $L = \{(x, y, z, t) \in \mathbb{R}^4 | x + z + t = 1, x + y + z = 2\}, K = \{(x, y, z, t) \in \mathbb{R}^4 | y + t = 3, x + t = 4\}$.

d) $L = \{(x, y, z, t) \in \mathbb{R}^4 | x + z + t = 1, x + y + z = 2, x - y + t = 2\}, K = \{(x, y, z, t) \in \mathbb{R}^4 | 2x + y + 2z + t = 4\}$.

Problem 5.6.21. Let V be an inner product space, let $U \subseteq V$ be an arbitrary subset and let $U_1, U_2 \subseteq V$ be subspaces. Show that $U^\perp$ is a subspace of V, and $V^\perp = 0$,respectively $0^\perp = V$. Further the following implication holds:
$U_1 \subseteq U_2 \Rightarrow U_1^\perp \supseteq U_2^\perp$.

6

Operators on inner product spaces.

6.1 Linear functionals and adjoints

A linear functional on a vector space V over the field $\mathbb{F}$ is a linear map $f : V \to \mathbb{F}$.

Example 6.1. $f : \mathbb{F}^3 \to \mathbb{F}$ given by $f(v_1, v_2, v_3) = 3v_1 + 4v_2 - 5v_3$ is a linear functional on $\mathbb{F}^3$.

Assume now that V is an inner product space. For fixed $v \in V$, the map $f : V \to \mathbb{F}$ given by $f(u) = \langle u, v\rangle$ is a linear functional. The next fundamental theorem shows that in case when V is a Hilbert space, then every linear continuous functional on V is of this form.
Recall that an inner product space is a Hilbert space if is complete, that is, every Cauchy sequence is convergent, i.e.,

$$\forall \epsilon > 0\, \exists n_\epsilon \in \mathbb{N} \text{ s.t. } \forall n, m > n_\epsilon \implies \|x_n - x_m\|_V < \epsilon$$

implies that (x_n) is convergent.

Theorem 6.2. *Suppose f is a linear continuous functional on the Hilbert space V.*

Then there is a unique vector $v \in V$ such that

$$f(u) = \langle u, v \rangle \ .$$

Proof. We will present the proof only in the finite dimensional case. We show first that there is a vector $v \in V$ such that $f(u) = \langle u, v \rangle$. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis of V. One has

$$\begin{aligned} f(u) &= f(\langle u, e_1 \rangle e_1 + \cdots + \langle u, e_n \rangle e_n) = \langle u, e_1 \rangle f(e_1) + \dots \langle u, e_n \rangle f(e_n) \\ &= \langle u, \overline{f(e_1)} e_1 + \cdots + \overline{f(e_n)} e_n \rangle \ , \end{aligned}$$

for every $u \in V$. It follows that the vector

$$v = \overline{f(e_1)} e_1 + \cdots + \overline{f(e_n)} e_n$$

satisfies $f(u) = \langle u, v \rangle$ for every $u \in V$.
It remains to prove the uniqueness of v. Suppose that there are $v_1, v_2 \in V$ such that

$$f(u) = \langle u, v_1 \rangle = \langle u, v_2 \rangle,$$

for every $u \in V$. It follows that

$$0 = \langle u, v_1 \rangle - \langle u, v_2 \rangle = \langle u, v_1 - v_2 \rangle \quad \forall \ \ u \in V$$

Taking $u = v_1 - v_2$ it follows that $v_1 = v_2$, so v is unique. □

Remark 6.3. Note that every linear functional on a finite dimensional Hilbert space is continuous. Even more, on every finite dimensional inner product space, the inner product defines a norm (metric) such that with the topology induced by this metric the inner product space is complete.

Let us consider another vector space W over $\mathbb{F}$, and an inner product on it, such that $(W, \langle\cdot,\cdot\rangle)$ is a Hilbert space.

Let $T \in L(V, W)$ be a continuous operator in the topologies induced by the norms $\|v\|_V = \sqrt{\langle v, v\rangle_V}$, respectively $\|w\|_W = \sqrt{\langle w, w\rangle_W}$, (as a continuous function in analysis). Now, we define the adjoint of T, as follows.

Fix $w \in W$ and consider the linear functional on V which maps v in $\langle T(v), w\rangle_W$. It follows that there exists a unique vector $T^*(w) \in V$ such that

$$\langle v, T^*(w)\rangle_V = \langle T(v), w\rangle_W \quad \forall v \in V.$$

The operator $T^\star : W \to V$ constructed above is called the **adjoint** of T.

Example 6.4. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by $T(x, y, z) = (y + 3z, 2x)$.

Its adjoint operator is $T^* : \mathbb{R}^2 \to \mathbb{R}^3$. Fix $(u, v) \in \mathbb{R}^2$. It follows

$$\begin{aligned}
\langle (x, y, z), T^*(u, v)\rangle &= \langle T(x, y, z), (u, v)\rangle \\
&= \langle (y + 3z, 2x), (u, v)\rangle \\
&= yu + 3zu + 2xv \\
&= \langle (x, y, z), (2v, u, 3u)\rangle
\end{aligned}$$

for all $(x, y, z) \in R^3$. This shows that

$$T^*(u, v) = (2v, u, 3u).$$

Note that in the example above T^* is not only a map from $\mathbb{R}^2$ to $\mathbb{R}^3$, but also a linear map.

We shall prove this in general. Let $T \in L(V, W)$, so we want to prove that $T^* \in L(W, V)$.

Let $w_1, w_1 \in W$. By definition one has:

$$\begin{aligned}\langle T(v), w_1 + w_2\rangle &= \langle T(v), w_1\rangle + \langle Tv, w_2\rangle \\ &= \langle v, T^*(w_1)\rangle + \langle v, T^*(w_2)\rangle \\ &= \langle v, T^*(w_1) + T^*(w_2)\rangle,\end{aligned}$$

which shows that $T^*(w_1) + T^*(w_2)$ plays the role of $T^*(w_1 + w_2)$.

By the uniqueness proved before, we have that

$$T^*(w_1) + T^*(w_2) = T^*(w_1 + w_2) \quad .$$

It remains to check the homogeneity of T^*. For $a \in \mathbb{F}$ one has

$$\begin{aligned}\langle T(v), aw\rangle &= \overline{a}\langle T(v), w\rangle \\ &= \overline{a}\langle v, T^*(w)\rangle \\ &= \langle v, aT^*(w)\rangle \ .\end{aligned}$$

This shows that $aT^*(w)$ plays the role of $T^*(aw)$, and again by the uniqueness of the adjoint we have that

$$aT^*(w) = T^*(aw) \ .$$

Thus T^* is a linear map, as claimed.

One can easily verify the following properties:

a) **additivity** $(S + T)^* = S^* + T^*$ for all $S, T \in L(V, W)$.

b) **conjugate homogeneity** $(aT)^* = \overline{a}T^*$ for all $a \in \mathbb{F}$ and $T \in L(V, W)$.

c) **adjoint of adjoint** $(T^*)^* = T$ for all $T \in L(V, W)$.

d) **identity** $I^* = I$, if $I = I_V$, $V = W$.

e) **products** $(ST)^* = T^*S^*$ for all $T \in L(V,W)$ and $S \in L(W,U)$.

