Přírodovědecká
fakulta

Univerzita Palackého
v Olomouci

LINEAR ALGEBRA LESSONS

Marek Jukl

Olomouc 2020

Univerzita Palackého v Olomouci
Přírodovědecká fakulta

LINEAR ALGEBRA LESSONS

Marek Jukl

Olomouc 2020

Oponenti:
doc. RNDr. Petr Emanovský, Ph.D.
doc. Mgr. Irena Hinterleitner, Ph.D.

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Vydání této publikace bylo financováno v rámci projektu *Univerzita Palackého jako komplexní vzdělávací instituce* CZ.02.2.69/0.0/0.0/16_015/0002337.

1. vydání

ISBN 978-80-244-5824-3

Foreword

Linear Algebra Lessons is a teaching material prepared for the course in linear algebra intended for undergraduate students of mathematics at the Science Faculty of Palacký University in Olomouc.

This text offers basic knowledge about Euclidean vector spaces, homomorphisms and endomorphisms of vector spaces and pseudoinversion. The text is subdivided in such a way that each subchapter corresponds basically to one lecture. Form the foundation stones of linear algebra, these sections are applied in other mathematical disciplines, such as mathematical analysis, geometry or mathematical statistics. They are also the basis of the mathematical apparatus applied to describe natural or social phenomena.

The text lays out basic concepts and does not aim to replace the lectures. On the contrary, it should provide students with sufficient basic knowledge so that the lectures can be devoted to building, or rather emphasizing the logical structure of algebra and its internal context. Each chapter contains educational goals, motivation and specific tasks, thereby giving students the opportunity to familiarize themselves with a given section before the relevant lecture and to determine the questions which the lecture is likely to particularly focus on. To record the findings provided by the lecture, students can use a blank sheet inserted after each subchapter.

Students with a deeper interest in linear algebra will welcome a list of further recommended literature.

The text offers color-coded wordings of definitions (red box) and mathematical theorems (grey box). The following icons will guide you through the text, which should help you work with it independently.

Objectives: At the beginning of each chapter you will find specifically formulated objectives. They will give you an overview of what you will understand after studying a given topic and what you will be able to do.

Motivation: A paragraph explaining why we are going to deal with a given problem at all. It should motivate you to study this particular passage.

Guide: A passage highlighting the links between the subject matter and other parts of the text as well as your previous knowledge. It can thus be seen as a kind of "contextualisation".

Remember: It should be used to alert you to a mistake commonly (and entirely unnecessarily) made by students.

Task: It is meant to encourage you to create something based on the study of a particular topic. The focus here is on the application of received knowledge.

? *Questions:* They check to what extent you understand the subject matter, whether you remember the essential information and whether you can apply it.

October 2020

Author

Contents

1 Euclidean Vector Space

1.1 Basic notions

Students are able to define an Euclidean vector space and to recognize a scalar product. They are able to define the distance function (metric) in an Euclidean vector space and to determine the norm of the vector and the angle of the vectors.

You know the term *scalar product of two vectors* from your secondary school but you started working with it using intuitively understood terms of *angle* and *length of a vector*. Here you will learn how to define the concept of a scalar product axiomatically and how to define the terms length and angle of vectors. You will also become familiar with the term *distance of two vectors*.

Definition 1.1 A vector space $\boldsymbol{V}$ over the field of real numbers $\mathbb{R}$ endowed with a mapping $\cdot : \boldsymbol{V} \times \boldsymbol{V} \to \mathbb{R}$ having the following properties:
$\forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{V},\ \forall t \in \mathbb{R}:$

1. $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$,
2. $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$,
3. $(t\boldsymbol{u}) \cdot \boldsymbol{v} = t(\boldsymbol{u} \cdot \boldsymbol{v})$,
4. $\boldsymbol{u} \neq \boldsymbol{o} \Rightarrow \boldsymbol{u} \cdot \boldsymbol{u} > 0$.

is called an *Euclidean vector space*. The mapping $\cdot : \boldsymbol{V} \times \boldsymbol{V} \to \mathbb{R}$ is called a *scalar product* (or a *dot product) in* $\boldsymbol{V}$. A real number $\boldsymbol{u} \cdot \boldsymbol{v}$ is called a *scalar product of vectors* $\boldsymbol{u}$ *and* $\boldsymbol{v}$.

Remark 1.2

- An Euclidean vector space may be considered as an ordered couple $(\boldsymbol{V}, \cdot)$. In a given real vector space, we may define different scalar products and thus we obtain different Euclidean vector spaces.
- With respect to the fact that scalars $(t, r, s \in \mathbb{R})$ and vectors $(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{V})$ are denoted by different symbols, we may denote the scalar product $\boldsymbol{u} \cdot \boldsymbol{v}$ only by $\boldsymbol{uv}$. It is distinguished from a scalar multiplication f.e. $t\boldsymbol{u}$.
- The scalar product of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ may be denoted also in the following ways: $<\boldsymbol{u}, \boldsymbol{v}>$, $\beta(\boldsymbol{u}, \boldsymbol{v})$ etc.

Let us remember that a scalar product is defined only for vector spaces over real numbers[1].

It is usual to denote a scalar product $\boldsymbol{uu}$ by $\boldsymbol{u}^2$.

Example 1.3 Is the following map "·" a scalar product in a vector space $\boldsymbol{V}$?

(i) $\boldsymbol{V} = \mathbb{R}^2$, $(x_1, x_2) \cdot (y_1, y_2) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$,
[yes]

(ii) $\boldsymbol{V} = \mathbb{R}^n$, $(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = \sum_{i=1}^{n} x_iy_i$,
[yes; this product is usually called a *standard scalar product* in $\mathbb{R}^n$]

(iii) $\boldsymbol{V} = \mathbb{R}^3$, $(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$,
[no]

(iv) $\boldsymbol{V} = \mathbb{R}^2$, $(x_1, x_2) \cdot (y_1, y_2) = 2x_1y_1 + x_1$,
[no]

(v) $\boldsymbol{V} = \mathbb{R}^3$, $(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = 2x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + 2x_3y_3$.
[yes]

Theorem 1.4 *Let $\boldsymbol{V}$ with scalar product $\cdot$ be an Euclidean vector space. Then any subspace $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ with a restriction $\cdot|\boldsymbol{W}\times\boldsymbol{W}$ is an Euclidean vector space.*

Corollary 1.5 *All notions defined for Euclidean vector spaces may be used in subspaces and need not be defined specially for subspaces. All theorems which hold for Euclidean spaces are true also for their subspaces and need not to be proved again.*

Theorem 1.6 *For any $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$ and any $t, r \in \mathbb{R}$ we have:*

1. $(t\boldsymbol{u})(r\boldsymbol{v}) = (tr)\boldsymbol{uv}$,
2. $\boldsymbol{uu} = 0 \Leftrightarrow \boldsymbol{u} = \boldsymbol{o}$,
3. $(\forall \boldsymbol{x} \in \boldsymbol{V}\colon \boldsymbol{xu} = 0) \Leftrightarrow \boldsymbol{u} = \boldsymbol{o}$,
4. $\boldsymbol{u} = \boldsymbol{v} \Leftrightarrow (\forall \boldsymbol{x} \in \boldsymbol{V}\colon \boldsymbol{xu} = \boldsymbol{xv})$.

Corollary 1.7 *For any vector $\boldsymbol{u} \in \boldsymbol{V}$ we have $\boldsymbol{uu} \geq 0$.*

[1]The generalisation of vector spaces over complex numbers is called a *unitary product*.

Ad 1.6 (4): Let us note that mere *existence* of $\boldsymbol{x}$ with $\boldsymbol{xu} = \boldsymbol{xv}$ does not imply the equality $\boldsymbol{u} = \boldsymbol{v}$. For example in $\mathbb{R}^2$ with the standard scalar product, we have $(1,1)\cdot(1,0) = (1,1)\cdot(\frac{1}{2},\frac{1}{2})$, but $(1,0) \neq (\frac{1}{2},\frac{1}{2})$.

Definition 1.8 Let $\boldsymbol{u} \in \boldsymbol{V}$. Then the number denoted by $\|\boldsymbol{u}\|$ and defined by

$$\|\boldsymbol{u}\| = \sqrt{\boldsymbol{uu}}.$$

is called a *norm* (or *lenght*) of a vector $\boldsymbol{u}$. In the case when $\|\boldsymbol{u}\| = 1$, a vector $\boldsymbol{u}$ is termed *normalised.*

Example 1.9 Write formula for the vector norm for each of scalar products in Exersice 1.3.

[Solution: Let us present a solution for a scalar product (i):
According to Definition 1.8, we may write for $\boldsymbol{u} = (u_1, u_2) \in \mathbb{R}^2$

$$\|\boldsymbol{u}\| = \sqrt{\boldsymbol{uu}} = \sqrt{(u_1,u_2)(u_1,u_2)} = \sqrt{2u_1u_1 + u_1u_2 + u_2u_1 + u_2u_2} = \\ = \sqrt{2u_1^2 + 2u_1u_2 + u_2^2}$$

Analogously, for s scalar product (ii) we obtain

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}.]$$

Theorem 1.10 *For any vectors $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$ and any $t \in \mathbb{R}$ the following holds*

1. $\|t\boldsymbol{u}\| = |t|\,\|\boldsymbol{u}\|$;
2. $\|\boldsymbol{u}\| = 0 \Leftrightarrow \boldsymbol{u} = \boldsymbol{o}$;
3. $\|\boldsymbol{u}\| > 0 \Leftrightarrow \boldsymbol{u} \neq \boldsymbol{o}$;
4. $\|\boldsymbol{u}\|\,\|\boldsymbol{v}\| \geq |\boldsymbol{uv}|$; *this inequality turns into equality if and only if* $\boldsymbol{u}, \boldsymbol{v}$ *are linearly independent vectors;*
5. $\|\boldsymbol{u}\| + \|\boldsymbol{v}\| \geq \|\boldsymbol{u} + \boldsymbol{v}\|$; *this inequality turns into equality if and only if there exists* $t \in \mathbb{R},\ t \geq 0$, *such that* $\boldsymbol{v} = t\boldsymbol{u}$ *or* $\boldsymbol{u} = t\boldsymbol{v}$;
6. $|\|\boldsymbol{u}\| - \|\boldsymbol{v}\|| \leq \|\boldsymbol{u} - \boldsymbol{v}\|$; *this inequality turns into equality if and only if there exists* $t \in \mathbb{R},\ t \geq 0$, *such that* $\boldsymbol{v} = t\boldsymbol{u}$ *or* $\boldsymbol{u} = t\boldsymbol{v}$.

The inequality 4 in Theorem 1.10 is called the *Cauchy* or *Schwarz inequality*; the inequality 5 is called the *triangle unequality.*

Definition 1.11 Let $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$. Then the number belonging to $< 0, \pi >$ denoted by $\measuredangle(\boldsymbol{u}, \boldsymbol{v})$ and defined by

1. $\boldsymbol{u} \neq \boldsymbol{o} \neq \boldsymbol{v}: \ \measuredangle(\boldsymbol{u}, \boldsymbol{v}) = \arccos \frac{\boldsymbol{uv}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}$,
2. $\boldsymbol{u} = \boldsymbol{o} \vee \boldsymbol{v} = \boldsymbol{o}: \ \measuredangle(\boldsymbol{u}, \boldsymbol{v}) = \frac{\pi}{2}$

is called an *angle of vectors* $\boldsymbol{u}$ *and* $\boldsymbol{v}$.

Corollary 1.12 *For any* $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$ *we have:*

1. $\measuredangle(\boldsymbol{u}, \boldsymbol{v}) = \frac{\pi}{2} \Leftrightarrow \boldsymbol{uv} = 0$,
2. $\boldsymbol{uv} = \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| \cos \measuredangle(\boldsymbol{u}, \boldsymbol{v})$.

Remark 1.13

- The symmetry of scalar product (property 1 of Definition 1.1) implies $\measuredangle(\boldsymbol{u}, \boldsymbol{v}) = \measuredangle(\boldsymbol{v}, \boldsymbol{u})$ for any $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{v}$.
- In the following section, the point 1 of Corollary 1.12 will be used for the formulation of a criterion of *orthogonality of two vectors*.
- In secondary schools, point 2 is usually used for the definition of a *scalar product* by the lenght of vectors and the angle between them. We can see that the axiomatic definition 1.1 of a scalar product is consistent with its intuitive definition presented in secondary schools.

Having defined notions of *lenght* of vector and *angle* between vectors, we introduce the notion *metric* or *distance* of two wectors.

In calculus, the notion *metric space* denotes any non-empty set M with a mapping $\rho \colon M \times M \to \mathbb{R}^+$ having the following properties:

1. $\forall x, y \in M, x \neq y \colon \rho(x, y) = \rho(y, x) > 0$,
2. $\forall x \in M \colon \rho(x, x) = 0$,
3. $\forall x, y, z \in M \colon \rho(x, y) + \rho(y, z) \geq \rho(y, z)$.

The mapping ρ is called a *metric on* M.

It is clear, that ρ is the natural generalisation of the notion of *distance between two points*, intuitively introduced in secondary schools for a plane or 3-dimensional space. The notion of *distance between two vectors* is not usually studied in secondary schools.

Theorem 1.14 *Let $\boldsymbol{V}$ be an euclidean vectors space. A mapping $\rho\colon \boldsymbol{V}\times\boldsymbol{V}\to\mathbb{R}^+$ defined by*

$$\forall \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{V}\colon\ \rho(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{v}-\boldsymbol{u}\| \tag{1.1}$$

is a metric on the set $\boldsymbol{V}$.

Definition 1.15 Let $\boldsymbol{V}$ with a scalar product $\cdot$ be an Euclidean vector space. Then a metric ρ on $\boldsymbol{V}$ defined by (1.1) is called a *metric induced by a scalar product* $\cdot$.

Remark 1.16 Any Euclidean vector space is also a metric space. On a given Euclidean vector space, other metrics may be also defined.

In the case of an arithmetical vector space $\mathbb{R}^n$ with the standard scalar product, the metric induced by this scalar product is given by (cf. example 1.9):

$$\rho((x_1,\dots,x_n),(y_1,\dots,y_n)) = \sqrt{\sum_{i=1}^{n}(y_i-x_i)^2}.$$

On the same vector space, we may defined also a metric ρ' by

$$\rho'((x_1,\dots,x_n),(y_1,\dots,y_n)) = \mathrm{Max}\{|y_i-x_i|\}_{i=1}^{n},$$

which is different from the metric induced by scalar product.

Let us consider an arithmetical space $\mathbb{R}^2$ with a scalar product defined by the relation (i) in Example 1.9. Then the relation for the metric induced by a scalar product can be written as

$$\rho((x_1,x_2),(y_1,y_2)) = \sqrt{2(y_1-x_1)^2+2(y_1-x_1)(y_2-x_2)+(y_2-x_2)^2}.$$

Notes:

1.2 Orthogonality in the Euclidean vector space

Students are able to define the notion of orthogonality of vectors of an Euclidean vector space. They are also able to construct an orthogonal complement of any subspace and to determine the orthogonality of two subspaces of an Euclidean vector space. Students can also construct an orthonormalisation to a given base. They know Euclidean formulas for the scalar product, the norm and the distance between two vectors.

Your secondary school surely gave you some idea about the *orthogonality* of two vectors, the orthogonality of a vector to a set of vectors and, in the case of one- and two-dimensional subspaces of a two- or three-dimensional space, of mutual orthogonality of subspaces. Here you will learn how to generally define the notion of orthogonality in Euclidean vector spaces to match your intuitive notion of a few specific cases. You will be able to demonstrate that in an orthonormal base, the formula for the scalar product of two vectors has a very simple form. You will also learn how to construct an orthonormal base in any Euclidean vector space.

Definition 1.17 Let $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$. Vectors $\boldsymbol{u}, \boldsymbol{v}$ are called *orthogonal* (or *perpendicular) vectors (which is denoted by* $\boldsymbol{u} \perp \boldsymbol{v}$*, if* $\measuredangle(\boldsymbol{u}, \boldsymbol{v}) = \frac{\pi}{2}$.

Remark 1.18 With regard to Remark 1.13 we can see that the relation *to be orthogonal* is symmetric on $\boldsymbol{V}$, i.e. there is no difference between $\boldsymbol{u} \perp \boldsymbol{v}$ and $\boldsymbol{v} \perp \boldsymbol{u}$.

From Corollary 1.12, we obtain a criterion of orthogonality of twoo vectors (it is used by some authors as the definition of orthogonality of twoo vectors).

Theorem 1.19 *Let* $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$*. Then* $\boldsymbol{v} \perp \boldsymbol{u} \Leftrightarrow \boldsymbol{u}\boldsymbol{v} = 0$.

Definition 1.20 Let $\mathcal{U} \subset \boldsymbol{V}$. We say that

1. $\mathcal{U}$ is an *orthogonal set of vectors* if

$$\forall \boldsymbol{u}, \boldsymbol{v} \in \mathcal{U} \colon \boldsymbol{u} \neq \boldsymbol{v} \Rightarrow \boldsymbol{u}\boldsymbol{v} = 0,$$

2. $\mathcal{U}$ is an *orthonormal set of vectors,* if $\mathcal{U}$ an orthogonal set of vectors and the lenght of every of these vectors is equal to one.

For finite subsets we have[2]:

Theorem 1.21 *Let $\mathcal{U} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_k\} \subset \boldsymbol{V}$. Then the following holds:*

1. *$\mathcal{U}$ is an orthogonal set of vectors if and only if*

$$\forall i, j \in \{1, 2, \dots, k\} \colon i \neq j \Rightarrow \boldsymbol{u}_i \boldsymbol{u}_j = 0,$$

2. *$\mathcal{U}$ is an orthonormal set of vectors if and only if*

$$\forall i, j \in \{1, 2, \dots, k\} \colon \boldsymbol{u}_i \boldsymbol{u}_j = \delta_{ij}.$$

Remark 1.22 Let us remark that orthogonality and orthonormality are always properties of a *set* of such vectors as a whole (not of individual vectors).

Now we will answer the question about the relation between orthogonality of a set of vectors and linear (in)dependence of such a set. We will also introduce a special term for bases whose elements form an orthonormal or orthogonal set of vectors.

Theorem 1.23 *Any orthogonal set of nonzero vectors of an Euclidean vector space is linearly independent.*

Corollary 1.24 *Any orthonormal set of vectors of an Euclidean vector space is linearly independent.*

Let us note in regard in Theorem 1.23 that the assumption that all vectors are nonzero cannot be omitted. Any set of vectors which contains zero vector is linearly dependent.

Definition 1.25 A basis $\mathcal{B}$ of an Euclidean vector space $\boldsymbol{V}$ is called

1. an *orthogonal basis* if $\mathcal{B}$ is an orthogonal set of vectors of $\boldsymbol{V}$,
2. an *orthonormal basis* if $\mathcal{B}$ is an orthonormal set of vectors of $\boldsymbol{V}$.

[2] A symbol δ_{ij}, called *Kronecker delta*, is defined for any natural numbers i, j in the following way:

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j, \\ 1 & \text{for } i = j. \end{cases}$$

From Theorem 1.23, it clearly follows:

Corollary 1.26

1. *Any orthogonal set of n nonzero vectors of $\boldsymbol{V}$ forms an orthogonal basis of $\boldsymbol{V}$,*
2. *Any orthonormal set of n vectors of $\boldsymbol{V}$ forms an orthonormal basis of $\boldsymbol{V}$.*

The following theorem follows from properties of the scalar product and from Theorem 1.21. It stipulates conditions when the formula for the scalar product (expressed in coordinates of given vectors) has the so-called *Euclidean* (or *Cartesian) form*. The simplicity of such a formula is the reason why orthonormal bases are preferred (see also Theorem 1.29 and its corollary).

The following theorem also implies that in the arithmetic Euclidean space $\mathbb{R}^n$ with the standard scalar product, the standad basis[3] is one of orthonormal bases .

Theorem 1.27 *Let $\mathcal{B}$ be a basis of a vector space $\boldsymbol{V}$. The basis $\mathcal{B}$ is orthonormal if and only if*

$$\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}, \{\boldsymbol{u}\}_{\mathcal{B}} = \{u_1, \ldots, u_n\}, \{\boldsymbol{v}\}_{\mathcal{B}} = \{v_1, \ldots, v_n\} \colon \boldsymbol{u}\boldsymbol{v} = \sum_{j=1}^{n} u_j v_j. \tag{1.2}$$

Let us recall that a symbol $\{\boldsymbol{x}\}_{\mathcal{B}}$ denotes coordinates of a vector $\boldsymbol{x}$ with respect to the basis $\mathcal{B}$.

Remark 1.28 The Euclidean formula (1.2) may be written also in the following form:

$$\boldsymbol{u}\boldsymbol{v} = \{\boldsymbol{u}\}_{\mathcal{B}} \{\boldsymbol{v}\}_{\mathcal{B}}^T.$$

Theorem 1.29 *Let $\mathcal{B}$ be a basis of a vector space $\boldsymbol{V}$. The basis $\mathcal{B}$ is orthonormal if and only if*

$$\forall \boldsymbol{u} \in \boldsymbol{V}, \{\boldsymbol{u}\}_{\mathcal{B}} = \{u_1, \ldots, u_n\} \colon \|\boldsymbol{u}\| = \sqrt{\sum_{j=1}^{n} u_j^2}.$$

[3] I.e. the basis formed by arithmetical vectors $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$.

Corollary 1.30 *Let $\mathcal{B}$ be a basis of a vector space $\boldsymbol{V}$. The basis $\mathcal{B}$ is orthonormal if and only if*

$$\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}, \{\boldsymbol{u}\}_{\mathcal{B}} = \{u_1, \dots, u_n\}, \{\boldsymbol{v}\}_{\mathcal{B}} = \{v_1, \dots, v_n\} \colon \rho(\boldsymbol{u}, \boldsymbol{v}) = \sqrt{\sum_{j=1}^{n} (v_j - u_j)^2}.$$

We do not yet know if there exists at least one orthonormal basis of an Euclidean vector space. By the mathematical induction (for the dimension of $\boldsymbol{V}$), try to verify the validity of the following theorem from which we can obtain the answer to this question (how?).

Theorem 1.31 *Any orthonormal basis of an subspace of an Euclidean vector space may be completed to the orthonormal basis of this space.*

Corollary 1.32 *In any Euclidean vector space, there exists at least one orthonormal basis.*

Example 1.33 Find at least one orthonormal basis of the Euclidean space described in Exercise 1.3, (v).
[Instruction: Let us consider an arbitrary nonzero vector $\bar{\boldsymbol{e}}_1$ z $\boldsymbol{V}$ and construct a set of vectors orthogonal to it. In this set, choose any nonzero vector $\bar{\boldsymbol{e}}_2$. Let us now construct a set of vectors orthogonal to $\bar{\boldsymbol{e}}_1$ as well as to $\bar{\boldsymbol{e}}_2$ and in this let us choose a vector $\bar{\boldsymbol{e}}_3$. Subsequently let us obtain an orthogonal basis of $\boldsymbol{V}$. Multiplying every vector $\bar{\boldsymbol{e}}_i$ by $\frac{1}{\|\bar{\boldsymbol{e}}_i\|}$ we get the demanded basis. For example, vectors $\boldsymbol{e}_1 = (\frac{\sqrt{2}}{2}, 0, 0)$, $\boldsymbol{e}_2 = (\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{3}, 0)$, $\boldsymbol{e}_3 = (0, 0, \frac{\sqrt{2}}{2})$ form one of solutions.]

Example 1.34 Let be given vectors from an arithmetical vector space $\mathbb{R}^3$:

$$\boldsymbol{e}_1 = (1, 1, 0), \ \boldsymbol{e}_2 = (0, 1, 1), \ \boldsymbol{e}_3 = (0, 0, 1).$$

Construct a scalar product $\cdot$ in $\mathbb{R}^3$ such that $\mathcal{B} = <\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3>$ is an orthonormal basis.
[Instruction: A formula for a scalar product of two vectors with respect to the basis $\mathcal{B}$ must have the form (1.2). Use a transformation of coordinates between basis $\mathcal{B}$ and the standard basis $\mathbb{R}^3$.
Solution:
$\boldsymbol{u} \cdot \boldsymbol{v} = 3u_1v_1 - 2u_1v_2 + u_1v_3 - 2u_2v_1 + 2u_2v_2 - u_2v_3 + u_3v_1 - u_3v_2 + u_3v_3.$]

The following theorem may be easily derived from the definition of transition matrix, matrix multiplication and relation (1.2).[4]

[4] By the symbol $(\mathcal{B}, \mathcal{C})$ we denote the transition matrix from the basis $\mathcal{B}$ to the basis $\mathcal{C}$.

Theorem 1.35 *Let $\mathcal{B}$ be an orthonormal basis and let $\mathcal{C}$ be an arbitrary basis of an Euclidean vector space. The basis $\mathcal{C}$ is orthonormal if and only if*

$$(\mathcal{B},\mathcal{C})(\mathcal{B},\mathcal{C})^T = E$$

Definition 1.36 A real square matrix A is called an *orthogonal matrix* if

$$AA^T = E.$$

Theorem 1.37 *A real square matrix of order n is orthogonal if and only if its rows form an orthonormal basis of an arithmetical vector space $\mathbb{R}^n$ with the standard scalar product.*

From theorems 1.35 and 1.37, it follows (how?):

Theorem 1.38 *A set of orthogonal matrices of a given order endowed with matrix multiplication forms a group which is a subgroup in the multiplicative group of real regular matrices of the same order.*

Let us recall that two basis of a given real vector space are called *consistently oriented* if the determinant of the transition matrix is a positive number. The following theorem (known also as a *Gram-Schmidt orthonormalisation* shows that to every basis of an Euclidean vector space, a certain, uniquely determined orthonormal basis may be constructed.

