MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
NATIONAL TECHNICAL UNIVERSITY
«KHARKIV POLYTECHNIC INSTITUTE»

L. V. Kurpa, K. I. Liubytska, V. M. Burlayenko

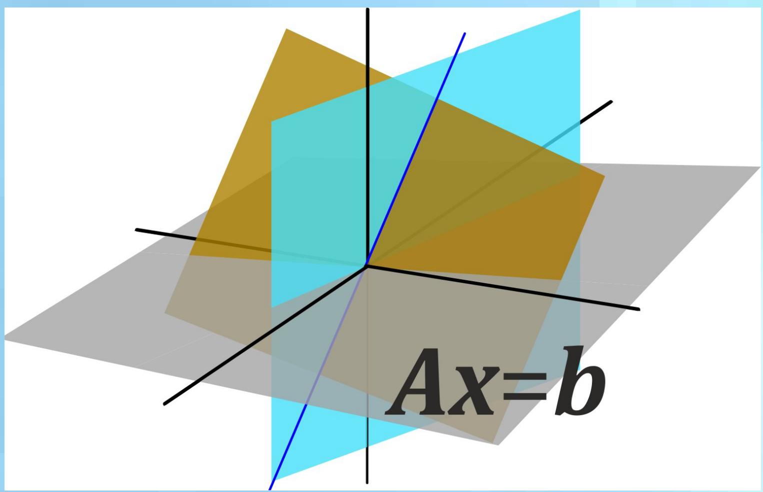
LINEAR ALGEBRA:
The Textbook for Engineering Students

EDUCATIONAL TEXTBOOK
for students of technical specialties in all forms of education

Л. В. Курпа, К. І. Любицька, В. М. Бурлаєнко

ЛІНІЙНА АЛГЕБРА:
Курс для студентів інженерно-технічних спеціальностей

НАВЧАЛЬНИЙ ПОСІБНИК
для студентів технічних спеціальностей усіх форм навчання



Kharkiv
NTU «KhPI»
2024

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

L. V. Kurpa, K. I. Liubytska, V. M. Burlayenko

LINEAR ALGEBRA:
The Textbook for Engineering Students

EDUCATIONAL TEXTBOOK
for students of technical specialties in all forms of education

Л. В. Курпа, К. І. Любицька, В. М. Бурлаєнко

ЛІНІЙНА АЛГЕБРА:
Курс для студентів інженерно-технічних спеціальностей

НАВЧАЛЬНИЙ ПОСІБНИК
для студентів технічних спеціальностей усіх форм навчання

ЗАТВЕРДЖЕНО
редакційно-видавничою
радою НТУ «ХПІ»,
протокол
№ 1 від 15. 02. 2024 р.

Харків
НТУ «ХПІ»
2024

УДК 512.64 (075)
К93

Reviewers:

N.D. Sizova, Professor, Dr. Phys.-Math. Sci., Professor of the Department of Computer Sciences and Information Technologies, M. Beketov National University of Urban Economy in Kharkiv

O.M. Lytvyn, Professor, Dr. Phys.-Math. Sci., Professor of the Department of Information Computer Technologies and Mathematics, Ukrainian Engineering and Pedagogical Academy

Kurpa L.V.

К 93 Лінійна алгебра: Курс для студентів інженерно-технічних спеціальностей : навчальний посібник для студентів технічних спеціальностей усіх форм навчання / Л. В. Курпа, К. І. Любицька, В. М. Бурлаєнко - Харків : НТУ «ХПІ», 2024. – 154 с. – Англійською мовою.

Посібник містить теоретичний матеріал з лінійної алгебри англійською мовою, який охоплює ключові поняття, твердження та формули. Він призначений для глибокого розуміння та розвитку навичок у застосуванні лінійної алгебри. Крім того, у посібнику наведено численні приклади для ілюстрації практичного застосування матеріалу, що полегшує засвоєння понять студентами технічних спеціальностей усіх форм навчання.

Розроблено для студентів технічних університетів, що вивчають курс лінійної алгебри англійською мовою. Цей посібник також корисний для іноземних студентів і викладачів, які шукають допомогу у розробці власних лекційних матеріалів для вищих технічних навчальних закладів.

ISBN 978-617-05-0479-1

Курпа Л.В.

К 93 Linear algebra: the Textbook for Engineering Students : educational textbook for students of technical specialties in all forms of education / L. V. Kurpa, K. I. Liubytska, V. M. Burlayenko - Kharkiv : NTU «KhPI», 2024. – 154 p. – in English.

This textbook provides theoretical content on linear algebra presented in English. It covers key concepts, statements, and formulas essential for a profound understanding and skill development in working with linear algebra. Additionally, numerous examples are included to illustrate the practical applications of the presented material, facilitating easier mastery of the concepts for students of technical specialties in all forms of education.

Tailored for students at technical universities who are taking a linear algebra course in English. It's also useful for foreign students and lecturers seeking assistance in developing their own lecture materials in higher technical educational institutions.

Fig. 7; Bibl. titles: 12

УДК 512.64 (075)

ISBN 978-617-05-0479-1

CONTENT

INTRODUCTION

Mathematical methods play a vital role in various scientific and engineering disciplines. Therefore, students, irrespective of their chosen field, require indispensably a solid theoretical foundation in mathematics to address real-world challenges effectively. Linear algebra provides a powerful toolkit for solving a wide range of problems across various disciplines, from physics and engineering to computer science and economics. Its methods are fundamental in analyzing and manipulating data, modeling physical systems, and understanding complex structures. Thus, proficiency in linear algebra is indispensable for students pursuing careers in these fields.

Moreover, in today's interconnected world, where collaboration knows no borders, the ability to communicate mathematical ideas effectively across linguistic and cultural boundaries is paramount. Ukrainian technical universities are increasingly engaged in collaborative research and projects with international partners. Hence, students equipped with knowledge of international mathematical terminology can seamlessly integrate into these collaborative environments. They can contribute meaningfully to interdisciplinary teams, draw upon a diverse range of perspectives, and tackle complex challenges with confidence. Therefore, fostering proficiency in international mathematical language not only enhances students' academic and professional prospects but also fosters global cooperation and innovation in science and engineering.

Hence, recognizing the crucial role of linear algebra in engineering education and the necessity of understanding international mathematical terminology, we felt inspired to write the textbook "Linear Algebra: A Textbook for Engineering Students". This textbook offers a mathematical course tailored for undergraduate students at technical universities. It has evolved from a series of lectures delivered by the authors over the past decade at the National Technical

University "Kharkov Polytechnic Institute", including students enrolled in mathematical courses taught in English.

The textbook consists of six chapters, each designed to provide a comprehensive understanding of linear algebra concepts. The opening chapter introduces fundamental theoretical principles concerning matrices and determinants, laying the groundwork for proficient solving and analysis of linear algebraic systems. Additionally, the chapter delves into the definition and operations with block matrices, offering a detailed examination of this essential topic. In the subsequent chapter, the focus shifts to the exploration of linear spaces, providing insights into their key concepts, properties, and applications in solving linear algebraic problems. This also includes a detailed explanation of coordinate transformation with a change of basis in a linear space and operations on subspaces. Moving forward, the third chapter serves as a thorough introduction to linear operators, shedding light on their significance and applications. As linear transformations play a central role in linear algebra, linear transformations are thoroughly considered in this chapter of the textbook. Additionally, this chapter introduces eigenvalues and eigenvectors, highlighting their significance in decomposing matrices into simpler formats and uncovering essential system characteristics. The fourth chapter focuses on concepts related to Euclidean space and orthonormal bases, providing a solid foundation for understanding geometric aspects within the linear algebra. It delves into topics such as inner products, orthogonality, and projections. Additionally, the chapter explores the Gram-Schmidt process for orthogonalization and the concept of minimization problem. In the fifth chapter, the discussion expands to encompass advanced topics in linear algebra. These include a detailed examination of adjoint and self-adjoint operators, unitary and orthogonal operators, and an in-depth exploration of their fundamental properties. This chapter goes deeper into the intricate aspects of these mathematical constructs, providing students with a comprehensive understanding of their applications and significance within the field of linear algebra. The final, sixth chapter considers bilinear and quadratic forms, providing a comprehensive exploration of these mathematical constructions. It covers various aspects, including their matrix representation, eigenvalues, and eigenvectors associated with bilinear and quadratic forms, discussions on positive definite, negative definite, and indefinite quadratic forms, as well as Sylvester's law of inertia for quadratic forms and diagonalization. Each chapter is enriched with numerous examples, enhancing the clarity and comprehension of the subject matter.

This book is recommended for students at technical universities enrolled in the Higher Mathematics course conducted in English, as well as for foreign students and universities lecturers seeking assistance in developing their own lecture materials. Additionally, it is valuable for anyone with an interest in acquiring knowledge of linear algebra using mathematical terminology in English.

Chapter 1. MATRICES AND DETERMINANTS

1.1. Basic Concepts of Matrices

The basic concepts of the matrices theory and determinants were presented in the course "Algebra and Analytical Geometry". In addition to these basic concepts, linear algebra studies finite, countable as well as infinite-dimensional vector spaces, linear operators, methods for finding their eigenvalues and eigenvectors, quadratic forms and methods for their reduction to canonical form, as well as many other topics that require knowledge of mathematics. operations with matrices and determinants. Therefore, we will begin this course by reviewing matrices and determinants.

First of all, let's remember what is called a matrix and what arithmetic operations with matrices can be performed.

Definition. A ***matrix*** *A* of size $m \times n$ is called a set of $m \cdot n$ elements a_{ij} written in the table with *m* rows and *n* columns, which has a form:

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

In more compact form the matrix is denoted as follows:

$$A = \left(a_{ij}\right)_{m,n}, \text{ where } i = 1...m, \ j = 1...n.$$

The individual elements a_{ij} are also called *entries* or *components of the matrix A*. The first subscript is the number of the *i*-th row and the second one is the number of the *j*-th column where the component a_{ij} is located. Note that here *m* is a total number of rows, while *n* is a total number of columns in the matrix *A*.

Definition. Matrices *A* and *B* are called ***equal*** if they have the same size and their corresponding components are equal.

Definition. A matrix with all zero components is called ***null matrix*** or ***zero matrix***.

Definition. A matrix is called a ***square*** of the order *n* if the number of rows coincides with the number of columns, i.e. $m = n$.

Definition. A square matrix is called ***diagonal*** if all the components except those located on the leading (main) diagonal ($a_{11}, a_{22}, \ldots, a_{nn}$) are equal to zero. The diagonal matrix is denoted as

$$A = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix} = diag(a_{11}, a_{22}, \ldots, a_{nn}).$$

Definition. A diagonal matrix in which all the components of the leading diagonal are equal to 1 is called an ***identity*** (*unit*) ***matrix*** and is usually denoted by the letter *E* or *I*.

When it is necessary to distinguish which size of identity matrix is being discussed, we will use the notation I_n for the $n \times n$ identity matrix.

Definition. The matrix A^T is called ***transposed*** to the matrix *A*, if it is obtained from the given matrix *A* by replacing the columns with rows or vice versa, i.e.

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

Definition. If $A = A^T, \Rightarrow a_{ij} = a_{ji},\ i = \overline{1,m},\ j = \overline{1,n}$ then the matrix is called ***symmetrical matrix.***

Definition. The matrix *A* is called a skew-symmetric matrix if

$$A = -A^T, \Rightarrow a_{ij} = -a_{ji}, \; i = \overline{1,m}, \; j = \overline{1,n}$$

Definition. A matrix is called a ***sparse*** matrix if its zero components predominate over non-zero components.

An example of the sparse matrix is any diagonal matrix.

A square matrix *A* of the *n*-th order is called *K*-diagonal (where *K* is a positive odd number) if

$$a_{ij} = 0 \text{ provided that } |i - j| > \frac{K-1}{2}$$

Example of a three-diagonal matrix:

$$\begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & \dots & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 & \dots & 0 \\ 0 & 0 & a_{43} & a_{44} & a_{45} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

1.2. Basic Operations on Matrices

Since the matrices are mathematical objects, it is naturally to introduce some algebraic operations on them such as addition, subtraction and multiplication.

Definition. ***The sum*** of two matrices *A* and *B* of the same size is called a matrix *C*=*A*+*B* with the components defined by the elementwise sum of the corresponding original matrices, i.e.

$$C = A + B, \quad c_{ij} = a_{ij} + b_{ij}, \; i = \overline{1,m}, \; j = \overline{1,n}$$

It is obvious that the matrix *C* has the same size as the original matrices.

Definition. ***The multiplication of the matrix*** *A* ***by a scalar*** α is called a matrix $C=\alpha A$ whose components are computed by multiplication of the corresponding components of the matrix *A* by the given scalar α, i.e.

$$c_{ij} = \alpha a_{ij}, \; i = \overline{1,m}, \; j = \overline{1,n}$$

<u>Definition.</u> ***The multiplication*** of the matrix *A* of size $m \times n$ by the matrix *B* of size $n \times p$ is called a matrix *C* = *AB* of the size $m \times p$ with components defined as follows:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \;\; i = \overline{1, m}, \;\; j = \overline{1, p}$$

Let us remain on some special cases of matrix multiplication.

1. *The rule of multiplication of diagonal matrices*.

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

That is, as a result of multiplying two diagonal matrices *A* and *B* we obtain a diagonal matrix *C*, whose diagonal components are calculated as a product of the diagonal components of the original matrices, i.e.

$$\mathrm{diag}(a_{11}, a_{22}, ..., a_{nn}) \cdot \mathrm{diag}(b_{11}, b_{22}, ..., b_{nn}) = \mathrm{diag}(a_{11}b_{11}, a_{22}b_{22}, ..., a_{nn}b_{nn})$$

2. *The rule of multiplying a matrix by a diagonal matrix*.

<u>*Rule A*:</u> multiplication of the diagonal matrix by the matrix on the left

$$\begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} d_{11}a_{11} & d_{11}a_{12} & \dots & d_{11}a_{1n} \\ d_{22}a_{21} & d_{22}a_{22} & \dots & d_{22}a_{2n} \\ \dots & \dots & \dots & \dots \\ d_{nn}a_{n1} & d_{nn}a_{n2} & \dots & d_{nn}a_{nn} \end{pmatrix}$$

That is, if matrix A is multiplied by a diagonal matrix on the left, we obtain a matrix whose rows are multiplied by the corresponding diagonal element located in the corresponding row.

Rule B: multiplication of the matrix by the diagonal matrix on the right

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix} \begin{pmatrix} d_{11} & 0 & ... & 0 \\ 0 & d_{22} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & d_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}d_{11} & a_{12}d_{22} & ... & a_{1n}d_{nn} \\ a_{21}d_{11} & a_{22}d_{22} & ... & a_{2n}d_{nn} \\ ... & ... & ... & ... \\ a_{n1}d_{11} & a_{n2}d_{22} & ... & a_{nn}d_{nn} \end{pmatrix}$$

If the matrix A is multiplied by the diagonal matrix on the right, it is equivalent to multiplying each column by the element of the diagonal matrix located in the corresponding column.

This result explains why $I = \text{diag}(1,1,...,1)$ is called an identity (unity) matrix.

1.3 Block matrices

If a matrix is very large and/or the matrix contains groups of the components that can be collected together based on some common properties, then a special algebraic construction called a ***block matrix*** is used instead of an extended matrix.

Definition. If all components of a matrix are matrices of certain dimensions then it is called a ***block matrix***, and the components of such matrix are called ***blocks***.

Agreement of the blocks means that all blocks located in one row of the block matrix have the same number of rows, and in one column – the same number of columns**.** The number of rows k and the number of columns l of the block matrix of size $m \times n$ form its format (or block size) $k \times l$. The next abbreviation is used for block matrix: $A = [A_{ij}]$, where the symbol A_{ij} denotes the block, i.e. a matrix located on the i-th row and j-th column.

Definition. Combining the components of the matrix into blocks is called

grouping, the reverse operation is a ***deployment***.

The purpose of grouping is to reduce the real size of the matrix and, as a consequence, to simplify the algebraic operations performed with it.

Two block matrices A and B are equal to each other, if the equality $A_{ij} = B_{ij}$ is valid for all the relevant blocks.

However, a matrix can be divided into blocks in many ways, that is the matrix can be partitioned with block matrices of different sizes.

An example block matrix is presented below:

$$A = \left(\begin{array}{c|ccc} 1 & 2 & 3 & 5 \\ 2 & 0 & 2 & 1 \\ 3 & 2 & 1 & 3 \\ \hline 4 & 3 & 2 & 1 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ 2 & 0 & 2 & 1 \\ \hline 3 & 2 & 1 & 3 \\ 4 & 3 & 2 & 1 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ \hline 2 & 0 & 2 & 1 \\ 3 & 2 & 1 & 3 \\ 4 & 3 & 2 & 1 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 5 \\ \hline 2 & 0 & 2 & 1 \\ 3 & 2 & 1 & 3 \\ \hline 4 & 3 & 2 & 1 \end{array}\right)$$

Let's consider the matrix $A = (5,5,5)$ with the same components. Let the matrices be $C = (5,5)$; $D = (5)$. Let's form block matrices $B = (C,D)$, $L = (D,C)$ and $K = (D,D,D)$. Matrices B and L have the same format but different block sizes, and matrices B, L and K have different formats, while they all are derived from matrix A.

Usually, to obtain a block matrix of the certain dimension, it is divided by a system of parallel lines (vertical and horizontal).

Example 1.1.

$$A = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \\ \hline 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{array}\right) = \begin{pmatrix} B & B \\ B & C \end{pmatrix}, \text{ where } B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},\ C = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}.$$

Square block matrix, which contains square blocks on the leading diagonal, and zero blocks outside of the leading diagonal, is called a ***block-***

diagonal matrix.

An example of a block-diagonal matrix is presented below:

$$\left(\begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ \hline 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{array}\right) = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}.$$

1.3.1 Addition and subtraction of block matrices

If the block matrices A and B have the same dimension and are partitioned in the same way, and A_{ij} and B_{ij} are their corresponding blocks of the same size, then to add (subtract) these matrices it is enough to add (subtract) the corresponding blocks of these matrices, i.e.

$$[A_{ij}] + [B_{ij}] = [A_{ij} + B_{ij}].$$

Remark. If the matrices have different sizes, then the assembly operation cannot be performed that is the addition or subtraction of the block matrices is impossible.

Example 1.2. Calculate: $C = 4A + 3B$, where

$$A = \begin{pmatrix} 3 & -2 & 0 \\ 1 & 2 & -3 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & 0 & 3 \\ -2 & 7 & 1 \end{pmatrix}$$

Solution.

$$\left.\begin{array}{l} 4A = \begin{pmatrix} 12 & -8 & 0 \\ 4 & 8 & -12 \end{pmatrix} \\ 3B = \begin{pmatrix} 12 & 0 & 9 \\ -6 & 21 & 3 \end{pmatrix} \end{array}\right] \Rightarrow \quad C = \begin{pmatrix} 24 & -8 & 9 \\ -2 & 29 & -9 \end{pmatrix}$$

Assume that the matrices are divided into blocks as follows:

$$A = (A_{11}, A_{12}), \qquad A_{11} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \quad A_{12} = \begin{pmatrix} -2 & 0 \\ 2 & -3 \end{pmatrix}$$

$$B = (B_{11}, B_{12}), \qquad B_{11} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \qquad B_{12} = \begin{pmatrix} 0 & 3 \\ 7 & 1 \end{pmatrix}$$

$$C = 4A + 3B = (4A_{11}, 4A_{12}) + (3B_{11}, 3B_{12}) = (4A_{11} + 3B_{11}, 4A_{12} + 3B_{12}) =$$

$$= \left(\begin{matrix} 12+2 \\ 4-6 \end{matrix} \quad \begin{pmatrix} -8 & 0 \\ 8 & -12 \end{pmatrix} + \begin{pmatrix} 0 & 9 \\ 21 & 3 \end{pmatrix} \right) = \begin{pmatrix} 24 & -8 & 9 \\ -2 & 29 & -9 \end{pmatrix}$$

1.3.2 Multiplication of block matrices

Above a multiplication operation was defined for matrices A and B under condition that the number of columns of matrix A is equal to the number of rows of matrix B. Then ***the product*** $A \cdot B$ is called a matrix $C = \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ c_{21} & \cdots & c_{2k} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mk} \end{pmatrix}$,

where $c_{ij} = \sum_{p=1}^{n} a_{ip} b_{pj}$, $A = [a_{ij}]_{\substack{i=1,m \\ j=1,n}}$, $B = [b_{ij}]_{\substack{i=1,n \\ j=1,k}}$, $C = [c_{ij}]_{\substack{i=1,m \\ j=1,k}}$.

In the case of block matrices, we have to state the following rule:

<u>Rule:</u> In order to multiply the block matrix $A = [A_{ij}]_{\substack{i=1,m \\ j=1,n}}$ format $m \times n$ by the block matrix $B = [B_{ij}]_{\substack{i=1,n \\ j=1,k}}$ format $n \times k$ with the corresponding sizes of blocks we should apply the following formula $C = [C_{ij}]_{\substack{i=1,m \\ j=1,k}}$, $\quad C_{ij} = \sum_{p=1}^{n} A_{ip} B_{pj}$

Agreement between block sizes means that all multiplications of the matrices used in these formulas are ***correct***.

<u>Remark</u>. If there is no agreement between the block sizes, then the block matrices need to be expanded and grouped in another way.

Example 1.3. Divide matrices A and B into blocks in different ways and multiply.

$$A=\left(\begin{array}{c|cc}1&1&3\\ \hline 0&2&4\\0&0&5\end{array}\right)\qquad B=\left(\begin{array}{c|cc}3&2&1\\ \hline 0&1&2\\0&0&3\end{array}\right)\qquad A\cdot B=\begin{pmatrix}3&3&12\\0&2&16\\0&0&15\end{pmatrix}$$

Solution.

a) $A=\begin{pmatrix}A_{11}&A_{12}\\A_{21}&A_{22}\end{pmatrix}, A_{11}=1,\quad A_{12}=(1,3),\ A_{21}=\begin{pmatrix}0\\0\end{pmatrix},\ A_{22}=\begin{pmatrix}2&4\\0&5\end{pmatrix},$

$B=\begin{pmatrix}B_{11}&B_{12}\\B_{21}&B_{22}\end{pmatrix},\ B_{11}=3,\quad B_{12}=(2,1),\ B_{21}=\begin{pmatrix}0\\0\end{pmatrix},\ B_{22}=\begin{pmatrix}1&2\\0&3\end{pmatrix}.$

$$A\cdot B=\begin{pmatrix}\underbrace{A_{11}B_{11}+A_{12}B_{21}}_{C_{11}}&\underbrace{A_{11}B_{12}+A_{12}B_{22}}_{C_{12}}\\ \underbrace{A_{21}B_{11}+A_{22}B_{21}}_{C_{21}}&\underbrace{A_{21}B_{12}+A_{22}B_{22}}_{C_{22}}\end{pmatrix}$$

$A_{11}B_{11}=(1)(3)=3,\qquad A_{12}B_{21}=(1\quad 3)\begin{pmatrix}0\\0\end{pmatrix}=0,\qquad \boxed{C_{11}=3}$

$A_{11}B_{12}=(1)(2\quad 1)=(2\quad 1),\ A_{12}B_{22}=(1\quad 3)\begin{pmatrix}1&2\\0&3\end{pmatrix}=(1\quad 11),\boxed{C_{12}=(3\quad 12)}$

$A_{21}B_{11}=\begin{pmatrix}0\\0\end{pmatrix}(3)=\begin{pmatrix}0\\0\end{pmatrix},\qquad A_{22}B_{21}=\begin{pmatrix}2&4\\0&5\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix},\ \boxed{C_{21}=\begin{pmatrix}0\\0\end{pmatrix}}$

$A_{21}B_{12}=\begin{pmatrix}0\\0\end{pmatrix}(2\quad 1)=\begin{pmatrix}0&0\\0&0\end{pmatrix},\ A_{22}B_{22}=\begin{pmatrix}2&4\\0&5\end{pmatrix}\begin{pmatrix}1&2\\0&3\end{pmatrix}=\begin{pmatrix}2&16\\0&15\end{pmatrix},$

$\boxed{C_{22}=\begin{pmatrix}2&16\\0&15\end{pmatrix}}$

$$A\cdot B=C=\begin{pmatrix}3&3&12\\0&2&16\\0&0&15\end{pmatrix}.$$

b) If the matrices are divided as:

$$A=\left(\begin{array}{cc|c}1&1&3\\0&2&4\\ \hline 0&0&5\end{array}\right),\quad B=\left(\begin{array}{c|cc}3&2&1\\0&1&2\\ \hline 0&0&3\end{array}\right),$$

then

$$A_{11}=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix},\qquad A_{12}=\begin{pmatrix}3\\ 4\end{pmatrix},\qquad A_{21}=\begin{pmatrix}0 & 0\end{pmatrix},\qquad A_{22}=5,$$

$$B_{11}=\begin{pmatrix}3\\ 0\end{pmatrix},\qquad B_{12}=\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix},\qquad B_{21}=0,\qquad B_{22}=\begin{pmatrix}0 & 3\end{pmatrix}.$$

$$C_{11}=A_{11}B_{11}+A_{12}B_{21}=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}\begin{pmatrix}3\\ 0\end{pmatrix}+\begin{pmatrix}3\\ 4\end{pmatrix}\cdot 0=\begin{pmatrix}3\\ 0\end{pmatrix}+\begin{pmatrix}0\\ 0\end{pmatrix}=\begin{pmatrix}3\\ 0\end{pmatrix}$$

$$C_{12}=A_{11}B_{12}+A_{12}B_{22}=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}+\begin{pmatrix}3\\ 4\end{pmatrix}\begin{pmatrix}0 & 3\end{pmatrix}=\begin{pmatrix}3 & 3\\ 2 & 4\end{pmatrix}+\begin{pmatrix}0 & 9\\ 0 & 12\end{pmatrix}=\begin{pmatrix}3 & 12\\ 2 & 16\end{pmatrix}$$

$$C_{21}=A_{21}B_{11}+A_{22}B_{21}=\begin{pmatrix}0 & 0\end{pmatrix}\begin{pmatrix}3\\ 0\end{pmatrix}+5\cdot 0=0$$

$$C_{22}=A_{21}B_{12}+A_{22}B_{22}=\begin{pmatrix}0 & 0\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}+5\cdot\begin{pmatrix}0 & 3\end{pmatrix}=\begin{pmatrix}0 & 0\end{pmatrix}+\begin{pmatrix}0 & 15\end{pmatrix}=\begin{pmatrix}0 & 15\end{pmatrix}$$

$$C=\begin{pmatrix}3 & 3 & 12\\ 0 & 2 & 16\\ 0 & 0 & 15\end{pmatrix}$$

1.4 The Rank of the Matrix and Rank Determination Methods

Definition. ***The rank*** of the matrix A ($\mathrm{Rg}A$) is a maximum order of its nontrivial minors.

The following statement is valid: elementary transformations of a matrix do not change its rank.

Definition. ***The elementary transformations*** are called the following ones:

1. Multiplication of any row (column) of the matrix by a non-zero number.
2. Addition of any row (column), previously multiplied by the non-zero number, to another row (column).
3. Interchanging two rows (columns).

4. Elimination of zero rows (columns) and elimination of the duplicate or proportional rows (columns) leaving only one of them.

Two matrices A and B are said to be ***similar*** or ***equivalent*** if there exist elementary transforms such that the matrix A follows from the matrix B and vice versa. Equivalent matrices have the same ranks.

The rank of a matrix can be found using two methods. The easiest of these methods is “converting matrix into row echelon form”.

i. Converting the matrix into row echelon form.

To find $\mathrm{Rg}A$ we can reduce the matrix to row echelon form, i.e. to the matrix that meets the following requirements:

- the first non-zero number from the left (“***leading coefficient***“ or “***pivot***”) is always to the right of the first non-zero number in the row above;

- rows consisting of all zeros are at the bottom of the matrix.

The matrix converted into row echelon form looks like this

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1r+1} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2r} & a_{2r+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr} & a_{rr+1} & \dots & a_{rn} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix},$$

where r leading coefficients a_{11}, a_{22}, ..., $a_{rr} \neq 0$.

Thus, the rank of the matrix is a number of the non-zero leading coefficients, i.e. $RgA = r$.

Example 1.4.

$$\begin{pmatrix} -1 & 2 & 1 \\ 3 & 2 & 4 \\ 2 & 4 & 6 \\ 4 & 0 & 4 \\ 0 & -5 & -5 \end{pmatrix} \sim \left\| \begin{matrix} r_4/4 \Rightarrow r_4 \\ r_5/(-5) \Rightarrow r_5 \end{matrix} \right\| \sim \begin{pmatrix} -1 & 2 & 1 \\ 3 & 2 & 4 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \sim \left\| \begin{matrix} r_4 \leftrightarrow r_1 \\ r_5 \leftrightarrow r_2 \\ r_1 + r_4 \Rightarrow r_3 \\ r_2 - 3r_4 \Rightarrow r_5 \\ r_3 - r_4 \Rightarrow r_4 \end{matrix} \right\| \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix} \sim$$

$$\sim \left\| \begin{matrix} r_3/2 \Rightarrow r_3 \\ r_4 - r_3 \Rightarrow r_4 \\ r_5 - r_3 \Rightarrow r_5 \end{matrix} \right\| \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \sim \left\| \begin{matrix} r_3 - r_2 \Rightarrow r_3 \\ r_5 \leftrightarrow r_3 \end{matrix} \right\| \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$\boxed{Rg\, A = 3}$$

ii. Using the method of fringing minors.

Theorem. Let the matrix A have a minor M of order r $(r \neq 0)$, and all minorities $(r+1)$-th order, fringing M, are zero, then the rank of the matrix A is equal to r.

Proof. By all bordering minors $(r+1)$-th order are zero, then, by the theorem on the base minor, all columns of the matrix are a linear combination of its "r" columns. That is, the maximum number of linearly independent columns is equal to $r \Rightarrow \mathrm{Rg}A = r$ ■

Example 1.5 Calculate the rank using the fringing minor's method

$$A = \begin{bmatrix} 2 & -4 & 3 & -3 & 5 \\ 1 & -2 & 1 & 5 & 3 \\ 1 & -2 & 4 & -34 & 0 \end{bmatrix}$$

One can notice that that minor $\begin{vmatrix} -4 & 3 \\ -2 & 1 \end{vmatrix} \neq 0$. Let's calculate its fringing minors. We can first interchange the 1st and 3rd columns. Then a non-zero minor will be in the upper left corner, and the resulting matrix B will be equivalent to A

($B \sim A$).

$$B = \begin{bmatrix} 3 & 4 & 3 & 3 & 5 \\ 1 & 2 & 1 & 5 & 3 \\ 4 & 2 & 1 & 34 & 0 \end{bmatrix}$$

Consider the fringing minors. There will be 3 minors:

$$\begin{vmatrix} -3 & -4 & 2 \\ 1 & -2 & 1 \\ 4 & -2 & 1 \end{vmatrix} = 6 - 4 - 16 + 16 - 6 + 4 = 0\,;$$

$$\begin{vmatrix} 3 & -4 & -3 \\ 1 & -2 & 5 \\ 4 & -2 & -34 \end{vmatrix} = 6 \cdot 34 + 6 - 80 - 24 + 30 - 34 \cdot 4 = 2 \cdot 34 + 36 - 104 =$$

$$= 68 + 36 - 104 = 0;$$

$$\begin{vmatrix} 3 & -4 & 5 \\ 1 & -2 & 3 \\ 4 & -2 & 0 \end{vmatrix} = -10 - 48 + 40 + 18 = 0\,.$$

So, $\mathrm{Rg}A = 2$.

1.5. Laplace's theorem

In the course of Algebra, we got acquainted with the concept of minor and algebraic cofactors of the element of a matrix.

The minor of an element a_{ij} is equal to the determinant of the matrix remaining after excluding the *i*-th row and *j*-th column containing this element, and is denoted as M_{ij} .

The algebraic cofactor of an element a_{ij} is a signed minor, i.e. it is defined by the formula: $A_{ij} = (-1)^{i+j} \cdot M_{ij}$.

Now we will generalize the definition of the minor of a matrix element and introduce the concept of the *k*-th order minor of the matrix.

Let's consider a square matrix A. We can choose any k different columns and rows of the given matrix $(k \leq n)$. The components that stand at the intersection of these k rows and k columns will form a matrix of the k-th order. The determinant of this matrix is called ***a minor of the k-th order*** of this matrix A.

The minor of the k -th order will be denoted as

$$m = m_{j_1 j_2 \ldots j_k}^{i_1 i_2 \ldots i_k} ,$$

where the lower subscripts indicate the numbers of the chosen k columns, and the upper ones indicate the numbers of the chosen k rows.

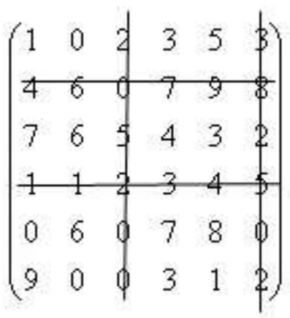
In particular, the minor of the n-th order of the matrix A with n rows and columns, i.e. $m_{12..n}^{12..n}$, is the determinant of this matrix: $m_{12..n}^{12..n} = \det A$.

Each element of the matrix is a minor of the first order.

If we cross out in the given matrix columns and rows which generate a minor of the k-th order, the remaining components form a square matrix of the $(n - k)$ order. The determinant of this $(n - k)$-th order matrix is called ***an additional minor*** to the minor m, which is generated by these k columns and k rows, and is denoted as $M = M_{j_1 j_2 \ldots j_k}^{i_1 i_2 \ldots i_k}$.

In particular, if the original minor of the first order is $m = m_j^i$, that is an element of the matrix a_{ij} , then the additional minor is written as $M=M_j^i=M_{ij}$.

Example 1.6.



$j_1 = 3,\qquad j_2 = 6,$

$i_1 = 2,\qquad i_2 = 4.$

$$m_{36}^{24} = \begin{vmatrix} 0 & 8 \\ 2 & 5 \end{vmatrix}$$

$$M_{36}^{24} = \begin{vmatrix} 1 & 0 & 3 & 5 \\ 7 & 6 & 4 & 3 \\ 0 & 6 & 7 & 8 \\ 9 & 0 & 3 & 1 \end{vmatrix}$$

<u>Definition</u>. ***Algebraic cofactors*** of minor *m* are called its additional minor multiplied by a factor: $(-1)^{i_1+i_2+...+i_k+j_1+j_2+...+j_k}$, i.e.

$$A_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k} = (-1)^{i_1+i_2+...+i_k+j_1+j_2+...+j_k} \cdot M_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k},$$

or

$$A_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k} = (-1)^{p} \cdot M_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k}, \text{ where } p = \sum_{i=1}^{k} (i_i + j_i).$$

Below we present the Laplace's theorem.

It is a rule that allows us to express the determinant of a matrix as a linear combination of determinants of lower order matrices.

Laplace's theorem. The *n*-th order determinant, Δ is equal to the sum of the products of all its minors of the *k*-th order, which are selected on *k* rows, multiplied by their algebraic cofactors.

$$\Delta_n = \sum m_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k} \cdot A_{j_1 j_2 \dots j_k}^{i_1 i_2 \dots i_k}.$$

Similarly, the theorem is formulated in the case of *k* selected columns.

<u>Note</u>. The Laplace's theorem allows reducing the calculation of the determinant of the *n*-th order to the calculation of several determinants of the *k*-th and $(n-k)$-th orders. As the order of the determinant increases, more new determinants appear. Therefore, the Laplace theorem is effective in the case when there exist many zero components in the determinant. Then *k* rows (or columns) can be chosen so that most of the *k*-th order minors located on these

rows will be equal to zero.

Example 1.7. Calculate the determinant:

$$\det A = \begin{vmatrix} -4 & 1 & 2 & -2 & 1 \\ 0 & 3 & 0 & 1 & -5 \\ 2 & -3 & 1 & -3 & 1 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 4 & 0 & 2 & 5 \end{vmatrix}$$

If we choose the 2-nd row and the 5-th one, then all the 2-nd order minors formed by the first and third columns with all the others will be zero. Therefore, it is necessary to take minors that are formed by the second, fourth and fifth columns. That is

$$\det A = \begin{vmatrix} 3 & 1 \\ 4 & 2 \end{vmatrix} \cdot (-1)^{2+5+2+4} \cdot \begin{vmatrix} -4 & 2 & 1 \\ 2 & 1 & 1 \\ -1 & 3 & 0 \end{vmatrix} + \begin{vmatrix} 3 & -5 \\ 4 & 5 \end{vmatrix} \cdot (-1)^{2+5+2+5} \cdot \begin{vmatrix} -4 & 2 & -2 \\ 2 & 1 & -3 \\ -1 & 3 & -1 \end{vmatrix} +$$

$$+ \begin{vmatrix} 1 & -5 \\ 2 & 5 \end{vmatrix} \cdot (-1)^{2+5+4+5} \cdot \begin{vmatrix} -4 & 1 & 2 \\ 2 & -3 & 1 \\ -1 & -1 & 3 \end{vmatrix} = (6-4)(-1)(6-2+1+12) +$$

$$+ 35 \cdot (4-12+6-2-36+4) + 15 \cdot (36-4-1-6-4-6) =$$

$$= -2 \cdot (17) + 35 \cdot (-36) + 15 \cdot 15 = 34 - 36 \cdot 35 + 225 = -1069.$$

Example 1.8. Calculate the determinant:

$$\begin{vmatrix} 0 & 5 & 2 & 0 \\ 8 & 3 & 5 & 4 \\ 7 & 2 & 4 & 1 \\ 0 & 4 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 5 & 2 \\ 4 & 1 \end{vmatrix} \cdot (-1)^{1+4+2+3} \cdot \begin{vmatrix} 8 & 4 \\ 7 & 1 \end{vmatrix} = -3 \cdot (-20) = 60.$$

If the leading diagonal in the determinant Δ is covered by square matrices without common components with determinants Δ_1 and Δ_2, and on one side of them all the components are equal to zero, then such determinant is called a quasi-triangular, and $\Delta = \Delta_1 \cdot \Delta_2$.