For the sake of completeness we prove the above statements. Let $v \in V$ and $w \in W$.

a) Let $S, T \in L(U,W)$. Then, $\langle (S+T)(v), w\rangle = \langle v, (S+T)^*(w)\rangle$. On the other hand

$\langle (S+T)(v), w\rangle = \langle S(v), w\rangle + \langle T(v), w\rangle = \langle v, S^*(w)\rangle + \langle v, T^*(w)\rangle = \langle v, (S^*+T^*)(w)\rangle$.

Hence, $(S+T)^* = S^* + T^*$.

b) Let $a \in \mathbb{F}$ and $T \in L(U,W)$. We have $\langle (aT)(v), w\rangle = \langle v, (aT)^*(w)\rangle$. But

$\langle (aT)(v), w\rangle = a\langle T(v), w\rangle = a\langle v, T^*(w) = \langle v, \overline{a}T^*(w)\rangle$.

Hence, $(aT)^* = \overline{a}T^*(w)$.

c) Let $T \in L(U,W)$. Then

$\langle w, T(v)\rangle = \overline{\langle T(v), w\rangle} = \overline{\langle v, T^*(w)\rangle} = \langle T^*(w), v\rangle = \langle w, (T^*)^*(v)\rangle$.

Hence, $(T^*)^* = T$.

d) Let $V = W$. We have $\langle v, I(w)\rangle = \langle v, w\rangle = \langle I(v), w\rangle = \langle v, I^*(w)\rangle$.

Hence $I = I^*$.

e) Let $T \in L(V,W)$ and $S \in L(W,U)$. Then for all $u \in U$ and $v \in V$ it holds:

$\langle T^*S^*(u), v\rangle = \langle S^*(u), ((T)^*)^*(v) = \langle S^*(u), T(v)\rangle = \langle u, (S^*)^*T(v)\rangle = \langle u, ST(v)\rangle = \overline{\langle ST(v), u\rangle} = \overline{\langle v, (ST)^*(u)\rangle} = \langle (ST)^*(u), v\rangle$.

Hence, $T^*S^* = (ST)^*$.

Proposition 6.5. *Suppose that $T \in L(V,W)$ is continuous. Then*

1. $\ker T^* = (\operatorname{im} T)^\perp$.

2. $\operatorname{im} T^* = (\ker T)^\perp$.

3. $\ker T = (\operatorname{im} T^*)^\perp$.

4. $\operatorname{im} T = (\ker T^*)^\perp$.

Proof. 1. Let $w \in W$. Then

$$\begin{aligned} w \in \ker T^* &\Leftrightarrow T^*(w) = 0 \\ &\Leftrightarrow \langle v, T^*(w)\rangle = 0 \quad \forall\, v \in V \\ &\Leftrightarrow \langle T(v), w\rangle = 0 \quad \forall v \in V \\ &\Leftrightarrow w \in (\operatorname{im} T)^{\perp} \quad , \end{aligned}$$

that is $\ker T^* = (\operatorname{im} T)^{\perp}$. If we take the orthogonal complement in both sides we get 4. Replacing T by T^* in 1 and 4 gives 3 and 2. $\square$

The conjugate transpose of a type (m, n)- matrix is an (n, m) matrix obtained by interchanging the rows and columns and taking the complex conjugate of each entry. The adjoint of a matrix (which is a linear transform between two finite dimensional spaces in the appropriate bases) is the conjugate transpose of that matrix as the next result shows.

Proposition 6.6. *Suppose that $T \in L(V, W)$. If $\{e_1, \ldots, e_n\}$, and $\{f_1, \ldots, f_m\}$ are orthonormal bases for V and W respectively, and we denote by M_T and M_{T^*} the matrices of T and T^* in these bases, then M_{T^*} is the conjugate transpose of M_T.*

Proof. The k^{th} column of M_T is obtained by writing $T(e_k)$ as linear combination of f_j's, the scalars used became the k^{th} column of M_T. Being the basis with f_j's orthonormal, it follows that

$$T(e_k) = \langle T(e_k), f_1\rangle f_1 + \cdots + \langle T(e_k), f_m\rangle f_m.$$

So on the position (k, j) of M_T we have $\langle T(e_k), f_j\rangle$. Replacing T with T^* and interchanging the roles played by e's and$f$'s, we see that the entry on the position (j, k) of M_{T^*} the entry is $\langle T^*(f_k), e_j\rangle$, which equals to $\langle f_k, T(e_j)\rangle$, which equals to $\overline{\langle T(e_j), f_k\rangle}$. In others words, M_{T^*} equals to the complex conjugate of M_T. $\square$

6.2 Normal operators

An operator on a Hilbert space is called normal if it commutes with its adjoint, that is

$$TT^* = T^*T\ .$$

Remark 6.7. *We will call a complex square matrix normal if it commutes with its conjugate transpose, that is $A \in \mathcal{M}_n(\mathbb{C})$ is normal iff*

$$AA^* = A^*A.$$

Here $A^* = \overline{A}^\top$ is the conjugate transpose of A.
It can be easily be observed that the matrix of a normal operator is a normal matrix.

Example 6.8. On $\mathbb{F}^2$ consider the operator which in the canonical basis has the matrix

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

This is a normal operator.

Indeed let $T : \mathbb{F}^2 \to \mathbb{F}^2$ be the operator whose matrix is A. Then $T(x,y) = (2x-3y, 3x+2y)$, thus $\langle T(x,y),(u,v)\rangle = (2x-3y)u + (3x+2y)v = (2v-3u)y + (3v+2u)x = \langle (x,y),(3v+2u, 2v-3u)\rangle$. Hence, the adjoint of T is $T^*(u,v) = (2u+3v, -3u+2v)$. It can easily be computed that $TT^*(u,v) = T^*T(u,v) = (13u, 13v)$, hence $TT^* = T^*T$.

Proposition 6.9. *An operator $T \in L(V)$ is normal operator iff*

$$\|T(v)\| = \|T^*(v)\| \textit{ for all } v \in V.$$

Proof. Let $T \in L(V)$.

$$\begin{aligned} T \text{ is normal} &\iff T^*T - TT^* = 0 \\ &\iff \langle (T^*T - TT^*)(v), v\rangle = 0 \text{ for all } v \in V \\ &\iff \langle T^*T(v), v\rangle = \langle TT^*(v), v\rangle \text{ for all } v \in V \\ &\iff \|T(v)\|^2 = \|T^*(v)\|^2 \text{ for all } v \in V. \end{aligned}$$

□

Theorem 6.10. *Let T be a normal operator on V and λ_0 be an eigenvalue of T.*

1. *The proper subspace $E(\lambda_0)$ is T^* invariant.*
2. *If v_0 is an eigenvector of T corresponding to the eigenvalue λ_0, then v_0 is an eigenvector of T^* corresponding to the eigenvalue $\overline{\lambda}_0$.*
3. *Let v, w be two eigenvectors corresponding to distinct eigenvalues λ, β. Then v, w are orthogonal.*

