ПОБЕДИТЕЛЬ

конкурса
рукописей учебников
МЭИ 2020/2021

MINISTRY OF SCIENCE AND HIGHER EDUCATION
OF RUSSIAN FEDERATION

NATIONAL RESEARCH UNIVERSITY ‘MPEI’

M.I. Besova, S.F. Kudin

BASIC COURSE OF LINEAR ALGEBRA AND ANALYTICAL GEOMETRY

Teaching guide
on a basic course ‘Higher mathematics. Linear algebra’

Moscow
Izdatel’stvo MEI
2022

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «МЭИ»

М.И. Бесова, С.Ф. Кудин

БАЗОВЫЙ КУРС ЛИНЕЙНОЙ АЛГЕБРЫ И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Учебное пособие

Москва
Издательство МЭИ
2022

УДК 512.1
ББК 22.143
Б 53

This teaching guide is approved by
the Educational Department of NRU 'MPEI'

The teaching guide was prepared at Higher Mathematics Department of NRU'MPEI

Reviewers: Dr.Sci. in Physics and Mathematics, professor of Mathematical and Computer Modelling Department ofNRU 'MPEI
Dr.Sci. in Physics and Mathematics, senior lecturer of MIPT V.Zh. Sakbaev

Besova, M.I.

Б 53 Basic Course of Linear Algebra and Analytical Geometry (teaching guide)/ M.I. Besova, S.F. Kudin. – Moscow, MPEI Publishing House (Izdatel'stvo MEI), 2022. – 140 p.

ISBN 978-5-7046-2582-7

The teaching guide focuses on the basic concepts, important theoretical and practical material for the course "Linear Algebra and Analytical Geometry". Detailed solutions of basic tasks for this course are given and tasks for independent implementation are proposed. The main purpose of the textbook is to help students, for whom Russian is not their native language, in learning the concepts of the theoretical provisions and in the presentation of methods for solving mathematical problems arising within the course. The sections of the teaching guide contain an introduction to the theory of matrices, methods for solving systems of linear algebraic equations, the main concepts of analytical geometry and the basics of the section "Linear operators". The textbook also includes solutions of the basic problems for the course "Linear Algebra and Analytical Geometry" in the MathCad system.

Intended for students of NRU "MPEI" studying in accordance with bachelors' educational programs for technical specialities.

УДК 512.1
ББК 22.143

ISBN 978-5-7046-2582-7

УДК 512.1
ББК 22.143
Б 53

*Утверждено учебным управлением НИУ «МЭИ»
в качестве учебного издания*

Подготовлено на кафедре высшей математики

Рецензенты: докт. физ.-мат. н., проф. каф. МКМ НИУ «МЭИ»
А.В. Перескоков,
докт. ф.-м.н., доц. МФТИ В.Ж. Сакбаев

Бесова, М.И.
Б 53 Базовый курс линейной алгебры и аналитической геометрии: учеб. пособие / М.И. Бесова, С.Ф. Кудин. − М.: Издательство МЭИ, 2022. − 140 с.

ISBN 978-5-7046-2582-7

В пособии изложены основные понятия, важный теоретический и практический материал по курсу «Линейная алгебра». Приведены подробные решения базовых задач по данному курсу и предложены задания для самостоятельного выполнения. Основная цель учебного пособия − помощь студентам, для которых русский язык не является родным, в усвоении теоретических положений курса и в изложении методики решения математических задач, возникающих в рамках курса. Разделы пособия содержат введение в теорию матриц, методы решения систем линейных алгебраических уравнений, основные положения аналитической геометрии и основы раздела «Линейные операторы». В пособие включены также расчеты основных задач курса «Линейная алгебра и аналитическая геометрия» в системе MathCad.

Предназначено для студентов НИУ «МЭИ», обучающихся по направлениям подготовки бакалавров технических специальностей.

УДК 512.1
ББК 22.143

Издано в авторской редакции

ISBN 978-5-7046-2582-7

Table of contents

BASIC COURSE OF LINEAR ALGEBRA AND ANALYTICAL GEOMETRY

Teaching guide

In this textbook the authors explain the basic concepts, the most important theoretical material and some guides to the problem solving for the course of linear algebra and analytical geometry. A book may help the foreign students to learn this course on their first year of study at Moscow Power Engineering Institution. The text contains basic definitions, formulae and principal laws of linear algebra and analytical geometry. The typical problems for each theme are described in detail and solved. After most of the chapters there are review questions and problems for the students that they should solve themselves. To make the studies more interesting, the facts from the history of mathematics and some creative tasks for the students are also included into this teaching guide. The main aim of this textbook is to explain the basic concepts of linear algebra and analytical geometry, to help the students in developing the skills of solving the problems and to help them in making their first steps in working with the mathematical software. The chapters of this teaching guide contatin three big themes – 'Introduction to matrix theory. Systems of linear equations', 'Analytical geometry' , 'Linear spaces and linear operators'. Some basic problems from the theme 'Introduction to matrix theory' are solved in Mathcad program; they are given in chapter 4. That helps the students to get acquainted with Mathcad software and responds the challenges of the practical branch of education because it is important for the students to learn how to work with specialized software and to practice their digital skills.

This teaching guide can be recommended as the basic textbook for the students of NRU 'MPEI' in English-speaking groups, and it also can be used as the additional teaching guide for all the technical specialties of the university.

INTRODUCTION

Linear algebra is one of the most important branches of Higher mathematics. We cannot imagine the modern life, with its digital challenges, without the concepts of a matrix theory that helps the computers to work with huge arrays, to organize the data, to build complicated mathematical models. Discovered in ancient times, the methods of solving the systems of linear equations help us to solve different problems nowadays almost in each sphere of our everyday life.

First textbooks in English on different courses of Higher mathematics were published by Nabebin A.A. in 1990s, however, since then at MPEI there were no full teaching guides dedicated to all the themes investigated in the course of linear algebra. The authors of this book often hear the requests from the students who come to Moscow Power Engineering University from the other countries – to find a good textbook in English which can help their studies easier. A language bareer is a typical problem for almost each student from the other country, so the authors have chosen the English language for their teaching guide (as the most ‘universal’ and widely-spread one) and they try to explain all the basic concepts of linear algebra in a clear and understandable language.

The course of linear algebra and analytical geometry introduced in this book, helps the students to learn the basic theoretical material, to practice solving the typical problems of the course. They can check themselves by the review questions (mostly they help the students to repeat the material of the previous paragraph in order to develop their skills of oral answers and to concentrate on the basic terms introduced in the theme). Some useful terms from each paragraph are introduced in Russian, and that helps the students from the other country to get acquainted with the special terms they will definitely meet in their future studies and to improve their vocabulary of Russian language. If the students have real interest in mathematics, they can read the inspiring biographies of great mathematicians and interesting facts about the history of mathematics. The authors hope that the students will enjoy the creative tasks that can help the students in developing their skills and knowledge and enjoy the educational process at the same time. And, of course, as the concept of a ‘digital textbook’ becomes increasingly popular, there is a simple and understandable guide for solving some basic problems of the matrix theory in chapter 4.

The authors hope that the process of learning the course of linear algebra will be an exciting journey into the world of Higher mathematics and wish all the students good luck in their future studies!

1. INTRODUCTION TO MATRIX THEORY. SYSTEMS OF LINEAR EQUATIONS

1.1. Matrices. Main definitions

A matrix is definitely a first thing to be introduced in the matrix theory. You have probably heard about it before. A matrix is a unique object that plays a significant role not only in mathematics but also in many other fields of science. Firstly we need it when we have to solve a complicated system of linear equations. Let us investigate matrices in terms of linear algebra.

A matrix is a rectangular array of numbers or symbols. (Sometimes a matrix can even include expressions). The most important parts of the matrix are rows and columns.

A row is an array of matrix elements arranged as a single horizontal line; ***a column*** is an array of matrix elements arranged as a single vertical line. Let us look at the example of a typical matrix.

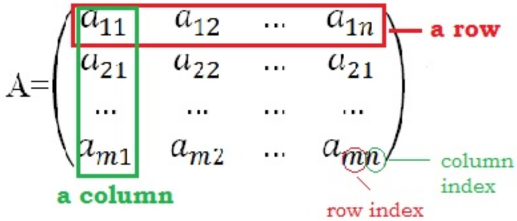
$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{21} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Each element of the matrix has its own index. It consists of 2 numbers: the first number indicates the row number of the element and the second number indicates the column number of this element. Further in the common form we would denote the element of a matrix A as a_{ij} where the index i indicates the row number and j indicates the column number.

So we can see the detailed description of any unspecified matrix.

This matrix A is called ***a matrix of dimension m × n*** or just ***m × n matrix*** (we should read it as 'm by n') as it has ***m*** rows and ***n*** columns.

There are many interesting private cases of matrices. One of the most useful examples is a ***square matrix***. The number of rows in a square matrix is equal to the number of its columns. (For example, 2×2, 3 × 3, 4 × 4…$n \times n$ matrices). Speaking about the square matrix, we introduce the term of a diagonal.



<u>The main diagonal of a matrix</u> is an array of matrix entries where the row index is equal to the column index, e.g. $a_{11}, a_{22}, \ldots a_{nn}$. So we can really see they are standing on the diagonal line from the upper left to the lower right corner of a matrix. It is also sometimes called ***major diagonal, principal diagonal, primary diagonal or leading diagonal.***

<u>The secondary diagonal of a square matrix</u> (let us take $m \times m$ matrix for our example) is an array of entries a_{ij} where $i + j = m + 1$. It is a diagonal line coming from the upper right corner to the lower left corner. It is also sometimes called ***counter diagonal, trailing diagonal, antidiagonal*** or even ***bad diagonal*** (although the authors consider that all the mathematical objects are good!). Let us investigate the 3×3 matrix A:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

We can see that a_{11}, a_{22}, a_{33} are entries of the main diagonal (the row index is equal to the column index in those entries) and a_{13}, a_{22}, a_{31} are entries of the secondary diagonal. The dimension of our matrix is 3×3 so the sum of row index and column index of the entry on a secondary diagonal should be equal to 4. So we can easily check that this condition is fulfilled with a_{13}, a_{22}, a_{31} .

One of the interesting cases is a ***zero matrix***: a matrix that has all the elements equal to zero. You can look at the example of a zero matrix M:

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Another important case is a ***diagonal matrix:*** a square matrix that has all the non-zero entries standing on the main diagonal while the other elements of this matrix are equal to zero. Below we give the example of a diagonal matrix *B*.

$$B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

One of the private cases of a diagonal matrix is the identity matrix: a square matrix, which has the only number 1 on the main diagonal; the other entries are equal to zero. Below we can see the 3×3 identity matrix I (usually we denote the identity matrix by this letter).

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

We can also investigate the ***vectors*** – the matrices which have only one column or one row. A row vector consists of only one row and a column vector consists of only one column. Below we see the example of a row vector D and a column vector *F*.

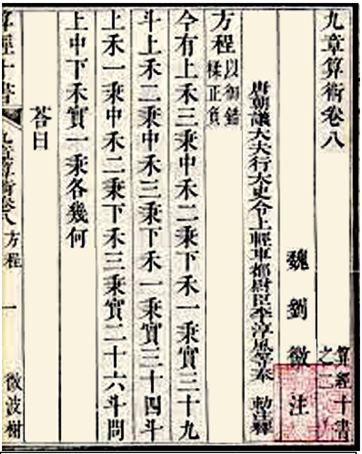
$$D = (d_{11}\, d_{12}\ \dots d_{1n})$$

$$F = \begin{pmatrix} f_{11} \\ f_{21} \\ \dots \\ f_{m1} \end{pmatrix}$$

We can see that the row vector *D* has a dimension $1\times n$ and a column vector *F* has a dimension $m\times 1$.

<u>History is a great teacher!</u> The first matrix-like structures were described in Ancient China. The book 'The Nine Chapters on the Mathematical Art' (written by several generations from 10-th century BCE to the 2-nd century BCE) was the first document that told us about the 'equations' similar to modern simultaneous linear systems of equations.

The solution method called 'Fang Cheng Shi' was a matrix method of solving the systems of linear equations known today as Gaussian elimination (we will talk about it later in the course). However, a Genevan mathematician Gabriel Cramer (1704–1752) and a German mathematician Carl Friedrich Gauss (1777–1855), with their methods of solving the linear equations, made a great contribution to forming a matrix theory only in the 18-th century – more than a thousand years after the investigations in Ancient China. The theory was developed in the 19-th and the 20-th century, and it got many useful applications.



A page from 'Nine Chapters on the Mathematical Art'

Creative task Find the information about the fields of science, everyday life or technology where the matrices are useful nowadays. Make a short report (not more than 2 minutes!)

Practice Russian

Ма́трица – a matrix
Строка́ – a row
Столбе́ц – a column
Диагона́ль – a diagonal
Квадра́тная ма́трица – a square matrix
Единичная ма́трица – an identity matrix

Review questions

1. What is a matrix? What types of matrices have you learned?
2. Describe an algorithm of matrix elements' numeration
3. What is a vector? Compare your previous knowledge about vectors (from school course of mathematics) and the new definition

1.2. Addition of matrices. Multiplication by a scalar

There is a range of operations we can do with matrices. Let us get acquainted with the process of ***matrix addition*** – it is very simple!

For our example we take 2 matrices, A and B.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \qquad (1)$$

We need to calculate the matrix C which will be the sum of those two matrices. So we need to take two members with both same row and column indexes from the matrices A and B, add the entry of a first matrix to the entry of a second matrix and put the result to the place of the entry with the same row and column index in the new matrix C. So $a_{ij} + b_{ij} = c_{ij}$.

$$C = A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} \end{pmatrix} = \\ = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$$

We can also subtract one matrix from the other matrix using the same rule.

$$D = A - B = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} & a_{13} - b_{13} \\ a_{21} - b_{21} & a_{22} - b_{22} & a_{23} - b_{23} \\ a_{31} - b_{31} & a_{32} - b_{32} & a_{33} - b_{33} \end{pmatrix} = \\ = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix}$$

Another simple operation is **multiplication by a scalar.** A ***scalar*** is a real number. So when we multiply a matrix by a scalar, we multiply each entry of matrix by this scalar and we put this entry multiplied by this scalar to the proper place of a resulting matrix.

For example, if we have a scalar λ and multiply it by the matrix A (see the matrix A in (1)):

$$X = \lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{pmatrix}$$

Let's come to some practical examples.

So we need to evaluate the matrix C. We add the entries with the same indexes in both matrices, A and B, and put the result into the new matrix.

$$C = \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix} + \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & 7 \\ 8 & -9 & 10 \end{pmatrix} =$$

$$= \begin{pmatrix} 3+1 & 2-2 & 1+0 \\ 0+4 & -1+5 & 4+7 \\ 1+8 & 2-9 & 6+10 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 1 \\ 4 & 4 & 11 \\ 9 & -7 & 16 \end{pmatrix}$$

So $C = \begin{pmatrix} 4 & 0 & 1 \\ 4 & 4 & 11 \\ 9 & -7 & 16 \end{pmatrix}$.

1. Subtract the matrix B from the matrix A (see the matrices in the previous task). Put the result into the matrix D.

We take the entries with the same indexes in both matrices, A and B, and subtract the entries of the matrix B from the corresponding entries of the matrix A and put the result into the new matrix.

$$D = \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix} - \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & 7 \\ 8 & -9 & 10 \end{pmatrix} =$$

$$= \begin{pmatrix} 3-1 & 2+2 & 1-0 \\ 0-4 & -1-5 & 4-7 \\ 1-8 & 2+9 & 6-10 \end{pmatrix} =$$

$$= \begin{pmatrix} 2 & 4 & 1 \\ -4 & -6 & -3 \\ -7 & 11 & -4 \end{pmatrix}$$

So $D = \begin{pmatrix} 2 & 4 & 1 \\ -4 & -6 & -3 \\ -7 & 11 & -4 \end{pmatrix}$.

2. Multiply the matrix A from the previous tasks by a scalar $\lambda = 5$.

We multiply each entry of the matrix A by a scalar $\lambda=5$.

$$X = \lambda A = 5 \cdot \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix} = \begin{pmatrix} 3 \cdot 5 & 2 \cdot 5 & 1 \cdot 5 \\ 0 \cdot 5 & -1 \cdot 5 & 4 \cdot 5 \\ 1 \cdot 5 & 2 \cdot 5 & 6 \cdot 5 \end{pmatrix} = \begin{pmatrix} 15 & 10 & 5 \\ 0 & -5 & 20 \\ 5 & 10 & 30 \end{pmatrix}$$

So $X = \begin{pmatrix} 15 & 10 & 5 \\ 0 & -5 & 20 \\ 5 & 10 & 30 \end{pmatrix}$.

Practice Russian

Сложéние мáтриц – matrix addition

Вычитáние мáтриц – matrix subtraction

Умножéние на число – multiplication by a scalar

Review questions

1. Describe the algorithm of matrices' addition.
2. Is the addition of matrices commutative? (Is the following expression true: $A + B = B + A$?)
3. Describe the algorithm of multiplication by a scalar

Practical tasks

Choose your variant and solve the task.

1. Add matrix A to the matrix B and put the result into the matrix C.

1. $A = \begin{pmatrix} 4 & 6 & -1 \\ 0 & -2 & 3 \\ 7 & 3 & 9 \end{pmatrix}$	$B = \begin{pmatrix} 6 & -5 & 7 \\ 3 & 8 & -3 \\ 5 & -1 & 1 \end{pmatrix}$
2. $A = \begin{pmatrix} -1 & 5 & 9 \\ 6 & -5 & 7 \\ 5 & -4 & 0 \end{pmatrix}$	$B = \begin{pmatrix} 8 & -1 & -4 \\ 2 & -7 & 2 \\ 9 & 5 & 6 \end{pmatrix}$
3. $A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix}$	$B = \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & 7 \\ 8 & -9 & 10 \end{pmatrix}$

4. $A=\begin{pmatrix}10 & 9 & -3\\ 8 & -2 & 7\\ 7 & 4 & 5\end{pmatrix}$	$B=\begin{pmatrix}8 & 9 & 0\\ 10 & 6 & 4\\ 5 & -1 & 2\end{pmatrix}$
5. $A=\begin{pmatrix}6 & -9 & 4\\ 10 & 2 & 3\\ 4 & -5 & 1\end{pmatrix}$	$B=\begin{pmatrix}-6 & 0 & 5\\ 2 & 3 & 4\\ 7 & 10 & 6\end{pmatrix}$
6. $A=\begin{pmatrix}9 & -7 & 3\\ 4 & -3 & 5\\ -1 & 7 & 0\end{pmatrix}$	$B=\begin{pmatrix}10 & 6 & -1\\ 7 & 8 & 4\\ 5 & -4 & 7\end{pmatrix}$
7. $A=\begin{pmatrix}1 & -3 & 10\\ 0 & 5 & -4\\ 7 & -8 & 6\end{pmatrix}$	$B=\begin{pmatrix}9 & 1 & 10\\ -2 & 3 & 8\\ 4 & 0 & -1\end{pmatrix}$
8. $A=\begin{pmatrix}-3 & 6 & 7\\ -7 & 1 & 9\\ -6 & 0 & 7\end{pmatrix}$	$B=\begin{pmatrix}9 & 10 & -3\\ -2 & 5 & 9\\ 1 & 2 & -5\end{pmatrix}$
9. $A=\begin{pmatrix}4 & -5 & 8\\ 6 & -2 & 3\\ 10 & 8 & 5\end{pmatrix}$	$B=\begin{pmatrix}0 & -3 & -2\\ 7 & 1 & -6\\ 2 & 3 & 1\end{pmatrix}$
10. $A=\begin{pmatrix}2 & 10 & 6\\ 4 & 10 & 7\\ 7 & 2 & -8\end{pmatrix}$	$B=\begin{pmatrix}6 & -7 & 9\\ 3 & 5 & 8\\ 8 & -6 & -7\end{pmatrix}$
11. $A=\begin{pmatrix}5 & 8 & -2\\ -5 & 6 & 2\\ 9 & 1 & 3\end{pmatrix}$	$B=\begin{pmatrix}4 & -1 & 9\\ 6 & 4 & 8\\ -10 & 1 & 6\end{pmatrix}$
12. $A=\begin{pmatrix}7 & 1 & 1\\ 2 & 5 & -9\\ 3 & 4 & 2\end{pmatrix}$	$B=\begin{pmatrix}0 & -1 & 0\\ 2 & 4 & 7\\ 7 & 2 & -8\end{pmatrix}$
13. $A=\begin{pmatrix}9 & 10 & 0\\ -5 & 6 & 1\\ -9 & 5 & 3\end{pmatrix}$	$B=\begin{pmatrix}2 & 7 & 1\\ 8 & 3 & -4\\ 0 & 5 & -2\end{pmatrix}$
14. $A=\begin{pmatrix}6 & 8 & -5\\ -7 & 4 & 0\\ 9 & 1 & -3\end{pmatrix}$	$B=\begin{pmatrix}9 & 7 & -1\\ 3 & 4 & -2\\ 5 & 6 & -7\end{pmatrix}$
15. $A=\begin{pmatrix}7 & -1 & 8\\ 5 & 3 & -4\\ 6 & 0 & 2\end{pmatrix}$	$B=\begin{pmatrix}3 & 9 & -2\\ 1 & -8 & 5\\ 4 & 7 & 2\end{pmatrix}$

1. Subtract matrix B from the matrix *A* and put the result into the matrix *C*. Take the matrices from the previous task.

2. Multiply the matrix *A* (see task 1) by a scalar λ. Take the matrix *A* from task 1. The scalars are given below.

Variant number	λ
1	3
2	–2
3	5
4	4
5	8
6	7
7	6
8	–5
9	10
10	2
11	–8
12	–4
13	–7
14	9
15	–3

1.3. Matrix multiplication

We have already learned how to multiply a matrix by a scalar – but what do we have to do when we need to multiply one matrix by the other one? Sometimes we need to multiply three or even more matrices at once…

First of all, we need to multiply the matrices consequently: firstly we multiply the first matrix by the second one, then the new resulting matrix is multiplied by the third one and so on. So the main algorithm we need is multiplication of two matrices.

To multiply two matrices we should remember the 'magic spell': 'Row by column and take the sum'. What does that mean?

For example, we have two matrices *A* and *B*:

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

So we need to get the matrix *C*. Let us look at the entry of the new matrix c_{ij} (for the common form we use the entry in the *i*-th row and the j-th column) and the algorithm of its evaluation:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots + a_{in}b_{nj}$$

We take the whole *i*-th row and the whole *j*-th column and multiply the first entry in the row by the first entry in the column, the second entry in the row by the second entry in the column and so on. So after multiplying all the corresponding members in one row and one column, we need to get a sum of all those pairwise products.

Let us look at the concrete example of the 3×3 matrices.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

To get c_{11} we take the first row of matrix *A* and the first column of matrix *B*.

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31}$$

To get c_{12} we take the first row of a matrix *A* and the second column of the matrix *B*.

$$c_{12} = a_{11}b_{21} + a_{12}b_{22} + a_{13}b_{32}$$

To get c_{23} we take the second row of a matrix *A* and the third column of a matrix *B*:

$$c_{23}=a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33}$$

It is important to mention that when we need to multiply two matrices, the number of rows in the first matrix should be equal to the number of columns in the second matrix.

Useful advice! For the first time you should write the resulting matrix you are going to get (in the common form). For example, we need to get 3×3 matrix *C* as a product of two 3×3 matrices *A* and *B* given above:

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$$

You should tell the algorithm to yourself: firstly I need to get the entry c_{11}. The row index of this entry is **1** so I take the **first row of the first matrix A**; the column index is also **1** so I take **the first column of the matrix B**. I multiply the first entry of the first row in the matrix A by the first entry of the first column in the matrix B. Then I multiply the second entry of the first row in the matrix A and the second entry of the first column in the matrix B (don't forget we are still working with the same row and column until we come to the last entries there). Then we repeat the procedure with the third entry in the first row of A and the third entry in the first column of B. I add the products of those first entries to the product of the second entries, get the result and then afterwards I add the product of the third entries. So we see that when the last entry of the first row and the last entry of the first column are pairwise multiplied and added to the other pairwise products for this row and this column, we come to the next entry of the resulting matrix C.

The next entry is c_{12}. So your words should sound like this: the row index is **1** so I take the **first row of *A*** and the column index is **2** so I take the **second column of the matrix *B***. Then all the operations described above should be repeated with the entries of the first row in A and the second row in B.

So two indexes of the new resulting matrix can help you to define the necessary row and columns in which the entries should be pairwise multiplied and then the sum of those pairwise products should be calculated.

<u>It's important!</u> When you calculate the product of two real numbers, you get used to the commutativity law: ***ab* = *ba***. Let us check whether this law is fulfilled for two matrices. So we have two 2×2 matrices, A and B and we should check whether $A\cdot B = B\cdot A$:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

So using the rule of multiplication, we get the matrix $C = A\cdot B$

$$C = A\cdot B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} =$$

$$= \begin{pmatrix} 1\cdot5+2\cdot7 & 1\cdot6+2\cdot8 \\ 3\cdot5+4\cdot7 & 3\cdot6+4\cdot8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

Then let us try to evaluate the matrix $D = B \cdot A$. We should check whether $C = D$ or not.

$$D = B \cdot A = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} =$$

$$= \begin{pmatrix} 5 \cdot 1 + 6 \cdot 3 & 5 \cdot 2 + 6 \cdot 4 \\ 7 \cdot 1 + 8 \cdot 3 & 7 \cdot 2 + 8 \cdot 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

So we can conclude that the multiplication of matrices is **not commutative:** $A \cdot B \neq B \cdot A$.

Remark **There are some private cases when $A \cdot B = B \cdot A$. The easiest way to illustrate one of those cases is to take two matrices of the same dimension where all the entries would be the same:**

a) $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $A \cdot B = B \cdot A$

b) $A = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix}$ $B = \begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix}$ $A \cdot B = B \cdot A$

One-minute task Your task is to check those cases and calculate the products $A \cdot B$ and $B \cdot A$ in cases a) and b).

We can also multiply more than two matrices but it is necessary to check their dimension!

For example, we have to multiply three matrices:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}$$

We remember that we can't speak about the commutativity so we need to multiply the matrices consequently: the first matrix by the second matrix. Then the resulting matrix of those two is multiplied by the third matrix.

So we have:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 2 \cdot 3 & 1 \cdot 3 + 2 \cdot 4 \\ 2 \cdot 2 + 3 \cdot 3 & 2 \cdot 3 + 3 \cdot 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} =$$

$$= \begin{pmatrix} 8 & 11 \\ 13 & 18 \end{pmatrix} \begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 8 \cdot 3 + 11 \cdot 5 & 8 \cdot 4 + 11 \cdot 6 \\ 13 \cdot 3 + 18 \cdot 5 & 13 \cdot 4 + 18 \cdot 6 \end{pmatrix} =$$

$$= \begin{pmatrix} 79 & 98 \\ 129 & 160 \end{pmatrix}$$

Creative task Now you can play the game. The student who begins the game should write any 2×2 matrices to the student sitting next to him/her and that second student should calculate the product of those matrices. Then the second student puts down the resulting matrix and then invents any new matrix he would like to write down. So the third student multiplies the previous resulting matrix (calculated by the second student) by the new matrix invented by the previous student. After putting down the new resulting matrix, the third student invents another matrix for the next student to multiply that new resulting matrix by the invented one. So the last student in the chain gives to the first student (who has begun the game) the matrix that he has got as the result of multiplication plus his own one. So the first student calculates the product and has to finish the game. You are allowed to use the calculator during the game: the numbers can be really impressive, especially for the last students in the chain.

Practice Russian

Умноже́ние ма́триц – matrix multiplication

Review questions

1. Describe the algorithm of matrix multiplication
2. Is the multiplication of matrices commutative?
3. Give your own examples of the private cases when $A\cdot B = B\cdot A$

Practical task

All the matrices for this task are already given above in the practical task 1 after the paragraph 1.2. However, this time you need to take the variant which comes after your own one: for example, if you have variant 1, you need to take variant 2 and so on; variant 15 should take the variant 1.

1. Multiply the matrix *A* by the matrix *B* and put the result into the matrix *C*. Then multiply the matrix *C* by the 3×3 identity matrix *I*.

1.4. Determinants of matrices

The determinant of a matrix is a useful mathematical object that helps us to solve complicated systems of linear equations; it is also widely used in analytical geometry (its application will be investigated later in our course). A determinant is a special number that can be evaluated from a square matrix.

Remark Don't forget that we work only with **square matrices** when we talk about calculation of a determinant!

So let us see how we can work with the determinants of matrices. We can have a look at the 3×3 matrix *A*.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

When we calculate the determinant of a matrix, first we put our matrix in straight lines (like the modulus sign). The denotation for the determinant of the matrix *A* is given below:

$$det\ A = |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

We need to calculate the value of this determinant.

1) The determinant of 2×2 matrix

For example, we have the 2×2 matrix *B*:

$$|B| = \begin{vmatrix} 2 & 6 \\ 7 & 1 \end{vmatrix}$$

Step 1 We multiply the elements of the main diagonal and get the result: $2 \cdot 1 = 2$

Step 2 We multiply the elements of the secondary diagonal and get the result: $6 \cdot 7 = 42$

Step 3 We subtract the product of the secondary diagonal's elements from the product of the main diagonal's elements: $2 - 42 = -40$.

So the value of the determinant for the matrix *B* is $|B| = -40$.

2) The determinant of a 3×3 matrix

The case of a determinant for a 3×3 matrix is very popular among the practical tasks for students. Moreover, it is a unique case of a determinant which has two convenient methods of evaluation. One method is rather universal and can be used for matrices with a higher dimension; the second one is used only for 3×3 matrices.

2.1. The row /column expansion method

Let us investigate the matrix D and its determinant:

$$D = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{pmatrix} \; |D| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{vmatrix}$$

Firstly, we need to introduce the definition of minors. What is a minor and how can we get it? Let us learn the rule. We take one entry of a matrix (in most cases we start from the very beginning so we take $d_{11} = 1$) and in our imagination we remove the whole row and the whole column where this entry is situated and take the four numbers that remain untouched as a new determinant. In our case the minor for d_{11} is $\begin{vmatrix} 7 & -1 \\ 6 & 1 \end{vmatrix}$.

However all the minors have the signs '+' or '–' and we put them chequerwise. For the minor of the first entry (in the scheme below its sign stands in the very first position) – in our case this entry is d_{11}- we always put '+', then for the minor of the next entry – for us this entry is d_{12} – the sign changes for '–'. For the second element in the first column the sign also is '–': if we investigate two elements standing next to each other, no matter in the row or in the column, the signs of their minors will always be different. In the scheme below we see the sign of each minor stands in the position of the element for which this minor is evaluated. We can see that for the elements in odd rows the signs of the minors begin with '+'; for even rows the signs begin with '–')

$$\begin{vmatrix} + & - & + & - & + & - & + & - & \dots \\ - & + & - & + & - & + & - & + & \dots \\ + & - & + & - & + & - & + & - & \dots \\ - & + & - & + & - & + & - & + & \dots \end{vmatrix}$$

So you can guess that we can put our minors into a special array. It is called the adjoint matrix and, if we speak about the given matrix D, the adjoint matrix is denoted as D^*. Let us put all the minors of all the entries there (we will denote the minor of the element d_{ij} as M_{ij}). This matrix is given below (pay attention to the signs!):

$$D^* = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{21} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} 7 & -1 \\ 6 & 1 \end{vmatrix} & -\begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 4 & 7 \\ 2 & 6 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 6 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 6 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 7 & -1 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 4 & -1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 4 & 7 \end{vmatrix} \end{pmatrix}$$

Now let us come to the row/column expansion method. It can be used for 3×3 matrices as well as for matrices of higher dimension. For determinant evaluation we should take any row or column of the determinant $|D|$.

For the row expansion we have a standard formula (the common form for the i-th row in the matrix with n columns):

$$|D| = d_{i1} \cdot M_{i1} + d_{i2} \cdot M_{i2} + \ldots + d_{in} M_{in}$$

For the column expansion we have a standard formula (the common form for the j-th column in the matrix with s rows)

$$|D| = d_{1j} \cdot M_{1j} + d_{2j} \cdot M_{2j} + \ldots + d_{sj} M_{sj}$$

Let us take the first row of our given matrix D. The formula for calculation of a determinant is given below:

$$|D| = d_{11} M_{11} + d_{12} M_{12} + d_{13} M_{13}$$

Let us put the concrete numbers and symbols from the matrix D and the adjoint matrix D^* (we should calculate the minors as a determinant of a 2×2 matrix):

$$|D| = 1 \cdot \begin{vmatrix} 7 & -1 \\ 6 & 1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 7 \\ 2 & 6 \end{vmatrix} =$$

$$= 1 \cdot (7 \cdot 1 - (-1 \cdot 6)) - 2 \cdot (4 \cdot 1 - (-1 \cdot 2)) + 3 \cdot (4 \cdot 6 - 7 \cdot 2) = 13 - 12 + 30 = 31$$

Let us check whether this algorithm works for the column expansion. So we can take the third column and put down the formula:

$$|D| = d_{13} M_{13} + d_{23} M_{23} + d_{33} M_{33}$$

When we take the concrete numbers we get:

$$|D| = 3\cdot\begin{vmatrix} 4 & 7 \\ 2 & 6 \end{vmatrix} - (-1)\cdot\begin{vmatrix} 1 & 2 \\ 2 & 6 \end{vmatrix} + 1\cdot\begin{vmatrix} 1 & 2 \\ 4 & 7 \end{vmatrix} =$$

$$= 3\cdot(4\cdot6 - 7\cdot2) + 1\cdot(1\cdot6 - 2\cdot2) + 1\cdot(1\cdot7 - 2\cdot4) = 30 + 2 - 1 = 31$$

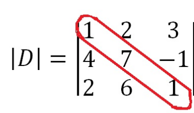
So we proved that the value of a determinant remains the same – no matter whether we take the row expansion or column expansion. We can work with any row or column in the matrix – it's important to pay attention to the minors' signs.