Theorem 1.39 *To any basis $\mathcal{U} = <\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n>$ of an Euclidean vector space $\boldsymbol{V}$, there exists exactly one orthonormal basis $\mathcal{V} = <\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n>$ of this space with following properties:*

1. *for every r, $r = 1, \ldots, n$, it holds: $[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r] = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r]$,*
2. *for every r, $r = 1, \ldots, n$, the following r-tuples $<\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r>$ a $<\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r>$ are consistely oriented basis of the respective subspaces.*

Let us illustrate the construction of an orthonormal basis whose existence is guaranteed by the previous theorem (the proof of Theorem 1.37 is based on this contruction. Try to prove it using mathematical induction for $n = \dim \boldsymbol{V}$):

Firstly, to a basis $\mathcal{U}$, let us construct an orthogonal basis $\mathcal{W} = <\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n>$ with properties 1 and 2 of the previous theorem. Then vectors of the basis $\mathcal{W}$ may be easily normalised and we thus obtain demanded basis $\mathcal{V}$.

Let us put

$$\boldsymbol{w}_1 = \boldsymbol{u}_1. \tag{1.3}$$

A vector $\boldsymbol{w}_2$ is searched in the following form:

$$\boldsymbol{w}_2 = \boldsymbol{u}_2 + t\boldsymbol{w}_1, \tag{1.4}$$

which tohether with (1.3) means that properties 1 and 2 are fulfilled for $r = 2$ [5]. Multiplying an equality (1.4) by a vector $\boldsymbol{w}_1$, we have

$$\boldsymbol{w}_1\boldsymbol{w}_2 = \boldsymbol{w}_1\boldsymbol{u}_2 + t(\boldsymbol{w}_1\boldsymbol{w}_1). \tag{1.5}$$

Because we want $\mathcal{W}$ to be orthogonal, we put $\boldsymbol{w}_1\boldsymbol{w}_2 = 0$. The equation

$$0 = \boldsymbol{w}_1\boldsymbol{u}_2 + t(\boldsymbol{w}_1\boldsymbol{w}_1)$$

with an unknown t has exactly one solution since $(\boldsymbol{w}_1\boldsymbol{w}_1) \neq 0$. Substituting t into an equality (1.4), we obtain a vector $\boldsymbol{w}_2$ which, together with $\boldsymbol{w}_1$, forms the demanded basis for $r = 2$.

Analogously, we put

$$\boldsymbol{w}_3 = \boldsymbol{u}_2 + t_2\boldsymbol{w}_2 + t_1\boldsymbol{w}_1. \tag{1.6}$$

Multiplying this equality by $\boldsymbol{w}_1$ and putting $\boldsymbol{w}_1\boldsymbol{w}_3 = 0$, we obtain a solvable equation with an unknown t_1 (because $\boldsymbol{w}_1 \perp \boldsymbol{w}_2$). Multiplying (1.6) by $\boldsymbol{w}_2$ and putting $\boldsymbol{w}_2\boldsymbol{w}_3 = 0$, we obtain a solvable equation with an unknown t_2, and we may construct a vector $\boldsymbol{w}_3$.

Proceeding further, we eventually obtain a vector $\boldsymbol{w}_n$ which, together with its predecessors, forms an orthogonal basis $\mathcal{W}$ with the demanded properties.

Example 1.40 Let be given a vector space $\mathbb{R}^3$ with a scalar product defined by

$$(x_1, x_2, x_3) \cdot (y_1, y_2, y_3) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 + 2x_3y_3.$$

Orthonormalize a basis $\mathcal{U} = <\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3>$[6], where

$$\boldsymbol{u}_1 = (1, 0, 0),\ \ \boldsymbol{u}_2 = (0, 1, 0), \boldsymbol{u}_3 = (0, 0, 1).$$

[Solution: $\boldsymbol{v}_1 = \frac{\sqrt{2}}{2}(1, 0, 0)$, $\boldsymbol{v}_2 = \sqrt{2}(-\frac{1}{2}, 1, 0)$, $\boldsymbol{v}_3 = \frac{\sqrt{2}}{2}(0, 0, 1)$.]

Definition 1.41 Let $\boldsymbol{u} \in \boldsymbol{V}$, $Q \subseteq \boldsymbol{V}$. A vector $\boldsymbol{u}$ is said to be *orthogonal to a set* Q (it will be denoted by $\boldsymbol{u} \perp Q$) if it is orthogonal to all vectors belonging to Q.

[5] An appropriate "candidate" for the transition matrix between bases $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$ and $\langle \boldsymbol{w}_1, \boldsymbol{w}_2 \rangle$ has the form $\left(\begin{smallmatrix} 1 & 0 \\ t & 1 \end{smallmatrix}\right)$. Thus, it is a regular matrix and has a positive determinant.

[6] You may see that $\mathcal{U}$ is not orthonormal.

If a set Q is also a subspace, we obtain the following useful criterion.

Theorem 1.42 *Let $\boldsymbol{u}\in\boldsymbol{V}$, $\boldsymbol{U}\subseteq\subseteq\boldsymbol{V}$. A vector $\boldsymbol{u}$ is orthogonal to $\boldsymbol{U}$ if and only if it is orthogonal to some (and thus to every) set of generators of a subspace $\boldsymbol{W}$.*

Definition 1.43 Let $Q\subseteq\boldsymbol{V}$. The set of all vectors of $\boldsymbol{V}$ which are orthogonal to Q is called an *orthogonal complement of the set Q in $\boldsymbol{V}$* and it is denoted by $Q^{\perp}$.

Corollary 1.44

1. *Let $Q\subseteq\boldsymbol{V}$. Then:*

$$Q^{\perp} = \{\boldsymbol{y} \in \boldsymbol{V};\ \forall \boldsymbol{x} \in Q : \boldsymbol{x}\boldsymbol{y} = 0\}.$$

2. *Let $Q \subseteq\subseteq \boldsymbol{V}$, $Q = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k]$. Then:*

$$Q^{\perp} = \{\boldsymbol{y} \in \boldsymbol{V};\ \forall i, i{=}1, \ldots, k : \boldsymbol{u}_i\boldsymbol{x} = 0\}.$$

Theorem 1.45 *Let $\boldsymbol{U} \subseteq\subseteq \boldsymbol{V}$. Then:*

1. $\dim \boldsymbol{U}^{\perp} = \dim \boldsymbol{V} - \dim \boldsymbol{U}$,
2. $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{U}^{\perp}$,
3. $\boldsymbol{U}^{\perp\perp} = \boldsymbol{U}$.

If $\boldsymbol{U} \subseteq\subseteq \boldsymbol{V}$, then it follows from the point 2 of the previous theorem that any vector $\boldsymbol{x} \in \boldsymbol{V}$ has one and only one expression in the form

$$\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{x}^{\perp}, \text{ where } \boldsymbol{x}^* \in \boldsymbol{U},\ \boldsymbol{x}^{\perp} \in \boldsymbol{U}^{\perp}.$$

This important result is used for the construction of the so-called *orthogonal projection of a vector onto a subspace.*

Theorem 1.46 *Let $\boldsymbol{U}, \boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$. Then:*

1. $(\boldsymbol{U} + \boldsymbol{W})^{\perp} = \boldsymbol{U}^{\perp} \cap \boldsymbol{W}^{\perp}$,
2. $(\boldsymbol{U} \cap \boldsymbol{W})^{\perp} = \boldsymbol{U}^{\perp} + \boldsymbol{W}^{\perp}$,
3. $\boldsymbol{U} \subseteq \boldsymbol{W} \Leftrightarrow \boldsymbol{W}^{\perp} \subseteq \boldsymbol{U}^{\perp}$.

The intuitive notion cultivated in secondary school plane and three-dimensional geometry lessons leads us to consider two lines as orthogonal if their direction vectors are orthogonal, i.e. if the direction of one of them belongs to the orthogonal complement of the direction of the second line. Similarly, a line is understood to be orthogonal to the plane if it is orthogonal to two concurrent lines of that plane, i.e. if its direction belongs to the orthogonal complement of the set of vectors associated with that plane. Finally, we consider two planes orthogonal if one of them contains a perpendicular line to the other, i.e. if the orthogonal complement of the set of vectors associated with the latter of them is contained in the set of vectors associated with the former. The following definition now seems natural.

Definition 1.47 Let $\boldsymbol{U}, \boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$. A vector $\boldsymbol{U}$ is said to be *orthogonal to a subspace* $\boldsymbol{W}$ which is denoted by $\boldsymbol{U} \perp \boldsymbol{W}$ if

$$\boldsymbol{U} \subseteq \boldsymbol{W}^{\perp} \vee \boldsymbol{W}^{\perp} \subseteq \boldsymbol{U}.$$

It follows from Theorems 1.45 and 1.46:

Corollary 1.48 *The relation "to be orthogonal" is a symmetric relation on the set of subspaces of a given vector space.*

We can speak of a pair of subspaces as of *orthogonal subspaces*.

Let us recall that a *(vector) hyperplane* of an n-dimensional vector space is its every $(n-1)$-dimensional subspace.

Theorem 1.49

1. *To any vector hyperplane $\boldsymbol{N} \subset \boldsymbol{V}$, there exists unique (up to a non-zero scalar multiplication) non-zero vector $\boldsymbol{n} \in \boldsymbol{V}$ such that*

$$\boldsymbol{N} = \{\boldsymbol{x} \in \boldsymbol{V}\colon\ \boldsymbol{x}\boldsymbol{n} = 0\}. \tag{1.7}$$

2. *To any non-zero vector $\boldsymbol{n} \in \boldsymbol{V}$, there exists a unique vector hyperplane $\boldsymbol{N} \subset \boldsymbol{V}$ such that (1.7) holds.*

Definition 1.50 Let $\boldsymbol{N}$ be an arbitrary hyperplane in $\boldsymbol{V}$. Then $\boldsymbol{N}^{\perp}$ is called a *normal direction of a vector hyperplane* $\boldsymbol{N}$ and any generator of the normal direction is called a *normal vector of a vector hyperplane* $\boldsymbol{N}$.

From Theorem 1.49, it follows:

Theorem 1.51 *Let $\boldsymbol{N} \subset \boldsymbol{V}$ be an arbitrary hyperplane and $\mathcal{B}$ some orthonormal basis of $\boldsymbol{V}$. Then for any normal vector $\boldsymbol{n}$ of this hyperplane it holds:*

$$\{\boldsymbol{n}\}_{\mathcal{B}} = (a_1, a_2, \ldots, a_n)$$

if and only if

$$\boldsymbol{N} = \{\boldsymbol{x} \in \boldsymbol{V}, \{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, x_2, \ldots, x_n) : \; a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0\}.$$

Example 1.52 Let $\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2]$ be a subspace of an Euclidean vector space $\boldsymbol{V}_3$ and let

$$\{\boldsymbol{w}_1\}_{\mathcal{B}} = (1, 2, 0), \; \{\boldsymbol{w}_2\}_{\mathcal{B}} = (1, 1, 1).$$

Construct an orthogonal complement $\boldsymbol{W}^{\perp}$.

[Instruction: use point 2 of Corollary 1.44. Solution: $\boldsymbol{W}^{\perp} = [(2, -1, -1)]$.

A subspace $\boldsymbol{W}$ is a vector hyperplane which (according to Theorem 1.51) may by expressed by the equation

$$2x_1 - x_2 - x_3 = 0.$$

Notes:

1.3 Distance and angle in Euclidean vector space

Students are able to define the terms distance of a vector and deviation of a vector from a subspace in the Euclidean vector space. They can define the terms orthogonal and exterior product. They are also able to construct an orthogonal projection of vectors onto a subspace, and to measure the distance and deviation of a vector from a given subspace. Students are able to use the method of least squares, are familiar with the properties of exterior and orthogonal products of vectors in the Euclidean vector space, and know how to work with them.

From the previous chapter, you know that an Euclidean vector space is a direct sum of any of its subspaces and its orthogonal complement. With use of decomposition of any vector, you will learn how to measure the distance and deviance of a vector from a given subspace. You will also become familiar with natural applications of this apparatus of linear algebra in geometry and in solving systems of linear equations. You will learn that the notions of mixed products and vector products that you worked with in physics lessons in your secondary school can be generalised also for higher dimensions.

1.3.1 Gram determinant, exterior product and orthogonal product

Before introducing the definitions of *distance* and *angle* between a vector and a subspace, it is useful to look at the notions of *Gram matrix, orthogonal product* and *exterior product.*

Definition 1.53 Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in \boldsymbol{V}$. Then

1. a matrix
$$\mathcal{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = \begin{pmatrix} \boldsymbol{u}_1\cdot\boldsymbol{u}_1 & \boldsymbol{u}_1\cdot\boldsymbol{u}_2 & \ldots & \boldsymbol{u}_1\cdot\boldsymbol{u}_k \\ \boldsymbol{u}_2\cdot\boldsymbol{u}_1 & \boldsymbol{u}_2\cdot\boldsymbol{u}_2 & \ldots & \boldsymbol{u}_2\cdot\boldsymbol{u}_k \\ \vdots & \vdots & & \vdots \\ \boldsymbol{u}_k\cdot\boldsymbol{u}_1 & \boldsymbol{u}_k\cdot\boldsymbol{u}_2 & \ldots & \boldsymbol{u}_k\cdot\boldsymbol{u}_k \end{pmatrix}$$
is called a *Gram matrix of vectors* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ and it is denoted by $\mathcal{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$.
2. a *Gram determinant* (or *Gramian) of vectors* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ is said to be a number defined by
$$G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = \det \mathcal{G}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k).$$

Remark 1.54 A Gram matrix is evidently a real symmetric matrix.

Theorem 1.55 *For any $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in \boldsymbol{V}$ it holds:*

1. $G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) \geq 0$,
2. $G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k) = 0$, *if and only if vectors* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ *are linearly dependent,*
3. $\forall \pi \in S_k : \ G(\boldsymbol{u}_{\pi(1)}, \dots, \boldsymbol{u}_{\pi(k)}) = G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$.

A symbol S_k denotes a set of all permutations of a set $\{1, 2, \dots, k\}$.

Let us recall that to *orient a vector space* $\boldsymbol{V}$ means denoting one of its bases as a positive basis. *Positive bases* are then all bases consistently oriented; the other bases are called *negative bases*. For any vector space, there are exactly two possible orientations.

Definition 1.56 Let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \boldsymbol{V}_n$ and let $\mathcal{B}$ be a positive orthonormal basis of an oriented vector space V_n.

If we denote $\{\boldsymbol{u}_i\}_{\mathcal{B}} = (u_{i1}, \dots, u_{in})$, $1 \leq i \leq n$, then an *exterior product of vectors* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ *(with respect to the basis* $\mathcal{B}$*)* is said to be a number denoted by $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]_{\mathcal{B}}$ and defined in the following way:

$$[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]_{\mathcal{B}} = \begin{vmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{vmatrix}.$$

It follows from Definition 1.56 that the value of an exterior product of given vectors formally depends on the choice of the basis. Let us take a close look at this dependence.

Let $\mathcal{B}, \mathcal{B}'$ be orthonormal bases. The definition of the product of matrices and the relation for the transformation of coordinates of vectors imply the following: having denoted coordinates of given vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ with respect to the basis $\mathcal{B}'$ in an analogical way as in definition 1.56, we may write

$$\begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{pmatrix} = \begin{pmatrix} u'_{11} & u'_{12} & \dots & u'_{1n} \\ u'_{21} & u'_{22} & \dots & u'_{2n} \\ \vdots & \vdots & & \vdots \\ u'_{n1} & u'_{n2} & \dots & u'_{nn} \end{pmatrix} (\mathcal{B}, \mathcal{B}').$$

Based on Theorem 1.35, we may formulate the following theorem.

Theorem 1.57

1. *The value of an exterior product of given vectors does not depend on the choice of a positive orthonormal basis.*
2. *A change of the orientation of a vector space implies that a value of an exterior product of given vectors turns into opposite number.*

Remark 1.58 In the chosen orientation of a vector space $\boldsymbol{V}$, the choice of the positive orthonormal basis is thus not important, which is why we can denote an exterior product only by $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$.

Theorem 1.59 *Let* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \boldsymbol{V}_n$. *Then it holds:*

1. $[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]^2 = G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$,
2. $\forall \pi \in S_n\colon [\boldsymbol{u}_{\pi(1)}, \dots, \boldsymbol{u}_{\pi(n)}] = \operatorname{sgn} \pi [\boldsymbol{u}_1, \dots, \boldsymbol{u}_n]$,
3. $\forall i,\ 1 \le i \le n, \forall \boldsymbol{u}_i, \boldsymbol{u}'_i \in V\colon$
$$[\boldsymbol{u}_1, \dots, (\boldsymbol{u}_i + \boldsymbol{u}'_i), \dots, \boldsymbol{u}_n] = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_n] + [\boldsymbol{u}_1, \dots, \boldsymbol{u}'_i, \dots, \boldsymbol{u}_n],$$
4. $\forall i,\ 1 \le i \le n, \forall c \in \mathbb{R}\colon [\boldsymbol{u}_1, \dots, c\boldsymbol{u}_i, \dots, \boldsymbol{u}_n] = c[\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_n]$.

Theorem 1.60 *Let* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1} \in \boldsymbol{V}_n$ *and let* $\boldsymbol{V}_n$ *be an oriented vector space. Then there exists exactly one vector* $\boldsymbol{u}^* \in V$ *such that it holds:*

1. $\boldsymbol{u}^* \perp \boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$,
2. $\|\boldsymbol{u}^*\| = \sqrt{G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1})}$,
3. *if* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$ *are linearly independent,* $<\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}, \boldsymbol{u}^*>$ *form a positive basis of a space* $\boldsymbol{V}$.

? If you choose a positive orthonormal basis $\mathcal{B} = <\boldsymbol{e}_1, \dots, \boldsymbol{e}_n>$ of a vector space $\boldsymbol{V}_n$, and put

$$\{\boldsymbol{u}_i\}_{\mathcal{B}} = (u_1, \dots, u_n),\ 1 \le i \le n-1,$$

followed by

$$\boldsymbol{u}^* = \begin{vmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n-1\ 1} & u_{n-1\ 2} & \dots & u_{n-1\ n} \\ \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_n \end{vmatrix}, \tag{1.8}$$

can you demonstrate that $\boldsymbol{u}^*$ meets the requirements of Theorem 1.60?

[Instruction: from (1.8), it follows that a vector $\boldsymbol{u}^*$ has its coordinates $(\mathcal{U}_{n1}, \dots, \mathcal{U}_{nn})$ in $\mathcal{B}$ where $\mathcal{U}_{ij}$ denotes an algebraic complement of the element in position (i, j) in a given matrix. If you then consider a matrix

$$\begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n-1\ 1} & u_{n-1\ 2} & \dots & u_{n-1\ n} \\ x_1 & x_2 & \dots & x_n \end{pmatrix},$$

it is evident that its determinant is equal to the scalar product $\boldsymbol{u}^* \cdot \boldsymbol{x}$ for any $\boldsymbol{x}, \{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, \dots, x_n)$.]

Definition 1.61 Let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1} \in \boldsymbol{V}_n$ and let $\boldsymbol{V}$ be oriented. Then a vector $\boldsymbol{u}^* \in \boldsymbol{V}$ meeting the requirements of Theorem 1.60 is called an *orthogonal product of vectors* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$ and is denoted by $\boldsymbol{u}_1 \times \boldsymbol{u}_2 \times \dots \times \boldsymbol{u}_{n-1}$.

Remark 1.62 Let be given vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$ of the oriented vector space $\boldsymbol{V}_n$.

- The orthogonal product is uniquely determined by vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$ and the choice of the orientation of a vector space $\boldsymbol{V}_n$.
- If a positive orthonormal basis of $\boldsymbol{V}$ is given, the orthogonal product $\boldsymbol{u}_1 \times \dots \times \boldsymbol{u}_{n-1}$ is equal to the symbolic determinant (1.8).

Theorem 1.63 *Let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$ be vectors of an oriented vector space $\boldsymbol{V}$. Then $\boldsymbol{u} = \boldsymbol{u}_1 \times \dots \times \boldsymbol{u}_{n-1}$ if and only if the following holds for every $\boldsymbol{x} \in \boldsymbol{V}$:*

$$[\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}, \boldsymbol{x}] = \boldsymbol{u} \cdot \boldsymbol{x}.$$

Theorem 1.64 *Let* $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1} \in \boldsymbol{V}_n$. *Then it holds:*

1. $\forall \pi \in S_{n-1}\colon \boldsymbol{u}_{\pi(1)} \times \dots \times \boldsymbol{u}_{\pi(n-1)} = \operatorname{sgn} \pi (\boldsymbol{u}_1 \times \dots \times \boldsymbol{u}_{n-1})$,

2. $\forall i,\ 1 \le i \le n{-}1, \forall \boldsymbol{u}_i, \boldsymbol{u}'_i \in V$:
   $\boldsymbol{u}_1 \times \dots \times (\boldsymbol{u}_i + \boldsymbol{u}'_i) \times \dots \times \boldsymbol{u}_{n-1} = \boldsymbol{u}_1 \times \dots \times \boldsymbol{u}_i \times \dots \times \boldsymbol{u}_{n-1} + \boldsymbol{u}_1 \times$
   $\times \dots \times \boldsymbol{u}'_i \times \dots \times \boldsymbol{u}_{n-1}$,

3. $\forall i,\ 1 \le i \le n{-}1, \forall c \in \mathbb{R}$:
   $\boldsymbol{u}_1 \times \dots \times c\boldsymbol{u}_i \times \dots \times \boldsymbol{u}_{n-1} = c(\boldsymbol{u}_1 \times \dots \times \boldsymbol{u}_i \times \dots \times \boldsymbol{u}_{n-1})$,

4. *if the orientation of a vector space* $\boldsymbol{V}$ *changes, the orthogonal product turns into the opposite vector.*

Definition 1.65 In a space $\boldsymbol{V}_3$, an exterior product of three vectors is called a *mixed product* while an orthogonal product of two vectors is called a *cross product.*

From Theorem 1.63, it follows:

Corollary 1.66 *For any vectors* $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3 \in \boldsymbol{V}_3$, *it holds:*

$$[\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3] = (\boldsymbol{u}_1 \times \boldsymbol{u}_2) \cdot \boldsymbol{u}_3.$$

Theorem 1.67 *Let* $\boldsymbol{N} \subset\subset \boldsymbol{V}$ *be a vector hyperplane and let* $<\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}>$ *be its arbitrary basis. Then a vector* $\boldsymbol{u}_1 \times \dots \times \boldsymbol{u}_{n-1}$ *is a normal vector of a hyperplane* $\boldsymbol{N}$.

Compare the properties of an orthogonal product of two vectors of $\boldsymbol{V}_3$ according to Theorem 1.60 with the properties you used to define a vector product in secondary school.

Example 1.68 In an Euclidean vector space $\boldsymbol{V}$, with respect to the positive orthonormal basis $\mathcal{B} = <\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4>$, it is given:

$$\{\boldsymbol{u}_1\}_{\mathcal{B}} = (1,0,1,0),\ \{\boldsymbol{u}_2\}_{\mathcal{B}} = (0,1,1,0),\ \{\boldsymbol{u}_3\}_{\mathcal{B}} = (1,1,1,1).$$

Compute an orthogonal product $\boldsymbol{u}_1 \times \boldsymbol{u}_2 \times \boldsymbol{u}_3$.
[Instruction: proceed according to Remark 1.62:
$\boldsymbol{u}_1 \times \boldsymbol{u}_2 \times \boldsymbol{u}_3 = \boldsymbol{e}_1 + \boldsymbol{e}_2 - \boldsymbol{e}_3 - \boldsymbol{e}_4$, which means that the orthogonal product has the following coordinates: $(1, 1, -1, -1)$.]

1.3.2 Distance and deviation in the Euclidean vector space

Let us go back to Theorem 1.45 according to which every vector $\boldsymbol{x} \in \boldsymbol{V}$ with respect to any subspace $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ can be written in one and only one way as

$$\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{x}^\perp, \text{ kde } \boldsymbol{x}^* \in \boldsymbol{W},\ \boldsymbol{x}^\perp \in \boldsymbol{W}^\perp. \tag{1.9}$$

Definition 1.69 Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$, let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$ and let $\boldsymbol{x}^*$, $\boldsymbol{x}^\perp$ be vectors complying with (1.9). Then a vector $\boldsymbol{x}^*$ is called an *orthogonal projection of a vector $\boldsymbol{x}$ onto a subspace $\boldsymbol{W}$*, while a vector is called a *perpendicular of a vector $\boldsymbol{x}$ to a subspace $\boldsymbol{W}$*.

Theorem 1.70 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds:*

1. *An orthogonal projection of a vector $\boldsymbol{x}$ onto a subspace $\boldsymbol{W}$ is equal to the perpendicular of a vector $\boldsymbol{x}$ to a subspace $\boldsymbol{W}^\perp$,*
2. *A perpendicular of a vector $\boldsymbol{x}$ to a subspace $\boldsymbol{W}$ is equal to the orthogonal projection of a vector $\boldsymbol{x}$ onto a subspace $\boldsymbol{W}^\perp$.*

Theorem 1.71 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds:*

$$\boldsymbol{x} \in \boldsymbol{W} \;\Leftrightarrow\; \boldsymbol{x}^* = \boldsymbol{x} \;\Leftrightarrow\; \boldsymbol{x}^\perp = \boldsymbol{o},$$

$$\boldsymbol{x} \perp \boldsymbol{W} \;\Leftrightarrow\; \boldsymbol{x}^* = \boldsymbol{o} \;\Leftrightarrow\; \boldsymbol{x}^\perp = \boldsymbol{x}.$$

? In a space $\boldsymbol{V}$, there exists exactly one vector for which $\boldsymbol{x} = \boldsymbol{x}^\perp = \boldsymbol{x}^*$. Which one is it?

[Use Theorem 1.71.]

Theorem 1.72 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds:*

$$\|\boldsymbol{x}\|^2 = \|\boldsymbol{x}^*\|^2 + \|\boldsymbol{x}^\perp\|^2.$$

Note: The *Pythagorean Theorem* thus holds true in Euclidean vector spaces.

The norm of the perpendicular will prove crucial for measuring the distance and deviation of a vector from a subspace.