Proof. Let $v \in E(\lambda_0)$. We have to prove that $T^*(v) \in E(\lambda_0)$.
Since $T(v) = \lambda_0 v$, we have

$$T(T^*(v)) = (TT^*)(v) = (T^*T)(v) = T^*(T(v)) = T^*(\lambda_0 v) = \lambda_0 T^*(v).$$

which show that $T^*(v) \in E(\lambda_0)$.
For the second statement in the theorem we have that $T(v_0) = \lambda_0 v_0$. Let $w \in E(\lambda_0)$. Then

$$\begin{aligned} \langle T^*(v_0), w\rangle &= \langle v_0, T(w)\rangle \\ &= \langle v_0, \lambda_0 w\rangle = \overline{\lambda}_0 \langle v_0, w\rangle \\ &= \langle \overline{\lambda}_0 v_0, w\rangle. \end{aligned}$$

This means that

$$\langle T^*(v_0) - \overline{\lambda}_0 v_0, w\rangle = 0 \ ,$$

for all $w \in E(\lambda_0)$. The first term in the inner product lives in $E(\lambda_0)$ by the previous statement. Take $w = T^*(v_0) - \overline{\lambda}_0 v_0$ and it follows that $T^*(v_0) = \overline{\lambda}_0 v_0$, i.e., the second assertion of the theorem holds true.

Now, we deal with the last statement. One has $T(v) = \lambda v$ and $T(\beta) = \beta w$. By the previous point $T^*(w) = \overline{\beta} w$, so

$$\langle T(v), w\rangle = \langle v, T^*(w)$$

(def. of adjoint), which implies $\lambda\langle v, w\rangle = \beta\langle v, w\rangle$. Since $\lambda \neq \beta$, it follows that $\langle v, w\rangle = 0$. □

Proposition 6.11. *If U is a T invariant subspace of V then $U^\perp$ is a T^* invariant subspace of V.*

Proof. We have

$$w \in U^\perp \ , \ v \in V \Longrightarrow w \in U^\perp \ , \ T(v) \in U \Longrightarrow \langle v, T^*(w)\rangle = \langle T(v), w\rangle.$$

That is $T^*(w) \in U^\perp$. □

A unitary space is an inner product space over $\mathbb{C}$.

Theorem 6.12. *Suppose that V is a finite dimensional unitary space, and $T \in L(V)$ is an operator. Then T is normal iff there exists an orthonormal basis B of V relative to which the matrix of T is diagonal.*

Proof. First suppose that T has a diagonal matrix. The matrix of T^* is the complex transpose, so it is again diagonal. Any two diagonal matrices commutes, that means that T is normal.

To prove the other direction suppose that T is normal. Then, there is a basis $\{e_1, \dots, e_n\}$ of V with respect to which the matrix of T is upper triangular, that is

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{n,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}.$$

We will show that the matrix A is actually a diagonal one.

We have

$$\|T(e_1)\| = \sqrt{|a_{1,1}|^2}$$

and

$$\|T^*(e_1)\| = \sqrt{|a_{1,1}|^2 + \dots + |a_{1,n}|^2}.$$

Because T is normal the norms are equal, so $a_{1,2} = \dots = a_{1,n} = 0$.

$$\|T(e_2)\| = \sqrt{|a_{1,2}|^2 + |a_{2,2}|^2} = \sqrt{|a_{2,2}|^2}$$

and

$$\|T^*(e_2)\| = \sqrt{|a_{2,2}|^2 + \dots + |a_{2,n}|^2}.$$

Because T is normal the norms are equal, so $a_{2,3} = \dots = a_{2,n} = 0$.

By continuing the procedure we obtain that for every $k \in \{1, \dots, n-1\}$ we have $a_{k,k+1} = \dots = a_{k,n} = 0$, hence A is diagonal.

□

Theorem 6.13. ***(Complex spectral theorem)*** *Suppose that V is a unitary space. Then T has an orthonormal basis consisting of eigenvectors iff T is normal.*

Proof. Induction on $n = \dim V$. The statement is obvious for $n = 1$. Suppose that this is true for all dimensions less than n. Let $T \in L(V)$. Then T has at least one eigenvalue λ. If $\dim E(\lambda) = n$ it is enough to construct an orthonormal basis of $E(\lambda)$. For $\dim E(\lambda) < n$, choose $E^{\perp}(\lambda)$, and we have $0 < \dim E^{\perp}(\lambda) < n$. Now $E(\lambda)$ is T^* invariant, so $E^{\perp}(\lambda)$ is T invariant. By the induction hypothesis, $E^{\perp}(\lambda)$ has an orthonormal basis consisting of eigenvectors of T. Add this to the orthonormal basis of $E(\lambda)$. The result is an orthonormal basis of V consisting of eigenvectors. $\square$

6.3 Isometries

An operator $T \in L(V)$ is called an isometry if

$$\|T(v)\| = \|v\| \text{ , for all } v \in V.$$

Example 6.14. Let I be the identity map of V (V complex vector space), and $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. The map λI is an isometry, since
$\|\lambda I(v)\| = \|\lambda v\| = |\lambda|\|v\| = \|v\|$.
If T is an isometry it easily follows that T is injective.
Indeed, assume the contrary, that is, there exists $u, v \in V$, $u \neq v$ such that $T(u) = T(v)$. Hence, $0 = \|T(u) - T(v)\| = \|T(u - v)\| = \|u - v\|$, contradiction with $u \neq v$.

Theorem 6.15. *Suppose $T \in L(V)$. The following are equivalent:*

1. *T is an isometry.*
2. *$\langle T(u), T(v)\rangle = \langle u, v\rangle$ for every $u, v \in V$.*
3. *$T^*T = I$.*
4. *$\{T(e_1), \ldots, T(e_m)\}$ is an orthonormal list for every orthonormal list $\{e_1, \ldots, e_m\}$.*

5. *There exists an orthonormal basis* $\{e_1, \ldots, e_n\}$ *of* V *such that* $(T\{e_1), \ldots, T(e_n)\}$ *is an orthonormal basis.*

6. T^* *is an isometry.*

7. $\langle T^*(u), T^*(v)\rangle = \langle u, v\rangle$ *for all* $u, v \in V$.

8. $TT^* = I$

9. $\{T^*(e_1), \ldots, T^*(e_m)\}$ *is an orthonormal list for every orthonormal list* $(e_1, \ldots, e_m)$.

10. *There exists an orthonormal basis* $\{e_1, \ldots, e_n\}$ *of* V *such that* $\{T^*(e_1), \ldots, T^*(e_n)\}$ *is an orthonormal basis.*

Proof. Suppose that 1 holds. Let $u, v \in V$. Then

$$\begin{aligned} \|u - v\|^2 = \|T(u - v)\|^2 & \\ &= \langle T(u) - T(v), T(u) - T(v)\rangle \\ &= \|T(u)\|^2 + \|T(v)\|^2 - 2\langle T(u), T(v)\rangle \\ &= \|u\|^2 + \|v\|^2 - 2\langle T(u), T(v)\rangle. \end{aligned}$$

On the other hand $\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle$.