Indeed, if

$$\Delta = \begin{vmatrix} \begin{bmatrix} a_{11} & \dots & a_{1k} \\ a_{21} & \dots & a_{2k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{bmatrix} & \begin{matrix} 0 & \dots & 0 \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{matrix} \\ \begin{matrix} a_{k+1,1} & \dots & a_{k+1,k} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nk} \end{matrix} & \begin{bmatrix} a_{k+1,k+1} & \dots & a_{k+1,n} \\ \dots & \dots & \dots \\ a_{n,k+1} & \dots & a_{nn} \end{bmatrix} \end{vmatrix} = \left| \begin{matrix} \text{Use Laplace's theorem ,} \\ \text{choosing first } k \text{ rows} \end{matrix} \right| =$$

$$= \begin{vmatrix} a_{11} & \dots & a_{1k} \\ a_{21} & \dots & a_{2k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \cdot (-1)^{1+2+\dots+k} \begin{vmatrix} a_{k+1,k+1} & \dots & a_{k+1,n} \\ \dots & \dots & \dots \\ a_{n,k+1} & \dots & a_{nn} \end{vmatrix} = \det A_1 \cdot \det A_2 = \Delta_1 \cdot \Delta_2 .$$

Chapter 2. LINEAR SPACES

2.1 Basic Concepts and Examples

Definition. A set of M elements $x, y, z, ...$ of any nature is called a ***linear space*** if:

i. The rule of "addition" is defined. It means that for any two elements $\bar{x}$ and $\bar{y}$ ($\bar{x} \in M \wedge \bar{y} \in M$) there corresponds the third element $\bar{z} \in M$, which is called the ***sum of the elements*** $\bar{x}$ and $\bar{y}$.

ii. The operation of multiplication by a number λ is defined (herewith $\lambda \in R, \lambda \in C$, or $\lambda \in$ another numerical set). It means that $\forall \bar{x} \in M$ and any number λ there corresponds the element $\bar{u} \in M$, which is called ***the product*** $\bar{u} = \lambda \bar{x}$, or $\bar{u} = \bar{x}\lambda$.

iii. The abovementioned two rules are subjected to the following axioms:

1°. Commutativity:

$$\bar{x} + \bar{y} = \bar{y} + \bar{x}, \ \forall \bar{x}, \bar{y} \in M$$

2°. Associativity:

$$\bar{x} + (\bar{y} + \bar{z}) = (\bar{x} + \bar{y}) + \bar{z}, \ \forall \bar{x}, \bar{y}, \bar{z} \in M$$

3°. A zero element exists

$\exists \bar{\theta}$ (zero space element) such that $\bar{x} + \bar{\theta} = \bar{x}, \ \forall \bar{x} \in M$.

4°. An opposite element exists

$\forall \bar{x} \in M \ \exists$ the opposite element $(-\bar{x}) \in M$ such that $\bar{x} + (-\bar{x}) = \bar{\theta}$.

5°. A unity element exists:

$\exists 1$ (unity element) such that $1 \cdot \bar{x} = \bar{x}$.

6°. Distributivity of multiplication with respect to the sum of scalar factors

$$(\lambda + \mu)\bar{x} = \lambda\bar{x} + \mu\bar{x} \quad 7^{o}. \ \lambda(\bar{x} + \bar{y}) = \lambda\bar{x} + \lambda\bar{y}.$$

8°. Distributivity of multiplication with respect to the product of scalar factors

$$\lambda(\mu\bar{x}) = (\lambda\mu)\bar{x}, \ \forall \bar{x} \in M, \ \forall \lambda, \mu \in R.$$

Consequences of the axioms:

1) ***The difference*** $\bar{x}-\bar{y}$ of two elements is called the element $\bar{z}\in M$, such that $\bar{x}=\bar{y}+\bar{z}$. It is easy to see that:

$$\bar{x}-\bar{y}=\bar{x}+(-\bar{y}).$$

Indeed, one can write that

$$\bar{y}+(\bar{x}+(-\bar{y}))=(\bar{y}+\bar{x})+(-\bar{y})=(-\bar{y})+(\bar{y}+\bar{x})=(-\bar{y})+\bar{y}+\bar{x}=$$

$$=(-\bar{y}+\bar{y})+\bar{x}=\bar{\theta}+\bar{x}=\bar{x}.$$

2) The uniqueness of the zero element.

Let's exist two zero elements: $\bar{\theta}_1$ and $\bar{\theta}_2$. Then by definition we have:

$$\bar{x}+\bar{\theta}_1=\bar{x},\ \forall\,\bar{x}\in M$$

$$\bar{x}+\bar{\theta}_2=\bar{x},\ \forall\,\bar{x}\in M$$

Put in the 1-st equation $\bar{x}=\bar{\theta}_2$ and in the second equation $\bar{x}=\bar{\theta}_1$. Then we get:

$$\left.\begin{array}{l}\bar{\theta}_2+\bar{\theta}_1=\bar{\theta}_2\\ \bar{\theta}_1+\bar{\theta}_2=\bar{\theta}_1\end{array}\right\}\quad\Rightarrow\quad\bar{\theta}_1=\bar{\theta}_2.$$

3) The uniqueness of the opposite element.

Let for some element $\bar{x}\in M$ exist two opposite elements $\bar{y}\in M$ and $\bar{z}\in M$. Then,

$$\left.\begin{array}{l}\bar{x}+\bar{y}=\bar{\theta}\\ \bar{x}+\bar{z}=\bar{\theta}\end{array}\right\}\quad\Rightarrow\quad\left.\begin{array}{l}\bar{y}=\bar{\theta}-\bar{x}\\ \bar{z}=\bar{\theta}-\bar{x}\end{array}\right\}\quad\Rightarrow\quad\bar{y}=\bar{z}.$$

4) The zero element of the space is equal to the product of any element $\bar{x}\in M$ by a number «0» that is $0\cdot\bar{x}=\bar{\theta}$

$$0\cdot\bar{x}=(0+0)\cdot\bar{x}=0\cdot\bar{x}+0\cdot\bar{x}$$

Let us add the opposite element «$-0\cdot\bar{x}$» to the left and right parts. Then we have:

$$\bar{\theta}=\bar{\theta}+0\cdot\bar{x}=0\cdot\bar{x}\quad\Rightarrow\quad 0\cdot\bar{x}=\bar{\theta}.$$

5) For any number $\alpha\in R\ (T)$ the product $\alpha\cdot\bar{\theta}=\bar{\theta}$.

Indeed, $\alpha \cdot \overline{\theta} = \alpha(\overline{\theta} + \overline{\theta})$. We can add again the opposite element «$-\alpha\overline{\theta}$» to the left and right part. Then:

$$-\alpha \cdot \overline{\theta} + \alpha \cdot \overline{\theta} = \alpha\overline{\theta} + \alpha\overline{\theta} + (-\alpha\overline{\theta})$$

$$\overline{\theta} = \alpha\overline{\theta} + \overline{\theta} \quad \Rightarrow \quad \overline{\theta} = \alpha\overline{\theta}.$$

6) If the product $\alpha\overline{x} = \overline{\theta}$, then or $\alpha = 0$, or $\overline{x} = \overline{\theta}$.

Indeed, let we have $\alpha \neq 0$, then $\overline{x} = 1 \cdot \overline{x} = \left(\frac{1}{\alpha} \cdot \alpha\right) \cdot \overline{x} = \frac{1}{\alpha}(\alpha\overline{x}) = \frac{1}{\alpha} \cdot \overline{\theta} = \overline{\theta} \Rightarrow$ $\overline{x} = \overline{\theta}$.

7) $\forall \overline{x}$, the product $(-1) \cdot \overline{x}$ is the opposite element to $\overline{x}$, i.e. $(-1) \cdot \overline{x} = -\overline{x}$.

Indeed, $1 \cdot \overline{x} + (-1) \cdot \overline{x} = (1 + (-1)) \cdot \overline{x} = 0 \cdot \overline{x} = \overline{\theta}$.

Some examples of linear spaces and presented below as follows:

Example 2.1**.** The set of all free vectors in three-dimensional space. The operations of addition and multiplication by a scalar have been defined earlier in the course of vector algebra. That is, the addition operation is defined by the parallelogram rule, and multiplication by a scalar λ is defined as the increase (decrease) of the vector length in $|\lambda|$ times. In this case, if $\lambda > 0$, the direction of the vector is preserved, and if $\lambda < 0$ it changes to the opposite.

Similar sets of vectors on the plane R^2 and on the straight line R^1 are also linear spaces.

Often the elements of linear spaces are called ***vectors***, and the linear spaces themselves as ***vector spaces***.

Example 2.2. Suppose that vectors in R^3 are given by the coordinates: $\overline{x} = (x_1, x_2, x_3)$, $\overline{y} = (y_1, y_2, y_3)$. Define the addition operation as $\overline{x} + \overline{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$, and the multiplication operation by a scalar as $\lambda\overline{x} = (\lambda x_1, \lambda x_2, \text{-}\lambda x_3)$.

It is easy to verify that addition and scalar multiplication operations are closed operations in R^3 . Indeed, we can write down

$$\left.\begin{array}{l}\bar{x}=(x_1,x_2,x_3)\\ \bar{y}=(y_1,y_2,y_3)\end{array}\right\} \quad\Rightarrow\quad \begin{array}{c}\bar{x}+\bar{y}=\bar{z}=(x_1+y_1,x_2+y_2,x_3+y_3)\in R^3\\ \lambda\bar{x}=(\lambda x_1,\lambda x_2,-\lambda x_3)\in R^3\end{array}$$

Let us check axioms 1-8.

1. $\bar{x}+\bar{y}=(x_1+y_1,x_2+y_2,x_3+y_3)=(y_1+x_1,y_2+x_2,y_3+x_3)=\bar{y}+\bar{x}$

2. $\bar{x}+(\bar{y}+\bar{z})=(\bar{x}+\bar{y})+\bar{z}$ – obviously.

3. $\exists\bar{\theta}=(0,0,0)$

4. $\exists\ -\bar{x}=(-x_1,-x_2,-x_3)$

5. $1\cdot\bar{x}=(x_1,x_2,-x_3)\neq\bar{x}$

This axiom does not hold, so other axioms can be left unchecked. Thus, this space does not belong to a linear space.

Example 2.3. Is *n*-dimensional space R^n a linear space? The addition and scalar multiplication operations are defined as:

$$\left.\begin{array}{l}\bar{x}=(x_1,x_2,...,x_n)\\ \bar{y}=(y_1,y_2,...,y_n)\end{array}\right\} \qquad \begin{array}{c}\bar{x}+\bar{y}=(x_1+y_1,x_2+y_2,...,x_n+y_n)\\ \lambda\cdot\bar{x}=(\lambda x_1,x_2,...,x_n)\end{array}$$

Solution. The mentioned operations are closed in R^n . Let's check axioms:

It is obviously,

1^o. $\bar{x}+\bar{y}=\bar{y}+\bar{x}$

2^o. $\bar{x}+(\bar{y}+\bar{z})=(\bar{x}+\bar{y})+\bar{z}$

3^o. $\exists\bar{\theta}=(0,0,...,0)$

4^o. $\exists-\bar{x}=(-x_1,-x_2,...,-x_n)$

5^o. $1\cdot\bar{x}=\bar{x}$

6^o. $(\lambda+\mu)\bar{x}=((\lambda+\mu)x_1,x_2,...,x_n)=(\lambda x_1,x_2,...,x_n)+(\mu x_1,0,...,0)=$

$=\lambda\bar{x}+\mu(x_1,0,...,0)\neq\lambda\bar{x}+\mu\bar{x}$

So the given space is not a linear space.

Example 2.4. Consider a set M of real functions that depend on one real variable. These functions are continuous and positive values $\forall t \in [a,b]$. We introduce operations of addition and multiplication by a number as:

$$\bar{x}(t) \oplus \bar{y}(t) = x(t) \cdot y(t)$$

$$\alpha \cdot \bar{x}(t) = [x(t)]^{\alpha}$$

Obviously, these operations are closed in the set M. Let us check the axioms:

1^o. $\quad \bar{x} \oplus \bar{y} = x(t) \cdot y(t) = y(t) \cdot x(t) = \bar{y} \oplus \bar{x}$

2^o. $\quad \bar{x} \oplus (\bar{y} \oplus \bar{z}) = x \cdot (y \cdot z) = (x \cdot y) \cdot z = (\bar{x} \oplus \bar{y}) \oplus \bar{z}$

3^o. $\quad \exists \bar{\theta}, \quad \bar{x} \oplus \bar{\theta} = \bar{x} \quad \Rightarrow \quad (\bar{x} \cdot \bar{\theta}) = \bar{x} \quad \Rightarrow \quad \bar{\theta} = 1$

$f(t) \equiv 1, \ \forall t \in [a,b]$.

4^o. $\quad \exists -\bar{x} = \bar{y}; \ \bar{x} \oplus \bar{y} = \bar{\theta} \quad \Rightarrow \quad \bar{x} \cdot \bar{y} = 1 \quad \Rightarrow \quad \bar{y} = \dfrac{1}{\bar{x}}$.

5^o. $\quad 1 \cdot \bar{x} = \bar{x}; 1 \cdot \bar{x} = x^1 = \bar{x}$

6^o. $\quad (\alpha + \beta)\bar{x} = \bar{x}^{\alpha + \beta} = x^{\alpha} \cdot x^{\beta} = x^{\alpha} \oplus x^{\beta} = \alpha\bar{x} \oplus \beta\bar{x}$

7^o. $\quad \alpha(\bar{x} \oplus \bar{y}) = (\bar{x} \oplus \bar{y})^{\alpha} = (x \cdot y)^{\alpha} = \bar{x}^{\alpha} \cdot \bar{y}^{\alpha} = \alpha\bar{x} \oplus \alpha\bar{y}$

8^o. $\quad \left.\begin{array}{l} \alpha(\beta\bar{x}) = \alpha\left(\bar{x}^{\beta}\right) = x^{\alpha\beta} \\ (\alpha\beta)\bar{x} = \bar{x}^{\alpha\beta} \end{array}\right] \quad \Rightarrow \quad \alpha(\beta\bar{x}) = (\alpha\beta)\bar{x}$

That is, it is a linear space.

Example 2.5. A set $\{P_n(t)\}$ of all algebraic polynomials of degree not exceeding the natural number n is a linear space.

However, the set of all only n -th degree polynomials is not a linear space, since the sum of two such polynomials may have a smaller degree.

2.2. Basis and Dimension of Linear Space

Let a linear space *M* be given on the set of real numbers *R*.

Definition 1. Elements of space *M* are called ***linearly dependent*** if there exist such arbitrary constants $\alpha, \beta, \gamma, ...$ that among of them at least one is nonzero but a linear combination of elements with these constants is zero element of the space *M*, i.e.

$$\alpha\overline{x} + \beta\overline{y} + \gamma\overline{z} + ... + \delta\overline{w} = \overline{\theta}, \quad .\alpha^2 + \beta^2 + \gamma^2 + ... + \delta^2 \neq 0$$

Definition 2. Elements $\overline{x}, \overline{y}, \overline{z}, ..., \overline{w}$ of space *M* are called ***linearly independent*** if their trivial linear combination is possible if and only if all arbitrary constants $\alpha, \beta, \gamma, ...$ are equal to zero, that is $\alpha = \beta = \gamma = ...\delta = 0$. Otherwise, these elements are called ***linearly dependent***.

Theorem. For the elements $\overline{x}, \overline{y}, \overline{z}, ..., \overline{w}$ of the linear space *M* to be linearly dependent it is necessary and sufficient that one of these elements was a linear combination of the others.

Necessity. Let the elements $\overline{x}, \overline{y}, \overline{z}, ..., \overline{w}$ be linearly dependent. This means that $\alpha\overline{x} + \beta\overline{y} + \gamma\overline{z} + ... + \delta\overline{w} = \overline{\theta}$, where at least one of the coefficients, for example, $\alpha \neq 0$. However, then, we can write $\overline{x} = -\frac{\beta}{\alpha}\overline{y} - \frac{\gamma}{\alpha}\overline{z} - ... - \frac{\delta}{\alpha}w$. So, this element is a linear combination of the others.

Sufficiency. Let one of the elements, for example, $\overline{x}$, be a linear combination of the others, i.e. $\overline{x} = \alpha_1\overline{y} + \alpha_2\overline{z} + ... + \alpha_k\overline{w}$, $\Rightarrow$ $1 \cdot \overline{x} - \alpha_1\overline{y} - \alpha_2\overline{z} - ... - \alpha_k\overline{w} = \overline{\theta}$. So, we have the situation when not all the coefficient are equal to zero, indeed the first coefficient is 1, i.e. $"1" \neq 0$. ■

Two elementary statements are true:

1. If among the elements $\bar{x},\bar{y},\bar{z},..,\bar{w}$ is a zero element, then these elements are linearly dependent.

2. If some elements of the set are linearly dependent, then all elements of this set are linearly dependent.

Example 2.6. Examine the linear dependence or linear independence of the set of matrices:

$$A_1 = \begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix}, \ A_2 = \begin{pmatrix} -1 & -2 \\ 2 & 5 \end{pmatrix}, \ A_3 = \begin{pmatrix} 2 & 1 \\ -4 & 2 \end{pmatrix}$$

Solution: Let's suppose that $\alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3 = O$, where O is a null matrix the same order as the matrices *A*. Then,

$$\alpha_1 \begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 & -2 \\ 2 & 5 \end{pmatrix} + \alpha_3 \begin{pmatrix} 2 & 1 \\ -4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} \alpha_1 - \alpha_2 + 2\alpha_3 & -2\alpha_2 + \alpha_3 \\ -2\alpha_1 + 2\alpha_2 - 4\alpha_3 & 3\alpha_1 + 5\alpha_2 + 2\alpha_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Hence, the system of equations with respect to unknown coefficients occurs

$$\begin{cases} \alpha_1 - \alpha_2 + 2\alpha_3 = 0 \\ -2\alpha_2 + \alpha_3 = 0 \\ -2\alpha_1 + 2\alpha_2 - 4\alpha_3 = 0 \\ 3\alpha_1 + 5\alpha_2 + 2\alpha_3 = 0 \end{cases}$$

Find the rank of the system matrix

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & -2 & 1 \\ -2 & 2 & -4 \\ 3 & 5 & 2 \end{pmatrix} \sim \begin{vmatrix} r_3 = r_3 + 2r_1 \\ r_4 = r_4 - 3r_1 \end{vmatrix} \sim \begin{pmatrix} 1 & -1 & 2 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 8 & -4 \end{pmatrix} \sim \left| r_4 = r_4 - 4r_2 \right| \sim \begin{pmatrix} 1 & -1 & 2 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We have 2 non-zero leading elements, i.e. the rank of matrix is 2, which is less than the number of unknown coefficients. Therefore, a non-zero solution exists, so the matrices form the linear dependent set.

Example 2.7. Examine the linear dependence or linear independence of the set of functions:

$$f_1(x) = x\,,\ f_2(x) = \sin x\,,\ f_3(x) = \cos x$$

Solution: Let's suppose that $\alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) = 0$, then, $\alpha_1 x + \alpha_2 \sin x + \alpha_3 \cos x = 0$

If $x = 0$ we have that $\alpha_3 = 0$, then, $\alpha_1 x + \alpha_2 \sin x = 0$.

The differentiation of this equation gets the equality: $\alpha_1 + \alpha_2 \cos x = 0$.

Substituting $x = 0$ and $x = \dfrac{\pi}{2}$ in the equality leads to the system of equations:

$$\begin{cases} \alpha_1 + \alpha_2 = 0 \\ \alpha_1 = 0 \end{cases} \Rightarrow \alpha_1 = 0 \text{ and } \alpha_2 = 0$$

Thereby, the linear combination of the function is equal to zero provided that $\alpha_1 = \alpha_2 = \alpha_3 = 0$, i.e. the functions form a linear independent set.

Definition. A set of linearly independent elements $e_1, e_2, ..., e_n$ of the space M is called ***the basis*** of this space, and any element of this space can be represented as a linear combination of basic elements, i.e. it holds the following equality:

$$\bar{x} = \sum_{i=1}^{n} x_i \bar{e}_i = x_1 \bar{e}_1 + x_2 \bar{e}_2 + ... + x_n \bar{e}_n \qquad (2.1)$$

Equation (2.1) is called ***the decomposition*** of the element $\bar{x}$ ***with respect to the basis*** $\{\bar{e}_i\}_{i=\overline{1,n}}$, and the coefficients x_i, $i = \overline{1,n}$ are called ***the coordinates*** of the element $\bar{x}$ in this base $\{\bar{e}_i\}_{i=\overline{1,n}}$. So any element $\bar{x}$ may be determined by the set of numbers $x_1, x_2, ..., x_n$.

Example 2.8. Prove that the matrix $A=\begin{pmatrix}1 & -2\\ -3 & 4\end{pmatrix}$ in the natural basis E_1, E_2, E_3, E_4 of the linear space $M(2\times2,R)$ has coordinates (1,-2,-3,4)

Solution: Let's compose the linear combination $A=\alpha_1E_1+\alpha_2E_2+\alpha_3E_3+\alpha_4E_4$. Then,

$$\begin{pmatrix}1 & -2\\ -3 & 4\end{pmatrix}=\alpha_1\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}+\alpha_2\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}+\alpha_3\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}+\alpha_4\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}=$$

$$\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}-2\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}-3\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}+4\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}=1E_1-2E_2-3E_3+4E_4$$

Example 2.9. Find coordinates of the vector given by the function $3x^2-2x+2\in R^3[x]$

(a) in the natural basis of the linear space $R^3[x]$

(b) in the basis of functions x^2, $x-1$, 1

Solution: The natural basis is formed by the set of functions: $e_1=1$, $e_2=x$, $e_3=x^2$. Then, the decomposition of the function is as follows:

$$3x^2-2x+2=2\cdot1-2\cdot x+3\cdot x^2=2e_1-2e_2+3e_3,$$

that is the coordinates are (2, -2,3).

Decompose the function in the basis of functions $b_1=x^2$, $b_2=x-1$, b_3 = 1:

$$3x^2-2x+2=3\cdot x^2-2\cdot(x-1)+0\cdot1=3b_1-2b_2+0b_3,$$

that is the coordinates are (3, -2,0).

Definition. A linear space M is called *n-dimensional* if it has n linearly independent elements, and any $(n+1)$ elements are linearly dependent.

In this case, the number "n" is called a ***dimension*** of space and is denoted as

dim M= n.

Further, we will assume that $n < \infty$. Such a vector space is called *finite-dimensional.*

Theorem. If V is a linear space of dimension n, then any n linearly independent elements $e_1, e_2, ..., e_n$ form its basis.

Proof. Let $\{\bar{e}_i\}_{i=\overline{1,n}}$ be any system of n linearly independent vectors of space R, and $\bar{x}$ is any element of R. According to the definition of n-dimensional space, the system of $(n+1)$ vectors is linearly dependent. So

$$\alpha_0 \bar{x} + \sum_{i=1}^{n} \alpha_i \bar{e}_i = \bar{\theta} \tag{2.2}$$

Note that $\alpha_0 \neq 0$, because otherwise the vectors $\{\bar{e}_i\}$ will be linearly dependent. Then it follows from (2.2) that element $\bar{x} = -\sum \frac{\alpha_i}{\alpha_0} \cdot \bar{e}_i$ is a linear combination of $\{\bar{e}_i\}_{i=\overline{1,n}}$. So, system of the elements $\{\bar{e}_i\}$ generates the basis. ■

Theorem. If a linear space V has a basis consisting of n elements, then $\dim V = n$.

Proof. Let $\{\bar{e}_i\}_{i=\overline{1,n}}$ be the basis of the space R. Choose any $(n+1)$ elements of this space $\bar{g}_1, \bar{g}_2, ..., \bar{g}_n, \bar{g}_{n+1}$ and decompose these elements in the basis $\{\bar{e}_i\}_{i=\overline{1,n}}$.

$$\bar{g}_1 = a_{11}\bar{e}_1 + a_{12}\bar{e}_2 + ... + a_{1n}\bar{e}_n$$

$$\bar{g}_2 = a_{21}\bar{e}_1 + a_{22}\bar{e}_2 + ... + a_{2n}\bar{e}_n$$

$$\dots$$

$$\bar{g}_n = a_{n1}\bar{e}_1 + a_{n2}\bar{e}_2 + ... + a_{nn}\bar{e}_n$$

$$\bar{g}_{n+1} = a_{n+1,1}\bar{e}_1 + a_{n+1,2}\bar{e}_2 + ... + a_{n+1,n}\bar{e}_n,$$

where $a_{ij} \in R$. Obviously, the linear relationship $\{\bar{g}_i\}_{i=\overline{1,n+1}}$ is equivalent to the linear dependence of the rows of the matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n+1,1} & a_{n+1,2} & \dots & a_{n+1,n} \end{pmatrix}.$$

But rows of the matrix A are obviously dependent because its $RgA \le n < n+1$. Therefore, at least one of the rows of this matrix will be a linear combination of the others (by the base minor theorem). Hence, the system of elements $\{\bar{g}_i\}_{i=\overline{1,n+1}}$ is linearly dependent. The theorem is proved. ■

2.3. The Transformation of Coordinates with a Change of Basis

Consider a linear space R^n and set of vectors $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$ in this space. To determine whether the system is the basis of this space, it is necessary to construct their linear combination and equate it to the zero element, that is

$$\alpha_1 \cdot \bar{a}_1 + \alpha_2 \cdot \bar{a}_2 + \dots + \alpha_n \cdot \bar{a}_n = \bar{\theta}, \tag{2.3}$$

α_1, α_2, …, α_n are arbitrary constants.

This equality is equivalent to a system of linear algebraic equations (SLAE). Indeed, if all vectors $\{\bar{a}_i\}$ are given by its coordinates

$$\begin{aligned} &\bar{a}_1 = (a_{11} \quad a_{12} \quad \dots \quad a_{1n}), \\ &\bar{a}_2 = (a_{21} \quad a_{22} \quad \dots \quad a_{2n}), \\ &\dots \\ &\bar{a}_n = (a_{n1} \quad a_{n2} \quad \dots \quad a_{nn}), \end{aligned}$$

then, the equality (2.3) can be written in expanded form as:

$$\begin{cases} \alpha_1 a_{11} + \alpha_2 a_{21} + \dots + \alpha_n a_{n1} = 0 \\ \alpha_1 a_{12} + \alpha_2 a_{22} + \dots + \alpha_n a_{n2} = 0 \\ \dots \\ \alpha_1 a_{1n} + \alpha_2 a_{2n} + \dots + \alpha_n a_{nn} = 0 \end{cases}$$

The matrix A of this system is composed of coordinate vectors $\{\bar{a}_i\}$, which are the columns of this matrix A, i.e. $(A \cdot \bar{\alpha} = \bar{0})$.

The system is homogeneous. The rank of the matrix $\mathrm{Rg}A$ coincides with number of linearly independent vectors.

Example 2.10. The system of vectors $\bar{a}_1 = (1,3,5,-1,-2)$, $\bar{a}_2 = (2,-1,-3,4,3)$, $\bar{a}_3 = (5,1,-1,7,4)$, $\bar{a}_4 = (7,7,9,1,3)$ are given in space R^5. Determine the maximum number of linearly independent vectors.

Solution. Let's form a vector equation similar to (2.3) as follows:

$$\alpha_1 \cdot \bar{a}_1 + \alpha_2 \bar{a}_2 + \alpha_3 \bar{a}_3 + \alpha_4 \bar{a}_4 = 0 \tag{2.4}$$

This equation is equivalent to the next system of equations:

$$\begin{cases} \alpha_1 + 2\alpha_2 + 5\alpha_3 + 7\alpha_4 = 0 \\ 3\alpha_1 - \alpha_2 + \alpha_3 + 7\alpha_4 = 0 \\ 5\alpha_1 - 3\alpha_2 - \alpha_3 + 9\alpha_4 = 0 \\ -\alpha_1 + 4\alpha_2 + 7\alpha_3 + \alpha_4 = 0 \\ -2\alpha_1 + 3\alpha_2 + 4\alpha_3 + 3\alpha_4 = 0 \end{cases} \quad \text{or } A \cdot \bar{\alpha} = \bar{0}\text{, where } \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix}.$$

Let us determine the rank of the matrix of this system:

$$A = \begin{pmatrix} 1 & 2 & 5 & 7 \\ 3 & -1 & 1 & 7 \\ 5 & -3 & -1 & 9 \\ -1 & 4 & 7 & 1 \\ -2 & 3 & 4 & 3 \end{pmatrix}$$

Note that the columns of the matrix A coincide with the coordinates of the vectors $\bar{a}_1, \bar{a}_2, \bar{a}_3, \bar{a}_4, \bar{a}_5$. So, further, we will study the linear independence of the system of vectors (i.e. finding the rank of the system of vectors) we need to compose the matrix A, the columns of which coincide with coordinates of the vectors.

Let's perform elementary transformations over rows of the matrix A to convert it into the row echelon form, i.e.

$$\left\|\begin{matrix} r_1\cdot(-3)+r_2\Rightarrow r_2 \\ r_3-5r_1\Rightarrow r_3 \\ r_4+r_1\Rightarrow r_4 \\ r_1\cdot 2+r_5\Rightarrow r_5 \end{matrix}\right\| \sim \begin{pmatrix} 1 & 2 & 5 & 7 \\ 0 & -7 & -14 & -14 \\ 0 & -13 & -26 & -26 \\ 0 & 6 & 12 & 8 \\ 0 & 7 & 14 & 17 \end{pmatrix} \sim \left\|\begin{matrix} r_2/(-7)\Rightarrow r_2 \\ r_3/(-13)\Rightarrow r_3 \\ r_4/2\Rightarrow r_4 \end{matrix}\right\| \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 5 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 2 \\ 0 & 3 & 6 & 4 \\ 0 & 7 & 14 & 17 \end{pmatrix} \sim \left\|\begin{matrix} r_1-2r_2\Rightarrow r_1 \\ r_4-3r_2\Rightarrow r_4 \\ r_5-2r_2\Rightarrow r_5 \end{matrix}\right\| \sim \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 3 \end{pmatrix} \sim \left\|\begin{matrix} r_3/(-2)\Rightarrow r_3 \\ r_4/3\Rightarrow r_4 \\ r_1-r_4\Rightarrow r_1 \end{matrix}\right\| \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad Rg\,A=3$$

Therefore, only 3 vectors are linearly independent, for example, $\bar{a}_1, \bar{a}_2, \bar{a}_4$. Now we can write down the corresponding homogeneous system of linear equations with respect to unknown coefficients:

$$\begin{cases} \alpha_1=-\alpha_3 \\ \alpha_2=-2\alpha_3-2\alpha_4 \\ \alpha_4=0 \end{cases} \quad \Rightarrow \quad \begin{cases} \alpha_1=-\alpha_3 \\ \alpha_2=-2\alpha_3 \\ \alpha_4=0 \end{cases}$$

Thus, the linear combination of the given four vectors takes the form:

$$(-\alpha_3)\bar{a}_1-2\alpha_3\bar{a}_2+\alpha_3\bar{a}_3=0.$$

or

$$a_1+2a_2-a_3=0 \;\Rightarrow\; \boxed{\bar{a}_3=\bar{a}_1+2\bar{a}_2}$$

Example 2.11. Prove that vectors $\bar{g}_1=(1;-3;2)$, $\bar{g}_2=(4;1;1)$ $\bar{g}_3=(2;4;-1)$ form a basis and decompose the vector $\bar{x}=(3;-2;3)$ in this basis.

Solution. Suppose that we have proved that $\{g_i\}_{i=\overline{1,3}}$ is the basis. Then $\bar{x} = (x_1; x_2; x_3) = (3;-2;3)$ can be represented by a vector equality:

$$\bar{x} = \alpha_1\bar{g}_1 + \alpha_2\bar{g}_2 + \alpha_3\bar{g}_3 \qquad (2.5)$$

The equation (2.5) corresponds to the SLAE, which can be written in matrix form as follows

$$G\bar{x}_g = \bar{x}, \qquad (2.6)$$

where G is a matrix whose columns coincide with the coordinates of vectors $\bar{g}_1, \bar{g}_2, \bar{g}_3$, and $\bar{x}$ is a matrix-column, and the vector $\bar{x}_g$ is a vector of unknown coefficients α_i (i = 1,2,3) in (2.5), i.e.

$$\bar{x}_g = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Let's find the solution of the system of equations (2.6) by using the Jordan-Gaussian method:

$$\left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ -3 & 1 & 4 & -2 \\ 2 & 1 & -1 & 3 \end{array}\right) \sim \left\| \begin{array}{l} r_1 \cdot 3 + r_2 \Rightarrow r_2 \\ r_3 - 2r_1 \Rightarrow r_3 \end{array} \right\| \sim \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & 13 & 10 & 7 \\ 0 & -7 & -5 & -3 \end{array}\right) \sim \left\| \begin{array}{l} r_3 \cdot 2 + r_2 \Rightarrow r_2 \\ r_2(-1) \Rightarrow r_3 \end{array} \right\| \sim$$

$$\sim \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & -1 & 0 & 1 \\ 0 & 7 & 5 & 3 \end{array}\right) \sim \left\| r_2 \cdot 7 + r_3 \Rightarrow r_3 \right\| \sim \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 5 & 10 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 4 & 2 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

As the rank of the matrix is $\mathrm{Rg}G = 3$, the vectors $\{g_i\}_{i=\overline{1,3}}$ form the basis. Therefore, following the Gauss method, we can find the unknown coefficients in the form:

$$\alpha_3 = 2;\ \alpha_2 = -1;\ \alpha_1 = 3 - 4\alpha_2 - 2\alpha_3 = 3 + 4 - 4 = 3;\ \alpha_1 = 3.$$

Thereby, the given vector $\bar{x}$ can be decomposed in the basis $\{g_i\}_{i=\overline{1,3}}$ as follows:

$$\overline{x} = 3\overline{g}_1 - \overline{g}_2 + 2\overline{g}_3 .$$

That is, if the basis is given by the system of vectors $G = (\overline{g}_1; \overline{g}_2; \overline{g}_3)$, then the vector $\overline{x}$ in this basis $\overline{x}_g = (\alpha_1; \alpha_2; \alpha_3)$ has coordinates $\overline{x}_g = (3;1;2)$.

In general case, if the basis is given by vectors $\overline{g}_1, \overline{g}_2, ..., \overline{g}_n$, and G is a non-singular square matrix constructed with columns of coordinates of this vectors then this matrix is called ***a matrix of the corresponding basis***. Hence, any vector $\overline{x}$ in this basis can be defined as:

$$\overline{x} = G\overline{x}_G \tag{2.7}$$

If another new basis H is specified in this space such that $\overline{h}_1, \overline{h}_2, ..., \overline{h}_n$ are linear independent vectors, and H is a matrix of the corresponding basis H, then, similarly to (2.7), the same vector $\overline{x}$ can be represented in the new basis in the form:

$$\overline{x} = H\overline{x}_H \tag{2.8}$$

For the same vector, expressions (2.7) and (2.8) are equal. Then, we can write down the transition of the vector coordinates from the "old" basis G to the "new" basis H as follows:

$$H\overline{x}_H = G\overline{x}_G \quad \Rightarrow \quad \overline{x}_H = H^{-1} \cdot G\overline{x}_G \tag{2.9}$$

The last formula determines the relationship between the coordinates of the vector $\overline{x}$ in the "old" and "new" bases.

Example 2.12. Two basis $G = (\overline{g}_1; \overline{g}_2)$ and $H = (\overline{h}_1; \overline{h}_2)$ are given. Coordinates of the basis vectors are $\overline{g}_1 = (3;4)$, $\overline{g}_2 = (-1;2)$, and $\overline{h}_1 = (1;1)$, $\overline{h}_2 = (5;4)$, respectively. Find the coordinates of the vector $\overline{x}$ in the new basis H (i.e. $\overline{x}_H$), if its coordinates in the basis G are $\overline{x}_G = (3;-2)$.

Solution. In accordance with the formula (2.9), we have to find:

$$\bar{x}_H = H^{-1} \cdot G\bar{x}_G$$

Construct the matrices *H* and *G*, respectively, as follows:

$$H = \begin{pmatrix} 1 & 5 \\ 1 & 4 \end{pmatrix}; \qquad \text{and} \quad G = \begin{pmatrix} 3 & -1 \\ 4 & 2 \end{pmatrix}$$

Then, find the inverse matrix H^{-1}:

$$H^{-1} = \frac{1}{-1} \cdot \begin{pmatrix} 4 & -5 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -4 & 5 \\ 1 & -1 \end{pmatrix}$$

By multiplying the matrices, we get

$$H^{-1}G = \begin{pmatrix} -4 & 5 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 14 \\ -1 & -3 \end{pmatrix}$$

Finally, the product of the three multiples gives the required coordinates of the vector in the new basis *H*:

$$\bar{x}_H = H^{-1} \cdot G\bar{x}_G = \begin{pmatrix} 8 & 14 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -4 \\ 3 \end{pmatrix}.$$

Sometimes the relation linking an "old" basis $G = (\bar{g}_1; \bar{g}_2; ...; \bar{g}_n)$ and a "new" basis $H = (\bar{h}_1; \bar{h}_2; ...; \bar{h}_n)$ are known instead of their basis vectors $\{\bar{g}_i\}$ and $\{\bar{h}_i\}$, that is an appropriate system of linear equations is given

$$\begin{cases} \bar{g}_1 = t_{11}\bar{h}_1 + t_{21}\bar{h}_2 + ... + t_{n1}\bar{h}_n \\ \bar{g}_2 = t_{12}\bar{h}_1 + t_{22}\bar{h}_2 + ... + t_{n2}\bar{h}_n \\ ... \\ \bar{g}_n = t_{1n}\bar{h}_1 + t_{2n}\bar{h}_2 + ... + t_{nn}\bar{h}_n \end{cases} \tag{2.10}$$

In the matrix form, the system (2.10) can be rewritten as:

$$G = H \cdot T \tag{2.11}$$

where

$$T = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \dots & \dots & \dots & \dots \\ t_{n1} & t_{n2} & \dots & t_{nn} \end{pmatrix} \tag{2.12}$$

Definition. The matrix T is called a ***transition matrix*** from the basis H (said to be a “new” basis) to the basis G (said to be an “old” basis).

One can notice the following:

1. The above transition matrix T may also be viewed as the matrix whose columns are, respectively, the coordinate column vectors of the “old” basis vectors $\{\overline{g}_i\}$ relative to the “new” basis H; namely,
$$T = \left(\{\overline{g}_1\}_H, \quad \{\overline{g}_2\}_H, \quad \dots, \quad \{\overline{g}_n\}_H\right)$$
2. Analogously, there is a transition matrix C from the “old” basis G to the “new” basis H. Similarly, C may be viewed as the matrix whose columns are, respectively, the coordinate column vectors of the “new” basis vectors $\{\overline{h}_i\}$ relative to the “old” basis G, i.e. $C = \left(\{\overline{h}_1\}_G, \quad \{\overline{h}_2\}_G, \quad \dots, \quad \{\overline{h}_n\}_G\right)$
3. Because the vectors $\overline{g}_1;\overline{g}_2;\dots;\overline{g}_n$ in the “new” basis H are linearly independent, the matrix T is invertible. Similarly, C is invertible due to the same reason for the vectors $\overline{h}_1;\overline{h}_2;\dots;\overline{h}_n$. In fact, we have that if T and C are the above transition matrices, then $C = T^{-1}$.

Taking into account the relation (2.11) we can rewrite the relation (2.9) as follows:

$$\overline{x}_H = \underbrace{H^{-1} \cdot H}_{=Identity\ marrix} \cdot T \cdot \overline{x}_G \Rightarrow \overline{x}_H = T \cdot \overline{x}_G \tag{2.13}$$

Thus, the coordinates of the vector in the "new" basis H via its coordinates in the "old" basis G are computed using the transition matrix T from the “new” basis to the “old” basis.

Example 2.13. The relationship between the two bases G and H in R^3 space is given by the system:

$$\begin{cases} g_1 = 4h_1 - 3h_2 + h_3, \\ g_2 = 2h_1 + 5h_2 - 3h_3, \\ g_3 = 7h_1 + 6h_2 + 2h_3. \end{cases}$$

Find the coordinates of the vector $\bar{x}$ in the basis *H*, if its coordinates in the basis *G* are known as $\bar{x}_G = (-3;-2;1)$.

