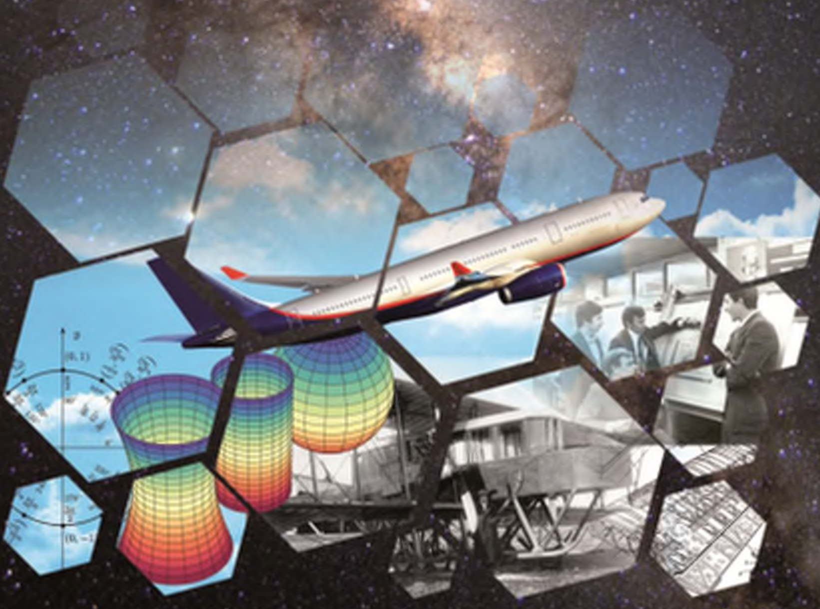
MOSCOW AVIATION INSTITUTE | NATIONAL RESEARCH UNIVERSITY

Textbook for Bachelor of Science Students

A.S. BORTAKOVSKIY
A.V. PANTELEEV
V.N. PANOVSKIY

LINEAR ALGEBRA AND ANALYTIC GEOMETRY



А. С. БОРТАКОВСКИЙ

А. В. ПАНТЕЛЕЕВ

В. Н. ПАНОВСКИЙ

ЛИНЕЙНАЯ АЛГЕБРА И АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ

A. S. BORTAKOVSKIY

A. V. PANTELEEV

V. N. PANOVSKIY

LINEAR ALGEBRA AND ANALYTIC GEOMETRY

MOSCOW AVIATION INSTITUTE
(NATIONAL RESEARCH UNIVERSITY)

A. S. BORTAKOVSKIY
A. V. PANTELEEV
V. N. PANOVSKIY

LINEAR ALGEBRA AND ANALYTIC GEOMETRY

Moscow
«Dobroe Slovo»
2019

МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

А. С. БОРТАКОВСКИЙ

А. В. ПАНТЕЛЕЕВ

В. Н. ПАНОВСКИЙ

ЛИНЕЙНАЯ АЛГЕБРА
И
АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ

Москва

Издательство «Доброе слово»

2019

УДК 512 (075.8)
ББК 22.143я73
Б82

Б82 Линейная алгебра и аналитическая геометрия: Учебное пособие / А.С. Бортаковский, А.В. Пантелеев, В.Н. Пановский. – М.: Издательство «Доброе слово», 2019. – 256 с.: ил.

ISBN 978-5-89796-649-4

Пособие предназначено для проведения практических занятий по курсу линейной алгебры и аналитической геометрии. Приведены основные понятия и методы решения задач по всем разделам курса: матрицы и определители, системы линейных алгебраических уравнений, квадратичные формы, линейные пространства, векторная алгебра, системы координат, преобразования плоскости и пространства, уравнения линий и поверхностей первого и второго порядков. В каждом разделе кратко изложены основные теоретические сведения, приведены решения типовых примеров и задачи для самостоятельного решения, в том числе зависящие от параметров m (порядкового номера учебной группы в потоке) и n (номера студента по списку группы).

Для студентов технических вузов и университетов.

B82 Linear Algebra and Analytic Geometry: textbook / A.S. Bortakovskiy, A.V. Panteleev, V.N. Panovskiy. – M.: «Dobroe Slovo», 2019. – 256 p.

Moscow Aviation Institute (National Research University). Moscow. E-mail: kaf805@mai.ru

Given textbook is written for student’s self-study of the course of linear algebra and analytic geometry. Material, that is described in this manual, covers all basic sections of linear algebra (including matrices and matrix operations, determinants, principal minors and matrix rank, inverse matrix, systems of ordinary linear equations, eigenvalues and eigenvectors, quadratic forms) and analytic geometry (including vector algebra, coordinate systems, algebraic lines and surfaces, linear spaces, mappings, and transformations). All material is supported by sufficient number of examples with detailed solutions and exercises depending on the parameters m (the sequence number of the group) and n (the student number in the group list).

For students of MAI International Bachelor's Degree Programs.

CONTENTS

PART I. LINEAR ALGEBRA

CHAPTER 1. MATRICES AND MATRIX OPERATIONS

1.1. NUMERICAL MATRICES

An $m \times n$ ***matrix*** A is a set of $m \cdot n$ numbers, represented by a rectangular array of m rows and n columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \text{ or } A = (a_{ij}),\ i = 1,\dots,m;\ j = 1,\dots,n.$$

Numbers, which form the matrix, are called ***matrix elements***: a_{ij} – element, that is placed on the intersection of the i-th row and the j-th column. Matrix elements are expected to be real numbers.

Example 1.1. Determine matrix sizes m and n:

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

□ Matrix A has sizes 3×2, matrix B – 2×4, c – 1×3, d – 2×1.■

Two matrices A and B are called ***equal*** ($A = B$) if they have the same sizes ($m \times n$), and their corresponding elements are equal:

$$a_{ij} = b_{ij},\quad i = 1,\dots,m;\quad j = 1,\dots,n.$$

In general case, a matrix (with sizes $m \times n$) is called ***rectangular***. In particular, if a matrix consists of a single column ($n = 1$) or a single row ($m = 1$), it is called ***column-matrix*** or ***row-matrix*** (or simply ***column*** or ***row***) respectively. Row-matrices and column-matrices are frequently denoted by lowercase letters (in example 1.1: c – row, d – column). A matrix of sizes 1×1 is simply a number (the only element of a matrix).

Matrix with the same number of rows (m) and columns (n) is called ***square matrix*** (of ***n-th order***). Elements $a_{11}, a_{22}, \dots, a_{nn}$ form the ***main diagonal*** of square matrix (dashed line on Fig. 1.1 which connects the upper-left corner of the matrix (element a_{11}) with the lower-right corner (element a_{nn})). The diagonal, which connects the lower-left corner (element a_{n1}) with the upper-right corner (element a_{1n}), is called ***secondary***.

$$\begin{pmatrix} a_{11} & & a_{1n} \\ & \ddots \; ⋰ & \\ a_{n1} & & a_{nn} \end{pmatrix}\begin{array}{l} \text{Secondary diagonal} \\ \\ \text{Main diagonal} \end{array}$$

Figure 1.1

Square matrix

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

with zero non-diagonal elements is called ***diagonal***, and denoted by $diag(a_{11}, a_{22}, \dots, a_{nn})$.

A special case of square matrix is a matrix

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

which is called ***identity*** (of n-th order) matrix. It is denoted by E (or E_n).

If all elements of a square matrix which are situated below (above) main diagonal are equal to zero, such a matrix is called ***upper-triangular*** (***lower-triangular***). Fig. 1.2 demonstrates diagonal and triangular matrices (now and later we will suppose that matrix's part denoted by O symbol consists of zero elements, and parts denoted by * symbol and lines consists of arbitrary elements). Notice that a diagonal matrix, particularly an identity matrix, is a lower and an upper-triangular simultaneously.

Diagonal Upper-triangular Lower-triangular

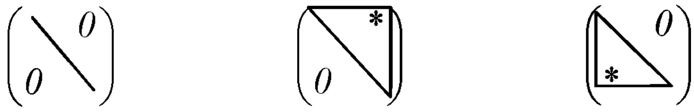


Figure 1.2

Matrix with all elements equal to zero is called ***zero matrix***.

Example 1.2. Define matrix types

$$A=\begin{pmatrix}0&0&0\\0&0&0\end{pmatrix},\ B=\begin{pmatrix}1&2&1\\0&4&5\\0&0&9\end{pmatrix},\ C=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ D=\begin{pmatrix}0&0\\0&0\end{pmatrix},\ E=\begin{pmatrix}1&0\\0&1\end{pmatrix},$$

$$F=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\quad G=\begin{pmatrix}1&0&0\\2&3&0\\4&5&6\end{pmatrix},\quad H=\begin{pmatrix}2&0&0\\0&2&0\\0&0&1\end{pmatrix}.$$

□

A – rectangular zero matrix of sizes 2×3;

B – 3rd order upper-triangular matrix;

C – 2nd order lower-triangular matrix;

D – 2nd order square zero matrix;

E – 2nd order identity matrix;

F – 3rd order identity matrix;

G – 3rd order lower-triangular matrix;

H – 3rd order diagonal matrix. ■

1.2. MATRIX OPERATIONS

1.2.1. Matrix Addition

Let $A=(a_{ij})$ and $B=(b_{ij})$ be matrices of the same sizes $m\times n$. Matrix $C=(c_{ij})$ of the same sizes $m\times n$ is called ***the sum of matrices*** A and B if its elements are equal to the sum of the corresponding elements of matrices A and B:

$$c_{ij}=a_{ij}+b_{ij},\ i=1,\ldots,m;\ j=1,\ldots,n.$$

The sum is denoted by $C = A + B$. Matrix addition is defined only for matrices of the same sizes and is calculated element-wise. From the definition it comes that *it is possible to sum only matrices of the same sizes*: e.g. it is impossible to find sums

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 \\ 6 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Example 1.3. Find the sum of two matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

□ Adding the corresponding elements, we get

$$C = \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}}_{(3\times 2)} + \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}}_{(3\times 2)} = \underbrace{\begin{pmatrix} 1+0 & 2+1 \\ 3+1 & 4+0 \\ 5+0 & 6+0 \end{pmatrix}}_{(3\times 2)} = \begin{pmatrix} 1 & 3 \\ 4 & 4 \\ 5 & 6 \end{pmatrix}. \ \blacksquare$$

1.2.2. Multiplication of Matrix by Number

A ***product of a matrix*** $A = (a_{ij})$ ***and a number*** λ is the matrix $C = (c_{ij})$ of the same sizes as matrix A which elements are equal to the product of number λ and the corresponding element of matrix A:

$$c_{ij} = \lambda \cdot a_{ij}, \quad i = 1,\dots,m; \quad j = 1,\dots,n.$$

Product is denoted by $\lambda \cdot A$ or $A \cdot \lambda$. Multiplication of a matrix by a number is done element-wise. It is possible to multiply any matrix by a number: each element should be multiplied by this number.

Example 1.4. Find the product of matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$ and number 2.

□ Multiplying each element of matrix A by 2 we get

$$C = 2\cdot A = A\cdot 2 = 2\cdot\begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix} = \begin{pmatrix}1\cdot 2 & 2\cdot 2\\ 3\cdot 2 & 4\cdot 2\\ 5\cdot 2 & 6\cdot 2\end{pmatrix} = \begin{pmatrix}2 & 4\\ 6 & 8\\ 10 & 12\end{pmatrix}. \blacksquare$$

Matrix $(-1)\cdot A$ is called ***opposite*** matrix of A and denoted by $(-A)$. Sum of matrixes B and $(-A)$ is called ***difference*** and denoted by $B - A$.

To find difference $B - A$ it is necessary to subtract elements of matrix A from the corresponding elements of matrix B. Subtraction is correct only for matrixes of the same sizes.

Example 1.5. Let

$$A = \begin{pmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{pmatrix}, \quad B = \begin{pmatrix}0 & 1\\ 1 & 0\\ 0 & 0\end{pmatrix}.$$

Find differences $B - A$ and $A - B$.

□ Subtracting the corresponding elements, we get

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

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

Linear matrix operations

There are two linear matrix operations:

1) matrix addition;

2) multiplication of a matrix by a number.

Properties of linear matrix operations coincide with the properties of addition (subtraction) of algebraic expressions (e.g. polynomials) and multiplication of an algebraic expression by a number.

For any matrices A, B, C of the same sizes and arbitrary numbers α, β the following equations are correct:

1) $A+B=B+A$;
2) $(A+B)+C=A+(B+C)$;
3) $\alpha\cdot(A+B)=\alpha\cdot A+\alpha\cdot B$;
4) $(\alpha+\beta)\cdot A=\alpha\cdot A+\beta\cdot A$;
5) $(\alpha\cdot\beta)\cdot A=\alpha\cdot(\beta\cdot A)$;
6) $1\cdot A=A$.

1.2.3. Matrix Multiplication

Let matrix $A=(a_{ij})$ of sizes $m\times p$ and $B=(b_{ij})$ of sizes $p\times n$. A matrix C of sizes $m\times n$ with elements c_{ij} that are calculated by the following formula

$$c_{ij}=a_{i1}\cdot b_{1j}+a_{i2}\cdot b_{2j}+\ldots+a_{ip}\cdot b_{pj},\quad i=1,\ldots,m;\ j=1,\ldots,n,$$

is called a ***product*** of matrices A and B and denoted by $C=AB$.

Multiplication of matrix A by matrix B is defined only for ***consistent*** matrices, i.e. matrices that satisfy the following property: number of columns of matrix A is equal to number of rows of matrix B:

$$\underset{m\times n}{C}=\underset{m\times p}{A}\cdot\underset{p\times n}{B}.$$

Let's consider the *procedure of finding a matrix product* in detail.

To find the element c_{ij} on the intersection of the i-th row and the j-th column of matrix C it is necessary to separate out the i-th row of matrix A and the j-th column of matrix B. They consist of the same number of elements because matrixes A and B are consistent.

Then it is required to find the sum of all pairwise products of the corresponding elements: the first element of the i-th row is multiplied by the first element of the j-th column, the second element of the i-th row is multiplied by the second element of the j-th column, and etc. Finally, the results are summed up.

In the product $A\cdot B$ matrix A is called the left-side multiplier for B, and it is said that matrix B is multiplied by matrix A ***from the left***. In a similar manner matrix

B is called the right-side multiplier for A, and it is said that matrix A is multiplied by matrix B ***from the right***.

Note that in general case $A \cdot B \neq B \cdot A$, but there are square matrices which product is unaffected by multiplier permutation.

Matrices A and B are called the ***permutation matrices*** if

$$A \cdot B = B \cdot A.$$

Permutation matrices can only be square matrices of the same order. In particular, it can be showed that diagonal matrices of the same order are permutation matrices.

For every square matrix A of order n the following equations are correct:

$$A \cdot E = E \cdot A = A$$

where E is an identity matrix of order n. In other words, an identity matrix and any square matrix of the same order are permutation matrices.

For every matrix A the following equations are correct

$$A \cdot O = O \quad \text{and} \quad O \cdot A = O$$

where O is a zero matrix of the appropriate order, i.e. a square zero matrix and any square matrix of the same order are permutation matrices.

Properties of matrix multiplication

Let λ be an arbitrary number; A, B, C – arbitrary matrices for which the operations of multiplication and addition on the left side are defined. Then the operations on the right side are defined and the following equations are correct:

1) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$;

2) $A \cdot (B + C) = A \cdot B + A \cdot C$;

3) $(A + B) \cdot C = A \cdot C + B \cdot C$;

4) $\lambda \cdot (A \cdot B) = (\lambda \cdot A) \cdot B$.

Example 1.6. Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$. Find products $A \cdot B$ and $B \cdot A$.

□ By the definition of matrix multiplication we get

$$\underbrace{A}_{2\times3}\cdot\underbrace{B}_{3\times2}=\begin{pmatrix}1&2&1\\0&1&2\end{pmatrix}\cdot\begin{pmatrix}1&0\\0&1\\1&1\end{pmatrix}=\begin{pmatrix}1\cdot1+2\cdot0+1\cdot1 & 1\cdot0+2\cdot1+1\cdot1\\0\cdot1+1\cdot0+2\cdot1 & 0\cdot0+1\cdot1+2\cdot1\end{pmatrix}=\underbrace{\begin{pmatrix}2&3\\2&3\end{pmatrix}}_{2\times2};$$

$$\underbrace{B}_{3\times2}\cdot\underbrace{A}_{2\times3}=\begin{pmatrix}1&0\\0&1\\1&1\end{pmatrix}\cdot\begin{pmatrix}1&2&1\\0&1&2\end{pmatrix}=\begin{pmatrix}1\cdot1+0\cdot0 & 1\cdot2+0\cdot1 & 1\cdot1+0\cdot2\\0\cdot1+1\cdot0 & 0\cdot2+1\cdot1 & 0\cdot1+1\cdot2\\1\cdot1+1\cdot0 & 1\cdot2+1\cdot1 & 1\cdot1+1\cdot2\end{pmatrix}=\underbrace{\begin{pmatrix}1&2&1\\0&1&2\\1&3&3\end{pmatrix}}_{3\times3}.$$

Both products $A\cdot B$ and $B\cdot A$ are defined, but they are matrices of different sizes, i.e. $A\cdot B\neq B\cdot A$. ■

Example 1.7. Let

$$A=\begin{pmatrix}1&2&1\\0&1&2\end{pmatrix},\quad x=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\quad b=\begin{pmatrix}1&2&3\end{pmatrix}.$$

Find the products $A\cdot x$, $b\cdot x$, $x\cdot b$.

□ By the definition of matrix multiplication we get

$$\underbrace{A}_{2\times3}\cdot\underbrace{x}_{3\times1}=\begin{pmatrix}1&2&1\\0&1&2\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1\cdot x_1+2\cdot x_2+1\cdot x_3\\0\cdot x_1+1\cdot x_2+2\cdot x_3\end{pmatrix}=\underbrace{\begin{pmatrix}x_1+2x_2+x_3\\x_2+2x_3\end{pmatrix}}_{2\times1};$$

$$\underbrace{b}_{1\times3}\cdot\underbrace{x}_{3\times1}=\begin{pmatrix}1&2&3\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\underbrace{\begin{pmatrix}1\cdot x_1+2\cdot x_2+3\cdot x_3\end{pmatrix}}_{1\times1}=x_1+2x_2+3x_3;$$

$$\underbrace{x}_{3\times1}\cdot\underbrace{b}_{1\times3}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\cdot\begin{pmatrix}1&2&3\end{pmatrix}=\underbrace{\begin{pmatrix}x_1&2x_1&3x_1\\x_2&2x_2&3x_2\\x_3&2x_3&3x_3\end{pmatrix}}_{3\times3}.\ ■$$

Example 1.8. Let $A=\begin{pmatrix}1&2\\3&4\end{pmatrix}$, $B=\begin{pmatrix}0&0\\1&1\end{pmatrix}$, $E=\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $O=\begin{pmatrix}0&0\\0&0\end{pmatrix}$. Find the products $A\cdot B$, $B\cdot A$, $A\cdot E$, $E\cdot A$, $B\cdot O$, $O\cdot B$.

□ All the matrices are the 2nd order square matrices. Hence, all products will be square matrices of the same order.

By the definition we get

$$A \cdot B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 0 + 2\cdot 1 & 1\cdot 0 + 2\cdot 1 \\ 3\cdot 0 + 4\cdot 1 & 3\cdot 0 + 4\cdot 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix};$$

$$B \cdot A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0\cdot 1 + 0\cdot 3 & 0\cdot 2 + 0\cdot 4 \\ 1\cdot 1 + 1\cdot 3 & 1\cdot 2 + 1\cdot 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 4 & 6 \end{pmatrix};$$

$$A \cdot E = E \cdot A = A;$$

$$A \cdot E = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 0 & 1\cdot 0 + 2\cdot 1 \\ 3\cdot 1 + 4\cdot 0 & 3\cdot 0 + 4\cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix};$$

$$E \cdot A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 0\cdot 3 & 1\cdot 2 + 0\cdot 4 \\ 0\cdot 1 + 1\cdot 3 & 0\cdot 2 + 1\cdot 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix};$$

$$A \cdot O = O \quad \text{and} \quad O \cdot A = O;$$

$$B \cdot O = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \; O \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \; \blacksquare$$

Example 1.9. Find the products $A \cdot B$ and $B \cdot A$:

a) $A = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$; b) $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}$;

c) $A = \begin{pmatrix} 6 & 1 \\ 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} -4 & -1 \\ -2 & 1 \end{pmatrix}$; d) $A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 3 \end{pmatrix}$.

□ a) The product $A \cdot B$ is a number:

$$A \cdot B = \underbrace{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}}_{1\times 3} \cdot \underbrace{\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}}_{3\times 1} = \begin{pmatrix} 1\cdot 4 + 2\cdot 5 + 3\cdot 6 \end{pmatrix} = \underbrace{(32)}_{1\times 1} = 32,$$

but the product $B \cdot A$ – is the 3$^{\text{rd}}$ order square matrix:

$$B \cdot A = \underbrace{\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}}_{3\times 1} \cdot \underbrace{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}}_{1\times 3} = \underbrace{\begin{pmatrix} 4\cdot 1 & 4\cdot 2 & 4\cdot 3 \\ 5\cdot 1 & 5\cdot 2 & 5\cdot 3 \\ 6\cdot 1 & 6\cdot 2 & 6\cdot 3 \end{pmatrix}}_{3\times 3} = \begin{pmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \end{pmatrix}.$$

It is obvious that $A \cdot B \neq B \cdot A$;

b) $$A\cdot B=\underbrace{\begin{pmatrix}1&2\\3&1\end{pmatrix}}_{2\times2}\cdot\underbrace{\begin{pmatrix}-1&3\\1&1\end{pmatrix}}_{2\times2}=\underbrace{\begin{pmatrix}1\cdot(-1)+2\cdot1&1\cdot3+2\cdot1\\3\cdot(-1)+1\cdot1&3\cdot3+1\cdot1\end{pmatrix}}_{2\times2}=\begin{pmatrix}1&5\\-2&10\end{pmatrix},$$

$$B\cdot A=\underbrace{\begin{pmatrix}-1&3\\1&1\end{pmatrix}}_{2\times2}\cdot\underbrace{\begin{pmatrix}1&2\\3&1\end{pmatrix}}_{2\times2}=\underbrace{\begin{pmatrix}(-1)\cdot1+3\cdot3&(-1)\cdot2+3\cdot1\\1\cdot1+1\cdot3&1\cdot2+1\cdot1\end{pmatrix}}_{2\times2}=\begin{pmatrix}8&1\\4&3\end{pmatrix}.$$

Both products are square matrices of the same order, but $A\cdot B\neq B\cdot A$;

c) $$A\cdot B=\underbrace{\begin{pmatrix}6&1\\2&1\end{pmatrix}}_{2\times2}\cdot\underbrace{\begin{pmatrix}-4&-1\\-2&1\end{pmatrix}}_{2\times2}=\underbrace{\begin{pmatrix}6\cdot(-4)+1\cdot(-2)&6\cdot(-1)+1\cdot1\\2\cdot(-4)+1\cdot(-2)&2\cdot(-1)+1\cdot1\end{pmatrix}}_{2\times2}=\begin{pmatrix}-26&-5\\-10&-1\end{pmatrix},$$

$$B\cdot A=\underbrace{\begin{pmatrix}-4&-1\\-2&1\end{pmatrix}}_{2\times2}\cdot\underbrace{\begin{pmatrix}6&1\\2&1\end{pmatrix}}_{2\times2}=\underbrace{\begin{pmatrix}(-4)\cdot6+(-1)\cdot2&(-4)\cdot1+(-1)\cdot1\\(-2)\cdot6+1\cdot2&(-2)\cdot1+1\cdot1\end{pmatrix}}_{2\times2}=\begin{pmatrix}-26&-5\\-10&-1\end{pmatrix}.$$

The results of multiplication are equal, i.e. $A\cdot B=B\cdot A$;

d) the product $A\cdot B$ cannot be found because the number of columns of matrix A (three) is not equal to the number of rows of matrix B (one). So, it is impossible to multiply matrix A by matrix B from the right. At the same time, it is possible to multiply matrix A by matrix B from the left:

$$B\cdot A=\underbrace{\begin{pmatrix}1&3\end{pmatrix}}_{1\times2}\cdot\underbrace{\begin{pmatrix}3&2&1\\0&1&2\end{pmatrix}}_{2\times3}=\underbrace{\begin{pmatrix}1\cdot3+3\cdot0&1\cdot2+3\cdot1&1\cdot1+3\cdot2\end{pmatrix}}_{1\times3}=\begin{pmatrix}3&5&7\end{pmatrix}.\ \blacksquare$$

Example 1.10. Find $(A\cdot B)\cdot C$, $A\cdot(B\cdot C)$, $A\cdot(B+C)$, $A\cdot B+A\cdot C$,

if $A=\begin{pmatrix}1&2\\3&4\end{pmatrix}$, $B=\begin{pmatrix}5&6\\7&8\end{pmatrix}$, $C=\begin{pmatrix}1&0\\0&2\end{pmatrix}$.

☐ Let's find

$$(A\cdot B)\cdot C=\left[\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\begin{pmatrix}5&6\\7&8\end{pmatrix}\right]\cdot\begin{pmatrix}1&0\\0&2\end{pmatrix}=\begin{pmatrix}19&22\\43&50\end{pmatrix}\cdot\begin{pmatrix}1&0\\0&2\end{pmatrix}=\begin{pmatrix}19&44\\43&100\end{pmatrix},$$

$$A\cdot(B\cdot C)=\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\left[\begin{pmatrix}5&6\\7&8\end{pmatrix}\cdot\begin{pmatrix}1&0\\0&2\end{pmatrix}\right]=\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\begin{pmatrix}5&12\\7&16\end{pmatrix}=\begin{pmatrix}19&44\\43&100\end{pmatrix},$$

$$A\cdot(B+C)=\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\left[\begin{pmatrix}5&6\\7&8\end{pmatrix}+\begin{pmatrix}1&0\\0&2\end{pmatrix}\right]=\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\begin{pmatrix}6&6\\7&10\end{pmatrix}=\begin{pmatrix}20&26\\46&58\end{pmatrix},$$

$$A\cdot B+A\cdot C=\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\begin{pmatrix}5&6\\7&8\end{pmatrix}+\begin{pmatrix}1&2\\3&4\end{pmatrix}\cdot\begin{pmatrix}1&0\\0&2\end{pmatrix}=\begin{pmatrix}19&22\\43&50\end{pmatrix}+\begin{pmatrix}1&4\\3&8\end{pmatrix}=\begin{pmatrix}20&26\\46&58\end{pmatrix}.$$

Note that $(A\cdot B)\cdot C=A\cdot(B\cdot C)$ and $A\cdot(B+C)=A\cdot B+A\cdot C$. ■

Power of matrix

Multiplication $A\cdot A$ (matrix A by itself) is defined for any square matrix A (of order n). So, it is possible to define any integer nonnegative ***power of a matrix***, as

$$A^0=E,\ A^1=A,\ A^2=A\cdot A,\ A^3=A^2\cdot A,\ldots,\ A^m=A^{m-1}\cdot A,\ \ldots\ .$$

Note that the ordinary properties of a power with natural index are correct:

$$A^k\cdot A^l=A^l\cdot A^k=A^{k+l},\quad \left(A^k\right)^l=A^{kl}.$$

Polynomial of matrix

Having defined the operations of matrix addition, multiplication by a number, and power of a matrix it is possible to get polynomial of matrix. Let $p_m(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m$ be a polynomial (power m) of variable x where A is a square matrix of order n. Expression

$$p_m(A)=a_0\underbrace{E}_{A^0}+a_1A+a_2A^2+\ldots+a_mA^m$$

is called the ***polynomial of a matrix*** A. Polynomial $p_m(A)$ is a square matrix of the n-th order.

Example 1.11. Find A^3 given that $A=\begin{pmatrix}1&2\\1&1\end{pmatrix}$.

□ By the definition of a power of a matrix we get

$$A^3=\begin{pmatrix}1&2\\1&1\end{pmatrix}^3=\begin{pmatrix}1&2\\1&1\end{pmatrix}\cdot\begin{pmatrix}1&2\\1&1\end{pmatrix}\cdot\begin{pmatrix}1&2\\1&1\end{pmatrix}=\begin{pmatrix}3&4\\2&3\end{pmatrix}\cdot\begin{pmatrix}1&2\\1&1\end{pmatrix}=\begin{pmatrix}7&10\\5&7\end{pmatrix}.$$ ■

Example 1.12. Find $p_2(A)$ given that $p_2(x)=x^2-5x+3$, $A=\begin{pmatrix} 2 & -1 \\ -3 & 3 \end{pmatrix}$.

□ Using the definition of the polynomial of a matrix:

$$p_2(A)=\begin{pmatrix} 2 & -1 \\ -3 & 3 \end{pmatrix}\cdot\begin{pmatrix} 2 & -1 \\ -3 & 3 \end{pmatrix}-5\cdot\begin{pmatrix} 2 & -1 \\ -3 & 3 \end{pmatrix}+3\cdot\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=$$

$$=\begin{pmatrix} 7 & -5 \\ -15 & 12 \end{pmatrix}-\begin{pmatrix} 10 & -5 \\ -15 & 15 \end{pmatrix}+\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \blacksquare$$

1.2.4. Matrix Transposition

For any matrix

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

a matrix

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

which can be obtained from matrix A by replacing its rows with the corresponding columns or the columns by the corresponding rows. This matrix is called a ***transposed matrix***.

To get matrix A^T from a given matrix A, the first row of matrix A is written as the first column of matrix A^T, the second row of matrix A is written as the second column of matrix A^T, and so on. This operation is called the ***transposition*** of matrix A.

A square matrix is called ***symmetric*** if

$$A^T=A,$$

and ***antisymmetric*** if

$$A^T=-A.$$

The elements of a symmetric matrix, placed symmetrically with respect to the main diagonal, are equal. The elements of an antisymmetric matrix, placed symmetrically with respect to the main diagonal, have opposite signs and all the diagonal elements are equal to zero.

Properties of transposition operation

Let λ be an arbitrary number, A, B – arbitrary matrices for which operations of matrix addition and multiplication on the left side are defined. Then the operations on the right side are defined as well and the following equations are correct:

1) $(\lambda \cdot A)^T = \lambda \cdot A^T$;

2) $(A+B)^T = A^T + B^T$;

3) $(A \cdot B)^T = B^T \cdot A^T$;

4) $(A^T)^T = A$.

Example 1.13. Find A^T, B^T, C^T, for

$$A = \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{pmatrix}}_{2\times 3}, \quad B = \begin{pmatrix} 0 & 4 & -5 \\ -4 & 0 & 6 \\ 5 & -6 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 3 \end{pmatrix}.$$

□ By the definition during the transposition the first row of matrix A becomes the first column of matrix A^T, the second row becomes the second column:

$$A^T = \underbrace{\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{pmatrix}}_{(3\times 2)}.$$

Similarly, we get

$$B^T = \begin{pmatrix} 0 & -4 & 5 \\ 4 & 0 & -6 \\ -5 & 6 & 0 \end{pmatrix}, \quad C^T = \begin{pmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 3 \end{pmatrix}.$$

As $B^T = -B$, it means that B is antisymmetric. As $C^T = C$, it means that C is symmetric. ■

Example 1.14. Find matrices $(\lambda \cdot A)^T$, $\lambda \cdot A^T$, $(A+B)^T$, $A^T + B^T$, $(A \cdot B)^T$, $B^T \cdot A^T$, $(A^T)^T$, for $\lambda = 2$, $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$.

□ We have

$$(2 \cdot A)^T = \left(2 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right)^T = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}^T = \begin{pmatrix} 2 & 6 \\ 4 & 8 \end{pmatrix},$$

$$(A+B)^T = \left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\right]^T = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}^T = \begin{pmatrix} 6 & 10 \\ 8 & 12 \end{pmatrix},$$

$$(A \cdot B)^T = \left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\right]^T = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}^T = \begin{pmatrix} 19 & 43 \\ 22 & 50 \end{pmatrix}.$$

Note that

$$2 \cdot A^T = 2 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = 2 \cdot \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 4 & 8 \end{pmatrix} = (2 \cdot A)^T,$$

$$A^T + B^T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 10 \\ 8 & 12 \end{pmatrix} = (A+B)^T,$$

$$B^T \cdot A^T = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 19 & 43 \\ 22 & 50 \end{pmatrix} = (A \cdot B)^T,$$

$$(A^T)^T = \left(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T\right)^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = A,$$

i.e. $(\lambda \cdot A)^T = \lambda \cdot A^T$, $(A+B)^T = A^T + B^T$, $(A \cdot B)^T = B^T \cdot A^T$, $(A^T)^T = A$. ■

1.2.5. Block Matrices and Block Matrix Operations

A numerical matrix A of sizes $m \times n$ divided by horizontal and vertical lines into blocks (cells), which represent matrices, is called a ***block*** (***cell***) ***matrix***.

The ***elements of a block matrix*** A are matrices A_{ij} of sizes $m_i \times n_j$, $i = 1,2,\ldots,p$, $j = 1,2,\ldots,q$, so that $m_1 + m_2 + \ldots + m_p = m$ and $n_1 + n_2 + \ldots + n_q = n$.

The operations of addition, multiplication by a number, and matrix multiplication for block matrices are performed by the same rules as for ordinary matrices, the only difference is that blocks are used instead of elements.

If numerical matrices A and B of same sizes are equally split into blocks $A=\left(A_{ij}\right)$ and $B=\left(B_{ij}\right)$, then their sum $C=A+B$ can be similarly split into blocks $C=\left(C_{ij}\right)$, so that for each block the following equation is correct: $C_{ij}=A_{ij}+B_{ij}$.

If a block matrix $A=\left(A_{ij}\right)$ is multiplied by a number, we get matrix $\lambda A=A\lambda=\left(\lambda A_{ij}\right)$.

During the transposition of block matrix, we should transpose both the block structure and the blocks, e.g.

$$A^T=\left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right)^T=\left(\begin{array}{c|c} A_{11}^T & A_{21}^T \\ \hline A_{12}^T & A_{22}^T \end{array}\right).$$

Example 1.15. We have the following block matrices

$$A=\left(\begin{array}{cc|c} 2 & 3 & 4 \\ \hline 3 & 4 & 5 \\ 4 & 5 & 6 \end{array}\right)=\left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \quad \text{and} \quad B=\left(\begin{array}{cc|c} 1 & 1 & 0 \\ \hline 2 & 1 & 2 \\ 3 & 0 & 1 \end{array}\right)=\left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right).$$

Find matrices $C=A+B$, $D=5B$, B^T.

□ Matrices A and B have blocks of equal sizes: blocks A_{11} and B_{11} have sizes $m_1\times n_1=1\times 2$; blocks A_{12} and B_{12} – $m_1\times n_2=1\times 1$; blocks A_{21} and B_{21} – $m_2\times n_1=2\times 2$; blocks A_{22} and B_{22} – $m_2\times n_2=2\times 1$.

Matrix $C=A+B$ will have blocks of the same sizes: $C=\left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right)$. For each block we find:

$$C_{11}=A_{11}+B_{11}=\begin{pmatrix}2 & 3\end{pmatrix}+\begin{pmatrix}1 & 1\end{pmatrix}=\begin{pmatrix}3 & 4\end{pmatrix}; \qquad C_{12}=A_{12}+B_{12}=(4)+(0)=(4);$$

$$C_{21}=A_{21}+B_{21}=\begin{pmatrix}3 & 4\\ 4 & 5\end{pmatrix}+\begin{pmatrix}2 & 1\\ 3 & 0\end{pmatrix}=\begin{pmatrix}5 & 5\\ 7 & 5\end{pmatrix}; \qquad C_{22}=A_{22}+B_{22}=\begin{pmatrix}5\\ 6\end{pmatrix}+\begin{pmatrix}2\\ 1\end{pmatrix}=\begin{pmatrix}7\\ 7\end{pmatrix}.$$

Hence, matrix C will have the following form:

$$C=\left(\begin{array}{cc|c} 3 & 4 & 4 \\ \hline 5 & 5 & 7 \\ 7 & 5 & 7 \end{array}\right)=\left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right).$$

Matrix $D=5B$ will have blocks of the same sizes as B:

$$D_{11}=5B_{11}=5\cdot\begin{pmatrix}1 & 1\end{pmatrix}=\begin{pmatrix}5 & 5\end{pmatrix}; \qquad D_{12}=5B_{12}=5\cdot(0)=(0);$$

$$D_{21}=5B_{21}=5\cdot\begin{pmatrix}2 & 1 \\ 3 & 0\end{pmatrix}=\begin{pmatrix}10 & 5 \\ 15 & 0\end{pmatrix}; \qquad D_{22}=5B_{22}=5\cdot\begin{pmatrix}2 \\ 1\end{pmatrix}=\begin{pmatrix}10 \\ 5\end{pmatrix}.$$

Therefore, matrix D will be

$$D=\left(\begin{array}{cc|c} 5 & 5 & 0 \\ \hline 10 & 5 & 10 \\ 15 & 0 & 5 \end{array}\right)=\left(\begin{array}{c|c} D_{11} & D_{12} \\ \hline D_{21} & D_{22} \end{array}\right).$$

By the rule of block matrix transposition we get

$$B^T=\left(\begin{array}{c|c} B_{11}^T & B_{21}^T \\ \hline B_{12}^T & B_{22}^T \end{array}\right)=\left(\begin{array}{c|cc} 1 & 2 & 3 \\ 1 & 1 & 0 \\ \hline 0 & 2 & 1 \end{array}\right). \blacksquare$$

Multiplication of block matrices

Block matrices A and B are called *consistent* if decomposition of matrix $A=(A_{ik})$ into blocks by columns is equal to the decomposition of matrix $B=(B_{kj})$ by rows, i.e. blocks A_{ik} have sizes $m_i\times p_k$, and blocks B_{kj} have sizes $p_k\times n_j$ $(k=1,2,\dots,s)$. Consistent block matrices' elements A_{ik} and B_{kj} are consistent matrices.

Product $C=A\cdot B$ of consistent *block matrices* A and B is a block matrix $C=(C_{ij})$ which elements are calculated by the following formula:

$$C_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\ldots+A_{is}B_{sj}.$$

It means that *block matrices that are divided into blocks in an appropriate way can be multiplied by the common way.* To get the block C_{ij} of the product, we need to separate out the i-th row of blocks of matrix A and the j-th column of blocks of matrix B. Then we should find the sum of pairwise products of the corresponding blocks: first block of the i-th row is multiplied by the first block of the j-th column, the second block of the i-th row is multiplied by the second block of the j-th row, and etc. Finally, the results of multiplication are summed up.

Example 1.16. We have block matrices

$$A = \left(\begin{array}{cc|c} 2 & 3 & 4 \\ \hline 3 & 4 & 5 \\ 4 & 5 & 6 \end{array}\right) = \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \quad \text{and} \quad B = \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 2 & 1 & 2 \\ \hline 3 & 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right).$$

Find the product $C = AB$.

□ Matrix A is divided into blocks: A_{11} of sizes $m_1 \times p_1 = 1\times 2$; A_{12} – $m_1 \times p_2 = 1\times 1$; A_{21} – $m_2 \times p_1 = 2\times 2$; A_{22} – $m_2 \times p_2 = 2\times 1$. Matrix B is divided into blocks: B_{11} of sizes $p_1 \times n_1 = 2\times 2$; B_{12} – $p_1 \times n_2 = 2\times 1$; B_{21} – $p_2 \times n_1 = 1\times 2$; B_{22} – $p_2 \times n_2 = 1\times 1$. Block matrices A and B are consistent. Matrix A is divided by columns into two and one (counting from the left), matrix B is divided by rows into two and one (counting from the above). Therefore, product AB is defined. Matrix $C = AB$ will have blocks $C = \left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right)$. For each block we get:

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} = \begin{pmatrix} 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} + (4)\cdot\begin{pmatrix} 3 & 0 \end{pmatrix} = \begin{pmatrix} 8 & 5 \end{pmatrix} + \begin{pmatrix} 12 & 0 \end{pmatrix} = \begin{pmatrix} 20 & 5 \end{pmatrix};$$

$$C_{12} = A_{11}B_{12} + A_{12}B_{22} = \begin{pmatrix} 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} + (4)\cdot(1) = (6) + (4) = (10);$$

$$C_{21} = A_{21}B_{11} + A_{22}B_{21} = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 5 \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 11 & 7 \\ 14 & 9 \end{pmatrix} + \begin{pmatrix} 15 & 0 \\ 18 & 0 \end{pmatrix} = \begin{pmatrix} 26 & 7 \\ 32 & 9 \end{pmatrix};$$

$$C_{22} = A_{21}B_{12} + A_{22}B_{22} = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 5 \\ 6 \end{pmatrix} \cdot (1) = \begin{pmatrix} 8 \\ 10 \end{pmatrix} + \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 13 \\ 16 \end{pmatrix}.$$

Hence, matrix C will be

$$C=\left(\begin{array}{cc|c} 20 & 5 & 10 \\ \hline 26 & 7 & 13 \\ 32 & 9 & 16 \end{array}\right)=\left(\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right). \blacksquare$$

1.2.6. Transforming a Matrix to Echelon Form

Elementary transformations of a matrix are the following transformations:

1. *Swapping two columns* (*rows*) *of matrix.*
2. *Multiplying the elements of a column* (*row*) *by the nonzero constant.*
3. *Addition of the elements of a column* (*row*) *multiplied by a constant to the elements of another column* (*row*).

Matrix B which is found from the initial matrix A by a finite amount of elementary transformations is called an ***equivalent matrix***. It is denoted by $A \sim B$.

A square matrix obtained by a finite amount of elementary transformations from an identity matrix, is called an ***elementary matrix***.

Echelon form of matrix:

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & \ast & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast \\ 0 & \cdots & 0 & 0 & 1 & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \ast & \cdots & \ast & \ast & \ast & \cdots & \ast \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \ddots & \vdots & \ast & \ast & \cdots & \ast \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \ast & \cdots & \ast \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (1.1)$$

The height of each "step" is a row, symbol "1" denotes unity elements of a matrix, symbol "$\ast$" denotes arbitrary elements, other elements are equal to zero.

Any matrix can be transformed into echelon form. It is *enough to use elementary transformations of matrix's rows*.

Remarks.

1. Matrix is also in echelon form if the elements denoted as "1" in (1.1), are arbitrary nonzero numbers.

2. It is considered that zero matrix is in echelon form.

Algorithm for transformation matrix to echelon form

To bring a matrix to echelon form (1.1) we need to make the following operations:

1. Choose a nonzero element (***pivot element***) in the first column. If the row with the pivot element (***pivot row***) is not the first, it should be placed on the first place (transformation of the I type). If the first column has no pivot element (all elements are equal to zero), this column is excluded and we continue the search of the pivot element in the remaining part of the matrix. Transformations finish when all columns are excluded or all elements in the remaining part of the matrix are equal to zero.

2. Divide the elements of the pivot row by the pivot element (II-type transformation). If the pivot row is the last, transformation procedure should be ended.

3. All the elements of the pivot row should be multiplied by a coefficient and added to all rows below (transformation of the III type). The value of the coefficient is chosen in order to nullify elements below the pivot elements.

4. After the exclusion of the row and the column that have the pivot element we return to step 1, and all operations should be applied to the remaining part of the matrix.

Example 1.17. Bring matrix to echelon form

$$A=\begin{pmatrix}3 & 9\\ 2 & 4\end{pmatrix},\quad B=\begin{pmatrix}0 & 2 & 3\\ 2 & 4 & 6\end{pmatrix},\quad C=\begin{pmatrix}2 & 4\\ 3 & 5\\ 6 & 7\end{pmatrix}.$$

□ In the first column of matrix A we choose the pivot element $a_{11}=3\neq 0$. Divide all elements of the row by $a_{11}=3$ (or multiply them by $\frac{1}{a_{11}}=\frac{1}{3}$):

$$A=\begin{pmatrix} \boxed{3} & 9 \\ 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}.$$

Add the first row multiplied by (–2) to the second row:

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{matrix} (-2) \\ \hookleftarrow \end{matrix} \sim \begin{pmatrix} 1 & 3 \\ 0 & -2 \end{pmatrix}.$$

The first row and the first column are now excluded from further examination. There is the only element (-2) in the remaining part of the matrix which is chosen as a pivot. Dividing of the last row by the pivot element we get matrix in echelon form

$$\begin{pmatrix} 1 & 3 \\ 0 & \boxed{-2} \end{pmatrix} \sim \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}.$$

Transformations are finished because the last pivot element is situated in the last row. Note that the obtained matrix is upper-triangular.

In the first column of matrix B we choose the pivot element $b_{21} = 2 \neq 0$, swap rows, and divide elements of the pivot row by the pivot element 2:

$$B=\begin{pmatrix} 0 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \sim \begin{pmatrix} \boxed{2} & 4 & 6 \\ 0 & 2 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \end{pmatrix}.$$

There is no need to make step 3 of the algorithm because of the zero element below the pivot element. We exclude the first row and the first column from the examination. There is the pivot element 2 in the remaining part. Dividing the second row by 2 we get the echelon form of the matrix

$$B \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & \boxed{2} & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1.5 \end{pmatrix}.$$

Transformations are finished, as the pivot row is the last.

In the first column of matrix C we chose the pivot element $c_{11} = 2 \neq 0$. The first row is the pivot. Divide its elements by $c_{11} = 2$ and get

$$C=\begin{pmatrix} \boxed{2} & 4 \\ 3 & 5 \\ 6 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 3 & 5 \\ 6 & 7 \end{pmatrix}.$$

To the second and to the third row we add the first one multiplied by (-3) and (-6) respectively:

$$\begin{pmatrix} \boxed{1} & 2 \\ 3 & 5 \\ 6 & 7 \end{pmatrix} \begin{matrix} (-3) & (-6) \\ & \\ & \end{matrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 0 & -5 \end{pmatrix}.$$

Pay attention that the obtained matrix is not in echelon form yet, as the second step is formed by two rows (second and third) of matrix. After the exclusion of the first row and the first column we search the pivot element in the remaining part. This element is (-1). We divide the second row by (-1), and add the pivot row, multiplied by 5, to the third row:

$$\begin{pmatrix} 1 & 2 \\ 0 & \boxed{-1} \\ 0 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & -5 \end{pmatrix} (5) \sim \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

We exclude the second row and the second column from the examination. Further transformations are impossible because all columns are excluded. The obtained matrix is in echelon form. ■

Algorithm for transformation matrix to reduced echelon form

If we continue making elementary row transformations, it is possible to simplify the matrix and transform it to ***reduced echelon form***:

$$\begin{pmatrix}
0 & \cdots & 0 & 1 & 0 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \ddots & \vdots & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{pmatrix} \qquad (1.2)$$

Symbol «1» denotes elements which are equal to one, symbol «$*$» denotes elements with arbitrary values, other elements are equal to zero. All the other elements in a column with «1» are equal to zero.

Example 1.18. Bring the matrix to reduced echelon form

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

□ Matrix is in echelon form.

Add the third row multiplied by (–1) to the first row, and the third row multiplied by (–2) to the second row:

$$A=\begin{pmatrix}0&1&1&1&1&1\\0&0&0&1&2&1\\0&0&0&0&1&0\\0&0&0&0&0&0\end{pmatrix}\begin{matrix}\\ \\(-2)\;(-1)\\ \\ \end{matrix}\sim\begin{pmatrix}0&1&1&1&0&1\\0&0&0&1&0&1\\0&0&0&0&1&0\\0&0&0&0&0&0\end{pmatrix}.$$

Now we will add the second row multiplied by (-1) to the first row. As the result, we will obtain matrix in reduced echelon form (1.2):

$$\begin{pmatrix}0&1&1&1&0&1\\0&0&0&1&0&1\\0&0&0&0&1&0\\0&0&0&0&0&0\end{pmatrix}\begin{matrix}\\(-1)\\ \\ \\ \end{matrix}\sim\begin{pmatrix}0&1&1&0&0&0\\0&0&0&1&0&1\\0&0&0&0&1&0\\0&0&0&0&0&0\end{pmatrix}.\ ■$$

Algorithm for transformation matrix to the simplest form

Any matrix by the elementary transformations (of rows and columns) can be reduced to the ***simplest form***:

$$\begin{pmatrix}1&\cdots&0&0&\cdots&0\\\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\0&\cdots&1&0&\cdots&0\\0&\cdots&0&0&\cdots&0\\\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\0&\cdots&0&0&\cdots&0\end{pmatrix}_{m\times n}=\begin{pmatrix}E_r&O\\O&O\end{pmatrix}. \qquad (1.3)$$

The upper-left corner of the matrix is an identity matrix of order r ($0 \le r \le \min\{m; n\}$), other elements are equal to zero. It is considered that zero matrix is always in the simplest form ($r = 0$).

Any matrix can be reduced to the simplest form by the elementary transformations of its rows and columns.

Example 1.19. Bring matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{pmatrix}$ to the simplest form.

☐ Let's choose the element $a_{11} = 1$ as the pivot element. Add the first row multiplied by (-2) to the second row:

$$A = \begin{pmatrix} \boxed{1} & 2 & 3 \\ 2 & 4 & 5 \end{pmatrix} \begin{matrix} (-2) \\ \hookleftarrow \end{matrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then we add the first column multiplied by (-2) to the second row and the first column multiplied by (-3) to the third row:

$$\begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

(-2)

(-3)

Multiply all elements of the last column by (-1), and switch it with the second column:

$$\begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ \underbrace{0 \quad 1}_{E_2} & & 0 \end{pmatrix}.$$

(-1)

Thus, the initial matrix A is reduced to the simplest form (1.3) by elementary transformations. ■

1.2.7. Trace

The ***trace*** of a square matrix is a sum of the elements of its main diagonal. A trace of a square matrix A of order n is denoted

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

For any square matrices A, B, C of the n-th order and vectors x, y of sizes $n \times 1$ the following properties are correct:

1) $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$;

2) $\operatorname{tr} A = \operatorname{tr} A^T$;

3) $\operatorname{tr}(A^T B) = \operatorname{tr}(B^T A) = \operatorname{tr}(AB^T) = \operatorname{tr}(BA^T)$;

4) $\operatorname{tr}(x \cdot y^T) = x^T \cdot y$;

5) $\operatorname{tr}(Axx^T) = x^T Ax$;

6) $\operatorname{tr}(ABC) = \operatorname{tr}(BCA) = \operatorname{tr}(CAB)$;

7) $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ij} = \operatorname{tr}(AB^T)$.

Matrix trace A is also denoted $\operatorname{sp} A$.

Example 1.20. Given that $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$, $C = \begin{pmatrix} 9 & 10 \\ 11 & 12 \end{pmatrix}$ and $x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $y = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, illustrate correctness of properties of the trace of a matrix.

☐ 1) $\operatorname{tr} A = 1 + 4 = 5$, $\operatorname{tr} B = 5 + 8 = 13$, $\operatorname{tr} A + \operatorname{tr} B = 18$,

$$\operatorname{tr}(A+B) = \operatorname{tr}\begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix} = 18;$$

2) $\operatorname{tr} A = 5$, $\quad \operatorname{tr} A^T = \operatorname{tr}\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = 5$;

3) $$\operatorname{tr}(A^T B) = \operatorname{tr}\left[\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix} 26 & 30 \\ 38 & 44 \end{pmatrix} = 26 + 44 = 70,$$

$$\operatorname{tr}(B^T A) = \operatorname{tr}\left[\begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix} 26 & 38 \\ 30 & 44 \end{pmatrix} = 26 + 44 = 70,$$

$$\operatorname{tr}(AB^T) = \operatorname{tr}\left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 5 & 7 \\ 6 & 8 \end{pmatrix}\right] = \operatorname{tr}\begin{pmatrix} 17 & 23 \\ 39 & 53 \end{pmatrix} = 17 + 53 = 70,$$

$$\operatorname{tr}\left(BA^T\right)=\operatorname{tr}\left[\begin{pmatrix}5 & 6\\ 7 & 8\end{pmatrix}\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}\right]=\operatorname{tr}\begin{pmatrix}17 & 39\\ 23 & 53\end{pmatrix}=17+53=70;$$

4) $$\operatorname{tr}\left(xy^T\right)=\operatorname{tr}\left[\begin{pmatrix}1\\ 2\end{pmatrix}\begin{pmatrix}3 & 4\end{pmatrix}\right]=\operatorname{tr}\begin{pmatrix}3 & 4\\ 6 & 8\end{pmatrix}=3+8=11,$$

$$x^T y=\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}3\\ 4\end{pmatrix}=1\cdot 3+2\cdot 4=11;$$

5) $$\operatorname{tr}\left(Axx^T\right)=\operatorname{tr}\left[\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}1\\ 2\end{pmatrix}\begin{pmatrix}1 & 2\end{pmatrix}\right]=\operatorname{tr}\left[\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}\right]=\operatorname{tr}\begin{pmatrix}5 & 10\\ 11 & 22\end{pmatrix}=27,$$

$$x^T Ax=\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}1\\ 2\end{pmatrix}=\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}5\\ 11\end{pmatrix}=27;$$

6) $$\operatorname{tr}\left(ABC\right)=\operatorname{tr}\left[\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}5 & 6\\ 7 & 8\end{pmatrix}\begin{pmatrix}9 & 10\\ 11 & 12\end{pmatrix}\right]=\operatorname{tr}\begin{pmatrix}413 & 454\\ 937 & 1030\end{pmatrix}=1443,$$

$$\operatorname{tr}\left(BCA\right)=\operatorname{tr}\left[\begin{pmatrix}5 & 6\\ 7 & 8\end{pmatrix}\begin{pmatrix}9 & 10\\ 11 & 12\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\right]=\operatorname{tr}\begin{pmatrix}477 & 710\\ 649 & 966\end{pmatrix}=1443,$$

$$\operatorname{tr}\left(CAB\right)=\operatorname{tr}\left[\begin{pmatrix}9 & 10\\ 11 & 12\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}5 & 6\\ 7 & 8\end{pmatrix}\right]=\operatorname{tr}\begin{pmatrix}601 & 698\\ 725 & 842\end{pmatrix}=1443;$$

7) $$\sum_{i=1}^{2}\sum_{j=1}^{2}a_{ij}b_{ij}=1\cdot 5+2\cdot 6+3\cdot 7+4\cdot 8=70=\operatorname{tr}\left(AB^T\right).\ \blacksquare$$

EXERCISES

1. For $A=\begin{pmatrix}m & n\\ n & m\end{pmatrix}$, $B=\begin{pmatrix}-n & m\\ n+m & n-m\end{pmatrix}$, find:

a) $A+2B$; b) $2A^T-B$; c) $A\cdot B-B\cdot A$; d) $(A-B)^2$; e) $\operatorname{tr}\left(A^T\cdot B\right)$;

f) $\operatorname{tr}\left(B^T\cdot A\right)$; g) $\operatorname{tr}\left(A\cdot B^T\right)$; h) $\operatorname{tr}\left(A+B\right)$; i) $\operatorname{tr}A+\operatorname{tr}B$.

2. Transform the matrix to echelon form: $\begin{pmatrix}1 & 1 & 1 & 1\\ m & m & n & n\\ n & n & m & m\end{pmatrix}$.

CHAPTER 2. DETERMINANTS AND THEIR PROPERTIES

2.1. INDUCTIVE DEFINITION

Let A be an $n \times n$ square matrix. The ***determinant*** of a square matrix A is the value, denoted by $\det A$, that is defined from A according to the following rules:

1. The determinant of the first order matrix $(n = 1)$ $A = (a_{11})$ is its only element: $\det(a_{11}) = a_{11}$.

2. The determinant of a matrix $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ of order $n > 1$ is a value

$$\det A = (-1)^{1+1} a_{11} M_{11} + (-1)^{1+2} a_{12} M_{12} + \ldots + (-1)^{1+n} a_{1n} M_{1n}, \qquad (2.1)$$

where M_{1j} is the determinant of the $(n-1) \times (n-1)$ matrix formed by deleting the first row and the j-th column of A.

The determinant is denoted by surrounding the matrix's elements with vertical bars:

$$\det A = |\, A \,| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

By this definition, we can talk about ***order of a determinant***, ***row*** or ***column of a determinant***, omitting "of a matrix". Thus, the first row of a determinant of order n is the first row $a_{11}, a_{12}, \ldots, a_{1n}$ of a square matrix of order n.

If the determinant of a square matrix is zero, the matrix is said to be ***singular***, if the determinant of a matrix is nonzero, the matrix is ***nonsingular***.

Calculation of the second–order determinant

By the definition, ***the second–order determinant*** is computed by the following formula:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \qquad (2.2)$$

The second–order determinant is the product of the elements on the main diagonal minus the product of the elements on the secondary diagonal (Fig. 2.1.).

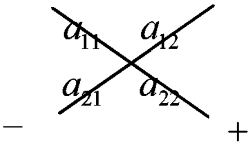


Figure 2.1

Calculation of the third–order determinant

By definition and formula (2.2), ***the third–order determinant*** can be evaluated by the following formula:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\cdot\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\cdot\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \qquad (2.3)$$

Determinant (2.3) is a sum of six components, each of them is a product of three elements from the different rows and columns of the matrix. Three of the components have the positive sign and another three have the negative sign.

To remember formula (2.3), ***the triangle's rule*** can be used: *add three products of the elements of the main diagonal and the elements in the vertexes of two triangles, having a side parallel to the main diagonal,* (Fig. 2.2, *a*), *and subtract three products of the elements of the secondary diagonal and in the vertexes of two triangles, having a side parallel to the secondary diagonal* (Fig. 2.2, *b*).

Figure 2.2

Sarrus' rule can also be used, then the determinant can be computed by the following scheme (Fig. 2.3): *write out the first two columns of the matrix to the right of the third column, so that you have five columns in a row, then add the products of* *the diagonals parallel to the main diagonal* (*going from top to bottom*) *and subtract the products of the diagonals parallel to the secondary diagonal* (*going from bottom to top*).

$$\left|\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right.$$

Figure 2.3

Example 2.1. Evaluate the determinants

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}, \quad \begin{vmatrix} 1 & 2 & 3 \\ 5 & 4 & 6 \\ 7 & -8 & -9 \end{vmatrix}, \quad \begin{vmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{vmatrix}.$$

□ Using (2.2) and (2.3), calculate

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1\cdot 4 - 2\cdot 3 = -2;$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 5 & 4 & 6 \\ 7 & -8 & -9 \end{vmatrix} = \left\langle \begin{matrix} 1 & 2 & 3 \\ 5 & 4 & 6 \\ 7 & -8 & -9 \end{matrix} \quad - \quad \begin{matrix} 1 & 2 & 3 \\ 5 & 4 & 6 \\ 7 & -8 & -9 \end{matrix} \right\rangle =$$

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

$$= -36 + 84 - 120 - 84 + 90 + 48 = -18$$

According to Sarrus' rule,

$$\begin{vmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{vmatrix} = \left\langle \begin{array}{ccc|cc} 2 & 2 & 1 & 2 & 2 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 & 2 \end{array} \right\rangle =$$

$$= 2\cdot 1\cdot 1 + 2\cdot 0\cdot 0 + 1\cdot 1\cdot 2 - 0\cdot 1\cdot 1 - 2\cdot 0\cdot 2 - 1\cdot 1\cdot 2 = 2. \blacksquare$$

2.2. COFACTOR EXPANSION FOR THE DETERMINANT

Let A be a square matrix of order n $(n>1)$. The (i,j)***-th minor*** of A, denoted M_{ij}, is the determinant of the $(n-1)\times(n-1)$ matrix formed by deleting the i-th column and the j-th row of A.

The (i,j)***-th cofactor*** A_{ij} of A is the minor M_{ij}, multiplied by $(-1)^{i+j}$:

$$A_{ij}=(-1)^{i+j}M_{ij}.$$

The determinant of A can be calculated as a sum of cofactors either along any row or column of the matrix multiplied by the elements that generated them:

$$\det A=\sum_{k=1}^{n}(-1)^{i+k}a_{ik}M_{ik}=\sum_{k=1}^{n}a_{ik}A_{ik}\quad (i\text{-th row expansion});$$

$$\det A=\sum_{k=1}^{n}(-1)^{k+j}a_{kj}M_{kj}=\sum_{k=1}^{n}a_{kj}A_{kj}\quad (j\text{-th column expansion}).$$

The determinant of a triangular matrix (upper triangular, lower triangular or a diagonal matrix) equals to the product of the elements of the main diagonal.

$$\Delta_n=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}=\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}=a_{11}\cdot a_{22}\cdot\ldots\cdot a_{nn}.$$

The determinant of any identity matrix equals to 1.

Example 2.2. Evaluate the determinant of

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

☐ Let's expand across the third row:

$$\det A=0\cdot A_{31}+0\cdot A_{32}+0\cdot A_{33}+5\cdot A_{34}=5\cdot(-1)^{3+4}\begin{vmatrix} 2 & 1 & 0 \\ 0 & 1 & 3 \\ -1 & 2 & 0 \end{vmatrix}=-5\cdot\begin{vmatrix} 2 & 1 & 0 \\ 0 & 1 & 3 \\ -1 & 2 & 0 \end{vmatrix}.$$

Next expand along the last column of the remaining third order determinant:

$$\det A = -5 \cdot \begin{vmatrix} 2 & 1 & 0 \\ 0 & 1 & 3 \\ -1 & 2 & 0 \end{vmatrix} = -5 \cdot \left(\; 0 \cdot A_{13} + 3 \cdot A_{23} + 0 \cdot A_{33} \; \right) = -5 \cdot 3 \cdot (-1)^{2+3} \cdot \begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix} .$$

The second order determinant is calculated according to the equation (2.2):

$$\det A = 15 \cdot \begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix} = 15 \cdot \left(2 \cdot 2 - (-1) \cdot 1\right) = 15 \cdot 5 = 75 . \; \blacksquare$$

2.3. PROPERTIES OF DETERMINANTS

2.3.1. Main Properties of Determinants

1. The determinant of the transpose of any square matrix is the same as the determinant of the original matrix: $\det A = \det\left(A^T\right)$. As a result, rows and columns of the determinant are "equal": any property that is true for the rows of a matrix would be true for the columns as well.

2. If a row of a matrix is zero (all the elements of the row are zero), then the determinant is zero: $\det(\dots o \dots) = 0$.

3. Interchanging any two columns of the matrix changes the sign of the determinant to the opposite one (asymmetric property):

$$\det(\dots a_j \dots a_k \dots) = -\det(\dots a_k \dots a_j \dots) .$$

4. If two rows of a matrix are equal, then the determinant is zero:

$$\det(\dots a_j \dots a_k \dots) = 0 \text{ if } a_j = a_k .$$

5. If two rows of a matrix are proportional, then the determinant is zero:

$$\det(\dots a_j \dots a_k \dots) = 0 \text{ if } a_j = \lambda a_k .$$

6. Multiplying a column by the constant multiplies the determinant by that constant:

$$\det(a_1 \dots \lambda \cdot a_j \dots a_n) = \lambda \cdot \det(a_1 \dots a_j \dots a_n) .$$

7. If the j-th column is written as the sum of the two columns $a_j + b_j$, then the determinant is the sum of two corresponding determinants, where j-th columns are a_j and b_j, respectively, and the other columns are the same:

$$\det(\dots a_j + b_j \dots) = \det(\dots a_j \dots) + \det(\dots b_j \dots).$$

8. The determinant is a linear function of each column:

$$\det(\dots \alpha \cdot a_j + \beta \cdot b_j \dots) = \alpha \cdot \det(\dots a_j \dots) + \beta \cdot \det(\dots b_j \dots).$$

9. If a scalar multiple of a column is added to another column, the value of the determinant is unchanged:

$$\det(\dots a_j + \lambda \cdot a_k \dots a_k \dots) = \det(\dots a_j \dots a_k \dots).$$

10. The sum of the products formed by multiplying each element of any column by the cofactors of corresponding elements of another column is zero:

$$\sum_{k=1}^{n} a_{ki} \cdot A_{kj} = 0 \text{ for } i \neq j.$$

From the formulas for row (column) expansion and Property 10, we have

$$\sum_{k=1}^{n} a_{ki} \cdot A_{kj} = \begin{cases} 0, & i \neq j, \\ \det A, & i = j, \end{cases} \qquad \sum_{k=1}^{n} a_{ik} \cdot A_{jk} = \begin{cases} 0, & i \neq j, \\ \det A, & i = j. \end{cases} \tag{2.4}$$

Let A be a square matrix. ***The adjoint matrix*** of A, denoted by A^+, is the square matrix of the same order where each element is the (j,i)-th cofactor of the matrix A: $a_{ij}^+ = A_{ji}$.

The adjoint matrix can be computed by the following procedure:

1) replace each element of the original matrix $A = (a_{ij})$ with corresponding cofactor $A_{ij} = (-1)^{i+j} M_{ij}$, thus obtaining the matrix (A_{ij});
2) find the adjoint matrix A^+, transposing the (A_{ij}) matrix.

From equation (2.4) it follows that $AA^+ = A^+ \cdot A = \det A \cdot E$, where E is the identity matrix of the same order as A.

Example 2.3. Given that $A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$, compare the determinant of A to the determinants of the following matrices: A^T; $B=\begin{pmatrix}2 & 1\\ 4 & 3\end{pmatrix}$; $C=\begin{pmatrix}3 & 4\\ 1 & 2\end{pmatrix}$; $D=\begin{pmatrix}1 & 2\\ 3\lambda & 4\lambda\end{pmatrix}$; $F=\begin{pmatrix}1+3\lambda & 2+4\lambda\\ 3 & 4\end{pmatrix}$, where λ is a certain scalar.

□ The determinant of matrix A was found in example 2.1: $\det A=-2$. Let's evaluate the other determinants, using formula (2.2):

$$\det\left(A^T\right)=\begin{vmatrix}1 & 3\\ 2 & 4\end{vmatrix}=1\cdot 4-3\cdot 2=-2=\det A,$$

according to Property 1;

$$\det B=\begin{vmatrix}2 & 1\\ 4 & 3\end{vmatrix}=2\cdot 3-1\cdot 4=2=-\det A,$$

according to Property 3, since matrix B is obtained from matrix A by switching the first and the second columns;

$$\det C=\begin{vmatrix}3 & 4\\ 1 & 2\end{vmatrix}=3\cdot 2-4\cdot 1=2=-\det A,$$

according to Property 3, since matrix C is obtained from matrix A by switching the first and the second rows;

$$\det D=\begin{vmatrix}1 & 2\\ 3\lambda & 4\lambda\end{vmatrix}=1\cdot 4\lambda-2\cdot 3\lambda=-2\lambda=\lambda\det A,$$

according to Property 6, since matrix D is obtained from matrix A by multiplying the second row by the constant λ;

$$\det F=\begin{vmatrix}1+3\lambda & 2+4\lambda\\ 3 & 4\end{vmatrix}=\left(1+3\lambda\right)\cdot 4-\left(2+4\lambda\right)\cdot 3=-2=\det A,$$

according to Property 9, since matrix F is obtained from matrix A by multiplying the second row by λ and adding the product to the first row. ■

Example 2.4. Given that $A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$, $B=\begin{pmatrix}1 & 2 & 3\\ 5 & 4 & 6\\ 7 & -8 & -9\end{pmatrix}$, find the corresponding adjoint matrices A^{+}, B^{+}.

☐ Let's calculate all the cofactors of matrix A:

$$A_{11}=(-1)^{1+1}\cdot 4=4,\qquad A_{12}=(-1)^{1+2}\cdot 3=-3,$$

$$A_{21}=(-1)^{2+1}\cdot 2=-2,\qquad A_{22}=(-1)^{2+2}\cdot 1=1.$$

Now we can find the adjoint matrix by transposing matrix (A_{ij}):

$$A^{+}=(A_{ij})^{T}=\begin{pmatrix}4 & -3\\ -2 & 1\end{pmatrix}^{T}=\begin{pmatrix}4 & -2\\ -3 & 1\end{pmatrix}.$$

Let's calculate the cofactors of matrix B:

$$B_{11}=(-1)^{1+1}\begin{vmatrix}4 & 6\\ -8 & -9\end{vmatrix}=12,\ B_{12}=(-1)^{1+2}\begin{vmatrix}5 & 6\\ 7 & -9\end{vmatrix}=87,\ B_{13}=(-1)^{1+3}\begin{vmatrix}5 & 4\\ 7 & -8\end{vmatrix}=-68,$$

$$B_{21}=(-1)^{2+1}\begin{vmatrix}2 & 3\\ -8 & -9\end{vmatrix}=-6,\ B_{22}=(-1)^{2+2}\begin{vmatrix}1 & 3\\ 7 & -9\end{vmatrix}=-30,\ B_{23}=(-1)^{2+3}\begin{vmatrix}1 & 2\\ 7 & -8\end{vmatrix}=22,$$

$$B_{31}=(-1)^{3+1}\begin{vmatrix}2 & 3\\ 4 & 6\end{vmatrix}=0,\ B_{32}=(-1)^{3+2}\begin{vmatrix}1 & 3\\ 5 & 6\end{vmatrix}=9,\ B_{33}=(-1)^{3+3}\begin{vmatrix}1 & 2\\ 5 & 4\end{vmatrix}=-6.$$

Then the adjoint matrix is found by transposing matrix (B_{ij}):

$$B^{+}=(B_{ij})^{T}=\begin{pmatrix}12 & 87 & -68\\ -6 & -30 & 22\\ 0 & 9 & -6\end{pmatrix}^{T}=\begin{pmatrix}12 & -6 & 0\\ 87 & -30 & 9\\ -68 & 22 & -6\end{pmatrix}.\ ■$$

2.3.2. Determinant of Matrix Product

Let A and B be square matrices of the same order. Then

$$\det(A\cdot B)=\det A\cdot\det B,$$

i.e. the determinant of a matrix product of square matrices equals to the product of their determinants.

Example 2.5. Calculate the determinant of the product of matrices:

$$A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix},\qquad B=\begin{pmatrix}1 & 3\\ 4 & 5\end{pmatrix}.$$

□ Let's evaluate the second order determinants of the matrices (see example 2.1): $\det A=-2$, $\det B=-7$. Using the property of the determinant of a matrix product, we get $\det(A\cdot B)=\det A\cdot\det B=(-2)\cdot(-7)=14$.