2.2. The triangle method

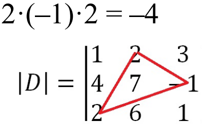
We use this method only for the calculation of a 3×3 matrix determinant. We take the matrix D from the previous example.

$$D = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{pmatrix} \quad |D| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{vmatrix}$$

Step 1. We take the main diagonal and multiply all three entries standing there: $d_{11}\cdot d_{22}\cdot d_{33} = 1\cdot7\cdot1 = 7$



Step 2. We form the first triangle. If a main diagonal can be presented as a straight line, then we take the parallel line to it. We see there are two parallel lines: the line with entries $d_{12} = 2$ and $d_{23} = -1$ and the line with entries $d_{21} = 4$ and $d_{32} = 6$. It doesn't matter whether we take the parallel line above the main diagonal or below the main diagonal for the first triangle – so let us take the line above. We need to make a triangle so we need to find the third distant point in the corner of a matrix (if we consider one entry to be a triangle-apex) for all the triangle apexes on that straight line: $d_{12} = 2$ and $d_{23} = -1$. As the straight line is in the upper right corner of a matrix, the third apex should be in the lower left corner of a matrix and it is d_{31}. We multiply all three entries that form the triangle: $d_{12}\cdot d_{23}\cdot d_{31} = 2\cdot(-1)\cdot2 = -4$



Step 3. We form the second triangle. We take the parallel line of entries lying below the main diagonal: d_{21} = 4 and d_{32} = 6. Then we should look for the third apex. The line is situated in the lower left corner of a matrix so the third triangle apex should be in the upper right corner: d_{13} = 3. We multiply all the entries that form the second triangle: $d_{21} \cdot d_{32} \cdot d_{13} = 4 \cdot 6 \cdot 3 = 72$

$$|D| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{vmatrix}$$

Step 4. We add all the products: the product of the main diagonal's entries, the product of the first triangle's entries and the product of the second triangle's entries:

$$d_{11} \cdot d_{22} \cdot d_{33} + d_{12} \cdot d_{23} \cdot d_{31} + d_{21} \cdot d_{32} \cdot d_{13} =$$
$$= (1 \cdot 7 \cdot 1) + (2 \cdot (-1) \cdot 2) + (4 \cdot 6 \cdot 3) = 7 - 4 + 72 = 75$$

Step 5. We take the secondary diagonal and multiply all three entries standing there: $d_{13} \cdot d_{22} \cdot d_{31} = 3 \cdot 7 \cdot 2 = 42$

$$|D| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{vmatrix}$$

Step 6. We form the third triangle. We take the parallel lines to the secondary diagonal. We see there are two parallel lines: the line with entries d_{12} = 2 and d_{21} = 4 and the line with entries d_{23} = –1 and d_{32} = 6. As well as for the main diagonal, it doesn't matter whether we take the parallel line above the main diagonal or below the main diagonal for the third triangle – so let us take the line above with the entries d_{12} and d_{21}. As the straight line is in the upper left corner of a matrix, the third apex should be in the lower right corner of a matrix and it is d_{33}. We multiply all three entries that form the triangle: $d_{12} \cdot d_{21} \cdot d_{33} = 2 \cdot 4 \cdot 1 = 8$.

$$|D| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{vmatrix}$$

Step 7. We form the fourth triangle. We take the parallel line of entries lying below the secondary diagonal: $d_{23} = -1$ and $d_{32} = 6$. The line is situated in the lower right corner of a matrix so the third triangle apex should be in the upper left corner: $d_{11} = 1$. We multiply all the entries that form the second triangle: $d_{23} \cdot d_{32} \cdot d_{13} = -1 \cdot 6 \cdot 1 = -6$

$$|D| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 7 & -1 \\ 2 & 6 & 1 \end{vmatrix}$$

Step 8. We summarize the product of the secondary diagonal's entries, the product of the third triangle's entries and the product of the fourth triangle's entries:

$$d_{13} \cdot d_{22} \cdot d_{31} + d_{12} \cdot d_{21} \cdot d_{33} + d_{23} \cdot d_{32} \cdot d_{11} =$$
$$= (3 \cdot 7 \cdot 2) + (2 \cdot 4 \cdot 1) + ((-1) \cdot 6 \cdot 1) = 42 + 8 - 6 = 44$$

Step 9. We subtract the product that we've got in step 8 from the product we've got in step 4 (the sum of the secondary diagonal's entries product and the products that refer to the third and fourth triangles are subtracted from the sum of the main diagonal's entries product and the products that refer to the first and the second triangles) and get the determinant value:

So let us calculate the value of a determinant |D|:

$$|D| = d_{11} \cdot d_{22} \cdot d_{33} + d_{12} \cdot d_{23} \cdot d_{31} + d_{21} \cdot d_{32} \cdot \mathrm{d}_{13} -$$
$$- (d_{13} \cdot d_{22} \cdot d_{31} + d_{12} \cdot d_{21} \cdot d_{33} + d_{23} \cdot d_{32} \cdot d_{11}) =$$
$$= d_{11} \cdot d_{22} \cdot d_{33} + d_{12} \cdot d_{23} \cdot d_{31} + d_{21} \cdot d_{32} \cdot d_{13} -$$
$$- d_{13} \cdot d_{22} \cdot d_{31} - d_{12} \cdot d_{21} \cdot d_{33} - d_{23} \cdot d_{32} \cdot d_{11} = 75 - 44 = 31$$

Now we see that both methods of row/column expansion and the triangle method give us the same results.

Practice Russian

Определитель (детерми́нант) ма́трицы– determinant of a matrix

Ме́тод разложе́ния по строке́/столбцу́ – row/column expansion method

Ме́тод треуго́льников – the triangle method

Review questions

1. Describe the algorithm of determinant calculation for 3x3 matrix(row/column expansion method)
2. Describe the algorithm of determinant calculation for a 3x3 matrix (the triangle method)
3. We have shown the example of determinant calculation using the expansion of the first row. Try to calculate the determinant of a matrix *D* using the column expansion (take any column you like).

Practical task

All the matrices for this task are already given above in the practical task 1 after the paragraph 1.2. However, this time you need to take your variant as your own one plus 2: for example, if you have variant 1, you need to take variant 3 and so on; variant 14 should take the variant 1, variant 15 should take the variant 2.

1. **Put down the matrix of minors for the matrix *A*.**
2. **Evaluate the determinants of the matrices *A* and *B*. You should use the method of row or column expansion (choose any row or column you like) and the triangle method.**

1.5. Inverse matrices

Before coming to the new theme of inverse matrices let us introduce the term of matrix transposition. A matrix transposition is a process when we change the rows and columns in the matrix: the first row becomes the first column (and vice versa), the second row becomes the second column, etc. For example, we take the matrix A below and transpose it; we denote the transposed matrix as A^T.

$$A = \begin{pmatrix} 4 & 7 & -2 \\ 3 & -5 & 8 \\ 1 & 10 & 25 \end{pmatrix}, \qquad A^T = \begin{pmatrix} 4 & 3 & 1 \\ 7 & -5 & 10 \\ -2 & 8 & 25 \end{pmatrix}$$

The inverse matrix of a given matrix A is a matrix A^{-1} which satisfies the following formula:

$$A \cdot A^{-1} = I \text{ (} I \text{ is an identity matrix)}$$

The inverse matrix is denoted as A^{-1} and the formula for its calculation is given below:

$$A^{-1} = \frac{1}{|A|}(A^*)^T$$

where A^* is a matrix of minors (see the paragraph 1.4 and the algorithm of its calculation).

Remark **We denote the inverse matrix as A^{-1} but in spite of the denotation please don't think that we get it after raising all the entries of a matrix A to the power of (–1)! Now you should get acquainted with the real formula of the inverse matrix – so don't make mistakes!**

There are 4 important properties for the inverse matrices:

1) $I^{-1}=I$; **2)** $(A^{-1})^{-1}=A$; **3)** $(A^{\mathrm{T}})^{-1}=(A^{-1})^{\mathrm{T}}$; **4)** $(AB)^{-1}=B^{-1}A^{-1}$

All those properties are rather easy to understand and to remember. The 4-th property is a little bit 'tricky' – however, let us pretend that the direct action is similar to the process of dressing up. The inverse action is taking the clothes off. So let *A* be a T-shirt and *B* be a coat: for the left-hand part we 'dress up': put on firstly a T-shirt *A* and then a coat B. For the inverse action (in the right-hand part) we have to 'undress: firstly, we take off the coat *B* and secondly, we take off the T-shirt *A*.

So let us investigate the matrix *A* described above:

$$A = \begin{pmatrix} 4 & 7 & -2 \\ 3 & -5 & 8 \\ 1 & 10 & 25 \end{pmatrix}$$

The step-by-step algorithm of the inverse matrix calculation is the following:

Step 1. We calculate the determinant of the given matrix. It should always be the first step because if the determinant of a matrix equals zero, the inverse matrix for the given one doesn't exist!

Let us calculate the determinant for the given matrix *A* using the method of first row expansion:

$$|A| = \begin{vmatrix} 4 & 7 & 2 \\ 3 & -5 & 8 \\ 1 & 10 & 25 \end{vmatrix} = 4\cdot\begin{vmatrix} -5 & 8 \\ 10 & 25 \end{vmatrix} - 7\cdot\begin{vmatrix} 3 & 8 \\ 1 & 25 \end{vmatrix} - 2\cdot\begin{vmatrix} 3 & -5 \\ 1 & 10 \end{vmatrix} =$$

$$= -820 - 469 - 70 = -1359$$

Step 2. We put down the matrix of cofactor for the given matrix A and evaluate those minors.

$$A^* = \begin{pmatrix} \begin{vmatrix} -5 & 8 \\ 10 & 25 \end{vmatrix} & -\begin{vmatrix} 3 & 8 \\ 1 & 25 \end{vmatrix} & \begin{vmatrix} 3 & -5 \\ 1 & 10 \end{vmatrix} \\ -\begin{vmatrix} 7 & 2 \\ 10 & 25 \end{vmatrix} & \begin{vmatrix} 4 & 2 \\ 1 & 25 \end{vmatrix} & -\begin{vmatrix} 4 & 7 \\ 1 & 10 \end{vmatrix} \\ \begin{vmatrix} 7 & 2 \\ -5 & 8 \end{vmatrix} & -\begin{vmatrix} 4 & 2 \\ 3 & 8 \end{vmatrix} & \begin{vmatrix} 4 & 7 \\ 3 & -5 \end{vmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} 205 & -67 & 35 \\ -155 & 98 & -33 \\ 66 & -26 & -41 \end{pmatrix}$$

Step 3. We transpose the matrix A^*: we change the rows and columns of this matrix.

$$(A^*)^T = \begin{pmatrix} 205 & -67 & 35 \\ -155 & 98 & -33 \\ 66 & -26 & -41 \end{pmatrix}^T = \begin{pmatrix} 205 & -155 & 66 \\ -67 & 98 & -26 \\ 35 & -33 & -41 \end{pmatrix}$$

Step 4. We calculate the inverse matrix using the formula:

$$A^{-1} = \frac{1}{|A|}(A^*)^T = \frac{1}{-1219}\begin{pmatrix} 205 & -155 & 66 \\ -67 & 98 & -26 \\ 35 & -33 & -41 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{205}{-1359} & \frac{155}{1359} & \frac{66}{-1359} \\ \frac{67}{1359} & \frac{98}{-1359} & \frac{26}{1359} \\ \frac{35}{-1359} & \frac{33}{1359} & \frac{41}{1359} \end{pmatrix}$$

Remark The numbers in the inverse matrix are not so 'beautiful' as we used to see in the practical tasks – however, for the inverse matrices it is a common practice. Don't be afraid of the big numbers and enjoy the process!

One-minute task Does the matrix *D* have the inverse matrix?

$$D = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Practice Russian

Транспонирование ма́трицы– transposition of a matrix

Транспонированная ма́трица – transposed matrix

Обра́тная ма́трица – the inverse matrix

Алгебраическое дополне́ние – cofactor

Review questions

1. Describe the algorithm of the inverse matrix calculation
2. What is the result of the inverse matrix of the product (you need to open the brackets and put down the right product in the answer using the inverse operation): $(AB)^{-1}$ – ?
3. What is an inverse matrix to an identity matrix?

Practical task

All the matrices for this task are already given above in the practical task 1 after the paragraph 1.2. However, this time you need to take your variant as your own one plus 2: for example, if you have variant 1, you need to take variant 3 and so on; variant 14 should take the variant 1, variant 15 should take the variant 2.

1. **Put down the inverse matrix for the matrix *B*.**

1.6. Rank of a matrix. The method of bordering minors.

In our life, we often hear the word 'rank' – speaking about the military service or some rating processes. In fact, we need this word to denote the position or status of somebody/something. In linear algebra the term 'rank' is also useful to define something like a 'matrix status', it depends on the dimension of a matrix and the number of linearly independent rows or columns.

The ***rank of a matrix*** is a highest order of non-zero minors of the matrix. We look at the given matrix and evaluate the maximal dimension of the minors we can 'cut' from this matrix, i.e. for the square matrix the maximal rank can (possibly – as we haven't checked it yet!) be equal to the dimension of this matrix.

It's important! If the number of rows inside the matrix is less than the number of columns then the maximal rank cannot be greater than the maximal number of rows.

If the number of columns inside the matrix is less than the number of rows then the maximal rank cannot be greater than the maximal number of columns.

(It is obvious – otherwise we can't get the 'square figure' made of entries and, however, all the minors are square). So for the $m \times n$ matrix we have the following

$$rk\,A \leq min\{m, n\}$$

The only matrix which has a zero rank is a zero matrix (all the entries in it are equal to zero). The rank of other matrices is not less than 1. Any column matrix or row matrix has a rank that equals 1. It's also clear that $rk(A) = rk(A)^T$ as during the transposition of a matrix all the minors are also transposed and so their values don't change.

When we evaluate the rank of a matrix, we should remember some important rules. For example, we have already found the non-zero minor of the r-th order but all the minors of the order $(r + 1)$ are equal to zero. Then the rank of this matrix equals r. Indeed, it is easy to learn that if we look for the minors of that matrix with the order higher than $(r + 1)$, they all will also be equal to zero. If we look at any minor of the order $(r + 2)$ and if we use, for example, the row expansion method (e.g. we take the first row – however, it can be any row or column of the matrix), we see that it's equal to zero as the $(r + 1)$ cofactor also equals zero. So this rule also works for the minors of the higher order.

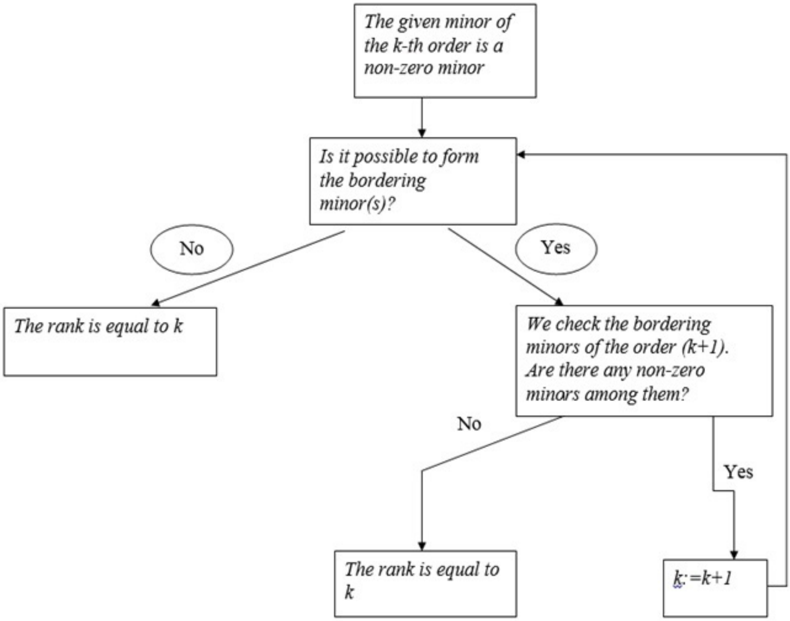
So we 'climb the stairs' from the minors of the lower order to the minors of the higher order: if we have already found the non-zero minor of the k-th order then we go to the evaluation of the minor of $(k + 1)$-th order (if they exist). If all of those minors are equal to zero then the rank equals k. Otherwise, the procedure goes on and we investigate the minors of the order $(k + 2)$ and so on. Using this method on practice for the minors of the order higher than 4, is rather time-taking and difficult. So they use the method of bordering minors.

A minor ***<u>borders</u>*** the given minor if the bordering minor has an order which is greater by one than the given minor and the bordering minor contains the rows and columns of the given one.

The main rule of the bordering minors' method is that during the rank evaluation procedure it is enough to calculate only the bordering minors on each step of the process. So if all the bordering minors are equal to zero then all the minors of that order will be equal to zero.

For example, if the given matrix has the dimension 4×4 and its rank equals 2. In order to get that result of the rank evaluation, at first we need to find a non-zero minor of the 2-nd order . Then we need to evaluate the minors of the third order. We have 16 minors of that order inside the whole matrix and only 4 bordering minors of that same order! The difference is great!

Let us introduce the scheme of the bordering minors' method:



<u>Example</u>. Find the rank of a matrix using the bordering minors' method

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

1) We need to find a non-zero 1×1 entry, i.e., any non-zero element of a matrix. Let us choose the first element in the first row, it's equal to 1. So we can tell $rk(A) \geq 1$.

2) We need to find a non-zero minor of a dimension 2×2 that includes the entry we have already investigated above. So the only 2×2 minor which includes that element is a minor of the order 2 in the upper left corner (it is marked in the matrix below)

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

Let us denote this minor as M_1. We evaluate: $M_1 = \begin{vmatrix} 1 & -3 \\ 0 & 1 \end{vmatrix} = 1$. So we have 2×2 non-zero bordering minor and we can tell $rk(A) \geq 2$.

3) We need to find a non-zero minor of a dimension 3×3 that includes the entry we have already investigated above. Let it also be the 3×3 minor in the left corner of a matrix and let us denote it as M_2.

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

So we have to calculate:

$$M_2 = \begin{vmatrix} 1 & -3 & -3 \\ 0 & 1 & 1 \\ -4 & 0 & 0 \end{vmatrix} = 1 \cdot 0 + 3 \cdot 4 - 3 \cdot 4 = 0$$

We see that we have a zero minor. However, we can check another bordering minor M_3, its entries are marked in the matrix below.

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

Thus, we can see that the bordering minor may not include the columns or rows that are standing next to the minor of the previous order, we can 'jump' over columns and rows but the necessary condition is that the previous minor should be inside the new one and it should not be broken into any parts.

$$M_3 = \begin{vmatrix} 1 & -3 & 1 \\ 0 & 1 & 3 \\ -4 & 0 & 7 \end{vmatrix} = 1 \cdot 7 + 3 \cdot 12 + 1 \cdot 4 = 47 \neq 0$$

So we can conclude that $rk\ A = 3$. It can't be more than 3 as we can't form the bordering minors around M_3 due to the dimension of the given matrix.

Practice Russian

Окаймляющий минóр – bordering minor

Ранг мáтрицы – rank of a matrix

Review questions

1. How do you understand the term 'a bordering minor'?
2. Describe the algorithm of finding a rank of a matrix

Practical task

1. Find the rank of a matrix using the minor method

$1.\ A = \begin{pmatrix} 3 & -4 & 9 & -5 \\ 1 & 0 & 7 & 8 \\ -6 & 2 & 3 & 0 \end{pmatrix}$
$2.\ A = \begin{pmatrix} -2 & 1 & 3 & 7 \\ 3 & -4 & 5 & -3 \\ 9 & -1 & 6 & 0 \end{pmatrix}$
$3.\ A = \begin{pmatrix} 6 & 8 & 7 & 1 \\ 9 & -3 & -4 & 2 \\ 0 & -2 & 3 & 5 \end{pmatrix}$
$4.\ A = \begin{pmatrix} 1 & 7 & 0 & 9 \\ 1 & 3 & 5 & 8 \\ 2 & 5 & -9 & 6 \end{pmatrix}$
$5.\ A = \begin{pmatrix} 1 & -5 & 4 & 6 \\ -1 & 7 & 9 & -8 \\ -4 & 0 & 3 & -2 \end{pmatrix}$
$6.\ A = \begin{pmatrix} 8 & -2 & 4 & 5 \\ 3 & -1 & 5 & -3 \\ -1 & -4 & 0 & -1 \end{pmatrix}$
$7.\ A = \begin{pmatrix} 3 & -1 & -1 & 9 \\ -3 & 2 & 1 & -3 \\ 5 & 9 & 8 & 2 \end{pmatrix}$

8. $A=\begin{pmatrix}-5 & 10 & 1 & 0\\ 7 & 2 & 8 & 3\\ 5 & 0 & 9 & 4\end{pmatrix}$
9. $A=\begin{pmatrix}3 & 4 & 7 & 8\\ 7 & 8 & -3 & -1\\ 1 & 5 & -4 & 2\end{pmatrix}$
10. $A=\begin{pmatrix}1 & 4 & -3 & 3\\ -3 & 1 & 6 & -4\\ 8 & 1 & -2 & 9\end{pmatrix}$
11. $A=\begin{pmatrix}5 & 8 & -2 & 7\\ -5 & 6 & 1 & -5\\ 0 & 2 & -3 & 2\end{pmatrix}$
12. $A=\begin{pmatrix}7 & -1 & 2 & 4\\ 4 & -5 & -3 & -1\\ 1 & 0 & -2 & 7\end{pmatrix}$
13. $A=\begin{pmatrix}8 & 9 & -5 & 3\\ -7 & 4 & 1 & 0\\ 2 & -10 & 5 & -2\end{pmatrix}$
14. $A=\begin{pmatrix}0 & -8 & 5 & 9\\ -7 & 5 & 10 & 1\\ -3 & 6 & -1 & -2\end{pmatrix}$
15. $A=\begin{pmatrix}-3 & 7 & 2 & -1\\ 2 & 4 & -9 & 0\\ -9 & 5 & 6 & 1\end{pmatrix}$

1.7. Gaussian method: finding the rank of a matrix and calculation of the inverse matrix

‘Gaussian method’, ‘Gaussian elimination’ – all those terms are named in honor of a prominent mathematician Johann Carl Friedrich Gauss (see p. 3 and the biography below in this chapter). However, before the introduction of Gaussian method let us explain the rules of elementary matrix transformation.

Elementary transformations inside the matrix include the following operations:

a) Rearrangement of two rows (columns) inside a matrix;

b) Multiplication of any row (column) of a matrix to some non-zero number;

c) Addition of any row (column) of the given matrix (it may be already multiplied by any non-zero number) to any other row or column in the matrix

If we get a matrix *B* as a result of the transformations inside the matrix A , then we tell that matrix *A* is equivalent to matrix *B* and so we put down $A \sim B$.

Remark **All three elementary transformations can be invertible: if we fulfil any of the transformations described above, we can return to the 'starting point' by making the other elementary transformations:**

For the operation **a):** to fulfil the inverse operation we need to return the rows (columns) into their initial places;

For the operation **b):** to fulfil the inverse operation we should multiply the transformed row (column) by the inverse number;

For the operation **c):** to fulfil the inverse operation we should take the row (column) that was earlier multiplied by a non-zero number; then we multiply this row(column) by the inverse number and we should add it to the row (column) which was transformed before that inverse process.

It's important! **The elementary transformations don't change the rank of a matrix.**

So when we have already learnt the phenomena of elementary transformation, let us introduce the algorithm of finding the inverse matrix using a Gaussian method:

Step 1. We put down the given matrix and to this matrix we join the identity matrix of the same dimension to the right of the given one. For example, if we have 3×3 given matrix then we put down the 3×3 identity matrix to the right of it. Usually we divide the given matrix from the identity matrix by a straight borderline:

$$\left(\begin{array}{ccc|ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array}\right)$$

Step 2. Using the elementary transformations, we should get the identity matrix in the left part (instead of the given matrix). Then in the right part we automatically should get the matrix which is inverse to the initial one. Below we denote the entries of the inverse matrix as a'_{ij}.

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & a'_{11} & a'_{12} & a'_{13} \\ 0 & 1 & 0 & a'_{21} & a'_{22} & a'_{23} \\ 0 & 0 & 1 & a'_{31} & a'_{32} & a'_{33} \end{array}\right)$$

Let us demonstrate a simple task. We need to calculate the inverse matrix to the given matrix *A*.

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

So we put down the matrix A and join the identity matrix as a part of a new common matrix

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -2 & -1 & 0 & 0 & 1 \end{array}\right)$$

We need to get an identity matrix in the left-hand part. So firstly we need to have a matrix of a so-called row echelon form – a form got by applying Gaussian elimination to a matrix. Let us introduce the definition of a leading coefficient.

A leading coefficient is a first non-zero entry of a matrix.

So for the echelon form of a matrix we have:

– All the rows that consist only of zeros, are situated on the bottom of the matrix.

– The leading coefficient (also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it (all of them are situated on the main diagonal).

Remark **Sometimes there is an additional condition that the leading coefficient should be 1, and there also exists a *reduced row echelon form* where:**

– **A matrix is already lead to a row echelon form;**

– **Each non-zero pivot in each row is 1**

– **The other elements in the row, except the pivots, are zeros).**

For our problem we should use the reduced echelon form as we need to get the identity matrix in the left-hand part. However, a row echelon form is a first step on this way.

So the echelon form looks like a 'ladder of zeros' under the main diagonal

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

Let's see how the elementary transformations help us to achieve our goal.

First of all, we have to make a zero instead of a first entry in a second row. We look at the whole matrix and we notice that the first row and the second row have '1' as the first element in the row – so if we subtract the first row from the second row, we would get zero instead of a first entry.

Let us put it down.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -2 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{II}-\text{I}\to\text{II}} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 1 & -2 & -1 & 0 & 0 & 1 \end{array}\right)$$

For a sign of elementary transformation we can use an arrow or a ' ~' sign. Above the sign we put down the information about the elementary transformations done on this step (in the transformation above we marked that we subtracted row I from row II and put the result into the second row). The other important moment is that in the result of the transformation we changed only the second row, and the first row is still the same. And, at last, we see that also we subtract the first row from the second row in our augmented matrix! That is the key peculiarity of the method that helps us to get the right result.

Let us look at the next step of our transformation. We need to get zero instead of a first entry in a third row. We see that the first row and the third row also have '1' as the first element in the row – so if we subtract the first row from the third row, we would get zero instead of a first entry. So we take the matrix got on the previous step and make another elementary operation:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 1 & -2 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{III}-\text{I}\to\text{III}} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 0 & -2 & -3 & -1 & 0 & 1 \end{array}\right)$$

Now we should put a zero instead of the second entry in a third row. We see that in the second row we have (-1) as the second entry. So if we multiply the second row by (-2) and add it to the third row, we would get the necessary result.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 1 & -2 & -3 & -1 & 0 & 1 \end{array}\right) \xrightarrow{\text{III+II}\cdot(-2)\to\text{III}} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right)$$

Now we can easily get ‘1’ on a main diagonal in the second row. We should just multiply it by ‘–1’.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right) \xrightarrow{\text{II}\cdot(-1)\to\text{II}} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right)$$

We need to get zeros everywhere except the main diagonal in the left-hand part of a matrix. Luckily, one zero in the first row was already given in the task as the second entry. So we have to get zeros instead of both numbers‘2’ – the third entries in the first and second rows. However, we need to save the results that we have already made on our way to the identity matrix. So the ideal row as a material for our transformations is the third row, where the leading coefficient in the left-hand part is 1 – so it is easy for us to multiply it by any number, and if we add this transformed row to any other row, the zeros won’t ‘spoil’ the result of addition. So we need to multiply the third row by (-2) and add it to the second row and to the first row.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right) \xrightarrow[\text{III}\cdot(-2)+\text{II}\to\text{II}]{\text{III}\cdot(-2)+\text{I}\to\text{I}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & -2 \\ 0 & 1 & 0 & -1 & -1 & -2 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{array}\right)$$

So we got the identity matrix in the left part of the matrix and now in the right-hand part we automatically got the inverse matrix.

The result is:

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

A task for everyone! Check the result by using the other method of the inverse matrix calculation.

Now let's see the peculiarities of finding the rank of a matrix. When we get the echelon form of the matrix, the formula looks like:

$$rkA = n - k$$

where n is the total number of rows in a matrix, k is a number of zero rows in a matrix (a zero row is a row containing only zeros; this row doesn't have any non-zero numbers).

For example, we have a matrix:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 3 & 6 & 9 \end{pmatrix}$$

We have one zero standing in the 'proper position' for the echelon form: it is the first entry in the second row. Now we need to make a zero instead of the first entry in the third row. We multiply the first row by (–3) and add it to the third row. So we get

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 3 & 6 & 9 \end{pmatrix} \xrightarrow{II\cdot(-1)\to II} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 0 \end{pmatrix}$$

We got one zero row. We came to an echelon form so

$$rkA = 3 - 1 = 2$$

When we put down a matrix with zero rows, further we can just remove this row, so the matrix will look like:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \end{pmatrix}$$

Of course, if all the elements of one row multiplied by the same number, give us the corresponding elements of the other row, we always get a zero row as a result of elementary transformations.

A linear combination of rows $r_1, r_{2, \ldots}, r_n$ in a matrix looks like:

$$\alpha_1 r_1 + \alpha_2 r_2 + \ldots + \alpha_n r_n$$

A linear combination of rows is ***trivial*** if all coefficients α_i are equal to zero at the same time. A trivial linear combination looks like a zero row (which can exist in the given matrix or can be got as the elementary transformations between the matrix rows).

A linear combination of rows is ***non-trivial*** if at least one of the coefficients α_i is not equal to zero.

The system of rows is ***linearly dependent***, if there is a non-trivial linear combination of rows, and this combination is equal to the zero row.

The system of rows is ***linearly independent***, if only trivial linear combination of rows is equal to the zero row.

There is a simple criterion to check whether the matrix has linearly dependent rows.

The rows in a matrix are linearly independent if the determinant of a matrix is not equal to zero. The rows of a matrix are linearly dependent if the determinant equals zero.

Johann Carl Friedrich Gauss

(1777–1855)

Inspiring biography *Johann Carl Friedrich Gauss* is one of the most famous German mathematicians. He was born in a poor family; his mother was illiterate and only could tell that her son was born eight days before the Feast of the Ascension (which is after 39 days after Easter). In 1799 Gauss calculated his precise date of birth by creating an algorithm that allowed to calculate the date of Easter for any year. However, at the age of three Johann Carl Friedrich corrected a mathematical mistake that his father had made, and at school he solved a problem of arithmetical series faster than 100 of his classmates using an interesting and rational method invented by himself. Gauss's talent impressed the Duke of Brunswick, who sent him to the Collegium Carolinum (now Braunschweig University of Technology), which he attended from 1792 to 1795, and to the University of Göttingen (from 1795 to 1798). The first famous discovery of Gauss occurred in 1796 when he created a ruler-and-compass construction for the solution of a regular polygon. It was a problem that had not been sold since its formulation in the times of Ancient Greece. During his life he made plenty of wonderful discoveries in the field of number theory, developed greatly the theory of complex numbers, laid the foundation of the linear algebra and invented the method of solving the systems of linear equations. He proved the Fundamental Theorem of Algebra, had researches in the theory of probability and statistics, differential geometry. He was interested in astronomy too and he held the post of Director of the astronomical observatory in Göttingen for many years. When the planetoid Ceres was in the process of being identified in the late 17th Century, Gauss made a prediction of its position which varied greatly from the predictions of most other astronomers of the time. But, when Ceres was finally discovered in 1801, it was almost exactly where Gauss had predicted!

<u>Practice Russian</u>

Ме́тод Га́усса – Gaussian method

Преобразова́ние Га́усса – Gaussian elimination

Элемента́рные преобразова́ния – elementary transformations

<u>Review questions</u>

1. What is an echelon form?
2. What is a row reduced echelon form?
3. How do we calculate the inverse matrix and how do we evaluate a rank of a matrix using the Gaussian method?

<u>Practical task</u>

1. Calculate the inverse matrix for a matrix A using the Gaussian method.

<u>Remark!</u> **Don't forget that you can rearrange rows and columns! For example, you can put the first row in a variant 14 instead of the second row (and, of course, the second row will be put instead of the first one) – and you will get a 'ready' zero as the first entry in the second row to get closer to the row echelon form. The third row in the variant 12 starts with the number '1' – you can change its place with the first row and you will get a 'ready' leading coefficient**

for the echelon form. In fact, to get the numbers '1' as the leading coefficients is very convenient: you can multiply this row by any number easily to get zeros in other rows as a result of elementary row transformations.

1. $A = \begin{pmatrix} 3 & -4 & 9 \\ 1 & 0 & 7 \\ -6 & 2 & 3 \end{pmatrix}$
2. $A = \begin{pmatrix} -2 & 1 & 3 \\ 3 & -4 & 5 \\ 9 & -1 & 6 \end{pmatrix}$
3. $A = \begin{pmatrix} 6 & 8 & 7 \\ 9 & -3 & -4 \\ 0 & -2 & 3 \end{pmatrix}$
4. $A = \begin{pmatrix} 1 & 7 & 0 \\ 1 & 3 & 5 \\ 2 & 5 & -9 \end{pmatrix}$
5. $A = \begin{pmatrix} 1 & -5 & 4 \\ -1 & 7 & 9 \\ -4 & 0 & 3 \end{pmatrix}$
6. $A = \begin{pmatrix} 8 & -2 & 4 \\ 3 & -1 & 5 \\ -1 & -4 & 0 \end{pmatrix}$
7. $A = \begin{pmatrix} 3 & -1 & -1 \\ -3 & 2 & 1 \\ 5 & 9 & 8 \end{pmatrix}$
8. $A = \begin{pmatrix} -5 & 10 & 1 \\ 7 & 2 & 8 \\ 5 & 0 & 9 \end{pmatrix}$
9. $A = \begin{pmatrix} 3 & 4 & 7 \\ 7 & 8 & -3 \\ 1 & 5 & -4 \end{pmatrix}$
10. $A = \begin{pmatrix} 1 & 4 & -3 \\ -3 & 1 & 6 \\ 8 & 1 & -2 \end{pmatrix}$
11. $A = \begin{pmatrix} 5 & 8 & -2 \\ -5 & 6 & 1 \\ 0 & 2 & -3 \end{pmatrix}$

12. $A = \begin{pmatrix} 7 & -1 & 2 \\ 4 & -5 & -3 \\ 1 & 0 & -2 \end{pmatrix}$
13. $A = \begin{pmatrix} 8 & 9 & -5 \\ -7 & 4 & 1 \\ 2 & -10 & 5 \end{pmatrix}$
14. $A = \begin{pmatrix} 0 & -8 & 5 \\ -7 & 5 & 10 \\ -3 & 6 & -1 \end{pmatrix}$
15. $A = \begin{pmatrix} -3 & 7 & 2 \\ 2 & 4 & -9 \\ -9 & 5 & 6 \end{pmatrix}$

1.8. Cramer's rule for systems of linear equations

Let us introduce the rule for solving the square systems of linear equations (a square system means that the number of variables in the system is equal to the number of equations in this system). It was invented by Gabriel Cramer, a Genevan mathematician who began his investigations of linear algebra long before the discoveries of Gauss (see the biography of G. Cramer at the end of this chapter).

If we have a square system of linear equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = c \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = d \end{cases}$$

Here $a_{11}, a_{12} \ldots a_{33}$ are some coefficients standing in front of the variables x_1, x_2, x_3 and b, c, d are just some numbers standing in the right-hand part.

The algorithm of the Cramer's rule is very simple:

Step 1. We put down ***the determinant of the system*** – the determinant which consists only from the coefficients standing in the left-hand part in front of the variables (for all the equations of the system). Usually we use a Greek letter Δ for the denotation of this determinant.

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \neq 0$$

Then we calculate the value of this determinant. We should notice that the determinant of the system should not be equal to zero. If we have a zero value then the Cramer's rule won't work.

Step 2. We compose a determinant Δ_1 – to use it further for the variable x_1. We take the determinant Δ and for the new determinant Δ_1 we remove the first column of Δ (the column of coefficients in front of x_1 in all the equations of the system) and put the column of the three right-hand part values instead:

$$\begin{pmatrix} b \\ c \\ d \end{pmatrix}$$

So the determinant Δ_1 looks like:

$$\Delta_1 = \begin{vmatrix} b & a_{12} & a_{13} \\ c & a_{22} & a_{23} \\ d & a_{32} & a_{33} \end{vmatrix}$$

Step 3. We make the similar operations and compose the determinants

Δ_2 and Δ_3. For the determinant Δ_2 we put the column of the right-hand parts instead of the second column of Δ (the column of coefficients in front of x_2). For the determinant Δ_3 we put the column of the right-hand parts instead of the third column of Δ (the column of coefficients in front of x_3). So we have:

$$\Delta_2 = \begin{vmatrix} a_{11} & b & a_{13} \\ a_{21} & c & a_{23} \\ a_{31} & d & a_{33} \end{vmatrix}$$

$$\Delta_3 = \begin{vmatrix} a_{11} & a_{12} & b \\ a_{21} & a_{22} & c \\ a_{31} & a_{32} & d \end{vmatrix}$$

Step 4. To learn the values of the variables x_1, x_2, x_3 we divide the determinants Δ_n (for our case n = 1, 2, 3) with corresponding indexes by the determinant of the system Δ:

$$x_1 = \frac{\Delta_1}{\Delta}; x_2 = \frac{\Delta_2}{\Delta}; x_3 = \frac{\Delta_3}{\Delta}$$

So we have got the answers of the given system.

Remark **You can use different denotations of the variables – for example, instead of x_1, x_2, x_3 you can put down x,y,z. Then it is better to denote the determinants with the indexes of the corresponding variables Δ_x, Δ_y, Δ_z. It provides better understanding of your calculations.**

Remark **You should have already guessed that if in one of the system's equations we don't have a full set of the variables – for example, for the system with three variables one of the equations looks like: $x_2 + 3x_3 = 4$, that means we have a zero coefficient in front of x_1. So we put a zero to its proper place in all the corresponding rows of our determinants.**

The Cramer's rule is rather simple so we suggest you to solve your practical tasks just right now!

Gabriel Cramer (1704–1752)

Inspiring biography *Gabriel Cramer,* a prominent mathematician, was born in the Republic of Geneva and showed real talent for mathematics since his early ages. He got the doctorate at the age of 18 and became a co-chair at the University of Geneva at the age of 20. Later he published a lot of works in the fields of geometry, probability theory, history of mathematics and also created the basement of linear algebra. He even had some published researches in philosophy! In 1750 he invented his famous Cramer's theorem for the algebraic curves and discovered the Cramer's rule described above.