Solution. Since $\bar{x}_H = T \cdot \bar{x}_G$,

where $T = \begin{pmatrix} 4 & 2 & 7 \\ -3 & 5 & 6 \\ 1 & -3 & 2 \end{pmatrix}$, we can write

$$\bar{x}_H = \begin{pmatrix} 4 & 2 & 7 \\ -3 & 5 & 6 \\ 1 & -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -9 \\ 5 \\ 5 \end{pmatrix}$$

So, $\overrightarrow{x_H} = \{-9;5;5\}$.

Let us consider other types of linear spaces.

Let K be a space of polynomials $\bar{x} = x(t)$ of degree not higher than 4: $\deg x(t) \le 4$, i.e. K is a space of polynomials of the form $x(t) = C + C_1 t + C_2 t^2 + C_3 t^3 + C_4 t^4$, where $C_i, (i = \overline{1,4}) \in R$, (i.e. C_i are the real numbers). Assume that the operations of addition and scalar multiplication are determined in the way usual for the linear space. Then these functions can be considered as vectors of the form: $\bar{x} = \{C_0, C_1, C_2, C_3, C_4\}$, i.e. they are elements of R^5 space, where the role of the basis plays the functions $g_1 = 1$; $g_2 = t$; $g_3 = t^2$; $g_4 = t^3$; and $g_5 = t^4$.

However, as a basis we can also take linearly independent polynomials of the other forms, e.g. $h_1 = 1$; $h_2 = (t-a)$; $h_3 = (t-a)^2$; $h_4 = (t-a)^3$; $h_5 = (t-a)^4$, where *a* is any real constant.

To find the coordinates of the vector $\bar{x}$ in the basis *H*, we can use Taylor's formula in the form of the fourth-order expansion:

$$x(t) = x(a) + \frac{x^{/}(a)}{1!}(t-a) + \frac{x^{//}(a)}{2!}(t-a)^2 + \frac{x^{///}(a)}{3!}(t-a)^3 + \frac{x^{IV}(a)}{4!}(t-a)^4.$$

Example 2.14 A polynomial $\bar{x} = 6 - 5t + t^2 - 3t^3 + 4t^4$ is given. Find the decomposition of this polynomial in the basis $g_1 = 1$; $g_2 = t + 2$; $g_3 = (t+2)^2$; $g_4 = (t+2)^3$; $g_5 = (t+2)^4$.

Solution. We present this polynomial in the form:

$$x(t) = a_1 + a_2(t+2) + a_3(t+2)^2 + a_4(t+2)^3 + a_5(t+2)^4.$$

To find the coefficients of this decomposition we use the Taylor's formula with known constant $a = -2$. Then, we calculate the required values

$$x(-2) = \underbrace{6 + 10 + 4}_{20} - \underbrace{3 \cdot (-8)}_{+24} + \underbrace{4 \cdot 16}_{+64} = 108$$

$$x'(t) = -5 + 2t - 9t^2 + 16t^3\big|_{t=-2} = -5 - 4 - 36 - 16 \cdot 8 = -173$$

$$x''(t) = 2 - 18t + 48t^2\big|_{t=-2} = 2 + 36 + 48 \cdot 4 = 230$$

$$x'''(t) = -18 + 96t\big|_{t=-2} = -18 - 192 = -210$$

$$x^{IV}(t) = 96$$

According to Taylor's formula we obtain that:

$$x(t) = 76 - 173(t+2) + \frac{230}{2}(t+2)^2 - \frac{210}{6}(t+2)^3 + \frac{96}{24}(t+2)^4 =$$

that is the coordinates of the vector in the new basis are

$$= \{76; -173; 115; -35; 4\}.$$

2.4. Subspaces

<u>Definition.</u> A set of elements $L \in R$ is called ***a subspace*** of linear space *R* if it is closed with respect to linear operations:

$1^0. \forall \vec{x}, \vec{y} \in L, \quad \Rightarrow \left(\vec{x} + \vec{y}\right) \in L$

$2^0. \forall \vec{x} \in L$ and $\alpha \in R(C), \quad \Rightarrow \alpha \cdot \vec{x} \in L$

It is easy to verify that the subspace L, which satisfies conditions 1^0 and 2^0, is also a linear space.

Indeed, all the axioms except for the axioms 3^0 and 4^0 are true as they are true $\forall \vec{x} \in R$.

Regarding the axioms 3^0 and 4^0, they follow from the consequences of the axioms:

$$0 \cdot \vec{x} = 0 \text{ and } -1 \cdot \vec{x} = -\vec{x}$$

Indeed using that $\lambda \cdot \vec{x} \in L$, we can choose $\lambda = 0$, then $0 \cdot \vec{x} = \vec{\theta}$, $\theta \in L$ if $\vec{x} \in L$ and if $\lambda = -1$, then one can obtain that $-1 \cdot \vec{x} = -\vec{x}$, $-\vec{x} \in L$ if $\vec{x} \in L$. That is L is a linear space.

<u>The simplest examples of subspaces are the following ones:</u>

1) Zero space, i.e. space that consists of only one $\vec{0}$-th element.

2) The whole space $\vec{R}$ is also a subspace of itself.

Both of these subspaces are called ***improper space***

3) Subset $\{P_n(t)\}$ of all algebraic polynomials of degree not exceeding $n \in N$ is a subspace in linear space $C_{[a,b]}$ of all continuous functions $x(t)$, on the segment $[a,b]$,.

4) Any plane P, which passes through the origin, forms a subspace of three-dimensional space R^3.

Indeed, this plane can be considered as a plane formed by two vectors $\vec{x}_1$ and $\vec{x}_2$, which come from the origin (straight line, if $\vec{x}_1 \parallel \overrightarrow{x_2}$). Obviously that $\forall \alpha_1$ and $\alpha_2 \in R$ vectors belong to the plane, i.e. $\alpha_1 \overrightarrow{x_1} + \alpha_2 \overrightarrow{x_2} \in P$ as shown in

Fig. 2.1.

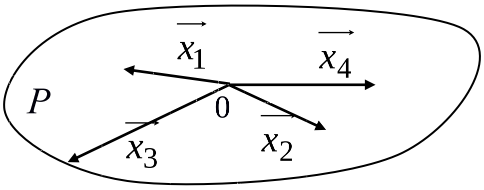


Fig. 2.1

2.5. Linear Spanning Set (Span)

Let's consider a vector space *V* over the field *K* and a set of vectors $\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}$, which belongs to this space *V*. If every vector in *V* can be expressed as a linear combination of $\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}$ then it could be said that this set of vectors $\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}$ form a *linear spanning set* of *V.*

Definition. The set of all linear combinations $\alpha_1 \overrightarrow{x_1} + \alpha_2 \overrightarrow{x_2} + ... \alpha_m \overrightarrow{x_m}$, where $\alpha_i \left(i = \overline{1, m}\right)$ are arbitrary real (or complex) numbers in *K* is called a ***linear spanning set*** (or a ***span***) of elements $\{x_n\}_{i=\overline{1,m}}$ and is denoted as $L\left(\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}\right)$ or $span\left(\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}\right)$.

For a span $L\left(\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}\right)$ the axioms of linear subspaces 1^0 and 2^0 are valid. Thus, any span is a subspace of linear space. Sometimes the span is called a subspace, which is generated by the elements $\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}$.

Definition. ***The dimension*** of a span is equal to the maximum number of linearly independent elements among the elements $\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}$ forming the span $L\left(\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}\right)$.

Definition. If a span $L\left(\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}\right) = V$ then we say that $\overrightarrow{x_1}, \overrightarrow{x_2}, ..., \overrightarrow{x_m}$ spans *V* and we call *V* *finite-dimensional*. A vector space that is not finite-dimensional is called *infinite-dimensional*.

Example 2.15. Consider a subspace which is generated by solutions of homogeneous SLAE with *m* equations and *n* unknowns.

In matrix form, this system is written such as $A\vec{x}=\vec{0}$, where

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

One can show that the set of solutions *M* of the homogeneous SLAE forms a subspace.

Indeed, if $\vec{x}_1 \in M$, i.e. $\overrightarrow{Ax_1}=\vec{0}$ and $\vec{x}_2 \in M$, i.e. $\overrightarrow{Ax_2}=\vec{0}$, then $x_1+x_2 \in M$, because $A(x_1+x_2)=\overrightarrow{Ax_1}+\overrightarrow{Ax_2}=\vec{0}$. Similarly, if $x_1 \in M$, $\overrightarrow{Ax_1}=\vec{0}$, then $\alpha \cdot \vec{x}_1 \in M$, $\forall \alpha \in R$, because $A(\alpha x_1)=\alpha \cdot \overrightarrow{Ax_1}=\vec{0}$.

To find the dimension of this subspace, we have to determine the number *k* of free variables in the system and the rank of the matrix $\mathrm{Rg}A$. Then, according to the Kronecker-Kapelly theorem, $\dim M = k = n - \mathrm{Rg}A$. At the same time, the fundamental system of solutions is a basis of this subspace.

Example 2.16. The subspace *L* is formed by vectors for which the following equations hold: $x_2=-x_1$, $x_5=2x_3$, $x_4=x_2$. Determine the dimension and basis of this subspace.

Solution. To find the dimension and basis of the subspace, we write an appropriate system of homogeneous equations

$$\begin{cases} x_1+x_2=0 \\ 2x_3-x_5=0 \\ x_2-x_4=0 \end{cases}, \text{ or, in the matrix form as } A\vec{x}=0$$

Solving this system of homogeneous equations with the Jordan-Gaussian method we can write:

$$A=\begin{pmatrix}1&1&0&0&0\\0&0&2&0&-1\\0&1&0&-1&0\end{pmatrix}\underset{r_2\leftrightarrow r_3}{\underset{r_1-r_3\to r_1}{}}\sim\begin{pmatrix}1&0&0&1&0\\0&1&0&-1&0\\0&0&2&0&-1\end{pmatrix}\Rightarrow\begin{cases}x_1=-x_4\\x_2=x_4\\2x_3=-x_5\end{cases}$$

Therefore, $\text{Rg}A=3$; and $\dim M=5-3=2$.

Since there are only two free variables, for example, x_4 and x_5. Then the fundamental system of solutions looks like:

	x_1	x_2	x_3	x_4	x_5
$\overrightarrow{e_1}$	-1	1	0	1	0
$\overrightarrow{e_2}$	0	0	-0.5	0	1

That is, the following vectors can be chosen as basis:

$$\overrightarrow{e_1}=\begin{pmatrix}-1\\1\\0\\1\\0\end{pmatrix},\text{ and }\overrightarrow{e_2}=\begin{pmatrix}0\\0\\-0.5\\0\\1\end{pmatrix}.$$

Therefore, any vector $\vec{x}$ (the system solution) can be presented in the form:

$$\vec{x}=C_1\overrightarrow{e_1}+C_2\overrightarrow{e_2}=C_1\begin{pmatrix}-1\\1\\0\\1\\0\end{pmatrix}+C_2\begin{pmatrix}0\\0\\-0.5\\0\\1\end{pmatrix}.$$

Example 2.17. Subspace $L\subset R^5$ is given by the SLAE:

$$\begin{cases}3x_1+x_2-2x_3-5x_4+3x_5=0\\2x_1+3x_2+x_3-2x_4-4x_5=0\\7x_1+7x_2-9x_4-5x_5=0\end{cases}$$

Find the dimension and basis of the subspace.

Solution. Let's build a matrix of this system and determine its rank.

$$A=\begin{pmatrix}3 & 1 & -2 & -5 & 3\\ 2 & 3 & 1 & -2 & -4\\ 7 & 7 & 0 & -9 & -5\end{pmatrix}_{r_1+2r_2\to r_1} \sim \begin{pmatrix}7 & 7 & 0 & -9 & -5\\ 2 & 3 & 1 & -2 & -4\\ 7 & 7 & 0 & -9 & -5\end{pmatrix} \sim$$

$$\sim \begin{pmatrix}7 & 7 & 0 & -9 & -5\\ 2 & 3 & 1 & -2 & -4\end{pmatrix}_{r_1/7\to r_1} \sim \begin{pmatrix}1 & 1 & 0 & \frac{-9}{7} & -\frac{5}{7}\\ 2 & 3 & 1 & -2 & -4\end{pmatrix}_{r_2-2r_1} \sim$$

$$\sim \begin{pmatrix}1 & 1 & 0 & -\frac{9}{7} & -\frac{5}{7}\\ 0 & 1 & 1 & \frac{4}{7} & -\frac{18}{7}\end{pmatrix} \sim \begin{pmatrix}1 & 0 & -1 & -\frac{13}{7} & \frac{13}{7}\\ 0 & 1 & 1 & \frac{4}{7} & -\frac{18}{7}\end{pmatrix}.$$

$\mathrm{Rg}A=2$, as a result, the number k of free variables is equal to $k=5-2=3$. If to choose x_3, x_4 and x_5 as free variables, then the variables x_1 and x_2 take the form:

$$x_2=-x_3-\frac{4}{7}x_4+\frac{18}{7}x_5 \text{ and } x_1=x_3+\frac{13}{7}x_4-\frac{13}{7}x_5.$$

To find the basis vectors (fundamental system of solutions) we give arbitrary values for the free variables x_3, x_4 and x_5 and get the corresponding values for x_1 and x_2 as follows:

	x_1	x_2	x_3	x_4	x_5
$\vec{e_1}$	1	-1	1	0	0
$\vec{e_2}$	13	-4	0	7	0
$\vec{e_3}$	-13	18	0	0	7

Therefore, the vectors $\vec{e_1},\vec{e_2},\vec{e_3}$ form the basis, and any vector the given linear subspace (as a general solution of the homogeneous system) can be presented in the form:

$$\vec{x} = C_1\vec{e_1} + C_2\vec{e_2} + C_3\vec{e_3} = C_1\begin{pmatrix}1\\-1\\1\\0\\0\end{pmatrix} + C_2\begin{pmatrix}13\\-4\\0\\7\\0\end{pmatrix} + C_3\begin{pmatrix}-13\\18\\0\\0\\7\end{pmatrix}.$$

Further, let's consider an inverse problem, i.e. let the subspace be given as a span of vectors. We need to build a SLAE that defines this subspace.

Example 2.18. A subspace $L \subset R^4$ is given as a span formed by the vectors $g_1 = \{1;0;2:-1\}$, $g_2 = \{3;-2;1:0\}$, $g_3 = \{1;-2;-3;2\}$. Write down a SLAE corresponding to this subspace.

Solution. First, determine whether the vectors are linearly independent. For this purpose, we construct the matrix using the coordinates of the vectors written as rows of the matrix. Then, we convert the matrix into a row echelon form:

$$\begin{pmatrix}1 & 0 & 2 & -1\\3 & -2 & 1 & 0\\1 & -2 & -3 & 2\end{pmatrix}_{\substack{r_2-3r_1\\r_3-r_1}} \sim \begin{pmatrix}1 & 0 & 2 & -1\\0 & -2 & -5 & 3\\0 & -2 & -5 & 3\end{pmatrix} \sim \begin{pmatrix}1 & 0 & 2 & -1\\0 & -2 & -5 & 3\end{pmatrix}_{r/-2} \sim$$

$$\sim \begin{pmatrix}1 & 0 & 2 & -1\\0 & 1 & \frac{5}{2} & -\frac{3}{2}\end{pmatrix}$$

Since the rank of the matrix is 2, one can say that among the three vectors forming a linear space, only two are linearly independent. That is, dim L=2.

These two obtained vectors are $\vec{h}_1 = (1,0,2,-1), \vec{h}_2 = \left(0,1,\frac{5}{2},-\frac{3}{2}\right)$, and they can be taken as a basis.

The general solution of the SLAE takes the form:

$$\vec{x} = C_1\vec{h_1} + C_2\vec{h_2} = C_1\begin{pmatrix}1\\0\\2\\-1\end{pmatrix} + C_2\begin{pmatrix}0\\1\\\frac{5}{2}\\-\frac{3}{2}\end{pmatrix} = \begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} \Rightarrow \begin{cases} x_1 = C_1, \\ x_2 = C_2, \\ x_3 = 2C_1 + \frac{5}{2}C_2, \\ x_4 = -C_1 - \frac{3}{2}C_2. \end{cases}$$

Excluding constants C_1 and C_2 from the resulting system, we have:

$$\begin{cases} x_3 = 2x_1 + \frac{5}{2}x_2, \\ x_4 = -x_1 - \frac{3}{2}x_2. \end{cases}$$

Thus, we have the following system of equations:

$$\begin{cases} 2x_1 + \frac{5}{2}x_2 - x_3 = 0 \\ -x_1 - \frac{3}{2}x_2 - x_4 = 0 \end{cases} \text{ or } \begin{cases} 4x_1 + 5x_2 - 2x_3 = 0, \\ 2x_1 + 3x_2 + 2x_4 = 0. \end{cases}$$

2.6. The Sum and Intersection of Subspaces

Definition. Let L_1 and L_2 be two subspaces of linear space *K*. The ***union*** (***sum***) of these subspaces L_1 and L_2 is called the set of all vectors (elements) of the form $\vec{x} + \vec{y}$, where $\vec{x} \in L_1$, $\vec{y} \in L_2$. The sum is denoted as $L_1 \cup L_2$ or $L_1 + L_2$, where $L_1 + L_2 = \{\bar{x} + \bar{y} \mid \bar{x} \in L_1 \text{ and } \bar{y} \in L_2\}$.

Definition. Let L_1 and L_2 be two subspaces of linear space *K*. The ***intersection*** of these subspaces L_1 and L_2 is called the set of all vectors (elements) that belongs to L_1 and L_2 simultaneously. The intersection is denoted as $L_1 \cap L_2$, where $L_1 \cap L_2 = \{\bar{v} \mid \bar{v} \in L_1 \text{ and } \bar{v} \in L_2\}$.

Theorem 1. The intersection $L_1 \cap L_2$ is a linear subspace.

Proof: Let $\vec{v} \in L_1 \cap L_2$, then $\lambda\vec{v} \in L_1$ because L_1 is a linear subspace, similarly $\lambda\vec{v} \in L_2 \Rightarrow \lambda\vec{v} \in (L_1 \cap L_2)$.

Let $\vec{x}, \vec{y} \in L_1 \cap L_2$, then $\vec{x}, \vec{y} \in L_1$ and $\vec{x}, \vec{y} \in L_2$, but then $\vec{x} + \vec{y} \in L_1$ (since L_1 is linear space), $\vec{x} + \vec{y} \in L_2 \Rightarrow \vec{x} + \vec{y} \in L_1 \cap L_2$. ■

A similar theorem is valid for the sum of subspaces.

Theorem 2. The sum $L_1 + L_2$ of linear subspaces is a linear subspace.

Proof. Let $\vec{v} = \vec{x} + \vec{y}$, where $\vec{x} \in L_1$, $\vec{y} \in L_2$, then $\lambda\vec{v} = \lambda\vec{x} + \lambda\vec{y}$, and $\lambda\vec{x} \in L_1$, $\lambda\vec{y} \in L_2$. ■

Theorem 3. The sum of the dimensions of subspaces L_1 and L_2 of a finite-dimensional linear space *R* is equal to the sum of the dimensions of the intersection of these subspaces and the dimension of the sum of these subspaces, i.e.

$$\dim L_1 + \dim L_2 = \dim(L_1 \cap L_2) + \dim(L_1 + L_2),$$

or

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2)$$

Proof. Let us denote intersection of L_1 and L_2 as $L_0 = L_1 \cap L_2$. The sum of L_1 and L_2 let us denote by L ($L = L_1 + L_2$). Suppose that L_0 is *k*-dimensional space. Let us choose the basis in it:

$$e_1,\ e_2, \ldots,\ e_k\,. \tag{2.14}$$

Let a supplement basis (2.14) to the basis in subspace L_1 be as

$$e_1,\ e_2, \ldots,\ e_k,\ g_1,\ g_2, \ldots,\ g_l \tag{2.15}$$

and to the base in subspace L_2 be as

$$e_1,\ e_2, \ldots,\ e_k,\ f_1,\ f_2, \ldots,\ f_m \tag{2.16}$$

Consider a set of elements:

$$g_1,\ g_2,\dots,\ g_l,\ e_1,\ e_2,\dots,\ e_k,\ f_1,\ f_2,\dots,\ f_m \qquad (2.17)$$

We can prove that the elements (2.17) are linearly independent.

Assume that some linear combination of elements (2.17) is a trivial:

$$\alpha_1 g_1 + \alpha_2 g_2 + \dots + \alpha_l g_l + \beta_1 e_1 + \dots + \beta_k e_k + \gamma_1 f_1 + \dots + \gamma_m f_m = 0, \qquad (2.18)$$

or

$$\alpha_1 g_1 + \alpha_2 g_2 + \dots + \alpha_l g_l + \beta_1 e_1 + \dots + \beta_k e_k = -\gamma_1 f_1 - \dots - \gamma_m f_m. \qquad (2.19)$$

The both left and right parts of the equation (2.19) belong to the intersection L_0 of the subspaces L_1 and L_2 because the left part is an element of L_1, and the right part is an element of L_2. However, the right-hand side of (2.19) is a linear combination of elements (2.15), i.e. there are arbitrary numbers $\lambda_1,\ \lambda_2,\dots,\ \lambda_k$ such that

$$-\gamma_1 f_1 - \gamma_2 f_2 - \dots - \gamma_m f_m = \lambda_1 e_1 + \dots + \lambda_k e_k \qquad (2.20)$$

Due to the linear independence of the basic elements (2.16), the equality (2.20) is possible if and only if all the coefficients $\gamma_1, \gamma_2, \dots \gamma_m, \lambda_1, \dots, \lambda_k$ equal to zero. Using (2.19), we get that

$$\alpha_1 g_1 + \alpha_2 g_2 + \dots + \alpha_l g_l + \beta_1 e_1 + \dots + \beta_k e_k = 0 \qquad (2.21)$$

Due to the linear independence of the basis vectors (2.17), the equality (2.21) is possible if and only if all the coefficients $\alpha_1, \alpha_2, \dots, \alpha_l, \beta_1, \dots \beta_k = 0$.

Thus, we found that the equality (2.19) is possible if and only if all the coefficients $\alpha_1, \dots, \alpha_l, \beta_1, \dots \beta_k, \gamma_1, \dots \gamma_m$ are zero, which proves the linear independence of the elements (2.17).

We have proven that any element x of the sum L is some linear combination of elements (2.17). Indeed, consider the element $\vec{x}$, which is represented as $\vec{x} = \sum_{i=1}^{l} \alpha_i g_i + \sum_{j=1}^{k} \beta_j e_j + \sum_{k=1}^{m} \gamma_k f_k$. The first two terms coincide with the element $x_1 \in L_1$, and the last term is equal to $x_2 \in L_2$. Whence it follows that $x \in L_1 + L_2$. That is $\dim(L_1 + L_2) = l + k + m$. Taking into account that $\dim(L_1 \cap L_2) = k$, $\dim L_1 = l + k$, $\dim L_2 = m + k$, we get the following expression:

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2).$$

The theorem is proved. ■

We have already mentioned that the intersection of two subspaces is all the vectors shared by both. If there are no vectors shared by both subspaces L_1 and L_2, meaning that $L_1 \cap L_2 = \vec{\theta}$, the sum $L_1 + L_2$ takes on a special name.

Definition. The space L is a ***direct sum*** of L_1 and L_2 and is denoted as $L_1 \oplus L_2$, if $L_1 \cap L_2 = \vec{\theta}$. That is an element $\vec{x} \in L$ can only be represented in the form $\vec{x} = \overrightarrow{x_1} + \vec{x}_2$, where $\overrightarrow{x_1} \in L_1$, $\vec{x}_2 \in L_2$.

In this case, theorem 3 leads to the formula:

$$\dim(L_1 \oplus L_2) = \dim L_1 + \dim L_2$$

Chapter 3. LINEAR OPERATORS

3.1. Concept of the Linear Operator

Consider the linear space K over the field of real numbers R.

Definition. Let D be some set of elements $\vec{x}$ of linear space K, i.e. $D \subset K$. If to each element $\vec{x}$ in set D there corresponds a certain element of space K ($\vec{y} \in M \subset K$) in accordance with a rule $\boldsymbol{A}$, then we can say that the ***operator A*** is specified, and it maps the elements from set D into the elements of set M.

The last statement can be written in the form $\boldsymbol{A}\vec{x} = \vec{y}$.

In a particular case, if K is a set of real numbers, then we deal with a real-value function $y = f(x)$. That is, the concept of the linear operator $\boldsymbol{A}$ is a generalization of a function definition as drawn in Figure:

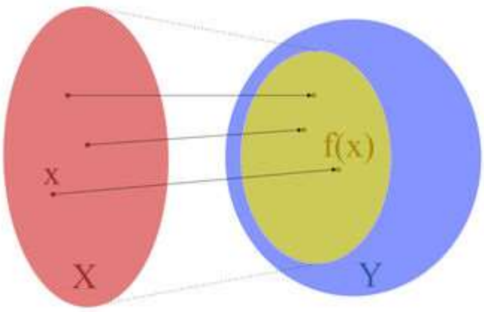


Fig. 3.1.

The set of all elements $\vec{x} \in K$, to which the operator $\boldsymbol{A}$ is applied, is called ***the definition domain*** of the operator and is denoted as D_A. In particular, D_A can coincide with the whole space $\vec{K}$.

The set of all elements $\vec{y} \in M$ is called ***the range of values*** of the operator $\boldsymbol{A}$ and is denoted as Δ_A.

An operator ***A*** is said to be given if:

1. a definition domain D_A is specified
2. a rule (law) according to which $\forall \vec{x} \in D_A$ there corresponds a certain

element (vector) $\vec{y} = \boldsymbol{A}\vec{x}$ is known

Two operators ***A*** and ***B*** are equal if:

1. $D_A = D_B$
2. $\boldsymbol{A}\vec{x} = \boldsymbol{B}\vec{x},\ \forall \vec{x} \in D$

Example 3.1. Let K be some space, and let the operator ***A*** be defined such that $\boldsymbol{A}\vec{x} = \theta,\ \forall \vec{x} \in K$. Such operator is called ***a cancellation or null operator.***

Example 3.2. Let K be an arbitrary space. The operator specified as $\boldsymbol{A}\vec{x} = \vec{x},\ \forall \vec{x} \in K$, is called ***identity operator***.

Example 3.3. If the operator ***A*** is defined as $\boldsymbol{A}\vec{x} = k\vec{x}, \forall \vec{x} \in K,\ k \in R$, then it is called ***the similarity operator***.

That is, the application of operator ***A*** to any element of space stretches (compresses) this vector *k* times.

Example 3.4. Let $K = C_{[a,b]}$ and $\vec{x} = x(t),\ \boldsymbol{A}\vec{x} = x'(t)$. In this case, we will write that $\boldsymbol{A} = \frac{d}{dt}$, and ***A*** is called ***the differentiation operator***.

Next, we consider operators that are given in the whole linear space *L*, and the range of values of the operator is $\Delta_A \subset L$.

Definition. An operator ***A*** given in a linear space is ***a linear operator*** if it satisfies the following conditions:

1. $\boldsymbol{A}(\vec{x_1} + \vec{x_2}) = \boldsymbol{A}\vec{x_1} + \boldsymbol{A}\vec{x_2},\ \forall \vec{x_1}, \vec{x_2} \in L$
2. $\boldsymbol{A}(\alpha\vec{x}) = \alpha\boldsymbol{A}\vec{x},\ \forall x \in L$

For example, we show that the differentiation operator is linear. Indeed, if $\boldsymbol{A}\vec{x} = x'(t)$, then we can write:

1) $\boldsymbol{A}(\vec{x_1} + \vec{x_2}) = (x_1 + x_2)' = x_1' + x_2' = \frac{d}{dt}x_1 + \frac{d}{dt}x_2 = \boldsymbol{A}\vec{x_1} + \boldsymbol{A}\vec{x_2}$

2) $A(\alpha x) = (\alpha x)' = \alpha x' = \alpha \cdot \frac{d}{dt} x = \alpha \cdot A\vec{x}$

3.2. Matrix Representation of the Linear Operator

Let a linear space *L* have dimension *n*, i.e. $\dim L = n$. We choose some basis $e_1, e_2, ..., e_n$ in *L*. Then $\forall \vec{x} \in L$ we have:

$$\vec{x} = \sum_{k=1}^{n} x_k \overrightarrow{e_k} \tag{3.1}$$

Suppose that a linear operator ***A*** is given in this space. We apply this operator to both the sides of (3.1):

$$A\vec{x} = A\left(\sum_{k=1}^{n} x_k \cdot e_k\right) = \sum_{k=1}^{n} x_k \cdot A\overrightarrow{e_k} \tag{3.2}$$

Thus, it follows from (3.2) that to specify the operator ***A*** it is enough to specify its value in the basis of vectors, i.e. $A\overrightarrow{e_1}, A\overrightarrow{e_2}, ..., A\overrightarrow{e_n}$.

Since $\left\{A\overrightarrow{e_i}\right\}_{i=\overline{1,n}}$ are vectors in space *L*, they can be uniquely represented in the form of decomposition in the basis of this space, i.e.

$$\begin{aligned} A\overrightarrow{e_1} &= a_{11}\overrightarrow{e_1} + a_{21}\overrightarrow{e_2} + ... + a_{n1}\overrightarrow{e_n} \\ A\overrightarrow{e_2} &= a_{12}\overrightarrow{e_1} + a_{22}\overrightarrow{e_2} + ... + a_{n2}\overrightarrow{e_n} \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ A\overrightarrow{e_n} &= a_{1n}\overrightarrow{e_1} + a_{2n}\overrightarrow{e_2} + ... + a_{nn}\overrightarrow{e_n} \end{aligned} \tag{3.3}$$

The relations (3.3) can be rewritten in the compact form as

$$A\overrightarrow{e_k} = \sum_{k=1}^{n} a_{ik}\overrightarrow{e_i}, k = \overline{1,n}, \tag{3.4}$$

where a_{ik} are the *i*-th coordinate of the *k*-the vector $A\overrightarrow{e_k}$ in the base $\{e_k\}_{k=\overline{1,n}}$.

Substituting (3.4) into (3.2), we obtain:

$$\boldsymbol{A}\vec{x}=\left(\sum_{k=1}^{n}x_k\sum_{k=1}^{n}a_{\text{ik}}\quad\vec{e_i}\right)=\sum_{i=1}^{n}\left(\sum_{k=1}^{n}x_k a_{\text{ik}}\vec{e_i}\right)=\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{\text{ik}}x_k\vec{e_i}\right). \qquad (3.5)$$

On the other hand, the vector $\boldsymbol{A}\vec{x}$, which is a result of the action of the operator applied to, has coordinates $y_1, y_2,\ldots, y_n$, in the base $\{e_k\}_{k=\overline{1,n}}$ i.e.

$$\boldsymbol{A}\vec{x}=\sum_{i=1}^{n}y_i\vec{e_i}\,. \qquad (3.6)$$

Comparing (3.5) and (3.6), we obtain:

$$\sum_{i=1}^{n}y_i\vec{e_i}=\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{\text{ik}}x_k\right)\vec{e_i}\,. \qquad (3.7)$$

Due to the uniqueness of the decomposition, it follows from (3.7):

$$y_i=\sum_{k=1}^{n}a_{\text{ik}}x_k\,,\forall\, i=\overline{1,n}\,. \qquad (3.8)$$

That is

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ y_2=a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n=a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n \end{cases} \qquad (3.9)$$

The matrix of the system (3.9) is as follows:

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \qquad (3.10)$$

Matrix (3.10) following from (3.9) is called ***a matrix of the linear operator A*** in the basis $\{e_k\}_{k=\overline{1,n}}$.

If matrix (3.10) is known, then, using formulas (3.3), we can find the vectors $\{\boldsymbol{A}\vec{e_i}\}$ and, as a result, we can find the vector $\boldsymbol{A}\vec{x}$ by formula (3.2).

That is, assigning the operator matrix is equivalent to assigning the operator itself. Conversely, by knowing the operator ***A*** and applying it to the basis vectors $\{e_i\}$, we can obtain the matrix of the operator ***A*** using the formula (3.3).

<u>*Conclusion*</u>. If a specific basis is given in a linear space, then, any matrix of the *n*-th order corresponds to a linear operator in this space and vice versa each linear operator in a linear space can be represented by a certain square matrix.

The system of equations (3.9) means that if $\{x_i\}_{i=\overline{1,n}}$ and $\{y_i\}_{i=\overline{1,n}}$ are coordinates of the vectors $\vec{x}$ and $\boldsymbol{A}\vec{x}$ in the basis, respectively, then, the coordinates of the second vector can be obtained by using a linear transformation (LT) with the matrix, which represents the matrix of the linear operator ***A*** in this basis.

That is, the application of a linear operator to a vector implies the application of a linear transformation to its coordinates.

In this respect, the concepts of linear transformation and linear operator are equivalent.

Example 3.5. Let $\boldsymbol{A} = \vec{0}$. Find the matrix of this operator. Since $\boldsymbol{A}\vec{x} = \theta$, we have

$$\begin{aligned}
\boldsymbol{A}\overrightarrow{e_1} &= 0\cdot\overrightarrow{e_1} + 0\cdot\overrightarrow{e_2} + ... + 0\cdot\overrightarrow{e_n},\\
\boldsymbol{A}\overrightarrow{e_2} &= 0\cdot\overrightarrow{e_1} + 0\cdot\overrightarrow{e_2} + ... + 0\cdot\overrightarrow{e_n},\\
&..\\
\boldsymbol{A}\overrightarrow{e_n} &= 0\cdot\overrightarrow{e_1} + 0\cdot\overrightarrow{e_2} + ... + 0\cdot\overrightarrow{e_n}.
\end{aligned}$$

Thus, the matrix of the null operator is a zero-matrix:

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

Example 3.6. Let $\boldsymbol{A} = E$, i.e $\boldsymbol{A}\vec{x} = \vec{x}$:

$$\boldsymbol{A}\overrightarrow{e_1} = \overrightarrow{e_1} = 1\cdot\overrightarrow{e_1} + 0\cdot\overrightarrow{e_2} + ... + 0\cdot\overrightarrow{e_n},$$
$$\boldsymbol{A}\overrightarrow{e_2} = \overrightarrow{e_2} = 0\cdot\overrightarrow{e_1} + 1\cdot\overrightarrow{e_2} + ... + 0\cdot\overrightarrow{e_n},$$
$$...$$
$$\boldsymbol{A}\overrightarrow{e_n} = \vec{e}_n = 0\cdot\overrightarrow{e_1} + 0\cdot\overrightarrow{e_2} + ... + 1\cdot\overrightarrow{e_n},$$

that is, the operator matrix has a form:

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

Thus, the matrix of the identical operator is an identity matrix.

Example 3.7. Let $\boldsymbol{A}$ be the rotation operator in the plane x_1ox_2, which rotates an element by an angle φ. Find the matrix of this operator in the natural basis $\{i, j\}$ (Fig. 3.2).

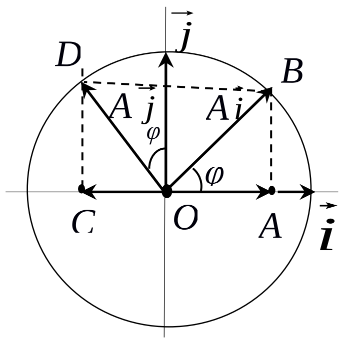


Fig. 3.2

Solution.

$$\vec{i}' = \boldsymbol{A}\vec{i} = \overrightarrow{OA} + \overrightarrow{AB} = \cos\varphi\cdot\vec{i} + \sin\varphi\cdot\vec{j}$$

$$\vec{j}' = \boldsymbol{A}\vec{j} = \overrightarrow{OC} + \overrightarrow{CD} = -\sin\varphi\cdot\vec{i} + \cos\varphi\cdot\vec{j}$$

Therefore, the rotation operator has a matrix in the form: $A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$.

Example 3.8. The operator $\boldsymbol{A}$ is given in space E^3. Geometric vectors are determined by the formula $\boldsymbol{A}\vec{x} = [\vec{a}, \vec{x}] - 3\vec{x}$. Prove the linearity of the operator and construct the matrix of the operator in the natural basis $\vec{i}, \vec{j}, \vec{k}$.

Solution.

1. Let us prove the linearity of this operator. Let $\vec{x}, \vec{y}$ be arbitrary elements of space E^3, then

$$\boldsymbol{A}(\vec{x}+\vec{y})=[\vec{a},\vec{x}+\vec{y}]-3(\vec{x}+\vec{y})=[\vec{a},\vec{x}]-3\vec{x}+[\vec{a},\vec{y}]-3\vec{y}=\boldsymbol{A}\vec{x}+\boldsymbol{A}\vec{y}$$

$$\boldsymbol{A}(\alpha\vec{x})=[\vec{a},\alpha\vec{x}]-3(\alpha\vec{x})=\alpha[a,\vec{x}]-\alpha\cdot 3\vec{x}=\alpha\boldsymbol{A}\vec{x}$$

That is, the given operator is a linear operator.

2. Choose as a vector $\vec{a}$ the following vector: $\vec{a}=\{1,-2,2\}$.

$$\boldsymbol{A}\vec{i}=[\vec{a},\vec{i}]-3\vec{i}=\begin{vmatrix} i & j & k \\ 1 & -2 & 2 \\ 1 & 0 & 0 \end{vmatrix}-3\vec{i}=\{0;2;2\}-3\vec{i}=\{-3;2;2\},$$

$$\boldsymbol{A}\vec{j}=[\vec{a},\vec{j}]-3\vec{j}=\begin{vmatrix} i & j & k \\ 1 & -2 & 2 \\ 0 & 1 & 0 \end{vmatrix}-3\vec{j}=-2\vec{i}-3\vec{j}+\vec{k}=\{-2;-3;1\},$$

$$\boldsymbol{A}\vec{k}=[\vec{a},\vec{k}]-3\vec{k}=\begin{vmatrix} i & j & k \\ 1 & -2 & 2 \\ 0 & 0 & 1 \end{vmatrix}-3\vec{k}=-2\vec{i}-\vec{j}-3\vec{k}=\{-2;-1;-3\}.$$

Then the matrix of the operator ***A*** has the following form:

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

Example 3.9. Find the matrix of the operator in the natural basis that corresponds to the mirror reflection of the point $C(x_1, y_1, z_1)$ with respect to the plane $2x-y+z=0$ (Fig. 3.3).

Solution. From the equation of a given plane $2x-y+z=0$, we obtain the normal vector $\vec{n}=\{2,-1,1\}$.