Theorem 1.73 *Let $\boldsymbol{N}$ be a vector hyperplane in $\boldsymbol{V}$ and let $\boldsymbol{n}$ be its arbitrary normal vector. Then for any vector $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:*

$$\|\boldsymbol{x}^{\perp}\| = \frac{|\boldsymbol{x}\cdot\boldsymbol{n}|}{\|\boldsymbol{n}\|}.$$

Theorem 1.74 *Let $\boldsymbol{W}$ be an arbitrary subspace in $\boldsymbol{V}$ and let $<\boldsymbol{u}_1,\dots,\boldsymbol{u}_k>$ be some of its bases. Then for any vector $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:*

$$\|\boldsymbol{x}^{\perp}\| = \sqrt{\frac{G(\boldsymbol{u}_1,\dots,\boldsymbol{u}_k,\boldsymbol{x})}{G(\boldsymbol{u}_1,\dots,\boldsymbol{u}_k)}}.$$

Example 1.75 In an Euclidean vector space $\boldsymbol{V}$, a vector $\boldsymbol{x}$ and a subspace $\boldsymbol{W} = [\boldsymbol{u}_1,\boldsymbol{u}_2]$ are given. Find the orthogonal projection of a vector $\boldsymbol{x}$ onto a subspace $\boldsymbol{W}$ and the perpendicular to $\boldsymbol{W}$ if in an orthonormal basis it is given:

$$\{\boldsymbol{x}\} = (5,3,1);\ \{\boldsymbol{u}_1\} = (1,2,0),\ \{\boldsymbol{u}_2\} = (1,1,1).$$

[Instruction: express a vector $\boldsymbol{x}$ according to (1.9) and a vector $\boldsymbol{x}^*$ as a linear combination of vectors $\boldsymbol{u}_1,\boldsymbol{u}_2$. Then scalarly multiply the obtained identity by a vector $\boldsymbol{u}_1$, considering that vectors $\boldsymbol{u}_1,\boldsymbol{x}^{\perp}$ are orthogonal. Proceed in the same way with a vector $\boldsymbol{u}_2$. You thus obtain a system of equations

$$\begin{aligned}\boldsymbol{x}\boldsymbol{u}_1 &= c_1(\boldsymbol{u}_1\boldsymbol{u}_1) + c_2(\boldsymbol{u}_2\boldsymbol{u}_1),\\ \boldsymbol{x}\boldsymbol{u}_2 &= c_1(\boldsymbol{u}_1\boldsymbol{u}_2) + c_2(\boldsymbol{u}_2\boldsymbol{u}_2).\end{aligned}$$

Upon solving it, you can find c_1, c_2 and thereby a vector $\boldsymbol{x}^*$.
Solution: $\{\boldsymbol{x}^*\} = (3,4,2),\ \boldsymbol{x}^{\perp} = (2,-1,-1).$]

Definition 1.76 Let $\boldsymbol{W}\subseteq\subseteq\boldsymbol{V}$, let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$ and let $\boldsymbol{x}^*$ be an orthogonal projection of a vector $\boldsymbol{x}$ onto $\boldsymbol{W}$. Then the *deviation of a vector $\boldsymbol{x}$ from a subspace $\boldsymbol{W}$* is understood to be a number denoted by $\measuredangle(\boldsymbol{x},\boldsymbol{W})$ and defined by the relation

$$\measuredangle(\boldsymbol{x},\boldsymbol{W}) = \measuredangle(\boldsymbol{x},\boldsymbol{x}^*).$$

Theorem 1.77 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds*

$$\measuredangle(\boldsymbol{x}, \boldsymbol{W}) \in \left\langle 0, \frac{\pi}{2} \right\rangle.$$

Theorem 1.78 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds*

1. $\measuredangle(\boldsymbol{x}, \boldsymbol{W}) = \frac{\pi}{2} \Leftrightarrow \boldsymbol{x} \perp \boldsymbol{W}$,
2. $\measuredangle(\boldsymbol{x}, \boldsymbol{W}) = 0 \Leftrightarrow \boldsymbol{x} \in \boldsymbol{W} \wedge \boldsymbol{x} \neq \boldsymbol{o}$.

Theorem 1.79 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds*

1. $\|\boldsymbol{x}^*\| = \|\boldsymbol{x}\| \cos \measuredangle(\boldsymbol{x}, \boldsymbol{W})$,
2. $\|\boldsymbol{x}^\perp\| = \|\boldsymbol{x}\| \sin \measuredangle(\boldsymbol{x}, \boldsymbol{W})$.

Now we will use Theorems 1.73 and 1.74:

Theorem 1.80 *Let $\boldsymbol{N}$ be a vector hyperplane in $\boldsymbol{V}$ and let $\boldsymbol{n}$ be its arbitrary normal vector. Then for any vector $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:*

$$\measuredangle(\boldsymbol{x}, \boldsymbol{N}) = \arcsin \frac{|\boldsymbol{x} \cdot \boldsymbol{n}|}{\|\boldsymbol{x}\| \|\boldsymbol{n}\|}.$$

Theorem 1.81 *Let $\boldsymbol{W}$ be an arbitrary subspace in $\boldsymbol{V}$ and let $<\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k>$ be some of its bases. Then for any vector $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:*

$$\measuredangle(\boldsymbol{x}, \boldsymbol{W}) = \arcsin \frac{\sqrt{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{x})}}{\|\boldsymbol{x}\| \sqrt{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)}}.$$

Example 1.82 In an Euclidean vector space $\boldsymbol{V}$, a vector $\boldsymbol{x}$ and a subspace $\boldsymbol{W} = [\boldsymbol{u}_1, \boldsymbol{u}_2]$ are given. Compute the deviation of a vector $\boldsymbol{x}$ of a subspace $\boldsymbol{W}$ if in an orthonormal basis, it is given:

$$\{\boldsymbol{x}\} = (5, 3, 1); \ \{\boldsymbol{u}_1\} = (1, 2, 0), \ \{\boldsymbol{u}_2\} = (1, 1, 1).$$

[Solution:$\measuredangle(\boldsymbol{x}, \boldsymbol{W}) = \arcsin \frac{6}{\sqrt{210}}$; compare it by a direct calculation using the result in Example 1.75.]

The next theorem clearly shows the meaning of the notion *distance of a vector from a subspace* defined in Definition 1.84.

Theorem 1.83 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{x} \in \boldsymbol{V}$. Then for any vector $\boldsymbol{y}$ of $\boldsymbol{W}$, it holds:*

$$\rho(\boldsymbol{x}, \boldsymbol{y}) \geq \rho(\boldsymbol{x}, \boldsymbol{x}^*),$$

where equality happens only if $\boldsymbol{y} = \boldsymbol{x}^*$.

Definition 1.84 Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$, let $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$ and let $\boldsymbol{x}^*$ be an orthogonal projection of a vector $\boldsymbol{x}$ onto $\boldsymbol{W}$. Then the *distance of a vector $\boldsymbol{x}$ from a subspace $\boldsymbol{W}$* is understood to be a number denoted by $\rho(\boldsymbol{x}, \boldsymbol{W})$ and defined by the relation

$$\rho(\boldsymbol{x}, \boldsymbol{W}) = \rho(\boldsymbol{x}, \boldsymbol{x}^*).$$

Corollary 1.85 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{x} \in \boldsymbol{V}$. Then it holds:*

$$\rho(\boldsymbol{x}, \boldsymbol{W}) = \mathrm{Min}\{\rho(\boldsymbol{x}, \boldsymbol{y})\}_{\boldsymbol{y} \in \boldsymbol{W}}.$$

Theorem 1.86 *Let $\boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ and $\boldsymbol{x}$ be a vector of $\boldsymbol{V}$. Then it holds:*

1. $\rho(\boldsymbol{x}, \boldsymbol{W}) = 0 \Leftrightarrow \boldsymbol{x} \in \boldsymbol{W}$,
2. $\rho(\boldsymbol{x}, \boldsymbol{W}) = \|\boldsymbol{x}\| \Leftrightarrow \boldsymbol{x} \perp \boldsymbol{W}$.

If you consider how to express the distance of a vector from a subspace using its perpendicular, then using Theorems 1.73 and 1.74 you obtain:

Theorem 1.87 *Let $\boldsymbol{N}$ be a vector hyperplane in $\boldsymbol{V}$ and let $\boldsymbol{n}$ be its arbitrary normal vector. Then for any vector $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:*

$$\rho(\boldsymbol{x}, \boldsymbol{N}) = \frac{|\boldsymbol{x} \cdot \boldsymbol{n}|}{\|\boldsymbol{n}\|}.$$

Theorem 1.88 *Let $\boldsymbol{W}$ be an arbitrary subspace in $\boldsymbol{V}$ and let $<\boldsymbol{u}_1, \dots, \boldsymbol{u}_k>$ be some of its bases. Then for any vector $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:*

$$\rho(\boldsymbol{x}, \boldsymbol{W}) = \sqrt{\frac{G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{x})}{G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)}}.$$

Example 1.89 In an Euclidean vector space $\boldsymbol{V}$, a vector $\boldsymbol{x}$ and a subspace $\boldsymbol{W} = [\boldsymbol{u}_1, \boldsymbol{u}_2]$ are given. Measure the distance of the vector $\boldsymbol{x}$ from the subspace $\boldsymbol{W}$ if in an orthonormal basis, it is given:

$$\{\boldsymbol{x}\} = (5, 3, 1);\ \{\boldsymbol{u}_1\} = (1, 2, 0),\ \{\boldsymbol{u}_2\} = (1, 1, 1).$$

[Solution: $\rho(\boldsymbol{x}, \boldsymbol{W}) = \sqrt{6}$; compare this distance by a direct calculation with use of the length of the perpendicular using the result in Example 1.75.]

At the end of this lesson, let us take notice of several natural applications of the theory of distances and deviations from the point of view of the theory of solving systems of linear equations and Euclidean geometry.

From the previous part of your linear algebra course, you know that a system of equations

$$\boldsymbol{A}(x_1, x_2, \dots, x_n)^T = (b_1, b_2, \dots, b_r)^T$$

is solvable if and only if the column vector of right-sides can be expressed as a linear combination of column vectors of the matrix of the given system; the coefficients of this linear combination then constitute the solution of the system.

If we consider a system of linear equations with real coefficients, it is possible to say that an ordered n-tuple $(x_1, \dots, x_n)$ is the solution of the system if and only if the distance between vectors $(\boldsymbol{A}(x_1, x_2, \dots, x_n)^T)$ and $(b_1, b_2, \dots, b_r)^T$ of an Euclidean vector space $\mathbb{R}^r$ is equal to zero.

In the case that the system is not solvable, it makes sense to search for such "substitution" $(x_1, \dots, x_n)$ for which the number

$$\rho((\boldsymbol{A}(x_1, x_2, \dots, x_n)^T), (b_1, b_2, \dots, b_r)^T)$$

is the smallest possible one. Based on Theorem 1.83, the next theorem describes a method called the *method of the smallest squares*[7], which shows how to obtain this approximate solution.

[7]The name of the method comes from the fact that the distance of two vectors in $\mathbb{R}^r$ with the standard scalar product is given by the relation mentioned in Corollary 1.30.

Theorem 1.90 *(method of the smallest squares)*
Let $\boldsymbol{A} \in \mathcal{M}_{r\times n}(\mathbb{R})$ and let $\boldsymbol{b} \in \mathcal{M}_{r\times 1}(\mathbb{R})$. If $\boldsymbol{b}^*$ is a orthogonal projection of a vector $\boldsymbol{b}$ onto a column subspace of a matrix $\boldsymbol{A}$ and $x_1, x_2, \ldots, x_n$ is the solution of the system of linear equations

$$\boldsymbol{A}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{b}^* \tag{1.10}$$

then for every $y_1, y_2, \ldots, y_n \in \mathbb{R}$ it holds:

$$\left\| \boldsymbol{A}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} - \boldsymbol{b} \right\| \geq \left\| \boldsymbol{A}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \boldsymbol{b} \right\|$$

where equality happens if and only if $y_1, y_2, \ldots, y_n \in \mathbb{R}$ solves the system of linear equations (1.10).

Example 1.91 A certain physical process is described by the functional dependence $y = f(x)$ of which it is known that it is linear. During an experiment the following values were found

x	0	1	2
$f(x)$	1	5	3

Using the method of the smallest squares, find the parameters in the functional rule of the given dependence so that it would best express the conducted experiment.

[Instruction: Assume the functional dependence as $y = ax + b$. By substituting the above pairs of values, you obtain a set of three linear equations with unknowns a, b, which is not solvable. Proceed further according to Theorem 1.90.
Solution: $y = x + 2$.]

Let us now consider an Euclidean space E with a subspace M_k given by a point A and and let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ be a basis of a set $\boldsymbol{V}(M)$ of vectors associated with this subspace (this set of vectors forms a vector subspace; it is called a *vector subspace associated with M_k* or a *direction subspace of M_k*)[8].

[8] Let us note that a set of vectors associated with a subspace M of the Euclidean space can be obtained, for instance, as a set of radius vectors leading from the point A to every point of the subspace.

Let us take a look at the following two tasks:

- measuring the distance of an arbitrary point B from a subspace M,
- measuring the deviation of any line p with a direction vector $\boldsymbol{s}$ from a subspace M.

In the first case, let us imagine an orthogonal projection B^* of a point B onto a subspace M; the distance $\rho(B, M)$ is defined in geometry as equal to the distance $\rho(B, B^*)$. If we consider a vector $B - A$, a vector $B^* - A$ is evidently its orthogonal projection onto the direction subspace $\boldsymbol{V}(M)$; thus it holds that $\rho(B, M) = \|B^* - B\| = \rho((B - A), \boldsymbol{V}(M))$.

In the second case, let us choose two different arbitrary points B, C on a line p. A vector $C - B = \boldsymbol{s}$ is a direction vector of this line. Let us construct their orthogonal projections C^*, B^*. In geometry, the deviation $\measuredangle(p, M)$ is defined as a deviation of lines BD and B^*D^* or it is equal to $\frac{\pi}{2}$ in the case $p \perp M$ (i.e. $B^* = D^*$). It is evident that the vector $C^* - B^*$ is an orthogonal projection of the vector $\boldsymbol{s}$ onto the direction subspace $\boldsymbol{V}(M)$; thus it holds that $\measuredangle(p, M) = = \measuredangle(\boldsymbol{s}, \boldsymbol{V}(M))$.

It follows that the corollary of Theorems 1.81 and 1.88 are the following theorems of Euclidean geometry.

Theorem 1.92 *Let M be an arbitrary subspace of the Euclidean space E defined by a point A and let $<\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k>$ be some of the bases of its direction subspace. Then for any point $B \in E$, it holds:*

$$\rho(B, M) = \sqrt{\frac{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, (B - A))}{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)}}.$$

Theorem 1.93 *Let M be an arbitrary subspace of the Euclidean space E defined by a point A and let $<\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k>$ be some of the bases of its direction subspace. Then for any line p in E with a direction vector $\boldsymbol{s}$, it holds:*

$$\measuredangle(p, M) = \arcsin \frac{\sqrt{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{s})}}{\|\boldsymbol{s}\| \sqrt{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)}}.$$

Both the exterior and the orthogonal vector product are also geometrically significant. Let us mention them at least for a three-dimensional Euclidean space E.

Theorem 1.94 *Let be given a parallelepiped $A_1B_1C_1D_1A_2B_2C_2D_2$ in the Euclidean space E_3. Then for its volume, it holds:*

$$V = |[(B_1 - A_1), (D_1 - A_1), (A_2 - A_1)]|.$$

Theorem 1.95 *Let be given a parallelogram $ABCD$ in the Euclidean space E_3. Then for its surface, it holds:*

$$S = \|(B - A) \times (D - A)\|.$$

Example 1.96 In an Euclidean space E_3, there are given points $A = [1, 1, 1]$, $B = [3, 1, 1]$, $C = [3, 1, 3]$. Measure the volume of a triangle ABC.

[Instruction: Use Theorem 1.95. Result: S=2]

Notes:

2 Homomorphisms of vector spaces

2.1 Basic notions

Students are able to define the term homomorphism of vector spaces. They can recognize mappings that are homomorphisms of vector spaces. They know the notions of monomorphism, epimorphism, isomorphism, endomorphism and automorphism, and are able to match a given homomorphism with these terms. Students also know the notions of image and kernel of homomorphism, and are able to find them. They can construct an analytic expression of homomorphism in selected bases and find a matrix of a homomorphism in any pair of bases.

From the last semester, you know the notion *vector space* – you know that it has to do with a set of elements (vectors) endowed with a field of scalars and a pair of mappings – with *addition of vectors* and *multiplication of vectors by scalars* (sc. scalar multiplication which must not be confused with the scalar product of two vectors). We will now show that of all the mappings which we can imagine between pairs of vector spaces (more precisely, between their sets of vectors) the important mappings are those that retain both mappings (they are called *homomorphisms*). You will see that to some extent, they "transfer" the structure of one space onto another, creating in some special cases of isomorphism a faithful "copy" of the former vector space. You will learn how to express these mappings using equations describing the coordinates of an image using the coordinates of the pattern. You will become familiar with how to represent a homomorphism using a matrix.

Definition 2.1 Let $(\boldsymbol{V}, +, T, \cdot)$ and $(\boldsymbol{W}, \oplus, T, \circ)$ be vector spaces[9]. A mapping $f\colon \boldsymbol{V} \to \boldsymbol{W}$ is called a *homomorphism of a vector space* $\boldsymbol{V}$ *to a vector space* $\boldsymbol{W}$ if it has the following properties:

1. $\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}$: $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) \oplus f(\boldsymbol{v})$,
2. $\forall \boldsymbol{u} \in \boldsymbol{V}, \forall t \in T$: $f(t \cdot \boldsymbol{u}) = t \circ f(\boldsymbol{u})$.

Remark 2.2

- If it is not stated clearly otherwise, we will continue to consider all vector spaces over the same field T;

[9] Let us take notice that both vector spaces have the same scalar field.

- if we know in every case to which vector space a given pair of vectors belongs, there is no danger of misunderstanding, so we can denote vector addition also in the latter space by the same symbol "+". We can just the same denote a multiplication of a vector by a scalar with the same symbol "$\cdot$", or we can leave out this dot altogether. For the same reason, we will also denote a zero vector in both spaces by the same symbol $\boldsymbol{o}$;
- we will denote all vector spaces only by the symbol of the corresponding set of vectors, i.e. instead of writing $(\boldsymbol{V}, +, T, \cdot)$, we will only write $\boldsymbol{V}$.

Notation 2.1 Let $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces. We will denote a set of all homomorphisms $\boldsymbol{V}$ to $\boldsymbol{W}$ as $\mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$.

Example 2.3 Let $\boldsymbol{V}_3, \boldsymbol{W}_3$ be vector spaces over $\mathbb{R}$ and let us consider a mapping $f\colon \boldsymbol{V}_3 \to \boldsymbol{W}_3$ given with respect to the selected bases $\mathcal{B}$, $\mathcal{C}$ by the rule:

$$\forall \boldsymbol{x} \in \boldsymbol{V}_3\colon \{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, x_2, x_3) \longmapsto \{f(\boldsymbol{x})\}_{\mathcal{C}} = (y_1, y_2, y_3),$$

where:

1. $y_1 = x_1$, $y_2 = 2x_1 + x_2$, $y_3 = 3x_1 - x_3$,
2. $y_1 = x_1^2$, $y_2 = 2x_1 + x_2$, $y_3 = 3x_1 - x_3$.
3. $y_1 = x_1$, $y_2 = 2x_1 + x_2 + 5$, $y_3 = 3x_1 - x_3$.

Is a mapping f a homomorphism in all the individual cases?
[1. yes; 2.,3. no.]

Definition 2.4 Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$;

1. if f is injective, it is called a *monomorphism*,
2. if f is surjective, it is called an *epimorphism*,
3. if f is bijective, it is called an *isomorphism*,
4. if $\boldsymbol{W} = \boldsymbol{V}$, f is called an *endomorphism of a vector space* $\boldsymbol{V}$,
5. if $\boldsymbol{W} = \boldsymbol{V}$ and f is bijective, f is called an *automorphism of a vector space* $\boldsymbol{V}$.

Definition 2.5 Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Then

1. an *image of homomorphism* f is understood to be a set denoted by $\mathrm{Im} f$ and defined as:

$$\mathrm{Im}\, f = \{\boldsymbol{y} \in \boldsymbol{W}; \exists \boldsymbol{x} \in \boldsymbol{V} : \boldsymbol{y} = f(\boldsymbol{x})\},$$

2. a *kernel of homomorphism* f is understood to be a set denoted by $\mathrm{Ker} f$ and defined as:

$$\mathrm{Ker}\, f = \{\boldsymbol{x} \in \boldsymbol{V}; f(\boldsymbol{x}) = \boldsymbol{o}\}.$$

Theorem 2.6 *Let* $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. *Then*

1. $\mathrm{Im}\, f \subseteq\subseteq \boldsymbol{W}$,
2. $\mathrm{Ker}\, f \subseteq\subseteq \boldsymbol{V}$.

Corollary 2.7 *Let* $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. *Then*

1. f *is an epimorphism* $\boldsymbol{V}$ *onto* $\mathrm{Im}\, f$,
2. *if* f *is a monomorphism* $\boldsymbol{V}$ *to* $\boldsymbol{W}$, *then it is an isomorphism* $\boldsymbol{V}$ *onto* $\mathrm{Im}\, f$.

Example 2.8 Let $\boldsymbol{V}_3, \boldsymbol{W}_4$ be vector spaces $\mathbb{R}$ and let us consider a mapping $f : \boldsymbol{V}_3 \to \boldsymbol{W}_4$ given with respect to the selected bases $\mathcal{B}, \mathcal{C}$ by the rule:

$$\forall \boldsymbol{x} \in \boldsymbol{V}_3 : \ \{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, x_2, x_3) \longmapsto \{f(\boldsymbol{x})\}_{\mathcal{C}} = (y_1, y_2, y_3, y_4),$$

where:

$$\begin{aligned} y_1 &= x_1 + 2x_2 - x_3 \\ y_2 &= x_1 + 5x_2 - 5x_3 \\ y_3 &= \phantom{x_1 + {}} 3x_2 - 4x_3 \\ y_4 &= x_1 + 8x_2 - 9x_3. \end{aligned}$$

Find its kernel and image.
[$\mathrm{Ker}\, f = \{\boldsymbol{x} \in \boldsymbol{V}, \{\boldsymbol{x}\}_{\mathcal{B}} \in [(-5, 4, 3)]\}$,
$\mathrm{Im}\, f = \{\boldsymbol{y} \in \boldsymbol{W}, \{\boldsymbol{y}\}_{\mathcal{C}} \in [(-2, -1, 1, 0), (1, 1, 0, 1)]\}$.]

Theorem 2.9 *Let* $\boldsymbol{V}$, $\boldsymbol{W}$ *be vector spaces. If* f *is an isomorphism* $\boldsymbol{V}$ *onto* $\boldsymbol{W}$, *then* f^{-1} *is an isomorphism* $\boldsymbol{W}$ *onto* $\boldsymbol{V}$.

Theorem 2.10 *Let* $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$*. Then it holds:*

1. f *is an epimorphism* $\boldsymbol{V}$ *onto* $\boldsymbol{W}$ *if and only* $\mathrm{Im}\, f = \boldsymbol{W}$*,*
2. f *is a monomorphism* $\boldsymbol{V}$ *to* $\boldsymbol{W}$ *if and only if* $\mathrm{Ker}\, f = \{\boldsymbol{o}\}$*.*

According to Definition 2.1, homomorphism preserves two basic mappings of a vector space $+$, $\cdot$. The next theorem shows that homomorphism preserves also any linear combination of vectors. This theorem will also show the relationships of the linear (in)dependence of origins and images.

Theorem 2.11 *Let* $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$*. Then it holds:*

1. *for every* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \boldsymbol{V}$ *and every* $t_1, \ldots, t_k \in T$*, it holds:*
$$f(t_1\boldsymbol{x}_1 + \cdots + t_k\boldsymbol{x}_k) = t_1 f(\boldsymbol{x}_1) + \cdots + t_k f(\boldsymbol{x}_k),$$
2. *for every* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \boldsymbol{V}$*, it holds: if* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ *are linearly dependent, then* $f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_k)$ *are also linearly dependent,*
3. *for every* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \boldsymbol{V}$*, it holds: if* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ *are linearly independent and if* f *is a monomorphism, then* $f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_k)$ *are also linearly independent.*

- Formulate a contrapositive of the statement in subsection 2.
- Find an example showing that the assumption that the homomorphism f is injective cannot be left out in subsection 3.

Corollary 2.12 *Let* $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$*,* $\boldsymbol{U} \subseteq\subseteq \boldsymbol{V}$*. Then it holds:*

1. *a set* $f(\boldsymbol{U})$ *is a subspace in* $\boldsymbol{W}$*,*
2. *if* $\mathcal{M}$ *is a set of generators of a subspace* $\boldsymbol{U}$*, then* $f(\mathcal{M})$ *is a set of generators of a subspace* $f(\boldsymbol{U})$*,*
3. *if* $\mathcal{G}$ *is a set of generators of a space* $\boldsymbol{V}$*, then* f *is an epimorphism if and only if* $f(\mathcal{G})$ *is a set of generators of a space* $\boldsymbol{W}$*,*
4. *if* $\mathcal{B}$ *is a basis of a space* $\boldsymbol{V}$*, then* f *is an isomorphism if and only if* $f(\mathcal{B})$ *is a basis of a space* $\boldsymbol{W}$*.*

Homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$, is, as a mapping, a set of ordered pairs $\{(\boldsymbol{x}, f(\boldsymbol{x})) \in \boldsymbol{V} \times \boldsymbol{W}, \boldsymbol{x} \in \boldsymbol{V}\}$. In the case of homomorphisms, however, it is not necessary (often even impossible) to define f by a list of all these ordered pairs – the important question is how many of these ordered pairs it is necessary to know so that it is possible to determine the mapping uniquely. The answer lies in the next *theorem on determination of homomorphism.*

Theorem 2.13 *Let $\boldsymbol{V}, \boldsymbol{W}$ be vector spaces. Then for every basis $<\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n>$ of a vector space $\boldsymbol{V}$ and for every n-tuple of vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ of a vector space $\boldsymbol{W}$, there exists one and only one homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$ with the property:*

$$f(\boldsymbol{v}_i) = \boldsymbol{w}_i,\ i = 1, \ldots, n.$$

Definition 2.14 Let $\boldsymbol{V}, \boldsymbol{W}$ be vector spaces. We say that these *vector spaces are isomorphic*, which we will denote as $\boldsymbol{V} \cong \boldsymbol{W}$ if there exists an isomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$.