Practice Russian

Пра́вило Кра́мера– Cramer's rule

Систе́ма лине́йных уравне́ний – the system of linear equations

Определитель систе́мы – the determinant of a system

Review questions

1. What are the restrictions for using the Cramer's rule?
2. Describe the algorithm of a Cramer's rule

Practical task

1. Solve the system using the Cramer's rule

1. $\begin{cases} 2x_1 - 4x_2 + 6x_3 = 4 \\ 6x_1 + 2x_2 - 2x_3 = 6 \\ 4x_1 + 10x_2 + 4x_3 = 18 \end{cases}$
2. $\begin{cases} 3x_1 + 6x_3 = 21 \\ 6x_1 - 3x_2 + 3x_3 = 9 \\ 3x_1 + 9x_2 - 3x_3 = 12 \end{cases}$
3. $\begin{cases} 2x_1 + 10x_2 + 4x_3 = 4 \\ 4x_1 + 6x_2 + 4x_3 = -6 \\ 2x_1 + 6x_2 + 8x_3 = -6 \end{cases}$
4. $\begin{cases} 4x_1 + 8x_2 - 12x_3 = 4 \\ 12x_1 + 8x_2 - 16x_3 = 0 \\ 8x_1 - 4x_2 = -4 \end{cases}$
5. $\begin{cases} 8x_1 + 4x_2 - 2x_3 = 0 \\ 2x_1 + 4x_2 + 2x_3 = 2 \\ x_2 - x_3 = -3 \end{cases}$
6. $\begin{cases} 5x_1 + 10x_2 - 5x_3 = 10 \\ 10x_1 - 15x_2 + 10x_3 = 15 \\ 15x_1 + 5x_2 + 5x_3 = 40 \end{cases}$
7. $\begin{cases} 3x_1 + 3x_2 + 3x_3 = 9 \\ 3x_1 + 6x_2 + 9x_3 = 6 \\ 3x_1 + 9x_2 + 18x_3 = -3 \end{cases}$
8. $\begin{cases} 4x_1 + 6x_2 - 2x_3 = 8 \\ 2x_1 + 2x_2 + 6x_3 = 10 \\ 6x_1 - 8x_2 + 2x_3 = 0 \end{cases}$
9. $\begin{cases} 3x_1 - 3x_2 + 6x_3 = -6 \\ 6x_1 - 6x_2 + 12x_3 = 12 \\ 9x_1 - 9x_2 + 18x_3 = 9 \end{cases}$
10. $\begin{cases} 8x_1 - 4x_2 - 12x_3 = 12 \\ 12x_1 + 16x_2 - 20x_3 = -32 \\ 32x_2 + 28x_3 = 68 \end{cases}$

11.	$\begin{cases} 4x_1 + 8x_2 - 2x_3 = 8 \\ -2x_1 + 6x_2 + 4x_3 = 14 \\ 6x_1 + 4x_2 - 4x_3 = -12 \end{cases}$
12.	$\begin{cases} 3x_1 + 15x_2 - 3x_3 = 0 \\ 6x_1 - 3x_2 + 3x_3 = 9 \\ 3x_1 + 6x_2 - 9x_3 = -6 \end{cases}$
13.	$\begin{cases} 12x_1 - 4x_2 - 4x_3 = -20 \\ 4x_1 + 12x_2 + 8x_3 = 8 \\ 20x_1 - 8x_2 + 16x_3 = -28 \end{cases}$
14.	$\begin{cases} 9x_1 - 6x_2 + 3x_3 = -30 \\ 6x_1 + 9x_2 - 12x_3 = 48 \\ 3x_1 - 12x_2 + 9x_3 = -54 \end{cases}$
15.	$\begin{cases} 15x_1 + 10x_2 + 5x_3 = 25 \\ 10x_1 + 15x_2 + 5x_3 = 5 \\ 10x_1 + 5x_2 + 15x_3 = 55 \end{cases}$

1.9. Gaussian elimination for systems of linear equations

We have already got acquainted with some key principles of Gaussian method based on the elementary row/column transformations of the matrix. Now we come to the next aspect of its application – it is extremely useful for solving the systems of linear equations.

If we have a system of linear equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = c \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = d \end{cases}$$

we can transform this system into the following matrix where we put down the coefficients standing in front of the variables and the numbers in the right-hand part. Our aim is to make the non-zero numbers only on the main diagonal using the Gaussian elimination – that will help us in finding the variables. Let us see how it works. Firstly, we form a matrix itself with the column of right-hand part numbers standing a little bit apart from the coefficients. Such a matrix made of the coefficient matrix and the right-hand part column, is called ***an augmented matrix***:

$$\left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b \\ a_{21} & a_{22} & a_{23} & c \\ a_{31} & a_{32} & a_{33} & d \end{array}\right)$$

Remark **Here we have the symbols *b, c, d* standing in the right-hand part. If we had symbols b_1, b_2, b_3 for the right-hand parts in each equation correspondently, we could put down the system in a matrix form $AX = B$ where A is the matrix of coefficients, X is a column of variables, B is a column of the right-hand parts. The augmented matrix in this case can be denoted as $(A|B)$.**

Then we need to get the following result with the help of the elementary row transformations:

$$\left(\begin{array}{ccc|c} a_{11}^* & 0 & 0 & e \\ 0 & a_{22}^* & 0 & f \\ 0 & 0 & a_{33}^* & g \end{array}\right),$$

where $a_{11}^*, a_{22}^*, a_{33}^*$ are the new coefficients that we have just got in the result of elementary transformations. And we also got the new numbers for the right-hand part of the system because of that same transformations.

So the resulting system will look like:

$$\begin{cases} a_{11}^* x_1 = e \\ a_{22}^* x_2 = f \\ a_{33}^* x_3 = g \end{cases}$$

We can also lead the resulting matrix to the row reduced echelon form:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & h \\ 0 & 1 & 0 & j \\ 0 & 0 & 1 & k \end{array}\right)$$

So we can have a simple system with the answers:

$$\begin{cases} x_1 = h \\ x_2 = j \\ x_3 = k \end{cases}$$

Example Solve the system of equations using the Gaussian method

$$\begin{cases} 7x_1 + 4x_2 - 8x_3 = 3 \\ 3x_1 - 2x_2 + 5x_3 = 7 \\ 5x_1 - 3x_2 - 4x_3 = -12 \end{cases}$$

We have already got acquainted with the rules of elementary row transformations so we won't put down in words any comments on the transformations we make on each step. However, we will put the transformations down schematically just right in the expressions below as we have already done before.

$$\left(\begin{array}{ccc|c} 7 & 4 & -8 & 3 \\ 3 & -2 & 5 & 7 \\ 5 & -3 & -4 & -12 \end{array}\right) \xrightarrow{\text{I}\cdot\left(-\frac{3}{7}\right)+\text{II}\to\text{II}} \left(\begin{array}{ccc|c} 7 & 4 & -8 & 3 \\ 0 & -\frac{26}{7} & \frac{59}{7} & \frac{40}{7} \\ 5 & -3 & -4 & -12 \end{array}\right) \xrightarrow{\text{I}\cdot\left(-\frac{5}{7}\right)+\text{III}\to\text{III}}$$

$$\xrightarrow{\text{I}\cdot\left(-\frac{5}{7}\right)+\text{III}\to\text{III}} \left(\begin{array}{ccc|c} 7 & 4 & -8 & 3 \\ 0 & -\frac{26}{7} & \frac{59}{7} & \frac{40}{7} \\ 0 & -\frac{41}{7} & \frac{12}{7} & -\frac{99}{7} \end{array}\right) \xrightarrow{\text{II}\cdot\left(-\frac{41}{26}\right)+\text{III}\to\text{III}}$$

$$\xrightarrow{\text{II}\cdot\left(-\frac{41}{26}\right)+\text{III}\to\text{III}} \left(\begin{array}{ccc|c} 7 & 4 & -8 & 3 \\ 0 & -\frac{26}{7} & \frac{59}{7} & \frac{40}{7} \\ 0 & 0 & -\frac{301}{26} & -\frac{301}{13} \end{array}\right) \xrightarrow[\text{III}:\left(-\frac{301}{26}\right)\to\text{III}]{\text{I}:(7)\to\text{I};\ \text{II}:\left(-\frac{26}{7}\right)\to\text{II}}$$

$$\left(\begin{array}{ccc|c} 1 & \frac{4}{7} & -\frac{8}{7} & \frac{3}{7} \\ 0 & 1 & -\frac{59}{26} & -\frac{20}{13} \\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow{\text{II}\cdot+\text{III}\cdot(\frac{59}{26})\to\text{II}}$$

$$\xrightarrow{\text{II}\cdot+\text{III}\cdot(\frac{59}{26})\to\text{II}} \left(\begin{array}{ccc|c} 1 & \frac{4}{7} & -\frac{8}{7} & \frac{3}{7} \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow{\text{I}+\text{II}\cdot(-\frac{4}{7})\to\text{I}} \left(\begin{array}{ccc|c} 1 & 0 & -\frac{8}{7} & -\frac{9}{7} \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow{\text{I}+\text{III}\cdot(\frac{8}{7})\to\text{I}}$$

$$\xrightarrow{\text{I}+\text{III}\cdot(\frac{8}{7})\to\text{I}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

So we get the resulting system:

$$\begin{cases} x_1 = 1 \\ x_2 = 3 \\ x_3 = 2 \end{cases}$$

Remark **In fact, it is not necessary to get '1' as the leading coefficients and zeros everywhere in the part of the coefficients except the pivots. You can stop earlier. For example, you can stop and make a new system as soon as you get the 'ladder of zeros' under the main diagonal (the echelon form):**

$$\begin{cases} \mathbf{7x_1 + 4x_2 - 8x_3 = 3} \\ \mathbf{-\frac{26}{7}x_2 + \frac{59}{7}x_3 = \frac{40}{7}} \\ \mathbf{-\frac{301}{26}x_3 = -\frac{301}{13}} \end{cases}$$

We investigated the system of linear equations with three variables and three equations in the system; however, this method is universal for any number of variables and equations inside one system.

Practice Russian

Систе́ма лине́йных уравне́ний – system of linear equations

Review questions

1. What is an augmented matrix?
2. Describe the algorithm of solving the system using the Gaussian method

Practical task

1. Solve the systems of the previous chapter (Cramer's rule) using the Gaussian method.

1.10. Linear dependence and independence of vectors. Bases.

We have already got acquainted with the vectors and their role in the matrix theory (Just let's come back to the very beginning and look at the row vectors and column vectors). Speaking about linear dependence and independence of vectors, we should firstly mention the term of a ***vector space.***

Let us introduce the vectors $b_1, b_2, \ldots b_n$ and some real numbers $\lambda_1, \lambda_2, \ldots \lambda_n$.

The vectors $b_1, b_2, \dots b_n$, are called ***linearly dependent*** if the linear combination of vectors $b_1, b_2, \dots b_n$ with the real numbers $\lambda_1, \lambda_2, \dots \lambda_n$ equals zero when at least one number among $\lambda_1, \lambda_2, \dots \lambda_n$ is a non-zero number:

$$\lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_n b_n = 0.$$

The vectors $b_1, b_2, \dots b_n$, are called ***linearly independent*** if the linear combination of vectors $b_1, b_2, \dots b_n$ and the real numbers $\lambda_1, \lambda_2, \dots \lambda_n$ equals zero only in the case when all the numbers $\lambda_1, \lambda_2, \dots \lambda_n$ are equal to zero:

$$\lambda_1 b_1 + \lambda_2 b_2 + \dots + \lambda_n b_n = 0.$$

There is another important criterion that helps us to check the linear (in)dependence of vectors. Let us explain at the example of 3-dimensional space.

We have an expression:

$$\begin{pmatrix} \lambda_{11}^1 \\ \lambda_{21}^1 \\ \lambda_{31}^1 \end{pmatrix} b_1 + \begin{pmatrix} \lambda_{11}^2 \\ \lambda_{21}^2 \\ \lambda_{31}^2 \end{pmatrix} b_2 + \begin{pmatrix} \lambda_{11}^3 \\ \lambda_{21}^3 \\ \lambda_{31}^3 \end{pmatrix} b_3 = 0$$

So we can put down a system of linear equations with variables b_1, b_2, b_3:

$$\begin{cases} \lambda_{11}^1 b_1 + \lambda_{11}^2 b_2 + \lambda_{11}^3 b_3 = 0 \\ \lambda_{21}^1 b_1 + \lambda_{21}^2 b_2 + \lambda_{21}^3 b_3 = 0 \\ \lambda_{31}^1 b_1 + \lambda_{31}^2 b_2 + \lambda_{31}^3 b_3 = 0 \end{cases}$$

Then we form a matrix C. The first row consists of the coefficients in front of a variable b_1 in all three equations of a system, i.e. it is a first column of coefficients; the second row is a row of coefficients in front of b_2 and the same rule for coefficients in front of b_3.

After that we look at the determinant of a matrix C. If it is not equal to zero, then all the vectors are linearly independent. If it is equal to zero, then there are linearly dependent vectors.

$$C = \begin{pmatrix} \lambda_{11}^1 & \lambda_{21}^1 & \lambda_{31}^1 \\ \lambda_{11}^2 & \lambda_{21}^2 & \lambda_{31}^2 \\ \lambda_{11}^3 & \lambda_{21}^3 & \lambda_{31}^3 \end{pmatrix}, \ |C| = \begin{vmatrix} \lambda_{11}^1 & \lambda_{21}^1 & \lambda_{31}^1 \\ \lambda_{11}^2 & \lambda_{21}^2 & \lambda_{31}^2 \\ \lambda_{11}^3 & \lambda_{21}^3 & \lambda_{31}^3 \end{vmatrix}$$

We can also check whether this set of vectors is linearly independent if we lead the matrix C to the row reduced echelon form – if we succeed and get no zero rows, then the vectors are linearly independent – and they're linearly dependent otherwise. Below you can see the example of a row reduced echelon form that shows us the linear independence of given vectors.

$$C = \begin{pmatrix} \lambda_{11}^1 & 0 & 0 \\ 0 & \lambda_{21}^2 & 0 \\ 0 & 0 & \lambda_{31}^3 \end{pmatrix}$$

<u>Remark</u> **There are some cases which help us to find out that the system of vectors is definitely linearly dependent:**

1. **One of the vectors in the given system is a zero vector**
2. **Two or more vectors in this system are equal to each other or the system contains proportional vectors: $b_1 = \alpha b_2$**
3. **When a vector is an obvious linear combination of some other vectors**
4. **If the space where a set of vectors is defined, is *n*-dimensional, then all the elements in a set of *(n* + 1*)* vectors can't be linearly independent!**

We say that linearly independent vectors $e_1, e_2, e_3 \ldots e_n$ form ***a basis*** (in Ancient Greece *'βάσις'* meant 'the foundation') if any vector d can be presented as the linear combination of vectors $e_1, e_2, e_3 \ldots e_n$ i.e. for any vector we can find such real numbers $\beta_1, \beta_2, \beta_3, \ldots, \beta_n$ that the following equality is true:

$$\boldsymbol{d = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3 + \ldots + \beta_n e_n}$$

The coefficients $\beta_1, \beta_2, \beta_3, \ldots, \beta_n$ here are called the coordinates of the vector d in the basis $\{ e_1, e_2, e_3 \ldots e_n \}$

<u>Theorem</u> *Any unspecified vector can be presented in the given basis, and this presentation is unique.*

Let us look at some simple problems.

<u>Task 1 (One-minute task)</u> There is a 3-dimensional vector space V^3. In this space we have a set of vectors:

$$a = \begin{pmatrix} 1 \\ 2 \\ 10 \end{pmatrix}, b = \begin{pmatrix} 5 \\ 8 \\ 11 \end{pmatrix}, c = \begin{pmatrix} 15 \\ -7 \\ 11 \end{pmatrix}, d = \begin{pmatrix} 6 \\ 9 \\ -27 \end{pmatrix}$$

Check the linear (in-)dependence of vectors.

Task 2 There is a resolution of a vector d in the basis $\{e_1, e_2, e_3\}$: $d = 5e_1 + 2e_2 - 3e_3$. Find the coordinates of the vector d in that basis

The coordinates of a vector in the basis are the coefficients standing in front of the vectors e_1, e_2, e_3 in this basis, so we have: $d = \{5; 2; -3\}_{\{e1, e2, e3\}}$

Task 3 Check whether the vectors form a basis in V^3

$$a = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 7 \\ 8 \end{pmatrix}, c = \begin{pmatrix} 9 \\ 17 \\ 11 \end{pmatrix}$$

For this task we need to create a matrix:

$$\begin{pmatrix} 1 & 3 & 5 \\ 4 & 7 & 8 \\ 9 & 17 & 11 \end{pmatrix}$$

We check whether its determinant is not equal to zero: $\begin{vmatrix} 1 & 3 & 5 \\ 4 & 7 & 8 \\ 9 & 17 & 11 \end{vmatrix} = 118 \neq 0 \rightarrow$ the vectors are linearly independent, the space is 3-dimensional , so they form a basis in V^3.

Task 4

Find the coordinates of a vector $\boldsymbol{x} = \begin{pmatrix} 8 \\ 6 \\ 9 \end{pmatrix}$ in a basis: $\mathbf{e_1} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}$, $\mathbf{e_2} = \begin{pmatrix} 3 \\ 4 \\ -7 \end{pmatrix}$, $\mathbf{e_3} = \begin{pmatrix} 3 \\ -6 \\ 6 \end{pmatrix}$.

If we present a vector $\mathbf{x}$ in that basis, the following expression is true:

$\mathbf{X} = \beta_1\mathbf{e_1} + \beta_2\mathbf{e_2} + \beta_3\mathbf{e_3}$

Let us put down the expression in the coordinate form.

$$\begin{pmatrix} 8 \\ 6 \\ 9 \end{pmatrix} = \beta_1 \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} + \beta_2 \begin{pmatrix} 3 \\ 4 \\ -7 \end{pmatrix} + \beta_3 \begin{pmatrix} 3 \\ -6 \\ 6 \end{pmatrix}$$

$$\begin{pmatrix} 8 \\ 6 \\ 9 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ 4\beta_1 \\ 5\beta_1 \end{pmatrix} + \begin{pmatrix} 3\beta_2 \\ 4\beta_2 \\ -7\beta_2 \end{pmatrix} + \begin{pmatrix} 3\beta_3 \\ -6\beta_3 \\ 6\beta_3 \end{pmatrix}$$

So we have:

$$8 = \beta_1 + 3\beta_2 + 3\beta_3$$
$$6 = 4\beta_1 + 4\beta_2 - 6\beta_3$$
$$9 = 5\beta_1 - 7\beta_2 + 6\beta_3$$

We can solve this system using the Gaussian method or Cramer's rule. Let us take the Cramer's rule.

$$\Delta = \begin{vmatrix} 1 & 3 & 3 \\ 4 & 4 & -6 \\ 5 & -7 & 6 \end{vmatrix} = -324; \quad \Delta_1 = \begin{vmatrix} 8 & 3 & 3 \\ 6 & 4 & -6 \\ 9 & -7 & 6 \end{vmatrix} = -648$$

$$\Delta_2 = \begin{vmatrix} 1 & 8 & 3 \\ 4 & 6 & -6 \\ 5 & 9 & 6 \end{vmatrix} = -324; \Delta_3 = \begin{vmatrix} 1 & 3 & 8 \\ 4 & 4 & 6 \\ 5 & -7 & 9 \end{vmatrix} = -324$$

$$\beta_1 = \frac{\Delta_1}{\Delta} = 2\ ; \ \beta_2 = \frac{\Delta_2}{\Delta} = 1; \ \beta_3 = \frac{\Delta_3}{\Delta} = 1$$

So $\mathbf{x} = 2\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$.

Remark **Often a problem is formulated like that: there are 3 vectors *a, b, c.* Show that they form a basis and find the coordinates of a vector *x* in that basis. It's obvious that firstly you should apply an algorithm from task 3 and after having proved that the vectors form a basis, you come to the next step described in the task 4.**

Practice Russian

Линéйно независимые вéкторы – linearly independent vectors
Линéйно зависимые вéкторы – linearly dependent vectors
Бáзис – a basis

Review questions

1. Which vectors are called linearly dependent and linearly independent?
2. Which obvious signs of linear dependence can a set of vectors have?
3. What is a basis? Which vectors form a basis?

Practical tasks

1. Prove that the vectors *a, b, c* **are linearly independent or linearly dependent and find the coordinates of a vector** *x* **in the basis** *a, b, c* **(in case of linear independence of** *a, b, c***)**

1. $a=\begin{pmatrix}1\\3\\5\end{pmatrix}, b=\begin{pmatrix}7\\9\\13\end{pmatrix}, c=\begin{pmatrix}23\\37\\78\end{pmatrix}$	$x=\begin{pmatrix}9\\15\\4\end{pmatrix}$
2. $a=\begin{pmatrix}10\\6\\-2\end{pmatrix}, b=\begin{pmatrix}11\\7\\-3\end{pmatrix}, c=\begin{pmatrix}17\\8\\18\end{pmatrix}$	$x=\begin{pmatrix}8\\5\\14\end{pmatrix}$
3. $a=\begin{pmatrix}-3\\4\\7\end{pmatrix}, b=\begin{pmatrix}-2\\6\\19\end{pmatrix}, c=\begin{pmatrix}1\\3\\28\end{pmatrix}$	$x=\begin{pmatrix}7\\-1\\12\end{pmatrix}$
4. $a=\begin{pmatrix}4\\-3\\5\end{pmatrix}, b=\begin{pmatrix}-3\\2\\1\end{pmatrix}, c=\begin{pmatrix}23\\7\\8\end{pmatrix}$	$x=\begin{pmatrix}10\\-5\\7\end{pmatrix}$
5. $a=\begin{pmatrix}-9\\4\\-2\end{pmatrix}, b=\begin{pmatrix}1\\19\\24\end{pmatrix}, c=\begin{pmatrix}-8\\7\\21\end{pmatrix}$	$x=\begin{pmatrix}16\\3\\19\end{pmatrix}$
6. $a=\begin{pmatrix}-1\\13\\6\end{pmatrix}, b=\begin{pmatrix}-2\\10\\16\end{pmatrix}, c=\begin{pmatrix}54\\7\\-8\end{pmatrix}$	$x=\begin{pmatrix}11\\1\\23\end{pmatrix}$
7. $a=\begin{pmatrix}-5\\2\\1\end{pmatrix}, b=\begin{pmatrix}-3\\2\\-1\end{pmatrix}, c=\begin{pmatrix}7\\2\\-9\end{pmatrix}$	$x=\begin{pmatrix}8\\-7\\5\end{pmatrix}$
8. $a=\begin{pmatrix}8\\-13\\16\end{pmatrix}, b=\begin{pmatrix}17\\3\\5\end{pmatrix}, c=\begin{pmatrix}2\\-7\\15\end{pmatrix}$	$x=\begin{pmatrix}9\\15\\-10\end{pmatrix}$
9. $a=\begin{pmatrix}7\\-9\\15\end{pmatrix}, b=\begin{pmatrix}-8\\-2\\1\end{pmatrix}, c=\begin{pmatrix}1\\3\\2\end{pmatrix}$	$x=\begin{pmatrix}7\\-12\\-5\end{pmatrix}$
10. $a=\begin{pmatrix}7\\-14\\9\end{pmatrix}, b=\begin{pmatrix}8\\-9\\11\end{pmatrix}, c=\begin{pmatrix}-10\\10\\7\end{pmatrix}$	$x=\begin{pmatrix}8\\-22\\-6\end{pmatrix}$
11. $a=\begin{pmatrix}9\\8\\-7\end{pmatrix}, b=\begin{pmatrix}1\\12\\-19\end{pmatrix}, c=\begin{pmatrix}8\\-7\\18\end{pmatrix}$	$x=\begin{pmatrix}4\\-11\\31\end{pmatrix}$

12. $a = \begin{pmatrix} -10 \\ -13 \\ 9 \end{pmatrix}, b = \begin{pmatrix} 5 \\ -9 \\ 3 \end{pmatrix}, c = \begin{pmatrix} 3 \\ 17 \\ 8 \end{pmatrix}$	$x = \begin{pmatrix} -7 \\ -16 \\ 2 \end{pmatrix}$
13 $a = \begin{pmatrix} 6 \\ -2 \\ 1 \end{pmatrix}, b = \begin{pmatrix} 7 \\ -8 \\ 10 \end{pmatrix}, c = \begin{pmatrix} 2 \\ -7 \\ 4 \end{pmatrix}$	$x = \begin{pmatrix} 9 \\ -8 \\ 7 \end{pmatrix}$
14. $a = \begin{pmatrix} 9 \\ 2 \\ -1 \end{pmatrix}, b = \begin{pmatrix} 8 \\ -7 \\ 1 \end{pmatrix}, c = \begin{pmatrix} 5 \\ 8 \\ -4 \end{pmatrix}$	$x = \begin{pmatrix} 20 \\ -1 \\ 4 \end{pmatrix}$
15. $a = \begin{pmatrix} -1 \\ 12 \\ 7 \end{pmatrix}, b = \begin{pmatrix} -3 \\ 2 \\ 5 \end{pmatrix}, c = \begin{pmatrix} 15 \\ 8 \\ -2 \end{pmatrix}$	$x = \begin{pmatrix} 7 \\ -8 \\ 3 \end{pmatrix}$

1.11. Consistent and inconsistent systems of linear equations

You have already learned a lot about the systems of linear equations. However, let us introduce other important terms.

Let us put down the system

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

Let us put down this system in matrix form:

$$\boldsymbol{AX = B}$$

The solution of this system looks like:

$$\boldsymbol{X = A^{-1}B}$$

Here A is a *m* x *n* matrix of the system's coefficients, *X* is a *n* x 1 column of variables and B is a *m* x 1 column of numbers standing in the right-hand part. Let us take the Gaussian method for the solution of a system (*). We form an augmented matrix.

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

There are some different cases for this system's solutions. Firstly we need to get an echelon form for a matrix.

1) As a result of the elementary row transformations, we can get a row:

$$\mathbf{(0\ 0\ 0\ \ldots 0\ |}\boldsymbol{b_i}\mathbf{)}$$

That means we will have an equation of the system:

$$\mathbf{0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b}$$

So this result shows that the equation has no solution and the system itself has no solution. Such a system without solutions is called ***an inconsistent system.***

2) If we don't have any contradictions in the row, then the horizontal part of the last 'pace' of a 'ladder' will cross the vertical line and it will be continued to the end of a matrix. Before we put down the system that we've got as a result of the elementary transformations, let us consider that the 'paces' were formed in the r columns where $r = rk(A)$. We leave the previously introduced denotations for the coefficients of a system and its right-hand part.

So we have:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r + a_{1,r+1}x_{r+1} + \cdots + a_{1n}x_n = b_1, \\ \quad a_{22}x_2 + \cdots + a_{2r}x_r + a_{2,r+1}x_{r+1} + \cdots + a_{2n}x_n = b_2, \\ \quad\quad \ddots \\ \quad\quad\quad a_{rr}x_r + a_{r,r+1}x_{r+1} + \cdots + a_{rn}x_n = b_r. \end{cases}$$

The variables that have 'paces' standing close to their coefficients, are called ***the basic variables*** (in our case they are: $x_1, \ldots, x_r$); all the other variables are called ***free variables***.

<u>Remark</u> **Sometimes we have to make the first variable in the system's equation to be a basic variable if we don't have the variable with the next 'proper' number.**

For example, we have a system:

$$\begin{cases} \boldsymbol{x_1 + x_2 + x_3 + x_4 = 4,} \\ \quad\quad \boldsymbol{x_3 - 2x_4 = 3.} \end{cases}$$

Here we have the basic variables x_1, x_3 as we don't have x_2 at all in the second equation of a system (which is, in most cases, the next basic variable when we have systems with two equations).

Here we also have 2 cases.

I. $\boxed{r = n}$, i.e. $\text{Rg}(A) = n$.

In this case, we do not have free variables so the system looks like:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \qquad\quad a_{22}x_2 + \cdots + a_{1n}x_n = b_2 \\ \qquad\qquad\qquad \ddots \\ \qquad\qquad\qquad\qquad a_{nn}x_n = b_n. \end{cases}$$

So all we have to do here is to find $x_n = b_n/a_{nn}$ from the last equation and to put its value into the next-to-last equation and to go on putting the resulting values to the equation above until the very first equation of a system. You have already got acquainted with that procedure. So we see that in this case all the variables are defined in a unique way. The given system is ***a determined system*** (it is sometimes called ***an independent system***

II. $\boxed{r < n}$, т. е. $\text{Rg}(A) < n$**.**

In this case we have $n - r$ free variables: $x_{r+1}, \ \dots, \ x_n$.

If we give some concrete values to those variables and put them into the right-hand part of a system, then for evaluation of basic variables we get the system which would be similar to the previous case. However, the difference is that the number r will play a role of a number n here. That means that the basic variables (with the condition that the free variables are already given) will be defined in a unique way. However, we can give any values to the free variables, and in the investigated case the system is undetermined. It is important to notice that such a system has ***an infinite number of solutions***.

However, when we put concrete numbers instead of free variables and then we define the values of basic variables, we get a solution of a system. So we can get any possible solution using this algorithm, can't we? Indeed, we can do that. Let us suppose that we have a specified solution of a given system. A given system is equivalent to the system led to an echelon form. That means, our solution also satisfies the system in the echelon form. If we suppose the free variables to be equal to the last (n-r) values of our solution, we will find the basic variables to be equal to those in our introduced solution, as they are defined in a unique way by the free variables.

So we have proved an important theorem.

If we have a system of linear equations, this system can be inconsistent; otherwise, in the case it is consistent, if *rk(A) = n*, it is ***a determined system***, and in the case *rk(A) < n*, it is ***underdetermined.***

In mathematics, a system of equations is considered ***overdetermined*** if there are more equations than unknowns. In most cases it doesn't have solutions. A system which has more unknowns than equations, is called ***underdetermined***.

Practice Russian

Совме́стная систе́ма – consistent system

Несовме́стная систе́ма – inconsistent system

Review questions

1. What are the definitions for an inconsistent and for a consistent system?
2. Describe the classification of systems by the number of their solutions.

1.12. Homogeneous systems of linear equations

A homogeneous system of linear equations is a system where all the equations have only zeros in the right-hand part (no non-zero numbers in the right-hand part):

$$AX = 0$$

Here A is a matrix of the system's coefficients, x is a column of variables – and zero in the right-hand part of this expression means that in *all* equations of the system each right-hand part is equal to zero.

Remark **A homogeneous system always has at least one solution, and this solution is a zero vector. Such a solution is called *a trivial solution*.**

So, if a homogeneous system has only one solution, this solution is a trivial one.

Let us firstly solve a system that has a unique solution:

$$\begin{cases} x + y = 0 \\ 2x - y = 0 \end{cases}$$

However, this unique solution is trivial: $x = 0$, $y = 0$. If we take the augmented matrix of a system, we have: $A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & -1 & 0 \end{pmatrix}$

If we lead the matrix to an echelon form (we multiply the first row by (-2) and add it to the second row; the result is put into a second row):

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

we see that $rk(A) = 2$. So we have just checked that the system has a unique solution (see the theorem above).

However, we need to have a universal algorithm of solving the system of linear equations even in case $rk(A) < n$.

Let us have a homogeneous system:

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

Let us consider from the very beginning that $\mathrm{rk}(A) = r < n$, where A is a matrix of system's coefficients. (We have already learned that in the case $r = n$ this system has only a trivial solution). Let us lead the system above to an echelon form and let us suppose that the 'paces of a ladder' have been formed in the *r* columns that come first. Then we have:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1r}x_r + a_{1,r+1}x_{r+1} + \cdots + a_{1n}x_n = 0 \\ \qquad a_{22}x_2 + \cdots + a_{2r}x_r + a_{2,r+1}x_{r+1} + \cdots + a_{2n}x_n = 0 \\ \qquad\qquad \ddots \\ \qquad\qquad a_{rr}x_r + a_{r,r+1}x_{r+1} + \cdots + a_{rn}x_n = 0 \end{cases}.$$

So $x_1, x_2, \ldots, x_r$ are the basic variables and $x_{r+1}, x_{r+2}, \ldots, x_n$ are the free variables. We have $n - r$ free variables.

If we give some numeric values to the free variables, then the values of basic variables will be defined in a unique way. In this case, let us suppose firstly:

$$x_{r+1} = 1,\ x_{r+2} = 0, \qquad x_{r+3} = 0, \ldots, x_n = 0.$$

We put those numbers into a system and so we find the values of basic variables. Let us get $x_1 = x_{11}^0,\ x_2 = x_{12}^0, \ldots,\ x_r = x_{1r}^0$ as a result. Then a vector

$$E_1 = (x_{11}^0, x_{12}^0, \ldots, x_{1r}^0, 1,0, \ldots ,0)^{\mathrm{T}}$$

is a solution of our system. Then let us suppose:

$$x_{r+1} = 0\ ,\ x_{r+2} = 1\ , x_{r+3} = 0\ , \dots ,\ x_n = 0.$$

In a similar way we define a vector

$$E_2 = (x_{21}^0\ , x_{22}^0, \dots , x_{2r}^0, 0,1,0, \dots ,0)^{\mathrm{T}},$$

That is also the solution of the system. If we continue this process, we put down on the last step:

$$x_{r+1} = 0\ ,\ x_{r+2} = 0\ , x_{r+3} = 0\ , \dots ,\ x_n = 1$$

and we get a solution:

$$E_{n-r} = \left(x_{n-r,1}^0\ , x_{n-r,2}^0, \dots , x_{n-r,r}^0, 0,0,0, \dots ,1\right)^{\mathrm{T}}.$$

Firstly, we need to check that the solutions that are found using this algorithm are unique. For this, we take a matrix:

$$\left(\begin{matrix} x_{11}^0 & x_{12}^0 & \dots & x_{1r}^0 \\ x_{21}^0 & x_{22}^0 & \dots & x_{2r}^0 \\ \dots & \dots & \dots & \dots \\ x_{n-r,1}^0 & x_{n-r,2}^0 & \dots & x_{n-r,r}^0 \end{matrix} \quad \boxed{\begin{matrix} 1\ 0\ 0 \dots 0 \\ 0\ 1\ 0 \dots 0 \\ \dots \dots \dots \dots \\ 0\ 0\ 0 \dots 1 \end{matrix}}\right),$$

(the structure of this matrix is obvious so we don’t give any comments). The framed minor of the order $n-r$ is equal to 1. Then the vectors $E_1, E_2, \dots , E_{n-r}$ are linearly independent.

Let $\tilde{X} = (\tilde{x}_1,\ \tilde{x}_2, \dots ,\ \tilde{x}_r,\ \tilde{x}_{r+1}, \dots ,\ \tilde{x}_n)^{\mathrm{T}}$ be any unspecified solution of a system. Let us investigate the solution

$$\tilde{X}^0 = \tilde{x}_{r+1}E_1 + \tilde{x}_{r+2}E_2 + \cdots + \tilde{x}_n E_{n-r}$$

simultaneously with the previous one.

It is obvious that for these two solutions the values of free variables are equal. So the values of the basic variables will also be equal! Thus, $\tilde{X} = \tilde{X}^0$, i.e. they are one solution. So we have just found out that any solution of a system is presented with the help of $E_1\ , E_2\ , \dots , E_{n-r}$. Therefore, those vectors form a basis in the space of system’s solutions – it is called the fundamental system of solutions.

If $X_{g.h.}$ is a denotation for a general solution of a homogeneous system (in other words, a set of all its solutions), then we have a structural formula:

$$\boxed{X_{g.h.} = C_1E_1 + C_2E_2 + \cdots + C_{n-r}E_{n-r}},$$

We have also discovered a fact that a number of variables n of any homogeneous system consists of a rank r of a system's matrix and a dimension k of the space of its solutions:

$$\boxed{k + r = n},$$

And this formula is also true for $r = n$, when a system has only a trivial solution.