Now, calculate the determinant by computing the matrix product:

$$A\cdot B=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\cdot\begin{pmatrix}1 & 3\\ 4 & 5\end{pmatrix}=\begin{pmatrix}9 & 13\\ 19 & 29\end{pmatrix}.$$

Hence, $\det(A\cdot B)=\begin{vmatrix}9 & 13\\ 19 & 29\end{vmatrix}=9\cdot 29-13\cdot 19=14$. The result is equal to the one obtained before. ■

2.3.3. Elementary Transformations

Definition-based evaluation of determinants is not generally applied to the large matrices ($n>3$), since the number of required operations, as well that the difficulty of the calculation, grows very quickly.

It is a much more efficient approach to use the properties of the determinant. The most important ones for evaluating determinants are Properties 3, 6, 9. These properties are called ***elementary transformations*** (*elementary row operations*).

- *Switching two rows (columns) of the determinant* reverses its sign.
- *Multiplying each element in a row (column) by a non-zero constant* multiplies the determinant by this constant.
- *Adding to each element of a row (column) a scalar multiple of a corresponding element of another row (column) of the determinant doesn't change the value of the determinant.*

Elementary transformations can be used to simplify the determinant, or to modify it so that it can be computed more easily.

Method of matrix reduction to triangular form

The method consists of two steps:

1) using elementary transformations reduce the determinant of a matrix to the triangular form;

2) calculate the determinant of a triangular matrix as a product of the diagonal elements.

Example 2.6. Calculate the determinant

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

by transforming it to a triangular form.

☐ 1. Let's use the elementary transformations to reduce the matrix to the triangular form. Choosing element $a_{11} = 1$ from the first row as a leading coefficient (a pivot), make all the other elements of the first column equal to zero. Add the first row times (-3) to the third row and add the first row times (-4) to the fourth row:

$$\det A = \begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & -8 & -10 \\ 0 & 1 & -10 & -13 \end{vmatrix}.$$

The value of the determinant doesn't change since we use the III-type elementary transformations.

Switch the second and the fourth rows of the determinant:

$$\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & -8 & -10 \\ 0 & 1 & -10 & -13 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & -8 & -10 \\ 0 & 3 & 0 & 1 \end{vmatrix}.$$

We reverse the sign of the determinant because we used the I-type elementary transformation.

Now choose entry $a_{22}=1$ as a leading coefficient and make element $a_{42}=3$ equal to zero by adding the second row times (-3) to the fourth one:

$$-\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & -8 & -10 \\ 0 & 3 & 0 & 1 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & -8 & -10 \\ 0 & 0 & 30 & 40 \end{vmatrix}.$$

Let's divide the third row by (-8), and the fourth row by 10, at the same time multiplying the determinant by $-80=(-8)\cdot 10$ in order to keep the equation balanced (II-type transformation):

$$-\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & -8 & -10 \\ 0 & 0 & 30 & 40 \end{vmatrix} = -(-80)\cdot\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & 1 & 1,25 \\ 0 & 0 & 3 & 4 \end{vmatrix} = 80\cdot\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & 1 & 1,25 \\ 0 & 0 & 3 & 4 \end{vmatrix}.$$

Let's choose $a_{33}=1$ as a leading coefficient and make $a_{43}=3$ equal to zero. Add the third row times (-3) to the fourth row:

$$80\cdot\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & 1 & 1,25 \\ 0 & 0 & 3 & 4 \end{vmatrix} = 80\cdot\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & 1 & 1,25 \\ 0 & 0 & 0 & 0,25 \end{vmatrix}.$$

Now we have an upper triangular matrix.

2. Evaluate the determinant of the upper triangular matrix by multiplying the elements of the main diagonal:

$$\det A = 80\cdot\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -10 & -13 \\ 0 & 0 & 1 & 1,25 \\ 0 & 0 & 0 & 0,25 \end{vmatrix} = 80\cdot 1\cdot 1\cdot 1\cdot 0,25 = 20. \blacksquare$$

Method of determinant order reduction

The method consists of two steps:

1) use the III-type elementary transformations to make all the elements of a row (column), except for one, equal to zero;

2) expand the determinant along this row (column), obtaining a determinant of decreased order. If the order of the new determinant $n > 1$, go to step 1, else finish the calculations.

Example 2.7. Evaluate the determinant

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

by reducing its order.

□ 1. Let's choose $a_{24} = 1$ as a leading coefficient, and make all the other elements of the second row equal to zero, using elementary transformations. Multiply the fourth column by (-3) and add it to the second one:

$$\begin{vmatrix} 1 & 0 & 3 & 4 \\ 0 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & -12 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 3 & -6 & 1 & 2 \\ 4 & -8 & 2 & 3 \end{vmatrix}.$$

2. Expand the determinant along the second row:

$$\begin{vmatrix} 1 & -12 & 3 & 4 \\ 0 & 0 & 0 & 1 \\ 3 & -6 & 1 & 2 \\ 4 & -8 & 2 & 3 \end{vmatrix} = 1 \cdot (-1)^{2+4} \cdot \begin{vmatrix} 1 & -12 & 3 \\ 3 & -6 & 1 \\ 4 & -8 & 2 \end{vmatrix}.$$

We now have a third order determinant.

Now let's multiply the second column by $0{,}5$, then we also have to multiply the determinant by 2):

$$\begin{vmatrix} 1 & -12 & 3 \\ 3 & -6 & 1 \\ 4 & -8 & 2 \end{vmatrix} = 2\cdot\begin{vmatrix} 1 & -6 & 3 \\ 3 & -3 & 1 \\ 4 & -4 & 2 \end{vmatrix}.$$

Add the first column to the second one:

$$2\cdot\begin{vmatrix} 1 & -6 & 3 \\ 3 & -3 & 1 \\ 4 & -4 & 2 \end{vmatrix} = 2\cdot\begin{vmatrix} 1 & -5 & 3 \\ 3 & 0 & 1 \\ 4 & 0 & 2 \end{vmatrix}.$$

Expand this determinant along the second column:

$$2\cdot\begin{vmatrix} 1 & -5 & 3 \\ 3 & 0 & 1 \\ 4 & 0 & 2 \end{vmatrix} = 2\cdot(-5)\cdot(-1)^{1+2}\cdot\begin{vmatrix} 3 & 1 \\ 4 & 2 \end{vmatrix} = 10\cdot\begin{vmatrix} 3 & 1 \\ 4 & 2 \end{vmatrix}.$$

We get the second order determinant.

Let's add the first row times (-2) to the second row:

$$10\cdot\begin{vmatrix} 3 & 1 \\ 4 & 2 \end{vmatrix} = 10\cdot\begin{vmatrix} 3 & 1 \\ -2 & 0 \end{vmatrix}.$$

Expand the determinant along the second row, getting the first order determinant, which value equals to its only element:

$$10\cdot\begin{vmatrix} 3 & 1 \\ -2 & 0 \end{vmatrix} = 10\cdot(-2)\cdot(-1)^{2+1}\cdot 1 = 20.$$

The result is equal to the one obtained in example 2.6. ■

EXERCISES

Evaluate the determinants:

a) $\begin{vmatrix} m+n & m-n \\ m-n & m+n \end{vmatrix}$; b) $\begin{vmatrix} m & m & n \\ m & m & m+n \\ n & m+n & 2n \end{vmatrix}$.

CHAPTER 3. MATRIX RANK

3.1. LINEAR DEPENDENCE AND LINEAR INDEPENDENCE OF MATRIX ROWS (COLUMNS)

In the following, we will call matrix-columns (matrix rows) simply ***columns*** (***rows***) and denote them by lowercase letters. Columns are ***equal*** if they have the same sizes and all the corresponding elements are equal.

Column A is called a ***linear combination*** of columns $A_1, A_2, \ldots, A_k$ of the same sizes, if

$$A = \alpha_1 \cdot A_1 + \alpha_2 \cdot A_2 + \ldots + \alpha_k \cdot A_k, \tag{3.1}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are arbitrary numbers. In that case we say that ***column*** A ***is decomposed into columns*** $A_1, A_2, \ldots, A_k$, and numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$ are called the ***decomposition coefficients***.

A linear combination $A = 0 \cdot A_1 + 0 \cdot A_2 + \ldots + 0 \cdot A_k$, where all the coefficients are equal to zero, is called ***trivial***.

If the columns in (3.1) are given by

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \; A_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \; \ldots \;, \; A_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix},$$

then the matrix equality (3.1) can be expressed in a form of element-to-element equalities

$$a_i = \alpha_1 \cdot a_{i1} + \alpha_2 \cdot a_{i2} + \ldots + \alpha_k \cdot a_{ik}, \qquad i = 1, \ldots, n.$$

A linear combination of rows of the same sizes is defined in a similar way.

A set of columns $A_1, A_2, \ldots, A_k$ of the same sizes is called a ***system of columns***. Any part of a system of columns system is called a ***subsystem***.

A system of k columns $A_1, A_2, \ldots, A_k$ is called ***linearly dependent***, if there exist such numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$, not all equal to zero, that

$$\alpha_1 \cdot A_1 + \alpha_2 \cdot A_2 + \ldots + \alpha_k \cdot A_k = o. \tag{3.2}$$

Hereinafter symbol o will denote a zero column of a corresponding sizes.

A system of k columns $A_1, A_2, \ldots, A_k$ is called ***linearly independent***, if equation (3.2) is correct only if $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, i.e. when the linear combination on the left side of equation (3.2) is trivial.

One column A_1 composes a system as well: for $A_1 = o$ the system is linearly dependent, and for $A_1 \neq o$ – linearly independent.

For rows (row matrices) we get similar definitions.

Example 3.1. By definition, determine linear dependence or linear independence of systems of columns:

$$\text{a) } A_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ A_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}; \qquad \text{b) } A_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ A_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

□ a) Columns $A_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ are linearly dependent, because we can compose a non-trivial linear combination, e.g., with coefficients $\alpha_1 = 2$, $\alpha_2 = -1$, which is equal to a zero column:

$$2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

b) columns $A_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ are linearly independent, because the equality $\alpha_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_2 \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, that matches the system $\begin{cases} 1 \cdot \alpha_1 = 0 \\ 2 \cdot \alpha_2 = 0 \end{cases}$, is correct only when $\alpha_1 = \alpha_2 = 0$.■

Properties of linearly dependent and linearly independent columns

The concepts of linear dependence and linear independence are defined for rows and columns in a similar manner. Hence, the properties of linear dependence and independence, given for columns, will be true for rows as well.

1. If there is a zero column in a system of columns, this system is linearly dependent.

2. If there are two equal columns in a system of columns, this system is linearly dependent.

3. If there are two proportional columns ($A_i = \lambda A_j$) in a system of columns, this system is linearly dependent.

4. A system of $k > 1$ columns is linearly dependent when and only when at least one of the columns is a linear combination of the others.

5. Any columns that are included in a linearly independent system, compose a linearly independent subsystem.

6. A system of columns that contains a linearly dependent subsystem, is linearly dependent.

7. If a system of columns $A_1, A_2, \ldots, A_k$ is linearly independent, but after the addition of column A becomes linearly dependent, then column A can be uniquely decomposed into columns $A_1, A_2, \ldots, A_k$, i.e. decomposition coefficients are single-valued.

8. Two nonzero columns A_1, A_2 compose a linearly dependent system, if they are proportional ($A_1 = \lambda A_2$), and a linearly independent system, if they are not proportional.

3.2. BASIS MINOR AND MATRIX RANK

Basis minor of a matrix. Computing the rank of a matrix

Let A be a $m \times n$ matrix, and k – a natural number not greater than m and n: $k \le \min\{\, m; n \,\}$. A ***minor of order k*** of matrix A is the determinant of a matrix of the k-th order, composed of the elements at the intersection of k arbitrarily chosen rows and k arbitrarily chosen columns of A.

Denoting minors, we will write the numbers of the chosen rows as superscripts and the numbers of the chosen columns as subscripts, in ascending order.

Example 3.3. Write down minors of different orders of the following matrices:

$$\text{a)}\quad A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}; \qquad \text{b)}\ B = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & 2 & 3 \\ 1 & 4 & 3 & 3 \end{pmatrix}.$$

□ a) Matrix A of sizes 2×3 has six minors of the first order, for example, $M_2^1 = \det(a_{12}) = 2$, and three minors of the second order, for example, $M_{23}^{12} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3$.

b) Matrix B of sizes 3×4 has 12 minors of the first order, e.g. $M_2^3 = \det(b_{32}) = 4$, and 18 minors of the second order, e.g. $M_{23}^{12} = \begin{vmatrix} 2 & 1 \\ 2 & 2 \end{vmatrix} = 2$, and four minors of the third order, e.g. $M_{134}^{123} = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 2 & 3 \\ 1 & 3 & 3 \end{vmatrix} = 0$. ■

Let A be a $m \times n$ matrix. A ***minor*** of A of order r is called ***basis***, if it is nonzero and all minors of order $(r+1)$ are equal to zero or do not exist.

The rank of a matrix is the order of its basis minor. The rank of a matrix A is denoted by $\operatorname{rg} A$. It can also be denoted by $\operatorname{Rg} A$, $\operatorname{rang} A$, $\operatorname{rank} A$.

A zero matrix doesn't have a basis minor. Thus, the rank of a zero matrix is, by definition, equal to zero.

If all minors of order k of a matrix are equal to zero, all minors of higher order are also equal to zero.

The rank of a matrix equals to the largest order of any nonzero minor of this matrix.

If a square matrix is nonsingular, its rank is equal to its order. If a square matrix is singular, its rank is less than its order.

The rank of a block matrix is computed as the rank of an ordinary (numerical) matrix, i.e. without paying attention to the block structure. In addition, the rank of a block matrix is not less than the ranks of its blocks:

$$\mathrm{rg}\left(A \mid B\right) \geq \mathrm{rg}\, A \quad \text{and} \quad \mathrm{rg}\left(A \mid B\right) \geq \mathrm{rg}\, B .$$

Example 3.4. Find all basis minors and ranks of the following matrices:

$$\text{a) } O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \qquad \text{b) } A = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}; \qquad \text{c) } B = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix};$$

$$\text{d) } C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix}; \qquad \text{e) } D = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 1 & 3 \end{pmatrix}; \qquad \text{f) } F = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 2 & 3 \end{pmatrix};$$

$$\text{g) } G = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}; \qquad \text{h) } H = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 0 & 2 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

□ a) Matrix O is zero, so all of its minors are equal to zero. A zero matrix doesn't have any basis minors, and it's rank equals to zero by the definition: $\mathrm{rg} O = 0$.

b) One of the first-order minors of matrix $A = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ is nonzero: $M_3^1 = 1$, and minors of the second order don't exist (since there's only one row). Hence, the minor M_3^1 is basis and the rank of this matrix is equal to 1.

c) For matrix $B = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ there are nonzero minors of the first order: $M_1^1 = 1$ and $M_2^1 = 2$. These minors are basis, because the only minor of the second order $M_{12}^{12}=\begin{vmatrix}1 & 2\\ 0 & 0\end{vmatrix}$ is equal to zero. Hence, $\mathrm{rg}\, B=1$.

d) All first-order minors minors of the first order of matrix $C=\begin{pmatrix}1 & 2 & 3\\ 2 & 4 & 6\end{pmatrix}$, which are equal to its elements, are nonzero, and all the minors of the second order are equal to zero, because rows of the matrix are proportional. Thus, the matrix has six basis minors and its rank is equal to 1.

e) Matrix D has a nonzero minor of the second order $M_{12}^{13}=\begin{vmatrix}1 & 2\\ 1 & 3\end{vmatrix}=1$, and minors of the third order don't exist (since there're only two columns). Hence, M_{12}^{13} – is the only basis minor and $\mathrm{rg}\, D=2$.

f) Matrix F has six nonzero minors of the second order: M_{12}^{12}, M_{13}^{12}, M_{23}^{12}, M_{12}^{23}, M_{13}^{23}, M_{23}^{23}, and the only third-order minor, i.e. the determinant, is equal to zero, since matrix has two equal rows (the first and the third ones). Hence, each of the mentioned minors of the second order is basis and the rank of the matrix is equal to 2.

g) The determinant of matrix G (i.e. the minor of the third order) is nonzero: $\det G=M_{123}^{123}=1\cdot 4\cdot 6\neq 0$. Hence, the minor M_{123}^{123} is basis and $\mathrm{rg}\, G=3$.

h) All third-order minors of this matrix are equal to zero, because the third row of these minors is zero. So, only a minor of the second order, situated in the first two rows of the matrix, can be basis. Searching through the six possible minors, we choose the nonzero ones:

$$M_{12}^{12}=M_{13}^{12}=\begin{vmatrix}1 & 2\\ 0 & 2\end{vmatrix},\quad M_{24}^{12}=M_{34}^{12}=\begin{vmatrix}2 & 0\\ 2 & 3\end{vmatrix}\ \text{and}\ M_{14}^{12}=\begin{vmatrix}1 & 0\\ 0 & 3\end{vmatrix}.$$

Each of these five minors is basis. Hence, the rank of this matrix is equal to 2. ■

Properties of basis minor and matrix rank

1. In an nonzero matrix A every row (column) is a linear combination of rows (columns), in which the basis minor is situated.

2. The determinant is equal to zero if and only if one of its rows (columns) is a linear combination of other rows (columns).

3. Applying elementary transformations to a matrix does not change its rank.

4. If a row (column) of a matrix is a linear combination of other rows (columns) of this matrix, this row (column) can be deleted from the matrix without changing its rank.

5. If a matrix is reduced to the simplest form (1.3), then $\operatorname{rg} A = \operatorname{rg} \Lambda = r$.

6. The rank of a matrix is equal to the maximum number of linearly independent rows of this matrix.

7. The maximum number of linearly independent rows of a matrix is equal to the maximum number of linearly independent columns:

$$\operatorname{rg} A = \operatorname{rg} A^T.$$

8. Elementary row transformations preserve linear dependence (or linear independence) of any system of columns of this matrix.

9. The rank of a matrix product is not larger than the ranks of factors:

$$\operatorname{rg}(AB) \le \min\{\operatorname{rg} A, \operatorname{rg} B\}.$$

10. If A is a nonsingular square matrix, then $\operatorname{rg}(AB) = \operatorname{rg} B$ and $\operatorname{rg}(CA) = \operatorname{rg} C$, i.e. the rank of a matrix does not change after multiplying it from the left or from the right by nonsingular square matrix.

11. The rank of a sum of matrices is not larger than the sum of the ranks of summands:

$$\operatorname{rg}(A+B) \le \operatorname{rg} A + \operatorname{rg} B.$$

3.3. METHODS FOR MATRIX RANK COMPUTATION

3.3.1. Method of Bounding Minors

Let A be an $m \times n$ matrix. We will say that minor $M^{i_1 i_2 \dots i_k i_{k+1}}_{j_1 j_2 \dots j_k j_{k+1}}$ of order $(k+1)$ ***bounds*** (contains) minor $M^{i_1 i_2 \dots i_k}_{j_1 j_2 \dots j_k}$ of order k.

Describing the method, we will write down indices of the chosen rows and columns, in which the minor is situated, without putting them in ascending order. In so doing, the minor at issue and the minor with indices put in order have equal absolute value and, maybe, are of different signs, but it is of no importance for the method of bounding minors, because we only want to find out the answer to the question: is the minor equal to zero or not.

1. Choose row i_1 and column j_1, so that the minor of the first order $M^{i_1}_{j_1} = a_{i_1 j_1}$ is nonzero. If it is possible, then $\operatorname{rg} A \geq 1$, else the process terminates and $\operatorname{rg} A = 0$.

2. Bound the minor $M^{i_1}_{j_1} \neq 0$ by adding another row $i_2 \neq i_1$ and another column $j_2 \neq j_1$ to the chosen i_1-th row and the j_1-th column, so that the minor $M^{i_1 i_2}_{j_1 j_2} = \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} \\ a_{i_2 j_1} & a_{i_2 j_2} \end{vmatrix} \neq 0$. If it is possible, then $\operatorname{rg} A \geq 2$, else the process should terminates and $\operatorname{rg} A = 1$.

3. Bound the minor $M^{i_1 i_2}_{j_1 j_2} \neq 0$ by adding to the previously chosen rows and columns another row i_3 and another column j_3 in order to get the minor $M^{i_1 i_2 i_3}_{j_1 j_2 j_3} \neq 0$. If it is possible, $\operatorname{rg} A \geq 3$, else the process terminates and $\operatorname{rg} A = 2$.

Continue the bounding process until it is terminated. Suppose we have found a nonzero minor of order r: $M^{i_1 i_2 \dots i_r}_{j_1 j_2 \dots j_r} \neq 0$, i.e. $\operatorname{rg} A \geq r$. But all the minors of order $(r+1)$, bounding it, are equal to zero $M^{i_1 i_2 \dots i_r i_{r+1}}_{j_1 j_2 \dots j_r j_{r+1}} = 0$ or do not exist (for $r = m$ or $r = n$). Then the process terminates and $\operatorname{rg} A = r$.

Example 3.5. Find ranks of matrices, using the method of bounding minors:

$$O=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix},\quad A=\begin{pmatrix}3 & 9\\ 2 & 4\end{pmatrix},\quad B=\begin{pmatrix}0 & 2 & 3\\ 2 & 4 & 6\end{pmatrix},\quad C=\begin{pmatrix}1 & 0 & 2 & 1 & 3\\ 2 & 0 & 1 & 1 & 2\\ 3 & 0 & 3 & 2 & 5\end{pmatrix}.$$

□ *Matrix* O:

1. This matrix does not have any nonzero minors of the first order, because all of its elements are equal to zero. Hence, $\operatorname{rg} O = 0$.

Matrix A:

1. Choose the first row ($i_1 = 1$) and the first column ($j_1 = 1$) of matrix A, at the intersection of which there is a nonzero element $a_{11} = 3 \neq 0$. We have minor $M_1^1 = 3 \neq 0$. Hence, $\operatorname{rg} A \geq 1$.

2. Add another row $i_2 = 2$ and another column $j_2 = 2$ to the previously chosen ones. We have a nonzero minor of the second order: $M_{12}^{12} = \det A = \begin{vmatrix}3 & 9\\ 2 & 4\end{vmatrix} = -6 \neq 0$.

Hence, $\operatorname{rg} A \geq 2$.

3. Since we have used all the rows and columns of matrix A, there are no minors bounding $M_{12}^{12} \neq 0$. Hence, $\operatorname{rg} A = 2$.

Matrix B:

1. Choose the first row and the second column of matrix at the intersection of which there is a nonzero element $b_{12} = 2 \neq 0$. We have a minor $M_2^1 = 2 \neq 0$. Hence, $\operatorname{rg} B \geq 1$.

2. Add the second row and the third column to the previously chosen ones. We have a minor of the second order: $M_{23}^{12} = \begin{vmatrix}2 & 3\\ 4 & 6\end{vmatrix} = 0$. The choice was unsuccessful, because we have got a zero minor. Let's take the first column instead of the third one. Then we have a nonzero minor of the second order: $M_{21}^{12} = \begin{vmatrix}2 & 0\\ 4 & 2\end{vmatrix} = 4 \neq 0$. Hence, $\operatorname{rg} B \geq 2$.

3. We have used all the rows of matrix B. There are no minors of the third order, thus $\operatorname{rg} B = 2$.

Matrix C:

1. Choose the first row ($i_1 = 1$) and the first column ($j_1 = 1$) of matrix C, at the intersection of which there is a nonzero element $a_{11} = 1 \neq 0$. We have minor $M_1^1 = 1 \neq 0$. Hence, $\operatorname{rg} C \geq 1$.

2. Add another row $i_2 = 2$ and another column $j_2 = 2$ to the previously chosen ones. We have a minor of the second order: $M_{12}^{12} = \begin{vmatrix} 1 & 0 \\ 2 & 0 \end{vmatrix}$. Choosing the second column was unsuccessful, because we have got a zero minor. Let's choose the third column ($j_2 = 3$) instead. Then we have a nonzero minor $M_{13}^{12} = \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3 \neq 0$. Hence, $\operatorname{rg} C \geq 2$.

3. Bound minor $M_{13}^{12} \neq 0$. There are three bounding minors:

$$M_{134}^{123} = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 3 & 3 & 2 \end{vmatrix} = 0, \quad M_{135}^{123} = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 3 & 5 \end{vmatrix} = 0, \quad M_{132}^{123} = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 3 & 3 & 0 \end{vmatrix} = 0.$$

All three determinants are equal to zero, since the third row is a sum of the first two. Thus, it's impossible to find a nonzero minor of the third order, i.e. the rank of matrix C is equal to 2. ■

3.3.2. Elementary Transformations Method

Let A be an $m \times n$ matrix. To calculate its rank we need to make the following steps.

1. Reduce the matrix to echelon form (see the method in section 1.2.6).
2. Calculate the number r of nonzero rows of the obtained matrix. This number is equal to the rank of matrix A.

This method is based on Property 8 (see section 3.2). The basis minor of a matrix in echelon form (1.1) is a minor

$$M=\begin{vmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix},$$

composed of columns containing unity elements (at the beginning of each "step"). This determinant of triangular form is nonzero (equals to 1), and each of its bounding minors (if it exists) is equal to zero, because it contains a zero row.

Example 3.6. Find ranks of matrices, using elementary transformations method

$$O=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},\ A=\begin{pmatrix} 3 & 9 \\ 2 & 4 \end{pmatrix},\quad B=\begin{pmatrix} 0 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix},$$

$$C=\begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 2 & 0 & 1 & 1 & 2 \\ 3 & 0 & 3 & 2 & 5 \end{pmatrix},\quad D=\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \\ 3 & 2 & 5 \end{pmatrix}.$$

□ *Matrix* O:

1. A zero matrix is already in echelon form (see definition in section 1.2.6).
2. A number of nonzero rows is equal to zero. Hence, $\operatorname{rg} O=0$.

Matrix A:

1. Reduce matrix A to echelon form (see example 1.18):

$$A\sim\begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}.$$

2. There are two nonzero rows in this matrix. Hence, $\operatorname{rg} A=2$.

Matrix B:

1. Reduce matrix B to echelon form (see example 1.18):

$$B \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1.5 \end{pmatrix}.$$

2. There are two nonzero rows in this matrix. Hence, $\operatorname{rg} B = 2$.

Matrix C:

1. Reduce matrix C to echelon form. Choose $a_{11} = 1$ as a pivot and make all the other elements of the first column equal to zero: add the first row, multiplied by (-2), to the second row, and the first row, multiplied by (-3), to the third one. We get matrix

$$C = \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 2 & 0 & 1 & 1 & 2 \\ 3 & 0 & 3 & 2 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 0 & -3 & -1 & -4 \\ 0 & 0 & -3 & -1 & -4 \end{pmatrix},$$

that has two equal rows. By Property 4 (see section 3.2), we delete one of the equal rows: $\begin{pmatrix} 1 & 0 & 2 & 1 & 3 \\ 0 & 0 & -3 & -1 & -4 \end{pmatrix}$. We got an echelon form of the matrix.

2. There are two nonzero rows in this matrix. Hence, $\operatorname{rg} C = 2$.

Matrix D:

1. Reduce matrix D to echelon form. We delete the zero row and choose $a_{11} = 1$ as a pivot element to make all other elements of the first column equal to zero:

$$D \sim \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \\ 3 & 2 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -3 \\ 0 & -1 & -1 \\ 0 & -4 & -4 \end{pmatrix}.$$

The last three rows of the matrix are proportional. By Property 4 (see section 3.2) we can delete two of them: $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \end{pmatrix}$. We have got an echelon form of the matrix.

2. There are two nonzero rows in this matrix. Hence, $\operatorname{rg} D = 2$.

Note, that $\operatorname{rg} C = \operatorname{rg} D$, because $D = C^T$ by Property 7 (see section 3.2). ■

EXERCISES

1. Calculate ranks of the matrices:

$$\text{a) } \begin{pmatrix} 1 & m & n & m \\ 2 & 1 & 1 & n \\ 3 & m+1 & n+1 & m+n \end{pmatrix}; \quad \text{b) } \begin{pmatrix} 1 & 2 & 1 & 2 \\ m & n & m & n \\ -m & n & -m & n \\ 1 & 2 & 1 & 2 \end{pmatrix}$$

using the method of bounding minors and elementary transformations method.

2. Calculate ranks of the matrices:

$$\text{a) } A=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \text{ b) } A=\begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}; \text{ c) } A=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \text{ d) } A=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \text{ e) } A=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix};$$

$$\text{f) } A=\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 3 & 3 & 4 \end{pmatrix}; \text{ g) } A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{pmatrix}; \text{ h) } A=\begin{pmatrix} 1 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 11 \end{pmatrix}.$$

3. Calculate ranks of the matrices using the elementary transformations method:

$$\text{a) } A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 5 & 6 \\ 3 & 6 & 13 & 16 \end{pmatrix}; \quad \text{b) } A=\begin{pmatrix} 25 & 31 & 17 & 43 \\ 75 & 94 & 53 & 132 \\ 75 & 94 & 54 & 134 \\ 25 & 32 & 20 & 48 \end{pmatrix}; \quad \text{c) } A=\begin{pmatrix} 0 & 2 & -4 \\ -1 & -4 & 5 \\ 3 & 1 & 7 \\ 0 & 5 & -10 \\ 2 & 3 & 0 \end{pmatrix};$$

$$\text{d) } A=\begin{pmatrix} 1 & 0 & 4 & -1 \\ 2 & 1 & 11 & 2 \\ 11 & 4 & 56 & 5 \\ 2 & -1 & 5 & -6 \end{pmatrix}; \quad \text{e) } A=\begin{pmatrix} 24 & 19 & 36 & 72 & -38 \\ 49 & 40 & 73 & 147 & -80 \\ 73 & 59 & 98 & 219 & -118 \\ 47 & 36 & 71 & 141 & -72 \end{pmatrix}.$$

4. Calculate ranks of the matrices using the method of bounding minors:

$$\text{a) } A=\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 5 & 6 \\ 3 & 6 & 13 & 16 \end{pmatrix}; \quad \text{b) } A=\begin{pmatrix} 1 & 0 & 2 & 0 \\ 4 & 3 & 1 & 2 \\ 6 & 3 & 5 & 2 \\ 5 & 3 & 3 & 2 \end{pmatrix}.$$

CHAPTER 4. INVERSE MATRIX

4.1. DEFINITION, EXISTENCE AND UNIQUENESS OF INVERSE MATRIX

Let's consider a problem of definition of an operation, opposite to the multiplication of matrices.

Let A be a square matrix of order n. Matrix A^{-1}, satisfying with the given matrix A to the equalities

$$A^{-1} \cdot A = A \cdot A^{-1} = E,$$

is called the ***inverse*** of A. Matrix A is called ***invertible***, if there exists an inverse matrix, otherwise it is called ***noninvertible***. By definition A and A^{-1} are permutation matrices.

From the definition it follows that if an inverse matrix A^{-1}exists, it is a square matrix of the same order as A.

A square matrix $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ *with nonzero determinant has an inverse matrix, which is unique*:

$$A^{-1} = \frac{1}{\det A} \cdot \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} = \frac{1}{\det A} \cdot A^{+}, \tag{4.1}$$

where $A^{+} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$ *is a transpose of a matrix composed of cofactors of matrix* A.

Matrix A^{+} is called the ***adjoint matrix*** of A (see section.2.3.1).

The operation of matrix inversion allows us to define an integer negative power of a matrix. For a nonsingular matrix A and any natural number n we have $A^{-n} = \left(A^{-1}\right)^n$.

4.2. PROPERTIES OF INVERSE MATRIX

The operation of matrix inversion has the following properties:

1. $\left(A^{-1}\right)^{-1} = A$,
2. $\left(A \cdot B\right)^{-1} = B^{-1} \cdot A^{-1}$,
3. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$,
4. $\det A^{-1} = \dfrac{1}{\det A}$,
5. $E^{-1} = E$,

if the operations in equalities $1-4$ have sense.

A matrix that is inverse to a nonsingular diagonal matrix is also diagonal:

$$\left[\, diag\left(a_{11}, a_{22}, \ldots, a_{nn}\right)\right]^{-1} = diag\left(\frac{1}{a_{11}}, \frac{1}{a_{22}}, \ldots, \frac{1}{a_{nn}}\right).$$

4.3. METHODS OF MATRIX INVERSION

Let A be a square matrix. We need to find the inverse matrix A^{-1}.

Algorithm for finding the inverse of a matrix using the adjoint matrix (first method)

1. Evaluate the determinant $\det A$ of the given matrix. If $\det A = 0$, the inverse matrix does not exist (matrix A is singular).

2. Calculate matrix (A_{ij}) of cofactors $A_{ij} = \left(-1\right)^{i+j} M_{ij}$ of matrix A.

3. Transposing matrix (A_{ij}), obtain the adjoint matrix $A^{+} = (A_{ij})^T$.

4. Compose the inverse matrix (4.1), by dividing all the elements of the adjoint matrix by the determinant $\det A$:

$$A^{-1} = \frac{1}{\det A} \cdot A^{+}.$$

Algorithm for finding the inverse of a matrix using elementary transformations (second method)

1. Compose a block matrix $(A \mid E)$, by adding an identity matrix of the same order to the right of A.

2. Using elementary row transformations of matrix $(A \mid E)$, reduce its left block to the simplest form Λ (1.3). In doing so, the block matrix takes on form $(\Lambda \mid S)$, where S is a square matrix, obtained from an identity matrix E by applying the elementary transformations.

3. If $\Lambda = E$, then block S is the inverse matrix, i.e. $S = A^{-1}$. If $\Lambda \neq E$, then matrix A is noninvertible.

For a nonsingular matrix A this method of finding the inverse matrix is illustrated by the following scheme:

$$(A \mid E) \xrightarrow{\text{Elementary row transformations}} (E \mid A^{-1})$$

For nonsingular square matrices of the second order $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ there is a simple rule for finding an inverse matrix, which follows from the first method:

1) switch the elements on the main diagonal;

2) change the signs of elements on the secondary diagonal;

3) divide the obtained matrix by the determinant $\det A = ad - bc \neq 0$:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{4.2}$$

Example 4.1. Given $A = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}$, find the inverse matrix.

□ *First method.*

1. Find the determinant $\det A = \begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} = 2 \neq 0$. Since the determinant is nonzero, matrix A is nonsingular and, therefore, has an inverse matrix.

2. Compose a matrix of cofactors: $(A_{ij})=\begin{pmatrix} 4 & -1 \\ -2 & 1 \end{pmatrix}$.

3. Transposing the matrix (A_{ij}), we get the adjoint matrix $A^{+}=(A_{ij})^{T}=\begin{pmatrix} 4 & -2 \\ -1 & 1 \end{pmatrix}$.

4. Dividing all the elements of the adjoint matrix by the determinant $\det A=2$, find the inverse matrix:

$$A^{-1}=\frac{1}{2}\cdot\begin{pmatrix} 4 & -2 \\ -1 & 1 \end{pmatrix}=\begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Let's check $A^{-1}A=\begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=E$.

Using the rule (4.2), for matrix $A=\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we get

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}=\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -1 & 1 \end{pmatrix}=\begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Note that $\det A^{-1}=\frac{1}{2}=\frac{1}{\det A}$.

Second method.

1. Compose a block matrix

$$(A\mid E)=\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right).$$

2. Applying elementary row transformations, reduce it to the simplest form $(E\mid A^{-1})$. Add the first row, multiplied by (-1), to the second row:

$$(A\mid E)=\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{array}\right)\sim\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 2 & -1 & 1 \end{array}\right).$$

Now add the second row, multiplied by (-1), to the first one:

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

To obtain an identity matrix in the left block we need to divide the second row by 2:

$$\left(\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 2 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} \underbrace{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}_{E_2} & \begin{matrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{matrix} \end{array}\right).$$

In the right block we have the inverse matrix $A^{-1} = \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$. ■

Example 4.2. Given $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix}$, find the inverse matrix.

□ *First method.*

1. Find the determinant $\det A = 2$.

2. Find the cofactors of matrix A:

$$A_{11} = (-1)^{1+1} \cdot \begin{vmatrix} 1 & 0 \\ 2 & 2 \end{vmatrix} = 2; \qquad A_{12} = (-1)^{1+2} \cdot \begin{vmatrix} 0 & 0 \\ 0 & 2 \end{vmatrix} = 0; \qquad A_{13} = (-1)^{1+3} \cdot \begin{vmatrix} 0 & 1 \\ 0 & 2 \end{vmatrix} = 0;$$

$$A_{21} = (-1)^{2+1} \cdot \begin{vmatrix} 2 & 1 \\ 2 & 2 \end{vmatrix} = -2; \qquad A_{22} = (-1)^{2+2} \cdot \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = 2; \qquad A_{23} = (-1)^{2+3} \cdot \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = -2;$$

$$A_{31} = (-1)^{3+1} \cdot \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} = -1; \qquad A_{32} = (-1)^{3+2} \cdot \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} = 0; \qquad A_{33} = (-1)^{3+3} \cdot \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1$$

and compose matrix $(A_{ij}) = \begin{pmatrix} 2 & 0 & 0 \\ -2 & 2 & -2 \\ -1 & 0 & 1 \end{pmatrix}$.

3. Transposing the matrix (A_{ij}), get the adjoint matrix $A^{+} = (A_{ij})^T = \begin{pmatrix} 2 & -2 & -1 \\ 0 & 2 & 0 \\ 0 & -2 & 1 \end{pmatrix}$.

4. Dividing all the elements of the adjoint matrix by the determinant $\det A = 2$, we get the inverse matrix:

$$A^{-1} = \frac{1}{\det A} A^{+} = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & -1 & \frac{1}{2} \end{pmatrix}.$$

Let's check the equality $A^{-1}A = E$: $\begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & -1 & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Second method.

1. Compose a block matrix $(A \mid E)$ by writing to the right of A an identity matrix of the same order:

$$(A \mid E) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 2 & 0 & 0 & 1 \end{array}\right).$$

2. Applying elementary row transformations, reduce it to the form $(E \mid A^{-1})$:

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

$$\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & \frac{1}{2} \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ \underbrace{0 \quad 0 \quad 1}_{E_3} & & & 0 & -1 & \frac{1}{2} \end{array}\right).$$

In the right block we have the inverse matrix $A^{-1} = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & -1 & \frac{1}{2} \end{pmatrix}$. ■

4.4. MATRIX EQUATIONS

Consider a matrix equation

$$A \cdot X = B, \tag{4.3}$$

where A and B are given matrices with the same number of rows (and matrix A is a square matrix). It is required to find matrix X that satisfies the equation (4.3).

If the determinant of matrix A is nonzero, then matrix equation (4.3) has a unique solution

$$X = A^{-1} \cdot B.$$

Let's also consider the following matrix equation

$$Y \cdot A = B, \tag{4.4}$$

where A and B are given matrices with the same number of columns (and matrix A is a square matrix). It is required to find matrix Y that satisfies the equation (4.4).

If the determinant of matrix A is nonzero, then equation (4.4) has a unique solution

$$Y = B \cdot A^{-1}.$$

Example 4.3. Given the matrices

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

solve equations: a) $A \cdot X = B$; b) $Y \cdot A = B$; c) $Y \cdot A = C$.

□ The inverse matrix $A^{-1} = \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ was found in example 4.1.

a) The solution of equation $A \cdot X = B$ is obtained by the multiplication of both parts of the equation by A^{-1} from the left:

$$X = A^{-1} \cdot B = \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 2 & 4 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

b) The equation has no solutions, because matrices A and B have different number of columns ($2 \neq 3$).

c) The solution of equation $YA = C$ is obtained by the multiplication of both parts of the equation by A^{-1} from the right:

$$Y = CA^{-1} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4 & -1 \\ 7 & -2 \end{pmatrix}. \ ■$$

Example 4.4. Solve the equation $A \cdot X \cdot B = C$, for

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

☐ Inverse matrices $A^{-1} = \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$ and $B^{-1} = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & -1 & \frac{1}{2} \end{pmatrix}$ were found in Examples 4.1 and 4.2, respectively.

Find the solution of the matrix equation by the formula

$$X = A^{-1} \cdot C \cdot B^{-1} = \begin{pmatrix} 2 & -1 \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & -1 & \frac{1}{2} \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 2 & 4 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & -1 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & -2 & 2 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}. \blacksquare$$

EXERCISES

1. Find inverse matrices for the given ones:

a) $\begin{pmatrix} m & n \\ -m & n \end{pmatrix}$; b) $\begin{pmatrix} m & n & -m \\ -n & 1 & 1 \\ m & -1 & 1 \end{pmatrix}$; c) $\begin{pmatrix} m & 1 & 2 \\ 0 & n & 3 \\ 0 & 0 & m+n \end{pmatrix}$.

2. Solve matrix equations:

a) $\begin{pmatrix} m & n \\ -m & n \end{pmatrix} \cdot X = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2 \cdot X$; b) $X \cdot \begin{pmatrix} m & n \\ -m & n \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2 \cdot X$;

c) $\begin{pmatrix} 1 & 1 \\ -m & n \end{pmatrix} \cdot X \cdot \begin{pmatrix} m & n \\ -n & m \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$; d) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & m \\ 1 & -n & 1 \end{pmatrix} \cdot X = \begin{pmatrix} 1 & 1 & 1 \\ m & m & m \\ n & n & n \end{pmatrix}$.

CHAPTER 5. SYSTEMS OF LINEAR ALGEBRAIC EQUATIONS

5.1. BASIC CONCEPTS AND DEFINITIONS

System of m ***linear algebraic equations with*** n ***unknowns*** is represented by the following formula

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m . \end{cases} \tag{5.1}$$

Numbers a_{ij}, $i = 1,\ldots,m$, $j = 1,\ldots,n$ are called ***coefficients of the system***; $b_1, b_2,\ldots,b_m$ – constant terms; $x_1, x_2,\ldots,x_n$ – ***unknowns***. The number of equations m can be less, more or equal to the number of unknowns n.

System solution is an ordered set of n numbers $(\alpha_1,\alpha_2,\ldots,\alpha_n)$ such that, if we substitute unknowns $x_1, x_2,\ldots,x_n$ with corresponding numbers $\alpha_1,\alpha_2,\ldots,\alpha_n$, then each equation of the system will be correct.

System is called ***consistent***, if it has at least one solution. If a system has no solutions, it is called inconsistent.

Consistent system is called ***determined***, if it has a unique solution, otherwise, if there is more than one solution, then system is called ***underdetermined***.

System (5.1) is called ***homogeneous***, if all constant terms are equal to zero:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \qquad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 . \end{cases} \tag{5.2}$$

Systems of general form (5.1) are called ***nonhomogeneous***.

System (5.1) is usually written in ***matrix form***. To do this, it is necessary to write the coefficients of the system as a ***coefficient matrix*** $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$, constant terms are written as a ***constant term column*** $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$, and unknowns – as an ***unknown column*** $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Matrix form of an nonhomogeneous system of equations (5.1) is given by

$$Ax = b, \tag{5.3}$$

and of a homogeneous system of equations (5.2):

$$Ax = o, \tag{5.4}$$

where symbol o on the right hand side denotes zero column of sizes $m \times 1$.

Matrix form (5.3) of a system of equations can be represented equivalently as the following:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \cdot x_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} \cdot x_2 + \ldots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \cdot x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Then the system solution is represented by a column $x = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ and satisfies the equation

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} \cdot \alpha_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} \cdot \alpha_2 + \ldots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \cdot \alpha_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \tag{5.5}$$

i.e. a constant term column is a linear combination of columns of a coefficient matrix.

5.2. CRAMER'S RULE

Consider the following case: the number of equations m is equal to the number of unknowns n ($m = n$), i.e. we have the following system

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ \qquad \cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n, \end{cases} \tag{5.6}$$

where coefficient matrix is a square matrix of the n-th order: $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$. Its determinant will be denoted by

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

Cramer's rule. *If the determinant Δ of the coefficient matrix of a system with n linear equations and n unknowns is nonzero, then the system has a unique solution, which is obtained by the following formulas*

$$x_i=\frac{\Delta_i}{\Delta}, \quad i=1,\ldots,n,$$

where Δ_i is the determinant of a matrix, obtained by the substitution of the i-th column with the constant term column, i.e. $\Delta_i=\begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2\,i-1} & b_2 & a_{2\,i+1} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{n\,i-1} & b_n & a_{n\,i+1} & \cdots & a_{nn} \end{vmatrix}$.

If $\Delta=0$ and at least one $\Delta_i\neq 0$, then the given system is inconsistent.

If $\Delta=\Delta_1=\ldots=\Delta_n=0$, two cases are possible: the given system can be both inconsistent or underdetermined.

Example 5.1. Solve the system of linear equations

$$\begin{cases} 2x_1+2x_2+x_3=9, \\ x_1+x_2 \qquad =3, \\ \quad 2x_2+x_3=7. \end{cases}$$

□ Let's compose the coefficient matrix $A=\begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$ and calculate its determinant

$\Delta = \begin{vmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{vmatrix} = 2+2-2=2$ (example 2.1). As the determinant is nonzero, the system has a unique solution (by Cramer's rule).

Find determinants Δ_i and unknowns x_i $(i=1,2,3)$:

$$\Delta_1 = \begin{vmatrix} 9 & 2 & 1 \\ 3 & 1 & 0 \\ 7 & 2 & 1 \end{vmatrix} = 9+6-7-6=2, \qquad x_1 = \frac{2}{2} = 1;$$

$$\Delta_2 = \begin{vmatrix} 2 & 9 & 1 \\ 1 & 3 & 0 \\ 0 & 7 & 1 \end{vmatrix} = 6+7-9=4, \qquad x_2 = \frac{4}{2} = 2;$$

$$\Delta_3 = \begin{vmatrix} 2 & 2 & 9 \\ 1 & 1 & 3 \\ 0 & 2 & 7 \end{vmatrix} = 14+18-12-14=6, \quad x_3 = \frac{6}{2} = 3. \ \blacksquare$$

5.3. SYSTEM OF LINEAR EQUATIONS CONSISTENCY CONDITION

Consider system (5.3) of m linear equations with n unknowns. Compose a block matrix by the addition of the constant term column to the right of matrix A. We obtain an ***augmented coefficient matrix***:

$$(A \mid b) = \left(\begin{array}{cccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right)_{m\times(n+1)} \tag{5.7}$$

This matrix contains the whole information about the system of equations except for the unknowns denotation.

Kronecker–Capelli theorem. *System* $Ax=b$ *is consistent if and only if the rank of the coefficient matrix is equal to the rank of the augmented coefficient matrix*: $\mathrm{rg}\, A = \mathrm{rg}\,(A \mid b)$.

Example 5.2. Has the system

$$\begin{cases} x_1 + 2x_3 + x_4 = 1, \\ 2x_1 + x_2 + x_4 = 0, \\ 3x_1 + x_2 + 2x_3 + 2x_4 = 2 \end{cases}$$

any solutions?

□ Let's compose coefficient and augmented coefficient matrices:

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 2 & 1 & 0 & 1 \\ 3 & 1 & 2 & 2 \end{pmatrix}, \qquad (A \mid b) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 1 & 1 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 2 & 2 & 2 \end{array}\right).$$

The rank of matrix A is equal to 2, because it has nonzero minors of the second order and its third row equals to the sum of the first and the second rows. Therefore, the third row can be excluded by the Property 4 (section 3.2), and the rank will remain the same.

The rank of the augmented matrix is equal to 3, because it has nonzero minor of the third order, e.g. the minor, composed from the first, second and the last columns of the augmented matrix:

$$M_{125}^{123} = \begin{vmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & 2 \end{vmatrix} = 2 + 2 - 3 = 1 \neq 0.$$

Hence, $\operatorname{rg} A \neq \operatorname{rg}(A \mid b)$ and the system is inconsistent (has no solutions). ■

5.4. GAUSS–JORDAN ALGORITHM FOR LINEAR EQUATIONS SYSTEM SOLUTION

Consider system (5.1) of m linear equations with n unknowns. To obtain solutions it is necessary to make the following steps:

1. Compose an augmented coefficient matrix (5.7):

$$(A \mid b) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right).$$

2. By the elementary row transformations of matrix $(A \mid b)$, reduce the matrix to echelon form (section 1.2.6). If the basis minor of matrix A is situated in the first r rows and r columns, we will obtain the following form:

$$\left(\tilde{A} \mid \tilde{b}\right)=\left(\begin{array}{ccccccc|c} 1 & \tilde{a}_{12} & \cdots & \tilde{a}_{1r} & \cdots & \tilde{a}_{1n} & & \tilde{b}_1 \\ 0 & 1 & \cdots & \tilde{a}_{2r} & \cdots & \tilde{a}_{2n} & & \tilde{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & \tilde{a}_{rn} & & \tilde{b}_r \\ 0 & 0 & \cdots & 0 & \cdots & 0 & & \tilde{b}_{r+1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & & 0 \end{array}\right). \tag{5.8}$$

3. Check the system's consistency. To do this it is necessary to find ranks of matrices A and $(A \mid b)$:

$$\operatorname{rg} A=\operatorname{rg} \tilde{A}=r \text{ – number of nonzero rows in matrix } \tilde{A};$$

$$\operatorname{rg}(A \mid b)=\operatorname{rg}\left(\tilde{A} \mid \tilde{b}\right)=\begin{cases} r+1, \text{ if } \tilde{b}_{r+1} \neq 0, \\ r, \text{ if } \tilde{b}_{r+1}=0. \end{cases}$$

If $\operatorname{rg} A \neq \operatorname{rg}(A \mid b)$ and $\tilde{b}_{r+1} \neq 0$, then the system has no solutions. The algorithm should be terminated.

If $\operatorname{rg} A=\operatorname{rg}(A \mid b)$ and $\tilde{b}_{r+1}=0$, then the system is consistent. The process should continue.

4. If the system is consistent ($\operatorname{rg} A=\operatorname{rg}(A \mid b)=r$), matrix $\left(\tilde{A} \mid \tilde{b}\right)$ should be reduced to ***echelon*** form (section 1.2.6). Using the elementary row transformations the matrix should be reduced to the form, in which each column (which is a part of the basis minor) has all elements equal to zero except for one (which is equal to 1). If the basis minor of matrix A is situated in the first r rows and first r columns, the matrix can be reduced to the following simplified form:

$$
(A' \mid b') = \left(\begin{array}{cccccccc|c}
1 & 0 & \cdots & 0 & a'_{1r+1} & \cdots & a'_{1n} & b'_1 \\
0 & 1 & \cdots & 0 & a'_{2r+1} & \cdots & a'_{2n} & b'_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & a'_{rr+1} & \cdots & a'_{rn} & b'_r \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0
\end{array}\right). \tag{5.9}
$$

Previous four steps are called the ***forward pass*** of Gauss–Jordan algorithm. As the result of the forward pass the initial system substantially simplifies to the form $A'x = b'$:

$$
\begin{cases}
x_1 + a'_{1\,r+1}x_{r+1} + \ldots + a'_{1\,n}x_n = b'_1, \\
\qquad \cdots \\
x_r + a'_{r\,r+1}x_{r+1} + \ldots + a'_{r\,n}x_n = b'_r.
\end{cases} \tag{5.10}
$$

5. By the simplified form (5.9) we divide all unknowns $x_1, x_2, \ldots, x_n$ into two groups: basis and free. Unknowns, which correspond to the columns, that form basis minor, are called ***basis variables***, other unknowns – ***free variables***.

For the system (5.10) basis variables are $x_1, x_2, \ldots, x_r$, free variables are $x_{r+1}, x_{r+2}, \ldots, x_n$. Denominate basis variables (5.10) by free ones:

$$
\begin{cases}
x_1 = b'_1 - a'_{1\,r+1}x_{r+1} - \ldots - a'_{1\,n}x_n, \\
\qquad \vdots \\
x_r = b'_r - a'_{r\,r+1}x_{r+1} - \ldots - a'_{r\,n}x_n.
\end{cases} \tag{5.11}
$$

If matrix rank r equals to the number of unknowns n ($r = \operatorname{rg} A = n$), then the left block of matrix (5.9) will be an identity matrix E_n:

$$
(A' \mid b') = \left(\begin{array}{cccc|c}
1 & 0 & \cdots & 0 & b'_1 \\
0 & 1 & \cdots & 0 & b'_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & b'_n
\end{array}\right).
$$

All unknowns $x_1, x_2, \ldots, x_n$ will be basis and formula (5.11) will define the unique solution of the system:

$$\begin{cases} x_1 = b'_1, \\ x_2 = b'_2, \\ \quad\vdots \\ x_n = b'_n. \end{cases} \tag{5.12}$$

If matrix rank is less than the number of unknowns ($\mathrm{rg}\, A < n$), then the system will have an infinite number of solutions, defined by formula (5.11), which will have the following properties:

- for any values of free variables $x_{r+1}, x_{r+2}, \ldots, x_n$ by the formula (5.11) we will obtain such values of basis variables, that the ordered set of numbers $x_1, x_2, \ldots, x_n$ will be the solution of the system (5.1);
- any solution $x_1, x_2, \ldots, x_n$ of system (5.1) will satisfy equalities (5.11).

Equalities (5.11), which denominate the basis variables by the free ones, are called the ***general solution*** of the system (5.1).

The solution, obtained by formula (5.11) with the exact values of free variables, is called the ***particular solution*** of the system (5.1).

The process of the solution of a consistent system (5.1) is terminated with obtaining the formula (5.11) of a general solution (in particular, the process is terminated with the definition of the exact solution (5.12)).

Step 5 is called the ***backward pass*** of Gauss–Jordan algorithm.

Example 5.3. Solve the systems of equations

$$\text{a)} \begin{cases} x_1 + 2x_2 - 2x_3 = 1, \\ x_1 + 3x_2 - 3x_3 = 1, \\ 3x_1 + x_2 - 2x_3 = 1, \\ 2x_1 - x_2 + x_3 = 3; \end{cases} \qquad \text{b)} \begin{cases} x_1 + 2x_2 - 2x_3 = 1, \\ x_1 + 3x_2 - 3x_3 = 1, \\ 3x_1 + x_2 - 2x_3 = 1, \\ 2x_1 - x_2 + x_3 = 2; \end{cases}$$

$$\text{c)} \begin{cases} x_1 + x_2 + 2x_3 = 4, \\ x_1 + 2x_2 + 3x_3 = 5; \end{cases} \qquad \text{d)} \begin{cases} x_1 + x_2 + 2x_3 + x_4 = 1, \\ 2x_1 + 3x_2 + x_4 = 0, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 1. \end{cases}$$

□ a) 1. Compose the augmented coefficient matrix: $(A \mid b)=\left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 1 & 3 & -3 & 1 \\ 3 & 1 & -2 & 1 \\ 2 & -1 & 1 & 3 \end{array}\right)$.

2. By the elementary transformations of rows of matrix $(A \mid b)$, we reduce it to echelon form. We choose $a_{11}=1\neq 0$ as a pivot element. We add the first row multiplied by (-1) to the second row, the first row multiplied by (-3) – to the third row, the first row multiplied by (-2) – to the fourth row:

$$(A \mid b)=\left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 1 & 3 & -3 & 1 \\ 3 & 1 & -2 & 1 \\ 2 & -1 & 1 & 3 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & 4 & -2 \\ 0 & -5 & 5 & 1 \end{array}\right).$$

The pivot element is $a_{22}=1\neq 0$. We add the second row multiplied by 5 to the third and to the fourth rows:

$$(A \mid b)\sim\left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & 4 & -2 \\ 0 & -5 & 5 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

The augmented matrix is reduced to echelon form.

3. We calculate the ranks of matrices: $\operatorname{rg} A=3$, $\operatorname{rg}(A \mid b)=4$. By the Kronecker–Capelli theorem the system is inconsistent. The last equation of the system has the following form: $0=1$ (incorrect equality). Thus, the system has no solutions.

$$\text{b)}\ \begin{cases} x_1+2x_2-2x_3=1, \\ x_1+3x_2-3x_3=1, \\ 3x_1+x_2-2x_3=1, \\ 2x_1-x_2+x_3=2. \end{cases}$$

1. Compose the augmented coefficient matrix: $(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 1 & 3 & -3 & 1 \\ 3 & 1 & -2 & 1 \\ 2 & -1 & 1 & 2 \end{array}\right)$.

The only difference from system "a" is the element $b_4 = 2$.

2. Reduce the augmented matrix to the echelon form by the repetition of the same steps from example "a":

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 1 & 3 & -3 & 1 \\ 3 & 1 & -2 & 1 \\ 2 & -1 & 1 & 2 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & 4 & -2 \\ 0 & -5 & 5 & 0 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right) = (\tilde{A} \mid \tilde{b}).$$

3. Calculate the ranks of matrices: $\mathrm{r} = \mathrm{rg}\, A = \mathrm{rg}(A \mid b) = 3$. By the Kronecker–Capelli theorem the system is consistent.

4. Reduce the matrix to simplified form. As the basis minor we choose M_{123}^{123}. We add the second row multiplied by (-2) to the first row, and then we add the third row to the second row:

$$(\tilde{A} \mid \tilde{b}) = \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right) = (A' \mid b').$$

5. The matrix rank r equals to the number of unknowns n ($r = 3 = n$). Thus, the system has a unique solution (all unknowns x_1, x_2, x_3 will be basis and there will be no free variables). From the simplified matrix $(A' \mid b')$ we obtain the unique solution: $x_1 = 1$; $x_2 = 2$; $x_3 = 2$, which is represented by the column $x = (1 \quad 2 \quad 2)^T$.

c) $\begin{cases} x_1 + x_2 + 2x_3 = 4, \\ x_1 + 2x_2 + 3x_3 = 5. \end{cases}$

1. Compose the augmented coefficient matrix: $(A \mid b) = \left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 1 & 2 & 3 & 5 \end{array}\right)$.

2. By the elementary transformations of rows of matrix $(A \mid b)$, we reduce it to the echelon form. We choose $a_{11} = 1 \neq 0$ as a pivot element. We add the first row multiplied by (-1) to the second row:

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 1 & 2 & 3 & 5 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 0 & 1 & 1 & 1 \end{array}\right) = (\tilde{A} \mid \tilde{b}).$$

The augmented coefficient matrix is now in echelon form.

3. Calculate the ranks of matrices: $\operatorname{rg} A = \operatorname{rg}(A \mid b) = 2$. By the Kronecker-Capelli theorem the system is consistent.

4. Reduce the matrix to simplified form. We choose M_{12}^{12} as the basis minor and add the second row multiplied by (-1) to the first row:

$$(\tilde{A} \mid \tilde{b}) = \left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 0 & 1 & 1 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 1 \end{array}\right) = (A' \mid b').$$

5. Variables x_1, x_2 are basis and x_3 is free. We write the general solution accordingly to the formula (5.11):

$$\begin{cases} x_1 = 3 - x_3, \\ x_2 = 1 - x_3. \end{cases}$$

The system has an infinite number of solutions. Let's find a particular solution, e.g. for $x_3 = 0$ we get $x_1 = 3$, $x_2 = 1$. Thus, the column $x = (3 \quad 1 \quad 0)^T$ is particular solution of the system.

d) $\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 1, \\ 2x_1 + 3x_2 + x_4 = 0, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 1. \end{cases}$

1. Compose the augmented coefficient matrix: $(A \mid b)=\left(\begin{array}{cccc|c}1 & 1 & 2 & 1 & 1\\ 2 & 3 & 0 & 1 & 0\\ 3 & 4 & 2 & 2 & 1\end{array}\right)$.

2. By the elementary transformations of rows of matrix $(A \mid b)$, we reduce it to the echelon form. We choose $a_{11}=1\neq 0$ as a pivot element. We add the first row multiplied by (-2) to the second row, and the first row multiplied by (-3) to the third row:

$$(A \mid b)=\left(\begin{array}{cccc|c}1 & 1 & 2 & 1 & 1\\ 2 & 3 & 0 & 1 & 0\\ 3 & 4 & 2 & 2 & 1\end{array}\right)\sim\left(\begin{array}{cccc|c}1 & 1 & 2 & 1 & 1\\ 0 & 1 & -4 & -1 & -2\\ 0 & 1 & -4 & -1 & -2\end{array}\right).$$

The pivot element $a_{22}=1\neq 0$. We add the second row multiplied by (-1) to the third row:

$$(A \mid b)\sim\left(\begin{array}{cccc|c}1 & 1 & 2 & 1 & 1\\ 0 & 1 & -4 & -1 & -2\\ 0 & 1 & -4 & -1 & -2\end{array}\right)\sim\left(\begin{array}{cccc|c}1 & 1 & 2 & 1 & 1\\ 0 & 1 & -4 & -1 & -2\\ 0 & 0 & 0 & 0 & 0\end{array}\right)=(\tilde{A} \mid \tilde{b}).$$

The augmented matrix is now in echelon form.

3. Calculate the ranks of matrices: $\operatorname{rg} A=\operatorname{rg}(A \mid b)=2$. By the Kronecker–Capelli theorem the system is consistent.

4. Now the matrix should be reduced to simplified form. We choose M_{12}^{12} as the basis minor. We add the second row multiplied by (-1) to the first row:

$$(\tilde{A} \mid \tilde{b})=\left(\begin{array}{cccc|c}1 & 1 & 2 & 1 & 1\\ 0 & 1 & -4 & -1 & -2\\ 0 & 0 & 0 & 0 & 0\end{array}\right)\sim\left(\begin{array}{cccc|c}1 & 0 & 6 & 2 & 3\\ 0 & 1 & -4 & -1 & -2\\ 0 & 0 & 0 & 0 & 0\end{array}\right)=(A' \mid b').$$

5. Variables x_1, x_2 are basis, and x_3, x_4 are free. The general expression can by obtained by the formula (5.11): $\begin{cases} x_1=3-6x_3-2x_4\,, \\ x_2=-2+4x_3+x_4\,. \end{cases}$

The system has an infinite number of solutions. Let's find a particular solution. For example, for $x_3 = x_4 = 0$ we have $x_1 = 3$, $x_2 = -2$. Thus, column $x = \begin{pmatrix} 3 & -2 & 0 & 0 \end{pmatrix}^T$ is a particular solution of the system. ■

5.5. HOMOGENEOUS SYSTEM GENERAL SOLUTION STRUCTURE

A homogeneous system of linear equations

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \quad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \quad \text{or} \quad Ax = o$$

is always consistent, because it has a ***trivial solution*** $x_1 = x_2 = \ldots = x_n = 0$ $(x = o)$.

If the rank of the matrix is equal to the number of unknowns ($\text{rg}\, A = n$), then the trivial solution is the only solution.

Suppose that $r = \text{rg}\, A < n$. Then a homogeneous system has an infinite number of solutions.

Note that an augmented coefficient matrix of a homogeneous system $(A \mid o)$ is reduced by elementary transformations to simplified form $(A' \mid o)$, i.e. $b_1' = b_2' = \ldots = b_r' = 0$ in (5.10). Thus, from (5.11) we obtain a ***general solution of a homogeneous system***:

$$\begin{cases} x_1 = -a'_{1\,r+1}x_{r+1} - \ldots - a'_{1\,n}x_n\,, \\ \qquad \vdots \\ x_r = -a'_{r\,r+1}x_{r+1} - \ldots - a'_{r\,n}x_n\,. \end{cases} \tag{5.13}$$

Properties of homogeneous systems solutions

1. If columns $\varphi_1, \varphi_2, \ldots, \varphi_k$ are solutions of a homogeneous system of equations, then any linear combination $\alpha_1\varphi_1 + \alpha_2\varphi_2 + \ldots + \alpha_k\varphi_k$ is also the solution of a homogeneous system.

2. If the rank of a homogeneous system coefficient matrix equals to r, then the system has $(n - r)$ linearly independent solutions.

Indeed, by the general solution formula (5.13) it is possible to find $(n-r)$ particular solutions $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$, assuming free variables equal to ***standard value sets*** (assuming that all free variables are equal to zero except for the one that equals to 1):

1) $x_{r+1}=1, x_{r+2}=0, \ldots, x_n=0$: $\quad \varphi_1=\begin{pmatrix}-a'_{1\,r+1} & \cdots & -a'_{r\,r+1} & 1 & 0 & \cdots & 0\end{pmatrix}^T$;

2) $x_{r+1}=0, x_{r+2}=1, \ldots, x_n=0$: $\quad \varphi_2=\begin{pmatrix}-a'_{1\,r+2} & \cdots & -a'_{r\,r+2} & 0 & 1 & \cdots & 0\end{pmatrix}^T$;

…

$n-r$) $x_{r+1}=0, x_{r+2}=0, \ldots, x_n=1$: $\varphi_{n-r}=\begin{pmatrix}-a'_{1\,n} & \cdots & -a'_{r\,n} & 0 & 0 & \cdots & 1\end{pmatrix}^T$.

As the result we will get $(n-r)$ solutions:

$$\varphi_1=\begin{pmatrix}-a'_{1\,r+1}\\ \vdots \\ -a'_{r\,r+1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\ \varphi_2=\begin{pmatrix}-a'_{1\,r+2}\\ \vdots \\ -a'_{r\,r+2}\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix},\ldots,\ \varphi_{n-r}=\begin{pmatrix}-a'_{1\,n}\\ \vdots \\ -a'_{r\,n}\\ 0\\ 0\\ \vdots\\ 1\end{pmatrix},$$

which are linearly independent.

Any combination of $(n-r)$ linearly independent solutions $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$ of a homogeneous system is called a ***fundamental system of solutions***.

Note that a fundamental system of solutions is defined ambiguously. A homogeneous system can have different fundamental systems of solutions, each consisting of the same number $(n-r)$ of linearly independent solutions.

Homogeneous system general solution structure. *If a set* $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$ *is a fundamental system of solutions of a homogeneous system* (5.4), *then the column*

$$x=C_1\cdot\varphi_1+C_2\cdot\varphi_2+\cdots+C_{n-r}\cdot\varphi_{n-r} \tag{5.14}$$

for any ***arbitrary values*** *of* C_1, $C_2, \ldots,$ C_{n-r} *is also a solution of system* (5.4), and vice versa, *for any solution* x *of this system it is possible to find such values of* C_1, $C_2, \ldots,$ C_{n-r}, *that make equality* (5.14) *correct.*

Matrix $\Phi = \begin{pmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_{n-r} \end{pmatrix}$, which columns compose a fundamental system of solutions of a homogeneous system, is called ***fundamental***. By the fundamental matrix, general solutions can be expressed in the following form

$$x = \Phi \cdot c,$$

where $c = \begin{pmatrix} C_1 & \cdots & C_{n-r} \end{pmatrix}^T$ – is a column of arbitrary constants.

Homogeneous system solution algorithm

1–5. Make the first five steps of Gauss–Jordan algorithm (section 5.4). At the same time there is no need in checking consistency of the system (because any homogeneous system has trivial solution), so step 3 can be skipped. Get formula (5.11) of a general solution, which will be in form (5.13).

If the rank r of a matrix equals to the number of unknowns n ($r = \operatorname{rg} A = n$), then the system has a unique trivial solution $x = o$ and the solution process terminates.

If the rank of a matrix is less than the number of unknowns ($\operatorname{rg} A < n$), then the system has an infinite number of solutions. The solution set structure will be found in the next steps.

6. Find the fundamental system of solutions $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$ of the homogeneous system. To do this, it is necessary to put the set of $(n-r)$ standard values (where all free variable are equal to zero except for one) consecutively into (5.13) (property 2 of homogeneous system solutions).

7. Write the general solution by the formula (5.14).

Example 5.4. Find fundamental systems of solutions and general solutions of homogeneous systems:

a) $\begin{cases} x_1 + x_2 + 2x_3 = 0, \\ x_1 + 2x_2 + 3x_3 = 0; \end{cases}$ b) $\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 0, \\ 2x_1 + 3x_2 + x_4 = 0, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 0. \end{cases}$

□ a) 1. Compose the augmented coefficient matrix: $(A \mid o)=\left(\begin{array}{ccc|c}1 & 1 & 2 & 0\\ 1 & 2 & 3 & 0\end{array}\right)$.

2–4. By the elementary transformations of rows reduce matrix $(A \mid o)$ to echelon and then to simplified form (example 5.3,"c"):

$$(A \mid o)=\left(\begin{array}{ccc|c}1 & 1 & 2 & 0\\ 1 & 2 & 3 & 0\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 1 & 2 & 0\\ 0 & 1 & 1 & 0\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 0 & 1 & 0\\ 0 & 1 & 1 & 0\end{array}\right)=(A' \mid o).$$

Step 3 is skipped.

5. Variables x_1, x_2 are basis and x_3 is free. Write formula (5.13) for the general solution of the homogeneous system

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

6. Find the fundamental system of solutions. As $n=3$ and $r=\operatorname{rg} A=2$, it is necessary to find $n-r=1$ linearly independent (i.e. nonzero) solutions. We put a standard value of the free variable into the formula of the general solution. If $x_3=1$, then $x_1=-1$, $x_2=-1$, i.e. the fundamental system of equations consists of a single column

$$\varphi_1=\begin{pmatrix}-1\\ -1\\ 1\end{pmatrix}.$$

7. Write the general solution of the homogeneous system by the formula (5.14):

$$x=C_1\cdot\begin{pmatrix}-1\\ -1\\ 1\end{pmatrix},$$

where C_1 is an arbitrary constant.

$$\text{b)}\ \begin{cases} x_1+x_2+2x_3+x_4=0\,, \\ 2x_1+3x_2+x_4=0\,, \\ 3x_1+4x_2+2x_3+2x_4=0\,. \end{cases}$$

1. Compose the augmented coefficient matrix $(A \mid o) = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 0 \\ 2 & 3 & 0 & 1 & 0 \\ 3 & 4 & 2 & 2 & 0 \end{array}\right)$.

2–4. By the elementary transformations of rows reduce matrix $(A \mid o)$ to echelon and then to simplified form (example 5.3,"d"):

$(A' \mid o) = \left(\begin{array}{cccc|c} 1 & 0 & 6 & 2 & 0 \\ 0 & 1 & -4 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$. Step 3 is skipped.

5. Variables x_1, x_2 are basis and x_3, x_4 are free. Write formula (5.13) for the general solution of the homogeneous system: $\begin{cases} x_1 = -6x_3 - 2x_4, \\ x_2 = 4x_3 + x_4. \end{cases}$

6. Find the fundamental system of solutions. As $n = 4$ and $r = \operatorname{rg} A = 2$, it is necessary to find $n - r = 2$ linearly independent solutions. Put standard value sets of free variables into the system:

- if $x_3 = 1$, $x_4 = 0$, then $x_1 = -6$, $x_2 = 4$;
- if $x_3 = 0$, $x_4 = 1$, then $x_1 = -2$, $x_2 = 1$.