Remark 2.15 Based on Theorem 2.9, we will see that the relation to be isomorphic is indeed symmetrical. The above definition is thus correct in this sense.

If vector spaces $(\boldsymbol{V}, +, T, \cdot)$ and $(\boldsymbol{W}, \oplus, T, \circ)$ are isomorphic, it implies more than just the fact that there exists a bijection between the sets $\boldsymbol{V}$ and $\boldsymbol{W}$ of their vectors. Let f be an isomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$. Now let us take a look at how to add two vectors $\boldsymbol{u}$, $\boldsymbol{w}$ belonging to $\boldsymbol{W}$. Since based on Theorem 2.9 and Definition 2.1

$$\boldsymbol{u} \oplus \boldsymbol{v} = f(f^{-1}(\boldsymbol{u})) \oplus f(f^{-1}(\boldsymbol{w})) = f(f^{-1}(\boldsymbol{u}) + f^{-1}(\boldsymbol{w})),$$

it is clear that this sum is fully determined by the operation $+$ vector addition in $\boldsymbol{V}$.

Similarly, you will find that also the multiplication $\circ$ of vectors by scalars in $\boldsymbol{W}$ is fully determined by the multiplication $\cdot$ in $\boldsymbol{V}$.

We can then conclude that two isomorphic vector spaces are their mutual "exact copies", which we express by saying that *two isomorphic vector spaces differ only in the names of their elements.*

Theorem 2.16 will give you a criterion when two vector spaces are isomorphic.

See that a mapping $\boldsymbol{V}_n \to T^n$, called a system of coordinates, which has been known to you from the previous part of your linear algebra course, is an isomorphism of the above vector spaces. Drawing on this fact, derive the validity of the next theorem!

[Instruction: use Theorem 2.9.]

Theorem 2.16 *Two vector spaces (over the same field) are isomorphic if and only if they have the same dimension.*

Definition 2.17 Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Let $\mathcal{B}$, $\mathcal{C}$ be, respectively, arbitrary bases of vector spaces $\boldsymbol{V}_n$ and $\boldsymbol{W}_m$, $\mathcal{B} = <\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n>$. If we denote

$$\{f(\boldsymbol{a}_i)\}_{\mathcal{C}} = (a_{i1}, a_{i2}, \ldots, a_{in}),\ i{=}1, \ldots, n,$$

then a matrix $(a_{ij}) \in \mathcal{M}_{n\times m}(T)$ is called the *matrix of homomorphism f with respect to the bases* $\mathcal{B}$, $\mathcal{C}$ and is denoted by $(f, \mathcal{B}, \mathcal{C})$.

Using Theorem 2.11 (1), you can verify the validity of the next theorem.

Theorem 2.18 *Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Let $\mathcal{B}$, $\mathcal{C}$ be, respectively, arbitrary bases of vector spaces $\boldsymbol{V}$ and $\boldsymbol{W}$. Then for any vector $\boldsymbol{x}$ from $\boldsymbol{V}$, it holds:*

$$\{f(\boldsymbol{x})\}_{\mathcal{C}} = \{\boldsymbol{x}\}_{\mathcal{B}}(f, \mathcal{B}, \mathcal{C}),$$

or:

if $\{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, \ldots, x_n)$, then for $f(\boldsymbol{x})$, it holds:

$$\{f(\boldsymbol{x})\}_{\mathcal{C}} = (y_1, \ldots, y_m) \Leftrightarrow \forall j, 1 \le j {\le} m\colon y_j = \sum_{i=1}^{n} a_{ij}x_i, \tag{2.1}$$

where $(a_{ij})_{n\times m} = (f, \mathcal{B}, \mathcal{C})$.

You already know that to determine a homomorphism f, it is not necessary to know all the ordered pairs $(\boldsymbol{x}, f(\boldsymbol{x}))$. If you know the images of vectors of some basis, you can find out the coordinates of the image of any of the vectors using the so-called *analytic expression of homomorphism with respect to the chosen pair of bases*, as we will further call the system of equalities (2.1).

As Theorem 2.13 shows, it does not evidently make a difference if you define a homomorphism by defining the images of the elements of a certain basis, using a matrix of homomorphism or using an analytic expression.

Theorem 2.19 *Let $\mathcal{B}$, $\mathcal{C}$ be, respectively, arbitrary bases of spaces $\boldsymbol{V}$, $\boldsymbol{W}$. Then a mapping $H_{\mathcal{BC}}\colon \mathrm{Hom}(\boldsymbol{V}_n, \boldsymbol{W}_m) \to \mathcal{M}_{n\times m}(T)$ defined by the relation*

$$\forall f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})\colon H_{\mathcal{BC}}(f) = (f, \mathcal{B}, \mathcal{C})$$

is a bijection of the above sets.

Example 2.20 Let $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces. Let $\mathcal{B} = <\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3>$ be a basis of a space $\boldsymbol{V}$ and let $\mathcal{C} = <\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3>$ be a basis of a space $\boldsymbol{W}$. Write an analytic expression of homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$ if it holds:

$$\begin{aligned} f(\boldsymbol{e}_1) + 2f(\boldsymbol{e}_2) &= \boldsymbol{b}_1 + 4\boldsymbol{b}_2 + 3\boldsymbol{b}_3 \\ f(\boldsymbol{e}_1) - f(\boldsymbol{e}_2) + f(\boldsymbol{e}_3) &= \boldsymbol{b}_1 + \boldsymbol{b}_2 + 2\boldsymbol{b}_3 \\ f(\boldsymbol{e}_2) + f(\boldsymbol{e}_3) &= \boldsymbol{b}_2 - \boldsymbol{b}_3. \end{aligned}$$

[Solution: the matrix of the searched homomorphism is $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.]

Theorem 2.21 *Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Let $\mathcal{B}, \mathcal{B}'$ or $\mathcal{C}, \mathcal{C}'$ be arbitrary bases of a vector space $\boldsymbol{V}$ or $\boldsymbol{W}$. Then it holds:*

$$(f, \mathcal{B}', \mathcal{C}') = (\mathcal{B}, \mathcal{B}')(f, \mathcal{B}, \mathcal{C})(\mathcal{C}', \mathcal{C}).$$

Theorem 2.22 *Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Let $\mathcal{B}, \mathcal{B}'$ or $\mathcal{C}, \mathcal{C}'$ be arbitrary bases of a vector space $\boldsymbol{V}$ or $\boldsymbol{W}$. Then it holds:*

$$\mathrm{h}(f, \mathcal{B}', \mathcal{C}') = \mathrm{h}(f, \mathcal{B}, \mathcal{C}) = \dim \mathrm{Im}\, f.$$

Note: The common rank of all the matrices of a given homomorphism f is said to be the *rank of a homomorphism f*.

Theorem 2.23 *Let $f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Then it holds:*

$$\dim \mathrm{Ker}\, f + \dim \mathrm{Im}\, f = \dim \boldsymbol{V}.$$

Corollary 2.24 *Let $f \in \operatorname{Hom}(\boldsymbol{V}, \boldsymbol{W})$ and let $\dim \boldsymbol{V} = \dim \boldsymbol{W}$. Then the following conditions are equivalent:*

1. *f is an isomorphism,*
2. *f je f is an epimorphism,*
3. *f is a monomorphism.*

Corollary 2.25 *Homomorphism is an isomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$ if and only if its matrix in one (and thus in every) pair of the bases of spaces $\boldsymbol{V}$ and $\boldsymbol{W}$ is regular.*

Theorem 2.26 *Let $f \in \operatorname{Hom}(\boldsymbol{V}_n, \boldsymbol{W}_m)$. Then there exist, respectively, bases $\mathcal{B}$, $\mathcal{C}$ of spaces $\boldsymbol{V}$, $\boldsymbol{W}$ such that the analytic expression of a homomorphism f in these pair of bases is written as:*

$$\begin{array}{lllll}
y_1 & = x_1 & & & \\
y_2 & = & x_2 & & \\
 & \vdots & & & \\
y_r & = & & x_r & \\
y_{r+1} & = & & 0 & \\
 & \vdots & & & \\
y_m & = & & 0 & ,
\end{array}$$

where r is the dimension of the image of a homomorphism f.

What form does the matrix of homomorphism have in the bases according to Theorem 2.26?

Example 2.27 Let $\boldsymbol{V}$ and $\boldsymbol{W}$ be vector spaces and let consider a homomorphism $f \colon \boldsymbol{V} \to \boldsymbol{W}$ given with respect to some chosen pair of bases by the following analytic expression:

$$\begin{aligned}
y_1 &= x_1 + 2x_2 - x_3 \\
y_2 &= x_1 + 5x_2 - 5x_3 \\
y_3 &= \phantom{x_1 +{}} 3x_2 - 4x_3 \\
y_4 &= x_1 + 8x_2 - 9x_3
\end{aligned}$$

Find such a pair of bases so that the analytic expression of a given homomorphism has the form described in Theorem 2.26.

[Instruction: Let us denote the searched bases as $\mathcal{B} = <\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3>$, $\mathcal{C} = <\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4>$, respectively. Consider that the kernel of a homomorphism f is a subspace in $\boldsymbol{V}$ whose every element is mapped on a zero vector. Thus, first find the basis of the kernel $\operatorname{Ker} f$. Its vectors form the elements of the basis $\mathcal{B}$ "from the back to the front". Complete the basis of $\operatorname{Ker} f$ arbitrarily in a basis of a space $\boldsymbol{V}$, which will give you a basis $\mathcal{B}$. In this specific case, $\dim \operatorname{Ker} f = 1$; thus, its basis is a vector $\boldsymbol{e}_3$ which we will then complete with vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ in the $\mathcal{B}$. Now map vectors of the basis $\mathcal{B}$ not belonging to the kernel $\operatorname{Ker} f$ in a homomorphism f. According to Theorem 2.12, these vectors generate $\operatorname{Im} f$ and their number is equal to the dimension of $\operatorname{Im} f$ according to Theorem 2.23 – they are thus linearly independent. Let us then denote $\boldsymbol{b}_1 = f(\boldsymbol{e}_1)$, $\boldsymbol{b}_2 = f(\boldsymbol{e}_2)$ and subsequently arbitrarily complete vectors $\boldsymbol{b}_1, \boldsymbol{b}_2$ in the basis $\mathcal{C}$C of a space $\boldsymbol{W}$.

Consider the definition of the matrix of homomorphism. Then it is clear that

$$(f, \mathcal{B}, \mathcal{C}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and the analytic expression thus has the required form.]

Notes:

2.2 Vector space of homomorphisms; composition of homomorphisms

Students are able to define the structure of a vector space on a set of homomorphisms. They can determine the matrix of a sum of homomorphisms and a scalar multiple of homomorphisms. They know the relation between a vector space of homomorphisms and the isomorphic structure of matrices, and are able to use the relations between these two isomorphic structures. Students can also compose homomorphisms and determine the matrix of a composition of homomorphisms.

In the previous lesson, you became familiar with a set of homomorphisms of a vector space $\boldsymbol{V}$ to a vector space $\boldsymbol{W}$. Now you will demonstrate how to naturally define an addition of homomorphisms and a scalar multiplication of homomorphism, thereby obtaining the structure of a *vector space* on a set $\operatorname{Hom}(\boldsymbol{V}, \boldsymbol{W})$. You will then learn the properties of a map assigning to every homomorphism its matrix with respect to the chosen basis. In the previous part of your linear algebra course, you learned that a set of matrices, along with matrix addition and scalar multiplication of the matrix, forms a vector space. You will now see that this vector space is isomorphic to the constructed vector space of homomorphisms.

You will then learn how to compose homomorphisms.

Described operations with homomorphisms will elucidate the *naturalness* of the definitions of matrix addition, matrix multiplication and scalar multiplication of the matrix with which you became familiar in the last semester.

2.2.1 Vector space of homomorphisms

Definition 2.28 Let $f, g \in \operatorname{Hom}(\boldsymbol{V}, \boldsymbol{W})$, $t \in T$. Then the *sum of homomorphisms f and g* denotes a mapping $f + g \colon \boldsymbol{V} \to \boldsymbol{W}$ defined by the relation

$$\forall \boldsymbol{x} \in \boldsymbol{V} \colon (f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}),$$

a *scalar t-multiple of a homomorphism f* denotes a mapping $tf \colon \boldsymbol{V} \to \boldsymbol{W}$ defined by the relation

$$\forall \boldsymbol{x} \in \boldsymbol{V} \colon (tf)(\boldsymbol{x}) = tf(\boldsymbol{x}).$$

Verify the validity of the axioms of a vector space for a set $\operatorname{Hom}(\boldsymbol{V}, \boldsymbol{W})$ along with addition of homomorphisms and scalar multiplication of a homomorphism.

Theorem 2.29 *A set* $\mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$*, along with addition of homomorphisms and scalar multiplication of a homomorphism, forms a vector space over a field* T*.*

Note: A zero element of a vector space $\mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$ is the so-called *zero homomorphism* o defined for every $\boldsymbol{x}$ from $\boldsymbol{V}$ by the relation $o(\boldsymbol{x}) = \boldsymbol{o}$.

Theorem 2.30 *Let* $f, g \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$*,* $t \in T$*. Then for arbitrary bases* $\mathcal{B}$*,* $\mathcal{C}$*, respectively, of spaces* $\boldsymbol{V}$*,* $\boldsymbol{W}$*, it holds:*

$$(f + g, \mathcal{B}, \mathcal{C}) = (f, \mathcal{B}, \mathcal{C}) + (g, \mathcal{B}, \mathcal{C}),$$
$$(tf, \mathcal{B}, \mathcal{C}) = t(f, \mathcal{B}, \mathcal{C}).$$

If you consider Theorem 2.19, you obtain:

Theorem 2.31 *Let* $\mathcal{B}$*,* $\mathcal{C}$ *be bases, respectively, of spaces* $\boldsymbol{V}$*,* $\boldsymbol{W}$*. Then a mapping* $H_{\mathcal{BC}}\colon \mathrm{Hom}(\boldsymbol{V}_n, \boldsymbol{W}_m) \to \mathcal{M}_{n\times m}(T)$ *defined by the relation*

$$\forall f \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})\colon H_{\mathcal{BC}}(f) = (f, \mathcal{B}, \mathcal{C})$$

is an isomorphism of vector spaces $\mathrm{Hom}(\boldsymbol{V}_n, \boldsymbol{W}_m)$ *a* $\mathcal{M}_{n\times m}(T)$.

For arbitrary admissable (i, j), denote a matrix of $\mathcal{M}_{n\times m}(T)$ whose elements are zero except precisely the element in position (i, j) which is equal to 1, by a symbol E_{ij}.

Every matrix of $\mathcal{M}_{n\times m}(T)$ can evidently be written as a linear combination of the matrices of a set

$$\mathcal{E} = <E_{11}, E_{12}, \ldots, E_{1n}, \ldots, E_{n1}, E_{n2}, \ldots, E_{nm}>,$$

which can be done in exactly one way. A set $\mathcal{E}$ is thus the basis of a vector space $\mathcal{M}_{n\times m}(T)$.

It is shown in the following example:

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 8 & 7 \end{pmatrix} = 1\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{E_{11}} +2\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{E_{12}} -1\underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}}_{E_{13}} +8\underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{E_{22}} +7\underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{E_{23}}.$$

Based on Theorem 2.31, you obtain:

Corollary 2.32 *Let $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces. Then it holds:*

1. $\dim \operatorname{Hom}(\boldsymbol{V}, \boldsymbol{W}) = \dim \boldsymbol{V} . \dim \boldsymbol{W}$,
2. *if $\mathcal{B}$, $\mathcal{C}$ are bases, respectively, in spaces $\boldsymbol{V}_n$, $\boldsymbol{W}_m$ and if we define for every i, j, $1 \leq i \leq n$, $1 \leq j \leq m$, a homomorphism e_{ij} by the relation $(e_{ij}, \mathcal{B}, \mathcal{C}) = E_{ij}$, then a set*

$$<e_{11}, e_{12}, \ldots, e_{1n}, \ldots, e_{n1}, e_{n2}, \ldots, e_{nm}>$$

is the basis of a vector space $\operatorname{Hom}(\boldsymbol{V}, \boldsymbol{W})$,

3. *if $\mathcal{B}$, $\mathcal{C}$ are bases, respectively, in spaces $\boldsymbol{V}_n$, $\boldsymbol{W}_m$, then it holds:*

$$(f, \mathcal{B}, \mathcal{C}) = (a_{ij})_{n \times m} \Longleftrightarrow f = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} a_{ij} e_{ij}.$$

Remark 2.33

- Drawing on Theorem 2.30, you can see that matrix addition is naturally determined by addition of homomorphisms. Similarly, you can see that scalar multiplication of matrices is naturally determined by a scalar multiplication of homomorphisms.
- Based on Corollary 2.32, subsection 3, it is evident that the elements of the matrix of homomorphism have also another significance – their significance lies in the fact that they may be considered as the coordinates of a given homomorphism in the basis according to subsection 2.

2.2.2 Composition of homomorphisms

The second part of this lesson is devoted to the composition of homomorphisms. We can compose two homomorphisms as any other mappings[10]. It is, however, important that compositing homomorphisms results in *homomorphism*.

Theorem 2.34 *Let $\boldsymbol{U}$, $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces. Then for any homomorphisms $f \in \mathrm{Hom}(\boldsymbol{U}, \boldsymbol{V})$, $g \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$, it holds that $f \circ g \in \mathrm{Hom}(\boldsymbol{U}, \boldsymbol{W})$.*

Theorem 2.35 *Let $\boldsymbol{U}$, $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces, $f \in \mathrm{Hom}(\boldsymbol{U}, \boldsymbol{V})$, $g \in \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$. Then if $\mathcal{B}$, $\mathcal{C}, \mathcal{D}$ are arbitrary bases, respectively, in spaces $\boldsymbol{U}$, $\boldsymbol{V}$, $\boldsymbol{W}$, it holds:*

$$(f \circ g, \mathcal{B}, \mathcal{D}) = (f, \mathcal{B}, \mathcal{C})(g, \mathcal{C}, \mathcal{D}).$$

Remark 2.36 Drawing on Theorem 2.35, you can see that matrix multiplication is naturally determined by the composition of homomorphisms.

Corollary 2.37 *Let $f \in \mathrm{Hom}(\boldsymbol{U}, \boldsymbol{V})$ be an isomorphism and let $\mathcal{B}$, $\mathcal{C}$ be arbitrary bases, respectively, in spaces $\boldsymbol{U}$, $\boldsymbol{V}$. Then it holds:*

$$(f^{-1}, \mathcal{C}, \mathcal{B}) = (f, \mathcal{B}, \mathcal{C})^{-1}.$$

Theorem 2.38 *Let $\boldsymbol{U}$, $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces. Then for arbitrary homomorphisms $f, g \in \mathrm{Hom}(\boldsymbol{U}, \boldsymbol{V})$, $h, k \in \mathrm{Hom}(\boldsymbol{V}, \boldsymbol{W})$ and $t \in T$, it holds:*

1. $(f + g) \circ h = f \circ h + g \circ h$,
2. $f \circ (h + k) = f \circ h + f \circ k$,
3. $(tf) \circ h = t(f \circ h) = f \circ (th)$.

What are the corollaries of Theorem 2.38 for operations with matrices?

[10] In this text, a combination of two mappings α, β will be consistently denoted as:

$$(\alpha \circ \beta)(x) = \beta(\alpha(x)).$$

Notes:

2.3 Endomorphisms of a vector space

Students are able to define a ring structure and a linear algebra structure on a set of endomorphisms in a vector space. Knowing the relation between a ring or linear algebra of endomorphisms nad the corresponding isomorphic matrix structures, they can make use of the relations between these isomorphic structures. They can also recognize the automorphism of a vector space and are able to construct a group structure in a set of automorphisms. They know the relation between this group and the corresponding matrix structure isomorphic to it. Students are also able to define the notion of the projection of a vector space and to distinguish a projection between different endomorphisms.

In one of the previous lessons (Definition 2.4), you became familiar with the notion of *endomorphism of a vector space* as a homomorphism of a vector space to itself. In the last lesson, you learned how to add homomorphisms, how to multiply them by a scalar, and how to compose them. Now you will demonstrate that endomorphisms of a given vector space form a ring. A mapping assigning to an endomorphism its matrix with respect to the chosen basis forms an isomorphism of this ring onto the matrix ring. You will further learn the structure of linear algebra endomorphisms, of a group of automorphisms and, in both of these cases, of matrix structures isomorphic to them.
You will become familiar with projections of a vector space onto a subspace such as endomorphisms mapping a vector space on some of its subspaces and thus on a lower dimension vector space[11], which endows projections with wide practical significance.

Notation 2.2 Let $\boldsymbol{V}$ be a vector space. Then a set of endomorphisms of a vector space $\boldsymbol{V}$ (and thus a set $\mathrm{Hom}(\boldsymbol{V}, \boldsymbol{V})$) will be denoted by $\mathrm{End}(\boldsymbol{V})$. The matrix of endomorphism f in the basis $\mathcal{B}$ will be denoted only by $(f, \mathcal{B})$.

Verify the validity of axioms of the ring for the set $\mathrm{End}(\boldsymbol{V})$ along with addition and composition of homomorphisms.

Theorem 2.39 *Let $\boldsymbol{V}$ be a vector space. A set $\mathrm{End}(\boldsymbol{V})$, along with addition of endomorphisms $+$ and composition of endomorphisms $\circ$, forms a ring with a unit which is an identical endomorphism id. This ring is not generally commutative.*

The next theorem is a corollary of Theorems 2.30 and 2.35.

[11] With the exception of the trivial case of identity.

Theorem 2.40 *Let $\mathcal{B}$ be an arbitrary basis of a space $\boldsymbol{V}$. Then the mapping $H_{\mathcal{B}}\colon \mathrm{End}(\boldsymbol{V}_n) \to \mathcal{M}_{n\times n}(T)$ defined by the relation*

$$\forall f \in \mathrm{End}(\boldsymbol{V})\colon H_{\mathcal{B}}(f) = (f, \mathcal{B})$$

is an isomorphism of the rings $(\mathrm{End}(\boldsymbol{V}_n), +, \circ)$ *a* $(\mathcal{M}_{n\times n}(T), +, \cdot)$.

Definition 2.41 Let A be a set, T be a field, and let be given mappings

$$+\colon A \times A \to A, \quad \circ\colon A \times A \to A, \quad \cdot\colon T \times A \to A,$$

where

1. A along with mappings $+, \cdot$ is a vector space over T,
2. A along with mappings $+, \circ$ is a ring with a unit element,
3. $\forall a, b \in A, \forall t \in T\colon t \cdot (a \circ b) = (t \cdot a) \circ b = a \circ (t \cdot b)$.

Then a set A with the above mappings is called a *linear algebra over a field T*. The *order of an algebra A* is understood to be the dimension of A as a vector space.

Definition 2.42 Let A and B be a linear algebra over the same field. Let us say that a *linear algebra A is isomorphic to a linear algebra B* if there is a mapping $H\colon A \to B$ which is simultaneously an isomorphism of A a B as both vector spaces and rings.

The following two theorems follow from Theorems 2.29, 2.31, 2.38, 2.39 and 2.40.

Theorem 2.43 *A set* $\mathrm{End}(\boldsymbol{V})$, *along with addition and composition of endomorphisms as well as a multiplication of an endomorphism by a scalar of T, forms a linear algebra over a field T whose order is equal to* $(\dim \boldsymbol{V})^2$.

Theorem 2.44 *Let $\mathcal{B}$ be some base of a space $\boldsymbol{V}_n$. Then a mapping $H_{\mathcal{B}}$ assigning to every endomorphism its matrix with respect to the basis $\mathcal{B}$ is an isomorphism of linear algebras* $\mathrm{End}(\boldsymbol{V})$ *and* $\mathcal{M}_{n\times n}(T)$.

Let us now have a look at a special case of endomorphisms, namely of automorphisms of a given vector space (see Definition 2.4).

Notation 2.3 Let $\boldsymbol{V}$ be a vector space. Then a set of automorphisms of a vector space $\boldsymbol{V}$ (i. e. a subset of the set $\mathrm{End}(\boldsymbol{V})$) will be denoted by $\mathrm{Aut}(\boldsymbol{V})$.

Verify the validity of group axioms for a set $\mathrm{Aut}(\boldsymbol{V})$ along with composition of the mapping.

[Among others, use Theorem 2.9.]

Theorem 2.45 *Let $\boldsymbol{V}$ be a vector space. The set $\mathrm{Aut}(\boldsymbol{V})$, along with composition of automorphisms, forms a group.*

Note: A group $(\mathrm{Aut}(\boldsymbol{V}), \circ)$ is called a *linear group of a vector space $\boldsymbol{V}$*.

From Corollary 2.25, it follows:

Theorem 2.46 *Endomorphism is an automorphism of a space $\boldsymbol{V}$ if and only if its matrix in one (and thus in every) basis of a space $\boldsymbol{V}$ is regular.*

Corollary 2.47 *A group of automorphisms of a vector space $\mathbf{V}_n$ is isomorphic to a multiplicative group of regular matrices of an order n over a field T. Let $\mathcal{B}$ be some basis of a space $\mathbf{V}_n$. Then a mapping $H_{\mathcal{B}}$ assigning to every endomorphism its matrix with respect to the basis $\mathcal{B}$ is an isomorphism of groups $(\mathrm{Aut}(\boldsymbol{V}), \circ)$ and $(\mathcal{L}_{n\times n}(T), \cdot)$.*

Using Corollary 2.12 and comparing the definitions of the matrix of a homomorphism and of the transition matrix, derive the validity of the next theoreom!

