ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение высшего профессионального образования
«**ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ**»

V.V. Konev

LINEAR ALGEBRA, VECTOR ALGEBRA AND ANALYTICAL GEOMETRY

TextBook

Рекомендовано в качестве учебного пособия
Редакционно-издательским советом
Томского политехнического университета

Издательство
Томского политехнического университета
2009

UDC 517

V.V. Konev. Linear Algebra, Vector Algebra and Analytical Geometry. Textbook. Tomsk: TPU Press, 2009, 114 pp.

This textbook consists of 3 parts devoted to the mathematical methods of Linear Algebra and Analytical Geometry based on the vector analysis technique. The basic concepts are explained by examples and illustrated by figures.

The textbook is helpful for students who want to understand and be able to use matrix operations, solve systems of linear equations, analyze relative positions of figures, transform coordinate systems, and so on.

The textbook is designed to English speaking students.

Reviewed by: V.A. Kilin, Professor of the Higher Mathematics Department, TPU, D.Sc.

PREFACE

This textbook is intended for students who have already studied basic mathematics and need to study the methods of higher mathematics. It covers three content areas: Linear Algebra, Vector Algebra and Analytical Geometry. Each part contains basic mathematical conceptions and explains new mathematical terms. Many useful examples and exercises are presented in the textbook. explained and illustrated by examples and exercises.
The Linear Algebra topics include matrix operations, determinants and systems of linear equations.
In the section “Vector Algebra”, a main attention is paid to the geometrical applications of vector operations. The vector approach is considered to be basic for discussion of classic problems of Analytical Geometry.
The author welcomes reader’s suggestions for improvement of future editions of this textbook.

CONTENTS

VECTOR ALGEBRA

ANALYTICAL GEOMETRY

LINEAR ALGEBRA

1. Matrices

Matrices allow us to operate with arrays consisting of many numbers, functions or mathematical statements, just as if we operate with several items.

Matrices have a wide application in different branches of knowledge, for instance, in mathematics, physics, computer science, and so on. Matrices allow us to solve systems of ordinary equations or sets of differential equations, to predict the values of physical quantities in quantum theory, to encrypt messages in the Internet, and so on.

In this chapter, we discuss the basic concepts of the matrix theory, introduce matrix characteristics, and study some matrix applications. The important propositions are proved and illustrated by examples.

1.1. Basic Definitions

A **matrix** is a rectangular array of numbers, algebraic symbols or mathematical functions, provided that such arrays are added and multiplied according to certain rules.

Matrices are denoted by upper case letters: A, B, C, …

The size of a matrix is given by the number of rows and the number of columns. A matrix with m rows and n columns is called an $m \times n$ matrix (pronounce m-by-n matrix). The numbers m and n are the dimensions of the matrix. Two matrices have the same size, if their dimensions are equal.

Examples:

3×2 matrix	2×3 matrix	2×2 matrix
$A=\begin{pmatrix} 2 & -7 \\ 1 & 0 \\ 3 & \mathbf{4} \end{pmatrix}$	$B=\begin{pmatrix} -1 & \mathbf{5} & 0 \\ 3 & 3 & 8 \end{pmatrix}$	$C=\begin{pmatrix} \sin x & -\cos x \\ \cos x & \sin x \end{pmatrix}$

Members of a matrix are called its **matrix elements** or **entries**. The entry in the i-th row and the j-th column of a matrix A is denoted by $a_{i,j}$ or $A_{i,j}$. The subscripts indicate the row first and the column second.

In the examples above, the boldface elements are $a_{3,2} = 4$ and $b_{1,2} = 5$.

A matrix with one row is called a **row matrix**: $\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \end{pmatrix}$.

A matrix with one column is called a **column matrix**: $\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \cdots \\ a_{m,1} \end{pmatrix}$.

In the general form, a matrix is written as follows:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,j} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,j} & \cdots & a_{m,n} \end{pmatrix}$$

A short form of this equality is $A = \| a_{i,j} \|$.

A **square matrix** has as many rows as columns, the number of which determines the order of the matrix, that is, an $n \times n$ matrix is the matrix of the n -th **order**.

1.2. Matrix Operations

Equality of Matrices

Two matrices, $A = \| a_{i,j} \|$ and $B = \| b_{i,j} \|$, are equal, if they have the same sizes and their elements are equal by pairs, that is,

$$A = B \quad \Leftrightarrow \quad a_{i,j} = b_{i,j}$$
for each pair of indexes $\{i, j\}$.

Scalar Multiplication

Any matrix A may be multiplied on the right or left by a scalar quantity λ. The product is the matrix $B = \lambda A$ (of the same size as A) such that

$$b_{i,j} = \lambda a_{i,j}$$

for each $\{i, j\}$.

To multiply a matrix by a scalar, multiply every matrix element by that scalar.

Example: Let $A = \begin{pmatrix} 2 & -3 & 0 \\ 1 & 4 & -1 \end{pmatrix}$. Then $5A = \begin{pmatrix} 10 & -15 & 0 \\ 5 & 20 & -5 \end{pmatrix}$.

The Sum of Matrices

If $A = \| a_{i,j} \|$ and $B = \| b_{i,j} \|$ are matrices of the same size, then the sum, $A + B$, is the matrix $C = \| c_{i,j} \|$ such that

$$c_{i,j} = a_{i,j} + b_{i,j}$$

for each pair $\{i, j\}$.

To add matrices, add the corresponding matrix elements.

Example: Let $A = \begin{pmatrix} 3 & 7 & 1 \\ -1 & 2 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 6 & -15 & 3 \\ 4 & 1 & 2 \end{pmatrix}$. Then

$$A + B = \begin{pmatrix} 3 & 7 & 1 \\ -1 & 2 & 0 \end{pmatrix} + \begin{pmatrix} 6 & -15 & 3 \\ 4 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 9 & -8 & 4 \\ 3 & 3 & 2 \end{pmatrix}.$$

Multiplication of a Row by a Column

Let A be a row matrix having as many elements as a column matrix B.
In order to multiply A by B, it is necessary to multiply the corresponding elements of the matrices and to add up the products. Symbolically,

$$AB = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \end{pmatrix} \begin{pmatrix} b_{1,1} \\ b_{2,1} \\ \cdots \\ b_{n,1} \end{pmatrix} = a_{1,1}b_{1,1} + a_{1,2}b_{2,1} + \ldots + a_{1,n}b_{n,1} = \sum_{k=1}^{n} a_{1,k}b_{k,1}$$

Thus, multiplying a row matrix by a column matrix we obtain a number. Later we will show that any number can be considered as an 1×1 matrix.

To multiply a two-row matrix $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \end{pmatrix}$ by the column matrix $B = (B_1) = \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{n,1} \end{pmatrix}$, we multiply each row of A by the column of B. In this case, the product AB is the following 2×1 matrix:

$$AB = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} (B_1) = \begin{pmatrix} A_1 B_1 \\ A_2 B_1 \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,1} + a_{1,2}b_{2,1} + \ldots + a_{1,n}b_{n,1} \\ a_{2,1}b_{1,1} + a_{2,2}b_{2,1} + \ldots + a_{2,n}b_{n,1} \end{pmatrix}.$$

Similarly, the multiplication of an m-row matrix by an n-column matrix generates the $m \times n$ matrix.

Matrix Multiplication

The product of two matrices, *A* and *B*, is defined, if and only if the number of elements in a row of A equals the number of ones in a column of B.
Let *A* be an $m\times l$ matrix and *B* be an $l\times n$ matrix. Then the product AB is the $m\times n$ matrix such that its entry in the *i*-th row and the *j*-th column is equal to the product of the *i*-th row of A and the *j*-th column of B. If we denote the rows of *A* by $A_1, A_2, \dots$ and the columns of *B* by $B_1, B_2, \dots$, then

$$C = AB = \begin{pmatrix} A_1 \\ A_2 \\ \dots \\ A_m \end{pmatrix} \cdot \begin{pmatrix} B_1 & B_2 & \cdots & B_n \end{pmatrix} = \begin{pmatrix} A_1B_1 & A_1B_2 & \cdots & A_1B_n \\ A_2B_1 & A_2B_2 & \cdots & A_2B_n \\ \dots & \dots & \dots & \dots \\ A_mB_1 & A_mB_2 & \cdots & A_mB_n \end{pmatrix}$$

To find the element $c_{i,j}$ in the *i*-th row and the *j*-th column of the matrix $C = AB$, multiply the *i*-th row of *A* by the *j*-th column of *B*:

$$c_{i,j} = A_i B_j = \sum_{k=1}^{l} a_{i,k} b_{k,j} \ .$$

Note 1: The symbolic notation A^2 means the product of two equal square matrices: $A^2 = A\cdot A$.
Similarly, $A^3 = A\cdot A\cdot A$,

$$A^n = \underbrace{A\cdot A\cdot \ldots \cdot A}_{n}.$$

Note 2: In general, the product of matrices is not commutative: $AB \neq BA$.

Examples:

1) For each of the following matrices,

$$B = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \end{pmatrix}, \quad \text{and} \quad F = \begin{pmatrix} 2 & 1 \\ 1-1 & 2+1 \end{pmatrix},$$

determine whether it equals the matrix $A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ or not.

Solution: The dimensions of both matrices, *C* and *D*, differ from ones of *A*. Therefore, $A \neq C$ and $A \neq D$.
There are two matrices, *B* and *F*, which consist of the same elements as *A* and have the same order. However, the corresponding entries of *A* and *B* are not equal in pairs, and so $A \neq B$.
The matrix *F* satisfies all conditions of matrix equality, that is, *A* = *F*.

2) Let $A=\begin{pmatrix}1 & 3\\ 2 & -4\end{pmatrix}$ and $B=\begin{pmatrix}4 & 10\\ 5 & 15\end{pmatrix}$.

Solve for X the matrix equation

$$X + 4\,A = B.$$

Solution:

$$X = B-4A=\begin{pmatrix}4 & 10\\ 5 & 15\end{pmatrix}-4\begin{pmatrix}1 & 3\\ 2 & -4\end{pmatrix}$$

$$=\begin{pmatrix}4 & 10\\ 5 & 15\end{pmatrix}+\begin{pmatrix}-4 & -12\\ -8 & 16\end{pmatrix}=\begin{pmatrix}0 & -2\\ -3 & 31\end{pmatrix}.$$

3) Given two matrices $A=\begin{pmatrix}1 & 2 & 3\end{pmatrix}$ and $B=\begin{pmatrix}5\\ -4\\ 0\end{pmatrix}$, find the matrix products AB and BA.

Solution:

$$AB=\begin{pmatrix}1 & 2 & 3\end{pmatrix}\cdot\begin{pmatrix}5\\ -4\\ 0\end{pmatrix}=1\cdot5+2\cdot(-4)+3\cdot0=-3,$$

$$BA=\begin{pmatrix}5\\ -4\\ 0\end{pmatrix}\cdot\begin{pmatrix}1 & 2 & 3\end{pmatrix}=\begin{pmatrix}5\cdot1 & 5\cdot2 & 5\cdot3\\ -4\cdot1 & -4\cdot2 & -4\cdot3\\ 0\cdot1 & 0\cdot2 & 0\cdot3\end{pmatrix}=\begin{pmatrix}5 & 10 & 15\\ -4 & -8 & -12\\ 0 & 0 & 0\end{pmatrix}.$$

4) Let $A=\begin{pmatrix}2 & 1\\ 0 & 3\end{pmatrix}$ and $B=\begin{pmatrix}3 & 5\\ 4 & -1\end{pmatrix}$. Find the difference between matrix products AB and BA.

Solution:

$$AB=\begin{pmatrix}2 & 1\\ 0 & 3\end{pmatrix}\cdot\begin{pmatrix}3 & 5\\ 4 & -1\end{pmatrix}=\begin{pmatrix}2\cdot3+1\cdot4 & 2\cdot5+1\cdot(-1)\\ 0\cdot3+3\cdot4 & 0\cdot5+3\cdot(-1)\end{pmatrix}=\begin{pmatrix}10 & 9\\ 12 & -3\end{pmatrix},$$

$$BA=\begin{pmatrix}3 & 5\\ 4 & -1\end{pmatrix}\cdot\begin{pmatrix}2 & 1\\ 0 & 3\end{pmatrix}=\begin{pmatrix}3\cdot2+5\cdot0 & 3\cdot1+5\cdot3\\ 4\cdot2+(-1)\cdot0 & 4\cdot1+(-1)\cdot3\end{pmatrix}=\begin{pmatrix}6 & 18\\ 8 & 1\end{pmatrix},$$

$$AB-BA=\begin{pmatrix}10 & 9\\ 12 & -3\end{pmatrix}-\begin{pmatrix}6 & 18\\ 8 & 1\end{pmatrix}=\begin{pmatrix}4 & -9\\ 4 & -4\end{pmatrix}.$$

5) Find A^{2002} , if $A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$.

Solution:

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

$$A^3 = A\cdot A^2=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}=\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix},$$

$$A^4 = A\cdot A^3=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix}=\begin{pmatrix}1 & 4\\ 0 & 1\end{pmatrix},\ \text{and so on.}$$

$$A^{2002}=\begin{pmatrix}1 & 2002\\ 0 & 1\end{pmatrix}.$$

1.3. Types of Matrices

In a square matrix $A=\|a_{i,j}\|$, the elements $a_{i,i}$, with i = 1, 2, 3, ... , are called the **diagonal matrix elements**. The set of the entries $a_{i,i}$ forms the leading (or principle) diagonal of the matrix.

A square matrix $A=\|a_{i,j}\|$ is called a **diagonal matrix**, if off-diagonal elements are equal to zero or, symbolically, $a_{i,j}=0$ for all $i\neq j$:

$$A=\begin{pmatrix}a_{1,1} & 0 & \cdots & 0\\ 0 & a_{2,2} & \cdots & 0\\ \vdots & \cdots & \cdots & \vdots\\ 0 & 0 & \cdots & a_{n,n}\end{pmatrix}.$$

Identity matrices *I* are square matrices such that

$$I\cdot A=A \quad\text{and}\quad A\cdot I=A.$$

Compare these matrix equalities with the corresponding property of real numbers: $1\cdot a=a$ and $a\cdot 1=a$.

Theorem: Any identity matrix *I* is a diagonal matrix whose diagonal elements are equal to unity:

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

This theorem is proved in the following section.

Examples:

1) It is not difficult to verify that

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Therefore, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity matrix of the second order.

2) Let $A = \| a_{i,j} \|$ be any 2×3 matrix. Then

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} \quad \text{and}$$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}.$$

A matrix is called a **zero-matrix** (0-matrix), if it consists of only zero elements: $a_{i,j} = 0$ for each $\{i, j\}$.

In a short form, a zero-matrix is written as 0:

$$0 \equiv \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

By the definition of a zero-matrix,

$$0 \cdot A = A \cdot 0 = 0$$

and

$$A + 0 = A,$$

that is, a zero-matrix has just the same properties as the number zero.

However, if the product of two matrices is equal to zero, it does not mean that at least one of the matrices is a zero-matrix.

For instance, both matrices, $A = \begin{pmatrix} 0 & 1 \\ 0 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 2 \\ 0 & 0 \end{pmatrix}$, are non-zero matrices, while their product is a zero-matrix:

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

A square matrix has a **triangular form**, if all its elements above or below the leading diagonal are zeros:

$$\text{all} \quad a_{i,j} = 0 \quad \text{for} \quad i > j \quad \text{or} \quad \text{for} \quad i < j.$$

Examples:

Upper-triangular matrix	**Lower-triangular matrix**
$A = \begin{pmatrix} 2 & a & 5 \\ 0 & b & c \\ 0 & 0 & 4 \end{pmatrix}$	$B = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 7 & 0 \\ -2 & 0 & 3 \end{pmatrix}$

Given an $m \times n$ matrix $A = \| a_{i,j} \|$, the **transpose of *A*** is the $n \times m$ matrix $A^T = \| a_{j,i} \|$ obtained from *A* by interchanging its rows and columns.

This means that the rows of the matrix *A* are the columns of the matrix A^T; and *vise versa*:

$$(A^T)_{i,j} = a_{j,i}.$$

For instance, the transpose of $A = \begin{pmatrix} 2 & -7 \\ 1 & 0 \\ 3 & 4 \end{pmatrix}$ is $A^T = \begin{pmatrix} 2 & 1 & 3 \\ -7 & 0 & 4 \end{pmatrix}$.

A square matrix $A = \| a_{i,j} \|$ is called a **symmetric** matrix, if *A* is equal to the transpose of *A*: $\quad A = A^T \quad \Leftrightarrow \quad a_{i,j} = a_{j,i}.$

The **examples** below illustrate the structure of symmetric matrices:

$$R = \begin{pmatrix} a & c \\ c & b \end{pmatrix} = R^T \quad \text{and} \quad S = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} = S^T.$$

A square matrix $A = \| a_{i,j} \|$ is called a **skew-symmetric** matrix, if *A* is equal to the opposite of its transpose:

$$a_{i,j} = -a_{j,i}.$$

The **example** below shows the structure of a skew-symmetric matrix:

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

1.4. Kronecker Delta Symbol

The Kronecker delta symbol is defined by the formula

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

The delta symbol cancels summation over one of the indexes in such expressions as

$$\sum_i a_i \delta_{i,j}\,, \qquad \sum_j a_j \delta_{i,j}\,,$$

$$\sum_i a_{i,j} \delta_{i,j}\,, \qquad \sum_j a_{i,j} \delta_{i,j}\,, \text{ and so on.}$$

For instance, the sum $\sum\limits_i a_i \delta_{i,j}$ may contain only one nonzero term $a_j \delta_{j,j} = a_j$, while all the other terms are equal to zero, because of $\delta_{i,j} = 0$ for any $i \neq j$.

$$\boxed{\begin{array}{c} \text{If} \quad k \le j \le n, \text{ then } \sum\limits_{i=k}^{n} a_i \delta_{i,j} = a_j. \\ \text{If} \quad j < k \quad \text{or} \quad j > n, \text{ then } \sum\limits_{i=k}^{n} a_i \delta_{i,j} = 0. \end{array}}$$

Likewise, if $k \le i \le n$, then $\sum\limits_{j=k}^{n} a_j \delta_{i,j} = a_i$.

Otherwise, if $i < k$ or $i > n$, then $\sum\limits_{j=k}^{n} a_j \delta_{i,j} = 0$.