As the result we have obtained the following fundamental system of equations

$$\varphi_1 = \begin{pmatrix} -6 \\ 4 \\ 1 \\ 0 \end{pmatrix}, \qquad \varphi_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

7. Write the general solution of the homogeneous system by the formula (5.14):

$$x = C_1 \cdot \begin{pmatrix} -6 \\ 4 \\ 1 \\ 0 \end{pmatrix} + C_2 \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Note that the fundamental system of equations can be obtained with another set of values of free variables, e.g. $x_3 = 1$, $x_4 = 2$ and $x_3 = 2$, $x_4 = 3$.

Then we will get another fundamental system of equations

$$\varphi_1 = \begin{pmatrix} -10 \\ 6 \\ 1 \\ 2 \end{pmatrix}, \quad \varphi_2 = \begin{pmatrix} -18 \\ 11 \\ 2 \\ 3 \end{pmatrix}$$

and the following general solution

$$x = C_1 \cdot \begin{pmatrix} -10 \\ 6 \\ 1 \\ 2 \end{pmatrix} + C_2 \cdot \begin{pmatrix} -18 \\ 11 \\ 2 \\ 3 \end{pmatrix}.$$

In spite of the difference, both formulas describe the same set of solutions. ■

5.6. NONHOMOGENEOUS SYSTEM GENERAL SOLUTION STRUCTURE

In section 5.4 there was the formula (5.11) of a system of linear equations general solution. Let's show another form, which represents the structure of a solution set.

Consider an nonhomogeneous system

$$Ax = b$$

and the ***corresponding*** homogeneous system

$$Ax = o.$$

Nonhomogeneous system general solution structure. *Let x^{p} be the particular solution of an nonhomogeneous system and $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$ compose the fundamental system of equations of the corresponding homogeneous system of equations. Then the following column*

$$x = x^{\text{p}} + C_1 \cdot \varphi_1 + C_2 \cdot \varphi_2 + \cdots + C_{n-r} \cdot \varphi_{n-r} \tag{5.15}$$

for any ***arbitrary values*** of C_1, $C_2, \ldots,$ C_{n-r} *is the solution of the nonhomogeneous system*, and vice versa, *for any solution* x *of this system it is possible to find such values of constants* C_1, $C_2, \ldots,$ C_{n-r}, *that make equality* (5.15) *correct.*

An nonhomogeneous system general solution is a sum of a particular solution of an nonhomogeneous system and a general solution of the corresponding homogeneous system:

$$x = \underbrace{\quad x^{\text{p}} \quad}_{\substack{\textit{nonhomogeneous system}\\ \textit{particular solution}}} + \underbrace{C_1 \cdot \varphi_1 + C_2 \cdot \varphi_2 + \cdots + C_{n-r} \cdot \varphi_{n-r}}_{\substack{\textit{homogeneous system}\\ \textit{general solution}}}.$$

Nonhomogeneous system solution algorithm

1–5. Make the first five steps of Gauss–Jordan algorithm (section 5.4) and get the nonhomogeneous system general solution formula in a form (5.11).

6. Find a particular solution x^{p} of the nonhomogeneous system by substituting free variables in (5.11) with zero.

7. Write formula (5.13) of the corresponding homogeneous system general solution and compose its fundamental system of solutions $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$. To do this it is necessary to put $(n-r)$ standard sets of values (where all free variables are equal to zero except for one) consecutively in (5.13).

8. Write the nonhomogeneous system general solution by the formula (5.15).

Note that the nonhomogeneous system solution ($Ax = b$) can be expressed with the fundamental matrix Φ of the corresponding homogeneous system $Ax = o$ in the following form

$$x = x^{\text{p}} + \Phi \cdot c,$$

where x^{p} is a particular solution of the nonhomogeneous system; $c = \begin{pmatrix} C_1 & \cdots & C_{n-r} \end{pmatrix}^T$ – arbitrary constants column.

Example 5.5. Find the nonhomogeneous system general solution structure (5.15):

a) $\begin{cases} x_1 + x_2 + 2x_3 = 4, \\ x_1 + 2x_2 + 3x_3 = 5; \end{cases}$ b) $\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 1, \\ 2x_1 + 3x_2 + x_4 = 0, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 1. \end{cases}$

□ a) 1–5. The first five steps of Gauss-Jordan algorithm were performed during the solution of example 5.3,"c". The following formula for the nonhomogeneous system general solution was obtained:

$$\begin{cases} x_1 = 3 - x_3\,, \\ x_2 = 1 - x_3\,. \end{cases}$$

Variables x_1, x_2 are basis; x_3 is free.

6. Assuming $x_3 = 0$, we obtain the particular solution of the nonhomogeneous system $x^{\text{p}} = \begin{pmatrix} 3 & 1 & 0 \end{pmatrix}^T$ (example 5.3,"c").

7. Find the fundamental system of equations of the corresponding homogeneous system (example 5.4,"a"): $\varphi_1 = \begin{pmatrix} -1 & -1 & 1 \end{pmatrix}^T$.

8. By the formula (5.15) write the nonhomogeneous system general solution

$$x = x^{\text{p}} + C_1 \cdot \varphi_1 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + C_1 \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

The initial solution set structure is found.

$$\text{b)} \begin{cases} x_1 + x_2 + 2x_3 + x_4 = 1\,, \\ 2x_1 + 3x_2 + x_4 = 0\,, \\ 3x_1 + 4x_2 + 2x_3 + 2x_4 = 1\,. \end{cases}$$

1–5. The first five steps of Gauss-Jordan algorithm were performed during the solution of example 5.3,"d". The following formula for the nonhomogeneous system general solution was obtained:

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

Variables x_1, x_2 are basis; x_3, x_4 – free.

6. Assuming $x_3 = 0$, $x_4 = 0$, we obtain the particular solution of the nonhomogeneous system

$$x^{\text{p}} = \begin{pmatrix} 3 & -2 & 0 & 0 \end{pmatrix}^T.$$

7. Find the fundamental system of solutions of the corresponding homogeneous system (example 5.4):

$$\varphi_1 = \begin{pmatrix}-6 & 4 & 1 & 0\end{pmatrix}^T, \qquad \varphi_2 = \begin{pmatrix}-2 & 1 & 0 & 1\end{pmatrix}^T.$$

8. Write the nonhomogeneous system general solution by the formula (5.15):

$$x = x^{\text{p}} + C_1 \cdot \varphi_1 + C_2 \cdot \varphi_2 = \begin{pmatrix}3\\-2\\0\\0\end{pmatrix} + C_1 \cdot \begin{pmatrix}-6\\4\\1\\0\end{pmatrix} + C_2 \cdot \begin{pmatrix}-2\\1\\0\\1\end{pmatrix}.$$

The initial solution set structure is found.

EXERCISES

1. Solve the system using Gauss–Jordan algorithm:

$$\begin{cases} x_1 + 2x_2 + 2x_3 = m, \\ 2x_1 + 4x_2 + 3x_3 = n. \end{cases}$$

2. Find the fundamental system of solutions and write the general solution structure:

$$\begin{cases} x_1 + x_2 + nx_3 + mx_4 = 0, \\ 2x_1 + 3x_2 + x_3 + x_4 = 0. \end{cases}$$

CHAPTER 6. EIGENVECTORS AND EIGENVALUES OF MATRICES

6.1. BASIC DEFINITIONS AND PROPERTIES

Let A be a square matrix of order n. A nonzero column $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, that satisfies

$$A \cdot x = \lambda \cdot x, \tag{6.1}$$

is called an ***eigenvector*** of matrix A.

The number λ in equality (6.1) is called an ***eigenvalue*** of matrix A. We say that x is an eigenvector ***corresponding*** to the eigenvalue λ.

Let's set a problem of matrix eigenvalues and eigenvectors calculation. Definition (6.1) can be rewritten as

$$(A - \lambda E) \cdot x = o,$$

where E is an identity matrix of order n. Hence, condition (6.1) is a homogeneous system of n linear algebraic equations with n unknowns $x_1, x_2, \dots, x_n$:

$$\begin{cases} (a_{11} - \lambda)\, x_1 + a_{12}\, x_2 + \ldots + a_{1n}\, x_n = 0, \\ a_{21}\, x_1 + (a_{22} - \lambda)\, x_2 + \ldots + a_{2n}\, x_n = 0, \\ \qquad \ldots \\ a_{n1}\, x_1 + a_{n2}\, x_2 + \ldots + (a_{nn} - \lambda)\, x_n = 0. \end{cases} \tag{6.2}$$

Since we are only interested in nontrivial solutions $(x \neq o)$ of the homogeneous system, then the determinant of the system matrix must be equal to zero:

$$\det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0. \tag{6.3}$$

Otherwise by Cramer's rule the system has a unique trivial solution.

The problem of eigenvalues calculation is reduced to the solution of the equation

$$\det(A-\lambda E)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}=0,$$

which is called the ***characteristic equation*** of matrix A.

The roots of the characteristic equation (6.3) *are the only eigenvalues of a matrix.*

By the fundamental theorem of algebra in the general case the characteristic equation has n complex roots (counted with multiplicity). Any square matrix has eigenvalues and eigenvectors.

Eigenvalues of a matrix are uniquely determined (counted with multiplicity) and eigenvectors are ambiguously determined. A set of all eigenvalues of a matrix (counted with multiplicity) is called its ***spectrum***. The spectrum of a matrix is called ***simple***, if all its eigenvalues are pairwise different (all roots of the characteristic equation are simple).

6.2. PROPERTIES OF EIGENVECTORS AND EIGENVALUES

Let A be a square matrix of order n.

1. Eigenvectors corresponding to different eigenvalues are linearly independent.

2. A nonzero linear combination of eigenvectors corresponding to one eigenvalue, is an eigenvector corresponding to the same eigenvalue.

3. Let $(A-\lambda E)^{+}$ be the adjoint matrix to the characteristic matrix $(A-\lambda E)$. If λ_0 is an eigenvalue of matrix A, then any nonzero column of matrix $(A-\lambda_0 E)^{+}$ is an eigenvector corresponding to the eigenvalue λ_0.

4. To extract the maximum linearly independent subsystem from the set of eigenvectors, it is necessary for all distinct eigenvalues $\lambda_1,\ldots,\lambda_k$ write in sequence

systems of linearly independent eigenvectors, in particular write in sequence the fundamental systems of solutions of homogeneous systems

$$(A-\lambda_1 E)\cdot x = o, \qquad (A-\lambda_2 E)\cdot x = o, \ldots, \qquad (A-\lambda_k E)\cdot x = o.$$

Algorithm for calculation of eigenvectors and eigenvalues

To calculate eigenvectors and eigenvalues of a square matrix A of order n we should make the following steps:

1. Compose the characteristic equation of the matrix

$$\Delta_A(\lambda) = \det(A-\lambda E).$$

2. Find all distinct roots $\lambda_1, \ldots, \lambda_k$ of the characteristic equation $\Delta_A(\lambda)=0$; it is not necessary to find multiplicities $n_1, n_2, \ldots, n_k$ ($n_1+n_2+\ldots+n_k=n$) of the roots.

3. For the root $\lambda=\lambda_1$ find the fundamental system of solutions $\varphi_1, \varphi_2, \ldots, \varphi_{n-r}$ ($r=\operatorname{rg}(A-\lambda_1 E)$) of a homogeneous system of equations

$$(A-\lambda_1 E)\cdot x = o.$$

To do this we can either use the algorithm for solving a homogeneous system or one of the methods for finding a fundamental matrix.

4. Write down linearly independent eigenvectors of A that correspond to the eigenvalue λ_1:

$$s_1 = C_1\cdot\varphi_1, s_2 = C_2\cdot\varphi_2, \ldots, s_{n-r} = C_{n-r}\cdot\varphi_{n-r}, \tag{6.4}$$

where $C_1, C_2, \ldots, C_{n-r}$ are nonzero arbitrary constants. A set of all eigenvectors corresponding to the eigenvalue λ_1 consists of nonzero columns that have the form $s = C_1\cdot\varphi_1 + C_2\cdot\varphi_2 + \ldots + C_{n-r}\cdot\varphi_{n-r}$. Hereinafter we will denote eigenvectors of a matrix by the letter s.

Repeat steps 3 and 4 for other eigenvalues $\lambda_2, \ldots, \lambda_k$.

Example 6.1. Find eigenvalues and eigenvectors of matrices:

$$A=\begin{pmatrix}1 & -2\\ 3 & 8\end{pmatrix},\ B=\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix},\ C=\begin{pmatrix}0 & 2\\ -2 & 0\end{pmatrix}.$$

□ *Matrix* A:

1. Compose the characteristic polynomial of the matrix:

$$\Delta_A(\lambda)=\begin{vmatrix} 1-\lambda & -2 \\ 3 & 8-\lambda \end{vmatrix}=(1-\lambda)(8-\lambda)+6=\lambda^2-9\lambda+8+6=\lambda^2-9\lambda+14.$$

2. Solve the characteristic equation:

$$\lambda^2-9\lambda+14=0 \quad\Rightarrow\quad \lambda_1=2,\ \lambda_2=7 \text{ (simple spectrum)}.$$

3^1. For the simple root $\lambda_1=2$ compose a homogeneous system of equations $(A-\lambda_1 E)\cdot x=o$:

$$\begin{pmatrix} 1-2 & -2 \\ 3 & 8-2 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad\Leftrightarrow\quad \begin{pmatrix} -1 & -2 \\ 3 & 6 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solve this system using Gauss–Jordan algorithm, reducing the augmented coefficient matrix to the simplified form:

$$\left(\begin{array}{cc|c} -1 & -2 & 0 \\ 3 & 6 & 0 \end{array}\right)\sim\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & 6 & 0 \end{array}\right)\sim\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

The rank of the system matrix is equal to 1 ($r=1$), the number of unknowns is $n=2$, hence a fundamental system of solutions consists of one $(n-r=1)$ solution.

Denominate the basis variable x_1 by the free one: $x_1=-2x_2$. Suppose $x_2=1$, and obtain the solution $\varphi_1=\begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

4^1. Write eigenvectors corresponding to the eigenvalue $\lambda_1=2$: $s_1=C_1\cdot\varphi_1$, where C_1 is a nonzero arbitrary constant: $s_1=C\cdot\varphi_1=C_1\cdot\begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

3^2. For the simple root $\lambda_2=7$ compose a homogeneous system of equations $(A-\lambda_2 E)\cdot x=o$:

$$\begin{pmatrix} 1-7 & -2 \\ 3 & 8-7 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad\Leftrightarrow\quad \begin{pmatrix} -6 & -2 \\ 3 & 1 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solve this system using Gauss–Jordan algorithm, reducing the augmented coefficient matrix to the simplified form:

$$\begin{pmatrix} -6 & -2 & | & 0 \\ 3 & 1 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 3 & 1 & | & 0 \\ -6 & -2 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \frac{1}{3} & | & 0 \\ -6 & -2 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & \frac{1}{3} & | & 0 \\ 0 & 0 & | & 0 \end{pmatrix}.$$

The rank of the system matrix is equal to 1 ($r=1$), the number of unknowns is $n=2$, hence a fundamental system of solutions consists of one $(n-r=1)$ solution.

Denominate the basis variable x_1 by the free one: $x_1 = -\dfrac{1}{3}x_2$. Suppose $x_2 = 1$, we have the solution $\varphi_2 = \begin{pmatrix} -\frac{1}{3} \\ 1 \end{pmatrix}$.

4^2. Write eigenvectors corresponding to the eigenvalue $\lambda_2 = 7$: $s_2 = C_2 \cdot \varphi_2$, where C_2 is a nonzero arbitrary constant: $s_2 = C_2 \cdot \varphi_2 = C_2 \cdot \begin{pmatrix} -\frac{1}{3} \\ 1 \end{pmatrix}$.

Matrix B:

1. Compose the characteristic polynomial of the matrix:

$$\Delta_B(\lambda) = |\, B - \lambda E \,| = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^3 + 2 - 3(1-\lambda) = -\lambda^3 + 3\lambda^2.$$

2. Solve the characteristic equation: $-\lambda^3 + 3\lambda^2 = 0 \quad \Rightarrow \quad \lambda_1 = 3,\ \lambda_2 = \lambda_3 = 0$ (spectrum).

3^1. For the simple root $\lambda_1 = 3$ compose a homogeneous system of equations $(B - \lambda_1 E) \cdot x = o$:

$$\begin{pmatrix} 1-3 & 1 & 1 \\ 1 & 1-3 & 1 \\ 1 & 1 & 1-3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{cases} -2x_1 + x_2 + x_3 = 0, \\ x_1 - 2x_2 + x_3 = 0, \\ x_1 + x_2 - 2x_3 = 0. \end{cases}$$

Solve this system using Gauss–Jordan algorithm, reducing the augmented coefficient matrix to the simplified form (pivot elements are in bold italics):

$$(B - \lambda_1 E \mid o) = \begin{pmatrix} -2 & 1 & 1 & | & 0 \\ 1 & -2 & 1 & | & 0 \\ \mathbf{1} & 1 & -2 & | & 0 \end{pmatrix} \sim \begin{pmatrix} \mathbf{1} & 1 & -2 & | & 0 \\ -2 & 1 & 1 & | & 0 \\ 1 & -2 & 1 & | & 0 \end{pmatrix} \sim$$

$$\sim\left(\begin{array}{ccc|c}1 & 1 & -2 & 0\\ 0 & \mathbf{3} & -3 & 0\\ 0 & -3 & 3 & 0\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 1 & -2 & 0\\ 0 & \mathbf{1} & -1 & 0\\ 0 & 0 & 0 & 0\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 0 & -1 & 0\\ 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 0\end{array}\right).$$

The rank of the system matrix is equal to 2 ($r=2$), the number of unknowns is $n=3$, hence a fundamental system of solutions consists of one $(n-r=1)$ solution.

Denominate basis variables x_1, x_2 by the free one x_3: $\begin{cases} x_1 = x_3, \\ x_2 = x_3, \end{cases}$

and, suppose $x_3 = 1$, we obtain the solution $\varphi = \begin{pmatrix}1\\1\\1\end{pmatrix}$.

4[1]. Calculate all eigenvectors corresponding to the eigenvalue $\lambda_1 = 3$ by formula $s = C_1 \cdot \varphi$, where C_1 is a nonzero arbitrary constant.

3[2]. For the double root $\lambda_2 = \lambda_3 = 0$ we have a homogeneous system $B \cdot x = o$. Solve it by using Jauss–Jordan algorithm:

$$(B \mid o) = \left(\begin{array}{ccc|c}1 & 1 & 1 & 0\\ 1 & 1 & 1 & 0\\ 1 & 1 & 1 & 0\end{array}\right)\sim\left(\begin{array}{ccc|c}1 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right).$$

The rank of the system matrix is equal to 1 ($r=1$), hence a fundamental system of solutions consists of two ($n-r=2$) solutions.

Denominate the basis variable x_1 by the free ones $x_1 = -x_2 - x_3$. Assuming standard value sets of free variables $x_2 = 1, x_3 = 0$ and $x_2 = 0$, $x_3 = 1$ we obtain two solutions: $\varphi_1 = \begin{pmatrix}-1\\1\\0\end{pmatrix}$, $\varphi_2 = \begin{pmatrix}-1\\0\\1\end{pmatrix}$.

4[2]. Write a set of eigenvectors corresponding to the eigenvalue $\lambda_2 = 0$: $s = C_1 \cdot \varphi_1 + C_2 \cdot \varphi_2$, where C_1, C_2 are arbitrary constant, not equal to zero at the same time.

In particular, for $C_1 = 0$, $C_2 = -1$ we have $s_1 = (1 \quad 0 \quad -1)^T$; for $C_1 = -1$, $C_2 = 0$ we have $s_2 = (1 \quad -1 \quad 0)^T$. Adding the eigenvector $s_3 = (1 \quad 1 \quad 1)^T$ corresponding to the eigenvalue $\lambda_1 = 3$ (see step 4^1 for $C_1 = 1$), to these eigenvectors, we find three linearly independent eigenvectors of matrix B:

$$s_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Matrix C:

1. Compose the characteristic polynomial of the matrix:

$$\Delta_C(\lambda) = |C - \lambda E| = \begin{vmatrix} -\lambda & 2 \\ -2 & -\lambda \end{vmatrix} = (-\lambda)(-\lambda) + 4 = \lambda^2 + 4.$$

2. Solve the characteristic equation:

$$\lambda^2 + 4 = 0 \quad \Rightarrow \quad \lambda_1 = 2i, \ \lambda_2 = -2i \text{ (simple spectrum)}.$$

3^1. For the simple root $\lambda_1 = 2i$ compose a homogeneous system of equations $(C - \lambda_1 E) \cdot x = o$:

$$\begin{pmatrix} -2i & 2 \\ -2 & -2i \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solve this system using Gauss-Jordan algorithm, reducing the augmented coefficient matrix to the simplified form:

$$\left(\begin{array}{cc|c} -2i & 2 & 0 \\ -2 & -2i & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & i & 0 \\ -2 & -2i & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & i & 0 \\ 0 & 0 & 0 \end{array}\right).$$

The rank of the system matrix is equal to 1 ($r = 1$), the number of unknowns is $n = 2$, hence a fundamental system of solutions consists of one $(n - r = 1)$ solution.

Denominate basis variable x_1 by the free one: $x_1 = -i\,x_2$. Suppose $x_2 = 1$, we obtain the solution $\varphi_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$.

4^1. Write eigenvectors corresponding to the eigenvalue $\lambda_1 = 2i$: $s_1 = C_1 \cdot \varphi_1$, where C_1 is an arbitrary nonzero complex number.

3^2. For the simple root $\lambda_2 = -2i$ compose a homogeneous system of equations $(C - \lambda_2 E)\cdot x = o$:

$$\begin{pmatrix} 2i & 2 \\ -2 & 2i \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solve this system using Gauss–Jordan algorithm, reducing the augmented coefficient matrix to the simplified form:

$$\left(\begin{array}{cc|c} 2i & 2 & 0 \\ -2 & 2i & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & -i & 0 \\ -2 & 2i & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & -i & 0 \\ 0 & 0 & 0 \end{array}\right).$$

The rank of the system matrix is equal to 1 ($r = 1$), the number of unknowns is $n = 2$, hence a fundamental system of solutions consists of one ($n - r = 1$) solution.

Denominate the basis variable x_1 by the free one: $x_1 = i x_2$. Suppose $x_2 = 1$, we have the solution $\varphi_2 = \begin{pmatrix} i \\ 1 \end{pmatrix}$.

4^2. Write eigenvectors corresponding to the eigenvalue $\lambda_2 = -2i$: $s_2 = C_2 \cdot \varphi_2$, where C_2 is an arbitrary nonzero complex number.

EXERCISES

Find the eigenvalues and corresponding eigenvectors of matrices:

a) $\begin{pmatrix} m & m \\ n & n \end{pmatrix}$; b) $\begin{pmatrix} n & n & n \\ m & m & m \\ 1 & 1 & 1 \end{pmatrix}$.

CHAPTER 7. QUADRATIC FORMS

7.1. DEFINITION

A ***quadratic form*** in variables $x_1,\dots,x_n$ is an expression given by

$$q(x)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\,, \tag{7.1}$$

where coefficients a_{ij}, not all equal to zero, satisfy the ***symmetry conditions*** $a_{ij}=a_{ji}$, $i=1,\dots,n$, $j=1,\dots,n$. We put this restriction without loss of generality, since a sum of two similar terms $a_{ij}x_ix_j+a_{ji}x_jx_i$ with unequal coefficients $a_{ij}\neq a_{ji}$ (for $i\neq j$) can always be replaced by a sum $a'_{ij}x_ix_j+a'_{ji}x_jx_i$ with equal coefficients, setting

$$a'_{ij}=a'_{ji}=\frac{a_{ij}+a_{ji}}{2}.$$

Let's consider ***real quadratic forms***, coefficients of which are real numbers and variables take real values.

Combining terms the quadratic form (7.1) can be rewritten as

$$q(x)=a_{11}x_1^2+2a_{12}x_1x_2+\dots+2a_{1n}x_1x_n+a_{22}x_2^2+2a_{23}x_2x_3+\dots+a_{nn}x_n^2. \tag{7.2}$$

This is a quadratic form with ***combined terms***.

A symmetric matrix $A=(a_{ij})$, made up from coefficients of the quadratic form (7.1) is called a ***matrix of the quadratic form***. The determinant of this matrix is called the ***discriminant***, and its rank is called the ***rank of a quadratic form***. A quadratic form is called ***singular*** if its matrix is singular ($\mathrm{rg}\,A<n$), otherwise if the matrix is nonsingular ($\mathrm{rg}\,A=n$), a quadratic form is called ***nonsingular***.

Composing a column matrix of variables $x=(x_1\ \cdots\ x_n)^T$ a quadratic form can be written in ***matrix*** form:

$$q(x)=x^T\cdot A\cdot x. \tag{7.3}$$

An important example of a quadratic form is the ***second differential of function*** $f(x)$ with vector argument $x=(x_1\ \cdots\ x_n)^T$:

$$d^2 f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} dx_i dx_j = dx^T \cdot \frac{d^2 f(x)}{dx^T dx} \cdot dx, \qquad (7.4)$$

where differentials $dx_1,\dots,dx_n$ are variables of the quadratic form; a matrix of second-order partial derivatives (Hessian matrix)

$$\frac{d^2 f(x)}{dx^T dx} = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}, \qquad (7.5)$$

computed for a fixed value of an argument, is a matrix of the quadratic form, and the differential of vector argument $dx = (dx_1 \ \cdots \ dx_n)^T$ is a column of its variables.

Example 7.1. For the function $f(x) = 2x_1^2 + x_1x_2 + x_2^2$ write the second differential $d^2 f(x)$ in matrix form (7.4).

□ The given function $f(x) = f(x_1,x_2)$ has two arguments x_1, x_2. Compose the matrix of second order differentials, i.e. the Hessian. First find first-order partial derivatives:

$$\frac{\partial f(x)}{\partial x_1} = 4x_1 + x_2; \qquad \frac{\partial f(x)}{\partial x_2} = x_1 + 2x_2,$$

and then – second-order partial derivatives:

$$\frac{\partial^2 f(x)}{\partial x_1^2} = 4; \qquad \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} = 1; \qquad \frac{\partial^2 f(x)}{\partial x_2^2} = 2.$$

Then by formula (7.5) compose the Hessian matrix $\left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right) = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$.

The second differential of function $f(x) = f(x_1,x_2)$ is a quadratic form (7.4) of differentials dx_1, dx_2: $d^2 f(x) = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} dx_i dx_j = 4dx_1^2 + 2dx_1dx_2 + 2dx_2^2 =$

$$= \begin{pmatrix} dx_1 & dx_2 \end{pmatrix} \cdot \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} = dx^T \cdot \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix} \cdot dx \ . \ ■$$

Any quadratic form $q(x) = x^T A\, x$ can be transformed to the ***canonical form***

$$q(Sy) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2$$

via linear non-degenerate change of variables $x = Sy$ $(\det S \neq 0)$ where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of matrix A.

7.2 DEFINITE AND INDEFINITE QUADRATIC FORMS

A real quadratic form $q(x) = x^T \cdot A \cdot x$ is called ***positive*** **(*negative*)** ***definite*** if $q(x) > 0$ $(q(x) < 0)$ for any $x \neq o$. Positive and negative definite forms are called ***definite.***

If a quadratic form takes on both positive and negative values, it is called ***indefinite***. Definite and indefinite quadratic forms are denoted by $q(x) > 0$, $q(x) < 0$, $q(x) \gtrless 0$, respectively.

A minor of A of order k is ***principal*** if it is obtained by deleting $(n - k)$ rows and the $(n - k)$ columns with the same numbers. The ***leading principal minor*** of A of order k is the minor of order k obtained by deleting the last $(n - k)$ rows and columns.

Sylvester's criterion. *A quadratic form $q(x) = x^T \cdot A \cdot x$ is positive definite if and only if all leading principal minors of its matrix are positive*:

$$\Delta_1 = a_{11} > 0, \ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \ \Delta_n = \det A > 0 .$$

A quadratic form is negative definite if and only if leading principal minors of its matrix change signs starting from the negative one:

$$\Delta_1 = a_{11} < 0, \ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \ (-1)^n \Delta_n = (-1)^n \det A > 0 .$$

A quadratic form is indefinite if at least one principal minor of even order is negative, or two principal minors of uneven order have different signs (*sufficient **criterion for indefiniteness of a quadratic form***).

Criterion for definiteness and indefiniteness of quadratic form by eigenvalues of its matrix

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be eigenvalues of matrix A, corresponding to a quadratic form $q(x) = x^T \cdot A \cdot x$. Eigenvalues of a real symmetric matrix are real.

1) A quadratic form $q(x) = x^T \cdot A \cdot x$ is positive definite if and only if all eigenvalues of its matrix are positive: $\lambda_1 > 0$, $\lambda_2 > 0, \ldots,$ $\lambda_n > 0$.

2) A quadratic form $q(x) = x^T \cdot A \cdot x$ is negative definite if and only if all eigenvalues of its matrix are negative: $\lambda_1 < 0$, $\lambda_2 < 0, \ldots,$ $\lambda_n < 0$.

3) A quadratic form $q(x) = x^T \cdot A \cdot x$ is indefinite if and only if its matrix has both positive and negative eigenvalues, i.e. $\lambda_i \cdot \lambda_j < 0$ for at least one pair of eigenvalues ($i \neq j$, $1 \leq i \leq n$, $1 \leq j \leq n$).

Example 7.2. Determine whether quadratic forms of the given matrices are positive definite, negative definite or indefinite

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}; \quad B = \begin{pmatrix} -2 & 2 \\ 2 & -5 \end{pmatrix}; \quad C = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix}.$$

□ The quadratic form $q(x) = x^T \cdot A \cdot x = x_1^2 + 2x_1x_2 + 2x_2^2$ is positive definite, since all leading principal minors of its matrix A are positive: $\Delta_1 = 1 > 0$, $\Delta_2 = 1 > 0$ (see Sylvester's criterion).

The quadratic form $q(x) = x^T \cdot B \cdot x = -2x_1^2 + 4x_1x_2 - 5x_2^2$ is negative definite, since leading principal minors of its matrix B change signs, starting with the negative one: $\Delta_1 = -2 < 0$, $\Delta_2 = 6 > 0$ (see Sylvester's criterion). Let's check this conclusion by examining eigenvalues of matrix B:

$$\det(B - \lambda E) = \begin{vmatrix} -2-\lambda & 2 \\ 2 & -5-\lambda \end{vmatrix} = 0 \quad \Leftrightarrow \quad \lambda^2 + 7\lambda + 6 = 0.$$

Hence, $\lambda_1 = -6$, $\lambda_2 = -1$. Since both eigenvalues are negative, the quadratic form is negative definite. Thus, the quadratic form is definite.

The quadratic form $q(x) = x^T \cdot C \cdot x = -x_1^2 + 2x_1x_2 + 2x_1x_3 - x_2^2 + 2x_2x_3 - 2x_3^2$ is neither positive nor negative definite, since its leading principal minors do not meet Sylvester's criterion: $\Delta_1 = -1 < 0$, $\Delta_2 = 0$, $\Delta_3 = 4 > 0$ (conditions (7.12) and (7.13) are not fulfilled). Let's calculate principal minors of this matrix:

$$M_1^1 = \Delta_1 = -1, \quad M_2^2 = -1, \quad M_3^3 = -2;$$

$$M_{12}^{12} = \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} = 0, \quad M_{13}^{13} = \begin{vmatrix} -1 & 1 \\ 1 & -2 \end{vmatrix} = 1, \quad M_{23}^{23} = \begin{vmatrix} -1 & 1 \\ 1 & -2 \end{vmatrix} = 1;$$

$$M_{123}^{123} = \Delta_3 = \det C = 4.$$

There are minors of uneven order that have different signs, e.g. $M_2^2 = -1 < 0$, $M_{123}^{123} = 4 > 0$. Hence, the quadratic form is indefinite (see criterion for indefiniteness). Let's check this conclusion by examining eigenvalues of matrix C:

$$\det(C - \lambda E) = \begin{vmatrix} -1-\lambda & 1 & 1 \\ 1 & -1-\lambda & 1 \\ 1 & 1 & -2-\lambda \end{vmatrix} = 0 \quad \Leftrightarrow \quad \lambda^3 + 4\lambda^2 + 2\lambda - 4 = 0.$$

Hence, $\lambda_1 = -2$, $\lambda_2 = -1 - \sqrt{3}$, $\lambda_3 = -1 + \sqrt{3}$. Since the matrix has both positive and negative eigenvalues ($\lambda_1 \cdot \lambda_3 = -2 \cdot (-1 + \sqrt{3}) < 0$), the quadratic form $q(x) = x^T \cdot C \cdot x$ is indefinite. ■

EXERCISES

Determine whether quadratic forms are positive definite, negative definite or indefinite (by Sylvester's criterion) and calculate norms of corresponding matrices:

a) $f(x) = m x_1^2 - n x_1 + n x_2^2 + m x_2 + x_1 x_2 + 2$;

b) $f(x) = -x_1^2 + 2n x_1 - 4x_2^2 - 4m x_2 - n^2 - m^2$.

PART II. ANALYTIC GEOMETRY

CHAPTER 8. VECTOR ALGEBRA

8.1. VECTORS AND VECTOR LINEAR OPERATIONS

8.1.1. Vector, Its Direction and Length

A ***vector*** is an ordered pair of points. The first point is called ***vector tail***, the second – ***vector head***. A distance between head and tail is called ***length***.

A vector with coincident tail and head is called ***zero vector***, its length equals to zero. If vector length is a positive value, then it is called ***nonzero vector***.

A nonzero vector can be defined as a ***directed segment***. One of its bounding points is considered as the first (vector tail), and another – as the second (vector head). Zero vector direction is, obviously, not determined.

A vector with the beginning in point A and ending in point B is denoted by $\overline{AB}$ and depicted by arrow, directed to vector head (Fig. 8.1). Vector tail is also called ***point of application***. Vector $\overline{AB}$ ***is applied to point*** A. The length of vector $\overline{AB}$ equals to the length of segment AB and it is denoted by $\left|\overline{AB}\right|$. With this notation, vector length is also called ***magnitude***, ***absolute value***.

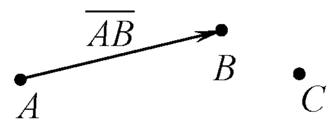


Figure 8.1

A zero vector, e.g. $\overline{CC}$, is denoted by symbol $\bar{o}$ and depicted by point (point C on Fig. 8.1).

A vector, which length equals to unit or assumed as a unit, is called ***unit vector***.

Nonzero vector $\overline{AB}$, beside directed segment, defines ***ray*** AB (with the beginning in point A) and ***line*** AB.

Two nonzero vectors are called ***collinear***, if they belong to one line or to two parallel lines, otherwise they are called ***noncollinear***. Vector collinearity is denoted by symbol $\|$. A zero vector is considered as collinear to any vector, because its direction is not defined. Any vector is collinear to itself.

Equally and oppositely directed nonzero collinear vectors are denoted by $\uparrow\uparrow$ and $\uparrow\downarrow$ accordingly.

Three nonzero vectors are called ***coplanar***, if they lie in the same or parallel planes, otherwise they are called ***noncoplanar***. Zero vector is coplanar to any other two vectors, because its direction is not defined.

Two vectors are ***equal***, if they:

a) are collinear and equally directed;

b) have equal lengths.

All zero vectors are equal to each other.

This definition of equality characterizes so-called ***free vectors***. Given free vector can be moved without change of its direction and length to any point of space (apply it to any point). As the result we will obtain vectors, which are equal to the given one.

It is possible to give equivalent definitions of collinearity and coplanarity.

Two nonzero vectors are called ***collinear***, if they lie on one line after application to the same point.

Three nonzero vectors are called ***coplanar***, if the lie in one plane after application to the same point.

Angle between nonzero vectors is an angle (not greater than π) between vectors, which are equal to them and applied to the same point.

Consider two nonzero vectors $\bar{a}$ and $\bar{b}$ (Fig.8.2). Construct equal vectors $\overline{OA}$ and $\overline{OB}$. In the plane, which contains rays OA and OB, we will obtain two angles AOB. The smaller one, which value φ is not greater than π ($0 \le \varphi \le \pi$), is taken as an angle between $\bar{a}$ and $\bar{b}$.

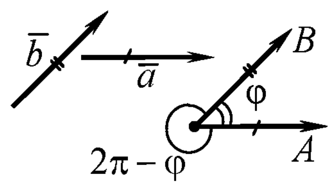


Figure 8.2

It is not possible to define angle between two vectors if at least one of them is zero, because zero vector direction is not defined. From the definitions it follows that angle between nonzero collinear vectors equals to zero (if vectors are equally directed) of equals to π (if they are oppositely directed).

Example 8.1. Consider triangle ABC; points L, M, N are midpoints of its sides. For vectors in Fig. 8.3, determine which of them are collinear, equally directed, oppositely directed and equal. Show angles between vectors $\overline{AM}$ and $\overline{AN}$, $\overline{MC}$ and $\overline{CL}$, $\overline{AM}$ and $\overline{MC}$, $\overline{CL}$ and $\overline{BL}$.

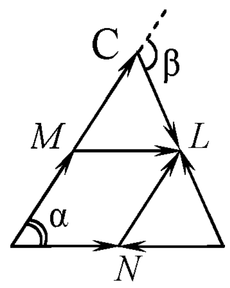


Figure 8.3

□ By the triangle mid-segment theorem we conclude that $ML \parallel AB$, $LN \parallel AC$. Thus vectors $\overline{AM}$, $\overline{MC}$, $\overline{NL}$ are collinear (because they lie on one or parallel lines), equally directed and have the same length, hence, they are equal: $\overline{AM} = \overline{MC} = \overline{NL}$. Similarly, $\overline{AN} = \overline{ML}$, $\overline{AN} \uparrow\downarrow \overline{BN}$, $\overline{BN} \uparrow\downarrow \overline{ML}$, $\overline{CL} \uparrow\downarrow \overline{BL}$. Vectors $\overline{AM}$ and $\overline{AN}$ form angle α, vectors $\overline{MC}$ and $\overline{CL}$ – angle β. The angle between vectors $\overline{AM}$ and $\overline{MC}$ equals to zero, because they are equally directed, and the angle between $\overline{CL}$ and $\overline{BL}$ equals to π, because they are oppositely directed. ■

8.1.2. Linear Operations on Vectors

A ***sum*** of two vectors $\overline{a}$ and $\overline{b}$ is a vector $\overline{OB} = \overline{a} + \overline{b}$ (Fig. 8.4, *a*), which tail coincides with the tail of vector $\overline{OA} = \overline{a}$, and head – with the head of vector $\overline{AB} = \overline{b}$ (*triangle rule*).

A ***product of nonzero vector*** $\overline{a}$ ***and real number*** λ $(\lambda \neq 0)$ is a vector $\lambda \cdot \overline{a}$, which satisfies the following conditions:

a) the length of vector $\lambda \cdot \overline{a}$ equals to $|\lambda| \cdot |\overline{a}|$, i.e. $|\lambda \overline{a}| = |\lambda| \cdot |\overline{a}|$;

b) vectors $\lambda \cdot \overline{a}$ and $\overline{a}$ are collinear ($\lambda \cdot \overline{a} \,||\, \overline{a}$);

c) vectors $\lambda \cdot \overline{a}$ and $\overline{a}$ are equally directed, if $\lambda > 0$, and oppositely directed, if $\lambda < 0$ (Fig. 8.4, *b*).

A product of a zero vector and any arbitrary number λ is a zero vector (by definitions): $\lambda \cdot \overline{o} = \overline{o}$; a product of any vector and zero is also a zero vector: $0 \cdot \overline{a} = \overline{o}$.

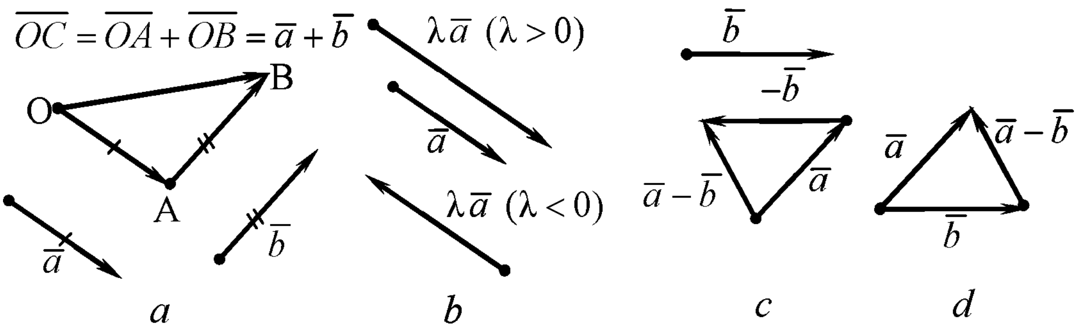


Figure 8.4

A vector $(-\overline{a})$ is called ***opposite*** to vector $\overline{a}$, if their sum equals to zero vector: $\overline{a} + (-\overline{a}) = \overline{o}$. The opposite vector $(-\overline{a})$ has length $|\overline{a}|$, and is collinear and oppositely directed to vector $\overline{a}$. A zero vector is opposite to itself. Note that $(-\overline{a}) = (-1) \cdot \overline{a}$.

The ***difference*** between vectors $\overline{a}$ and $\overline{b}$ is the sum of vector $\overline{a}$ and vector $(-\overline{b})$ opposite to vector $\overline{b}$: $\overline{a} - \overline{b} = \overline{a} + (-\overline{b})$ (Fig. 8.4, *c*). In other words, the difference $\overline{a} - \overline{b}$ of vectors $\overline{a}$ and $\overline{b}$ – is a vector, which sum with $\overline{b}$ gives vector $\overline{a}$ (Fig. 8.4, *d*).

Addition and multiplication by number operations are called ***linear operations on vectors***.

Vector $\overline{a}$ is called a ***linear combination*** of vectors $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_k$, if it can be expressed in the following form

$$\overline{a} = \alpha_1 \overline{a}_1 + \alpha_2 \overline{a}_2 + \ldots + \alpha_k \overline{a}_k,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ – are some numbers. In this case it is said, that ***vector*** $\overline{a}$ ***is decomposed by vectors*** $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_k$, numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$ are called ***decomposition coefficients***.

To find a sum of several vectors you should construct a polyline from vectors, which are equal to the given ones, by applying the second vector to the first vector head, the third vector to the second vector head and so on. Then the ***locking*** vector, which connects the first vector tail with the last vector head, equals to the sum of all vectors of polyline (*polyline rule*).

Example 8.2. For vectors on Fig. 8.3 find the following sums and differences: $\overline{BN} + \overline{AM}$; $\overline{AM} - \overline{BL}$; $\overline{AN} + \overline{AM}$; $\overline{BN} + \overline{AM} + \overline{CL}$. Decompose vector $\overline{AC}$ by vectors $\overline{BN}$ and $\overline{BL}$.

□ Taking into account that $\overline{AM} = \overline{NL}$, by the triangle rule we obtain $\overline{BN} + \overline{AM} = \overline{BN} + \overline{NL} = \overline{BL}$.

Since $\overline{BL} = -\overline{CL}$ and $\overline{AM} = \overline{MC}$, then $\overline{AM} - \overline{BL} = \overline{MC} + \overline{CL} = \overline{ML}$.

Since $\overline{ML} = \overline{AN}$, then by the triangle rule $\overline{AN} + \overline{AM} = \overline{AM} + \overline{ML} = \overline{AL}$.

Since $\overline{BN} + \overline{AM} = \overline{BL}$ and $\overline{CL} = -\overline{BL}$, we obtain

$$\overline{BN} + \overline{AM} + \overline{CL} = \underbrace{\left(\overline{BN} + \overline{AM}\right)}_{\overline{BL}} + \underbrace{\overline{CL}}_{-\overline{BL}} = \overline{BL} - \overline{BL} = \overline{o}.$$

Since $\overline{BA} + \overline{AC} = \overline{BC}$, $\overline{BA} = 2 \cdot \overline{BN}$, $\overline{BC} = 2 \cdot \overline{BL}$, then $\overline{AC} = -2 \cdot \overline{BN} + 2 \cdot \overline{BL}$. ■

8.2. ORTHOGONAL PROJECTIONS OF VECTORS

An ***orthogonal*** (***direct***) ***projection of point*** A ***to line*** l is a foot of perpendicular A_l, constructed from point A to line l (Fig. 8.5, *a*). An ***orthogonal*** (***direct***) ***projection of point*** A ***to plane*** π is a foot of perpendicular A_π, constructed from point A to plane π (Fig. 8.5, *b*).

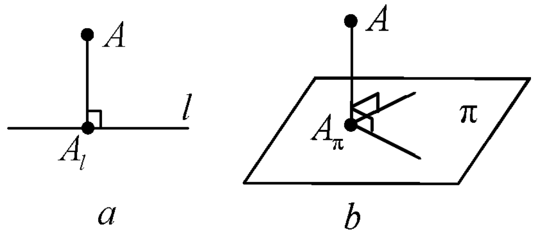


Figure 8.5

An ***orthogonal projection of vector*** $\overline{a} = \overline{AB}$ ***to line*** l is vector $\overline{a}_l = \overline{A_l B_l}$, which tail is the orthogonal projection A_l of point A and head is the orthogonal projection B_l of point B (Fig. 8.6, *a* – plane case, Fig. 8.6, *b* – space case). An orthogonal projection of vector $\overline{a}$ to line l will be denoted by $\overline{proj}\,\overline{a}$.

An ***orthogonal projection of vector*** $\overline{a}$ ***to axis, formed by vector*** $\overline{e} \neq \overline{o}$, is its orthogonal projection to line, which contains vector $\overline{e}$. This projection will be denoted by $\overline{proj_{\overline{e}}}\overline{a}$.

An ***orthogonal projection of vector*** $\overline{a} = \overline{AB}$ ***to plane*** π is vector $\overline{a}_\pi = \overline{A_\pi B_\pi}$, which tail is the orthogonal projection A_π of point A to plane π and head is the orthogonal projection B_π of point B (Fig. 8.6, *c*). An orthogonal projection of vector $\overline{a}$ to plane π will be denoted by $\overline{proj_\pi}\,\overline{a}$.

The difference between vectors $\overline{a}$ and its orthogonal projection is the ***orthogonal component of vector*** $\overline{a}$ ***relative to line*** ($\overline{a}_{\perp\overline{e}} = \overline{a}_{\perp l}$ on Fig. 8.6, *a*) or ***plane*** ($\overline{a}_{\perp\pi}$ on Fig. 8.6, *c*).

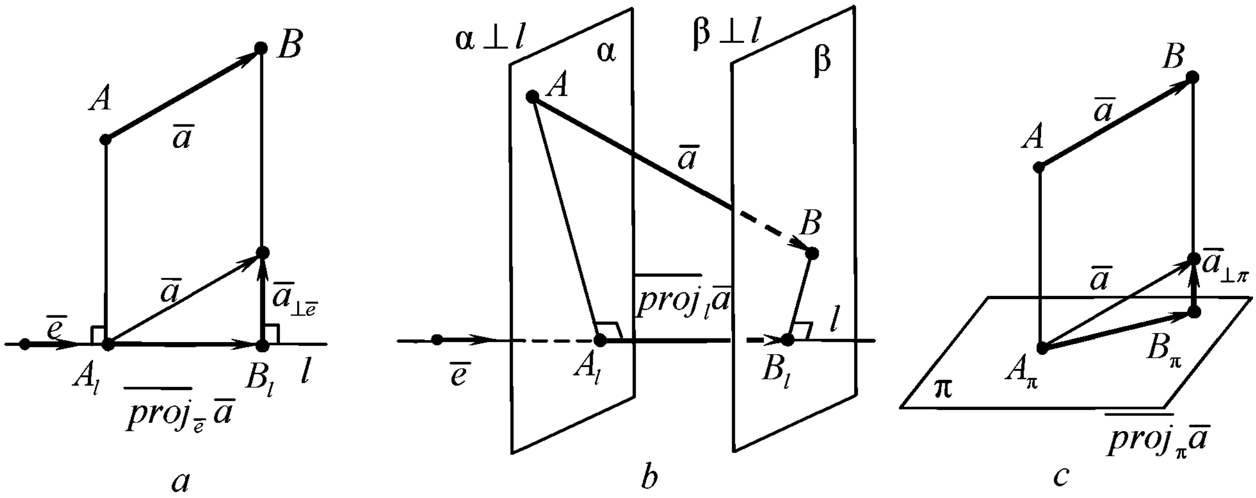


Figure 8.6

Algebraic value of projection length

Let φ be an ***angle between nonzero vector $\overline{a}$ and axis***, formed by vector $\overline{e} \neq \overline{o}$, i.e. angle between nonzero vectors $\overline{a}$ and $\overline{e}$.

The ***algebraic value of length of vector $\overline{a}$ orthogonal projection to axis, formed by vector*** $\overline{e} \neq \overline{o}$ is the length of its orthogonal projection $\overline{proj_{\overline{e}}\overline{a}}$, taken with positive sign if angle φ is not greater than $\frac{\pi}{2}$, and with negative sign if angle φ is greater than $\frac{\pi}{2}$ (Fig. 8.7).

Properties of projection length algebraic values:

- *Algebraic value of projection length of vector sum equals to the sum of summands algebraic values of orthogonal projection lengths.*

- *Algebraic value of orthogonal projection length of vector and number product equals to the product of this number and algebraic value of this vector orthogonal projection length.*

$$proj_{\bar e}\bar a = \left|\overline{proj_{\bar e}\bar a}\right| = \left|\bar a\right|\cdot\cos\varphi \qquad proj_{\bar e}\bar b = -\left|\overline{proj_{\bar e}\bar b}\right| = \left|\bar b\right|\cdot\cos\psi$$

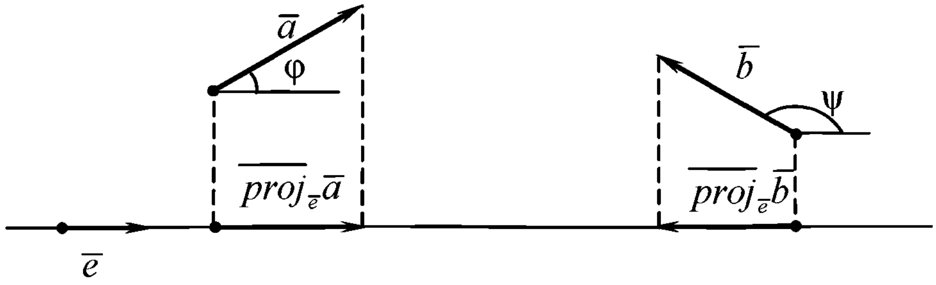


Figure 8.7

Example 8.3. Bases AB and CD of equal-sided trapezium $ABCD$ are equal to a and b accordingly; point M is the middle point of BC (Fig. 8.8).

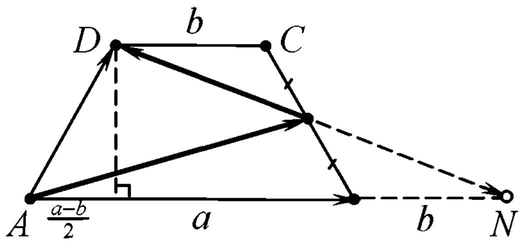


Figure 8.8

Find algebraic values of orthogonal projection lengths of vectors $\overline{AM}$ and $\overline{MD}$ to axis, formed by vector $\overline{AB}$.

□ Let DL be trapezium height, N –intersection point of lines AB и DM.

By the property of equal-sided trapezium: $AL = \dfrac{a-b}{2}$; from the equality of triangles CDM and BNM: $BN = CD = b$. Denote required algebraic values of orthogonal projection lengths by $x = proj_{\overline{AB}}\overline{AM}$, $y = proj_{\overline{AB}}\overline{MD}$. From the equalities $\overline{AM}+\overline{MD}=\overline{AD}$, $\overline{AM}-\overline{MD}=\overline{AM}+\overline{MN}=\overline{AN}$ and Property 1 we have:

$$proj_{\overline{AB}}\left(\overline{AM}+\overline{MD}\right) = proj_{\overline{AB}}\overline{AM} + proj_{\overline{AB}}\overline{MD} = proj_{\overline{AB}}\overline{AD}, \text{ i.e. } x+y=\frac{a-b}{2};$$

$$proj_{\overline{AB}}\left(\overline{AM}-\overline{MD}\right) = proj_{\overline{AB}}\overline{AM} - proj_{\overline{AB}}\overline{MD} = proj_{\overline{AB}}\overline{AN}, \text{ i.e. } x-y=a+b.$$

Solving the system $\begin{cases} x+y=\dfrac{a-b}{2}, \\ x-y=a+b, \end{cases}$ we obtain $\begin{cases} x=\dfrac{3a+b}{4}, \\ y=-\dfrac{a+3b}{4}, \end{cases}$ i.e.

$proj_{\overline{AB}}\overline{AM}=\dfrac{3a+b}{4}$, $proj_{\overline{AB}}\overline{MD}=-\dfrac{a+3b}{4}$. ■

8.3. BASIS AND VECTOR COORDINATES

8.3.1. Basis on Line. Vector Coordinate on Line

A ***basis on line*** is any nonlinear vector $\overline{e}$ on this line (Fig. 8.9). This vector $\overline{e}$ is called ***basis***.

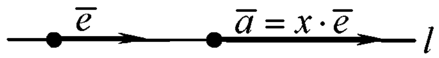


Figure 8.9

Theorem of vector decomposition on line. *Any vector $\overline{a}$, which is collinear to the line, can be decomposed by basis $\overline{e}$ on this line, i.e. represented in a form $\overline{a}=x\cdot\overline{e}$, where number x is uniquely defined.*

The coefficient x in decomposition is called ***vector coordinate $\overline{a}$ relative to basis $\overline{e}$***. All nonzero vectors equally directed with vector $\overline{e}$ have positive coordinates, and oppositely directed – negative. A zero vector coordinate equals to zero.

Example 8.4. Given vectors $\overline{a}=-2\cdot\overline{e}$ and $\overline{b}=4\cdot\overline{e}$, parallel to axes, formed by vector $\overline{e}\neq\overline{o}$. Find coordinates of vectors $\overline{a}+\overline{b}$; $-\overline{b}$; $\overline{a}-\overline{b}$; $3\cdot\overline{a}+2\cdot\overline{b}$ relative to basis $\overline{e}$, and coordinate of vector $\overline{a}+\overline{b}$ relative to basis $\overline{b}$.

□ By the property of collinear vectors we find decompositions by basis $\overline{e}$:

$$\overline{a}+\overline{b}=-2\cdot\overline{e}+4\cdot\overline{e}=(-2+4)\cdot\overline{e}=2\cdot\overline{e};$$

$$-\overline{b}=(-1)\cdot\overline{b}=(-1)\cdot4\cdot\overline{e}=-4\cdot\overline{e};$$

$$\overline{a}-\overline{b}=-2\cdot\overline{e}-4\cdot\overline{e}=(-2-4)\cdot\overline{e}=-6\cdot\overline{e};$$

$$3\cdot\overline{a}+2\cdot\overline{b}=3\cdot(-2\cdot\overline{e})+2\cdot(4\cdot\overline{e})=\left[3\cdot(-2)+2\cdot4\right]\cdot\overline{e}=2\cdot\overline{e}.$$

Thus $\bar{a}+\bar{b}=2\cdot\bar{e}=\frac{1}{2}\cdot\bar{b}$. Note, that vector $\bar{a}+\bar{b}$, relative to basis $\bar{e}$, has coordinate equal to 2, and relative to basis $\bar{b}$ – coordinate equal to $\frac{1}{2}$, i.e. vector has unequal coordinates relative to different bases. ■

8.3.2. Basis on Plane. Vector Coordinates on Plane

A ***basis on plane*** is a system of two noncollinear vectors $\bar{e}_1,\bar{e}_2$ of this plane, taken in specific order (Fig. 8.10). These vectors $\bar{e}_1,\bar{e}_2$ are called ***basis***.

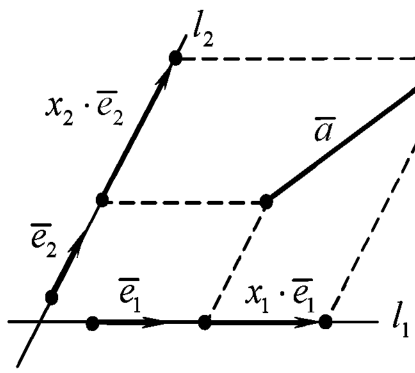


Figure 8.10

Theorem of vector decomposition on plane. *Any vector $\bar{a}$ of a plane can be decomposed by basis $\bar{e}_1,\bar{e}_2$ on this plane, i.e. it can be represented in a form $\bar{a}=x_1\cdot\bar{e}_1+x_2\cdot\bar{e}_2$, where numbers x_1 and x_2 are uniquely defined.*

Coefficients x_1 and x_2 in decomposition are called ***coordinates of vector $\bar{a}$ relative to basis*** $\bar{e}_1,\bar{e}_2$ (number is called ***abscissa*** and x_2 – ***ordinate*** of vector $\bar{a}$), e.g. numbers 2 and -3 are coordinates of vector $\bar{a}=2\cdot\bar{e}_1-3\cdot\bar{e}_2$ ($x_1=2-$ abscissa, $x_2=-3$ – ordinate).

A basis on plane is called ***right*** (or an ordered pair of noncollinear vectors is called ***right*** pair), if the shortest turn from the first vector to the second one is counterclockwise (this direction is assumed as positive). Basis vectors $\bar{e}_1,\bar{e}_2$ (Fig. 8.11, *a*) of right basis are ordered as thumb and forefinger of right hand (if we look at palm).

A ***left*** basis on plane (***left pair***) is such a basis, that the shortest turn from vector $\overline{e}_1$ to vector $\overline{e}_2$ is clockwise (this direction is assumed as negative). Basis vectors $\overline{e}_1, \overline{e}_2$ (Fig. 8.11, *b*) of left basis are ordered as thumb and forefinger of left hand (if we look at palm).

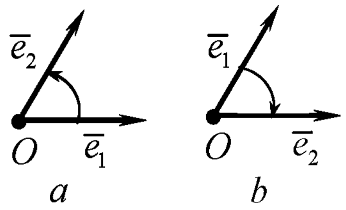


Figure 8.11

8.3.3. Basis in Space. Vector Coordinates in Space

A ***basis in space*** is a system of three noncoplanar vectors $\overline{e}_1, \overline{e}_2, \overline{e}_3$, taken in specific order (Fig. 8.13). These vectors $\overline{e}_1, \overline{e}_2, \overline{e}_3$ are called ***basis***.

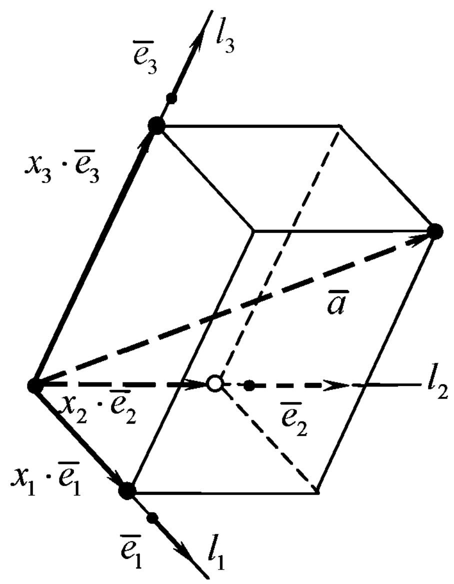


Figure 8.13

Theorem of vector decomposition in space. *Any vector $\overline{a}$ can be decomposed by basis $\overline{e}_1, \overline{e}_2, \overline{e}_3$ in space*, i.e. *represented in a form $\overline{a} = x_1 \cdot \overline{e}_1 + x_2 \cdot \overline{e}_2 + x_3 \cdot \overline{e}_3$, where numbers x_1, x_2, x_3 are uniquely defined.*

Coefficients x_1, x_2, x_3 in decomposition are called ***coordinates of vector $\overline{a}$ relative to basis*** $\overline{e}_1, \overline{e}_2, \overline{e}_3$ (number x_1 is called ***abscissa***, x_2 – ***ordinate***, x_3 – ***applicate*** of vector $\overline{a}$), e.g. numbers 3, 2, -1 are coordinates of vector $\overline{a}=3\cdot\overline{e}_1+2\cdot\overline{e}_2-\overline{e}_3$ ($x_1=3$ – abscissa, $x_2=2$ – ordinate, $x_3=-1$ – applicate).

A basis in space is called ***right*** (ordered triplet of noncoplanar vectors is called ***right*** triplet), if looking from the head of the third vector the shortest turn from the first vector to the second one is counterclockwise (Fig. 8.14, *a*). If the described turn is clockwise, then the basis is called ***left*** (ordered triplet of noncoplanar vectors is called ***left*** triplet) (Fig. 8.14, *b*).

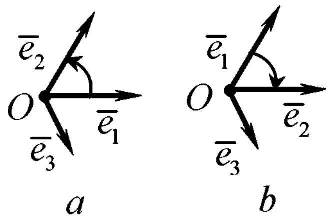


Figure 8.14

8.3.4. Linear Operations in Coordinate Form

Theorems of vector basis decomposition determine one-to-one correspondence between a set of vectors in space and a set of their coordinates in current basis, to be exact: between vectors on line and real numbers, between vectors on plane and ordered pairs of numbers, between vectors in space and ordered triplets of numbers.

For example, in fixed basis $(\overline{e})=(\overline{e}_1,\overline{e}_2,\overline{e}_3)$ for vector $\overline{a}=x_1\cdot\overline{e}_1+x_2\cdot\overline{e}_2+x_3\cdot\overline{e}_3$ there is a uniquely specified ordered triplet of numbers x_1,x_2,x_3, and vice versa, any for any ordered triplet of numbers x_1,x_2,x_3 there is a vector $\overline{a}=x_1\cdot\overline{e}_1+x_2\cdot\overline{e}_2+x_3\cdot\overline{e}_3$, i.e.

$$\overline{a}\underset{(\overline{e})}{\leftrightarrow}(x_1,x_2,x_3).$$

Example: if vector $\overline{a}$ in basis $(\overline{e})=(\overline{e}_1,\overline{e}_2,\overline{e}_3)$ has decomposition $\overline{a}=2\cdot\overline{e}_1-3\cdot\overline{e}_2+4\cdot\overline{e}_3$, then this vector corresponds to triplet $(2,-3,4)$ and vice versa.

A zero vector in any basis corresponds to a zero triplet $(0,0,0)$.

It is convenient to represent vector coordinates as a column-matrix (or row-matrix), which are called ***coordinate columns*** (***coordinate rows***).

In basis $(\overline{e})=(\overline{e}_1,\overline{e}_2,\overline{e}_3)$ vector $\overline{a}=x_1\cdot\overline{e}_1+x_2\cdot\overline{e}_2+x_3\cdot\overline{e}_3$ corresponds to coordinate column $\underset{(\overline{e})}{a}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$. Basis notation $(\overline{e})$ can be omitted, if it does not lead to ambiguity.

Vector linear operations correspond to coordinate columns linear operations, e.g. if in basis $(\overline{e})$ vectors $\overline{a}$ and $\overline{b}$ correspond to vector columns a and b, then their linear combination $\overline{c}=\alpha\cdot\overline{a}+\beta\cdot\overline{b}$ corresponds to coordinate column $c=\alpha\cdot a+\beta\cdot b$, i.e. *coordinate column of vectors' linear combination equals to linear combination of its coordinate columns.*

Note: concepts of linear dependence and linear independence of systems of columns with all properties transfer to vectors and coordinate columns.

Example 8.5. Vectors $\overline{a}$ and $\overline{b}$ relative to basis $\overline{e}_1,\overline{e}_2,\overline{e}_3$ have coordinates $2,0,-3$ and $4,2,-1$ accordingly.

Find coordinates of vectors $\overline{a}+\overline{b}$, $\overline{a}-\overline{b}$, $3\cdot\overline{a}+2\cdot\overline{b}$ relative to the same basis.

□ Write basis decompositions of the given vectors:

$$\overline{a}=2\cdot\overline{e}_1+0\cdot\overline{e}_2-3\cdot\overline{e}_3;\qquad \overline{b}=4\cdot\overline{e}_1+2\cdot\overline{e}_2-1\cdot\overline{e}_3.$$

Using the properties of linear operations, find basis decomposition of the given vectors:

$$\overline{a}+\overline{b}=(2+4)\cdot\overline{e}_1+(0+2)\cdot\overline{e}_2+(-3-1)\cdot\overline{e}_3=6\cdot\overline{e}_1+2\cdot\overline{e}_2-4\cdot\overline{e}_3;$$

$$\overline{a}-\overline{b}=(2-4)\cdot\overline{e}_1+(0-2)\cdot\overline{e}_2+(-3+1)\cdot\overline{e}_3=-2\cdot\overline{e}_1-2\cdot\overline{e}_2-2\cdot\overline{e}_3;$$

$$3\cdot\overline{a}+2\cdot\overline{b}=3\cdot(2\cdot\overline{e}_1+0\cdot\overline{e}_2-3\cdot\overline{e}_3)+2\cdot(4\cdot\overline{e}_1+2\cdot\overline{e}_2-1\cdot\overline{e}_3)=$$

$$=(3\cdot2+2\cdot4)\cdot\overline{e}_1+(3\cdot0+2\cdot2)\cdot\overline{e}_2+[3\cdot(-3)+2\cdot(-1)]\cdot\overline{e}_3=14\cdot\overline{e}_1+4\cdot\overline{e}_2-11\cdot\overline{e}_3.$$

Thus, vectors $\overline{a}+\overline{b}$, $\overline{a}-\overline{b}$, $3\cdot\overline{a}+2\cdot\overline{b}$ have coordinates $6, 2, -4$; $-2, -2, -2$; $14, 4, -11$ accordingly.

Let's find obtained coordinates using the matrix notation. Vectors $\overline{a}$ and $\overline{b}$ (in given basis) corresponds to the following coordinate columns

$$a=\begin{pmatrix}2\\0\\-3\end{pmatrix},\quad b=\begin{pmatrix}4\\2\\-1\end{pmatrix}.$$

Find coordinate columns of vectors $\overline{a}+\overline{b}$, $\overline{a}-\overline{b}$, $3\cdot\overline{a}+2\cdot\overline{b}$:

$$a+b=\begin{pmatrix}2\\0\\-3\end{pmatrix}+\begin{pmatrix}4\\2\\-1\end{pmatrix}=\begin{pmatrix}6\\2\\-4\end{pmatrix};\qquad a-b=\begin{pmatrix}2\\0\\-3\end{pmatrix}-\begin{pmatrix}4\\2\\-1\end{pmatrix}=\begin{pmatrix}-2\\-2\\-2\end{pmatrix};$$

$$3\cdot a+2\cdot b=3\cdot\begin{pmatrix}2\\0\\-3\end{pmatrix}+2\cdot\begin{pmatrix}4\\2\\-1\end{pmatrix}=\begin{pmatrix}14\\4\\-11\end{pmatrix}.$$

Results are the same. ■

8.3.5. Orthogonal and Orthonormal Bases

Two vectors are called ***orthogonal*** (***perpendicular***), if the angle between them is the right angle (value φ equals to $\frac{\pi}{2}$).

A system of vectors is called ***orthogonal***, if all forming vectors are pairwise orthogonal. A system of vectors if called ***orthonormal***, if it is orthogonal and the length of each vector equals to unit.

Standard basis on line, plane and in space

Bases on line, plane and in space are not uniquely defined. Some of them, which are more convenient to use, are accepted as standard.

Standard basis on line is unit vector $\overline{i}$ on the given line (Fig. 8.15, *a*). *Any vector $\overline{a}$, which is collinear to the given line, can be decomposed by the standard basis on line ($\overline{e}=\overline{i}$), i.e. represented in form $\overline{a}=x\cdot\overline{i}$.*

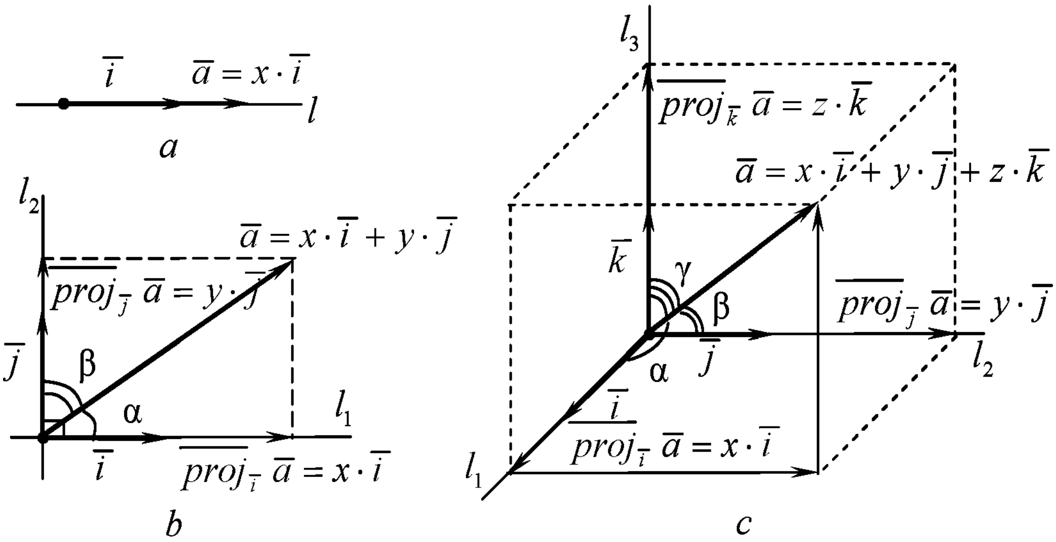


Figure 8.15

A ***standard basis on plane*** is an ordered pair of unit and perpendicular vectors $\overline{i}, \overline{j}$ on the given plane (Fig. 8.15, *b*). *Any vector $\overline{a}$ on the given plane can be decomposed by the standard basis on plane ($\overline{e}_1 = \overline{i}$, $\overline{e}_2 = \overline{j}$), i.e. represented in a form $\overline{a} = x \cdot \overline{i} + y \cdot \overline{j}$.*

A ***standard basis in space*** is an ordered triplet of unit and pairwise perpendicular vectors $\overline{i}, \overline{j}, \overline{k}$ (Fig. 8.15, *c*). The first basis vector $\overline{i}$ in Fig. 8.15, *c* is directed perpendicularly to the figure's plane (towards the reader). *Any vector $\overline{a}$ in space can be decomposed by standard bases in space ($\overline{e}_1 = \overline{i}$, $\overline{e}_2 = \overline{j}$, $\overline{e}_3 = \overline{k}$), i.e. represented in form $\overline{a} = x \cdot \overline{i} + y \cdot \overline{j} + z \cdot \overline{k}$.*

Standard bases on plane and in space are orthonormal right bases.