Remark **In this paragraph we have given a method of solution for a *specified* system. However, it is necessary to understand that for 'finding' the general solution of a system we can take any fundamental system of solutions. We can take any unspecified (but non-zero!!!) numbers instead of '1' that we have taken for the solution above. The values of the free variables can be any numbers you wish to put – however, the 'framed' minor (see the example above) should not be equal to zero.**

Example Find the solution of a system:

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

$$\begin{pmatrix} 1 & 2 & -1 & 4 \\ 1 & 5 & -3 & 2 \end{pmatrix} \xrightarrow{I\cdot(-1)+II\rightarrow II} \begin{pmatrix} 1 & 2 & -1 & 4 \\ 0 & 3 & -2 & -2 \end{pmatrix}$$

So we have a system afterwards:

$$\begin{cases} x_1 + 2x_2 - x_3 + 4x_4 = 0 \\ \qquad 3x_2 - 2x_3 - 2x_4 = 0 \end{cases}$$

x_1, x_2 are basic variables; x_3, x_4 are free variables.

Let us suppose x_3=3, x_4=0. The system will look like:

$$\begin{cases} x_1 + 2x_2 - 3 = 0 \\ \qquad 3x_2 - 6 = 0 \end{cases}$$

$$\begin{cases} x_1 + 2x_2 = 3 \\ \qquad x_2 = 2 \end{cases}$$

$$\begin{cases} x_1 = -1 \\ x_2 = 2 \end{cases}$$

Let us now suppose $x_3 = 0$, $x_4 = 3$. We have the following system:

$$\begin{cases} x_1 + 2x_2 + 4 = 0 \\ \qquad 3x_2 - 6 = 0 \end{cases}$$

$$\begin{cases} x_1 = -8 \\ x_2 = 2 \end{cases}$$

So we get a fundamental system of solutions:

$$E_1 = \begin{pmatrix} -1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, E_2 = \begin{pmatrix} -8 \\ 2 \\ 0 \\ 3 \end{pmatrix},$$

$$X_{g.h.} = C_1 \begin{pmatrix} -1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} -8 \\ 2 \\ 0 \\ 3 \end{pmatrix}; \ C_1, C_2 \in R$$

Practice Russian

Одноро́дная систе́ма лине́йных уравне́ний – homogeneous system of linear equations

Фундамента́льная систе́ма реше́ний – fundamental system of solutions

Practical task

1. Solve the homogeneous system

1. $\begin{cases} 2x_1 - 4x_2 + 6x_3 - 5x_4 = 0 \\ 10x_2 + 4x_3 - 6x_4 = 0 \end{cases}$

2. $\begin{cases} 3x_1 + 7x_2 + 6x_3 - 9x_4 = 0 \\ 9x_2 - 3x_3 + 6x_4 = 0 \end{cases}$

3. $\begin{cases} x_1 + 3x_2 + 4x_3 - 5x_4 = 0 \\ 6x_2 + 8x_3 - x_4 = 0 \end{cases}$

4. $\begin{cases} 4x_1 + 7x_2 - 12x_3 + 2x_4 = 0 \\ 4x_2 + 8x_3 - 2x_4 = 0 \end{cases}$

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

6. $\begin{cases} 5x_1 + 10x_2 - 9x_3 + x_4 = 0 \\ 15x_2 + 10x_3 + 15x_4 = 0 \end{cases}$

7. $\begin{cases} 3x_1 + 7x_2 + 2x_3 - 2x_4 = 0 \\ 9x_2 + 18x_3 + 3x_4 = 0 \end{cases}$

8. $\begin{cases} 4x_1 + 6x_2 - 2x_3 + 7x_4 = 0 \\ 6x_1 - 8x_2 + 2x_3 = 0 \end{cases}$

9. $\begin{cases} x_1 - x_2 + 2x_3 - 4x_4 = 0 \\ -6x_2 + 12x_3 - x_4 = 0 \end{cases}$

10. $\begin{cases} 8x_1 - 4x_2 - x_3 + 7x_4 = 0 \\ 16x_2 - 20x_3 + 5x_4 = 0 \end{cases}$

11. $\begin{cases} x_1 + 8x_2 - x_3 + 3x_4 = 0 \\ x_2 + 4x_3 - 5x_4 = 0 \end{cases}$

12. $\begin{cases} x_1 + 15x_2 - 2x_3 + x_4 = 0 \\ -x_2 + x_3 + 4x_4 = 0 \end{cases}$

13. $\begin{cases} 2x_1 - x_2 - 3x_3 + 8x_4 = 0 \\ 12x_2 + 8x_3 - 7x_4 = 0 \end{cases}$

14. $\begin{cases} 4x_1 - x_2 + 3x_3 + x_4 = 0 \\ 9x_2 - 12x_3 + 5x_4 = 0 \end{cases}$

15. $\begin{cases} 15x_1 + 2x_2 + 5x_3 - 7x_4 = 0 \\ 10x_2 + 2x_3 - 6x_4 = 0 \end{cases}$

1.13. Inhomogeneous systems of linear equations

Let us introduce the other type of systems – ***an inhomogeneous system of linear equations.*** In this system in the right-hand part of at least one equation we have a non-zero number.

Let us investigate the inhomogeneous system of linear equations.

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

Any its concrete solution is called ***a partial solution,*** and we denote it as $X_{p.i.}$ (partial solution of an inhomogeneous system). If in this system all the right-hand parts are equal to zero, we get a homogeneous system that is called ***allied homogeneous system.***

Theorem (about a structure of general solution of an inhomogeneous system)

A general solution of a consistent inhomogeneous system consists of its fixed (partial) solution and of an allied homogeneous system's general solution

Proof We put down our inhomogeneous system in a matrix form $AX = B$. Then an allied homogeneous system is $AX = O$. Let us fix the partial solution of an inhomogeneous system X_1. So $AX_1 = B$. If X_0 is any solution of an allied system, i.e. $AX_0 = O$, then we will have $A(X_1 + X_0) = A(X_1) + A(X_0) = B + O = B$, so. $X_1 + X_0$ also will be the solution of an inhomogeneous system. We suppose now that X_2 is an unspecified solution of a given system, and let us look at the difference $X_2 - X_1 = X^0$. Then $AX^0 = A(X_2 - X_1) = A(X_2) - A(X_1) = B - B = O$. We get: X^0 is a solution of an allied system and, therefore, we have $X_2 = X_1 + X^0$. If we put any possible solutions of an allied system instead of X in the expression $X_1 + X$, then we will get any possible solutions of a given inhomogeneous system – and no solutions will be missed! This fact is described in a structural formula:

$$\boxed{X_{g.i.} = X_{p.i.} + X_{g.h.}},$$

here $X_{g.i.}$ is a denotation for a general solution of an inhomogeneous system. In the other form it looks like:

$$\boxed{X_{g.i.} = X_{p.i.} + C_1E_1 + C_2E_2 + \cdots + C_{n-r}E_{n-r}}. \ \blacksquare$$

Kronecker-Capelli theorem

The system of linear equations $AX = B$ is consistent if and only if the rank of an augmented matrix of a given system is equal to a rank of a given system: $\mathrm{Rg}(A|B) = \mathrm{Rg}(A)$.

Task 1. Solve the inhomogeneous system of linear equations.

$$\begin{cases} x_1 + 2x_2 - 3x_3 + x_4 = 10 \\ \quad x_2 + 3x_3 + 4x_4 = 8 \end{cases}$$

In this case, we don't need to lead a matrix to an echelon form. We have basic variables x_1, x_2 and free variables x_3, x_4.

First of all, we need to find a partial solution of an inhomogeneous system. We make all the free variables to be equal to zero, we leave only the basic variables and right-hand parts in the equations of a system.

$$\begin{cases} x_1 + 2x_2 = 10 \\ \quad x_2 = 8 \end{cases}$$

$$\begin{cases} x_1 = -6 \\ x_2 = 8 \end{cases}$$

So $X_{p.i.} = \begin{pmatrix} -6 \\ 8 \\ 0 \\ 0 \end{pmatrix}$

Then we put down an allied homogeneous system and use the algorithm described above for the theme 'Homogeneous systems of linear equations'.

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

1. $x_3 = 1, x_4 = 0$

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

$$\begin{cases} x_1 - 6 - 3 = 0 \\ x_2 = -3 \end{cases}$$

$$\begin{cases} x_1 = 9 \\ x_2 = -3 \end{cases}$$

$$E_1 = \begin{pmatrix} 9 \\ -3 \\ 1 \\ 0 \end{pmatrix}$$

2. $x_3 = 0, x_4 = 1$

$$\begin{cases} x_1 + 2x_2 + 1 = 0 \\ x_2 + 4 = 0 \end{cases}$$

$$\begin{cases} x_1 - 8 + 1 = 0 \\ x_2 = -4 \end{cases}$$

$$\begin{cases} x_1 = -7 \\ x_2 = -4 \end{cases}$$

$$E_2 = \begin{pmatrix} -7 \\ -4 \\ 0 \\ 1 \end{pmatrix}$$

$$X_{g.i.} = X_{p.i.} + C_1E_1 + C_2E_2$$

$$X_{g.i.} = \begin{pmatrix} -6 \\ 8 \\ 0 \\ 0 \end{pmatrix} + C_1\begin{pmatrix} 9 \\ -3 \\ 1 \\ 0 \end{pmatrix} + C_2\begin{pmatrix} -7 \\ -4 \\ 0 \\ 1 \end{pmatrix}$$

Remark There is no factor C_i in front of the column of a partial solution. Try to explain this fact.

Practice Russian

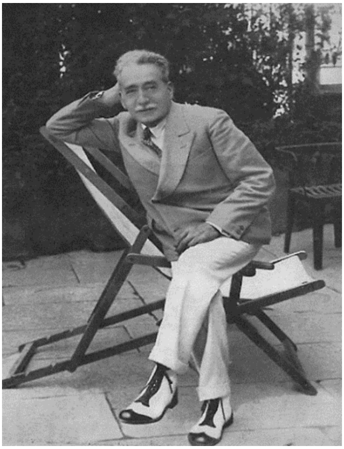
Неоднорóдная систéма линéйных уравнéний – inhomogeneous system of linear equations

Союзная однорóдная систéма – allied homogeneous system

Leopold Kronecker (1823–1891)

Inspiring biography *Leopold Kronecker* was born in Liegnitz, Prussia (now Legnica, Poland) in 1823. Kronecker became a student at Berlin University in 1841. He showed great interest in studying mathematics as well as in astronomy, meteorology and chemistry. He was especially interested in philosophy studying the philosophical works of Descartes, Leibniz, Kant, Spinoza and Hegel. He spent a summer of 1843 at the University of Bonn, which he went to because of his interest in astronomy rather than mathematics, he then went to the University of Breslau for the winter semester of 1843–44. Kronecker spent a year at Breslau before returning to Berlin for the winter semester of 1844–45. In Berlin he worked on his doctoral thesis on algebraic number theory under Dirichlet's supervision. The title of his thesis was *On complex units* and he became a PhD in 1845. (Kronecker was questioned at his oral examination on a wide range of topics including the theory of probability as applied to astronomical observations, the theory of definite integrals, series and differential equations, as well as on Greek, and the history of philosophy). Kronecker's works were in the theory of equations and higher algebra, the theory of algebraic equations, and the theory of algebraic numbers. However, he liked to investigate only the integers and a finite number of steps. Kronecker is well known for his remark: '*God created the integers, all else is the work of man*'. He was a professor of Berlin University; however, he accepted the honours of Paris Academy, St. Petersburg University, Royal Society of London for Improving Natural Knowledge.

Inspiring biography *Alfredo Capelli* was born in 1855 in Milan, Lombardo-Veneto (now Italy). He attended the University of Rome and in 1878, he wrote his thesis dedicated to the theory of groups that had a great success. In 1880s he published some brilliant works about solving the systems of linear equations – for example, he and Gabrieri showed that a system of equations having rank k is equivalent to a triangular system with exactly k nonzero diagonal terms. He was a professor at University of Palermo and later got a chair of algebra at the University of Naples where he stayed for the rest of his life.



Alfredo Capelli (1855–1910)

Practical task

1. **Solve the homogeneous system**

1. $\begin{cases} 2x_1 - 4x_2 + 6x_3 - 5x_4 = 4 \\ 10x_2 + 4x_3 - 6x_4 = 7 \end{cases}$

2. $\begin{cases} 3x_1 + 7x_2 + 6x_3 - 9x_4 = 3 \\ 9x_2 - 7x_3 + 6x_4 = 8 \end{cases}$

3. $\begin{cases} x_1 + 3x_2 + 4x_3 - 5x_4 = 4 \\ 6x_2 + 8x_3 - x_4 = 2 \end{cases}$

4. $\begin{cases} 4x_1 + 7x_2 - 12x_3 + 2x_4 = 9 \\ 4x_2 + 8x_3 - 2x_4 = 6 \end{cases}$

5. $\begin{cases} 5x_1 + 7x_2 - 2x_3 + 4x_4 = 10 \\ 4x_2 + 2x_3 - 5x_4 = 8 \end{cases}$

6. $\begin{cases} 5x_1 + 10x_2 - 9x_3 + x_4 = 3 \\ 15x_2 + 10x_3 + 15x_4 = 12 \end{cases}$

7. $\begin{cases} 3x_1 + 7x_2 + 2x_3 - 2x_4 = 5 \\ 9x_2 + 18x_3 + 3x_4 = 9 \end{cases}$

8. $\begin{cases} 4x_1 + 6x_2 - 2x_3 + 7x_4 = 7 \\ 6x_1 - 8x_2 + 2x_3 = 6 \end{cases}$

9. $\begin{cases} x_1 - x_2 + 2x_3 - 4x_4 = 4 \\ -6x_2 + 12x_3 - x_4 = 2 \end{cases}$

10. $\begin{cases} 8x_1 - 4x_2 - x_3 + 7x_4 = 10 \\ 16x_2 - 20x_3 + 5x_4 = 20 \end{cases}$

11. $\begin{cases} x_1 + 8x_2 - x_3 + 3x_4 = 1 \\ x_2 + 4x_3 - 5x_4 = 4 \end{cases}$

12. $\begin{cases} x_1 + 15x_2 - 2x_3 + x_4 = 8 \\ -x_2 + x_3 + 4x_4 = 5 \end{cases}$

13. $\begin{cases} 2x_1 - x_2 - 3x_3 + 8x_4 = 12 \\ 12x_2 + 8x_3 - 7x_4 = 6 \end{cases}$

14. $\begin{cases} 4x_1 - x_2 + 3x_3 + x_4 = 4 \\ 9x_2 - 12x_3 + 5x_4 = 1 \end{cases}$

15. $\begin{cases} 15x_1 + 2x_2 + 5x_3 - 7x_4 = 7 \\ 10x_2 + 2x_3 - 6x_4 = 2 \end{cases}$

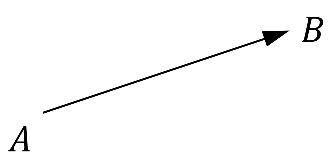
Creative task We have just got acquainted with an interesting theme and we have come to the end of the chapter. It is time to show your talents! You can take any concrete theme, formula or a definition from this chapter and create something interesting dedicated to this theme! It can be a poem, a crossword puzzle, a painting or a pencil sketch, a fairy-tale or even a song – however, your creative work should rely to the matrix theory and systems of linear equations. Share your achievements with friends and remember: it is not a contest, it is a festival, and each participant is already a winner! This task is not necessary – however, we hope all of you have something positive to say, to draw or to sing!

2. ANALYTICAL GEOMETRY

Analytical geometry is a branch of mathematics in which we investigate geometry with the use of algebraic methods. The coordinate system is the basic object of the analytical geometry. The system of coordinates lets us establish the connection between geometrical images and algebraic constructions, e.g. equations, inequalities, etc. In further material, we would work with Euclidean geometry (i.e. the geometry that you have already learned at school). That is why we will not discuss the details of objects' definitions, such as 'a point', 'a straight line', 'an angle', 'a length', 'a volume', etc. However, we should note that we would have one unit of measurement for the length – one for the whole chapter. So we would denote lengths, surface areas and volumes by dimensionless numbers. (Thus, we won't wonder why, for example, the length is equal to the volume. We use only the numerical values!)

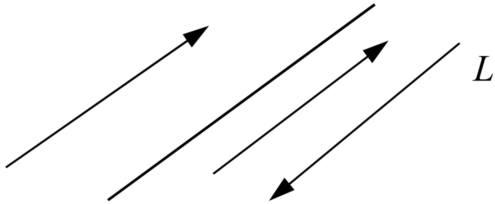
2.1. Vectors. Basic definitions and elementary operations

For example, we have two given points A and B. We draw a line from A to B and so we get a segment AB. We put 'an arrow' at the end of the segment (close to the point B). This new object is called a ***geometric vector*** or just a ***vector*** $\overrightarrow{AB}$. This word came from Latin language and in those ancient times it meant 'a transporter, a carrier'.

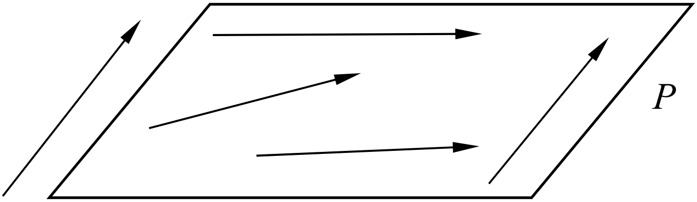


More precisely, $\overrightarrow{AB}$ is a fixed vector and it is considered that it is fixed at the point A. However, we won't put down the word 'fixed'. The length of the segment AB is called ***a vector modulus*** and is denoted as $|\overrightarrow{AB}|$. If the point A coincides with the vector B then this vector is called ***a zero vector***. Obviously, it is the only vector that has a modulus which is equal to zero. We can't draw a precise picture of a zero vector. We don't consider that depicting it as a point is a right way. *A point is not a vector.* If $|\overrightarrow{AB}| = 1$ then the vector is called ***a unitary vector.***

Collinearity The set of vectors is called ***collinear*** if all of those vectors are parallel to a straight line L (we don't exclude the case when the vector 'lies' on this straight line). They also say at this case that the straight line L is collinear to those vectors. According to the definition, it is considered that each straight line is collinear to a zero vector. Thus, if we join a zero vector to a collinear set or remove this vector from that set, we say that this operation doesn't influence the collinearity.



Coplanarity A set of vector is called ***coplanar*** if all those vectors are parallel to a plane P. It is obvious that the collinear set of vectors is also coplanar.



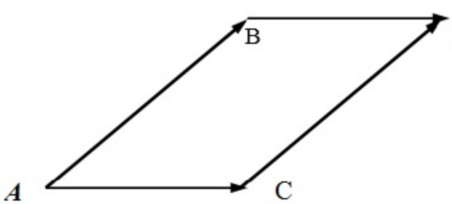
Two vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are called ***the equal vectors*** if 3 conditions are fulfilled:

1) *The vectors* $\overrightarrow{AB}$ *and* $\overrightarrow{CD}$ *are collinear*;

2) *They have the same direction*;

3) $|\overrightarrow{AB}| = |\overrightarrow{CD}|$.

You can look at the examples of equal vectors in the picture below.



We can see from this picture that $\overrightarrow{AB}$ and $\overrightarrow{CD}$ are equal only in the case when $\overrightarrow{AC}$ and $\overrightarrow{BD}$ are also equal vectors. However, the graphical obviosity is not a real proof! We will not prove the fact of vectors' equality but we will introduce the axiom which would further allow us to use this fact without restrictions.

Equality axiom: for any 4 points A,B,C,D the following expression is always true:

$$\overrightarrow{AB} = \overrightarrow{CD} \quad \Leftrightarrow \quad \overrightarrow{AC} = \overrightarrow{BD}.$$

We should notice that in this axiom we don't claim that those 4 points should be different! There are simple but important statements which come out of this axiom.

***Statement* 1.** *All the fixed zero vectors are equal to each other.*

Indeed, let A and B be any unspecified points. Zero vector fixed in the point A, is denoted as $\overrightarrow{AA}$, zero vector fixed in the point B, is denoted as $\overrightarrow{BB}$. Then, using the axiom and the obvious equality, we can tell that $\overrightarrow{AB} = \overrightarrow{AB} \Rightarrow \overrightarrow{AA} = \overrightarrow{BB}$.

This statement gives us an opportunity to speak about a zero vector in all the cases – without dependence on the place where it is fixed. Let us denote a zero vector: $\vec{\mathbf{0}}$.

***Statement* 2.** If $\overrightarrow{AB} = \overrightarrow{AC}$ then the point B coincides with the point *C*.

Indeed, from the equality axiom we get: $\overrightarrow{AB} = \overrightarrow{AC} \Rightarrow \overrightarrow{AA} = \overrightarrow{BC}$. However, $\overrightarrow{AA} = \vec{\mathbf{0}}$. Then B and C can't be different points.

A fixed vector $\overrightarrow{BA}$ is called ***the opposite vector to*** $\overrightarrow{\mathbf{AB}}$ and it is denoted as $-\overrightarrow{AB}$. So $\overrightarrow{BA} = -\overrightarrow{AB}$.

We can show that if $\overrightarrow{AB} = \overrightarrow{CD}$, then $\overrightarrow{BA} = \overrightarrow{DC}$. Indeed, it comes from the following statements:

$\overrightarrow{AB} = \overrightarrow{CD} \Rightarrow \overrightarrow{CD} = \overrightarrow{AB} \Rightarrow \overrightarrow{CA} = \overrightarrow{DB} \Rightarrow \overrightarrow{DB} = \overrightarrow{CA} \Rightarrow \overrightarrow{DC} = \overrightarrow{BA} \Rightarrow \overrightarrow{BA} = \overrightarrow{DC}$. Here we have used the axiom of vectors' equality twice.

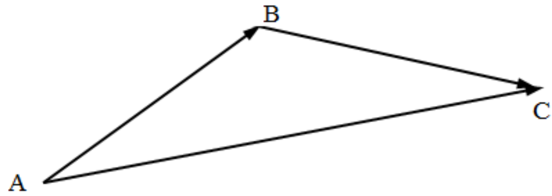
Thus, we have the following: if there is a vector $\vec{\boldsymbol{a}} = \{\overrightarrow{AB}, \overrightarrow{CD}, \overrightarrow{PE}, \dots\}$, then we can correctly define a vector $\vec{\boldsymbol{b}} = \{\overrightarrow{BA}, \overrightarrow{DC}, \overrightarrow{EP}, \dots\}$, as inside the brackets we also have vectors that are equal to each other. Vector $\vec{\boldsymbol{b}}$ we will denote as $-\vec{\boldsymbol{a}}$ and we will call it the ***opposite vector*** to $\vec{\boldsymbol{a}}$.

The addition of vectors.

The axiom of addition: For each of the points A, B, C, the following equality is true:

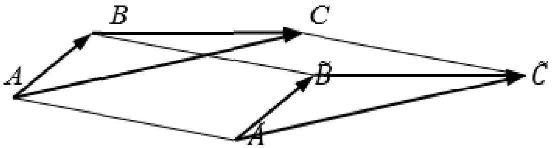
$$\boxed{\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC}}$$

If all the vectors which form the equality above, are non-zero vectors, then the axiom of addition is equivalent to the ***triangle rule*** for vectors' addition. The illustration of this rule is depicted below: if we draw the vector $\overrightarrow{AB}$, then we draw the vector $\overrightarrow{BC}$ coming from the ending point of the vector $\overrightarrow{AB}$. The vector $\overrightarrow{AC}$ is a vector of the sum for this two vectors $\overrightarrow{AB}$ and $\overrightarrow{BC}$: the starting point of $\overrightarrow{AC}$ coincides with the starting point of $\overrightarrow{AB}$ (the first vector of the sum) and the ending point of $\overrightarrow{AC}$ coincides with the ending point of $\overrightarrow{BC}$ (the second vector of the sum).



For example, there are two vectors $\vec{\boldsymbol{a}}$ и $\vec{\boldsymbol{b}}$. Let us put the vector $\vec{\boldsymbol{a}}$ from the point A, and let the point B be its ending point. Then $\vec{\boldsymbol{a}}=\overrightarrow{AB}$. Then we put the vector $\vec{\boldsymbol{b}}$ from the point B, and let the point C be its ending point, so that $\vec{\boldsymbol{b}}=\overrightarrow{BC}$. *The sum* $\vec{\boldsymbol{a}}+\vec{\boldsymbol{b}}$ of vectors $\vec{\boldsymbol{a}}$ и $\vec{\boldsymbol{b}}$ is a vector $\vec{\boldsymbol{c}}$, which can be depicted as the vector $\overrightarrow{AC}$. We should put down: $\vec{\boldsymbol{c}}=\vec{\boldsymbol{a}}+\vec{\boldsymbol{b}}$.

The result of the vector addition doesn't depend on the choice of the starting point. If $\overrightarrow{AB}=\overrightarrow{\tilde{A}\tilde{B}}$ and $\overrightarrow{BC}=\overrightarrow{\tilde{B}\tilde{C}}$, then $\overrightarrow{AC}=\overrightarrow{\tilde{A}\tilde{C}}$. In the picture below we can see the illustration for that law.



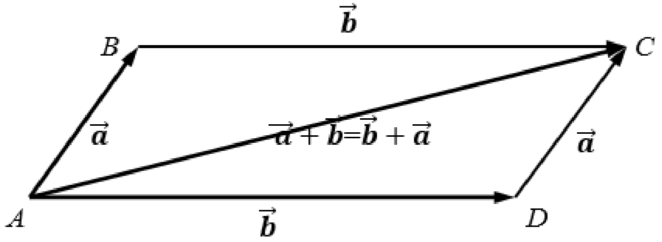
If we don't use the picture and describe that fact analytically, we get the necessary result in a following way:

$$\left.\begin{matrix}\overrightarrow{AB}=\overrightarrow{\tilde{A}\tilde{B}}\\ \overrightarrow{BC}=\overrightarrow{\tilde{B}\tilde{C}}\end{matrix}\right\} \Rightarrow \left.\begin{matrix}\overrightarrow{A\tilde{A}}=\overrightarrow{B\tilde{B}}\\ \overrightarrow{B\tilde{B}}=\overrightarrow{C\tilde{C}}\end{matrix}\right\} \Rightarrow \overrightarrow{A\tilde{A}}=\overrightarrow{C\tilde{C}} \Rightarrow \overrightarrow{AC}=\overrightarrow{\tilde{A}\tilde{C}}.$$

Here we use the equality axiom several times.

Addition properties

1) $\boxed{\vec{\boldsymbol{a}}+\vec{\boldsymbol{b}}=\vec{\boldsymbol{b}}+\vec{\boldsymbol{a}}}$, i.e. the vectors' addition is commutative. We can illustrate it geometrically using the well-known parallelogram rule. We won't prove the rule because it is quite obvious but it is highly important to investigate the illustration in the picture below.



Here $\vec{\boldsymbol{a}}=\overrightarrow{AB}=\overrightarrow{DC}$ and $\vec{\boldsymbol{b}}=\overrightarrow{BC}=\overrightarrow{AD}$.

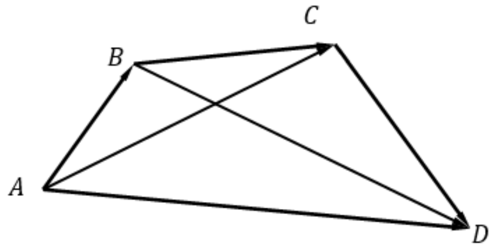
In fact, we need to prove that $\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{BC}+\overrightarrow{AB}$.

Firstly, $\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC}$. Then we create a vector $\vec{\boldsymbol{b}}=\overrightarrow{AD}$ and we draw a vector $\vec{\boldsymbol{a}}$ from the point D. Then let $\tilde{C}$ be its ending point, i.e. $\vec{\boldsymbol{a}}=\overrightarrow{D\tilde{C}}$. In this case $\overrightarrow{BC}+\overrightarrow{AB}=\vec{\boldsymbol{b}}+\vec{\boldsymbol{a}}=\overrightarrow{AD}+\overrightarrow{D\tilde{C}}=\overrightarrow{A\tilde{C}}$. Then we have

$$\left.\begin{matrix}\overrightarrow{AD}=\overrightarrow{BC}\\ \overrightarrow{AB}=\overrightarrow{D\tilde{C}}\end{matrix}\right\} \Rightarrow \left.\begin{matrix}\overrightarrow{AB}=\overrightarrow{DC}\\ \overrightarrow{AB}=\overrightarrow{D\tilde{C}}\end{matrix}\right\} \Rightarrow \overrightarrow{DC}=\overrightarrow{D\tilde{C}}$$ and that means that (with the use of statement 2 above), C и $\tilde{C}$ is the same point. Then $\overrightarrow{AC}=\overrightarrow{A\tilde{C}}$. The property is proved.

2) $\boxed{(\vec{\boldsymbol{a}}+\vec{\boldsymbol{b}})+\vec{\boldsymbol{c}}=\vec{\boldsymbol{a}}+(\vec{\boldsymbol{b}}+\vec{\boldsymbol{c}})}$, i.e. the addition of vectors is associative. In fact, this property allows us not to put down any brackets in similar statements.

Let us denote: $\vec{\boldsymbol{a}} = \overrightarrow{AB},\ \vec{\boldsymbol{b}} = \overrightarrow{BC}, \vec{\boldsymbol{c}} = \overrightarrow{CD}$; then, on one hand, we have $\left(\overrightarrow{AB} + \overrightarrow{BC}\right) + \overrightarrow{CD} = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD}$, and on the other hand, we get $\overrightarrow{AB} + \left(\overrightarrow{BC} + \overrightarrow{CD}\right) = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}$. The property is proved and below you can look at the illustration of that same property.



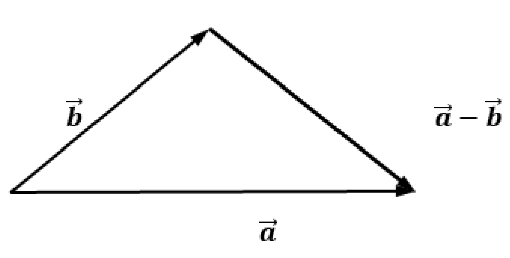
3) $\boxed{\vec{\boldsymbol{a}} + \vec{\mathbf{0}} = \vec{\boldsymbol{a}}}$**,** i.e. the addition of a zero vector to any vector doesn't change this vector. The proof is obvious:

$\vec{\boldsymbol{a}} + \vec{\mathbf{0}} = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} = \vec{\boldsymbol{a}}$. Here we used the addition axiom.

4) $\boxed{\vec{\boldsymbol{a}} + (-\vec{\boldsymbol{a}}) = \vec{\mathbf{0}}}$, i. e. if any vector is added to its opposite vector, we get zero vector as a result. Indeed, $\vec{\boldsymbol{a}} + (-\vec{\boldsymbol{a}}) = \overrightarrow{AB} + \left(-\overrightarrow{AB}\right) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \vec{\mathbf{0}}$.

The ***<u>subtraction of vectors</u>*** is defined in a following way: for any $\vec{a}$ and for any $\vec{b}$: $\left(\vec{\boldsymbol{a}} - \vec{\boldsymbol{b}} := \vec{\boldsymbol{a}} + \left(-\vec{\boldsymbol{b}}\right)\right)$, where the denotation (:=) can be read as «is equal according to the definition». From that definition we get $\vec{\boldsymbol{a}} - \vec{\boldsymbol{a}} = \vec{\mathbf{0}}$ and it is quite obvious.

<u>Remark</u> From this statement we can formulate the geometrical rule for subtraction. To get the difference $\vec{\boldsymbol{a}} - \vec{\boldsymbol{b}}$ **we should add the vector** $\overrightarrow{-\boldsymbol{b}}$ **to the vector** $\vec{\boldsymbol{a}}$**.**



Vectors' multiplication by a scalar.

Let us introduce the vector $\vec{a} \neq \vec{0}$ and the scalar $\alpha \neq 0$. **The product of a scalar α and the vector $\vec{a}$** is such a vector $\vec{b}$ *(*we put down: $\vec{b} = \alpha\vec{a}$*)* that satisfies all the three conditions:

I) $|\vec{b}| = |\alpha||\vec{a}|$;

II) the vectors $\vec{a}$ *and* $\vec{b}$ are collinear;

III) the vectors $\vec{a}$ *u* $\vec{b}$ have the same direction if $\alpha > 0$ and they have the opposite direction if $\alpha < 0$.

In all the other cases we suppose that $\alpha\vec{a} = \vec{0}$.

Let us investigate the main properties of this operation (α and β are scalars, $\vec{a}$ and $\vec{b}$ are vectors):

1) $(\alpha\beta)\vec{a} = \alpha(\beta\vec{a})$;

2) $(\alpha + \beta)\vec{a} = \alpha\vec{a} + \beta\vec{a}$;

3) $\alpha(\vec{a} + \vec{b}) = \alpha\vec{a} + \alpha\vec{b}$;

4) $1 \cdot \vec{a} = \vec{a}$.

It is better to join into one list all the properties of the operations introduced above:

1) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$;

2) $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$;

3) $\vec{a} + \vec{0} = \vec{a}$;

4) $\vec{a} + (-\vec{a}) = \vec{0}$;

5) $(\alpha\beta)\vec{a} = \alpha(\beta\vec{a})$;

6) $(\alpha + \beta)\vec{a} = \alpha\vec{a} + \beta\vec{a}$;

7) $\alpha(\vec{a} + \vec{b}) = \alpha\vec{a} + \alpha\vec{b}$;

8) $1 \cdot \vec{a} = \vec{a}$.

A set of all vectors together with the operations of addition and multiplication by a scalar , is called ***a space of geometrical vectors.*** We can also speak in terms of that space about the linear dependence or independence of vectors, linear combination of vectors, etc.

Lemma If $\vec{\boldsymbol{a}} \neq \vec{\mathbf{0}}$ then the vectors $\vec{\boldsymbol{a}}$ и $\vec{\boldsymbol{b}}$ are collinear if and only if there is a scalar α such that $\vec{\boldsymbol{b}} = \alpha\vec{\boldsymbol{a}}$**.** (α is defined in a univalent way). We will not prove this statement.

Corollary Vectors $\vec{\boldsymbol{a}}$ и $\vec{\boldsymbol{b}}$ are collinear if and only if they are linearly dependent.

In a similar way we can state that 3 vectors are coplanar if and only if they are linearly dependent.

The last aspect to be investigated is the calculation of vectors' coordinates. Let us find the coordinates of a vector $\vec{\boldsymbol{a}} = \overrightarrow{AB}$, which starts at the point $A(x_A, y_A, z_A)$ and has an ending point $B(x_B, y_B, z_B)$. May the point O be the origin of coordinates. Then $\overrightarrow{OA}$ and $\overrightarrow{OB}$ are radius vectors and thus, $\overrightarrow{OA} = \{x_A, y_A, z_A\}$ and $\overrightarrow{OB} = \{x_B, y_B, z_B\}$. Then from the equality $\overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OB}$ we get $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA}$. As in the process of vector subtraction we subtract the corresponding coordinates of vectors, we get the formula:

$$\boxed{\overrightarrow{AB} = \{x_B - x_A, y_B - y_A, z_B - z_A\}}$$

So we subtract the 'starting' coordinates from the 'ending' coordinates.

Practice Russian

Ве́ктор – a vector

Сложе́ние ве́кторов – addition of vectors

Вычита́ние ве́кторов – subtraction of vectors

Умноже́ние ве́ктора на число – multiplication of a vector by a scalar

Review questions

1. Describe all the ways of vector addition
2. How do we get the coordinates of a vector $\overrightarrow{AB}$ if we have the coordinates of the points A and B?