The equation of a straight line passing through the point C and normal to the plane is $\frac{x-x_1}{2}=\frac{y-y_1}{-1}=\frac{z-z_1}{1}=t$

In the parametric form, these equations are rewritten as:

$$\begin{cases} x = 2t + x_1 \\ y = -t + y_1 \\ z = t + z_1 \end{cases}$$

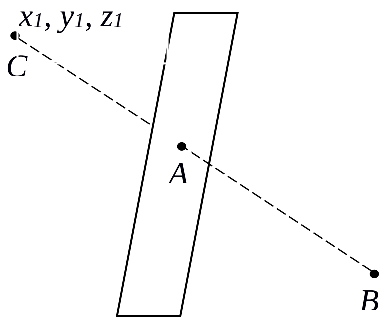


Fig. 3.3

The point A of intersection of the plane and the straight line can be found as follows:

$$4t + 2x_1 + t - y_1 + t + z_1 = 0$$

$$6t = y_1 - 2x_1 - z_1$$

$$t = \frac{1}{6}(y_1 - 2x_1 - z_1)$$

$$x_A = \frac{1}{3}(y_1 - 2x_1 - z_1) + x_1 = \frac{1}{3}x_1 + \frac{1}{3}y_1 - \frac{1}{3}z_1$$

$$y_A = -\frac{1}{6}(y_1 - 2x_1 - z_1) + y_1 = \frac{1}{3}x_1 + \frac{5}{6}y_1 + \frac{1}{6}z_1$$

$$z_A = \frac{1}{6}(y_1 - 2x_1 - z_1) + z_1 = -\frac{1}{3}x_1 + \frac{1}{6}y_1 + \frac{5}{6}z_1$$

Let us find a point B symmetric to point A with respect to the given plane, knowing that A is the middle of the CB (Fig. 3.2):

$$x_A = \frac{x_C + x_B}{2};\ x_B = 2x_A - x_1 = \frac{2}{3}x_1 - x_1 + \frac{2}{3}y_1 - \frac{2}{3}z_1 = -\frac{1}{3}x_1 + \frac{2}{3}y_1 - \frac{2}{3}z_1$$

$$y_B = 2y_A - y_1 = \frac{2}{3}x_1 + \frac{5}{3}y_1 + \frac{1}{3}z_1 - y_1 = \frac{2}{3}x_1 + \frac{2}{3}y_1 + \frac{1}{3}z_1$$

$$z_B = 2z_A - z_1 = -\frac{2}{3}x_1 + \frac{1}{3}y_1 + \frac{5}{3}z_1 - z_1 = -\frac{2}{3}x_1 + \frac{1}{3}y_1 + \frac{2}{3}z_1$$

Therefore,

$$\boldsymbol{A}\vec{x} = \vec{y} = \left(-\frac{1}{3}x_1 + \frac{2}{3}y_1 - \frac{2}{3}z_1; \frac{2}{3}x_1 + \frac{2}{3}y_1 + \frac{1}{3}z_1; -\frac{2}{3}x_1 + \frac{1}{3}y_1 + \frac{2}{3}z_1\right) =$$

$$= \frac{1}{3}(-x_1 + 2y_1 - 2z_1; 2x_1 + 2y_1 + z_1; -2x_1 + y_1 + 2z_1).$$

To find the matrix of this operator in natural basis $\{\vec{i}, \vec{j}, \vec{k}\}$, we find

$\boldsymbol{A}\vec{i} = \frac{1}{3}(-1;2;-2)$, $\boldsymbol{A}\vec{j} = \frac{1}{3}(2;2;1)$, $\boldsymbol{A}\vec{k} = \frac{1}{3}(-2;1;2)$. Then the matrix of the operator has a form:

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

Example 3.10. Check the linearity and compose the matrix of the differential operator $\boldsymbol{A}$ given in the space of polynomials $P(t)$. The degree of polynomials is $P(t) \leq 2$, and the basis is formed by a set of functions: $1, (t-1), (t-1)^2$. The operator is defined in the form: $\boldsymbol{A}P(t) = t^2 \cdot P''(t) + 3P(t)$.

Solution. The polynomial is not higher than the second degree as a result it has a general form $P(t) = a_0 + a_1 t + a_2 t^2$.

First, we prove the linearity of this operator.

Let $P_1(t), P_2(t) \in L$, then:

a) $\boldsymbol{A}(P_1 + P_2) = t^2 \cdot (P_1 + P_2)^{//} + 3(P_1 + P_2) = t^2 P_1^{//} + t^2 P_2^{//} + 3P_1 + 3P_2 =$

$$= \underbrace{t^2 P_1^{//} + 3P_1(t)}_{\boldsymbol{A}P_1(t)} + \underbrace{t^2 P_2^{//} + 3P_2(t)}_{\boldsymbol{A}P_2(t)} = \boldsymbol{A}P_1 + \boldsymbol{A}P_2$$

b) $\boldsymbol{A}(\alpha P(t)) = t^2 \cdot (\alpha P(t))^{//} + 3(\alpha P(t)) = \alpha\left(t^2 P^{//}(t) + 3P(t)\right) = \alpha \cdot \boldsymbol{A}P(t)$

Indeed, the operator is linear.

To obtain the matrix A of the operator in the basis $1, (t-1), (t-1)^2$, we apply this operator to each basis vector:

$$\boldsymbol{A}(1) = t^2 \cdot 0 + 3 = 3 = \{3;0;0\}$$

$$\boldsymbol{A}(t-1) = t^2 \cdot 0 + 3t = 3 = \{0;3;0\}$$

$$\boldsymbol{A}(t-1)^2 = t^2\cdot 2 + 3(t-1)^2 = \underbrace{2\left(t^2-2t+1\right)}_{(t-1)^2} + 4t - 2 + 3(t-1)^2 =$$

$$= 5(t-1)^2 + 4(t-1) + 2 = \{2;4;5\}.$$

Therefore, the operator matrix has a form:

$$A = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{pmatrix}.$$

Example 3.11. Construct a matrix of an integral operator acting in the space *L*, given by a set of the functions: $1, \cos t, \sin t$. Check its linearity if the operator is given as

$$\boldsymbol{A}x(t) = \int_0^{\pi/2} \sin(t+4u)\cdot x(u)du$$

Herein *u* is an integration variable.

Solution. Let's prove the linearity of the operator.

a) $$\boldsymbol{A}\left(\vec{x}(t)+\vec{y}(t)\right) = \int_0^{\pi/2} \sin(t+4u)\cdot\left(\vec{x}(u)+\vec{y}(u)\right)du = \int_0^{\pi/2} \sin(t+4u)\cdot\vec{x}(u)du +$$

$$+\int_0^{\pi/2} \sin(t+4u)\cdot\vec{y}(u)du = \boldsymbol{A}\vec{x}(t)+\boldsymbol{A}\vec{y}(t)$$

b) $$\boldsymbol{A}\left(\alpha\cdot\vec{x}(t)\right) = \alpha\int_0^{\pi/2} \sin(t+4u)\cdot\vec{x}(u)du = \alpha\cdot\boldsymbol{A}\vec{x}(t)$$

To find the operator matrix, we apply operator to the basis vectors $1, \cos t, \sin t$.

$$\boldsymbol{A}\cdot 1 = \int_0^{\pi/2} \sin(t+4u)du = -\frac{1}{4}\cos(t+4u)\Big|_0^{\frac{\pi}{2}} = -\frac{1}{4}(\cos t - \cos t) = 0 = \{0;0;0\}.$$

$$\boldsymbol{A}\cdot\cos t = \int_0^{\pi/2} \sin(t+4u)\cdot\cos u du = \frac{1}{2}\int_0^{\pi/2} \left(\sin(t+5u)+\sin(t+3u)\right)du =$$

$$=\frac{1}{2}\left(-\frac{1}{5}\cos(t+5u)-\frac{1}{3}\cos(t+3u)\right)\Bigg|_0^{\frac{\pi}{2}}=-\frac{1}{2}\left(\frac{1}{5}\cos\left(t+\frac{5}{2}\pi\right)-\frac{1}{2}\cos t+\right.$$

$$\left.+\frac{1}{3}\cos\left(t+\frac{3}{2}\pi\right)-\frac{1}{3}\cos t\right)=\frac{1}{2}\left(\frac{5}{6}\cos t-\frac{1}{5}\sin t+\frac{1}{3}\sin t\right)=-\frac{1}{2}\left(\cos t+\frac{2}{15}\sin t\right)=$$

$$=\left\{0;-\frac{1}{2};-\frac{1}{15}\right\}.$$

$$\boldsymbol{A}\cdot\sin t=\int_0^{\pi/2}\sin(t+4u)\cdot\sin u\,du=\frac{1}{2}\int_0^{\pi/2}(\cos(\alpha+\beta)+\cos(\alpha-\beta))du=$$

$$=\frac{1}{2}\int_0^{\pi/2}(\cos(t+5u)+\cos(t+3u))du=\frac{1}{2}\left(\frac{1}{5}\sin(t+5u)+\frac{1}{3}\sin(t+3u)\right)\Bigg|_0^{\frac{\pi}{2}}=$$

$$=\frac{1}{2}\left(\frac{1}{5}\sin\left(t+\frac{\pi}{2}\right)-\frac{1}{5}\sin t-\frac{2}{3}\cos t\right)=\frac{1}{2}\left(\frac{1}{5}\cos t-\frac{1}{5}\sin t-\frac{2}{3}\cos t\right)=$$

$$=\frac{1}{2}\left(\frac{7}{15}\cos t-\frac{1}{5}\sin t\right)=\left\{0;-\frac{7}{30};\frac{1}{10}\right\}.$$

Thus, the matrix of the operator has a form:

$$A=\begin{pmatrix}0 & 0 & 0\\ 0 & -\frac{1}{2} & -\frac{7}{30}\\ 0 & -\frac{1}{15} & \frac{1}{10}\end{pmatrix}.$$

Example 3.12. Verify that the function f is a linear operator in a linear space R^2, if for any $x=(x_1,x_2)\in R^2$, the function is given by relation

$$f(x)=(x_2-x_1,3x_1+x_2).$$

Solution.

a) Let's take two different elements of the space $x=(x_1,x_2)$ and $y=(y_1,y_2)\in R^2$. Then

$$f(x+y)= f((x_1+y_1),(x_2+y_2))=$$

$$=((x_2+y_2)-(x_1+y_1),3(x_1+y_1)+(x_2+y_2))=$$

$$=(x_2+y_2-x_1-y_1,3x_1+3y_1+x_2+y_2);$$

$$f(x)+ f(y)=(x_2-x_1,3x_1+x_2)+(y_2-y_1,3y_1+y_2)=$$

$$=(x_2-x_1+y_2-y_1,3x_1+x_2+3y_1+y_2),$$

so, we have that $f(x+y)= f(x)+ f(y)$.

Further, $f(\alpha x)= f((\alpha x_1,\alpha x_2))=(\alpha x_2-\alpha x_1,3\alpha x_1+\alpha x_2)$,

$$\alpha f(x)=\alpha(x_2-x_1,3x_1+x_2)=(\alpha x_2-\alpha x_1,3\alpha x_1+\alpha x_2),$$

so,

$$f(\alpha x)=\alpha f(x).$$

Thus, *f* is a linear operator in the linear space R^2.

Example 3.13. The linear operator ***A*** has a matrix in the natural basis of the linear space of polynomials $P(t)$, which are not higher than the first order: $A=\begin{pmatrix}2 & 4\\ 3 & 1\end{pmatrix}$, $g(t)=t-3$. Find the operator $\boldsymbol{A}g(t)$.

Solution. The natural basis of the linear space of polynomials $P(t)$ not higher than the first order is presented by the functions (1, t). The vector $g(t)$ in this basis has coordinates $\vec{x}$ = (–3, 1). Then

$$\vec{y}=A\vec{x}=\begin{pmatrix}2 & 4\\ 3 & 1\end{pmatrix}\begin{pmatrix}-3\\ 1\end{pmatrix}=\begin{pmatrix}-2\\ -8\end{pmatrix}.$$

So, $\mathbf{A}g(t)=-2-8t$.

3.3. Matrix Transformation of a Linear Operator with Changing a Basis

Suppose that a linear operator ***A*** is given in an *n*-dimensional linear space.

That is

$$\boldsymbol{A}\vec{x} = \vec{y}.$$

Let's choose the basis $\{g_j\}_{j=\overline{1,n}}$, in which the operator $\boldsymbol{A}$ is represented by the matrix A_G. This matrix is constructed as a result of applying the operator $\boldsymbol{A}$ to the basis vectors $\{g_i\}$ and their subsequent decompositions in the basis $\{g_i\}$, i.e.

$$\begin{aligned}
&\boldsymbol{A}\overrightarrow{g_1} = a_{11}\overrightarrow{g_1} + a_{21}\overrightarrow{g_2} + ... + a_{n1}\overrightarrow{g_n} \\
&\boldsymbol{A}\overrightarrow{g_2} = a_{12}\overrightarrow{g_1} + a_{22}\overrightarrow{g_2} + ... + a_{n2}\overrightarrow{g_n} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\boldsymbol{A}\overrightarrow{g_n} = a_{1n}\overrightarrow{g_1} + a_{2n}\overrightarrow{g_2} + ... + a_{nn}\overrightarrow{g_n}
\end{aligned}.$$

Hence, we have

$$A_G = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Suppose that H is a new basis in this space. We need to find the matrix that represents this operator $\boldsymbol{A}_H$ in the new basis H.

Let's establish the link between the "old" and "new" bases by the following relations:

$$\begin{aligned}
&\overrightarrow{g_1} = t_{11}\overrightarrow{h_1} + t_{21}\overrightarrow{h_2} + ... + t_{n1}\overrightarrow{h_n} \\
&\overrightarrow{g_2} = t_{12}\overrightarrow{h_1} + t_{22}\overrightarrow{h_2} + ... + t_{n2}\overrightarrow{h_n} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\overrightarrow{g_n} = t_{1n}\overrightarrow{h_1} + t_{2n}\overrightarrow{h_2} + ... + t_{nn}\overrightarrow{h_n},
\end{aligned}$$

One can rewrite this system of equations in the matrix form as

$$G = H \cdot T,$$

where $G = (g_1, g_2, ..., g_n)$ and $H = (h_1, h_2, h_3, \cdots h_n)$, and the transition matrix T from basis H to basis G is determined by

$$T = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \dots & \dots & \dots & \dots \\ t_{n1} & t_{n2} & \dots & t_{nn} \end{pmatrix}.$$

Taking into account that

$$\vec{y}_G = \boldsymbol{A}_G \vec{x}_G \text{ , and } \vec{y}_H = \boldsymbol{T} \vec{y}_G, \tag{3.11}$$

we get

$$\vec{y}_H = \boldsymbol{T}\boldsymbol{A}_G \vec{x}_G. \tag{3.12}$$

The vector $\vec{x}_G$ in the basis G is related to the vector $\vec{x}_H$ in the basis H via the transition matrix as follows (2.13):

$$\vec{x}_H = \boldsymbol{T}\vec{x}_G, \ \Rightarrow \ \vec{x}_G = \boldsymbol{T}^{-1}\vec{x}_H \tag{3.13}$$

Substituting (3.13) into (3.12), we obtain the expression

$$\vec{y}_H = \boxed{TA_G T^{-1}} \vec{x}_H .$$

Thereby, the matrix representing the operator with changing a basis (***change-of-basis matrix***) from the "old" basis to the "new" basis is given by the formula:

$$\boxed{A_H = TA_G T^{-1}} \tag{3.14}$$

Example 3.14. Let the linear operator be given in two-dimensional space with the basis $\vec{g}_1 = \{1;-2\}$, $\vec{g}_2 = \{2;-3\}$. The matrix of the operator in this basis is $A_G = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$. Find the operator matrix in the natural basis.

Solution. Following the task conditions, the matrix representing the operator in the basis $\vec{g}_1, \vec{g}_2$ is $G = \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix}$.

The operator matrix in the natural (new) basis of the vectors $\vec{i} = (1;0)$,

$\vec{j} = (0;1)$ takes the form: $H = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

The relationship between the old and new bases is defined as

$$g_1 = \vec{i} - 2\vec{j}, \; g_2 = 2\vec{i} - 3\vec{j},$$

then, the transition matrix is $T = \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix}$.

Find the inverse matrix T^{-1}. Since the determinant of the matrix T is $\det T = 1$, we have $T^{-1} = \begin{pmatrix} -3 & -2 \\ 2 & 1 \end{pmatrix}$.

Finally, in accordance with (3.14), the change-of-basis matrix takes the form:

$$A_H = \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} -3 & -2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & -3 \end{pmatrix}\begin{pmatrix} -3 & -2 \\ 8 & 5 \end{pmatrix} = \begin{pmatrix} 13 & 8 \\ -18 & -11 \end{pmatrix}.$$

Example 3.15. Find the matrix of the operator in the basis *H*, i.e. A_H, if the operator matrix in the basis *G* is known as $A_G = \begin{pmatrix} 6 & -1 \\ -2 & 5 \end{pmatrix}$. The bases $G = (g_1, g_2)$ and $H = (h_1, h_2)$ are given by corresponding vectors $\vec{g}_1 = (3;4)$, $\vec{g}_2 = (-1;2), \vec{h}_1 = (1;1)$, and $\vec{h}_2 = (5;4)$

Solution. Let's find the transition matrix *T* from basis *G* to basis *H*:

$$G = H \cdot T \Rightarrow \boxed{T = H^{-1}G},$$

where the matrices are given as $G = \begin{pmatrix} 3 & -1 \\ 4 & 2 \end{pmatrix}$, $H = \begin{pmatrix} 1 & 5 \\ 1 & 4 \end{pmatrix}$, respectively.

Find the inverse matrix H^{-1}:

$$\det H = -1, \text{ then, } H^{-1} = -\begin{pmatrix} 4 & -5 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -4 & 5 \\ 1 & -1 \end{pmatrix},$$

Thereby,

$$T = H^{-1}G, \Rightarrow T = \begin{pmatrix} -4 & 5 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} -12+20 & 4+10 \\ 3-4 & -1-2 \end{pmatrix} = \begin{pmatrix} 8 & 14 \\ -1 & -3 \end{pmatrix}$$

Similarly, find the inverse matrix T^{-1}:

$$\det T = -10, \text{ then, } T^{-1} = -\frac{1}{10}\begin{pmatrix} -3 & -14 \\ 1 & 8 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 3 & 14 \\ -1 & -8 \end{pmatrix}$$

Following the formula (3.14), the change-of-basis matrix is computed as

$$A_H = TA_GT^{-1}, \Rightarrow$$

$$A_H = \frac{1}{10}\begin{pmatrix} 8 & 14 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 6 & -1 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 3 & 14 \\ -1 & -8 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 8 & 14 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} 19 & 92 \\ -11 & -68 \end{pmatrix} =$$

$$= \frac{1}{10}\begin{pmatrix} 152-154 & 736-952 \\ 14 & 204-92 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} -2 & -216 \\ 14 & 112 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -1 & -108 \\ 7 & 56 \end{pmatrix}.$$

Similarity of the matrices

The matrices A_G and $TA_GT^{-1} = A_H$, where T is the non-singular matrix, represent the same operator in different bases *H* and *G*, respectively.

These matrices A_G and $A_H = T \cdot A_G T^{-1}$ are called ***similar*** matrices or A_H is said to be obtained from A_G by a similarity transformation.

One of the important properties of such matrices is the equality of their determinants. Indeed,

$$\det A_H = \det\left(T \cdot A_G \cdot T^{-1}\right) =$$

$$= \det T \cdot \det A_G \cdot \det T^{-1} = \det T \cdot \frac{1}{\det T} \cdot \det A_G = \det A_G.$$

Thus, the determinant of the operator matrix does not depend on the choice of the basis.

Theorem. Two matrices represent the same linear operator if and only if the matrices are similar.

That is, all the matrix representations of a linear operator ***A*** form an equivalence class of similar matrices.

3.4. Eigenvectors and Eigenvalues of Linear Operators

Suppose that a linear operator ***A*** is given in the linear space *K* over the field of real numbers *R*.

Definition. A nonzero vector $\vec{x} \in K$ is called ***an eigenvector*** of the operator ***A*** with corresponding ***eigenvalue*** λ if the following equation holds:

$$\boldsymbol{A}\vec{x} = \lambda\vec{x} \tag{3.15}$$

It should be noted that a zero vector cannot be an eigenvector, but zero can be an eigenvalue. Also, if zero is an eigenvalue for an operator ***A***, then ***A*** is not a one-to-one mapping.

Example 3.16 Let $\boldsymbol{A} = 0$. Then $\forall x \in K$:

$$0 \cdot \vec{x} = \vec{0}, \text{ i.e.}$$

$$0 \cdot \vec{x} = 0 \cdot \vec{x},$$

that is, the null operator has an "0" eigenvalue:

$$\lambda = 0 \ \ \forall x \in K$$

In this respect, further we will understand that an eigenvector is a nonzero vector $\vec{x} \neq \vec{0}$, such that $\boldsymbol{A}\vec{x} = \lambda\vec{x}$.

Theorem. A set of all eigenvectors corresponding to the same eigenvalue λ forms a subspace *L* of the space *K*.

Proof. Let *L* be the set of all eigenvectors of operator ***A*** with eigenvalue λ. Let's consider two arbitrary vectors $\vec{x}_1$ and $\vec{x}_2$. Then, there exist $\boldsymbol{A}\vec{x}_1 = \lambda\vec{x}_1$ and $\boldsymbol{A}\vec{x}_2 = \lambda\vec{x}_2$. Summarizing them it gives $\boldsymbol{A}\vec{x}_1 + \boldsymbol{A}\vec{x}_2 = \lambda\vec{x}_1 + \lambda\vec{x}_2 = \lambda(\vec{x}_1 + \vec{x}_2)$. Since ***A*** is a linear operator, then $\boldsymbol{A}\vec{x}_1 + \boldsymbol{A}\vec{x}_2 = \boldsymbol{A}(\vec{x}_1 + \vec{x}_2)$. So $\boldsymbol{A}(\vec{x}_1 + \vec{x}_2) = \lambda(\vec{x}_1 + \vec{x}_2)$, i.e. $(\vec{x}_1 + \vec{x}_2) \in L$.

Similarly, we can check the second axiom of subspace. Really, if $\vec{x} \in L$, then $\boldsymbol{A}\vec{x} = \lambda\vec{x}$ and $\forall \alpha \in R$, we have: $\alpha\boldsymbol{A}\vec{x} = \alpha\lambda\vec{x}$, i.e. $\boldsymbol{A}(\alpha\vec{x}) = \lambda(\alpha\vec{x})$, therefore $\alpha\vec{x} \in L$. The theorem is proved. ■

Definition. The subspace of all eigenvectors of operator ***A***, which share the same eigenvalue λ is called an ***eigenspace*** denoted as $E_\lambda(\boldsymbol{A})$

Note. Every linear combination of the eigenvectors with the same eigenvalue λ is an eigenvector of the operator with this eigenvalue. In simple terms, any sum of eigenvectors is again an eigenvector if they share the same eigenvalue.

The number of times that any given root λ_i appears in the collection of eigenvalues is called its *multiplicity*.

Lemma. If ***A*** is a linear operator represented by an $n \times n$ matrix A, then the dimension of the eigenspace $\dim(E_\lambda(\boldsymbol{A})) \leq m$, where λ is an eigenvalue of A of multiplicity m.

Definition. The set of all eigenvalues of operator ***A*** is called a ***spectrum of the operator.***

Eigenvectors and eigenvalues finding problem:

Suppose the basis $\{e_i\}_{i=\overline{1,n}}$ in the space K is give, and the operator matrix in this basis is known $A = \{a_{ij}\}_{i,j=\overline{1,n}}$. If the vector $\vec{x} \in K$ has coordinates $\vec{x} = \{x_1, x_2, ..., x_n\}^T$ in this basis, the coordinates of the vector $\boldsymbol{A}\vec{x}$ can be found and written in the matrix form as follows:

$$\boldsymbol{A}\vec{x} = A \cdot X = \{a_{ij}\}X$$

In accordance with $\boldsymbol{A}\vec{x} = \lambda\vec{x}$, we can write down the following matrix form of the eigenvalue-eigenvector equation:

$$A \cdot X = \lambda I \cdot X \ , \tag{3.16}$$

where I is an identity matrix.

Hence, the equation (3.16) may be written in the form:

$$(A - \lambda I)X = 0 \tag{3.17}$$

or

$$\left(\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = 0$$

Then, we have

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = 0. \tag{3.18}$$

Thereby, to find eigenvectors and eigenvalues of the linear operator, we have to solve the following homogeneous system:

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots + a_{2n}x_n = 0 \\ \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)x_n = 0 \end{cases} \tag{3.19}$$

The homogeneous system (3.19) with n equations and n variables has a nonzero solution if and only if its matrix is singular, i.e. we require that

$$D(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0 \text{ or } \det(A - \lambda I) = 0 \tag{3.20}$$

Equation (3.20) is called ***a characteristic equation***, and the left-hand side of this equation is a polynomial in the variable λ called the ***characteristic polynomial***.

The following theorem is valid.

Theorem. The eigenvalues of a linear operator coincide with the roots of the characteristic polynomial.

Next, we need to find the eigenvectors after the eigenvalues have been computed. For this purpose, we substitute each computed eigenvalue into (3.19) to find a nonzero solution of the system, which is associated with the coordinates of the eigenvector appropriate to this eigenvalue.

Example 3.17. Find eigenvalues and eigenvectors of the linear operator which is presented by a matrix $A = \begin{pmatrix} 6 & 2 \\ 3 & 7 \end{pmatrix}$.

Solution. Write the equation (3.18) for the given matrix *A* and unknown coordinates of the vector $\vec{x} = \{x_1, x_2\}^T$:

$$\begin{pmatrix} 6-\lambda & 2 \\ 3 & 7-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

Then, compose the characteristic equation (3.20)

$$\det(A - \lambda I) = \begin{vmatrix} 6-\lambda & 2 \\ 3 & 7-\lambda \end{vmatrix} = 0,$$

Computing the determinant as usual, the result is

$(6-\lambda)(7-\lambda) - 6 = 0 \Rightarrow 42 - 7\lambda - 6\lambda + \lambda^2 - 6 = 0, \Rightarrow \lambda^2 - 13\lambda + 36 = 0,$

Solving this equation, we find the roots:

$$\lambda_1 = 9, \text{ and } \lambda_2 = 4$$

They are eigenvalues of the operator presented by the matrix *A*.

Now we need to find the basic eigenvectors for each λ. First we will find the eigenvectors for $\lambda_1 = 9$. We wish to find all vectors $X \neq 0$ such that $AX = 9X$. These are the solutions to $(A - 9I)X = 0$, i.e.

$$\begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Rightarrow \begin{cases} -3x_1 + 2x_2 = 0 \\ 3x_1 - 2x_2 = 0 \end{cases}$$

It follows from the solution of the system:

$$3x_1 = 2x_2, \Rightarrow x_2 = \frac{3}{2}x_1.$$

Then, any vector of the form $\vec{x}_1 = (1;3/2)x_1$ is the eigenvector corresponding to the eigenvalue $\lambda_1 = 9$. If we assign $x_1 = 2$, the eigenvector is $\vec{x}_1 = (2;3)$.

Analogously, for the second eigenvalue $\lambda_2 = 4$. If follows from the solution of $(A - 4I)X = 0$:

$$\begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0, \quad \Rightarrow 2x_1 + 2x_2 = 0, \Rightarrow x_1 = -x_2.$$

Thus, any vector of the form $\vec{x}_2 = (-1;1)x_2$ is the eigenvector, e.g. at $x_2 = 1$ $\vec{x}_2 = (-1;1)$ is the eigenvector corresponding to the eigenvalue $\lambda_2 = 4$.

Note. If the matrix operator A is an $n \times n$ matrix, then the characteristic polynomial of the operator $\boldsymbol{A}$ will have degree n. Since, the characteristic equation (3.20) is a polynomial of the n-th degree with respect to λ, let's denote it as $P(\lambda)$.

It should be noticed that finding the eigenvalues can be computationally challenging and could be done using a computer in most cases. In addition, the roots of characteristic polynomials can be both real and complex.

According to the basic theorem of algebra, any polynomial of the n-th degree has at least one root, which is either real or complex. Thereby, we can notice that

1. A linear operator in a complex finite-dimensional linear space has always

at least one eigenvector;

2. If a linear operator is defined in a real linear space, then the polynomial $P(\lambda)$ has real coefficients. Moreover, if *n* is odd, then the characteristic equation has at least one real root, as a result, the linear operator has at least one eigenvector.

3. A linear operator in an *n*-dimensional space cannot have more than *n* various eigenvalues, because the characteristic equation is *n*-th degree.

4. Eigenvectors $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n$ of linear operator $\boldsymbol{A}$ corresponding to pairwise distinct eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ are linearly independent.

Proof. (We use the method of mathematical induction).

Obviously, for $m = 1$ the statement is true. Suppose that it is valid for $(m - 1)$ eigenvectors of the operator $\boldsymbol{A}$, and check it for m eigenvectors.

Let's assume the opposite. Let m eigenvectors be linearly dependent, i.e.

$$\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2 + ... + \alpha_m \vec{x}_m = \vec{0}, \qquad (3.21)$$

and at least one of the coefficients in (3.21) is not equal to zero, for example, $\alpha_1 \neq 0$. Applying the operator $\boldsymbol{A}$ to (3.21), we get for each term:

$$\boldsymbol{A}(\alpha_1 \vec{x}_1) = \lambda_1(\alpha_1 \vec{x}_1),\ \boldsymbol{A}(\alpha_2 \vec{x}_2) = \lambda_2(\alpha_2 \vec{x}_2),\ \ldots,\ \boldsymbol{A}(\alpha_m \vec{x}_m) = \lambda_m(\alpha_m \vec{x}_m)$$

By the properties of the linear operator, we can write

$$\alpha_1 \lambda_1 \vec{x}_1 + \alpha_2 \lambda_2 \vec{x}_2 + ... + \alpha_m \lambda_m \vec{x}_m = \vec{0}. \qquad (3.22)$$

Multiplying (3.21) by λ_m and subtracting it from (3.22), we get

$$\alpha_1(\lambda_1 - \lambda_m)\vec{x}_1 + \alpha_2(\lambda_2 - \lambda_m)\vec{x}_2 + ... + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)\vec{x}_{m-1} = \vec{0}. \qquad (3.23)$$

Due to the assumption of the statement, (m-1) eigenvectors are linearly independent, that is all the coefficients in (3.23) must be equal to zero. Therefore, $\alpha_1(\lambda_1 - \lambda_m) = 0$, but it contradicts the assumption that $\alpha_1 \neq 0$. So, the assertion is proved.■

5. If all n roots of the characteristic polynomial $\lambda_1, \lambda_2, ..., \lambda_n$ are distinct, then the corresponding eigenvectors $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n$ are linearly independent and they can be taken as a basis of the n-dimensional space K.

A diagonal operator matrix

If we have a basis that consists of eigenvectors of an $n \times n$ matrix A, then the representation of matrix A with respect to that basis is diagonal.

Indeed let's construct an operator matrix in a basis of its eigenvectors $(\vec{x}_1, \vec{x}_2, ..., \vec{x}_n)$. We need to apply consequently the operator to each eigenvector, then, we get

$$\begin{cases} \boldsymbol{A}\vec{x}_1 = \lambda_1 \vec{x}_1 \\ \boldsymbol{A}\vec{x}_2 = \lambda_2 \vec{x}_2 \\ ... \\ \boldsymbol{A}\vec{x}_n = \lambda_n \vec{x}_n \end{cases} \Rightarrow \begin{cases} \boldsymbol{A}x_1 = \lambda_1 \vec{x}_1 + 0 \cdot \vec{x}_2 + ... + 0 \cdot \vec{x}_n \\ \boldsymbol{A}x_2 = 0 \cdot \vec{x}_1 + \lambda_2 \vec{x}_2 + ... + 0 \cdot \vec{x}_n \\ ... \\ \boldsymbol{A}x_n = 0 \cdot \vec{x}_1 + 0 \cdot \vec{x}_2 + ... + \lambda_n \vec{x}_n \end{cases}$$

That is, the operator matrix is diagonal matrix whose diagonal elements are the eigenvalues listed in the same order as the corresponding eigenvectors:

$$A = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_n \end{bmatrix} = \text{diag}\{\lambda_1; \lambda_2; \dots ; \lambda_n\} \qquad (3.21)$$

In this case, the characteristic polynomial of the operator $\boldsymbol{A}$ is calculated easy as a product of the following linear factors:

$$P(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)...(\lambda_n - \lambda).$$

So, a diagonal operator matrix has an advantage from a computational viewpoint.

It should be noticed that any operator matrix A given in any basis $\{\vec{e}_i\}$ can be reduced to a diagonal form. We will do so using the representation of the operator matrix A with respect to a new basis formed by its eigenvectors $\{\vec{x}_i\}$. If the eigenvectors $\{\vec{x}_i\}$ of the operator matrix A have been calculated, then, each basis vector of the "new" basis $\{\vec{x}_i\}$ can be expressed as a linear combination of the vectors of the "old" basis $\{\vec{e}_i\}$ in accordance with (2.11) as follows: $H = GT^{-1}$. Herewith the inverse matrix T^{-1} is a transition matrix C from the "old" basis to the "new" basis, i.e. $T^{-1} = C = \left(\{\vec{x}_1\}_E, \{\vec{x}_2\}_E, ..., \{\vec{x}_n\}_E\right)$.

Thus, the matrix of the operator $\boldsymbol{A}$ in the basis of eigenvectors $\{\vec{x}_i\}$ and that in the given basis $\{\vec{e}_i\}$ are connected by the formula:

$$A_x = C^{-1} A_e C \tag{3.22}$$

For instance, the eigenvectors of the operator $\boldsymbol{A}$ found as $\vec{x}_1 = (2;3)$ and $\vec{x}_2 = (-1;1)$ in the previous example, form the transition matrix C in the form: $C = \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}$. Inverting this matrix gives us $C^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix}$. Finally, the diagonal form of the matrix A in the basis of these eigenvectors is as follows:

$$A_e = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 6 & 2 \\ 3 & 7 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 45 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}.$$

Example 3.17. Linear operator $\boldsymbol{A}$ in a natural basis has a matrix

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

Find the matrix B of this operator in the basis of its eigenvectors, as well as the matrix C of the transition to this basis.

Solution. First of all, find the eigenvalues and eigenvectors of the operator. To this end, we form a characteristic equation:

$$\det(A-\lambda I)=\begin{vmatrix}1-\lambda & 1 & 1\\ 1 & 1-\lambda & 1\\ 1 & 1 & 1-\lambda\end{vmatrix}=0.$$

$$\begin{vmatrix}1-\lambda & 1 & 1\\ 1 & 1-\lambda & 1\\ 1 & 1 & 1-\lambda\end{vmatrix}_{\substack{r_2-r_1\\ r_3-r_2}}=\begin{vmatrix}1-\lambda & 1 & 1\\ \lambda & -\lambda & 0\\ 0 & \lambda & -\lambda\end{vmatrix}=\lambda^2\cdot\begin{vmatrix}1-\lambda & 1 & 1\\ 1 & -1 & 0\\ 0 & 1 & -1\end{vmatrix}_{r_1+r_2}=$$

$$=\lambda^2\cdot\begin{vmatrix}2-\lambda & 0 & 1\\ 1 & 0 & -1\\ 0 & 1 & -1\end{vmatrix}=-\lambda^2\cdot\begin{vmatrix}2-\lambda & 1\\ 1 & -1\end{vmatrix}=-\lambda^2(\lambda-3)$$

$\lambda_{1,2}=0$; $\lambda_3=3$ are eigenvalues of the operator.

Find the corresponding eigenvectors for $\lambda_{1,2}=0$. After substitution eigenvalue into equation (3.17), we obtain: $\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=0$.

We solve the obtained SLAE. Performing an elementary transformations with a matrix (A-λI), we come to the equivalent matrix:

$$\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}\sim\begin{pmatrix}1 & 1 & 1\end{pmatrix}.$$

Thus, the homogeneous system is reduced to the following equation:

$$x_1+x_2+x_3=0,\Rightarrow\ x_1=-(x_2+x_3).$$

Let us find the fundamental system of solutions (FSS). Assigning x_2 and x_3 as free variables, we can compose two eigenvectors:

FSS	x_1	x_2	x_3	
$\vec{e}_1$	-1	1	0	$\rightarrow$ eigenvector
$\vec{e}_2$	-1	0	1	$\rightarrow$ eigenvector

Thus, the corresponding eigenvectors $\vec{e}_1$ and $\vec{e}_2$ are obtained.

Similarly, we find eigenvectors that correspond to eigenvalue $\lambda_3 = 3$. After substitution $\lambda_3 = 3$ into the system of equations (3.17) we obtain:

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

Solving the SLAE:

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}_{\substack{r_2 \to r_1 \\ 2r_2 + r_1 \\ r_3 - r_2}} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}.$$

Thus, the system is reduced to the following one:

$$\begin{cases} x_2 = x_3, \\ x_1 = 2x_2 - x_3. \end{cases} \Rightarrow \begin{cases} x_2 = x_3, \\ x_1 = x_3. \end{cases}$$

The variable x_3 is free. Assigning its value equal to 1, then, we have $x_1 = 1$, $x_2 = 1$. Thus, the eigenvector $\vec{e}_3$ is defined as

$$\vec{e}_3 = (1;1;1).$$

Let's form a matrix from eigenvectors, i.e. a matrix of transition from a natural basis to a basis of eigenvectors:

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

Then, the matrix A in the basis of the eigenvectors is defined as $B = C^{-1}AC$.

$$B = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

One can that the matrix is diagonal with eigenvalues on the leading diagonal.

Also, we can directly build the matrix B without the transition matrix. As described earlier, we need to obtain the vectors $\boldsymbol{A}\vec{e}_1$, $\boldsymbol{A}\vec{e}_2$, $\boldsymbol{A}\vec{e}_3$ as follows:

$$\boldsymbol{A}\vec{e}_1 = \lambda_1 \cdot \vec{e}_1 = 0 \cdot \vec{e}_1 = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3,$$

$$\boldsymbol{A}\vec{e}_2 = \lambda_2 \cdot \vec{e}_2 = 0 \cdot \vec{e}_2 = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3,$$

$$\boldsymbol{A}\vec{e}_3 = \lambda_3 \cdot \vec{e}_3 = 3 \cdot \vec{e}_3 = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 3 \cdot \vec{e}_3,$$

Thereby, we get

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

3.5. Operations with Linear Operators and their Matrices

Let operators $\boldsymbol{A}$ and $\boldsymbol{B}$ be given in the linear space *K*. The operators $\boldsymbol{A}$ and $\boldsymbol{B}$ are called ***equal*** if the following equality holds

$$\boldsymbol{A}\vec{x} = \boldsymbol{B}\vec{x}, \quad \forall \vec{x} \in K.$$

Let $A = \lfloor a_{\text{ij}} \rfloor$ and $B = \lfloor b_{\text{ij}} \rfloor$ are matrices of these operators in some basis $\{\vec{e}_i\}_{i=\overline{1,n}}$. Since, one can present

$$\boldsymbol{A}\vec{e}_j = a_{1j}\vec{e}_1 + a_{2j}\vec{e}_2 + ... + a_{nj}\vec{e}_n,$$

$$\boldsymbol{B}\vec{e}_j = b_{1j}\vec{e}_1 + b_{2j}\vec{e}_2 + ... + b_{nj}\vec{e}_n,$$

and $\boldsymbol{A}\vec{e}_j = \boldsymbol{B}\vec{e}_j$, we get that $a_{ij} = b_{ij}$, $\forall\, i, j = \overline{1,n}$.