Examples:

$$\sum_{i=1}^{100} i^2 \delta_{i,3} = 1^2 \cdot 0 + 2^2 \cdot 0 + 3^2 \cdot 1 + 4^2 \cdot 0 + \ldots = 9.$$

$$\sum_{k=1}^{20} 2^k \delta_{k,10} = 2^{10} = 1024, \quad \text{however} \quad \sum_{k=1}^{5} 2^k \delta_{k,10} = 0.$$

Now we can easily prove the above-mentioned **theorem of identity matrix**:

$$I = \| \delta_{i,j} \| = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

The theorem states that the $I = \|\delta_{i,j}\|$ is an identity matrix. Therefore, we have to prove that $A \cdot I = A$ for any matrix A.

Proof: Let A be an arbitrary $m \times n$ matrix and $\|\delta_{i,j}\|$ be the square matrix of the *n-th* order. Then the matrix product $A \cdot I$ is the matrix of the same size as A.

By the definition of the matrix product and in view of the properties of the delta symbol, we obtain that

$$(A \cdot I)_{i,j} = \sum_{k=1}^{n} a_{i,k}\delta_{k,j} = a_{i,j}$$

for each pair of indexes $\{i, j\}$.

The equality of the corresponding matrix elements implies the equality of the matrices: $A \cdot I = A$.

1.5. Properties of Matrix Operations

Properties involving Addition

1. For any matrix A there exists the opposite matrix $(-A)$ such that
$$A + (-A) = A - A = 0.$$
2. If A and B are matrices of the same size, then
$$A + B = B + A.$$
3. If A, B, and C are matrices of the same size, then
$$(A + B) + C = A + (B + C).$$
4. The transpose of the matrix sum is the sum of the transpose of the matrices:
$$(A + B)^T = A^T + B^T.$$

The above properties of matrices result from the properties of real numbers. The proofs are left to the reader.

Properties involving Multiplication

1. Let A be a matrix. If λ and μ are scalar quantities, then
$$\lambda(\mu A) = (\lambda \mu)A.$$
2. Let A and B be two matrices such that the product AB is defined. If λ is a scalar quantity, then
$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$
3. Let A, B, and C be three matrices such that all necessary multiplications are appropriate. Then
$$(AB)C = A(BC).$$

4. Let A and B be two matrices such that the product AB is defined. Then
$$(AB)^T = B^T A^T .$$
5. If A and B are two diagonal matrices of the same order, then
$$AB = BA .$$

Properties 1) and 2) simply result from the properties of real numbers and the definition of the scalar multiplication.

To prove Property 3, we have to show that the corresponding elements of the two matrices, $(AB)C$ and $A(BC)$, are equal.

By the definition, the matrix element in the *i*-th row and the *k*-th column of the matrix AB is
$$(AB)_{i,k} = \sum_l a_{i,l} b_{l,k} .$$
The matrix element in the *i*-th row and the *j*-th column of the matrix $(AB)C$ can be expressed as
$$((AB)C)_{i,j} = \sum_k (AB)_{i,k} C_{k,j} = \sum_k \sum_l a_{i,l} b_{l,k} c_{k,j} .$$
By changing the order of summation, we obtain
$$((AB)C)_{i,j} = \sum_l \sum_k a_{i,l} b_{l,k} c_{k,j} = \sum_l a_{i,l} \sum_k b_{l,k} c_{k,j}$$
$$= \sum_l a_{i,l} (BC)_{l,j} = (A(BC))_{i,j} .$$
The equality of the corresponding matrix elements is satisfied that implies the equality of the matrices: $(AB)C = A(BC)$.

To demonstrate Property 4, we transform the entry in the *i*-th row and the *j*-th column of the matrix $(AB)^T$. In view of the definition of the transpose of matrix,
$$(AB)^T{}_{i,j} = (AB)_{j,i} = \sum_k a_{j,k} b_{k,i}$$
$$= \sum_k A^T{}_{k,j} B^T{}_{i,k} = \sum_k B^T{}_{i,k} A^T{}_{k,j} = (B^T A^T)_{i,j} .$$
Thus, $(AB)^T$ and $(B^T A^T)$ obey the conditions of equality of matrices.

Property 5 is based on the following reasons: 1) diagonal matrices are symmetric ones; 2) the product of diagonal matrices is a diagonal matrix. Therefore, we need only to show that $(AB)_{i,i} = (BA)_{i,i}$. Indeed,
$$(AB)_{i,i} = \sum_k a_{i,k} b_{k,i} = \sum_k a_{k,i} b_{i,k} = \sum_k b_{i,k} a_{k,i} = (BA)_{i,i} .$$

Properties involving Addition and Multiplication

1. Let A, B, and C be three matrices such that the corresponding products and sums are defined. Then

$$A(B+C) = AB + AC\,,$$

$$(A+B)C = AC + BC\,.$$

2. Let A and B be two matrices of the same size. If λ is a scalar, then

$$\lambda(A+B) = \lambda A + \lambda B\,.$$

To prove Property 1, consider the element on the *i*-th row and the *j*-th column of the matrix $A(B+C)$.

By the definition of the matrix product and in view of the addition properties, we have

$$\begin{aligned}(A(B+C))_{i,j} &= \sum_k a_{i,k}(B+C)_{k,j} = \sum_k a_{i,k}(b_{k,j}+c_{k,j}) \\ &= \sum_k a_{i,k}b_{k,j} + \sum_k a_{i,k}c_{k,j} = (AB)_{i,j} + (AC)_{i,j} = (AB+AC)_{i,j}\end{aligned}$$

for each pair of indexes $\{i, j\}$.

Therefore, the matrices $A(B + C)$ and $(AB + AC)$ are equal.

The equality of the matrices $(A + B)C$ and $(AC + BC)$ can be proven in a similar way:

$$\begin{aligned}((A+B)C)_{i,j} &= \sum_k (A+B)_{i,k} C_{k,j} = \sum_k (a_{i,k}+b_{i,k})c_{k,j} \\ &= \sum_k a_{i,k}c_{k,j} + \sum_k b_{i,k}c_{k,j} = (AC)_{i,j} + (BC)_{i,j} = (AC+BC)_{i,j}.\end{aligned}$$

The corresponding matrix elements are equal by pairs. Hence, the matrices are equal.

Property 2 results from the properties of real numbers. The proof can be performed by the reader.

Short Summary: Operations with matrices, such as addition and multiplication, have similar properties as that with usual real numbers. Numerical matrices of the first order can be interpreted as usual real numbers, that is, $\| a_{1,1} \| \equiv a_{1,1}$. The set of matrices is a generalization of the set of real numbers.

Examples:

1) Let $A=\begin{pmatrix}1 & 2\end{pmatrix}$, $B=\begin{pmatrix}3 & -1\\ 0 & 4\end{pmatrix}$, and $C=\begin{pmatrix}-2 & 4 & 1\\ 5 & 0 & 2\end{pmatrix}$.

By a straightforward procedure, show that $(AB)C = A(BC)$.

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

$$(A\cdot B)C=\begin{pmatrix}3 & 7\end{pmatrix}\cdot\begin{pmatrix}-2 & 4 & 1\\ 5 & 0 & 2\end{pmatrix}=\begin{pmatrix}29 & 12 & 17\end{pmatrix},$$

$$B\cdot C=\begin{pmatrix}3 & -1\\ 0 & 4\end{pmatrix}\cdot\begin{pmatrix}-2 & 4 & 1\\ 5 & 0 & 2\end{pmatrix}=\begin{pmatrix}-11 & 12 & 1\\ 20 & 0 & 8\end{pmatrix},$$

$$A(B\cdot C)=\begin{pmatrix}1 & 2\end{pmatrix}\cdot\begin{pmatrix}-11 & 12 & 1\\ 20 & 0 & 8\end{pmatrix}=\begin{pmatrix}29 & 12 & 17\end{pmatrix}=(A\cdot B)C\,.$$

2) Let $A=\|a_{i,j}\|$ and $B=\|b_{i,j}\|$ be two matrices of the second order.

Verify the identity $(AB)^T=B^TA^T$.

Solution: Find the matrix product of A and B and the transpose of AB:

$$A\cdot B=\begin{pmatrix}a_{1,1} & a_{1,2}\\ a_{2,1} & a_{2,2}\end{pmatrix}\cdot\begin{pmatrix}b_{1,1} & b_{1,2}\\ b_{2,1} & b_{2,2}\end{pmatrix}=\begin{pmatrix}a_{1,1}b_{1,1}+a_{1,2}b_{2,1} & a_{1,1}b_{1,2}+a_{1,2}b_{2,2}\\ a_{2,1}b_{1,1}+a_{2,2}b_{2,1} & a_{2,1}b_{1,2}+a_{2,2}b_{2,2}\end{pmatrix},$$

$$(A\cdot B)^T=\begin{pmatrix}a_{1,1}b_{1,1}+a_{1,2}b_{2,1} & a_{2,1}b_{1,1}+a_{2,2}b_{2,1}\\ a_{1,1}b_{1,2}+a_{1,2}b_{2,2} & a_{2,1}b_{1,2}+a_{2,2}b_{2,2}\end{pmatrix}.$$

Then find the matrix product B^TA^T to see that $(AB)^T=B^TA^T$:

$$B^TA^T=\begin{pmatrix}b_{1,1} & b_{2,1}\\ b_{1,2} & b_{2,2}\end{pmatrix}\cdot\begin{pmatrix}a_{1,1} & a_{2,1}\\ a_{1,2} & a_{2,2}\end{pmatrix}=\begin{pmatrix}b_{1,1}a_{1,1}+b_{2,1}a_{1,2} & b_{1,1}a_{2,1}+b_{2,1}a_{2,2}\\ b_{1,2}a_{1,1}+b_{2,2}a_{1,2} & b_{1,2}a_{2,1}+b_{2,2}a_{2,2}\end{pmatrix}.$$

3) Let $f(x)=2x-x^2+1$ and $A=\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix}$. Find $f(A)$.

Solution: The matrix-analogue of the number 1 is the identity matrix I. Therefore,

$$f(A)=2A-A^2+I=2\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix}-\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix}\cdot\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix}+\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

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

2. DETERMINANTS

2.1. Permutations and Transpositions

A **permutation** of elements of a set of ordered elements is any one-to-one transformation of the set onto itself.

Let S be the ordered set of the natural numbers from 1 to n:

$$S = \{1, 2, 3, \ldots, n\}.$$

A permutation of S is the set of the same numbers arranged in a particular way:

$$\{1, 2, 3, \ldots, n\} \quad \Rightarrow \quad \{i_1, i_2, i_3, \ldots, i_n\}.$$

A permutation is called a **transposition**, if the order of two elements of the set is changed but all other elements remain fixed.

Example of a permutation:

$$\{1, 2, 3, 4\} \quad \Rightarrow \quad \{2, 4, 1, 3\}$$

Example of a transposition:

$$\{1, 2, 3, 4\} \quad \Rightarrow \quad \{\mathbf{4}, 2, 3, \mathbf{1}\}$$

Every permutation of ordered elements can be expressed through a sequence of several transpositions. For instance, permutation $\{2, 4, 1, 3\}$ can be presented by the sequence of the following transpositions:

$$\{\mathbf{1}, 2, 3, \mathbf{4}\} \quad \Rightarrow \quad \{\mathbf{3}, \mathbf{2}, 1, 4\} \quad \Rightarrow \quad \{2, \mathbf{3}, 1, \mathbf{4}\} \quad \Rightarrow \quad \{2, 4, 1, 3\}$$

It is said that a permutation of S contains the **inversion** of elements i_j and i_k, if

$$j < k \quad \text{and} \quad i_j > i_k.$$

The total number of inversions determines the **inversion parity** of the permutation that takes on two values: either even or odd.

A permutation is called an **even** permutation, if it contains an even number of inversions. This means that an even permutation is formed by an even number of transpositions of S.

An **odd permutation** contains an odd number of inversions.

This means that an odd permutation is a sequence of an odd number of transpositions of S.

Example: The permutation {2, 4, 1, 3} of $\{1, 2, 3, 4\}$ is odd, since it contains three inversions:

2 and 1, 4 and 1, 4 and 3.

Theorem 1

Any transposition changes the inversion parity of a permutation.

Proof: It is not difficult to see that the transposition of neighboring elements i_j and i_{j+1} changes the inversion parity of a given permutation.

The transposition of elements i_j and i_{j+k} can be expressed as the sequence of $(2k-1)$ transpositions. Really, by k transpositions of the element i_j with the neighboring element on the right of i_j we get the permutation $\{\cdots, i_{j+k}, i_j, \cdots\}$:

$$\overset{k \text{ transpositions}}{\{\cdots, i_j, \underbrace{\cdots, i_{j+k},}_{k \text{ elements}} \cdots\}}$$

Then, by k - 1 transpositions of the element i_{j+k} with the neighboring element on the left of i_{j+k}, we get the desired permutation $\{\cdots, i_j, \cdots, i_{j+k}, \cdots\}$:

$$\overset{k\text{-1 transpositions}}{\{\cdots, \cancel{i_j}, \underbrace{\cdots,}_{k\text{-1 elements}} i_{j+k}, i_j, \cdots\}}$$

The total number $k+(k-1)=2k-1$ of the transpositions is an odd number, and hence the inversion parity of the permutation is changed.

Theorem 2

Given the set $S=\{1, 2, 3, \ldots, n\}$, there are n! different permutations of *S*.

Proof: Consider an arbitrary permutation of *S*.

The first position can be displaced by any of *n* elements.

The second position can be displaced by any of the remaining $n-1$ elements.

The third position can be displaced by any of the remaining $n-2$ elements, and so on.

The *n*-th position can be displaced by the rest single element.

Therefore, there are $n(n-1)(n-2)\ldots1 = n!$ ways to get a new permutation of the elements of *S*.

Example:

The set $S = \{1, 2, 3\}$ consists of three elements, and so the number of different permutations is $3! = 6$:

$$\{1, 2, 3\}, \quad \{2, 3, 1\}, \quad \{3, 1, 2\},$$
$$\{3, 2, 1\}, \quad \{2, 1, 3\}, \quad \{1, 3, 2\}.$$

a) The permutations

$$\{1, 2, 3\}, \ \{2, 3, 1\} \text{ and } \{3, 1, 2\}$$

are even, since each of them is a sequence of an even number of transpositions of the elements of S:

$$\{\mathbf{1}, 2, \mathbf{3}\} \quad \to \quad \{\mathbf{3}, \mathbf{2}, 1\} \quad \to \quad \{2, 3, 1\},$$
$$\{\mathbf{1}, \mathbf{2}, 3\} \quad \to \quad \{\mathbf{2}, 1, \mathbf{3}\} \quad \to \quad \{3, 1, 2\}.$$

In terms of inversions, the permutations $\{1, 2, 3\}$, $\{2, 3, 1\}$ and $\{3, 1, 2\}$ are even, since each of them contains an even number of inversions of the elements. For instance, the permutation $\{2, 3, 1\}$ contains two inversions of the elements:

2 and 1, since 2 is on the left from 1, and $2 > 1$,
3 and 1, since 3 is on the left from 1, and $3 > 1$.

b) Likewise, the permutations

$$\{3, 2, 1\}, \quad \{2, 1, 3\} \quad \text{and} \quad \{1, 3, 2\}$$

are odd, since each of them is a sequence of an odd number of transpositions of the elements of S. In particular, the permutation $\{3, 2, 1\}$ is the transposition of the elements 1 and 3 of S.

In terms of inversions, the permutation $\{3, 2, 1\}$ is odd, since it contains the odd number of the inversions:

3 and 2, since 3 is on the left from 2 and $3 > 2$,
3 and 1, since 3 is on the left from 1 and $3 > 1$,
2 and 1, since 2 is on the left from 1 and $2 > 1$.

The permutation $\{2, 1, 3\}$ contains the inversion of the elements 2 and 1.

The permutation $\{1, 3, 2\}$ contains the inversion of the elements 3 and 2.

2.2. General Definition

Let $A = \| a_{i,j} \|$ be a square matrix of the order *n*, and let $S = \{1, 2, \cdots, n\}$ be the ordered set of the first *n* natural numbers.
Consider the following product of *n* matrix elements:

$$a_{1,k_1} a_{2,k_2} \ldots a_{n,k_n} (-1)^{P\{k_1, k_2, \ldots, k_n\}}, \qquad (1)$$

where $\{k_1, k_2, \cdots, k_n\}$ is a permutation of *S*, and $P\{k_1, k_2, \cdots, k_n\}$ is the inversion parity of the permutation $\{k_1, k_2, \cdots, k_n\}$. That is, $(-1)^P = 1$ for an even permutation, and $(-1)^P = -1$ for an odd one:

$$(-1)^{P\{k_1, k_2, \ldots, k_n\}} = \operatorname{sign}\{k_1, k_2, \ldots, k_n\}.$$

Expression (1) is the product of matrix elements such that each row and each column of *A* is presented by one and only one its element. According to Theorem 2, there are *n*! different permutations of *S*, each of which generates the product of type (1).
The sum of products (1) over all possible permutations $\{k_1, k_2, \cdots, k_n\}$ is called the **determinant** of the matrix *A*:

$$\det A = \sum_{\{k_1, k_2, \ldots, k_n\}} a_{1,k_1} a_{2,k_2} \ldots a_{n,k_n} (-1)^{P\{k_1, k_2, \ldots, k_n\}}. \qquad (2)$$

It is denoted by the array between vertical bars:

$$\det A = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \cdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}. \qquad (3)$$

Sum (2) contains *n*! terms (1) with even and odd permutations, fifty-fifty.

The determinant is very important characteristic of the matrix. As a rule, it is important only whether the determinant of a given matrix equals zero or not. For instance, the inverse matrix of *A* exists if and only if $\det A \neq 0$.

Do not confuse the determinant of a matrix with the matrix itself!

While a numerical matrix *A* is an array of numbers, $\det A$ is some single number but not an array of numbers.

Particular cases

1. A matrix of the first order contains only one element. The determinant of that matrix is equal to the matrix element itself: $\det\|a_{1,1}\| = a_{1,1}$.

2. Let A be a square matrix of the second order: $A = \begin{pmatrix} a_{1,1} & a_{1,,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$.