Practical task

1. Find a vector $\overrightarrow{AB}$ with the given coordinates of A and B:

A	B
1. $\begin{pmatrix}1\\3\\5\end{pmatrix}$	$\begin{pmatrix}9\\15\\4\end{pmatrix}$
2. $\begin{pmatrix}10\\6\\-2\end{pmatrix}$	$\begin{pmatrix}8\\5\\14\end{pmatrix}$
3. $\begin{pmatrix}-3\\4\\7\end{pmatrix}$	$\begin{pmatrix}7\\-1\\12\end{pmatrix}$
4. $\begin{pmatrix}4\\-3\\5\end{pmatrix}$	$\begin{pmatrix}10\\-5\\7\end{pmatrix}$
5. $\begin{pmatrix}-9\\4\\-2\end{pmatrix}$	$\begin{pmatrix}16\\3\\19\end{pmatrix}$
6. $\begin{pmatrix}-1\\13\\6\end{pmatrix}$	$\begin{pmatrix}11\\1\\23\end{pmatrix}$
7. $\begin{pmatrix}-5\\2\\1\end{pmatrix}$	$\begin{pmatrix}8\\-7\\5\end{pmatrix}$
8. $\begin{pmatrix}8\\-13\\16\end{pmatrix}$	$\begin{pmatrix}9\\15\\-10\end{pmatrix}$
9. $\begin{pmatrix}7\\-9\\15\end{pmatrix}$	$\begin{pmatrix}7\\-12\\-5\end{pmatrix}$
10. $\begin{pmatrix}7\\-14\\9\end{pmatrix}$	$\begin{pmatrix}8\\-22\\-6\end{pmatrix}$
11. $\begin{pmatrix}9\\8\\-7\end{pmatrix}$	$\begin{pmatrix}4\\-11\\31\end{pmatrix}$
12. $\begin{pmatrix}-10\\-13\\9\end{pmatrix}$	$\begin{pmatrix}-7\\-16\\2\end{pmatrix}$

13. $\begin{pmatrix} 6 \\ -2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 9 \\ -8 \\ 7 \end{pmatrix}$
14. $\begin{pmatrix} 9 \\ 2 \\ -1 \end{pmatrix}$	$x = \begin{pmatrix} 20 \\ -1 \\ 4 \end{pmatrix}$
15. $\begin{pmatrix} -1 \\ 12 \\ 7 \end{pmatrix}$	$x = \begin{pmatrix} 7 \\ -8 \\ 3 \end{pmatrix}$

2.2. Scalar product of vectors

Let us remind you two definitions:

The length (the modulus) of a vector is the numeric value of a given vector. In 3-dimensional space for the length of a vector $\vec{a}$ $\{a_x, a_y, a_z\}$ we have a formula:

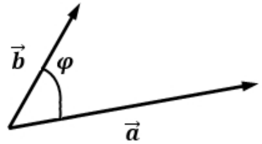
$$|\vec{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}.$$

The angle between two vectors $\vec{a}$ and $\vec{b}$ is a minimal value from two angles that we get when we lead the vector $\vec{a}$ and the vector $\vec{b}$ to a common starting point. In our explanations we would denote this angle with the letter φ.

A scalar product of vector $\vec{a}$ and vector $\vec{b}$ is a number which is denoted by $(\vec{a}, \vec{b})$ and equal to

$$(\vec{a}, \vec{b}) = |\vec{a}||\vec{b}| \cos \varphi,$$

where $0 \leq \varphi \leq \pi$ is an angle between the vectors $\vec{a}$ and $\vec{b}$, if both of the vectors are not equal to zero. In all the other cases the scalar product equals zero.



Remark ***The result of a vector product is always a number, not a vector!***

The vectors $\vec{a}$ и $\vec{b}$ are called ***the orthogonal vectors*** if their scalar product equals zero: $(\vec{a}, \vec{b}) = 0$.

Remark **If at least one of the vectors is a zero vector then the vectors are always orthogonal.**

Vectors' orthogonality criterion

The vectors $\vec{a} \neq \vec{0}$ and $\vec{b} \neq \vec{0}$ are orthogonal if and only if the angle between them equals $\varphi = 90^0$.

Proof. If $\varphi = 90^0$, then $\cos 90^0 = 0 \Rightarrow (\vec{a}, \vec{b}) = 0$, so the vectors are orthogonal. And vice versa, if $(\vec{a}, \vec{b}) = 0$ then $|\vec{a}||\vec{b}| \cos\varphi = 0$. However, according to the theorem, $|\vec{a}||\vec{b}| \neq 0$. Thus, $\cos\varphi = 0 \Rightarrow \varphi = 90^0$ (we mean that $0 \leq \varphi \leq 180^0$).

The main properties of a scalar multiplication:

1) $(\vec{a}, \vec{b}) = (\vec{b}, \vec{a})$ – commutativity;

2) $(\vec{a}, \vec{b} + \vec{c}) = (\vec{a}, \vec{b}) + (\vec{a}, \vec{c})$ – linearity (in terms of addition);

3) $(\alpha\vec{a}, \vec{b}) = \alpha(\vec{a}, \vec{b})$ – linearity (in terms of multiplication);

4) $(\vec{a}, \vec{a}) \geq 0$, considering $(\vec{a}, \vec{a}) = 0 \Leftrightarrow \vec{a} = \vec{0}$ (positive definiteness and nondegeneracy)

One-minute task **Try to prove the properties 1 and 4.**

If we look at the vectors $\vec{a}\ \{a_x, a_y, a_z\}$ and $\vec{b}\{b_x, b_y, b_z\}$:

$$\boxed{(\vec{a}, \vec{b}) = a_x b_x + a_y b_y + a_z b_z}. \qquad (*)$$

From this formula we take many important formulae.

Firstly, as we have already learned, $(\vec{a}, \vec{a}) = |\vec{a}|^2$, and then

$$\boxed{|\vec{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}}.$$

As we have already learnt, the formula above describes the ***length of a vector.***

This expression comes from (*) in the case when $a_x = b_x$, $a_y = b_y$, $a_z = b_z$. If φ is an angle between two vectors, then from the definition: $(\vec{a}, \vec{b}) = |\vec{a}||\vec{b}|\cos\varphi$ and formula (*) we get:

$$\boxed{\cos\varphi = \frac{a_x b_x + a_y b_y + a_z b_z}{\sqrt{a_x^2 + a_y^2 + a_z^2}\sqrt{b_x^2 + b_y^2 + b_z^2}}}.$$

And the vectors' orthogonality condition is put down as:

$$\boxed{a_x b_x + a_y b_y + a_z b_z = 0}.$$

Practice Russian

Скалярное произведе́ние ве́кторов – scalar product of vectors

Review questions

1. How can we find an angle between two vectors (using the material of this paragraph)?
2. How can we understand that two vectors are orthogonal?
3. Name all the properties of a scalar product for vectors

Practical tasks

1. Give at least three examples of the orthogonal vectors.

2. Find the cosine of an angle between 2 vectors

1. $\vec{a}\{1, 0, 3\}$; $\vec{b}\{2, 4, -1\}$
2. $\vec{a}\{-1, 2, 5\}$; $\vec{b}\{7, 3, 0\}$
3. $\vec{a}\{4, 5, 3\}$; $\vec{b}\{-6, 5, 1\}$
4. $\vec{a}\{2, 8, 10\}$; $\vec{b}\{9, 2, -3\}$
5. $\vec{a}\{1, 4, 9\}$; $\vec{b}\{-2, 7, 3\}$
6. $\vec{a}\{-8, 10, 4\}$; $\vec{b}\{12, -6, 5\}$
7. $\vec{a}\{-1, 9, 7\}$; $\vec{b}\{3, -4, 11\}$
8. $\vec{a}\{3, 0, 8\}$; $\vec{b}\{4,-4,0\}$
9. $\vec{a}\{6,7,2\}$; $\vec{b}\{3,9,-5\}$
10. $\vec{a}\{0, 0, 17\}$; $\vec{b}\{16, 9, -7\}$
11. $\vec{a}\{-5, 6, -7\}$; $\vec{b}\{-9, 14, 8\}$
12. $\vec{a}\{7, 4, 5\}$; $\vec{b}\{10, -4, 9\}$
13. $\vec{a}\{3, 4, 6\}$; $\vec{b}\{12, 9, -10\}$
14. $\vec{a}\{5, 5, 4\}$; $\vec{b}\{-1, 2, -8\}$
15. $\vec{a}\{11, 9, -7\}$; $\vec{b}\{7, 8, 9\}$

2.3. Vector product

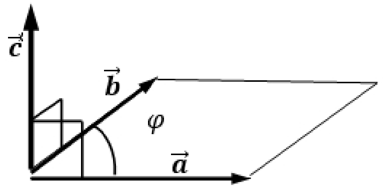
A vector product of a vector $\vec{a}$ by the vector $\vec{b}$ is a vector $\vec{c}$ which is denoted as $[\vec{a}, \vec{b}]$ and satisfies the following conditions:

a) $|\vec{c}| = |\vec{a}||\vec{b}| \sin\varphi,$ where φ is an angle between these two vectors, $0 \le \varphi \le \pi$;

b) $\vec{c} \perp \vec{a}$ and $\vec{c} \perp \vec{b}$ if $\vec{a}$ and $\vec{b}$ are not collinear;

c) If the condition b) is fulfilled then we see the rotation from the end of vector $\vec{c}$ from the vector $\vec{a}$ to the vector $\vec{b}$ by the angle φ going counter clock wright.

Remark As in the case of a scalar product, if the vector $\vec{a}$ or/and the vector $\vec{b}$ equals $\vec{0}$, then the angle φ is not defined and can be considered to be equal to any unspecified number.



The criterion of vectors' collinearity

The vectors $\vec{a}$ and $\vec{b}$ are collinear if and only if $[\vec{a}, \vec{b}] = \vec{0}.$

Proof Let us firstly introduce $\vec{c} = [\vec{a}, \vec{b}] = \vec{0}.$ Then $|\vec{c}| = |\vec{a}||\vec{b}| \sin\varphi = 0.$ So we have $|\vec{a}| = 0$ or $|\vec{b}| = 0,$ or $\sin\varphi = 0.$ Thus, $\vec{a} = \vec{0},$ or $\vec{b} = \vec{0},$ or $\varphi = 0$ or $\varphi = \pi$. In all the cases we get the same conclusion that the vectors $\vec{a}$ and $\vec{b}$ *are collinear*. Vice versa, let us suppose that vectors are collinear. In this case we have a zero vector among them, and then $|\vec{c}| = 0$ and $\vec{c} = \vec{0},$ or $\varphi = 0$ or $\varphi = \pi,$ and then $\sin\varphi = 0 \Rightarrow |\vec{c}| = 0 \Rightarrow \vec{c} = \vec{0}.$ ■

From this criterion $[\vec{a}, \vec{a}] = \vec{0}.$

In the picture which illustrates the definition of a vector product, a parallelogram is depicted. It is clear that if we denote its total area as S then

$$\boxed{|\vec{c}| = S}.$$

This property is called ***<u>the geometrical sense of a vector product modulus.</u>***

There are some useful expressions coming out from the definition.

$$[\vec{\boldsymbol{i}},\vec{\boldsymbol{j}}] = \vec{\boldsymbol{k}}, \quad [\vec{\boldsymbol{j}},\vec{\boldsymbol{k}}] = \vec{\boldsymbol{i}}, \quad [\vec{\boldsymbol{k}},\vec{\boldsymbol{i}}] = \vec{\boldsymbol{j}}, \quad [\vec{\boldsymbol{j}},\vec{\boldsymbol{i}}] = -\vec{\boldsymbol{k}},$$
$$[\vec{\boldsymbol{k}},\vec{\boldsymbol{j}}] = -\vec{\boldsymbol{i}}, \quad [\vec{\boldsymbol{i}},\vec{\boldsymbol{k}}] = -\vec{\boldsymbol{j}}.$$

<u>The main properties of a vector product</u>

1) $[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}] = -[\vec{\boldsymbol{b}},\vec{\boldsymbol{a}}]$ – (anticommutativity);

2) $[\alpha\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}] = \alpha[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}]$ – (linearity of multiplication by a scalar);

3) $[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}+\vec{\boldsymbol{c}}] = [\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}] + [\vec{\boldsymbol{a}},\vec{\boldsymbol{c}}]$ – (linearity of addition).

<u>Remark</u> The property **1)** makes the properties **2)** и **3)** correct for the cofactor $\vec{\boldsymbol{b}}$ as well. So $\alpha[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}] = -\alpha[\vec{\boldsymbol{b}},\vec{\boldsymbol{a}}] = -[\alpha\vec{\boldsymbol{b}},\vec{\boldsymbol{a}}] = [\vec{\boldsymbol{a}},\alpha\vec{\boldsymbol{b}}]$, i.e. $[\vec{\boldsymbol{a}},\alpha\vec{\boldsymbol{b}}] = \alpha[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}]$.

From these properties we can make a formula for the vector product via the coordinates in the Cartesian coordinate system. Let us present

$$\vec{\boldsymbol{a}} = a_x\vec{\boldsymbol{i}} + a_y\vec{\boldsymbol{j}} + a_z\vec{\boldsymbol{k}} \quad \text{and} \quad \vec{\boldsymbol{b}} = b_x\vec{\boldsymbol{i}} + b_y\vec{\boldsymbol{j}} + b_z\vec{\boldsymbol{k}}.$$

Here the vectors $\vec{i},\vec{j},\vec{k}$ are the ***unit vectors – the vectors that have a length equal to 1. They are used to show the directions in space (it is mostly used when we have the collinear vectors: the unit vectors show their directions).*** The formula for calculation of the unit vectors is

$$\hat{x} = \frac{\vec{x}}{|\vec{x}|},$$

where $\hat{x}$ is a unit vector, which is collinear to the given vector $\vec{x}$.

Then the following chain of equalities is true:

$$[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}] = [a_x\vec{\boldsymbol{\imath}} + a_y\vec{\boldsymbol{\jmath}} + a_z\vec{\boldsymbol{k}},\ b_x\vec{\boldsymbol{\imath}} + b_y\vec{\boldsymbol{\jmath}} + b_z\vec{\boldsymbol{k}}] =$$
$$= a_x b_x[\vec{\boldsymbol{\imath}},\vec{\boldsymbol{\imath}}] + a_x b_y[\vec{\boldsymbol{\imath}},\vec{\boldsymbol{\jmath}}] + a_x b_z[\vec{\boldsymbol{\imath}},\vec{\boldsymbol{k}}] +$$
$$+ a_y b_x[\vec{\boldsymbol{\jmath}},\vec{\boldsymbol{\imath}}] + a_y b_y[\vec{\boldsymbol{\jmath}},\vec{\boldsymbol{\jmath}}] + a_y b_z[\vec{\boldsymbol{\jmath}},\vec{\boldsymbol{k}}] + a_z b_x[\vec{\boldsymbol{k}},\vec{\boldsymbol{\imath}}] +$$
$$+ a_z b_y[\vec{\boldsymbol{k}},\vec{\boldsymbol{\jmath}}] + a_z b_z[\vec{\boldsymbol{k}},\vec{\boldsymbol{k}}] =$$
$$= (a_y b_z - a_z b_y)\vec{\boldsymbol{\imath}} - (a_x b_z - a_z b_x)\vec{\boldsymbol{\jmath}} + (a_x b_y - a_y b_x)\vec{\boldsymbol{k}}.$$

It’s not so hard to conclude that the last expression in the chain above can be presented in a following way:

$(a_y b_z - a_z b_y)\vec{\boldsymbol{\imath}} - (a_x b_z - a_z b_x)\vec{\boldsymbol{\jmath}} + (a_x b_y - a_y b_x)\vec{\boldsymbol{k}} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix}\vec{\boldsymbol{\imath}} - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix}\vec{\boldsymbol{\jmath}} + + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix}\vec{\boldsymbol{k}}$, and that coincides with evaluating the determinant by the first row

$$\begin{vmatrix} \vec{\boldsymbol{\imath}} & \vec{\boldsymbol{\jmath}} & \vec{\boldsymbol{k}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}.$$

Thus, finally we have

$$\boxed{[\vec{\boldsymbol{a}},\vec{\boldsymbol{b}}] = \begin{vmatrix} \vec{\boldsymbol{\imath}} & \vec{\boldsymbol{\jmath}} & \vec{\boldsymbol{k}} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}}.$$

Remark **We shouldn’t be afraid of the fact that we have vectors standing in the first row. We multiply them by numbers and not by each other.**

Practice Russian

Ве́кторное произведе́ние – a vector product

Единичный орт – a unit vector

Review questions

1. Formulate the properties of a vector product
2. How do we calculate the vector product if the vectors are given in the coordinate form?
3. What is a result of the vector product – a vector or a number?
4. What is a criterion for vectors’ collinearity?

Practical task

1. Find the vector product of vectors $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{b}}$ given in the practical task 2 of the previous chapter.

2.4. Triple scalar product

We have already got acquainted with both ways of multiplication for the case of two vectors. How should we act when we have three of them? The answer is hidden in the ***triple scalar product.***

A triple scalar product of 3 vectors $\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}$ is a scalar value

$$V = (\vec{\boldsymbol{a}}, [\vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}])$$

If we have the following expressions in the Cartesian system:

$$\vec{\boldsymbol{a}} = a_x\vec{\boldsymbol{\imath}} + a_y\vec{\boldsymbol{\jmath}} + a_z\vec{\boldsymbol{k}}, \quad \vec{\boldsymbol{b}} = b_x\vec{\boldsymbol{\imath}} + b_y\vec{\boldsymbol{\jmath}} + b_z\vec{\boldsymbol{k}}, \quad \vec{\boldsymbol{c}} = c_x\vec{\boldsymbol{\imath}} + c_y\vec{\boldsymbol{\jmath}} + c_z\vec{\boldsymbol{k}},$$

then we get the following formula for V:

$$V = a_x\begin{vmatrix} b_y & b_z \\ c_y & c_z \end{vmatrix} - a_y\begin{vmatrix} b_x & b_z \\ c_x & c_z \end{vmatrix} + a_z\begin{vmatrix} b_x & b_y \\ c_x & c_y \end{vmatrix} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}.$$

So we have the following determinant.

$$\boxed{V = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}}.$$

Let us calculate this determinant by the third row:

$$V = c_x\begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} - c_y\begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} + c_z\begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} = ([\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}], \vec{\boldsymbol{c}}).$$

This equality shows that $(\vec{\boldsymbol{a}}, [\vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}]) = ([\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}], \vec{\boldsymbol{c}})$ and putting the 'square' brackets inside the 'round' ones doesn't influence the result. So we have finally:

$$V = \vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}}.$$

Vectors' coplanarity criterion

The three vectors $\vec{a},\ \vec{b},\ \vec{c}$ are coplanar if and only if their triple scalar product equals zero: $\vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} = 0$.

Proof If the vectors $\vec{\boldsymbol{a}},\ \vec{\boldsymbol{b}},\ \vec{\boldsymbol{c}}$ are coplanar, then they are linearly dependent $\Rightarrow$ then the rows of the determinant are also linearly dependent $\Rightarrow$ $\vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} = 0$. Vice versa, if $\vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} = 0$, the rank of the determinant's matrix is less than 3 $\Rightarrow$ its rows are linearly dependent $\Rightarrow$ the vectors $\vec{\boldsymbol{a}},\ \vec{\boldsymbol{b}},\ \vec{\boldsymbol{c}}$ are linearly dependent $\Rightarrow$ these vectors are coplanar.

The vectors $\vec{\boldsymbol{a}},\ \vec{\boldsymbol{b}},\ \vec{\boldsymbol{c}}$ are called the arguments of triple scalar product

The properties of the triple scalar product can be easily got out of the determinants' properties:

1) **Linearity for each argument.** For example,

$$(\vec{\boldsymbol{a}}_1 + \vec{\boldsymbol{a}}_2)\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} = \vec{\boldsymbol{a}}_1\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} + \vec{\boldsymbol{a}}_2\vec{\boldsymbol{b}}\vec{\boldsymbol{c}},$$

For each scalar α the following expression is true:

$$(\alpha\vec{\boldsymbol{a}})\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} = \vec{\boldsymbol{a}}\left(\alpha\vec{\boldsymbol{b}}\right)\vec{\boldsymbol{c}} = \vec{\boldsymbol{a}}\vec{\boldsymbol{b}}(\alpha\vec{\boldsymbol{c}}) = \alpha(\vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}});$$

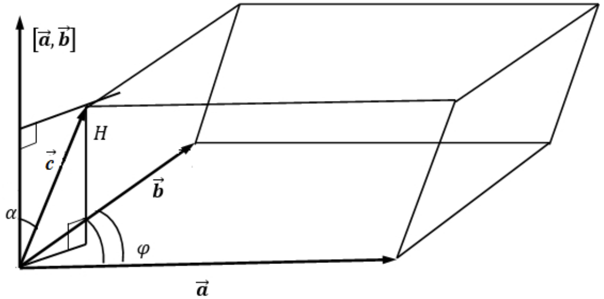
2) **Anticommutativity for each argument.** For example,

$$\vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}} = -\vec{\boldsymbol{b}}\vec{\boldsymbol{a}}\vec{\boldsymbol{c}} = \vec{\boldsymbol{b}}\vec{\boldsymbol{c}}\vec{\boldsymbol{a}} = -\vec{\boldsymbol{c}}\vec{\boldsymbol{b}}\vec{\boldsymbol{a}} = \cdots \text{ , etc.}$$

The geometrical sense of a triple scalar product

Let the vectors $\vec{\boldsymbol{a}},\ \vec{\boldsymbol{b}},\ \vec{\boldsymbol{c}}$ be non-coplanar and let them be put from one starting point A. Let us denote: V_p is the cubage of the parallelepiped that is based on these vectors. Then

$$V_p = \left|\vec{\boldsymbol{a}}\vec{\boldsymbol{b}}\vec{\boldsymbol{c}}\right|.$$



Indeed, $V_p = S_p H$, where S_p is a total area of a parallelogram lying in the basement of a parallelepiped, and H is its altitude. On the other hand, from geometrical properties we have: $H = |\vec{c}||\cos\alpha|$ and $S_p = |\vec{a}||\vec{b}|\sin\varphi$ (we took $|\cos\alpha|$, not $\cos\alpha$ – considering the fact that α can be also more then 90̊, and, however, the length of an altitude should be a positive number). Thus, $V_p = |\vec{a}||\vec{b}|\sin\varphi \cdot |\vec{c}||\cos\alpha| = |[\vec{a},\vec{b}]| \cdot |\vec{c}||\cos\alpha| = \left||[\vec{a},\vec{b}]| \cdot |\vec{c}|\cos\alpha\right| = |([\vec{a},\vec{b}],\vec{c})| = |\vec{a}\vec{b}\vec{c}|$.

Remark **In the case when $\vec{a}$, $\vec{b}$, $\vec{c}$, are coplanar, the parallelepiped transforms into a flat figure and we can't speak about its cubage. However, if we say that in this case its cubage is equal to zero, then the formula $V_p = |\vec{a}\vec{b}\vec{c}|$ is also true.**

Practice Russian

Сме́шанное произведе́ние – a triple scalar product

Объём – cubage

Пло́щадь – total area

Review questions

1. How do we calculate the triple scalar product?
2. What are the main properties of a triple scalar product?
3. What is the main geometrical application for a triple scalar product?

Practical task

1. Find the cubage of a parallelepiped based on the following vectors:

1. $\vec{a}$ {1, 0,3}; $\vec{b}${2, 4, –1}; $\vec{c}${–1, 2, 5}
2. $\vec{a}$ {–1, 2,5}; $\vec{b}${7, 3, 0}; $\vec{c}$ {4, 5,3}
3. $\vec{a}$ {4, 5,3}; $\vec{b}${–6, 5, 1}; $\vec{c}${2, 8, 10}
4. $\vec{a}$ {2, 8, 10}; $\vec{b}${9, 2, –3}; $\vec{c}$ {1,4,9}
5. $\vec{a}$ {1, 4, 9}; $\vec{b}${–2, 7, 3}; $\vec{c}$ {–8, 10, 4}
6. $\vec{a}${–8, 10, 4}; $\vec{b}${12, –6, 5}; $\vec{c}${–1, 9, 7}
7. $\vec{a}${–1, 9, 7}; $\vec{b}${3, –4, 11}; $\vec{c}${3, 0, 8}
8. $\vec{a}${3, 0, 8}; $\vec{b}${4, –4, 0}; $\vec{c}${6, 7, 2}
9. $\vec{a}${6, 7, 2}; $\vec{b}${3, 9, –5}; $\vec{c}$ {0, 0, 17}
10. $\vec{a}${0, 0, 17}; $\vec{b}${16, 9, –7}; $\vec{c}${–5, 6, –7}

11. $\vec{\boldsymbol{a}}\{-5, 6, -7\}$; $\vec{\boldsymbol{b}}\{-9, 14, 8\}$; $\vec{\boldsymbol{c}}$ $\{7, 4, 5\}$
12. $\vec{\boldsymbol{a}}$ $\{7, 4, 5\}$; $\vec{\boldsymbol{b}}\{10, -4, 9\}$; $\vec{\boldsymbol{c}}$ $\{3, 4, 6\}$;
13. $\vec{\boldsymbol{a}}\{3, 4, 6\}$; $\vec{\boldsymbol{b}}\{12, 9, -10\}$; $\vec{\boldsymbol{c}}$ $\{5, 5, 4\}$
14. $\vec{\boldsymbol{a}}\{5, 5, 4\}$; $\vec{\boldsymbol{b}}\{-1, 2, -8\}$; $\vec{\boldsymbol{c}}$ $\{11, 9, -7\}$
15. $\vec{\boldsymbol{a}}\{11, 9, -7\}$; $\vec{\boldsymbol{b}}\{7, 8, 9\}$, $\vec{\boldsymbol{c}}$ $\{4, 5, 3\}$

2.5. Planes. Plane equation, mutual location of planes

We would consider that the Cartesian system $Oxyz$ is fixed in space, and the coordinates of points, vectors, etc. , are given in that system. Let us investigate a plane P. Let the vector $\vec{\boldsymbol{n}} = \{A,\ B,\ C\} \neq \vec{\boldsymbol{0}}$ be orthogonal to that plane (see the picture) – it is usually called the normal vector. Let us introduce the point $M_0(x_0, y_0, z_0) \in P$. In this case it's obvious that the point $M(x, y, z) \in P \Leftrightarrow \vec{\boldsymbol{n}} \perp \overrightarrow{M_0M} \Leftrightarrow \left(\vec{\boldsymbol{n}},\ \overrightarrow{M_0M}\right) = 0$. If we put the last equality in the coordinate form, we get

$$\boxed{A(x - x_0) + B(y - y_0) + C(z - z_0) = 0}. \qquad (*)$$

The equation $(*)$ is called the equation of a plane going through the point $M_0(x_0, y_0, z_0)$. So only the coordinates of the plane P satisfy the equation (*) (and they're unique!)

The inverse statement is also true: for any unspecified numbers x_0, y_0, z_0, and for the numbers A, B, C *(if they are not equal to zero at the same time!),* a set of points that satisfy the equation (*), fills the plane coming through the point $M_0(x_0, y_0, z_0)$ and orthogonal of the vector $\vec{\boldsymbol{n}} = \{A,\ B,\ C\}$.

We open the brackets in the equation (*) and get

$$Ax + By + Cz - Ax_0 - By_0 - Cz_0 = 0.$$

Let us introduce $D = -Ax_0 - By_0 - Cz_0$. Then the equation $(*)$ can be put down in the form

$$Ax + By + Cz + D = 0.$$

And the inverse statement: if we have the equation $Ax + By + Cz + D = 0$, where not all of the numbers A, B, C are equal to zero. First of all, this equation has solutions. Indeed, let us suppose $A \neq 0$. Then, for example, the numbers $x = -D/A$, $y = 0$, $z = 0$ satisfy this equation. The other similar cases 0 you can investigate yourself. And now let the three numbers x_0, y_0, z_0 be its solution. That means $Ax_0 + By_0 + + Cz_0 + D = 0$. In other words, $D = -Ax_0 - By_0 - Cz_0$. We put the value of D into the equation and get: $Ax + By + Cz - Ax_0 - By_0 - Cz_0 = 0$. After getting the common factors we put down $A(x - x_0) + + B(y - y_0) + C(z - z_0) = 0$, i.e. the equation of a plane going through the point $M_0(x_0, y_0, z_0)$ *orthogonally to the vector* $\vec{\boldsymbol{n}} = \{A,\ B,\ C\}$.

So any plane can be given by the equation $Ax + By + Cz + D = 0$. And vice versa, any equation of a type $Ax + By + Cz + D = 0$ with a condition that not all from the numbers A, B, C are equal to zero, defines a plane in space.

$$Ax + By + Cz + D = 0, \text{ where } A^2 + B^2 + C^2 \neq 0$$

So that is the reason why the equation is usually called a general plane equation. The expression $A^2 + B^2 + C^2 \neq 0$ is a short version of the fact that not all of the numbers A, B, C are equal to zero.

At last, we once again remind you the geometrical sense of the coefficients A, B, C: the vector $\vec{n} = \{A,\ B,\ C\}$ is orthogonal to the $Ax + By + Cz + D = 0$.

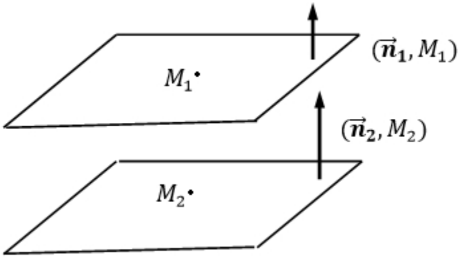
Remark A vector $\vec{\boldsymbol{n}}$ is also called a normal vector to the plane P. It is clear that it doesn't matter which vector should be chosen for the plane equation. All those vectors are collinear and, therefore, they are orthogonal to the given plane. It also has its reflection in the fact that if we multiply the plane equation by any non-zero number , we come to the equivalent equation,i.e. to the same equation of that same plane.

As the plane is given in a unique way by its normal vector and a point, in further explanations we would denote a plane as $(\vec{n}, M)$.

Mutual location of 2 planes in space

There are only two variants of mutual location for 2 planes in space: intersection or parallel planes.

A pair of parallel planes



Let $\vec{\boldsymbol{n}}_1 = \{A_1, B_1, C_1\}$, $M_1(x_1, y_1, z_1)$ and $\vec{\boldsymbol{n}}_2 = \{A_2, B_2, C_2\}$, $M_2(x_2, y_2, z_2)$.

Then $(\vec{\boldsymbol{n}}_1, M_1) \parallel (\vec{\boldsymbol{n}}_2, M_2) \Leftrightarrow \vec{\boldsymbol{n}}_1$ and $\vec{\boldsymbol{n}}_2$ are collinear $\Leftrightarrow$ there is to be found such a number α that $\vec{\boldsymbol{n}}_2 = \alpha\vec{\boldsymbol{n}}_1 \Leftrightarrow A_2 = \alpha A_1, B_2 = \alpha B_1, C_2 = \alpha C_1$. If we remove α from those equalities, we finally get a proportion:

$$\boxed{\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1}}$$

The same condition can be put down in the other way: $[\vec{\boldsymbol{n}}_1, \vec{\boldsymbol{n}}_2] = \vec{\mathbf{0}}.$ It is also can be put down in the following form:

$$\text{Rg}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = 1.$$

However, there is also a case of coinciding planes. The conditions above are also fulfilled when the planes coincide. So, in fact, we deal with one plane. The equation for one of the planes has a form $A_1x + B_1y + C_1z + D_1 = 0$, for the second plane $\alpha A_1x + \alpha B_1y + \alpha C_1z + D_2 = 0$. Those planes coincide if, for example, the coordinates of the point M_1 satisfy both equations. That means, two equalities are true: $A_1x_1 + B_1y_1 + C_1z_1 + D_1 = 0$ and $\alpha A_1x_1 + \alpha B_1y_1 + \alpha C_1z_1 + D_2 = 0$. From the first equality we have $A_1x_1 + B_1y_1 + C_1z_1 = -D_1$. The second can be put down as $\alpha(A_1x_1 + B_1y_1 + C_1z_1) + D_2 = 0$. Thus, we get $-\alpha D_1 + + D_2 = 0 \Leftrightarrow D_2 = \alpha D_1$, i.e., $D_2/D_1 = \alpha$. Finally,

$$\boxed{\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1} = \frac{D_2}{D_1}}$$

is a criterion of *coinciding planes*. It is equivalent to the expression:

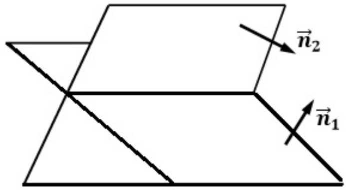
$$\mathrm{Rg}\begin{pmatrix} A_1 & B_1 & C_1 & D_1 \\ A_2 & B_2 & C_2 & D_2 \end{pmatrix} = 1.$$

If we have

$$\boxed{\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1} \neq \frac{D_2}{D_1}}$$

Then we get the criterion for the *parallel planes*.

Intersection of planes

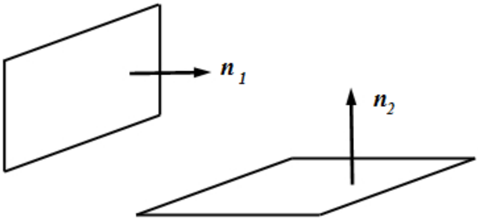


From the analysis above we can conclude that the criterion of 2 planes' intersection has the following form:

$$\mathrm{Rg}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = 2.$$

We should notice that the planes are perpendicular to each other in the case when their normal vectors are perpendicular to each other (and in that case only!) That means $(\vec{\boldsymbol{n}}_1, \vec{\boldsymbol{n}}_2) = 0$ or, in coordinate form,

$$\boxed{A_1A_2 + B_1B_2 + C_1C_2}$$



Practice Russian

Плóскость – a plane
Пересечéние – the intersection

Review questions

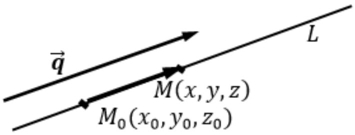
1. How does the general equation of a plane look like?
2. How many types of mutual locations does exist for two planes?

Practical task

1. Write an equation of plane that contains point A, if vector AB is a normal vector to that plane. Coordinates of points A and B are written in the task for a paragraph 2.1.

2.6. Straight line equation

Let us define a straight line L in space.



Any vector $\vec{\boldsymbol{q}} = \{l, m, n\} \neq \vec{\mathbf{0}}$, which is collinear to L, is called a direction vector of L.

Parametric straight line equations

The point $M_0(x_0, y_0, z_0) \in L$, and $\vec{\boldsymbol{q}} = \{l, m, n\} \neq \vec{\mathbf{0}}$ **is its direction vector**. Then the point $M(x, y, z) \in L$ so the vectors $\vec{\boldsymbol{q}}$ and $\overrightarrow{M_0M}$ are collinear**.** Thus, we can find such a number t_m, for which the following equality is true: $\overrightarrow{M_0M} = t_m\vec{\boldsymbol{q}}$.

On the other hand, it's obvious that if *t* is any unspecified number and $\overrightarrow{M_0M} = t\vec{\boldsymbol{q}}$**,** then the point $M \in L$. So the straight line L consists only of those points M, for which the following equality is fulfilled:

$$\boxed{\overrightarrow{M_0M} = t\vec{\boldsymbol{q}}},$$

Where the number t 'runs along' the complete number scale. This expression is called a vector form of the straight line equation.

We can put it down in the coordinate form:

$$\{x - x_0, y - y_0, z - z_0\} = \{tl, tm, tn\}.$$

If we set the coordinates of vectors in the right-hand part and left-hand part of this equality, equal to each other, we get $x - x_0 = tl$, $y - y_0 = tm$, $z - z_0 = tn$. Finally, we can put down these equalities as a system of equations

$$\boxed{\begin{cases} x = x_0 + tl, \\ y = y_0 + tm, \\ z = z_0 + tn, \end{cases}} \qquad t \in \mathbb{R},$$

And this system is called parametrical equations of a straight line. The title is given thanks to the variable t, which can be investigated as an independent parameter, which can take any number values. If we give a concrete number to the number t, e.g. $t = t_1$, we get three numbers $x_1 = x_0 + t_1 l$, $y_1 = y_0 + t_1 m$, $z_1 = z_0 + t_1 n$, and those numbers are the coordinates of the point $M_1(x_1, y_1, z_1)$ which lies on the straight line L. If we take all the values of t consequently, every time we will get the coordinates of the points lying on the straight line L and so we won't miss any of those points. So on the straight line L there are only those points whose coordinates satisfy the parametric equations with different values of t.

Canonic straight-line equations

We exclude the parameter t from the parametric equations of a straight line (in all of the three equations). So we get $t = (x - x_0)/l$, $t = (y - y_0)/m$, $t = (z - z_0)/n$. Thus, we have a double proportion

$$\boxed{\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}},$$

And only the coordinates of the points on the straight line L satisfy that proportion. This double proportion is called ***the canonic equations of a straight line.***

Remark **In the canonic equations of a straight line you can notice some expressions of the type** $(x - x_0)/0$, for example, when $l = 0$. We should undertand that there is no division by zero actually. We only mean that $x = x_0$. It is especially clear in the parametric form for the case when $l = 0$. In a geometrical sense, this situation means that all the straight line lies in the plane which is defined by the equation $x = x_0$.