That is, the equal operators have the same matrices in the same basis.

<u>Definition</u>. ***The sum*** of two linear operators is a linear operator $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$, which is defined according to the rule:

$$\boldsymbol{C}\vec{x} = (\boldsymbol{A} + \boldsymbol{B})\vec{x} = \boldsymbol{A}\vec{x} + \boldsymbol{B}\vec{x}.$$

It is easy to prove that the operator $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ is linear if both the operators $\boldsymbol{A}$ and $\boldsymbol{B}$ are linear operators.

Let's consider the linear operator matrix $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ in a basis $\{\vec{e}_i\}_{i=\overline{1,n}}$.

Suppose the operator $\boldsymbol{A}$ is represented by a matrix $A = \lfloor a_{ij} \rfloor$, the columns of which are the coordinates of the vectors $\boldsymbol{A}\vec{e}_j$ in this basis. Similarly, the operator $\boldsymbol{B}$ is represented by a matrix $B = \lfloor b_{ij} \rfloor$ with the vectors $\boldsymbol{B}\vec{e}_j$ in the same basis. Then, the matrix of the operator $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ can be composed in the form:

$$(\boldsymbol{A}+\boldsymbol{B})\vec{e}_j = \boldsymbol{A}\vec{e}_j + \boldsymbol{B}\vec{e}_j = a_{1j}\vec{e}_1 + a_{2j}\vec{e}_2 + \ldots + a_{nj}\vec{e}_n + b_{1j}\vec{e}_1 + b_{2j}\vec{e}_2 + \ldots + b_{nj}\vec{e} =$$
$$= (a_{1j} + b_{1j})\vec{e}_1 + (a_{2j} + b_{2j})\vec{e}_2 + \ldots + (a_{nj} + b_{nj})\vec{e}_n .$$

It is obvious that the matrix of the operator $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ in the basis $\{\vec{e}_i\}_{i=\overline{1,n}}$ is a sum of the matrices of these operators in the same basis.

<u>Definition</u>. ***The product of*** two linear operators $\boldsymbol{A} \cdot \boldsymbol{B}$ is a linear operator $\boldsymbol{C}$ such that

$$\boldsymbol{C}\vec{x} = \boldsymbol{A}(\boldsymbol{B}\,\vec{x}).$$

It is defined that $\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{B}$ is a linear operator if both operators $\boldsymbol{A}$ and $\boldsymbol{B}$ are linear (here without proof).

The matrix of the operator C is equal to the product of the matrices corresponding operators, i.e. $C = A \cdot B$.

If $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{I}$, then it is to be said that $\boldsymbol{B}$ is an ***inverse operator*** to $\boldsymbol{A}$ and is denoted as $\boldsymbol{B} = \boldsymbol{A}^{-1}$.

Using the matrix representation of the operators as matrices A and B, respectively, in a given basis, one can say that the operator $\boldsymbol{I}$ is presented by an identity matrix. That is,

$$A \cdot B = I \;\Rightarrow\; B = A^{-1}.$$

<u>Definition</u>. For any vector $\vec{x} \in R^n$, a linear operator $\boldsymbol{I}$ represented by the $n \times n$ identity matrix I, which maps every vector $\vec{x}$ into itself is called an ***identity operator***, that is

$$\boldsymbol{I}\,\vec{x} = \vec{x}$$

<u>*Note.*</u> the inverse operator has a matrix inverted to the operator matrix A. Obviously, this matrix is non-singular.

3.6. Simple Structure Operator

If the linear operator $\boldsymbol{A}$ has *n* linearly independent eigenvectors in an *n*-dimensional space *K*, then it is called ***a simple structure operator***

<u>*A sufficient condition of a simple structure operator*</u>:

If roots of the characteristic equation of a linear operator are distinct, the operator has a simple structure.

Indeed, in the case of *n* distinct eigenvalues of an operator, it has, pairwise *n* distinct eigenvectors, which are linearly independent ones. So they can form a new *n*-dimensional basis. In this basis, the matrix of a linear operator is diagonal.

In other words, an operator matrix is *diagonalizable* if there exists a basis of eigenvectors.

Recall that the *multiplicity* of an eigenvalue λ is the number of times that it occurs as a root of the characteristic polynomial. Let’s consider now the following lemma.

Lemma. If λ is an eigenvalue of a linear operator ***A***, then the number of linearly independent λ-eigenvectors (associated with eigenvalue λ) is never more than the multiplicity of λ.

We now use this fact to provide a theorem:

Theorem. To provide the existence of the eigenvectors basis, it is necessary and enough that to each eigenvalue there exist so many linearly independent eigenvectors, what is equal to multiplicity of this eigenvalue.

Since the linear operator has similar matrices in different bases, we can use the theorem to provide a *diagonalizability condition of the operator matrix*:

Theorem. Let $\boldsymbol{A}$ be a linear operator represented by an $n \times n$ matrix A. Then, A is *diagonalizable* if and only if for each eigenvalue λ of A, the dimension of the operator eigenspace $dim\,(E_\lambda(\boldsymbol{A}))$ is equal to the multiplicity of λ .

Example 3.18. Check a *diagonalizability* of the linear operator matrix given in the natural basis of linear space R^2 in the form:

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

Solution. Find the eigenvalues of the operator $\boldsymbol{A}$.

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) + 1 = \lambda^2 - 4\lambda + 4 = (\lambda - 2)^2 = 0.$$

So, $\lambda = 2$ – eigenvalue of operator $\boldsymbol{A}$ of multiplicity 2.

The corresponding eigenvectors follow from solving the SLAE

$$\begin{pmatrix} 1-2 & 1 \\ -1 & 3-2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0. \Rightarrow \begin{cases} -x_1 + x_2 = 0 \\ -x_1 + x_2 = 0 \end{cases} \Rightarrow \quad x_1 = x_2.$$

Thus, $\vec{x}_1 = (1;1)$ is an eigenvector corresponding to $\lambda = 2$. That is, we have that the fundamental system of solutions consists of one vector. Then all other solutions can be expressed through $\vec{x}_1$, which implies that any 2 eigenvectors will be linearly dependent. Thus, in linear space R^2 there is no eigenbasis of the linear operator $\boldsymbol{A}$. Thus, the linear operator $\boldsymbol{A}$ is not an operator of simple structure.

Chapter 4. EUCLIDEAN SPACE AND ORTHONORMAL BASIS

4.1 The Concept of Euclidean Space

Definition. A ***Euclidean space*** is a finite-dimensional vector space over the reals R, with a ***scalar product*** defined such that to each pair of elements of this space $\vec{x}$ and $\vec{y}$ matches a scalar denoted by $(\vec{x},\vec{y})$. The properties of the scalar product comply with the following axioms:

1. $(\vec{x},\vec{y}) = (\vec{y},\vec{x})$
2. $(\alpha\vec{x},\vec{y}) = \alpha(\vec{x},\vec{y})$
3. $(\vec{x}+\vec{z},\vec{y}) = (\vec{x},\vec{y}) + (\vec{z},\vec{y})$
4. $(\vec{x},\vec{x}) > 0$, if $\vec{x} \neq 0$

Consequences of the axioms:

1. $(\vec{x},\lambda\vec{y}) = \overline{\lambda}(\vec{x},\vec{y})$

   Indeed, $(\vec{x},\lambda\vec{y}) = \overline{(\lambda\vec{y},\vec{x})} = \overline{\lambda}\left(\overline{\vec{y},\vec{x}}\right) = \overline{\lambda}\left(\overline{\overline{\vec{x},\vec{y}}}\right) = \overline{\lambda}(\vec{x},\vec{y})$

2. $(\vec{x},\vec{y}+\vec{z}) = (\vec{x},\vec{y}) + (\vec{x},\vec{z})$

   Indeed, $(\vec{x},\vec{y}+\vec{z}) = (\vec{y}+\vec{z},\vec{x}) = (\vec{y},\vec{x}) + (\vec{z},\vec{x}) = (\vec{x},\vec{y}) + (\vec{x},\vec{z})$

Examples of scalar products in different spaces.

1. In vector spaces of real numbers R^2 and R^3, the scalar product is given by

$$(\vec{x},\vec{y}) = |\vec{x}||\vec{y}|\cdot\cos\left(\widehat{\vec{x},\vec{y}}\right)$$

All axioms can easily be verified as done in the course of analytical geometry. One can notice that this is specifically the case when a Cartesian coordinate system has been chosen, as, in this case, the scalar product of two vectors is the dot product or scalar product of their coordinate vectors.

2. A space continuous functions in the closed interval $[a, b]$ denoted as

$C[a,b]$, the scalar product is defined by

$$(x(t), y(t)) = \int_a^b x(t) \cdot y(t) dt$$

All axioms are fulfilled due to the properties of the definite integral. It should also be noted that in this space one can introduce the scalar product using other definitions. For example, one can present the scalar product in $C[a,b]$ in the form:

$$(\vec{x}, \vec{y}) = \int_a^b \psi^2(t) \cdot \vec{x}(t) \cdot \vec{y}(t) dt,$$

where $\psi(t)$ – an arbitrary nonzero function continuous on [a, b].

Example 4.1. Consider an n-dimensional vector space over a field of complex numbers. Let's choose the basis in this space $\{l_i\}_{i=\overline{1,n}}$. Then

$$\vec{x} = x_1 \vec{l}_1 + x_2 \vec{l}_2 + \ldots + x_n \vec{l}_n$$
$$\vec{y} = y_1 \vec{l}_1 + y_2 \vec{l}_2 + \ldots + y_n \vec{l}_n$$

Based on the properties of the scalar product:

$$(\vec{x}, \vec{y}) = \sum_{i=1}^{n} x_i \vec{l}_i \sum_{k=1}^{n} y_k \vec{l}_k = \sum_{i=1}^{n} \sum_{k=1}^{n} x_i y_k \cdot (\vec{l}_i, \vec{l}_k)$$

We denote $(\vec{l}_i, \vec{l}_k) = \alpha_{ik}$, where $\alpha_{ii} \neq 0$ in accordance with the axiom 4.

Hence,

$$(\vec{x}, \vec{y}) = \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{ik} x_i y_k \qquad (4.1)$$

The relation (4.1) is a general form of the scalar product in a finite-dimensional space, expressed via the coordinates of vectors.

4.2. Orthogonality and Modulus of the Vector

Let E be an arbitrary Euclidean space.

Definition. Vectors $\vec{x} \in E$ and $\vec{y} \in E$ are called ***orthogonal*** if their scalar product is zero, i.e.

$$(\vec{x}, \vec{y}) = 0.$$

In this case, we write $\vec{x} \perp \vec{y}$.

Definition. ***The modulus of the vector*** $\vec{x} \in E$ is a non-negative real number, which is determined by the formula:

$$|\vec{x}| = \sqrt{(\vec{x}, \vec{x})} \tag{4.2}$$

In the spaces R^2 and R^3, the orthogonality of vectors means their perpendicularity, and their moduli are their lengths.

Example 4.2. Let $E = T_n$ be an Euclidean space defined over the field of complex numbers. Then orthogonal vectors $\vec{x}$ and $\vec{y}$ satisfy the ratio

$$\sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{ik} \vec{x}_i \vec{y}_k = 0 \tag{4.3}$$

In particular, if

$$\alpha_{ik} = \begin{cases} 0, i \neq k \\ 1, i = k \end{cases}$$

then the relation (4.3) takes the form:

$$\sum_{k=1}^{n} \vec{x}_k \vec{y}_k = 0 \tag{4.4}$$

Then, the vector modulus can be calculated by

$$|\vec{x}| = \sqrt{\sum_{k=1}^{n} |x_k^2|} \tag{4.5}$$

In the space $E = R^n$ instead of formulas (4.4) and (4.5) we obtain:

$$\sum_{k=1}^{n} x_k y_k = 0 \tag{4.6}$$

$$|\vec{x}| = \sqrt{\sum_{k=1}^{n} x_k^2} \tag{4.7}$$

Example 4.3. In the space $C[a,b]$ the orthogonality of the elements means that $\left(\vec{x} \perp \vec{y}\right) \Rightarrow \int_a^b p(t) \cdot x(t) \cdot y(t) dt = 0$, where $p(t) > 0, \forall t \in [a,b]$.

In a general case, the modulus of an element of an arbitrary Euclidean space is called ***the norm*** and is denoted as

$$\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})}.$$

In the spaces R^2 and R^3, the norm coincides with the length of the vector $\vec{x}$, i.e. $\|\vec{x}\| = |\vec{x}|$. In the space $C[a,b]$ the norm of elements is defined as:

$$\|x(t)\| = \left[\int_a^b \rho(t) \cdot x^2(t) \cdot dt\right]^{1/2}$$

or if $\rho(t) = 1$, then

$$\|x\| = \left[\int_a^b x^2(t) \cdot dt\right]^{1/2}.$$

From the definition of the norm it follows that

1) $\|\vec{x}\| > 0$, at $x \neq 0$ and $\|\vec{x}\| = 0$, only when $\vec{x} = \vec{\theta}$.
2) $\|\lambda \vec{x}\| = |\lambda| \|\vec{x}\|$.

If $\|\vec{x}\| = 1$ then the vector $\vec{x}$ is called ***normalized***.

Obviously, any nonzero vector can be normalized by multiplying it by a factor $\lambda = \frac{1}{\|x\|}$. Then $\vec{y} = \frac{\vec{x}}{\|\vec{x}\|}$, and $\|\vec{y}\| = 1$.

4.3. Schwartz and Cauchy-Bunyakovsky Inequality

Let E coincide with a three-dimensional Euclidean space ($E=R^3$) with an scalar product:

$$\left|\left(\vec{x}, \vec{y}\right)\right| = |\vec{x}| \cdot |\vec{y}| \cdot \cos\left(\widehat{\vec{x}, \vec{y}}\right) \tag{4.8}$$

Taking into account that $\left|\cos\left(\widehat{\vec{x},\vec{y}}\right)\right| \le 1$, it follows from (4.8):

$$\left|\left(\vec{x},\vec{y}\right)\right| = \left|\vec{x}\right| \cdot \left|\vec{y}\right| \cdot \cos\left(\widehat{\vec{x},\vec{y}}\right) \le \left|\vec{x}\right| \cdot \left|\vec{y}\right|, \tag{4.9}$$

that is

$$\left|\left(\vec{x},\vec{y}\right)\right| \le \left|\vec{x}\right| \cdot \left|\vec{y}\right| \tag{4.10}$$

We show that inequality (4.10) is valid in any Euclidean space. For this purpose we will take arbitrary elements $\vec{x} \in E, \vec{y} \in E$ and any scalar $\alpha \in R$. Then,

$$\left|\vec{x} - \alpha \vec{y}\right|^2 = \left(\vec{x} - \alpha\vec{y}, \vec{x} - \alpha\vec{y}\right) \ge 0 \tag{4.11}$$

The left-hand side of the inequality can be expanded in the form:

$$\left(\vec{x},\vec{x}\right) - \left(\vec{x},\alpha\vec{y}\right) - \left(\alpha\vec{y},\vec{x}\right) + \left(\alpha\vec{y},\alpha\vec{y}\right) \ge 0,$$

that is

$$\left(\vec{x},\vec{x}\right) - \overline{\alpha}\left(\vec{x},\vec{y}\right) - \alpha\left(\vec{y},\vec{x}\right) + \alpha\overline{\alpha}\left(\vec{y},\vec{y}\right) \ge 0,$$

or

$$\left|\vec{x}\right|^2 - \overline{\alpha}\left(\vec{x},\vec{y}\right) - \alpha\left(\vec{y},\vec{x}\right) + \left|\alpha\right|^2\left|\vec{y}\right|^2 \ge 0. \tag{4.12}$$

This inequality is valid $\forall \alpha$. Let's choose $\alpha = \dfrac{\left(\vec{x},\vec{y}\right)}{\left|\vec{y}\right|^2}$, then the inequality (4.12) takes the form:

$$\left|\vec{x}\right|^2 - \frac{\overline{\left(\vec{x},\vec{y}\right)}}{\left|\vec{y}\right|^2}\left(\vec{x},\vec{y}\right) - \frac{\left(\vec{x},\vec{y}\right)}{\left|\vec{y}\right|^2}\overline{\left(\vec{x},\vec{y}\right)} + \frac{\left|\left(\vec{x},\vec{y}\right)\right|^2}{\left|\vec{y}\right|^4}\left|\vec{y}\right|^2 \ge 0,$$

or

$$\left|\vec{x}\right|^2 - \frac{\left|\left(\vec{x},\vec{y}\right)\right|^2}{\left|\vec{y}\right|^2} \ge 0,$$

that is

$$\left|\left(\vec{x},\vec{y}\right)\right|^2 \le \left|\vec{x}\right|^2 \cdot \left|\vec{y}\right|^2 \qquad (4.13)$$

Taking into account that $\left|\vec{x}\right|^2 = \left\|\vec{x}\right\|, \left|\vec{y}\right|^2 = \left\|\vec{y}\right\|$, the inequality (4.13) can be rewritten as:

$$\left|\left(\vec{x},\vec{y}\right)\right|^2 \le \left\|\vec{x}\right\|^2 \cdot \left\|\vec{y}\right\|^2 \qquad (4.14)$$

This inequality in any Euclidean space is called ***the Schwartz inequality***.

Let's write it in another form, extracting the square root of both parts

$$\left|\left(\vec{x},\vec{y}\right)\right| \le \left\|\vec{x}\right\| \cdot \left\|\vec{y}\right\| \qquad (4.15)$$

In an *n*-dimensional Euclidean space with *a natural basis*, this inequality will be written as follows:

$$\left|\sum_{k=1}^{n} x_k \cdot y_k\right| \le \sqrt{\sum_{k=1}^{n} x_k^2 \cdot \sum_{k=1}^{n} y_k^2} \qquad (4.16)$$

The inequality (4.16) is called ***Cauchy inequality***.

In space $C[a,b]$, the inequality (4.15) takes the form:

$$\left|\int_a^b x(t)\cdot y(t)dt\right| \le= \sqrt{\int_a^b [x(t)]^2 dt} \cdot \sqrt{\int_a^b [y(t)]^2 dt} \qquad (4.17)$$

The inequality (4.17) is called ***Bunyakovsky's inequality***.

4.4 Orthogonal and Orthonormal Basis. Gram-Schmidt procedure.

Definition. The basis of Euclidean space $\left\{\vec{l}_i\right\}_{i=\overline{1,n}}$ is called ***orthogonal*** if the scalar products of distinct basis vectors are zero,

$$(\vec{l}_i, \vec{l}_j) = 0 \text{ at } i \ne k .$$

If, in addition, the modulus of each basis vector is one, $\left\|\vec{l}_i\right\| = 1, \forall i = \overline{1,n}$, then the basis is called ***orthonormal (ONB)***, i.e.

$$(\vec{l}_i, \vec{l}_j) = \delta_{ij} = \begin{cases} 0, \text{ if } i \neq k \\ 1, \text{ if } i = k \end{cases}.$$

Lemma. Pairwise orthogonal nonzero vectors are linearly independent.

Proof. Let the vectors $\left\{\overrightarrow{x_i}\right\}_{i=\overline{1,m}}$ be pairwise orthogonal, i.e. $\left(\overrightarrow{x_i}, \overrightarrow{x_k}\right) = 0$ if $i \neq k$. At the same time all $x_i \neq 0$.

Suppose that

$$\sum_{i=1}^{m} \alpha_i \cdot \overrightarrow{x_i} = 0 \tag{4.18}$$

Taking the scalar products for both of parts (4.18) with vectors $\overrightarrow{x_k}$ $\left(k = \overline{1,n}\right)$ and accounting the properties of the scalar product, we get that

$$\sum_{i=1}^{m} \alpha_1 \left(\overrightarrow{x_i}, \overrightarrow{x_k}\right) = 0, \forall k = \overline{1,n}.$$

Let k = 1, then

$$\sum_{i=1}^{m} \alpha_1 (x_i, x_1) = 0 \Rightarrow \left\| (x_i, x_1) = \left(\overrightarrow{x_1}, \overrightarrow{x_1}\right) \right\| \Rightarrow \alpha_1 (x_1, x_1) = 0 \Rightarrow \left\| (x_1, x_1) > 0 \right\| \Rightarrow \boxed{\alpha_1 = 0}$$

Similarly, we can show that all α_i in (4.18) is zero, i.e. $\left\{\overrightarrow{x_i}\right\}_{i=\overline{1,m}}$ are linearly independent. ■

Theorem. Every Euclidean space has an orthonormal basis.

Proof. Let E be an n-dimensional linear space. Then there always are n linearly independent vectors $\left\{\vec{l}_i\right\}_{i=\overline{1,n}}$. Let's denote $\vec{l}_1 = \vec{l}_1^{\,*}$ and construct a vector $\vec{l}_2^{\,*} = \vec{l}_2 + \alpha_{11} \vec{l}_1^{\,*}$. Herewith, it should be noticed that $\vec{l}_2^{\,*} \neq 0$ as the system of vectors $\left\{\vec{l}_i\right\}_{i=\overline{1,n}}$ is linearly independent. Coefficient α_{11} choose so that the vectors $\vec{l}_1^*$ and $\vec{l}_2^*$ were orthogonal, i.e.

$$\left(\vec{l}_1^*,\vec{l}_2^*\right)=0 \Rightarrow \left(\vec{l}_2,\vec{l}_1^*\right)+\alpha_{11}\left(\vec{l}_1^*,\vec{l}_1^*\right)=0 \Rightarrow \alpha_{11}=-\frac{\left(\vec{l}_2,\vec{l}_1^*\right)}{\left(\vec{l}_1^*,\vec{l}_1^*\right)}=-\frac{\left(\vec{l}_2,\vec{l}_1\right)}{\left(\vec{l}_1^*,\vec{l}_1^*\right)}.$$

Next, let's compose a vector $\vec{l}_3^*$:

$$\vec{l}_3^*=\vec{l}_3+\alpha_{21}\vec{l}_1^*+\alpha_{22}\vec{l}_2^*,$$

where $\vec{l}_3^*\neq 0$, since $\vec{l}_1^*,\vec{l}_2^*$ and $\vec{l}_3$ – are linearly independent.

Let's choose α_{21} and α_{22} so that the vector $\vec{l}_3^*$ will be orthogonal to the vectors $\vec{l}_1^*$ and $\vec{l}_2^*$, i.e.

$$\left(\vec{l}_3^*,\vec{l}_1^*\right)=\left(\vec{l}_3,\vec{l}_1^*\right)+\alpha_{21}\left(\vec{l}_1^*,\vec{l}_1^*\right)+\alpha_{22}\underbrace{\left(\vec{l}_2^*,\vec{l}_1^*\right)}_{=0}=0,$$

$$\left(\vec{l}_3,\vec{l}_1^*\right)+\alpha_{21}\left(\vec{l}_1^*,\vec{l}_1^*\right)=0 \Rightarrow \alpha_{21}=-\frac{\left(\vec{l}_3,\vec{l}_1^*\right)}{\left(\vec{l}_1^*,\vec{l}_1^*\right)}$$

Similarly,

$$\left(\vec{l}_3^*,\vec{l}_2^*\right)=\left(\vec{l}_3,\vec{l}_2^*\right)+\alpha_{22}\left(\vec{l}_2^*,\vec{l}_2^*\right)=0, \Rightarrow \alpha_{22}=-\frac{\left(\vec{l}_3,\vec{l}_2^*\right)}{\left(\vec{l}_2^*,\vec{l}_2^*\right)}.$$

If (n-1) vectors pairwise orthogonal are constructed in the same way, then the vector $\vec{l}_n^*$ can be chosen in the form:

$$\vec{l}_n^*=\vec{l}_n+\alpha_{n-1,1}\vec{l}_1^*+\alpha_{n-1,2}\vec{l}_2^*+\ldots+\alpha_{n-1,n-1}\vec{l}_{n-1}^*,$$

where $\vec{l}_n^*\neq 0$, since $\vec{l}_n$ and $\left\{\vec{l}_i^*\right\}_{i=\overline{1,n-1}}$ are linearly independent.

We will demand $\vec{l}_n^*$ to be orthogonal to all the other vectors $\left\{\vec{l}_i^*\right\}_{i=\overline{1,n-1}}$. That is, we have the following (n - 1) conditions:

$$\begin{cases}\left(\vec{l}_n^*,\vec{l}_1^*\right)=0,\\ \left(\vec{l}_n^*,\vec{l}_2^*\right)=0,\\ \ldots\ldots.\\ \left(\vec{l}_n^*,\vec{l}_{n-1}^*\right)=0.\end{cases}$$

All unknown coefficients will be determined from this system as follows:

$$\alpha_{n-1,k} = -\frac{\left(\vec{l}_n, \vec{l}_k^*\right)}{\left(\vec{l}_k^*, \vec{l}_k^*\right)}, k = \overline{1, n-1}.$$

Thus, for an arbitrary basis $\left\{\vec{l}_i\right\}_{i=\overline{1,n}}$ it is always possible to construct *n* pairwise orthogonal vectors, which by virtue of the lemma will be linearly independent.

If each vector of the orthogonal basis $\left\{\vec{l}_i^*\right\}_{i=\overline{1,n}}$ is divided by its norm, we obtain an ***orthonormal basis***:

$$\frac{\vec{l}_1^*}{\left\|\vec{l}_1^*\right\|}; \frac{\vec{l}_2^*}{\left\|\vec{l}_2^*\right\|}; \ldots; \frac{\vec{l}_n^*}{\left\|\vec{l}_n^*\right\|}.$$

That is, there exists a set of orthonormal linearly independent vectors which span a particular Euclidean space. ■

Note:

1. A Euclidean space has more than one orthonormal basis.

2. The algorithm described above in the proof for orthonormalizing a set of vectors in a Euclidean space is called ***the orthogonalization process (or Gram–Schmidt orthogonalization)***.

The scalar $\alpha_{i-1,j} = -\frac{\left(\vec{l}_i, \vec{l}_j\right)}{\left(\vec{l}_j, \vec{l}_j\right)}$ is called the *Fourier coefficient* of $\vec{l}_i$ with respect to $\vec{l}_j$. The basic operation related to finding an orthogonal vector is called *orthogonal projection*. This operator projects the vector $\vec{l}_i$ orthogonally onto the line spanned by vector $\vec{l}_j$ as $proj_{\vec{l}_j} \vec{l}_i = \frac{\left(\vec{l}_i, \vec{l}_j\right)}{\left(\vec{l}_j, \vec{l}_j\right)} \vec{l}_j$.

Example 4.3. Find the orthogonal basis in the space of polynomials not higher than the second degree, which are defined on the segment [-1; 1].

Solution. As a starting point we take the natural basis as a system of the following functions:

$$\vec{l}_0 = 1,\ \vec{l}_1 = t,\ \vec{l}_2 = t^2.$$

Then, we will find a vector $\vec{l}_1^{\,*}$ orthogonal to the first basis vector $\vec{l}_0^{\,*} = \vec{l}_0$ as $\vec{l}_1^{\,*} = \vec{l}_1 + \alpha_{10}\vec{l}_0^{\,*}$. That is, using orthogonal projection (taking the **scalar** product of $\vec{l}_1^{\,*}$ and $\vec{l}_0^{\,*}$ and equating it to 0) we get

$$\underbrace{\int_{-1}^{1} 1\cdot t\,dt}_{=0} + \alpha_{10}\int_{-1}^{1} dt = 0 \Rightarrow \alpha_{10}\cdot 2 = 0 \Rightarrow \alpha_{10} = -\frac{\left(\vec{l}_1, \vec{l}_0^{\,*}\right)}{\left(\vec{l}_0^{\,*}, \vec{l}_0^{\,*}\right)} = 0, \Rightarrow \vec{l}_1^{\,*} = \vec{l}_1 = t.$$

Similarly, $\vec{l}_2^{\,*} = \vec{l}_2 + \alpha_{20}\vec{l}_0^{\,*} + \alpha_{21}\vec{l}_1^{\,*}$. Then, appropriate orthogonal projections lead to

$$\left(\vec{l}_2, \vec{l}_0^{\,*}\right) = \int_{-1}^{1} t^2 dt = \left.\frac{t^3}{3}\right|_{-1}^{1} = \frac{2}{3};\text{ and } \left(\vec{l}_0^{\,*}, \vec{l}_0^{\,*}\right) = \int_{-1}^{1} 1\cdot 1 dt = t\Big|_{-1}^{1} = 2;\text{ i.e.}$$

$$\alpha_{20} = -\frac{\left(\vec{l}_2, \vec{l}_0^{\,*}\right)}{\left(\vec{l}_0^{\,*}, \vec{l}_0^{\,*}\right)} = -\frac{1}{3};$$

$$\left(\vec{l}_2, \vec{l}_1^{\,*}\right) = \int_{-1}^{1} t^3 dt = \left.\frac{t^4}{4}\right|_{-1}^{1} = 0; \Rightarrow \alpha_{21} = -\frac{\left(\vec{l}_2, \vec{l}_1^{\,*}\right)}{\left(\vec{l}_1^{\,*}, \vec{l}_1^{\,*}\right)} = 0.$$

Thereby, we have a second basis vector $\vec{l}_2^{\,*} = t^2 - \frac{1}{3}$.

Summarizing all the vectors we get the orthogonal basis in the form:

$$\vec{l}_0^{\,*} = 1,$$

$$\vec{l}_1^{\,*} = t,$$

$$\vec{l}_2^{\,*} = t^2 - \frac{1}{3}.$$

Let's find the norms of these vectors:

$$\|l_0^*\| = \sqrt{\int_{-1}^{1} dt} = \sqrt{2},$$

$$\|l_1^*\| = \sqrt{\int_{-1}^{1} t^2 dt} = \sqrt{\frac{2}{3}},$$

$$\|l_2^*\| = \sqrt{\int_{-1}^{1}\left(t^2 - \frac{1}{3}\right)^2 dt} = \sqrt{\int_{-1}^{1}\left(t^4 - \frac{2}{3}t^2 + \frac{1}{9}\right)} = \sqrt{2\left(\frac{1}{5} - \frac{1}{9}\right)} = \sqrt{2\left(\frac{4}{45}\right)} = \frac{2}{3}\sqrt{\frac{2}{5}};$$

Finally, the vectors $\frac{1}{\sqrt{2}}$; $\frac{t\cdot\sqrt{3}}{\sqrt{2}}$; and $\frac{\left(t^2 - \frac{1}{3}\right)\cdot 3\sqrt{5}}{2\sqrt{2}}$ form an orthonormal basis of the space of polynomials not higher than the second degree defined on the segment [-1; 1].

4.5. Orthogonal Complements

Definition. If $\vec{y} \in E$ is an element of a given linear space such that satisfies the condition $(\vec{y}, \vec{x}) = 0$, i.e. $\vec{y} \perp \vec{x}\ \ \forall \vec{x} \in L$, where L is a subspace of the space E, then it is said that the vector $\vec{y}$ is orthogonal to subspace L. The set of all elements $\vec{y} \in E$, orthogonal to L, is called an ***orthogonal complement*** of the subspace L and is denoted as $L^{\perp}$ (read "*L perpendicular*").

Lemma. If L is subspace of the space E, i.e. $L \subset E$, then its orthogonal complement also forms a subspace of the space E, i.e. $L^{\perp} \subset E$.

Proof. Take any two elements $\overrightarrow{y_1}$ and $\overrightarrow{y_2}$ in the $L^{\perp}$. Then $(\overrightarrow{y_1}, \vec{x}) = 0\ \Lambda\ (\overrightarrow{y_2}, \vec{x}) = 0, \forall \vec{x} \in L$. Consider the case $\overrightarrow{y_1} + \overrightarrow{y_2}$:

$(\overrightarrow{y_1} + \overrightarrow{y_2}, \vec{x}) = (\overrightarrow{y_1}, \vec{x}) + (\overrightarrow{y_2}, \vec{x}) = 0$, i.e. $\overrightarrow{y_1} + \overrightarrow{y_2} \in L^{\perp}$.

Similarly, $\alpha \cdot \overrightarrow{y_1} \in L^{\perp}$, since $(\alpha \cdot \overrightarrow{y_1}, \vec{x}) = \alpha(\overrightarrow{y_1}, \vec{x}) = 0, \forall \vec{x} \in L$. ■

Example 4.4. Let L be a one-dimensional subspace of two-dimensional space V^2 (plane). Then $L^{\perp}$ is a line in this plane perpendicular to the line L.

Example 4.5. Let L be a plane in three-dimensional space V^3 that passes through the origin, then $L^{\perp}$ is a perpendicular to this plane and is also passing through the origin.

Let E be an arbitrary space, and L be its subspace and $L^{\perp}$ be its orthogonal complement. We also suppose that an element $\vec{x} \in L$ and $\vec{x} \in L^{\perp}$ exists. That means $(\vec{x}, \vec{x}) = 0, \Rightarrow \vec{x} = 0$. So,

$$L \cap L^{\perp} = \vec{0}.$$

Theorem. In any Euclidean space, the sum of the dimensions of the subspaces L and $L^{\perp}$ always equals to the dimension of the whole space.

Proof. Suppose that a subspace L is given in space E, and its dimension is equal to k, i.e. $\dim L = k$. Also, let the dimension of its orthogonal complement be equal to m, $\dim L^{\perp} = m$.

Suppose orthonormal bases in L and $L^{\perp}$ are given by vectors $\overrightarrow{g_1}, \overrightarrow{g_2}, ..., \overrightarrow{g_k}$ and $\overrightarrow{f_1}, \overrightarrow{f_2}, ..., \overrightarrow{f_m}$, respectively. Let $\vec{x}$ be an arbitrary vector $\vec{x} \in E^n$. We can construct a new vector $\vec{y}$ as follows:

$$\vec{y} = \sum_{i=1}^{k} (\vec{x}, \overrightarrow{g_i}) \overrightarrow{g_i} \tag{4.19}$$

Obviously, this vector belongs to the same space, $\vec{y} \in L$.

In the same way, we construct the other vector

$$\vec{z} = \vec{x} - \vec{y} = \vec{x} - \sum_{i=1}^{k} (\vec{x}, \overrightarrow{g_i}) \overrightarrow{g_i} \tag{4.20}$$

Taking **scalar** products of each side of the equality (4.20) with respect to the vector $\overrightarrow{g_j}\left(j=\overline{1,k}\right)$, we get

$$\left(\vec{z},\overrightarrow{g_j}\right)=\left(\vec{x},\overrightarrow{g_j}\right)-\left(\sum_{i=1}^{k}\left(\vec{x},\overrightarrow{g_i}\right)\right)\left(\overrightarrow{g_i},\overrightarrow{g_j}\right)=\left(\vec{x},\overrightarrow{g_j}\right)-\sum_{i=1}^{k}\left(\vec{x},\overrightarrow{g_i}\right)\cdot\underbrace{\left(\overrightarrow{g_i},\overrightarrow{g_j}\right)}_{\delta_{ij}}=$$

$$=\left(\vec{x},\overrightarrow{g_j}\right)-\left(\vec{x},\overrightarrow{g_j}\right)=0\,.$$

Thus, the element $\vec{z}$ is orthogonal to each of the vectors of the basis $\overrightarrow{g_1},\overrightarrow{g_2},...,\vec{g}_k$. Then any vector $\vec{y}\in L$ satisfies condition $\vec{z}\perp\vec{y}$. That means $\vec{z}\perp L$, i.e. $\vec{z}\in L^{\perp}$.

In this case $\vec{z}$ can be represented in the basis of vectors $\overrightarrow{f_1},\overrightarrow{f_2},...,\overrightarrow{f_m}$ as:

$$\vec{z}=\sum_{l=1}^{m}\eta_l\cdot\overrightarrow{f_l}\,, \tag{4.21}$$

where $\eta_l=\left(\vec{z},\overrightarrow{f_l}\right)$ are coefficients and $\{f_l\}_{l=\overline{1,m}}$ are vectors of the orthonormal basis.

Let's denote $\left(\vec{x},\overrightarrow{g_j}\right)=\xi_i$, then

$$\vec{x}=\vec{z}+\vec{y}\Rightarrow$$

$$\vec{x}=\sum_{l=1}^{m}\eta_l\cdot\overrightarrow{f_l}+\sum_{i=1}^{k}\xi_i\cdot\overrightarrow{g_i} \tag{4.22}$$

Thus, the element $\vec{x}\in E^n$ can be decomposed in the basis vectors:

$$\overrightarrow{g_1},\overrightarrow{g_2},...,\overrightarrow{g_k},\overrightarrow{f_1},\overrightarrow{f_2},...,\overrightarrow{f_m} \tag{4.23}$$

The vectors from (4.23) are pairwise orthogonal, i.e. $\left(\overrightarrow{g_i}\cdot\overrightarrow{f_j}\right)=0,\forall i=\overline{1,k}$ and $\forall j=\overline{1,m}$, as they are basis vectors of $\{g_i\}_{i=\overline{1,k}}$ and $\left\{\overrightarrow{f_j}\right\}_{j=\overline{1,m}}$.

Thus, the vectors (4.23) are linearly independent, so they form a basis in E^n and their common number is $k+m=n$. That is $\dim L+\dim L^{\perp}=\dim E^n$ ■

The latter means that E^n is a direct sum of subspaces L and $L^{\perp}$ and is denoted as:

$$E_n = L \oplus L^{\perp}$$

Definition. A space R is ***a direct sum of subspaces R₁ and R₂***, if $\forall x \in R$ is a decomposition $\vec{x} = \vec{x}_1 + \vec{x}_2$, where $\vec{x}_1 \in R_1, \vec{x}_2 \in R_2$. In doing so, this decomposition is the only one.

Remark. Analogously to the three-dimensional space R^3, in the case of arbitrary Euclidean space E^n, a vector $\vec{y}$, which is represented by expression (4.19), is called the *orthogonal projection* of the vector $\vec{x}$ on subspace L, and the vector $\vec{z} = \vec{x} - \vec{y}$ is *orthogonal projection* of the element $\vec{x}$ on subspace $L^{\perp}$ (see

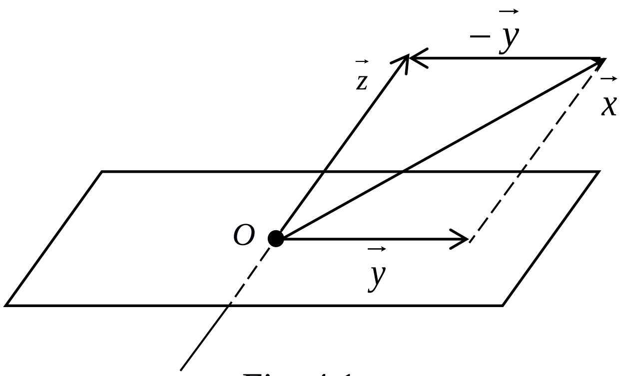


Fig. 4.1

Figure 4.1)

Example 4.6. A subspace L is formed by a spanning set of vectors $a_1 = (2;-1;3;-2), a_2 = (4;-2;5;1)$, $a_3 = (2;-1;1;8)$. Find the basis of orthogonal complement $L^{\perp}$.