There exist the following two permutations of $\{1, 2\}$: $\{1, 2\}$ and $\{2, 1\}$.
The permutation $\{1, 2\}$ is even, since it does not contain any inversions, while the permutation $\{2, 1\}$ is odd, since two elements form the inversion. These permutations generate two products of the elements with opposite signs,

$$+a_{1,1}a_{2,2} \quad \text{and} \quad -a_{1,2}a_{2,1},$$

the sum of which gives the determinant of A:

$$\boxed{\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}$$

3. If a matrix has the third order then we have to consider all possible permutations of the set $\{1, 2, 3\}$. There exist the following six permutations of $\{1, 2, 3\}$:

$$\{1, 2, 3\}, \qquad \{2, 3, 1\}, \qquad \{3, 1, 2\},$$

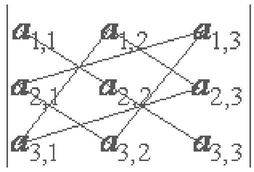
$$\{3, 2, 1\}, \qquad \{2, 1, 3\}, \qquad \{1, 3, 2\}.$$

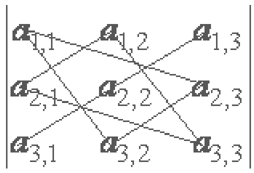
The permutations $\{1, 2, 3\}$, $\{2, 3, 1\}$, and $\{3, 1, 2\}$ are even since each of them contains an even number of inversions of elements.
The permutations $\{3, 2, 1\}$, $\{2, 1, 3\}$, and $\{1, 3, 2\}$ are odd since there are odd numbers of inversions of elements in these permutations. (See details in the above example.)
Therefore,

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2}$$

$$- a_{1,3}a_{2,2}a_{3,1} - a_{1,2}a_{2,1}a_{3,3} - a_{1,1}a_{2,3}a_{3,2}$$

To remember this formula, apply the Sarrus Rule which is shown in the figure below.





$$+a_{1,1}a_{2,2}a_{3,3} \qquad -a_{1,3}a_{2,2}a_{3,1}$$
$$+a_{1,2}a_{2,3}a_{3,1} \qquad -a_{1,2}a_{2,1}a_{3,3}$$
$$+a_{1,3}a_{2,1}a_{3,2} \qquad -a_{1,1}a_{2,3}a_{3,2}$$

The elements on a diagonal or at the vertices of a triangular form the product of three elements. If the base of the triangle is parallel to the leading diagonal of the matrix, the product keeps the sign; otherwise, the product changes the sign.

2.3. Properties of Determinants

1. The determinant of the transpose of A is equal to the determinant of the given matrix A:

$$\det A^T = \det A\,.$$

Proof: This property results from the determinant definition since both determinants consist of the same terms.

2. Multiplying any row or column of a determinant by a number λ, multiplies the determinant by that number:

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda\cdot a_{i1} & \lambda\cdot a_{i2} & \cdots & \lambda\cdot a_{ij} & \cdots & \lambda\cdot a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} = \lambda\cdot \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

This means that the common factor of a row (column) can be taken out.

Proof: Every term of the sum

$$\det A = \sum_{\{i_1,i_2,\ldots,i_n\}} a_{1,i_1}a_{2,i_2}\ldots a_{n,i_n}(-1)^{P\{i_1,i_2,\ldots,i_n\}}$$

contains one and only one element of a row and a column of the matrix. Therefore, if the row (or column) is multiplied by a number, each term is multiplied by that common factor.

3. The determinant changes the sign, if two rows (or columns) of a matrix are interchanged:

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} = - \begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}$$

Proof: By Theorem 1, any transposition changes the inversion parity of a given permutation. Therefore, each term of the sum

$$\det A = \sum_{\{i_1, i_2, \ldots, i_n\}} a_{1,i_1} a_{2,i_2} \ldots a_{n,i_n} (-1)^{P\{i_1, i_2, \ldots, i_n\}}$$

changes its sign.

4. If a matrix has a zero-row or zero-column, then the determinant is equal to zero:

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ 0 & 0 & 0 & \cdots & 0 \\ a_{j,1} & a_{j,2} & a_{j,3} & \cdots & a_{j,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0$$

Proof: Every product of the sum

$$\sum_{\{i_1, i_2, \ldots, i_n\}} a_{1,i_1} a_{2,i_2} \ldots a_{n,i_n} (-1)^{P\{i_1, i_2, \ldots, i_n\}} = \det A$$

contains a zero factor and so equals zero.

5. If a matrix has two equal rows (or columns) then the determinant is equal to zero:

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0$$

Proof: Let two identical rows (or columns) be interchanged. Then, by Property 3, the determinant changes the sign. On the other hand, the rows (or columns) are equal, and hence the determinant keeps its value:

$$\det A = -\det A \quad \Rightarrow \quad \det A = 0.$$

6. If two rows (or columns) of a matrix are proportional to each other then the determinant is equal to zero:

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ ca_{i,1} & ca_{i,2} & ca_{i,3} & \cdots & ca_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0$$

Proof: Multiplying the *i*-th row of the matrix by the constant of proportionality we obtain the determinant with equal rows.

7. If each element of a row (column) of a determinant is the sum of two entries then

$$\begin{vmatrix} \boldsymbol{a}_{1,1} & \boldsymbol{a}_{1,2} & \boldsymbol{a}_{1,3} & \cdots & \boldsymbol{a}_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{a}_{k,1}+\boldsymbol{b}_{k,1} & \boldsymbol{a}_{k,2}+\boldsymbol{b}_{k,2} & \boldsymbol{a}_{k,3}+\boldsymbol{b}_{k,3} & \cdots & \boldsymbol{a}_{k,n}+\boldsymbol{b}_{k,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{a}_{n,1} & \boldsymbol{a}_{n,2} & \boldsymbol{a}_{n,3} & \cdots & \boldsymbol{a}_{n,n} \end{vmatrix} =$$

$$\begin{vmatrix} \boldsymbol{a}_{1,1} & \boldsymbol{a}_{1,2} & \boldsymbol{a}_{1,3} & \cdots & \boldsymbol{a}_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{a}_{k,1} & \boldsymbol{a}_{k,2} & \boldsymbol{a}_{k,3} & \cdots & \boldsymbol{a}_{k,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{a}_{n,1} & \boldsymbol{a}_{n,2} & \boldsymbol{a}_{n,3} & \cdots & \boldsymbol{a}_{n,n} \end{vmatrix} + \begin{vmatrix} \boldsymbol{a}_{1,1} & \boldsymbol{a}_{1,2} & \boldsymbol{a}_{1,3} & \cdots & \boldsymbol{a}_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{b}_{k,1} & \boldsymbol{b}_{k,2} & \boldsymbol{b}_{k,3} & \cdots & \boldsymbol{b}_{k,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \boldsymbol{a}_{n,1} & \boldsymbol{a}_{n,2} & \boldsymbol{a}_{n,3} & \cdots & \boldsymbol{a}_{n,n} \end{vmatrix}$$

Proof:

$$\sum_{\{i_1,i_2,\ldots,i_n\}} a_{1,i_1} \ldots (a_{k,i_k} + b_{k,i_k}) \ldots a_{n,i_n} (-1)^{P\{i_1,i_2,\ldots,i_n\}} =$$

$$= \sum_{\{i_1,i_2,\ldots,i_n\}} \ldots a_{k,i_k} \ldots \quad + \sum_{\{i_1,i_2,\ldots,i_n\}} \ldots b_{k,i_k} \ldots$$

8. A determinant holds its value, if a row (column) is multiplied by a number and then is added to another one:

$$\begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kn} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix} = \begin{vmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{k1}+ca_{i1} & a_{k2}+ca_{i2} & \cdots & \cdots & a_{kn}+ca_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}$$

Proof: The determinant on the right hand can be expressed as the sum of two determinants, one of which contains two proportional rows. Therefore, the determinant equals zero.

9. Let A and B be square matrices of the same order. Then the determinant of their product is equal to the product of the determinants:

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

10. The determinant of a triangular matrix is equal to the product of the elements on the principle diagonal:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{n,n} \end{vmatrix} = a_{1,1} \cdot a_{2,2} \cdot a_{3,3} \cdot \ldots \cdot a_{n,n} .$$

In particular, the determinant of an identity matrix I equals the unity.

Proof: First, there is only $a_{1,1}$ which is a non-zero element in the first column Therefore, sum (2) consists of zero terms for all values of i_1 except for $i_1 = 1$.

Next, we have to ignore the first row and choose a non-zero element on the second column. Only the element $a_{2,2}$ satisfies these conditions, and so we can set $i_2 = 2$ in sum (2).

Likewise, on the third column we can take only the element $a_{3,3}$ to get a non-zero product of elements and so on.

Therefore, all appropriate permutations of indexes give zero products of elements, except for the product of the elements on the principle diagonal.

Examples:

1) Let $A=\begin{pmatrix} \sin x & \cos x \\ -\cos x & \sin x \end{pmatrix}$. Find $\det A$.

Solution:

$$\det A=\begin{vmatrix} \sin x & \cos x \\ -\cos x & \sin x \end{vmatrix}=\sin^2 x+\cos^2 x=1.$$

2) Let $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Verify that $\det A=\det A^T$.

Solution:

$$\det A=\begin{vmatrix} a & b \\ c & d \end{vmatrix}=ad-bc \quad \text{and} \quad \det A^T=\begin{vmatrix} a & c \\ b & d \end{vmatrix}=ad-bc.$$

3) Evaluate $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$.

Solution:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} \overset{\substack{r_2\to r_2-2r_1 \\ r_3\to r_3-3r_1}}{=} \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 4 & 2 & 0 \end{vmatrix}=2\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 2 & 1 & 0 \end{vmatrix}=0.$$

4) Let $A=\begin{pmatrix} 5 & 0 \\ 1 & 2 \end{pmatrix}$ and $B=\begin{pmatrix} 7 & 1 \\ 3 & 2 \end{pmatrix}$. Verify that $\det AB=\det A\cdot\det B$.

Solution:

$$\det A=\begin{vmatrix} 5 & 0 \\ 1 & 2 \end{vmatrix}=10, \quad \det B=\begin{vmatrix} 7 & 1 \\ 3 & 2 \end{vmatrix}=11,$$

$$\det A\cdot\det B=110.$$

$$AB=\begin{pmatrix} 5 & 0 \\ 1 & 2 \end{pmatrix}\cdot\begin{pmatrix} 7 & 1 \\ 3 & 2 \end{pmatrix}=\begin{pmatrix} 35 & 5 \\ 13 & 5 \end{pmatrix},$$

$$\det AB=\begin{vmatrix} 35 & 5 \\ 13 & 5 \end{vmatrix}=5\cdot\begin{vmatrix} 35 & 1 \\ 13 & 1 \end{vmatrix}=5(35-13)=110.$$

5) Evaluate $\det A^{1000}$, if $A = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$.

Solution: Note that

$$\det A^{1000} = (\det A)^{1000}.$$

Then

$$\det A = \begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} = 3 - 2 = 1 \qquad \Rightarrow \qquad \det A^{1000} = 1^{1000} = 1.$$

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

Calculate: (a) $\det A$, (b) $\det A^3$, (c) $\det(2A)$, (d) $\det(-3A)$, (e) $\det(A - 2I)$.

Solution: (a) The determinant of a matrix in the triangular form equals the product of the principle diagonal elements. Therefore,

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

(b) The determinant of the product of matrices is equal to the product of their determinants, and so

$$\det A^3 = (\det A)^3 = (-2)^3 = -8.$$

(c) Let I be the identity matrix of the third order. Then

$$\det(2A) = \det(2I) \cdot \det A = 2^3(-2) = -16.$$

(d) Likewise,

$$\det(-3A) = \det(-3I) \cdot \det A = (-3)^3(-2) = 54.$$

(e) Simplify the matrix $(A - 2I)$:

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

The determinant of this matrix equals zero: $\det(A - 2I) = 0$.

2.4. Calculation of Determinants

Methods of determinant calculation are based on the properties of determinants. Here we consider two methods which being combined together result in the most efficient computing technique.

2.4.1. Expanding a determinant by a row or column

Before formulating the theorem, let us introduce a few definitions.
Let *A* be a square matrix of the order *n*. By removing the *i*-th row and *j*-th column, we obtain a submatrix of *A*, having the order $(n-1)$. The determinant of that submatrix is called the minor of the element $a_{i,j}$, which is denoted by $M_{i,j}$.

$$M_{i,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix}$$

The **cofactor** of the element $a_{i,j}$ is defined as the minor $M_{i,j}$ with the sign $(-1)^{i+j}$. It is denoted by the symbol $A_{i,j}$:

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

The following **theorem** gives a systematic procedure of determinant calculation.

> The determinant of a matrix *A* equals the sum of the products of elements of any row of *A* and the corresponding cofactors:
>
> $$\det A = a_{i,1} A_{i,1} + a_{i,2} A_{i,2} + \ldots + a_{i,n} A_{i,n} = \sum_{j=1}^{n} a_{i,j} A_{i,j}$$

The above theorem is known as the expansion of the determinant according to its *i*-th row.

Proof: By the definition, $\det A$ is the algebraic sum of the products $a_{1,k_1} a_{2,k_2} \ldots a_{n,k_n}$ taken with the signs

$$\operatorname{sign}\{k_1,k_2,\ldots,k_n\} \equiv (-1)^{P\{k_1,k_2,\ldots,k_n\}}$$

over all possible permutations $\{k_1,k_2,\ldots,k_n\}$, that is,

$$\det A = \sum_{\{k_1,k_2,\ldots,k_n\}} a_{1,k_1} a_{2,k_2} \ldots a_{n,k_n} (-1)^{P\{k_1,k_2,\ldots,k_n\}} .$$

Each product $a_{1,k_1} a_{2,k_2} \ldots a_{n,k_n}$ contains the element $a_{i,j}$ on the *i*-th row and *j*-th column. Therefore, by regrouping the terms, the above sum can be expressed as the linear combination of the elements $a_{i,j}$ ($j = 1,2,\cdots,n$):

$$\det A = a_{i,1}\, A_{i,1} + a_{i,2}\, A_{i,2} + \ldots + a_{i,n}\, A_{i,n} .$$

Here

$$A_{i,j} = \sum_{\{k_1,\cdots,k_{i-1},k_i=j,k_{i+1},\cdots,k_n\}} a_{1,k_1} a_{2,k_2} \cdots a_{i-1,k_{i-1}} a_{i+1,k_{i+1}} \ldots a_{n,k_n}$$

$$\cdot \operatorname{sign}\{k_1,\cdots,k_{i-1},j,k_{i+1},\cdots,k_n\}.$$

By the theorem of inversion parity of a permutation,

$$\operatorname{sign}\{k_1,\cdots,k_{i-1},j,k_{i+1},\cdots,k_n\} = (-1)^{i-1} \operatorname{sign}\{j,k_1,\cdots,k_{i-1},k_{i+1},\cdots,k_n\} .$$

There are $(j-1)$ inversions of *j* in the permutation $\{j,k_1,\cdots,k_n\}$, and so

$$\operatorname{sign}\{j,k_1,\cdots,k_{i-1},k_{i+1},\cdots,k_n\} = (-1)^{j-1} \operatorname{sign}\{k_1,\cdots,k_{i-1},k_{i+1},\cdots,k_n\},$$

$$\operatorname{sign}\{k_1,\cdots,k_{i-1},j,k_{i+1},\cdots,k_n\} = (-1)^{i+j} \operatorname{sign}\{k_1,\cdots,k_{i-1},k_{i+1},\cdots,k_n\} .$$

However, $\displaystyle\sum_{\{k_1,\cdots,k_{i-1},k_{i+1},\cdots,k_n\}} \cdots a_{i-1,k_{i-1}} a_{i+1,k_{i+1}} \ldots \ \operatorname{sign}\{\cdots,k_{i-1},k_{i+1},\cdots\} = M_{i,j}$

is the minor of the element $a_{i,j}$.

Therefore, $A_{i,j} = (-1)^{i+j} M_{i,j}$ is the cofactor of the element $a_{i,j}$.

Since both matrices, *A* and the transpose of *A*, have equal determinants, the theorem can be formulated in terms of expanding a determinant by a column:

> The determinant of a matrix *A* equals the sum of the products of elements of any column of *A* and the corresponding cofactors:
>
> $$\det A = a_{1,j}\, A_{1,j} + a_{2,j}\, A_{2,j} + \ldots + a_{n,j}\, A_{n,j}$$
>
> $$= \sum_{i=1}^{n} a_{i,j}\, A_{i,j}$$

Due to the theorem, a given determinant of the order *n* is reduced to *n* determinants of the order (*n* –1).

Examples:

1) Expand the determinant of the matrix $A = \| a_{ij} \|$ of the order 3 by (i) the first row; (ii) the second column. Compare the results.

Solution:

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

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

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

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

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

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

Both results are identically equal.

2) Calculate the determinant $\begin{vmatrix} 2 & -5 & 3 \\ 1 & 4 & 0 \\ -3 & 7 & 5 \end{vmatrix}$, by its expansion according to the first row and the second column.

Solution: The expansion by the first row yields

$$\begin{vmatrix} 2 & -5 & 3 \\ 1 & 4 & 0 \\ -3 & 7 & 5 \end{vmatrix} = 2 \begin{vmatrix} 4 & 0 \\ 7 & 5 \end{vmatrix} - (-5) \begin{vmatrix} 1 & 0 \\ -3 & 5 \end{vmatrix} + 3 \begin{vmatrix} 1 & 4 \\ -3 & 7 \end{vmatrix}$$

$$= 2 \cdot 4 \cdot 5 + 5 \cdot 1 \cdot 5 + 3(7 + 12) = 122,$$

Now expand the determinant according to the second column:

$$\begin{vmatrix} 2 & -5 & 3 \\ 1 & 4 & 0 \\ -3 & 7 & 5 \end{vmatrix} = -(-5) \begin{vmatrix} 1 & 0 \\ -3 & 5 \end{vmatrix} + 4 \begin{vmatrix} 2 & 3 \\ -3 & 5 \end{vmatrix} - 7 \begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix}$$

$$= 5(5 + 0) + 4(10 + 9) - 7(0 - 3) = 122.$$

2. Evaluation of determinants by elementary operations on matrices

By means of elementary row and column operations, a matrix can be reduced to the triangular form, the determinant of which is equal to the product of the diagonal elements.

Let us define the **elementary operations**.

In view of the properties of determinants, any techniques which are developed for rows may be also applied to columns.

In order to calculate a determinant one may:

1. **Interchange two rows**.
   As a result, the determinant changes its sign.
2. **Multiply a row by a nonzero number**.
   As a consequence of this operation, the determinant is multiplied by that number.
3. **Add a row multiplied by a number to another row**.
   By this operation, the determinant holds its value.

We can also use the elementary operations to get some row or column consisting of zero elements except for one element, and then to expand the determinant by that row (or column).

Examples:

1) Let $A = \begin{pmatrix} 2 & -4 & 1 \\ -3 & 2 & 5 \\ 1 & 2 & 3 \end{pmatrix}$.

By elementary row and column operations on the matrix, reduce the matrix to the triangular form and calculate $\det A$.