General equation of a straight line

Let two planes be given in space:

$$P_1: \quad A_1x + B_1y + C_1z + D_1 = 0,$$
$$P_2: \quad A_2x + B_2y + C_2z + D_2 = 0.$$

Further we will see that all those planes are intersected by the common straight line and define that straight line only in the case when

$$\operatorname{rk}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = 2.$$

When this condition is fulfilled, then the point lies on the straight line of two planes' intersection if and only if the coordinates of the point satisfy both the equation of a plane P_1 and the equation of a plane P_2. In other words, they should satisfy the system

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases}$$

This system is called a general equation of a straight line. They also say that the straight line is defined by *a pair of intersecting planes.*

Remark **The common equation of a straight line contains a widest variety of combinations as *an endless amount* of planes can go through the given line. Any pair of them defines this straight line**.

Practice Russian

Прямáя – a straight line

Review questions

1. How do we formulate the parametric equation for a straight line?
2. How do we formulate the canonic equation for a straight line?
3. What is a general equation of a straight line and how do we get it?

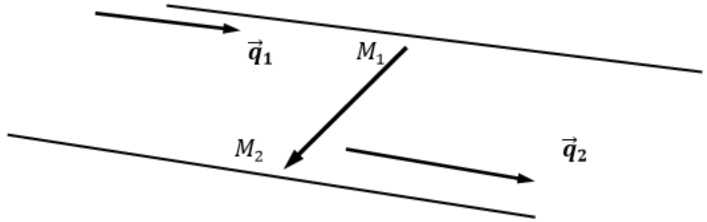
2.7. Mutual location of straight lines and planes in space

As a plane or a straight line are given by their normal vector (or a direction vector correspondently) and a point, then we would further denote a plane as $(\vec{\boldsymbol{n}}, M)$**,** and a straight line as $(\vec{\boldsymbol{q}}, M)$**.** (Probably, we will use some special indexes for them).

A pair of straight lines in space

Two straight lines in space can be parallel or they can intersect each other and they also can be crossed.

a) It is clear that the parallelism of two straight lines $(\vec{\boldsymbol{q}}_1, M_1)$ и $(\vec{\boldsymbol{q}}_2, M_2)$ **is** equivalent to the fact that their direction vectors $\vec{\boldsymbol{q}}_1$ and $\vec{\boldsymbol{q}}_2$ are collinear.



Let $\vec{\boldsymbol{q}}_1 = \{l_1, m_1, n_1\}$ and $\vec{\boldsymbol{q}}_2 = \{l_2, m_2, n_2\}$**.** The collinearity criterion can be presented in several equivalent forms:

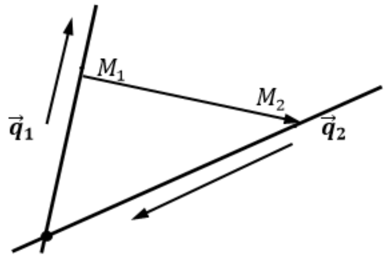
$$[\vec{\boldsymbol{q}}_1, \vec{\boldsymbol{q}}_2] = \vec{\mathbf{0}} \;\Leftrightarrow\; \frac{l_1}{l_2} = \frac{m_1}{m_2} = \frac{n_1}{n_2} \;\Leftrightarrow\; \mathrm{Rg}\begin{pmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \end{pmatrix} = 1.$$

If there are two parallel lines then we can choose for them a common direction vector $\vec{q} = \{l, m, n\}$, that is collinear to the vectors $\vec{q}_1$ and $\vec{q}_2$ or it can even coincide with one of them. Then all the conditions are fulfilled in a trivial way. We should only investigate the question about coinciding or non-coinciding straight lines. However, it is absolutely obvious that $\left[\vec{\boldsymbol{q}}, \overrightarrow{M_1M_2}\right] = \vec{\mathbf{0}}$ – is a condition for *coinciding straight lines*, and $\left[\vec{\boldsymbol{q}}, \overrightarrow{M_1M_2}\right] \neq \vec{\mathbf{0}}$ is a criterion for *parallel lines*.

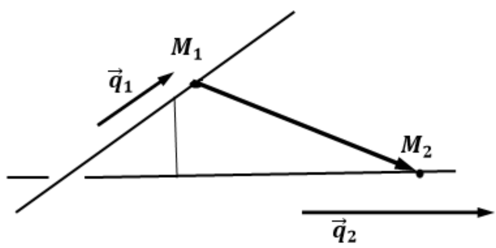
B) Two straight lines intersect only in the case when $\vec{\boldsymbol{q}}_1$**,** $\vec{\boldsymbol{q}}_2$ and $\overrightarrow{M_1M_2}$ are coplanar but $\vec{\boldsymbol{q}}_1$ and $\vec{\boldsymbol{q}}_2$ are non-collinear. It means that triple scalar product is equal to zero: $\vec{\boldsymbol{q}}_1\vec{\boldsymbol{q}}_2\overrightarrow{M_1M_2} = 0$ but $[\vec{\boldsymbol{q}}_1, \vec{\boldsymbol{q}}_2] \neq \vec{\mathbf{0}}$.

With the help of logical element ∧ (we read it as 'and'), we can put down this condition in the following form:

$$\left(\vec{q}_1\vec{q}_2\overrightarrow{M_1M_2}=0\right)\wedge\left([\vec{q}_1,\vec{q}_2]\neq\vec{0}\right)$$



C) Let us remind that two straight lines are called ***the crossing lines*** if they are not parallel and they don't intersect each other.



We can deny either parallelism or intersection by a unique condition:

$$\vec{q}_1\vec{q}_2\overrightarrow{M_1M_2}\neq 0$$

Indeed, this condition guarantees the impossibility of intersection as for the intersection we need the following formula to be fulfilled: $\vec{q}_1\vec{q}_2\overrightarrow{M_1M_2}=0$. And, besides, from the expression comes that the vectors $\vec{q}_1$ and $\vec{q}_2$ are non-collinear and so the straight lines can't be parallel.

Remark **The condition for two straight lines to be parallel in space: it is necessary and sufficient for their direction vectors to be perpendicular to each other: $(\vec{q}_1,\vec{q}_2)=0$ or**

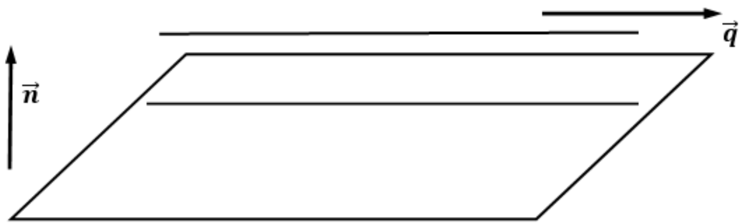
$$l_1l_2+m_1m_2+n_1n_2=0.$$

Straight lines and planes in space

A straight line in space is either parallel to a plane, or intersects a plane.

A parallelism means that $\vec{\boldsymbol{q}} = \{l, m, n\} \perp \vec{\boldsymbol{n}} = \{A, B, C\}$. We can tell in other words, that

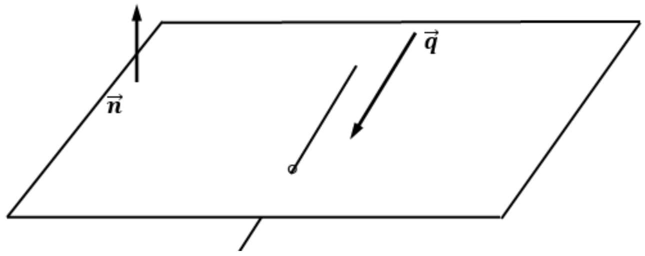
$$\boxed{Al + Bm + Cn = 0}$$



That also includes a private case when the whole straight line is situated inside the plane.

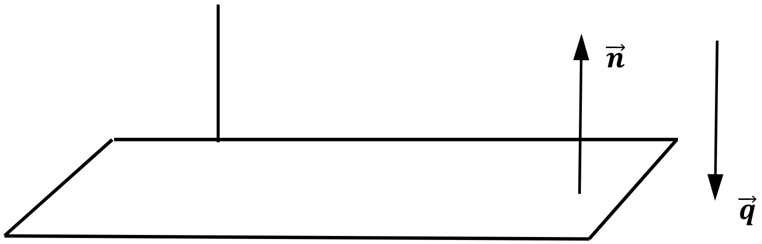
The intersection is, in fact, denying of a parallelism. That's why we can put down the criterion of a parallelism in a following form:

$$\boxed{Al + Bm + Cn \neq 0}.$$



We can tell that the criterion for perpendicularity of a straight line and a plane is a fact of collinearity of vectors $\vec{\boldsymbol{q}}$ and $\vec{\boldsymbol{n}}$, and that can be written in the form of a double proportion:

$$\boxed{\frac{A}{l} = \frac{B}{m} = \frac{C}{n}}$$



2.8. Useful applications

In this section we would give the basic examples of solving problems which investigate straight lines and planes in space.

1) There are two given points $M_1(x_1, y_1, z_1)$ and $M_2(x_2, y_2, z_2)$. Put down the canonic equation of a straight line that comes via these points.

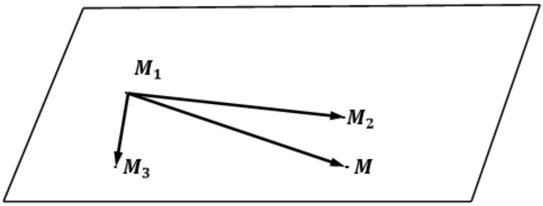
Solution As both points M_1 and M_2 lie on one straight line, then the vector $\overrightarrow{M_1M_2} = \{x_2 - x_1, y_2 - y_1, z_2 - z_1\}$ can be defined as the direction vector. We can choose any of these two points as a point lying on a line. For example, we take M_1.

So we get

$$\boxed{\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}}$$

2) Put down the equation of a plane which goes via the three given points $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, $M_3(x_3, y_3, z_3)$ (the points don't lie on one straight line).

Solution The point $M(x, y, z)$ belongs to a plane which is defined by the three given points if and only if the vectors $\overrightarrow{M_1M_2} = \{x_2 - x_1,\ y_2 - y_1,\ z_2 - z_1\}$, $\overrightarrow{M_1M_3} = \{x_3 - x_1,\ y_3 - y_1,\ z_3 - z_1\}$ and $\overrightarrow{M_1M} = \{x - x_1,\ y - y_1,\ z - z_1\}$ are coplanar.



Thus, $\overrightarrow{M_1M_2}\,\overrightarrow{M_1M_3}\,\overrightarrow{M_1M} = 0$, and that means the following determinant is equal to zero:

$$\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix} = 0.$$

If we calculate this determinant by the first row then we get ***the general plane equation*** after some simplifications

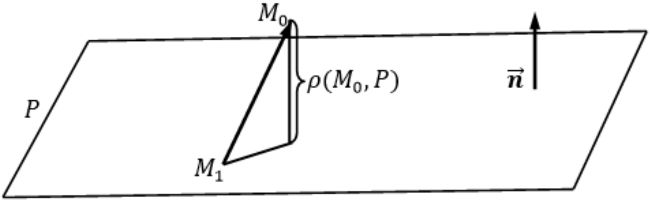
$$\boxed{\begin{vmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{vmatrix} = 0}$$

3) Find the distance between the point $M_0(x_0, y_0, z_0)$ and a plane

$$\text{P:}\quad Ax + By + Cz + D = 0.$$

Solution For the investigated distance let us introduce the denotation $\rho(M_0, P)$. Let now $M_1(x_1, y_1, z_1)$ be any unspecified point of a plane. Then we use the formula where $\vec{\boldsymbol{n}} = \{A, B, C\}$ is a vector which is normal to the surface P.

$$\rho(M_0, P) = \frac{\left|\left(\vec{\boldsymbol{n}}, \overrightarrow{M_1M_0}\right)\right|}{|\vec{\boldsymbol{n}}|} = \frac{|A(x_0 - x_1) + B(y_0 - y_1) + C(z_0 - z_1)|}{\sqrt{A^2 + B^2 + C^2}}.$$



If we open the brackets in the numerator of this fraction and if we consider that $D = -Ax_1 - By_1 - Cz_1$, (as the point M_1 belongs to the plane), then we get the final result:

$$\rho(M_0, P) = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

4) Find the distance between parallel planes

Solution As the planes are parallel, then as a normal vector we can take the same vector $\vec{n} = \{A, B, C\}$. Then let the equations of planes be P_1: $Ax + By + Cz + D_1 = 0$ and P_2: $Ax + By + Cz + D_2 = 0$. We denote the distance between those planes as $\rho(P_1, P_2)$. This value is calculated as the distance between the unspecified point $M_0 \in P_1$ and the plane P_2. However,

$$\rho(M_0, P_2) = \frac{|Ax_0 + By_0 + Cz_0 + D_2|}{\sqrt{A^2 + B^2 + C^2}}$$

And as $M_0 \in P_1$, then $Ax_0 + By_0 + Cz_0 = -D_1$. If we put that into the formula above, we get:

$$\rho(P_1, P_2) = \frac{|D_2 - D_1|}{\sqrt{A^2 + B^2 + C^2}}$$

These are only a few algorithms for solving problems with straight lines and planes. However, in other cases it is much better to investigate the conditions of the problems thoroughly than to look for ready step-by-step solutions.

Creative task We have come to the end of the chapter. Show your creative talents again like you have already done after the end of previous chapter! Now your theme is ‘Analytical geometry’. You can choose any forms you like, just enjoy the process!

This task is not necessary – however, we hope you will have much joy and fun doing this task!

3. LINEAR SPACES AND OPERATORS

This chapter contains very important concepts of linear algebra. However, here you will meet less practical tasks and less review questions. Your main aim is to learn all the basic properties described in each paragraph. You should be ready to answer about them to your lecturer orally or in written form. This chapter contains a lot of definitions, properties and formulae so the authors leave the material for your self-control so you should learn to understand the necessary information, to analyze it and to prepare well-structured answers.

3.1. Linear space

A linear space $\mathscr{L}$ is a set of elements $\overline{\boldsymbol{a}}, \overline{\boldsymbol{b}}, \overline{\boldsymbol{c}}, \dots, \overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{z}}, \dots$ (the type of those elements is unspecified) for which we can introduce two operations: addition, which is denoted by the sign $+$, and multiplication by a scalar (usually denoted as $\cdot$ but very often we don't put any special signs for it, like, for example, in the expression *ab*).

We also have to check that $\forall \overline{\boldsymbol{a}} \in \mathscr{L}\ \forall \overline{\boldsymbol{b}} \in \mathscr{L} (\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}} \in \mathscr{L})$ and $\forall \alpha \in R\ \forall \overline{\boldsymbol{a}} \in \mathscr{L} (\alpha \overline{\boldsymbol{a}} \in \mathscr{L})$, i.e. we have the elements from the same space $\mathscr{L}$ as the results of both operations (it is called ***an operation closure***). Besides, the following ***axioms of linear space*** should be fulfilled for any elements $\overline{\boldsymbol{a}}, \overline{\boldsymbol{b}}, \overline{\boldsymbol{c}}$ and for any numbers α, β**:**

1) $\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}} = \overline{\boldsymbol{b}} + \overline{\boldsymbol{a}}$ – *the commutativity axiom;*

2) $(\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}}) + \overline{\boldsymbol{c}} = \overline{\boldsymbol{a}} + (\overline{\boldsymbol{b}} + \overline{\boldsymbol{c}})$ – *the associativity axiom;*

3) There is to be found such an element $\overline{\boldsymbol{\theta}}$, which is called a neutral element, that for each element $\overline{\boldsymbol{a}} \in \mathscr{L}$ the following equality is fulfilled: $\overline{\boldsymbol{a}} + \overline{\boldsymbol{\theta}} = \overline{\boldsymbol{a}}$;

4) For each element $\overline{\boldsymbol{a}}$ in the set $\mathscr{L}$ we can find the inverse element $\overline{\boldsymbol{b}}$, so that $\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}} = \overline{\boldsymbol{\theta}}$;

5) $(\alpha\beta)\overline{\boldsymbol{a}} = \alpha(\beta\overline{\boldsymbol{a}})$;

6) $(\alpha + \beta)\overline{\boldsymbol{a}} = \alpha\overline{\boldsymbol{a}} + \beta\overline{\boldsymbol{a}}$;

7) $\alpha(\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}}) = \alpha\overline{\boldsymbol{a}} + \alpha\overline{\boldsymbol{b}}$;

8) $1\overline{\boldsymbol{a}} = \overline{\boldsymbol{a}}$. ■

Remark 1 **The elements of linear spaces are often called vectors, and the space itself is sometimes called a vector space. The numbers are also sometimes called scalars.**

Remark 2 **We should notice at the very beginning that any set is not a linear space yet until the rules of addition and multiplication by a scalar are not defined for that space. This rules are, in fact, created by ourselves and they can be formulated in any unspecified way. However, it is strictly necessary for those rules to satisfy the conditions of closure and axioms of linear space.**

One-minute task Do you know any other sets (except the set of vectors that have just been given as an example) that can be called linear spaces?

Simple properties of linear spaces

Let $\mathscr{L}$ be a linear space. The properties formulated below, can seem obvious to us in the cases when we investigate some well-known objects, for example, geometrical vectors or matrices. However, we shouldn't forget that we work with an abstract linear space and all we have is 8 axioms that are true for each case of linear space. For example, we shouldn't look inattentively at the operation of addition and the sign '+'. In linear space it is just an action which can be denoted in different ways, depending on the case. For example, the sign for the operation can look like: , $*$; $\therefore$; $\otimes$; $\diamond$; $\oplus$, or in some other ways.

So let us come to the properties:

1) *There is only one neutral element in the linear space*

Indeed, if $\overline{\boldsymbol{\theta}}_{\mathbf{1}}$ and $\overline{\boldsymbol{\theta}}_{\mathbf{2}}$ are both neutral elements, then $\overline{\boldsymbol{\theta}}_{\mathbf{2}} + \overline{\boldsymbol{\theta}}_{\mathbf{1}} = \overline{\boldsymbol{\theta}}_{\mathbf{2}}$ as $\overline{\boldsymbol{\theta}}_{\mathbf{1}}$ is a neutral element; $\overline{\boldsymbol{\theta}}_{\mathbf{1}} + \overline{\boldsymbol{\theta}}_{\mathbf{2}} = \overline{\boldsymbol{\theta}}_{\mathbf{1}}$ because $\overline{\boldsymbol{\theta}}_{\mathbf{2}}$ is a neutral element. However, using the commutativity axiom, we get $\overline{\boldsymbol{\theta}}_{\mathbf{2}} + \overline{\boldsymbol{\theta}}_{\mathbf{1}} = \overline{\boldsymbol{\theta}}_{\mathbf{1}} + \overline{\boldsymbol{\theta}}_{\mathbf{2}} \Rightarrow \overline{\boldsymbol{\theta}}_{\mathbf{1}} = \overline{\boldsymbol{\theta}}_{\mathbf{2}}$. As the neutral element occurred to be unique, we would denote it as $\overline{0}$ **,** and we would call it *zero* (or *zero element*).

2) *Each element* $\overline{\boldsymbol{a}}$ *in linear space has a unique inverse element.*

Indeed, let us suppose that $\overline{\boldsymbol{b}}_{\mathbf{1}}$ and $\overline{\boldsymbol{b}}_{\mathbf{2}}$ are the inverse elements to $\overline{\boldsymbol{a}}$**.** Then $\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}}_{\mathbf{1}} = \overline{\mathbf{0}}$ and $\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}}_{\mathbf{2}} = \overline{\mathbf{0}}$**.** The sequence of equalities shows us that

$$\overline{\boldsymbol{b}}_{\mathbf{1}} = \overline{\boldsymbol{b}}_{\mathbf{1}} + \overline{\mathbf{0}} = \overline{\boldsymbol{b}}_{\mathbf{1}} + \left(\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}}_{\mathbf{2}}\right) = \left(\overline{\boldsymbol{b}}_{\mathbf{1}} + \overline{\boldsymbol{a}}\right) + \overline{\boldsymbol{b}}_{\mathbf{2}} = \left(\overline{\boldsymbol{a}} + \overline{\boldsymbol{b}}_{\mathbf{1}}\right) + \overline{\boldsymbol{b}}_{\mathbf{2}} = \overline{= \mathbf{0}} + \overline{\boldsymbol{b}}_{\mathbf{2}} = \overline{\boldsymbol{b}}_{\mathbf{2}} + \overline{\mathbf{0}} = \overline{\boldsymbol{b}}_{\mathbf{2}}. \text{ So } \overline{\boldsymbol{b}}_{\mathbf{1}} = \overline{\boldsymbol{b}}_{\mathbf{2}}.$$

So the inverse element turned out to be unique and we denote it as $(-\overline{\boldsymbol{a}})$. If we take the commutativity axiom, we get the following law: if $\overline{\boldsymbol{a}} + (-\overline{\boldsymbol{a}}) = \overline{\mathbf{0}}$, then $(-\overline{\boldsymbol{a}}) + \overline{\boldsymbol{a}} = \overline{\mathbf{0}}$. Thus, $\overline{\boldsymbol{a}}$ is the inverse element for $(-\overline{\boldsymbol{a}})$. If we recall the denotation of the inverse element, we get $\overline{\boldsymbol{a}} = -(-\overline{\boldsymbol{a}})$.

3) *The equation* $\overline{\boldsymbol{x}} + \overline{\boldsymbol{a}} = \overline{\boldsymbol{b}}$ *in linear space for every* $\overline{\boldsymbol{a}}$ *and* $\overline{\boldsymbol{b}}$ *has a solution, and this solution is unique.*

Firstly, we should notice that $\left(\overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}})\right) + \overline{\boldsymbol{a}} = \overline{\boldsymbol{b}} + \left((-\overline{\boldsymbol{a}}) + \overline{\boldsymbol{a}}\right) = \overline{\boldsymbol{b}} + \overline{\mathbf{0}} = \overline{\boldsymbol{b}}$. So $\overline{\boldsymbol{x}} = \overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}})$ is a solution of this equation. Now let us take $\overline{\boldsymbol{x}}_{\mathbf{0}}$ for its solution. Then $\overline{\boldsymbol{x}}_{\mathbf{0}} + \overline{\boldsymbol{a}} = \overline{\boldsymbol{b}}$. Now we have the following expression: $(\overline{\boldsymbol{x}}_{\mathbf{0}} + \overline{\boldsymbol{a}}) + (-\overline{\boldsymbol{a}}) = \overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}}) \Rightarrow \overline{\boldsymbol{x}}_{\mathbf{0}} + \left(\overline{\boldsymbol{a}} + (-\overline{\boldsymbol{a}})\right) = \overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}}) \Rightarrow \overline{\boldsymbol{x}}_{\mathbf{0}} + \overline{\mathbf{0}} = \overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}})$ and, finally, $\overline{\boldsymbol{x}}_{\mathbf{0}} = \overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}})$. So we have the solution, this solution is unique and equals $\overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}})$.

According to the definition, we take $\overline{\boldsymbol{b}} - \overline{\boldsymbol{a}} = \overline{\boldsymbol{b}} + (-\overline{\boldsymbol{a}})$ and call it the difference between the elements $\overline{\boldsymbol{b}}$ and $\overline{\boldsymbol{a}}$. The operation '$(-)$' we call ***subtraction***. By the way, $\overline{\boldsymbol{a}} - \overline{\boldsymbol{a}} = \overline{\boldsymbol{a}} + (-\overline{\boldsymbol{a}}) = \overline{\mathbf{0}}$, which is rather obvious.

Remark **In a similar way, we have the unique solution of the equation** $\overline{\boldsymbol{a}} + \overline{\boldsymbol{y}} = \overline{\boldsymbol{b}}$ **and this solution is an element** $\overline{\boldsymbol{y}} = \overline{\boldsymbol{b}} - \overline{\boldsymbol{a}}$.

4) *In a linear space for each element* $\overline{\boldsymbol{a}}$ *the following equality is true*: $0 \cdot \overline{\boldsymbol{a}} = \overline{\mathbf{0}}$.

The unique solution of the equation $\overline{\boldsymbol{a}} + \overline{\boldsymbol{x}} = \overline{\boldsymbol{a}}$ is obviously $\overline{\boldsymbol{x}} = \overline{\mathbf{0}}$. On the other hand, we have

$\overline{\boldsymbol{a}} + 0 \cdot \overline{\boldsymbol{a}} = 1 \cdot \overline{\boldsymbol{a}} + 0 \cdot \overline{\boldsymbol{a}} = (1 + \mathbf{0}) \cdot \overline{\boldsymbol{a}} = 1 \cdot \overline{\boldsymbol{a}} = \overline{\boldsymbol{a}}$.

Thus, $0 \cdot \overline{\boldsymbol{a}}$ is also a solution. So $0 \cdot \overline{\boldsymbol{a}} = \overline{\mathbf{0}}$.

5) *In a linear space for every number* α *the following equality is true:* $\alpha \cdot \overline{\mathbf{0}} = \overline{\mathbf{0}}$.

The unique solution of the $\alpha \cdot \overline{\mathbf{0}} + \overline{\boldsymbol{x}} = \alpha \cdot \overline{\mathbf{0}}$ is, obviously, $\overline{\boldsymbol{x}} = \overline{\mathbf{0}}$. On the other hand,

$$\alpha \cdot \overline{\mathbf{0}} + \alpha \cdot \overline{\mathbf{0}} = \alpha \cdot (\overline{\mathbf{0}} + \overline{\mathbf{0}}) = \alpha \cdot \overline{\mathbf{0}}.$$

So $\alpha \cdot \overline{\mathbf{0}} = \overline{\mathbf{0}}$.

6) *If* $\alpha \cdot \overline{\boldsymbol{a}} = \overline{\mathbf{0}}$, *then* $\alpha = 0$ *or/and* $\overline{\boldsymbol{a}} = \overline{\mathbf{0}}$.

If $\alpha = 0$, then this property comes from the property **4**). Let us suppose that $\alpha \neq 0$. We multiply both parts of the equation $\alpha \cdot \overline{\boldsymbol{a}} = \overline{\mathbf{0}}$ by the number $1/\alpha$. We get that $1/\alpha \cdot (\alpha \cdot \overline{\boldsymbol{a}}) = 1/\alpha \cdot \overline{\mathbf{0}}$. According to the property **5)** the right-hand part of this equality equals $\overline{\mathbf{0}}$. In the left-hand part we have $1/\alpha\cdot(\alpha\cdot\overline{\boldsymbol{a}}) = (1/\alpha\cdot\alpha)\cdot\overline{\boldsymbol{a}} = 1\cdot\overline{\boldsymbol{a}} = \overline{\boldsymbol{a}}.$ So we have $\overline{\boldsymbol{a}} = \overline{\mathbf{0}}$.

7) $(-\mathbf{1})\cdot\overline{\boldsymbol{a}} = -\overline{\boldsymbol{a}}$ *for each element* $\overline{\boldsymbol{a}}.$

Indeed, the element $-\bar{a}$ is a unique solution of the equation $\overline{\boldsymbol{a}} + \overline{\boldsymbol{x}} = \overline{\mathbf{0}}.$ On the other hand,

$$\overline{\boldsymbol{a}} + \big((-\mathbf{1})\cdot\overline{\boldsymbol{a}}\big) = \mathbf{1}\cdot\overline{\boldsymbol{a}} + \big((-\mathbf{1})\cdot\overline{\boldsymbol{a}}\big) = (\mathbf{1}-\mathbf{1})\cdot\overline{\boldsymbol{a}} = \mathbf{0}\cdot\overline{\boldsymbol{a}} = \overline{\mathbf{0}}.$$

Thus, $(-\mathbf{1})\cdot\overline{\boldsymbol{a}} = -\overline{\boldsymbol{a}}$.

From this property we get the other rules of working with signs. For example,

$$-((-\mathbf{1})\cdot\overline{\boldsymbol{a}}) = (-\mathbf{1})\cdot((-\mathbf{1})\cdot\overline{\boldsymbol{a}}) = \big((-\mathbf{1})\cdot(-\mathbf{1})\big)\cdot\overline{\boldsymbol{a}} = \mathbf{1}\cdot\overline{\boldsymbol{a}} = \overline{\boldsymbol{a}},$$

Or $(-\mathbf{1})\cdot(-\overline{\boldsymbol{a}}) = (-\mathbf{1})\cdot\big((-\mathbf{1})\cdot\overline{\boldsymbol{a}}\big) = \overline{\boldsymbol{a}}$ and so on.

So the rules for the signs are just the same rules we got used to study at school.

8) $((\alpha\neq 0)\wedge(\alpha\overline{\boldsymbol{a}} = \alpha\overline{\boldsymbol{b}})) \Rightarrow (\overline{\boldsymbol{a}} = \overline{\boldsymbol{b}})$ – *the first rule of simplification*

We have an equality: $\alpha\overline{\boldsymbol{a}} = \alpha\overline{\boldsymbol{b}} \Leftrightarrow \alpha\overline{\boldsymbol{a}} - \alpha\overline{\boldsymbol{b}} = \overline{\mathbf{0}} \Leftrightarrow$ $\Leftrightarrow \alpha(\overline{\boldsymbol{a}} - \overline{\boldsymbol{b}}) = \overline{\mathbf{0}}.$ And as $\alpha\neq 0$, then $\overline{\boldsymbol{a}} - \overline{\boldsymbol{b}} = \overline{\mathbf{0}} \Leftrightarrow \overline{\boldsymbol{a}} = \overline{\boldsymbol{b}}$.

9) $(\overline{\boldsymbol{a}}\neq\overline{\mathbf{0}})\wedge(\alpha\overline{\boldsymbol{a}} = \beta\overline{\boldsymbol{a}})) \Leftrightarrow (\alpha=\beta)$ – *the second rule of simplification.*

We have an equality: $\alpha\overline{\boldsymbol{a}} = \beta\overline{\boldsymbol{a}} \Leftrightarrow \alpha\overline{\boldsymbol{a}} - \beta\overline{\boldsymbol{a}} = \overline{\mathbf{0}} \Leftrightarrow$ $\Leftrightarrow (\alpha-\beta)\overline{\boldsymbol{a}} = \overline{\mathbf{0}}.$ And as $\overline{\boldsymbol{a}}\neq\overline{\mathbf{0}},$ then $\alpha-\beta = 0 \Leftrightarrow \alpha=\beta$.

Remark **From the proved properties we can conclude that in the unspecified linear space we can add and multiply the elements by a scalar exactly as in the geometrical vectors' space. However, the types of a linear space's elements can be different. In the unspecified linear space we don't have any geometrical factors or clear evidence. That makes the difference of investigation for linear spaces.**

History is a great teacher! We have just learned the term 'zero element' – and would you like to learn how old is the number 'zero'? By 1770 BC in Ancient Egypt zero was used in the shape of the hieroglyph called *'nfr'* which looked like a heart connected with trachea – that hieroglyph meant ‘beautiful, good’. It was used to indicate the base level in drawings and so all the distances were marked above or below this sign. In Pre-Columbian America different civilizations had different counting systems; the earliest date that was drawn on the ancient stela, was 8-th December, 36 BC. It is interesting to learn that ancient Maia used the same denotation for zero and infinity as it meant ‘the beginning’, ‘the main cause’. In Ancient Greece zero wasn’t considered a number for a long time – as well as in the Medieval Europe it wasn’t a number for a long time. However, still there are discussions whether the modern symbol 0 came from a Greek lettet ‘o’ – ‘omikron’ – or from a sign in Ancient India which looked like a large point or a circle painted inside. In India that symbol was firstly written in 876 AC; zero as a number was called *śūnyaḥ* which meant ‘emptiness’, and the symbol described above was called *‘śūnya-binduḥ’* – ‘the point of emptiness’. From India via the Arabian countries this number came to Western Europe.

Fig. 1 One of the ancient Maia’s drawings depicting a zero on their counting system – ‘an empty shell’

<u>Practice Russian</u>
Лине́йное простра́нство – a linear space

3.2. The concept of an operator in linear space

Let $\mathscr{L}$ be a linear space.

<u>The operator $\mathcal{U}$ from $\mathscr{L}$ to $\mathscr{L}$</u> (we put it down in a following way: $\mathcal{U}$: $\mathscr{L} \longrightarrow \mathscr{L}$) is *any rule*, where for each element $\overline{x} \in \mathscr{L}$ a concretely defined element $\overline{y} \in \mathscr{L}$ is put in accordance.
Thus, an operator is an action or a set of actions that we should do with the element $\overline{x}$ (according to some special rules) to get the element $\overline{y}$. We call $\overline{y}$ ***<u>the image of $\overline{x}$,</u>*** and $\overline{x}$ is ***<u>a pre-image of $\overline{y}$</u>*** while the

oprator $\mathcal{U}$ is in action. The fact that the concrete element $\bar{y}$ **(and $\bar{y}$ only)** is an image of a concrete element $\bar{x}$ **(and $\bar{x}$ only)** is put down as an equality: $\bar{y} = \mathcal{U}(\bar{x})$. The operator makes an element $\bar{x}$ of a linear space to the element $\bar{y}$ of that same space. This can be presented as a scheme:

$$\bar{x} \longrightarrow \boxed{\mathcal{U}} \longrightarrow \bar{y}.$$

We have the vector $\bar{x}$ at the input and a vector $\bar{y}$ at the output.

Let us define the basic 'arithmetic' rules for the operators.

Addition The operator $\mathcal{W}$ is called ***the sum of operators*** $\mathcal{V}$ and $\mathcal{U}$ ($\mathcal{W} = \mathcal{V} + \mathcal{U}$), if $\forall\ \bar{x} \in \mathscr{L}$

$$\mathcal{W}(\bar{x}) = \mathcal{V}(\bar{x}) + \mathcal{U}(\bar{x}).$$

Multiplication by a scalar The operator $\mathcal{V} = \alpha\mathcal{U}$, if $\forall\ \bar{x} \in \mathscr{L}$:

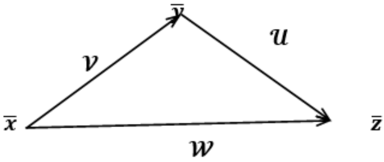
$\mathcal{V}(\bar{x}) = \alpha\,\mathcal{U}(\bar{x})$.

A product of operators The operator $\mathcal{W} = \mathcal{U}\mathcal{V}$ is called a product of operators $\mathcal{V}$ and $\mathcal{U}$, if $\forall\ \bar{x} \in \mathscr{L}$

$$\mathcal{W}(\bar{x}) = \mathcal{U}(\mathcal{V}(\bar{x})).$$

This equality should be interpreted in a following way: firstly the operator $\mathcal{V}$ influences the element $\bar{x}$; we get the element $\bar{y} = \mathcal{V}(\bar{x})$ as the result; then $\bar{y}$ is influenced by the operator $\mathcal{U}$, so we get $\bar{z} = \mathcal{U}(\bar{y})$. So $\mathcal{W}(\bar{x}) = \bar{z}$.

We can illustrate it by the diagram:



A product of operators has an associativity property. The equality $(\mathcal{W}\mathcal{U})\mathcal{V} = \mathcal{W}(\mathcal{U}\mathcal{V})$ is true for any three operators. On one hand, $\forall\ \bar{x} \in \mathscr{L}$

$$\big((\mathcal{W}\mathcal{U})\mathcal{V}\big)(\bar{x}) = (\mathcal{W}\mathcal{U})\big(\mathcal{V}(\bar{x})\big) = \mathcal{W}\Big(\mathcal{U}\big(\mathcal{V}(\bar{x})\big)\Big).$$

On the other hand, we get

$$(\boldsymbol{\mathcal{W}}(\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}}))(\overline{\boldsymbol{x}}) = \boldsymbol{\mathcal{W}}((\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}})(\overline{\boldsymbol{x}})) = \boldsymbol{\mathcal{W}}\left(\boldsymbol{\mathcal{U}}(\boldsymbol{\mathcal{V}}(\overline{\boldsymbol{x}}))\right),$$

So we get the same result. That's why $(\boldsymbol{\mathcal{W}}\boldsymbol{\mathcal{U}})\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{W}}(\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}})$. According to that, we have no need to put any brackets.