In standard basis length of vector equals to square root of its component sum:

$$|\overline{a}| = \sqrt{x^2} \text{ (on line)};$$

$$|\overline{a}| = \sqrt{x^2 + y^2} \text{ (on plane)};$$

$$|\overline{a}| = \sqrt{x^2 + y^2 + z^2} \text{ (in space)}.$$

Direction Cosines

In standard bases on plane and in space it is convenient to describe the direction of nonzero vector $\overline{a}$ by the angles between the vector and basis vectors: α – the angle between $\overline{a}$ and the first basis vector $\overline{i}$; β – the angle between $\overline{a}$ and the second basis vector $\overline{j}$ (Fig. 8.15, *b*); γ – the angle between $\overline{a}$ and the third basis vector $\overline{k}$ (Fig. 8.15, *c*). It is sufficient to take into account angles cosines, which are called ***direction cosines of vector*** $\overline{a}$ (in standard basis).

Coordinates of unit vector $\overline{e}$, equally directed with vector $\overline{a}$ on plane, are equal to direction cosines of vector $\overline{a}$:

$$\overline{e} = \frac{\overline{a}}{|\overline{a}|} = \cos\alpha \cdot \overline{i} + \cos\beta \cdot \overline{j},$$

i.e. $x = \cos\alpha$, $y = \cos\beta$. Values of direction cosines satisfy the following condition: $\cos^2\alpha + \cos^2\beta = 1$.

Coordinates of unit vector $\overline{e}$, equally directed with vector $\overline{a}$ in space, are equal to direction cosines of vector $\overline{a}$:

$$\overline{e} = \frac{\overline{a}}{|\overline{a}|} = \cos\alpha \cdot \overline{i} + \cos\beta \cdot \overline{j} + \cos\gamma \cdot \overline{k},$$

i.e. $x = \cos\alpha$, $y = \cos\beta$, $z = \cos\gamma$ and $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

Example 8.6. Find lengths and direction cosines of vectors $\overline{a} = 3 \cdot \overline{i} - \sqrt{3} \cdot \overline{j}$ and $\overline{b} = \overline{i} - 2 \cdot \overline{j} + 2 \cdot \overline{k}$.

□ Vector $\overline{a} = 3 \cdot \overline{i} - \sqrt{3} \cdot \overline{j}$ is defined relative to standard basis $\overline{i}, \overline{j}$ on plane.

By coordinates $x = 3$, $y = -\sqrt{3}$ of vector $\overline{a}$ find its length by the formula (8.1): $|\overline{a}| = \sqrt{3^2 + (-\sqrt{3})^2} = 2\sqrt{3}$.

Dividing vector $\overline{a}$ by its length we find the unit vector, equally directed with vector $\overline{a}$: $\dfrac{\overline{a}}{|\overline{a}|} = \dfrac{3}{2\sqrt{3}} \cdot \overline{i} - \dfrac{\sqrt{3}}{2\sqrt{3}} \cdot \overline{j} = \dfrac{\sqrt{3}}{2} \cdot \overline{i} - \dfrac{1}{2} \cdot \overline{j}$.

According to (8.3), its coordinates are direction cosines $\cos\alpha = \frac{\sqrt{3}}{2}$; $\cos\beta = -\frac{1}{2}$. So, vector $\overline{a}$ forms the following angles with basis vectors $\overline{i}$ and $\overline{j}$: $\alpha = \frac{\pi}{6}$ and $\beta = \frac{2\pi}{3}$.

Vector $\overline{b} = \overline{i} - 2 \cdot \overline{j} + 2 \cdot \overline{k}$ is defined relative to standard basis $\overline{i}, \overline{j}, \overline{k}$ in space.

By coordinates $x = 1$, $y = -2$, $z = 2$ of vector $\overline{b}$ find its length by the formula (8.2): $|\overline{b}| = \sqrt{1^2 + (-2)^2 + 2^2} = 3$.

Dividing vector $\overline{b}$ by its length we find the unit vector, equally directed with vector $\overline{b}$: $\frac{\overline{b}}{|\overline{b}|} = \frac{1}{3} \cdot \overline{i} - \frac{2}{3} \cdot \overline{j} + \frac{2}{3} \cdot \overline{k}$.

According to (8.4), its coordinates are direction cosines: $\cos\alpha = \frac{1}{3}$; $\cos\beta = -\frac{2}{3}$; $\cos\gamma = \frac{2}{3}$. ■

8.4. SCALAR PRODUCT OF VECTORS

A ***scalar product*** of two nonzero vectors is a number, equal to the product of their lengths and cosine of angle between them. If at least one vector is zero, then the angle between them is not defined, and product is assumed to be equal to zero. Scalar product of vectors $\overline{a}$ and $\overline{b}$ is denoted by

$$(\overline{a}, \overline{b}) = |\overline{a}| \cdot |\overline{b}| \cdot \cos\varphi, \tag{8.5}$$

where φ is a value of the angle between $\overline{a}$ and $\overline{b}$ (Fig. 8.2 in section 8.1.1). A scalar product $(\overline{a}, \overline{a}) = |\overline{a}|^2$ is called ***scalar square***.

A *scalar product of two nonzero vectors* $\overline{a}$ *and* $\overline{b}$ *equals to the product of vector* $\overline{b}$ *length and algebraic value of orthogonal projection of vector* $\overline{a}$ *to axis, formed by vector* $\overline{b}$ (Fig. 8.16):

$$(\overline{a}, \overline{b}) = |\overline{b}| \cdot proj_{\overline{b}}\overline{a}. \tag{8.6}$$

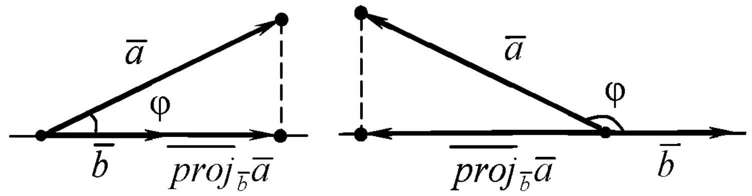


Figure 8.16

This formula remains correct if $\overline{a} = \overline{o}$, because $proj_{\overline{b}}\overline{o} = 0$.

In other words, a *scalar product of nonzero vectors $\overline{a}$ and $\overline{b}$ equals to the product of vector $\overline{a}$ length and algebraic value of orthogonal projection length of vector $\overline{b}$ to axis, formed by vector $\overline{a}$*:

$$(\overline{a},\overline{b}) = |\overline{a}| \cdot proj_{\overline{a}}\overline{b}.$$

Example 8.7. Find scalar products $(\overline{a},\overline{b})$, $(\overline{b},\overline{a})$, $(\overline{a},\overline{c})$, $(\overline{b},\overline{c})$, $(\overline{a},\overline{d})$, $(\overline{b},\overline{d})$, $(\overline{c},\overline{d})$, if it is known, that $|\overline{a}|=1$, $|\overline{b}|=2$, $|\overline{c}|=4$, $|\overline{d}|=1$, angle φ between vectors $\overline{a}$ and $\overline{b}$ equals to $\frac{\pi}{3}$, $\overline{c}\uparrow\downarrow\overline{b}$, and vector $\overline{d}$ forms angle with vector $\overline{a}$ equal to $\delta = \frac{5\pi}{6}$ (Fig. 8.17).

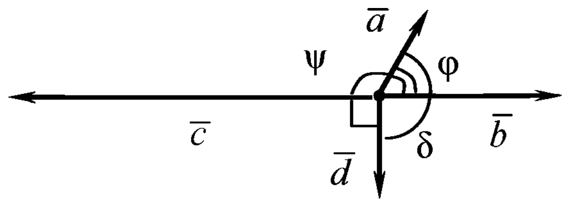


Figure 8.17

□ By the definitions find $(\overline{a},\overline{b}) = |\overline{a}| \cdot |\overline{b}| \cdot \cos\varphi = 1 \cdot 2 \cdot \cos\frac{\pi}{3} = 1$; $(\overline{b},\overline{a}) = |\overline{b}| \cdot |\overline{a}| \cdot \cos\varphi = 2 \cdot 1 \cdot \cos\frac{\pi}{3} = 1$. Since vectors $\overline{b}$ and $\overline{c}$ are oppositely directed, then angle ψ between vectors $\overline{a}$ and $\overline{c}$ equals to $\frac{2\pi}{3}$, so $(\overline{a},\overline{c}) = |\overline{a}| \cdot |\overline{c}| \cdot \cos\psi = 1 \cdot 4 \cdot \cos\frac{2\pi}{3} = -2$.

Angle between oppositely directed vectors $\overline{b}$ and $\overline{c}$ equals to π, so $(\overline{b},\overline{c}) = |\overline{b}| \cdot |\overline{c}| \cdot \cos\pi = 2 \cdot 4 \cdot \cos\pi = -8$.

Vector $\overline{d}$ is orthogonal to vector $\overline{b}$ (and vector $\overline{c}$), because value of angle between them equals to $\frac{5\pi}{6} - \frac{\pi}{3} = \frac{\pi}{2}$ and $\cos\frac{\pi}{2} = 0$, so $(\overline{b},\overline{d}) = (\overline{c},\overline{d}) = 0$.

Angle δ between vectors $\overline{a}$ and $\overline{d}$ equals to $\frac{5\pi}{6}$, so $(\overline{a},\overline{d}) = 1 \cdot 1 \cdot \cos\frac{5\pi}{6} = -\frac{\sqrt{3}}{2}$. ■

Algebraic Properties of Scalar Product

For any vectors $\overline{a}$, $\overline{b}$, $\overline{c}$ and any real number λ:

1. $(\overline{a},\overline{b}) = (\overline{b},\overline{a})$;
2. $(\overline{a} + \overline{b},\overline{c}) = (\overline{a},\overline{c}) + (\overline{b},\overline{c})$;
3. $(\lambda \cdot \overline{a},\overline{b}) = \lambda \cdot (\overline{a},\overline{b})$;
4. $(\overline{a},\overline{a}) \geq 0$, and from the equality $(\overline{a},\overline{a}) = 0$ follows that $\overline{a} = \overline{o}$.

Geometric Properties of Scalar Product

1. Length of vector $\overline{a}$ is calculated by formula $|\overline{a}| = \sqrt{(\overline{a},\overline{a})}$.
2. Value of angle φ between two nonzero vectors is calculated by formula:

$$\cos\varphi = \frac{(\overline{a},\overline{b})}{|\overline{a}| \cdot |\overline{b}|} = \frac{(\overline{a},\overline{b})}{\sqrt{(\overline{a},\overline{a})} \cdot \sqrt{(\overline{b},\overline{b})}}.$$

3. Algebraic value of orthogonal projection length of vector $\overline{a}$ to axis, formed by vector $\overline{b} \neq \overline{o}$: $proj_{\overline{b}}\, \overline{a} = \frac{(\overline{a},\overline{b})}{|\overline{b}|} = \frac{(\overline{a},\overline{b})}{\sqrt{(\overline{b},\overline{b})}}$.

4. Orthogonal projection of vector $\overline{a}$ to axis, formed by vector $\overline{b} \neq \overline{o}$: $\overline{proj_{\overline{b}}\overline{a}} = \frac{(\overline{a},\overline{b})}{(\overline{b},\overline{b})} \cdot \overline{b}$. If axis is formed by unit vector $\overline{e}$, then $\overline{proj_{\overline{e}}\,\overline{a}} = (\overline{a},\overline{e}) \cdot \overline{e}$.

Scalar Product in Orthonormal Basis

In orthonormal basis a scalar product of vectors equals to the sum of the products of its corresponding elements:

1) if vectors $\overline{a}$ and $\overline{b}$ relative to orthonormal basis on plane have coordinates x_a, y_a and x_b, y_b accordingly, then the scalar product of these vectors is calculated by the formula

$$(\overline{a},\overline{b}) = x_a \cdot x_b + y_a \cdot y_b; \tag{8.7}$$

2) if vectors $\overline{a}$ and $\overline{b}$ relative to orthonormal basis in space have coordinates x_a, y_a, z_a and x_b, y_b, z_b accordingly, then the scalar product of these vectors is calculated by the formula

$$(\overline{a},\overline{b}) = x_a \cdot x_b + y_a \cdot y_b + z_a \cdot z_b. \tag{8.8}$$

Coordinates of vector $\overline{a}$ in orthonormal bases equal to scalar products of this vector and according basis vectors:

$$x_a = (\overline{a},\overline{i}), \quad y_a = (\overline{a},\overline{j}), \quad z_a = (\overline{a},\overline{k}).$$

Example 8.8. Given vectors $\overline{a} = \overline{i} - 2\cdot\overline{j} + 2\cdot\overline{k}$, $\overline{b} = 2\cdot\overline{i} + 3\cdot\overline{j} + 2\cdot\overline{k}$, $\overline{c} = \overline{j} - \overline{k}$, find scalar products $(\overline{a},\overline{b})$, $(\overline{a},\overline{c})$, $(\overline{b},\overline{c})$, $(\overline{a},\overline{i})$, $(\overline{a},\overline{j})$, $(\overline{a},\overline{k})$, lengths of vectors $|\overline{a}|$, $|\overline{b}|$, $|\overline{c}|$, angles $\varphi_{\overline{a}\overline{b}}$, $\varphi_{\overline{a}\overline{c}}$ between vectors $\overline{a}$ and $\overline{b}$, $\overline{a}$ and $\overline{c}$ accordingly, and projection $\overline{proj_{\overline{c}}\overline{a}}$ and algebraic value $proj_{\overline{c}}\overline{a}$ of the projection length.

☐ By the geometric Properties 1 – 4 and (8.8), obtain:

$$(\overline{a},\overline{b}) = (1\cdot\overline{i} - 2\cdot\overline{j} + 2\cdot\overline{k}, 2\cdot\overline{i} + 3\cdot\overline{j} + 2\cdot\overline{k}) = 1\cdot 2 + (-2)\cdot 3 + 2\cdot 2 = 0;$$

$$(\overline{a},\overline{c}) = (1\cdot\overline{i} - 2\cdot\overline{j} + 2\cdot\overline{k}, 0\cdot\overline{i} + 1\cdot\overline{j} - 1\cdot\overline{k}) = 1\cdot 0 + (-2)\cdot 1 + 2\cdot(-1) = -4;$$

$$(\overline{b},\overline{c}) = (2\cdot\overline{i} + 3\cdot\overline{j} + 2\cdot\overline{k}, 0\cdot\overline{i} + 1\cdot\overline{j} - 1\cdot\overline{k}) = 2\cdot 0 + 3\cdot 1 + 2\cdot(-1) = 1;$$

$$(\overline{a},\overline{i}) = (1\cdot\overline{i} - 2\cdot\overline{j} + 2\cdot\overline{k}, 1\cdot\overline{i} + 0\cdot\overline{j} + 0\cdot\overline{k}) = 1\cdot 1 + (-2)\cdot 0 + 2\cdot 0 = 1;$$

$$(\overline{a},\overline{j}) = (1\cdot\overline{i} - 2\cdot\overline{j} + 2\cdot\overline{k}, 0\cdot\overline{i} + 1\cdot\overline{j} + 0\cdot\overline{k}) = 1\cdot 0 + (-2)\cdot 1 + 2\cdot 0 = -2;$$

$$(\overline{a},\overline{k}) = (1\cdot\overline{i} - 2\cdot\overline{j} + 2\cdot\overline{k}, 0\cdot\overline{i} + 0\cdot\overline{j} + 1\cdot\overline{k}) = 1\cdot 0 + (-2)\cdot 0 + 2\cdot 1 = 2;$$

$$|\overline{a}|=\sqrt{(\overline{a},\overline{a})}=\sqrt{1^2+(-2)^2+2^2}=3; \quad |\overline{b}|=\sqrt{(\overline{b},\overline{b})}=\sqrt{2^2+3^2+2^2}=\sqrt{17};$$

$$|\overline{c}|=\sqrt{(\overline{c},\overline{c})}=\sqrt{0^2+1^2+(-1)^2}=\sqrt{2};$$

$$\cos\varphi_{\overline{a}\overline{b}}=\frac{(\overline{a},\overline{b})}{|\overline{a}||\overline{b}|}=0 \quad \Rightarrow \quad \varphi_{\overline{a}\overline{b}}=\frac{\pi}{2} \text{ (vectors } \overline{a} \text{ and } \overline{b} \text{ are orthogonal);}$$

$$\cos\varphi_{\overline{a}\overline{c}}=\frac{(\overline{a},\overline{c})}{|\overline{a}||\overline{c}|}=\frac{-4}{3\cdot\sqrt{2}}=-\frac{2\sqrt{2}}{3} \quad \Rightarrow \quad \varphi_{\overline{a}\overline{c}}=\arccos\left(-\frac{2\sqrt{2}}{3}\right);$$

$$proj_{\overline{c}}\overline{a}=\frac{(\overline{a},\overline{c})}{|\overline{c}|}=\frac{-4}{\sqrt{2}}=-2\sqrt{2}; \quad \overline{proj_{\overline{c}}\overline{a}}=\frac{(\overline{a},\overline{c})}{(\overline{c},\overline{c})}\cdot\overline{c}=\frac{-4}{2}\cdot\left(\overline{j}-\overline{k}\right)=-2\overline{j}+2\overline{k}. \ \blacksquare$$

8.5. OUTER PRODUCT OF VECTORS

A vector $\overline{c}$ is called an ***outer product of noncollinear vectors*** $\overline{a}$ and $\overline{b}$, if:

1) its length is equal to the product of vectors $\overline{a}$ and $\overline{b}$ lengths and sine of angle between them: $|\overline{c}|=|\overline{a}|\cdot|\overline{b}|\cdot\sin\varphi$ (Fig. 8.18);
2) vector $\overline{c}$ is orthogonal to vectors $\overline{a}$ and $\overline{b}$;
3) vectors $\overline{a},\overline{b},\overline{c}$ (in the given order) are right triplet.

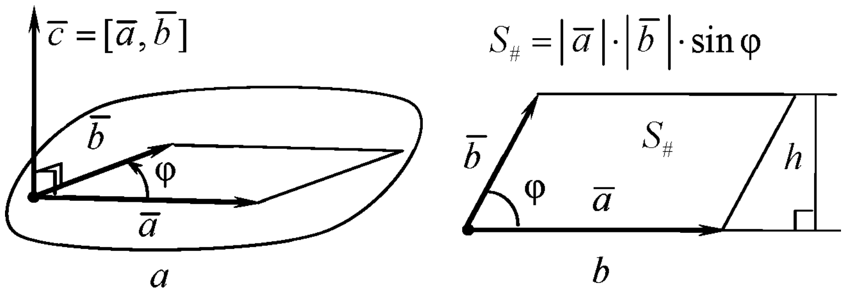


Figure 8.18

An outer product of collinear vectors (in particular, if at least one of them is zero vector) equals to zero vector. The outer product is denoted as $\overline{c}=[\overline{a},\overline{b}]$ (or $\overline{a}\times\overline{b}$).

Algebraic Properties of Outer Product

For any vectors $\overline{a}$, $\overline{b}$, $\overline{c}$ and any real number λ:

1. $[\overline{a},\overline{b}]=-[\overline{b},\overline{a}]$;

2. $[\overline{a}+\overline{b},\overline{c}]=[\overline{a},\overline{c}]+[\overline{b},\overline{c}]$;

3. $[\lambda\cdot\overline{a},\overline{b}]=\lambda\cdot[\overline{a},\overline{b}]$.

Geometric Properties of Outer Product

1. The absolute value of vectors' outer product numerically equals to the area of a parallelogram, constructed on these vectors (Fig. 8.18, *b*).

2. An outer product equals to zero vector if and only if the vectors are collinear, i.e. $[\overline{a},\overline{b}]=\overline{o} \Leftrightarrow \overline{a}\parallel\overline{b}$, in particular $[\overline{a},\overline{a}]=\overline{o}$.

Outer Product of Vectors in Orthonormal Basis

Consider right orthonormal (standard) basis in space $\overline{i},\overline{j},\overline{k}$ (sect. 8.3.5). Outer products of basis vectors are found by the definition:

$$[\overline{i},\overline{j}]=\overline{k};\quad [\overline{j},\overline{k}]=\overline{i};\quad [\overline{k},\overline{i}]=\overline{j};\quad [\overline{j},\overline{i}]=-\overline{k};\quad [\overline{k},\overline{j}]=-\overline{i};\quad [\overline{i},\overline{k}]=-\overline{j};$$

$$[\overline{i},\overline{i}]=[\overline{j},\overline{j}]=[\overline{k},\overline{k}]=\overline{o}.$$

Outer product calculation formula. If vectors $\overline{a}$ and $\overline{b}$ in right orthonormal basis $\overline{i},\overline{j},\overline{k}$ have coordinates x_a, y_a, z_a and x_b, y_b, z_b accordingly, then the outer product of these vectors is calculated by the formula

$$[\overline{a},\overline{b}]=\overline{i}\cdot\begin{vmatrix} y_a & z_a \\ y_b & z_b \end{vmatrix}-\overline{j}\cdot\begin{vmatrix} x_a & z_a \\ x_b & z_b \end{vmatrix}+\overline{k}\cdot\begin{vmatrix} x_a & y_a \\ x_b & y_b \end{vmatrix}=\begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ x_a & y_a & z_a \\ x_b & y_b & z_b \end{vmatrix}. \qquad (8.9)$$

Example 8.9. Parallelogram *ABCD* is constructed on vectors $\overline{AB}=\overline{i}+2\cdot\overline{j}+2\cdot\overline{k}$, $\overline{AD}=3\cdot\overline{i}-2\cdot\overline{j}+\overline{k}$ (Fig. 8.19). Find:

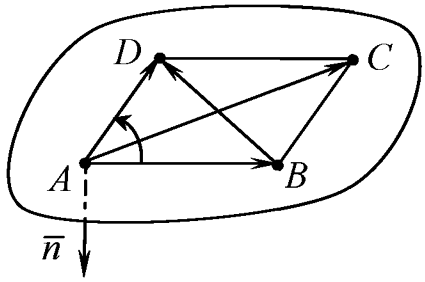


Figure 8.19

a) outer products $[\overline{AB}, \overline{AD}]$ and $[\overline{AC}, \overline{BD}]$;

b) area of parallelogram $ABCD$;

c) direction cosines of vector $\overline{n}$, perpendicular to the plane, which contains $ABCD$, and that form the left triplet $\overline{AB}, \overline{AD}, \overline{n}$.

□ a) Outer product $[\overline{AB}, \overline{AD}]$ is calculated by the formula (8.9):

$$[\overline{AB}, \overline{AD}] = \begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ 1 & 2 & 2 \\ 3 & -2 & 1 \end{vmatrix} = \overline{i} \cdot \begin{vmatrix} 2 & 2 \\ -2 & 1 \end{vmatrix} - \overline{j} \cdot \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} + \overline{k} \cdot \begin{vmatrix} 1 & 2 \\ 3 & -2 \end{vmatrix} = 6 \cdot \overline{i} + 5 \cdot \overline{j} - 8 \cdot \overline{k}.$$

Outer product $[\overline{AC}, \overline{BD}]$ is determined by the algebraic properties:

$$[\overline{AC}, \overline{BD}] = [\overline{AB} + \overline{AD}, \overline{AD} - \overline{AB}] = [\overline{AB}, \overline{AD}] - \underbrace{[\overline{AB}, \overline{AB}]}_{\overline{o}} + \underbrace{[\overline{AD}, \overline{AD}]}_{\overline{o}} - \underbrace{[\overline{AD}, \overline{AB}]}_{-[\overline{AB}, \overline{AD}]} =$$

$$= [\overline{AB}, \overline{AD}] + [\overline{AB}, \overline{AD}] = 2 \cdot [\overline{AB}, \overline{AD}].$$

Consequently, $[\overline{AC}, \overline{BD}] = 2 \cdot \left(6 \cdot \overline{i} + 5 \cdot \overline{j} - 8 \cdot \overline{k}\right) = 12 \cdot \overline{i} + 10 \cdot \overline{j} - 16 \cdot \overline{k}$.

b) Area of parallelogram $ABCD$ is found as an absolute value of product $[\overline{AB}, \overline{AD}]$:

$$S_{\#} = \left|[\overline{AB}, \overline{AD}]\right| = \left|6 \cdot \overline{i} + 5 \cdot \overline{j} - 8 \cdot \overline{k}\right| = \sqrt{6^2 + 5^2 + (-8)^2} = 5\sqrt{5}.$$

c) Vector, that is oppositely directed with vector $[\overline{AB}, \overline{AD}]$, satisfies enumerated conditions, thus

$$\overline{n} = -[\overline{AB}, \overline{AD}] = -\left(6 \cdot \overline{i} + 5 \cdot \overline{j} - 8 \cdot \overline{k}\right) = -6 \cdot \overline{i} - 5 \cdot \overline{j} + 8 \cdot \overline{k}.$$

Dividing this vector by its length $|\overline{n}| = |[\overline{AB}, \overline{AD}]| = 5\sqrt{5}$, we will obtain the unit vector: $\frac{\overline{n}}{|\overline{n}|} = \frac{-6\cdot\overline{i} - 5\cdot\overline{j} + 8\cdot\overline{k}}{5\sqrt{5}} = -\frac{6}{5\sqrt{5}}\cdot\overline{i} - \frac{5}{5\sqrt{5}}\cdot\overline{j} + \frac{8}{5\sqrt{5}}\cdot\overline{k}$. According to (8.4), its coordinates are direction cosines: $\cos\alpha = \frac{-6}{5\sqrt{5}}$, $\cos\beta = \frac{-1}{\sqrt{5}}$, $\cos\gamma = \frac{8}{5\sqrt{5}}$. ■

8.6. COMPOSITIONAL PRODUCT OF VECTORS

A ***compositional product of vectors*** $\overline{a}, \overline{b}, \overline{c}$ is a number $(\overline{a}, [\overline{b}, \overline{c}])$, equal to the scalar product of vector $\overline{a}$ and the outer product of vectors $\overline{b}$ и $\overline{c}$. A compositional product is denoted by $(\overline{a}, \overline{b}, \overline{c})$.

Geometric Properties of Compositional Product

1. The absolute value of compositional product of noncoplanar vectors $\overline{a}$, $\overline{b}$, $\overline{c}$ equals to the volume $V_{\#\overline{a},\overline{b},\overline{c}}$ of a parallelepiped, constructed on these vectors. A product $(\overline{a}, \overline{b}, \overline{c})$ is positive, if vector triplet $\overline{a}, \overline{b}, \overline{c}$ is right, and negative, if vector triplet $\overline{a}, \overline{b}, \overline{c}$ is left.
2. A compositional product $(\overline{a}, \overline{b}, \overline{c})$ equals to zero if and only if vectors $\overline{a}, \overline{b}, \overline{c}$ are coplanar, i.e.:

$$(\overline{a}, \overline{b}, \overline{c}) = 0 \quad \Leftrightarrow \quad \text{vectors } \overline{a}, \overline{b}, \overline{c} \text{ are coplanar.}$$

Algebraic Properties of Compositional Product

1. A swap of two multipliers in compositional product changes the sign to the opposite one:

$$(\overline{a}, \overline{b}, \overline{c}) = -(\overline{b}, \overline{a}, \overline{c}), \quad (\overline{a}, \overline{b}, \overline{c}) = -(\overline{c}, \overline{b}, \overline{a}), \quad (\overline{a}, \overline{b}, \overline{c}) = -(\overline{a}, \overline{c}, \overline{b});$$

a cycle (round) swap of multipliers does not change the product:

$$(\overline{a}, \overline{b}, \overline{c}) = (\overline{b}, \overline{c}, \overline{a}) = (\overline{c}, \overline{a}, \overline{b}).$$

2. A compositional product is linear by any multiplier.

Compositional Product of Vectors in Orthonormal Basis

Compositional product calculation formula. If vectors $\overline{a}, \overline{b}, \overline{c}$ in right orthonormal basis $\overline{i}, \overline{j}, \overline{k}$ have coordinates x_a, y_a, z_a; x_b, y_b, z_b; x_c, y_c, z_c accordingly, then the compositional product of these vectors can be calculated by the formula

$$(\overline{a}, \overline{b}, \overline{c}) = \begin{vmatrix} x_a & y_a & z_a \\ x_b & y_b & z_b \\ x_c & y_c & z_c \end{vmatrix}. \tag{8.10}$$

Example 8.10. Parallelepiped $ABCDA_1B_1C_1D_1$ is constructed on vectors $\overline{AB} = \overline{i} + 2\cdot\overline{j} + 2\cdot\overline{k}$; $\overline{AD} = 3\cdot\overline{i} - 2\cdot\overline{j} + \overline{k}$; $\overline{AA_1} = 2\cdot\overline{i} - 1\cdot\overline{j} + 3\cdot\overline{k}$ (Fig. 8.20). Find:

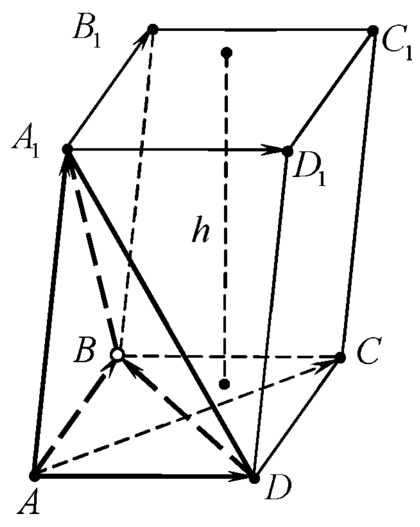


Figure 8.20

a) compositional product $(\overline{AB}, \overline{AD}, \overline{AA_1})$;

b) orientation of triplet $\overline{AB}, \overline{AD}, \overline{AA_1}$;

c) volume of parallelepiped $ABCDA_1B_1C_1D_1$;

d) volume of triangular pyramid $ABDA_1$;

e) height h of the parallelepiped (distance between planes of bases $ABCD$ and $A_1B_1C_1D_1$).

□ a) Compositional product $(\overline{AB}, \overline{AD}, \overline{AA_1})$ is found by formula (8.10):

$$(\overline{AB},\overline{AD},\overline{AA_1})=\begin{vmatrix}1 & 2 & 2\\ 3 & -2 & 1\\ 2 & -1 & 3\end{vmatrix}=-17.$$

b) Since the product is negative, then the triplet $\overline{AB},\overline{AD},\overline{AA_1}$ is left (by the first geometric property of compositional product).

c) Volume $V_{\#}$ of parallelepiped $ABCDA_1B_1C_1D_1$ equals to the absolute value of the compositional product (by the first geometric property of compositional product): $V_{\#}=\left|(\overline{AB},\overline{AD},\overline{AA_1})\right|=\left|-17\right|=17$.

d) Volume V of triangular pyramid $ABDA_1$ equals to one sixth of parallelepiped's volume $V_{\#}$. Indeed, their heights are equal, and area S_{base} of pyramid base equals to half of parallelogram $ABCD$ area $S_{\#}$. So $V=\frac{1}{3}\cdot S_{base}\cdot h=\frac{1}{3}\cdot\frac{1}{2}\cdot S_{\#}\cdot h=\frac{1}{6}\cdot V_{\#}$ and $V_{\#}=\left|(\overline{AB},\overline{AD},\overline{AA_1})\right|=17$ and then $V=\frac{1}{6}\cdot V_{\#}=\frac{17}{6}$.

e) Height h of the parallelepiped is obtained by formula $h=\frac{V_{\#}}{S_{\#}}$, where $S_{\#}$ is the area of parallelogram $ABCD$. Since $V_{\#}=17$ and $S_{\#}=5\sqrt{5}$ (example 8.11), then $h=\frac{17}{5\sqrt{5}}$. ■

8.7. METRIC APPLICATIONS OF VECTOR PRODUCTS

It is assumed that coordinates of vectors $\overline{a},\overline{b},\overline{c}$, which are given in formulas, are found relative to standard basis $\overline{i},\overline{j},\overline{k}$ in space:

$$\overline{a}=x_a\cdot\overline{i}+y_a\cdot\overline{j}+z_a\cdot\overline{k},$$

$$\overline{b}=x_b\cdot\overline{i}+y_b\cdot\overline{j}+z_b\cdot\overline{k},$$

$$\overline{c}=x_c\cdot\overline{i}+y_c\cdot\overline{j}+z_c\cdot\overline{k}.$$

Remember, that in standard basis scalar, outer and compositional products of vectors are calculated by formulas (8.8)–(8.10):

$$(\overline{a},\overline{b}) = x_a \cdot x_b + y_a \cdot y_b + z_a \cdot z_b;$$

$$[\overline{a},\overline{b}] = \begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ x_a & y_a & z_a \\ x_b & y_b & z_b \end{vmatrix};$$

$$(\overline{a},\overline{b},\overline{c}) = \begin{vmatrix} x_a & y_a & z_a \\ x_b & y_b & z_b \\ x_c & y_c & z_c \end{vmatrix}.$$

1. A vector $\overline{a} = \overline{o}$ if and only if

$$(\overline{a},\overline{a}) = 0 \quad \Leftrightarrow \quad x_a^2 + y_a^2 + z_a^2 = 0 \quad \Leftrightarrow \quad x_a = y_a = z_a = 0.$$

2. Nonzero vectors $\overline{a}$ and $\overline{b}$ are orthogonal if and only if

$$(\overline{a},\overline{b}) = 0 \quad \Leftrightarrow \quad x_a \cdot x_b + y_a \cdot y_b + z_a \cdot z_b = 0.$$

3. Vectors $\overline{a}$ and $\overline{b}$ are collinear if and only if

$$[\overline{a},\overline{b}] = \overline{o} \quad \Leftrightarrow \quad \begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ x_a & y_a & z_a \\ x_b & y_b & z_b \end{vmatrix} = \overline{o}.$$

4. Vectors $\overline{a}$, $\overline{b}$, $\overline{c}$ are coplanar if and only if

$$(\overline{a},\overline{b},\overline{c}) = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x_a & y_a & z_a \\ x_b & y_b & z_b \\ x_c & y_c & z_c \end{vmatrix} = 0.$$

5. The length of a vector $\overline{a} \neq \overline{o}$ is calculated by formula

$$|\overline{a}| = \sqrt{(\overline{a},\overline{a})} = \sqrt{x_a^2 + y_a^2 + z_a^2}.$$

6. The angle φ between two nonzero vectors $\overline{a}$ and $\overline{b}$ is calculated by formula

$$\cos\varphi=\frac{(\bar a,\bar b)}{\sqrt{(\bar a,\bar a)}\cdot\sqrt{(\bar b,\bar b)}}=\frac{x_a\cdot x_b+y_a\cdot y_b+z_a\cdot z_b}{\sqrt{x_a^2+y_a^2+z_a^2}\cdot\sqrt{x_b^2+y_b^2+z_b^2}}.$$

7. The algebraic value of the orthogonal projection of vector $\bar a$ on axis, formed by vector $\bar b\neq\bar o$, is calculated by formula

$$proj_{\bar b}\,\bar a=\frac{(\bar a,\bar b)}{|\bar b|}=\frac{x_a\cdot x_b+y_a\cdot y_b+z_a\cdot z_b}{\sqrt{x_b^2+y_b^2+z_b^2}}.$$

8. The orthogonal projection of vector $\bar a$ on axis, formed by vector $\bar b\neq\bar o$:

$$\overline{proj_{\bar b}\bar a}=\frac{(\bar a,\bar b)}{(\bar b,\bar b)}\cdot\bar b=\frac{x_a\cdot x_b+y_a\cdot y_b+z_a\cdot z_b}{x_b^2+y_b^2+z_b^2}\cdot\left(x_b\cdot\bar i+y_b\cdot\bar j+z_b\cdot\bar k\right).$$

9. Direction cosines of vector $\bar a$ are found by formulas

$$\cos\alpha=\frac{(\bar a,\bar i)}{|\bar a|}=\frac{x_a}{\sqrt{x_a^2+y_a^2+z_a^2}};\qquad \cos\beta=\frac{(\bar a,\bar j)}{|\bar a|}=\frac{y_a}{\sqrt{x_a^2+y_a^2+z_a^2}};$$

$$\cos\gamma=\frac{(\bar a,\bar k)}{|\bar a|}=\frac{z_a}{\sqrt{x_a^2+y_a^2+z_a^2}}.$$

10. A unit vector $\bar e$, equally directed with vector $\bar a$, is found by formula

$$\bar e=\frac{\bar a}{|\bar a|}=\bar i\cdot\cos\alpha+\bar j\cdot\cos\beta+\bar k\cdot\cos\gamma.$$

11. Area $S_{\#\bar a,\bar b}$ of a parallelogram, constructed on vectors $\bar a$ and $\bar b$, is calculated by formula: $S_{\#\bar a,\bar b}=\left|[\bar a,\bar b]\right|$. Area S_{ABC} of triangle ABC equals to one half of area $S_{\#\overline{AB},\overline{AC}}$ of a parallelogram, constructed on vectors $\overline{AB}$ and $\overline{AC}$, i.e. $S_{ABC}=\frac{1}{2}\cdot S_{\#\overline{AB},\overline{AC}}$.

12. Volume $V_{\#\bar a,\bar b,\bar c}$ of a parallelepiped, constructed on vectors $\bar a,\bar b,\bar c$, is calculated by formula $V_{\#\bar a,\bar b,\bar c}=\left|(\bar a,\bar b,\bar c)\right|$. Volume V_{ABCD} of triangular pyramid $ABCD$ equals to one sixth of volume $V_{\#\overline{AB},\overline{AC},\overline{AD}}$ of a parallelepiped, constructed on vectors $\overline{AB}$, $\overline{AC}$, $\overline{AD}$, i.e. $V_{ABCD}=\frac{1}{6}\cdot V_{\#\overline{AB},\overline{AC},\overline{AD}}$.

13. A triplet of noncoplanar vectors $\bar{a}, \bar{b}, \bar{c}$ is right (left) if and only if $(\bar{a}, \bar{b}, \bar{c}) > 0$ $((\bar{a}, \bar{b}, \bar{c}) < 0)$.

14. The height h of a parallelogram, constructed on vectors $\bar{a}, \bar{b}$, is calculated by formula (Fig. 8.18, *b*)

$$h = \frac{S_{\#\bar{a},\bar{b}}}{|\bar{a}|} = \frac{|[\bar{a}, \bar{b}]|}{\sqrt{(\bar{a}, \bar{a})}}.$$

15. The height h of a parallelepiped, constructed on vectors $\bar{a}, \bar{b}, \bar{c}$, is calculated by formula

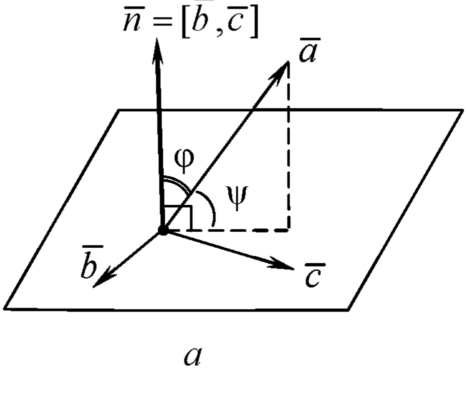
$$h = \frac{V_{\#\bar{a},\bar{b},\bar{c}}}{S_{\#\bar{b},\bar{c}}} = \frac{|(\bar{a}, \bar{b}, \bar{c})|}{|[\bar{b}, \bar{c}]|}.$$

16. The angle ψ between vector $\bar{a}$ and a plane, containing vectors $\bar{b}$ and $\bar{c}$, completes the angle φ between vector $\bar{a}$ and vector $\bar{n} = [\bar{b}, \bar{c}]$ (which is perpendicular to the plane (Fig. 8.21, *a*)) to the right angle, and is calculated by formula

$$\sin\psi = |\cos\varphi| = \frac{|(\bar{a}, \bar{b}, \bar{c})|}{|\bar{a}| \cdot |[\bar{b}, \bar{c}]|}.$$

17. The angle δ between plane, containing vectors $\bar{a}, \bar{b}$ and $\bar{c}, \bar{d}$ accordingly, is calculated as the angle between vectors $\bar{m} = [\bar{a}, \bar{b}]$, $\bar{n} = [\bar{c}, \bar{d}]$, that are perpendicular to the given planes (Fig. 8.21, *b*) by formula

$$\cos\delta = \frac{|([\bar{a}, \bar{b}], [\bar{c}, \bar{d}])|}{|[\bar{a}, \bar{b}]| \cdot |[\bar{c}, \bar{d}]|}.$$



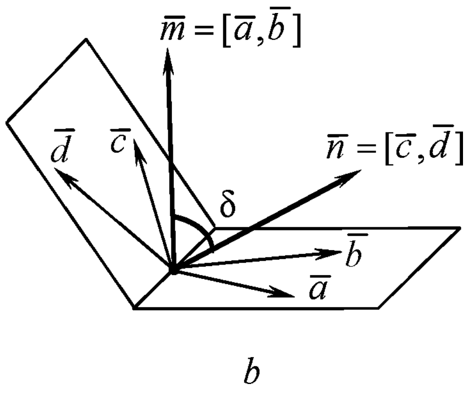


Figure 8.21

Given properties 1–3, 5–11, 14 are also applied to vectors on plane, assuming their applicates equal to zero.

Example 8.11. Triangle OAB is constructed on vectors $\overline{OA} = 4\cdot\overline{i} + 3\cdot\overline{j}$ and $\overline{OB} = 12\cdot\overline{i} - 5\cdot\overline{j}$ (Fig. 8.22). Find:

a) lengths of sides of the triangle;

b) value of angle AOB;

c) area of the triangle;

d) coordinates of vector $\overline{BH}$ (in standard basis), where BH is the height of the triangle.

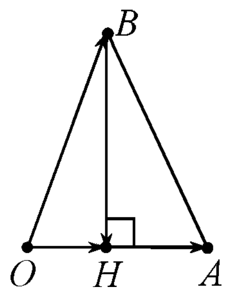


Figure 8.22

□ a) Lengths of sides OA and OB are found by the Property 5:

$$\left|\overline{OA}\right| = \sqrt{(\overline{OA},\overline{OA})} = \sqrt{4^2+3^2} = 5; \quad \left|\overline{OB}\right| = \sqrt{(\overline{OB},\overline{OB})} = \sqrt{12^2+(-5)^2} = 13.$$

To find the length of side AB, obtain coordinates of vector $\overline{AB} = \overline{OB} - \overline{OA} = 12\cdot\overline{i} - 5\cdot\overline{j} - \left(4\cdot\overline{i} + 3\cdot\overline{j}\right) = 8\cdot\overline{i} - 8\cdot\overline{j}$, and then its length:

$$\left|\overline{AB}\right| = \sqrt{(\overline{AB},\overline{AB})} = \sqrt{8^2+(-8)^2} = 8\sqrt{2}.$$

b) Value φ of angle AOB find by Property 6:

$$\cos\varphi = \frac{(\overline{OA},\overline{OB})}{\sqrt{(\overline{OA},\overline{OA})}\cdot\sqrt{(\overline{OB},\overline{OB})}} = \frac{4\cdot12+3\cdot(-5)}{5\cdot13} = \frac{33}{65}.$$

Consequently, $\varphi = \arccos\frac{33}{65}$.

c) Area S of triangle OAB equals to one half of area of a parallelogram, constructed on vectors $\overline{OA}$ and $\overline{OB}$: $S = \frac{1}{2}S_{\#\overline{OA},\overline{OB}}$ (Property 11). To find area of the parallelogram, add zero applicate to vectors $\overline{OA}$ and $\overline{OB}$, i.e. $\overline{OA}=4\cdot\overline{i}+3\cdot\overline{j}+0\cdot\overline{k}$; $\overline{OB}=12\cdot\overline{i}-5\cdot\overline{j}+0\cdot\overline{k}$, and calculate their outer product:

$$[\overline{OA},\overline{OB}]=\begin{vmatrix}\overline{i} & \overline{j} & \overline{k}\\ 4 & 3 & 0\\ 12 & -5 & 0\end{vmatrix}=\overline{i}\cdot\begin{vmatrix}3 & 0\\ -5 & 0\end{vmatrix}-\overline{j}\cdot\begin{vmatrix}4 & 0\\ 12 & 0\end{vmatrix}+\overline{k}\cdot\begin{vmatrix}4 & 3\\ 12 & -5\end{vmatrix}=0\cdot\overline{i}-0\cdot\overline{j}+(-56)\cdot\overline{k}.$$

Then $S_{\#\overline{OA},\overline{OB}}=\left|[\overline{OA},\overline{OB}]\right|=\left|0\cdot\overline{i}+0\cdot\overline{j}+(-56)\cdot\overline{k}\right|=\sqrt{0^2+0^2+(-56)^2}=56$.

So, triangle area $S=\frac{1}{2}\cdot 56=28$.

d) Find vector $\overline{BH}=\overline{OH}-\overline{OB}$. Projection $\overline{OH}$ of vector $\overline{OB}$ on axis, formed by vector $\overline{OA}$, we obtain by Property 8:

$$\overline{OH}=\frac{(\overline{OB},\overline{OA})}{(\overline{OA},\overline{OA})}\cdot\overline{OA}=\frac{12\cdot 4+(-5)\cdot 3}{25}\cdot\left(4\cdot\overline{i}+3\cdot\overline{j}\right)=\frac{132}{25}\cdot\overline{i}+\frac{99}{25}\cdot\overline{j}.$$

From this $\overline{BH}=\frac{132}{25}\cdot\overline{i}+\frac{99}{25}\cdot\overline{j}-\left(12\cdot\overline{i}-5\cdot\overline{j}\right)=-\frac{168}{25}\cdot\overline{i}+\frac{224}{25}\cdot\overline{j}$. Consequently, its coordinates are $-\frac{168}{25}, \frac{224}{25}$. Find the length of this vector, i.e. triangle height:

$\left|\overline{BH}\right|=\sqrt{\left(-\frac{168}{25}\right)^2+\left(\frac{224}{25}\right)^2}=\frac{56}{5}$. Note that triangle area $S=28$, so the height can be calculated by formula $BH=\frac{2\cdot S}{OA}=\frac{2\cdot 28}{5}=\frac{56}{5}$. The results are the same. ■

Example 8.12. Triangular pyramid $OABC$ is constructed on vectors $\overline{OA}=1\cdot\overline{i}+3\cdot\overline{j}-1\cdot\overline{k}$, $\overline{OB}=2\cdot\overline{i}+1\cdot\overline{j}-2\cdot\overline{k}$, $\overline{OC}=3\cdot\overline{i}-2\cdot\overline{j}+4\cdot\overline{k}$ (Fig. 8.23).

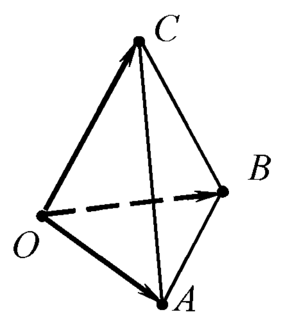


Figure 8.23

Find:

a) lengths of edges OA, OB, OC;

b) value φ of angle AOC;

c) area S_{OAC} of triangle OAC;

d) volume of pyramid $OABC$;

e) height h_B of pyramid, dropped from vertex B;

f) height h_a of triangle OAC, dropped from vertex A;

g) angle ψ between edge OA and the plane of side OBC;

h) value δ between planes of sides OAC and OBC;

i) direction cosines of vector $\overline{OB}$;

j) algebraic value of orthogonal projection of vector $\overline{OA}$ on the direction of vector $\overline{OB}$;

k) orthogonal projection of vector $\overline{OA}$ on line, which is perpendicular to side OBC;

l) unit vector $\overline{e}$ (ort), equally directed with vector $\overline{AB}$;

m) vector $\overline{a}$ with the length equal to the length of vector $\overline{AB}$ and equally directed with vector $\overline{AC}$.

□ a) Lengths of edges OA, OB and OC are calculated by Property 5:

$$|\overline{OA}| = \sqrt{(\overline{OA}, \overline{OA})} = \sqrt{1^2 + 3^2 + (-1)^2} = \sqrt{11};$$

$$|\overline{OB}| = \sqrt{(\overline{OB}, \overline{OB})} = \sqrt{2^2 + 1^2 + (-2)^2} = 3;$$

$$|\overline{OC}| = \sqrt{(\overline{OC}, \overline{OC})} = \sqrt{3^2 + (-2)^2 + 4^2} = \sqrt{29}.$$

b) Value φ of angle AOC is found as the angle between vectors $\overline{OA}$ and $\overline{OC}$ by Property 6:

$$\cos\varphi = \frac{(\overline{OA}, \overline{OC})}{|\overline{OA}| \cdot |\overline{OC}|} = \frac{1 \cdot 3 + 3 \cdot (-2) + (-1) \cdot 4}{\sqrt{11} \cdot \sqrt{29}} = -\frac{7}{\sqrt{319}},$$

i.e. $\varphi = \pi - \arccos\frac{7}{\sqrt{319}}$.

c) First we calculate area of parallelogram, constructed on vectors $\overline{OA}$ and $\overline{OC}$ by Property11. To do this we find outer product

$$[\overline{OA},\overline{OC}]=\begin{vmatrix}\overline{i} & \overline{j} & \overline{k}\\ 1 & 3 & -1\\ 3 & -2 & 4\end{vmatrix}=10\cdot\overline{i}-7\cdot\overline{j}-11\cdot\overline{k},$$

and then its absolute value: $S_{\#\overline{OA},\overline{OC}}=\left|[\overline{OA},\overline{OC}]\right|=\sqrt{10^2+(-7)^2+(-11)^2}=\sqrt{270}$.

Required area of the triangle equals to one half of the previously obtained area: $S_{OAC}=\frac{1}{2}\cdot S_{\#\overline{OA},\overline{OC}}=\frac{\sqrt{270}}{2}$.

d) By Property 12 find the volume $V_{\#\overline{OA},\overline{OB},\overline{OC}}$ of the parallelepiped, constructed on vectors $\overline{OA}$, $\overline{OB}$, $\overline{OC}$:

$$(\overline{OA},\overline{OB},\overline{OC})=\begin{vmatrix}1 & 3 & -1\\ 2 & 1 & -2\\ 3 & -2 & 4\end{vmatrix}=-35 \quad\Rightarrow$$

$$\Rightarrow\quad V_{\#\overline{OA},\overline{OB},\overline{OC}}=\left|(\overline{OA},\overline{OB},\overline{OC})\right|=\left|-35\right|=35.$$

Required pyramid volume is six times smaller: $V_{OABC}=\frac{1}{6}\cdot V_{\#\overline{OA},\overline{OB},\overline{OC}}=\frac{35}{6}$.

e) Height h_B of pyramid is found by Property 15:

$$h_B=\frac{V_{\#\overline{OA},\overline{OB},\overline{OC}}}{S_{\#\overline{OA},\overline{OC}}}=\frac{35}{\sqrt{270}}.$$

f) Height h_a of triangle OAC, dropped from vertex A is found by Property 14:

$$h_a=\frac{S_{\#\overline{OA},\overline{OC}}}{\left|\overline{OC}\right|}=\frac{\sqrt{270}}{\sqrt{29}}.$$

g) At first we obtain vector $\overline{n}$, which is perpendicular to side OBC:

$$\overline{n}=[\overline{OB},\overline{OC}]=\begin{vmatrix}\overline{i} & \overline{j} & \overline{k}\\ 2 & 1 & -2\\ 3 & -2 & 4\end{vmatrix}=0\cdot\overline{i}-14\cdot\overline{j}-7\cdot\overline{k}.$$

Then we calculate angle ψ between vector $\overline{OA}$ and the plane of side OBC by the Property 16:

$$\sin\psi = \frac{\left|(\overline{OA},\overline{OB},\overline{OC})\right|}{\left|\overline{OA}\right|\cdot\left|[\overline{OB},\overline{OC}]\right|} = \frac{\left|(\overline{OA},\overline{n})\right|}{\left|\overline{OA}\right|\cdot\left|\overline{n}\right|} =$$

$$= \frac{\left|1\cdot 0 + 3\cdot(-14) + (-1)\cdot(-7)\right|}{\sqrt{11}\cdot\sqrt{(-14)^2+(-7)^2}} = \frac{35}{\sqrt{11}\cdot 7\cdot\sqrt{5}} = \frac{\sqrt{5}}{\sqrt{11}},$$

i.e. $\psi = \arcsin\frac{\sqrt{5}}{\sqrt{11}}$.

h) Find vector $\overline{m}$, which is perpendicular to the plane of side OAC:

$$\overline{m} = [\overline{OA},\overline{OC}] = \begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ 1 & 3 & -1 \\ 3 & -2 & 4 \end{vmatrix} = 10\cdot\overline{i} - 7\cdot\overline{j} - 11\cdot\overline{k}.$$

Vector $\overline{n}$, which is perpendicular to side OBC, was found in "g". Required angle δ is calculated by Property 17:

$$\cos\delta = \frac{\left|([\overline{OA},\overline{OC}],[\overline{OB},\overline{OC}])\right|}{\left|[\overline{OA},\overline{OC}]\right|\cdot\left|[\overline{OB},\overline{OC}]\right|} = \frac{\left|(\overline{m},\overline{n})\right|}{\left|\overline{m}\right|\cdot\left|\overline{n}\right|} =$$

$$= \frac{\left|10\cdot 0 + (-7)\cdot(-14) + (-11)\cdot(-7)\right|}{\sqrt{10^2+(-7)^2+(-11)^2}\cdot 7\cdot\sqrt{5}} = \frac{5}{3\sqrt{6}},$$

i.e. $\delta = \arccos\frac{5}{3\sqrt{6}}$.

i) Direction cosines of vector $\overline{OB}$ are calculated by Property 9:

$$\cos\alpha = \frac{(\overline{OB},\overline{i})}{\left|\overline{OB}\right|} = \frac{2\cdot 1 + 1\cdot 0 + (-2)\cdot 0}{\sqrt{2^2+1^2+(-2)^2}} = \frac{2}{3};$$

$$\cos\beta = \frac{(\overline{OB},\overline{j})}{\left|\overline{OB}\right|} = \frac{1}{3}; \qquad \cos\gamma = \frac{(\overline{OB},\overline{k})}{\left|\overline{OB}\right|} = \frac{-2}{3}.$$

Note, that $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = \left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 = 1$.

j) Algebraic value $proj_{\overline{OB}}\,\overline{OA}$ of projection length is found by Property 7 ($\bar{a} = \overline{OA}$, $\bar{b} = \overline{OB}$):

$$proj_{\overline{OB}}\,\overline{OA} = \frac{(\overline{OA}, \overline{OB})}{\left|\overline{OB}\right|} = \frac{1\cdot 2 + 3\cdot 1 + (-1)\cdot(-2)}{3} = \frac{7}{3}.$$

k) Required orthogonal projection $\overline{proj_{\bar{n}}\,\overline{OA}}$ is obtained by Property 8 ($\bar{a} = \overline{OA}$, $\bar{b} = \bar{n}$), using vector $\bar{n}$, which was found in "g":

$$\overline{proj_{\bar{n}}\,\overline{OA}} = \frac{(\overline{OA}, \bar{n})}{(\bar{n}, \bar{n})}\cdot\bar{n} = \frac{1\cdot 0 + 3\cdot(-14) + (-1)\cdot(-7)}{0^2 + (-14)^2 + (-7)^2}\cdot\left(0\cdot\bar{i} + (-14)\cdot\bar{j} + (-7)\cdot\bar{k}\right) =$$

$$= 2\cdot\bar{j} + \bar{k}.$$

l) Obtain coordinates of vector $\overline{AB}$ and its length:

$$\overline{AB} = \overline{OB} - \overline{OA} = \left(2\cdot\bar{i} + 1\cdot\bar{j} - 2\cdot\bar{k}\right) - \left(1\cdot\bar{i} + 3\cdot\bar{j} - 1\cdot\bar{k}\right) = 1\cdot\bar{i} + 4\cdot\bar{j} - 1\cdot\bar{k};$$

$$\left|\overline{AB}\right| = \sqrt{1^2 + 4^2 + (-1)^2} = 3\sqrt{2},$$

and then the required vector $\bar{e} = \frac{\overline{AB}}{\left|\overline{AB}\right|} = \frac{1}{3\sqrt{2}}\cdot\bar{i} + \frac{4}{3\sqrt{2}}\cdot\bar{j} - \frac{1}{3\sqrt{2}}\cdot\bar{k}$.

m) Find coordinates of vector $\overline{AC}$ and its length:

$$\overline{AC} = \overline{OC} - \overline{OA} = \left(3\cdot\bar{i} - 2\cdot\bar{j} + 4\cdot\bar{k}\right) - \left(1\cdot\bar{i} + 3\cdot\bar{j} - 1\cdot\bar{k}\right) = 2\cdot\bar{i} - 5\cdot\bar{j} + 5\cdot\bar{k};$$

$$\left|\overline{AC}\right| = \sqrt{2^2 + (-5)^2 + 5^2} = 3\sqrt{6},$$

and then the required vector

$$\bar{a} = \frac{\left|\overline{AB}\right|}{\left|\overline{AC}\right|}\cdot\overline{AC} = \frac{3\sqrt{2}}{3\sqrt{6}}\cdot\left(2\cdot\bar{i} - 5\cdot\bar{j} + 5\cdot\bar{k}\right) = \frac{2}{\sqrt{3}}\cdot\bar{i} - \frac{5}{\sqrt{3}}\cdot\bar{j} + \frac{5}{\sqrt{3}}\cdot\bar{k}. \blacksquare$$

EXERCISES

1. Consider vectors $\overline{a} = 2\cdot\overline{i} - n\cdot\overline{j}$; $\overline{b} = m\cdot\overline{i} + 3\cdot\overline{j}$. Decompose vector $\overline{i}$ by vectors $\overline{a}$ and $\overline{b}$. Find:

a) coordinates of vector $\overline{c} = 2\cdot\overline{a} - 3\cdot\overline{b}$ in standard basis;

b) length and direction cosines of vector $\overline{c}$.

2. Consider vectors $\overline{a} = n\cdot\overline{i} - \overline{j} + m\cdot\overline{k}$; $\overline{b} = 2\cdot\overline{i} + m\cdot\overline{j} + \overline{k}$; $\overline{c} = 4\cdot\overline{i} + +n\cdot\overline{j} - 3\cdot\overline{k}$. Decompose vector $\overline{i}$ by vectors $\overline{a}, \overline{b}, \overline{c}$. Find:

a) coordinates of vector $\overline{d} = \overline{a} + 2\cdot\overline{b} - 3\cdot\overline{c}$ in standard basis;

b) length and direction cosines of vector $\overline{d}$.

3. Consider vectors $\overline{a} = 2\cdot\overline{i} - n\cdot\overline{j}$; $\overline{b} = m\cdot\overline{i} + 3\cdot\overline{j}$. Decompose vector $\overline{i}$ by vectors $\overline{a}$ and $\overline{b}$. Find:

a) products $(\overline{a},\overline{b})$, $(\overline{a},\overline{a})$, $(\overline{b},\overline{b})$;

b) orthogonal projections $\overline{proj_{\overline{b}}\overline{a}}$, $\overline{proj_{\overline{c}}\overline{a}}$ of vector $\overline{a}$;

c) algebraic values $proj_{\overline{b}}\overline{a}$ and $proj_{\overline{a}}\overline{b}$ of orthogonal projection lengths;

d) angle φ between vectors $\overline{a}$ and $\overline{b}$;

e) area of parallelogram $S_{\#\overline{a},\overline{b}}$, which is built on vectors $\overline{a}$ and $\overline{b}$.

4. Consider vectors $\overline{a} = n\cdot\overline{i} - \overline{j} + m\cdot\overline{k}$; $\overline{b} = 2\cdot\overline{i} + m\cdot\overline{j} + \overline{k}$; $\overline{c} = 4\cdot\overline{i} + +n\cdot\overline{j} - 3\cdot\overline{k}$. Decompose vector $\overline{i}$ by vectors $\overline{a}, \overline{b}, \overline{c}$. Find:

a) products $(\overline{a},\overline{b})$, $[\overline{a},\overline{b}]$, $(\overline{a},\overline{b},\overline{c})$, determine orientation of the triplet $\overline{a}, \overline{b}, \overline{c}$;

b) orthogonal projections $\overline{proj_{\overline{b}}\overline{a}}$, $\overline{proj_{\overline{c}}\overline{a}}$ of vector $\overline{a}$;

c) algebraic values $proj_{\overline{b}}\overline{a}$ and $proj_{\overline{a}}\overline{b}$ of orthogonal projection lengths;

d) angle φ between vectors $\overline{a}$ and $\overline{b}$;

e) angle ψ between vector $\overline{a}$ and plane, which contains vectors $\overline{b}$ and $\overline{c}$;

f) area of parallelogram $S_{\#\overline{a},\overline{b}}$, which is built on vectors $\overline{a}$ and $\overline{b}$;

g) volume of parallelepiped $V_{\#\overline{a},\overline{b},\overline{c}}$, which is built on vectors $\overline{a}, \overline{b}, \overline{c}$.

CHAPTER 9. COORDINATE SYSTEMS

9.1. CARTESIAN COORDINATE SYSTEMS

9.1.1. Cartesian Coordinates of Vectors and Points

Let O be a fixed point in space. An set of point O and a basis is called an ***affine coordinate system*** and point O is called its ***origin***. Lines passing through the origin in the direction of basis vectors are called ***coordinate axes***.

For any point A in a given affine coordinate system we can consider vector $\overline{OA}$, its tail being the origin and head – point A (Fig. 9.1–9.3). This vector is called a ***position*** or ***radius vector*** of point A.

An affine coordinate system is called ***Cartesian (rectangular)*** if its basis is orthonormal (see Section 8.3.5).

The coordinates of a vector in a Cartesian coordinate system are the coefficients of its decomposition by standard basis (see Section 8.3.5).

The coordinates of a point A ***in a Cartesian coordinate system*** are the coordinates of its position vector $\overline{OA}$ in standard basis. In space these are coefficients x, y, z of decomposition $\overline{OA} = x \cdot \overline{i} + y \cdot \overline{j} + z \cdot \overline{k}$, on plane – coefficients x, y of decomposition $\overline{OA} = x \cdot \overline{i} + y \cdot \overline{j}$, on line – coefficient x of decomposition $\overline{OA} = x \cdot \overline{i}$. Denotations $A(x,y,z)$, $A(x,y)$, $A(x)$ are used, respectively. Cartesian coordinates of a point (or its position vector) can be represented by a coordinate column:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ in space,} \qquad \begin{pmatrix} x \\ y \end{pmatrix} \text{ on plane.}$$

Choosing standards bases (see Section 8.3.5), we obtain:

$O\overline{i}$ – ***Cartesian coordinate system on a line*** – is represented by point O and unit vector $\overline{i}$ on a line. Points O and A (Fig. 9.1) on axis Ox are denoted by $O(0)$ and $A(1)$;

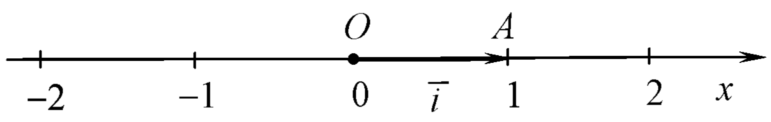


Figure 9.1

$O\bar{i}\,\bar{j}$ – ***Cartesian coordinate system on a plane*** – is represented by point O and two mutually perpendicular unit vectors $\bar{i}$ and $\bar{j}$ on a plane (vector $\bar{i}$ is the first basis vector and $\bar{j}$ is the second one; $\bar{i}$, $\bar{j}$ is the right pair of vectors). Axes Ox (abscissa) and Oy (ordinate) divide the plane into 4 parts, called ***quadrants*** (Fig. 9.2), e.g. point $A(1,1)$ belongs to the I quadrant;

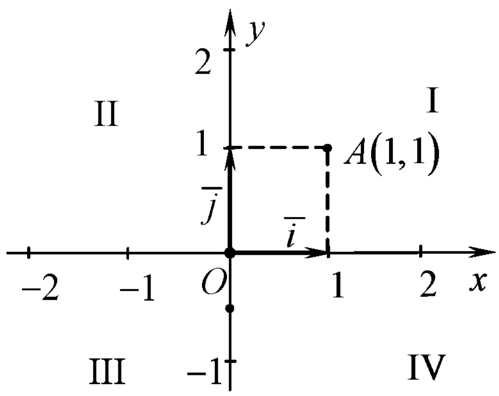


Figure 9.2

Coordinates of vectors and points in a Cartesian coordinate system are called ***Cartesian coordinates***.

$O\bar{i}\,\bar{j}\,\bar{k}$ – ***Cartesian coordinate system in space*** – is represented by point O and three pairwise perpendicular unit vectors $\bar{i},\bar{j},\bar{k}$ (vector $\bar{i}$ is the first basis vector, $\bar{j}$ is the second and $\bar{k}$ is the third one; $\bar{i},\bar{j},\bar{k}$ is the right triplet of vectors). Axes are denoted by Ox – abscissa, Oy – ordinate, Oz – applicate.

Coordinate planes Oxy, Oxz, Oyz, passing through pairs of axes, divide space into 8 ***octants*** (Fig. 9.3), e.g. point $A(1,2,2)$ belongs to the I octant.

Cartesian coordinate systems can also be denoted by the origin and the axes, e.g. Ox, Oxy, $Oxyz$.

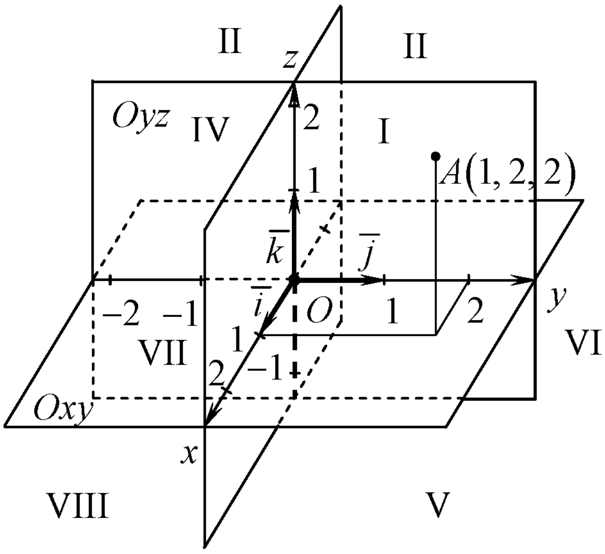


Figure 9.3

To find the coordinates of a vector $\overline{AB}$ *with the tail in the point* $A(x_A, y_A, z_A)$ *and the head in the point* $B(x_B, y_B, z_B)$, *we should subtract the coordinates of its tail from the corresponding coordinates of its head*:

$$\overline{AB} = (x_B - x_A)\cdot\overline{i} + (y_B - y_A)\cdot\overline{j} + (z_B - z_A)\cdot\overline{k}\,.$$

This rule also holds for Cartesian coordinate systems on a plane and on a line.

Coordinates of a point M *that divides a segment* AB *in the ratio of* $\dfrac{AM}{MB} = \dfrac{\beta}{\alpha}$ ($\alpha > 0$, $\beta > 0$), *are found by the coordinates of its endpoints* $A(x_A, y_A, z_A)$ *and* $B(x_B, y_B, z_B)$ (see Section 2.1.1):

$$M\left(\frac{\alpha\cdot x_A + \beta\cdot x_B}{\alpha+\beta}; \frac{\alpha\cdot y_A + \beta\cdot y_B}{\alpha+\beta}; \frac{\alpha\cdot z_A + \beta\cdot z_B}{\alpha+\beta}\right). \tag{9.1}$$

In particular:

- point $M\left(\dfrac{x_A + x_B}{2}; \dfrac{y_A + y_B}{2}; \dfrac{z_A + z_B}{2}\right)$ is the midpoint of a segment AB;

- point $M\left(\frac{x_A+x_B+x_C}{3};\frac{y_A+y_B+y_C}{3};\frac{z_A+z_B+z_C}{3}\right)$ is the intersection point of triangle ABC's medians .

Similar formulas are true for coordinates of points on a plane and on a line.

In a Cartesian coordinate system the distance AB between points $A(x_A, y_A, z_A)$ and $B(x_B, y_B, z_B)$ is obtained by the formula

$$AB=\sqrt{(x_B-x_A)^2+(y_B-y_A)^2+(z_B-z_A)^2}\,. \tag{9.2}$$

For the coordinate plane and the coordinate line, respectively:

$$AB=\sqrt{(x_B-x_A)^2+(y_B-y_A)^2}\,;\qquad AB=|\,x_B-x_A\,|.$$

If Cartesian coordinates of vertexes $A(x_A, y_A)$, $B(x_B, y_B)$, $C(x_C, y_C)$ of triangle ABC on a plane are given, its area is calculated by the formula $S_{ABC}=|\,S^{\wedge}_{ABC}\,|$, where

$$S^{\wedge}_{ABC}=\frac{1}{2}\cdot\begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_C & y_C & 1 \end{vmatrix}. \tag{9.3}$$

If Cartesian coordinates of vertexes $A(x_A, y_A, z_A)$, $B(x_B, y_B, z_B)$, $C(x_C, y_C, z_C)$, $D(x_D, y_D, z_D)$ of triangular pyramid $ABCD$ are given, its volume is calculated by the formula $V_{ABCD}=|\,V^{\wedge}_{ABCD}\,|$, where

$$V^{\wedge}_{ABCD}=\frac{1}{6}\cdot\begin{vmatrix} x_A & y_A & z_A & 1 \\ x_B & y_B & z_B & 1 \\ x_C & y_C & z_C & 1 \\ x_D & y_D & z_D & 1 \end{vmatrix}. \tag{9.4}$$

Example 9.1. Given Cartesian coordinates of vertexes $A(1,1)$, $B(4,5)$, $C(13,6)$ of the triangle ABC (see Figure 9.4), find

a) the length of the median AM;

b) the length of the angle bisector AL;

c) the height h_a dropped from the vertex A.