Solution:

$$\det A = \begin{vmatrix} 2 & -4 & 1 \\ -3 & 2 & 5 \\ 1 & 2 & 3 \end{vmatrix} \overset{\substack{r_1 \to r_1 - 2r_3 \\ r_2 \to r_2 + 3r_3}}{=} \begin{vmatrix} 0 & -8 & -5 \\ 0 & 8 & 14 \\ 1 & 2 & 3 \end{vmatrix}$$

$$\overset{r_1 \leftrightarrow r_3}{=} -\begin{vmatrix} 1 & 2 & 3 \\ 0 & 8 & 14 \\ 0 & -8 & -5 \end{vmatrix} \overset{r_3 \to r_3 + r_2}{=} -\begin{vmatrix} 1 & 2 & 3 \\ 0 & 8 & 14 \\ 0 & 0 & 9 \end{vmatrix}.$$

The determinant of the matrix in the triangular form is equal to the product of the elements on the principle diagonal. Therefore,

$$\det A = -1 \cdot 8 \cdot 9 = -72.$$

2) Evaluate the determinant of the matrix

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

Solution: First, transform the first row via elementary column operations.
Keeping the first and last columns, subtract the first column multiplied by 5 from the second one, and add the first column multiplied by 2 to the third one:

$$\det A=\begin{vmatrix}1 & 5 & -2 & 0\\ 3 & 1 & 6 & -1\\ 7 & 0 & 1 & 3\\ 4 & 5 & 2 & 1\end{vmatrix}\overset{\substack{c_2\to c_2-5c_1\\ c_3\to c_3+2c_1}}{=}\begin{vmatrix}1 & 0 & 0 & 0\\ 3 & -14 & 12 & -1\\ 7 & -35 & 15 & 3\\ 4 & -15 & 10 & 1\end{vmatrix}.$$

Then expand the determinant by the first row:

$$\det A=\begin{vmatrix}-14 & 12 & -1\\ -35 & 15 & 3\\ -15 & 10 & 1\end{vmatrix}$$

Transform the third column by adding the third row to the first one and subtracting the third row multiplied by 3 from the second row:

$$\begin{vmatrix}-14 & 12 & -1\\ -35 & 15 & 3\\ -15 & 10 & 1\end{vmatrix}\overset{\substack{r_1\to r_1+r_3\\ r_2\to r_2-3r_3}}{=}\begin{vmatrix}-29 & 22 & 0\\ 10 & -15 & 0\\ -15 & 10 & 1\end{vmatrix}$$

Expand the determinant by the third column:

$$\det A=\begin{vmatrix}-29 & 22\\ 10 & -15\end{vmatrix}$$

We can still take out the common factor 5 from the last row:

$$\det A=5\begin{vmatrix}-29 & 22\\ 2 & -3\end{vmatrix}=5((-29)\cdot(-3)-22\cdot 2)=105.$$

3. Inverse Matrices

Let A be a square matrix.

A matrix A^{-1} is called an **inverse matrix** of A if

$$A^{-1}A = AA^{-1} = I\,,$$

where I is an identity matrix.

If the determinant of a matrix is equal to zero, then the matrix is called **singular**; otherwise, if $\det A \neq 0$, the matrix A is called **regular**.

If each element of a square matrix A is replaced by its cofactor, then the transpose of the matrix obtained is called the **adjoint matrix** of A:

$$\operatorname{adj} A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n} \end{pmatrix}^T = \begin{pmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1,n} & A_{2,n} & \cdots & A_{n,n} \end{pmatrix}.$$

3.1. Three Lemmas

Lemma 1: Given a square matrix A of the order n, the sum of the products of the elements of any row (or column) and the cofactors of another row (column) is equal to zero:

$$\sum_{k=1}^{n} a_{i,k} A_{j,k} = 0, \qquad (i \neq j) \tag{1}$$

and

$$\sum_{k=1}^{n} a_{k,i} A_{k,j} = 0, \qquad (i \neq j)\,. \tag{2}$$

Proof: To prove (1), consider an auxiliary matrix $\tilde{A}$ that is obtained from the matrix A by replacing the j-th row with the i-th one:

$$A = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j,1} & a_{j,2} & a_{j,3} & \cdots & a_{j,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \quad \Rightarrow \quad \tilde{A} = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots & a_{i,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Expand $\det \tilde{A}$ by the j-th row:

$$\det \tilde{A} = \sum_{k=1}^{n} a_{j,k} \tilde{A}_{j,k} = \sum_{k=1}^{n} a_{i,k} \tilde{A}_{j,k}\,.$$

The only difference between matrices $\tilde{A}$ and A is the j-th row. However, the cofactors $\tilde{A}_{j,k}$ do not depend on the elements on the j-th row, and so $\tilde{A}_{j,k} = A_{j,k}$, which implies

$$\det \tilde{A} = \sum_{k=1}^{n} a_{i,k} A_{j,k} .$$

On the other hand, the matrix $\tilde{A}$ has two equal rows. Therefore, by the properties of determinants,

$$\det \tilde{A} = \sum_{k=1}^{n} a_{i,k} A_{j,k} = 0, \qquad (i \neq j) .$$

Statement (2) can be proven in a similar way.

Lemma 2: The matrix products $A \cdot \operatorname{adj} A$ and $\operatorname{adj} A \cdot A$ are diagonal matrices, that is,

$$(A \cdot \operatorname{adj} A)_{i,j} = 0 \qquad (i \neq j) ,$$
$$(\text{ad}j\, A \cdot A)_{j,i} = 0 \qquad (i \neq j) .$$

Proof: If $i \neq j$ then, by Lemma 1,

$$\sum_{k=1}^{n} a_{i,k} A_{j,k} = 0 \quad \Rightarrow \quad \sum_{k=1}^{n} a_{i,k} A^{T}_{k,j} = 0 \quad \Rightarrow \quad (A \cdot \operatorname{adj} A)_{i,j} = 0 ,$$

and

$$\sum_{k=1}^{n} a_{k,i} A_{k,j} = 0 \quad \Rightarrow \quad \sum_{k=1}^{n} A^{T}_{j,k}\, a_{k,i} = 0 \quad \Rightarrow \quad (\operatorname{adj} A \cdot A)_{j,i} = 0 .$$

Lemma 3: The diagonal elements of the matrices $A \cdot \operatorname{adj} A$ and $\operatorname{adj} A \cdot A$ are equal to the determinant of the matrix A:

$$(A \cdot \operatorname{adj} A)_{i,i} = (\operatorname{adj} A \cdot A)_{i,i} = \det A .$$

Proof: By the theorem of expansion of determinants according to a row,

$$\det A = \sum_{k=1}^{n} a_{i,k} A_{i,k} = \sum_{k=1}^{n} a_{i,k} A^{T}_{k,i} = (A \cdot \operatorname{adj} A)_{i,i} .$$

Likewise, the theorem of expansion determinants by a column yields

$$\det A = \sum_{k=1}^{n} a_{k,i} A_{k,i} = \sum_{k=1}^{n} A^{T}_{i,k}\, a_{k,i} = (\operatorname{adj} A \cdot A)_{i,i} .$$

Hence, the lemma.

3.2. Theorem of Inverse Matrix

> For any regular matrix A there exists the unique inverse matrix:
> $$A^{-1} = \frac{1}{\det A}\operatorname{adj} A .$$
> Any singular matrix has no an inverse matrix.

Proof:

1. Assume that there exists an inverse of matrix A. Then
$$A A^{-1} = I \qquad \Rightarrow \qquad \det A \cdot \det A^{-1} = 1,$$
and hence $\det A \neq 0$.
Therefore, singular matrices have no inverse matrices.
2. Assume that each of the matrices, A^{-1} and B^{-1}, is an inverse of A:
$$A A^{-1} = A^{-1}A = I \quad \text{and} \quad AB^{-1} = B^{-1}A = I .$$
Then
$$B^{-1} = B^{-1}I = B^{-1}AA^{-1} = (B^{-1}A)A^{-1} = IA^{-1} = A^{-1}.$$
Therefore, there exists the unique inverse of A.
3. Find the inverse of matrix A.
By the Lemma 2,
$$(A \cdot \operatorname{adj} A)_{i,j} = 0, \quad \text{if} \quad i \neq j .$$
By the Lemma 3,
$$\frac{1}{\det A}(A \cdot \operatorname{adj} A)_{i,i} = 1.$$
Combining the above equalities, we obtain
$$\frac{1}{\det A}(A \cdot \operatorname{adj} A)_{i,j} = \delta_{i,j} ,$$
where the delta symbol $\delta_{i,j}$ denotes the matrix elements of an identity matrix.
Therefore,
$$A \cdot \frac{1}{\det A}\operatorname{adj} A = I .$$
Likewise, $(\frac{1}{\det A}\operatorname{adj} A) \cdot A = I$, and hence,
$$\frac{1}{\det A}\operatorname{adj} A = A^{-1}.$$

3.2.1. Examples of Calculations of Inverse Matrices

Example 1: Given the matrix $A=\begin{pmatrix}3 & 4\\ 1 & 2\end{pmatrix}$, find the inverse of *A*.

Solution: First, calculate the determinant:

$$\det A=\begin{vmatrix}3 & 4\\ 1 & 2\end{vmatrix}=6-4=2.$$

Next, find the cofactors of all elements:

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

Then, find the adjoint matrix of *A*:

$$\operatorname{adj}A=\begin{pmatrix}A_{1,1} & A_{1,2}\\ A_{2,1} & A_{2,2}\end{pmatrix}^T=\begin{pmatrix}2 & -1\\ -4 & 3\end{pmatrix}^T=\begin{pmatrix}2 & -4\\ -1 & 3\end{pmatrix}.$$

Finally, obtain

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix}2 & -4\\ -1 & 3\end{pmatrix}=\frac{1}{2}\begin{pmatrix}2 & -4\\ -1 & 3\end{pmatrix}=\begin{pmatrix}1 & -2\\ -\frac{1}{2} & \frac{3}{2}\end{pmatrix}.$$

Verification:

$$AA^{-1}=\frac{1}{2}\begin{pmatrix}3 & 4\\ 1 & 2\end{pmatrix}\cdot\begin{pmatrix}2 & -4\\ -1 & 3\end{pmatrix}=\frac{1}{2}\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}=I,$$

and

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

Example 2: Let $A=\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}$. Find the inverse of *A*.

Solution: Calculate the determinant:

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

Therefore, the given matrix is singular, and so it has no the inverse of *A*.

Example 3: Let $A=\begin{pmatrix}1 & -2 & 3\\ 0 & 4 & -1\\ 5 & 0 & 1\end{pmatrix}$. Find the inverse of matrix *A*.

Solution:

1) To calculate the determinant of *A*, add the first row doubled to the second row. Then expand the determinant by the second column.

$$\det A=\begin{vmatrix}1 & -2 & 3\\ 0 & 4 & -1\\ 5 & 0 & 1\end{vmatrix}\overset{r_2\to r_2+2r_1}{=}\begin{vmatrix}1 & -2 & 3\\ 2 & 0 & 5\\ 5 & 0 & 1\end{vmatrix}=(-2)\cdot(-1)^{1+2}\begin{vmatrix}2 & 5\\ 5 & 1\end{vmatrix}=-46.$$

2) Find the cofactors of the elements of the matrix.

$$A_{1,1}=(-1)^{1+1}\begin{vmatrix}4 & -1\\ 0 & 1\end{vmatrix}=4,\qquad A_{1,2}=(-1)^{1+2}\begin{vmatrix}0 & -1\\ 5 & 1\end{vmatrix}=-5,$$

$$A_{1,3}=(-1)^{1+3}\begin{vmatrix}0 & 4\\ 5 & 0\end{vmatrix}=-20,\qquad A_{2,1}=(-1)^{2+1}\begin{vmatrix}-2 & 3\\ 0 & 1\end{vmatrix}=2,$$

$$A_{2,2}=(-1)^{2+2}\begin{vmatrix}1 & 3\\ 5 & 1\end{vmatrix}=-14,\qquad A_{2,3}=(-1)^{2+3}\begin{vmatrix}1 & -2\\ 5 & 0\end{vmatrix}=-10,$$

$$A_{3,1}=(-1)^{3+1}\begin{vmatrix}-2 & 3\\ 4 & -1\end{vmatrix}=-10,\quad A_{3,2}=(-1)^{3+2}\begin{vmatrix}1 & 3\\ 0 & -1\end{vmatrix}=1,$$

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

3) Write down the adjoint matrix of *A*.

$$\operatorname{adj}A=\begin{pmatrix}4 & -5 & -20\\ 2 & -14 & -10\\ -10 & 1 & 4\end{pmatrix}^T=\begin{pmatrix}4 & 2 & -10\\ -5 & -14 & 1\\ -20 & -10 & 4\end{pmatrix}.$$

4) The inverse of matrix *A* is

$$A^{-1}=-\frac{1}{46}\begin{pmatrix}4 & 2 & -10\\ -5 & -14 & 1\\ -20 & -10 & 4\end{pmatrix}.$$

5) **Verification**:

$$AA^{-1} = -\frac{1}{46}\begin{pmatrix} 1 & -2 & 3 \\ 0 & 4 & -1 \\ 5 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & 2 & -10 \\ -5 & -14 & 1 \\ -20 & -10 & 4 \end{pmatrix}$$

$$= -\frac{1}{46}\begin{pmatrix} -46 & 0 & 0 \\ 0 & -46 & 0 \\ 0 & 0 & -46 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

Likewise,

$$A^{-1}A = -\frac{1}{46}\begin{pmatrix} -46 & 0 & 0 \\ 0 & -46 & 0 \\ 0 & 0 & -46 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

Example 4: Let $A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 4 & 1 \\ 3 & 1 \end{pmatrix}$.

Solve for X the matrix equation

$$X\,A = B.$$

Solution: Note that $\det A = \begin{vmatrix} 3 & 5 \\ 1 & 2 \end{vmatrix} = 1 \neq 0$, that is, A is a regular matrix. Therefore, there exists the inverse of A:

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

Find the inverse of matrix A.

$$\operatorname{adj} A = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}^T = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} \qquad \Rightarrow$$

$$A^{-1} = \frac{1}{\det A}\operatorname{adj} A = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}.$$

Thus,

$$X = \begin{pmatrix} 4 & 1 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 7 & -17 \\ 5 & -12 \end{pmatrix}.$$

Verification:

$$X \cdot A = \begin{pmatrix} 7 & -17 \\ 5 & -12 \end{pmatrix} \cdot \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 3 & 1 \end{pmatrix} \equiv B.$$

3.3. Calculation of Inverse Matrices by Elementary Transformations

Let A be a regular matrix.
The inverse of A can be found by means of the elementary transformations of the following extended matrix

$$(A\,|\,I)=\left(\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,n} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,n} & 0 & \cdots & 1 \end{array}\right),$$

where I is the identity matrix of the corresponding order.
By making use of elementary row operations we have to transform the extended matrix to the form (I | B). Then $B=A^{-1}$.
The following elementary transformations (involving only rows) can be applied:

1) Multiplying a row by a nonzero number.
2) Adding a row to another row.

Example: Let $A=\begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$. Find the inverse of A.

Solution: Consider the extended matrix $(A\,|\,I)=\left(\begin{array}{cc|cc} 3 & 4 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right)$.

Multiply the second row by the number 2 and then subtract the result from the first row:

$$\left(\begin{array}{cc|cc} 3 & 4 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right)\to\left(\begin{array}{cc|cc} 1 & 0 & 1 & -2 \\ 1 & 2 & 0 & 1 \end{array}\right).$$

Subtract the first row from the second one:

$$\left(\begin{array}{cc|cc} 1 & 0 & 1 & -2 \\ 1 & 2 & 0 & 1 \end{array}\right)\to\left(\begin{array}{cc|cc} 1 & 0 & 1 & -2 \\ 0 & 2 & -1 & 3 \end{array}\right).$$

Divide the second row by 2:

$$\left(\begin{array}{cc|cc} 1 & 0 & 1 & -2 \\ 0 & 2 & -1 & 3 \end{array}\right)\to\left(\begin{array}{cc|cc} 1 & 0 & 1 & -2 \\ 0 & 1 & -\frac{1}{2} & \frac{3}{2} \end{array}\right).$$

The desired form is obtained and hence,

$$A^{-1}=\begin{pmatrix} 1 & -2 \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix}=\frac{1}{2}\begin{pmatrix} 2 & -4 \\ -1 & 3 \end{pmatrix}.$$

4. Systems of Linear Equations

4.1. Matrix Rank

An $m \times n$ matrix *A* is said to be the matrix of **rank** *r*, if

– there exists at least one regular submatrix of order r;
– every submatrix of a higher order is singular.

According to the definition,

$$\operatorname{rank} A \le \min\{m, n\}.$$

The rank of a matrix can be evaluated by applying just those elementary row and column operations which are used to simplify determinants, that is,

1. Interchanging two rows or columns.
2. Multiplying a row (column) by a nonzero number.
3. Multiplying a row (column) by a number and adding the result to another row (column).

If a row or column consists of zeros then it can be omitted.
These operations are said to be **elementary transformations** of a matrix.

Theorem: If a matrix $\tilde{A}$ is obtained from *A* by elementary transformations then $\operatorname{rank} \tilde{A} = \operatorname{rank} A$.

Proof: Interchanging two rows or two columns of a matrix changes the sign of the determinant.
Multiplying a row (column) by a nonzero number multiplies the determinant by that number.
Adding a row (column) to another one holds the magnitude of the determinant.
Therefore, all singular submatrices are transformed into singular submatrices, and regular submatrices are transformed into regular submatrices. Hence, the theorem.
By elementary transformations of a matrix we try to obtain as many zeros as possible to reduce the matrix to the echelon form:

For instance,

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

is the matrix of the reduced row echelon form.
The number of the rows gives the rank of *A*: rank *A* = 3.

Examples:

1) Let $A = \begin{pmatrix} 4 & -1 & 5 & 2 \\ 1 & 0 & -3 & 3 \\ 7 & -2 & 13 & 1 \\ 3 & -1 & 8 & -2 \end{pmatrix}$.

Find the rank of *A*.