<u>The identity operator $\boldsymbol{\mathcal{I}}$</u> is defined by the rule: $\forall\, \overline{\boldsymbol{x}} \in \mathscr{L}\, (\boldsymbol{\mathcal{I}}\,(\overline{\boldsymbol{x}}) = \overline{\boldsymbol{x}}\,)$. For the *zero operator* $\boldsymbol{\mathcal{O}}$ we have a rule: $\forall\, \overline{\boldsymbol{x}} \in \mathscr{L}\, (\boldsymbol{\mathcal{O}}\,(\overline{\boldsymbol{x}}) = \overline{\boldsymbol{0}}\,)$. Заметим, что тождественный оператор in the process of operators' multiplication plays a role of a 'pivot', i.e. $\forall\, \overline{\boldsymbol{x}} \in \mathscr{L}\,(\,(\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{I}})\,(\overline{\boldsymbol{x}}) = \boldsymbol{\mathcal{U}}(\boldsymbol{\mathcal{I}}(\overline{\boldsymbol{x}})) = \boldsymbol{\mathcal{U}}(\overline{\boldsymbol{x}})$. Thus, $\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{U}}$. In a similar way we get $\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{U}}$.

The operator $\boldsymbol{\mathcal{U}}$ is called ***<u>the inverse operator</u>*** for the operator $\boldsymbol{\mathcal{V}}$ if $\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{I}}$.

The operator that has an inverse one, is called ***<u>the invertible operator.</u>*** It is clear that this term can be used as a pairwise one: if the operator $\boldsymbol{\mathcal{U}}$ is the inverse operator for $\boldsymbol{\mathcal{V}}$, then , vice versa, $\boldsymbol{\mathcal{V}}$ is an inverse operator for $\boldsymbol{\mathcal{U}}$.

According to the associativity rule, an operator can have only one inverse operator. Indeed, if $\boldsymbol{\mathcal{U}}_1$ and $\boldsymbol{\mathcal{U}}_2$ are inverse for the operator $\boldsymbol{\mathcal{V}}$, then

$$\boldsymbol{\mathcal{U}}_1 = \boldsymbol{\mathcal{U}}_1\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{U}}_1(\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{U}}_2) = (\boldsymbol{\mathcal{U}}_1\boldsymbol{\mathcal{V}})\boldsymbol{\mathcal{U}}_2 = \boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{U}}_2 = \boldsymbol{\mathcal{U}}_2.$$

For the inverse operator to $\boldsymbol{\mathcal{U}}$ we have a denotation $\boldsymbol{\mathcal{U}}^{-1}$.

If $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$ are invertible operators then the product $\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}}$ is also an invertible operator. The following expression is fulfilled: $(\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}})^{-1} = \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{\mathcal{U}}^{-1}$. Indeed, $\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}}\boldsymbol{\mathcal{V}}^{-1}\boldsymbol{\mathcal{U}}^{-1} = \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{U}}^{-1} = \boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{U}}^{-1} = \boldsymbol{\mathcal{I}}$. In a similar way we get: $\boldsymbol{\mathcal{V}}^{-1}\boldsymbol{\mathcal{U}}^{-1}\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{\mathcal{I}}\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{V}}^{-1}\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{I}}$. It is quite obvious that $(\boldsymbol{\mathcal{U}}^{-1})^{-1} = \boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{I}}^{-1} = \boldsymbol{\mathcal{I}}$.

<u>The degree of an operator</u>

The multiplication operation allows us to introduce the concept of the operator's degree. The operation of operators' multiplication allows to introduce the concept of operator's degree. Let $\boldsymbol{\mathcal{U}}$ be any unspecified operator and let $n \geq 1$ be a natural number. Then, using the definition, we get

$$\boldsymbol{\mathcal{U}}^n = \begin{cases} \boldsymbol{\mathcal{U}}, & n = 1; \\ \underbrace{\boldsymbol{\mathcal{U}}\boldsymbol{\mathcal{U}}\ldots\boldsymbol{\mathcal{U}}}_{n}, & n \geq 2. \end{cases}$$

According to the definition, we suppose that for any operator $\mathcal{U}^0 = \mathcal{I}$. We can further introduce the negative degrees but we can do it only for the invertible operators. We can do it using the following algorithm: if $\mathcal{U}$ is an invertible operator and the integer number $n < 0$, then we suppose

$$\mathcal{U}^n = (\mathcal{U}^{-1})^{-n}.$$

It's not hard to check that for every pair of integer numbers n and m the following equalities are true: **1**) $\mathcal{U}^{n+m} = \mathcal{U}^n\mathcal{U}^m$; **2**) $(\mathcal{U}^n)^m = \mathcal{U}^{nm}$. We should only remember that if at least one of the degrees is negative then the operator should be an invertible one. If n or m is zero then $\mathcal{U} \neq \mathcal{O}$.

We call $\mathcal{U}$ and $\mathcal{V}$ ***the commuting operators*** if $\mathcal{U}\mathcal{V} = \mathcal{V}\mathcal{U}$.

For the operators of that type the following equality is true:

$$(\mathcal{U}\mathcal{V})^n = \mathcal{U}^n\mathcal{V}^n.$$

At last, we should notice that a set of all the operators which act in the given space, together with the operations of addition and multiplication by a scalar, comprises a linear space itself. As for the operators' product, we can tell that this operation is *external* for the given linear space and doesn't have any influence on the given linear space.

Remark We should note that the operators are also called transformations, functions or functionals. Often it depends from the type of a space where they take place. For example, a functional is an operator in the functional space. In our book we will mostly use the term 'operator'.

Practice Russian

Обра́тный опера́тор – the inverse operator

3.3. Linear operators

We come to the basic idea of a lecture. However, the concepts introduced above will be very useful for the further material

The operator $\mathcal{A}$: $\mathcal{L} \longrightarrow \mathcal{L}$ is called ***a linear operator*** if 2 properties are fulfilled

1) $\mathcal{A}(\bar{x} + \bar{y}) = \mathcal{A}(\bar{x}) + \mathcal{A}(\bar{y})$,

2) $\mathcal{A}(\alpha\bar{x}) = \alpha\mathcal{A}(\bar{x})$, which take place for every elements $\overline{\boldsymbol{x}}$ and $\overline{\boldsymbol{y}}$ of a given functional space and for each scalar α.

If at least one of these conditions is not fulfilled for at least one element of a linear space or for at least one scalar, then such an operator is not a linear operator. The operator which is not a linear operator , is called ***a non-linear operator.***

If the operator $\boldsymbol{\mathcal{A}}$ is linear then $\boldsymbol{\mathcal{A}}(\overline{\mathbf{0}}) = \overline{\mathbf{0}}$. **We can tell that it is a necessary criterion of a linearity. In a case when this criterion is not fulfilled, we can tell in advance that the investigated operator is non-linear.** So let $\boldsymbol{\mathcal{A}}$ be a linear operator and $\boldsymbol{\mathcal{A}}(\overline{\mathbf{0}}) = \overline{\boldsymbol{a}}$. As the operator is linear, $\overline{\boldsymbol{a}} = \boldsymbol{\mathcal{A}}(\overline{\mathbf{0}}) = \boldsymbol{\mathcal{A}}(\overline{\mathbf{0}} + \overline{\mathbf{0}}) = \boldsymbol{\mathcal{A}}(\overline{\mathbf{0}}) + \boldsymbol{\mathcal{A}}(\overline{\mathbf{0}}) = \overline{\boldsymbol{a}} + \overline{\boldsymbol{a}}$. From that we get: $\overline{\boldsymbol{a}} = \overline{\boldsymbol{a}} - \overline{\boldsymbol{a}} = \overline{\mathbf{0}}$.

Statement Let $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ be linear operators and let α be any unspecified number. Then the operators $\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}}$, $\alpha\boldsymbol{\mathcal{A}}$, $\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{A}}^{-1}$ (if $\boldsymbol{\mathcal{A}}^{-1}$ exists) are linear operators.

You can prove this statement yourself.

Some examples of linear operators

1) A zero operator

$$\boldsymbol{\mathcal{O}}(\overline{\boldsymbol{x}} + \overline{\boldsymbol{y}}) = \overline{\mathbf{0}} = \overline{\mathbf{0}} + \overline{\mathbf{0}} = \boldsymbol{\mathcal{O}}(\overline{\boldsymbol{x}}) + \boldsymbol{\mathcal{O}}(\overline{\boldsymbol{y}}),$$
$$\boldsymbol{\mathcal{O}}(\alpha\overline{\boldsymbol{x}}) = \overline{\mathbf{0}} = \alpha \cdot \overline{\mathbf{0}} = \alpha\boldsymbol{\mathcal{O}}(\overline{\boldsymbol{x}}).$$

1) An identity operator

$\boldsymbol{\mathcal{I}}(\overline{\boldsymbol{x}} + \overline{\boldsymbol{y}}) = \overline{\boldsymbol{x}} + \overline{\boldsymbol{y}} = \boldsymbol{\mathcal{I}}(\overline{\boldsymbol{x}}) + \boldsymbol{\mathcal{I}}(\overline{\boldsymbol{y}}),$

$\boldsymbol{\mathcal{I}}(\alpha\overline{\boldsymbol{x}}) = \alpha\overline{\boldsymbol{x}} = \alpha\boldsymbol{\mathcal{I}}(\overline{\boldsymbol{x}}).$

3) Let us introduce the operator $\boldsymbol{\mathcal{A}}$ in the space $\mathbb{R}^n$ using the following rule. Let A be the matrix of the n-th order. Then for $\forall X \in \mathbb{R}^{\boldsymbol{n}}$ we take $\boldsymbol{\mathcal{A}}(X) = A\,X$. Now let us show that the operator $\boldsymbol{\mathcal{A}}$ is linear:

$$\boldsymbol{\mathcal{A}}(X + Y) = A(X + Y) = A\,X + A\,Y = \boldsymbol{\mathcal{A}}(X) + \boldsymbol{\mathcal{A}}(Y),$$

$$\boldsymbol{\mathcal{A}}(\alpha X) = A(\alpha X) = \alpha AX = \alpha\boldsymbol{\mathcal{A}}(X).$$

4) Let us fix the vector $\vec{\boldsymbol{a}}$ in the space of geometrical vectors. Let the operator $\boldsymbol{\mathcal{A}}$ act according to the rule:

$$\boldsymbol{\mathcal{A}}(\vec{\boldsymbol{x}}) = [\vec{\boldsymbol{a}}, \vec{\boldsymbol{x}}],$$

so this expression means that under the influence of $\mathcal{A}$ every given vector is transformed into a vector product of a fixed vector by a given vector. Its linearity comes from the properties of a vector product. Indeed,

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

$$\mathcal{A}(\alpha\vec{x}) = [\vec{a},\alpha\vec{x}] = \alpha[\vec{a},\vec{x}] = \alpha\mathcal{A}(\vec{x}).$$

5) In the space $\mathscr{P}_n$ (or $\mathscr{P}$) let us investigate the operator of differentiation $\mathcal{A} = d/dx$. The operator acts according to the rule: if $\boldsymbol{P}(x)$ is a polynomial, then $\mathcal{A}(\boldsymbol{P}) = d\boldsymbol{P}(x)/dx$. The linearity of an operator comes from the rule that says 'the derivative of a sum is a sum of derivatives' and 'the number can be pulled out of the derivation sign'.

There are also many examples of the operators that fulfill some geometrical actions, like the operator $\mathcal{A}_\varphi$ of a plane rotation by the angle φ around the point O**;** the operator $\mathcal{A}_P$ of orthogonal projection to the plane **P,** etc., and the linearity of those operators is proved using the geometrical concepts as well as analytical ones.

Practice Russian

Линéйный оперáтор – a linear operator

3.4. A matrix of a linear operator

From this point we will consider that $\mathcal{L} = \mathcal{L}^n$ is an n-dimensional linear space.

Let $\mathcal{A}$: $\mathcal{L}^n \longrightarrow \mathcal{L}^n$ be a linear operator. Let us fix in our space the basis $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_n$. Further we introduce $\bar{x} \in \mathcal{L}^n$. A vector $\bar{x}$ can be presented in this basis. Let us suggest that $\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n$ is a presentation in the basis. Then, due to the linearity of $\mathcal{A}$**,** we get:

$$\mathcal{A}(\bar{x}) = \mathcal{A}(x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n) =$$
$$= x_1\mathcal{A}(\bar{e}_1) + x_2\mathcal{A}(\bar{e}_2) + \cdots + x_n\mathcal{A}(\bar{e}_n).$$

The equality $\mathcal{A}(\bar{x}) = x_1\mathcal{A}(\bar{e}_1) + x_2\mathcal{A}(\bar{e}_2) + \cdots + x_n\mathcal{A}(\bar{e}_n)$ shows that the operator will be completely defined, i.e. we can learn the image of any vector if we know the images of basis vectors: $\mathcal{A}(\bar{e}_1)$, $\mathcal{A}(\bar{e}_2), \dots, \mathcal{A}(\bar{e}_n)$.

Now we should note that the images of basis vectors are themselves the elements of the linear space. So they are able to be presented as the vectors of the given basis. Let us suppose that we already know the result of the transformation for the new basis and that they have a form:

$$\begin{gathered}\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_1) = a_{11}\bar{\boldsymbol{e}}_1 + a_{21}\bar{\boldsymbol{e}}_2 + \dots + a_{n1}\bar{\boldsymbol{e}}_n, \\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_2) = a_{12}\bar{\boldsymbol{e}}_1 + a_{22}\bar{\boldsymbol{e}}_2 + \dots + a_{n2}\bar{\boldsymbol{e}}_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_n) = a_{1n}\bar{\boldsymbol{e}}_1 + \boldsymbol{a}_{2n}\bar{\boldsymbol{e}}_2 + \dots + a_{nn}\bar{\boldsymbol{e}}_n.\end{gathered}$$

The presentation of the image j of that basis vector will look like that:

$$\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_j) = a_{1j}\bar{\boldsymbol{e}}_1 + a_{2j}\bar{\boldsymbol{e}}_2 + \dots + a_{nj}\bar{\boldsymbol{e}}_n.$$

The numbers $a_{1j}, a_{2j}, \dots, a_{nj}$ are the coordinates of vector $\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_j)$ in the given basis.

<u>A matrix A = $\|a_{ij}\|$ of a linear operator $\boldsymbol{\mathcal{A}}$</u> in the given basis $\bar{\boldsymbol{e}}_1, \bar{\boldsymbol{e}}_2, \dots, \bar{\boldsymbol{e}}_n$ is a square matrix of the order n, and in its j– th column the coordinates of the vector $\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_j)$ for its presentation in this basis are situated.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}.$$

In the case of a fixed basis, each operator has only one matrix, i.e. each of the vectors $\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_1), \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_2), \dots, \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_n)$ can be presented in the basis in a unique way.

<u>Conclusion</u>: **If we have a fixed basis in a linear space, then for each operator we can take the corresponding matrix, i.e. a matrix of the operator in that basis.**

It is clear that for the different operators we have different corresponding matrices. Indeed, if a pair of operators had the same matrices that meant the coinciding images of different matrices. That would cause the coincidence of any unspecified vectors' images, i.e. the equality of the operators.

Remark **We consider the following rule to be obvious: $\mathcal{A}$ and $\mathcal{B}$ are equal ($\mathcal{A} = \mathcal{B}$), if $\forall \overline{\boldsymbol{x}} \in \mathcal{L}(\mathcal{A}(\overline{\boldsymbol{x}}) = \mathcal{B}(\overline{\boldsymbol{x}}))$.**

Statement If A is an unspecified square matrix of the dimension n, then there is a unique linear operator $\mathcal{A}$, so that this square matrix is a matrix of this operator in the given basis.

Proof Let

$$A = \begin{pmatrix} a_{11}\, a_{12} \cdots a_{1j} \cdots a_{1n} \\ a_{21}\, a_{22} \cdots a_{2j} \cdots a_{2n} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{i1}\, a_{i2} \cdots a_{ij} \cdots a_{in} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{n1}\, a_{n2} \cdots a_{nj} \cdots a_{nn} \end{pmatrix}$$

be a given matrix. The basis $\overline{\boldsymbol{e}}_{\mathbf{1}},\ \overline{\boldsymbol{e}}_{\mathbf{2}}, \ldots,\ \overline{\boldsymbol{e}}_{\boldsymbol{n}}$ is fixed. We would introduce n vectors in a following way:

$$\begin{aligned} \overline{\boldsymbol{g}}_{\mathbf{1}} &= a_{11}\overline{\boldsymbol{e}}_{\mathbf{1}} + a_{21}\overline{\boldsymbol{e}}_{\mathbf{2}} + \cdots + a_{n1}\overline{\boldsymbol{e}}_{\boldsymbol{n}}, \\ \overline{\boldsymbol{g}}_{\mathbf{2}} &= a_{12}\overline{\boldsymbol{e}}_{\mathbf{1}} + a_{22}\overline{\boldsymbol{e}}_{\mathbf{2}} + \cdots + a_{n2}\overline{\boldsymbol{e}}_{\boldsymbol{n}}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \overline{\boldsymbol{g}}_{\boldsymbol{n}} &= a_{1n}\overline{\boldsymbol{e}}_{\mathbf{1}} + a_{2n}\overline{\boldsymbol{e}}_{\mathbf{2}} + \cdots + a_{nn}\overline{\boldsymbol{e}}_{\boldsymbol{n}}. \end{aligned}$$

We have a rule that defines the action of the operator $\mathcal{A}$. If $\overline{\boldsymbol{x}} = x_1\overline{\boldsymbol{e}}_{\mathbf{1}} + x_2\overline{\boldsymbol{e}}_{\mathbf{2}} + \cdots + x_n\overline{\boldsymbol{e}}_{\boldsymbol{n}}$, then

$$\mathcal{A}(\overline{\boldsymbol{x}}) = x_1\overline{\boldsymbol{g}}_{\mathbf{1}} + x_2\overline{\boldsymbol{g}}_{\mathbf{2}} + \cdots + x_n\overline{\boldsymbol{g}}_{\boldsymbol{n}} = \overline{\boldsymbol{z}}.$$

Here it is important to tell that the operator $\mathcal{A}$, with the use of a fixed basis, puts a corresponding element $\overline{\boldsymbol{z}}$ to every element $\overline{\boldsymbol{x}}$, and $\overline{\boldsymbol{z}}$ doesn't depend on any basis.

Let us check the linearity of that operator. Then we introduce $\overline{\boldsymbol{y}} = y_1\overline{\boldsymbol{e}}_{\mathbf{1}} + y_2\overline{\boldsymbol{e}}_{\mathbf{2}} + \cdots + y_n\overline{\boldsymbol{e}}_{\boldsymbol{n}}$ and α (α is a number). So we have

$$\overline{\boldsymbol{x}} + \overline{\boldsymbol{y}} = (x_1 + y_1)\overline{\boldsymbol{e}}_{\mathbf{1}} + (x_2 + y_2)\overline{\boldsymbol{e}}_{\mathbf{2}} + \cdots + (x_n + y_n)\overline{\boldsymbol{e}}_{\boldsymbol{n}}$$

and

$$\mathcal{A}(\overline{\boldsymbol{x}} + \overline{\boldsymbol{y}}) = (x_1 + y_1)\overline{\boldsymbol{g}}_{\mathbf{1}} + (x_2 + y_2)\overline{\boldsymbol{g}}_{\mathbf{2}} + \cdots + (x_n + y_n)\overline{\boldsymbol{g}}_{\boldsymbol{n}}.$$

On the other hand,

$$\begin{aligned} \mathcal{A}(\overline{\boldsymbol{x}}) &= x_1\overline{\boldsymbol{g}}_{\mathbf{1}} + x_2\overline{\boldsymbol{g}}_{\mathbf{2}} + \cdots + x_n\overline{\boldsymbol{g}}_{\boldsymbol{n}}, \\ \mathcal{A}(\overline{\boldsymbol{y}}) &= y_1\overline{\boldsymbol{g}}_{\mathbf{1}} + y_2\overline{\boldsymbol{g}}_{\mathbf{2}} + \cdots + y_n\overline{\boldsymbol{g}}_{\boldsymbol{n}}. \end{aligned}$$

After the addition (between each other) of the two equalities above and simplifying, we get

$$\mathcal{A}(\bar{x}) + \mathcal{A}(\bar{y}) = (x_1 + y_1)\bar{g}_1 + (x_2 + y_2)\bar{g}_2 + \cdots + (x_n + y_n)\bar{g}_n.$$

So $\mathcal{A}(\bar{x} + \bar{y}) = \mathcal{A}(\bar{x}) + \mathcal{A}(\bar{y})$.

Further we have

$$\mathcal{A}(\alpha\bar{x}) = (\alpha x_1)\bar{g}_1 + (\alpha x_2)\bar{g}_2 + \cdots + (\alpha x_n)\bar{g}_n = \\ = \alpha(x_1\bar{g}_1 + x_2\bar{g}_2 + \cdots + +x_n\bar{g}_n) = \alpha\mathcal{A}(\bar{x}).$$

The linearity is proved.

The fact that the matrix A is the matrix of the given operator in the basis $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$, can be proved in a rather simple way. Indeed,

$$\bar{e}_1 = 1 \cdot \bar{e}_1 + 0 \cdot \bar{e}_2 + \cdots + 0 \cdot \bar{e}_n$$

and so

$$\mathcal{A}(\bar{e}_1) = 1 \cdot \bar{g}_1 + 0 \cdot \bar{g}_2 + \cdots + 0 \cdot \bar{g}_n = \bar{g}_1 = a_{11}\bar{e}_1 + a_{21}\bar{e}_2 + \cdots + a_{n1}\bar{e}_n,$$

i.e.

$$\mathcal{A}(\bar{e}_1) = a_{11}\bar{e}_1 + a_{21}\bar{e}_2 + \cdots + a_{n1}\bar{e}_n;$$

and afterwards we get:

$$\bar{e}_2 = 0 \cdot \bar{e}_1 + 1 \cdot \bar{e}_2 + \cdots + 0 \cdot \bar{e}_n$$

and so that means

$$\mathcal{A}(\bar{e}_2) = 0 \cdot \bar{g}_1 + 1 \cdot \bar{g}_2 + \cdots + 0 \cdot \bar{g}_n = \\ = \bar{g}_2 = a_{12}\bar{e}_1 + a_{22}\bar{e}_2 + \cdots + a_{n2}\bar{e}_n,$$

i.e.

$$\mathcal{A}(\bar{e}_2) = a_{12}\bar{e}_1 + a_{22}\bar{e}_2 + \cdots + a_{n2}\bar{e}_n, \ldots$$

and so on.

At last, the uniqueness of the created operator comes out from the fact that the equality of matrices leads to the equality of the operators.

Conclusion **In a linear space $\mathcal{L}^n$ for a fixed basis, between all the linear operators and their matrices in this basis there is a *reciprocation (one-to-one correspondence).* In accordance with it, each operator has a unique corresponding matrix, and vice versa: each matrix has a unique linear operator, for which this matrix is the matrix of this operator in the given basis.**

This reciprocation is denoted in a following way: $\boldsymbol{\mathcal{A}} \leftrightarrow \boldsymbol{A}$.

Example Find the matrix of an operator $\boldsymbol{\mathcal{A}}(\vec{\boldsymbol{x}}) = [\vec{\boldsymbol{a}}, \vec{\boldsymbol{x}}]$ in the basis $\vec{\boldsymbol{\imath}}, \vec{\boldsymbol{\jmath}}, \vec{\boldsymbol{k}}$, where $\vec{\boldsymbol{a}} = a_x\vec{\boldsymbol{\imath}} + a_y\vec{\boldsymbol{\jmath}} + a_z\vec{\boldsymbol{k}}$ is a given vector.

So we have:

$$\boldsymbol{\mathcal{A}}(\vec{\imath}) = \left[a_x\vec{\boldsymbol{\imath}} + a_y\vec{\boldsymbol{\jmath}} + a_z\vec{\boldsymbol{k}}, \vec{\boldsymbol{\imath}}\right] = a_x[\vec{\boldsymbol{\imath}}, \vec{\boldsymbol{\imath}}] + a_y[\vec{\boldsymbol{\jmath}}, \vec{\boldsymbol{\imath}}] + a_z[\vec{\boldsymbol{k}}, \vec{\boldsymbol{\imath}}] =$$
$$= -a_y\vec{\boldsymbol{k}} + a_z\vec{\boldsymbol{\jmath}} \quad \Rightarrow \boldsymbol{\mathcal{A}}(\vec{\imath}) = 0\vec{\boldsymbol{\imath}} + a_z\vec{\boldsymbol{\jmath}} - a_y\vec{\boldsymbol{k}};$$

$$\boldsymbol{\mathcal{A}}(\vec{\jmath}) = \left[a_x\vec{\boldsymbol{\imath}} + a_y\vec{\boldsymbol{\jmath}} + a_z\vec{\boldsymbol{k}}, \vec{\boldsymbol{\jmath}}\right] = a_x[\vec{\boldsymbol{\imath}}, \vec{\boldsymbol{\jmath}}] + a_y[\vec{\boldsymbol{\jmath}}, \vec{\boldsymbol{\jmath}}] + a_z[\vec{\boldsymbol{k}}, \vec{\boldsymbol{\jmath}}] =$$
$$= a_x\vec{\boldsymbol{k}} - a_z\vec{\boldsymbol{\imath}} \quad \Rightarrow \boldsymbol{\mathcal{A}}(\vec{\jmath}) = -a_z\vec{\boldsymbol{\imath}} + 0\vec{\boldsymbol{\jmath}} + a_x\vec{\boldsymbol{k}};$$
$$\boldsymbol{\mathcal{A}}(\vec{\boldsymbol{k}}) = \left[a_x\vec{\boldsymbol{\imath}} + a_y\vec{\boldsymbol{\jmath}} + a_z\vec{\boldsymbol{k}}, \vec{\boldsymbol{k}}\right] = a_x[\vec{\boldsymbol{\imath}}, \vec{\boldsymbol{k}}] + a_y[\vec{\boldsymbol{\jmath}}, \vec{\boldsymbol{k}}] + a_z[\vec{\boldsymbol{k}}, \vec{\boldsymbol{k}}] =$$
$$= -a_x\vec{\boldsymbol{\jmath}} + a_y\vec{\boldsymbol{\imath}} \quad \Rightarrow \boldsymbol{\mathcal{A}}(\vec{\boldsymbol{k}}) = a_y\vec{\boldsymbol{\imath}} - a_x\vec{\boldsymbol{\jmath}} + 0\vec{\boldsymbol{k}}.$$

We get the following result:

$$A = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix}.$$

Besides, we can tell that the zero operator $\boldsymbol{\mathcal{O}}$ in any basis has a zero matrix O, because

$$\boldsymbol{\mathcal{O}}(\bar{\boldsymbol{e}}_1) = 0 \cdot \bar{\boldsymbol{e}}_1 + 0 \cdot \bar{\boldsymbol{e}}_2 + \cdots + 0 \cdot \bar{\boldsymbol{e}}_n,$$
$$\boldsymbol{\mathcal{O}}(\vec{\boldsymbol{e}}_2) = 0 \cdot \bar{\boldsymbol{e}}_1 + 0 \cdot \bar{\boldsymbol{e}}_2 + \cdots + 0 \cdot \bar{\boldsymbol{e}}_n, \Rightarrow \boldsymbol{\mathcal{O}} \leftrightarrow O$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$
$$\boldsymbol{\mathcal{O}}(\bar{\boldsymbol{e}}_n) = 0 \cdot \bar{\boldsymbol{e}}_1 + 0 \cdot \bar{\boldsymbol{e}}_2 + \cdots + 0 \cdot \bar{\boldsymbol{e}}_n,$$

and the identity operator $\boldsymbol{\mathcal{I}}$ always has a corresponding matrix E which is an identity matrix.

$$\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{e}}_1) = 1 \cdot \bar{\boldsymbol{e}}_1 + 0 \cdot \bar{\boldsymbol{e}}_2 + \cdots + 0 \cdot \bar{\boldsymbol{e}}_n,$$
$$\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{e}}_2) = 0 \cdot \bar{\boldsymbol{e}}_1 + 1 \cdot \bar{\boldsymbol{e}}_2 + \cdots + 0 \cdot \bar{\boldsymbol{e}}_n, \qquad \Rightarrow \boldsymbol{\mathcal{I}} \leftrightarrow E$$
$$\cdot\ \cdot,$$
$$\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{e}}_n) = 0 \cdot \bar{\boldsymbol{e}}_1 + 0 \cdot \bar{\boldsymbol{e}}_2 + \cdots + 1 \cdot \bar{\boldsymbol{e}}_n.$$

If $\boldsymbol{\mathcal{A}} \leftrightarrow A$, $\boldsymbol{\mathcal{B}} \leftrightarrow B$ and α is a number, then

$$\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}} \leftrightarrow A + B, \qquad \alpha\,\boldsymbol{\mathcal{A}} \leftrightarrow \alpha A, \qquad \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{B}} \leftrightarrow AB, \qquad \boldsymbol{\mathcal{A}}^{-1} \leftrightarrow A^{-1}.$$

(The last one is true if the inverse operator exists).

For example, we can state that any operator expression

$$\alpha \boldsymbol{\mathcal{A}}^n \boldsymbol{\mathcal{B}}^m \dots \boldsymbol{\mathcal{C}}^l + \beta \boldsymbol{\mathcal{H}}^p \boldsymbol{\mathcal{F}}^q \dots \boldsymbol{\mathcal{D}}^r + \dots + \gamma \boldsymbol{\mathcal{G}}^s \boldsymbol{\mathcal{Q}}^k \dots \boldsymbol{\mathcal{T}}^t$$

comes into a matrix expression

$$\alpha A^n B^m \dots C^l + \beta H^p F^q \dots D^r + \dots + \gamma G^s Q^k \dots T^t.$$

Example Find a matrix of an operator

$$\boldsymbol{\mathcal{A}}^2 - 3\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{B}} + 2\boldsymbol{\mathcal{B}}^2, \quad \text{if } \boldsymbol{\mathcal{A}} \leftrightarrow A, \quad \boldsymbol{\mathcal{B}} \leftrightarrow B;$$

where
$A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 7 & 1 \end{pmatrix}$ are the matrices of the given operators in some basis of $\boldsymbol{\mathcal{L}}^2$.

Solution Firstly, we come to the correspondence

$$\boldsymbol{\mathcal{A}}^2 - 3\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{B}} + 2\boldsymbol{\mathcal{B}}^2 \leftrightarrow A^2 - 3AB + 2B^2.$$

Then

$$A^2 = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 6 \\ 4 & 7 \end{pmatrix},$$
$$B^2 = \begin{pmatrix} 2 & 0 \\ 7 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 7 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 23 & 1 \end{pmatrix},$$
$$AB = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 7 & 1 \end{pmatrix} = \begin{pmatrix} 23 & 3 \\ 11 & 1 \end{pmatrix}.$$

And so we have $A^2 - 3AB + 2B^2 = \begin{pmatrix} 7 & 6 \\ 4 & 7 \end{pmatrix} - 3\begin{pmatrix} 23 & 3 \\ 11 & 1 \end{pmatrix} + 2\begin{pmatrix} 4 & 0 \\ 23 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 6 \\ 4 & 7 \end{pmatrix} - \begin{pmatrix} 69 & 9 \\ 33 & 3 \end{pmatrix} + \begin{pmatrix} 8 & 0 \\ 46 & 2 \end{pmatrix} = \begin{pmatrix} -54 & -3 \\ 17 & 6 \end{pmatrix}$

We got the answer.

3.5. The general operator equation and its matrix form

Let $\boldsymbol{\mathcal{A}}$: $\boldsymbol{\mathcal{L}}^n \longrightarrow \boldsymbol{\mathcal{L}}^n$ be a linear operator. We denote: $\bar{\boldsymbol{y}}$ is an image of a vector $\bar{\boldsymbol{x}}$ при действии оператора. Then the expression $\bar{\boldsymbol{y}} = \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{x}})$ is called ***a general operator equation.***

Let us choose the basis $\bar{\boldsymbol{e}}_1, \bar{\boldsymbol{e}}_2, \dots, \bar{\boldsymbol{e}}_n$ **in** $\boldsymbol{\mathcal{L}}^n$. Let $A = \|a_{ij}\|$ – be the matrix of the operator $\boldsymbol{\mathcal{A}}$ and the following expressions take place:

$$\bar{\boldsymbol{x}} = x_1\bar{\boldsymbol{e}}_1 + x_2\bar{\boldsymbol{e}}_2 + \dots + x_n\bar{\boldsymbol{e}}_n, \quad \bar{\boldsymbol{y}} = y_1\bar{\boldsymbol{e}}_1 + y_2\bar{\boldsymbol{e}}_2 + \dots + y_n\bar{\boldsymbol{e}}_n.$$

Let $\langle \boldsymbol{e} \rangle = \langle \bar{\boldsymbol{e}}_{\mathbf{1}},\ \bar{\boldsymbol{e}}_{\mathbf{2}}, \dots,\ \bar{\boldsymbol{e}}_{\boldsymbol{n}} \rangle$ be row matrix of vectors; then we denote as:

$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ the column matrices of the vectors $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}}$;

Let $\langle \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{1}}),\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{2}}), \dots,\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\boldsymbol{n}}) \rangle$ be the vector row matrix of the basis vectors' images.

If we use simple calculations, it's not hard to show that

$$\langle \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{1}}),\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{2}}), \dots,\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\boldsymbol{n}}) \rangle = \langle \bar{\boldsymbol{e}}_{\mathbf{1}},\ \bar{\boldsymbol{e}}_{\mathbf{2}}, \dots, \bar{\boldsymbol{e}}_{\boldsymbol{n}} \rangle \begin{pmatrix} a_{11}\, a_{12} \cdots a_{1j} \cdots a_{1n} \\ a_{21}\, a_{22} \cdots a_{2j} \cdots a_{2n} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{i1}\, a_{i2} \cdots a_{ij} \cdots a_{in} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{n1}\, a_{n2} \cdots a_{nj} \cdots a_{nn} \end{pmatrix} = \langle \boldsymbol{e} \rangle A,$$

So $\bar{\boldsymbol{y}} = \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{x}}) \quad \Leftrightarrow \quad y_1\bar{\boldsymbol{e}}_{\mathbf{1}} + y_2\bar{\boldsymbol{e}}_{\mathbf{2}} + \cdots + y_n\bar{\boldsymbol{e}}_{\boldsymbol{n}} = x_1\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{1}}) + x_2\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{2}}) + \cdots + x_n\boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\boldsymbol{n}})$

$$\Updownarrow$$

$$\langle \boldsymbol{e} \rangle Y = \langle \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{1}}),\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\mathbf{2}}), \dots,\ \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{e}}_{\boldsymbol{n}}) \rangle X = (\langle \boldsymbol{e} \rangle A)X.$$

And at last, as $(\langle \boldsymbol{e} \rangle A)X = \langle \boldsymbol{e} \rangle (AX)$ due to the linear independence of the basis vectors, we get

$$\boxed{Y = AX}$$

That is the relation that we needed to get. If we have a fixed basis, it is equivalent to the operator $\bar{\boldsymbol{y}} = \boldsymbol{\mathcal{A}}(\bar{\boldsymbol{x}})$**.** That is the reason why the operator is often given by its matrix. It is easier because the action of the operator in this case is just reduced to a matrix multiplication.

In a full form the formula $Y = AX$ is equivalent to n equalities

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n. \end{cases}$$

3.6. Kernel, defect, image and rank of an operator

Let $\mathcal{A}\colon \mathcal{L}^n \longrightarrow \mathcal{L}^n$ be a linear operator.

<u>The kernel of an operator $\mathcal{A}$</u> is a set of all vectors in $\boldsymbol{\mathcal{L}^n}$ that have a vector $\overline{\mathbf{0}}$ as their image.

Usually a kernel of an operator $\mathcal{A}$ is denoted as: $\mathcal{K}er\mathcal{A}$. So

$\mathcal{K}er\mathcal{A} = \{\overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{L}^n}\colon \ \mathcal{A}(\overline{\boldsymbol{x}}) = \overline{\mathbf{0}}\}.$

If $\mathcal{K}er\mathcal{A} = \{\overline{\mathbf{0}}\}$, then it is called ***<u>a trivial kernel.</u>***

<u>A defect of an operator $\mathcal{A}$</u> is a dimension of its kernel.

If a defect is denoted as k, then $\dim(\mathcal{K}er\mathcal{A}) = k$.

<u>The image of an operator $\mathcal{A}$</u> is a set of all its values. The image is denoted as $\mathcal{I}m\,\mathcal{A}$.

Let's try to read the formula below.