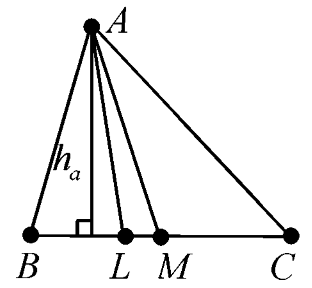


Figure 9.4

☐ a) By formula (9.1) calculate the coordinates of the point M – the midpoint of the side BC: $M\left(\frac{4+13}{2}, \frac{5+6}{2}\right)$, i.e. $M\left(\frac{17}{2}, \frac{11}{2}\right)$. Using a special case of formula (9.2) for a plane, compute the length of the median:

$$AM = \sqrt{\left(\frac{17}{2} - 1\right)^2 + \left(\frac{11}{2} - 1\right)^2} = \frac{\sqrt{306}}{2}.$$

b) Calculate the coordinates of the point L that divides the side BC in the ratio $BL:LC = AB:AC$ (the angle bisector theorem). Since $AB = \sqrt{(4-1)^2 + (5-1)^2} = 5$ and $AC = \sqrt{(13-1)^2 + (6-1)^2} = 13$, by formula (9.1), taking into account that $BL:LC = 5:13 \;\Rightarrow\; \alpha = 13,\; \beta = 5$, we find $L\left(\frac{13\cdot 4 + 5\cdot 13}{13+5}, \frac{13\cdot 5 + 5\cdot 6}{13+5}\right)$, i.e. $L\left(\frac{13}{2}, \frac{95}{18}\right)$. Compute the length of the angle bisector:

$$AL = \sqrt{\left(\frac{13}{2} - 1\right)^2 + \left(\frac{95}{18} - 1\right)^2} = \frac{11\cdot\sqrt{130}}{18}.$$

c) By formula (9.3) find: $S_{ABC}^{\wedge} = \frac{1}{2}\cdot \begin{vmatrix} 1 & 1 & 1 \\ 4 & 5 & 1 \\ 13 & 6 & 1 \end{vmatrix} = -\frac{33}{2}.$

Hence, the area of the triangle ABC $S_{ABC} = \left| \hat{S}_{ABC} \right| = \frac{33}{2}$, then $h_a = \frac{2 \cdot S_{ABC}}{BC} = \frac{33}{\sqrt{82}}$, since $BC = \sqrt{(13-4)^2 + (6-5)^2} = \sqrt{82}$. ■

Example 9.2. Given Cartesian coordinates of vertexes $A(1,1,3)$, $B(3,5,4)$, $C(-1,3,2)$, $D(5,3,-1)$ of the triangular pyramid $ABCD$, find:

a) the length of the segment DM connecting the vertex D of the pyramid and the point M of intersection of medians of the face ABC;

b) the volume V_{ABCD} of the pyramid.

□ a) Find the coordinates of the point M (the intersection of medians of the triangle ABC) by using a special case of formula (9.1):

$$M\left(\frac{1+3+(-1)}{3}; \frac{1+5+3}{3}; \frac{3+4+2}{3}\right), \text{ i.e. } M(1,3,3).$$

By formula (9.2) calculate

$$DM = \sqrt{(1-5)^2 + (3-3)^2 + (3+1)^2} = 4\sqrt{2}.$$

b) Find the volume of the pyramid $ABCD$. By formula (9.4), subtracting the first row from the others and expanding the determinant across the last row (see Section 2.2), we obtain

$$\hat{V}_{ABCD} = \frac{1}{6}\begin{vmatrix} 1 & 1 & 3 & 1 \\ 3 & 5 & 4 & 1 \\ -1 & 3 & 2 & 1 \\ 5 & 3 & -1 & 1 \end{vmatrix} = \frac{1}{6}\begin{vmatrix} 1 & 1 & 3 & 1 \\ 2 & 4 & 1 & 0 \\ -2 & 2 & -1 & 0 \\ 4 & 2 & -4 & 0 \end{vmatrix} = \frac{1}{6} \cdot (-1)^{1+4} \cdot 1 \cdot \begin{vmatrix} 2 & 4 & 1 \\ -2 & 2 & -1 \\ 4 & 2 & -4 \end{vmatrix} =$$

$$= -\frac{1}{6} \cdot (-16 - 16 - 4 - 8 - 32 + 4) = 12.$$

Hence, $V_{ABCD} = \left| \hat{V}_{ABCD} \right| = 12$. ■

9.1.2. Cartesian Coordinate Transformations on Plane and in Space

Let's give formulas connecting coordinates of a point during the transition from one Cartesian coordinate system to another one. Consider three types of transformations:

a) *translation*;

b) *rotation*;

c) *reflection across abscissa* (changing direction of the ordinate axis to the opposite one).

Coordinates x, y of a point in the old coordinate system $O\overline{i}\,\overline{j}$ and coordinates x', y' in the new coordinate system $O'\overline{i}'\overline{j}'$ are related to one another by the following formulas:

a) After the translation of a coordinate system (Fig. 9.5, *a*) by a vector $\overline{s} = \overline{OO'} = x_s \cdot \overline{i} + y_s \cdot \overline{j}$:

$$\begin{cases} x = x_s + x', \\ y = y_s + y'. \end{cases}$$

b) After the rotation of a coordinate system by an angle φ (Fig. 9.5, *b*):

$$\begin{cases} x = x' \cdot \cos\varphi - y' \cdot \sin\varphi, \\ y = x' \cdot \sin\varphi + y' \cdot \cos\varphi. \end{cases}$$

c) After the reflection across abscissa (changing direction of the ordinate axis to the opposite one) (see Figure 9.5, *c)*:

$$\begin{cases} x = x', \\ y = -y'. \end{cases}$$

Any transformation of a Cartesian coordinate system on a plane can be reduced to a composition of transformations, each of them being a translation, rotation or reflection across an axis.

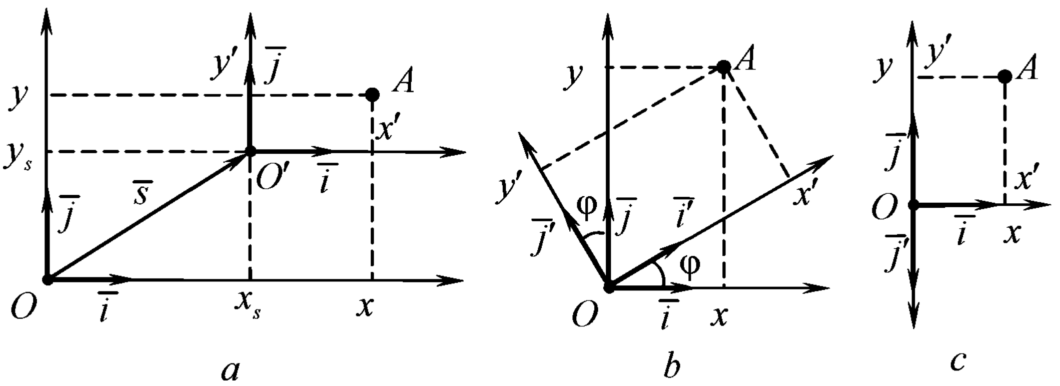


Figure 9.5

Suppose we have two Cartesian coordinate systems on a plane: $O\bar{i}\,\bar{j}$ and $O'\bar{i}'\bar{j}'$. Formulas, connecting old and new coordinates of a point, take the form:

- for coordinates systems with the same orientation (i.e. transitions between right and right or left and left coordinate systems) (Fig. 9.6, *a*):

$$\begin{cases} x = x_s + x' \cdot \cos\varphi - y' \cdot \sin\varphi, \\ y = y_s + x' \cdot \sin\varphi + y' \cdot \cos\varphi, \end{cases} \tag{9.5}$$

- for coordinates systems with different orientations (Fig. 9.6, *b*):

$$\begin{cases} x = x_s + x' \cdot \cos\varphi + y' \cdot \sin\varphi, \\ y = y_s + x' \cdot \sin\varphi - y' \cdot \cos\varphi. \end{cases} \tag{9.6}$$

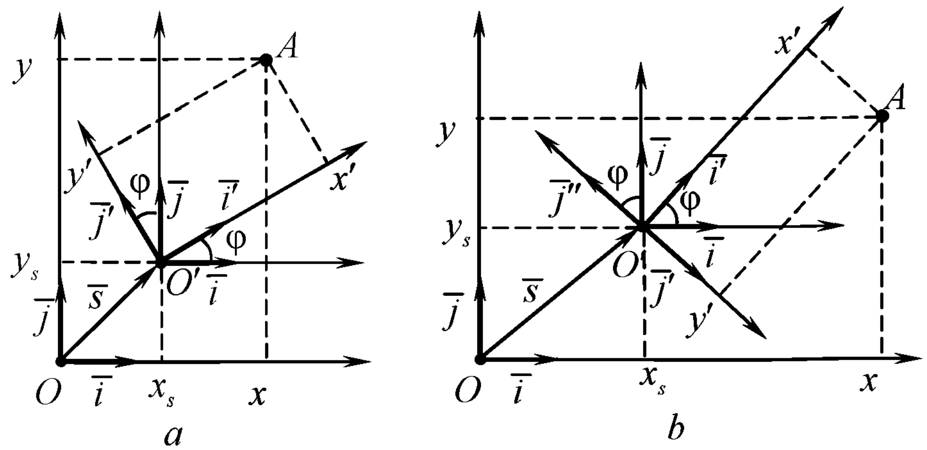


Figure 9.6

For the above transformations of point coordinates, new coordinates are expressed via the old ones by the following formulas:

a) $\begin{cases} x' = x - x_s, \\ y' = y - y_s, \end{cases}$ b) $\begin{cases} x' = x \cdot \cos\varphi + y \cdot \sin\varphi, \\ y' = -x \cdot \sin\varphi + y \cdot \cos\varphi, \end{cases}$ c) $\begin{cases} x' = x, \\ y' = -y. \end{cases}$

For transformation (9.5) similar formulas take the form:

$$\begin{cases} x' = (x - x_s)\cdot\cos\varphi + (y - y_s)\cdot\sin\varphi, \\ y' = -(x - x_s)\cdot\sin\varphi + (y - y_s)\cdot\cos\varphi. \end{cases} \qquad (9.7)$$

For $x_s = 0, y_s = 0$ and $\varphi = \frac{\pi}{2}$ from formula (9.6) we obtain the transformation

$$\begin{cases} x = y', \\ y = x', \end{cases}$$

that changes the names of axes (reflection across the line containing the bisector of the first coordinate angle).

Transformations of Cartesian Coordinates in Space

Consider three types of transformations of a Cartesian coordinate system:

a) *translation*;

b) *rotation around an axis*;

c) *reflection on a plane* (changing direction of one axis to the opposite one).

Coordinates x, y, z of a point in the old coordinate system $O\bar{i}\,\bar{j}\,\bar{k}$ and coordinates x', y', z' in the new coordinate system $O'\bar{i}'\bar{j}'\bar{k}'$ are related to one another by the following formulas

a) After the translation of a coordinate system by an origin translation vector $\bar{s} = \overline{OO'} = x_s\cdot\bar{i} + y_s\cdot\bar{j} + z_s\cdot\bar{k}$:

$$\begin{cases} x = x_s + x', \\ y = y_s + y', \\ z = z_s + z'. \end{cases}$$

b) After the rotation of the coordinate system by an angle φ around the applicate axis:

$$\begin{cases} x = x'\cdot\cos\varphi - y'\cdot\sin\varphi, \\ y = x'\cdot\sin\varphi + y'\cdot\cos\varphi, \\ z = z'. \end{cases}$$

It's obvious that a coordinate system on the plane Oxy is rotated by an angle φ during this transformation.

c) After the reflection on the plane Oxy (changing direction of the applicate axis to the opposite one):

$$\begin{cases} x = x', \\ y = y', \\ z = -z'. \end{cases}$$

Reflections on other coordinate planes are defined similarly (changing direction of the abscissa or ordinate axis to the opposite one).

Any transformation of a Cartesian coordinate system in space can be reduced to a composition of transformations, each of them being a translation, rotation around an axis or reflection on a coordinate plane.

In particular, for a composition of rotation by an angle φ around the Oz axis and translation by a vector $\overline{s} = \overline{OO'} = x_s \cdot \overline{i} + y_s \cdot \overline{j} + z_s \cdot \overline{k}$ coordinate transformation formulas take the form:

$$\begin{cases} x = x_s + x' \cdot \cos\varphi - y' \cdot \sin\varphi, \\ y = y_s + x' \cdot \sin\varphi + y' \cdot \cos\varphi, \\ \qquad z = z_s + z'. \end{cases} \tag{9.8}$$

Formulas for expressing new coordinates of points via the old ones take the form:

$$\begin{cases} x' = (x - x_s) \cdot \cos\varphi + (y - y_s) \cdot \sin\varphi, \\ y' = -(x - x_s) \cdot \sin\varphi + (y - y_s) \cdot \cos\varphi, \\ \qquad z' = z - z_s. \end{cases} \tag{9.9}$$

Similar formulas can be written for other compositions of transformations, e.g. to obtain formulas of coordinate transformations for *a composition of rotation by an angle* φ *around the abscissa axis and translation by a vector* $\overline{s} = \overline{OO'} = x_s \cdot \overline{i} + y_s \cdot \overline{j} + z_s \cdot \overline{k}$, we should write formulas (9.8) or (9.9), making a cyclic interchange of letters x to y, y to z, z to x:

$$\begin{cases} \qquad x = x_s + x', \\ y = y_s + y' \cdot \cos\varphi - z' \cdot \sin\varphi, \\ z = z_s + y' \cdot \sin\varphi + z' \cdot \cos\varphi \end{cases} \quad \text{or} \quad \begin{cases} \qquad x' = x - x_s, \\ y' = (y - y_s) \cdot \cos\varphi + (z - z_s) \cdot \sin\varphi, \\ z' = -(y - y_s) \cdot \sin\varphi + (z - z_s) \cdot \cos\varphi. \end{cases} \tag{9.10}$$

Example 9.3. The point A in old coordinate system $O\overline{i}\,\overline{j}$ has coordinates $x=3$, $y=4$. New Cartesian coordinate system $O'\overline{i}'\overline{j}'$ is obtained from the old one by transition by the vector $\overline{s}=2\cdot\overline{i}+\overline{j}$ and rotation by the angle $\varphi=\frac{\pi}{3}$. Find coordinates of the point $A(x',y')$ in new coordinate system.

□ Since $x_s=2$, $y_s=1$, by formulas (9.7) we obtain:

$$x'=(3-2)\cdot\cos\tfrac{\pi}{3}+(4-1)\cdot\sin\tfrac{\pi}{3}=\frac{1}{2}+\frac{3\sqrt{3}}{2}=\frac{1+3\sqrt{3}}{2};$$

$$y'=-(3-2)\cdot\sin\tfrac{\pi}{3}+(4-1)\cdot\cos\tfrac{\pi}{3}=-\frac{\sqrt{3}}{2}+\frac{3}{2}=\frac{3-\sqrt{3}}{2}.\ \blacksquare$$

Example 9.4. The point A in old coordinate system $O\overline{i}\,\overline{j}\,\overline{k}$ has coordinates $x=3$, $y=4$, $z=5$. New Cartesian coordinate system $O'\overline{i}'\overline{j}'\overline{k}'$ is obtained from the old one by transition by the vector $\overline{s}=2\cdot\overline{i}+3\cdot\overline{j}+\overline{k}$ and rotation by the angle $\varphi=\frac{\pi}{3}$ around the abscissa. Find coordinates of the point $A(x',y',z')$ in new coordinate system.

□ Since $x_s=2$, $y_s=3$, $z_s=1$, by formulas (9.10) obtain:

$$x'=3-2=1;$$

$$y'=(4-3)\cdot\cos\tfrac{\pi}{3}+(5-1)\cdot\sin\tfrac{\pi}{3}=\frac{1}{2}+\frac{4\sqrt{3}}{2}=\frac{1+4\sqrt{3}}{2};$$

$$z'=-(4-3)\cdot\sin\tfrac{\pi}{3}+(5-1)\cdot\cos\tfrac{\pi}{3}=-\frac{\sqrt{3}}{2}+\frac{4}{2}=\frac{4-\sqrt{3}}{2}.\ \blacksquare$$

9.2. POLAR COORDINATE SYSTEM

A polar coordinate system on plane is an aggregate of point O, called the ***pole***, and ray Ox, called the ***polar axis***. Also a ***scale interval*** is given to measure distances from points on a plane to the pole. As a rule, vector $\overline{i}$ on a polar axis, applied to point O, is chosen and its length is taken as range of the scale interval, while its direction specifies positive direction on the polar axis (Fig. 9.7, *a*).

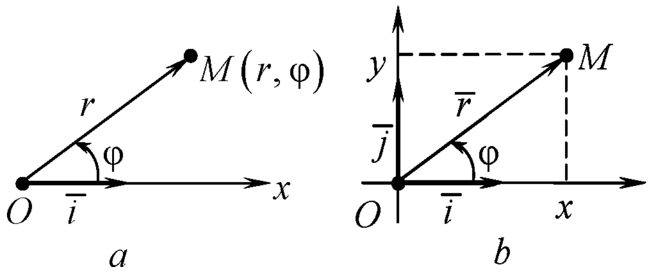


Figure 9.7

The position of point M in a polar coordinate system is determined by the distance r (***radius***) from point M to the pole, i.e. $r = |\overline{OM}|$, and the angle φ (***polar angle***, or ***azimuth***) between the polar axis and vector $\overline{OM}$. Radius and polar angle make ***polar coordinates*** of point M, written as $M(r, \varphi)$. Polar angle is expressed in radians and is measured from the polar axis:

- in positive direction (counterclockwise) if the angle value is positive;
- in negative direction (clockwise) is the angle value is negative.

Radius is defined for any point on a plane and takes on non-negative values $r \geq 0$. Polar angle φ is defined for any point on a plane, except for the pole O, and takes on values $-\pi < \varphi \leq \pi$, called ***principal values of the polar angle***.

A polar coordinates system $Or\varphi$ can be associated with a Cartesian coordinate system $O\bar{i}\,\bar{j}$, origin O of which coincides with the pole, and the abscissa axis (more exactly, positive abscissa semi-axis) – with the polar axis. The ordinate axis is added perpendicular to the abscissa axis so that a Cartesian coordinate system is obtained (Fig. 9.7, *b*). Lengths of basis vectors are determined by the scale interval on the polar axis.

Vice versa, if a right-handed Cartesian system is given on plane, we can obtain a polar coordinate system (associated with the given Cartesian one) by assuming positive abscissa semi-axis as the polar axis.

Let's give formulas for converting polar coordinates r, φ of point M, not the same as point O, to Cartesian coordinates x, y. By Figure 9.7, *b* we obtain:

$$\begin{cases} x = r \cdot \cos\varphi, \\ y = r \cdot \sin\varphi. \end{cases} \qquad (9.11)$$

These formulas allow us to find *Cartesian coordinates* by given *polar coordinates*.

The reverse conversion is performed by the formulas:

$$\begin{cases} r = \sqrt{x^2 + y^2}\,, \\ \cos\varphi = \dfrac{x}{r} = \dfrac{x}{\sqrt{x^2 + y^2}}, \\ \sin\varphi = \dfrac{y}{r} = \dfrac{y}{\sqrt{x^2 + y^2}}. \end{cases} \tag{9.12}$$

Two last equalities give the polar angle with accuracy up to summands $2\pi n$, where $n \in \mathbb{Z}$. For $x \neq 0$ it follows that $\tan\varphi = \dfrac{y}{x}$. The principal value of the polar angle φ $(-\pi < \varphi \le \pi)$ is found by the formulas (see Figure 9.8):

$$\varphi = \begin{cases} \arctan\dfrac{y}{x}, & x > 0\,, \\ \pi + \arctan\dfrac{y}{x}, & x < 0,\, y \ge 0, \\ -\pi + \arctan\dfrac{y}{x}, & x < 0,\, y < 0, \\ \dfrac{\pi}{2}, & x = 0,\, y > 0\,, \\ -\dfrac{\pi}{2}, & x = 0,\, y < 0\,. \end{cases}$$

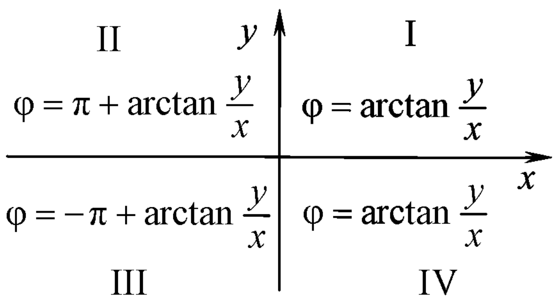


Figure 9.8

The principal value of the polar angle can be chosen differently, i.e. $0 \le \varphi < 2\pi$.

Example 9.5. In the polar coordinate system $Or\varphi$:

a) sketch ***coordinate lines*** $r=1$, $r=2$, $r=3$, $\varphi=\frac{\pi}{4}$, $\varphi=\frac{\pi}{2}$, $\varphi=\frac{3\pi}{4}$;

b) plot points M_1, M_2 with polar coordinates $r_1=3$, $\varphi_1=\frac{9\pi}{4}$, $r_2=3$, $\varphi_2=-\frac{7\pi}{4}$. Find principal values of these points' polar angles;

c) find Cartesian coordinates of points M_1, M_2;

d) find polar coordinates of point A, given its Cartesian coordinates $A(-3,4)$.

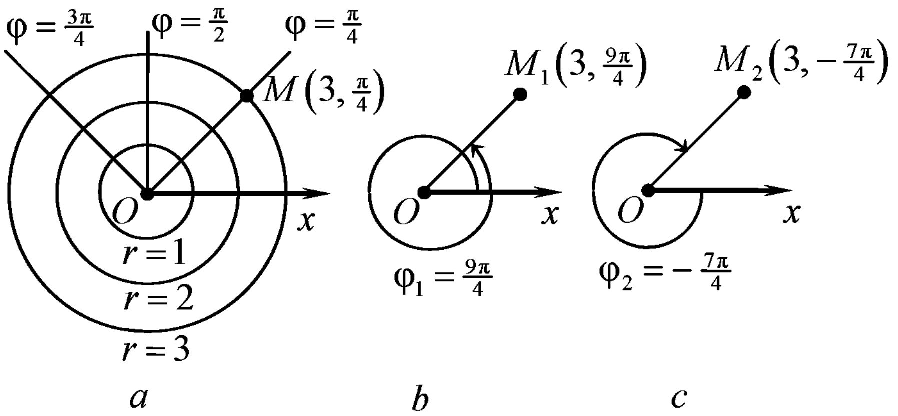


Figure 9.9

□ a) Coordinate lines $r=1$, $r=2$, $r=3$ are circles of respective radiuses, and lines $\varphi=\frac{\pi}{4}$, $\varphi=\frac{\pi}{2}$, $\varphi=\frac{3\pi}{4}$ are rays (Fig. 9.9,*a*).

b) Let's plot points $M_1\left(3,\frac{9\pi}{4}\right)$ and $M_2\left(3,-\frac{7\pi}{4}\right)$ (Fig. 9.9, *b,c*). Their coordinates have different polar angles, but the same principal value $\varphi=\frac{\pi}{4}$. Hence, it is the same point that coincides with point $M\left(3,\frac{\pi}{4}\right)$, plotted on Fig. 9.9,*a*.

c) Taking into account step 'b', find Cartesian coordinates of point M. By formulas (9.11) obtain:

$$x=r\cdot\cos\varphi=3\cdot\cos\frac{\pi}{4}=\frac{3\sqrt{2}}{2};\ y=r\cdot\sin\varphi=3\cdot\sin\frac{\pi}{4}=\frac{3\sqrt{2}}{2},\text{ i.e. } M\left(\frac{3\sqrt{2}}{2},\frac{3\sqrt{2}}{2}\right).$$

d) For point $A(-3,4)$ by formulas (9.12) find radius $r_A=\sqrt{(-3)^2+4^2}=5$, and, taking into account Fig. 9.8, the principal value of the polar angle $\varphi_A=\pi+\arctan\left(\frac{-3}{4}\right)=\pi-\arctan\frac{3}{4}$. ■

The distance between two points $A(r_A,\varphi_A)$ and $B(r_B,\varphi_B)$ (the length of segment AB on Fig. 9.10) is calculated by the formula

$$AB=\sqrt{r_A^2+r_B^2-2\cdot r_A\cdot r_B\cdot\cos(\varphi_B-\varphi_A)},$$

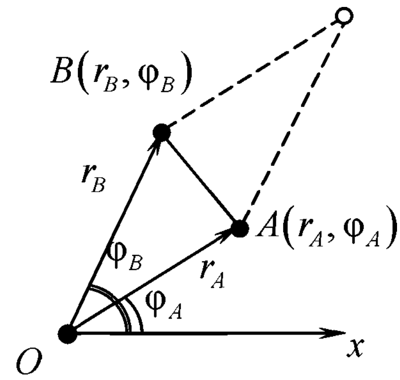


Figure 9.10

and the area $S_{\#\overline{OA},\overline{OB}}$ of a parallelogram constructed on vectors $\overline{OA}$ and $\overline{OB}$ – by the formula

$$S_{\#\overline{OA},\overline{OB}}=r_A\cdot r_B\cdot\sin|\varphi_B-\varphi_A|.$$

Example 9.6. Given polar coordinates $\varphi_A=\frac{\pi}{3}$, $r_A=4$ and $\varphi_B=\frac{2\pi}{3}$, $r_B=2$ of points A and B (Fig. 9.11), find:

a) the scalar product $(\overline{OA},\overline{OB})$;

b) the length of the interval AB;

c) the area of a parallelogram constructed on vectors $\overline{OA}$ and $\overline{OB}$;

d) the area S_{OAB} of the triangle OAB;

e) the coordinates of the midpoint C of the interval AB in the Cartesian coordinate system, related to the given polar one.

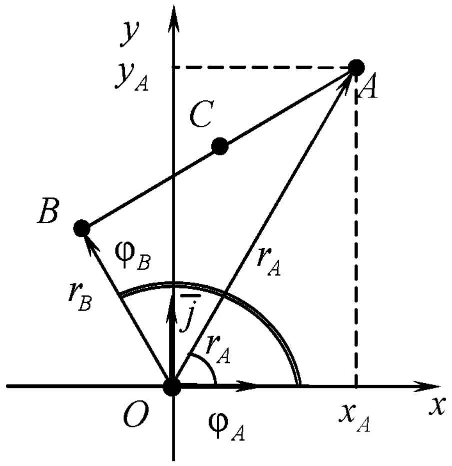


Figure 9.11

□ a) By the definition of scalar product find:

$$(\overline{OA},\overline{OB}) = \left|\overline{OA}\right| \cdot \left|\overline{OB}\right| \cdot \cos\psi = r_A \cdot r_B \cdot \cos(\varphi_B - \varphi_A) = 4 \cdot 2 \cdot \cos\frac{\pi}{3} = 4 .$$

b) Calculate the length of the interval:

$$AB = \sqrt{r_A^2 + r_B^2 - 2 \cdot r_A \cdot r_B \cdot \cos(\varphi_B - \varphi_A)} = \sqrt{4^2 + 2^2 - 2 \cdot 4 \cdot 2 \cdot \tfrac{1}{2}} = 2 \cdot \sqrt{3} .$$

c) Find the area of a parallelogram constructed on vectors $\overline{OA}$ and $\overline{OB}$:

$$S_{\#\overline{OA},\overline{OB}} = r_A \cdot r_B \cdot \sin\left|\varphi_B - \varphi_A\right| = 2 \cdot 4 \cdot \sin\frac{\pi}{3} = 4\sqrt{3} .$$

d) The area of the triangle OAB is calculated like a half of the area of a parallelogram constructed on vectors $\overline{OA}$ and $\overline{OB}$:

$$S_{OAB} = \frac{1}{2} \cdot S_{\#\overline{OA},\overline{OB}} = \frac{1}{2} \cdot 4\sqrt{3} = 2\sqrt{3} .$$

e) By formulas (9.11) find Cartesian coordinates of points A and B:

$$x_A = r_A \cdot \cos\varphi_A = 4 \cdot \frac{1}{2} = 2; \qquad y_A = r_A \cdot \sin\varphi_A = 4 \cdot \frac{\sqrt{3}}{2} = 2\sqrt{3} ;$$

$$x_B = r_B \cdot \cos\varphi_B = 2 \cdot \left(-\frac{1}{2}\right) = -1; \qquad y_B = r_B \cdot \sin\varphi_B = 2 \cdot \frac{\sqrt{3}}{2} = \sqrt{3} ,$$

and then coordinates of the midpoint C of the interval AB:

$$x_C = \frac{x_A + x_B}{2} = \frac{2 + (-1)}{2} = \frac{1}{2}; \qquad y_C = \frac{y_A + y_B}{2} = \frac{2\sqrt{3} + \sqrt{3}}{2} = \frac{3\sqrt{3}}{2} . \quad ■$$

9.3. CYLINDRICAL COORDINATE SYSTEM

To introduce a cylindrical coordinate system in space we need to:

- choose a plane (***reference plane***) and define on it a polar coordinate system with pole O and polar axis Ox.
- draw axis Oz (***applicate axis***) through point O perpendicular to the reference plane and choose its direction so that increase of polar angle, seen from positive direction of axis Oz, happens counterclockwise (Fig. 9.12, *a*).

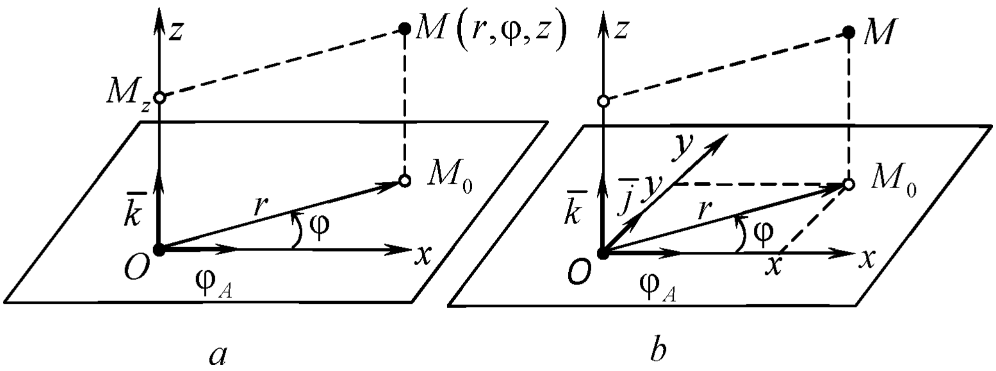


Figure 9.12

Cylindrical coordinates of point M is an ordered triplet of numbers r, φ, z – ***radius*** ($r \geq 0$), ***azimuth*** ($-\pi < \varphi \leq \pi$) and ***height*** ($-\infty < z < +\infty$). The polar angle of points that belong to the applicate axis is not determined, they are defined by height and zero radius.

A cylindrical coordinate system $Or\varphi z$ can be associated with a Cartesian coordinate system $O\bar{i}\,\bar{j}\,\bar{k}$ (Fig. 9.12, *b*), the origin of which coincides with the origin of the cylindrical coordinate system and basis vectors $\bar{i}, \bar{k}$ – with unit vectors on the polar axis and the applicate axis, respectively, and basis vector $\bar{j}$ is chosen in such way that triplet $\bar{i}, \bar{j}, \bar{k}$ is right (giving us a standard basis).

Vice versa, if a right-handed Cartesian system is given in space, we can obtain a cylindrical coordinate system (*associated with the given Cartesian one*) by assuming positive abscissa semi-axis as the polar axis.

Since applicate z of point M in a Cartesian coordinate system and height z in a cylindrical coordinate system are the same, formulas that relate Cartesian coordinates x, y, z of point M and its cylindrical coordinates r, φ, z, take the form:

$$\begin{cases} x = r \cdot \cos\varphi, \\ y = r \cdot \sin\varphi, \\ z = z. \end{cases} \tag{9.13}$$

These formulas allow us to find *Cartesian coordinates* by the given *cylindrical* ones.

The reverse conversion is performed by the formulas:

$$\begin{cases} r = \sqrt{x^2 + y^2}, \\ \cos\varphi = \dfrac{x}{\sqrt{x^2 + y^2}}, \\ \sin\varphi = \dfrac{y}{\sqrt{x^2 + y^2}}, \\ z = z. \end{cases} \tag{9.14}$$

The principal value of the azimuth φ $(-\pi < \varphi \leq \pi)$ is found by the formulas given on Fig. 9.8.

These formulas allow us to find *cylindrical coordinates* by the given *Cartesian* ones.

Example 9.7. In the cylindrical coordinate system $Or\phi z$:

a) built coordinate surfaces $r = R$, $\varphi = 0, \varphi = \varphi_0$, $z = 0$, $z = z_0$;

b) find cylindrical coordinates of point A by the given Cartesian coordinates $A(4, -3, 2)$;

c) find Cartesian coordinates of point B by the given cylindrical coordinates: $r_B = 2$, $\varphi_B = \frac{2\pi}{3}$, $z_B = 1$.

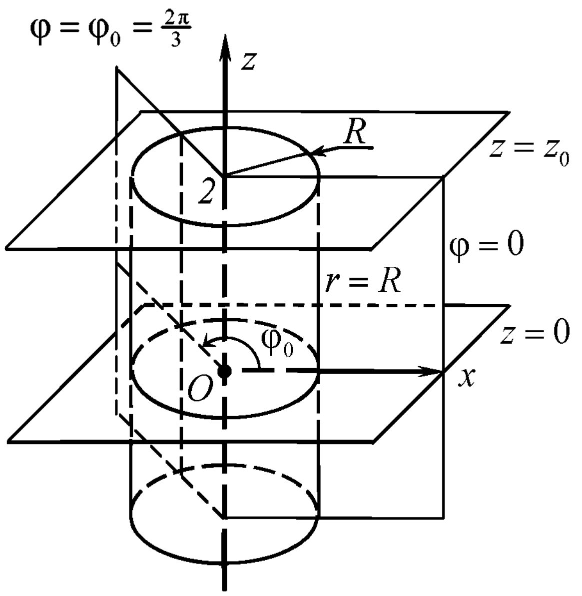


Figure 9.13

□ a) Coordinates surface $r = R$, i.e. geometric locus of points $M(R,\varphi,z)$ with a fixed value of radius $r = R$, is a right circular cylinder, the axis of which coincides with the applicate axis (Fig. 9.13). It explains the name of cylindrical coordinate system. Coordinate surface $\varphi = \varphi_0$, i.e. geometric locus of points $M(r,\varphi_0,z)$ with a fixed value of azimuth $\varphi = \varphi_0$, is a half-plane bound by the applicate axis (Figure 9.13 shows half-planes $\varphi = 0$ and $\varphi = \varphi_0 = \frac{2\pi}{3}$). Coordinate surface $z = z_0$, i.e. geometric locus of points $M(r,\varphi,z_0)$ with a fixed value of height $z = z_0$, is a plane perpendicular to the applicate axis (Figure 9.13 shows planes $z = 0$ and $z = 2$).

b) Find cylindrical coordinates of point $A(4,-3,2)$. The height $z_A = 2$, the radius and the azimuth are found by formulas (9.14) taking into account Fig. 9.8:

$$r_A = \sqrt{x_A^2 + y_A^2} = \sqrt{4^2 + (-3)^2} = 5;\ \varphi_A = \arctan\frac{y_A}{x_A} = \arctan\frac{-3}{4} = -\arctan\frac{3}{4};\ z_A = 2,$$

since $-\pi < \varphi \le \pi$ and orthogonal projection of point A on coordinate plane Oxy (reference plane) belongs to quadrant IV.

c) Find Cartesian coordinates of point B. By formulas (9.13) calculate (see Example 9.6):

$$x_B = r_B \cdot \cos\varphi_B = 2\cdot\left(-\tfrac{1}{2}\right) = -1;\quad y_B = r_B \cdot \sin\varphi_B = 2\cdot\tfrac{\sqrt{3}}{2} = \sqrt{3};\quad z_B = 1.\ ■$$

9.4. SPHERICAL COORDINATE SYSTEM

To introduce a spherical coordinate system in space we need to:

- choose a plane (***reference plane***) and define on it a polar coordinate system with pole O (***origin of the spherical coordinate system***) and polar axis Ox.
- draw axis Oz (***applicate axis***) through point O perpendicular to the reference plane and choose its direction so that increase of polar angle, seen from positive direction of axis Oz, happens counterclockwise (Fig. 9.14, *a*).

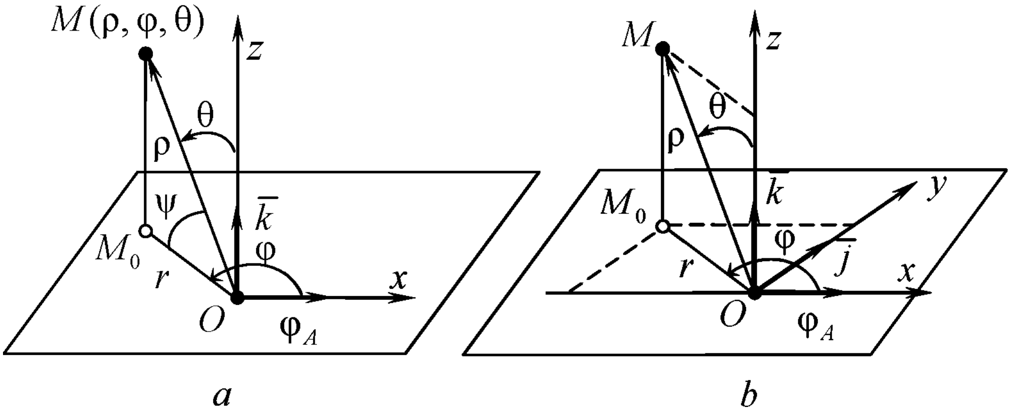


Figure 9.14

Spherical coordinates of point M is an ordered triplet of numbers ρ, φ, θ – ***radius*** ($\rho \geq 0$), ***polar angle*** ($-\pi < \varphi \leq \pi$) and ***azimuthal angle*** ($0 \leq \theta \leq \pi$). The polar angle of points belonging to the applicate axis is not determined, they are defined by radius ρ and azimuthal angle $\theta = 0$ for the positive part of axis Oz and $\theta = \pi$ for its negative part. The origin is defined by zero value of radius ρ. Sometimes angle $\psi = \frac{\pi}{2} - \theta$, taking on values $-\frac{\pi}{2} \leq \psi \leq \frac{\pi}{2}$, is called the azimuthal angle instead of angle θ.

A spherical coordinate system $O\rho\varphi\theta$ can be associated with a Cartesian coordinate system $O\bar{i}\,\bar{j}\,\bar{k}$ (Fig. 9.14, *b*), the origin of which coincides with the origin of the spherical coordinate system and basis vectors $\bar{i}, \bar{k}$ - with unit vectors on the polar axis Ox and the applicate axis Oz, respectively, and basis vector $\bar{j}$ is chosen so that triplet $\bar{i}, \bar{j}, \bar{k}$ is right (giving us a standard basis).

Vice versa, if a right-handed Cartesian system is given in space, we can obtain a spherical coordinate system (*associated with the given Cartesian one*) by assuming positive abscissa semi-axis as the polar axis.

Formulas that relate Cartesian coordinates x, y, z of point M and its spherical coordinates ρ, φ, θ follow from Fig. 9.14, *b*:

$$\begin{cases} x = \rho \cdot \cos\varphi \cdot \sin\theta, \\ y = \rho \cdot \sin\varphi \cdot \sin\theta, \\ z = \rho \cdot \cos\theta. \end{cases} \tag{9.15}$$

This formulas allow us to find *Cartesian coordinates* by the given *spherical* ones.

The reverse conversion is performed by the formulas:

$$\begin{cases} \rho = \sqrt{x^2 + y^2 + z^2}, \\ \cos\varphi = \dfrac{x}{\sqrt{x^2 + y^2}}, \\ \sin\varphi = \dfrac{y}{\sqrt{x^2 + y^2}}, \\ \theta = \arccos\dfrac{z}{\rho} = \arccos\dfrac{z}{\sqrt{x^2 + y^2 + z^2}}. \end{cases} \tag{9.16}$$

Formulas (9.16) allow us to calculate the polar angle φ with accuracy up to summands $2\pi n$, where $n \in Z$. For $x \neq 0$ it follows that $\tan\varphi = \dfrac{y}{x}$. The principal value of the polar angle φ $(-\pi < \varphi \leq \pi)$ is found by the formulas given on Fig. 9.8.

Example 9.8. In the spherical coordinate system $O\rho\varphi\theta$:

a) built coordinate surfaces $\rho = R$, $\varphi = \varphi_0$, $\theta = \theta_0$ $(0 < \theta_0 < \pi)$;

b) find spherical coordinates ρ, φ, θ of point A by the given Cartesian coordinates $A(4, -3, 12)$;

c) find Cartesian coordinates x, y, z of point B by the given spherical coordinates: $\rho = 4$; $\varphi = \dfrac{2\pi}{3}$, $\theta = \dfrac{3\pi}{4}$.

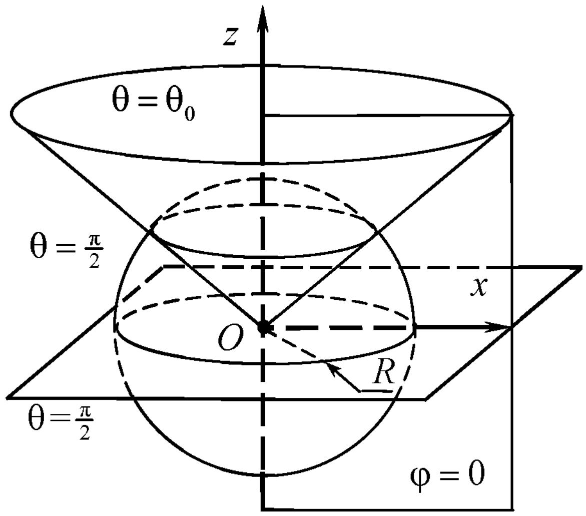


Figure 9.15

☐ a) Coordinate surface $\rho = R$, i.e. geometric locus of points $M(R, \varphi, \theta)$ with a fixed value of radius $\rho = R$, is a sphere with the center in the origin (Fig. 9.15). It explains the name of spherical coordinate system. Coordinate surface $\varphi = \varphi_0$, i.e. geometric locus of points $M(\rho, \varphi_0, \theta)$ with a fixed value of polar angle $\varphi = \varphi_0$, is a half-plane bound by the applicate axis (Fig. 9.15 shows half-plane $\varphi = 0$). Coordinate surface $\theta = \theta_0$, i.e. geometric locus of points $M(\rho, \varphi, \theta_0)$ with a fixed value of azimuthal angle $\theta = \theta_0 \neq \frac{\pi}{2}$, is a cone, axis of which coincides with the applicate axis and vertex – with the origin. For $\theta = \frac{\pi}{2}$ we obtain the reference plane. Fig. 9.15 shows cone $\theta = \theta_0 \neq \frac{\pi}{2}$ and reference plane $\theta = \frac{\pi}{2}$.

b) Find spherical coordinates of point $A(4, -3, 12)$. By formulas (9.16), taking into account Fig. 9.8 (see Example 9.6), we obtain:

$$\rho = \sqrt{4^2 + (-3)^2 + 12^2} = 13; \quad \varphi = -\arctan\frac{3}{4}; \quad \theta = \arccos\frac{12}{13}.$$

c) By formulas (9.15) we obtain

$$x = \rho \cdot \cos\varphi \cdot \sin\theta = 4 \cdot \left(-\tfrac{1}{2}\right) \cdot \left(\tfrac{\sqrt{2}}{2}\right) = -\sqrt{2}; \quad y = \rho \cdot \sin\varphi \cdot \sin\theta = 4 \cdot \left(\tfrac{\sqrt{3}}{2}\right) \cdot \left(\tfrac{\sqrt{2}}{2}\right) = \sqrt{6};$$

$$z = \rho \cdot \cos\theta = 4 \cdot \left(-\tfrac{\sqrt{2}}{2}\right) = -2\sqrt{2}. \blacksquare$$

EXERCISES

1. Consider coordinates $x = n$, $y = m$, $z = n + 2$ of point A in old coordinate system $Oxyz$. The new one $O'x'y'z'$ is obtained by the translation by vector $\overline{s} = 3n \cdot \overline{i} - m \cdot \overline{j} + 2 \cdot \overline{k}$ and rotation by angle $\varphi = \frac{\pi}{6}$ around applicate axis. Find coordinates x', y', z' of point A in new coordinate system.

2. Given the polar coordinates $r_A = m$, $\varphi_A = \frac{\pi}{30}n$ and $r_B = m + 5$, $\varphi_B = \frac{\pi}{60}n$ of points A and B find the length of segment AB, area of triangle OAB and coordinates of the middle point of AB in Cartesian coordinate system Oxy, associated with the given polar coordinate system $Or\varphi$.

3. Given the cylindrical coordinate system $Or\varphi z$ and Cartesian coordinate system $Oxyz$, associated to it, find:

a) cylindrical coordinates of the point A, if its Cartesian coordinates are $A(2n, -3m, m+n)$;

b) Cartesian coordinates of the point B, if its cylindrical coordinates are: $r_B = 3n$, $\varphi_B = \frac{\pi}{30}n$, $z_B = m$.

4. Given the spherical coordinates system $O\rho\varphi\theta$ and Cartesian coordinate system $Oxyz$, associated with it, find:

a) spherical coordinates of the point A, if its Cartesian coordinates are $A(2n, -3m, m+n)$;

b) Cartesian coordinates of the point B, if its spherical coordinates are: $\rho_B = n + 4$, $\varphi_B = \frac{\pi}{2} + \frac{\pi}{30}n$, $\theta_B = \frac{\pi}{6}$.

CHAPTER 10. ALGEBRAIC PLANE CURVES

10.1. FIRST–ORDER CURVES (LINES ON PLANE)

10.1.1. Main Types of Equations for Lines on Plane

A nonzero vector $\overline{n}$, perpendicular to the given line, is called a ***normal vector*** (or simply '***normal***') to this line. A ***direction vector*** of a line is a nonzero vector, ***collinear with this line***, i.e. belonging to the line or parallel to it. Two lines are called ***collinear***, if they are parallel or coincident.

General equation of a line on plane:

$$A\cdot x+B\cdot y+C=0,\quad A^2+B^2\neq 0. \tag{10.1}$$

Way of representation: the line passes through point $M_0(x_0,y_0)$ perpendicular to the vector $\overline{n}=A\cdot\overline{i}+B\cdot\overline{j}$ (Fig. 10.1,*a*).

Geometric sense of coefficients: leading coefficients A, B are the coordinates of the normal $\overline{n}=A\cdot\overline{i}+B\cdot\overline{j}$; constant term $C=-Ax_0-By_0$.

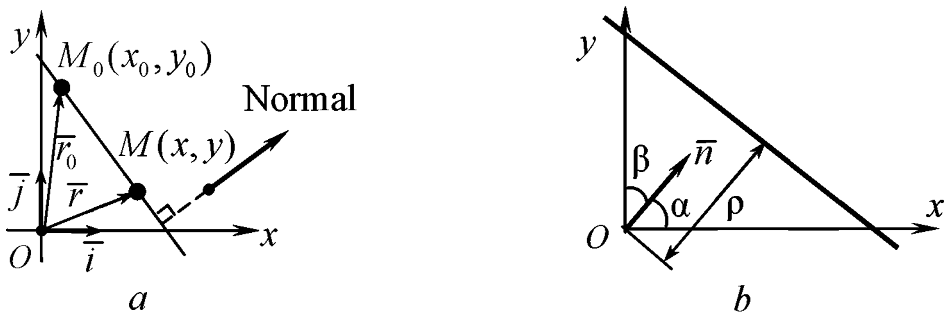


Figure 10.1

Denoting radius-vectors of points $M_0(x_0,y_0)$ and $M(x,y)$ by $\overline{r}_0$ and $\overline{r}$, respectively, we can write a ***vector equation of a line on plane***, passing through point $M_0(x_0,y_0)$ perpendicular to the normal $\overline{n}=A\cdot\overline{i}+B\cdot\overline{j}$:

$$(\overline{r}-\overline{r}_0,\overline{n})=0.$$

The scalar product is equal to zero, representing perpendicularity condition of vectors $\overline{r}-\overline{r}_0$ and $\overline{n}$ (see Section 9.7). In coordinate form the equation takes the following form:

$$A\cdot(x-x_0)+B\cdot(y-y_0)=0. \qquad (10.2)$$

***Normalized equation of a line on plane*:**

$$x\cdot\cos\alpha+y\cdot\cos\beta-\rho=0, \qquad \rho\geq 0. \qquad (10.3)$$

Way of representation: the line passes through point $M_0(x_0,y_0)$ perpendicular to the vector $\overline{n}=\cos\alpha\cdot\overline{i}+\cos\beta\cdot\overline{j}$ (Fig. 10.1,*a*).

Geometric sense of coefficients: leading coefficients $\cos\alpha$, $\cos\beta$ are direction cosines of the normal $\overline{n}=\cos\alpha\cdot\overline{i}+\cos\beta\cdot\overline{j}$; constant term $\rho=\dfrac{|Ax_0+By_0|}{\sqrt{A^2+B^2}}$ is the distance from the origin to the line (Fig. 10.1,*b*).

***Vector parametric equation of a line on plane*:**

$$\overline{r}=\overline{r}_0+t\cdot\overline{p}, \quad t\in\mathbb{R}, \quad \overline{p}\neq\overline{o}. \qquad (10.4)$$

Way of representation: the line passes through point $M_0(x_0,y_0)$, which is defined by radius-vector $\overline{r}_0$, collinear with the direction vector $\overline{p}\neq\overline{o}$ (Fig. 10.2).

Parameter t in equation (10.4) has the following *geometric sense*: the value of t is proportional to the distance between the initial point M_0 and point M, defined by radius-vector $\overline{r}$.

Physical sense of parameter t: it is time in uniform rectilinear motion from point M along the line. For $t=0$ point M coincides with the initial point M_0 ($\overline{r}=\overline{r}_0$), when t is increasing, the motion happens in the direction determined by the direction vector $\overline{p}$.

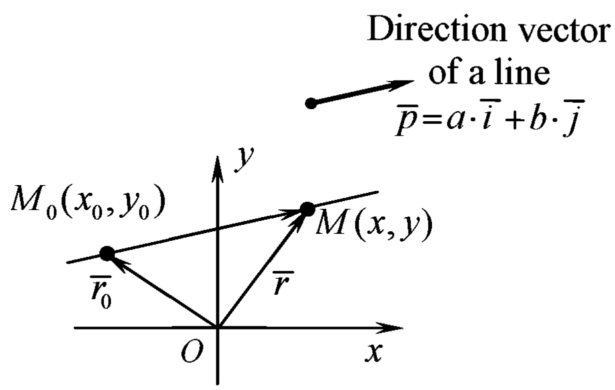


Figure 10.2

***Parametric equation of a line on plane*:**

$$\begin{cases} x = x_0 + a \cdot t, \\ y = y_0 + b \cdot t, \end{cases} \quad t \in \mathbb{R}, \quad a^2 + b^2 \neq 0. \tag{10.5}$$

Way of representation: the line passes through point $M_0(x_0, y_0, z_0)$ collinear with the vector $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j}$ (Fig. 10.3).

Geometric sense of coefficients: a and b are coordinates of direction vector of the line $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j}$; x_0, y_0 are coordinates of point $M_0(x_0, y_0, z_0)$ that belongs to the line. Parameter t has the same sense as in equation (10.4).

Note that equation (10.5) is a coordinate form of equation (10.4).

***Canonical equation of a line on plane*:**

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}, \quad a^2 + b^2 \neq 0. \tag{10.6}$$

Way of representation: the line passes through point $M_0(x_0, y_0, z_0)$ collinear with the vector $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j}$ (Fig. 10.3).

Geometric sense of coefficients: a and b are coordinates of direction vector of the line $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j}$; x_0, y_0 are coordinates of point $M_0(x_0, y_0, z_0)$ that belongs to the line.

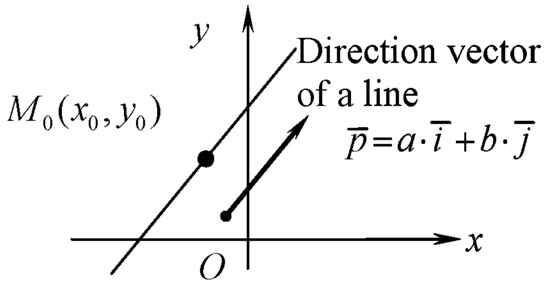


Figure 10.3

One of the denominators a or b in equation (10.6) can be equal to zero, in this case we assume the corresponding numerator equal to zero:

$\dfrac{x - x_0}{0} = \dfrac{y - y_0}{b} \Leftrightarrow x = x_0$ – equation of a line parallel to the ordinate axis;

$\dfrac{x - x_0}{a} = \dfrac{y - y_0}{0} \Leftrightarrow y = y_0$ – equation of a line parallel to the abscissa axis.

***Affine equation of a line on plane passing through two given points*:**

$$\overline{r} = (1-t)\cdot\overline{r_0} + t\cdot\overline{r_1}, \quad t\in\mathbb{R}. \tag{10.7}$$

Equation (10.7) can be rewritten in coordinate form:

$$\begin{cases} x = (1-t)\cdot x_0 + t\cdot x_1, \\ y = (1-t)\cdot y_0 + t\cdot y_1, \end{cases} \quad t\in\mathbb{R}.$$

Way of representation: the line passes through two given points $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$ which are defined by radius vectors $\overline{r_0}$ and $\overline{r_1}$, respectively (Fig. 10.4). Radius vector $\overline{r}$ defines the position of point $M(x, y, z)$ that belongs to the line.

Geometric sense of coefficients: x_0, y_0 and x_1, y_1 are coordinates of points $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$, through which the line (10.7) passes. Parameter t in equation (10.7) defines the position of point $M(x, y, z)$ that belongs to the line, e.g., for $t = 0$ point M coincides with point $\overline{r} = \overline{r_0}$, and for $t = 1$ – with point M_1 ($\overline{r} = \overline{r_1}$).

***Equation of a line on plane passing through two given points $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$*:**

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0}. \tag{10.8}$$

Way of representation: the line passes through two given points $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$ (Fig. 10.4).

Geometric sense of coefficients: x_0, y_0 and x_1, y_1 are coordinates of points $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$, through which the line (10.8) passes. Like in the canonical equation, one of the denominators in (10.8) can be equal to zero, assuming the corresponding numerator equal to zero as well.

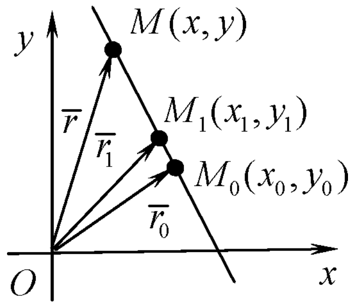


Figure 10.4

***Two intercept form for the equation of a line*:**

$$\frac{x}{x_1}+\frac{y}{y_1}=1,\ x_1\neq 0,\ y_1\neq 0. \qquad (10.9)$$

Way of representation: the line passes through two given points $X_1(x_1,0)$ and $Y_1(0,y_1)$ (Fig. 10.5).

Geometric sense of coefficients: line (10.9) intercepts coordinate axes, cutting off segments x_1 on the abscissa axis and y_1 on the ordinate axis.

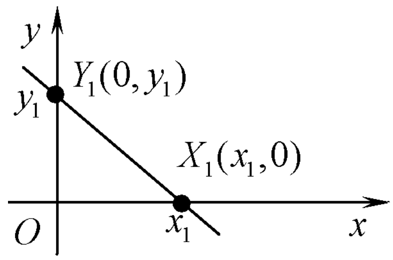


Figure 10.5

Slope-intercept form for the equation of a line (equation solved for y):

$$y=k\cdot x+y_0,\ \ k=\tan\alpha. \qquad (10.10)$$

Way of representation: the line passes through point $Y_0(0,y_0)$ including angle α ($0\leq\alpha<\pi$, $\alpha\neq\frac{\pi}{2}$) to positive direction of the abscissa axis (Fig. 10.6).

Geometric sense of coefficients: k is the slope of the line and y_0 is the ordinate of point $Y_0(0,y_0)$, through which the line (10.10) passes.

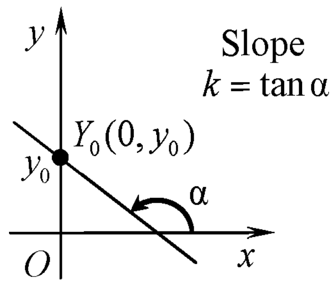


Figure 10.6

If the line passes through the given point $M_0(x_0, y_0)$, we use the slope-intercept form of the equation as: $y - y_0 = k \cdot (x - x_0)$.

Ways of Converting from One Form of Line Equation to Another

1. To convert from the general equation of a line (10.1) to the normalized one (10.3) it is sufficient to divide both parts of the general equation by the length of the normal $|\overline{n}| = \sqrt{A^2 + B^2}$, if the constant term is negative ($C < 0$), or divide by its opposite $-|\overline{n}| = -\sqrt{A^2 + B^2}$, if the constant term is non-negative ($C \geq 0$).

2. To convert from the general equation of a line (10.1) to the canonical one (10.6) we should make the following steps:

1) find any solution (x_0, y_0) to equation $A \cdot x + B \cdot y + C = 0$, thus defining coordinates of point $M_0(x_0, y_0)$ that belongs to the line;

2) find any nonzero solution (a, b) of homogeneous equation $A \cdot a + B \cdot b = 0$, thus defining coordinates a, b of the direction vector $\overline{p}$, in particular we can assume $a = B$, $b = -A$;

3) write canonical equation (10.6).

3. To convert from the canonical equation of a line to the general one it's sufficient to transpose all terms of equation (10.6) to the left part:

$$\frac{x - x_0}{a} - \frac{y - y_0}{b} = 0 \quad \Leftrightarrow \quad \frac{1}{a} \cdot x - \frac{1}{b} \cdot y + \frac{y_0}{b} - \frac{x_0}{a} = 0 .$$

Obtained equation (for $a \neq 0$, $b \neq 0$) takes on form (10.1) where $A = \frac{1}{a}$, $B = -\frac{1}{b}$, $C = \frac{y_0}{b} - \frac{x_0}{a}$.

4. To convert from the canonical equation to the parametric one, assume left and right part of equation (10.6) equal to parameter t and write obtained double equation in form of a system (10.5):

$$\frac{x - x_0}{a} = t = \frac{y - y_0}{b} \quad \Leftrightarrow \quad \begin{cases} x = x_0 + a \cdot t, \\ y = y_0 + b \cdot t, \end{cases} \quad t \in \mathbb{R}.$$

5. It is possible to convert the general equation of a line (10.1) to the two intercept form (10.9) if all coefficients of general equation are not equal to zero. To do this, it is necessary to transpose the constant term to the right part of the equation: $A \cdot x + B \cdot y = -C$, and then divide both part of the equation by $-C$: $\frac{A}{-C} \cdot x + \frac{B}{-C} \cdot y = 1$. Denoting $x_1 = -\frac{C}{A}$, $y_1 = -\frac{C}{B}$, obtain the intercept equation (10.9): $\frac{x}{x_1} + \frac{y}{y_1} = 1$.

6. To convert from the general equation of a line (10.1) $A \cdot x + B \cdot y + C = 0$ to slope-intercept form (10.10), it is necessary to solve the general equation for the unknown y:

$$y = -\frac{A}{B} \cdot x - \frac{C}{B} \quad \Leftrightarrow \quad y = k \cdot x + y_0,$$

where $k = -\frac{A}{B}$, $y_0 = -\frac{C}{B}$. This conversion is possible if $B \neq 0$.

Example 10.1. Given points $K(1,2)$ and on $L(5,0)$ coordinate plane Oxy (in Cartesian coordinate system), write the equation for the perpendicular bisector of segment KL (Fig. 10.7).

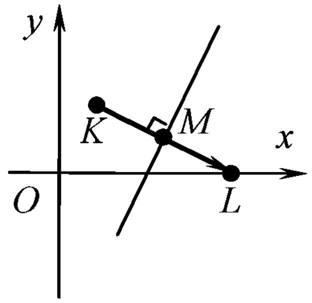


Figure 10.7

□ The perpendicular bisector, by definition, passes through the midpoint of segment KL perpendicular to it. Let's find coordinates of the midpoint M of segment KL (see a special case of formula (9.1) in Section 9.1.1): $M\left(\frac{1+5}{2}, \frac{2+0}{2}\right)$, i.e. $M(3,1)$.

Vector $\overline{KL}$ can be taken as the normal to the perpendicular bisector. Find coordinates of this vector by subtracting coordinates of its tail from the corresponding coordinates of its head:

$$\overline{KL} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} = \overline{n}.$$

Hence, equation (10.1) of the required line is given by $4 \cdot x - 2 \cdot y + C = 0$.

Now we only have to find the value of the constant term C. Since point $M(3,1)$ belongs to the line, its coordinates $x = 3, y = 1$ must satisfy the equation of this line, hence, $4 \cdot 3 - 2 \cdot 1 + C = 0$. Thus $C = -10$.

Hence, the perpendicular bisector is determined by the following equation:

$$4 \cdot x - 2 \cdot y - 10 = 0 \quad \Leftrightarrow \quad 2 \cdot x - y - 5 = 0.$$

The equation of this line can also be obtained in form (10.2), by inserting coordinates of the normal $\overline{n} = (4 \quad -2)^T$ and point $M(3,1)$:

$$4 \cdot (x-3) - 2 \cdot (y-1) = 0.$$

The solution is obtained analytically, without using graphic representation (see Fig. 10.7). Plots in analytic geometry, as a rule, serve only as illustrations to solutions. ■

Example 10.2. Given line l, represented by equation $x-3\cdot y+3=0$, and point $M(5,6)$ (Fig. 10.8) on coordinate plane Oxy (in Cartesian coordinate system), it is required:

a) write parametric equation of line m passing through point M perpendicular to the given line;

b) find orthogonal projection M_l of point M to line l;

c) find coordinates of point M', symmetrical to point M with respect to line l.

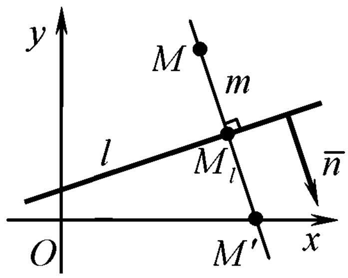


Figure 10.8

□ a) Normal $\overline{n}$ to line l is the direction vector $\overline{p}$ to line m. Coordinates of the normal can be found from the general equation of line l: $\overline{n}=1\cdot\overline{i}-3\cdot\overline{j}$, then $\overline{p}=1\cdot\overline{i}-3\cdot\overline{j}$, $x_0=5$, $y_0=6$. Write the parametric equation (10.5) of line m:

$$\begin{cases} x=5+1\cdot t, \\ y=6+(-3)\cdot t, \end{cases} \quad t\in\mathbb{R}.$$

b) Projection M_l of point M is the intersection point of lines m and l. Let's find its coordinates. To do this, insert expressions of coordinates $x=5+t$, $y=6-3\cdot t$ from the parametrical equation of line m into the equation of line l: $x-3\cdot y+3=0$. We obtain the equation:

$$\underbrace{5+t}_{x}-3\cdot\underbrace{(6-3\cdot t)}_{y}+3=0 \quad\Leftrightarrow\quad 10\cdot t-10=0 \quad\Leftrightarrow\quad t=1.$$

The value $t=1$ of the parameter is corresponded by the point with coordinates $x=5+1=6$, $y=6-3\cdot 1=3$. Thus, the point in question is $M_l(6,3)$.

c) During step "a" we wrote the parametrical equation of line m. From this equation for $t=0$ we obtain point M, for $t=1$ – point M_l, thus the point in question M' is obtained for $t=2$, since by symmetry $M'M_l = M_lM$. Find the coordinates of the required point:

$$M'(5+1\cdot 2, 6+(-3)\cdot 2), \text{ i.e. } M'(7,0). \blacksquare$$

10.1.2. Geometric Relationships of Lines on Plane

Let two lines l_1 and l_2 be given by their general equations

$$l_1:\ A_1\cdot x + B_1\cdot y + C_1 = 0; \qquad l_2:\ A_2\cdot x + B_2\cdot y + C_2 = 0$$

or equations in slope-intercept form:

$$l_1:\ y = k_1\cdot x + b_1; \qquad l_2:\ y = k_2\cdot x + b_2.$$

Geometric relationship of two lines on plane can be assessed by coefficients of their equations with the help the following criteria:

- *parallel* lines:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2} \quad \text{or} \quad k_1 = k_2,\ b_1 \neq b_2;$$

- *coincident* lines:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} \quad \text{or} \quad k_1 = k_2,\ b_1 = b_2;$$

- *collinear* lines:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \quad \text{or} \quad k_1 = k_2;$$

- *intersecting* lines:

$$A_1\cdot B_2 \neq A_2\cdot B_1 \quad \text{or} \quad k_1 \neq k_2;$$

- *perpendicular* lines:

$$A_1\cdot A_2 + B_1\cdot B_2 = 0 \quad \text{or} \quad k_1\cdot k_2 = -1.$$

If two lines intersect, coordinates of the common point can be obtained by solving one of the systems:

$$\begin{cases} A_1 \cdot x + B_1 \cdot y + C_1 = 0, \\ A_2 \cdot x + B_2 \cdot y + C_2 = 0 \end{cases} \quad \text{or} \quad \begin{cases} y = k_1 \cdot x + b_1, \\ y = k_2 \cdot x + b_2. \end{cases}$$

Example 10.5. Find geometric relationships of each pair of lines (intersection, parallelism, incidence, perpendicularity, if lines intersect, find their common point):

a) $2 \cdot x - y + 3 = 0$, $-4 \cdot x + 2 \cdot y - 6 = 0$;

b) $2 \cdot x + 3 \cdot y - 6 = 0$, $4 \cdot x + 6 \cdot y + 3 = 0$;

c) $3 \cdot x - 2 \cdot y + 1 = 0$, $4 \cdot x + 6 \cdot y - 16 = 0$;

d) $x + 2 \cdot y - 3 = 0$, $x - 4 \cdot y + 3 = 0$;

e) $y = -4 \cdot x + 1$, $y = -4 \cdot x - 3$;

f) $y = -x + 1$, $-2 \cdot x + 2 \cdot y - 6 = 0$.

□ a) Since $A_1 = 2$, $B_1 = -1$, $C_1 = 3$, $A_2 = -4$, $B_2 = 2$, $C_2 = -6$ and $\frac{A_1}{A_2} = \frac{2}{-4} = -\frac{1}{2}$, $\frac{B_1}{B_2} = -\frac{1}{2}$, $\frac{C_1}{C_2} = \frac{3}{-6} = -\frac{1}{2}$, then $\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = -\frac{1}{2}$. Hence, lines are coincident.

b) Since $A_1 = 2$, $B_1 = 3$, $C_1 = -6$, $A_2 = 4$, $B_2 = 6$, $C_2 = 3$ and $\frac{A_1}{A_2} = \frac{1}{2}$, $\frac{B_1}{B_2} = \frac{1}{2}$, $\frac{C_1}{C_2} = -2$, then $\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}$. Hence, lines are parallel.

c) Since $A_1 = 3$, $B_1 = -2$, $C_1 = 1$, $A_2 = 4$, $B_2 = 6$, $C_2 = -16$, then $A_1 \cdot B_2 = 3 \cdot 6 = 18$ and $A_2 \cdot B_1 = 4 \cdot (-2) = -8$. Hence, $A_1 \cdot B_2 \neq A_2 \cdot B_1$ and lines intersect. Since $A_1 \cdot A_2 + B_1 \cdot B_2 = 3 \cdot 4 + (-2) \cdot 6 = 0$, lines are perpendicular. Coordinates of the intersection point $(1;2)$ satisfy the system of equations $\begin{cases} 3 \cdot x - 2 \cdot y + 1 = 0, \\ 4 \cdot x + 6 \cdot y - 16 = 0. \end{cases}$

d) Since $A_1 = 1$, $B_1 = 2$, $C_1 = -3$, $A_2 = 1$, $B_2 = -4$, $C_2 = 3$, then $A_1 \cdot B_2 \neq A_2 \cdot B_1$ and $A_1 \cdot A_2 + B_1 \cdot B_2 = 1 \cdot 1 + 2 \cdot (-4) = -7 \neq 0$. Hence, the lines intersect, but they are not perpendicular. Coordinates of the intersection point $(1;1)$ satisfy the system of equations $\begin{cases} x+2\cdot y-3=0, \\ x-4\cdot y+3=0. \end{cases}$

e) Since $k_1=-4$, $b_1=1$, $k_2=-4$, $b_2=-3$, then $k_1=k_2$ и $b_1 \neq b_2$. Hence, lines are parallel.

f) For the first line we have $k_1=-1$, $b_1=1$. Solving the second equation for y, we obtain equation $y=x+3$, i.e. $k_2=1$, $b_2=3$. Since $k_1 \neq k_2$, lines intersect. Since lines are perpendicular. Coordinates of the intersection point $(-1;2)$ satisfy the system of equations $\begin{cases} y=-x+1, \\ y=x+3. \end{cases}$ ■

10.1.3. Metric Applications of Line Equations on Plane

Let's give formulas for calculating lengths of a line segments (distances) and values of angles by the equations of lines that form them.

An angle between two lines on plane is the angle between their direction vectors. By this definition we get not one, but two adjacent supplementary angles that add up to π. In elementary geometry, as a rule, the smaller of the two angles is chosen, i.e. value φ of an angle between two lines satisfies the condition $0 \le \varphi \le \frac{\pi}{2}$.

1. The distance d from point $M^*(x^*, y^*)$ to line $A\cdot x+B\cdot y+C=0$ (Fig. 10.9, a) is calculated by the formula $d=\dfrac{|A\cdot x^*+B\cdot y^*+C|}{\sqrt{A^2+B^2}}$.