Suppose now that 2 holds. Then

$$\langle (T^*T - I)(u), v\rangle = \langle T(u), T(v)\rangle - \langle u, v\rangle = 0 \ .$$

for every $u, v \in V$. Take $v = (T^*T - I)(u)$ and it follows that $T^*T - I = 0$, i.e. 3.

Suppose 3 holds. Let $(e_1 \ldots e_m)$ be an orthonormal list of vectors in V. Then

$$\langle T(e_j), T(e_k)\rangle = \langle T^*T(e_j), e_k\rangle = \langle e_j, e_k\rangle,$$

i.e. 4 holds. Obviously 4 implies 5.

Suppose 5 holds. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of V such that $\{T(e_1), \ldots, T(e_n)\}$ is orthonormal basis. For $v \in V$

$$\begin{aligned}
\|T(v)\|^2 &= \|T(\langle v, e_1\rangle e_1 + \cdots + \langle v, e_n\rangle e_n)\|^2 \\
&= \|\langle v, e_1\rangle T(e_1) + \cdots + \langle v, e_n\rangle T(e_n)\|^2 \\
&= |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 \\
&= \|v\|^2 .
\end{aligned}$$

Taking square roots we see that T is an isometry. We have now $1 \Longrightarrow 2 \Longrightarrow 3 \Longrightarrow 4 \Longrightarrow 5 \Longrightarrow 1$. Replacing T by T^* we see that 6 through 10 are equivalent. We need only to prove the equivalence of one assertion in the first group with one in the second group.

$3 \Leftrightarrow 8$ which is easy to see since $TT^* = I \Rightarrow$
$TT^*(u) = u$, $\forall u \in V \Rightarrow (TT^*)(T(u)) = T(u)$, $\forall u \in V$, or equivalently $T((T^*T)(u)) = T(u), \forall u \in V$, T is injective, hence $T^*T = I$.
Conversely, $T^*T = I \Rightarrow T^*T(u) = u$, $\forall u \in V \Rightarrow (T^*T)(T^*(u)) = T^*(u)$, $\forall u \in V$, or equivalently $T^*((TT^*)(u)) = T^*(u), \forall u \in V$, T^* is injective, hence $TT^* = I$. □

Remark 6.16. Recall that a real square matrix A is called orthogonal iff $AA^\top = A^\top A = I$. A complex square matrix B is called unitary if $BB^* = B^*B = I$, where B^* is the conjugate transpose of B, that is $B^* = \overline{B}^\top$. It can easily be observed that the matrix of an isometry on a real (complex) finite dimensional inner product space is an orthogonal (unitary) matrix.

The last theorem shows that every isometry is a normal operator. So, the characterizations of normal operators can be used to give a complete description of isometries.

Theorem 6.17. *Suppose that V is a complex inner product space and $T \in L(V)$.*

Then T is an isometry iff there is an orthonormal basis of T consisting of eigenvectors of T which correspond to eigenvalues having modulus 1.

Proof. Suppose that there is an othonormal basis $\{e_1, \ldots, e_n\}$ consisting of eigenvectors whose corresponding eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ have absolute value 1. It follows that for every $v \in V$

$$\begin{aligned} T(v) &= \langle v, e_1 \rangle T(e_1) + \cdots + \langle v, e_n \rangle T(e_n) \\ &= \lambda_1 \langle v, e_1 \rangle e_1 + \cdots + \lambda_n \langle v, e_n \rangle e_n. \end{aligned}$$

Thus $\|T(v)\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2 = \|v\|^2$ that is

$$\|T(v)\| = \|v\|.$$

Now we are going to prove the other direction. Suppose T is an isometry. By the complex spectral theorem there is an orthonormal basis of V consisting of eigenvectors $\{e_1, \ldots, e_n\}$.. Let e_j, $j \in \{1, \ldots, n\}$ be such a vector, associated to an eigenvalue λ_j. It follows that

$$|\lambda_j| \|e_j\| = \|\lambda_j e_j\| = \|T(e_j)\| = \|e_j\|,$$

hence $|\lambda_j| = 1$, for all $j \in \{1, \ldots, n\}$. □

Finally we state the following important theorem concerning on the form of the matrix of an isometry.

Theorem 6.18. *Suppose that V is a real inner product space and $T \in L(V)$. Then T is an isometry iff there exist an orthonormal basis of V with respect to which T has a block diagonal matrix where each block on the diagonal matrix is a* $(1,1)$ *matrix containing* 1 *or* -1*, or a* $(2,2)$ *matrix of the form*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

with $\theta \in (0, \pi)$.

Proof. The eigenvalues of T have modulus 1, hence are the form 1, -1 or $\cos\theta \pm \sin\theta$. On the other hand, the matrix of T is similar to a diagonal matrix whose diagonal entries are the eigenvalues. □

6.4 Self adjoint operators

An operator $T \in L(V)$ is called self-adjoint if $T = T^*$ that is $\langle T(v), w\rangle = \langle v, T(w)\rangle$ for all $v, w \in V$.

Remark 6.19. Obviously a self adjoint operator $T \in L(V)$ is normal since in this case holds

$$TT^* = T^*T^* = T^*T.$$

Example 6.20. Let T be an operator on $\mathbb{F}^2$ whose matrix with respect to the standard basis is

$$\begin{pmatrix} 2 & b \\ 3 & 5 \end{pmatrix}.$$

Then T is self-adjoint iff $b = 3$.

Indeed, for $(x, y) \in \mathbb{F}^2$ one has $T(x, y) = (2x + by, 3x + 5y)$, hence for $(u, v) \in \mathbb{F}^2$ it holds

$$\langle T(x, y), (u, v)\rangle = (2x + by)u + (3x + 5y)v = \langle (x, y), (2u + 3v, bu + 5v)\rangle.$$

Thus $T^*(x, y) = (2x + 3y, bx + 5y)$.
In conclusion T is self adjoint, i.e. $T = T^*$ if $b = 3$.
It can easily be verified that the sum of two self adjoint operators and the product of an self adjoint operator by a real scalar is an self-adjoint operator.

Indeed, let $S, T \in L(V)$ be two self adjoint operators. Then $(S+T)^* = S^* + T^* = S + T$, hence $S+T$ is self adjoint. On the other hand for their product we have $(ST)^* = T^*S^* = TS$. Hence TS is self adjoint iff $ST = TS$. Let now $a \in \mathbb{R}$. Then $(aT)^* = \overline{a}T^* = aT$, hence aT is self adjoint.

Remark 6.21. When $\mathbb{F} = \mathbb{C}$ the adjoint on $L(V)$ plays a similar role to complex conjugation on $\mathbb{C}$. A complex number is real iff $z = \overline{z}$. Thus for a self adjoint operator T the sum $T + T^*$ is analogous to a real number. The analogy is reflected in some important properties of a self-adjoint operator, beginning with its eigenvalues.