$\mathcal{I}m\,\mathcal{A} = \{\overline{\boldsymbol{y}} \in \boldsymbol{\mathcal{L}^n}\colon \ \exists\, \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{L}^n}(\mathcal{A}(\overline{\boldsymbol{x}}) = \overline{\boldsymbol{y}})\}.$

Here we have just put down the fact that the vector $\overline{\boldsymbol{y}}$ belongs to the kernel of the operator only in the case when it has at least one pre-image.

<u>The rank of the operator $\mathcal{A}$</u> is a dimension of its image.

If a rank of $\mathcal{A}$ is denoted by the letter r, then $\dim(\mathcal{I}m\,\mathcal{A}) = r$.

The following important equality is true:

$\dim(\mathcal{K}er\mathcal{A}) + \dim(\mathcal{I}m\,\mathcal{A}) = n.$

$$\Updownarrow$$

$$\boxed{k + r = n}$$

In fact, the common approach to the definition of a kernel and the image of the operator $\mathcal{A}$ can be explained in a following way. We fix any basis in space and we find the operator's matrix A in it. Then the equation $\mathcal{A}(\overline{\boldsymbol{x}}) = \overline{\mathbf{0}}$ will be equivalent to the matrix equation $AX = O$. In a full form it looks likea square system of linear equations with the matrix A. The space of this system's solutions is the kernel of the operator. However, each vector will be put down as a column of its coordinates in the given basis. Each fundamental system of solutions is a basis in the space of solutions, i.e. in the kernel of the operator. So the number of vectors in the fundamental system of solutions is k – the dimension of a kernel.

The number $(n-k)=rkA$**.** The basis columns of a matrix A are linearly independent and their number is equal to $(n-k)$. The vector $Y\in \mathcal{Im}\,\boldsymbol{\mathcal{A}}$ in the case (and in that case only!), when the system of equations $AX=Y$ has a solution. Using a theorem of Kronecker-Capelli it is possible only in the case when $\mathrm{rk}A=\mathrm{rk}(A|Y)$. The last formula means that Y should be a linear combination of basis. So the basis columns are the vectors of the basis of an operator's image, and those vectors are represented in the coordinate form. The image itself consists of any possible combinations of the basis columns. Besides, we proved that $r=rkA$ and we also found that $k+r=n$.

3.7. Transformation of vector coordinates and operator's matrix for the other basis

Transfer matrix

Let $\boldsymbol{\mathcal{L}}^n$ be a linear space. Let us choose two bases in this space. $\langle\boldsymbol{e}\rangle=\langle\bar{\boldsymbol{e}}_1,\bar{\boldsymbol{e}}_2,\dots,\bar{\boldsymbol{e}}_n\rangle$ и $\langle\boldsymbol{e}'\rangle=\langle\bar{\boldsymbol{e}}'_1,\bar{\boldsymbol{e}}'_2,\dots,\bar{\boldsymbol{e}}'_n\rangle$.

As $\langle\boldsymbol{e}\rangle$ is a basis, then the vectors of a basis $\langle\boldsymbol{e}'\rangle$ can be presented in a basis $\langle\boldsymbol{e}\rangle$**.** Let those transformations have a form:

$$\begin{aligned}\bar{\boldsymbol{e}}'_1&=c_{11}\bar{\boldsymbol{e}}_1+c_{21}\bar{\boldsymbol{e}}_2+\dots+c_{n1}\bar{\boldsymbol{e}}_n,\\ \bar{\boldsymbol{e}}'_2&=c_{12}\bar{\boldsymbol{e}}_1+c_{22}\bar{\boldsymbol{e}}_2+\dots+c_{n2}\bar{\boldsymbol{e}}_n,\\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots, \qquad (*)\\ \bar{\boldsymbol{e}}'_n&=c_{1n}\bar{\boldsymbol{e}}_1+c_{2n}\bar{\boldsymbol{e}}_2+\dots+c_{nn}\bar{\boldsymbol{e}}_n.\end{aligned}$$

In the matrix form the expressions $(*)$ look like: $\langle\boldsymbol{e}'\rangle=\langle\boldsymbol{e}\rangle C$. Here

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & c_{2n}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{in}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ c_{n1} & c_{n2} & \cdots & c_{nj} & \cdots & c_{nn}\end{pmatrix}$$

– is a so-called transfer matrix from a basis $\langle\boldsymbol{e}\rangle$ to a basis $\langle\boldsymbol{e}'\rangle$**.** In the j – th column of this matrix there are the coordinates of a vector $\bar{\boldsymbol{e}}'_j$ for its presentation in the basis $\langle\boldsymbol{e}\rangle$.

For example, if

$$\bar{\boldsymbol{e}}_1' = 2\bar{\boldsymbol{e}}_1 + 3\bar{\boldsymbol{e}}_2 - 7\bar{\boldsymbol{e}}_3,$$

$$\bar{\boldsymbol{e}}_2' = 5\bar{\boldsymbol{e}}_1 + 8\bar{\boldsymbol{e}}_2 + 9\bar{\boldsymbol{e}}_3,$$

$$\bar{\boldsymbol{e}}_3' = 4\bar{\boldsymbol{e}}_1 - 3\bar{\boldsymbol{e}}_2 - \bar{\boldsymbol{e}}_3,$$

then

$$C = \begin{pmatrix} 2 & 5 & 4 \\ 3 & 8 & -3 \\ -7 & 9 & -1 \end{pmatrix}.$$

Remark **So we see that the coefficients standing in front of the vector $\bar{\boldsymbol{e}}_1$ in all of the three equation above are now the elements of the first row in the matrix *C*. The coefficients in front of the $\bar{\boldsymbol{e}}_2$ are the elements of the second row in C, etc. That means that the column of the coefficients in front of the same vector in all the equations is a corresponding row in the transfer matrix.**

The transformation of vector coordinates and the operator's matrix in the other basis

Let $\langle \boldsymbol{e} \rangle = \langle \bar{\boldsymbol{e}}_1, \bar{\boldsymbol{e}}_2, \dots, \bar{\boldsymbol{e}}_n \rangle$ and $\langle \boldsymbol{e}' \rangle = \langle \bar{\boldsymbol{e}}_1', \bar{\boldsymbol{e}}_2', \dots, \bar{\boldsymbol{e}}_n' \rangle$ be two bases, $\bar{\boldsymbol{x}}$ be a vector and $\mathcal{A}$ is a linear operator. A vector $\bar{\boldsymbol{x}}$ can be presented either in the first, or in the second basis. The same goes on the operator's matrix. It can be put down in both bases. Let us denote the columns of the coordinates of a vector $\bar{\boldsymbol{x}}$ in $\langle \boldsymbol{e} \rangle$ and $\langle \boldsymbol{e}' \rangle$ in a following way:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad X' = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$$

Then $\bar{\boldsymbol{x}} = \langle \boldsymbol{e} \rangle X = \langle \boldsymbol{e}' \rangle X'$. If C is a transfer matrix from a basis $\bar{\boldsymbol{e}}_1, \bar{\boldsymbol{e}}_2, \dots, \bar{\boldsymbol{e}}_n$ to a basis $\bar{\boldsymbol{e}}_1', \bar{\boldsymbol{e}}_2', \dots, \bar{\boldsymbol{e}}_n'$, then $\langle \boldsymbol{e}' \rangle = \langle \boldsymbol{e} \rangle C$. So we have $\langle \boldsymbol{e} \rangle X = \langle \boldsymbol{e} \rangle (CX')$.

At last, as the vectors in the basis are linearly independent, we get

$$\boxed{X = CX'} . \qquad \text{or} \qquad \boxed{X' = C^{-1}X}$$

Now let A and A' be the matrices of the operator $\boldsymbol{\mathcal{A}}$ in the corresponding bases. Then on one hand, we have $\langle \boldsymbol{\mathcal{A}}\langle \boldsymbol{e}\rangle\rangle = \langle \boldsymbol{e}\rangle A$, and on the other hand, $\langle \boldsymbol{\mathcal{A}}\langle \boldsymbol{e}'\rangle\rangle = \langle \boldsymbol{e}'\rangle A'$. However, it is not hard to notice that $\langle \boldsymbol{\mathcal{A}}\langle \boldsymbol{e}'\rangle\rangle = \langle \boldsymbol{\mathcal{A}}\langle \boldsymbol{e}\rangle\rangle C$. That's why $\langle \boldsymbol{\mathcal{A}}\langle \boldsymbol{e}\rangle\rangle C = \langle \boldsymbol{e}'\rangle A'$. And so $\langle \boldsymbol{e}\rangle AC =$ $= \langle \boldsymbol{e}\rangle C\, A'$. Therefore, $AC = C\, A' \Leftrightarrow A' = C^{-1}AC$. So we got a formula:

$$\boxed{A' = C^{-1}AC}$$

and that is a formula of a matrix transformation for the other basis.

Practice Russian

Ма́трица перехо́да – transfer matrix

Practical task

1) This task is not very easy but it is rather creative. You have to create a 3x3 matrix with any numbers you like. Then you have to think about its coordinates in the 'old' basis (be careful with dimensions!) and to create a new basis. Your task is to get a matrix A' in the new basis.

3.8. Eigenvalues and eigenvectors of linear operators

Let $\boldsymbol{\mathcal{L}}^n$ be a linear space and let $\boldsymbol{\mathcal{A}}$ be a linear operator in that space.

A vector $\overline{\boldsymbol{x}} \neq \overline{\boldsymbol{0}}$ is called ***the eigenvector of an operator*** $\boldsymbol{\mathcal{A}}$, if we can find such a number λ that we can satisfy the equality $\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) = \lambda\overline{\boldsymbol{x}}$. The number λ is called ***the eigenvalue of a given operator*** when the condition above is fulfilled.

So the eigenvector of an operator is such a non-zero vector which has an image proportional to the vector itself or a zero image. So all non-zero vectors of an operator's kernel are its eigenvectors with the eigenvalue $\lambda = 0$.

Further, when an eigenvector has an eigenvalue λ, we will say that it *belongs* to this value.

The properties of eigenvalues and eigenvectors.

1) An eigenvector $\overline{\boldsymbol{x}}$ can't belong to different eigenvalues. Indeed, from $\lambda_1\overline{\boldsymbol{x}} = \lambda_2\overline{\boldsymbol{x}}$ we get $\lambda_1 = \lambda_2$, т. к. $\overline{\boldsymbol{x}} \neq \overline{\boldsymbol{0}}$.

2) Let us investigate a set of all the eigenvectors that belong to an eigenvalue λ, and we add there a vector $\overline{\boldsymbol{0}}$. All this set of vectors we denote as $\mathcal{N}_\lambda$. We would like to emphasize once again that $\overline{\boldsymbol{0}} \in \boldsymbol{\mathcal{N}}_\lambda$ but it is not an eigenvector.

Let us introduce the definition of a ***subspace*** – it is a space that is totally situated in the other space, whose points or elements are all situated in the parent space; it also inherits all the properties of a parent space.

A set $\mathcal{N}_\lambda$ will be a subspace in $\boldsymbol{\mathcal{L}}^n$. To prove it, it is just enough to mention that $\forall\, \overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{N}}_\lambda$ we have $\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) = \lambda\overline{\boldsymbol{x}} \Leftrightarrow \boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) = (\lambda\boldsymbol{\mathcal{I}})\overline{\boldsymbol{x}} \Leftrightarrow \boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) - (\lambda\boldsymbol{\mathcal{I}})\overline{\boldsymbol{x}} = \overline{\boldsymbol{0}} \Leftrightarrow (\boldsymbol{\mathcal{A}} - \lambda\boldsymbol{\mathcal{I}})\overline{\boldsymbol{x}} = \overline{\boldsymbol{0}} \Leftrightarrow \overline{\boldsymbol{x}} \in \mathcal{Ker}(\boldsymbol{\mathcal{A}} - \lambda\boldsymbol{\mathcal{I}})$**.** So $\boldsymbol{\mathcal{N}}_\lambda = \mathcal{Ker}(\boldsymbol{\mathcal{A}} - \lambda\boldsymbol{\mathcal{I}})$**.** And a kernel of an operator is a subspace (we can tell that $\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{0}}) = \overline{\boldsymbol{0}} = \lambda\overline{\boldsymbol{0}}$).

$\boldsymbol{\mathcal{N}}_\lambda$ is usually called ***the eigensubspace of an operator.*** *There is an important property:* if $\overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{N}}_\lambda$, then $\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) = \lambda\overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{N}}_\lambda$**.** Such subspaces are called the invariant subspaces of an operator.

The subspace $\boldsymbol{\mathcal{H}} \subset \boldsymbol{\mathcal{L}}^n$ is called ***the invariant subspace of an operator*** $\boldsymbol{\mathcal{A}}$, if from $\overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}$ we can conclude $\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) \in \boldsymbol{\mathcal{H}}$**.** ■

The importance of those subspaces is in the fact that the operator woks in them independently – it doesn't depend on its properties in the other parts of space.

If an operator is investigated only on its invariant subspace $\mathcal{H}$, then we will denote it as $\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}$ and call it a constrint of the operator $\boldsymbol{\mathcal{A}}$ on $\boldsymbol{\mathcal{H}}$**.** So, according to the definition, we have:

$$\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\overline{\boldsymbol{x}}) = \boldsymbol{\mathcal{A}}(\overline{\boldsymbol{x}}) \qquad \forall\overline{\boldsymbol{x}} \in \boldsymbol{\mathcal{H}}.$$

If vectors $\overline{\boldsymbol{x}}_1, \overline{\boldsymbol{x}}_2, \dots, \overline{\boldsymbol{x}}_p$ belong to p different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_p$, then they are *linearly independent.*

If an operator $\boldsymbol{\mathcal{A}}$ has n different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ then the vectors $\overline{\boldsymbol{s}}_1, \overline{\boldsymbol{s}}_2, \dots, \overline{\boldsymbol{s}}_n$ that belong to those eigenvalues, are linearly independent and can be taken as a basis.

A basis that consists only of the eigenvectors of an operator, is called ***the eigenbasis*** of that operator.

Remark **Here it does not matter whether the eigenvalues of an operator are differ from each other or not.**

A matrix has a diagonal form in the eigenbasis, and the eigenvalues of an operator stand on the main diagonal. Indeed, if the eigenbasis consists of the vectors $\overline{\boldsymbol{s}}_1, \overline{\boldsymbol{s}}_2, \dots, \overline{\boldsymbol{s}}_n$ with the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ correspondently, then

$$\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{s}}_1) = \lambda_1 \cdot \overline{\boldsymbol{s}}_1 + 0 \cdot \overline{\boldsymbol{s}}_2 + \cdots + 0 \cdot \overline{\boldsymbol{s}}_n,$$
$$\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{s}}_2) = 0 \cdot \overline{\boldsymbol{s}}_1 + \lambda_2 \cdot \overline{\boldsymbol{s}}_2 + \cdots + 0 \cdot \overline{\boldsymbol{s}}_n,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$\boldsymbol{\mathcal{A}}(\overline{\boldsymbol{s}}_n) = 0 \cdot \overline{\boldsymbol{s}}_1 + 0 \cdot \overline{\boldsymbol{s}}_2 + \cdots + \lambda_n \cdot \overline{\boldsymbol{s}}_n,$$

and so we have

$$A_s = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

The inverse property is also true: if in a basis the operator's matrix is diagonal, then it is the eigenbasis, and there are eigenvalues on a main diagonal.

Practice Russian

Со́бственные ве́кторы – the eigenvectors

Со́бственные значе́ния – the eigenvalues

3.9. The characteristic polynomial and its invariance

Let us describe the general method of finding the eigenvalues and eigenvectors that belong to those eigenvalues.

The number λ and a vector $\bar{x} \neq \bar{0}$ are the eigenvalue and the eigenvector of an operator $\mathcal{A}$ only in the case when $\bar{\boldsymbol{x}} \in \mathcal{K}er(\boldsymbol{\mathcal{A}} - \lambda\boldsymbol{\mathcal{I}})$**,** i.e. $\bar{\boldsymbol{x}}$ satisfies the equation $(\boldsymbol{\mathcal{A}} - \lambda\boldsymbol{\mathcal{I}})\bar{\boldsymbol{x}} = \bar{\boldsymbol{0}}$**.** If we fix a basis in the space and denote an operator's matrix in this basis as A, the column of coordinates for the vector $\bar{x}$ will be denoted as через X , then , considering the fact that in any basis a matrix of the identity operator $\boldsymbol{\mathcal{I}}$ is an identity matrix **I,** the operator equation will take a form of the equivalent matrix equation

$$(A - \lambda I)X = O.$$

This is a square homogeneous system of linear equations. That is why the number λ will be the eigenvalue of an operator if and only if we have a non-trivial consistent system. It is possible only in the case when $|A - \lambda I| = 0$. This formula is called ***a characteristic equation.***

In a detailed form it looks like:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

So the number λ is the eigenvalue of an operator only when it is a root of the characteristic equation.

If we look at the determinant standing in the left-hand part of a characteristic equation, then it is not hard to understand that it turns out to be a $\boldsymbol{P}_n(\lambda)$ with the higherst order n:

$$\boldsymbol{P}_{\boldsymbol{n}}(\lambda) = (-1)^n\lambda^n + \cdots + |A|.$$

This polynomial is called a characteristic polynomial. Its roots (and its roots only!) are the eigenvalues of an operator.

Let λ_0 be one of that roots. Then, in order to find the eigenvectors that belong to it, we get a system of homogeneous equations $(A - \lambda_0 I)X = O$. The general solution of this system fills in the eigensubspace $\boldsymbol{\mathcal{N}}_{\lambda_0}$ of an operator. Each fundamental system of solutions will be a basis in this space. However, that same basis is investigated when we find the eigenvectors. If we know a basis, then the eigenvectors are known.

There is another question: does the characteristic polynomial change when we change the basis? The answer is: no, it is the same in all the bases. This property is called ***the invariance*** of a polynomial. Let us prove this fact.

Let the matrix A' be the matrix of an operator in the other basis and let C be the transfer matrix for the presentation in that basis. Then we get the sequence of equalities:

$$|A' - \lambda I| = |C^{-1}AC - \lambda C^{-1}C| = |C^{-1}(A - \lambda I)C| =$$
$$= |C^{-1}||A - \lambda I| \cdot |C| =$$
$$= |C^{-1}||C||A - \lambda I| = 1 \cdot |A - \lambda I| = |A - \lambda I|.$$

So the invariance is proved.

Remark **All the coefficients of a characteristic equation are invariant in terms of changing the basis. The determinant of the corresponding matrix is invariant, too.**

Example Find the eigenvalues and eigenvectors of a matrix.

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

$$|A - \lambda I| = 0$$

Let us form a matrix $A - \lambda I$:

$$\begin{pmatrix} 3-\lambda & 1 & 2 \\ 1 & 3-\lambda & 0 \\ 2 & 0 & 3-\lambda \end{pmatrix}$$

After calculation of a determinant we get:

$$\lambda_1 = 3 + \sqrt{5};\ \lambda_2 = 3;\ \lambda_3 = 3 - \sqrt{5}$$

We put down a matrix form: $A\varphi = \lambda\varphi$ where φ is an eigenvector[1]

$$(A - \lambda I)\ \varphi = 0$$

For the first eigenvector φ_1 we have:

$$(A - \lambda_1 I)\ \varphi_1 = 0$$

$$\begin{pmatrix} 3-(3+\sqrt{5}) & 1 & 2 \\ 1 & 3-(3+\sqrt{5}) & 0 \\ 2 & 0 & 3-(3+\sqrt{5}) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

From this we get:

$$\begin{cases} -\sqrt{5}x + y + 2z = 0 \\ x - \sqrt{5}y = 0 \\ 2x - \sqrt{5}z = 0 \end{cases}$$

We know in advance that this system will have an infinite number of solutions so one variable can be taken as a parameter. We will have 3 variables: x, y, z. Let $z = t$

We take 2 linearly independent equations of a system.

$$\begin{cases} x - \sqrt{5}y = 0 \\ 2x - \sqrt{5}t = 0 \end{cases}$$

So

$$2\sqrt{5}y - \sqrt{5}t = 0;\ y = 0{,}5t;\quad x - 0{,}5\sqrt{5}t = 0;\ x = \frac{\sqrt{5}}{2}t$$

We get a system:

$$\begin{cases} x = \frac{\sqrt{5}}{2}t \\ y = \frac{1}{2}t \\ z = t \end{cases};\ \varphi_1 = \begin{pmatrix} \frac{\sqrt{5}}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

---

[1] Previously we used the denotation 'X' for an eigenvector; however, for the practical tasks it is easier to use Greek letters because they are not so widely-used as Latin ones and we can easily recognize the denotations

Now we investigate the second eigenvector φ_2

$$(A - \lambda_2 I)\ \varphi_2 = 0$$

$$\begin{pmatrix} 3-3 & 1 & 2 \\ 1 & 3-3 & 0 \\ 2 & 0 & 3-3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

$$\begin{cases} y + 2z = 0 \\ x = 0 \\ z = t \end{cases};\quad \begin{cases} y = -2t \\ x = 0 \\ z = t \end{cases};\quad \varphi_2 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}$$

The next step:

$$(A - \lambda_3 I)\ \varphi_3 = 0$$

$$\begin{pmatrix} 3-(3-\sqrt{5}) & 1 & 2 \\ 1 & 3-(3-\sqrt{5}) & 0 \\ 2 & 0 & 3-(3-\sqrt{5}) \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$$

$$\begin{cases} \sqrt{5}x + y + z = 0 \\ x + \sqrt{5}y = 0 \\ 2x + \sqrt{5}z = 0 \end{cases};\quad \begin{cases} x + \sqrt{5}y = 0 \\ 2x + \sqrt{5}z = 0 \\ z = t \end{cases}$$

$$\begin{cases} x = -\frac{\sqrt{5}}{2}t \\ y = \frac{1}{2}t \\ z = t \end{cases};\quad \varphi_1 = \begin{pmatrix} -\frac{\sqrt{5}}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

Practice Russian

Характеристический многочлéн – characteristic polynomial
Характеристическое уравнение – characteristic equation

Creative task

You have just read the last chapter of a theoretical material. However, this part is rather difficult for some 'freestyle' creative tasks. So we suggest you to make a crossword puzzle based on the main definition and terms that you have met in this chapter. However, if you want to, you can also choose any other form!

4. SOME PROBLEMS OF A MATRIX THEORY SOLVED IN MATHCAD PROGRAM

In this chapter we will get acquainted with the interface, main commands and basic properties of Mathcad program and we will learn how to use its apparatus for solving some typical problems of the matrix theory. We will present some practical tasks and in the process of getting the solutions of those problems, we will show use some algorithms of work with Mathcad software. In this chapter we will not give any personal tasks; however, you can train yourselves in using the program or to fulfil some of your personal tasks in the previous chapters according to the recommendations of your lecturer. In our examples we will work mostly with the graphic interface of the program.

Task 1

a) Add matrix A to a matrix B.

b) Multiply the result by a scalar λ.

c) Multiply the matrices A and B

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix}, B = \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & 7 \\ 8 & -9 & 10 \end{pmatrix},$$

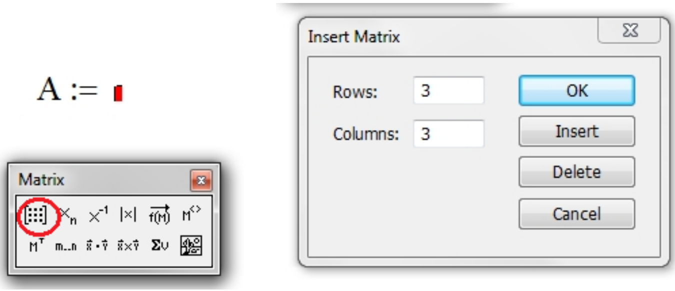
$$\lambda = 5$$

Firstly, we need to find a toolbar ‘Vector and matrix’:

Afterwards we type in the blank field a letter A and put down a colon sign[2]. That will automatically assign a value to your introduced object. After you see: ‘A:=’ on the display and there appears

[2] You can also use the equality sign on your keyboard to assign a value; however, it is a more multi-functional button so further we will speak about the colon button when we need to assign some values

a place to assign a value, you open the 'Vector and matrix' toolbar and choose the sign for a matrix creation, then you choose the number of rows and columns:



So you put down the numerical values of the elements into a matrix *A*. Then you should repeat the same procedure to create a matrix *B*:

$$A := \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix} \qquad B := \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & 7 \\ 8 & -9 & 10 \end{pmatrix}$$

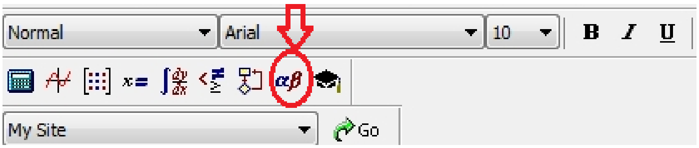
Then you should type the resulting matrix *C*

$$C: = A + B =$$

When you put down the equality sign (after the sum, without the colon sign), the result will be calculated automatically.

$$C := A + B = \begin{pmatrix} 4 & 0 & 1 \\ 4 & 4 & 11 \\ 9 & -7 & 16 \end{pmatrix}$$

Then you should introduce the coefficient λ. You need to find a toolbar with Greek symbols:



You type a letter λ and so you assign the value for it with the colon sign again: all you have to do is just to type a proper number. Then you introduce the matrix D:

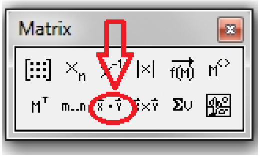
$$D: = \lambda \cdot A$$

Here and afterwards or the operation of multiplication you should use the symbol * on your keyboard. Then all you have to do is to put the equality sign and to get the result:

$$\lambda := 5 \qquad D := \lambda \cdot A = \begin{pmatrix} 15 & 10 & 5 \\ 0 & -5 & 20 \\ 5 & 10 & 30 \end{pmatrix}$$

Then you have to multiply 2 matrices A and B. You can just put down $A \cdot B$ by the symbol * we've just introduced. However, there is another useful button on a toolbar.

You need to find a 'Dot product' on a 'Vector and matrix Toolbar':



So you will see a dot sign and two places to put down the two matrices on the display. You put down A and B into the empty places and then you put the equality sign. The result will appear automatically.

$$A \cdot B = \begin{pmatrix} 19 & -5 & 24 \\ 28 & -41 & 33 \\ 57 & -46 & 74 \end{pmatrix}$$

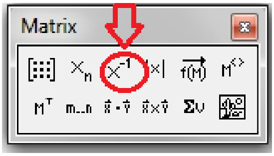
Your first task is to do the same algorithm with different numeric values (you should change the elements of a matrix and a value of λ.

<u>Task 2</u> Calculate the inverse matrix to a matrix A from a previous task

You introduce the matrix A:

$$A: = \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1 & 4 \\ 1 & 2 & 6 \end{pmatrix}$$

Then you put down the letter *A* below and click the 'Inverse matrix' button on a 'Vector and matrix' toolbar:

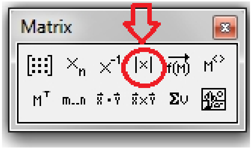


Afterwards you just put down the equality sign and you get the ready inverse matrix!

$$A^{-1} = \begin{pmatrix} 0.424 & 0.303 & -0.273 \\ -0.121 & -0.515 & 0.364 \\ -0.03 & 0.121 & 0.091 \end{pmatrix}$$

Task 3 Find the determinant for the matrix *A* from the previous task.

You just have to use the 'Determinant' button on a 'Vector and matrix toolbar':



So after cliking this command you will get the determinant sign on the main screen and all you have to do is to put the letter *A* into the sign of the determinant and to put down the equality sign:

$$|A| = -33$$

Solving the systems of linear equations using the Mathcad software

There are some useful commands that help us to solve the systems of linear equations in Mathcad.

The 'Given-Find' unit

Helps to define the column of unknown variables. Firstly, we assign zero values to the variables x_0, x_1, x_2 (those are just the initial values to introduce the variables and they will not influence the result). Then we type the command 'Given' and put down the given system of equations (no special brackets demanded). For this system we don't use the simple equality sign. We use the other special operator *'Equal To'*: **'Ctrl+='** on your keyboard (it is depicted as the bold-font equality sign in the program) – it allows us to get the necessary result in spite of having assigned the zero values to x_1, x_2, x_3. After that, we put down the word ‘Find’ and put into the brackets the variable (or variables) that we need to find. So when we write: ‘Find (x)’, the column of the results for x_0, x_1, x_2 will be calculated automatically:

$$x_0 := 0 \qquad x_1 := 0 \qquad x_2 := 0$$

Given

$$7x_0 + 4x_1 - 8x_2 = 3$$

$$3x_0 - 2x_1 + 5x_2 = 7$$

$$5x_0 - 3x_1 - 4x_2 = -12$$

$$\text{Find}(x) = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

We can also use the *‘Given-Find’* unit for a matrix form of the system. Then we just should define a matrix *A*, a matrix *B* and give the initial values for a column of variables *X*. Then we put down the general equation in the matrix form: *AX* = *B*. Again, we do not use the simple equality sign – we put down the operator ‘equal to’ (see below). Then we type ‘Find (*X*)’, and the result appears automatically.

Given

$$A := \begin{pmatrix} 7 & 4 & -8 \\ 3 & -2 & 5 \\ 5 & -3 & -4 \end{pmatrix} \qquad B := \begin{pmatrix} 3 \\ 7 \\ -12 \end{pmatrix}$$

$$X := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad A \cdot X = B \qquad \text{Find}(X) = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

‘Lsolve’ command For this command we should previously put down the matrix *A* and the column of the right-hand parts *B* (we investigate the system of linear equations *AX* = *B*)

Then we put down *lsolve (A, b)* and the program calculates the result, i.e. the column vector of the solutions of the system.

$$A := \begin{pmatrix} 7 & 4 & -8 \\ 3 & -2 & 5 \\ 5 & -3 & -4 \end{pmatrix} \qquad B := \begin{pmatrix} 3 \\ 7 \\ -12 \end{pmatrix}$$

$$\text{lsolve}(A,B) = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

‘Augment’, ‘rref’ and ‘submatrix’ commands

We use the command *‘augment’* to form the augmented matrix that consists of the coefficients’ matrix of a system and a column of the right-hand parts of a system.

So we firstly put the command *ORIGIN:= 1* – that means the numeration for the rows of the matrix will start from the number ‘1’.

Then we input the matrices *A* and *B*. Afterwards we create a new augmented matrix: we may call it *Au* and put down: *Au:= augment(A, B)*. After that we put down ‘*Au*=’ and the augmented matrix is generated by the program.

Then we should use the program operator *‘rref’* that uses the Gaussian method transforming directly the coefficient matrix according to the method and then solving the system of equations consequently. We create a new matrix to get a result in it: *C:= rref(Au)*.

Then we use the function ‘*submatrix*’ that helps us to take a submatrix of a necessary dimension from a larger matrix. We use the *X* letter to get a necessary result for our column of variables *X*.

The common form of a command is: $submatrix(M, r_i, r_j, c_i, c_j)$ where M is a matrix from which we ‘cut’ a submatrix; r_i, r_j are the numbers of rows and c_i, c_j are the numbers of columns. So r_i is the first number of a row that is going to be included into the matrix and r_j is the last number of a row that is going to be included to the matrix. The similar rule is for columns c_i, c_j. It is clear that in both cases $i \leq j$. However, there is an important peculiarity: in many versions of Mathcad the default settings assign the number '0' to a first row and a first column so the *submatrix* command will look like that:

$$submatrix(M, r_{i-1}, r_{j-1}, c_{i-1}, c_{j-1})$$

where *i,j* are the real indexes of rows and columns (they start from 1). In the screenshot below we use the default setting when the program assigns a zero number to the first row/column.

The command submatrix for our task looks is the following:

X:= submatrix (C, 0, 2, 3, 3) (we take row 1 and 3 and the 4-th column for our submatrix. You have already guessed that it is the last column of answers)

The whole code looks like:

$$A := \begin{pmatrix} 7 & 4 & -8 \\ 3 & -2 & 5 \\ 5 & -3 & -4 \end{pmatrix} \qquad B := \begin{pmatrix} 3 \\ 7 \\ -12 \end{pmatrix}$$

$$Au := augment(A,B) = \begin{pmatrix} 7 & 4 & -8 & 3 \\ 3 & -2 & 5 & 7 \\ 5 & -3 & -4 & -12 \end{pmatrix}$$

$$C := rref(Au) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$x := submatrix(C,0,2,3,3) = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

5. CONCLUSION

Linear algebra is really an important part of Higher Mathematics Course at the university. You got acquainted only with the basic concepts of this branch – the material of this textbook is optimized for one academic semester. However, in your other courses at the technical university, you will often use the acquired knowledge and you will also broaden your horizons in linear algebra. Three themes presented in the book – matrix theory, analytical geometry and theory of linear spaces – together make the necessary basis for your further studies. Some theorems are given without a proof – it is really a good exercise to try proving them yourselves.

To get prepared for your examination tasks, it is very important to read once again all the basic definitions and theorems given in the book. It is necessary to clear out all the terms, formulae and algorithms that are difficult for your understanding – do not hesitate to ask your university lecturers to explain some ways of solving the typical problems. To check yourselves, you should answer the review questions given after each paragraph.

Some typical problems solved in Mathcad program, are only a small part of your possibilities. Try to get used to this software – however, you can use any other mathematical packages you like. In a career of an engineer it is extremely important to be a highly-qualified specialist, and nowadays that also means using the professional computer programs and to have an ability to find the necessary information as soon as possible.

We hope you have enjoyed reading about prominent mathematicians and their biographies inspired you for the new researches and your desire to learn more and more. We also hope it was interesting to fulfil the creative tasks. If you have some ideas to share with the authors, please contact us by e-mail: BesovaMI@mpei.ru. We also will be glad if you send us your creative works dedicated to our main themes (you have the 'freestyle' creative tasks at the end of each chapter, except the last one). Good luck in your future studies! And don't forget about the wonderful words of a famous Russian scientist Mikhail Lomonosov (1711–1765): 'Mathematics is already good because it brings your mind in order'!

BIBLIOGRAPHY

Basic literature

1. Beklemishev D.V. A course of analytical geometry and linear algebra (in Russian): Teaching guide. St. Petersburg, Lan', 2015

2. Gorlach B.A., Rostova E.P. Linear algebra and analytical geometry (in Russian). St. Petersburg, Lan', 2020

3. Proskuryakov I.V. Problem book on Linear Algebra (in Russian). St. Petersburg, Lan', 2021

4. Umnov A.E. Analytical geometry and linear algebra (in Russian). Moscow, MFTI, 2011.

5. Zimina O.V., Kirillov A.I., Salnikova T.A. Problem book. Higher mathematics (in Russian). Moscow, Fizmatlit, 2005

Additional literature

1. Electronic source: MacTutor History of Mathematics Archive (The website of School of Mathematics and Statistics; University of St. Andrews, Scotland)

https://mathhistory.st-andrews.ac.uk/Biographies

2. Electronic source: PTC Mathcad Help Guide

http://support.ptc.com/help/mathcad/r6.0/en/

Educational edition

Margarita I. **Besova**

Sergey F. **Kudin**

BASIC COURSE OF LINEAR ALGEBRA AND ANALYTICAL GEOMETRY

Teaching guide

Computer layout L.I. Vecelovsky

Signed for printing 17.01.2022. Offset printing Format 60×90/16
Printer's sheets 8,75 Edition of 100 copies Edit. № 21u-138 Offer №

Paste-up layout is prepared in Editing Department of NRU 'MPEI'
111250, Moscow, Krasnokazarmennaya str., 14
Printed in the Printing office of NRU 'MPEI'
111250, Moscow, Krasnokazarmennaya str., 13

Учебное издание

Бесова Маргарита Ильинична

Кудин Сергей Фёдорович

БАЗОВЫЙ КУРС ЛИНЕЙНОЙ АЛГЕБРЫ И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Учебное пособие

Компьютерная верстка Л.И. Веселовский

Подписано в печать 17.01.2022. Печать офсетная Формат 60×90/16
Печ. л. 8,75 Тираж 100 экз. Изд. № 21у-138 Заказ № ___

Оригинал-макет подготовлен в РИО НИУ «МЭИ».
111250, г. Москва, ул. Красноказарменная, д. 14.
Отпечатано в типографии НИУ «МЭИ».
111250, г. Москва, ул. Красноказарменная, д. 13

ДЛЯ ЗАМЕТОК