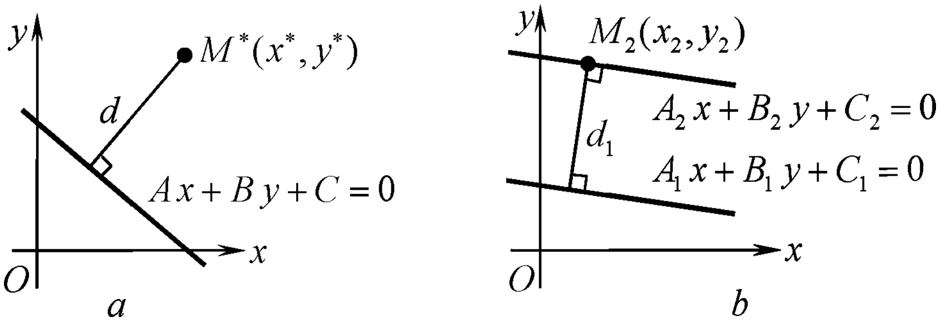


Figure 10.9

2. The distance between two parallel lines $A_1 \cdot x + B_1 \cdot y + C_1 = 0$ and $A_2 \cdot x + B_2 \cdot y + C_2 = 0$ (Fig.10.9, *b*) is calculated as the distance d_1 from point $M_2(x_2, y_2)$, coordinates of which satisfy equation $A_2 \cdot x_2 + B_2 \cdot y_2 + C_2 = 0$, to line $A_1 \cdot x + B_1 \cdot y + C_1 = 0$ by formula

$$d_1 = \frac{|A_1 \cdot x_2 + B_1 \cdot y_2 + C_1|}{\sqrt{A_1^2 + B_1^2}}.$$

3. The acute angle φ between two lines l_1 and l_2 is found by formulas

a) $\cos\varphi = \dfrac{|a_1 \cdot a_2 + b_1 \cdot b_2|}{\sqrt{a_1^2 + b_1^2} \cdot \sqrt{a_2^2 + b_2^2}}$, if $\overline{p}_1 = a_1 \cdot \overline{i} + b_1 \cdot \overline{j}$ and $\overline{p}_2 = a_2 \cdot \overline{i} + b_2 \cdot \overline{j}$ are direction vectors of lines l_1 and l_2, respectively (if lines are given by the canonical or parametric equations (Fig. 10.10, *a*));

b) $\cos\varphi = \dfrac{|A_1 \cdot A_2 + B_1 \cdot B_2|}{\sqrt{A_1^2 + B_1^2} \cdot \sqrt{A_2^2 + B_2^2}}$, if $\overline{n}_1 = A_1 \cdot \overline{i} + B_1 \cdot \overline{j}$ and $\overline{n}_2 = A_2 \cdot \overline{i} +$ $+ B_2 \cdot \overline{j}$ are normals to lines $l_1$and l_2, respectively (if lines are given by the general equations (Fig. 10.10, *a*));

c) $\tan\varphi = \left|\dfrac{k_1 - k_2}{1 + k_1 \cdot k_2}\right|$, $k_1 \cdot k_2 \neq -1$, if $k_1 = \tan\alpha_1$ and $k_2 = \tan\alpha_2$ are slopes of lines l_1 and l_2 respectively (if lines are given by the equations in slope-intercept form (Fig. 10.10, *b*)). If $k_1 \cdot k_2 = -1$, then $\varphi = \frac{\pi}{2}$, since lines are perpendicular (see Section 10.1.2).

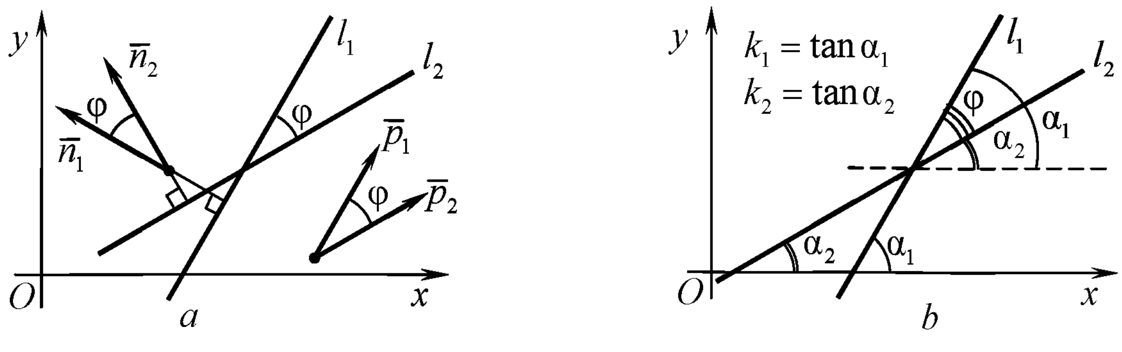


Figure 10.10

Example 10.6. Find:

a) the distance from point $M^*(1,-2)$ to line $3\cdot x+4\cdot y+10=0$;

b) the distance between parallel lines $2\cdot x+3\cdot y-6=0$ and $4\cdot x+6\cdot y+2=0$;

c) the acute angle between lines l_1: $3\cdot x-y-3=0$ and l_2: $x-2\cdot y+4=0$;

d) the acute angle between lines $\frac{x-1}{-1}=\frac{y-3}{-3}$ and $\frac{x-4}{1}=\frac{y-2}{-2}$;

e) the acute angle between lines $y=3\cdot x-1$ and $y=-2\cdot x+2$.

□ a) Let's use the first formula of metric applications ($x^*=1$, $y^*=-2$, $A=3$, $B=4$, $C=10$): $d=\frac{|A\cdot x^*+B\cdot y^*+C|}{\sqrt{A^2+B^2}}=\frac{|3\cdot 1+4\cdot(-2)+10|}{\sqrt{3^2+4^2}}=\frac{5}{5}=1$.

b) Let's choose an arbitrary point $M_2(x_2,y_2)$ on the second line $4\cdot x+6\cdot y+2=0$, e.g., point $M_2(1,-1)$. Then by the second formula of metric applications we obtain (for $A_1=2$, $B_1=3$, $C_1=-6$, $x_2=1$, $y_2=-1$):

$$d_1=\frac{|A_1\cdot x_2+B_1\cdot y_2+C_1|}{\sqrt{A_1^2+B_1^2}}=\frac{|2\cdot 1+3\cdot(-1)-6|}{\sqrt{2^2+3^3}}=\frac{7}{\sqrt{13}}.$$

c) By the general equations of lines find normals

$$\overline{n}_1=A_1\cdot\overline{i}+B_1\cdot\overline{j}=3\cdot\overline{i}-1\cdot\overline{j},\quad \overline{n}_2=A_2\cdot\overline{i}+B_2\cdot\overline{j}=1\cdot\overline{i}-2\cdot\overline{j},$$

and angle φ between the lines by the third formula of metric applications (case "b") (for $A_1=3$, $B_1=-1$, $A_1=3$, $B_2=-2$):

$$\cos\varphi=\frac{|A_1\cdot A_2+B_1\cdot B_2|}{\sqrt{A_1^2+B_1^2}\cdot\sqrt{A_2^2+B_2^2}}=\frac{|3\cdot 1+(-1)\cdot(-2)|}{\sqrt{3^2+(-1)^2}\cdot\sqrt{1^2+(-2)^2}}=\frac{5}{5\cdot\sqrt{2}}=\frac{1}{\sqrt{2}}\Rightarrow\varphi=\frac{\pi}{4}.$$

d) The lines are given by the canonical equations. By coefficients of the equations find direction vectors $\overline{p}_1=a_1\cdot\overline{i}+b_1\cdot\overline{j}=-1\cdot\overline{i}-3\cdot\overline{j}$, $\overline{p}_2=a_2\cdot\overline{i}+b_2\cdot\overline{j}=1\cdot\overline{i}-2\cdot\overline{j}$, and then – angle φ between the lines by the third formula of metric applications (case "a") (for $a_1=-1$, $b_1=-3$, $a_2=1$, $b_2=-2$):

$$\cos\varphi = \frac{|a_1 \cdot a_2 + b_1 \cdot b_2|}{\sqrt{a_1^2 + b_1^2} \cdot \sqrt{a_2^2 + b_2^2}} = \frac{|(-1)\cdot 1 + (-3)\cdot(-2)|}{\sqrt{(-1)^2 + (-3)^2} \cdot \sqrt{1^2 + (-2)^2}} = \frac{5}{\sqrt{10}\cdot\sqrt{5}} = \frac{1}{\sqrt{2}} \Rightarrow$$

$\varphi = \frac{\pi}{4}$.

e) By the equations of lines find their slopes: $k_1 = 3$, $k_2 = -2$, and then – angle φ between the lines by the third formula of metric applications (case "c"): $\tan\varphi = \left|\frac{k_1 - k_2}{1 + k_1 \cdot k_2}\right| = \left|\frac{3 - (-2)}{1 + 3\cdot(-2)}\right| = 1$, i.e. $\varphi = \frac{\pi}{4}$. ■

Example 10.7. Write the equation of a line passing through point $y_0 = 1$ on the ordinate axis and forming angle $\frac{\pi}{4}$ with line $y = \frac{1}{2}\cdot x + 1$.

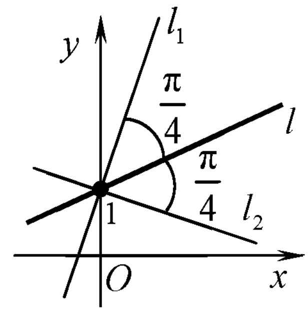


Figure 10.11

□ The required line (with slope k) makes acute angle $\varphi = \frac{\pi}{4}$ with the given line l (with slope $\frac{1}{2}$), $\tan\varphi = 1$. By the third formula of metric applications (case "c"), taking into account that φ is an acute angle, compose the equation and simplify it:

$$1 = \left|\frac{k - \frac{1}{2}}{1 + k\cdot\frac{1}{2}}\right| \quad\Leftrightarrow\quad \frac{k - \frac{1}{2}}{1 + k\cdot\frac{1}{2}} = \pm 1 \quad\Leftrightarrow\quad \left[\begin{array}{l} k - \frac{1}{2} = 1 + \frac{1}{2}\cdot k, \\ k - \frac{1}{2} = -1 - \frac{1}{2}\cdot k. \end{array}\right.$$

We obtain two solutions: $k_1 = 3$ or $k_2 = -\frac{1}{3}$. Hence, taking into account (10.10) for $y_0 = 1$, there are two lines that satisfy the given problem (Fig. 10.13) – l_1: $y = 3\cdot x + 1$ or l_2: $y = -\frac{1}{3}\cdot x + 1$. Note, that these lines are mutually perpendicular, since condition $k_1 \cdot k_2 = 3\cdot(-\frac{1}{3}) = -1$ is satisfied. ■

10.2. SECOND–ORDER CURVES

10.2.1. Classification of Second–Order Curves

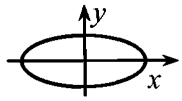
A second-order algebraic curve is the locus of points in plane which in some affine coordinate system Oxy can be given by an equation in form

$$a_{11} \cdot x^2 + 2 \cdot a_{12} \cdot x \cdot y + a_{22} \cdot y^2 + 2 \cdot a_1 \cdot x + 2 \cdot a_2 \cdot y + a_0 = 0,$$

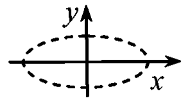
where leading coefficients a_{11}, a_{12}, a_{22} are not equal to zero at the same time ($a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$). For every second-order algebraic curve there exists a Cartesian coordinate system Oxy, in which the equation takes on the simplest (***canonical***) form. This coordinate system, as well as the equation, are called ***canonical***.

Canonical Equations of Second–Order Curves

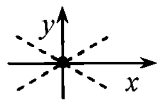
1) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ – equation of an ellipse;



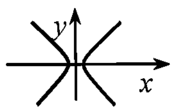
2) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$ – equation of an imaginary ellipse;



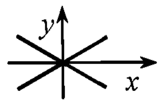
3) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$ – equation of a pair of imaginary intersecting lines;



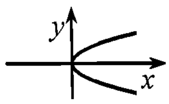
4) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ – equation of a hyperbola;



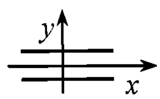
5) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$ – equation of a pair of intersecting lines;



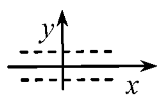
6) $y^2 = 2 \cdot p \cdot x$ – equation of a parabola;



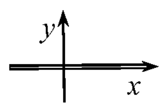
7) $y^2 - b^2 = 0$ – equation of a pair of parallel lines;



8) $y^2 + b^2 = 0$ – equation of a pair of imaginary parallel lines;



9) $y^2 = 0$ – equation of a pair of coincident lines.



In these equations $a > 0$, $b > 0$, $p > 0$, where $a \geq b$ in equations 1–3.

Lines (1),(4)–(7),(9) are called ***real***, and lines (2),(3),(8) – ***imaginary***. Real lines are sketched out in canonical coordinate systems. Imaginary lines are hatched only for illustration.

A second-order curve is called a ***central conic*** if it has a unique center of symmetry. Otherwise, if a center of symmetry does not exist or is not unique, a line is called ***non-central***. Central conics are ellipses (real and imaginary), hyperbola, a pair of intersecting lines (real and imaginary). Other curves are non-central.

10.2.2. Ellipse

An ***ellipse*** is a locus of points on plane for each of which the sum of distances to two given points F_1 and F_2 is constant ($2a$), and bigger than the distance ($2c$) between these given points (Fig. 10.12, *a*). Points F_1 and F_2 are called ***focal points*** (***foci***), the distance between them $2c = F_1F_2$ – ***focal distance***, midpoint O of segment F_1F_2 – ***center*** of the ellipse. Segments F_1M and F_2M that connect an arbitrary point M of the ellipse with its foci are called ***focal radiuses*** of point M.

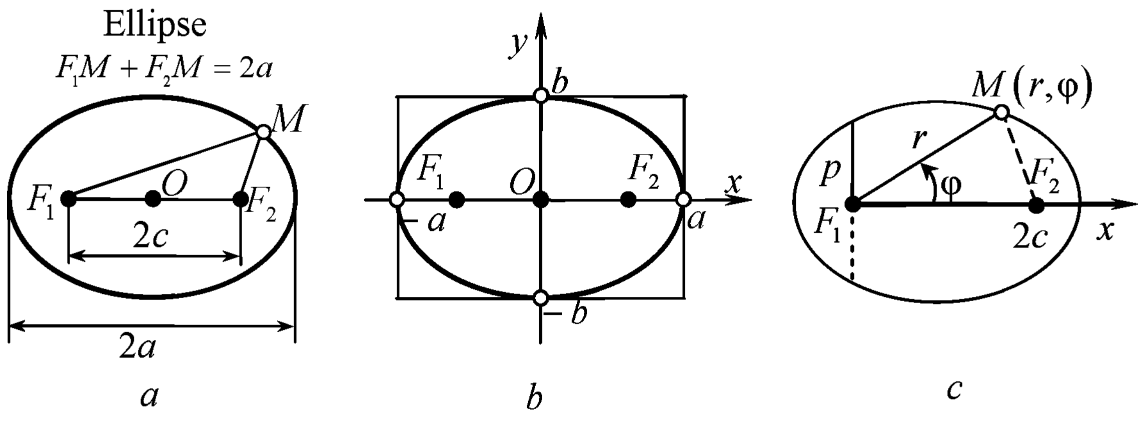


Figure 10.12

Proportion $e = \dfrac{c}{a}$ is called an ***eccentricity*** of an ellipse. By definition ($2a > 2c$) it follows that $0 \leq e < 1$. The bigger is e, the more elongated an ellipse gets. For $e = 0$, i.e. for $c = 0$, foci F_1 and F_2, as well as the center O coincide, and ellipse is a *circle of radius* a.

In a canonical coordinate system, chosen as illustrated on Fig. 10.12, *b*, an ellipse can be given by ***canonical equation***

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,$$

where $b=\sqrt{a^2-c^2}$.

Coordinate axes (of the canonical coordinate system) are the ***axes of symmetry*** of the ellipse (called ***principal axes***, the larger of these two axes is called the ***major axis***, the smaller – the ***minor axis***), and its center – ***center of symmetry***. Segments a and b are called ***semi-major*** and ***semi-minor axis*** of the ellipse, respectively, proportion $k=\frac{b}{a}\le 1$ is called an ***aspect ratio***. Lines $x=\pm a$, $y=\pm b$ bound on the coordinate plane the ***principal rectangle***, inside of which the ellipse is situated (see Fig.10.12, *b*). Points where coordinate axes cross the ellipse are called the ***vertices*** of the ellipse.

Parametrical equation of an ellipse in a canonical coordinate system takes the form:

$$\begin{cases} x=a\cdot\cos t, \\ y=b\cdot\sin t, \end{cases} \quad 0\le t<2\pi.$$

Equation of an ellipse in a polar coordinate system $F_1 r\varphi$ (Fig. 10.12, *c*) takes the form:

$$r=\frac{p}{1-e\cdot\cos\varphi},$$

where $p=\frac{b^2}{a}$ is the ***focal parameter*** of an ellipse, $0\le e<1$.

Equation $\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$, $a\ge b$, defines an ellipse with the center in point $O'(x_0,y_0)$, axes of which are parallel to the coordinate axes (Fig. 10.13, *a*). This equation can be reduced to the canonical one by translation. For $a<b$ this equation defines an ellipse, foci of which are situated on an axis parallel to the Oy axis

(Fig. 10.13, *b*). In this case the equation can be reduced to the canonical one by translation and changing the names of coordinate axes (see Section 9.1.2).

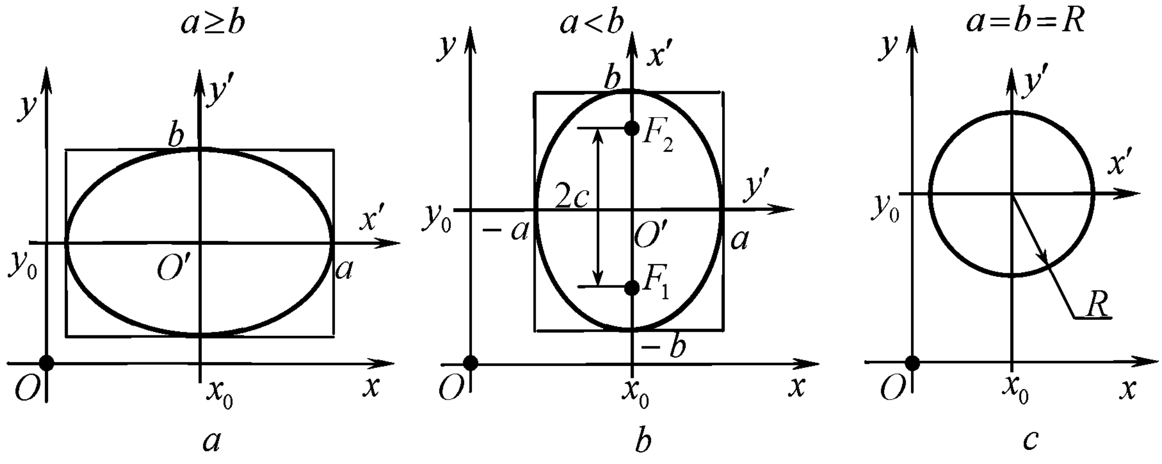


Figure 10.13

For $a = b = R$ equation $(x - x_0)^2 + (y - y_0)^2 = R^2$ defines a ***circle*** of radius R and center in point $O'(x_0, y_0)$ (Fig. 10.13, *c*).

Example 10.9. Sketch ellipses

a) $\dfrac{x^2}{2^2} + \dfrac{y^2}{1^2} = 1$;

b) $\dfrac{(x-1)^2}{2^2} + \dfrac{(y-2)^2}{1^2} = 1$;

c) $\dfrac{(x-3)^2}{1^2} + \dfrac{(y+1)^2}{2^2} = 1$

in the given (Oxy) and canonical $(O'x'y')$ coordinate systems. Find semi-axes, focal distance, eccentricity, aspect ratio and focal parameter.

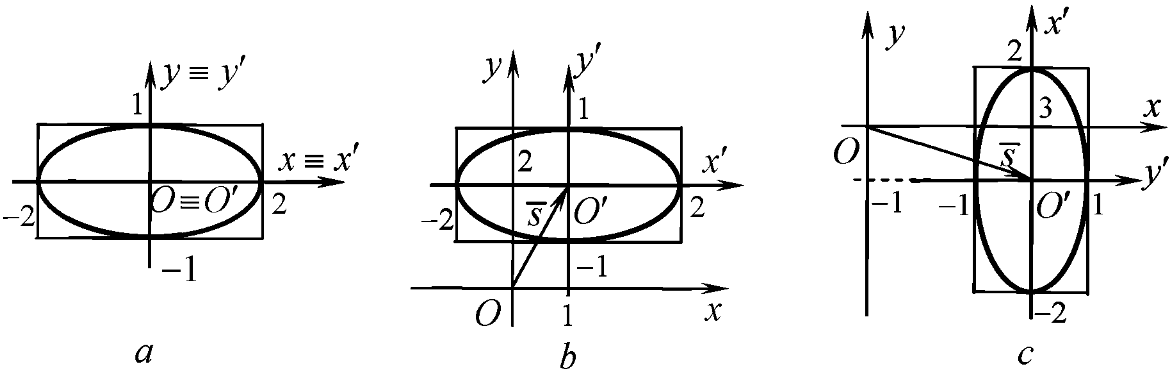


Figure 10.14

□ a) Coordinate system Oxy is canonical, since the given equation is in canonical form. By the equation define semi-axes: $a=2$ is the semi-major axis, $b=1$ – the semi-minor axis. Built the principal rectangle with sides $2a=4$, $2b=2$ and center in the origin (Fig. 10.14, *a*). Taking into account the symmetry of the ellipse, inscribe it into the principal rectangle. Find the aspect ratio $k=\frac{b}{a}=\frac{1}{2}$; focal distance $2\cdot c=2\cdot\sqrt{a^2-b^2}=2\cdot\sqrt{2^2-1^2}=2\sqrt{3}$; eccentricity $e=\frac{c}{a}=\frac{\sqrt{3}}{2}$; focal parameter $p=\frac{b^2}{a}=\frac{1^2}{2}=\frac{1}{2}$.

b) Comparing the given equation to the equation of an ellipse $\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$, we obtain $x_0=1$, $y_0=2$, $a=2$, $b=1$. Taking into account Fig. 10.13, *a*, sketch the given ellipse in the given and canonical coordinate systems (Fig. 10.14, *b*).

Note, that the canonical coordinate system $O'x'y'$ is obtained from the given one after translating it by the vector $\overline{s}=\overline{i}+2\cdot\overline{j}$. In other words, change of unknowns $x=1+x'$, $y=2+y'$ converts the equation to the canonical form: $\frac{(x')^2}{2^2}+\frac{(y')^2}{1^2}=1$. Since the canonical equation of the ellipse is the same as in "a", all the other parameters of these ellipses are the same: $k=\frac{1}{2}$; $2\cdot c=2\sqrt{3}$; $e=\frac{\sqrt{3}}{2}$; $p=\frac{1}{2}$.

c) Comparing the given equation to the equation of an ellipse $\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$, we obtain $x_0=3$, $y_0=-1$, $a=1$, $b=2$. Taking into account Fig. 10.13, *a*, sketch the given ellipse in the given and canonical coordinate systems (Fig. 10.14, *c*).

Note, that the canonical coordinate system $O'x'y'$ is obtained from the given one after translating it by the vector $\overline{s}=3\cdot\overline{i}-\overline{j}$ and changing the names of the axes. In other words, change of unknowns $x=3+y'$, $y=-1+x'$ converts the equation to the canonical form: $\frac{(x')^2}{2^2}+\frac{(y')^2}{1^2}=1$. Since the canonical equation of the ellipse is the same as in "a", all the other parameters of these ellipses are the same: $k=\frac{1}{2}$; $2\cdot c=2\sqrt{3}$; $e=\frac{\sqrt{3}}{2}$; $p=\frac{1}{2}$. ■

10.2.3. Hyperbola

A ***hyperbola*** is a locus of points on plane for each of which the module of the difference of distances to two given points F_1 and F_2 is constant $(2a)$ and smaller than the distance between these given points $(2c)$ (Fig. 10.15,*a*).

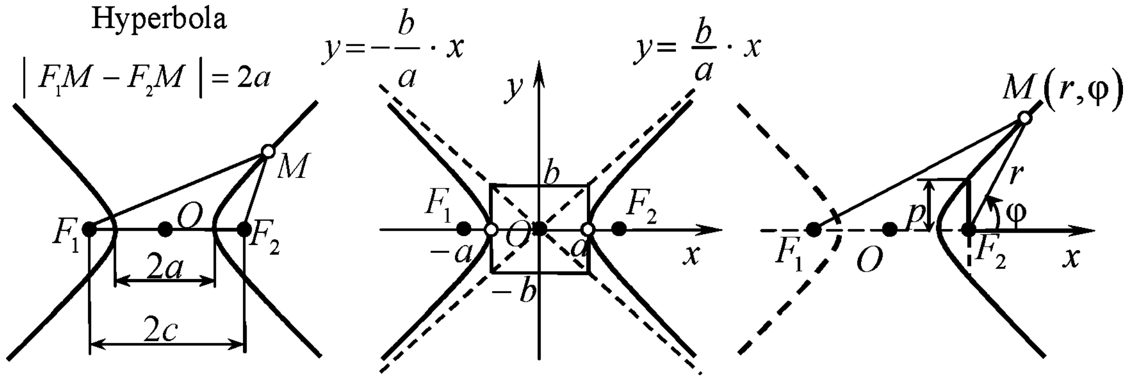


Figure 10.15

Points F_1 and F_2 are called the ***foci*** of the hyperbola, the distance $2c=F_1F_2$ between them – ***focal distance***, midpoint O of segment F_1F_2 – ***center*** of the hyperbola. Segments F_1M and F_2M that connect an arbitrary point M of the hyperbola with its foci are called ***focal radiuses*** of point M. Proportion $e=\frac{c}{a}$ is called an ***eccentricity*** of the hyperbola. By definition $(2a<2c)$ it follows that $e>1$. Eccentricity e define the form of the hyperbola. Bigger values of e correspond to hyperbolas with wider branches, while values closer to 1 correspond to hyperbolas with more narrow branches.

In a canonical coordinate system, chosen as illustrated on Fig. 10.15, *b*, a hyperbola can be given by ***canonical equation***

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1,$$

where $b=\sqrt{c^2-a^2}$.

Coordinate axes (of the canonical coordinate system) are the ***axes of symmetry*** of the hyperbola (called ***principal axes***), and its center – ***center of symmetry***, a – ***real semi-axis***, b – ***imaginary semi-axis*** of the hyperbola. Lines $x=\pm a$, $y=\pm b$ bound on the coordinate plane the ***principal rectangle***, outside of which the hyperbola is situated (see Fig. 10.15, *b*). Points where coordinate axes cross the ellipse are called the ***vertices*** of the hyperbola. Lines $y=\pm\frac{b}{a}\cdot x$ that contain the diagonals of the main rectangle, are called the ***asymptotes*** of the hyperbola (see Fig. 10.15, *b*).

The ***equation*** of the right branch ***of a hyperbola in a polar coordinate system*** $F_2 r\varphi$ (Fig.10.15,*c*) may be written as $r=\frac{p}{1-e\cdot\cos\varphi}$, where $p=\frac{b^2}{a}$ is the ***focal parameter*** of the hyperbola, $e>1$.

Parametrical equation of a hyperbola in a canonical coordinate system takes the form:

$$\begin{cases} x=a\cdot\cosh t, \\ y=b\cdot\sinh t, \end{cases} \quad t\in\mathbb{R},$$

where $\cosh t=\frac{e^t+e^{-t}}{2}$ is hyperbolic cosine and $\sinh t=\frac{e^t-e^{-t}}{2}$ is hyperbolic sine.

Equation $\frac{(x-x_0)^2}{a^2}-\frac{(y-y_0)^2}{b^2}=1$ defines a hyperbola with the center in point $O'(x_0,y_0)$, axes of which are parallel to the coordinate axes (Fig. 10.16,*a*). This equation can be reduced to the canonical one by translation.

Equation $-\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$ defines a ***conjugate hyperbola*** (Fig.10.16, *b*) with the center in point $O'(x_0,y_0)$. This equation can be reduced to the canonical one by translation and changing the names of coordinate axes (see Section 9.1.2).

$$\frac{(x-x_0)^2}{a^2}-\frac{(y-y_0)^2}{b^2}=1 \qquad -\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$$

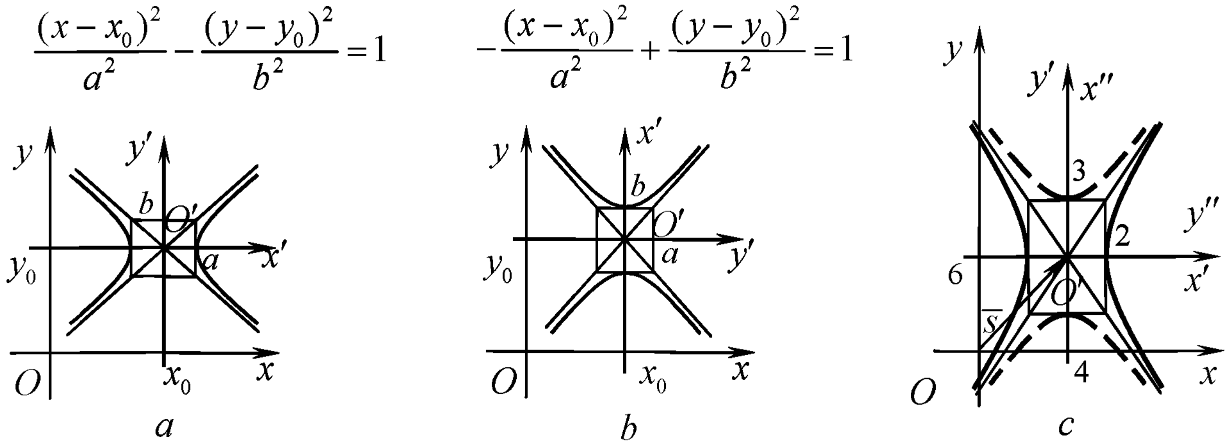


Figure 10.16

Example 10.10. Sketch hyperbolas

a) $\frac{(x-4)^2}{2^2}-\frac{(y-6)^2}{3^2}=1$;

b) $\frac{(y-6)^2}{3^2}-\frac{(x-4)^2}{2^2}=1$

in the given (Oxy) and canonical coordinate systems. Find semi-axes, focal distance, eccentricity, focal parameter, and equations of the asymptotes.

☐ a) Comparing the given equation to the equation of a hyperbola $\frac{(x-x_0)^2}{a^2}-\frac{(y-y_0)^2}{b^2}=1$, we obtain $x_0=4$, $y_0=6$, $a=2$, $b=3$. Taking into account Fig.10.18,*a*, built the principal rectangle with sides $2a=4$, $2b=6$ and center in the origin of the canonical coordinate system. Draw the asymptotes by extending the diagonals of the principal rectangle. Build the hyperbola, taking into account its symmetry with respect to the coordinate (drawn in full line of Fig.10.16, *c*).

Note, that the canonical coordinate system $O'x'y'$ is obtained from the given one after translating it by the vector $\overline{s}=4\cdot\overline{i}+6\cdot\overline{j}$. In other words, change of unknowns $x=4+x'$, $y=6+y'$ converts the equation to the canonical form: $\frac{(x')^2}{2^2}-\frac{(y')^2}{3^2}=1$. Calculate the focal distance $2\cdot c=2\cdot\sqrt{a^2+b^2}=$ $=2\cdot\sqrt{2^2+3^2}=2\sqrt{13}$; eccentricity $e=\frac{c}{a}=\frac{\sqrt{13}}{2}$; focal parameter $p=\frac{b^2}{a}=\frac{3^2}{2}=4{,}5$.

Write equations of the asymptotes $y - y_0 = \pm\frac{b}{a}\cdot(x - x_0)$, i.e. $y - 6 = \pm\frac{3}{2}\cdot(x - 4)$.

b) Comparing the given equation to the equation of a hyperbola $-\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1$, obtain parameters $x_0 = 4$, $y_0 = 6$, $a = 2$, $b = 3$ of a hyperbola, conjugate with the one built in “a”. Taking into account Fig.10.16, b, build the principal rectangle and the asymptotes as in “a”, and then build the conjugate hyperbola (hatched on Fig.10.16, *c*).

Note, that the canonical coordinate system $O'x''y''$ is obtained from the given one after translating it by the vector $\overline{s} = 4\cdot\overline{i} + 6\cdot\overline{j}$ and changing the names of the coordinate axes. In other words, change of unknowns $x = 4 + y''$, $y = 6 + x''$ converts the equation to the canonical form: $\frac{(x'')^2}{3^2} - \frac{(y'')^2}{2^2} = 1$ (here $a = 3$, $b = 2$). Calculate the focal distance $2\cdot c = 2\cdot\sqrt{a^2 + b^2} = 2\cdot\sqrt{3^2 + 2^2} = 2\sqrt{13}$; eccentricity $e = \frac{c}{a} = \frac{\sqrt{13}}{3}$; focal parameter $p = \frac{b^2}{a} = \frac{2^2}{3} = \frac{4}{3}$. Equations of the asymptotes are the same as in “a”. ■

10.2.4. Parabola

A ***parabola*** is a locus of points on plane that are equidistant from a given point F and a given line d that does not pass through this point. Point F is are called the ***focus*** of the parabola, line d – ***directrix*** of the parabola, midpoint O of the perpendicular dropped from the focus on the directrix – ***vertex*** of the parabola, distance p between the focus and the directrix – ***parameter*** of the parabola, and distance $\frac{p}{2}$ between the vertex and the focus – ***focal length*** (Fig. 10.17, *a*). Parameter p of the parabola define its form. Bigger values of p correspond to parabolas with wider branches, while values closer to 0 correspond to parabolas with more narrow branches.

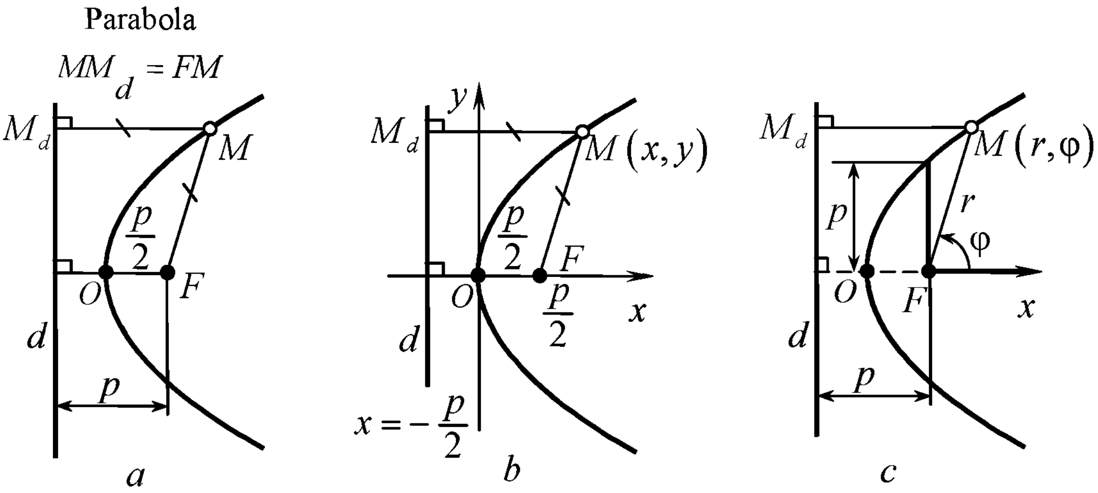


Figure 10.17

A line, perpendicular to the directrix and passing through the focus, is called the ***axis*** of the parabola (***focal axis***). Segment FM that connects an arbitrary point M of the parabola with its focus is called a ***focal radius*** of point M. ***Eccentricity*** of a parabola equals one by definition ($e=1$).

In a canonical coordinate system, chosen as illustrated on Fig. 10.17, *b*, a parabola can be given by ***canonical equation***

$$y^2 = 2\cdot p\cdot x.$$

In this coordinate system equation of the directrix is $x=-\frac{p}{2}$, coordinates of the focus are $F\left(\frac{p}{2},0\right)$. The axes of the canonical coordinate system are called the ***principal axes*** of the parabola.

Equation of a parabola in a polar coordinate system $Fr\varphi$ (Fig. 10.17, *c*) takes the form:

$$r = \frac{p}{1-e\cdot\cos\varphi},$$

where p is the parameter of a parabola, $e=1$ is its eccentricity.

Equation $(y-y_0)^2 = 2\cdot p\cdot(x-x_0)$, $p\neq 0$, defines a parabola with vertex $O'(x_0,y_0)$, the axis of which is parallel to the abscissa axis: for $p>0$ the directions of axes Ox and $O'x'$ are the same (Fig. 10.18, *a*), and for $p<0$ they are opposite

(Fig. 10.18, *b*). This equation can be reduced to the canonical one by translation (and changing the direction of the abscissa axis if $p<0$).

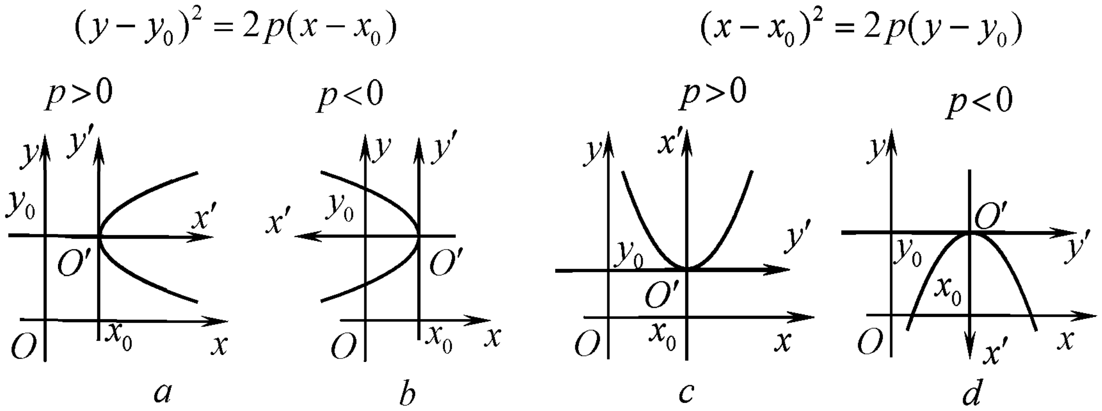


Figure 10.18

Equation $(x-x_0)^2 = 2\cdot p\cdot(y-y_0)$, $p \neq 0$, also defines a parabola with vertex $O'(x_0, y_0)$, the axis of which is parallel to the ordinate axis: for $p>0$ the directions of axes Oy and $O'x'$ are the same (Fig. 10.18, *c*), and for $p<0$ they are opposite (Fig. 10.18, *d*). This equation can be reduced to the canonical one by translation, changing the names of the coordinate axes (and changing the direction of the ordinate axis if $p<0$).

Example 10.11. Sketch parabolas

a) $y^2 = 2\cdot x$; b) $(y-1)^2 = -2\cdot(x-2)$; c) $(x-2)^2 = 2\cdot(y+1)$;

in the given (Oxy) and canonical ($O'x'y'$) coordinate systems. Find the parameter of the parabola, coordinates of the focus and equation of the directrix

□ a) Coordinate system Oxy is canonical, since the given equation is in canonical form. From the equation obtain the parameter $p=1$. Build the parabola, taking into account its symmetry with respect to the abscissa axis (Fig. 10.19, *a*). Coordinates of the focus are $x_F = \frac{p}{2} = \frac{1}{2}$, $y_F = 0$, i.e. $F\left(\frac{1}{2}, 0\right)$. Write the equation of the directrix $x = -\frac{p}{2}$, i.e. $x = -\frac{1}{2}$.

b) Comparing the given equation to the equation of a parabola $(y-y_0)^2 = 2\cdot p\cdot(x-x_0)$, we obtain $x_0 = 2$, $y_0 = 1$, $p = -1 < 0$. Taking into account

Fig. 10.20, b, build a parabola, symmetric with respect to axis $O'x'$ (Fig. 10.19, b).

Note, that the canonical coordinate system $O'x'y'$ is obtained from the given one after translating it by the vector $\overline{s} = 2\cdot\overline{i} + \overline{j}$ and changing the direction of the abscissa axis. In other words, change of unknowns $x = 2 - x'$, $y = 1 + y'$ converts the equation to the canonical form: $(y')^2 = 2\cdot 1\cdot x'$. Since the canonical equation of the parabola is the same as in "a", the value of the parameter, the equation of the directrix $x' = -\frac{1}{2}$ and coordinates $x'_F = \frac{p}{2} = \frac{1}{2}$, $y'_F = 0$ of the focus are the same as the ones obtained in "a".

c) Comparing the given equation to the equation of a parabola $(x - x_0)^2 = 2\cdot p\cdot(y - y_0)$, we obtain $x_0 = 2$, $y_0 = -1$, $p = 1 > 0$. Taking into account Fig. 10.18, c, build a parabola, symmetric with respect to axis $O'x'$ (Fig. 10.19, c).

Note, that the canonical coordinate system $O'x'y'$ is obtained from the given one after translating it by the vector $\overline{s} = 2\cdot\overline{i} - \overline{j}$ and changing the names of the coordinate axes. In other words, change of unknowns $x = 2 + y'$, $y = -1 + x'$ converts the equation to the canonical form: $(y')^2 = 2\cdot 1\cdot x'$. Since the canonical equation of the parabola is the same as in "a" and "b", the equation of the directrix $x' = -\frac{1}{2}$ and coordinates $x'_F = \frac{p}{2} = \frac{1}{2}$, $y'_F = 0$ of the focus are the same as the ones obtained in "a" and "b". ■

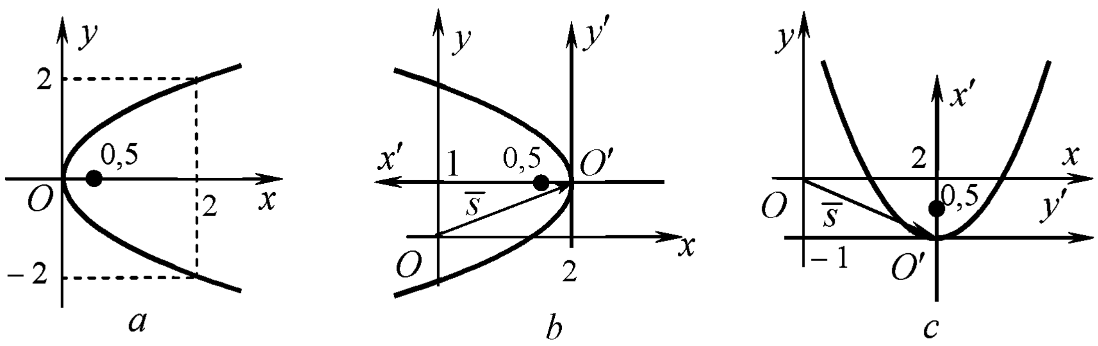


Figure 10.19

EXERCISES

1. For the line, which passes through the points $A(1, 4)$ and $B(2, 0)$, compose:

a) general equation; b) parametric equation; c) canonical equation;

d) intercept equation;slope-intercept equation.

2. Find information about positional relationship of each pair of lines (are they skew, intersecting, parallel, equal, perpendicular, if they are intersecting find their mutual point):

a) $x+y-3=0$, $2\cdot x+3\cdot y-8=0$;

b) $y=5\cdot x-24$, $y=-0{,}2\cdot x+2$;

c) $\begin{cases} x=5+4\cdot t, \\ y=-2-2\cdot t, \end{cases}$ $\dfrac{x-1}{-2}=\dfrac{y-7}{1}$;

d) $4\cdot x+5\cdot y-6=0$, $\dfrac{x+6}{5}=\dfrac{y-6}{-4}$.

3. On coordinate plane Oxy sketch ellipses:

a) $\dfrac{(x-n)^2}{(m+n)^2}+\dfrac{(y+m)^2}{n^2}=1$; b) $\dfrac{(x-n)^2}{m^2}+\dfrac{(y-m)^2}{(m+n)^2}=1$.

For each ellipse find its focal distance, aspect ratio, focal parameter and eccentricity, coordinates of center, focuses and vertexes.

4. On coordinate plane Oxy sketch hyperbolas:

a) $\dfrac{(x+n)^2}{m^2}-\dfrac{(y-m)^2}{n^2}=1$; b) $\dfrac{(y+n)^2}{m^2}-\dfrac{(x+m)^2}{n^2}=1$.

For each hyperbola find its focal distance, focal parameter and eccentricity, coordinates of center, focuses and vertexes, equations of asymptotes.

5. On coordinate plane Oxy sketch parabolas:

a) $(y-m)^2=2\cdot n\cdot x$; b) $(y+m)^2=2\cdot(n-x)$; c) $(x-m)^2=2\cdot m\cdot(y+n)$.

For each parabola find its parameter, coordinates of vertex and focus, equation of directrix.

CHAPTER 11. ALGEBRAIC LINES AND SURFACES IN SPACE

11.1. FIRST–ORDER SURFACES (PLANES)

11.1.1. Main Types of Plane Equations

A nonzero vector $\bar{n}$, which is perpendicular to the given plane, is called ***normal vector*** (or simply ***normal***) for this plane.

Recall, that three of more vectors are called ***coplanar***, if there exists a plane, that they are parallel to. We will call this plane ***coplanar to the given vectors***.

Direction vectors of a plane are two noncollinear vectors, which are ***coplanar to the given plane***, i.e. they belong to the plane or they are parallel to it.

General (point-normal) equation of a plane:

$$A \cdot x + B \cdot y + C \cdot z + D = 0, \quad A^2 + B^2 + C^2 \neq 0. \tag{11.1}$$

Way of representation: plane passes through point $M_0(x_0, y_0, z_0)$ and it is perpendicular to vector $\bar{n} = A \cdot \bar{i} + B \cdot \bar{j} + C \cdot \bar{k}$ (fig.11.1, *a*).

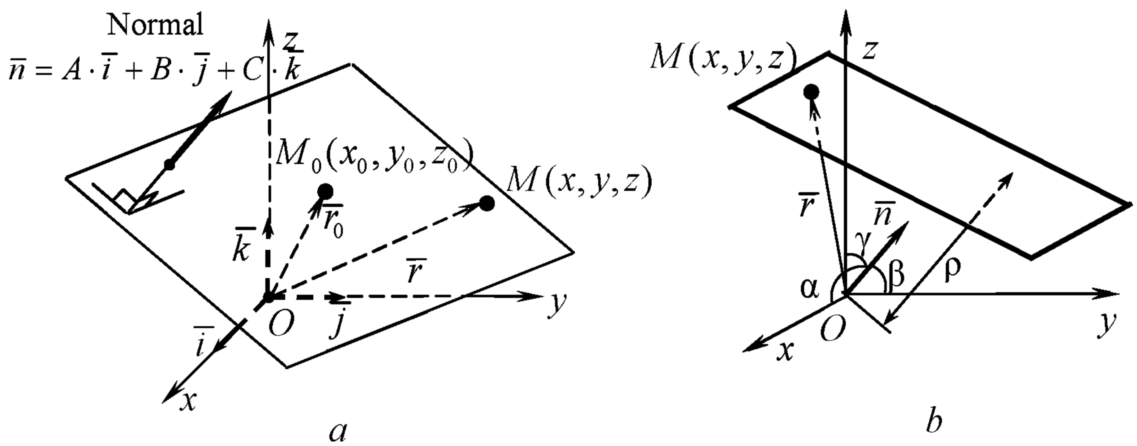


Figure 11.1

Geometric sense of coefficients: leading coefficients A, B, C are coordinates of the normal $\bar{n} = A \cdot \bar{i} + B \cdot \bar{j} + C \cdot \bar{k}$; constant term $D = -A \cdot x_0 - B \cdot y_0 - C \cdot z_0$. Denoting radius vectors of points $M_0(x_0, y_0, z_0)$ and $M(x,y,z)$ by $\bar{r}_0$ and $\bar{r}$ accordingly, it is possible to write ***vector equation of a plane***, which passes through the point $M_0(x_0, y_0, z_0)$ and which is perpendicular to normal $\bar{n} = A \cdot \bar{i} + B \cdot \bar{j} + C \cdot \bar{k}$:

$$(\bar{r}-\bar{r}_0,\bar{n})=0.$$

Right zero part of the scalar product denominates perpendicularity condition of vectors $\bar{r}-\bar{r}_0$ and $\bar{n}$ (Section 8.7). In coordinate form equation can be expressed in the following form:

$$A\cdot(x-x_0)+B\cdot(y-y_0)+C\cdot(z-z_0)=0. \qquad (11.2)$$

Normalized equation of a plane:

$$x\cdot\cos\alpha+y\cdot\cos\beta+z\cdot\cos\gamma-\rho=0,\ \ \rho\geq 0. \qquad (11.3)$$

Way of representation: plane passes through the point $M_0(x_0,y_0,z_0)$ and it is perpendicular to the vector $\bar{n}=\cos\alpha\cdot\bar{i}+\cos\beta\cdot\bar{j}+\cos\gamma\cdot\bar{k}$ (Fig. 11.1, *a*).

Geometric sense of coefficients: leading coefficients $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are direction cosines of normal $\bar{n}=\cos\alpha\cdot\bar{i}+\ \ +\cos\beta\cdot\bar{j}+\cos\gamma\cdot\bar{k}$; constant term $\rho=\dfrac{|A\cdot x_0+B\cdot y_0+C\cdot z_0|}{\sqrt{A^2+B^2+C^2}}$ is the distance between the coordinate origin and the plane (Fig. 11.1, *b*).

Vector parametric equation of a plane:

$$\bar{r}=\bar{r}_0+t_1\cdot\bar{p}_1+t_2\cdot\bar{p}_2,\quad t_1,t_2\in\mathbb{R},\quad [\bar{p}_1,\bar{p}_2]\neq\bar{o}. \qquad (11.4)$$

Way of representation: plane passes through the point $M_0(x_0,y_0,z_0)$, which is defined by radius vector $\bar{r}_0$, and it is coplanar to two direction vectors $\bar{p}_1$, $\bar{p}_2$ (Fig.11.2). Parameters t_1, t_2 in equation (11.4) have the following geometric sense: values t_1, t_2 are proportional to the distance between the given point $M_0(x_0,y_0,z_0)$ and the point $M(x,y,z)$, which is defined by radius vector $\bar{r}$. If $t_1=t_2=0$, then the point $M(x,y,z)$ coincides with the given point $M_0(x_0,y_0,z_0)$: $\bar{r}=\bar{r}_0$. Increase of t_1 (or t_2) results in the shift of the point $M(x,y,z)$ to the direction defined by vector $\bar{p}_1$ (or $\bar{p}_2$), and decrease of t_1 (or t_2) – to the opposite direction.

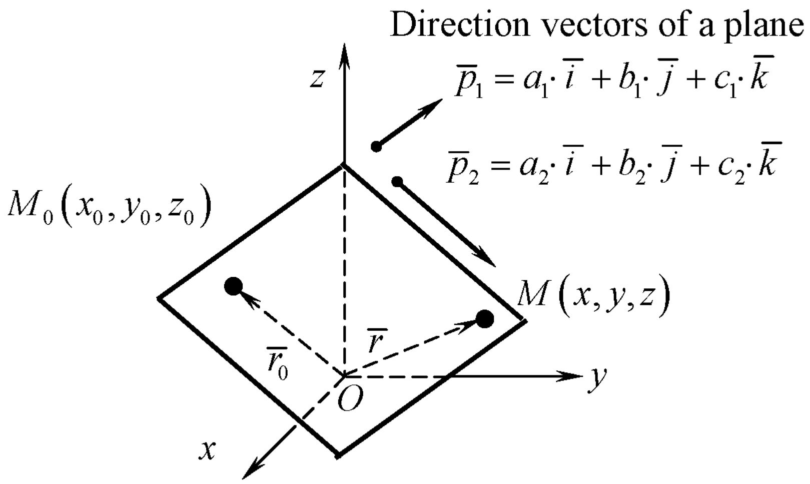


Figure 11.2

Parametric equation of a plane:

$$\begin{cases} x = x_0 + a_1 \cdot t_1 + a_2 \cdot t_2 \,, \\ y = y_0 + b_1 \cdot t_1 + b_2 \cdot t_2 \,, \\ z = z_0 + c_1 \cdot t_1 + c_2 \cdot t_2 \,, \end{cases} \quad t_1, t_2 \in \mathbb{R}, \quad \operatorname{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2\,. \tag{11.5}$$

Way of representation: plane passes through the point $M_0(x_0, y_0, z_0)$ and it is coplanar to two noncollinear vectors $\overline{p}_1 = a_1 \cdot \overline{i} + b_1 \cdot \overline{j} + c_1 \cdot \overline{k}$ and $\overline{p}_2 = a_2 \cdot \overline{i} + b_2 \cdot \overline{j} + c_2 \cdot \overline{k}$ (Fig.11.2).

Geometric sense of coefficients: a_1, b_1, c_1 and a_2, b_2, c_2 are coordinates of direction vectors $\overline{p}_1 = a_1 \cdot \overline{i} + b_1 \cdot \overline{j} + c_1 \cdot \overline{k}$, $\overline{p}_2 = a_2 \cdot \overline{i} + b_2 \cdot \overline{j} + \; + c_2 \cdot \overline{k}$, and x_0, y_0, z_0 are coordinates of the point $M_0(x_0, y_0, z_0)$, which belongs to the plane. Parameters t_1, t_2 have the same sense as in equation (11.4).

Note that the equation (11.5) is a coordinate form of the equation (11.4).

Equation of a plane, which passes through the given point and is coplanar to two noncollinear vectors:

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} = 0, \quad \operatorname{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2\,. \tag{11.6}$$

Way of plane representation and geometric sense of coefficients in equation (11.6) are the same as in equation (11.5). Conditions $[\overline{p}_1, \overline{p}_2] \neq \overline{o}$ in (11.4) and

$\operatorname{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2$ in (11.5), (11.6) denominate noncollinearity property of vectors $\bar{p}_1$ and $\bar{p}_2$.

Affine equation of a plane, which passes through the given three points:

$$\bar{r} = (1 - t_1 - t_2) \cdot \bar{r}_0 + t_1 \cdot \bar{r}_1 + t_2 \cdot \bar{r}_2, \quad t_1, t_2 \in \mathbb{R}. \tag{11.7}$$

Equation (11.7) can be rewritten in coordinate form:

$$\begin{cases} x = (1 - t_1 - t_2) \cdot x_0 + t_1 \cdot x_1 + t_2 \cdot x_2, \\ y = (1 - t_1 - t_2) \cdot y_0 + t_1 \cdot y_1 + t_2 \cdot y_2, \\ z = (1 - t_1 - t_2) \cdot z_0 + t_1 \cdot z_1 + t_2 \cdot z_2, \end{cases} \quad t_1, t_2 \in \mathbb{R}. \tag{11.8}$$

Way of representation: plane passes through the three given points $M_0(x_0, y_0, z_0)$, $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, which are defined by radius vectors $\bar{r}_0$, $\bar{r}_1$ and $\bar{r}_2$ accordingly (Fig.11.3, *a*). Radius vector $\bar{r}$ defines the position of point $M(x, y, z)$, which belongs to the plane.

Geometric sense of coefficients: x_0, y_0, z_0; x_1, y_1, z_1; x_2, y_2, z_2 – coordinates of points $M_0(x_0, y_0, z_0)$, $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$, through which the plane (11.8) passes. Parameters t_1, t_2 in equation (11.7) define the position of point $M(x, y, z)$, which belongs to the plane, e.g. if $t_1 = 0$, $t_2 = 1$, then M coincides with M_2, and if $t_1 = 1$, $t_2 = 0$ – with M_1.

Equation of a plane, which passes through three given points:

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{vmatrix} = 0, \quad \operatorname{rg}\begin{pmatrix} x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{pmatrix} = 2. \tag{11.9}$$

Way of plane representation and geometric sense of coefficients in equation (11.9) are the same as in equation (11.8).

Intercept equation of a plane:

$$\frac{x}{x_1} + \frac{y}{y_1} + \frac{z}{z_1} = 1, \quad x_1 \neq 0, \ y_1 \neq 0, \ z_1 \neq 0. \tag{11.10}$$

Way of representation: plane passes through three given points $X_1(x_1, 0, 0)$, $Y_1(0, y_1, 0)$ и $Z_1(0, 0, z_1)$, and $x_1 \neq 0$, $y_1 \neq 0$, $z_1 \neq 0$ (Fig. 11.3, *b*).

Geometric sense of coefficients: plane (11.10) cuts off "segments" on coordinate axes: x_1 on abscissa axis, y_1 on ordinate axis, z_1 on applicate axis.

Ways of Equation Type Transformation

1. To transform general equation of a plane (11.1) into normalized equation (11.3) it is sufficient to divide both parts of general equation by the length of normal $|\bar{n}|=\sqrt{A^2+B^2+C^2}$ (if constant term is negative $D<0$) or by opposite quantity $-|\bar{n}|= =-\sqrt{A^2+B^2+C^2}$ (if constant term is nonnegative $D\geq 0$).

2. To transform general equation of a plane (11.1) to parametric equation (11.5) it is necessary to make the following steps:

1) find any solution (x_0,y_0,z_0) of equation $A\cdot x+B\cdot y+C\cdot z+D=0$, defining the coordinates of a point $M_0(x_0,y_0,z_0)$, which belongs to the plane;

2) find any two linearly independent solutions (a_1,b_1,c_1), (a_2,b_2,c_2) of homogeneous equation $A\cdot a+B\cdot b+C\cdot c=0$, defining the coordinates a_1,b_1,c_1 and a_2,b_2,c_2 of direction vectors $\bar{p}_1$ and $\bar{p}_2$ of the plane;

3) write parametric equation (11.5).

3. To transform parametric equation into general, it is sufficient to write the equation (11.6) and expand the determinant or to find normal as the outer product of direction vectors (Section 8.5):

$$\bar{n}=[\bar{p}_1,\bar{p}_2]=\begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}=\underbrace{\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}}_{A}\cdot\bar{i}-\underbrace{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}_{B}\cdot\bar{j}+\underbrace{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}_{C}\cdot\bar{k},$$

and write general equation of a plane in a form (11.2):

$$A\cdot(x-x_0)+B\cdot(y-y_0)+C\cdot(z-z_0)=0.$$

4. Transformation of general equation of a plane (11.1) into intercept equation (11.10) is possible if all coefficients of general equation are nonzero. To do this transformation it is necessary to transfer constant term to the right part of equation: $A\cdot x+B\cdot y+C\cdot z=-D$ and then divide both parts of equation by $-D$:

$\frac{A}{-D}\cdot x+\frac{B}{-D}\cdot y+\frac{C}{-D}\cdot z=1$. Denoting $x_1=-\frac{D}{A}$, $y_1=-\frac{D}{B}$, $z_1=-\frac{D}{C}$, we will get intercept equation (11.10): $\frac{x}{x_1}+\frac{y}{y_1}+\frac{z}{z_1}=1$.

Example 11.2. In coordinate space $Oxyz$ (in Cartesian coordinate system) the following points are given: $K(2,3,4)$, $L(6,-3,4)$, $M(-4,6,-4)$. Find:

a) general equation of a plane, which contains triangle KLM;

b) intercept equation of triangle KLM plane;

c) points of plane and coordinate axes intersection.

□ a) Compose the equation (11.9):

$$\begin{vmatrix} x-2 & y-3 & z-4 \\ 6-2 & -3-3 & 4-4 \\ -4-2 & 6-3 & -4-4 \end{vmatrix}=0 \quad\Leftrightarrow\quad \begin{vmatrix} x-2 & y-3 & z-4 \\ 4 & -6 & 0 \\ -6 & 3 & -8 \end{vmatrix}=0.$$

By the determinant expansion and similar term simplification we get

$$48\cdot(x-2)+32\cdot(y-3)-24\cdot(z-4)=0 \quad\Leftrightarrow\quad 6\cdot x+4\cdot y-3\cdot z-12=0.$$

b) By transferring the constant term of general equation to its right part and dividing by 12: $\frac{x}{2}+\frac{y}{3}+\frac{z}{-4}=1$ we have obtained intercept equation.

c) By the intercept equation we find that the plane passes through the following points: $X(2,0,0)$, $Y(0,3,0)$, $Z(0,0,-4)$ on coordinate axes. ■

11.1.2. Planes Positional Relationships

Consider two planes π_1 and π_2, which are defined by the following general equations:

$$\pi_1:\ A_1\cdot x+B_1\cdot y+C_1\cdot z+D_1=0; \qquad \pi_2:\ A_2\cdot x+B_2\cdot y+C_2\cdot z+D_2=0.$$

Information about plane positional relationship can be obtained from the coefficients in their equations by the following criteria of:

- plane *parallelism*:

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

- plane *equality*:

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

- plane *parallelism or equality*:

$$\mathrm{rg}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix}=1;$$

- plane *intersection*:

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

- plane *perpendicularity*:

$$A_1\cdot A_2+B_1\cdot B_2+C_1\cdot C_2=0.$$

If planes intersect, then the coordinates of their common points can be found as the solution of the following system of equations:

$$\begin{cases} A_1\cdot x+B_1\cdot y+C_1\cdot z+D_1=0, \\ A_2\cdot x+B_2\cdot y+C_2\cdot z+D_2=0. \end{cases}$$

This system has an infinite number of solutions, which form the plane intersection line.

Example 11.3. Describe the positional relationship of each pair of planes (intersection, parallelism, equality, perpendicularity):

a) $2\cdot x-y-4\cdot z+3=0$, $-4\cdot x+2\cdot y+8\cdot z-6=0$;

b) $2\cdot x+3\cdot y+z-6=0$, $4\cdot x+6\cdot y+2\cdot z+3=0$;

c) $3\cdot x-2\cdot y+z+1=0$, $4\cdot x+5\cdot y-2\cdot z-1=0$;

d) $3\cdot x-2\cdot y+z+1=0$, $4\cdot x+5\cdot y+2\cdot z-1=0$.

□ a) Since $A_1=2$, $B_1=-1$, $C_1=-4$, $D_1=3$, $A_2=-4$, $B_2=2$, $C_2=8$, $D_2=-6$ and $\frac{A_1}{A_2}=\frac{2}{-4}=-\frac{1}{2}$, $\frac{B_1}{B_2}=-\frac{1}{2}$, $\frac{C_1}{C_2}=\frac{-4}{8}=-\frac{1}{2}$, $\frac{D_1}{D_2}=\frac{3}{-6}=-\frac{1}{2}$, then $\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}=\frac{D_1}{D_2}=-\frac{1}{2}$. Hence, these planes coincide (they are equal).

b) Since $A_1 = 2$, $B_1 = 3$, $C_1 = 1$, $D_1 = -6$, $A_2 = 4$, $B_2 = 6$, $C_2 = 2$, $D_2 = 3$ and $\frac{A_1}{A_2} = \frac{1}{2}$, $\frac{B_1}{B_2} = \frac{1}{2}$, $\frac{C_1}{C_2} = \frac{1}{2}$, $\frac{D_1}{D_2} = -2$, then $\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} \neq \frac{D_1}{D_2}$. Hence, planes are parallel.

c) Since $A_1 = 3$, $B_1 = -2$, $C_1 = 1$, $D_1 = 1$, $A_2 = 4$, $B_2 = 5$, $C_2 = -2$, $D_2 = -1$, then $\operatorname{rg}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 3 & -2 & 1 \\ 4 & 5 & -2 \end{pmatrix} = 2$. Hence, planes intersect. Since $A_1 \cdot A_2 + B_1 \cdot B_2 + C_1 \cdot C_2 = 3 \cdot 4 + (-2) \cdot 5 + 1 \cdot (-2) = 0$, then planes are perpendicular.

d) Since $A_1 = 3$, $B_1 = -2$, $C_1 = 1$, $D_1 = 1$, $A_2 = 4$, $B_2 = 5$, $C_2 = 2$, $D_2 = -1$, то $\operatorname{rg}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 3 & -2 & 1 \\ 4 & 5 & 2 \end{pmatrix} = 2$. Hence, planes intersect. Since $A_1 \cdot A_2 + B_1 \cdot B_2 + C_1 \cdot C_2 = 3 \cdot 4 + (-2) \cdot 5 + 1 \cdot 2 = 4 \neq 0$, then planes are not perpendicular. ■

11.1.3. Metric Applications of Plane Equations

Let's list formulas for segment lengths (distances) and values of angles calculation by equations of their forming planes.

Angle between two planes can be determined as angle between their normal vectors (on Fig. 11.3 normal vectors of planes π_1 and π_2 are denoted by $\bar{n}_1$, $\bar{n}_2$ accordingly). By this definition we have two adjacent supplementary angles, which complement each other to π. In elementary geometry usually the smallest angle is chosen from two adjacent angles, i.e. value of angle φ between two planes satisfy the following condition $0 \le \varphi \le \frac{\pi}{2}$.

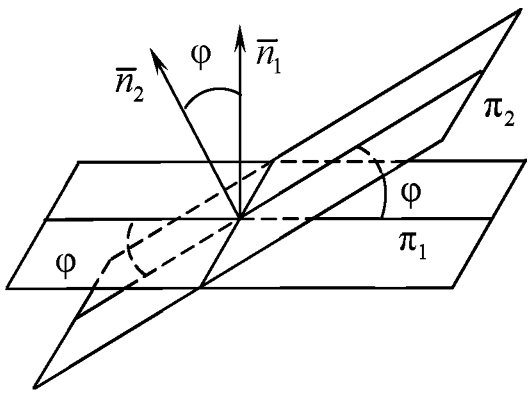


Figure 11.3

1. Distance d between point $M^*(x^*, y^*, z^*)$ and plane $A \cdot x + B \cdot y + C \cdot z + D = 0$ is calculated by the following formula (Fig. 11.4, *a*)

$$d = \frac{\left| A \cdot x^* + B \cdot y^* + C \cdot z^* + D \right|}{\sqrt{A^2 + B^2 + C^2}}.$$

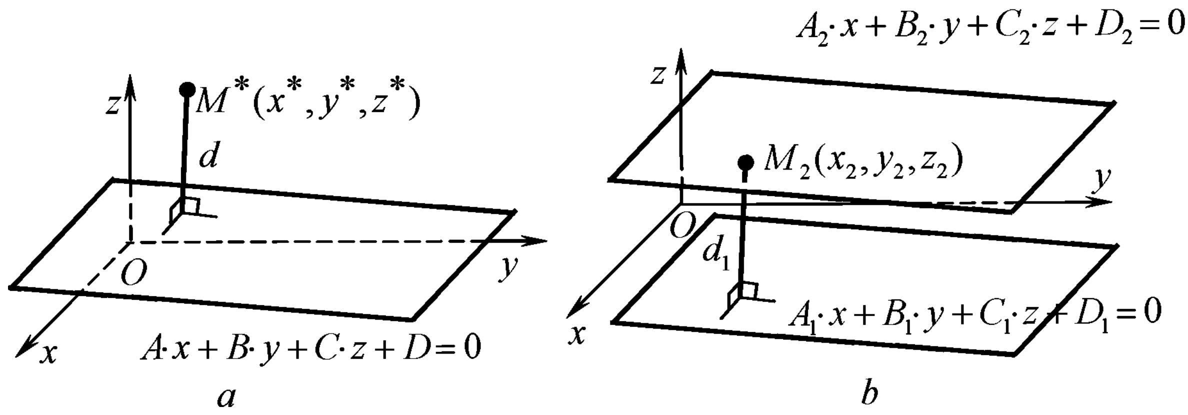


Figure 11.4

2. Distance between parallel planes $A_1 \cdot x + B_1 \cdot y + C_1 \cdot z + D_1 = 0$ and $A_2 \cdot x + B_2 \cdot y + C_2 \cdot z + D_2 = 0$ is calculated as distance d_1 between point $M_2(x_2, y_2, z_2)$, which coordinates satisfy the equation $A_2 \cdot x + B_2 \cdot y + C_2 \cdot z + D_2 = 0$, and plane $A_1 \cdot x + B_1 \cdot y + C_1 \cdot z + D_1 = 0$ by the following formula (Fig. 11.4, *b*)

$$d_1 = \frac{\left| A_1 \cdot x_2 + B_1 \cdot y_2 + C_1 \cdot z_2 + D_1 \right|}{\sqrt{A_1^2 + B_1^2 + C_1^2}}.$$

3. Acute angle φ between two planes

π_1: $A_1 \cdot x + B_1 \cdot y + C_1 \cdot z + D_1 = 0$ and π_2: $A_2 \cdot x + B_2 \cdot y + C_2 \cdot z + D_2 = 0$

is calculated by the following formula

$$\cos\varphi = \frac{\left| A_1 \cdot A_2 + B_1 \cdot B_2 + C_1 \cdot C_2 \right|}{\sqrt{A_1^2 + B_1^2 + C_1^2} \cdot \sqrt{A_2^2 + B_2^2 + C_2^2}},$$

where $\overline{n}_1 = A_1 \cdot \overline{i} + B_1 \cdot \overline{j} + C_1 \cdot \overline{k}$ and $\overline{n}_2 = A_2 \cdot \overline{i} + B_2 \cdot \overline{j} + C_2 \cdot \overline{k}$ are normals to planes π_1 and π_2 accordingly (Fig. 11.5).

Example 11.4. In coordinate space *Oxyz* the following vertexes of triangle pyramid *OABC* are given: $A(1,3,-1)$, $B(2,1,-2)$, $C(4,2,-6)$. Find:

a) general equation of a plane, which contains side ABC;

b) distance d between vertex C and side OAB;

c) value of angle φ between sides ABC and OAB.

□ a) By formula (11.9) compose equation of the plane π_{ABC}, which passes through points A,B,C:
$$\begin{vmatrix} x-1 & y-3 & z-(-1) \\ 2-1 & 1-3 & -2-(-1) \\ 4-1 & 2-3 & -6-(-1) \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x-1 & y-3 & z+1 \\ 1 & -2 & -1 \\ 3 & -1 & -5 \end{vmatrix} = 0.$$

Expanding the determinant by the first row we get
$$9\cdot(x-1)+2\cdot(y-3)+5\cdot(z+1)=0 \quad \Leftrightarrow \quad 9\cdot x+2\cdot y+5\cdot z-10=0.$$

The required equation is obtained.

b) To find distance d we compose equation of the plane, which passes through points O, A,B:
$$\begin{vmatrix} x-0 & y-0 & z-0 \\ 1-0 & 3-0 & -1-0 \\ 2-0 & 1-0 & -2-0 \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x & y & z \\ 1 & 3 & -1 \\ 2 & 1 & -2 \end{vmatrix} = 0 \quad \Leftrightarrow \quad x+z=0.$$

By the first formula of metric applications for $M^* \equiv C$ we have:
$$d = \frac{|1\cdot 4+0\cdot 2+1\cdot(-6)+0|}{\sqrt{1^2+0^2+1^2}} = \frac{2}{\sqrt{2}} = \sqrt{2}.$$

c) Acute angle φ between planes $9\cdot x+2\cdot y+5\cdot z-10=0$ and $x+z=0$ is found by the third formula of metric applications:
$$\cos\varphi = \frac{|9\cdot 1+2\cdot 0+5\cdot 1|}{\sqrt{9^2+2^2+5^2}\cdot\sqrt{1^2+0^2+1^2}} = \frac{14}{\sqrt{220}} = \frac{7}{\sqrt{55}}.$$

Consequently, $\varphi = \arccos\frac{7}{\sqrt{55}}$. ■

11.2. LINES IN SPACE

11.2.1. Main Types of Line Equations in Space

Direction vector of line is a nonzero vector, which is ***collinear to the given line***, i.e. this vector belongs or is parallel to the line. Two lines are called ***collinear***, if they are parallel or coincident.