Remark 6.22. Recall that a complex square matrix A is called hermitian iff $A = A^*$, where A^* is the conjugate transpose of A, that is $A^* = \overline{A}^\top$. If A is a square matrix with real entries, then A is called symmetric iff $A = A^\top$. It can easily be observed that matrix of a self adjoint operator on a complex (real) inner product space is hermitian (symmetric).

Proposition 6.23. *The following statements hold.*

- *Every eigenvalue of a self-adjoint operator is real.*
- *Let v, w eigenvectors corresponding to distinct eigenvalues. Then $\langle v, w\rangle = 0$.*

Proof. Suppose that T is a self-adjoint operator on V. Let λ be an eigenvalue of T, and v be an eigenvector, that is $T(v) = \lambda v$. Then

$$\begin{aligned}
\lambda\|v\|^2 &= \langle \lambda v, v\rangle \\
&= \langle T(v), v\rangle \\
&= \langle v, T(v)\rangle \quad \text{(because } T \text{ is self-adjoint)} \\
&= \langle v, \lambda v\rangle \\
&= \overline{\lambda}\|v\|^2.
\end{aligned}$$

Thus $\lambda = \overline{\lambda}$, i.e., λ is real. $\square$

The next assertion cames from the fact that a self-adjoint operator is normal.

Theorem 6.24. *Let $T \in L(V)$, where V is an inner product space. The following statements are equivalent.*

1. *T is self-adjoint.*
2. *There exists an orthonormal basis of V relative to which the matrix of T is diagonal with real entries.*

Proof. Assume that T is self adjoint. Since T is normal there exists exists an orthonormal basis of V relative to which the matrix of M_T of the operator is upper triangular. But the matrix of T^* in this basis is $M_{T^*} = {M_T}^*$, and from $T = T^*$ one has $M_T = M_T^*$, hence M_T is diagonal, and also the diagonal are formed by real entries.
Conversely, let M_T a diagonal matrix of T, with real entries in some orthonormal basis. Then $M_T = M_T^\top$, hence $M_T = M_{T^*}$ or equivalently $T = T^*$. $\square$

6.5 Problems

Problem 6.5.1. Suppose that A is a complex matrix with real eigenvalues which can be diagonalized by a unitary matrix. Prove that A must be hermitian.

Problem 6.5.2. Prove or give a counter example: the product of two self adjoint operators on a finite dimensional inner product space is self adjoint.

Problem 6.5.3. Show that an upper triangular matrix is normal if and only if it is diagonal.

Problem 6.5.4. Suppose $p \in L(V)$ is such that $p^2 = p$. Prove that p is an orthogonal projection if and only if p is self adjoint.

Problem 6.5.5. Show that if V is a real inner product space, then the set of self adjoint operators on V is a subspace of $L(V)$. Show that if V is a complex inner product space, then the set of self-adjoint operators on V is not a subspace of $L(V)$.

Problem 6.5.6. Show that if $\dim V \geq 2$ then the set of normal operators on V is not a subspace of $L(V)$.

Problem 6.5.7. Let A be a normal matrix. Prove that A is unitary if and only if all its eigenvalues λ satisfy $|\lambda| = 1$.

Problem 6.5.8. Let $X \in \mathcal{M}_n(\mathbb{C})$ be any complex matrix and put $A = I_n - 2XX^*$. Prove that A is both hermitian and unitary. Deduce that $A = A^{-1}$.

Problem 6.5.9. Suppose V is a complex inner product space and $T \in L(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self adjoint and $T^2 = T$.

Problem 6.5.10. Let A be a normal matrix. Show that A is hermitian if and only if all its eigenvalues are real.

Problem 6.5.11. Prove that if $T \in L(V)$ is normal, then

$$\operatorname{im} T = \operatorname{im} T^*.$$

and

$$\ker T^k = \ker T$$

$$\operatorname{im} T^k = \operatorname{im} T$$

for every positive integer k.

Problem 6.5.12. A complex matrix A is called skew-hermitian if $A^* = -A$. Prove the following statements.

a) A skew-hermitian matrix is normal.

b) The eigenvalues of a skew-hermitian matrix are purely imaginary, that is, have the real part 0.

c) A normal matrix is skew-hermitian if all its eigenvalues are purely imaginary.

Problem 6.5.13. Suppose V is a complex inner product space. An operator $S \in L(V)$ is called a square root of $T \in L(V)$ if $S^2 = T$. We denote $S = \sqrt{T}$. Prove that every normal operator on V has a square root.

Problem 6.5.14. Prove or disprove: e identity operator on $\mathbb{F}^2$ has infinitely many self adjoint square roots.

Problem 6.5.15. Let $T, S \in L(V)$ be isometries and $R \in L(V)$ a positive operator, (that is $\langle R(v), v\rangle \geq 0$ for all $v \in V$), such that $T = SR$. Prove that $R = \sqrt{T^*T}$.

Problem 6.5.16. Let $\mathbb{R}_2[X]$ be the inner product space of polynomials with degree at most 2, with the scalar product

$$\langle p, q\rangle = \int_0^1 p(t)q(t)dt.$$

Let $T \in L(\mathbb{R}_2[X])$, $T(ax^2 + bx + c) = bx$.

a) Show that the matrix of T with respect to the given basis is hermitian.

b) Show that T is not self-adjoint.

(Note that there is no contradiction between these statements because the basis in the first statement is not orthonormal.)

Problem 6.5.17. Prove that a normal operator on a complex inner-product space is self-adjoint if and only if all its eigenvalues are real.

7

Elements of geometry

7.1 Quadratic forms

Consider the n-dimensional space $\mathbb{R}^n$ and denote by $x = (x_1, \ldots, x_n)$ the coordinates of a vector $x \in \mathbb{R}^n$ with respect to the canonical basis $E = \{e_1, \ldots, e_n\}$. A **quadratic form** is a map $Q : \mathbb{R}^n \to \mathbb{R}$

$$Q(x) = a_{11}x_1^2 + \ldots a_{nn}x_n^2 + 2a_{12}x_1x_2 + \cdots + 2a_{ij}x_ix_j + \ldots 2a_{n-1,n}x_{n-1}x_n,$$

where the coefficients a_{ij} are all real.

Thus, quadratic forms are homogenous polynomials of degree two in a number of variables.

Using matrix multiplication, we can write Q in a compact form as

$$Q(x) = X^\top AX,$$

where

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{and} \qquad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}.$$

The symmetric matrix A (notice that $a_{ij} = a_{ji}$) is be called the matrix of the quadratic form. Being symmetric (and real), A it is the matrix of a self-adjoint operator with respect to the basis E. This operator, that we call T, is diagonalizable and there exists a basis $B = \{b_1, \dots, b_n\}$ formed by eigenvectors with respect to which T has a diagonal matrix consisting of eigenvalues (also denoted by T)

$$T = diag\{\lambda_1. \dots, \lambda_n\}.$$

Let C be the transition matrix from E to B and

$$X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

the coordinates of the initial vector written in B. We have that

$$X = CX'$$

Knowing that $T = C^{-1}AC$, and that $C^{-1} = C^\top$ we can compute that

$$\begin{aligned} Q(x) &= X^\top AX \\ &= (CX')^\top A\,(CX') \\ &= X'^\top C^\top ACX' \\ &= X'^\top TX' \\ &= \lambda_1 x'^2_1 + \dots + \lambda_n x'^2_n, \end{aligned}$$

and we say that we have reduced Q to its canonical form

$$Q(x) = \lambda_1 {x'}_1^2 + \cdots + \lambda_n {x'}_n^2.$$

This is called the geometric method.