Solution: Subtract the first and fourth rows from the third one:

$$A = \begin{pmatrix} 4 & -1 & 5 & 2 \\ 1 & 0 & -3 & 3 \\ 7 & -2 & 13 & 1 \\ 3 & -1 & 8 & -2 \end{pmatrix} \; r_3 \to r_3 - r_1 - r_4 \; \begin{pmatrix} 4 & -1 & 5 & 2 \\ 1 & 0 & -3 & 3 \\ 0 & 0 & 0 & 1 \\ 3 & -1 & 8 & -2 \end{pmatrix}.$$

Add the third row multiplied by suitable numbers to the other rows:

$$\begin{pmatrix} 4 & -1 & 5 & 2 \\ 1 & 0 & -3 & 3 \\ 0 & 0 & 0 & 1 \\ 3 & -1 & 8 & -2 \end{pmatrix} \begin{matrix} r_1 \to r_1 - 2r_3 \\ r_2 \to r_2 - 3r_3 \\ r_4 \to r_4 + 2r_3 \end{matrix} \begin{pmatrix} 4 & -1 & 5 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 3 & -1 & 8 & 0 \end{pmatrix}.$$

Subtracting the first row from the fourth row and then adding the second row to the fourth one we obtain

$$\begin{pmatrix} 4 & -1 & 5 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 3 & 0 \end{pmatrix} r_4 \to r_4 + r_2 \begin{pmatrix} 4 & -1 & 5 & 0 \\ 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Any further transformations are not necessary, because the determinant of the order 4 is equal to zero, but there is a submatrix of the third order the determinant of which is nonzero:

$$\begin{vmatrix} -1 & 5 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{vmatrix} = (-1) \cdot (-3) \cdot 1 = 3.$$

Hence, rank A = 3.

4.2. Main Definitions

Consider a system of *m* linear equations with *n* unknowns:

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

Here a_{ij} are numerical coefficients; b_i are constants $(i = 1,2,\cdots,m)$ and x_j are unknowns $(j = 1,2,\cdots,n)$.

A **solution of system** (1) is a set of values of the unknowns x_j that reduces all equations (1) to identities. If there exists a solution of simultaneous equations then the system is called **consistent**; otherwise, the system is **inconsistent**.

Using matrix multiplications, system of equations (1) can be represented by a single matrix equation

$$AX = B,$$

where A is the coefficient matrix consisting of a_{ij}; the column matrix $B = \| b_{i,1} \|$ is called the **non-homogeneous term**; X is the column matrix, whose elements are the unknowns x_j:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}.$$

If the non-homogeneous term B is equal to zero, then the linear system is called the **homogeneous system**:

$$AX = 0.$$

Two linear systems are called **equivalent**, if they have the same solution set.

Elementary transformations of the linear system is the process of obtaining an equivalent linear system from the given system by the following operations:

1) Interchange of two equations.
2) Multiplication of an equation by a nonzero number.
3) Addition of an equation multiplied by a constant to another equation.

Each of the above operations generates an equivalent linear system.

Two linear systems of equations are equivalent if one of them can be obtained from another by the elementary transformations.

Applying the linear transformations we try to find an equivalent system which can be easier solved.

4.3. Gaussian Elimination

Consider the augmented matrix of system (1):

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

There is one-to-one correspondence between the elementary transformations of the linear system and linear row operations on the augmented matrix. Indeed:

– Interchanging two equations of the system corresponds to interchanging the rows of the augmented matrix.
– Multiplication of an equation by a nonzero number corresponds to multiplication of the row by that number.
– Addition of two equations of the system corresponds to addition of the rows of the matrix.

The main idea is the following.

First, transform the augmented matrix to the upper triangle form or row echelon form:

$$(A|B) \quad \Rightarrow \quad \left(\begin{array}{cccccc|c} \tilde{a}_{11} & \cdot & \cdot & \cdots & \cdot & \cdots & \tilde{b}_1 \\ 0 & \tilde{a}_{22} & \cdot & \cdots & \cdot & \cdots & \tilde{b}_2 \\ 0 & 0 & \tilde{a}_{33} & \cdots & \cdot & \cdots & \tilde{b}_3 \\ \vdots & \cdots & \cdots & \ddots & \cdot & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{a}_{rr} & \cdots & \tilde{b}_r \\ 0 & 0 & 0 & \cdots & \cdot & \cdots & \vdots \end{array}\right)$$

Then write down the linear system corresponding to the augmented matrix in the triangle form or reduced row echelon form. This system is equivalent to the given system but it has a simpler form.

Finally, solve the system obtained by the method of back substitution. If it is necessary, assign parametric values to some unknowns.

This systematic procedure of solving systems of linear equations by elementary row operations is known as **Gaussian elimination**.

4.3.1. Some Examples

1) Solve the system below by Gaussian elimination:

$$\begin{cases} 2x_1 - x_2 + 5x_3 = 10 \\ x_1 + x_2 - 3x_3 = -2 \\ 2x_1 + 4x_2 + x_3 = 1 \end{cases}$$

Solution: Reduce the augmented matrix to a triangle form:

$$\left(\begin{array}{ccc|c} 2 & -1 & 5 & 10 \\ 1 & 1 & -3 & -2 \\ 2 & 4 & 1 & 1 \end{array}\right) \begin{array}{c} r_2 \to 2r_2 \end{array} \left(\begin{array}{ccc|c} 2 & -1 & 5 & 10 \\ 2 & 2 & -6 & -4 \\ 2 & 4 & 1 & 1 \end{array}\right)$$

$$\begin{array}{c} r_3 \to r_3 - r_2 \\ r_2 \to r_2 - r_1 \end{array} \left(\begin{array}{ccc|c} 2 & -1 & 5 & 10 \\ 0 & 3 & -11 & -14 \\ 0 & 2 & 7 & 5 \end{array}\right) r_3 \to 3r_3 \left(\begin{array}{ccc|c} 2 & -1 & 5 & 10 \\ 0 & 3 & -11 & -14 \\ 0 & 6 & 21 & 15 \end{array}\right)$$

$$r_3 \to r_3 - 2r_2 \left(\begin{array}{ccc|c} 2 & -1 & 5 & 10 \\ 0 & 3 & -11 & -14 \\ 0 & 0 & 43 & 43 \end{array}\right) r_3 \to \frac{r_3}{43} \left(\begin{array}{ccc|c} 2 & -1 & 5 & 10 \\ 0 & 3 & -11 & -14 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

The later matrix corresponds to the system

$$\begin{cases} 2x_1 - x_2 \ \ + 5x_3 \ = 10 \\ \qquad 3x_2 - 11x_3 = -14 \\ \qquad\qquad\quad x_3 = 1 \end{cases}$$

which is equivalent to the initial system.
Now the solution can be easily found:

$$3x_2 = -14 + 11x_3 = -3 \quad \Rightarrow \quad x_2 = -1,$$
$$2x_1 = 10 + x_2 - 5x_1 = 4 \quad \Rightarrow \quad x_1 = 2.$$

Thus we obtain the solution $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ of the given sysytem.

It is not difficult to verify the values of the unknowns satisfy all the given equations:

$$\begin{cases} 2x_1 - x_2 + 5x_3 = 2 \cdot 2 - (-1) + 5 \cdot 1 \equiv 10, \\ x_1 + x_2 - 3x_3 = 2 + (-1) - 3 \cdot 1 \equiv -2, \\ 2x_1 + 4x_2 + x_3 = 2 \cdot 2 + 4 \cdot (-1) + 1 \equiv 1. \end{cases}$$

2) Find all solutions of the system of equations via Gaussian elimination

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

Solution: The system can be represented by the augmented matrix. Apply the linear row operations:

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 2 & -1 & -1 & -2 \\ 4 & 1 & -3 & 5 \end{array}\right) \begin{array}{c} r_3 \to r_3 - 2r_2 \\ r_2 \to r_2 - 2r_1 \end{array} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -3 & 1 & -2 \\ 0 & 3 & -1 & 9 \end{array}\right)$$

$$r_3 \to r_3 + r_2 \left(\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -3 & 1 & -2 \\ 0 & 0 & 0 & 7 \end{array}\right).$$

The third row corresponds to the equation

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 7,$$

which evidently has no solutions.
Therefore, the given system is inconsistent.

3) Use Gaussian elimination to solve the system of equations.

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

Solution: By elementary transformations, the augmented matrix can be reduced to the row echelon form:

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & -2 & 0 \\ 2 & 1 & -1 & 1 & -2 \\ -1 & 1 & -3 & 1 & 4 \end{array}\right) \begin{array}{c} r_3 \to r_3 + r_1 \\ r_2 \to r_2 - 2r_1 \end{array} \left(\begin{array}{cccc|c} 1 & 1 & -1 & -2 & 0 \\ 0 & -1 & 1 & 5 & -2 \\ 0 & 2 & -4 & -1 & 4 \end{array}\right)$$

$$r_3 \to r_3 + 2r_2 \left(\begin{array}{cccc|c} 1 & 1 & -1 & -2 & 0 \\ 0 & -1 & 1 & 5 & -2 \\ 0 & 0 & -2 & 9 & 0 \end{array}\right).$$

The reduced matrix has the rank 3 and corresponds to the following system of linear equations:

$$\begin{cases} x_1 + x_2 - x_3 - 2x_4 = 0 \\ \quad - x_2 + x_3 + 5x_4 = -2 \\ \qquad -2x_3 + 9x_4 = 0 \end{cases}$$

The variable x_4 is considered to be an arbitrary parameter *c*, regardless of the value of which the remaining values of x_1, x_2, and x_3 reduce all equations of the given system to identities.
From the last equation we find

$$x_3 = \frac{9}{2}x_4 .$$

Then we obtain

$$x_2 = 2 + x_3 + 5x_4 = 2 + \frac{9}{2}x_4 + 5x_4 = 2 + \frac{19}{2}x_4 ,$$

$$x_1 = x_3 - x_2 + 2x_4 = \frac{9}{2}x_4 - \frac{19}{2}x_4 - 2 + 2x_4 = -2 - 3x_4 .$$

The general solution of the system

$$X = \begin{pmatrix} -2-3c \\ 2+\frac{19}{2}c \\ \frac{9}{2}c \\ c \end{pmatrix}$$

depends on the arbitrary parameter *c*. Any particular value of *c* gives a particular solution of the system. Assigning, for instance, the zero values to the parameter *c*, we obtain a particular solution $X_1 = \begin{pmatrix} -2 \\ 2 \\ 0 \\ 0 \end{pmatrix}$.

Setting *c* = 2 we obtain another particular solution $X_2 = \begin{pmatrix} -8 \\ 21 \\ 9 \\ 2 \end{pmatrix}$.

Conclusion: The given system has an infinite number of solutions.

Solution check: Let us verify that the set of values

$$x_1 = -2-3c, \qquad x_2 = 2+\frac{19}{2}c, \qquad x_3 = \frac{9}{2}c, \qquad x_4 = c$$

satisfies the given system of equations:

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

$$\begin{cases} -2-3c+2+\frac{19}{2}c-\frac{9}{2}c-2c = 0 \\ -4-6c+2+\frac{19}{2}c-\frac{9}{2}c+c = -2 \\ 2+3c+2+\frac{19}{2}c-\frac{27}{2}c+c = 4 \end{cases} \Rightarrow \begin{cases} 0 \equiv 0 \\ -2 \equiv -2 \\ 4 \equiv 4 \end{cases}$$

That is true.

4.4. Homogeneous Systems of Linear Equations

A homogeneous system of linear equations has the following form;

$$AX = 0, \qquad (2)$$

where A is the coefficient matrix, and X is the column matrix of the unknowns.

Evidently, any homogeneous system has the particular solution $X = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

which is called the **trivial** solution.

> **Theorem**:
> If X_1 and X_2 are solutions of a homogeneous system then a linear combinations of the solutions
> $$c_1X_1 + c_2X_2$$
> is also a solution of the system.

Proof: By the conditions of the theorem,

$$AX_1 = 0 \quad \text{and} \quad AX_2 = 0.$$

For any constants c_1 and c_2

$$c_1AX_1 = 0 \quad \Rightarrow \quad A(c_1X_1) = 0,$$

$$c_2AX_2 = 0 \quad \Rightarrow \quad A(c_2X_2) = 0.$$

Adding together the above identities we obtain

$$A(c_1X_1)+A(c_2X_2)=0,$$

which implies

$$A(c_1X_1+c_2X_2)=0.$$

Hence, the theorem.

4.4.1. Examples

1) Use Gaussian elimination to solve the following homogeneous system of equations.

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

Solution: By elementary transformations, the coefficient matrix can be reduced to the row echelon form

$$\begin{pmatrix} 1 & -1 & -1 & 3 \\ 1 & 1 & -2 & 1 \\ 4 & -2 & 4 & 1 \end{pmatrix} \begin{matrix} r_2\to r_2-r_1 \\ r_3\to r_3-4r_1 \end{matrix} \begin{pmatrix} 1 & -1 & -1 & 3 \\ 0 & 2 & -1 & -2 \\ 0 & 2 & 8 & -11 \end{pmatrix}$$

$$r_3\to r_3-r_2 \begin{pmatrix} 1 & -1 & -1 & 3 \\ 0 & 2 & -1 & -2 \\ 0 & 0 & 9 & -9 \end{pmatrix}.$$

The rank of this matrix equals 3, and so the system with four unknowns has an infinite number of solutions depending on one free variable. If we choose x_4 as the free variable and set $x_4=c$, then the leading unknowns, x_1, x_2 and x_3, are expressed through the parameter c. The above matrix corresponds to the following homogeneous system

$$\begin{cases} x_1-x_2-x_3+3c=0 \\ \quad 2x_2-x_3-2c=0 \\ \quad\quad 9x_3-9c=0 \end{cases}$$

The last equation implies $x_3=c$.

Using the method of back substitution we obtain

$$2x_2=x_3+2c=3c \qquad \Rightarrow \qquad x_2=\frac{3}{2}c,$$

$$x_1=x_2+x_3-3c=\frac{3}{2}c+c-3c=-\frac{1}{2}c.$$

Therefore, the general solution of the system is

$$X=\begin{pmatrix}-\frac{1}{2}c\\ \frac{3}{2}c\\ c\\ c\end{pmatrix}=c\begin{pmatrix}-\frac{1}{2}\\ \frac{3}{2}\\ 1\\ 1\end{pmatrix}.$$

To obtain a particular solution X_1 we have to assign some numerical value to the parameter c. If we set $c = 4$, then

$$X_1=\begin{pmatrix}-2\\ 6\\ 4\\ 4\end{pmatrix}$$

Solution check: The set of values of the unknowns

$$x_1=-\frac{1}{2}c,\quad x_2=\frac{3}{2}c,\quad x_3=c,\quad x_4=c$$

reduces equations of the given linear system to the identities:

$$\begin{cases}-\frac{1}{2}c-\frac{3}{2}c-c+3c=0\\ -\frac{1}{2}c+\frac{3}{2}c-2c+c=0\\ -\frac{4}{2}c-\frac{6}{2}c+4c+c=0\end{cases}\quad\Rightarrow\quad\begin{cases}0\equiv 0\\ 0\equiv 0\\ 0\equiv 0\end{cases}$$

2) Let $A=\begin{pmatrix}1&-1&-1&1\\ 2&-2&1&1\\ 5&-5&-2&4\end{pmatrix}$.

Find the solution of the homogeneous system of linear equations

$$AX=0.$$

Solution: Transform the coefficient matrix to the row echelon form:

$$\begin{pmatrix}1&-1&-1&1\\ 2&-2&1&1\\ 5&-5&-2&4\end{pmatrix}\begin{matrix}r_2\to r_2-2r_1\\ r_3\to r_3-2r_2-r_1\end{matrix}\begin{pmatrix}1&-1&-1&1\\ 0&0&3&-1\\ 0&0&-3&1\end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -3 & 1 \end{pmatrix} r_3 \to r_3 + r_2 \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since $\operatorname{rank} A = 2$, we have to choose two unknowns as the leading unknowns and to assign parametric values to the remaining unknowns. Setting $x_2 = c_1$ and $x_3 = c_2$ we obtain the following linear system:

$$\begin{cases} 2x_1 - c_1 - c_2 + x_4 = 0 \\ \qquad\qquad 3c_2 - x_4 = 0 \end{cases}$$

Therefore,

$$x_4 = 3c_2 \quad \text{and} \quad x_1 = \frac{1}{2}(c_1 + c_2 - 3c_2) = \frac{1}{2}c_1 - c_2.$$

Thus, the given system has the following general solution:

$$X = \begin{pmatrix} \frac{1}{2}c_1 - c_2 \\ c_1 \\ c_2 \\ 3c_2 \end{pmatrix}.$$

In view of the matrix properties, the general solution can be also expressed as the linear combination of particular solutions:

$$X = \begin{pmatrix} \frac{1}{2}c_1 \\ c_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -c_2 \\ 0 \\ c_2 \\ 3c_2 \end{pmatrix} = c_1 \begin{pmatrix} 1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 3 \end{pmatrix}.$$

The particular solutions $X_1 = \begin{pmatrix} 1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $X_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 3 \end{pmatrix}$ form the system of solutions which is called the fundamental set of solutions.

Thus,

$$X = c_1 X_1 + c_2 X_2.$$

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

i) Solve the following homogeneous system of linear equations
$$AX = 0.$$
ii) Explain why there are no solutions, an infinite number of solutions, or exactly one solution.

Solution: Note that any homogeneous system is consistent and has at least the trivial solution.
Transform the coefficient matrix to the triangular or row echelon form.

$$\begin{pmatrix} -1 & 1 & -1 \\ 3 & -1 & -1 \\ 2 & 2 & -3 \end{pmatrix} \begin{matrix} r_2 \to r_2 + 3r_1 \\ r_3 \to r_3 + 2r_1 \end{matrix} \begin{pmatrix} -1 & 1 & -1 \\ 0 & 2 & -3 \\ 0 & 4 & -5 \end{pmatrix} r_3 \to r_3 - 2r_2 \begin{pmatrix} -1 & 1 & -1 \\ 0 & 2 & -3 \\ 0 & 0 & 1 \end{pmatrix}.$$

The rank of A equals 3. Therefore, there are no free variables and the system

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

has the trivial solution $x_1 = x_2 = x_3 = 0$, only.

4.5. Cramer's Rule

There is a particular case when the solution of a system of linear equations can be written in the explicit form. The corresponding theorem is known as Cramer's Rule whose importance is determined by its applications in theoretical investigations.

Cramer's Rule: Let
$$A\,X = B \tag{3}$$
be a system of n linear equations with n unknowns.

If the coefficient matrix A is regular, then the system is consistent and has a unique solution set $\{x_1, x_2, \dots, x_n\}$ which is represented by the formula:

$$x_i = \frac{D_i}{D}, \qquad i = 1, 2, \dots, n, \tag{4}$$

where $D = \det A$; D_i is the determinant of the matrix obtained by replacing the i-th column of A with the column matrix B:

$$D_i = \begin{vmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & & \cdots & & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{vmatrix}.$$

Proof: We have to prove the following statements:

1) a solution is unique;
2) formulas (4) follow from system (3);
3) formulas (4) yield system (3).

Since $\det A \neq 0$, there exists the inverse of *A*. Therefore, matrix equality (3) implies

$$X = A^{-1}B. \qquad (5)$$

By the theorem of inverse matrix, for any regular matrix *A* there exists a unique inverse matrix

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

that proves the uniqueness of solution (5).