Solution. Let us check whether all vectors are linearly independent, and define the dimension of the subspace, $\dim L$.

$$\begin{pmatrix} 2 & -1 & 3 & -2 \\ 4 & -2 & 5 & 1 \\ 2 & -1 & 1 & 8 \end{pmatrix} \begin{matrix} r_1 \times (-2) + r_2 \to r_1 \\ r_3 - r_1 \to r_1 \\ \end{matrix} \sim \begin{pmatrix} 2 & -1 & 3 & -2 \\ 0 & 0 & -1 & 5 \\ 0 & 0 & -2 & 10 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & 3 & -2 \\ 0 & 0 & -1 & 5 \\ 0 & 0 & -1 & 5 \end{pmatrix}$$

.

That is only two vectors, for example, $\overrightarrow{a_1}$ and $\overrightarrow{a_2}$ are linearly independent, and they can be taken as a basis of this subspace *L*.

Let $\vec{x} = (\xi_1, \xi_2, \xi_3, \xi_4) \in L^{\perp}$, then in accordance with the orthogonal complement's properties, we have

$$\begin{cases} (\vec{x}, \overrightarrow{a_1}) = 0 \\ (\vec{x}, \overrightarrow{a_2}) = 0 \end{cases} \Rightarrow \begin{cases} 2\xi_1 - \xi_2 + 3\xi_3 - 2\xi_4 = 0 \\ -\xi_3 + 5\xi_4 = 0 \end{cases} \Rightarrow \begin{cases} \xi_3 = 5\xi_4 \\ \xi_2 = 2\xi_1 + 3\xi_3 - 2\xi_4 \end{cases}$$

$$\begin{cases} \xi_3 = 5\xi_4 \\ \xi_2 = 2\xi_1 + 3\xi_3 - 2\xi_4 \end{cases} \Rightarrow \begin{cases} \xi_3 = 5\xi_4 \\ \xi_2 = 2\xi_1 + 3 \cdot 5\xi_4 - 2\xi_4 = 2\xi_1 + 13\xi_4 \end{cases}$$

In this homogeneous system of SLAE, two free variables are assigned as ξ_1 and ξ_4. That is the dimension of the subspace is $\underline{\dim L = 2}$. The FSS associated with this system can be found as

	ξ_1	ξ_2	ξ_3	ξ_4
e_1	1	2	0	0
e_2	0	13	5	1

Hence, the vectors $\overrightarrow{e_1} = (1;2;0;0)$ and $\overrightarrow{e_2} = (0;13;5;1)$ form a basis in $L^{\perp}$.

Example 4.7. A subspace $L \subset R^4$ is given by the system of homogeneous equations as $\begin{cases} \xi_1 + 2\xi_2 + 3\xi_3 - \xi_4 = 0 \\ \xi_1 - \xi_2 + \xi_3 + 2\xi_4 = 0 \\ \xi_1 + 5\xi_2 + 5\xi_3 - 4\xi_4 = 0 \end{cases}$.

Find the basis of the orthogonal complement $L^{\perp}$.

Solution. Compose a matrix of homogeneous SLAE and calculate its rank by reducing this matrix to the row echelon form as follows:

$$\begin{pmatrix} 1 & 2 & 3 & -1 \\ 1 & -1 & 1 & 2 \\ 1 & 5 & 5 & -4 \end{pmatrix} \begin{matrix} \\ r_2 - r_1 \\ r_3 - r_1 \end{matrix} \sim \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & -3 & -2 & 3 \\ 0 & 3 & 2 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & -1 \\ 0 & 3 & 2 & -3 \end{pmatrix} \sim$$

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

Hence, $\mathrm{Rg}A = 2$.

It means that the system is reduced to two equations and has two linearly independent solutions, as a result, $\dim L = 2$ and $\dim L^{\perp} = 2$.

Determine the solution of the reduced system:

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 + 2\xi_4 = 0 \\ 3\xi_2 + 2\xi_3 - 3\xi_4 = 0 \end{cases}$$

Suppose that ξ_3 and ξ_4 are free variables. Then

$$\xi_2 = \frac{1}{3}(-2\xi_3 + 3\xi_4)$$

$$\xi_1 = \xi_2 + \xi_3 + 2\xi_4 = \frac{1}{3}(-2\xi_3 + 3\xi_4) + \xi_3 + 2\xi_4 = \frac{1}{3}\xi_3 + 3\xi_4$$

The FSS is calculated as

	ξ_1	ξ_2	ξ_3	ξ_4
e_1	$\frac{1}{3}$	$-\frac{2}{3}$	1	0
e_2	3	1	0	1

That is, the basis vectors of L are $\overrightarrow{e_1}$ and $\overrightarrow{e_2}$ such that

$$\overrightarrow{e_1} = (1;-2;3;0) \text{ and } \overrightarrow{e_2} = (3;1;0;1)$$

To find the basis of the orthogonal complement $L^{\perp}$, we have to write:

$$\begin{cases} (\vec{x}, \overrightarrow{e_1}) = 0 \\ (\vec{x}, \overrightarrow{e_2}) = 0 \end{cases}, \text{ i.e. } \begin{cases} \xi_1 - 2\xi_2 + 3\xi_3 = 0 \\ 3\xi_1 + \xi_2 + \xi_4 = 0 \end{cases}.$$

Let's find the FSS of this system of homogeneous equations

$$\begin{pmatrix} 1 & -2 & 3 & 0 \\ 3 & 1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & 0 \\ 0 & 7 & -9 & 1 \end{pmatrix} \Rightarrow \begin{matrix} \xi_1 - 2\xi_2 + 3\xi_3 = 0 \\ 7\xi_2 - 9\xi_3 + \xi_4 = 0 \end{matrix} \Rightarrow \begin{matrix} \xi_4 = -7\xi_2 + 9\xi_3 \\ \xi_1 = 2\xi_2 - 3\xi_3 \end{matrix}$$

	ξ_1	ξ_2	ξ_3	ξ_4
$\overrightarrow{b_1} = \overrightarrow{e_1}$	2	1	0	-7
$\overrightarrow{b_2} = \overrightarrow{e_2}$	-3	0	1	9

That is, the vectors $\overrightarrow{b_1}(2;1;0;-7)$, $\overrightarrow{b_2}(-3;0;1;9)$ form a basis of $L^{\perp}$.

4.6. The Gram determinant

Suppose that vectors $\{\overrightarrow{g_i}\}_{i=\overline{1,k}}$ are given in the linear space *E*. The *Gram matrix* (or *Gramian matrix*, *Gramian*) is a matrix whose entries are given by the scalar product $\Gamma = (g_{ij}) = (\overrightarrow{g_i}, \overrightarrow{g_j}), \forall i, j = \overline{1,k}$.

Then, the determinant of this matrix is called the *Gram determinant* and looks like

$$\det\Gamma = \Delta = \begin{vmatrix} (\overrightarrow{g_1}, \overrightarrow{g_1}) & (\overrightarrow{g_1}, \overrightarrow{g_2}) & \dots & (\overrightarrow{g_1}, \overrightarrow{g_n}) \\ (\overrightarrow{g_2}, \overrightarrow{g_1}) & (\overrightarrow{g_2}, \overrightarrow{g_2}) & \dots & (\overrightarrow{g_2}, \overrightarrow{g_n}) \\ \dots & \dots & \dots & \dots \\ (\overrightarrow{g_k}, \overrightarrow{g_1}) & (\overrightarrow{g_k}, \overrightarrow{g_2}) & \dots & (\overrightarrow{g_k}, \overrightarrow{g_k}) \end{vmatrix}, \quad (4.24)$$

Suppose that the vectors $\{\overrightarrow{g_i}\}_{j=\overline{1,k}}$ are linearly dependent. For instance, a vector $\overrightarrow{g_k}$ is a linear combination of vectors $\overrightarrow{g_1}, \overrightarrow{g_2}, ..., \overrightarrow{g_k}$ as a result the *k*-th column of the determinant will be a linear combination of the other its columns. That is, the Gram determinant is equal to zero, $\Delta = 0$.

Theorem. If the vectors $\{\overrightarrow{g_i}\}_{i=\overline{1,k}}$ are linearly independent, then Gram's determinant of these vectors is nonzero, i.e. $\Delta \neq 0$.

Proof. Suppose the opposite, that the Gram's determinant $\Delta = 0$. Then, based on the properties of the determinant, one of the columns (rows) is a linear combination of the others. Therefore, there is a nonzero set of numbers $\alpha_1, \alpha_2, .., \alpha_n$, such that

$$\sum_{i=1}^{n} \alpha_i \cdot \overrightarrow{g_i} = \vec{0}, \tag{4.25}$$

where $\overrightarrow{g_i}$ is a vector whose coordinates coincide with the entries of the *i*-th row of the determinant (4.24).

Equation (4.25) is equivalent to *k* equalities in the form:

$$\alpha_1 \left(\overrightarrow{g_1}, \overrightarrow{g_1}\right) + \alpha_2 \left(\overrightarrow{g_2}, \overrightarrow{g_1}\right) + .. + \alpha_k \left(\overrightarrow{g_k}, \overrightarrow{g_1}\right) = 0,$$

or

$$\left(\left(\alpha_1 \overrightarrow{g_1} + \alpha_2 \overrightarrow{g_2} + ... + \alpha_k \overrightarrow{g_k}\right), \overrightarrow{g_1}\right) = 0 \tag{4.26}$$

Let's denote the sum of *k* terms as a new vector:

$$\sum_{i=1}^{k} \alpha_i \vec{g}_i = \vec{y} \tag{4.27}$$

It is obvious that $\vec{y} \in L$.

Then, the equality (4.26) takes the form:

$$\left(\vec{y}, \overrightarrow{g_1}\right) = 0 \tag{4.28}$$

That is, the vectors are orthogonal $\vec{y} \perp \overrightarrow{g_1}$.

Similarly, we can write the other equalities, i.e. the following system occurs:

$$\begin{cases} \alpha_1 \left(\overrightarrow{g_1}, \overrightarrow{g_2}\right) + \alpha_2 \left(\overrightarrow{g_2}, \overrightarrow{g_2}\right) + ... + \alpha_k \left(\overrightarrow{g_k}, \overrightarrow{g_2}\right) = 0 \\ ... \\ \alpha_1 \left(\overrightarrow{g_1}, \overrightarrow{g_k}\right) + \alpha_2 \left(\overrightarrow{g_2}, \overrightarrow{g_k}\right) + ... + \alpha_k \left(\overrightarrow{g_k}, \overrightarrow{g_k}\right) = 0 \end{cases} \tag{4.29}$$

Further,

$$\begin{cases} \sum_{i=1}^{k} \alpha_i \left(\overrightarrow{g_i}, \overrightarrow{g_2}\right) = 0 \\ \ldots \\ \sum_{i=1}^{k} \alpha_i \left(\overrightarrow{g_i}, \overrightarrow{g_k}\right) = 0 \end{cases}, \qquad (4.30)$$

Finally, we take the form:

$$\begin{cases} \left(\vec{y}, \overrightarrow{g_2}\right) = 0 \\ \left(\vec{y}, \overrightarrow{g_3}\right) = 0 \\ \ldots \\ \left(\vec{y}, \overrightarrow{g_k}\right) = 0 \end{cases} \qquad (4.31)$$

It follows from the equations (4.28) and (4.31) that

$$\vec{y} \perp \overrightarrow{g_i}, \qquad \forall\, i = \overline{1,k}$$

Thus, if L is a set of spanning vectors $\{g_i\}_{i=\overline{1,k}}$, then $\vec{y} \perp L$ and $\vec{y} \perp L^{\perp}$. On the other hand, it follows from (4.27) that $\vec{y} \in L$. So,

$$\left(\vec{y} \perp L\right) \wedge \left(\vec{y} \perp L^{\perp}\right) \Rightarrow \vec{y} = \vec{0}$$

With this fact, it follows from (4.30) that $\sum_{i=1}^{k} \alpha_i^{\,2} \neq 0$. It means that the vectors $\{g_i\}_{i=\overline{1,k}}$ are linearly dependent. However, this contradicts with the conditions of the theorem. ■

4.7. Orthogonal Projection

Let's consider two subspaces L and $L^{\perp}$ of the space E^n, where the dimension of the subspace L is $\dim L = k$. Suppose $\vec{x} \in E^n$ is a given nonzero vector of this space, and suppose $\vec{y}$ is another vector. We seek **for** the orthogonal projection of $\vec{y}$ onto the subspace L. Let a basis of L be given by the vectors $\{g_i\}_{i=\overline{1,k}}$. Then the vector $\vec{y}$ can be represented there in the form:

$$\vec{y} = \sum_{i=1}^{k} \xi_i \overrightarrow{g_i} \tag{4.32}$$

So, the problem is reduced to finding the coefficients of decomposition ξ_i in (4.32). Recall if the basis $\{g_i\}_{i=\overline{1,k}}$ is orthonormal, then we have that $\xi_i = \left(\vec{x}, \overrightarrow{g_i}\right),\ i = \overline{1,k}$. However, in our case, we have chosen an arbitrary basis as a result the problem turns into a more general one.

Let $\vec{z}$ be projection of vector $\vec{x}$ along $L^{\perp}$. Then $\vec{z} \perp L$, i.e. $\vec{z} \perp \overrightarrow{g_i}$, $\forall i = \overline{1,k}$. That is $\left(z, \overrightarrow{g_i}\right) = 0$, $\forall i = \overline{1,k}$, or $\left(\vec{x} - \vec{y}, g_i\right) = 0$, or $\left(\vec{x}, \overrightarrow{g_i}\right) = \left(\vec{y}, \overrightarrow{g_i}\right)$. By assigning *i* with values $1,2,...,k$ consequently, we get

$$\begin{cases} \left(\vec{x}, \vec{g}_1\right) = \left(\vec{y}, \vec{g}_1\right) \\ \left(\vec{x}, \vec{g}_2\right) = \left(\vec{y}, \vec{g}_2\right) \\ \left(\vec{x}, \vec{g}_k\right) = \left(\vec{y}, \vec{g}_k\right) \end{cases} \tag{4.33}$$

Substituting (4.32) into the system (4.33), we can write

$$\begin{cases} \xi_1\left(\vec{g}_1, \vec{g}_1\right) + \xi_2\left(\vec{g}_2, \vec{g}_1\right) + \ldots + \xi_k\left(\vec{g}_k, \vec{g}_1\right) = \left(\vec{x}, \vec{g}_1\right) \\ \xi_1\left(\vec{g}_1, \vec{g}_2\right) + \xi_2\left(\vec{g}_2, \vec{g}_2\right) + \ldots + \xi_k\left(\vec{g}_k, \vec{g}_2\right) = \left(\vec{x}, \vec{g}_2\right) \\ \xi_1\left(\vec{g}_1, \vec{g}_k\right) + \xi_2\left(\vec{g}_2, \vec{g}_k\right) + \ldots + \xi_k\left(\vec{g}_k, \vec{g}_k\right) = \left(\vec{x}, \vec{g}_k\right) \end{cases} \tag{4.34}$$

Since the vectors $\{g_i\}_{i=\overline{1,k}}$ are known, all the **scalar** products $\left(\vec{g}_i, \vec{g}_j\right)$ are known except for the coefficients $\{\xi_i\}_{i=\overline{1,k}}$. One can see that the determinant of the system (4.34) is the Gram determinant which is constructed by linearly independent vectors $\{g_i\}_{i=\overline{1,k}}$, i.e. the determinant $\Delta \neq 0$. Therefore, the system (4.34) has a unique solution. After finding this solution, the coefficients $\{\xi_i\}_{i=\overline{1,k}}$ will be known, in turn, the vector $\vec{y}$ will be found.

Suppose that the basis $\{g_i\}_{i=\overline{1,k}}$ is orthogonal, i.e. $\left(\vec{g}_i, \vec{g}_j\right) = 0 \ \ \forall\, i \neq j$.

Thus, the system (4.34) takes the form:

$$\begin{cases} \xi_1\left(\vec{g}_1, \vec{g}_1\right) = \left(\vec{x}, \vec{g}_1\right) \\ \xi_2\left(\vec{g}_2, \vec{g}_2\right) = \left(\vec{x}, \vec{g}_2\right) \\ \xi_k\left(\vec{g}_k, \vec{g}_k\right) = \left(\vec{x}, \vec{g}_k\right) \end{cases} \Rightarrow \xi_i = \frac{\left(\vec{x}, \vec{g}_i\right)}{\left\|\vec{g}_i\right\|^2}, \forall\ i = \overline{1,k}\ .$$

That is, if the basis is orthonormal, then, the coefficients ξ_i are determined by a simpler formula such that $\xi_i = \left(\vec{x}, \vec{g}_i\right)$.

Example 4.8. Find the orthogonal projection of the vector $\vec{x} = (4,-1,-3,4)$ given in the arithmetic space R^4 onto the space L and also its projection onto the orthogonal complement $L^{\perp}$ (i.e. so-called an *orthogonal component*), if it is known that L is spanned by the vectors $\vec{x}_1 = (1;1;1;1)$, $\vec{x}_2 = (1;2;2;-1)$, $\vec{x}_3 = (1;0;0;3)$.

Solution. To find the basis of the subspace L, we use common procedure, when the matrix, whose rows are the given vectors, is reduced to a row echelon form:

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

Hence, the matrix rank is $\mathrm{RgA} = 2$. Also, we choose the vectors $\vec{g}_1 = (1;0;0;3)$ and $\vec{g}_2 = (0;1;1;-2)$ as basis vectors.

Then, the orthogonal projection of the vector $\vec{x}$ onto the subspace L is

$$\vec{y} = proj_L\,\vec{x} = \alpha_1 \cdot \vec{g_1} + \alpha_2 \cdot \vec{g_2}.$$

Coefficients α_1 and α_2 satisfy the system:

$$\left(\vec{z}, \vec{g_i}\right) = 0, \text{ where } \vec{z} = \vec{x} - \vec{y}, \text{ i.e. } \left(\vec{y}, \vec{g_i}\right) = \left(\vec{x}, \vec{g_i}\right),$$

or

$$\begin{cases} \alpha_1(\overrightarrow{g_1},\overrightarrow{g_1})+\alpha_2(\overrightarrow{g_2},\overrightarrow{g_1})=(\vec{x},\overrightarrow{g_1}) \\ \alpha_1(\overrightarrow{g_1},\overrightarrow{g_2})+\alpha_2(\overrightarrow{g_2},\overrightarrow{g_2})=(\vec{x},\overrightarrow{g_2}) \end{cases}$$

Finding the scalar products,

$$(\overrightarrow{g_1},\overrightarrow{g_1})=10, \qquad (\overrightarrow{g_1},\overrightarrow{g_2})=-6, \qquad (\overrightarrow{g_2},\overrightarrow{g_2})=6.$$

$$(\vec{x},\overrightarrow{g_1})=16, \ (\vec{x},\overrightarrow{g_2})=-12$$

allow us to form the system of equations

$$\begin{cases} 10\alpha_1-6\alpha_2=16 \\ -6\alpha_1+6\alpha_2=-12 \end{cases} \Rightarrow \begin{cases} 5\alpha_1-3\alpha_2=8 \\ \alpha_1-\alpha_2=2 \end{cases} \Rightarrow \begin{cases} \alpha_1=2+\alpha_2 \\ 5(2+\alpha_2)-3\alpha_2=8 \end{cases}$$

$$\Rightarrow \begin{cases} \alpha_1=2+\alpha_2 \\ 2\alpha_2=-2 \end{cases} \Rightarrow \begin{cases} \alpha_1=1 \\ \alpha_2=-1 \end{cases}$$

Hence,

$$proj_L\vec{x}=\vec{y}=(1,0,0,3)-(0,1,1,-2)=(1,-1,-1,5),$$

and

$$proj_{L^\perp}\vec{x}=\vec{z}=\vec{x}-\vec{y}=(4,-1,-3,4)-(1,-1,-1,5)=(3,0,-2,-1).$$

Example 4.9. Find the orthogonal projection $\vec{y}$ and $proj_{L^\perp}\vec{x}=\vec{z}$ of the vector $\vec{x}=(1,0,0,3)$ onto the subspace *L*, which is given by SLAE:

$$\begin{cases} \xi_1+\xi_2+2\xi_4=0 \\ \xi_2+\xi_3+\xi_4=0 \\ 2\xi_1-3\xi_2-\xi_3+\xi_4=0 \end{cases}$$

To find the basis of *L*, we need to define the FSS of the given system. Reducing the coefficient matrix to the row echelon form, we can find the rank:

$$A=\begin{pmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 2 & -3 & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & -5 & -1 & -3 \end{pmatrix} \underset{5r_2+r_3}{\overset{r_1-r_2}{\sim}} \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 4 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

$$\text{Rg}A=3.$$

Hence,

$$\begin{cases} 2\xi_3 + \xi_4 = 0 \\ \xi_2 = -\xi_3 - \xi_4 \\ \xi_1 = \xi_3 - \xi_4 \end{cases} \Rightarrow \begin{cases} \xi_4 = -2\xi_3 \\ \xi_2 = -\xi_3 + 2\xi_3 = \xi_3 \\ \xi_1 = \xi_3 + 2\xi_3 = 3\xi_3 \end{cases}$$

Then, assigning a value 1 to the free unknown ξ_3, we can calculate

	ξ_1	ξ_2	ξ_3	ξ_4
$\overrightarrow{a_1}$	3	1	1	-2

So, $\vec{y} = \alpha \cdot \overrightarrow{a_1}$.

Taking into account that $(\vec{y}, \overrightarrow{a_1}) = (\vec{x}, \overrightarrow{a_1})$, we get:

$$\alpha(9+1+1+4) = 3-6;$$

$$15\alpha = -3; \; \alpha = \frac{-3}{15} = -\frac{1}{5}.$$

Therefore,

$$proj_L \vec{x} = \vec{y} = -\frac{1}{5}(3,1,1,-2),$$

and

$$proj_{L^\perp} \vec{x} = \vec{z} = \vec{x} - \vec{y} = (1,0,0,3) + \frac{1}{5}(3,1,1,-2) = \left(\frac{8}{5}, \frac{1}{5}, \frac{1}{5}, \frac{13}{5}\right).$$

4.8. Orthogonal projection and minimization problem

Let us now apply the **scalar** product to the following minimization problem: Given a subspace $L \subset E$ and an arbitrary vector $\vec{x} \in E$, we have to find among all vectors $\vec{y} \in L$ such one that is closest to $\vec{x}$, i.e. to make the distance between the vectors $\vec{x}$ and $\vec{y}$ as small as possible: $\|\vec{x} - \vec{y}\| \to \min$.

The next proposition shows that an orthogonal projection $\vec{x}$ onto the subspace L, i.e. $\vec{y} = proj_L \vec{x}$ is the closest point in L to the vector $\vec{x}$ and that this minimum is, in fact, unique.

Let $\vec{y} = proj_L \vec{x}$, and $\vec{y_1}$ is any vector in L. Then,

$$\vec{x} - \vec{y_1} = \left(\vec{x} - \vec{y}\right) + \left(\vec{y} - \vec{y_1}\right) \qquad (4.35)$$

$\left(\vec{y} - \vec{y_1}\right) \in L$, and as $\vec{z} = \vec{x} - \vec{y} = proj_{L^\perp} \vec{x}$, we have that $\left(\vec{x} - \vec{y}\right) \in L^\perp$, as a result this vector $\left(\vec{x} - \vec{y}\right) \perp L$ and, also, $\left(\vec{x} - \vec{y}\right) \perp \left(\vec{y} - \vec{y_1}\right)$. Thereby, $\left(\left(\vec{x} - \vec{y}\right), \left(\vec{y} - \vec{y_1}\right)\right) = 0$.

Taking **scalar** products of each side of the equality (4.35) with respect to the vector $\vec{x} - \vec{y_1}$, we get:

$$\left(\vec{x} - \vec{y_1}\right)^2 = \left(\left(\vec{x} - \vec{y}\right) + \left(\vec{y} - \vec{y_1}\right), \left(\vec{x} - \vec{y}\right) + \left(\vec{y} - \vec{y_1}\right)\right) = \left(\vec{x} - \vec{y}\right)^2 + \left(\vec{y} - \vec{y_1}\right)^2$$

or

$$\left\|\vec{x} - \vec{y_1}\right\|^2 = \left\|\vec{x} - \vec{y}\right\|^2 + \left\|\vec{y} - \vec{y_1}\right\|^2$$

As $\vec{y} \neq \vec{y_1}$, we can claim that $\left\|\vec{y} - \vec{y_1}\right\|^2 > 0$.

Hence,

$$\left\|\vec{x} - \vec{y_1}\right\|^2 > \left\|\vec{x} - \vec{y}\right\|^2.$$

That is, the distance between the vectors $\vec{x}$ and $\vec{y}$ $\left\|\vec{x} - \vec{y}\right\|$ is smallest, if $\vec{y} = proj_L \vec{x}$. This fact is obvious in space R^3, as seen in Fig. 4.2.

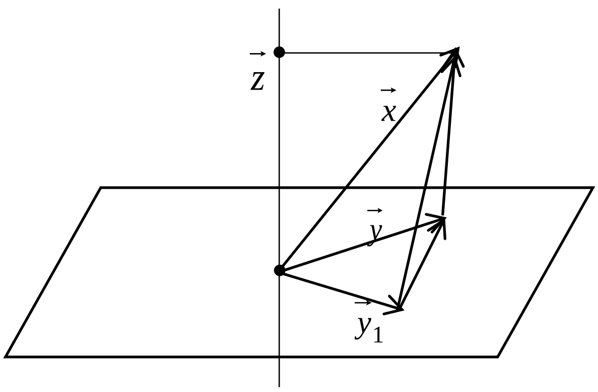


Fig. 4.2

Chapter 5. LINEAR OPERATORS IN EUCLIDEAN SPACE

5.1. Adjoint Operator

Definition. An operator $\boldsymbol{A}^{*}$ is called adjoint to an operator in space E^n if the following condition is satisfied:

$$\left(\boldsymbol{A}\,\vec{x}\,,\vec{y}\right)=\left(\vec{x}\,,\boldsymbol{A}^{*}\,\vec{y}\right). \tag{5.1}$$

for any $\vec{x}$ and $\vec{y}$ of this space E^n.

Note. Adjoint operators mimic the behavior of the transpose matrix on real Euclidean space. Recall that the transpose A^{T} of a real $m \times n$ matrix A satisfies

$$(A\,x, y) = (x, A^T y)$$

for all $x \in R^n$ and $y \in R^m$, where $(*,*)$ is the Euclidean scalar product, i.e. the dot product.

Indeed, in the matrix form, the scalar product of two vectors in a given basis can be written using the Gram's matrix Γ as follows:

$$\left(\vec{x}\,,\vec{y}\right)=X^T\Gamma Y.$$

With this relation, one can rewrite (5.1) in the form:

$$(AX)^T\Gamma Y = X^T\Gamma A^* Y$$

In accordance with the properties of the transpose operation we have

$$X^T A^T\Gamma Y = X^T\Gamma A^* Y$$

That is, the following equality between the matrices occurs

$$A^T\Gamma = \Gamma A^*,$$

Hence,

$$A^* = \Gamma^{-1}A^T\Gamma \tag{5.2}$$

In the case of an orthonormal basis in real Euclidean space, we have that the Gram matrix coincides with an identity matrix, i.e. $\Gamma = I$, then (5.2) takes the form:

$$A^* = A^{\mathrm{T}} \tag{5.3}$$

Theorem. For any linear operator $\boldsymbol{A}$ in the space E^n there exists an adjoint operator $\boldsymbol{A}^*$, and this operator is unique.

The adjoint operator $\boldsymbol{A}^*$ is linear, and its matrix $A^* = \{a_{ik}^*\}$ with respect to any *orthonormal basis* can be deduced from the matrix of the operator $\boldsymbol{A}$ given in the same basis according to the rule:

1. $A^* = \bar{A}^{\mathrm{T}}$ or $a_{ik}^* = \bar{a}_{ki}$ (the conjugate transpose in the case of a complex space);
2. $A^* = A^{\mathrm{T}}$ or $a_{ik}^* = a_{ki}$ (the transpose in the case of a real space).

Proof. We prove this fact for a real space. Let us choose an orthonormal basis $\{g_i\}_{i=\overline{1,n}}$ in the space E^n, where an linear operator $\boldsymbol{A}$ is given, and let $\{\xi_i\}_{i=\overline{1,n}}$ and $\{\eta_i\}_{i=\overline{1,n}}$ be the coordinates of the vectors $\vec{x}$ and $\vec{y}$ in this basis.

Making the scalar product of the vectors $\boldsymbol{A}\,\vec{x}$ and $\vec{y}$, we get

$$\left(\boldsymbol{A}\,\vec{x}\,,\vec{y}\right) = \left(\boldsymbol{A}\sum_{i=1}^{n}\xi_i\,\vec{g_i}\,,\sum_{j=1}^{n}\eta_j\,\vec{g_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\eta_j\left(\boldsymbol{A}\,\vec{g_i}\,,\vec{g_j}\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\,\eta_j\left(a_{1i}\,\vec{g_1} + a_{2i}\,\vec{g_2} + \cdots + a_{ni}\,\vec{g_n}\,,\vec{g_j}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\eta_j a_{ji}$$

Here $A = \{a_{ij}\}_{\substack{i=\overline{1,n}\\ j=\overline{1,n}}}$ is the matrix representation of the operator $\boldsymbol{A}$ in the orthonormal basis $\{g_i\}_{i=\overline{1,n}}$. Also, we transpose this matrix as $A^T = \{a_{ji}\}_{\substack{i=\overline{1,n}\\ j=\overline{1,n}}}$.

One can prove that the matrix A^T corresponds to a linear operator $\boldsymbol{A}^T$ that is an adjoint of $\boldsymbol{A}$, i.e. $\boldsymbol{A}^T = \boldsymbol{A}^*$. For this purpose, let us consider the **dot** product:

$$\left(\vec{x}\,,\boldsymbol{A}^T\,\vec{y}\right) = \left(\sum_{i=1}^{n}\xi_i\,\vec{g_i},\sum_{j=1}^{n}\eta_j\boldsymbol{A}^T\,\vec{g_i}\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\,\eta_j\left(\vec{g_i}\,,a_{j1}\,\vec{g_1} + a_{j2}\,\vec{g_2} + \cdots + a_{jn}\,\vec{g_n}\right) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \xi_i \eta_j (\overrightarrow{g_i}, a_{ji} \overrightarrow{g_i}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \xi_i \eta_j a_{ji}$$

So, $(\boldsymbol{A}\,\vec{x}\,,\vec{y}) = (\vec{x}\,,\boldsymbol{A}^T\,\vec{y}) \Rightarrow \boldsymbol{A}^T = \boldsymbol{A}^*$. ■

<u>*Remark.*</u> It should be noted that this theorem is not valid if a space has *infinite dimension*. It is also emphasized that no such simple relationship exists between the matrices representing A and A^* if the basis is *not orthonormal*. Otherwise, if the basis is *not orthonormal*, the matrix of an adjoint operator is defined as

1. $A^* = \Gamma^{-1}\bar{A}^{\mathrm{T}}\Gamma$ (in the case of a complex space);
2. $A^* = \Gamma^{-1}A^{\mathrm{T}}\Gamma$ (in the case of a real space).

Thus, it is seen one useful property of orthonormal bases.

<u>Properties of an adjoint operator</u>

Let E^n be a real Euclidean space, then

1) for given an identity operator, we have

$$\boldsymbol{I}^* = \boldsymbol{I}$$

According to the properties of identity matrix, the scalar product of appropriate vectors leads us to the following result:

$$(\boldsymbol{I}\,\vec{x}\,,\vec{y}) = (\vec{x},\vec{y}) = (\vec{x},\boldsymbol{I}\,\vec{y}) \Rightarrow \boldsymbol{I} = \boldsymbol{I}^*$$

2) An adjoint of an adjoint linear operator is the linear operator itself

$$(\boldsymbol{A}^*)^* = \boldsymbol{A}$$

Indeed, the scalar product of appropriate vectors gives us

$$(\boldsymbol{A}\,\vec{x}\,,\vec{y}) = (\vec{x}\,,\boldsymbol{A}^*\,\vec{y}) = (\boldsymbol{A}^*\,\vec{y}\,,\vec{x}) = (\vec{y}\,,(\boldsymbol{A}^*)^*\,\vec{x}) = ((\boldsymbol{A}^*)^*\,\vec{x}\,,\vec{y})$$

Equating both sides of the equality, we have $\boldsymbol{A} = (\boldsymbol{A}^*)^*$

It follows from this fact that twice transposed matrix coincides with the matrix itself, i.e. $(A^T)^T = A$

3) $(\boldsymbol{A}+\boldsymbol{B})^* = \boldsymbol{A}^* + \boldsymbol{B}^*$

Indeed, using the properties of the scalar product of appropriate vectors, we obtain

$$\left(\vec{x}\,,(\boldsymbol{A}+\boldsymbol{B})^*\,\vec{y}\right) = \left((\boldsymbol{A}+\boldsymbol{B})\,\vec{x}\,,\vec{y}\right) = \left(\boldsymbol{A}\,\vec{x}\,,\vec{y}\right) + \left(\boldsymbol{B}\,\vec{x}\,,\vec{y}\right) =$$

$$= \left(\vec{x},\boldsymbol{A}^*\,\vec{y}\right) + \left(\vec{x},\boldsymbol{B}^*\,\vec{y}\right) = \left(\vec{x}\,,\boldsymbol{A}^*\,\vec{y} + \boldsymbol{B}^*\,\vec{y}\right) = \left(\vec{x}\,,(\boldsymbol{A}^*+\boldsymbol{B}^*)\,\vec{y}\right)$$

Hence, the matrix transposed of the sum of two matrices is equal to the sum of their transposed matrices, i.e. $(A+B)^T = A^T + B^T$

4) $(\boldsymbol{A}\cdot\boldsymbol{B})^* = \boldsymbol{A}^*\cdot\boldsymbol{B}^*$

Indeed, by making the **scalar** product of appropriate vectors and performing transformations, we have

$$\left(\vec{x}\,,(\boldsymbol{A}\cdot\boldsymbol{B})^*\,\vec{y}\right) = \left((\boldsymbol{A}\cdot\boldsymbol{B})\,\vec{x}\,,\vec{y}\right) = \left(\boldsymbol{A}\left(\boldsymbol{B}\,\vec{x}\right),\vec{y}\right) = \left(\boldsymbol{B}\,\vec{x}\,,\boldsymbol{A}^*\tilde{y}\right) =$$

$$\left(\vec{x},\boldsymbol{B}^*\left(\boldsymbol{A}^*\,\vec{y}\right)\right) = \left(\vec{x},\boldsymbol{B}^*\boldsymbol{A}^*\,\vec{y}\right), \Rightarrow (\boldsymbol{A}\cdot\boldsymbol{B})^* = \boldsymbol{A}^*\cdot\boldsymbol{B}^*$$

Therefore, the matrix transposed of a product of two matrices is equal to the product of the transposed matrices in reverse order: $(AB)^T = B^T\cdot A^T$.

5) If a linear operator $\boldsymbol{A}^{-1}$ inverse to a linear operator $\boldsymbol{A}$ exists, then $(\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^*)^{-1}$.

Indeed, as $\boldsymbol{A}\cdot\boldsymbol{A}^{-1} = \boldsymbol{I}$, in turn, as $\boldsymbol{I} = \boldsymbol{I}^*$ and

$\boldsymbol{I}^* = (\boldsymbol{A}\cdot\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^{-1})^*\cdot\boldsymbol{A}^* = \boldsymbol{I}$.

We have $(\boldsymbol{A}^{-1})^*\cdot\boldsymbol{A}^* = \boldsymbol{I}$, that means the fact

$$(\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^*)^{-1}.$$

Therefore, the following statement is true for matrices $(A^{-1})^T = (A^T)^{-1}$

Example 5.1. A linear operator $\boldsymbol{A}$ in the basis $\overrightarrow{a_1} = (3,-1)$, $\overrightarrow{a_2} = (2,1)$ has a matrix $A_a = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$. Find the matrix of the adjoint operator $\boldsymbol{A}_e^*$ in the natural basis and $\boldsymbol{A}_e^*$ in the current basis.

Solution. Construct a matrix $\boldsymbol{A}$ in the orthonormal basis, using the formula

$A_e = C \cdot A_a \cdot C^{-1}$ where C is the matrix of the transition from the basis $\{\overrightarrow{e_i}\}_{i=1,2}$ to $\{\overrightarrow{a_i}\}_{i=1,2}$. This transition matrix has a form:

$$\left.\begin{matrix} \overrightarrow{a_1} = 3\,\overrightarrow{e_1} - e_2 \\ \overrightarrow{a_2} = 2\,\overrightarrow{e_1} + e_2 \end{matrix}\right\} \Rightarrow C = \begin{vmatrix} 3 & 2 \\ -1 & 1 \end{vmatrix},$$

then

$$C^{-1} = \frac{1}{5}\begin{vmatrix} 1 & -2 \\ 1 & 3 \end{vmatrix},$$

$$A_e = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & -3 \\ 0 & 0 \end{pmatrix}$$

$$= \frac{1}{5}\begin{pmatrix} -3 & -9 \\ 1 & 3 \end{pmatrix}$$

$$A_e^* = \frac{1}{5}\begin{pmatrix} -3 & 1 \\ -9 & 3 \end{pmatrix}$$

Let us now construct the matrix A_a^* using the formula:

$$A_e^* = C \cdot A_a^* \cdot C^{-1} \Rightarrow A_a^* = C^{-1} \cdot A_e^* \cdot C$$

$$A_a^* = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix} \cdot \frac{1}{5}\begin{pmatrix} -3 & 1 \\ -9 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} -10 & -5 \\ -30 & -15 \end{pmatrix} =$$

$$= -\frac{1}{5}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} =$$

$$= -\frac{1}{5}\begin{pmatrix} -10 & -5 \\ 20 & 10 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -4 & -2 \end{pmatrix}.$$

5.2. Unitary and Orthogonal Operators

Definition: An operator ***U*** on a Euclidean space E^n is called a ***unitary operator*** (the underlying field is complex (Hilbert space)) or ***orthogonal operator*** (the underlying field is real) if the operator maps orthonormal bases to orthonormal bases.

One can demonstrate that if ***U*** is a *unitary* (*orthogonal*) *operator* in a space E^n and the vectors $\{\overrightarrow{g_i}\}_{i=\overline{1,n}}$ form an orthonormal basis in this space, i.e.