Theorem 2.48 *Let f be an endomorphism and let $\mathcal{B}$ be some basis in a space $\boldsymbol{V}$. Then f is an automorphism of a vector space $\boldsymbol{V}$ if and only if the set $\mathcal{C}$, $\mathcal{C} = f(\mathcal{B})$, is the basis of a space $\boldsymbol{V}$. In such case, it holds:*

$$(\mathcal{B}, \mathcal{C}) = (f, \mathcal{B}).$$

If you imagine an intuitively understood notion of the projection of a 3-dimensional vector space onto some of its 2-dimensional subspace parallel to the chosen direction, you can see that the next definition is its natural generalisation.

Definition 2.49 Let $\boldsymbol{U}, \boldsymbol{W} \subseteq\subseteq \boldsymbol{V}$ be such that $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$. Then a mapping denoted as p_W^U and defined by

$$\forall \boldsymbol{x} \in \boldsymbol{V},\ \boldsymbol{x} = \boldsymbol{x}_W + \boldsymbol{x}_U,\ \boldsymbol{x}_W \in \boldsymbol{W}, \boldsymbol{x}_U \in \boldsymbol{U}\colon\ p_W^U(\boldsymbol{x}) = \boldsymbol{x}_W,$$

is called a *projection of a vector space* $\boldsymbol{V}$ *onto a subspace* $\boldsymbol{W}$ *parallel to a subspace* $\boldsymbol{U}$.

Remark 2.50 Since the sum $\boldsymbol{V} = \boldsymbol{U} \oplus \boldsymbol{W}$ is direct, the mapping p_W^U is defined correctly.

Theorem 2.51 *Let p_W^U be a projection of a vector space $\boldsymbol{V}$. Then for every $\boldsymbol{x}$ from $\boldsymbol{V}$ it holds:*

1. $p_W^U(\boldsymbol{x}) = \boldsymbol{o}$ *if and only if* $\boldsymbol{x} \in \boldsymbol{U}$,
2. $p_W^U(\boldsymbol{x}) = \boldsymbol{x}$ *if and only if* $\boldsymbol{x} \in \boldsymbol{W}$.

Corollary 2.52 *Every projection p_W^U of a space $\boldsymbol{V}$ is an surjection $\boldsymbol{V}$ onto $\boldsymbol{W}$.*

Corollary 2.53 *Let p_W^U be a projection of a vector space $\boldsymbol{V}$. Then it holds:*

1. $p_W^U = \mathrm{id}_V \Leftrightarrow \boldsymbol{U} = \{\boldsymbol{o}\} \Leftrightarrow \boldsymbol{W} = \boldsymbol{V}$,
2. $p_W^U = o \Leftrightarrow \boldsymbol{W} = \{\boldsymbol{o}\} \Leftrightarrow \boldsymbol{U} = \boldsymbol{V}$.

Theorem 2.54 *Every projection of a space $\boldsymbol{V}$ is an endomorphism of a space $\boldsymbol{V}$.*

Theorem 2.55 *Let p be a projection of a space $\boldsymbol{V}$. Then p is a projection $\boldsymbol{V}$ onto $\operatorname{Im} p$ parallel to $\operatorname{Ker} p$.*

Theorem 2.56 *Let p be an endomorphism of a space $\boldsymbol{V}$. Then p is a projection if and only if it holds:*

1. $\boldsymbol{V} = \operatorname{Ker} p \oplus \operatorname{Im} p$,
2. $p|\operatorname{Im} p = \mathrm{id}_{\operatorname{Im} p}$

Theorem 2.57 *Let p be an endomorphism of a space $\boldsymbol{V}$. Then p is a projection if and only if it holds:*

$$p \circ p = p.$$

Corollary 2.58 *Let p be an endomorphism of a space $\boldsymbol{V}$. Then p is a projection if and only if in an arbitrary (and thus in every) basis $\mathcal{B}$ of a space $\boldsymbol{V}$, it holds:*

$$(p, \mathcal{B})^2 = (p, \mathcal{B}).$$

Example 2.59 Find a projection p of a space $\boldsymbol{V}$ over $\mathbb{R}$ for which it holds:

$$p(\boldsymbol{u}_i) = \boldsymbol{v}_i, \;\; i = 1, 2,$$

if in the chosen basis $\mathcal{B}$ of a space $\boldsymbol{V}$, it is given:

$$\{\boldsymbol{u}_1\}_{\mathcal{B}}=(1,2,1,-1),\ \{\boldsymbol{u}_2\}_{\mathcal{B}}=(3,0,0,1),\ \{\boldsymbol{v}_1\}_{\mathcal{B}}=(1,2,0,0),\ \{\boldsymbol{v}_2\}_{\mathcal{B}}=(1,1,1,1).$$

[Instruction: consider that vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ belong to $\operatorname{Im} p$ and use Theorem 2.56.

Solution:

$$(p, \mathcal{B}) = \frac{1}{11}\begin{pmatrix} 3 & 4 & 2 & 2 \\ 4 & 9 & -1 & -1 \\ 2 & -1 & 5 & 5 \\ 2 & -1 & 5 & 5 \end{pmatrix}.]$$

Notes:

2.4 Eigenvalues and eigenspaces of endomorphisms of vector spaces

Students can define the terms eigenvalue and eigenvector of an endomorphism of a vector space. They can also find eigenvalues and eigenspaces for a given endomorphism. They know the relation between the multiplicity of an eigenvalue as a root of characteristic polynomial and the dimension of an eigenspace. Students know how to apply criteria to make sure that endomorphism is diagonalisable. The can define and find eigenvalues and eigenspaces of a square matrix.

In a number of areas of mathematics (for instance, in geometry, mathematical analysis and statistics) and their applications, it is important for a given endomorphism to know vectors determining the same direction as their images. We will call such non-zero vectors eigenvectors and in this chapter, we will learn how to look for them. You will learn that a set of vectors that are mapped through an endomorphism on its certain scalar multiple forms a subspace and that in some cases, a vector space is equal to the direct sum of such subspaces.

Definition 2.60 Let f be an endomorphism of a vector space $\boldsymbol{V}$. If it holds for a scalar $\lambda \in T$ and a non-zero vector $\boldsymbol{x} \in \boldsymbol{V}$

$$f(x) = \lambda \boldsymbol{x},$$

we say that λ is an *eigenvalue of an endomorphism* f while $\boldsymbol{x}$ is an *eigenvector of an endomorphism* f *corresponding to an eigenvalue* λ.

A set of all eigenvalues of an endomorphism f is called the *spectrum of an endomorphism* f and is denoted by $\mathrm{Spec} f$.

Note: The terms *characteristic value* and *characteristic vector* are also used.

Notation 2.4 Let f be an endomorphism in a vector space $\boldsymbol{V}$ and let λ be some of its eigenvalues. Then the symbol $\boldsymbol{N}_\lambda$ denotes the following set:

$$\boldsymbol{N}_\lambda = \{\boldsymbol{x} \in \boldsymbol{V}; f(\boldsymbol{x}) = \lambda \boldsymbol{x}\}. \tag{2.2}$$

Theorem 2.61 *Let f be an endomorphism in a vector space $\boldsymbol{V}$ and let λ be some of its eigenvalues. Then it holds:*

1. $\boldsymbol{N}_\lambda \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{N}_\lambda = \mathrm{Ker}(f - \lambda\,\mathrm{id})$,
2. *if $\mathcal{B}$ is some basis of $\boldsymbol{V}$, then $\boldsymbol{x} \in \boldsymbol{V}$ is an eigenvector of endomorphism f corresponding to λ if and only if its coordinates in the basis $\mathcal{B}$ are a non-trivial solution of a system of linear homogeneous equations with a matrix $(f, \mathcal{B})^T - \lambda\boldsymbol{E}$.*

Definition 2.62 Let f be an endomorphism in a vector space $\boldsymbol{V}$ and let λ be some of its eigenvalues. Then a set $\boldsymbol{N}_\lambda$ defined by relation (2.2) is called an *eigenspace of an endomorphism f corresponding to an eigenvalue λ*.

Corollary 2.63 *Let f be an endomorphism in a vector space $\boldsymbol{V}$, let λ be some of its eigenvalues and let $\mathcal{B}$ be an arbitrary basis.*
If we denote $(f, \mathcal{B}) = (a_{ij})_{n \times n}$, then a vector $\boldsymbol{x}$, $\{\boldsymbol{x}\}_\mathcal{B} = (x_1, \dots, x_n)$, belongs to an eigenspace $\boldsymbol{N}_\lambda$ if and only if

$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{21}x_2 + \cdots + a_{n1}x_n &= 0 \\ a_{12}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{n2}x_n &= 0 \\ &\vdots \\ a_{1n}x_1 + a_{2n}x_2 + \cdots + (a_{nn} - \lambda)x_n &= 0 \end{aligned} \tag{2.3}$$

The previous theorem gives an instruction how to find eigenvectors of an endomorphism f for a given λ. It remains to find eigenvalues of a given endomorphism. If we consider that eigenvectors are non-zero, what we are looking for is a non-trivial solution of a system of equations (2.3). From the previous part of your linear algebra course, you know that its existence is equivalent to a singularity of the matrix of the system. Hence, Theorem 2.65 gives instruction for finding eigenvalues λ.

Definition 2.64 Let f be an endomorphism in a vector space $\boldsymbol{V}$ and let $\mathcal{B}$ be an arbitrary basis of this space. Then a *characteristic polynomial of an endomorphism f* is understood to be a polynomial $ch_f(x) \in T[x]$ defined by the relation

$$ch_f(x) = \det(\,(f, \mathcal{B}) - x\boldsymbol{E}\,). \tag{2.4}$$

Make sure that the characteristic polynomial of a given endomorphism does not depend on the choice of a basis $\mathcal{B}$, namely that Definition 2.64 is correct in this sense.

[Instruction: use Theorem 2.21.]

Theorem 2.65 *Let f be an endomorphism in a vector space $\boldsymbol{V}$. Then the spectrum of endomorphism f is equal to a set of roots of its characteristic polynomial.*

If we go back also to Theorem 2.46, then we have:

Corollary 2.66 *Let f be an endomorphism in a vector space $\boldsymbol{V}$. Then the spectrum of an endomorphism f is a set of exactly those $\lambda \in T$ for which an endomorphism $f - \lambda \mathrm{id}$ is not an automorphism of a space $\boldsymbol{V}$.*

Example 2.67 In a certain basis $\mathcal{B}$ of a space $\boldsymbol{V}$ over $\mathbb{R}$, let be given an endomorphism f by a matrix $\boldsymbol{A}$ and an endomorphism g by a matrix $\boldsymbol{B}$.

$$A = \begin{pmatrix} 3 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}.$$

Find eigenvalues and eigenspaces of both endomorphisms.

[First compute the characteristic polynomial of any of given endomorphism according to (2.4). For an endomorphism f, we obtain $ch_f(x) = -(x-1)^2(x-2)$, i.e. $\mathrm{Spec} f = \{1, 2\}$. Then for every eigenvalue, assemble a system of linear equations (2.3) and solve it. Thus you obtain an eigenspace for every eigenvalue.

Solution: For an endomorphism f, we obtain: $\mathrm{Spec}\, f = \{1, 2\}$, $\boldsymbol{N}_1 = [(1,0,2), (0,1,1)]$, $\boldsymbol{N}_2 = [(1,0,1)]$. For an endomorphism g: $\mathrm{Spec}\, g = \{-1, 3\}$, $\boldsymbol{N}_{-1} = [(-2,1,0)]$, $\boldsymbol{N}_3 = [(1,-1,1)]$.]

Theorem 2.68 *Let $\lambda_1, \dots, \lambda_r$ be mutually different eigenvalues of an endomorphism f in some vector space $\boldsymbol{V}$. If we denote $\boldsymbol{N}_1, \dots, \boldsymbol{N}_r$ as corresponding eigenspaces, it holds:*

$$\boldsymbol{N}_1 + \cdots + \boldsymbol{N}_r = \boldsymbol{N}_1 \oplus \cdots \oplus \boldsymbol{N}_r.$$

Corollary 2.69

1. *Every eigenvector corresponds to one and only one eigenvalue of a given endomorphism.*
2. *Eigenvectors corresponding to different eigenvalues of the same endomorphism are linearly independent.*

Theorem 2.70 *Let f be an endomorphism in a vector space $\boldsymbol{V}$, let λ be some of its eigenvalues and let n_λ be its multiplicity as a root of the characteristic polynomial $ch_f(x)$. Then for the dimension of the eigenspace $\boldsymbol{N}_\lambda$, it holds:*

$$\dim \boldsymbol{N}_\lambda \leq n_\lambda.$$

Take notice that the dimension of an eigenspace does not actually have to be equal to mutiplicity – see endomorphism g in Example 2.67.

Definition 2.71 An endomorphism f in a vector space $\boldsymbol{V}$ is called *diagonalisable* if there exists a basis $\mathcal{B}$ of a space $\boldsymbol{V}$ such that the matrix $(f, \mathcal{B})$ is diagonal.

A question arises which endomorphisms are diagonalisable. The following three theorems bring certain criteria which we can used here. From the second of these theorems and from Example 2.67, it follows that there are endomorphisms *which are not diagonalisable* (an endomorphism f is diagonalisable, while an endomorphism g is not).

It thus generally does not hold that for every endomorphism there exists a basis such that the matrix of an endomorphism over this basis is diagonal.

Theorem 2.72 *An endomorphism f in a vector space $\boldsymbol{V}$ is diagonalisable if and only if there exists a basis $\mathcal{B}$ of a space $\boldsymbol{V}$ formed by eigenvectors of an endomorphism f.*
A diagonal of a matrix $(f, \mathcal{B})$ is in this case formed by eigenvalues of an endomorphism f; each of them is situated on a diagonal as many times as corresponding to the multiplicity of root of the characteristic polynomial $ch_f(x)$.

Theorem 2.73 *An endomorphism f in a vector space $\boldsymbol{V}$ is diagonalisable if and only if a vector space $\boldsymbol{V}$ is equal to the sum of all eigenspaces of endomorphism f.*

Theorem 2.74 *If an endomorphism f is diagonalisable, then for each of its eigenspace $\boldsymbol{N}_\lambda$, it holds that the dimension of $\boldsymbol{N}_\lambda$ is equal to the multiplicity λ as a root of the characteristic polynomial.*
For endomorphisms in vector spaces over $\mathbb{C}$, the converse theorem also holds.

Remark 2.75 Analogously[12], we can arrive at the notions of *eigenvector, eigenvalue* and *eigenspace of a matrix* $\boldsymbol{A} \in \mathcal{M}_{n\times n}(T)$:

- If for a scalar $\lambda \in T$ and a non-zero vector $\boldsymbol{x} \in T^n$, it holds

  $$\boldsymbol{x}.\boldsymbol{A} = \lambda\boldsymbol{x},$$

  we say that λ is an *eigenvalue of a matrix* $\boldsymbol{A}$ and $\boldsymbol{x}$ is an *eigenvector of a matrix* $\boldsymbol{A}$ *corresponding to an eigenvalue* λ.

- A subspace $\boldsymbol{N}_\lambda \subseteq\subseteq T^n$,

  $$\boldsymbol{N}_\lambda = \{\boldsymbol{x} \in T^n; \boldsymbol{x}.\boldsymbol{A} = \lambda\boldsymbol{x}\},$$

  is called an *eigenspace of a matrix* $\boldsymbol{A}$*corresponding to an eigenvalue* λ.

[12]It is enough to consider an arithmetic vector space $\boldsymbol{V} = T^n$ and to define for the chosen matrix $\boldsymbol{A}$ an endomorphism f by the rule $f(\boldsymbol{x}) = \boldsymbol{x}.\boldsymbol{A}$ which enables us to transfer the properties of eigenvectors/eigenvalues/eigenspaces of an endomorphism onto these notions used for matrices.

Notes:

2.5 Homomorphisms of Euclidean vector spaces

Students are able to define the notions of orthogonal projection and orthogonal homomorphism (isometries). They can determine an orthogonality of a given projection or homomorphism. They are able to apply the properties of orthogonal projections or homomorphisms when constructing these mappings. They know the relation between a group of orthogonal automorphisms and the corresponding multiplicative matrix group. They know necessary and sufficient conditions for a given mapping to be an orthogonal homomorphism, and are able to distinguish isomorphic Euclidean vector spaces.

In Chapter 1.2, you learned that an Euclidean vector space is equal to the direct sum of its arbitrary subspace and its orthogonal complement. If you use the information in Chapter 2.3, you can examine projections whose kernel and image are mutually orthogonal complements. These projections, called *orthogonal projections*, have a number of applications in geometry and other areas of mathematics.

In Chapter 2.1, you became familiar with the isomorphism of vector spaces and could see that in this case, the second of a pair of isomorphic vector spaces is only a "copy" of the first one because isomorphism determines the elements as well as both operations of the second vector space. In the case of Euclidean vector spaces, we will introduce the notion of the *orthogonal isomorphism (isometry* or *Euclidean isomorphism)* which also preserves scalar product apart from addition of vectors and multiplication of vectors by a scalar. You will see that in the case when two Euclidean vector spaces are orthogonally isomorphic, the latter of them is again a "copy" of the former one.

2.5.1 Orthogonal projection

Let us go back to Definition 2.49 and Theorem 1.45. Then it is evident that it holds:

Theorem 2.76 *Let $\boldsymbol{W}$ be a subspace of an Euclidean vector space $\boldsymbol{V}$. Then a projection $p_W^{W^\perp}$ assigns to every vector $\boldsymbol{x}$ of $\boldsymbol{V}$ its orthogonal projection onto a subspace $\boldsymbol{W}$.*

Definition 2.77 Let $\boldsymbol{W}$ be a subspace of an Euclidean vector space $\boldsymbol{V}$. Then a projection $p_W^{W^\perp}$ is called an *orthogonal projection of a space $\boldsymbol{V}$ onto a subspace $\boldsymbol{W}$* and is denoted by p_W.

A mapping assigning to each vector of $\boldsymbol{V}$ its orthogonal projection onto a subspace $\boldsymbol{W}$ – an *orthogonal projection* – is thus a special case of projection (and thus an endomorphism) – it is a projection $\boldsymbol{V}$ onto $\boldsymbol{W}$ parallel to $\boldsymbol{W}^{\perp}$.

Theorem 2.78 *Let p be an arbitrary projection of an Euclidean vector space $\boldsymbol{V}$ onto some of its subspaces. Then p is an orthogonal projection if and only if*

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}\colon p(\boldsymbol{x}) \cdot \boldsymbol{y} = \boldsymbol{x} \cdot p(\boldsymbol{y}).$$

Lemma 2.79 *Let p be an endomorphism of an Euclidean vector space $\boldsymbol{V}$ and let $\mathcal{B}$ be an arbitrary orthonormal basis. Then it holds:*

$$[\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}\colon p(\boldsymbol{x}) \cdot \boldsymbol{y} = \boldsymbol{x} \cdot p(\boldsymbol{y})] \Leftrightarrow [(p, \mathcal{B})^T = (p, \mathcal{B})].$$

Theorem 2.80 *Let p be an arbitrary projection of an Euclidean vector space $\boldsymbol{V}$ onto some of its subspaces. Then p is an orthogonal projection if and only if in some (and thus in every) orthonormal basis $\mathcal{B}$ of a space $\boldsymbol{V}$, it holds:*

$$(p, \mathcal{B})^T = (p, \mathcal{B}).$$

Example 2.81 Find an orthogonal projection p of a space $\boldsymbol{V}$ onto a subspace $\boldsymbol{W} = [\boldsymbol{v}_1, \boldsymbol{v}_2]$ if in the chosen orthonormal basis $\mathcal{B}$ of a space $\boldsymbol{V}$, it is given:

$$\{\boldsymbol{v}_1\} = (1, 2, 0, 0), \ \{\boldsymbol{v}_2\} = (1, 1, 1, 1).$$

[Instruction: consider what is the kernel and image of the searched projection and where the vectors belonging to its kernel and image are mapped. It is also possible to solve this example differently using, for instance, Corollary 2.58 and Theorem 2.80 – verify it!
Solution:

$$(p, \mathcal{B}) = \frac{1}{11} \begin{pmatrix} 3 & 4 & 2 & 2 \\ 4 & 9 & -1 & -1 \\ 2 & -1 & 5 & 5 \\ 2 & -1 & 5 & 5 \end{pmatrix}.]$$

2.5.2 Orthogonal homomorphisms

Definition 2.82 Let $(\boldsymbol{V}, \cdot)$ and $(\boldsymbol{W}, \odot)$ be Euclidean vector spaces. A homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$ is called *orthogonal* (or *isometric* or *Euclidean*) if it holds:

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}\colon \boldsymbol{x} \cdot \boldsymbol{y} = f(\boldsymbol{x}) \odot f(\boldsymbol{y}).$$

Remark 2.83 If there is no danger of misunderstanding, we denote a scalar product in different Euclidean vector spaces by the same symbol "·", or we leave it without any notation altogether.

Consider Definition 1.8 of the vector norm, Definition 1.11 of the angle between vectors, and Definition 1.15 of the distance of vectors. Then you can easily derive the following corollary from Definition 2.82. You will later see that subsections (1) and (3) are not only necessary but also sufficient condition of the orthogonality of homomorphism.
Will it be the same in the case of subsection (2)? [No; explain why not!]
Theorem 2.85 then follows from subsection 2.84 (3).

Corollary 2.84 *Let $f\colon \boldsymbol{V} \to \boldsymbol{W}$ be an orthogonal homomorphism. Then for every $\boldsymbol{x}, \boldsymbol{y}$ of $\boldsymbol{V}$, it holds:*

1. $\|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|$,
2. $\measuredangle(f(\boldsymbol{x}), f(\boldsymbol{y})) = \measuredangle(\boldsymbol{x}, \boldsymbol{y})$,
3. $\rho(f(\boldsymbol{x}), f(\boldsymbol{y})) = \rho(\boldsymbol{x}, \boldsymbol{y})$.

Theorem 2.85 *Every orthogonal homomorphism is a monomorphism.*

Hence from this theorem and from Theorem 2.24, it follows:

Theorem 2.86 *If* $\dim \boldsymbol{V} = \dim \boldsymbol{W}$*, then every orthogonal homomorphism* $\boldsymbol{V} \to \boldsymbol{W}$ *is an isomorphism* $\boldsymbol{V}$ *onto* $\boldsymbol{W}$*. Especially, every orthogonal endomorphism of a vector space is an automorphism of this space.*

Remark 2.87 An example of an orthogonal isomorphism is the Cartesian coordinate system of an Euclidean vector space (it is an orthogonal isomorphism $\boldsymbol{V}_n$ onto $\mathbb{R}^n$ endowed with the standard scalar product).

You will also easily find that every orthogonal isomorphism $\boldsymbol{V}_n$ onto $\mathbb{R}^n$ endowed with the standard scalar product is a Cartesian coordinate system.

Let us go back to Corollary 2.12 (4). You should also realise that due to Corollary 2.84, an orthogonal homomorphism maps an orthonormal set of vectors once again onto an orthonormal set. Finally, if some homomorphism maps an orthonormal basis once again onto an orthonormal basis, then with respect to Theorem 1.27, you can easily find by a direct calculation that the next theorem is indeed an equivalence.

Theorem 2.88 *Let f be a homomorphism $\boldsymbol{V}$ to $\boldsymbol{W}$ and let $\mathcal{B}$ be an arbitrary orthonormal basis of a space $\boldsymbol{V}$. Then f is an orthogonal isomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$ if and only if $f(\mathcal{B})$ is the orthonormal basis of a space $\boldsymbol{W}$.*

Corollary 2.89 *Let $\boldsymbol{U}$, $\boldsymbol{V}$, $\boldsymbol{W}$ be Euclidean vector spaces. Then it holds:*

1. *if f is an orthogonal isomorphism $\boldsymbol{U}$ onto $\boldsymbol{V}$, then f^{-1} is an orthogonal isomorphism $\boldsymbol{V}$ onto $\boldsymbol{U}$,*
2. *if f is an orthogonal homomorphism $\boldsymbol{U}$ to $\boldsymbol{V}$ and g is an orthogonal homomorphism $\boldsymbol{V}$ to $\boldsymbol{W}$, then $f \circ g$ is an orthogonal homomorphism $\boldsymbol{U}$ to $\boldsymbol{W}$.*

It is natural to ask how easy it is to find whether a given homomorphism is or is not orthogonal. If you consider again Theorem 2.88, the definition of the matrix of homomorphism 2.17, definition of matrix multiplication as well as the Cartesian formula for the scalar product, you can easily verify the validity of the following criterion.

Theorem 2.90 *Let f be a homomorphism $\boldsymbol{V}$ to $\boldsymbol{W}$, and let $\mathcal{B}$, $\mathcal{C}$ be arbitrary orthonormal bases, respectively, of spaces $\boldsymbol{V}$, $\boldsymbol{W}$. Then f is an orthogonal homomorphism $\boldsymbol{V}$ to $\boldsymbol{W}$ if and only if*

$$(f, \mathcal{B}, \mathcal{C})(f, \mathcal{B}, \mathcal{C})^T = E.$$

Remember that the assumption of *orthonormality* cannot be left out in any of the bases!