The quadratic form is called

- positive definite if $Q(x) > 0$ for every $x \in \mathbb{R}^n \setminus \{0\}$
- negative definite if $Q(x) < 0$ for every $x \in \mathbb{R}^n \setminus \{0\}$.

We can characterize the positive definiteness of a quadratic form in terms of the diagonal minors of its matrix

$$D_1 = a_{11}, D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}, \ldots, D_n = \det A.$$

We have the following criteria:

- Q is positive definite iff $D_i > 0$ for every $i = \overline{1, n}$
- Q is negative definite iff $(-1)^i D_i > 0$ for every $i = \overline{1, n}$.

7.2 Quadrics

The general equation of a quadric is

$$\begin{aligned} &a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz \\ &2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = 0. \end{aligned}$$

From a geometric point of view, quadrics, which are also called **quadric surfaces**, are two-dimensional surfaces defined as the locus of zeros of a second degree polynomial in x, y and z. Maybe the most prominent example of a quadric is the sphere (the spherical surface).

The type is determined by the quadratic form that contains all terms of degree two

$$Q = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz.$$

We distinguish, based on the sign of the eigenvalues of the matrix of Q, between: ellipsoids, elliptic or hyperbolic paraboloids, hyperboloids with one or two sheets, cones and cylinders.

We study how to reduce the general equations of a quadric to a canonical form.

We reduce Q to a canonical form using the geometric method.

Consider the matrix A associated to Q. Being symmetric, A has real eigenvalues $\lambda_1, \lambda_2, \lambda_3$. If they are distinct, the corresponding eigenvectors are orthogonal (if not we apply the Gram-Schmidt algorithm). Thus, we obtain three orthogonal unit vectors $\{b_1, b_2, b_3\}$, a basis in $\mathbb{R}^3$.

Let R be the transition matrix from $\{i, j, k\}$ to the new basis $\{b_1, b_2, b_3\}$. We recall from previous chapters that R has the three vectors b_1, b_2, b_3 as its columns

$$R = [b_1 | b_2 | b_3] \, .$$

Now, we compute $\det R$ and check whether

$$\det R = 1.$$

If necessary, i.e., if .$\det R = -1$, we must change one of the vectors by its opposite (for example take $R = [-b_1 | b_2 | b_3]$). This assures that the matrix R defines a **rotation**, the new basis being obtained from the original one by this rotation. Let (x, y, z) and (x', y', z') be the coordinates of the same point in the original basis and in the new one, we have

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

We know that with respect to the new coordinates

$$Q = \lambda_1 x'^2 + \lambda_1 y'^2 + \lambda_n z'^2,$$

and thus, the equation of the quadric reduces to the simpler form

$$\lambda_1 x'^2 + \lambda_1 y'^2 + \lambda_n z'^2 + 2a'_{14}x' + 2a'_{24}y' + 2a'_{34}z' + a_{44} = 0.$$

To obtain the canonical form of the quadric we still have to perform another transformation, namely a **translation**. To complete this step we investigate three cases: (A) when A has three nonzero eigenvalues, (B) when one eigenvalue is zero and (C) when two eigenvalues are equal to zero.

(A) For $\lambda_i \neq 0$ we obtain

$$\lambda_1(x' - x_0)^2 + \lambda_2(y' - y_0)^2 + \lambda_3(z' - z_0)^2 + a'_{44} = 0$$

Consider the translation defined by

$$\begin{aligned} x'' &= x' - x_0, \\ y'' &= y' - y_0, \\ z'' &= z' - z_0. \end{aligned}$$

In the new coordinates the equation of the quadric reduces to the canonical form

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \lambda_3 z''^2 + a'_{44} = 0.$$

The cases (B) and (C) can be treated similarly.

We end this section by showing plots of the different quadric surfaces, starting with the degenrate cases, namely the cone and the cylinder.

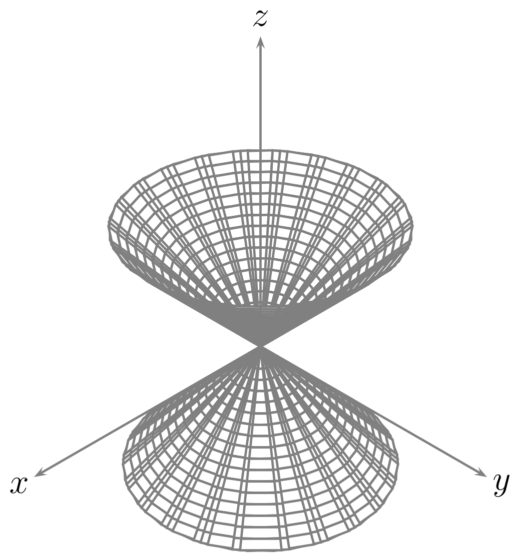

Figure 7.1: Cone $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$

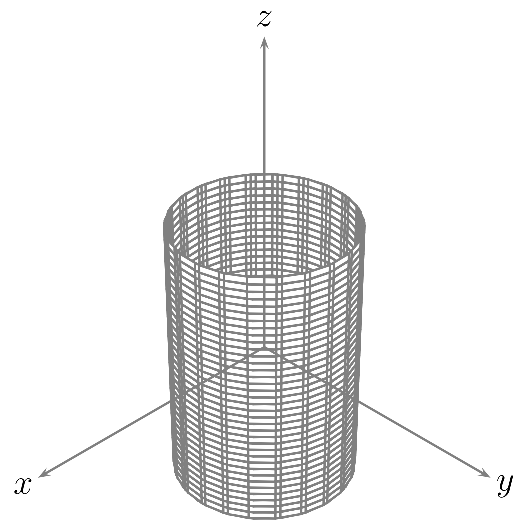

Figure 7.2: Cylinder $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

The nondegenerate quadic surfaces are shown on the next page.