$(\overrightarrow{g_r}, \overrightarrow{g_s}) = \begin{cases} 1, s = r \\ 0, s \neq r \end{cases}$, then the vectors $\{\boldsymbol{U}\,\overrightarrow{g_i}\}_{i=\overline{1,n}}$ form an orthonormal basis as well, that is

$$(\boldsymbol{U}\,\overrightarrow{g_r}, \boldsymbol{U}\,\overrightarrow{g_s}) = \begin{cases} 1, s = r \\ 0, s \neq r \end{cases} \tag{5.4}$$

Let $U = \{u_{\mathrm{ik}}\}$ represent the matrix of the operator $\boldsymbol{U}$ in the basis $\{\overrightarrow{g_i}\}_{i=\overline{1,n}}$. Then,

$$\boldsymbol{U}\,\overrightarrow{g_r} = \sum_{i=1}^{n} u_{ir}\,\overrightarrow{g_i}\,, \qquad (r = \overline{1,n})$$

Similarly

$$\boldsymbol{U}\,\overrightarrow{g_s} = \sum_{j=1}^{n} u_{js}\,\overrightarrow{g_j}\,, \qquad (s = \overline{1,n})$$

Then, since given the vectors $\{\overrightarrow{g_i}\}_{i=\overline{1,n}}$ as an orthonormal basis, we obtain

$$(\boldsymbol{U}\,\overrightarrow{g_r}, \boldsymbol{U}\,\overrightarrow{g_s}) = \left(\sum_{i=1}^{n} u_{ir}\,\overrightarrow{g_i}\,, \sum_{j=1}^{n} u_{js}\,\overrightarrow{g_j} \right) = \sum_{i=1}^{n} u_{ir}\bar{u}_{is}.$$

Since $\bar{u}_{is}$ are either complex conjugate or transpose values, the last equality takes the form

$$\sum_{i=1}^{n} u_{ir}\bar{u}_{is} = \begin{cases} 0, s \neq r \\ 1, s = r \end{cases}$$

The expansion of this sum results in two equations:

$$\begin{gathered} u_{1r} \cdot u_{1s} + u_{2r} \cdot u_{2s} + \cdots + u_{nr} \cdot u_{ns} = 0, (r \neq s) \\ {u_{1r}}^2 + {u_{2r}}^2 + \cdots + {u_{nr}}^2 = 1, (r = s) \end{gathered} \tag{5.5}$$

Thus, it has been proven that

$$(\boldsymbol{U}\,\overrightarrow{g_r}, \boldsymbol{U}\,\overrightarrow{g_s}) = (\overrightarrow{g_r}, \overrightarrow{g_s}) = \begin{cases} 0, s \neq r \\ 1, s = r \end{cases} = \delta_{rs} \ \blacksquare$$

The next theorem gives alternative characterizations of these operators.

Theorem 1. The following conditions on a *unitary* (*orthogonal*) *operator* $\boldsymbol{U}$ are equivalent:

(**i**) if $\boldsymbol{U}\boldsymbol{U}^* = \boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{I}$, then $\boldsymbol{U}^*$ coincides with the inverse of the operator $\boldsymbol{U}$, i.e.

$$\boldsymbol{U}^* = \boldsymbol{U}^{-1} \tag{5.6}$$

(**ii**) $\boldsymbol{U}$ preserves **scalar** products, that is, for every $\vec{x}, \vec{y} \in E^n$, the equality takes place:

$$(\boldsymbol{U}\vec{x}, \boldsymbol{U}\vec{y}) = (\vec{x}, \vec{y}) \tag{5.7}$$

(**iii**) $\boldsymbol{U}$ preserves **norm (length),** that is, for every $\vec{x} \in E^n$,

$$\|\boldsymbol{U}\vec{x}\| = \|\vec{x}\| \tag{5.8}$$

(i) *Proof.*

$$(\boldsymbol{U}\,\vec{x}\,, \boldsymbol{U}\,\vec{y}) = (\vec{x}\,, \vec{y}) \Rightarrow .(\vec{x}\,, \boldsymbol{U}^*\boldsymbol{U}\,\vec{y}) = (\vec{x}\,, \boldsymbol{I}\,\vec{y})\,, \forall\, \vec{x}\,, \vec{y}$$

Then, we have that

$$\boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{I} \Rightarrow .\boldsymbol{U}^* = \boldsymbol{U}^{-1}\ ■$$

(ii) *Proof.* Indeed, let $\{\vec{e_i}\}_{i=\overline{1,n}}$ be an orthonormal basis. The operator action gives $\boldsymbol{U}\,\vec{e_i} = \vec{e_i}'$, where $\{\vec{e_i}'\}_{i=\overline{1,n}}$ is also an orthonormal basis. Take any two elements $\vec{x} = \sum_{i=1}^{n} \xi_i\, \vec{e_i}$ and $\vec{y} = \sum_{j=1}^{n} \eta_j\, \vec{e_j}$. Then:

$$(\boldsymbol{U}\,\vec{x}\,, \boldsymbol{U}\,\vec{y}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \xi_i \overline{\eta_j} (\boldsymbol{U}\,\vec{e_i}\,, \boldsymbol{U}\,\vec{e_j}) = \sum_{i=1}^{n}\sum_{j=1}^{n} \xi_i \overline{\eta_j} \left(\vec{e_i}'\,, \vec{e_j}'\right) =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \xi_i \overline{\eta_j} = (\vec{x}\,, \vec{y}).\ ■$$

(iii) *Proof.* It follows from the previous proof, considering the scalar product as a multiplication of the vector $\boldsymbol{U}\,\vec{x}$ by itself.

By the Theorem 1, we obtain the following results.

Theorem 2. A complex matrix U represents a unitary operator $\boldsymbol{U}$ (relative to an orthonormal basis) if and only if $U^* = U^{-1}$.

Theorem 3. A real matrix U represents an orthogonal operator $\boldsymbol{U}$ (relative to an orthonormal basis) if and only if $U^T = U^{-1}$.

The above theorems motivate the following definitions:

Definition: A complex matrix U for which $U^* = U^{-1}$ is called a *unitary matrix*.

In other words, an invertible complex square matrix U is unitary if its conjugate transpose U^* is also its inverse, that is, if $U^* = U^{-1}$.

Definition: A real matrix U for which $U^T = U^{-1}$ is called an *orthogonal matrix*.

Note: The entries of the unitary (orthogonal) matrix of a unitary (orthogonal) operator $\boldsymbol{U}$

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix}$$

satisfy the properties (5.5).

That is, the sum of the products of the elements of any two columns is equal to 0, but multiplication of the column by itself gives 1.

Thus, if the columns of the unitary (orthogonal) matrix are considered as vectors, then these vectors will form an orthonormal basis.

Remark. If $\boldsymbol{U}$ is an orthogonal operator, then $\boldsymbol{U}^*$ will also be an *orthogonal operator*.

Indeed,

$$\left(\boldsymbol{U}^* \vec{x}, \boldsymbol{U}^* \vec{y}\right) = \left(\vec{x}, \boldsymbol{U}^{**}(\boldsymbol{U}^* y)\right) = \left(\vec{x}, \boldsymbol{U}\boldsymbol{U}^{-1} \vec{y}\right) = \left(\vec{x}, \boldsymbol{I} \vec{y}\right) = \left(\vec{x}, \vec{y}\right).$$

Also, in the case of an orthogonal matrix, the columns of the matrix U^T (i.e.

rows of the matrix *U*) also form an *orthonormal system of vectors*, i.e.

$$\sum_{s=1}^{n} u_{is} \cdot u_{ks} = 0\,, \qquad i \neq k\,,$$

$$\sum_{k=1}^{n} u_{ik}^2 = 1, \quad \forall i = \overline{1,n}.$$

Thus, the matrix of an orthogonal operator in any orthonormal basis is *orthogonal*.

An example of the orthogonal matrix is the standard matrix for the *counterclockwise rotation* of R^2 through an angle θ, i.e.

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

This matrix is orthogonal for all choices of θ since

$$A^T A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

Also, the *reflection matrix* that **maps** each point into its symmetric image about the *x*-axis is orthogonal.

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

Indeed,

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

The *determinant* of an orthogonal matrix is equal to ± 1.

Indeed, it follows from equality $U \cdot U^T = I$ that $det(U \cdot U^T) = det\, U \cdot det\, U^T = det\, I = 1$. Since $\det U = \det U^T$, then we have $(\det U)^2 = 1$ as a result, $\det U = \pm 1$

Continuing the previous examples, one can show that the determinants associated with those matrices are

$$\det A = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = +1$$

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

Theorem. The *eigenvalues* of an orthogonal operator are equal to ± 1

Proof. Let $\vec{x}$ be an eigenvector, and λ is the corresponding eigenvalue of the orthogonal operator ***U***. Then

$$(\vec{x}, \vec{x}) = (\boldsymbol{U}\vec{x}, \boldsymbol{U}\vec{x}) = (\lambda\vec{x}, \lambda\vec{x}) = |\lambda|^2(\vec{x}, \vec{x})$$

Since $(\vec{x}, \vec{x}) \neq 0$, it follows from the last equality $(\vec{x}, \vec{x}) = |\lambda|^2(\vec{x}, \vec{x})$ that $|\lambda|^2 = 1 \Rightarrow \lambda = \pm 1$. ■

<u>*Remark.*</u> The matrix of transition from one orthonormal basis to another orthonormal basis is orthogonal.

Indeed, let $\{e_i\}_{i=\overline{1,n}}$ and $\{e_i^*\}_{i=1,n}$ be two orthonormal basis. Then

$$\begin{cases} \overrightarrow{e_1^*} = a_{11}\vec{e}_1 + a_{21}\overrightarrow{e_2} + \cdots + a_{n1}\vec{e}_n \\ \overrightarrow{e_2^*} = a_{12}e_1 + a_{22}\overrightarrow{e_2} + \cdots + a_{n2}\overrightarrow{e_n} \\ \dots\dots\dots. \\ \overrightarrow{e_n} = a_{1n}\overrightarrow{e_1} + a_{2n}\overrightarrow{e_2} + \cdots + a_{nn}\overrightarrow{e_n} \end{cases}.$$

The transition matrix has the following form:

$$T = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Let us consider scalar products (e_i^*, e_k^*). For example, $(e_1^*, e_2^*) = 0$ written in the expanded form, leads to

$$a_{11} \cdot a_{12} + a_{21} \cdot a_{22} + \cdots + a_{n1} \cdot a_{n2} = 0,$$

and the expansion of $(e_1^*, e_1^*) = 1$ gives

$$a_{11}^2 + a_{21}^2 + \cdots + a_{n1}^2 = 1$$

Similarly, for any $\overrightarrow{e_i^*}$ and $\overrightarrow{e_k^*}$, i.e. $\left(\overrightarrow{e_i^*}, \overrightarrow{e_k^*}\right)$ at $i \neq k$, we get

$$\sum_{j=1}^{n} a_{ij} \cdot a_{kj} = 0,$$

and $\left(\overrightarrow{e_i^*}, \overrightarrow{e_k^*}\right)$ at $i = k$ gives

$$\sum_{j=1}^{n} a_{ij}^2 = 1$$

Therefore, the matrix T is orthogonal.

5.3. Self-adjoint Operators

<u>Definition</u>. An operator $\boldsymbol{A}$ that coincides with its adjoint, i.e. $\boldsymbol{A}^* = \boldsymbol{A}$ is called a ***self-adjoint*** (or ***hermitian***) ***operator***.

If $\boldsymbol{A}$ is a self-adjoint operator, then $\forall \vec{x}, \vec{y} \in E^n$ the following equality holds:

$$\left(\boldsymbol{A}\vec{x}, \vec{y}\right) = \left(\vec{x}, \boldsymbol{A}\vec{y}\right) \tag{5.9}$$

Let $A = [a_{ik}]_{i,k=\overline{1,n}}$ be a matrix of a self-adjoint operator in an orthonormal basis. Then,

1. if a space is over the field of real numbers, the matrix satisfies $A = A^T$ or $a_{ik} = a_{ki}$.

This matrix is called ***symmetric***.

2. if a space is over the field of complex numbers, the matrix is a conjugate transpose of A, i.e. $A = A^*$ or $a_{ik} = \bar{a}_{ki}$.

This matrix is called ***Hermitian***.

The following properties of a self-adjoint operator are important:

Theorem 1. All the roots of the characteristic polynomial of a self-adjoint operator $\boldsymbol{A}$ are *real*, i.e. the eigenvalues of a self-adjoint operator are *real*.

Proof. Let λ be an eigenvalue of a self-adjoint operator $\boldsymbol{A}$ and $\vec{x}$ be the corresponding eigenvector, i.e.

$$\boldsymbol{A}\vec{x} = \lambda\vec{x}, \text{ where } \vec{x} \neq 0.$$

Because $(\boldsymbol{A}\vec{x}, \vec{x}) = (\vec{x}, \boldsymbol{A}\vec{x})$, then $(\lambda\vec{x}, \vec{x}) = \bar{\lambda}(\vec{x}, \vec{x})$. Given that $(\vec{x}, \vec{x}) \neq 0 \Rightarrow \lambda = \bar{\lambda}$, i.e. λ is a real number. ■

Theorem 2. Eigenvectors of a self-adjoint operator $\boldsymbol{A}$, which correspond to distinct eigenvalues, are *orthogonal*.

Proof. Let λ_1 and λ_2 be distinct eigenvalues of a self-adjoint operator $\boldsymbol{A}$, and $\vec{x}_1$, $\vec{x}_2$ are the corresponding eigenvectors.

Since $(\boldsymbol{A}\vec{x_1}, \vec{x_2}) = (\vec{x_1}, \boldsymbol{A}\vec{x_2})$ and $\boldsymbol{A}\vec{x}_1 = \lambda_1\vec{x}_1$, $\boldsymbol{A}\vec{x}_2 = \lambda_2\vec{x}_2$, then
$\lambda_1(\vec{x}_1, \vec{x}_2) = \lambda_2(\vec{x}_1, \vec{x}_2), \Rightarrow (\lambda_1 - \lambda_2)(\vec{x}_1, \vec{x}_2) = 0,$
It follows from this equality that since $\lambda_1 \neq \lambda_2$, we have $(\vec{x}_1, \vec{x}_2) = 0$.
That is, $\vec{x}_1$ and $\vec{x}_2$ are the orthogonal vectors. ■

Theorem 3. A self-adjoint operator has a simple structure. (Without proof).

Theorem 4. For any self-adjoint operator in Euclidean space, there is an orthonormal basis composed of the eigenvectors of this operator.

Proof. Let an eigenvalue λ of this operator have a multiplicity of "*k*". Since $\boldsymbol{A}$ is an operator of simple structure, this eigenvalue corresponds to "*k*" linearly independent eigenvectors. These vectors form a subspace of dimension "*k*". We choose an orthogonal basis in this subspace. Eigenvectors corresponding to other eigenvalues will be orthogonal to this subspace. Doing the same with the vectors corresponding to other eigenvalues accounting for their multiplicity, we obtain an orthonormal basis of the whole space. ■

Theorem 5. A matrix of a self-adjoint operator in some orthonormal basis is represented by a diagonal matrix relative to this basis (In other words, a matrix of a self-adjoint operator is diagonalizable)

Proof. Let λ_1 be one of eigenvalue of a self-adjoint operator $\boldsymbol{A}$. By Theorem 1, λ_1 is a real number. Let $\vec{e}_1$ be an eigenvector corresponding to this eigenvalue, i.e. $\boldsymbol{A}\vec{e}_1 = \lambda_1 \vec{e}_1$. The vector $\vec{e}_1$ can be considered as a unit length vector, otherwise it could be replaced by a unit eigenvector $\frac{\vec{e}_1}{\|\vec{e}_1\|}$ associated with the same eigenvalue.

We denote as R_1 a one-dimensional subspace generated by the vector $\vec{e}_1$. Its orthogonal complement $R_1^{\perp}$ will be invariant with respect to the operator $\boldsymbol{A}$.

Recall that subspace $R_1 \subset R$ is called ***invariant*** with respect to a linear operator $\boldsymbol{A}$ if the image $\boldsymbol{A}\vec{x}$ of each vector $\vec{x} \in R_1$ also belongs R_1, i.e., if $\vec{x} \in R_1 \Rightarrow \boldsymbol{A}\vec{x} \in R_1$. The operator $\boldsymbol{A}$ remains, of course, to be self-adjoint.

Let λ_2 be another real eigenvalue of the operator $\boldsymbol{A}$ in the subspace $R_1^{\perp}$. The corresponding eigenvector is denoted by $\vec{e_2}$, then

$$\boldsymbol{A}\vec{e}_2 = \lambda_2 \vec{e}_2.$$

Let R_2 be an invariant subspace generated by the vectors $\vec{e}_1$ and $\vec{e_2}$, then the subspace $R_2^{\perp}$ will also be invariant relatively $\boldsymbol{A}$. Continue this construction, we find *n* pairwise orthogonal, and hence linearly independent unit eigenvectors of the operator $\boldsymbol{A}$. In the basis consisting of these vectors, the matrix A of the operator $\boldsymbol{A}$ is reduced to a diagonal form:

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \blacksquare$$

<u>Algorithm for constructing an orthonormal basis of a self-adjoint operator</u>

1. Compose the characteristic equation of the linear operator $\det(A - \lambda I) = 0$.

2. Find all eigenvalues.

3. Find associated eigenvectors.

4. With obtained eigenvectors, construct the orthonormal basis.

Example 5.2. Construct an orthonormal basis using the eigenvectors of a linear operator, which is given by a matrix A_f in an orthonormal basis of vectors $\overrightarrow{f_1}, \overrightarrow{f_2}, \overrightarrow{f_3}$, as

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

Compose the transition matrix to a new basis and the matrix of the operator A_e in a new basis.

Solution. Find the eigenvalues and eigenvectors of the operator $\boldsymbol{A}$. For this purpose, we write a characteristic equation: $det(A_f - \lambda I) = 0$, i.e.

$$\begin{vmatrix} 2/3-\lambda & 2/3 & -1/3 \\ 2/3 & -1/3-\lambda & 2/3 \\ -1/3 & 2/3 & 2/3-\lambda \end{vmatrix} = 0, \Rightarrow -\lambda^3 + \lambda^2 + \lambda - 1 = 0,$$

Solving the equation, we get the eigenvalues: λ_1= -1,λ_2=λ_3= 1, which are distinct reals. Making the solution of the system of equations $(A - \lambda I) \cdot X = 0$ for each eigenvalue, we find eigenvectors associated with it:

λ_1= -1, then

$$\begin{pmatrix} 5/3 & 2/3 & -1/3 \\ 2/3 & 2/3 & 2/3 \\ -1/3 & 2/3 & 5/3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Solving this system, we get the first eigenvector as $\overrightarrow{e_1} = (1; -2; 1)$. Then,

λ_2=λ_3= 1 (the multiplicity is 2)

$$\begin{pmatrix} -1/3 & 2/3 & -1/3 \\ 2/3 & -4/3 & 2/3 \\ -1/3 & 2/3 & -1/3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The solution of this system leads to the second and third eigenvectors in the form: $\overrightarrow{e_2} = (2;1;0)$ and $\overrightarrow{e_3} = (-1;0;1)$, respectively.

The vectors $\overrightarrow{e_1}$, $\overrightarrow{e_2}$, and $\overrightarrow{e_3}$ are linearly independent, i.e. they can form a basis. In doing so, one can see that pairwise orthogonal $\overrightarrow{e_1} \perp \overrightarrow{e_2}$ and $\overrightarrow{e_1} \perp \overrightarrow{e_3}$, but $\overrightarrow{e_2}$ is not perpendicular to $\overrightarrow{e_3}$.

Applying orthogonalzation procedure, we can construct an orthonormal basis:

$\overrightarrow{g_1} = \overrightarrow{e_1} = (1;-2;1)$,

$\overrightarrow{g_2} = \overrightarrow{e_2} = (2;1;0)$,

$\overrightarrow{g_3} = \overrightarrow{e_3} + \alpha_{32}\,\overrightarrow{g_2}$, where $\alpha_{32} = -\frac{(\overrightarrow{e_3},\overrightarrow{g_2})}{(\overrightarrow{g_2},\overrightarrow{g_2})}$, then $\overrightarrow{g_3} = (-1/5; 2/5; 1/5)$

These vectors are orthogonal, but are not still normalized. Their lengths (norms) are

$$\|\overrightarrow{g_1}\| = \sqrt{6}; \|\overrightarrow{g_2}\| = \sqrt{5}; \|\overrightarrow{g_3}\| = \sqrt{6}/5.$$

Finally, we obtain an orthonormal basis formed by the vectors:

$$\overrightarrow{e_1}^* = \frac{\sqrt{6}}{6}(1;-2;1)$$

$$\overrightarrow{e_2}^* = \frac{\sqrt{5}}{5}(2;1;0)$$

$$\overrightarrow{e_3}^* = \frac{\sqrt{6}}{6}(-1;2;5)$$

The transition matrix from the basis of vectors $\left\{\overrightarrow{f_i}\right\}_{i=1,2,3}$ to the basis of eigenvectors $\left\{\overrightarrow{e_i}^*\right\}_{i=1,2,3}$ has a form:

$$C = \begin{pmatrix} \frac{\sqrt{6}}{6} & \frac{2\sqrt{5}}{5} & -\frac{\sqrt{6}}{6} \\ -\frac{2\sqrt{6}}{6} & \frac{\sqrt{5}}{5} & \frac{2\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} & 0 & \frac{5\sqrt{6}}{6} \end{pmatrix}$$

The matrix of the operator in the basis of eigenvectors takes the form:

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

We can check this matrix by using the change of basis formula as follows:

$$A_e = C^T \cdot A \cdot C =$$

$$\begin{pmatrix} \frac{\sqrt{6}}{6} & \frac{2\sqrt{5}}{5} & -\frac{\sqrt{6}}{6} \\ -\frac{2\sqrt{6}}{6} & \frac{\sqrt{5}}{5} & \frac{2\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} & 0 & \frac{5\sqrt{6}}{6} \end{pmatrix}^T \cdot \frac{1}{3}\begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} \frac{\sqrt{6}}{6} & \frac{2\sqrt{5}}{5} & -\frac{\sqrt{6}}{6} \\ -\frac{2\sqrt{6}}{6} & \frac{\sqrt{5}}{5} & \frac{2\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} & 0 & \frac{5\sqrt{6}}{6} \end{pmatrix} =$$

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

5.4. Spectral Decomposition of a Self-adjoint Operator

Let's consider a self-adjoint operator $\boldsymbol{A}$ in a Euclidean space E^n. Let $\vec{e}_1$, ..., $\vec{e}_n$ be an orthonormal basis in this space formed by eigenvectors of the operator $\boldsymbol{A}$ associated with eigenvalues λ_1,...,λ_n. An arbitrary vector $\vec{x} \in E^n$ can be decomposed along the basis of vectors $\{\vec{e}_i\}_{i=\overline{1,n}}$.

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n = \sum_{j=1}^{n} x_j\vec{e}_j$$

Making consequently scalar products of vectors $\vec{x}$ and $\vec{e}_j$, where $j = 1, \dots, n$, we can express that $(\vec{x}, \vec{e}_j) = x_j$. That is, one can write

$$\vec{x} = \sum_{j=1}^{n} (\vec{x}, \vec{e}_j)\vec{e}_j$$

Since, for any basis vector $\vec{e}_j$, $j = 1, \dots, n$, the action of the operator is $\boldsymbol{A}(\vec{e}_j) = \lambda_j\vec{e}_j$, where λ_j is an eigenvalue, then the image of the vector one can be presented as follows:

$$\boldsymbol{A}\,\vec{x} = \boldsymbol{A}\left(\sum_{j=1}^{n}(\vec{x},\vec{e}_j)\vec{e}_j\right) = \sum_{j=1}^{n}(\vec{x},\vec{e}_j)\boldsymbol{A}\vec{e}_j = \sum_{j=1}^{n}\lambda_j(\vec{x},\vec{e}_j)\vec{e}_j$$

The expression $(\vec{x},\vec{e}_j)\vec{e}_j$ corresponds an orthogonal projection of the vector $\vec{x}$ onto a one-dimensional eigenspace of the operator $\boldsymbol{A}$ generated by the eigenvector $\vec{e}_j$. Thus, we can introduce an *orthogonal projection operator* denoted as $\boldsymbol{P}_j$ such that

$$\boldsymbol{P}_j\vec{x} = (\vec{x},\vec{e}_j)\vec{e}_j.$$

The operator is called a ***projector*** on the subspace generated by an eigenvector $\vec{e}_j$.

It follows from the feature of the scalar product that the projector is a self-adjoint operator. Indeed,

$$(\boldsymbol{P}_j\vec{x},\vec{y}) = \left((\vec{x},\vec{e}_j)\vec{e}_j,\vec{y}\right) = (\vec{x},\vec{e}_j)(\vec{e}_j,\vec{y}), \text{ and}$$

$$(\vec{x},\boldsymbol{P}_j\vec{y}) = (\vec{x},(\vec{y},\vec{e}_j)\vec{e}_j) = (\overline{\vec{y},\vec{e}_j})(\vec{x},\vec{e}_j) = (\vec{e}_j,\vec{y})(\vec{x},\vec{e}_j)$$

Whence,

$$(\boldsymbol{P}_j\vec{x},\vec{y}) = (\vec{x},\boldsymbol{P}_j\vec{y})$$

<u>*Properties of the projector*</u>*:*

1. $\boldsymbol{P}_j{}^2 = \boldsymbol{P}_j$
2. $\boldsymbol{P}_j\boldsymbol{P}_k = 0$, for $k \neq j$

Indeed,

$$(\boldsymbol{P}_j\boldsymbol{P}_k)\vec{x} = \boldsymbol{P}_j(\boldsymbol{P}_k\vec{x}) = \boldsymbol{P}_j(\vec{x},\vec{e}_k)\vec{e}_k = (\vec{x},\vec{e}_k)\boldsymbol{P}_j\vec{e}_k = (\vec{x},\vec{e}_k)(\vec{e}_k,\vec{e}_j)\vec{e}_j =$$

$$= \begin{cases}(\vec{x},\vec{e}_j)\vec{e}_j = \boldsymbol{P}_j & if\ k = j\\ 0 & if\ k \neq j\end{cases}$$

With the projector, the image of the vector one can be presented in the form:

$$\boldsymbol{A}\,\vec{x} = \sum_{j=1}^{n} \lambda_j \boldsymbol{P}_j \vec{x}$$

Therefore, we have

$$\boldsymbol{A} = \sum_{j=1}^{n} \lambda_j \boldsymbol{P}_j = \lambda_1 \boldsymbol{P}_1 + \lambda_2 \boldsymbol{P}_2 + \cdots + x_n \boldsymbol{P}_n \tag{5.10}$$

This expansion (5.10) is called the ***spectral decomposition of a self-adjoint operator***.

According to the feature of the projector, we have

$$\boldsymbol{A}^2 = \sum_{j=1}^{n} {\lambda_j}^2 \boldsymbol{P}_j$$

In general, for any number $s > 0$, it is valid that

$$\boldsymbol{A}^s = \sum_{j=1}^{n} {\lambda_j}^s \boldsymbol{P}_j$$

Let's consider a polynomial of the i-th order with respect to the parameter λ, i.e. $p(\lambda) = \sum_{j=1}^{n} a_j \lambda^j$. Then, this polynomial with respect to the operator $\boldsymbol{A}$ has a form:

$$p(\boldsymbol{A}) = \sum_{j=1}^{n} a_j \boldsymbol{A}^j = \sum_{j=1}^{n} a_j \sum_{k=1}^{n} {\lambda_k}^j \boldsymbol{P}_k = \sum_{k=1}^{n} \sum_{j=1}^{n} a_j {\lambda_k}^j \boldsymbol{P}_k = \sum_{k=1}^{n} p(\lambda_k) \boldsymbol{P}_k \tag{5.11}$$

Hamilton-Kelly theorem. If $\boldsymbol{A}$ is a self-adjoint operator and $p(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I})$ is a characteristic polynomial of this operator, then

$$p(\boldsymbol{A}) = 0$$

Proof. If the operator $\boldsymbol{A}$ is a self-adjoint operator and $\lambda_1, ..., \lambda_n$ are its eigenvalues, then they are roots of the characteristic equation, that is, $p(\lambda_i) = 0$.

Hence, it follows from (5.11) that $p(\boldsymbol{A}) = 0$. ■

Considering a matrix representation of the operator, the theorem is valid for its matric. Herewith, the polynomial $p(\lambda)$ of the variable λ is called ***annihilating polynomial*** of any square matrix *A*, and the polynomial with respect to the matrix takes the form similar to (5.11):

$$p(A) = \sum_{j=1}^{n} p(\lambda_j)\boldsymbol{P}_j.$$

Example 5.3. Given a matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, show that the characteristic equation of the matrix *A* is an annihilating polynomial of it.

Solution. Find the characteristic polynomial of the matrix:

$$det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = 3\lambda^2 - \lambda^3.$$

Substituting the matrix *A* in this expression instead of the variable λ, we obtain

$$3\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^2 - \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^3 = 3\begin{pmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{pmatrix} - \begin{pmatrix} 9 & 9 & 9 \\ 9 & 9 & 9 \\ 9 & 9 & 9 \end{pmatrix} = \vec{0}$$

What was necessary to prove.

Positive Definite and Positive Operators

Definition: A self-adjoint operator $\boldsymbol{A}$ on a space E^n is called ***positive*** (*semi-definite* or *non-negative*) if $(\boldsymbol{A}\vec{x}, \vec{x}) \geq 0$ for every $\vec{x} \in E^n$.

Definition: A self-adjoint operator $\boldsymbol{A}$ on a space E^n is called ***positive definite*** if $(\boldsymbol{A}\vec{x}, \vec{x}) > 0$ for every $\vec{x} \in E^n$ except for $\vec{x} = 0$.

Properties of positive definite and positive operators

1. A self-adjoint operator $\boldsymbol{A}$ is *positive* if and only if all its eigenvalues are non-negative (positive).
2. A self-adjoint operator $\boldsymbol{A}$ is *positive definite* if and only if all its eigenvalues are strictly positive.
3. A self-adjoint operator $\boldsymbol{A}$ is *positive / positive definite* if there exist $\boldsymbol{A} = \boldsymbol{B}^2$ for some *self-adjoint* / *nonsingular self-adjoint* operator $\boldsymbol{B}$.

   (i.e. every non-negative number has a unique non-negative square root)
4. A self-adjoint operator $\boldsymbol{A}$ is *positive / positive definite* if there exist $\boldsymbol{A} = \boldsymbol{S}^*\boldsymbol{S}$ for some *operator* / *nonsingular operator* $\boldsymbol{S}$.

   (It is an analogy with complex numbers, i.e. a complex number z is nonnegative if and only if has the form $z = \bar{w}w$ for some complex number w).

Polar decomposition

Theorem: Any operator $\boldsymbol{A}$ in a Euclidean space E^n can be presented as factorization of the form

$$\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{P}$$

In the case of a real space, $\boldsymbol{U}$ is an orthogonal operator and $\boldsymbol{P}$ is a positive semi-definite self-adjoint operator with symmetric matrix; or $\boldsymbol{U}$ is a unitary operator and $\boldsymbol{P}$ is a positive semi-definite self-adjoint operator with Hermitian matrix in the complex case.

Proof: Let's consider a positive self-adjoint operator $\boldsymbol{D}$ such that $\boldsymbol{D} = \boldsymbol{A}^*\boldsymbol{A}$ for some operator $\boldsymbol{A}$. Given the self-adjoint operator $\boldsymbol{D}$, there exists a self-adjoint operator $\boldsymbol{P}$ such that $\boldsymbol{D} = \boldsymbol{P}^*\boldsymbol{P}$. Let's compose an operator $\boldsymbol{U} = \boldsymbol{A}\,\boldsymbol{P}^{-1}$. One can show that this operator is unitary/orthogonal. Indeed,

$$\boldsymbol{U}^*\boldsymbol{U} = (\boldsymbol{A}\,\boldsymbol{P}^{-1})^*\boldsymbol{A}\,\boldsymbol{P}^{-1} = (\boldsymbol{P}^{-1})^*\underbrace{\boldsymbol{A}^*\boldsymbol{A}}_{\boldsymbol{D}}\,\boldsymbol{P}^{-1} = (\boldsymbol{P}^{-1})^*\underbrace{\boldsymbol{P}^*\boldsymbol{P}}_{\boldsymbol{D}}\,\boldsymbol{P}^{-1} =$$

$$(\boldsymbol{P}\,\boldsymbol{P}^{-1})^*\boldsymbol{P}\,\boldsymbol{P}^{-1} = \boldsymbol{I}^*\boldsymbol{I} = \boldsymbol{I}$$

Thus, $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{P}$. ■

Remark 1. If the operator $\boldsymbol{A}$ is *nonsingular* then $\boldsymbol{P}$ is either positive definite symmetric operator in the real case or positive definite Hermitian operator in the complex case.

Remark 2. The polar decomposition should be considered as an analogy between set of complex numbers C and a Euclidean space. First, we recall the polar form of a complex number $z = |z|e^{i\theta}$, where $|z|$ is the absolute value or modulus of z and $e^{i\theta}$ lies on the unit circle in R^2. Then, in terms of an operator $\boldsymbol{A} \in E^n$, a unitary/orthogonal operator $\boldsymbol{U}$ takes the role of $e^{i\theta}$, and $|\boldsymbol{A}|$ takes the role of the modulus. As seen, if $\boldsymbol{A}^*\boldsymbol{A} \geq 0$ so that $|\boldsymbol{A}| = (\boldsymbol{A}^*\boldsymbol{A})^{1/2}$ exists and satisfies $|\boldsymbol{A}| \geq 0$ as well, i.e. a positive self-adjoint operator $\boldsymbol{P} = (\boldsymbol{A}^*\boldsymbol{A})^{1/2}$.

Intuitively, we can imagine this decomposition via the factorization of **matricide** associated with the operators, namely, if a real $n \times n$ matrix A is interpreted as a linear transformation of *n*-dimensional space R^n, then, the polar decomposition separates it into a *rotation* or *reflection* U of the space R^n, and a scaling of the space along a set of *n* orthogonal axes by P.

Chapter 6. BILINEAR AND QUADRATIC FORMS

6.1. Basic concepts of bilinear functions (forms)

In previous chapters, we have studied linear operators that map from the vector space into the vector space. The scalar field is the simplest of all nontrivial vector spaces. Given a vector space with a scalar field, then a linear mapping from the vector space into its scalar field is called a *linear functional* on a vector space.

Definition. It is said that ***a linear functional*** $f(\vec{x})$ is given on a vector space R if a vector $\vec{x} \in R$ and a scalar $f(\vec{x})$ the following conditions are fulfilled:

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}),$$

$$f(\alpha\vec{x}) = \alpha f(\vec{x}),$$

where $\vec{x}, \vec{y}$ are arbitrary vectors of R, and α is any real number.

Example 6.1. Let R be the vector space of n-tuples, which we write as column vectors with coordinates $\{x_1, x_2, \dots, x_n\}$. Then, any linear functional in the space of row vectors has the representation:

$$f(x_1, x_2, \dots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

Historically, the formal expression on the right was termed ***a linear form***.

Definition. A function of two variables $\Phi(\vec{x}, \vec{y})$ given in a vector space R over is called ***a bilinear functional*** (***form***) if for a fixed $\vec{x} \in R$ it is a linear function with respect to the vector $\vec{y} \in R$, and for a fixed $\vec{y}$ it is a linear function with respect to $\vec{x}$.

If $\Phi(\vec{x}, \vec{y})$ is a bilinear functional, then for $\vec{x}_1$, $\vec{x}_2$, $\vec{y}_1$, $\vec{y}_2$ $\in R$ and an arbitrary real α the following conditions are valid:

$$\Phi(\vec{x}_1 + \vec{x}_2, \vec{y}) = \Phi(\vec{x}_1, \vec{y}) + \Phi(\vec{x}_2, \vec{y}),$$

$$\Phi(\alpha\vec{x},\vec{y}) = \alpha\Phi(\vec{x},\vec{y}),$$
$$\Phi(\vec{x},\vec{y}_1 + \vec{y}_2) = \Phi(\vec{x},\vec{y}_1) + \Phi(\vec{x},\vec{y}_2),$$
$$\Phi(\vec{x},\alpha\vec{y}) = \alpha\Phi(\vec{x},\vec{y})$$

An example of a bilinear functional is the dot product of vectors $(\vec{x},\vec{y})$ on R^n, i.e.

$$(\vec{x},\vec{y}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Let us consider a bilinear functional using coordinates of the vectors. Let a space *R* be given by the basis vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. It is obvious that any vector of the space can be decomposed along the basis vectors, e.g. $\vec{x},\vec{y} \in R$ have a form:

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n = \sum_{i=1}^{n} x_i\vec{e}_i\,,$$
$$\vec{y} = y_1\vec{e}_1 + y_2\vec{e}_2 + \dots + y_n\vec{e}_n = \sum_{i=1}^{n} y_i\vec{e}_i\,.$$

Then, the bilinear functional can be expressed as follows:

$$\Phi(\vec{x},\vec{y}) = \Phi(x_1\vec{e}_1 + x_2\vec{e}_2 + \dots + x_n\vec{e}_n, y_1\vec{e}_1 + y_2\vec{e}_2 + \dots + y_n\vec{e}_n) = \tag{6.1}$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x_jy_j\Phi(\vec{e}_j,\vec{e}_j) = \sum_{i,j=1}^{n} \Phi(\vec{e}_j,\vec{e}_j)\,x_jy_j = \sum_{i,j=1}^{n} a_{ij}x_jy_j$$

where the coefficients $a_{ij} = \Phi(\vec{e}_j,\vec{e}_j)$ depend only on the basis vectors, but are regardless of the vectors $\vec{x}$ and $\vec{y}$ themselves.

It is said that in the given basis, the bilinear functional $\Phi(\vec{x},\vec{y})$ is represented by ***a bilinear form***

$$\Phi(\vec{x},\vec{y}) = \sum_{i,j=1}^{n} a_{ij}x_jy_j$$

Definition. The matrix $A = [a_{ij}]_{i,j=\overline{1,\dots,n}}$ is called a ***matrix of the bilinear form*** relative to the given basis.

Definition. The bilinear form is called ***symmetric*** if for any $\vec{x}, \vec{y} \in R$, an equality

$$\Phi(\vec{x}, \vec{y}) = \Phi(\vec{y}, \vec{x})$$

is valid.

In this case, we have $a_{ij} = a_{ji}$, i.e. the matrix of symmetric bilinear form in any basis is *symmetric*.

Definition. The bilinear function is called ***skew-symmetric*** if for any $\vec{x}, \vec{y} \in R$, an equality occurs:

$$\Phi(\vec{x}, \vec{y}) = -\Phi(\vec{y}, \vec{x})$$

In this case, we have $a_{ij} = -a_{ji}$, i.e. the matrix of skew-symmetric bilinear form in any basis is *skew-symmetric*.

Definition. A bilinear form is *alternating* if and only if its coordinate matrix is skew-symmetric and the diagonal entries are all zero, i.e. $a_{ii} = 0$, $\forall i = j$.

6.2. Quadratic forms. Basic concepts

Definition. A ***quadratic form*** is a bilinear symmetric form for $\vec{x} = \vec{y}$, i.e. $\Phi(\vec{x}, \vec{x})$. Therefore, a ***quadratic form*** of *n* variables $x_1, x_2, ..., x_n$ is a polynomial of these variables such that every term has degree two:

$$\Phi(\vec{x}, \vec{x}) = Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} x_i x_k = \sum_{i,k=1}^{n} a_{ik} x_i x_k, \tag{6.2}$$

Herein, we take into account that $a_{ik} = a_{ki}$.