From Theorem 2.90 and Corollary 2.89, it follows:

Corollary 2.91

1. *A set of orthogonal automorphisms of an Euclidean vector space* $\boldsymbol{V}$*, along with composition of homomorphisms, forms a group which is a subgroup of a group of automorphisms of a vector space* $\boldsymbol{V}$*.*
2. *A group of orthogonal automorphisms*[13] *is isomorphic to a multiplicative group of orthogonal matrices of order* n*.*
   If $\mathcal{B}$ *is an orthonormal basis of a space* $\boldsymbol{V}_n$*, then a mapping* $H_{\mathcal{B}}$ *assigning to every endomorphism its matrix with respect to the basis* $\mathcal{B}$ *is an isomorphism of the above groups.*

The following four theorems provides other necessary and sufficient conditions for a homomorphism or a mapping to be an orthogonal homo-, or an isomorphism. Compare their contents with the conditions of orthogonality of homomorphism in Definition 2.82 and with Corollary 2.84.

Using the identity $(\boldsymbol{x}+\boldsymbol{y})(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{x}\boldsymbol{x} + 2\boldsymbol{x}\boldsymbol{y} + \boldsymbol{y}\boldsymbol{y}$, verify the validity of the first of the following theorems!

[Instruction: from this identity, derive the relation for the norm of the sum of vectors and the norm of the sum of their images.]

Theorem 2.92 *Let* f *be a homomorphism* $\boldsymbol{V}$ *to* $\boldsymbol{W}$*. Then* f *is an orthogonal homomorphism* $\boldsymbol{V}$ *to* $\boldsymbol{W}$ *if and only if for every* $\boldsymbol{x}$ *of* $\boldsymbol{V}$*, it holds:*

$$\|f(\boldsymbol{x})\| = \|\boldsymbol{x}\|.$$

Theorem 2.93 *Let* f *be a homomorphism* $\boldsymbol{V}$ *to* $\boldsymbol{W}$*. Then* f *is an orthogonal homomorphism* $\boldsymbol{V}$ *to* $\boldsymbol{W}$ *if and only if for every* $\boldsymbol{x}$*,* $\boldsymbol{y}$ *of* $\boldsymbol{V}$*, it holds:*

$$\rho(f(\boldsymbol{x}), f(\boldsymbol{y})) = \rho(\boldsymbol{x}, \boldsymbol{y}).$$

Theorem 2.94 *Let* f *be a bijection* $\boldsymbol{V}$ *onto* $\boldsymbol{W}$*. Then* f *is an orthogonal isomorphism* $\boldsymbol{V}$ *onto* $\boldsymbol{W}$ *if and only if for every* $\boldsymbol{x}$*,* $\boldsymbol{y}$ *of* $\boldsymbol{V}$*, it holds:*

$$f(\boldsymbol{x})f(\boldsymbol{y}) = \boldsymbol{x}\boldsymbol{y}.$$

[13] A group of orthogonal automorphisms is called an *orthogonal* (or *isometric) group of a given vector space*; it is thus a subgroup of the so-called linear group of a space $\boldsymbol{V}_n$ – cf. Theorem 2.45.

Theorem 2.95 *Let f be a bijection $\boldsymbol{V}$ onto $\boldsymbol{W}$. Then f is an orthogonal isomomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$ if and only if it holds:*

1. $f(\boldsymbol{o}) = \boldsymbol{o}$,
2. $\forall \boldsymbol{x},\ \boldsymbol{y} \in \boldsymbol{V}$: $\rho(f(\boldsymbol{x}), f(\boldsymbol{y})) = \rho(\boldsymbol{x}, \boldsymbol{y})$.

Example 2.96 Find all orthogonal homomorphisms $f\colon \boldsymbol{V} \to \boldsymbol{W}$ for which

$$\boldsymbol{u} \mapsto \boldsymbol{v},$$

if in the chosen orthonormal bases $\mathcal{B}$, $\mathcal{C}$, respectively, of spaces $\boldsymbol{V}$, $\boldsymbol{W}$, it is given:

$$\{\boldsymbol{u}\}_{\mathcal{B}} = (-\sqrt{2}, \sqrt{2}),\ \ \{\boldsymbol{v}\}_{\mathcal{C}} = (0, -2).$$

[Instruction: consider the validity of Theorem 2.90.
Solution: there exist exactly two orthogonal homomorphisms given by the following matrices

$$\frac{1}{2}\begin{pmatrix} \sqrt{2} & \sqrt{2} \\ \sqrt{2} & -\sqrt{2} \end{pmatrix}, \text{ resp. } \ \frac{1}{2}\begin{pmatrix} -\sqrt{2} & \sqrt{2} \\ -\sqrt{2} & -\sqrt{2} \end{pmatrix}]$$

Definition 2.97 Let $\boldsymbol{V}, \boldsymbol{W}$ be Euclidean vector spaces. Let us say that the above *Euclidean vector spaces are isomorphic* if there exists an orthogonal isomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$.

The attribute *Euclidean* in the phrase *Euclidean vector spaces are isomorphic* needs to be emphasised. Precisely to distinguish this homomorphism from a common homomorphism of vector spaces, the term *orthogonally* (or *Euclidean) isomorphic vector spaces* or *isometric vector spaces* is sometimes also being used.

Remark 2.98 Drawing on Corollary 2.89, we can see that the relation *to be orthogonally isomorphic* is indeed symmetric. The newly introduced definition is thus correct in this sense.

You already know that if vector spaces $(\boldsymbol{V}, +, T, \cdot)$ and $(\boldsymbol{W}, \oplus, T, \circ)$ are isomorphic, it does not only mean that there is a bijection between sets $\boldsymbol{V}$ and $\boldsymbol{W}$ of their vectors: in Section 2.1, we showed that also addition of vectors $\oplus$ and multiplication of vectors by the scalar $\circ$ in the second of the spaces are fully determined by the corresponding operations in the first space. This is what you were talking about when you were saying that apart from naming vectors, there is no need to distinguish between these vector spaces.

Let us now consider two isomorphic *Euclidean* vector spaces

$$((\boldsymbol{V}, +, \mathbb{R}, \cdot), \bullet) \quad \text{and} \quad ((\boldsymbol{W}, \oplus, \mathbb{R}, \circ), \odot)$$

and let f is a corresponding orthogonal isomorphism $\boldsymbol{V}$ onto $\boldsymbol{W}$. Let us now look at scalar product of two vectors belonging to $\boldsymbol{W}$. Since drawing on Corollary 2.89 and Definition 2.82, we can write

$$\boldsymbol{u} \odot \boldsymbol{w} = f(f^{-1}(\boldsymbol{u})) \odot f(f^{-1}(\boldsymbol{w})) = f^{-1}(\boldsymbol{u}) \bullet f^{-1}(\boldsymbol{w}),$$

you can see that this scalar product is fully determined by the scalar product $\bullet$ in $\boldsymbol{V}$. It is then possible to say that (also) two isomorphic Euclidean vector spaces are their mutual "exact copies". The phrase *two isomorphic Euclidean vector spaces differ from one another only in the names of their elements* is thus being justifiably used.

From Definition 2.97, it is clear that if two Euclidean vector spaces are orthogonally isomorphic, they are isomorphic as vector spaces. Hence, according to Theorem 2.16, they have the same dimension. Justify the converse statement! [Instruction: look at Remark 2.87.] You will thus prove the next theorem (compare it with Theorem 2.16 !).

Theorem 2.99 *Two Euclidean vector spaces are isomorphic if and only if they have the same dimension.*

Notes:

Notes:

3 Factor vector spaces

Students know the notion of factor vector space. They can tell whether two vectors are congruent modulo a subspace and are able to construct a factorization to a given subspace and to find its basis and determine its dimension. They are able to apply the theorem on homomorphism to determine all homomorphic images of the chosen vector space and to use the theory of factor vector spaces to construct an affine space and its subspaces.

In your algebra course, you became familiar with the relation of equivalence on a set and learned how to construct a decomposition (factorisation) of this set to the chosen equivalence. Using this knowledge, you will now learn how to assign to any chosen subspace of a vector space the relation of equivalence *(congruence)* so that it will be possible to construct a vector space structure in the set of the formed classes *(linear manifolds)*.

You will see that the theorem on homomorphism of sets can be extended to vector spaces. You will learn that the affine space, known to you from geometry, can be obtained also through factorization of a vector space according to its subspaces.

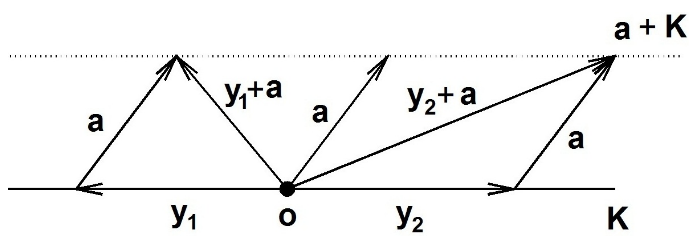
The symbol $\boldsymbol{V}$ will continue to denote an arbitrary n-dimensional vector space over a field T.

Definition 3.1 Let $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{a} \in \boldsymbol{V}$. Then a set denoted by $\boldsymbol{a} + \boldsymbol{K}$ defined by the relation

$$\boldsymbol{a} + \boldsymbol{K} = \{\boldsymbol{x} \in \boldsymbol{V}\,;\ \exists \boldsymbol{y} \in \boldsymbol{K} \colon \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{y}\}$$

is called a *linear manifold of a space $\boldsymbol{V}$ determined by a vector $\boldsymbol{a}$ parallel to $\boldsymbol{K}$*.

Remark 3.2 Let us consider a 2-dimensional vector space $\boldsymbol{V}$, its one-dimensional subspace (direction) $\boldsymbol{K}$ and an arbitrary vector $\boldsymbol{a} \in \boldsymbol{V}$ (see the figure below). Then a linear manifold $\boldsymbol{a} + \boldsymbol{K}$ is equal to a set of vectors whose end points lie on the dotted line.



Remark 3.3

1. Directly from Definition 3.1, it follows that for every $\boldsymbol{x}$ of $\boldsymbol{V}$, it holds:

$$\boldsymbol{x} \in \boldsymbol{a} + \boldsymbol{K} \Leftrightarrow \boldsymbol{x} - \boldsymbol{a} \in \boldsymbol{K}.$$

2. A manifold $\boldsymbol{a}+\boldsymbol{K}$ should not be generally identified with a subspace $[\boldsymbol{a}]+\boldsymbol{K}$. (Consider this! What would a subspace $[\boldsymbol{a}] + \boldsymbol{K}$ be equal to in the case described in Remark 3.2?)

We will now define a certain relation for every subspace $\boldsymbol{K}$ on a set $\boldsymbol{V}$. We will then show that it is exactly the equivalence which decomposes a set $\boldsymbol{V}$ into a set of exactly all linear manifolds parallel to $\boldsymbol{K}$.

Definition 3.4 Let $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{V}$. We say that a *vector $\boldsymbol{a}$ is congruent to a vector $\boldsymbol{b}$ modulo $\boldsymbol{K}$* which we denote as $\boldsymbol{a} \equiv \boldsymbol{b} \,(\mathrm{mod}\, \boldsymbol{K})$ if a vector $\boldsymbol{b} - \boldsymbol{a}$ belongs to $\boldsymbol{K}$.

It needs to be specified with respect to which subspace $\boldsymbol{K}$ the considered vectors are congruent – if we simultaneously work with several different subspaces in $\boldsymbol{V}$, we always must consistently add "mod $\boldsymbol{K}$" to the symbol $\boldsymbol{a} \equiv \boldsymbol{b}$. Only if it is evident which subspace we have in mind, we can write only $\boldsymbol{a} \equiv \boldsymbol{b}$.

Derive the validity of the following lemma.

Lemma 3.5 *Let $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$. Then it holds:*

1. $\forall \boldsymbol{a} \in \boldsymbol{V}$: $\boldsymbol{a} \equiv \boldsymbol{a} \,(\mathrm{mod}\, \boldsymbol{K})$,
2. $\forall \boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{V}$: $\boldsymbol{a} \equiv \boldsymbol{b} \,(\mathrm{mod}\, \boldsymbol{K}) \Rightarrow \boldsymbol{b} \equiv \boldsymbol{a} \,(\mathrm{mod}\, \boldsymbol{K})$,
3. $\forall \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \boldsymbol{V}$: $(\boldsymbol{a} \equiv \boldsymbol{b} \,(\mathrm{mod}\, \boldsymbol{K}) \wedge \boldsymbol{b} \equiv \boldsymbol{c} \,(\mathrm{mod}\, \boldsymbol{K})) \Rightarrow \boldsymbol{a} \equiv \boldsymbol{c} \,(\mathrm{mod}\, \boldsymbol{K})$.

Thus:

Theorem 3.6 *The relation "to be congruent modulo $\boldsymbol{K}$" is for every subspace $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$ a relation of equivalence on a set $\boldsymbol{V}$.*

Theorem 3.7 *Let $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{a}, \boldsymbol{x} \in \boldsymbol{V}$. Then it holds:*

$$\boldsymbol{x} \in \boldsymbol{a} + \boldsymbol{K} \Leftrightarrow \boldsymbol{x} \equiv \boldsymbol{a} \,(\mathrm{mod}\, \boldsymbol{K}).$$

Let us take another look at how to obtain a quotient set of the relation of equivalence[14]. In this way, you obtain the already mentioned important statement:

Corollary 3.8 *Let* $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$, $\boldsymbol{a} \in \boldsymbol{V}$. *Then a linear manifold* $\boldsymbol{a}+\boldsymbol{K}$ *is an equivalence class of a set* $\boldsymbol{V}$ *of the relation "to be congruent modulo* $\boldsymbol{K}$*" determined by an element* $\boldsymbol{a}$.

Based on this corollary, Definition 3.4 and the already familiar properties of quotient sets (thus, in our case, a quotient of a set $\boldsymbol{V}$ by the $\equiv (\mathrm{mod}\, \boldsymbol{K})$), you obtain the following statement:

Theorem 3.9 *Let* $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$. *Then for every* $\boldsymbol{a}, \boldsymbol{b}$ *of* $\boldsymbol{V}$ *it holds:*

1. $\forall \boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{V}$: $(\boldsymbol{a} + \boldsymbol{K} = \boldsymbol{b} + \boldsymbol{K}) \Leftrightarrow (\boldsymbol{b} - \boldsymbol{a}) \in \boldsymbol{K}$,
2. $\forall \boldsymbol{a} \in \boldsymbol{V}$: $(\boldsymbol{a} + \boldsymbol{K} = \boldsymbol{K}) \Leftrightarrow \boldsymbol{a} \in \boldsymbol{K}$,
3. $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a} \in \boldsymbol{V}$: $(\boldsymbol{x} \in \boldsymbol{a} + \boldsymbol{K} \wedge \boldsymbol{y} \in \boldsymbol{a} + \boldsymbol{K}) \Leftrightarrow (\boldsymbol{y} - \boldsymbol{x}) \in \boldsymbol{K}$.

Choose $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$. Then you can construct a *quotient set* of $\boldsymbol{V}$ by the equivalence relation $\equiv (\mathrm{mod}\, \boldsymbol{K})$ – a set of exactly all manifolds of a space $\boldsymbol{V}$ parallel to $\boldsymbol{K}$. Let us denote this so-called *factor set* by the symbol $\boldsymbol{V}/\boldsymbol{K}$. It then holds:

$$\boldsymbol{V}/\boldsymbol{K} = \{\{\boldsymbol{x} \in \boldsymbol{V};\ \exists \boldsymbol{y} \in \boldsymbol{K} \colon \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{y}\}, \boldsymbol{a} \in \boldsymbol{V}\} = \{\{\boldsymbol{a} + \boldsymbol{K}\}, \boldsymbol{a} \in \boldsymbol{V}\}.$$

On this set, we want to construct a *vector space structure over* T – that is, we want to define an addition of manifolds and a multiplication of the manifold by a scalar:

(i) Let us choose $\boldsymbol{a} + \boldsymbol{K},\ \boldsymbol{b} + \boldsymbol{K} \in \boldsymbol{V}/\boldsymbol{K}$ and let us put

$$(\boldsymbol{a} + \boldsymbol{K}) + (\boldsymbol{b} + \boldsymbol{K}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{K}. \tag{3.1}$$

(ii) Let us choose $\boldsymbol{a} + \boldsymbol{K} \in \boldsymbol{V}/\boldsymbol{K}$, $t \in T$ and let us put[15]

$$t \cdot (\boldsymbol{a} + \boldsymbol{K}) = (t \cdot \boldsymbol{a}) + \boldsymbol{K}. \tag{3.2}$$

[14] An equivalence class defined by, for instance, an element $\boldsymbol{a}$ is a set of exactly all elements which are equivalent to it.

[15] We usually leave out the symbol "$\cdot$" and write only $t(\boldsymbol{a} + \boldsymbol{K})$.

Using relations (3.1) and (3.2), demonstrate that the definitions of operations $+, \cdot$ are correct; that is, that the result of both operations does not depend on the choice of vectors determining given manifolds (the so-called *manifold representative)* – a manifold $\boldsymbol{c}+\boldsymbol{K}$ can be after all determined also by a different vector $\bar{\boldsymbol{c}}$.

[Instruction: between two vectors $\boldsymbol{c}$, $\bar{\boldsymbol{c}}$ determining the same manifold, there must be a relation following from Theorem 3.9 (1). Then it is enough to compare the manifolds that are understood to be the result of the sum of manifolds or of the t-multiple of the manifold.]

Definition 3.10 Let $\boldsymbol{K}\subseteq\subseteq\boldsymbol{V}$ and let be given $\boldsymbol{a}+\boldsymbol{K}$, $\boldsymbol{b}+\boldsymbol{K} \in \boldsymbol{V}/\boldsymbol{K}$ and $t \in T$. Then

1. the *sum of linear manifolds* $\boldsymbol{a}+\boldsymbol{K}$ *and* $\boldsymbol{b}+\boldsymbol{K}$ is understood to be a linear manifold denoted by $(\boldsymbol{a}+\boldsymbol{K})+(\boldsymbol{b}+\boldsymbol{K})$ and defined by the relation (3.1),
2. the *(scalar) t-multiple of a linear manifold* $\boldsymbol{a}+\boldsymbol{K}$ is understood to be a linear manifold denoted by $t\cdot(\boldsymbol{a}+\boldsymbol{K})$ and defined by the relation (3.2).

Verify if $(\boldsymbol{V}/\boldsymbol{K}, +, T, \cdot)$ meets the axioms of a vector space where a zero manifold is a manifold $\boldsymbol{o}+\boldsymbol{K}$ while the opposite manifold to a linear manifold $\boldsymbol{a}+\boldsymbol{K}$ is a manifold $((-\boldsymbol{a})+\boldsymbol{K})$. Derive for which vectors $\boldsymbol{b}$ a manifold $\boldsymbol{b}+\boldsymbol{K}$ is a zero manifold[16].

Theorem 3.11 *Let $\boldsymbol{K}\subseteq\subseteq\boldsymbol{V}$. Then a set $\boldsymbol{V}/\boldsymbol{K}$, along with addition of linear manifolds and multiplication of a linear manifold by a scalar from T, forms a vector space over a field T.*

Definition 3.12 Let be given $\boldsymbol{K}\subseteq\subseteq\boldsymbol{V}$. Then a vector space $(\boldsymbol{V}/\boldsymbol{K}, +, T, \cdot)$ is called a *factor vector space of a vector space $\boldsymbol{V}$ to a subspace $\boldsymbol{K}$* (or a *factorisation of a vector space $\boldsymbol{V}$ to a subspace $\boldsymbol{K}$* or a *quotient of the vector space $\boldsymbol{V}$ by a subspace $\boldsymbol{K}$*).

Now let us take a closer look at the relation between a vector space $\boldsymbol{V}$ and its factorisation $\boldsymbol{V}/\boldsymbol{K}$.

[16][Solution: $\boldsymbol{b}+\boldsymbol{K}$ is a zero manifold if and only if $\boldsymbol{b} \in \boldsymbol{K}$.]

From the theory of quotient sets, you know that a mapping ν assigning to an element of a given set an equivalence class determined by this element is a surjection of a given set onto its quotient (so-called *natural* or *canonical mapping*). Furthermore, if you take into consideration Definition 3.10, you can see that a natural mapping is also a homomorphism $\boldsymbol{V}$ onto $\boldsymbol{V}/\boldsymbol{K}$ – we are thus ready to formulate Definition 3.13.

Let us subsequently take another look at the theorem on homomorphism of sets according to which: (1) every surjection f of the chosen set V onto a certain set W induces on a set V an equivalence relation $\approx$ making the elements with the same image in a mapping f equivalent; (2) a surjection f can be decomposed in exactly one way into a natural mapping ν of a set V onto its quotient set $V/\approx$ and a bijection g of this quotient set onto a set W.

Now let us choose two vector spaces $\boldsymbol{V}$, $\boldsymbol{W}$ and an epimorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$, and let us apply the theorem on homomorphism of sets. Since f is a homomorphism, we can easily find that:

$$\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}\colon (\boldsymbol{u} \approx \boldsymbol{v}) \Leftrightarrow (\boldsymbol{u} - \boldsymbol{v} \in \operatorname{Ker} f) \Leftrightarrow (\boldsymbol{u} \equiv \boldsymbol{v} (\operatorname{mod} \operatorname{Ker} f)),$$

i.e. this induced equivalence on $\boldsymbol{V}$ is exactly the relation $\equiv (\operatorname{mod} \operatorname{Ker} f)$ according to Definition 3.4. Using Definition 3.10 of operations with linear manifolds, we can see that a bijection g for which $g \circ \nu_{\operatorname{Ker} f} = f$, preserves these operations and is thus a bijective homomorphism.

In summary, you have found that the following definition is correct and that Theorem 3.14 is valid.

Definition 3.13 Let be given $\boldsymbol{K} \subseteq\subseteq \boldsymbol{V}$. A mapping $\nu_K\colon \boldsymbol{V} \to \boldsymbol{V}/\boldsymbol{K}$ defined by the relation

$$\forall \boldsymbol{a} \in \boldsymbol{V}\colon \nu_K(\boldsymbol{a}) = \boldsymbol{a} + \boldsymbol{K}$$

is called a *natural* (or *canonical) homomorphism corresponding to a factorisation* $\boldsymbol{V}/\boldsymbol{K}$.

Theorem 3.14 (on homomorphism of vector spaces)
Let $\boldsymbol{V}$, $\boldsymbol{W}$ be vector spaces. Then for every epimorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$, there exists one and only one isomorphism $g\colon \boldsymbol{V}/\operatorname{Ker} f \to \boldsymbol{W}$ such that

$$g \circ \nu_{\operatorname{Ker} f} = f,$$

or the following diagram commutes:

$$\begin{tikzcd}
\mathbf{V} \arrow[r, "\nu_{\operatorname{Ker} f}"] \arrow[d, "f"'] & \mathbf{V}/\operatorname{Ker} f \arrow[dl, dashed, "g"] \\
\mathbf{W} &
\end{tikzcd}$$

If you replace in the previous theorem the word *epimorphism* with the word *homomorphism* and the word *isomorphism* with the word *monomorphism*, you again obtain a valid, although more general statement. Explain why it is so!

[Instruction: see Corollary 2.7]

The above theorem makes it possible to describe all – apart from isomorphisms – homomorphic images of a given vector space:

Corollary 3.15 *Apart from isomorphism, a set of all homomorphic images of a given vector space is equal to the set of all its factorisations according to its respective subspaces.*

We will now focus on determining the dimension and finding the basis of a factorisation $\boldsymbol{V}/\boldsymbol{K}$. For determining the dimension, it is essential that

$$\operatorname{Im} \nu_K = \boldsymbol{V}/\boldsymbol{K}.$$

Then using Theorem 2.23, you obtain Theorem 3.16.

To find some bases of a factorisation $\boldsymbol{V}/\boldsymbol{K}$, we will proceed from the evident fact that a linear manifold determined by a linear combination of vectors is a linear combination of manifolds defined by these vectors with the same coefficients [why?]. Hence, you can see that linear manifolds determined by vectors of an arbitrary basis of a space $\boldsymbol{V}$ form a set of generators of a factorisation $\boldsymbol{V}/\boldsymbol{K}$. Realize which manifolds are determined by vectors belonging to a subspace $\boldsymbol{K}$ and construct a basis $\mathcal{B}$ of a space $\boldsymbol{V}$ by complementing an arbitrary basis of a subspace $\boldsymbol{K}$. The number of non-zero linear manifolds determined by elements of such a basis $\mathcal{B}$ is then according to Theorem 3.16 equal to the dimension of a factorisation $\boldsymbol{V}/\boldsymbol{K}$. The validity of Theorem 3.17 has thus been confirmed.

Theorem 3.16 *Let be given $\boldsymbol{K} \subseteq\subsetneq \boldsymbol{V}$. Then it holds:*

$$\dim \boldsymbol{V}/\boldsymbol{K} = \dim \boldsymbol{V} - \dim \boldsymbol{K}.$$

Theorem 3.17 *Let be given $\boldsymbol{K} \subseteq\subsetneq \boldsymbol{V}$. Let $\langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-k} \rangle$ be an arbitrary system of vectors complementing some basis of a subspace $\boldsymbol{K}$ to a basis of a space $\boldsymbol{V}$. Then*

$$\langle \boldsymbol{e}_1 + \boldsymbol{K}, \ldots, \boldsymbol{e}_{n-k} + \boldsymbol{K} \rangle$$

forms the basis of a factor vector space $\boldsymbol{V}/\boldsymbol{K}$.