The *i*-th row of $\operatorname{adj} A$ is formed by the cofactors $A_{1,i}$, $A_{2,i}$, …, $A_{n,i}$ of the elements in the *i*-th column of the matrix *A*. The equality (5) implies

$$x_i = (A^{-1}B)_i = \frac{1}{D}\begin{pmatrix} A_{1,i} & A_{2,i} & \cdots & A_{n,i} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{D}\sum_{k=1}^{n} A_{k,i}\, b_k .$$

The sum on the right side is the expansion of the determinant D_i in terms of the elements in the *i*-th column. Hence, we have obtained the desired formula:

$$x_i = \frac{D_i}{D}.$$

Now prove that the set $\{x_1, x_2, \ldots, x_n\}$, with $x_i = \frac{1}{D}\sum_{k=1}^{n} A_{k,i}\, b_k$, implies system (3).

Multiply both sides of this equality by $D\, a_{j,i}$ and then sum the result over *i*:

$$D\sum_{i=1}^{n} a_{j,i}\, x_i = \sum_{i=1}^{n}\sum_{k=1}^{n} A_{k,i}\, a_{j,i}\, b_k .$$

Interchange the order of summation in the expression on the right side.

$$D\sum_{i=1}^{n} a_{j,i}\, x_i = \sum_{k=1}^{n} b_k \sum_{i=1}^{n} A_{k,i}\, a_{j,i}\,. \qquad (6)$$

In view of the theorem of inverse matrix,

$$\sum_{i=1}^{n} A_{k,i} a_{j,i} = \delta_{kj} \det A,$$

where δ_{kj} is the Kronecker delta.

The Kronecker delta takes away the summation over k in expression (6):

$$D\sum_{i=1}^{n} a_{j,i}\, x_i = D\sum_{k=1}^{n} b_k \delta_{kj} = Db_j\,.$$

Hence, we have the desired linear system of equations:

$$\sum_{i=1}^{n} a_{j,i}\, x_i = b_j \qquad (i = 1, 2, \ldots, n).$$

The theorem is proven.

Example: Use Cramer's Rule to solve the following system of linear equations.

$$\begin{cases} 2x_1 - x_2 + 5x_3 = 10 \\ x_1 + x_2 - 3x_3 = -2 \\ 2x_1 + 4x_2 + x_3 = 1 \end{cases}$$

Solution:

$$D = \begin{vmatrix} 2 & -1 & 5 \\ 1 & 1 & -3 \\ 2 & 4 & 1 \end{vmatrix} \overset{\substack{r_1 \to r_1 - 2r_2 \\ r_3 \to r_3 - r_1}}{=} \begin{vmatrix} 0 & -3 & 11 \\ 1 & 1 & -3 \\ 0 & 5 & -4 \end{vmatrix} = -\begin{vmatrix} -3 & 11 \\ 5 & -4 \end{vmatrix} = 43,$$

$$D_1 = \begin{vmatrix} 10 & -1 & 5 \\ -2 & 1 & -3 \\ 1 & 4 & 1 \end{vmatrix} \overset{\substack{r_1 \to r_1 + 5r_2 \\ r_2 \to r_2 + r_3}}{=} \begin{vmatrix} 0 & 4 & -10 \\ 0 & 9 & -1 \\ 1 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 4 & -10 \\ 9 & -1 \end{vmatrix} = 86,$$

$$D_2 = \begin{vmatrix} 2 & 10 & 5 \\ 1 & -2 & -3 \\ 2 & 1 & 1 \end{vmatrix} \overset{\substack{r_1 \to r_1 - r_3 \\ r_3 \to r_3 - 2r_1}}{=} \begin{vmatrix} 0 & 9 & 4 \\ 1 & -2 & -3 \\ 0 & 5 & 7 \end{vmatrix} = -\begin{vmatrix} 9 & 4 \\ 5 & 7 \end{vmatrix} = -43,$$

$$D_3 = \begin{vmatrix} 2 & -1 & 10 \\ 1 & 1 & -2 \\ 2 & 4 & 1 \end{vmatrix} \overset{\substack{r_1 \to r_1 - 2r_2 \\ r_3 \to r_3 - r_1}}{=} \begin{vmatrix} 0 & -5 & 9 \\ 1 & 1 & -2 \\ 0 & 2 & 5 \end{vmatrix} = -\begin{vmatrix} -5 & 9 \\ 2 & 5 \end{vmatrix} = 43.$$

Therefore,

$$x_1 = \frac{D_1}{D} = \frac{-86}{43} = -2, \quad x_2 = \frac{D_2}{D} = \frac{-43}{43} = -1, \quad x_3 = \frac{D_3}{D} = \frac{43}{43} = 1.$$

Compare this solution with that obtained by Gaussian elimination in Example 1, p. 44.

4.6. Cramer’s General Rule

Cramer's General Rule formulates the existence condition of a solution for any given system of linear equations.

Cramer's General Rule: A system of m linear equations with n unknowns

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

is consistent if and only if the rank of the augmented matrix is equal to the rank of the coefficient matrix.

Proof: Let A be the coefficient matrix and let $\overline{A}$ be the augmented matrix of the given system. We have to prove that

(a) If the system is consistent, then $\operatorname{rank} A = \operatorname{rank} \overline{A}$.

(b) If $\operatorname{rank} A = \operatorname{rank} \overline{A}$, then the system is consistent.

To prove statement (a), we have to assume that the system is consistent.
Consider the augmented matrix $\overline{A}$:

$$\overline{A} = \left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,1} & b_1 \\ \vdots & \cdots & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} & b_m \end{array} \right).$$

If we subtract the first column multiplied by x_1, the second column multiplied by x_2, and so on from the last column, then we obtain the matrix of the same rank as $\overline{A}$ (by the theorem of matrix rank):

$$\operatorname{rank} \overline{A} = \operatorname{rank}\left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,n} & b_1 - (a_{1,1}x_1 + \ldots + a_{1,n}x_n) \\ \vdots & \cdots & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} & b_m - (a_{m,1}x_1 + \ldots + a_{m,n}x_n) \end{array}\right).$$

Since the system is consistent, each elements of the last column equals zero. Therefore,

$$\operatorname{rank} \overline{A} = \operatorname{rank}\left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,n} & 0 \\ \vdots & \cdots & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} & 0 \end{array}\right) = \operatorname{rank}\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \cdots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} = \operatorname{rank} A.$$

(b) Now suppose that $\operatorname{rank} A = \operatorname{rank} \overline{A} = r$. It means there exists a nonsingular $r \times r$ submatrix $\tilde{A}$ of the matrix *A*, by which we select *r* leading unknowns and assign parametric values to the remaining $(n - r)$ free unknowns. The reduced system of linear equations is equivalent to the initial system, and, by Cramer's Rule, it has a unique solution for each set of values of the free unknowns.
Hence, the theorem.

Corollary:

i) If $\operatorname{rank} A = \operatorname{rank} \overline{A}$ and equals the number *n* of unknowns, then the solution of the system is unique.

ii) If $\operatorname{rank} A = \operatorname{rank} \overline{A} < n$ then there exist an infinite number of solutions of the given system.

Statement i) follows from the Cramer's Rule.
If $\operatorname{rank} A = \operatorname{rank} \overline{A} = r < n$ then the given system is equivalent to the system of *r* linear equations with *r* leading unknowns.
An infinite number of the values of the remaining (*n* - *r*) unknowns leads to an infinite number of solutions.

Examples:

1. The system of linear equations is given below. Formulate the conditions on *a, b,* and *c*, making the system

$$\begin{cases} x_1 + 2x_2 + 3x_3 = a \\ 4x_1 + 5x_2 + 6x_3 = b \\ 7x_1 + 8x_2 + 9x_3 = c \end{cases}$$

to be inconsistent.

Solution: Consider the augmented matrix and transform it to the reduced row echelon form.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & a \\ 4 & 5 & 6 & b \\ 7 & 8 & 9 & c \end{array}\right) \begin{array}{c} r_2 \to r_2 - 2r_1 \\ r_3 \to r_3 - 3r_1 \end{array} \left(\begin{array}{ccc|c} 1 & 2 & 3 & a \\ 2 & 1 & 0 & b-2a \\ 4 & 2 & 0 & c-3a \end{array}\right)$$

$$r_3 \to r_3 - 2r_2 \left(\begin{array}{ccc|c} 1 & 2 & 3 & a \\ 2 & 1 & 0 & b-2a \\ 0 & 0 & 0 & c-3a-2b \end{array}\right).$$

The system is inconsistent, if

$$c - 3a - 2b \neq 0 .$$

Otherwise, one of unknowns is a parametric variable, and the system has an infinite number of solutions.

2. Given the reduced row echelon form of the augmented matrix,

$$\overline{A} = \left(\begin{array}{cccc|c} 1 & 3 & -1 & 5 & 2 \\ 0 & 7 & 0 & 2 & 4 \\ 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right),$$

find the number of solutions of the corresponding system. It is not necessary to solve the system.

Solution: The rank of the coefficient matrix equals the rank of the augmented matrix and equals the number of the unknowns. Hence, by the Corollary to Cramer's General Rule, the solution is unique.

3. Let a system of linear equations be given by the augmented matrix

$$\overline{A} = \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 0 & 5 & 6 & 7 & 2 \\ 0 & 0 & 8 & 9 & 3 \\ 0 & 0 & 0 & 0 & a \end{array}\right).$$

How many solutions has the system?

Solution: If $a \neq 0$ then $\operatorname{rank} \overline{A} = 4$ while $\operatorname{rank} A = 3$. By Cramer's General Rule, the system is inconsistent, and so it has no solutions.
If $a = 0$ then $\operatorname{rank} A = \operatorname{rank} \overline{A} = 3$, while the number of the unknowns is $n = 4$. So one of the unknowns has to be considered as a parameter, and the system has a solution for each value of that parameter. Hence, the system has an infinite number of solutions.

VECTOR ALGEBRA

5. Vectors

5.1. Basic Definitions

A **three-dimensional vector** in some coordinate system is an ordered triplet of numbers that obeys certain rules of addition and multiplication, and that are transformed under rotation of a coordinate system just as the coordinates of a point.

The members of the triplet are called the **coordinates of the vector**.

Likewise, one can define an *n*-dimensional vector.

Usually, vectors are denoted by boldface letters: $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, …. The notation

$$\boldsymbol{a} = \{a_1, a_2, a_3\}$$

means that the numbers a_1, a_2 and a_3 are the coordinates of the vector $\boldsymbol{a}$ in a three-dimensional coordinate system.

Two vectors, $\boldsymbol{a} = \{a_1, a_2, a_3\}$ and $\boldsymbol{b} = \{b_1, b_2, b_3\}$, are **equal**, if their coordinates are respectively equal, that is,

$$\boldsymbol{a} = \boldsymbol{b} \qquad \Leftrightarrow \qquad \begin{cases} a_1 = b_1, \\ a_2 = b_2, \\ a_3 = b_3. \end{cases}$$

Note that a vector equality is equivalent to the system of three scalar equalities for the coordinates of the vector.

Linear vector operations include the multiplication of a vector by a scalar quantity and the addition of vectors.

If a vector $\boldsymbol{a} = \{a_1, a_2, a_3\}$ is multiplied by a scalar λ, then $\boldsymbol{b} = \lambda\boldsymbol{a}$ is the vector such that

$$\begin{cases} b_1 = \lambda a_1, \\ b_2 = \lambda a_2, \\ b_3 = \lambda a_3. \end{cases}$$

The **sum** of two vectors $\boldsymbol{a} = \{a_1, a_2, a_3\}$ and $\boldsymbol{b} = \{b_1, b_2, b_1\}$ is the vector

$$\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} = \{a_1 + b_1, a_2 + b_2, a_3 + b_3\}.$$

The **difference** between two vectors is defined in terms of addition:

$$\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b} = \boldsymbol{a} + (-\boldsymbol{b}).$$

Therefore,

$$\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b} \qquad \Leftrightarrow \qquad c_1 = a_1 - b_1, \quad c_2 = a_2 - b_2, \quad c_3 = a_3 - b_3.$$

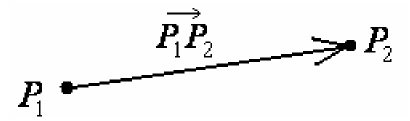
5.2. Geometric Interpretation

5.2.1. Vectors in Three-Dimensional Space

Consider a rectangular coordinate system.
Let $\boldsymbol{a} = \{a_1, a_2, a_3\}$ be a given vector, and P_1 and P_2 be two points with the coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2), respectively.
The points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ can be selected so as to satisfy conditions

$$a_1 = x_2 - x_1, \qquad a_2 = y_2 - y_1, \qquad a_3 = z_2 - z_1.$$

Therefore, vector $\boldsymbol{a}$ can be interpreted as the directed line segment $\overrightarrow{P_1P_2}$ from P_1 to P_2:

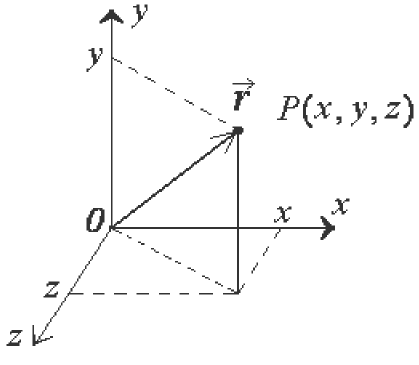


The coordinates of $\overrightarrow{P_1P_2}$ are equal to the differences between the corresponding coordinates of the points $P_2(x_2, y_2, z_2)$ and $P_1(x_1, y_1, z_1)$:

$$\overrightarrow{P_1P_2} = \{x_2 - x_1, y_2 - y_1, z_2 - z_1\}.$$

The point P_1 is the base of $\overrightarrow{P_1P_2}$ and P_2 is the head. The base of a vector is also called the vector tail or the origin of the vector.

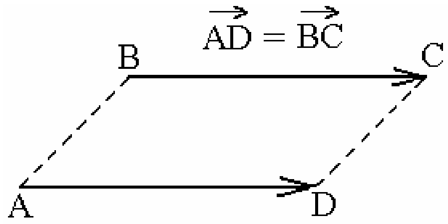
The **length of a vector** $\overrightarrow{P_1P_2}$ is defined as the length of the line segment joining P_1 and P_2.
Note that a vector is a quantity possessing both magnitude and direction at once. The boldface letter $\boldsymbol{a}$ represents a vector quantity, while $a = |\boldsymbol{a}|$ is the magnitude of the vector $\boldsymbol{a}$, that is, a is a scalar quantity entirely defined by a numerical value.
If a vector joins the origin of the coordinate system with a point $P(x, y, z)$, then it is called the **radius-vector** of the point P and denoted as $\boldsymbol{r}$.

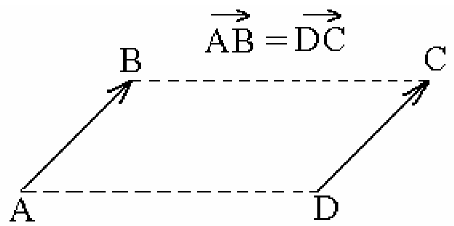


5.2.2. Linear Vector Operations

Equality of Vectors

By parallel translation, equal vectors should coincide with each other:

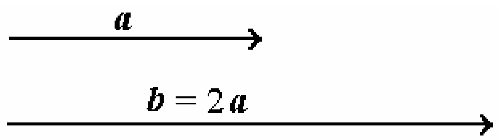




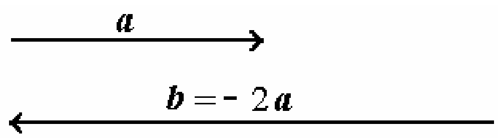
Scalar Multiplication

The length of the vector $\boldsymbol{b} = \lambda \boldsymbol{a}$ is $b = |\lambda| a$.

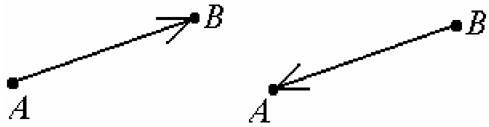
If $\lambda > 0$ then $\boldsymbol{b}$ is a vector of the same direction as $\boldsymbol{a}$:



If $\lambda < 0$ then vector $\boldsymbol{b} = \lambda \boldsymbol{a}$ has the opposite direction with respect to $\boldsymbol{a}$:

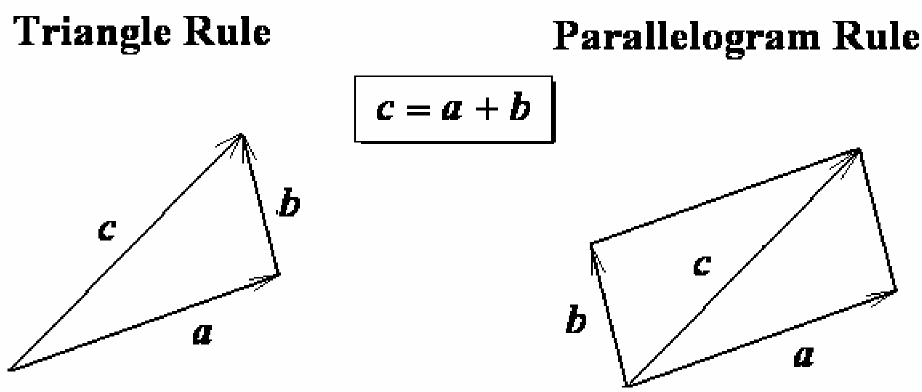


The opposite vector of $\overrightarrow{AB}$ is the vector $\overrightarrow{BA} = -\overrightarrow{AB}$.



The length of a **unit vector** equals unity. If $\boldsymbol{a}$ is a non-zero vector then $\boldsymbol{u} = \dfrac{\boldsymbol{a}}{a}$ is the unit vector in the direction of $\boldsymbol{a}$.

The Sum of Two Vectors

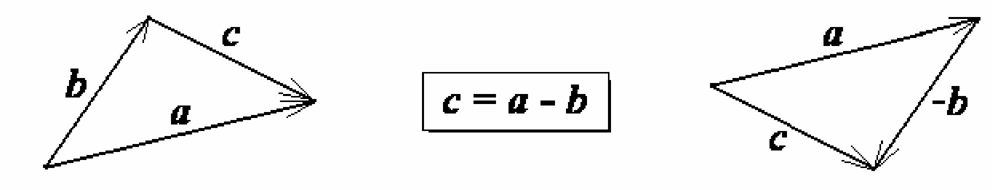


The Difference Between Two Vectors

In order to subtract a vector ***b*** from ***a***, add the opposite of ***b*** to ***a***:

$$\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b} = \boldsymbol{a} + (-\boldsymbol{b}).$$

Thus, the difference between two vectors ***a*** and ***b*** is the vector $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$ such that $\boldsymbol{c} + \boldsymbol{b} = \boldsymbol{a}$.