If the variables $x_1, x_2, ..., x_n$ are considered as coordinates of the vector $\vec{x}$ given in a linear *n*-dimensional space with a basis $\{e_i\}_{i=\overline{1,n}}$, then the quadratic form can be defined as a scalar function of this vector, i.e.

$$Q(x_1, x_2, \ldots, x_n) = F(\vec{x}) \tag{6.3}$$

<u>Definition</u>. A matrix $A = (a_{ik})_{i,k=\overline{1,\ldots,n}}$ is called a *matrix of quadratic form* in a given basis of the space, and the matrix is symmetric, $a_{ik} = a_{ki}$, in any basis.

The quadratic form can be written in a compact matrix form:

$$Q(x_1, x_2, \ldots, x_n) = x_1(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) +$$

$$+x_2(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n) + \cdots + x_n(a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n) =$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}x_i x_k = (x_1, x_2, \ldots, x_n)\begin{pmatrix} a_{11} & a_{12}/2 & \ldots & a_{1n}/2 \\ a_{12}/2 & a_{22} & \ldots & a_{2n}/2 \\ \ldots & \ldots & \ldots & \ldots \\ a_{1n}/2 & a_{2n}/2 & \ldots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$$

or

$$Q(x_1, x_2, \ldots, x_n) = X^T \cdot A \cdot X, \tag{6.4}$$

where the coordinates of the vector are a column, $X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$.

Example 6.2. Compose a matrix *A* of the quadratic form such as $Q = 3x_1^2 + 2x_1x_2 - x_1x_3 + 2x_2^2 + 6x_2x_3 - 5x_3^2$.

$$A = \begin{pmatrix} 3 & 1 & -\frac{1}{2} \\ 1 & 2 & 3 \\ -\frac{1}{2} & 3 & -5 \end{pmatrix}$$

If two quadratic forms have the same matrices, which differ from each other only by the denotation of the variables, these forms are called ***equal***.

6.3. Change of Basis

We now answer the question, how does a matrix representing a quadratic form transform when a new basis is selected?

Let $\{x_i\}_{i=\overline{1,n}}$ and $\{y_j\}_{j=\overline{1,n}}$ be the coordinates of the same vector $\vec{x}$ in two distinct bases $\{\vec{g}_i\}_{i=\overline{1,n}}$ and $\{\vec{h}_j\}_{j=\overline{1,n}}$ in an *n*-dimensional linear space. The transition from the basis $\{\vec{g}_i\}_{i=\overline{1,n}}$ to a basis $\{\vec{h}_j\}_{j=\overline{1,n}}$ is given by the following relations:

$$\begin{cases} \overrightarrow{h_1} = C_{11}\,\overrightarrow{g_1} + C_{21}\,\overrightarrow{g_2} + \ldots + C_{n1}\,\overrightarrow{g_n} \\ \overrightarrow{h_2} = C_{12}\,\overrightarrow{g_1} + C_{22}\,\overrightarrow{g_2} + \ldots + C_{n2}\,\overrightarrow{g_n} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \overrightarrow{h_n} = C_{1n}\,\overrightarrow{g_1} + C_{2n}\,\overrightarrow{g_2} + \ldots + C_{nn}\,\overrightarrow{g_n} \end{cases} \tag{6.5}$$

Then, the coordinates of the vector $\vec{x}$ in the basis $G = \{\vec{g}_i\}_{i=\overline{1,n}}$ are related to its coordinates in the basis $H = \{\vec{h}_i\}_{i=\overline{1,n}}$ as follows:

$$\overrightarrow{x_G} = C\,\overrightarrow{x_H} = C\,\vec{y}, \Rightarrow \overrightarrow{x_H} = \vec{y} \tag{6.6}$$

where $\begin{pmatrix} C_{11} & C_{12} & \ldots & C_{1n} \\ C_{21} & C_{22} & \ldots & C_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ C_{n1} & C_{n2} & \ldots & C_{n2} \end{pmatrix}$ is the matrix of transition from basis $\{g_i\}_{i=\overline{1,n}}$ to basis $\{h_j\}_{j=\overline{1,n}}$, and $\overrightarrow{x_G} = X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\overrightarrow{x_H} = \vec{y} = Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

Expanding (6.6), we get

$$\begin{cases} x_1 = C_{11}y_1 + C_{12}y_2 + \ldots + C_{1n}y_n \\ x_2 = C_{21}y_1 + C_{22}y_2 + \ldots + C_{2n}y_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = C_{n1}y_1 + C_{n2}y_2 + \ldots + C_{nn}y_n \end{cases} \tag{6.7}$$

Then, we can talk not about the transformation of the basis by formulas (6.5), but about the linear transformation of variables $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ of the given quadratic form $F(\vec{x})$.

The expression (6.7) in a matrix form looks like:

$$X = C \cdot Y.$$

Given the formulas $F(\vec{x}_G) = X^T A_G X$ and $X^T = (CY)^T = Y^T C^T$, we obtain

$$F(\vec{x}_G) = Y^T C^T A_G C Y$$

On the other hand,

$$F(\overrightarrow{x_H}) = X^T \cdot A_H \cdot X = F(\vec{y}) = Y^T A_H Y$$

Since the quadratic form does not depend on the denotation of variables, then equating the right-hand sides of the quadratic form expression, we get

$$A_H = C^T \cdot A_G \cdot C \tag{6.8}$$

The formula (6.8) presents the transformation of the quadratic form matrix

Let us consider the matrix $A_H = C^T A_G C$. Since ${A_H}^T = (C^T A_G C)^T = (A_G C)^T \cdot (C^T)^T = C^T \cdot {A_G}^T \cdot C = C^T \cdot A_G \cdot C$, then $A_H = {A_H}^T$, therefore, the matrix is a symmetric matrix that defines the quadratic form.

It follows from the last equality that

$$det\, A_H = (det\, C)^2 \cdot det\, A_G \tag{6.9}$$

Since C is a transition matrix, it is non-singular, and $(\det C)^2 > 0$ as a result, the matrices $A_H,\ A_G$ always have the same signs.

Example 6.3. Write the expression of the quadratic form $F(\vec{x}) = x_1^2 + 4x_1x_2 + 2x_2^2$ in a new basis of the vectors $\overrightarrow{h_1} = (1,3),\ \overrightarrow{h_2} = (-1,2)$.

Solution. The matrix of the quadratic form in the initial basis is $A_G = A_e = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$. Then, compose the transition matrix from the “old” basis to a new basis *H*. Thus, $\mathrm{C} = \begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix}$. Finally, according to rule (6.8), the matrix of the quadratic form looks like

$$A_H = C^T A_G C = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 31 & 9 \\ 9 & 1 \end{pmatrix}.$$

In turn, the quadratic form in the new basis is

$$F(\vec{y}) = 31y_1^2 + 18y_1y_2 + y_2^2.$$

6.4. Classification of Quadratic Forms

Definition. The rank of a quadratic form matrix is called a ***rank of the***

quadratic form.

If Rg$A = n$ (n is a space dimension), then the quadratic form is called ***non-singular***, otherwise it is ***singular***.

Definition. The ***canonical quadratic form*** $F(\vec{x})$ is a type of a quadratic form, whose $a_{ik} = 0$, if $i \neq k$, and $a_{ii} \neq 0$, if $i = k$, i.e. the form does not contain the products of distinct variables and can be presented in the form:

$$F(\vec{x}) = \lambda_1 {x_1}^2 + \lambda_2 {x_2}^2 + \ldots + \lambda_n {x_n}^2.$$

The basis, where such form occurs is called a ***canonical basis***.

Definition. The ***normal form*** of the quadratic form $F(\vec{x})$ is its canonical form, in which the coefficients before the squares of the variables (excluding zeros) are equal to ± 1.

Definition. A quadratic form $F(\vec{x})$ is ***positive definite*** if $\forall \vec{x} \neq \vec{0}$, $F(\vec{x}) > 0$.

Definition. A quadratic form $F(\vec{x})$ is ***negative definite*** if $\forall \vec{x} \neq \vec{0}$, $F(\vec{x}) < 0$.

Theorem (Sylvester's law of inertia for quadratic forms). For any way of reducing a quadratic form

$$F(\vec{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_k$$

with real coefficients $a_{ij} \in R$ to a sum of squares

$$\sum_{i=1}^{n} b_i y_i^2$$

by a linear change of variables $\vec{x} = C\,\vec{y}$, where C is a non-singular matrix with real coefficients, the number of the coefficients b_i of a given sign in the *canonical quadratic form* is an invariant of $F(\vec{x})$, i.e. does not depend on a particular choice of diagonalizing basis.

Expressed geometrically, the law of inertia says that all maximal subspaces on which the restriction of the quadratic form is *positive definite* (respectively, *negative definite*) have the same dimension. These dimensions are the positive and negative indices of inertia.

Theorem (Sylvester's criterion). A quadratic form $F(\vec{x})$ is positive definite if and only if all minors taken from the top left corner of the quadratic form matrix are positive, i.e.

$$D_k > 0, \forall k = \overline{1, n}, \text{ where } D_k = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}$$

Theorem. A quadratic form $F(\vec{x})$ is negative definite if and only if the signs of all minors taken from the top left corner of the quadratic form matrix D_i alternate, starting with a minus:

$$D_1 < 0,\ D_2 > 0,\ \ D_3 < 0, \dots$$

Example 6.4. Investigate the significance of the quadratic form $F(\vec{x}) = -x_1{}^2 - 6x_2{}^2 - 6x_3{}^2 + 12x_1x_2 - 12x_1x_3 + 6x_2x_3$.

Solution. Find a positive(negative-)definiteness of a quadratic form given by the matrix in the form:

$$A = \begin{pmatrix} -11 & 6 & -6 \\ 6 & -6 & 3 \\ -6 & 3 & -6 \end{pmatrix}$$

$$D_1 = -11 < 0, D_2 = \begin{vmatrix} -11 & 6 \\ 6 & -6 \end{vmatrix} = 30 > 0, D_3 = \begin{vmatrix} -11 & 6 & -6 \\ 6 & -6 & 3 \\ -6 & 3 & -6 \end{vmatrix} = -81 <$$

0.

Since the signs of the minors alternate from minus, the quadratic form is negative.

6.5. Lagrange Reduction of Quadratic Form to Canonical form

Theorem. Any quadratic form can be reduced to a canonical form by some non-singular linear transformation. That is, there is a basis in which this form is reduced to the sum of squares.

Lagrange's method consists in the following. One may assume that not all the coefficients in (6.2) are zero. Therefore, two cases are possible.

1. A quadratic form contains the square of at least one variable x_i, i.e. $\exists a_{ii} \neq 0$

In this case, we need to group all the terms containing this variable, and complete the square so that the remaining terms do not contain the variable x_i. After that, the remaining terms form a quadratic form of the (n-1)-th order. After finite number of similar steps one can reduce the form to a sum of squares.

Example 6.5. Reduce the quadratic form to the canonical form:

$$F = 9x_1^2 + x_1x_2 + 6x_1x_3 + x_2^2 + x_4^2 - 4x_2x_3 + 2x_2x_4 - 8x_3x_4$$

Solution. This quadratic form contains squares of x_1, x_2, x_4. Select any of them, for example x_2, and group all the members that contain this variable:

$$(x_2^2 + x_1x_2 - 4x_2x_3 + 2x_2x_4) + 9x_1^2 + 6x_1x_3 + x_4^2 - 8x_3x_4 + 4x_4^2 - 8x_3x_4.$$

Complete the square with respect to x_2:

$$\underbrace{\left(x_2 + \frac{1}{2}x_1 - 2x_3 + x_4\right)}_{y_1}^2 - \frac{1}{4}x_1^2 - 4x_3^2 - x_4^2 + 8x_1x_3 + 4x_3x_4 - x_1x_4.$$

Then you can write $F(\vec{x})$ as:

$$F = y_1^2 + F_1(\vec{x}),$$

where

$$F_1(\vec{x}) = \frac{35}{4}x_1^2 - 4x_3^2 + 8x_1x_3 - 4x_3x_4 - x_1x_4 = -4\underbrace{\left(x_3 - x_1 + \frac{1}{2}x_4\right)}_{y_2}^2 + \frac{51}{4}x_1^2 + x_4^2 - 5x_1x_4,$$

or

$$F_1 = -4y_2^2 + F_2(\overrightarrow{x}),$$

where

$$F_2(\overrightarrow{x}) = \frac{51}{4}x_1^2 + x_4^2 - 5x_1x_4 = \underbrace{\left(x_4 - \frac{5}{2}x_1\right)}_{y_3} + \frac{13}{2}\underbrace{x_1^2}_{y_4}.$$

Finally, $F(\overrightarrow{x})$ looks like:

$$F(\overrightarrow{y}) = y_1^2 - 4y_2^2 + y_3^2 + \frac{13}{2}y_4^2,$$

where we denote:

$$\begin{cases} y_1 = x_2 + \frac{1}{2}x_1 - 2x_3 + x_4 \\ y_2 = x_3 - x_1 + \frac{1}{2}x_4 \\ y_3 = x_4 - \frac{5}{2}x_1 \\ y_4 = x_1 \end{cases} \tag{6.10}$$

The system (6.10) establishes a connection between the new coordinates and the old ones. One can find a matrix of transition from the old basis to the new one by expressing the old variables through new ones:

$$\begin{cases} x_1 = y_4 \\ x_4 = y_3 + \frac{5}{2}y_4 \\ x_3 = y_2 + y_4 - \frac{1}{2}\left(y_3 + \frac{5}{2}y_4\right) = y_2 - \frac{1}{2}y_3 - \frac{1}{4}y_4 \\ x_2 = y_1 - \frac{1}{2}y_4 + 2\left(y_2 - \frac{1}{2}y_3 - \frac{1}{4}y_4\right) - y_3 - \frac{5}{2}y_4 = \\ \qquad = y_1 + 2y_2 - 2y_3\left(-\frac{1}{2} + \frac{1}{2} + \frac{5}{2}\right) \end{cases}$$

or

$$\begin{cases} x_1 = y_4 \\ x_2 = y_1 + 2y_2 - 2y_3 - \frac{7}{2}y_4 \\ x_3 = y_2 - \frac{1}{2}y_2 - \frac{1}{4}y_4 \\ x_4 = y_3 + \frac{5}{2}y_4 \end{cases} \Rightarrow C = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & -2 & -\frac{7}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{5}{2} \end{pmatrix}$$

Let's check result using the formula $A_H = C^T \cdot A_G \cdot C$.

$$A_H = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & -2 & -\frac{1}{2} & 1 \\ 1 & -\frac{7}{2} & -\frac{1}{4} & \frac{5}{2} \end{pmatrix} \cdot \begin{pmatrix} 9 & \frac{1}{2} & 3 & 0 \\ \frac{1}{2} & 1 & -2 & 1 \\ 3 & -2 & 0 & -4 \\ 0 & 1 & -4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 2 & -2 & -\frac{7}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{5}{2} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{13}{2} \end{pmatrix}$$

2. A quadratic form has no squares of variables, i.e. all $a_{ii} = 0$, and some of the coefficients $a_{ij} \neq 0$, for example $a_{12} \neq 0$.

In this case, first of all, a replacement is introduced:

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \\ \dots\dots\dots\dots\dots \\ x_n = y_n \end{cases} \qquad (6.11)$$

System (6.11) is a transition to a new basis in which the quadratic form will have squares of variables. That is, we can reduce the form using the previous scheme.

Example 6.6. Reduce $F = x_1x_2 - 2x_1x_3 + 4x_2x_3$ to the canonical form.

Solution. Introduce the replacement (6.11):

$$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2. \\ x_3 = y_3 \end{cases}$$

Then, the quadratic form with respect to new variables looks like:

$$F(\vec{y}) = y_1^2 - y_2^2 - 2(y_1 + y_2) \cdot y_3 + 4(y_1 - y_2)y_3$$

$$= y_1^2 - y_2^2 + 2y_1y_3 - 6y_2y_3$$

Next, combining the variables and completing the squares, we get:

$$F = \underbrace{(y_1^2 + 2y_1y_3 + y_3^2)}_{(y_1+y_3)^2} - y_3^2 - y_2^2 - 6y_2y_3 =$$

herein, new variables are introduced such that

$$= \underbrace{(y_1 + y_3)}_{z_1}^2 - \underbrace{(y_2^2 + 6y_3y_2 + 9_3^2)}_{y_2+3y_3=z_2} + 8\underbrace{y_3^2}_{z^3} = z_1^2 - z_2^2 + 8z_3^2$$

The final relationship between the coordinates is

$$\begin{cases} z_1 = y_1 + y_3 \\ z_2 = y_2 + 3y_3 \\ z_3 = y_3 \end{cases} \Rightarrow \begin{cases} y_3 = z_3 \\ y_2 = z_2 - 3z_3 \\ y_1 = z_1 - z_3 \end{cases} \Rightarrow \begin{cases} x_3 = z_3 \\ x_2 = z_1 - z_2 + 2z_3, \\ x_1 = z_1 + z_2 - 4z_3 \end{cases}$$

or

$$\begin{cases} x_1 = z_1 + z_2 - 4z_3 \\ x_2 = z_1 - z_2 + 2z_3. \\ x_3 = z_3 \end{cases}$$

It follows from this system, the transition matrix takes a form

$$C = \begin{pmatrix} 1 & 1 & -4 \\ 1 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Let's check it by using the formula: $A_e = C^T \cdot A \cdot C$.

$$A_e = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ -4 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & 2 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -4 \\ 1 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

6.6. Quadratic Forms and Principal Axes

Let us be given a quadratic form $F(\vec{x})$ in the Euclidean space with an orthonormal basis, $\{f_k\}_{k=\overline{1,n}}$. Let $x_1, x_2, x_3, .., x_n$ be the coordinates of the vector $\vec{x}$ in this basis. Since the quadratic form matrix is symmetric, this matrix can be considered as a matrix of the self-adjoint operator *A* in the orthonormal basis.

It is known that the matrix of a self-adjoint operator takes a diagonal form in an orthonormal basis $\{e_k\}_{k=\overline{1,n}}$ generated by its eigenvectors. So, if we choose a transformation $\vec{x} = P\vec{y}$ providing *orthogonally diagonalize matrix A*, then a new quadratic form will be $\vec{y}^T A_e \vec{y}$, where A_e is a diagonal matrix with the eigenvalues λ_i of *A* on the main diagonal, that is,

$$\vec{x}^T A\vec{x} = \vec{y}^T A_e \vec{y} = \{y_1 \quad y_2 \quad \cdots \quad y_n\} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Bmatrix}$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$$

Thus, we have the following result, called *the principal axes theorem*:

Theorem: If *A* is a symmetric $n \times n$ matrix, then there is an *orthogonal change of variable* that transforms the quadratic form $\vec{x}^T A\vec{x}$ into a quadratic form $\vec{y}^T D\vec{y}$ with no cross terms. Specifically, if *P orthogonally diagonalizes A*, then making the change of variable $\vec{x} = P\vec{y}$ in the quadratic form $\vec{x}^T A\vec{x}$ yields the quadratic form

$$F(\vec{x}) = \vec{x}^T A\vec{x} = \vec{y}^T A_e \vec{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \qquad (6.12)$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of *A* corresponding to the eigenvectors that form the successive columns of *P*, and $y_1, y_2, \ldots, y_n$ are the coordinates of the vector $\vec{x}$ in the new basis.

The eigenvectors associated with the eigenvalues $\lambda_i\left(i=\overline{1,n}\right)$ form lines called the *principal axes of the quadratic form*.

The transition from the natural basis to the principal axes basis is fulfilled by the orthogonal matrix *P*. Thus, the matrix *A* of quadratic form is related to the matrix A_e in the basis of eigenvectors by the formula:

$$A_e = P^T \cdot A \cdot P.$$

It follows that the reduction of a quadratic form to the principal axes coincides with the algorithm of diagonalization of Hermitian matrices. Thus, to reduce a quadratic form to the canonical form using an orthogonal transformation, it is necessary to perform the following steps:

1. Compose a matrix *A* of square form
2. Find the eigenvalues of this matrix and write the form in the principal axes.
3. Create an orthonormal basis $\{e_k\}_{k=1}^{n}$ using eigenvectors associated with the known eigenvalues.
4. Compose the orthogonal matrix *P* of the transition from the natural basis to the basis $\{e_k\}_{k=1}$, where the quadratic form matrix is diagonal, i.e. the quadratic form takes a form (6.12) with respect to the principal axes.

Example 6.7. Find the orthogonal transformation that reduces a quadratic form $F\left(\vec{x}\right) = 17x_1^2 + 14x_2^2 + 14x_3^2 - 4x_1x_2 - 4x_1x_3 - 8x_2x_3$ to the canonical form and write down it.

Solution.

1. Let's compose the matrix *A*:

$$A = \begin{pmatrix} 17 & -2 & -2 \\ -2 & 14 & -4 \\ -2 & -4 & 14 \end{pmatrix}.$$

2. Find eigenvalues and eigenvectors of this matrix. Let's compose the

homogeneous system of equations $A\vec{x} = \lambda\vec{x}$, i.e.

$$\begin{cases}(17-\lambda)x_1 - 2x_2 - 2x_3 = 0,\\ -2x_1 + (14-\lambda)x_2 - 4x_3 = 0,\\ -2x_1 + -4x_2 + (14-\lambda)x_3 = 0.\end{cases}$$

Let's construct the characteristic equation: $p(\lambda) = det(A - \lambda I) = 0$

$$p(\lambda) = \begin{vmatrix} 17-\lambda & -2 & -2 \\ -2 & 14-\lambda & -4 \\ -2 & -4 & (14-\lambda) \end{vmatrix}_{r_3 - r_2} = 0$$

$$\begin{vmatrix} 17-\lambda & -2 & -2 \\ -2 & 14-\lambda & -4 \\ -2 & -18+\lambda & 18+\lambda \end{vmatrix} = (18-\lambda)\begin{vmatrix} 17-\lambda & -2 & -2 \\ -2 & 14-\lambda & -4 \\ 0 & -1 & 1 \end{vmatrix}_{t_2 + t_3} =$$

$$= (18-\lambda)\begin{vmatrix} 17-\lambda & -4 & -2 \\ -2 & 10-\lambda & -4 \\ 0 & 0 & 0 \end{vmatrix} = (18-\lambda)\begin{vmatrix} 17-\lambda & -4 \\ -2 & 10-\lambda \end{vmatrix} =$$

$$(18-\lambda)\big((17-\lambda)(10-\lambda) - 8\big) = (18-\lambda)(\lambda^2 - 27\lambda + 162) = 0 \Rightarrow$$

$$\lambda_1 = 9, \lambda_2 = \lambda_3 = 18$$

3. Find the eigenvectors:

$\underline{\lambda_1 = 9}$

After substituting this first eigenvalue in the homogeneous system, we find the solution of the system as follows:

$$\begin{pmatrix} 8 & -2 & -2 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix} \sim \begin{pmatrix} 4 & -1 & -1 \\ -2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}_{\substack{r_2 - r_3 \\ r_1 + 2r_3}} \sim \begin{pmatrix} 0 & -9 & 9 \\ 0 & 9 & -9 \\ -2 & -4 & 5 \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & -1 \\ -2 & -4 & 5 \end{pmatrix}$$

$x_2 = x_3$

$-2x_1 = 4x_2 - 5x_3 \Rightarrow -2x_1 = 4x_3 - 5x_3 = -x_3 \Rightarrow x_1 = \frac{x_3}{2}$

	x_1	x_2	x_3
$\vec{e_1}$	1	2	2

That is, $\vec{e_1} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ is the eigenvector associated with $\lambda_1 = 9$.

Similarly,

$\underline{\lambda_2 = 18}$:

$$\begin{pmatrix} -1 & -2 & -2 \\ -2 & -4 & -4 \\ -2 & -4 & -4 \end{pmatrix} \sim (1\ 2\ 2) \Rightarrow x_1 = -2x_2 - 2x_3$$

	x_1	x_2	x_3
$\overrightarrow{e_2}$	-2	1	0
$\overrightarrow{e_3}$	-2	0	1

That is, $\overrightarrow{e_2} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$ and $\overrightarrow{e_3} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ are the eigenvectors associated with $\lambda_2 = 18$.

The system of vectors $\overrightarrow{e_1}\,\overrightarrow{e_2}\,\overrightarrow{e_3}$ is not orthonormal, i.e. the vectors are not pairwise orthogonal, e.g. $\left(\overrightarrow{e_2}, \overrightarrow{e_3}\right) = 4 \neq 0$ and are not unit length.

4. Let's orthogonalize $\overrightarrow{e_2}$ and $\overrightarrow{e_3}$, using the Gramm-Schmidt process:

$\overrightarrow{e_2^*} = \overrightarrow{e_2}$;

$\overrightarrow{e_3^*} = \overrightarrow{e_3} + \alpha \cdot \overrightarrow{e_2^*} \Rightarrow \left(\overrightarrow{e_3^*}, \overrightarrow{e_2^*}\right) = \left(\overrightarrow{e_3}, \overrightarrow{e_2^*}\right) + \alpha \cdot \left(\overrightarrow{e_2^*}, \overrightarrow{e_2^*}\right) = 0.$

$$\alpha = -\frac{\left(\overrightarrow{e_3}, \overrightarrow{e_2^*}\right)}{\left(\overrightarrow{e_2^*}, \overrightarrow{e_2^*}\right)} = -\frac{4}{5};$$

$$\overrightarrow{e_3^*} = (-2{,}0{,}1) - \frac{4}{5}(-2{,}1{,}0) = \left(-2 + \frac{8}{5}; -\frac{4}{5}; 1\right) = \left(-\frac{2}{5}; -\frac{4}{5}; 1\right).$$

Thereby, the system of vectors $\overrightarrow{e_1}\, e_2^* e_3^*$ is orthogonal. Let's make it orthonormal:

$$\overrightarrow{g_1} = \frac{\overrightarrow{e_1}}{\|\overrightarrow{e_1}\|} = \left(\frac{1}{3}; \frac{2}{3}; \frac{2}{3}\right),$$

$$\overrightarrow{g_2} = \frac{\overrightarrow{e_2}}{\|\overrightarrow{e_2}\|} = \frac{1}{\sqrt{5}}(-2; 1; 0),$$

$$\overrightarrow{g_3} = \frac{\overrightarrow{e_3}}{\|\overrightarrow{e_3}\|} = \frac{1}{\underbrace{\sqrt{\frac{4}{25} + \frac{16}{25} + 1}}_{3/\sqrt{5}}}\left(-\frac{2}{5}; -\frac{4}{5}; 1\right) = \left(-\frac{2}{3\sqrt{5}}; -\frac{4}{3\sqrt{5}}; \frac{5}{3\sqrt{5}}\right).$$

5. Construct a matrix of transition from the old basis to a new basis of the eigenvectors. So, each column of the matrix is the appropriate eigenvector:

$$P = \begin{pmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{pmatrix}$$

$$\overrightarrow{g_1} \quad \overrightarrow{g_2} \quad \overrightarrow{g_3}$$

Let $y_1 y_2 y_3$ be the coordinates of the vector $\vec{x}$ in the basis $\overrightarrow{g_1}$, $\overrightarrow{g_2}$, $\overrightarrow{g_3}$. Then

$$\vec{x} = P \cdot \vec{y},$$

or

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$x_1 = \frac{1}{3}y_1 - \frac{2}{5}y_2 - \frac{2}{3\sqrt{5}}y_3$$

$$x_2 = \frac{2}{3}y_1 + \frac{1}{5}y_2 + \frac{4}{3\sqrt{5}}y_3$$

$$x_3 = \frac{2}{3}y_1 + \frac{5}{3\sqrt{5}}y_3$$

Finally, the quadratic form in the principal axes looks like:

$$F(\vec{y}) = 9y_1^2 + 18y_2^2 + 18y_3^2$$

One can verify the diagonal matrix of the quadratic form by using the transformation of the matrices: $A_e = P^T \cdot A \cdot P$.

6.7. Simultaneous reduction of two quadratic forms to the canonical form

The problem of reducing simultaneously two quadratic forms to the canonical form does not always have a solution. But it can be done if certain conditions are fulfilled.

Theorem. If $F_1(\vec{x}) = X^T AX$ is an arbitrary, and $F_2(\vec{x}) = X^T BX$ is positive define quadratic forms, then there is a non-singular transformation that reduces both the forms to the canonical form, herewith the form $F_2(\vec{x})$ is a normalized form.

Proof. By Lagrange's theorem, any quadratic form can always be reduced to a diagonal form. According to the Sylvester's criterion and the laws of inertia, any positive define quadratic form in a canonical form has all eigenvalues $\lambda_i > 0$. In addition, if all $\lambda_i > 0$ then we can find a transformation that all λ_i will be equal to unity. That is, there is a transformation *C* such that

$$B_H = C^T BC = I \tag{6.13}$$

If we apply this transformation to the first quadratic form $F_1(\vec{x})$, its matrix will change according to the rule (6.8), i.e.

$$A_H = C^T AC. \tag{6.14}$$

Next, find the orthogonal transformation *D*, which diagonalizes A_H as follows:

$$A_e = D^T A_H D = diag(\dots).$$

Then, we apply the same orthogonal transformation to the quadratic form $F_2(\vec{x})$. In this case, the matrix $B_H = I$. Thus, we get

$$B_e = D^T B_H D = D^T ID = A_e = D^T D = I.$$

Therefore, since the orthogonal transformation does not change the unit matrix, the simultaneous transformation of quadratic forms $F_1(\vec{x})$ and $F_2(\vec{x})$ is possible. This is what was to be proven. ■

Find the orthogonal transformation *D*. Since the orthogonal transformation takes place at the second step, the quadratic form is considered in the base *H*, i.e.

$$det(A_H - \lambda I) = 0.$$

Substituting (6.13) and (6.14) in this equation, we get

$$det(C^T AC - \lambda C^T BC) = 0,$$

$$det(C^T(A - \lambda B)C) = 0.$$

According to the properties of the determinant:

$$detC^2 \cdot det(A - \lambda B) = 0.$$

Since *C* is non-singular matrix, then $detC \neq 0$ and

$$det(A - \lambda B) = 0 \qquad (6.15)$$

Let's write down the corresponding SLAE for search of eigenvectors:

$$(A_H - \lambda I)X_H = \vec{0}$$

$$(C^T AC - \lambda C^T BC)X_H = C^T(A - \lambda B)CX_H = C^T(A - \lambda B)X_G = \vec{0}.$$

So, we got the system in the initial basis:

$$(A - \lambda B)X = 0 \qquad (6.16)$$

Thus, for the simultaneous reduction of two quadratic forms to the canonical form, it is necessary to solve successively the problems (6.15) and (6.16).

Example 6.8. Reduce simultaneously two quadratic forms to the canonical form:

$$F_1 = 8x^2 - 26xy + 21y^2, \quad F_2 = 10x^2 - 34xy + 29y^2.$$

Solution. Let's write the matrices of these quadratic forms:

$$A_1 = \begin{pmatrix} 8 & -13 \\ -13 & 21 \end{pmatrix}, A_2 = \begin{pmatrix} 10 & -17 \\ -17 & 29 \end{pmatrix}.$$

F_2 is a positive definite form. Indeed,

$$D_1 = 10 > 0 \text{ and } D_2 = \begin{vmatrix} 10 & -17 \\ -17 & 29 \end{vmatrix} = 1 > 0.$$

Thus, by the appropriate transformations we get a diagonal form of F_1, and a normalized from of F_2.

Solve the equation (6.15), bearing in mind that $A = A_1,\ B = A_2$ are the corresponding matrices of the quadratic forms. Then,

$$det(A-\lambda B)=0 \quad \Rightarrow \quad \begin{vmatrix} 8-10\lambda & -13+17\lambda \\ -13+17\lambda & 21-29\lambda \end{vmatrix}=0.$$

The solutions are the values $\lambda_1 = 1$ and $\lambda_2 = -1$.

Thus, in the new basis, the quadratic forms look like:

$$F_1' = \lambda_1 {x_1}^2 + \lambda_2 {y_1}^2 = {x_1}^2 - {y_1}^2 \text{ and } F_2{}' = {x_1}^2 + {y_1}^2$$

Now, we need to find the basis that allows diagonalizing these forms. To do it, we have to solve the system of equations (6.16) at

$\underline{\lambda_1 = 1}$:

$$\begin{pmatrix} -2 & 4 \\ 4 & -8 \end{pmatrix} \sim (-1 \quad 2) \Rightarrow x_1 = 2x_2$$

	x_1	x_2
$\vec{e_1}$	2	1

$\vec{e_1} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is the first eigenvector;

and, similarly, at

$\underline{\lambda_2 = -1}$:

$$\begin{pmatrix} 18 & -30 \\ -30 & 50 \end{pmatrix} \sim (-3 \quad 5) \Rightarrow 3x_1 = 5x_2$$

	x_1	x_2
$\vec{e_2}$	5	3

$\vec{e_2} = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$ is the second eigenvector.

Compose a matrix of transition, where their columns are the eigenvectors:

$$C = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}.$$

Then, the transformation between the “old” and “new” coordinates looks like:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \Rightarrow \begin{cases} x = 2x_1 + 5y_1 \\ y = x_1 + 3y_1 \end{cases}$$

It could be verified by direct substituting the above written transform to the quadratic forms в F_1 and F_2 to get F_1' and F_2'.

Appendix 1

A short English-Ukrainian vocabulary

A

adjoint	приєднаний, спряжений
algebraic cofactor	алгебраїчне доповнення
associative	асоціативний

B

basis	базис
– arbitrary	– довільний базис
– natural	– канонічний базис
– orthonormal	– ортонормальний базис
bilinear form	білінійна форма
– alternating	– знакозміна
– quadratic	– квадратична
– symmetric	– симетрична
– skew-symmetric	– кососиметрична
block matrix	блочна матриця
bracket	дужка

C

column	стовпчик
commutative	комутативний
complete square	виділити повний квадрат

D

decomposition	розкладання
determinant	визначник
diagonalizable	діагоналізований
dimension	вимір
direct sum	пряма сума
distributive	дистрибутивний
domain	область

	E
eigenspace	власний простір
eigenvalue	власне значення
eigenvector	власний вектор
	F
form	алгебраїчний об'єкт у вигляді поліноміального виразу змінних
	G
Gram – Schmidt orthogonalization	ортогоналізація Грама – Шмідта
– Gram determinant	визначник Грама
– Gram matrix	матриця Грама
	H
homogeneous	однорідний
	I
intersection	перетин
invertible	невироджений
	L
linear span	лінійна оболонка
linear	
– dependency	лінійна залежність
– independency	лінійна незалежність
	M
matrix	матриця
– block	– блочна матриця
– change-of-basis	– матриця перетворення
– Gram	– матриця Грама
– identity	– одинична матриця
– inverse	– обернена матриця
– Hermitian	– Ермітова матриця
– reflection	– матриця відзеркалення
– sparse	– розріджена матриця

– square	– квадратна матриця
– transpose	– транспонована матриця
– transition	– матриця переходу
– triangular	– трикутна матриця
minor	мінор
– additional	– додатковий
– basic	– головний
multiplicity	кратність

O

operator	оператор
– adjoint	– спряжений оператор
– nonsingular	– несингулярний оператор
– orthogonal	– ортогональний оператор
– positive define	– додатно визначений оператор
– projection	– оператор проєктування
– self-adjoint	– самоспряжений оператор
– semi-definite	– напіввизначений оператор
– unitary	– унітарний оператор
orthogonal	
– complement	– ортогональне доповнення
– projection	– ортогональна проєкція

P

pairwise	попарно
pivot	ведучий елемент
principal axis	головні осі
polynomial	
– annihilating	– анігіляційний поліном
– characteristic	– характеристичний поліном
– n-th degree	– поліном n-го порядку

	Q
quadratic form	квадратична форма
	R
rank	ранг
– matrix	– ранг матриці
row	рядок
– reduced row echelon form	– зведена канонічна форма
	S
simultaneously	одночасно
skew-symmetric	кососиметричний
spectral decomposition	спектральне розкладання
space	простір
– Euclidean	– простір Евкліда
– linear	– лінійний простір
– normed	– нормований простір
– vector	– векторний простір
subspace	підпростір
	T
transformation	перетворення
	U
union	об’єднання
	V
vector algebra	векторна алгебра

REFERENCES

1. Andrilli S. Hecker D. *Elementary linear algebra*. Academic Press, Oxford, 5th Ed., 2016. - 793 p.
2. Strang G. *Introduction to linear algebra*. Cambridge Press, Wellesley, 5th Ed., 2017. - 584 p.
3. Lay D. C., Lay S. R., McDonald J. J. *Linear algebra and its applications*. Pearson Education, Inc., New York, 5th Ed., 2016. - 579 p.
4. Norman D., Wolczuk D. Introduction *to linear algebra for science and engineering*. Pearson Education, Inc., Boston, 2nd Ed., 2016. - 550 p.
5. Kirkwood J. R., Kirkwood B. H. *Elementary linear algebra*. CRC Press. Taylor & Francis Group, Boca Raton, 2018. – 323 p.
6. Howard A., Rorres C. *Elementary linear algebra*. Wiley & Sons, Inc., Hoboken, 11th Ed., 2014. – 802 p.
7. Bronson R., Costa G.B., Saccoman J.T. *Linear algebra: algorithms, applications, and techniques.* Academic Press, Waltham, 3rd Ed., 2014. – 519 p.
8. Lipschutz S., Lipson M.L. *Linear algebr.* McGraw-Hill Education, New York, 6th Ed., 2018. – 432 p.
9. Lal A. K., Pati S. *Linear algebra through matrices*. Indian Institute of Technology Kanpur, Draft, 2018, - 245 p. Internet access: https://home.iitk.ac.in/~arlal/MTH102/la.pdf
10. Liniyna alhebra. Zbirka zavdanʹ ta metodyka rozv″yazannya: navchalʹno-metodychnyy posibnyk / L. P. Dzyubak, S. P. Ihlin, H.B. Linnyk, I. O. Morachkovsʹka. – KH.: NTU "KHPI", 2013. – 240 p.
11. Liniyna alhebra:Kurs lektsiy [Elektronnyy resurs]: kurs lekts. dlya stud. spetsialʹnosti 122 «Komp″yuterni nauky» / KPI im. Ihorya Sikorsʹkoho; uklad. YU. YE. Bokhonov. – Elektronni tekstovi dani.–Kyiv: KPI im. Ihorya Sikorsʹkoho, 2022. – 214 p.
12. O.M.Romaniv. Liniyna alhebra. Chastyna 2 : pidruchnyk. Lʹviv: Vydavetsʹ I.E.Chyzhykov. - 2014. – 279 p.

Навчальне видання

КУРПА Лідія Василівна
ЛЮБИЦЬКА Катерина Ігорівна
БУРЛАЄНКО В’ячеслав Миколайович

ЛІНІЙНА АЛГЕБРА:
Курс для студентів інженерно-технічних спеціальностей

Навчальний посібник
для студентів технічних спеціальностей
усіх форм навчання

Англійською мовою

Відповідальний за випуск проф. Галина ТИМЧЕНКО
Роботу до видання рекомендував проф. Дмитро БРЕСЛАВСЬКИЙ

В авторській редакції

План 2024 р., поз. 38

Підп. до друку 12. 03. 2024 р.
Гарнітура Times New Roman.

Видавничий центр НТУ «ХПІ»
Свідоцтво про державну реєстраціюДК № 5478 від 21.08.2017 р.
61002, Харків, вул. Кирпичова, 2

Електронне видання