General equation of a line in space:

$$\begin{cases} A_1 \cdot x + B_1 \cdot y + C_1 \cdot z + D_1 = 0, \\ A_2 \cdot x + B_2 \cdot y + C_2 \cdot z + D_2 = 0, \end{cases} \quad \text{rg}\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = 2. \qquad (11.11)$$

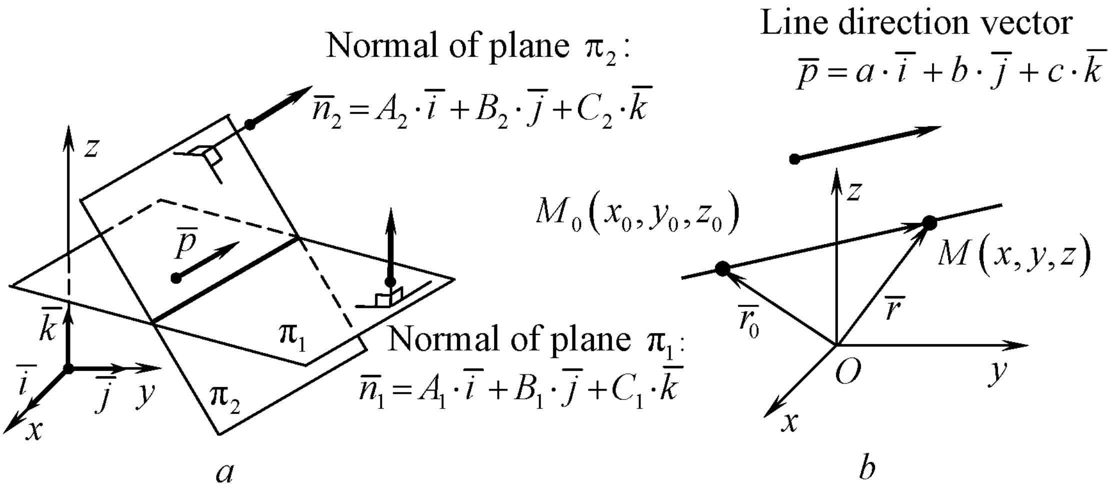


Figure 11.5

Way of representation: line is defined as an intersection line between two planes (Fig.11.5, *a*):

$$\pi_1: \ A_1 \cdot x + B_1 \cdot y + C_1 \cdot z + D_1 = 0; \qquad \pi_2: \ A_2 \cdot x + B_2 \cdot y + C_2 \cdot z + D_2 = 0.$$

Geometric sense of coefficients: A_1, B_1, C_1 and A_2, B_2, C_2 are coordinates of normals $\overline{n}_1 = A_1 \cdot \overline{i} + B_1 \cdot \overline{j} + C_1 \cdot \overline{k}$ and $\overline{n}_2 = A_2 \cdot \overline{i} + B_2 \cdot \overline{j} + C_2 \cdot \overline{k}$ of planes π_1 and π_2 accordingly. The equality of matrix rang to two in (11.11) denominates the condition of noncollinearity of normals (it equivalent to the plane intersection condition, Section 11.1.2).

Vector parametric equation of a line in space:

$$\overline{r} = \overline{r}_0 + t \cdot \overline{p}, \quad t \in \mathbb{R}, \quad \overline{p} \neq \overline{o}. \qquad (11.12)$$

Way of representation: line passes through the point $M_0(x_0, y_0, z_0)$, which is defined by radius vector $\overline{r}_0$, and is collinear to direction vector $\overline{p} \neq \overline{o}$ (Fig.11.5, *b*).

Geometric sense of parameter t in equation (11.12): value of t is proportional to the distance between initial point M_0 and point M, which is determined by radius vector $\overline{r}$.

Physical sense of parameter t: it denotes time of uniform rectilinear motion of point M. If $t=0$, then point M coincides with the initial point M_0, increase of t denominates motion with the direction defined by vector $\overline{p}$.

Parametric equation of a line in space:

$$\begin{cases} x = x_0 + a \cdot t, \\ y = y_0 + b \cdot t, \quad t \in \mathbb{R}, \quad a^2 + b^2 + c^2 \neq 0. \\ z = z_0 + c \cdot t, \end{cases} \tag{11.13}$$

Way of representation: line passes through the point $M_0(x_0, y_0, z_0)$ and is collinear to vector $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j} + c \cdot \overline{k}$ (Fig. 11.7, b).

Geometric sense of coefficients: coefficients a, b, c are coordinates of direction vector $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j} + c \cdot \overline{k}$ of a line and x_0, y_0, z_0 are coordinates of the point $M_0(x_0, y_0, z_0)$, which belongs to the line. Parameter t has the same meaning as in equation (11.12).

Note that the equation (11.13) is a coordinate form of the equation (11.12).

Canonical equation of a line in space:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}, \qquad a^2 + b^2 + c^2 \neq 0. \tag{11.14}$$

Way of representation: line passes through the point $M_0(x_0, y_0, z_0)$ and it is collinear to vector $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j} + c \cdot \overline{k}$ (Fig. 11.7, *b*).

Geometric sense of coefficients: coefficients a, b, c are coordinates of direction vector $\overline{p} = a \cdot \overline{i} + b \cdot \overline{j} + c \cdot \overline{k}$ of a line and x_0, y_0, z_0 are coordinates of the point $M_0(x_0, y_0, z_0)$, which belongs to the line. Parameter t has the same meaning as in previous equations (11.12), (11.13).

One or two from three denominators of fractions in (11.14) can be equal to zero, at that it is considered that the according numerator is equal to zero, e.g:

a) $\dfrac{x - x_0}{0} = \dfrac{y - y_0}{0} = \dfrac{z - z_0}{c}$ – equations of a line, which is parallel to the applicate axis (Fig. 11.6, *a*), i.e. $\begin{cases} x = x_0, \\ y = y_0; \end{cases}$

b) $\dfrac{x-x_0}{a}=\dfrac{y-y_0}{b}=\dfrac{z-z_0}{0}$ – equation of a line, which is parallel to coordinate plane Oxy (Fig. 11.6, *b*), i.e. $\begin{cases} z=z_0, \\ \dfrac{x-x_0}{a}=\dfrac{y-y_0}{b}. \end{cases}$

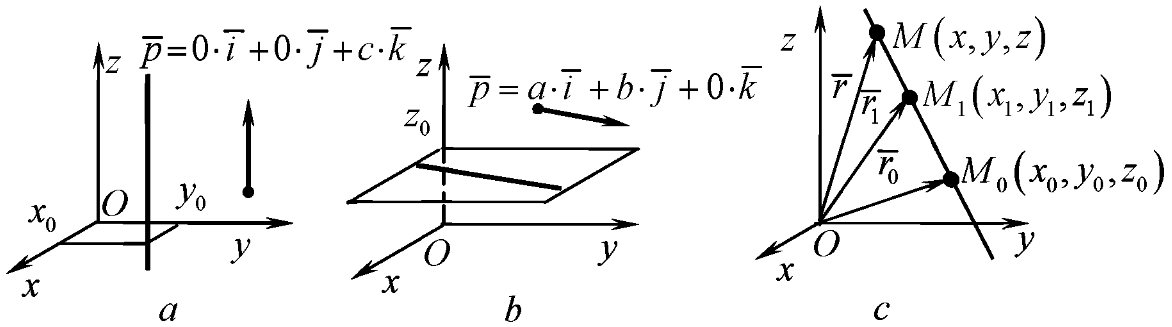


Figure 11.6

***Affine equation of a line, which passes through two given points, in space*:**

$$\overline{r}=(1-t)\cdot\overline{r}_0+t\cdot\overline{r}_1,\quad t\in\mathbb{R}. \tag{11.15}$$

Equation (11.15) can be written in coordinate form:

$$\begin{cases} x=(1-t)\cdot x_0+t\cdot x_1, \\ y=(1-t)\cdot y_0+t\cdot y_1, \quad t\in\mathbb{R}. \\ z=(1-t)\cdot z_0+t\cdot z_1, \end{cases}$$

Way of representation: line passes through two given points $M_0(x_0,y_0,z_0)$ and $M_1(x_1,y_1,z_1)$, which are defined by two radius vectors $\overline{r}_0$ and $\overline{r}_1$ accordingly (Fig.11.6, *c*). Radius vector $\overline{r}$ defines the location of point $M(x,y,z)$, which belongs to the line.

Geometric sense of coefficients: x_0, y_0, z_0 and x_1, y_1, z_1 are coordinates of points $M_0(x_0,y_0,z_0)$ and $M_1(x_1,y_1,z_1)$, through which the line passes (11.15). Parameter t in equation (11.15) defines the location of point $M(x,y,z)$, which belongs to the line, e.g. if $t=0$ then point M coincides with the point M_0 ($\overline{r}=\overline{r}_0$), and if $t=1$ – with the point M_1 ($\overline{r}=\overline{r}_1$).

***Equation of a line, which passes through two given points $M_0(x_0,y_0,z_0)$ and $M_1(x_1,y_1,z_1)$ in space*:**

$$\frac{x-x_0}{x_1-x_0}=\frac{y-y_0}{y_1-y_0}=\frac{z-z_0}{z_1-z_0}. \qquad (11.16)$$

Way of representation: line passes through two given points $M_0(x_0,y_0,z_0)$ and $M_1(x_1,y_1,z_1)$ (Fig.11.6, *c*).

Geometric sense of coefficients: x_0,y_0,z_0 and x_1,y_1,z_1 are coordinates of points $M_0(x_0,y_0,z_0)$ and $M_1(x_1,y_1,z_1)$, through which the line passes (11.16). As in canonical equation, one or two from three denominators of fractions in (11.16) can be equal to zero, at that it is considered that the according numerator is equal to zero.

Ways of Equation Type Transformation

1. To transform general equation of a line in space (11.11) into canonical equation (11.14) it is necessary to make the following steps:

1) find any solution (x_0,y_0,z_0) of the system $\begin{cases} A_1\cdot x+B_1\cdot y+C_1\cdot z+D_1=0, \\ A_2\cdot x+B_2\cdot y+C_2\cdot z+D_2=0, \end{cases}$

thus defining coordinates of point $M_0(x_0,y_0,z_0)$, which belongs to the line;

2) find any nonzero solution (a,b,c) of homogeneous system $\begin{cases} A_1\cdot a+B_1\cdot b+C_1\cdot c=0, \\ A_2\cdot a+B_2\cdot b+C_2\cdot c=0, \end{cases}$ thus defining coordinates a,b,c of direction vector $\overline{p}$, or find direction vector $\overline{p}$ as outer product of normals $\overline{n}_1=A_1\cdot\overline{i}+B_1\cdot\overline{j}+C_1\cdot\overline{k}$, $\overline{n}_2=A_2\cdot\overline{i}+B_2\cdot\overline{j}+C_2\cdot\overline{k}$ of given planes:

$$\overline{p}=[\overline{n}_1,\overline{n}_2]=a\cdot\overline{i}+b\cdot\overline{j}+c\cdot\overline{k}=\begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix}.$$

3) write canonical equation (11.14).

2. To transform canonical equation into general one, it is sufficient to write the double equality (11.14) as a system

$$\begin{cases} \dfrac{x-x_0}{a}=\dfrac{y-y_0}{b}, \\ \dfrac{y-y_0}{b}=\dfrac{z-z_0}{c} \end{cases}$$

and reduce similar terms.

3. To transform canonical equation (11.14) into parametric equation (11.13), it is necessary to equate each fraction in equation (11.14) to parameter t and write obtained equalities as a system (11.13):

$$\frac{x-x_0}{a}=\frac{y-y_0}{b}=\frac{z-z_0}{c}=t \quad \Leftrightarrow \quad \begin{cases} x=x_0+a\cdot t, \\ y=y_0+b\cdot t, \quad t\in\mathbb{R}. \\ z=z_0+c\cdot t, \end{cases}$$

Example 11.5. In coordinate space $Oxyz$ the following vertexes of triangle are given: $A(1,2,3)$, $B(3,0,2)$, $C(7,4,6)$ (fig.11.7). Find:

a) general equation of a line, which contains altitude AH of triangle;

b) canonical equation of a line, which contains altitude AH of triangle;

c) general equation of a line, which contains bisectrix AL of triangle;

d) parametric equation of a line, which contains median AM of triangle.

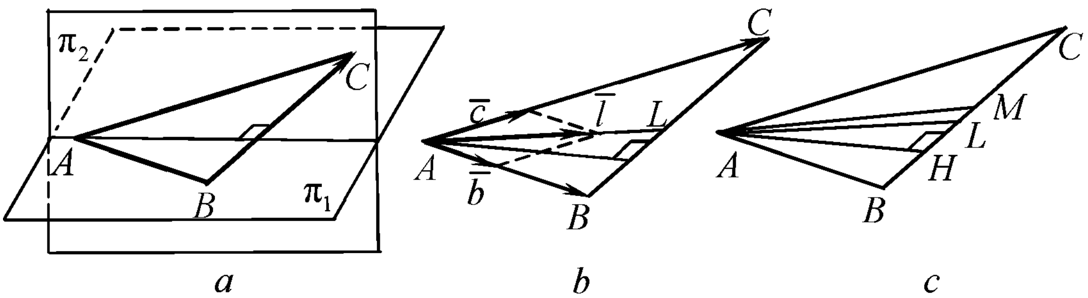


Figure 11.7

□ a) Line AH is an intersection line of two planes: π_1 of triangle ABC and π_2, which passes through the point A and which is perpendicular to vector $\overline{BC}$ (Fig.11.7, *a*). By the formula (11.9) compose the equation of the plane π_1, which passes through the points A, B, C:

$$\begin{vmatrix} x-1 & y-2 & z-3 \\ 3-1 & 0-2 & 2-3 \\ 7-1 & 4-2 & 6-3 \end{vmatrix} = \begin{vmatrix} x-1 & y-2 & z-3 \\ 2 & -2 & -1 \\ 6 & 2 & 3 \end{vmatrix} = 0 \quad \Leftrightarrow \quad x+3\cdot y-4\cdot z+5=0.$$

By the formula (11.2) compose the equation of the plane π_2, which passes through the point A and which is perpendicular to vector $\overline{BC}=(7-3)\cdot\overline{i}+(4-0)\cdot\overline{j}+(6-2)\cdot\overline{k}= =4\cdot\overline{i}+4\cdot\overline{j}+4\cdot\overline{k}$:

$$4\cdot(x-1)+4\cdot(y-2)+4\cdot(z-3)=0 \quad \Leftrightarrow \quad x+y+z-6=0.$$

Consequently, general equation (11.11) of the line AH has the following view:

$$\begin{cases} x+3\cdot y-4\cdot z+5=0, \\ x+y+z-6=0. \end{cases}$$

b) General equation a the line AH was obtained in "a". To transform general equation into canonical one it is necessary to:

- find any solution (x_0, y_0, z_0) of the system, e.g. $x_0=1$, $y_0=2$, $z_0=3$ (coordinates of the vertex $A(1,2,3)$);
- find direction vector $\bar{p}$ of the line as outer product of normals $\bar{n}_1=1\cdot\bar{i}+3\cdot\bar{j}-4\cdot\bar{k}$, $\bar{n}_2=1\cdot\bar{i}+1\cdot\bar{j}+1\cdot\bar{k}$ of the given planes:

$$\bar{p}=[\bar{n}_1,\bar{n}_2]=\begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ 1 & 3 & -4 \\ 1 & 1 & 1 \end{vmatrix}=7\cdot\bar{i}-5\cdot\bar{j}-2\cdot\bar{k};$$

- write canonical equation (11.14): $\dfrac{x-1}{7}=\dfrac{y-2}{-5}=\dfrac{z-3}{-2}$.

c) First compose canonical equation of the line AL. To do this we should find direction vector $\bar{l}$ of this line (Fig. 11.7, *b*). Taking into account, that diagonal of rhombus is a bisectrix, we get $\bar{l}=\bar{b}+\bar{c}$, where $\bar{b}$ and $\bar{c}$ are unit vectors with the same direction as vectors $\overline{AB}$ and $\overline{AC}$ accordingly. Find

$$\overline{AB}=2\cdot\bar{i}-2\cdot\bar{j}-1\cdot\bar{k}, \quad \left|\overline{AB}\right|=3, \quad \bar{b}=\frac{\overline{AB}}{\left|\overline{AB}\right|}=\frac{2}{3}\cdot\bar{i}-\frac{2}{3}\cdot\bar{j}-\frac{1}{3}\cdot\bar{k};$$

$$\overline{AC}=6\cdot\bar{i}+2\cdot\bar{j}+3\cdot\bar{k}, \quad \left|\overline{AC}\right|=7, \quad \bar{c}=\frac{\overline{AC}}{\left|\overline{AC}\right|}=\frac{6}{7}\cdot\bar{i}+\frac{2}{7}\cdot\bar{j}+\frac{3}{7}\cdot\bar{k};$$

$$\bar{l}=\bar{b}+\bar{c}=\left(\frac{2}{3}\cdot\bar{i}-\frac{2}{3}\cdot\bar{j}-\frac{1}{3}\cdot\bar{k}\right)+\left(\frac{6}{7}\cdot\bar{i}+\frac{2}{7}\cdot\bar{j}+\frac{3}{7}\cdot\bar{k}\right)=\frac{32}{21}\cdot\bar{i}-\frac{8}{21}\cdot\bar{j}+\frac{2}{21}\cdot\bar{k}.$$

Compose canonical equation of the line AL: $\dfrac{x-1}{\frac{32}{21}}=\dfrac{y-2}{-\frac{8}{21}}=\dfrac{z-3}{\frac{2}{21}}$.

Writing double equality as a system, we get general equation of the line AL:

$$\begin{cases} \dfrac{x-1}{\frac{32}{21}} = \dfrac{y-2}{-\frac{8}{21}}, \\ \dfrac{y-2}{-\frac{8}{21}} = \dfrac{z-3}{\frac{2}{21}}, \end{cases} \Leftrightarrow \quad \begin{cases} x + 4 \cdot y - 9 = 0, \\ y + 4 \cdot z - 14 = 0. \end{cases}$$

d) Find coordinates of the middle point M of BC: $M(5,2,4)$. Compose equation (11.16) of the line AM (Fig.11.7, *c*):

$$\frac{x-1}{5-1} = \frac{y-2}{2-2} = \frac{z-3}{4-3} \quad \Leftrightarrow \quad \frac{x-1}{4} = \frac{y-2}{0} = \frac{z-3}{1}.$$

Transfer the obtained equation into parametric one, by equating each fraction to parameter t:

$$\frac{x-1}{4} = \frac{y-2}{0} = \frac{z-3}{1} = t \quad \Leftrightarrow \quad \begin{cases} x = 1 + 4 \cdot t, \\ y = 2, \\ z = 3 + 1 \cdot t, \end{cases} \quad t \in \mathbb{R}. \ \blacksquare$$

11.2.2. Positional Relationships of Lines in Space

Consider two lines l_1 and l_2 specified by their canonical equations:

$$l_1 : \frac{x - x_1}{a_1} = \frac{y - y_1}{b_1} = \frac{z - z_1}{c_1}, \quad l_2 : \frac{x - x_2}{a_2} = \frac{y - y_2}{b_2} = \frac{z - z_2}{c_2},$$

where x_1, y_1, z_1 and x_2, y_2, z_2 are coordinates of points $M_1(x_1, y_1, z_1)$ and $M_2(x_2, y_2, z_2)$, which belong to lines l_1 and l_2 accordingly; a_1, b_1, c_1 and a_2, b_2, c_2 are coordinates of direction vectors $\overline{p}_1 = a_1 \cdot \overline{i} + b_1 \cdot \overline{j} + c_1 \cdot \overline{k}$ and $\overline{p}_2 = a_2 \cdot \overline{i} + b_2 \cdot \overline{j} + c_2 \cdot \overline{k}$ of these lines (Fig. 11.8).

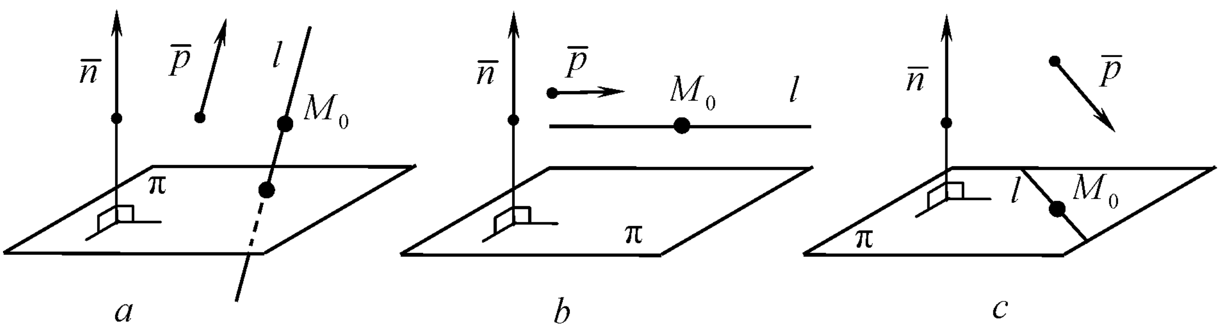


Figure 11.8

Information about line positional relationship can be obtained from the number of linearly independent rows of matrix

$$\begin{pmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}, \tag{11.17}$$

and by the following criteria:

- *skew* lines:

$$\mathrm{rg}\begin{pmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 3;$$

- *parallel* lines:

$$\mathrm{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 1 \text{ and } \mathrm{rg}\begin{pmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \end{pmatrix} = 2;$$

- *equal* lines:

$$\mathrm{rg}\begin{pmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 1;$$

- *collinear* lines:

$$\mathrm{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 1;$$

- *intersecting* lines:

$$\mathrm{rg}\begin{pmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \mathrm{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = 2;$$

- *perpendicular* lines:

$$a_1 \cdot a_2 + b_1 \cdot b_2 + c_1 \cdot c_2 = 0.$$

If lines intersect, then the coordinates of intersection point can be found as the solution of the following system of equations:

$$\begin{cases} \dfrac{x-x_1}{a_1} = \dfrac{y-y_1}{b_1} = \dfrac{z-z_1}{c_1}, \\ \dfrac{x-x_2}{a_2} = \dfrac{y-y_2}{b_2} = \dfrac{z-z_2}{c_2}. \end{cases}$$

Example 11.6. Find information about positional relationship of each pair of lines (are they skew, intersecting, parallel, equal, perpendicular, if they are intersecting find their mutual point):

a) $\dfrac{x-1}{2} = \dfrac{y-2}{-3} = \dfrac{z-3}{4}$, $\dfrac{x+1}{4} = \dfrac{y-1}{0} = \dfrac{z+2}{-2}$;

b) $\dfrac{x-1}{2} = \dfrac{y-2}{-3} = \dfrac{z-3}{4}$, $\dfrac{x+1}{-4} = \dfrac{y-1}{6} = \dfrac{z+2}{-8}$;

c) $\dfrac{x-3}{2} = \dfrac{y+2}{-3} = \dfrac{z-7}{4}$, $\dfrac{x+3}{-4} = \dfrac{y-7}{6} = \dfrac{z+5}{-8}$;

d) $\dfrac{x-3}{2} = \dfrac{y+2}{-3} = \dfrac{z-7}{4}$, $\dfrac{x-2}{5} = \dfrac{y+2}{-6} = \dfrac{z-5}{10}$.

□ a) By the coefficients of line equations we compose the matrix (11.17):

$$\begin{pmatrix} x_1-x_2 & y_1-y_2 & z_1-z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \begin{pmatrix} 1-(-1) & 2-1 & 3-(-2) \\ 2 & -3 & 4 \\ 4 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 5 \\ 2 & -3 & 4 \\ 4 & 0 & -2 \end{pmatrix}.$$

Since $\operatorname{rg}\begin{pmatrix} 2 & 1 & 5 \\ 2 & -3 & 4 \\ 4 & 0 & -2 \end{pmatrix} = 3$, the given lines are skew. As $a_1 \cdot a_2 + b_1 \cdot b_2 + c_1 \cdot c_2 = 2 \cdot 4 + (-3) \cdot 0 + 4 \cdot (-2) = 0$, they are perpendicular.

b) By the coefficients of line equations we compose the matrix (11.17):

$$\begin{pmatrix} x_1-x_2 & y_1-y_2 & z_1-z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \begin{pmatrix} 1-(-1) & 2-1 & 3-(-2) \\ 2 & -3 & 4 \\ -4 & 6 & -8 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 5 \\ 2 & -3 & 4 \\ -4 & 6 & -8 \end{pmatrix}.$$

Since $\operatorname{rg}\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 2 & -3 & 4 \\ -4 & 6 & -8 \end{pmatrix} = 1$ and $\operatorname{rg}\begin{pmatrix} x_1-x_2 & y_1-y_2 & z_1-z_2 \\ a_1 & b_1 & c_1 \end{pmatrix} =$ $\operatorname{rg}\begin{pmatrix} 2 & 1 & 5 \\ 2 & -3 & 4 \end{pmatrix} = 2$, the lines are parallel.

c) By the coefficients of line equations we compose the matrix (11.17):

$$\begin{pmatrix} x_1-x_2 & y_1-y_2 & z_1-z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \begin{pmatrix} 3-(-3) & -2-7 & 7-(-5) \\ 2 & -3 & 4 \\ -4 & 6 & -8 \end{pmatrix} = \begin{pmatrix} 6 & -9 & 12 \\ 2 & -3 & 4 \\ -4 & 6 & -8 \end{pmatrix}.$$

Since $\mathrm{rg}\begin{pmatrix} 6 & -9 & 12 \\ 2 & -3 & 4 \\ -4 & 6 & -8 \end{pmatrix} = 1$, the lines are equal.

d) By the coefficients of line equations we compose the matrix (11.17):

$$\begin{pmatrix} x_1-x_2 & y_1-y_2 & z_1-z_2 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \begin{pmatrix} 3-2 & -2-(-2) & 7-5 \\ 2 & -3 & 4 \\ 5 & -6 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -3 & 4 \\ 5 & -6 & 10 \end{pmatrix}.$$

Since $\mathrm{rg}\begin{pmatrix} 1 & 0 & 2 \\ 2 & -3 & 4 \\ 5 & -6 & 10 \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 2 & -3 & 4 \\ 5 & -6 & 10 \end{pmatrix} = 2$, the lines intersect. Find the coordinates of intersection point as the solution of the system:

$$\begin{cases} \dfrac{x-3}{2} = \dfrac{y+2}{-3} = \dfrac{z-7}{4}, \\ \dfrac{x-2}{5} = \dfrac{y+2}{-6} = \dfrac{z-5}{10}, \end{cases} \Leftrightarrow \begin{cases} -3\cdot x+9 = 2\cdot y+4, \\ 4\cdot y+8 = -3\cdot z+21, \\ -6\cdot x+12 = 5\cdot y+10, \\ 10\cdot y+20 = -6\cdot z+30, \end{cases} \Leftrightarrow \begin{cases} 3\cdot x+2\cdot y=5, \\ 4\cdot y+3\cdot z=13, \\ 6\cdot x+5\cdot y=2, \\ 5\cdot y+3\cdot z=5. \end{cases}$$

Subtracting the second equation from the last one we obtain $y=-8$. Substituting $y=-8$ in the first two equations we find $x=7$, $z=15$. Hence, the only common point of the lines has coordinates $x=7$, $y=-8$, $z=15$. ■

11.2.3. Positional Relationships of Line and Plane

Consider line l and plane π specified by the following equations:

$$l:\ \frac{x-x_0}{a} = \frac{y-y_0}{b} = \frac{z-z_0}{c}; \qquad \pi:\ A\cdot x+B\cdot y+C\cdot z+D=0,$$

i.e. line l passes through the point $M_0(x_0, y_0, z_0)$ and is collinear to vector $\overline{p} = a\cdot\overline{i} + b\cdot\overline{j} + c\cdot\overline{k}$ and plane π is perpendicular to vector $\overline{n} = A\cdot\overline{i} +$

$+B\cdot\overline{j}+C\cdot\overline{k}$ (Fig. 11.9). Information about positional relationship of line and plane can be obtained from scalar product of vectors $(\overline{p},\overline{n})=a\cdot A+b\cdot B+c\cdot C$ and by the following criteria:

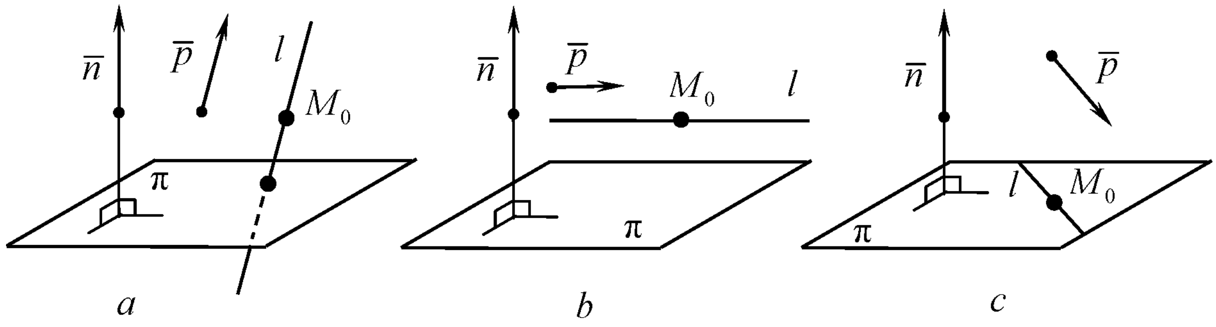


Figure 11.9

- *intersection* of a line l and a plane π (Fig. 11.9, *a*):

$$a\cdot A+b\cdot B+c\cdot C\neq 0;$$

- *perpendicularity* of a line l and a plane π:

$$\mathrm{rg}\begin{pmatrix} a & b & c \\ A & B & C \end{pmatrix}=1;$$

- *parallelism* of a line l and a plane π (Fig. 11.9, *b*):

$$\begin{cases} a\cdot A+b\cdot B+c\cdot C=0, \\ A\cdot x_0+B\cdot y_0+C\cdot z_0+D\neq 0; \end{cases}$$

- *belonging* of a line l to a plane π (Fig. 11.9, *c*):

$$\begin{cases} a\cdot A+b\cdot B+c\cdot C=0, \\ A\cdot x_0+B\cdot y_0+C\cdot z_0+D=0. \end{cases}$$

In case of intersection it is convenient to use parametric equation of a line to find coordinates of a common point. Substituting the following expressions

$$x=x_0+a\cdot t,\quad y=y_0+b\cdot t,\quad z=z_0+c\cdot t \qquad (11.18)$$

in equation of a plane $A\cdot x+B\cdot y+C\cdot z+D=0$, it is possible to calculate value of parameter t^* for intersection point, and then the coordinates of the required point by formula (11.18) assuming $t=t^*$.

Example 11.7. Get information about positional relationship of each pair of line and plane (are they intersecting, perpendicular, parallel, if the line belongs to a plane, in case of intersection find mutual point):

a) $\dfrac{x-1}{1}=\dfrac{y-2}{-6}=\dfrac{z+3}{4}$, $2\cdot x+3\cdot y+4\cdot z+4=0$;

b) $\dfrac{x-1}{1}=\dfrac{y-2}{-6}=\dfrac{z+3}{4}$, $2\cdot x+3\cdot y+4\cdot z+1=0$;

c) $\dfrac{x-1}{1}=\dfrac{y-2}{-3}=\dfrac{z+3}{4}$, $2\cdot x+3\cdot y+4\cdot z-5=0$;

d) $\dfrac{x-1}{1}=\dfrac{y+2}{-3}=\dfrac{z+3}{-4}$, $2\cdot x-6\cdot y-8\cdot z+14=0$.

□ a) By the coefficients in equations define $a=1$, $b=-6$, $c=4$, $x_0=1$, $y_0=2$, $z_0=-3$, $A=2$, $B=3$, $C=4$, $D=4$. Since

$$\begin{cases} a\cdot A+b\cdot B+c\cdot C=0, \\ A\cdot x_0+B\cdot y_0+C\cdot z_0+D=0, \end{cases} \Leftrightarrow \begin{cases} 1\cdot 2+(-6)\cdot 3+4\cdot 4=0, \\ 2\cdot 1+3\cdot 2+4\cdot(-3)+4=0, \end{cases}$$

the line belongs to the plane.

b) By the coefficients in equations define $a=1$, $b=-6$, $c=4$, $x_0=1$, $y_0=2$, $z_0=-3$, $A=2$, $B=3$, $C=4$, $D=1$. Since

$$\begin{cases} a\cdot A+b\cdot B+c\cdot C=0, \\ A\cdot x_0+B\cdot y_0+C\cdot z_0+D\neq 0, \end{cases} \Leftrightarrow \begin{cases} 1\cdot 2+(-6)\cdot 3+4\cdot 4=0, \\ 2\cdot 1+3\cdot 2+4\cdot(-3)+1\neq 0, \end{cases}$$

the line is parallel to the plane.

c) By the coefficients in equations define $a=1$, $b=-3$, $c=4$, $x_0=1$, $y_0=2$, $z_0=-3$, $A=2$, $B=3$, $C=4$, $D=-5$. Since

$$a\cdot A+b\cdot B+c\cdot C\neq 0 \quad \Leftrightarrow \quad 1\cdot 2+(-3)\cdot 3+4\cdot 4\neq 0,$$

the line intersects the plane. Since $\operatorname{rg}\begin{pmatrix} 1 & -3 & 4 \\ 2 & 3 & 4 \end{pmatrix}=2\neq 1$, the line is not perpendicular to the plane. Find the coordinates of intersection point. Substituting $x=1+1\cdot t$, $y=2-3\cdot t$, $z=-3+4\cdot t$, in plane equation we get

$$2\cdot(1+1\cdot t)+3\cdot(2-3\cdot t)+4\cdot(-3+4\cdot t)-5=0 \quad \Leftrightarrow \quad t^*=1.$$

Then, coordinates of the common point are $x=1+1\cdot t^*=2$, $y=2-3\cdot t^*=-1$, $z=-3+4\cdot t^*=1$.

d) By the coefficients in equations define $a=1$, $b=-3$, $c=-4$, $x_0=1$, $y_0=-2$, $z_0=-3$, $A=2$, $B=-6$, $C=-8$, $D=14$. Since

$$\operatorname{rg}\begin{pmatrix} 1 & -3 & -4 \\ 2 & -6 & -8 \end{pmatrix}=1,$$

the line is perpendicular to the plane. Find the coordinates of intersection point. Substituting $x=1+1\cdot t$, $y=-2-3\cdot t$, $z=-3-4\cdot t$ in plane equation we get

$$2\cdot(1+1\cdot t)-6\cdot(-2-3\cdot t)-8\cdot(-3-4\cdot t)+14=0 \quad \Leftrightarrow \quad t^*=-1.$$

Then, coordinates of the common point are $x=1+1\cdot t^*=0$, $y=-2-3\cdot t^*=1$, $z=-3-4\cdot t^*=1$.■

11.2.4. Metric Applications of Line Equations in Space

Angle between a line l ***and a plane*** π is defined as an angle between a line l and its orthogonal projection l_{proj} to a plane (Fig. 11.10, *a*). From two adjacent angles φ and φ' usually chose the smallest one, i.e. $0\le\varphi\le\frac{\pi}{2}$. If line l is perpendicular to plane (its orthogonal projection is a point), then angle equals to $\frac{\pi}{2}$.

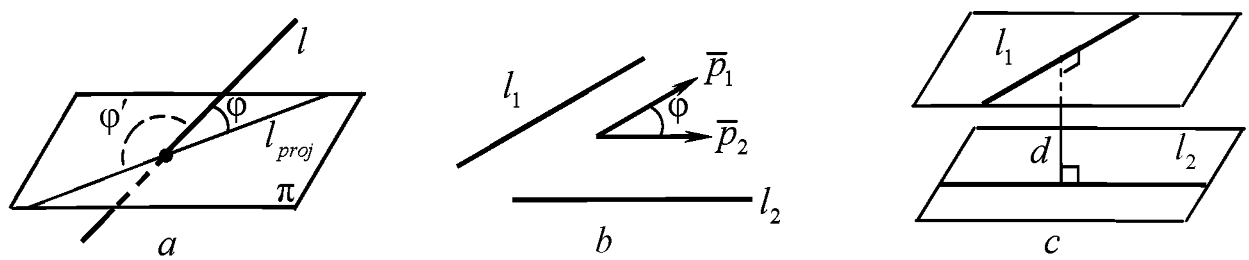


Figure 11.10

Angle between lines is defined as an angle between their direction vectors (Fig.11.10, *b*).

Distance between skew lines is the length of their mutual perpendicular (Fig.11.10, *c*), i.e. the smallest distance between the points of these lines.

1. Distance d from point $M_1(x_1, y_1, z_1)$ to line $\dfrac{x-x_0}{a} = \,=\dfrac{y-y_0}{b} = \dfrac{z-z_0}{c}$

(Fig.11.11, *a*) is calculated by the following formula

$$d = \frac{\sqrt{\begin{vmatrix} x_1-x_0 & y_1-y_0 \\ a & b \end{vmatrix}^2 + \begin{vmatrix} y_1-y_0 & z_1-z_0 \\ b & c \end{vmatrix}^2 + \begin{vmatrix} x_1-x_0 & z_1-z_0 \\ a & c \end{vmatrix}^2}}{\sqrt{a^2+b^2+c^2}}.$$

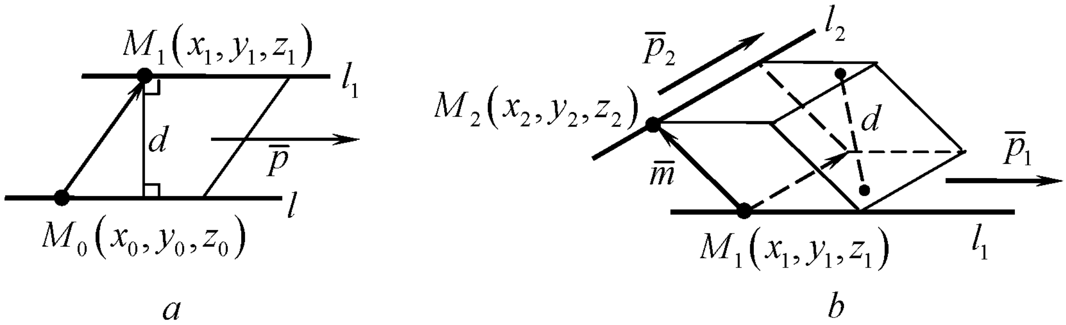


Figure 11.11

By this formula it is also possible to find distance between two parallel lines $\dfrac{x-x_0}{a} = \dfrac{y-y_0}{b} = \dfrac{z-z_0}{c}$ and $\dfrac{x-x_1}{a_1} = \dfrac{y-y_1}{b_1} = \dfrac{z-z_1}{c_1}$, coordinates of which direction vectors are proportional: $\dfrac{a}{a_1} = \dfrac{b}{b_1} = \dfrac{c}{c_1}$ (Fig. 11.11, *a*).

2. Distance d between skew lines (Fig. 11.11, *b*)

$$\frac{x-x_1}{a_1} = \frac{y-y_1}{b_1} = \frac{z-z_1}{c_1} \quad \text{and} \quad \frac{x-x_2}{a_2} = \frac{y-y_2}{b_2} = \frac{z-z_2}{c_2}$$

is calculated by the following formula: $d = \dfrac{|(\overline{m}, \overline{p}_1, \overline{p}_2)|}{|[\overline{p}_1, \overline{p}_2]|}$,

where $(\overline{m}, \overline{p}_1, \overline{p}_2) = \begin{vmatrix} x_2-x_1 & y_2-y_1 & z_2-z_1 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} \neq 0$, $[\overline{p}_1, \overline{p}_2] = \begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix}$ are compositional and outer products of vectors $\overline{m} = (x_2-x_1)\cdot\overline{i} + +(y_2-y_1)\cdot\overline{j} + (z_2-z_1)\cdot\overline{k}$, $\overline{p}_1 = a_1\cdot\overline{i} + b_1\cdot\overline{j} + c_1\cdot\overline{k}$, $\overline{p}_2 = a_2\cdot\overline{i} + b_2\cdot\overline{j} + c_2\cdot\overline{k}$.

3. Angle φ between two lines

$$\frac{x-x_1}{a_1}=\frac{y-y_1}{b_1}=\frac{z-z_1}{c_1} \quad \text{and} \quad \frac{x-x_2}{a_2}=\frac{y-y_2}{b_2}=\frac{z-z_2}{c_2}$$

is calculated by the following formula

$$\cos\varphi=\frac{|a_1\cdot a_2+b_1\cdot b_2+c_1\cdot c_2|}{\sqrt{a_1^2+b_1^2+c_1^2}\cdot\sqrt{a_2^2+b_2^2+c_2^2}}.$$

4. Angle φ between line $\frac{x-x_0}{a}=\frac{y-y_0}{b}=\frac{z-z_0}{c}$ and plane $A\cdot x+B\cdot y+C\cdot z+D=0$ is calculated by the following formula

$$\sin\varphi=\frac{|a\cdot A+b\cdot B+c\cdot C|}{\sqrt{a^2+b^2+c^2}\cdot\sqrt{A^2+B^2+C^2}}.$$

Example 11.8. In coordinate space $Oxyz$ (in Cartesian coordinate system) the following vertexes of triangle are given: $A(1,2,3)$, $B(3,0,2)$, $C(7,4,6)$ (Fig. 11.7).

Find:

a) equation of the line BC;

b) altitude h of the triangle, dropped to the side BC;

c) distance d from the line BC and the abscissa axis;

d) value of acute angle φ between these lines;

e) value of angle ψ between the abscissa axis and the plane of triangle ABC.

☐ a) Compose the equation (11.16) of a line, which passes through the points $B(3,0,2)$, $C(7,4,6)$:

$$\frac{x-3}{7-3}=\frac{y-0}{4-0}=\frac{z-2}{6-2} \Leftrightarrow \frac{x-3}{4}=\frac{y}{4}=\frac{z-2}{4} \Leftrightarrow \frac{x-3}{1}=\frac{y}{1}=\frac{z-2}{1}.$$

b) Required altitude h is found by the first formula of metric applications, assuming $x_0=3$, $y_0=0$, $z_0=2$, $x_1=1$, $y_1=2$, $z_1=3$, $a=b=c=1$:

$$h=\frac{\sqrt{\begin{vmatrix}-2 & 2\\ 1 & 1\end{vmatrix}^2+\begin{vmatrix}2 & 1\\ 1 & 1\end{vmatrix}^2+\begin{vmatrix}-2 & 1\\ 1 & 1\end{vmatrix}^2}}{\sqrt{1^2+1^2+1^2}}=\frac{\sqrt{16+1+9}}{\sqrt{3}}=\frac{\sqrt{26}}{\sqrt{3}}.$$

c) Canonical equation of the abscissa axis has the following form $\frac{x}{1}=\frac{y}{0}=\frac{z}{0}$, since the axis passes through the point $O(0,0,0)$, and $\overline{i}$ is its direction vector. Canonical equation of the line BC was obtained in "a": $\frac{x-3}{1}=\frac{y}{1}=\frac{z-2}{1}$.

Assuming $\overline{m}=\overline{OB}=(3-0)\cdot\overline{i}+(0-0)\cdot\overline{j}+(2-0)\cdot\overline{k}=3\cdot\overline{i}+0\cdot\overline{j}+2\cdot\overline{k}$, $\overline{p}_1=1\cdot\overline{i}+$ $+0\cdot\overline{j}+0\cdot\overline{k}$, $\overline{p}_2=1\cdot\overline{i}+1\cdot\overline{j}+1\cdot\overline{k}$, by the second formula of metric applications:

$$(\overline{m},\overline{p}_1,\overline{p}_2)=\begin{vmatrix} 3 & 0 & 2 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix}=2, \quad [\overline{p}_1,\overline{p}_2]=\begin{vmatrix} \overline{i} & \overline{j} & \overline{k} \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix}=0\cdot\overline{i}-1\cdot\overline{j}+1\cdot\overline{k},$$

$$d=\frac{|(\overline{m},\overline{p}_1,\overline{p}_2)|}{|[\overline{p}_1,\overline{p}_2]|}=\frac{2}{\sqrt{0^2+(-1)^2+1^2}}=\sqrt{2}.$$

d) Acute angle φ is obtained by the third formula of metric applications:

$$\cos\varphi=\frac{|(\overline{p}_1,\overline{p}_2)|}{|\overline{p}_1|\cdot|\overline{p}_2|}=\frac{|1\cdot1+0\cdot1+0\cdot1|}{\sqrt{1^2+0^2+0^2}\cdot\sqrt{1^2+1^2+1^2}}=\frac{1}{\sqrt{3}}, \text{ i.e. } \varphi=\arccos\frac{1}{\sqrt{3}}.$$

e) Equation of a plane π_1, which passes through the points A,B,C, was obtained in example 11.5 "a": $x+3\cdot y-4\cdot z+5=0$. Acute angle ψ between the abscissa axis $\frac{x}{1}=\frac{y}{0}=\frac{z}{0}$ and the plane $x+3\cdot y-4\cdot z+5=0$ is obtained by the fourth formula of metric applications:

$$\sin\psi=\frac{|a\cdot A+b\cdot B+c\cdot C|}{\sqrt{a^2+b^2+c^2}\cdot\sqrt{A^2+B^2+C^2}}=\frac{|1\cdot1+0\cdot3+0\cdot(-4)|}{\sqrt{1^2+0^2+0^2}\cdot\sqrt{1^2+3^2+(-4)^2}}=\frac{1}{\sqrt{26}},$$

i.e. $\psi=\arcsin\frac{1}{\sqrt{26}}$. ■

11.3. QUADRIC SURFACES

11.3.1. Classification of Quadric Surfaces

Algebraic surface of the second order (***quadric surface***) is a locus of points in space, which can be represented in some affine coordinate system $Oxyz$ by the following equation

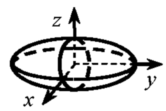
$$a_{11}\cdot x^2 + a_{22}\cdot y^2 + a_{33}\cdot z^2 + 2\cdot a_{12}\cdot x\cdot y + 2\cdot a_{13}\cdot x\cdot z + 2\cdot a_{23}\cdot y\cdot z +$$

$$+2\cdot a_1\cdot x + 2\cdot a_2\cdot y + 2\cdot a_3\cdot z + a_0 = 0,$$

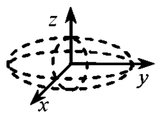
where the leading coefficients a_{11}, a_{12}, a_{13}, a_{22}, a_{23}, a_{33} are not all equal to zero simultaneously. For any quadric surface there exists a rectangular coordinate system $Oxyz$, in which the equation has the simplest (***canonical***) view. This system is called ***canonical***, and equation is also called ***canonical***.

Canonical Equations of Quadric Surfaces

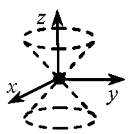
1) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$ – ellipsoid equation;



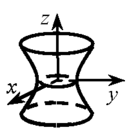
2) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=-1$ – imaginary ellipsoid equation;



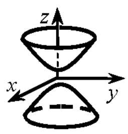
3) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=0$ – imaginary cone equation;



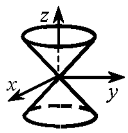
4) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1$ – one-sheet hyperboloid equation;



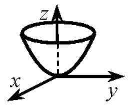
5) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=-1$ – two-sheet hyperboloid equation;



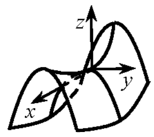
6) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=0$ – cone equation;



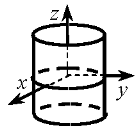
7) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=2\cdot z$ – elliptic paraboloid equation;



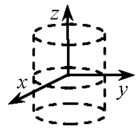
8) $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=2\cdot z$ – hyperbolic paraboloid equation;



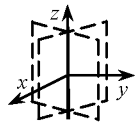
9) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ – elliptic cylinder equation;



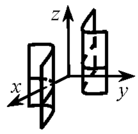
10) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=-1$ – imaginary elliptic cylinder equation;



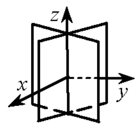
11) $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=0$ – pair of imaginary planes equation;



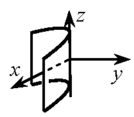
12) $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ – hyperbolic cylinder equation;



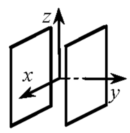
13) $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=0$ – pair of intersecting planes equation;



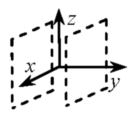
14) $y^2=2\cdot p\cdot x$ – parabolic cylinder equation;



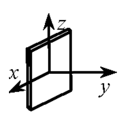
15) $y^2-b^2=0$ – pair of parallel planes equation;



16) $y^2+b^2=0$– pair of imaginary parallel planes equation;



17) $y^2=0$ – pair of equal planes equation.



In equations $a>0$, $b>0$, $c>0$, $p>0$ and $a\ge b\ge c$ in equations 1–3; $a\ge b$ in equations 4–7, 9–11.

11.3.2. Ellipsoids

Ellipsoid is a surface, which is defined in some rectangular coordinate system $Oxyz$ by the following canonical equation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \tag{11.19}$$

where a, b,c are positive parameters, which satisfy the inequalities $a \geq b \geq c$.

If a point $M(x,y,z)$ belongs to an ellipsoid (11.19), then coordinates $(\pm x,\pm y,\pm z)$ for any combination of signs also satisfy the equation (11.19). It is the reason why ellipsoid is symmetric relative to coordinate planes, coordinate axes and the coordinate origin. The origin of coordinates is called a ***center*** of ellipsoid. Six points $(\pm a,0,0)$, $(0,\pm b,0)$, $(0,0,\pm c)$ of intersection of ellipsoid and coordinate axes are called its ***vertexes***, and three segments of coordinate axes, which connect its vertexes are called ***axes*** of ellipsoid. Axes of ellipsoid, which belong to coordinate axes Ox, Oy, Oz, have lengths $2\cdot a, 2\cdot b, 2\cdot c$ accordingly. If $a > b > c$, then the number a is called ***semi-major axis***, number b – ***semi-mean axis***, number c – ***semi-minor axis*** of ellipsoid. If semi-axes do not satisfy the conditions $a \geq b \geq c$, then the equations (11.19) are not canonical. However, by renaming of the unknowns it is always possible to make the inequalities $a \geq b \geq c$ correct.

Plane sections give an opportunity to get a rough idea about the form of an ellipsoid (Fig. 11.12, *a*). , e.g. assuming $z=0$ in equation (11.19), we get the equation $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ of an intersection line of ellipsoid and coordinate plane Oxy. This equation on plane Oxy defines an ellipse (Section 10.2.2). Intersection lines of ellipsoid with other coordinate planes are also ellipses. They are called the ***principal profiles*** (***principal ellipses***) of ellipsoid.

Planes $x=\pm a$, $y=\pm b$, $z=\pm c$ define in space ***principal rectangular parallelepiped***, inside which an ellipsoid is situated (Fig. 11.12, *b*). Sides of the parallelepiped touch ellipsoid in its vertexes.

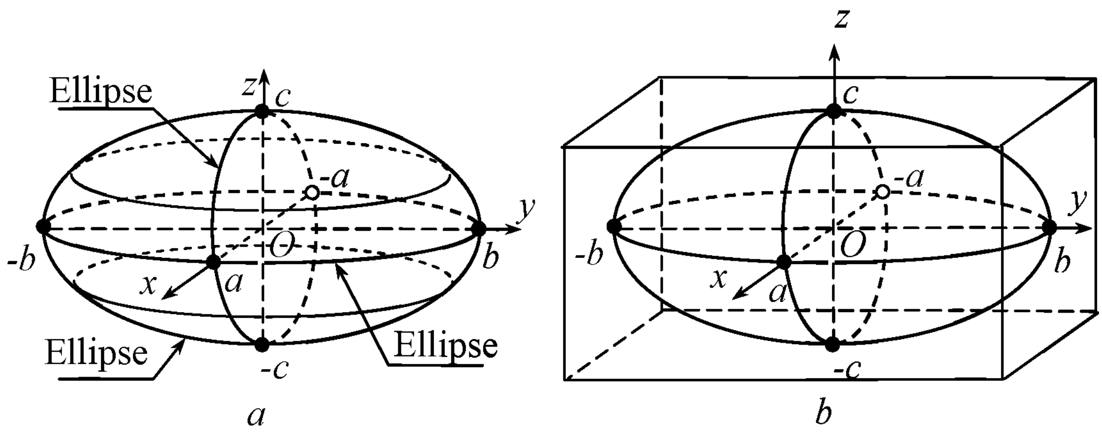


Figure 11.12

Ellipsoid, which semi-axes are pairwise different ($a > b > c$), is called ***three-axial*** (or ***general***). Ellipsoid with two equal semi-axes is called ***ellipsoid of revolution***, e.g. if $a = b$, then such surface can be obtained by the rotation of ellipse $\frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ (which is defined in plane Oyz) around axis Oz. If all semi-axes of ellipsoid are equal ($a = b = c = R$), then it represents a ***sphere*** $x^2 + y^2 + z^2 = R^2$ of radius R.

11.3.3. Hyperboloids

One-sheet hyperboloid is a surface, which is defined in some Cartesian coordinate system $Oxyz$ by the following canonical equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \tag{11.20}$$

Two-sheet hyperboloid is a surface, which is defined in some Cartesian coordinate system $Oxyz$ by the following canonical equations

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \tag{11.21}$$

In equations (11.20), (11.21) a, b, c are positive parameters ($a \geq b$), which specify hyperboloid.

The origin of coordinates is called ***center*** of hyperboloid. Points of intersection of hyperboloid and coordinate axes are called its ***vertexes***. These are four points of intersection $(\pm a,0,0)$, $(0,\pm b,0)$ for one-sheet hyperboloid (11.20) and two points $(0,0,\pm c)$ for two-sheet hyperboloid (11.21). Three segments of coordinate axes, which connect hyperboloid vertexes are called ***axes of hyperboloid***. Hyperboloid axes, which belong to coordinate axes Ox, Oy, are called ***lateral axes*** of hyperboloids, and axis, which belongs to applicate axis Oz, – ***longitudinal axis*** of hyperboloid. Numbers a, b, c, which are equal to a half of axis length, are called ***semi-axes*** of hyperboloid.

Plane sections give an opportunity to get a rough idea about the form of an *one-sheet hyperboloid*, e.g. assuming $z=0$ in equation (11.20), we get equation $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ of an intersection line of one-sheet hyperboloid and coordinate plane Oxy. This equation on plane Oxy defines ellipse (Section 10.2.2), which is called ***throat ellipse***. Intersection lines of a one-sheet hyperboloid and other coordinate planes are hyperbolas. They are called ***principal hyperbolas***, e.g. assuming $x=0$ we get principal hyperbola $\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$ and assuming $y=0$ – principal hyperbola $\frac{x^2}{a^2}-\frac{z^2}{c^2}=1$.

One-sheet hyperboloid can be expressed as a surface, that is formed by ellipses, which vertexes are situated on principal hyperbolas (Fig. 11.13, *a*). Section of one-sheet hyperboloid with a plane, which is parallel to applicate axis and which has the only common point with the throat ellipse (i.e. which touches it), is a pair of lines, which intersect in a tangency point, e.g. assuming $x=\pm a$ in equation (11.20), we get equation $\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$ of two intersection lines (Fig. 11.13, *a*).

Plane sections give an opportunity to get a rough idea about the form of an *two-sheet hyperboloid*. Sections of a two-sheet hyperboloid with coordinate planes Oyz and Oxz are hyperbolas (***principal hyperbolas***) and with planes, which are parallel to the plane Oxy are ellipses. Two-sheet hyperboloid can be expressed as a surface, that is formed by ellipses, which vertexes lie on principal hyperbolas (Fig.11.13, *b*).

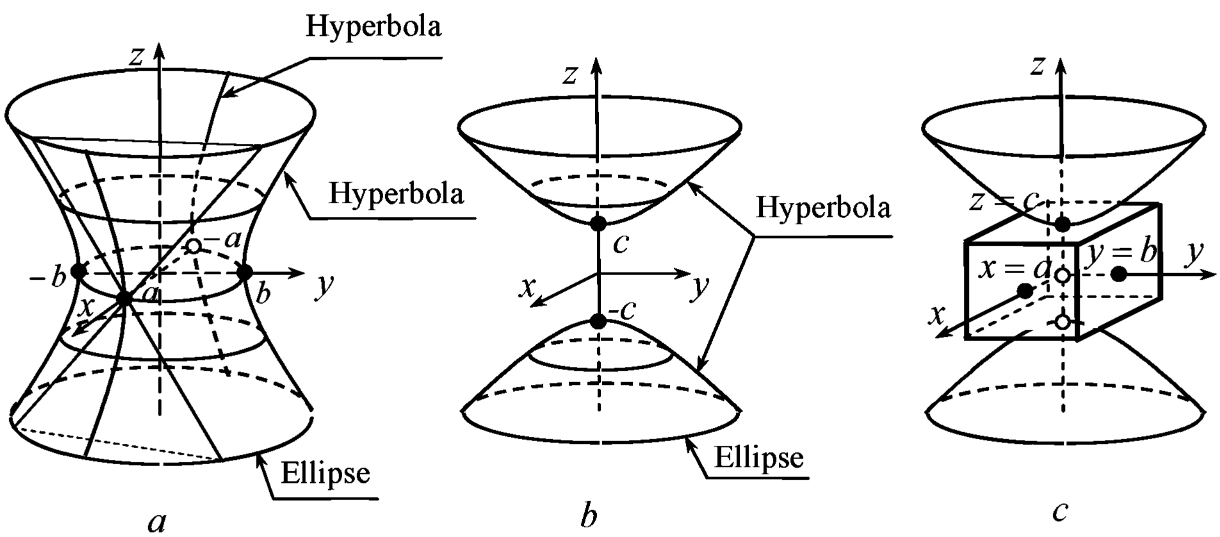


Figure 11.13

Planes $x=\pm a$, $y=\pm b$, $z=\pm c$ define in space ***principal rectangular parallelepiped***. Two sides ($z=\pm c$) of parallelepiped touch two-sheet hyperboloid in its vertexes (Fig. 11.13, *c*).

Hyperboloid with different latitude axes ($a\neq b$), is called ***three-axial*** (or ***general***). Hyperboloid with equal latitude axes ($a=b$) is called ***hyperboloid of revolution***. One-sheet and two-sheet hyperboloids can be obtained by the rotation of hyperbola $\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$ or conjugate hyperbola $\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1$ accordingly around the axis *Oz* (Section 10.2.3).

11.3.4. Cones

Cone is a surface, which is defined in some Cartesian coordinate system *Oxyz* by the following canonical equation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0, \qquad (11.22)$$

where a, b, c are positive parameters ($a\geq b$.), which specify cone.

The origin of coordinates is called the ***center*** of cone (fig.11.14), point O – ***vertex*** of cone (11.22), and any ray OM, which belongs to the cone, – its ***generator***.

Plane sections give an opportunity to get a rough idea about the form of a cone, e.g. sections of a cone with coordinate planes Oxz, Oyz are pairs of intersecting lines, which satisfy the planes equations $\frac{x^2}{a^2}-\frac{z^2}{c^2}=0$ (for $y=0$) and $\frac{y^2}{b^2}-\frac{z^2}{c^2}=0$ (for $x=0$) accordingly. Sections of cone with planes parallel to the plane Oxy, are ellipses. Cone can be expressed as a surface, that is formed by ellipses, which centers lie on the applicate axis and which vertexes belong to coordinate planes Oxz and Oyz (Fig. 11.14).

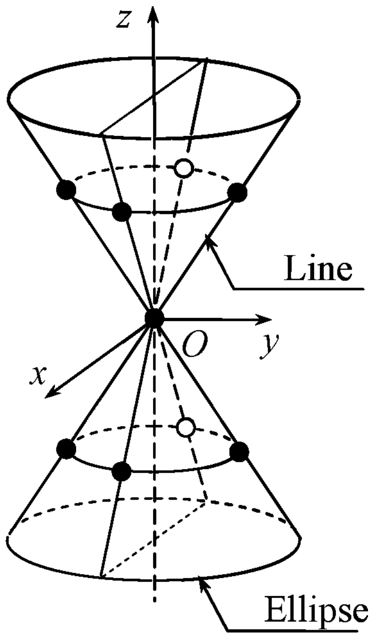


Figure 11.14

If $a=b$, then all sections of cone by planes $z=h$ $(h\neq 0)$ are circumference. Such cone is called ***right circular cone***. It can be obtained by the rotation of a line $z=\frac{c}{b}\cdot y$ (***generator***) around the applicate axis.

11.3.5. Paraboloids

Elliptic paraboloid is a surface, which is determined in some Cartesian coordinate system $Oxyz$ by the following canonical equation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=2\cdot z. \tag{11.23}$$

Hyperbolic paraboloid is a surface, which is defined in some Cartesian coordinate system $Oxyz$ by the following canonical equation

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=2\cdot z. \tag{11.24}$$

In equations (11.23), (11.24) a are b positive parameters (for elliptic paraboloid $a \geq b$), which specify paraboloids.

The origin of coordinates is called a ***vertex*** of each paraboloid ((11.23) or (11.24)).

Plane sections give an opportunity to get a rough idea about the form of an *elliptic paraboloid*, e.g. plane Oxz intersects elliptic paraboloid (11.23) along the line, which in this plane is given by equation $\frac{x^2}{a^2}=2\cdot z$, which is equivalent to the equation $x^2=2\cdot p\cdot z$ of parabola with focal parameter $p=a^2$. Section of paraboloid with plane Oyz is obtained by assuming $x=0$ in equation (11.23): $\frac{y^2}{b^2}=2\cdot z$. This equation is equivalent to the equation $y^2=2\cdot q\cdot z$ of parabola with focal parameter $q=b^2$. These sections are called ***principal parabolas*** of elliptic paraboloid (11.23). Sections of paraboloid with planes, which are parallel to plane Oxy, are ellipses. Elliptic paraboloid can be expressed as a surface, which is formed by ellipses, which vertexes lies on principal parabolas (Fig. 11.15, *a*).

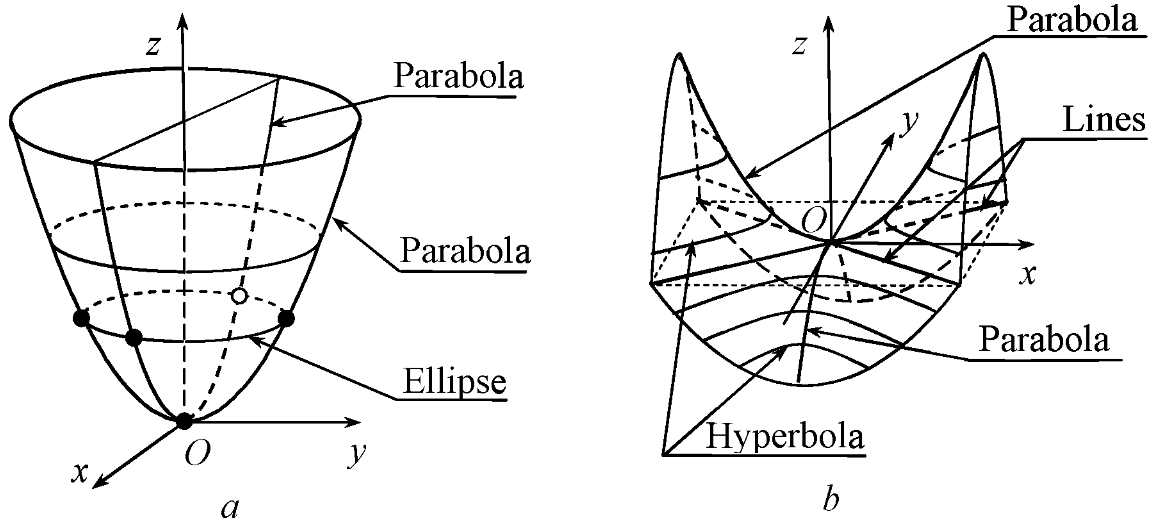


Figure 11.15

Elliptic paraboloid with $a=b$ is called ***paraboloid of revolution***. It can be obtained by the rotation of parabola $\mathbb{R}^{2\times3}$ (where $q=a^2=b^2$) around the axis Oz.

Plane sections give an opportunity to get a rough idea about the form of a *hyperbolic paraboloid*, e.g. sections of hyperbolic paraboloid with coordinate planes Oxz and Oyz are parabolas (***principal parabolas***) $x^2=2pz$ and $y^2=-2\cdot q\cdot z$ with parameters $p=a^2$ or $q=b^2$ accordingly. Since symmetry axes of principal parabolas are directed in opposite sides, hyperbolic paraboloid is called ***saddle surface***. Section of hyperbolic paraboloid with plane Oxy is a pair of line, which intersect in the origin, and section with a plane, which is parallel to the plane Oxy, is hyperbola.

Hyperbolic paraboloid can be expressed as a surface, which is formed by hyperbolas (including the "cross" from their asymptotes), which vertexes lie on principal parabolas (Fig. 11.15, *b*).

EXERCISES

1. Plane passes through the points $A(1,2,3)$, $B(-1,3,1)$, $C(3,-4,0)$. For the given plane find: a) general equation; b) parametric equation.

2. Find information about positional relationship of each pair of planes (are they skew, intersecting, parallel, equal, perpendicular, if they are intersecting find their mutual point):

a) $2\cdot x+2\cdot y+4\cdot z-12=0$, $\quad 3\cdot x-6\cdot y+1=0$;
b) $3\cdot x-2\cdot y-3\cdot z+5=0$, $\quad 9\cdot x-6\cdot y-9\cdot z-5=0$;
c) $2\cdot x-y-z-3=0$, $\quad 10\cdot x-5\cdot y-5\cdot z-15=0$;
d) $2\cdot x-y+4\cdot z-3=0$, $\quad x-6\cdot y-2\cdot z-1=0$.

3. Find information about positional relationship of each pair of lines (are they skew, intersecting, parallel, equal, perpendicular, if they are intersecting find their mutual point):

a) $\begin{cases} x=1+2\cdot t, \\ y=7+t, \quad t\in\mathbb{R}, \\ z=3+4\cdot t, \end{cases}$ $\quad \begin{cases} x=6+3\cdot t, \\ y=-1-2\cdot t, \; t\in\mathbb{R}; \\ z=-2+t, \end{cases}$

b) $\begin{cases} 2\cdot x+3\cdot y+2\cdot z=0, \\ x+z-8=0, \end{cases}$ $\quad \begin{cases} z-4=0, \\ 2\cdot x+3\cdot z-7=0; \end{cases}$

c) $\begin{cases} x = 9 \cdot t, \\ y = 5 \cdot t, \quad t \in \mathbb{R}, \\ z = -3 + t, \end{cases}$ $\qquad \begin{cases} 2 \cdot x - 3 \cdot y - 3 \cdot z - 9 = 0, \\ x - 2 \cdot y + z + 3 = 0; \end{cases}$

d) $\dfrac{x}{1} = \dfrac{y+2}{0} = \dfrac{z-3}{2}$, $\qquad \dfrac{x-4}{2} = \dfrac{y-1}{3} = \dfrac{z-6}{-1}$

e) $\dfrac{x}{-1} = \dfrac{y+8}{4} = \dfrac{z+3}{3}$, $\qquad \begin{cases} x + y - z = 0, \\ 2 \cdot x - y + 2 \cdot z = 0. \end{cases}$

4. Get information about positional relationship of each pair of line and plane (are they intersecting, perpendicular, parallel, if the line belongs to a plane, in case of intersection find mutual point):

a) $\dfrac{x-12}{4} = \dfrac{y-9}{3} = \dfrac{z-1}{1}$, $\qquad 3 \cdot x + 5 \cdot y - z - 2 = 0;$

b) $\begin{cases} x - 3 \cdot y + 2 \cdot z + 3 = 0, \\ 2 \cdot x + z - 3 = 0, \end{cases}$ $\qquad x - y - 2 \cdot z + 3 = 0;$

c) $\begin{cases} x = -1 + 2 \cdot t, \\ y = 3 + 4 \cdot t, \quad t \in \mathbb{R}; \\ z = 3 \cdot t; \end{cases}$ $\qquad 3 \cdot x - 3 \cdot y + 2 \cdot z - 5 = 0;$

d) $\begin{cases} x + 2 \cdot y + 3 \cdot z + 8 = 0, \\ 5 \cdot x + 3 \cdot y + z - 16 = 0, \end{cases}$ $\qquad 2 \cdot x - y - 4 \cdot z - 24 = 0.$

5. Define surface names and compose according canonical equations of the given algebraic surfaces of the second order written in Cartesian coordinate system:

a) $x^2 + y^2 - z^2 - 2 \cdot x - 2 \cdot y + 2 \cdot z = 0;$
b) $x^2 - y^2 - z^2 - 2 \cdot y - 1 = 0;$
c) $x^2 - 4 \cdot x + z + 3 = 0;$
d) $2 \cdot x^2 + 9 \cdot y^2 + 2 \cdot z^2 - 4 \cdot x \cdot y + 4 \cdot y \cdot z - 1 = 0;$
e) $3 \cdot x^2 + 3 \cdot y^2 + 3 \cdot z^2 - 8 \cdot x \cdot y - 6 \cdot y \cdot z = 0;$
g) $2 \cdot x^2 + 2 \cdot y^2 + z^2 - 10 \cdot x \cdot y + 20 \cdot x - 8 \cdot y + 29 = 0;$
h) $16 \cdot x^2 + 9 \cdot y^2 - z^2 - 24 \cdot x \cdot y - 9 \cdot x - 12 \cdot y + 4 \cdot z + 71 = 0.$

6. Define surface names and compose according canonical equations of the given algebraic surfaces of the second order written in Cartesian coordinate system:

a) $m \cdot x^2 + n \cdot y^2 - z^2 - 2 \cdot m \cdot n \cdot x - 2 \cdot m \cdot n \cdot y - 2 \cdot m \cdot n \cdot z + m^2 \cdot n \cdot (2 - n) = 0;$

b) $n \cdot x^2 + n \cdot y^2 + \dfrac{2}{m+n} \cdot z^2 + 2 \cdot m \cdot x \cdot y + 2 \cdot x \cdot z + 2 \cdot y \cdot z + m - n = 0.$

CHAPTER 12. LINEAR (VECTOR) SPACES

12.1. DEFINITION AND EXAMPLES OF LINEAR SPACES

Axioms of Linear Spaces

Linear (***vector***) ***space*** is a set $\mathbf{V}$ of arbitrary elements, that are called ***vectors***, in which the operations of vector addition and multiplication by a number are defined, i.e. any two vectors $\mathbf{u}$ and $\mathbf{v}$ have a corresponding vector $\mathbf{u}+\mathbf{v}$, which is called ***sum of vectors*** $\mathbf{u}$ and $\mathbf{v}$, any vector $\mathbf{v}$ and any number λ have a corresponding vector $\lambda\mathbf{v}$, which is called ***product of vector*** $\mathbf{v}$ ***and number*** λ, that the following conditions are satisfied:

1) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u} \quad \forall\mathbf{u},\mathbf{v}\in\mathbf{V}$; (addition commutativity)

2) $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w} \quad \forall\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbf{V}$; (addition associativity)

3) there exists such an element $\boldsymbol{o}\in\mathbf{V}$, which is called ***zero vector***, that $\mathbf{v}+\boldsymbol{o}=\mathbf{v} \quad \forall\mathbf{v}\in\mathbf{V}$;

4) for any vector $\mathbf{v}$ there exist such a vector $(-\mathbf{v})\in\mathbf{V}$, which is called ***opposite*** to vector $\mathbf{v}$, that $\mathbf{v}+(-\mathbf{v})=\boldsymbol{o}$;

5) $\lambda(\mathbf{u}+\mathbf{v})=\lambda\mathbf{u}+\lambda\mathbf{v} \quad \forall\mathbf{u},\mathbf{v}\in\mathbf{V},\ \forall\lambda\in\mathbb{R}$;

6) $(\lambda+\mu)\mathbf{v}=\lambda\mathbf{v}+\mu\mathbf{v} \quad \forall\mathbf{v}\in\mathbf{V},\ \forall\lambda,\mu\in\mathbb{R}$;

7) $\lambda(\mu\mathbf{v})=(\lambda\mu)\mathbf{v} \quad \forall\mathbf{v}\in\mathbf{V},\ \forall\lambda,\mu\in\mathbb{R}$;

8) $1\cdot\mathbf{v}=\mathbf{v} \quad \forall\mathbf{v}\in\mathbf{V}$.