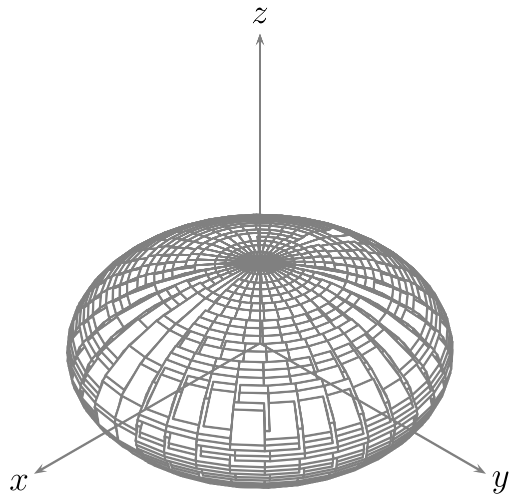

Figure 7.3: Ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

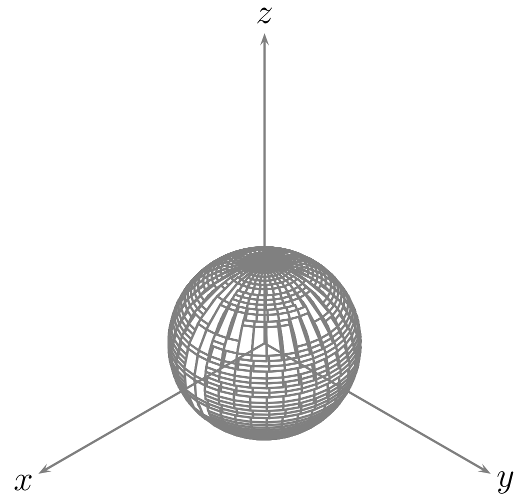

Figure 7.4: Sphere $x^2 + y^2 + z^2 = 1$

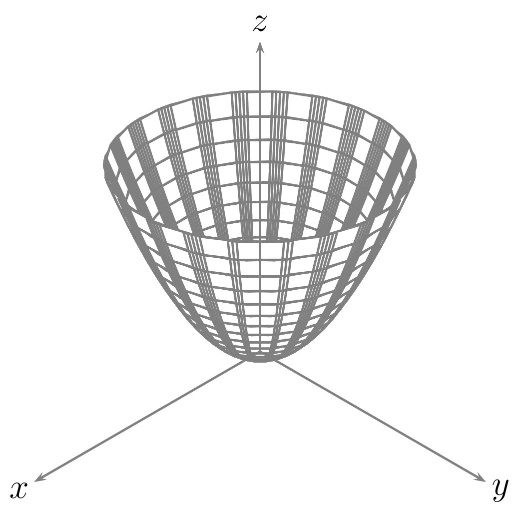

Figure 7.5: Elliptic paraboloid $\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0$

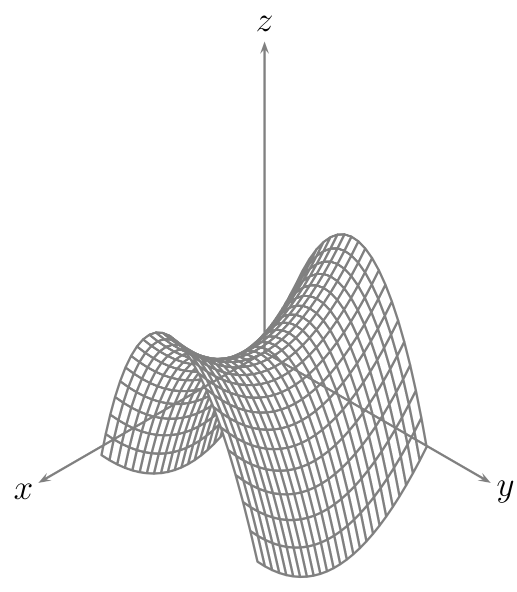

Figure 7.6: Hyperboic paraboloid $\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$

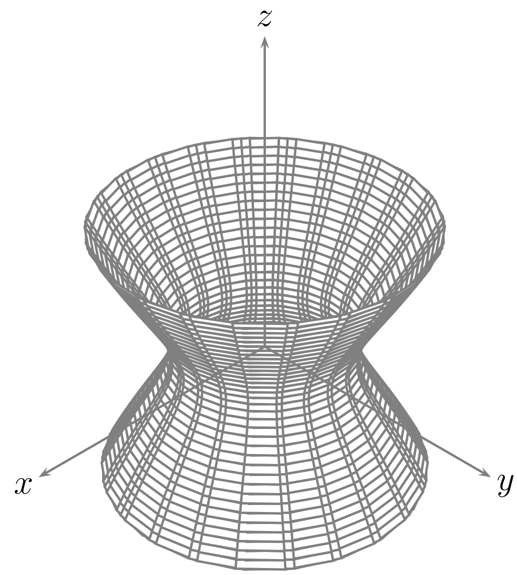

Figure 7.7: Hyperboloid of one sheet
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$

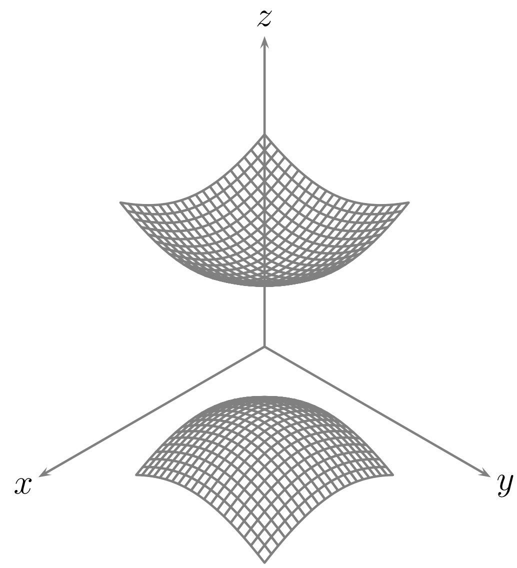

Figure 7.8: Hyperboloid of two sheets
$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1$

7.3 Conics

Studied since the time of ancient greek geometers, **conic sections** (or just conics) are obtained, as their name shows, by intersecting a cone with a sectioning plane. They have played a crucial role in the development of modern science, especially in astronomy. Also, we point out the fact that the circle is a conic section, a special case of ellipse.

The general equation of a conic is

$$a_{11}x^2+2a_{12}xy+a_{22}y^2+2a_{13}x+2a_{23}y+a_{33}=0.$$

The following two determinants obtained from the coefficients of the conic play a crucial role in the classification of conics

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} \qquad \text{and} \qquad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}.$$

Notice that the second determinant corresponds to the quadratic form defined by the first three terms.

Conical sections can be classified as follows:

Degenerate conics, for which $\mathbf{\Delta = 0}$. These include:two intersecting lines (when $D_2 < 0$), two parallel lines or one line (when $D_2 = 0$) and one point (when $D_2 > 0$).

Nondegenerate conics, for which $\mathbf{\Delta = 0}$. Depending on D_2 we distinguish between the

Ellipse $(D_2 > 0)$ whose canonical equation is $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$,

Parabola $(D_2 = 0)$ whose canonical equation is $y^2 - 2ax = 0$,

Hyperbola $(D_2 < 0)$ whose canonical equation is $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$.

A graphical representation of each of these curves is given below.