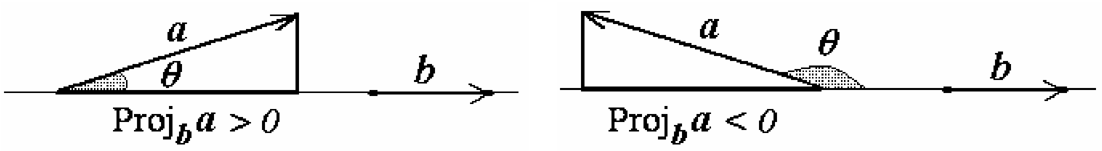


5.2.3. Projection of a Vector in a Given Direction

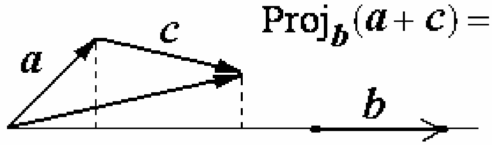
Let θ be an angle between two vectors, ***a*** and ***b***.
The quantity

$$\mathrm{Proj}_{\boldsymbol{b}}\boldsymbol{a} = a\cos\theta \tag{1}$$

is called the **projection** of ***a*** on ***b***.
If θ is an acute angle then the projection is positive.
If θ is an obtuse angle then the projection is negative.

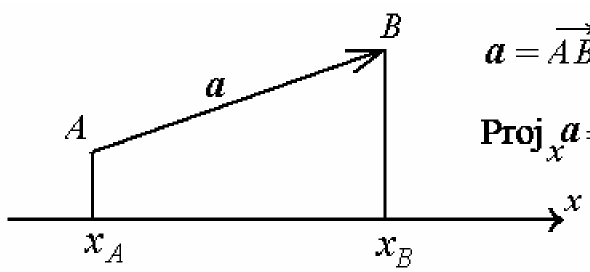


One can easily prove that



$$\mathrm{Proj}_{\boldsymbol{b}}(\boldsymbol{a} + \boldsymbol{c}) = \mathrm{Proj}_{\boldsymbol{b}}\boldsymbol{a} + \mathrm{Proj}_{\boldsymbol{b}}\boldsymbol{c}$$

If the direction is determined by the *x*-axis, then the projection of ***a*** onto the *x*-axis equals the difference between the coordinates of the endpoints of the vector:



$$\boldsymbol{a} = \overrightarrow{AB}$$

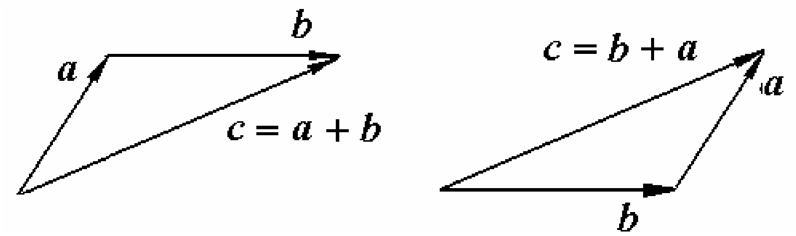
$$\mathrm{Proj}_{x}\boldsymbol{a} = x_B - x_A$$

5.2.4. Properties of Linear Vector Operations

All the below formulated properties are based on the properties of real numbers, and they are result of the definitions of linear vector operations. Proofs can be easily performed by the reader.

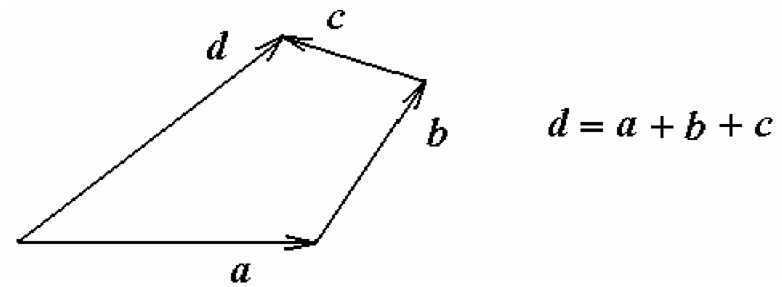
1) The commutative law for addition:

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$$



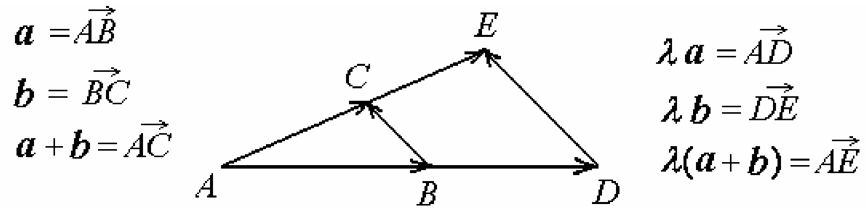
2) The associative law for addition:

$$\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c}$$



3) The distributive laws for multiplication over addition:

$$\lambda(\boldsymbol{a}+\boldsymbol{b}) = \lambda\,\boldsymbol{a} + \lambda\,\boldsymbol{b}, \qquad (\lambda+\mu)\boldsymbol{a} = \lambda\,\boldsymbol{a} + \mu\,\boldsymbol{a}.$$



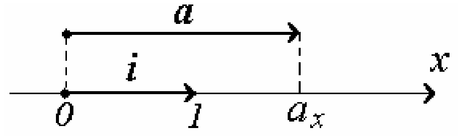
5.3. Decomposition of Vectors into Components

5.3.1. Rectangular Orthogonal Basis

1) Let $\boldsymbol{i} = \{1, 0, 0\}$ be the unit vector in the positive direction of the x-axis. Any vector $\boldsymbol{a} = \{a_x, 0, 0\}$ can be expressed as

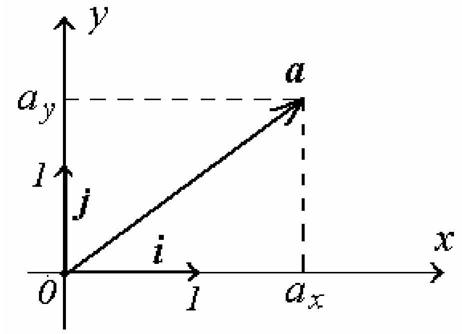
$$\boldsymbol{a} = \{a_x, 0, 0\} = a_x\{1, 0, 0\} = a_x\,\boldsymbol{i}.$$

The vector $\boldsymbol{i}$ is said to be a basis in an one-dimensional space of vectors.



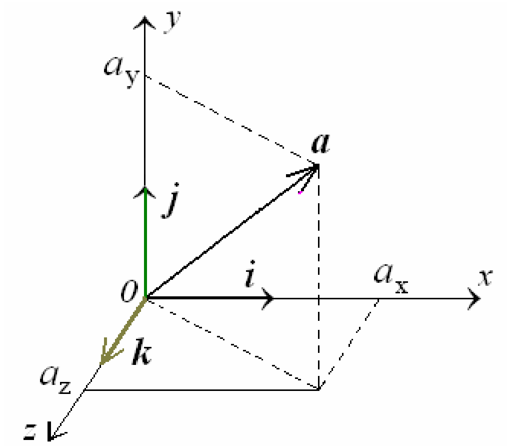
2) Let $\boldsymbol{i}$ = {1, 0, 0} and $\boldsymbol{j}$ = {0, 1, 0} be two unit vectors in the positive directions of the *x*-axis and *y*-axis, respectively.
Any vector $\boldsymbol{a} = \{a_x, a_y, 0\}$ can be expressed as

$$\boldsymbol{a} = \{a_x, a_y, 0\} = a_x\{1,0,0\} + a_y\{0,1,0\} = a_x\,\boldsymbol{i} + a_y\,\boldsymbol{j}\,.$$



They say that $\boldsymbol{i}$ and $\boldsymbol{j}$ are the basis vectors in a two-dimensional space of vectors.

3) Let $\boldsymbol{i}$ = {1, 0, 0}, $\boldsymbol{j}$ = {0, 1, 0}, and $\boldsymbol{k}$ = {0, 0, 1} be three mutually orthogonal unit vectors in the positive directions of the Cartesian coordinate axes.



Any vector $\boldsymbol{a} = \{a_x, a_y, a_z\}$ can be expressed as the linear combination of the vectors $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$:

$$\boldsymbol{a} = \{a_x, a_y, a_z\} = a_x\{1,0,0\} + a_y\{0,1,0\} + a_z\{0,0,1\}\,.$$

Therefore, we obtain the resolution of an arbitrary vector $\boldsymbol{a}$

$$\boldsymbol{a} = a_x\,\boldsymbol{i} + a_y\,\boldsymbol{j} + a_z\,\boldsymbol{k} \qquad (2)$$

over the orthogonal basis of vectors, where quantities a_x, a_y and a_z are called the coordinates of the vector $\boldsymbol{a}$ with respect to this basis.

5.3.2. Linear Dependence of Vectors

Let $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ be any vectors and $\lambda_1, \lambda_2, \ldots, \lambda_n$ be numbers.
The expression of the form

$$\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \ldots + \lambda_n \boldsymbol{a}_n$$

is called a **linear combination** of the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$.
If there exists a non-trivial solution of the homogeneous vector equation

$$\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \ldots + \lambda_n \boldsymbol{a}_n = 0 \qquad (3)$$

with respect to $\lambda_1, \lambda_2, \ldots, \lambda_n$, then it is said that $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is the set of linear **dependent vectors**.
Otherwise, if equation (3) has only the trivial solution

$$\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0,$$

then $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$ is called the set of linear **independent vectors**.
In other words, the set of vectors is linear dependent, if one of the vectors can be expressed as a linear combination of the other vectors of the set.
For instance, if $\lambda_1 \neq 0$, then

$$\boldsymbol{a}_1 = -\frac{1}{\lambda_1}(\lambda_2 \boldsymbol{a}_2 + \ldots + \lambda_n \boldsymbol{a}_n).$$

Theorem:

1) Any two non-zero vectors are linear dependent, if and only if they are collinear.
2) Any three non-zero vectors are linear dependent, if and only if they are coplanar.
3) Any four vectors in a three-dimensional space are linear dependent.

Important note: The theorem states that two non-collinear vectors are linear independent, and three non-coplanar vectors are linear independent.

Proof:

1) Two vectors, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, are linear dependent, if the equation

$$\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 = 0$$

has a non-zero solution with respect to λ_1 and λ_2.

In this case, $\lambda_2 \boldsymbol{a}_2$ is the opposite vector of $\lambda_1 \boldsymbol{a}_1$, that is, $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are collinear vectors.

Hence, any two collinear vectors are linear dependent, and any two non-collinear vectors are linear independent.

2) Consider a set of three vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$. In the coordinate form, the vector equation

$$\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \lambda_3 \boldsymbol{a}_3 = 0$$

can be expressed as the homogeneous system of the following linear equations:

$$\begin{cases} \lambda_1 a_{11} + \lambda_2 a_{21} + \lambda_3 a_{31} = 0, \\ \lambda_1 a_{12} + \lambda_2 a_{22} + \lambda_3 a_{32} = 0, \\ \lambda_1 a_{13} + \lambda_2 a_{23} + \lambda_3 a_{33} = 0. \end{cases}$$

At first, let us assume that the vectors $\boldsymbol{a}_1 = \{a_{11}, a_{12}, a_{13}\}$, $\boldsymbol{a}_2 = \{a_{21}, a_{22}, a_{23}\}$ and $\boldsymbol{a}_3 = \{a_{31}, a_{32}, a_{33}\}$ are coplanar.
Then there exists a coordinate system in which

$$a_{13} = a_{23} = a_{33} = 0\,.$$

Therefore, the above homogeneous system is reduced to the system of two linear equations with three unknowns λ_1, λ_2 and λ_3, and hence, has a non-zero solution.
Thus, a set of three coplanar vectors is linear dependent.
Assume now that vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ are non-coplanar.
A linear combination of vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ is a vector lying in the same plane as $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. Hence, $\boldsymbol{a}_3$ cannot be expressed as a linear combination of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, and so a set of three non-coplanar vectors is linear independent.
3) In case of four vectors, the equation

$$\lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2 + \lambda_3 \boldsymbol{a}_3 + \lambda_4 \boldsymbol{a}_4 = 0\,,$$

is equivalent to the homogeneous system of three linear equations with four unknowns λ_1, λ_2, λ_3, and λ_4. Such system has an infinite number of solutions. Hence, any set of four vectors is linear dependent.
A set of *n* linear independent vectors is called a **basis** in the *n*-dimension space of vectors. Therefore, any three non-coplanar vectors form the basis in the three-dimensional space of vectors, that is, any vector $\boldsymbol{d}$ can be expressed as a linear combination of the basis vectors:

$$\boldsymbol{d} = d_1 \boldsymbol{a}_1 + d_2 \boldsymbol{a}_2 + d_3 \boldsymbol{a}_3\,.$$

This formula generalizes the concept of the rectangular vector basis $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ to an arbitrary set $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ of non-coplanar vectors in a three-dimensional space. The numbers d_1, d_2 and d_3 are called the coordinates of $\boldsymbol{d}$ in the basis of vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$.

5.3.3. Vector Bases

Let $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ and $\{\tilde{\boldsymbol{a}}_1, \tilde{\boldsymbol{a}}_2, \tilde{\boldsymbol{a}}_3\}$ be two different bases in a three-dimensional space of vectors. By the theorem of linear independent vectors,

$$\boldsymbol{d} = d_1 \boldsymbol{a}_1 + d_2 \boldsymbol{a}_2 + d_3 \boldsymbol{a}_3 \qquad (4)$$

and

$$\boldsymbol{d} = \tilde{d}_1 \tilde{\boldsymbol{a}}_1 + \tilde{d}_2 \tilde{\boldsymbol{a}}_2 + \tilde{d}_3 \tilde{\boldsymbol{a}}_3 \qquad (5)$$

for an arbitrary vector $\boldsymbol{d}$.

In order to find relations between the coordinates of $\boldsymbol{d}$ in these bases, we need to resolve vectors $\tilde{\boldsymbol{a}}_1$, $\tilde{\boldsymbol{a}}_2$ and $\tilde{\boldsymbol{a}}_3$ into the basis vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$:

$$\begin{cases} \tilde{\boldsymbol{a}}_1 = a_{11}\boldsymbol{a}_1 + a_{12}\boldsymbol{a}_2 + a_{13}\boldsymbol{a}_3 \\ \tilde{\boldsymbol{a}}_2 = a_{21}\boldsymbol{a}_1 + a_{22}\boldsymbol{a}_2 + a_{23}\boldsymbol{a}_3 \\ \tilde{\boldsymbol{a}}_3 = a_{31}\boldsymbol{a}_1 + a_{32}\boldsymbol{a}_2 + a_{33}\boldsymbol{a}_3 \end{cases}$$

Coefficients a_{ij} of the linear combinations are the coordinates of the vectors $\tilde{\boldsymbol{a}}_i$ in the basis of vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$.

Substituting these expressions in equality (5) and combining the similar terms we obtain

$$\boldsymbol{d} = \sum_{k=1}^{3} (a_{1k}\tilde{d}_1 + a_{2k}\tilde{d}_2 + a_{3k}\tilde{d}_3)\boldsymbol{a}_k .$$

In view of equality (4), we get the transformation formulas of the coordinates of a vector from one basis to another:

$$\begin{cases} d_1 = a_{11}\tilde{d}_1 + a_{21}\tilde{d}_2 + a_{31}\tilde{d}_3 \\ d_2 = a_{12}\tilde{d}_1 + a_{22}\tilde{d}_2 + a_{32}\tilde{d}_3 \\ d_3 = a_{13}\tilde{d}_1 + a_{23}\tilde{d}_2 + a_{33}\tilde{d}_3 \end{cases}$$

A transformation of a rectangular basis by rotation of the coordinate system is considered in section 5.7.

Example: Let be given the vector resolution $\boldsymbol{d} = 4\boldsymbol{i} + 7\boldsymbol{j} + \boldsymbol{k}$ and the basis vectors $\tilde{\boldsymbol{a}}_1 = \{3,0,-5\}$, $\tilde{\boldsymbol{a}}_2 = \{1,1,-8\}$ and $\tilde{\boldsymbol{a}}_3 = \{7,-2,0\}$.

To find the coordinates of $\boldsymbol{d}$ in the basis of $\tilde{\boldsymbol{a}}_1$, $\tilde{\boldsymbol{a}}_2$ and $\tilde{\boldsymbol{a}}_3$, we have to solve the system of linear equations:

$$\begin{cases} 4 = 3\tilde{d}_1 + \tilde{d}_2 + 7\tilde{d}_3 \\ 7 = 0\cdot\tilde{d}_1 + \tilde{d}_2 - 2\tilde{d}_3 \\ 1 = 5\tilde{d}_1 - 8\tilde{d}_2 + 0\cdot\tilde{d}_3 \end{cases} \quad \Rightarrow \quad \tilde{d}_1 = 5,\ \tilde{d}_2 = 3 \text{ and } \tilde{d}_3 = -2 .$$

5.4. Scalar Product of Vectors

Assume that vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are given by their coordinates in a rectangular coordinate system:

$$\boldsymbol{a}=\{a_x,a_y,a_z\} \quad \text{and} \quad \boldsymbol{b}=\{b_x,b_y,b_z\}.$$

The **scalar product** $\boldsymbol{a}\cdot\boldsymbol{b}$ is a number that equals the sum of the products of the corresponding coordinates:

$$\boldsymbol{a}\cdot\boldsymbol{b}=a_xb_x+a_yb_y+a_zb_z. \qquad (6)$$

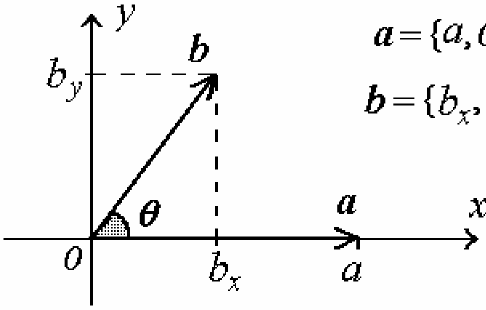
The scalar product is also known as inner or dot product. It is also denoted as $(\boldsymbol{a}, \boldsymbol{b})$.

Theorem: If θ is the angle between vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ then

$$\boldsymbol{a}\cdot\boldsymbol{b}=a\,b\cos\theta, \qquad (7)$$

where a and b are the lengths of the vectors.