Remark 3.18 One of the applications of the theory of factor vector spaces is the construction of the *affine space* (which you know, probably in a different form, from geometry):
If a vector space $\boldsymbol{V}$ is given, we can construct an affine space $\mathcal{A}$ with an associated vector space $\boldsymbol{V}$, i.e. $\mathcal{A} = \mathcal{A}(\boldsymbol{V})$ in the following way:

- *points of an affine space* $\mathcal{A}$ will be understood to be just all one-element subsets in $\boldsymbol{V}$ (i.e. all factorisations $\boldsymbol{V}$ to a trivial subspace) – as seen, for instance, in $A = \{\boldsymbol{a}\}$, $B = \{\boldsymbol{b}\}$, where $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{V}$, are points in $\mathcal{A}$;
- *lines of an affine space* $\mathcal{A}$ will be understood to be just all factorisations $\boldsymbol{V}$ to its one-dimensional subspaces, as seen, for instance, in $p = \boldsymbol{a} + \boldsymbol{K}$, $q = \boldsymbol{b} + \boldsymbol{L}$, where $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{V}$, $\boldsymbol{K}, \boldsymbol{L} \subseteq\subseteq \boldsymbol{V}$, $\dim \boldsymbol{K} = \dim \boldsymbol{L} = 1$ are lines in $\mathcal{A}$; p is a line defined by a point $A = \{\boldsymbol{a}\}$ and a direction $\boldsymbol{K}$, q is a line defined by a point $B = \{\boldsymbol{b}\}$ and a direction $\boldsymbol{L}$;

  generally:
- *k-dimensional affine subspaces of an affine space* $\mathcal{A}$, $0 \leq k \leq n$, will be understood to be just all factorisations $\boldsymbol{V}$ to its k-dimensional subspaces, as seen, for instance, in $\mathcal{K} = \boldsymbol{a} + \boldsymbol{K}$, $\mathcal{L} = \boldsymbol{b} + \boldsymbol{L}$, where $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{V}$, $\boldsymbol{K}, \boldsymbol{L} \subseteq\subseteq \boldsymbol{V}$, $\dim \boldsymbol{K} = \dim \boldsymbol{L} = k$ are k-dimensional subspaces in $\mathcal{A}$; $\mathcal{K}$ is a subspace defined by a point $A = \{\boldsymbol{a}\}$ with a direction subspace $\boldsymbol{K}$, and $\mathcal{L}$ is a subspace defined by a point $B = \{\boldsymbol{b}\}$ with a direction subspace $\boldsymbol{L}$;
- a *relation of incidence* is a set inclusion, which means that a point $A = \{\boldsymbol{a}\}$ lies in a subspace $\mathcal{L} = \boldsymbol{b} + \boldsymbol{L}$ if and only if $\{\boldsymbol{a}\} \subseteq \boldsymbol{b} + \boldsymbol{L}$, i.e. $\boldsymbol{a} - \boldsymbol{b} \in \boldsymbol{L}$.

The figure accompanying Remark 3.2 thus presumably shows a dotted line defined by a point $A = \{\boldsymbol{a}\}$ and a direction $\boldsymbol{K}$, with a point $Y_1 = \{\boldsymbol{y}_1 + \boldsymbol{a}\}$ being one of its points because $\boldsymbol{y}_1 \in \boldsymbol{K}$.

Example 3.19 In a vector space $\boldsymbol{V}_4$, there are given a subspace $\boldsymbol{K} = [\boldsymbol{u}, \boldsymbol{v}]$ and a vector $\boldsymbol{x} \in \boldsymbol{V}$:

$$\{\boldsymbol{u}\} = (1, 2, 1, 1), \{\boldsymbol{v}\} = (2, 0, 0, 1) \textit{ and } \{\boldsymbol{x}\} = (1, -2, 2, -1).$$

Determine the dimension of a factor vector space $\boldsymbol{V}/\boldsymbol{K}$ and find at least one of its bases; then determine the coordinates of a manifold $\boldsymbol{x} + \boldsymbol{K}$ in this basis.

[Instruction: For the computation of the dimension, use Theorem 3.16. To construct a basis $\mathcal{C}$ of the factorisation, proceed according to the guide prior to Theorem 3.16. Consider that a manifold $\boldsymbol{x} + \boldsymbol{K}$ has in a basis $\mathcal{C} = \langle \boldsymbol{e}_1 + \boldsymbol{K}, \boldsymbol{e}_2 + \boldsymbol{K} \rangle$ coordinates x_1, x_2 if and only if $\boldsymbol{x} - (x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2) \in \boldsymbol{K}$.
Solution:

- $\dim \boldsymbol{V}/\boldsymbol{K} = 2$,

- $\mathcal{C}$ is, for instance, $\langle (0,0,1,0)+\boldsymbol{K}, (0,0,0,1)+\boldsymbol{K} \rangle$,
- $\{\boldsymbol{x}+\boldsymbol{K}\}_{\mathcal{C}} = (3,-1)$.]

Notes:

4 Dual vector spaces

Students can define a dual vector space and know the notion of linear form. On a set of linear forms, they are able to define a vector space structure. They know the notion of dual basis and are able to construct a basis dual to a given basis of a vector space. Finally, they can assign a linear form to a vector from a given vector space and *vice versa*.

In Chapter 2, you became familiar with a special kind of mapping between vector spaces over the same field of scalars, namely with *homomorphisms of vector spaces*. We will now look at a special case of homomorphism. Since every field can be seen as a (1-dimensional) vector space over itself, for a vector space $\boldsymbol{V}$ over a field T, we can consider homomorphisms $\mathrm{Hom}(\boldsymbol{V}, T)$ called *linear forms*. It is their properties that will be discussed in this chapter.

The symbol $\boldsymbol{V}$ denotes an arbitrary n-dimensional vector space over a field T.

Definition 4.1 Let $\boldsymbol{V}$ be a vector space over T. Then

1. a vector space $\mathrm{Hom}(\boldsymbol{V}, T)$ is called a *dual vector space of a space* $\boldsymbol{V}$ and is denoted by $\widetilde{\boldsymbol{V}}$;
2. every element f of $\widetilde{\boldsymbol{V}}$ is called a *linear form on* $\boldsymbol{V}$.

Remark 4.2

- A linear form on $\boldsymbol{V}$ is then every mapping $f\colon \boldsymbol{V} \to T$ with the following properties:

  1. $\forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}\colon f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$,
  2. $\forall \boldsymbol{u} \in \boldsymbol{V},\ \forall t \in T\colon f(t\boldsymbol{u}) = tf(\boldsymbol{u})$.

- A dual vector space $\widetilde{\boldsymbol{V}}$ is a set of exactly all linear forms on $\boldsymbol{V}$ along with addition of linear forms and multiplication of a linear form by a scalar (See Definition 2.28).
- The notion of *linear form* is also used in the theory of polynomials where it is understood to be a homogeneous first degree polynomial. As you will see, a linear form in the sense of our definition is exactly such mapping $\boldsymbol{V} \to T$ whose analytic expression is a homogeneous first degree polynomial – it is thus a linear form in the sense of the theory of polynomials where indefinites are the coordinates of a vector from $\boldsymbol{V}$.

Since a linear form is a special case of homomorphism, many of its properties are only a specialisation of the general terms and of one of the theorems on homomorphism which you learned in Chapter 2. We will present only some of these statements (find the corresponding theorems in Chapter 2!).

Theorem 4.3 *Let $\boldsymbol{V}$ be a vector space. Then it holds:*

$$\dim \widetilde{\boldsymbol{V}} = \dim \boldsymbol{V}.$$

Theorem 4.4 *For every basis $\mathcal{B} = \langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_n \rangle$ of a vector space $\boldsymbol{V}$ and for every ordered n-tuple $(a_1, \dots, a_n)$ of scalars from T, there exists one and only one linear form f on $\boldsymbol{V}$ with the property:*

$$f(\boldsymbol{u}_i) = a_i, \ \ i = 1, \dots, n. \tag{4.1}$$

Remark 4.5 As for the choice of a basis of T as a vector space, we implicitly assume that the chosen basis is $\langle 1 \rangle$.

The matrix of a linear form f with respect to a basis $\mathcal{B}$, which is clearly a column vector $(a_1, \dots, a_n)^T$ with elements given by the relation (4.1), is denoted only by $(f, \mathcal{B})$.

Theorem 4.6 *Let f be a linear form on $\boldsymbol{V}$ and let $\mathcal{B}$ be an arbitrary basis of a space $\boldsymbol{V}$. Then for every $\boldsymbol{x}$ from $\boldsymbol{V}$, it holds:*

$$f(\boldsymbol{x}) = \{\boldsymbol{x}\}_{\mathcal{B}}(f, \mathcal{B}),$$

or if $\{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, \dots, x_n)$ and $\mathcal{B} = \langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_n \rangle$, then

$$f(\boldsymbol{x}) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n,$$

where $f(\boldsymbol{u}_i) = a_i, \ \ i = 1, \dots, n.$

Theorem 4.7 *A dual vector space $\widetilde{\boldsymbol{V}_n}$ is isomorphic to an arithmetic vector space T^n.*

If $\mathcal{B}$ is a basis of a space $\boldsymbol{V}$, then a mapping $H_{\mathcal{B}}$ defined by the relation

$$\forall f \in \widetilde{\boldsymbol{V}} \colon H_{\mathcal{B}}(f) = (f, \mathcal{B})$$

is an isomorphism of a vector space $\widetilde{\boldsymbol{V}_n}$ onto T^n.

The following theorem follows also from the theory of solving systems of linear homogeneous equations.

Theorem 4.8 *Let f be a linear form on $\boldsymbol{V}$. Then it holds:*

1. *$f = o$, if and only if* $\operatorname{Ker} f = \boldsymbol{V}$,
2. *$f \neq o$, if and only if* $\dim \operatorname{Ker} f = n - 1$.

For a space $\operatorname{Hom}(\boldsymbol{V}_n, \boldsymbol{W}_m)$, you can construct a certain basis – see Corollary 2.32. Now let us construct its special case for a space $\widetilde{\boldsymbol{V}}$. Let us choose a basis $\mathcal{B}$ of a space $\boldsymbol{V}$,

$$\mathcal{B} = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \rangle.$$

Since in our case $m = 1$, we will denote the elements, i.e. linear forms, forming the respective basis of a space $\widetilde{\boldsymbol{V}}$ not by $\langle e_{11}, e_{21}, \ldots, e_{n1} \rangle$, but (respectively) by

$$\langle \tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \ldots, \tilde{\boldsymbol{e}}_n \rangle. \tag{4.2}$$

The above linear forms (so-called *coordinate linear forms)* fulfil the relation (verify it!)

$$\forall i,\ 1 \leq i \leq n \colon \tilde{\boldsymbol{e}}_i(\boldsymbol{e}_k) = \delta_{ik},\ \ k = 1, \ldots, n. \tag{4.3}$$

The term *coordinate linear form* follows from the following series of equalities where for arbitrary $\boldsymbol{x} \in \boldsymbol{V}$, $\{\boldsymbol{x}\}_{\mathcal{B}} = (x_1, \ldots, x_n)$, and chosen i, $i = 1, \ldots, n$, we can write:

$$\tilde{\boldsymbol{e}}_i(\boldsymbol{x}) = \tilde{\boldsymbol{e}}_i\Big(\sum_{k=1}^{n} x_k \boldsymbol{e}_k\Big) = \sum_{k=1}^{n} x_k\, \tilde{\boldsymbol{e}}_i(\boldsymbol{e}_k) \overset{(4.3)}{=} x_i.$$

Definition 4.9 Let $\mathcal{B} = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \rangle$ be a basis of a space $\boldsymbol{V}$. Then a system of linear forms $\langle \tilde{\boldsymbol{e}}_1, \tilde{\boldsymbol{e}}_2, \ldots, \tilde{\boldsymbol{e}}_n \rangle$ defined by the relation (4.3) is called a *basis of a space $\widetilde{\boldsymbol{V}}$ dual* (or *reciprocal) for a basis $\mathcal{B}$* and will be denoted by $\widetilde{\mathcal{B}}$.

From Theorem 4.4, it follows:

Theorem 4.10 *Let $\mathcal{B} = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_2 \rangle$ be a basis of a space $\boldsymbol{V}$ and let $\mathcal{G} = \langle g_1, \ldots, g_n \rangle$ be a basis of a space $\widetilde{\boldsymbol{V}}$. Then a basis $\mathcal{G}$ is a basis dual for a basis $\mathcal{B}$ if and only if*

$$\forall i,\ 1 \leq i \leq n \colon g_i(\boldsymbol{e}_k) = \delta_{ik},\ \ k = 1, \ldots, n.$$

We have shown how to assign a basis of a space $\widetilde{\boldsymbol{V}}$ to a given basis of a space $\boldsymbol{V}$. Does that mean that the set of all the bases of a dual space $\widetilde{\boldsymbol{V}}$ is exhausted by this procedure?

Using the definition of the transition matrix and the relations (4.3), you obtain the following theorem:

Theorem 4.11 *Let $\mathcal{B}$, $\mathcal{C}$ be bases of a space $\boldsymbol{V}$. Then it holds:*

$$(\widetilde{\mathcal{C}}, \widetilde{\mathcal{B}}) = (\mathcal{B}, \mathcal{C})^T.$$

Corollary 4.12 *Every basis of a space $\widetilde{\boldsymbol{V}}$ is dual for one and only one basis of a space $\boldsymbol{V}$.*

From Corollary 2.32 (3), it follows:

Theorem 4.13 *If $\mathcal{B}$ is a basis of a space $\boldsymbol{V}$, then for an arbitrary linear form f, it holds:*

$$(f, \mathcal{B}) = (a_1, \dots, a_n) \Leftrightarrow \{f\}_{\widetilde{\mathcal{B}}} = (a_1, \dots, a_n).$$

In other words: *In the chosen basis, a linear form f has the analytic expression $f(\boldsymbol{x}) = a_1x_1 + \dots + a_nx_n$ if and only if $f = a_1\tilde{\boldsymbol{e}}_1 + \dots + a_n\tilde{\boldsymbol{e}}_n$.*

Corollary 4.14 *If $\mathcal{B}$ is a basis of a space $\boldsymbol{V}$, then a mapping $\beta\colon \boldsymbol{V} \to \widetilde{\boldsymbol{V}}$ defined by the rule*

$$\forall \boldsymbol{a} \in \boldsymbol{V}\colon\ \beta(\boldsymbol{a}) = f \Leftrightarrow \{f\}_{\widetilde{\mathcal{B}}} = \{\boldsymbol{a}\}_{\mathcal{B}}$$

is an isomorphism of a vector space $\boldsymbol{V}$ onto $\widetilde{\boldsymbol{V}}$.

Note: For instance, if a vector $\boldsymbol{a}$ in the chosen basis has the coordinates $(1, 8, -2)$, then an isomorphism β assigns to it a linear form f with the analytic expression $f(\boldsymbol{x}) = x_1 + 8x_2 - 2x_3$.

Theorem 4.15 *Let f, g be linear forms on $\boldsymbol{V}$. Then it holds that* $\operatorname{Ker} f \subseteq \operatorname{Ker} g$ *if and only if there exists $c \in T$ such that $g = cf$.*

Remember that the order of the forms cannot be interchanged – in the case that g is a zero form and f a non-zero one, it is not possible to express f as a c-multiple of a form g.

Example 4.16 On a vector space $\boldsymbol{V}_3$, there are given linear forms g_1, g_2, g_3. Decide whether $\mathcal{G} = \langle g_1, g_2, g_3 \rangle$ is a basis of a space $\widetilde{\boldsymbol{V}}$ and if so, find a basis of a space $\boldsymbol{V}$ for which a basis $\mathcal{G}$ is dual if in the chosen basis $\mathcal{B}$ of a space $\boldsymbol{V}$, it is given:

$$
\begin{aligned}
g_1(\boldsymbol{x}) &= x_1 + 2x_2 \\
g_2(\boldsymbol{x}) &= x_1 - x_2 + x_3 \\
g_3(\boldsymbol{x}) &= 2x_1 + x_2.
\end{aligned}
$$

[Instruction: In order to find if $\mathcal{G}$ forms a basis, realise the meaning of coefficients of the analytic expression of the linear form according to Theorem 4.13. Then proceed as in case of, for instance, arithmetic vectors.
In order to find a basis in $\boldsymbol{V}$ for which $\mathcal{G}$ is dual, use Theorem 4.10.

Solution: Yes, it is a basis. Namely, it is a basis dual for a basis $\mathcal{C} = \langle \boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3 \rangle$, where

$$\{\boldsymbol{c}_1\}_{\mathcal{B}} = (-\tfrac{1}{3}, \tfrac{2}{3}, 1), \quad \{\boldsymbol{c}_2\}_{\mathcal{B}} = (0, 0, 1), \quad \{\boldsymbol{c}_3\}_{\mathcal{B}} = (\tfrac{2}{3}, -\tfrac{1}{3}, -1).]$$

Notes:

5 Pseudo-inverse matrices and homomorphisms

Students can define the notion of pseudo-inverse matrix (both general and Moore–Penrose one) and are able to find pseudo-inverse matrices to a given matrix. They know the connections between the theory of pseudo-inverse and the theory of homomorphisms or the theory of systems of linear equations. They can apply the theory of pseudo-inverse when constructing an optimal approximate solution of a system of linear equations.

In the previous part of your linear algebra course, you became familiar with the notion of inverse matrix to a given regular matrix. In this lesson, you will learn how to generalise the above notion for an arbitrary matrix. You will also learn how the properties of pseudo-inverse matrices relate to the solvability of systems of linear equations, and will become familiar with another method to approximately solve systems of linear equations. In addition, you will learn how to generalise the notion of inverse homomorphism also for any homomorphism.

5.1 Pseudo-inverse matrices

In case when $\boldsymbol{A}$ is a regular square matrix over a filed T, there exists a matrix inverse with respect to it, denoted by the symbol $\boldsymbol{A}^{-1}$. It is a (one and only) matrix with the property:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}. \tag{5.1}$$

In this chapter, we will properly generalise the concept of inverse matrix by introducing the notion of *pseudoinverse matrix* – regardless of not only the regularity or singularity of square matrices but also regardless of any matrix type.

Let $\boldsymbol{A} \in \mathcal{M}_{n\times n}(T)$ be a regular matrix and let us consider a following system of linear equations:

$$\boldsymbol{A}\boldsymbol{x}^T = \boldsymbol{c}^T, \tag{5.2}$$

where $\boldsymbol{c} \in T^n$.

The (unique) solution of this system is evidently a vector $\boldsymbol{x} = (x_1, x_2, \dots x_n)$,

$$\boldsymbol{x}^T = \boldsymbol{A}^{-1}\boldsymbol{c}^T. \tag{5.3}$$

The relation (5.1) further implies the identity

$$\boldsymbol{A}\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}. \tag{5.4}$$

Let us continue to consider a general matrix $\boldsymbol{A}$, namely

$$\boldsymbol{A} \in \mathcal{M}_{m\times n}(T), \tag{5.5}$$

and a system of m linear equations with n unknowns

$$\boldsymbol{A}\boldsymbol{x}^T = \boldsymbol{c}^T, \tag{5.6}$$

where $\boldsymbol{c} \in T^m$. This system is solvable for some $\boldsymbol{c}$ while remaining unsolvable for other $\boldsymbol{c}$[17].

We have seen that in the case of regular matrices, there exists a (one and only) matrix $\boldsymbol{B}$ for every matrix $\boldsymbol{A}$ with the property that $\boldsymbol{ABA} = \boldsymbol{A}$. We have also seen that it means that (for any $\boldsymbol{c} \in T^n$) a vector in the form $\boldsymbol{x}^T = \boldsymbol{B}\boldsymbol{c}^T$ is the solution of the system (5.6).

This leads us to the following consideration:

(i) Let $\boldsymbol{A} \in \mathcal{M}_{m\times n}(T)$ and let there exists a $\boldsymbol{B} \in \mathcal{M}_{n\times m}(T)$ such that

$$\boldsymbol{ABA} = \boldsymbol{A}. \tag{5.7}$$

Let further $\boldsymbol{c} \in T^m$ such that the system (5.6) is solvable for it – let $\boldsymbol{u} \in T^n$ be some of its solutions, i.e. $\boldsymbol{A}\boldsymbol{u}^T = \boldsymbol{c}^T$. Then we can write:

$$\boldsymbol{A}(\boldsymbol{B}\boldsymbol{c}^T) = (\boldsymbol{AB})\boldsymbol{c}^T = (\boldsymbol{AB})(\boldsymbol{A}\boldsymbol{u}^T) = (\boldsymbol{ABA})\boldsymbol{u}^T \stackrel{(5.7)}{=} \boldsymbol{A}\boldsymbol{u}^T = \boldsymbol{c}^T.$$

We see that a vector $\boldsymbol{x}$, $\boldsymbol{x}^T = \boldsymbol{B}\boldsymbol{c}^T$, is a (another) solution of the system (5.6).

(ii) Now let us consider that to a matrix $\boldsymbol{A} \in \mathcal{M}_{m\times n}(T)$, there exists a matrix $\boldsymbol{B} \in \mathcal{M}_{n\times m}(T)$ such that for every $\boldsymbol{c} \in T^m$ for which the system (5.6) is solvable, an ordered n-tuple $\boldsymbol{x}$, $\boldsymbol{x}^T = \boldsymbol{B}\boldsymbol{c}^T$, be one of its solutions. Does the identity (5.7) hold true in this case?

For an arbitrary $i = 1, \dots, n$ let us denote by the symbol $\boldsymbol{e}_i$ an ordered n-tuple with an i-component equal to 1 and other components equal to 0. By multiplying an arbitrary matrix by a vector $\boldsymbol{e}_i$ from the left, we obtain its i-th row. By multiplying and arbitrary matrix by a vector $\boldsymbol{e}_i^T$ from the right, we obtain its i-th column.

If we denote a column i of a matrix $\boldsymbol{A}$ by $\boldsymbol{a}^{(i)}$, it then holds

$$\boldsymbol{A}\boldsymbol{e}_i^T = \boldsymbol{a}^{(i)}, \tag{5.8}$$

i.e. for $\boldsymbol{c}^T = \boldsymbol{a}^{(i)}$, the system (5.6) is solvable. According to our assumption, however, it means that it holds that

$$\boldsymbol{A}(\boldsymbol{B}\boldsymbol{a}^{(i)}) = \boldsymbol{a}^{(i)}, \tag{5.9}$$

for $i = 1, \dots, n$ which means that we can write:

$$(\boldsymbol{ABA})\boldsymbol{e}_i^T = (\boldsymbol{AB})(\boldsymbol{A}\boldsymbol{e}_i^T) = (\boldsymbol{AB})\boldsymbol{a}^{(i)} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{a}^{(i)}) \stackrel{(5.9)}{=} \boldsymbol{a}^{(i)}.$$

[17] See the Frobenius theorem.

The obtained system of equalities $(\boldsymbol{ABA})\boldsymbol{e}_i^T = \boldsymbol{a}^{(i)}$, $1 \le i \le n$, implies that all the columns in matrices $\boldsymbol{ABA}$ and $\boldsymbol{A}$ are the same, which means that the two matrices are equal. It can thus be said that (5.7) holds.

Theorem 5.1 *Let $\boldsymbol{A} \in \mathcal{M}_{m\times n}(T)$. Then it holds:*

1. *if there exists a matrix $\boldsymbol{B} \in \mathcal{M}_{n\times m}(T)$ such that $\boldsymbol{ABA} = \boldsymbol{A}$, then it holds that for every $\boldsymbol{c} \in T^m$ for which a system $\boldsymbol{Ax}^T = \boldsymbol{c}^T$ is solvable, an n-tuple $\boldsymbol{x}$, $\boldsymbol{x}^T = \boldsymbol{Bc}^T$, belongs to a set of its solutions;*
2. *if there exists a matrix $\boldsymbol{B} \in \mathcal{M}_{n\times m}(T)$ such that for every $\boldsymbol{c} \in T^m$ for which a system $\boldsymbol{Ax}^T = \boldsymbol{c}^T$ is solvable, an n-tuple $\boldsymbol{x}$, $\boldsymbol{x}^T = \boldsymbol{Bc}^T$, belongs to a set of its solutions, then it holds that $\boldsymbol{ABA} = \boldsymbol{A}$.*

From the above considerations it follows that the notion of *pseudo-inverse matrix* can be fittingly defined in the following way:

Definition 5.2 Let a matrix $\boldsymbol{A} \in \mathcal{M}_{m\times n}(T)$ be given. Then a *pseudo-inverse matrix to a matrix $\boldsymbol{A}$* is understood to be every matrix $\boldsymbol{A}^-$, $\boldsymbol{A}^- \in \mathcal{M}_{n\times m}(T)$, with the property:

$$\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A} = \boldsymbol{A}. \tag{5.10}$$

Note: the symbol $\boldsymbol{A}^-$ is read as "$\boldsymbol{A}$ minus"; apart from the term pseudo-inverse matrix, two other terms, *generalised inversion* and *g-inversion*, are sometimes also used.

Remark 5.3 From the definition of pseudo-inverse matrix, it follows that in the case when $\boldsymbol{A}$ is a regular matrix, there exists exactly one matrix $\boldsymbol{A}^-$ and it holds that $\boldsymbol{A}^- = \boldsymbol{A}^{-1}$.

In the case when $\boldsymbol{A}$ is a zero matrix of the $m \times n$ type, a matrix $\boldsymbol{A}^-$ is any matrix of the $n \times m$ type.

A pseudo-inverse matrix to a given matrix (if it exists) does not generally have to be unique, which is why the symbol minus does not denote a mapping $\mathcal{M}_{m\times n}(T)$ to $\mathcal{M}_{n\times m}(T)$.

Is there a pseudo-inverse matrix to every matrix? Let us consider a matrix $\boldsymbol{A} \in \mathcal{M}_{m\times n}(T)$. Then based on Theorems 2.26 and 2.21, there exists a matrix $\boldsymbol{D} = (d_{ij})$ of the same type such that for every i, j, $1 \le i \le m$, $1 \le j \le n$, it holds

$$(i{\neq}j \Rightarrow d_{ij} = 0) \quad \text{and} \quad d_{ii} \in \{0, 1\},$$

and regular matrices $\boldsymbol{B} \in \mathcal{M}_{m\times m}(T)$, $\boldsymbol{C} \in \mathcal{M}_{n\times n}(T)$ with

$$\boldsymbol{A} = \boldsymbol{BDC}. \tag{5.11}$$

If we consider the definition of matrix multiplication, we will easily find that it holds (think it over!):

$$\boldsymbol{D} = \boldsymbol{D}\,\boldsymbol{D}^T\boldsymbol{D}. \tag{5.12}$$

Now it is evident that if we put

$$\boldsymbol{X} = \boldsymbol{C}^{-1}\boldsymbol{D}^T\boldsymbol{B}^{-1},$$

then by using (5.11) and (5.12), we obtain:

$$\begin{aligned}\boldsymbol{AXA} &= (\boldsymbol{BDC})(\boldsymbol{C}^{-1}\,\boldsymbol{D}^T\boldsymbol{B}^{-1})(\boldsymbol{BDC}) = (\boldsymbol{BD})(\boldsymbol{CC}^{-1})\,\boldsymbol{D}^T(\boldsymbol{B}^{-1}\boldsymbol{B})(\boldsymbol{DC}) = \\ &= \boldsymbol{B}(\boldsymbol{D}\,\boldsymbol{D}^T\boldsymbol{D})\boldsymbol{C} = \boldsymbol{BDC} = \boldsymbol{A}.\end{aligned}$$

A matrix $\boldsymbol{X}$ thus meets the requirements put on a pseudo-inverse matrix $\boldsymbol{A}^-$ (see (5.10)).