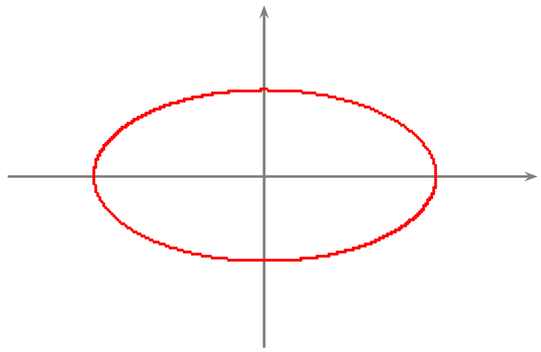

Figure 7.9: Ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

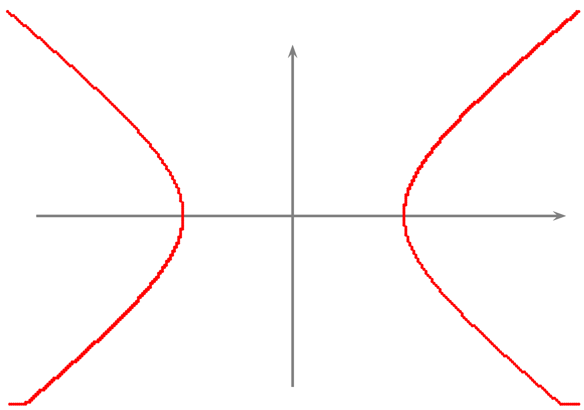

Figure 7.10: Hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

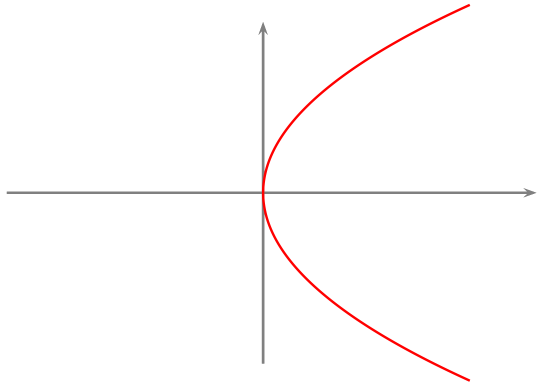

Figure 7.11: Parabola $y^2 - 2px = 0$

The reduction of a conic section to its canonical form is very similar with the procedure that we have presented in the last section when dealing with quadrics. Again, we must perform a rotation and a translation. We show how the reduction can be performed by means of an example.

Example 7.1. Find the canonical form of $5x^2 + 4xy + 8y^2 - 32x - 56y + 80 = 0$.

The matrix of the quadratic form of this conic is

$$\begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$$

and its eigenvalues are the roots of $\lambda^2 - 13\lambda + 36 = 0$. So $\lambda_1 = 9$ and $\lambda_2 = 4$, while two normed eigenvectors are $v_1 = \frac{1}{\sqrt{5}}(1, 2)$ and $v_2 = \frac{1}{\sqrt{5}}(-2, 1)$ respectively. The rotation matrix is thus

$$R = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix},$$

and we can check that $\det R = 1$.

Now, the relation between the old and the new coordinates is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = R \begin{pmatrix} x' \\ y' \end{pmatrix}$$

that is

$$\begin{aligned} x &= \frac{1}{\sqrt{5}}\left(x' - 2y'\right) \\ y &= \frac{1}{\sqrt{5}}\left(2x' + y'\right). \end{aligned}$$

By substituting these expressions in the initial equation we get

$$9x'^2 + 4y'^2 - \frac{144\sqrt{5}}{5}x' + \frac{8\sqrt{5}}{5}y' + 80 = 0.$$

To see the translation that we need to perform we rewrite the above equation as follows

$$\begin{aligned} &9\left(x'^2 - 2\frac{8\sqrt{5}}{5}x' + \left(\frac{8\sqrt{5}}{5}\right)^2\right) + 4\left(y'^2 + 2\frac{\sqrt{5}}{5}y' + \left(\frac{\sqrt{5}}{5}\right)^2\right) \\ &-9\left(\frac{8\sqrt{5}}{5}\right)^2 - 4\left(\frac{\sqrt{5}}{5}\right)^2 + 80 = 0. \end{aligned}$$

Finally, we obtain

$$9\left(x' - \frac{8\sqrt{5}}{5}\right)^2 + 4\left(y' + \frac{\sqrt{5}}{5}\right)^2 - 30 = 0.$$

Thus, the translation $x'' = x' - \frac{8\sqrt{5}}{5}, y'' = y' + \frac{\sqrt{5}}{5}$ reduces the conic to the canonical form

$$\frac{3}{10}x''^2 + \frac{2}{15}y''^2 = 1.$$

7.4 Problems

Problem 7.4.1. Find the canonical form of the following quadric surfaces:

a) $2y^2 + 4xy - 8xz - 6x + 8y + 8 = 0$,

b) $3x^2 + y^2 + z^2 - 2x - 4z - 4 = 0$,

c) $xz = y$,

d) $x^2 + y^2 + 5z^2 - 6xy + 2xz - 2yz - 4x + 8y - 12z + 14 = 0$

Problem 7.4.2. Find the canonical form of the following conics:

a) $x^2 - 6xy + 9y^2 + 20 = 0$,

b) $3y^2 - 4xy - 2y + 4x - 3 = 0$,

c) $5x^2 + 6xy + 2y^2 + 2y - 1 = 0$,

d) $xy = 1$.

Problem 7.4.3. Given the ellipsoid

$$(E) : \frac{x^2}{4} + \frac{y^2}{3} + z^2 - 1 = 0,$$

find the value of the parameter p for which the straight line

$$\begin{cases} x = z + p \\ y = z + 2 \end{cases}$$

is tangent to (E).

Problem 7.4.4. Find the equation of the plane that is tangent to the sphere

$$x^2 + y^2 + z^2 = 1$$

at the point $M\left(\frac{1}{2}, \frac{1}{2}, \frac{\sqrt{2}}{2}\right)$.

Problem 7.4.5. Find the equations of the planes that contain the line

$$\begin{cases} x = 1 \\ y = 0 \end{cases}$$

and are tangent to the sphere

$$(x-1)^2 + y^2 + z^2 = 1.$$

Problem 7.4.6. Consider the hyperboloid of one sheet

$$\frac{x^2}{16} + \frac{y^2}{9} - \frac{z^2}{4} = 1.$$

Determine the straight lines that belong to this surface and pass through the point $P(4, 3, -2)$.

Problem 7.4.7. Find the equation of the circle whose center is located in $C\,(1,2)$ and whose radius is $R=2$.

Problem 7.4.8. Determine the center and the radius of the circle of equation

$$x^2+y^2+2x-4y-4=0.$$

Problem 7.4.9. Write the equation of the circle that passes through the points $A\,(1,1)\,,B\,(1,5)\,,C\,(4,1)$.

Bibliography

[1] S. Axler: *Linear Algebra Done Right*, Springer, 2-nd edition, 2004.

[2] D. Cîmpean, D. Inoan, I. Raşa, *An Invitation to Linear Algebra and Analytic Geometry*, Editura Mediamira, Cluj-Napoca, 2009.

[3] V. Pop, I. Raşa, *Linear Algebra with Application to Markov Chains*, Editura Mediamira, Cluj-Napoca, 2005.

[4] I. V. Proskuryakov, *Problems in Linear Algebra*, MIR Publishing House, Moskow, 1978.

[5] D. Robinson: *A Course in Linear Algebra with Applications*, World Scientific Publishing, 2-nd Edition, 2006.