Conditions 1–8 are called the ***axioms of linear space***. Equality sign between vectors means that it is the same element of set V on both sides of equation. Such vectors are called ***equal***.

Linear space is a nonempty set, because it necessarily has zero vector.

Operations of vector addition and multiplication of a vector by a number are called ***linear vector operations***.

Difference of vectors $\mathbf{u}$ and $\mathbf{v}$ is a sum of vector $\mathbf{u}$ and opposite vector $-v$; it is denoted as follows: $\mathbf{u}-\mathbf{v}=\mathbf{u}+(-\mathbf{v})$.

Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are called ***collinear*** (***proportional***), if there exists such a number λ, that $\mathbf{v} = \lambda\mathbf{u}$. Collinearity notion is applicable to any finite number of vectors. Zero vector $\boldsymbol{o}$ is collinear to any vector.

In the definition of linear space operation of vector multiplication by number is determined for real numbers. Such space is called ***linear space over the field of real numbers***, or simply ***real linear space***. If we substitute the field of real numbers $\mathbb{R}$ with the field of complex numbers $\mathbb{C}$, then we will obtain ***linear space over the field of complex numbers***, or simply ***complex linear space***.

Further, if there is no additional information, real linear spaces will be considered. In some cases for simplicity we will omit the word "linear", because all spaces considered in this section are linear.

Examples of Linear Spaces

1. Consider $\{\, \boldsymbol{o} \,\}$ – set, which consists of the only zero element, with operations $\boldsymbol{o} + \boldsymbol{o} = \boldsymbol{o}$ and $\lambda\boldsymbol{o} = \boldsymbol{o}$. For these operation axioms 1–8 are satisfied. Consequently, set $\{\, \boldsymbol{o} \,\}$ is a linear set over any numerical field. Such space is called ***zero*** space.

2. Consider V_1, V_2, V_3 – sets of geometric vectors (directed segments) on line, plane and in space accordingly with ordinary operations of vector addition and vector multiplication by a number. From the elementary geometry we get that all axioms 1–8 of linear space are satisfied. Consequently, sets V_1, V_2, V_3 are real linear spaces. Instead of free vectors we can consider corresponding sets of radius vectors, e.g. sets of vectors on plane, which have common tail, i.e. are applied to a fixed point of a plane, which is a real linear space.

Set of unit radius vectors does not form a linear space, because any sum of these vectors does not belong to the considered set.

3. Consider $\mathbb{R}^n$ – set of matrix-columns of sizes $n \times 1$ with operations of matrix addition and matrix multiplication by a number. Axioms 1–8 of linear space are satisfied for this space. Zero vector in this set is zero column $o = \begin{pmatrix} 0 & \cdots & 0 \end{pmatrix}^T$. Consequently, set $\mathbb{R}^n$ is real linear space.

Similarly, set $\mathbb{C}^n$ of columns of sizes $n \times 1$ with complex elements is a complex linear space.

Set of matrix-columns with nonnegative real elements is not a linear space, because it has no opposite elements.

4. Consider $\{Ax = o\}$ – set of solutions of homogeneous system $Ax = o$ of linear algebraic equations with n unknowns (where A is the matrix of system), as set of matrix-columns of sizes $n \times 1$ with operations of matrix addition and matrix multiplication by a number. Note, that this operations are determined on set $\{Ax = o\}$. From Property 1 of homogeneous system solutions it follows that sum of two homogenous system solutions and product of its multiplication by a number are also solutions of the system, i.e. they belong to the set $\{Ax = o\}$. Axioms of linear space for columns are satisfied (previous example). Thus, set of homogeneous system solutions is a real linear space.

Set $\{Ax = b\}$ of inhomogeneous system $Ax = b$ solutions ($b \neq o$), is not a linear space, because it has no zero element ($x = o$ is not the solution of inhomogeneous system).

5. Consider $\mathbb{R}^{m \times n}$ – set of matrices of sizes $m \times n$ with operation of matrix addition and matrix multiplication by a number. Axioms 1–8 of linear space for this set are satisfied. Zero element is a zero matrix O of corresponding sizes. Consequently, set $\mathbb{R}^{m \times n}$ is linear space.

6. Consider $P(\mathbb{C})$ – set of polynomials of the only variable and with complex coefficients. Operations of polynomial addition and multiplication by a number, which is considered as zero order polynomial, are defined and they satisfy axioms 1–8 (in particular, zero vector is a polynomial, which is identically equal to zero). Therefore, set $P(\mathbb{C})$ is a linear space over the field of complex numbers.

Set $P(\mathbb{R})$ of polynomials with real coefficients is also a linear space (obviously, over the field of real numbers).

Set $P_n(\mathbb{R})$ of polynomials of order not greater than n with real numbers is also a real linear space. Note, that operation of polynomial addition is defined on his set, because order of polynomial sum does not exceed order of its summands.

Set of polynomials of order n is not a linear space, because sum of such two polynomials can be a polynomial of lower order, which does not belong to the considered set.

Set of all polynomials of order not greater than n with positive coefficients is also not a linear space, because multiplication of such polynomial by negative number results into a polynomial, which does not belong to the considered set.

7. Consider $C(\mathbb{R})$ – set of real functions, determined and continuous on $\mathbb{R}$. Sum $(f+g)$ of functions f, g and product λf of multiplication of function $P_n(\mathbb{R})$ by real number λ are defined by the following equalities: $(f+g)(x)=f(x)+g(x)$, $(\lambda f)(x)=\lambda\cdot f(x)$ for all $x\in\mathbb{R}$. These operations are defined on $C(\mathbb{R})$, as sum of continuous function and product of multiplication of continuous function by a number are continuous functions, i.e. elements of $C(\mathbb{R})$. Let's check the correctness of linear space axioms. From the commutativity of real number addition it follows the correctness of the following equality $f(x)+g(x)=g(x)+f(x)$ for any $x\in\mathbb{R}$. Therefore $f+g=g+f$, i.e. axiom 1 is satisfied. Axiom 2 follows similarly from the addition associativity. Function $o(x)$, which is indentically equal to zero, can be considered as zero element (it is obviously continuous). For any function f the following equality is correct: $f(x)+o(x)=f(x)$, i.e. the axiom 3 is satisfied. The opposite element for function f is function $(-f)(x)=-f(x)$. Then $f+(-f)=o$ (axiom 4 is satisfied). Axioms 5, 6 follow from addition and multiplication by a number distributivity and axiom 7 – from associativity of multiplication by a number. The last axiom is satisfied, as multiplication of function by unit do not change the function: $1 \cdot f(x) = f(x)$ for any $x \in \mathbb{R}$, т.е. $1 \cdot f = f$. Thus, the considered set $C(\mathbb{R})$ with defined operations is a real linear space.

Similarly it can be proved that $C^1(\mathbb{R})$, $C^2(\mathbb{R}),\ldots$, $C^m(\mathbb{R}),\ldots$ (sets of functions with continuous derivatives of the first, second and etc. order accordingly) are linear spaces.

12.2. LINEAR DEPENDENCE AND LINEAR INDEPENDENCE OF VECTORS

Notions of Linear Dependence and Linear Independence of Vectors

For elements of a linear space operations of multiplication by a number and addition are defined. With these operations algebraic expressions can be composed.

Vector $\mathbf{v}$ is called a ***linear combination*** of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, if

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_k \mathbf{v}_k, \tag{12.1}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are some numbers. In this case it is said that ***vector*** $\mathbf{v}$ ***is decomposed by vectors*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ (***vector*** $\mathbf{v}$ ***is linearly expressed by vectors*** $v_1, v_2, \ldots, v_k$) and numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$ are called ***decomposition coefficients***. Linear combination with zero coefficients $\mathbf{v} = 0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + \ldots + 0 \cdot \mathbf{v}_k$ is called ***trivial***.

Set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ from $\mathbf{V}$ is called ***system of vectors***, and any part of the system – ***subsystem***.

System of k vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called ***linearly dependent***, if there exist such numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$, not all equal to zero simultaneously, that the following equation is correct

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_k \mathbf{v}_k = \boldsymbol{o}, \tag{12.2}$$

i.e. their linear combination is a zero vector.

System of k vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called ***linearly independent***, if equality (12.2) is possible only when $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$, i.e. linear combination in the left part of (12.2) is trivial. One vector $\mathbf{v}_1$ also forms the system: if $\mathbf{v}_1 = \boldsymbol{o}$ – linearly dependent, if $\mathbf{v}_1 \neq \boldsymbol{o}$ – linearly independent. ***Rank of the vector system*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is a maximum number of linearly independent vectors of the system; it is denoted by $\mathrm{rg}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$.

Property of Linearly Dependent and Linearly Independent Vectors

1. If system of vectors has zero vector, then it is linearly dependent.
2. If system of vectors has two equal vectors, then it is linearly dependent.
3. If system of vector has two proportional (collinear) vectors ($\mathbf{v}_i = \lambda \mathbf{v}_j$), then it is linearly dependent.
4. System of $k > 1$ vectors is linearly dependent if and only if there is at least one vector, which is a linear combination of others.
5. Any vectors, which are part of the linearly independent system, form linearly independent subsystem.
6. System of vectors, which has linearly dependent subsystem, is linearly dependent.
7. If system of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is linearly independent and after addition of vector $\mathbf{v}$ it becomes linearly dependent, then vector $\mathbf{v}$ can be uniquely decomposed by vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, i.e. decomposition coefficients (12.1) are defined unambiguously.
8. Let any vector of system $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_l$ can be decomposed by vectors of system $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, i.e. $\mathbf{u}_i = \sum_{j=1}^{k} a_{ji} \mathbf{v}_j$, $i = 1, \ldots, l$ (it is said that **system of vectors** $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_l$ **is linearly expressed by system of vectors** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$). Then if $l > k$, the system of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_l$ is linearly dependent.

Consider system of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ of real linear space $\mathbf{V}$ (i.e. over the field of real numbers $\mathbb{R}$). Set of all linear combinations of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called their ***linear span*** and it is denoted by

$$\mathrm{Span}(v_1, v_2, ..., v_k) = \{\ v:\ v = a_1 v_1 + a_2 v_2 + ... + a_k v_k;\ a_i \in \mathbb{R},\ \ i = 1, ..., k\ \}.$$

Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are called ***generators of linear span*** $\mathrm{Span}(v_1, v_2, ..., v_k)$. Also linear span can be denoted by $\mathrm{Lin}(v_1, v_2, ..., v_k)$.

Example 12.2. In space V_2 of radius vectors on plane (the second example of linear spaces) consider two noncollinear vectors $\bar{a} = \overline{OA}$ and $\bar{b} = \overline{OB}$. Find $\mathrm{Span}(\bar{a}, \bar{b})$.

□ Any radius vector $\bar{c} = \overline{OC}$ can be decomposed by two noncollinear vectors of this plane, i.e. it can be expressed as linear combination $\bar{c} = \alpha \cdot \bar{a} + \beta \cdot \bar{b}$, where $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$. Consequently, set of all possible linear combinations of $\bar{a}$ and $\bar{b}$ coincides with the whole space V_2 of radius vectors on a plane, i.e. $\mathrm{Span}(\bar{a}, \bar{b}) = V_2$. ■

Example 12.3. Prove that in space $P_2(\mathbb{R})$ of polynomials, which order is not greater than two (the sixth example of linear spaces), polynomials $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$ are linearly independent. Decompose polynomial $p(x) = (x+1)^2$ by $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$.

□ Compose linear combination of the given polynomials and equate it to zero (to zero element – polynomial identically equal to zero):

$$\lambda_1 p_0(x) + \lambda_2 p_1(x) + \lambda_3 p_2(x) = \lambda_1 \cdot 1 + \lambda_2 x + \lambda_3 x^2 = 0.$$

Identical equality to zero of a polynomial is possible in only one case: when all its coefficients are equal to zero, i.e. $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Consequently, considered polynomials are linearly independent. Write the given polynomial $p(x)$ as linear combination of polynomials $p_0(x)$, $p_1(x)$, $p_2(x)$:

$$p(x) = (x+1)^2 = x^2 + 2x + 1 = 1 \cdot p_0(x) + 2 \cdot p_1(x) + 1 \cdot p_2(x). ■$$

12.3. DIMENSIONALITY AND BASIS OF LINEAR SPACE

Dimensionality and Basis Notions

Linear space $\mathbf{V}$ is called ***n-dimensional***, if there exists a system of n linearly independent vectors and each system of bigger number of vectors is linearly dependent.

Number n is called ***dimensionality*** (***number of dimensions***) of linear space $\mathbf{V}$ and it is denoted by $\dim \mathbf{V}$. In other words, dimensionality of a space is a maximum number of linearly independent vectors of this space. If such number exists, then space is called ***finite-dimensional***. If for any natural number n in space $\mathbf{V}$ there exists a system of n linearly independent vectors, then this space is called ***infinite-dimensional*** (it is denoted by $\dim \mathbf{V} = \infty$). Further, if there is no additional information we will consider finite-dimensional spaces.

Basis of n-dimensional linear space is an ordered set of n linearly independent vectors (***basis vectors***). Basis of linear space is defined ambiguously, e.g. if $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is a basis of $\mathbf{V}$, then the system of vectors $\lambda \mathbf{e}_1$, $\lambda \mathbf{e}_2, \ldots, \lambda \mathbf{e}_n$ for any $\lambda \neq 0$ is also a basis of $\mathbf{V}$.

Number of basis vectors in different bases of the same finite-dimensional space is obviously the same, because this number equals to dimensionality of this space.

In some spaces, which often appear in different applications, one of possible bases, that is the most convenient from practical point of view, is called ***standard***.

Properties of Basis

1. If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is basis of n-dimensional linear space $\mathbf{V}$, then any vector $\mathbf{v} \in \mathbf{V}$ can be expressed as linear combination of these basis vectors:

$$\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \ldots + v_n \mathbf{e}_n \tag{12.3}$$

and moreover this expression is unique, i.e. the coefficients $v_1, v_2, \ldots, v_n$ are define unambiguously. In other words, any vector of the space can be uniquely decomposed by basis.

2. If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is basis of space $\mathbf{V}$, then $\mathbf{V} = \operatorname{Span}(e_1, e_2, \ldots, e_n)$, i.e. linear space is linear span of its basis vectors.

3. If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is linearly independent vector system of linear space $\mathbf{V}$ and any vector $\mathbf{v} \in \mathbf{V}$ can be expressed as a linear combination $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \ldots + v_n\mathbf{e}_n$, then space $\mathbf{V}$ has dimensionality n and system $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is its basis.

4. Any linearly independent system of k vectors of n-dimensional linear space $(1 \le k < n)$ can be complemented to the basis of this space.

Examples of Linear Spaces Bases

1. Zero linear space $\{\, \boldsymbol{o} \,\}$ has no linearly independent vectors. Therefore, its dimensionality equals to zero: $\dim\{\, \boldsymbol{o} \,\} = 0$. This space has no basis.

2. Spaces V_1, V_2, V_3 have dimensionalities 1, 2, 3 accordingly. Indeed, any nonzero vector of space V_1 forms linearly independent system (by definition), and any two nonzero vectors of V_1 are collinear, i.e. linearly dependent (example 12.1). Consequently, $\dim V_1 = 1$, and basis of this space V_1 is any nonzero vector. Similarly, it can be proved that $\dim V_2 = 2$ and $\dim V_3 = 3$. Basis of space V_2 is any ordered set of noncollinear vectors (one of them is assumed as the first basis vector, another – as the second). Basis of V_3 is an ordered triplet of noncoplanar vectors. Standard basis of V_1 is unit vector $\bar{i}$ on a line. Standard basis of V_2 is basis $\bar{i}$, $\bar{j}$, which consists of two mutually perpendicular unit vectors of plane. Standard basis in V_3 is basis $\bar{i}$, $\bar{j}$, $\bar{k}$, which consists of three unit pairwise perpendicular vectors, which form the right triplet.

3. In space $\mathbb{R}^n$ it is easy to find system of n linearly independent columns, e.g. columns of identity matrix, which are linearly independent

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots,\ e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Consequently, $\dim \mathbb{R}^n = n$. Space $\mathbb{R}^n$ is called ***n-dimensional real arithmetical space***. The given set of bases is considered as standard basis of space $\mathbb{R}^n$. Similarly, it can be proved, that $\dim \mathbb{C}^n = n$, therefore space $\mathbb{C}^n$ is called ***n-dimensional complex arithmetical space***.

4. Recall, that any solution of homogeneous system $Ax = o$ of linear equations with n unknowns can be expressed in a form $x = C_1\varphi_1 + C_2\varphi_2 + \ldots + C_{n-r}\varphi_{n-r}$, where $r = \operatorname{rg} A$ and φ_1, $\varphi_2, \ldots$, φ_{n-r} is fundamental system of solutions. Consequently, $\{Ax = o\} = \operatorname{Span}(\varphi_1, \varphi_2, \ldots, \varphi_{n-r})$, i.e. basis of space $\{Ax = o\}$ of homogeneous system solutions is its fundamental system of solutions, dimensionality of such space $\dim\{\, Ax = o \,\} = n - r$.

5. In space $\mathbb{R}^{2\times 3}$ of matrices of sizes 2×3 it is possible to choose 6 matrices:

$$e_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\quad e_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},\quad e_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$
$$e_4 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},\quad e_5 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\quad e_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which are linearly independent. Indeed, their linear combination

$$\alpha_1 \cdot e_1 + \alpha_2 \cdot e_2 + \alpha_3 \cdot e_3 + \alpha_4 \cdot e_4 + \alpha_5 \cdot e_5 + \alpha_6 \cdot e_6 = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_4 & \alpha_5 & \alpha_6 \end{pmatrix} \tag{12.4}$$

equals to zero matrix only in trivial case ($\alpha_1 = \alpha_2 = \ldots = \alpha_6 = 0$). Reading the equality (12.4) from right to left, we conclude, that any matrix from $\mathbb{R}^{2\times 3}$ can be linearly expressed via the chosen 6 matrices, i.e. $\mathbb{R}^{2\times 3} = \operatorname{Span}(e_1, e_2, \ldots, e_6)$. Consequently, $\dim \mathbb{R}^{2\times 3} = 2\cdot 3 = 6$, and matrices e_1, $e_2, \ldots$, e_6 are (standard) basis of this space. Similarly it can be proved, that $\dim \mathbb{R}^{m\times n} = m\cdot n$.

6. For any natural number n in space of polynomials $P(\mathbb{C})$ with complex coefficients it is possible to find n linearly independent elements, e.g. polynomials $e_1 = 1$, $e_2 = z$, $e_3 = z^2, \ldots$, $e_n = z^{n-1}$ are linearly independent, because its linear combination $a_1 \cdot e_1 + a_2 \cdot e_2 + \ldots + a_n \cdot e_n = a_1 + a_2 z + \ldots + a_n z^{n-1}$ equals to zero polynomial ($o(z) \equiv 0$) only in trivial case ($a_1 = a_2 = \ldots = a_n = 0$). Since this system of polynomials is linearly independent for any natural n, the space $P(\mathbb{C})$ is infinite-dimensional.

Similarly, we make a conclusion about infinite dimensionality of space $P(\mathbb{R})$ of polynomials with real coefficients.

Space $P_n(\mathbb{R})$ of polynomials of order not greater than n is finite-dimensional. Indeed, vectors $e_1 = 1$, $e_2 = x$, $e_3 = x^2, \ldots,$ $e_{n+1} = x^n$ form (standard) basis of this space, because they are linearly independent and any polynomial from $P_n(\mathbb{R})$ can be expressed as linear combinations of these vectors:

$$a_n x^n + \ldots + a_1 x + a_0 = a_0 \cdot e_1 + a_1 \cdot e_2 + \ldots + a_n \cdot e_{n+1}$$

Consequently, $\dim P_n(\mathbb{R}) = n+1$.

7. In space $T_\omega(\mathbb{R})$ of trigonometric binomials (with frequency $\omega \neq 0$) with real coefficients, basis is formed by monomials $e_1(t) = \sin\omega t$, $e_2(t) = \cos\omega t$. They are linearly independent, because identical equality $a\sin\omega t + b\cos\omega t \equiv 0$ is possible only in trivial case ($a = b = 0$). Any function $f(t) = a\sin\omega t + b\cos\omega t$ linearly expresses via basis binomials: $f(t) = a\, e_1(t) + b\, e_2(t)$. Hence, $\dim T_\omega(\mathbb{R}) = 2$.

12.4. COORDINATES AND COORDINATES TRANSFORMATIONS

Coordinates of Vectors in the Given Basis

Consider $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ as basis of linear space $\mathbf{V}$. Then each vector $\mathbf{v} \in \mathbf{V}$ can be decomposed by basis (Property 1 in Section 12.3), i.e. expressed in a form $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \ldots + v_n\mathbf{e}_n$, and coefficients $v_1, v_2, \ldots, v_n$ in decomposition are uniquely defined. These coefficients $v_1, v_2, \ldots, v_n$ are called ***coordinates of vector*** $\mathbf{v}$ in basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ (or relative to basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$).

Coordinates $v_1, v_2, \ldots, v_n$ of vector $\mathbf{v}$ is an ordered set of numbers, which is represented as a matrix-column $v = (v_1 \ \cdots \ v_n)^T$, is called ***coordinate column of vector*** $\mathbf{v}$ (in the given basis).

Vector and its coordinate column is denoted by the same letter – bold and light font accordingly.

If basis (as an ordered set of vectors) is expressed as a symbolic matrix-row $(\mathbf{e}) = (\mathbf{e}_1, \ldots, \mathbf{e}_n) = (\mathbf{e}_1 \quad \cdots \quad \mathbf{e}_n)$, then the decomposition of vector $\mathbf{v}$ by basis $(\mathbf{e})$ can be written in the following form:

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \ldots + v_n\mathbf{e}_n = (\mathbf{e}_1 \quad \cdots \quad \mathbf{e}_n)\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (\mathbf{e})\, v\,.$$

There multiplication of the symbolic matrix-row $(\mathbf{e})$ by numerical matrix-column v is calculated by the rules of matrix multiplication.

If it is necessary, when there are different bases in question, notation of a basis, relative to which a coordinate column was obtained, can be specified, e.g. $\underset{(\mathbf{e})}{v}$ – coordinate column of vector $\mathbf{v}$ relative to basis $(\mathbf{e}) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$.

By the Property 1 (Section 12.3) it follows, that *equal vectors has equal corresponding coordinates* (*in the same basis*), and vice versa, *if corresponding coordinates of vector are equal, such vectors are equal too*.

Linear Operations in Coordinate Form

Consider $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ – basis of linear space $\mathbf{V}$, vectors $\mathbf{u}$ and $\mathbf{v}$ in this basis have the following coordinates $u = (u_1 \quad \cdots \quad u_n)^T$ and $v = (v_1 \quad \cdots \quad v_n)^T$ accordingly, i.e.

$$\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \ldots + u_n\mathbf{e}_n, \qquad \mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \ldots + v_n\mathbf{e}_n. \qquad (12.5)$$

During the addition vector coordinates are summed up:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1)\,\mathbf{e}_1 + (u_2 + v_2)\,\mathbf{e}_2 + \ldots + (u_n + v_n)\,\mathbf{e}_n. \qquad (12.6)$$

During the multiplication by a number all coordinates are multiplied by this number:

$$\lambda\mathbf{v} = (\lambda v_1)\,\mathbf{e}_1 + (\lambda v_2)\,\mathbf{e}_2 + \ldots + (\lambda v_n)\,\mathbf{e}_n. \qquad (12.7)$$

In other words, *sum of vectors* $\mathbf{u}+\mathbf{v}$ *has coordinates* $u+v$, and *product* $\lambda\mathbf{v}$ *has coordinates* λv. Certainly, all coordinates are obtained in the same basis $(\mathbf{e})=(\mathbf{e}_1,...,\mathbf{e}_n)$.

Coordinates Transformation during Basis Change

Consider two bases of space $\mathbf{V}$: $(\mathbf{e})=(\mathbf{e}_1,...,\mathbf{e}_n)$ and $(\mathbf{e}')=(\mathbf{e}'_1,\mathbf{e}'_2,...,\mathbf{e}'_n)$. Basis $(\mathbf{e})$ we will call "old" and basis $(\mathbf{e}')$ – "new", then decomposition of each vector of new basis by the old one is given by:

$$\mathbf{e}'_i = s_{1i}\,\mathbf{e}_1 + s_{2i}\,\mathbf{e}_2 + ... + s_{ni}\,\mathbf{e}_n, \quad i=1,2,...,n. \tag{12.8}$$

Writing by columns the coordinates of vectors $(\mathbf{e}'_1,\mathbf{e}'_2,...,\mathbf{e}'_n)$ in basis $(\mathbf{e})$ it is possible to compose matrix:

$$S=\begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{pmatrix}. \tag{12.9}$$

Square matrix S, which is composed from coordinate columns of vectors from new basis $(\mathbf{e}')$ decomposed by old basis $(\mathbf{e})$, is called ***transition matrix*** from old basis to new one. By the transition matrix (12.9) formulas (12.8) can be rewritten as:

$$(\mathbf{e}'_1 \ \cdots \ \mathbf{e}'_n) = (\mathbf{e}_1 \ \cdots \ \mathbf{e}_n)\cdot S \text{ or simpler } (\mathbf{e}')=(\mathbf{e})\cdot S. \tag{12.10}$$

Multiplication of symbolic matrix-row $(\mathbf{e})$ by transition matrix S in (12.10) is calculated by the rules of matrix multiplication.

Consider in basis $(\mathbf{e})$ vector $\mathbf{v}$ with coordinates $v_1,v_2,\ldots,v_n$, and in $(\mathbf{e}')$ with coordinates $v'_1,v'_2,\ldots,v'_n$, i.e.

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + ... + v_n\mathbf{e}_n = v'_1\mathbf{e}'_1 + v'_2\mathbf{e}'_2 + ... + v'_n\mathbf{e}'_n$$

or simpler, $\mathbf{v}=(\mathbf{e})\,v=(\mathbf{e}')\,v'$.

Coordinate column of vector in old basis is obtained as the result of multiplication of transition matrix by the coordinate column of vector in new basis:

$$\underset{(\mathbf{e})}{v} = S \underset{(\mathbf{e}')}{v'} \quad \text{or} \quad \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{pmatrix} \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix}. \tag{12.11}$$

Properties of Transition Matrix

1. Consider three bases $(\mathbf{e})$, $(\mathbf{f})$, $(\mathbf{g})$ of space $\mathbf{V}$ and the following transition matrices: $\underset{(\mathbf{e})\to(\mathbf{f})}{S}$ from basis $(\mathbf{e})$ to basis $(\mathbf{f})$; $\underset{(\mathbf{f})\to(\mathbf{g})}{S}$ from basis $(\mathbf{f})$ to basis $(\mathbf{g})$; $\underset{(\mathbf{e})\to(\mathbf{g})}{S}$ from basis $(\mathbf{e})$ to basis $(\mathbf{g})$. Then

$$\underset{(\mathbf{e})\to(\mathbf{g})}{S} = \underset{(\mathbf{e})\to(\mathbf{f})}{S} \cdot \underset{(\mathbf{f})\to(\mathbf{g})}{S}. \tag{12.12}$$

2. If S is transition matrix from basis $(\mathbf{e})$ to basis $(\mathbf{f})$, then the matrix S is invertible and inverse matrix S^{-1} is transition matrix from basis $(\mathbf{f})$ to basis $(\mathbf{e})$. Coordinates of vector $\mathbf{v}$ in bases $(\mathbf{e})$ and $(\mathbf{f})$ are connected by the following formulas:

$$\underset{(\mathbf{e})}{v} = S \underset{(\mathbf{f})}{v}, \qquad \underset{(\mathbf{f})}{v} = S^{-1} \underset{(\mathbf{e})}{v}.$$

3. Any invertible square matrix of n-th order can be a transition matrix from one basis of n-dimensional linear space to another one.

Example 12.5. In two dimensional arithmetical space $\mathbb{R}^2$ there are two bases: $\mathbf{f}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, $\mathbf{f}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{g}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{g}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Find transition matrix $\underset{(\mathbf{f})\to(\mathbf{g})}{S}$ from basis $(\mathbf{f})$ to basis $(\mathbf{g})$ and coordinates of vector $\mathbf{v} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$ in each basis.

□ Consider standard basis $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ of space $\mathbb{R}^2$ (Section 12.3). Find coordinates of vectors $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{g}_1$, $\mathbf{g}_2$ in standard basis. Decompose vector $\mathbf{f}_1$:

$$\mathbf{f}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3 \cdot \mathbf{e}_1 + 2 \cdot \mathbf{e}_2, \text{ i.e. } f_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

In standard basis $(\mathbf{e})$ of space R^2 coordinate column f_1 coincides with $\mathbf{f}_1$. For other vectors similarly we get $f_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $g_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $g_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. From the coordinate columns we compose the transition matrix (12.9) from standard basis $(\mathbf{e})$ to the given bases $(\mathbf{f})$ and $(\mathbf{g})$: $\underset{(\mathbf{e})\to(\mathbf{f})}{S} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$, $\underset{(\mathbf{e})\to(\mathbf{g})}{S} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}$.

By the Property 1 of transition matrices we obtain $\underset{(\mathbf{f})\to(\mathbf{g})}{S} = \underset{(\mathbf{f})\to(\mathbf{e})}{S} \cdot \underset{(\mathbf{e})\to(\mathbf{g})}{S}$. By Property 2: $\underset{(\mathbf{f})\to(\mathbf{e})}{S} = \underset{(\mathbf{e})\to(\mathbf{f})}{S^{-1}}$. Therefore

$$\underset{(\mathbf{f})\to(\mathbf{g})}{S} = \underset{(\mathbf{e})\to(\mathbf{f})}{S^{-1}} \cdot \underset{(\mathbf{e})\to(\mathbf{g})}{S} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 4 & 5 \end{pmatrix}.$$

In standard basis $(\mathbf{e})$ of space $\mathbb{R}^2$ coordinate column $\underset{(\mathbf{e})}{v} = \begin{pmatrix} 6 \\ 9 \end{pmatrix}$ coincides with vector $\mathbf{v}$. Find coordinates of this vector in basis $(\mathbf{f})$ (by Property 2 of transition matrices):

$$\underset{(\mathbf{f})}{v} = \underset{(e)\to(\mathbf{f})}{S^{-1}} \underset{(\mathbf{e})}{v} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \begin{pmatrix} -3 \\ 15 \end{pmatrix}.$$

Indeed, the following decomposition is correct

$$v = \begin{pmatrix} 6 \\ 9 \end{pmatrix} = -3\begin{pmatrix} 3 \\ 2 \end{pmatrix} + 15\begin{pmatrix} 1 \\ 1 \end{pmatrix} = -3 \cdot f_1 + 15 \cdot f_2.$$

Find coordinates of vector $\mathbf{v}$ in basis $(\mathbf{g})$ in several ways:

$$\underset{(\mathbf{g})}{v} = \underset{(\mathbf{f})\to(\mathbf{g})}{S^{-1}} \underset{(\mathbf{f})}{v} = \begin{pmatrix} -1 & -2 \\ 4 & 5 \end{pmatrix}^{-1} \begin{pmatrix} -3 \\ 15 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 5 & 2 \\ -4 & -1 \end{pmatrix} \begin{pmatrix} -3 \\ 15 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix};$$

$$\underset{(\mathbf{g})}{v} = \underset{(\mathbf{e})\to(\mathbf{g})}{S^{-1}} \underset{(\mathbf{e})}{v} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

The obtained result confirms the decomposition:

$$v = \begin{pmatrix} 6 \\ 9 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 2 \end{pmatrix} + (-1)\begin{pmatrix} -1 \\ 1 \end{pmatrix} = 5 \cdot g_1 + (-1) \cdot g_2. \blacksquare$$

12.5. SUBSPACES OF LINEAR SPACE

12.5.1. Definition of Linear Subspace

Nonempty subset $\mathbf{L}$ of linear space $\mathbf{V}$ is called ***linear subspace*** of space $\mathbf{V}$, if:

1) $\mathbf{u}+\mathbf{v}\in\mathbf{L}\quad \forall\,\mathbf{u},\mathbf{v}\in\mathbf{L}$ (subspace is *closed relative to addition operation*);

2) $\lambda\mathbf{v}\in\mathbf{L}\quad \forall\ \mathbf{v}\in\mathbf{L}$ for any number λ (subspace is *closed relative to multiplication by a number operation*).

For linear subspace denotation we will use the following structure $\mathbf{L}\triangleleft\mathbf{V}$, and the word "linear" will be omitted for simplicity.

Note, that conditions 1, 2 in definition can be substituted with only one condition: $\lambda\mathbf{u}+\mu\mathbf{v}\in\mathbf{L}$ for any vectors $\mathbf{u},\mathbf{v}\in\mathbf{L}$ and any numbers λ and μ. Certainly, here and in definition it we speak about arbitrary numbers from the same numerical field, over which the linear space $\mathbf{V}$ is defined (Section 12.1.1).

Properties of Linear Subspaces

1. Any linear space $\mathbf{V}$ has two linear subspaces:

a) the space itself $\mathbf{V}$, i.e. $\mathbf{V}\triangleleft\mathbf{V}$;

b) zero subspace $\{\,o\,\}$, which consists of a unique zero vector of $\mathbf{V}$, i.e. $\{\,o\,\}\triangleleft\mathbf{V}$. These subspaces are called ***improper***, and the rest – ***proper***.

2. Any subspace $\mathbf{L}$ of linear space $\mathbf{V}$ is its subset: $\mathbf{L}\triangleleft\mathbf{V}\ \Rightarrow\ \mathbf{L}\subset\mathbf{V}$, but not every subset $M\subset\mathbf{V}$ is linear subset, because it can be unclosed relative to linear operations.

3. Subspace $\mathbf{L}$ of linear space $\mathbf{V}$ is linear space with the same operations of vector addition and vector multiplication by a number as in space $\mathbf{V}$, because they satisfy axioms 1–8 (see Section 12.1). Thus, it is possible to talk about subspace dimensionality, basis and etc.

4. Dimensionality of any subspace $\mathbf{L}$ of linear space $\mathbf{V}$ does not exceed dimensionality of $\mathbf{V}$: $\dim\mathbf{L}\le\dim\mathbf{V}$. If the dimensionality of subspace $\mathbf{L}\triangleleft\mathbf{V}$ equals to dimensionality of finite-dimensional space $\mathbf{V}$ ($\dim \mathbf{L} = \dim \mathbf{V}$), then the subspace coincides with the given space: $\mathbf{L} = \mathbf{V}$.

5. For any subset M of linear space $\mathbf{V}$ its linear span $\mathrm{Lin}(M)$ is subspace of $\mathbf{V}$ and $M \subset \mathrm{Lin}(M) \triangleleft \mathbf{V}$.

6. Linear span $\mathrm{Lin}(L)$ of subspace $\mathbf{L} \triangleleft \mathbf{V}$ coincides with the subspace $\mathbf{L}$, i.e. $\mathrm{Lin}(L) = L$.

12.5.2. Examples of Linear Subspaces

1. Space $\{ \boldsymbol{o} \}$, which consists of a unique zero vector of space $\mathbf{V}$, is a subspace, i.e. $\{ \boldsymbol{o} \} \triangleleft \mathbf{V}$.

2. Consider V_1, V_2, V_3 – sets of geometric vectors (directed segments) on a line, plane and in space accordingly. If the line belongs the plane, then $V_1 \triangleleft V_2 \triangleleft V_3$. On the contrary, set of unit vectors is not a linear subspace, because its multiplication by a number, which is not equal to 1, will result into vector, which does not belong to the initial subspace.

3. In n-dimensional arithmetical space $\mathbb{R}^n$ consider set L of "semi-zero" columns $x = (x_1 \ \cdots \ x_m \ 0 \ \cdots \ 0)^T$ with last $(n-m)$ elements equal to zero. Sum of such "semi-zero" columns is a column of the same form, i.e. addition operation is closed in L. Multiplication of "semi-zero" column by a number will result into "semi-zero" column, i.e. multiplication operation is closed in L. Therefore, $L \triangleleft \mathbb{R}^n$ and $\dim L = m$.

On the contrary, subset of nonzero columns in $\mathbb{R}^n$ is not a linear subspace, because multiplication by a zero results into zero column, which does not belong to the considered set. Examples of other subspaces of $\mathbb{R}^n$ are listed further.

4. Space $\{Ax = o\}$ of homogeneous system solutions with n unknowns is a subspace of n-dimensional arithmetical space $\mathbb{R}^n$. Dimensionality of this subspace is defined by a matrix of system: $\dim\{Ax = o\} = n - \mathrm{rg}\, A$.

Set $\{Ax=b\}$ of inhomogeneous system solutions ($b \neq o$) is not a subspace of $\mathbb{R}^n$, because the sum of two solutions of inhomogeneous system will not be the solution of this system.

5. In a space $\mathbb{R}^{n\times n}$ of square matrices of order n consider two subsets: set $\mathbb{R}^{n\times n}_{\text{sym}}$ of symmetric matrices and set $\mathbb{R}^{n\times n}_{\text{skew}}$ of skew-symmetric matrices. Sum of symmetric matrices is a symmetric matrix, i.e. addition operation is closed in $\mathbb{R}^{n\times n}_{\text{sym}}$. Multiplication of symmetric matrix by a number also results in symmetric matrix, i.e. multiplication operation is closed in $\mathbb{R}^{n\times n}_{\text{sym}}$. Consequently, set of symmetric matrices is a subspace of square matrices space, i.e $\mathbb{R}^{n\times n}_{\text{sym}} \lhd \mathbb{R}^{n\times n}$. It is easy to find the dimensionality of this subspace. Standard basis is formed by n matrices with the only nonzero element (it is equal to 1) on the main diagonal, and matrices with two nonzero elements (they are equal to 1), which are placed symmetrically relative to the main diagonal. In total there will be $n+(n-1)+\ldots+2+1=\dfrac{n(n+1)}{2}$ matrices in the basis. Consequently, $\dim \mathbb{R}^{n\times n}_{\text{sym}} = \dfrac{n(n+1)}{2}$. Similarly we find, that $\mathbb{R}^{n\times n}_{\text{skew}} \lhd \mathbb{R}^{n\times n}$ and $\dim \mathbb{R}^{n\times n}_{\text{skew}} = \dfrac{n(n-1)}{2}$.

Set of singular square matrices of n-th order is not a subspace of $\mathbb{R}^{n\times n}$, because the sum of two singular matrices can be not singular matrix in $\mathbb{R}^{2\times 2}$, e.g. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

6. In space of polynomials $P(\mathbb{R})$ with real coefficients it is possible to show the following sequence of subspaces $P_0(\mathbb{R}) \lhd P_1(\mathbb{R}) \lhd P_2(\mathbb{R}) \lhd \ldots \lhd P_n(\mathbb{R}) \lhd \ldots \lhd P(\mathbb{R})$.

Set of even polynomials ($p(-x)=p(x)$) is a linear subspace of $P(\mathbb{R})$, because the sum of even polynomials and multiplication of even polynomials by a number will be even polynomials.

Set of odd polynomials ($p(-x)=-p(x)$) is also a linear space. Set of polynomials with real roots is not a linear subspace, because addition of such polynomials can result into a polynomial with no real roots, e.g. $(x^2 - x) + (x+1) = x^2 + 1$.

7. In space $C(\mathbb{R})$ it is possible to show the following sequence of subspaces: $C(\mathbb{R}) \rhd C^1(\mathbb{R}) \rhd C^2(\mathbb{R}) \rhd \ldots \rhd C^m(\mathbb{R}) \rhd \ldots$.

Polynomials from $P(\mathbb{R})$ can be considered as functions, determined over $\mathbb{R}$. Since polynomial is a continuous function with derivatives of any order, it is possible to write: $P(\mathbb{R}) \lhd C(\mathbb{R})$ and $P_n(\mathbb{R}) \lhd C^m(\mathbb{R})$ $\forall\, m, n \in \mathbb{N}$.

Space of trigonometric binomials $T_\omega(\mathbb{R})$ is subspace of $C^m(\mathbb{R})$, because derivatives of any order of function $f(t) = \; = a\sin\omega t + b\cos\omega t$ are continuous, i.e. $T_\omega(\mathbb{R}) \lhd C^m(\mathbb{R})$ $\forall\, m \in \mathbb{N}$.

Set of continuous periodic function is not a subspace of $C(\mathbb{R})$, because the sum of two periodic functions can be a aperiodic function, e.g. $\sin t + \sin(\pi t)$.

EXERCISES

1. Prove that for the given linear space the system of vectors $(\mathbf{e})$ form valid basis. Decompose vector $(\mathbf{v})$ by this basis:

a) space $\mathbb{R}^2$: $\mathbf{e}_1 = \begin{pmatrix} 1 \\ m \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 2 \\ m+1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 3 \\ n \end{pmatrix}$;

b) space $\mathbb{R}^3$: $\mathbf{e}_1 = \begin{pmatrix} m \\ 1 \\ 1 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} m+1 \\ 2 \\ 1 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} m+2 \\ 2 \\ 1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} n \\ 3 \\ 2 \end{pmatrix}$;

c) space of polynomials P_2 of degree not exceeding 2: $\mathbf{e}_1(x) = x + m$, $\mathbf{e}_2(x) = x^2 - 1$, $\mathbf{e}_3(x) = x - m - 1$, $\mathbf{v}(x) = n \cdot x^2$.

2. Find transition matrix S from basis $(\mathbf{f})$ to basis $(\mathbf{g})$:

a) space $\mathbb{R}^2$: $\mathbf{f}_1 = \begin{pmatrix} m \\ 1 \end{pmatrix}, \mathbf{f}_2 = \begin{pmatrix} m+1 \\ 1 \end{pmatrix}, \mathbf{g}_1 = \begin{pmatrix} 1 \\ n \end{pmatrix}, \mathbf{g}_2 = \begin{pmatrix} 1 \\ n+1 \end{pmatrix}$;

b) space of symmetric matrices of the 2^{nd} order: $\mathbf{f}_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\mathbf{f}_2=\begin{pmatrix}0&0\\0&1\end{pmatrix},\mathbf{f}_3=\begin{pmatrix}0&1\\1&0\end{pmatrix},\mathbf{g}_1=\begin{pmatrix}m&0\\0&0\end{pmatrix},\mathbf{g}_2=\begin{pmatrix}0&0\\0&n\end{pmatrix},\mathbf{g}_3=\begin{pmatrix}m&m+n\\m+n&n\end{pmatrix}$;

c) space of polynomials P_2 of degree not exceeding 2: $\mathbf{f}_1(x)=1$, $\mathbf{f}_2(x)=x$, $\mathbf{f}_3(x)=x^2$, $\mathbf{g}_1(x)=x+m$, $\mathbf{g}_2(x)=x^2-n$, $\mathbf{g}_3(x)=x-m-1$.

3. Find dimensionality and basis of the given subspaces of $\mathbb{R}^4$:

a) $\{Ax=0\}$ – set of solutions of the system: $\begin{cases} x_1+x_2+nx_3+mx_4=0, \\ mx_1+nx_2+2x_3+3x_4=0; \end{cases}$

b) $\text{Lin}(a_1,a_2,a_3)$ – linear span of vectors $a_1=(1\quad 1\quad m\quad -n)^T$, $a_2=(0\quad 2\quad 0\quad m)^T$, $a_3=(1\quad 3\quad m\quad m-n)^T$.

4. Find transition matrix S from basis $(\mathbf{f})$ to basis $(\mathbf{g})$:

a) $f_1=\begin{pmatrix}3\\2\end{pmatrix}$, $f_2=\begin{pmatrix}2\\1\end{pmatrix}$, $g_1=\begin{pmatrix}2\\-2\end{pmatrix}$, $g_2=\begin{pmatrix}-1\\6\end{pmatrix}$;

b) $f_1=\begin{pmatrix}1\\2\\1\end{pmatrix}$, $f_2=\begin{pmatrix}2\\3\\3\end{pmatrix}$, $f_3=\begin{pmatrix}3\\7\\1\end{pmatrix}$, $g_1=\begin{pmatrix}3\\1\\4\end{pmatrix}$, $g_2=\begin{pmatrix}5\\2\\1\end{pmatrix}$, $g_3=\begin{pmatrix}1\\1\\-6\end{pmatrix}$.

5. In the space $\mathbb{R}^{2\times 2}$ of square second-order matrices, a set $\{AX=XA\}$ of matrices that are permutable with a matrix $A=\begin{pmatrix}m&n\\m-n&m+n\end{pmatrix}$ is given. Show that this set is a linear subspace in $\mathbb{R}^{2\times 2}$, find its dimensionality and basis.

CHAPTER 13. LINEAR MAPPINGS AND TRANSFORMATIONS

13.1. LINEAR MAPPINGS

13.1.1. Definition of Linear Mappings

Let's show main definitions, connected with the notion of mapping (function, operator).

Consider V and W – given sets. It is said, that ***mapping*** (***function***) f ***is defined*** on set V, if every element $v \in V$ has a unique corresponding element $f(v)$ of set W. Such a correspondence is called a ***mapping of a set*** V ***into a set*** W and it is denoted by $f: V \to W$ or $V \xrightarrow{f} W$.

If mapping f for an element $v \in V$ return a corresponding element $w \in W$, i.e. $w = f(v)$, element w is called ***image*** of v, and v – ***original*** of w.

Two mappings $f: V \to W$ and $g: V \to W$ are called ***equal***, if $f(v) = g(v)$ $\forall\, v \in V$.

Mapping $f: V \to W$ is called:

• ***injective***, if different elements of V have different images: $v_1 \neq v_2 \Rightarrow f(v_1) \neq f(v_2)$;

• ***surjective***, if for any element from W there is at least one original: $\forall w \in W\ \exists\, v \in V : w = f(v)$;

• ***bijective*** (***unambiguous***), if it is injective and surjective simultaneously.

Surjective mapping is also called a ***mapping of a set*** V ***to a set*** W.

Composition of mappings $g: U \to V$ and $f: V \to W$ is a mapping $f \circ g: U \to W$, which is defined by the equality $(f \circ g)(u) = f(g(u))$.

Mapping $\mathcal{E}_V : V \to V$ is called ***identical***, if any element of set V is associated with itself: $\mathcal{E}_V(v) = v\ \ \forall\, v \in V$.

Mapping $f^{-1}: W \to V$ is called ***inverse*** for the mapping $f: V \to W$, if $f^{-1} \circ f = \mathcal{E}_V : V \to V$ and $f \circ f^{-1} = \mathcal{E}_W : W \to W$. Mapping f is called ***invertible***, if there exists an inverse mapping for it. Necessity and sufficiency condition of invertibility is bijectivity (unambiguity) of a mapping.

Consider $\mathbf{V}$ and $\mathbf{W}$ – linear space (over the same numerical field). Mapping $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ is called ***linear***, if:

1) $\mathcal{A}(\mathbf{v}_1 + \mathbf{v}_2) = \mathcal{A}(\mathbf{v}_1) + \mathcal{A}(\mathbf{v}_2)$ $\forall\, \mathbf{v}_1 \in \mathbf{V}$, $\forall\, \mathbf{v}_2 \in \mathbf{V}$;

2) $\mathcal{A}(\lambda \cdot \mathbf{v}) = \lambda \cdot \mathcal{A}(\mathbf{v})$ $\forall\, \mathbf{v} \in \mathbf{V}$ for any number λ (from the given numerical field).

Condition 1 is called ***additivity*** of mapping, and condition 2 – ***homogeneity***. Space $\mathbf{V}$ is called ***space of originals***, and space $\mathbf{W}$ – ***space of images***.

Note, that conditions of additivity and homogeneity can be substituted with a unique condition of ***linearity***:

$$\mathcal{A}(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) = \lambda_1 \mathcal{A}(\mathbf{v}_1) + \lambda_2 \mathcal{A}(\mathbf{v}_2) \quad \forall\, \mathbf{v}_1 \in \mathbf{V},\ \forall\, \mathbf{v}_2 \in \mathbf{V}$$

for any numbers λ_1 and λ_2 from the given numerical field.

13.1.2. Properties of Linear Mappings

Consider $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ – linear mapping.

1. Linear mapping $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ associates zero element $\boldsymbol{o}_{\mathbf{V}}$ of $\mathbf{V}$ with zero elements $\boldsymbol{o}_{\mathbf{W}}$ of $\mathbf{W}$.

2. Linear mapping of linear combination is a linear combination of images: $\mathcal{A}\left(\sum_{i=1}^{k} \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^{k} \lambda_i \mathcal{A}(\mathbf{v}_i)$.

3. If vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly dependent, then their images are linearly dependent.

4. Consider $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ – surjective mapping of space $\mathbf{V}$ onto space $\mathbf{W}$ and vectors $\mathbf{w}_1, \ldots, \mathbf{w}_k$ of space $\mathbf{W}$, which form linearly independent system. Then in $\mathbf{V}$ there exists such linearly independent system of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, that $\mathcal{A}(\mathbf{v}_i) = \mathbf{w}_i$, $i = 1, \ldots, k$.

5. During linear surjective mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ of finite-dimensional space the dimensionality of image space does not exceed the dimensionality of original space, i.e. $\dim\mathbf{W}\le\dim\mathbf{V}$.

6. Composition of linear mapping is a linear mapping too.

7. If linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is invertible (unambiguous), then the inverse mapping $\mathcal{A}^{-1}:\mathbf{W}\to\mathbf{V}$ is linear.

8. Linear mapping of finite-dimensional space is unambiguously defined by images of basis vectors.

Linear Operations with Linear Mappings

Sum of mappings $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ and $\mathcal{B}:\mathbf{V}\to\mathbf{W}$ is a mapping $(\mathcal{A}+\mathcal{B}):\mathbf{V}\to\mathbf{W}$, which is defined by the following equality $(\mathcal{A}+\mathcal{B})(\mathbf{v})=\mathcal{A}(\mathbf{v})+\mathcal{B}(\mathbf{v})$ for any $\mathbf{v}\in\mathbf{V}$.

Product of mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ ***and number*** λ is a mapping $(\lambda\cdot\mathcal{A}):\mathbf{V}\to\mathbf{W}$, which is defined by the following equality $(\lambda\cdot\mathcal{A})(\mathbf{v})=\lambda\cdot\mathcal{A}(\mathbf{v})$ for any $\mathbf{v}\in\mathbf{V}$.

Sum of linear mapping and product of linear mapping and a number are linear mappings.

13.1.3. Examples of Linear Mappings

1. Denote by $\boldsymbol{O}:\mathbf{V}\to\mathbf{W}$ a zero mapping, which associates any vector $\mathbf{v}\in\mathbf{V}$ with zero element $\boldsymbol{o}_{\mathrm{W}}$ of space $\mathbf{W}$. Conditions of additivity and homogeneity of such mapping are, obviously, satisfied. This mapping is neither injective (different originals $\mathbf{v}_1$ and $\mathbf{v}_2$ are associated with the same image $\boldsymbol{o}_{\mathrm{W}}$) nor surjective (from all vectors of $\mathbf{W}$ only zero element has an original). Therefore, zero mapping is not bijective and consequently it is not invertible.

2. Consider in n-dimensional linear space $\mathbf{V}$ basis $\mathbf{e}_1,\ldots,\mathbf{e}_n$. Denote by æ:$\mathbf{V}\to\mathbb{R}^n$ mapping, which associates every vector $\mathbf{v}$ with its coordinate column $v=(v_1\ \cdots\ v_n)^T$ relative to the given basis. This mapping is linear, because during the addition of vectors of the same basis their coordinates are also summed up and during the multiplication of vector by a number coordinates of this vector are also multiplied by that number (Section 12.4). This mapping is injective (different vectors have different coordinates in the same basis) and surjective (for any column $v = (v_1 \quad \cdots \quad v_n)^T \in \mathbb{R}^n$ there exists an original $v = v_1\mathbf{e}_1 + \ldots + v_n\mathbf{e}_n$). Therefore, mapping æ is bijective and consequently invertible. On the other hand, mapping, which associates every vector $\mathbf{v} \in \mathbf{V}$ with column $v = (v_1 + 1 \quad \cdots \quad v_n + 1)^T \in \mathbb{R}^n$, is not linear, because the image of zero vector $\boldsymbol{o}_{\mathrm{V}} \in \mathbf{V}$ for such mapping is a column $(1 \quad \cdots \quad 1)^T \neq o$, which is not equal to zero.

3. Consider $P_n(\mathbb{R})$ and $P_{n-1}(\mathbb{R})$ – spaces of polynomials with real coefficients of order not greater than n and $(n-1)$ accordingly. Denote by $\mathcal{D}(p(x)) = \dfrac{dp(x)}{dx}$ polynomial derivative $p(x) \in P_n(\mathbb{R})$. Then the mapping (differentiation operator) $\mathcal{D}: P_n(\mathbb{R}) \to P_{n-1}(\mathbb{R})$ associates every polynomial $p(x) \in P_n(\mathbb{R})$ with its derivative, i.e. polynomial from $P_{n-1}(\mathbb{R})$. This operator is linear, because derivative of a sum equals to a sum of derivatives and derivative of a product of number and function equals to a product of derivative and that number. Differentiation operator is not injective (two polynomials with different constant terms have the same derivative) and it is surjective (for any polynomial $p_{n-1}(x)$ there is an original – polynomial from the set of primitives $\int p_{n-1}(x)\,dx + C$, where C is arbitrary constant). Therefore differentiation operator is not bijective and consequently it is noninvertible.

Integration operator $\mathcal{I}: P_{n-1}(\mathbb{R}) \to P_n(\mathbb{R})$, which associates polynomial $p_{n-1}(x) \in P_{n-1}(\mathbb{R})$ with polynomial

$$p_n(x) = \int_0^x p_{n-1}(t)\,dt,$$

is also linear (by properties of integral). This operator is injective (from the equality of images by the differentiation by the upper limit of integration we will obtain the equality of originals) and it is not surjective (polynomial with nonzero constant term has no original). Therefore, integration operator is not bijective and consequently it is noninvertible.

13.1.4. Matrix of Linear Mapping

Consider $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ – linear mapping of n-dimensional space $\mathbf{V}$ onto m-dimensional space $\mathbf{W}$. Fix in space $\mathbf{V}$ arbitrary basis $(\mathbf{e})=(\mathbf{e}_1,...,\mathbf{e}_n)$, and in space $\mathbf{W}$ basis $(\mathbf{f})=(\mathbf{f}_1,...,\mathbf{f}_m)$. Linear mapping is unambiguously determined by images of basis vectors (Property 8). Decompose images $\mathcal{A}(\mathbf{e}_i)$, $i=1,...,n$ of basis vectors $(\mathbf{e})$ by basis $(\mathbf{f})$:

$$\mathcal{A}(\mathbf{e}_i)=\sum_{j=1}^{m}a_{ji}\mathbf{f}_j\,,\quad i=1,...,n\,.$$

From the coordinate columns $\mathcal{A}(\mathbf{e}_1),\ldots,\mathcal{A}(\mathbf{e}_n)$ relative to basis $(\mathbf{f})$ compose matrix of sizes $m\times n$:

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}. \tag{13.1}$$

It is called ***matrix of linear mapping*** $\mathcal{A}$ ***in bases*** $(\mathbf{e})$ ***and*** $(\mathbf{f})$, or ***relative to bases*** $(\mathbf{e})$ ***and*** $(\mathbf{f})$. Matrix of mapping is also denoted by $\underset{(\mathbf{e}),(\mathbf{f})}{A}$, to emphasize its dependency on the chosen bases.

Matrix of mapping associates coordinates of image $\mathbf{w}=\mathcal{A}(\mathbf{v})$ and original $\mathbf{v}$. *If* $v=(v_1\ \cdots\ v_n)^T$ *is coordinate column of* $\mathbf{v}$, *and* $w=(w_1\ \cdots\ w_m)^T$ *is coordinate column of* $\mathbf{w}$ (i.e. $\mathbf{v}=v_1\mathbf{e}_1+...+v_n\mathbf{e}_n$ and $\mathbf{w}=w_1\mathbf{f}_1+...+w_m\mathbf{f}_m$), *then*

$$\begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad\Leftrightarrow\quad w=Av\,, \tag{13.2}$$

where A *is matrix* (13.1) *of mapping* $\mathcal{A}$.

To find matrix of mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ it is necessary to make the following steps:

1) specify bases $(\mathbf{e}) = (\mathbf{e}_1,...,\mathbf{e}_n)$ and $(\mathbf{f}) = (\mathbf{f}_1,...,\mathbf{f}_m)$ of spaces $\mathbf{V}$ and $\mathbf{W}$;

2) find image $\mathcal{A}(\mathbf{e}_1)$ of the first basis vector and decompose it by basis $(\mathbf{f})$. Obtained coordinates are written to the first column of matrix (13.1) of mapping $\mathcal{A}$;

3) find image $\mathcal{A}(\mathbf{e}_2)$ of the second basis vector and decompose it by basis $(\mathbf{f})$. Obtained coordinate are written to the second column of matrix (13.1) of mapping and etc. In the last column of matrix (13.1) we should write the coordinates of image $\mathcal{A}(\mathbf{e}_n)$ of the last basis vector.

Properties of Linear Mapping Matrices

For fixed bases of linear spaces:

1) *matrix of sum of linear mappings equals to the sum of their matrices*;

2) *matrix of multiplication of a matrix by a number equals to a product of mapping matrix and the same number*;

3) *matrix of an inverse mapping is inverse matrix of the mapping*;

4) *matrix of mapping composition* $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ *equals to the product of mapping matrices:* $C = BA$.

13.1.5. Kernel and Image of Linear Mapping

Kernel of linear mapping $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ is a set of such vectors $\mathbf{v} \in \mathbf{V}$, that $\mathcal{A}(\mathbf{v}) = \boldsymbol{o}_{\mathbf{W}}$, i.e. set of vectors from $\mathbf{V}$, which are associated with zero vector of space $\mathbf{W}$. Kernel of mapping $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ is denoted by:

$$\mathbf{Ker}\,\mathcal{A} = \{\, \mathbf{v} : \mathbf{v} \in \mathbf{V}, \mathcal{A}(\mathbf{v}) = \boldsymbol{o}_{\mathbf{W}} \}.$$

Image of linear mapping $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ is set of images $\mathcal{A}(\mathbf{v})$ of all vectors $\mathbf{v}$ from $\mathbf{V}$. Image of mapping $\mathcal{A}: \mathbf{V} \to \mathbf{W}$ is denoted by $\mathbf{Im}\,\mathcal{A}$ or $\mathcal{A}(\mathbf{V})$:

$$\mathbf{Im}\,\mathcal{A} = \mathcal{A}(\mathbf{V}) = \{\, \mathbf{w} : \mathbf{w} = \mathcal{A}(\mathbf{v}),\ \forall\, \mathbf{v} \in \mathbf{V} \}.$$

Note, that symbol $\mathbf{Im}\,\mathcal{A}$ should be distinguished from $\operatorname{Im} z$ – imaginary part of a complex number.

Examples of Kernels and Images of Linear Mappings

1. Kernel of zero mapping $\mathcal{O}:\mathbf{V}\to\mathbf{W}$ is whole space $\mathbf{V}$ and image consists of the only zero vector, i.e. $\mathbf{Ker}\,\mathcal{O}=\mathbf{V}$, $\mathbf{Im}\,\mathcal{O}=\{\,o_{\mathrm{W}}\,\}$.

2. Consider mapping æ:$\mathbf{V}\to\mathbb{R}^n$, which associates every vector $\mathbf{v}$ of n-dimensional linear space $\mathbf{V}$ its coordinate column $v=(v_1 \quad \cdots \quad v_n)^T$ relative to the given basis $\mathbf{e}_1,\ldots,\mathbf{e}_n$. Kernel of this mapping is zero vector $\boldsymbol{o}_{\mathrm{V}}$ of space $\mathbf{V}$, because it is the only vector, that has zero coordinate column æ$(\boldsymbol{o}_{\mathrm{V}})=o\in\mathbb{R}^n$. Image of mapping æ equals to the whole space $\mathbb{R}^n$, since the mapping is surjective (every column of $\mathbb{R}^n$ is a coordinate column of some vector in space $\mathbf{V}$).

3. Consider mapping $proj_{\bar{i}}:V_3\to\mathbb{R}$, which associates every vector $\bar{v}$ of three-dimensional space V_3 of geometric vectors with algebraic value $proj_{\bar{i}}(\bar{v})=(\bar{v},\bar{i})$ of its orthogonal projection to the axis, which is formed by vector $\bar{i}$, i.e. to abscissa axis $\bar{v}$. Kernel of this mapping is set of vectors $\mathrm{Lin}(\bar{j},\bar{k})$, which are perpendicular to vector $\bar{i}$. The image is the whole set of real numbers $\mathbb{R}$.

4. Consider mapping $\mathcal{D}:P_n(\mathbb{R})\to P_{n-1}(\mathbb{R})$, which associates every polynomial or order not greater than n with its derivative. Kernel of this mapping is set $P_0(\mathbb{R})$ of zero-ordered polynomials and image is whole set $P_{n-1}(\mathbb{R})$.

Properties of Kernel and Image of Linear Mappings

1. Kernel of any linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is a subspace: $\{\boldsymbol{o}_{\mathrm{V}}\}\triangleleft\mathbf{Ker}\,\mathcal{A}\triangleleft\mathbf{V}$.

2. Image of any linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is a subspace: $\mathbf{Im}\,\mathcal{A}\triangleleft\mathbf{W}$.

Since kernel and image of linear mapping are linear subspaces (Properties 1 and 2), it is possible to speak of their dimensionalities.

Nullity of linear mapping is dimensionality of its kernel: $d=\dim(\mathbf{Ker}\,\mathcal{A})$, and ***rank of linear mapping*** is dimensionality of its image: $\mathrm{rg}\,\mathcal{A}=r=\dim(\mathbf{Im}\,\mathcal{A})$.

3. Rank of linear mapping equals to the rank of its matrix (defined relative to any basis).

4. Linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is injective if and only if $\mathbf{Ker}\,\mathcal{A}=\{\,\boldsymbol{o}_{\mathrm{V}}\,\}$, in other words, when nullity of mapping equals to zero: $d=\dim\left(\mathbf{Ker}\,\mathcal{A}\right)=0$.

5. Linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is surjective if and only if $\mathbf{Im}\,\mathcal{A}=\mathbf{W}$, in other words, when mapping image rank equals to image space dimensionality: $r=\dim\left(\mathbf{Im}\,\mathcal{A}\right)=\dim\mathbf{W}$.

6. Linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is bijective (and invertible) if and only if $\mathbf{Ker}\,\mathcal{A}=\{\boldsymbol{o}_{\mathrm{V}}\}$ and $\mathbf{Im}\,\mathcal{A}=\mathbf{W}$ simultaneously.

7. Sum of kernel and image dimensionalities of any linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ equals to the dimensionality of originals:

$$\dim\left(\mathbf{Ker}\,\mathcal{A}\right)+\dim\left(\mathbf{Im}\,\mathcal{A}\right)=\dim\mathbf{V}. \tag{13.3}$$

8. Linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{W}$ is bijective (and invertible) if and only if its matrix is invertible (determined for any basis). Invertible linear mappings are also called **nonsingular** (meaning non-singularity of its matrix).

13.2. LINEAR TRANSFORMATIONS (OPERATORS)

13.2.1. Definition and Examples of Linear Transformations

Linear transformation (***linear operator***) of linear space $\mathbf{V}$ is linear mapping $\mathcal{A}:\mathbf{V}\to\mathbf{V}$ of space $\mathbf{V}$ onto itself.

Since linear transformation is a particular case of linear mapping, all properties of linear mapping are applied to it (injectivity, surjectivity, bijectivity, invertibility, kernel, image, nullity and rank and other notions).

Matrix of linear transformation $\mathcal{A}:\mathbf{V}\to\mathbf{V}$ ***in basis*** $\mathbf{e}_1,\ldots,\mathbf{e}_n$ ***of space*** $\mathbf{V}$ is a square matrix A, formed by coordinate columns of basis vector images $\mathcal{A}(\mathbf{e}_1),\ldots,\mathcal{A}(\mathbf{e}_n)$, which are found relative to basis $\mathbf{e}_1,\ldots,\mathbf{e}_n$.

Matrix of bijective linear transformation is invertible, i.e. nonsingular. Therefore, bijective (invertible) transformations are called ***nonsingular***.

13.2.2. Matrices of Linear Transformation Relative to Different Bases

Let's show the connection between matrices of linear transformation relative to different bases.

Let linear transformation $\mathcal{A}:\mathbf{V}\to\mathbf{V}$ *have in basis* $(\mathbf{e})=(\mathbf{e}_1,\ldots,\mathbf{e}_n)$ *matrix* $\underset{(\mathbf{e})}{A}$ *and in basis* $(\mathbf{f})=(\mathbf{f}_1,\ldots,\mathbf{f}_n)$ – *matrix* $\underset{(\mathbf{f})}{A}$. *If* S *is transition matrix from basis* $(\mathbf{e})$ *to basis* $(\mathbf{f})$ *then*

$$\underset{(\mathbf{f})}{A}=S^{-1}\underset{(\mathbf{e})}{A}S. \tag{13.4}$$

This formula demonstrates, that matrices of linear transformation in different bases are similar (sect. 6.2). And vice versa, every pair of similar matrices are matrices of some linear transformations, which were found relative to different bases.

EXERCISES

1. Find kernel and image of the linear transformation $\mathcal{A}:\mathbb{R}^3\to\mathbb{R}^3$, which matrix in standard basis of space $\mathbb{R}^3$ has the following matrix

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

Determine, whether this transformation is injective, surjective, bijective, invertible.

2. Linear transformation $\mathcal{A}:\mathbb{R}^2\to\mathbb{R}^2$ in basis $a_1=\begin{pmatrix}1\\2\end{pmatrix}$, $a_2=\begin{pmatrix}2\\3\end{pmatrix}$ has matrix $A=\begin{pmatrix}3&5\\4&3\end{pmatrix}$ and linear transformation $\mathcal{B}:\mathbb{R}^2\to\mathbb{R}^2$ in basis $b_1=\begin{pmatrix}3\\1\end{pmatrix}, b_2=\begin{pmatrix}4\\2\end{pmatrix}$ has matrix $B=\begin{pmatrix}4&6\\6&9\end{pmatrix}$. Find matrix of transformation $\mathcal{A}+\mathcal{B}$ in basis b_1,b_2.

REFERENCES

1. Beklemishev D.V. Kurs analiticheskoj geometrii i linejnoj algebry. – M.: Nauka, 1984. – 320 p.

2. Bortakovskii A.S., Panteleev A.V. Linejnaja algebra v primerah i zadachah. – M.: Vysshaja shkola, 2010. – 592 p.

3. Bortakovskii A.S., Panteleev A.V. Analiticheskaja geometrija v primerah i zadachah. – M.: Vysshaja shkola, 205. – 592 p.

4. Bortakovskii A.S., Panteleev A.V. Praktikum po linejnoj algebre i analiticheskoj geometrii. – M.: Vysshaja shkola, 2007. – 352 p.

5. Bortakovskii A.S., Panteleev A.V. Prakticheskij kurs linejnoj algebry i analiticheskoj geometrii. – M.: Universitetskaja kniga; Logos, 2008. – 328 p.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Беклемишев Д.В.* Курс аналитической геометрии и линейной алгебры. – М.: Наука, 1984.

2. *Бортаковский А.С., Пантелеев А.В.* Аналитическая геометрия в примерах и задачах. – М.: Высшая школа, 2005.

3. *Бортаковский А.С., Пантелеев А.В.* Линейная алгебра в примерах и задачах. – М.: Высшая школа, 2010.

4. *Бортаковский А.С., Пантелеев А.В.* Практикум по линейной алгебре и аналитической геометрии. – М.: Высшая школа, 2007.

5. *Бортаковский А.С., Пантелеев А.В.* Практический курс линейной алгебры и аналитической геометрии. – М.: Университетская книга; Логос, 2008.

Учебное издание

Бортаковский Александр Сергеевич

Пантелеев Андрей Владимирович

Пановский Валентин Николаевич

ЛИНЕЙНАЯ АЛГЕБРА И АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ

УЧЕБНОЕ ПОСОБИЕ

Издательство «Доброе слово»
Заказ книг: http://www.dobroeslovo.info

Подписано в печать: 25.03.2019
П.л. 16. Формат 60×90/16
Тираж 100 экз.