Theorem 5.4 *To every matrix there exists at least one pseudo-inverse matrix.*

What implications does the above theorem have for the solving of linear equations?

Example 5.5 A matrix $\boldsymbol{A}$ z $\mathcal{M}_{3\times 3}(\mathbb{R})$ is given:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 2 & -1 & 5 \end{pmatrix}.$$

Find pseudo-inverse matrices to $\boldsymbol{A}$.

[Instruction: Denote the searched matrix by $\boldsymbol{A}^- = (b_{ij})_{3\times 3}$. From relation (5.10), you obtain a system of linear equations for elements b_{ij}. You have surely noticed that a matrix $\boldsymbol{A}$ is singular, which means that there exists more than one pseudo-inverse matrix.

Solution: A pseudo-inverse matrix to a matrix $\boldsymbol{A}$ is every matrix having the following form:

$$\boldsymbol{A}^- = \begin{pmatrix} 1 - 3b_{31} - 2b_{13} & -3b_{32} + b_{13} & b_{13} \\ -2b_{23} - b_{31} & 1 + b_{23} - b_{32} & b_{23} \\ b_{31} & b_{32} & 0 \end{pmatrix}, \qquad b_{13}, b_{23}, b_{31}, b_{32} \in \mathbb{R}.]$$

5.2 Moore-Penrose pseudoinverse. Optimal approximate solution of systems of linear equations

Students can define the Moore-Penrose pseudo-inverse matrix. They are able to find a Moore-Penrose pseudo-inverse matrix to a given matrix. Students can define an optimal approximate solution of systems of linear equations, know its properties and are able to apply this knowledge to find the optimal approximate solution for a given system of equations. They are familiar with the Moore-Penrose pseudo-inverse homomorphism and are able to construct it for a given homomorphism. Students know the connections between the theory of Moore-Penrose pseudoinverse and the theory of projections of vector spaces onto a subspace.

In the previous chapter, you became familiar with the notion of pseudo-inverse matrix which is not generally unique to a given matrix. In this chapter, you will become familiar with a special case of pseudoinversion in the case of real matrices – a *Moore-Penrose matrix* which is uniquely assigned to a given matrix. You will see how to use the above matrix to look for certain approximate solutions of systems of linear equations (you have already become familiar with one method, see Theorem 1.90). You will also learn how to generalise the notion of inverse homomorphism to see that in the case of Euclidean vector spaces, it is possible to construct it not only to isomorphisms.

5.2.1 Moore-Penrose pseudo-inverse matrix

Definition 5.6 Let be given a matrix $\boldsymbol{A} \in \mathcal{M}_{m\times n}(\mathbb{R})$. Then a *Moore-Penrose pseudo-inverse matrix to a martix* $\boldsymbol{A}$ is understood to be a matrix $\boldsymbol{A}^{+}$, $\boldsymbol{A}^{+} \in \mathcal{M}_{n\times m}(\mathbb{R})$, with the following properties:

1. $\boldsymbol{A}^{+}$ is a pseudo-inverse matrix to $\boldsymbol{A}$ (tj. $\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{A} = \boldsymbol{A}$),
2. $\boldsymbol{A}$ is a pseudo-inverse matrix to $\boldsymbol{A}^{+}$ (tj. $\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{A}^{+} = \boldsymbol{A}^{+}$),
3. matrices $\boldsymbol{A}\boldsymbol{A}^{+}$ i $\boldsymbol{A}^{+}\boldsymbol{A}$ are symmetric.

Note: the symbol $\boldsymbol{A}^{+}$ is read as "$\boldsymbol{A}$ *plus*"; for matrix $\boldsymbol{A}^{+}$, only the term the *Moore-Penrose matrix* is sometimes used.

Realise that we define the Moore-Penrose matrix only for matrices over real numbers.

What significance do the matrix products $\boldsymbol{P}{=}\boldsymbol{A}\boldsymbol{A}^{+}$ and $\boldsymbol{Q}{=}\boldsymbol{A}^{+}\boldsymbol{A}$ have? Using subsections 1 and 2 of Definition 5.6, you will easily find that matrices

$\boldsymbol{P}$, $\boldsymbol{Q}$ are idempotent, which means that with respect to subsection 3, the above definition and according to Corollary 2.58 and Theorem 2.80, these matrices represent matrices of orthogonal projections p, q, respectively, of spaces $\mathbb{R}^m$, $\mathbb{R}^n$, respectively, with the standard scalar product.
In an arithmetic vector space, a homomorphism whose image is equal to the row subspace of a given matrix $\boldsymbol{C}$, is given by the following formula [why?]:

$$f(\boldsymbol{x}) = \boldsymbol{x}\,\boldsymbol{C}.$$

With regard to Theorem 2.35, you can see that $\operatorname{Im} p$ is included in the row subspace of a matrix $\boldsymbol{A}^+$ and $\operatorname{Im} q$ is included in the row subspace of a matrix $\boldsymbol{A}$.

You know from the linear algebra course that the rank of the product of arbitrary matrices is lower than or equal to the rank of any of the matrices[18]. Using Definition 5.2, try to verify that specially for the product of a matrix and its pseudo-inverse matrix, this particular relation of ranks becomes an equality. This is why the images of both projections are equal to the above row subspaces.

Lemma 5.7 *Let $\boldsymbol{A}$ be a real matrix to which there exists a matrix $\boldsymbol{A}^+$. Then it holds:*

1. *$(\boldsymbol{A}\boldsymbol{A}^+)$ is a matrix of an orthogonal projection of space $\mathbb{R}^m$ onto a row subspace of a matrix $\boldsymbol{A}^+$,*
2. *$(\boldsymbol{A}^+\boldsymbol{A})$ is a matrix of an orthogonal projection of space $\mathbb{R}^n$ onto a row subspace of a matrix $\boldsymbol{A}$.*

If you consider that by transposition, a row subspace of a given matrix turns into a column one, it then is possible to derive the following from the definition of Moore-Penrose matrix:

Lemma 5.8 *Let $\boldsymbol{A}$ be a real matrix to which there exists a matrix $\boldsymbol{A}^+$. Then it holds:*

1. *a column subspace of a matrix $\boldsymbol{A}$ is equal to the row subspace of a matrix $\boldsymbol{A}^+$,*
2. *a column subspace of a matrix $\boldsymbol{A}^+$ is equal to the row subspace of a matrix $\boldsymbol{A}$.*

[18] If you have forgotten this fact, derive it from, for instance, Theorem 2.35.

Remark 5.9 Let a matrix $\boldsymbol{A} \in \mathcal{M}_{m\times n}(\mathbb{R})$ be given. Now let us describe the construction of a matrix $\boldsymbol{A}^{+}$. Let us denote the basis of column subspace of a matrix $\boldsymbol{A}$ by $\mathcal{B} = \langle \boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)} \ldots \boldsymbol{b}^{(r)}\rangle$ and let us denote a matrix formed by the above columns by $\boldsymbol{F}$, i.e. $\boldsymbol{F} \in \mathcal{M}_{m\times r}(\mathbb{R})$. Let us then express any of columns $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)} \ldots \boldsymbol{a}^{(n)}$ of a matrix $\boldsymbol{A}$ as a linear combination of elements of a basis $\mathcal{B}$:

$$\begin{aligned}
\boldsymbol{a}^{(1)} &= g_{11}\boldsymbol{b}^{(1)} + g_{12}\boldsymbol{b}^{(2)} + \cdots + g_{1r}\boldsymbol{b}^{(r)}\\
\boldsymbol{a}^{(2)} &= g_{21}\boldsymbol{b}^{(1)} + g_{22}\boldsymbol{b}^{(2)} + \cdots + g_{2r}\boldsymbol{b}^{(1)}\\
\ldots\\
\boldsymbol{a}^{(n)} &= g_{n1}\boldsymbol{b}^{(1)} + g_{n2}\boldsymbol{b}^{(1)} + \cdots + g_{nr}\boldsymbol{b}^{(1)}
\end{aligned}$$

and let us construct a matrix $\boldsymbol{G} = (g_{ij})$, $\boldsymbol{G} \in \mathcal{M}_{n\times r}(\mathbb{R})$.

You can see that the Moore-Penrose matrix to a matrix $\boldsymbol{A}$ is given by the relation

$$\boldsymbol{A}^{+} = \boldsymbol{G}(\boldsymbol{G}^{T}\boldsymbol{G})^{-1}(\boldsymbol{F}^{T}\boldsymbol{F})^{-1}\boldsymbol{F}^{T}. \tag{5.13}$$

Is the Moore-Penrose matrix to a given matrix unique? If we consider a real matrix $\boldsymbol{A}$ of the $m \times n$ type, an orthogonal projection of a space $\mathbb{R}^n$ onto its row subspace and an orthogonal projection of a space $\mathbb{R}^m$ onto its column space are thereby without a doubt unequivocally determined – that means that their matrices $\boldsymbol{Q}$, $\boldsymbol{P}$ are also unequivocally determined in accordance with Lemmas 5.7 and 5.8. If $\boldsymbol{A}^{*}$ is also a Moore-Penrose matrix to a matrix $\boldsymbol{A}$, it will have to hold

$$(\boldsymbol{A}^{+}\boldsymbol{A}) = \boldsymbol{Q} = (\boldsymbol{A}^{*}\boldsymbol{A}) \quad \text{and} \quad (\boldsymbol{A}\boldsymbol{A}^{+}) = \boldsymbol{P} = (\boldsymbol{A}\boldsymbol{A}^{*}),$$

from which you can derive $\boldsymbol{A}^{*} = \boldsymbol{A}^{+}$, using subsections 1 and 2 of Definition 5.2 (try it!).

Theorem 5.10 *To every real matrix, there exists one and only Moore-Penrose pseudo-inverse matrix.*

In Section 1.3, you learned how to find an approximate solution of the system of linear equations

$$\boldsymbol{A}(x_1, x_2, \ldots, x_n)^{T} = (b_1, b_2, \ldots, b_r)^{T},$$

which we called the method of the smallest squares (see Theorem 1.90). Let us recall that we tried to find such arithmetic vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ for which the value

$$\rho((\boldsymbol{A}(x_1, x_2, \ldots, x_n)^{T}), (b_1, b_2, \ldots, b_r)^{T})$$

would be lowest possible.

Now we will show another method of approximate solution of systems of linear equations which will try to minimise not only the value $\|\boldsymbol{A}\boldsymbol{x}^T - \boldsymbol{b}^T\|$ but also the length of this approximate solution, i.e. the value $\|\boldsymbol{x}\|$. This approximate solution can be defined in the following way:

Definition 5.11 Let be given a system of linear equations

$$\boldsymbol{A}\boldsymbol{x}^T = \boldsymbol{b}^T, \tag{5.14}$$

where $\boldsymbol{A} \in \mathcal{M}_{m\times n}(\mathbb{R})$, $\boldsymbol{b} \in \mathbb{R}^m$. An arithmetic vector $\boldsymbol{x}_0 \in \mathbb{R}^n$ is called an *optimal approximate solution of a system of linear equations* (5.14) if:

1. for every $\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^n$, it holds: $\left\|\boldsymbol{A}\boldsymbol{x}^T - \boldsymbol{b}^T\right\| \geq \left\|\boldsymbol{A}\boldsymbol{x}_0^T - \boldsymbol{b}^T\right\|$,
2. if for some $\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x} \neq \boldsymbol{x}_0$, it holds: $\left\|\boldsymbol{A}\boldsymbol{x}^T - \boldsymbol{b}^T\right\| = \left\|\boldsymbol{A}\boldsymbol{x}_0^T - \boldsymbol{b}^T\right\|$, then $\|\boldsymbol{x}_0\| < \|\boldsymbol{x}\|$.

For a $\boldsymbol{b}$ for which (5.14) is solvable, the n-tuple $\boldsymbol{x}^T{=}\boldsymbol{A}^-\boldsymbol{b}^T$ is one of its solutions according to Theorem 5.1. What role in particular does the solution $\boldsymbol{x}_0^T{=}\boldsymbol{A}^+\boldsymbol{b}^T$ play? If $\boldsymbol{x}$ is also the solution of the system (5.14), it is possible to write:

$$\boldsymbol{x}_0 = \boldsymbol{b}(\boldsymbol{A}^+)^T = (\boldsymbol{x}\boldsymbol{A}^T)(\boldsymbol{A}^+)^T = \boldsymbol{x}(\boldsymbol{A}^+\boldsymbol{A}), \quad \text{i.e. } \boldsymbol{x}_0 = q(\boldsymbol{x}).$$

As we know, q is an orthogonal projection, which means that according to Theorem 1.72, it holds that $\boldsymbol{x} \neq \boldsymbol{x}_0$: $\|\boldsymbol{x}_0\| < \|\boldsymbol{x}\|$. This means that $\boldsymbol{x}_0 = \boldsymbol{b}(\boldsymbol{A}^+)^T$ meets both requirements of Definition 5.11.[19]

And what significance does $\boldsymbol{x}_0 = \boldsymbol{b}(\boldsymbol{A}^+)^T$ have in an general case (that is, when $\boldsymbol{b}$ is an arbitrary vector from $\mathbb{R}^m$ and thus even such when the system is unsolvable)? For the product $\boldsymbol{A}\boldsymbol{x}_0^T$, we can write:

$$\boldsymbol{x}_0\boldsymbol{A}^T = \boldsymbol{b}(\boldsymbol{A}^+)^T\boldsymbol{A}^T = \boldsymbol{b}(\boldsymbol{A}\boldsymbol{A}^+)^T = \boldsymbol{b}(\boldsymbol{A}\boldsymbol{A}^+), \quad \text{i.e. } (\boldsymbol{x}_0\boldsymbol{A}^T) = p(\boldsymbol{b}),$$

which means – since p is an orthogonal projection – that according to Lemmas 5.8 and 5.7, $(\boldsymbol{x}_0\boldsymbol{A}^T)$ belongs to the column subspace of a matrix $\boldsymbol{A}$. It also means that according to Theorem 1.83, requirement 1 of Definition 5.11.

Similar considerations will also show that for any $\boldsymbol{x}$ for which $\|\boldsymbol{A}\boldsymbol{x}^T - \boldsymbol{b}\| = \|\boldsymbol{A}\boldsymbol{x}_0^T - \boldsymbol{b}\|$ (which implies that $\boldsymbol{x}\boldsymbol{A}^T = \boldsymbol{x}_0\boldsymbol{A}^T$) the vector $\boldsymbol{x}_0$ meets also requirement 2 (try it!).

[19] The first requirement is met trivially.

Theorem 5.12 *Let (5.14) be a system of linear equations with a matrix* $\boldsymbol{A} \in \mathcal{M}_{m\times n}(\mathbb{R})$*. Then an arithmetic vector* $\boldsymbol{x}_0 \in \mathbb{R}^n$ *given by the relation*

$$\boldsymbol{x}_0^T = \boldsymbol{A}^+ \boldsymbol{b}^T \tag{5.15}$$

is the optimal approximate solution of the system of linear equations (5.14).

Theorem 5.13 *If the system of linear equations (5.14) is solvable, then an arithmetic vector* $\boldsymbol{x}_0$ *given by the relation (5.15) is such solution of this system that has the smallest possible length of all its solutions.*

Note: This solution of a given system is called an *optimal* or *minimal solution*; for this solution, the value $\sum_{i=1}^{n} x_i^2$ is the smallest possible of all the solutions.

5.2.2 Moore-Penrose homomorphism

As you saw in Chapter 2.1, there exists an inverse mapping only for these homomorphisms that are isomorphisms. Let us then generalise the notion of isomorphism so that it is possible to construct a certain generalised mapping for every homomorphism.

Let us consider Euclidean vector spaces $\boldsymbol{V}$, $\boldsymbol{W}$ and a homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$. Let us further denote:

$$\boldsymbol{V}_{(1)} = \operatorname{Ker} f,\ \ \boldsymbol{V}_{(2)} = \boldsymbol{V}_{(1)}^{\perp},\ \ \boldsymbol{W}_{(2)} = \operatorname{Im} f,\ \ \boldsymbol{W}_{(1)} = \boldsymbol{W}_{(2)}^{\perp}. \tag{5.16}$$

Since $\boldsymbol{V} = \boldsymbol{V}_{(1)} \oplus \boldsymbol{V}_{(2)}$ (and f is a homomorphism), it holds for the restriction of f to $\boldsymbol{V}_{(2)}$ that $\operatorname{Im}(f|\boldsymbol{V}_{(2)}) = \operatorname{Im} f$ and $\operatorname{Ker} f|\boldsymbol{V}_{(2)} = \{\boldsymbol{o}\}$, which means that $f|\boldsymbol{V}_{(2)}$ is an isomorphism $\boldsymbol{V}_{(2)}$ onto $\boldsymbol{W}_{(2)}$.

Let us consider an orthogonal projection $p_{W_{(2)}}$ of a space $\boldsymbol{W}$ onto a subspace $\boldsymbol{W}_{(2)}$.

Therefore the following composition gives a homomorphism $f^{+}\colon \boldsymbol{W} \to \boldsymbol{V}$:

$$f^{+} = p_{W_{(2)}} \circ (f|\boldsymbol{V}_{(2)})^{-1} \tag{5.17}$$

A homomorphism f^{+} can be equivalently expressed by the formula:

$\forall \boldsymbol{x} \in \boldsymbol{W},\ \boldsymbol{x}{=}\boldsymbol{x}_1{+}\boldsymbol{x}_2,\ \boldsymbol{x}_1 \in \boldsymbol{W}_1, \boldsymbol{x}_2 \in \boldsymbol{W}_{(2)}:$

$$(f^{+}(\boldsymbol{x}) = \boldsymbol{y}) \Leftrightarrow (f(\boldsymbol{y}) = \boldsymbol{x}_2,\ \boldsymbol{y} \in \boldsymbol{V}_{(2)}). \tag{5.18}$$

Definition 5.14 Let $\boldsymbol{V}, \boldsymbol{W}$ be Euclidean vector spaces and let a homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$ be given. Then a homomorphism $f^{+}\colon \boldsymbol{W} \to \boldsymbol{V}$ defined by the relation (5.17) is called a *Moore-Penrose pseudo-inverse homomorphism of a homomorphism f.*

Similarly to the case of the Moore-Penrose matrix, we can refer to this type of homomorphism as only the *Moore-Penrose homomorphism.*

Remark 5.15 From the considerations prior to Definition 5.14, it follows that f^{+} is an epimorphism if and only if f is a monomorphism and that it is a monomorphism if and only if f is an epimorphism. It is evident that f^{+} is an inverse homomorphism to f, i.e. .$f^{+} = f^{-1}$ if and only if f is an isomorphism.

Theorem 5.16 *Let $\boldsymbol{V}$, $\boldsymbol{W}$ be Euclidean vector spaces. For every homomorphism $f\colon \boldsymbol{V} \to \boldsymbol{W}$ there exists one and only one Moore-Penrose pseudo-inverse homomorphism $f^{+}\colon \boldsymbol{W} \to \boldsymbol{V}$.*

Remark 5.17 The Moore-Penrose pseudo-inverse homomorphism depends on the choice of the scalar products in both vector spaces – if we change some of them, we generally obtain a different homomorphism f^+ to the chosen homomorphism f (cf. Remark 1.2).

By use of the relations (5.16) and (5.17), or (5.18), demonstrate that for the chosen homomorphism f and its Moore-Penrose homomorphism f^+, it holds:

- $f \circ f^+ \circ f = f$,
- $f^+ \circ f \circ f^+ = f^+$,
- $f \circ f^+$ is an orthogonal projection of $\boldsymbol{V}$ onto $\boldsymbol{V}_{(2)}$,
- $f^+ \circ\ f$ is an orthogonal projection of $\boldsymbol{W}$ onto $\boldsymbol{W}_{(2)}$.

Hence and from Definition 5.6[20] the validity of the theorem describing a matrix of the Moore-Penrose homomorphism is already evident (why is the assumption of the orthonormality of the bases necessary?).

Theorem 5.18 *Let $\boldsymbol{V}$, $\boldsymbol{W}$ be Euclidean vector spaces and let $\mathcal{B}$, $\mathcal{C}$ be, respectively, orthonormal bases of these spaces. Then for every homomorphism $f : \boldsymbol{V} \to \boldsymbol{W}$, it holds:*

$$(f^+, \mathcal{C}, \mathcal{B}) = (f, \mathcal{B}, \mathcal{C})^+. \tag{5.19}$$

Example 5.19 Let a system of linear equations be given:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 5 \\ x_1 + \ \ x_2 + 2x_3 + 3x_4 &= 2 \\ 2x_1 + 3x_2 + 5x_3 + 7x_4 &= 3 \end{aligned}$$

Find its optimal approximate solution.

[Instruction: Denote the matrix of the system by $\boldsymbol{A}$ and the vector of right sides by $\boldsymbol{b}$. The optimal approximate solution is given by the relation (5.15). First find a matrix $\boldsymbol{A}^+$ – proceed according to Remark 5.9:

Since $\mathrm{h}(\boldsymbol{A}) = r = 2$, it is possible to choose

$$\boldsymbol{F} = (\boldsymbol{b}^{(1)}, \boldsymbol{b}^{(2)}) = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 2 & 3 \end{pmatrix},$$

[20] Also by use of, for example, theorems 2.35 and 2.80.

and then

$$\boldsymbol{G} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

We will further find that

$$(\boldsymbol{F}^T\boldsymbol{F})^{-1} = \frac{1}{3}\begin{pmatrix} 14 & -9 \\ -9 & 6 \end{pmatrix}, \; (\boldsymbol{G}^T\boldsymbol{G})^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

Substituting to the relation (5.13), you obtain

$$\boldsymbol{A}^{+} = \frac{1}{9}\begin{pmatrix} -7 & 8 & 1 \\ 10 & -11 & -1 \\ 3 & -3 & 0 \\ -4 & 5 & 1 \end{pmatrix}.$$

Substituting to the relation (5.15), you obtain the following result:

$$\boldsymbol{x}_0^T = \frac{1}{9}\begin{pmatrix} -7 & 8 & 1 \\ 10 & -11 & -1 \\ 3 & -3 & 0 \\ -4 & 5 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix},$$

which means that the optimal approximate solution is the ordered triple

$$\boldsymbol{x}_0 = (x_1, x_2, x_3) = \left(-\frac{16}{9}, \frac{25}{9}, -\frac{7}{9}\right).$$

Note: for the given system, it holds that $\mathrm{h}(\boldsymbol{A}) = 2$, $\mathrm{h}(\boldsymbol{A}|\boldsymbol{b}) = 3$, which means that the system is not solvable.]

Notes:

Notes:

References

In this bibliography, the interested can find further information on the topic of this textbook and related topics.

[1] Bican, L.: *Lineární algebra a geometrie.* 2. edition Praha: Academia, 2009, 303 p. ISBN 978-80-200-1707-9.

[2] Bican, L.: *Lineární algebra v úlohách.* Praha: Státní pedagogické nakladatelství, 1982, 303 p.

[3] Gantmacher, F. R.: *Teorija matric.* 4. edition Moskva: Nauka, 1988, 548 p.

[4] Gel̆fand, I.: *Lectures on linear algebra.* New York: Dover Publications, 1989, 185 p. ISBN 04-866-6082-6.

[4] Hefferon, J.: *Linear algebra.* Colchester, 2017, 525 p. ISBN 978-1944325114.

[5] Jukl, M.: *Lineární operátory.* Olomouc: Univerzita Palackého, Přírodovědecká fakulta, 2001, 107 p. ISBN 80-244-0342-0.

[6] Jukl, M.: *Lineární algebra: euklidovské vektorové prostory : homomorfizmy vektorových prostorů.* 2. edition Olomouc: Univerzita Palackého v Olomouci, 2010, 179 p. ISBN 978-80-244-2522-1.

[7] Naylor, W., Sell, G.: *Linear Operator Theory in Engineering and Science.* New York: Springer, 1982, 624 p. ISBN 978-1-4612-5773-8

[8] Rao, C. R., Mitra, K. S.: *Generalized Inverse of Matrices and Its Application*, New York 1971 Generalized inverse of matrices and its applications. 2. edition New York: Wiley, 1971, xiv, 240 p. ISBN 04-717-0821-6.

[9] Zlatoš, P.: *Lineárna algebra a geometria: cesta z troch rozmerov s presahmi do príbuzných odborov.* Bratislava: Marenčin PT, 2011, 741 p. ISBN 978-80-8114-111-9.

Doc. RNDr. Marek Jukl, Ph.D.

Linear Algebra Lessons

Výkonný redaktor Mgr. Miriam Delongová
Odpovědný redaktor Bc. Otakar Loutocký
Technická redakce autor
Návrh a grafické zpracování obálky Bc. Karina Pavlíková, Ivana Perůtková

Publikace neprošla ve VUP redakční jazykovou úpravou

Vydala a vytiskla Univerzita Palackého v Olomouci
Křížkovského 8, 771 47 Olomouc
www.vydavatelstvi.upol.cz
www.e-shop.upol.cz
vup@upol.cz

1. vydání
Olomouc 2020
Neprodejná publikace
ISBN 978-80-244-5824-3
VUP 2020/0394