Proof: Let us choose a rectangular coordinate system such that
- both vectors, $\boldsymbol{a}$ and $\boldsymbol{b}$, lie in the x,y-plane;
- the x-axis is directed along the vector $\boldsymbol{a}$.



$$\boldsymbol{a}=\{a,0,0\}$$

$$\boldsymbol{b}=\{b_x,b_y,0\}$$

Since $a_x=a$, $a_y=a_z=0$, and $b_x=b\cos\theta$, we obtain the desired result.

The theorem states that

$$\boldsymbol{a}\cdot\boldsymbol{b}=a\,\mathrm{Pr}\,oj_{\boldsymbol{a}}\boldsymbol{b}=b\,\mathrm{Pr}\,oj_{\boldsymbol{b}}\boldsymbol{a},$$

$$\cos\theta=\frac{\boldsymbol{a}}{a}\cdot\frac{\boldsymbol{b}}{b},$$

where $\dfrac{\boldsymbol{a}}{a}$ and $\dfrac{\boldsymbol{b}}{b}$ are the unit vectors in the directions of $\boldsymbol{a}$ and $\boldsymbol{b}$, correspondingly.

If $\boldsymbol{a}\perp\boldsymbol{b}$ then $\cos\theta=\cos\dfrac{\pi}{2}=0$, which implies the following orthogonality condition of the vectors $\boldsymbol{a}=\{a_x,a_y,a_z\}$ and $\boldsymbol{b}=\{b_x,b_y,b_z\}$:

$$a_xb_x+a_yb_y+a_zb_z=0.$$

If $\boldsymbol{b}=\boldsymbol{a}$ then $\theta=0$, $\cos\theta=1$, and so

$$\boldsymbol{a}\cdot\boldsymbol{a} = a^2 = a_x^2 + a_y^2 + a_z^2 .$$

Therefore, the length of the vector ***a*** is expressed as

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

Applying formulas (6) and (7) we find the cosine of the angle between vectors ***a*** and ***b***:

$$\boxed{\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{a\,b} = \frac{a_x b_x + a_y b_y + a_z b_z}{\sqrt{a_x^2 + a_y^2 + a_z^2}\sqrt{b_x^2 + b_y^2 + b_z^2}}}$$

The most important applications of the scalar product are related with finding the angle between vectors.

5.4.1. Properties of the Scalar Product

The below properties are based on the definition of the scalar product. They can be easily proved by the reader.

1) The scalar product is commutative:
$$\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{b}\cdot\boldsymbol{a} .$$
2) The scalar product is distributive:
$$(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{c} = \boldsymbol{a}\cdot\boldsymbol{c} + \boldsymbol{b}\cdot\boldsymbol{c} .$$
3) If the scalar product of two non-zero vectors equals zero, then the vectors are perpendicular; and *vice versa*, if two vectors are perpendicular then their scalar product equals zero:
$$\boldsymbol{a}\perp\boldsymbol{b} \quad \Leftrightarrow \quad \boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{0} .$$

5.4.2. Examples

Example 1: Let ***i*** = {1, 0, 0}, ***j*** = {0, 1, 0}, and ***k*** = {0, 0, 1} be three basis vectors in a rectangular Cartesian coordinate system. Then

$$\boldsymbol{i}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{j} = \boldsymbol{k}\cdot\boldsymbol{k} = 1,$$
$$\boldsymbol{i}\cdot\boldsymbol{j} = \boldsymbol{i}\cdot\boldsymbol{k} = \boldsymbol{j}\cdot\boldsymbol{k} = 0 .$$

Example 2: If ***a*** = {2, –1, 3}, ***b*** = {5, 7, 4} and θ is the angle between vectors ***a*** and ***b***, then

$$\boldsymbol{a}\cdot\boldsymbol{b} = 2\cdot 5 + (-1)\cdot 7 + 3\cdot 4 = 15,$$
$$a = |\boldsymbol{a}| = \sqrt{\boldsymbol{a}\cdot\boldsymbol{a}} = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14} ,$$
$$b = |\boldsymbol{b}| = \sqrt{\boldsymbol{b}\cdot\boldsymbol{b}} = \sqrt{5^2 + 7^2 + 4^2} = \sqrt{80} = 4\sqrt{5} ,$$

$$\cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}|\cdot|\boldsymbol{b}|} = \frac{15}{4\sqrt{14}\sqrt{5}} = \frac{3}{56}\sqrt{70},$$

Example 3: Find the angle between two vectors $\boldsymbol{a} = \{3, 2, -5\}$ and $\boldsymbol{b} = \{5, 7, 4\}$.
Solution:

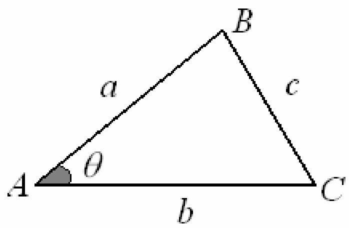
$$\boldsymbol{a}\cdot\boldsymbol{b} = 3\cdot 1 + 4\cdot 3 + (-5)\cdot 3 = 0$$

Since the scalar product is equal to zero, the vectors are orthogonal.

Example 4: Let $\boldsymbol{p} = \boldsymbol{a} + \boldsymbol{b}$ and $\boldsymbol{q} = \boldsymbol{a} - \boldsymbol{b}$. Simplify the scalar product of the vectors $\boldsymbol{p}$ and $\boldsymbol{q}$.
Solution:

$$\boldsymbol{p}\cdot\boldsymbol{q} = (\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b}) = \boldsymbol{a}^2 - \boldsymbol{a}\cdot\boldsymbol{b} + \boldsymbol{b}\cdot\boldsymbol{a} - \boldsymbol{b}^2 = a^2 - b^2.$$

Example 5: Given two sides *AB* and *AC* of the triangle *ABC* and the angle θ between these side, find the third side of the triangle.



Solution: Let us denote $\boldsymbol{a} = \overrightarrow{AB}$, $\boldsymbol{b} = \overrightarrow{AC}$ and $\boldsymbol{c} = \overrightarrow{CB}$. Then

$$\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b} \quad\Rightarrow\quad \boldsymbol{c}^2 = (\boldsymbol{a}-\boldsymbol{b})^2 = \boldsymbol{a}^2 + \boldsymbol{b}^2 - 2\boldsymbol{a}\cdot\boldsymbol{b} \quad\Rightarrow$$

$$c^2 = a^2 + b^2 - 2ab\cos\theta, \qquad c = \sqrt{c^2}.$$

5.4.3. Direction Cosines

Let α, β and γ be the angles between a unit vector **u** and the axes of a rectangular coordinate system. The cosines of these angles are called the **direction cosines**.

Theorem: In a rectangular coordinate system, the coordinates u_x, u_y, and u_z of a unit vector $\boldsymbol{u} = \{u_x, u_y, u_z\}$ are equal to the direction cosines.

The theorem follows from the definition of the scalar product. The scalar product of the unit vectors $\boldsymbol{u} = \{u_x, u_y, u_z\}$ and $\boldsymbol{i} = \{1, 0, 0\}$ can be written as

$$\boldsymbol{u}\cdot\boldsymbol{i} = u_x \quad\text{and}\quad \boldsymbol{u}\cdot\boldsymbol{i} = |\boldsymbol{u}|\cdot|\boldsymbol{i}|\cdot\cos\alpha = \cos\alpha,$$

and so $u_x = \cos\alpha$.

Likewise,

$$\boldsymbol{u}\cdot\boldsymbol{j}=u_y=\cos\beta \quad \text{and} \quad \boldsymbol{u}\cdot\boldsymbol{k}=u_z=\cos\gamma\,,$$

which required to be proved.
By the definition of a unit vector,

$$|\boldsymbol{u}|^2=u_x^2+u_y^2+u_z^2=1.$$

Therefore,

$$\cos^2\alpha+\cos^2\beta+\cos^2\gamma=1.$$

The direction cosines of an arbitrary vector ***a*** can be expressed as

$$\cos\alpha=\frac{a_x}{a}, \qquad \cos\beta=\frac{a_y}{a},\ \cos\gamma=\frac{a_z}{a}.$$

5.5. Vector Product

Given the vectors $\boldsymbol{a}=\{a_x,a_y,a_z\}$ and $\boldsymbol{b}=\{b_x,b_y,b_z\}$ in a rectangular coordinate system, the vector product $\boldsymbol{a}\times\boldsymbol{b}$ is the vector, which is defined by the formula

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}, \tag{8}$$

where ***i***, ***j*** and ***k*** are the unit vectors of the rectangular coordinate basis.
The vector product is also known as cross product. It is also denoted as $[\boldsymbol{a},\boldsymbol{b}]$.
Expanding the determinant by the first row we obtain

$$\boldsymbol{a}\times\boldsymbol{b}=(a_yb_z-a_zb_y)\boldsymbol{i}+(a_zb_x-a_xb_x)\boldsymbol{j}+(a_xb_y-a_yb_x)\boldsymbol{k}\,.$$

Theorem: Let ***a*** and ***b*** be two non-parallel vectors. Then

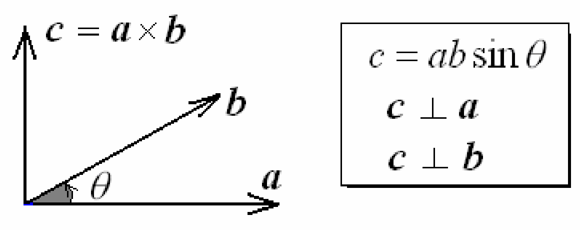
i) the vector $\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{b}$ is orthogonal to both ***a*** and ***b***;

ii) the length of ***c*** is expressed by the formula

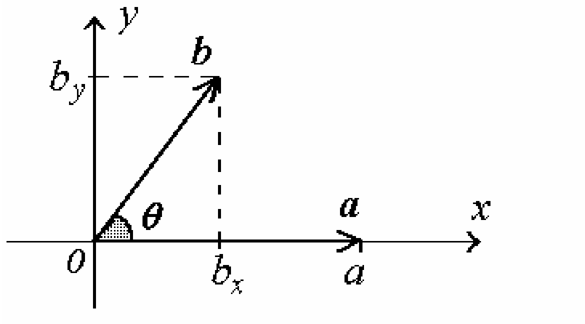
$$c=ab\sin\theta,$$

where θ is the angle between ***a*** and ***b***;

iii) the set of vectors {***a***, ***b***, ***c***} is a right-handed triplet as it is shown in the figure below.



Proof: Let the rectangular coordinate system be chosen such that both vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ lie in the x,y-plane, and the x-axis is directed along $\boldsymbol{a}$.



Then $\boldsymbol{a}=\{a, 0, 0\}$ and $\boldsymbol{b}=\{b\cos\theta, b\sin\theta, 0\}$.
Therefore,

$$\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a & 0 & 0 \\ b\cos\theta & b\sin\theta & 0 \end{vmatrix}=ab\sin\theta\,\boldsymbol{k}.$$

Therefore, $|\boldsymbol{c}|=a\sin\theta$ and $\boldsymbol{c}$ is directed along the z-axis which is perpendicular to the x,y-plane. Hence, the theorem.

5.5.1. Properties of the Vector Product

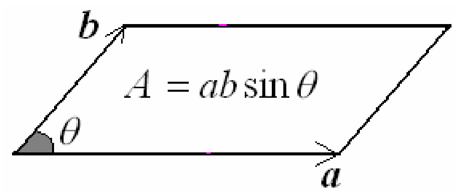
1) The vector product is anti-commutative:

$$\boldsymbol{a}\times\boldsymbol{b}=-\boldsymbol{b}\times\boldsymbol{a}.$$

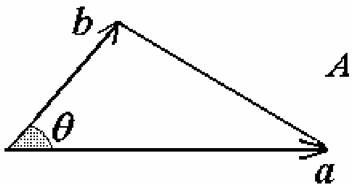
2) The vector product is distributive:

$$(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c}.$$

3) The length of the vector $\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{b}$ is equal to the area of the parallelogram with adjacent sides a and b.



Corollary: The area of the triangle with adjacent sides a and b is given by formula



$$A_{\triangle}=\frac{1}{2}|\boldsymbol{a}\times\boldsymbol{b}|$$

4) The vector product of two collinear vectors equals zero.

Properties 1) and 2) follow from the properties of determinants. Indeed,

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ a_x & a_y & a_z\\ b_x & b_y & b_z\end{vmatrix}=-\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ b_x & b_y & b_z\\ a_x & a_y & a_z\end{vmatrix}=-\boldsymbol{b}\times\boldsymbol{a},$$

$$(\boldsymbol{a}+\boldsymbol{b})\times\boldsymbol{c}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ a_x+b_x & a_y+b_y & a_z+b_z\\ c_x & c_y & c_z\end{vmatrix}$$

$$=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ a_x & a_y & a_z\\ c_x & c_y & c_z\end{vmatrix}+\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ b_x & b_y & b_z\\ c_x & c_y & c_z\end{vmatrix}=\boldsymbol{a}\times\boldsymbol{c}+\boldsymbol{b}\times\boldsymbol{c}.$$

Property 3) follows from the theorem of vector product.
Property 4) is quite evident.

5.5.2. Examples

1) Let $\boldsymbol{i}$ = {1, 0, 0}, $\boldsymbol{j}$ = {0, 1, 0}, and $\boldsymbol{k}$ = {0, 0, 1} be three basis vectors of the rectangular Cartesian coordinate system. By the definition of the vector product,

$$\boldsymbol{i}\times\boldsymbol{j}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ 1 & 0 & 0\\ 0 & 1 & 0\end{vmatrix}=\boldsymbol{k},\quad \boldsymbol{k}\times\boldsymbol{i}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ 0 & 0 & 1\\ 1 & 0 & 0\end{vmatrix}=\boldsymbol{j},\quad \boldsymbol{j}\times\boldsymbol{k}=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{vmatrix}=\boldsymbol{i}.$$

That is,

$$\boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k},\quad \boldsymbol{k}\times\boldsymbol{i}=\boldsymbol{j},\quad \boldsymbol{j}\times\boldsymbol{k}=\boldsymbol{i}.$$

2) Let $\boldsymbol{p}$ = $\boldsymbol{a}$ + $\boldsymbol{b}$ and $\boldsymbol{q}$ = $\boldsymbol{a}$ – $\boldsymbol{b}$. Simplify the vector product of the vectors $\boldsymbol{p}$ and $\boldsymbol{q}$.

Solution:

$$\boldsymbol{p}\times\boldsymbol{q}=(\boldsymbol{a}+\boldsymbol{b})\times(\boldsymbol{a}-\boldsymbol{b})$$

$$=\boldsymbol{a}\times\boldsymbol{a}-\boldsymbol{a}\times\boldsymbol{b}+\boldsymbol{b}\times\boldsymbol{a}-\boldsymbol{b}\times\boldsymbol{b}=2\boldsymbol{b}\times\boldsymbol{a}.$$

3) Let ABC be a triangle with the vertices at the points A(1, 0, –2), B(1, 5, 0) and C(0, 4, –1). Find the area A of the triangle.

Solution: Consider the vectors $\boldsymbol{a}=\overrightarrow{AB}=\{0,5,2\}$ and $\boldsymbol{b}=\overrightarrow{AC}=\{-1,4,1\}$.

By the properties of the vector product,

$$A = \frac{1}{2}|\boldsymbol{a}\times\boldsymbol{b}|.$$

Find the vector product:

$$\boldsymbol{a}\times\boldsymbol{b} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 0 & 5 & 2 \\ -1 & 4 & 1 \end{vmatrix} = -3\boldsymbol{i} - 2\boldsymbol{j} + 5\boldsymbol{k}.$$

Therefore,

$$A = \frac{1}{2}\sqrt{(-3)^2 + (-2)^2 + 5^2} = \frac{\sqrt{38}}{2}.$$

5.6. Scalar Triple Product

The scalar product and the vector product may be combined into the **scalar triple product** (or mixed product):

$$([\boldsymbol{a},\boldsymbol{b}],\boldsymbol{c}) = (\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}.$$

Theorem: Given three vectors $\boldsymbol{a} = \{a_x, a_y, a_z\}$, $\boldsymbol{b} = \{b_x, b_y, b_z\}$ and $\boldsymbol{c} = \{c_x, c_y, c_z\}$ in some rectangular coordinate system, the scalar triple product is defined by the formula

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}. \tag{9}$$

Proof: Carrying out the scalar product of the vectors

$$\boldsymbol{a}\times\boldsymbol{b} = (a_y b_z - a_z b_y)\boldsymbol{i} + (a_z b_x - a_x b_x)\boldsymbol{j} + (a_x b_y - a_y b_x)\boldsymbol{k}$$

and

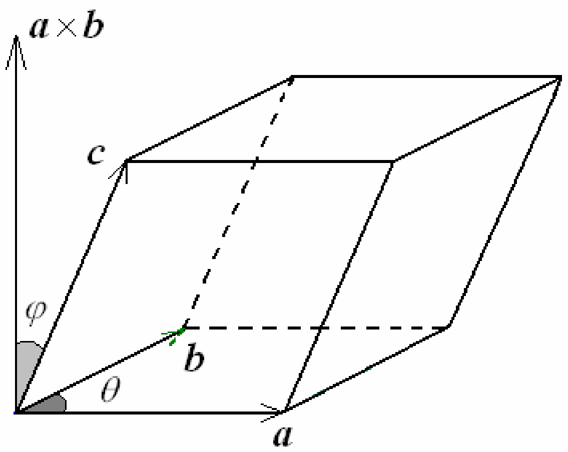
$$\boldsymbol{c} = c_x\boldsymbol{i} + c_y\boldsymbol{j} + c_z\boldsymbol{k}$$

we obtain

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot c = (a_y b_z - a_z b_y)c_x + (a_z b_x - a_x b_x)c_y + (a_x b_y - a_y b_x)c_z$$

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

Geometric Interpretation. The absolute value of the number $(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}$ is the volume of a parallelepiped formed by the vectors $\boldsymbol{a},\boldsymbol{b}$ and $\boldsymbol{c}$ as it is shown in the figure below.



Indeed, the volume of the parallelepiped is equal to the product of the area of the base and the height.
By the theorem of scalar product,

$$(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c} = |\boldsymbol{a}\times\boldsymbol{b}|\cdot|\boldsymbol{c}|\cos\varphi .$$

The quantity $|\boldsymbol{a}\times\boldsymbol{b}|$ equals the area of the parallelogram, and the product $|\boldsymbol{c}|\cos\varphi$ equals the height of the parallelepiped.

Corollary 1: If three vectors are coplanar then the scalar triple product is equal to zero.

Corollary 2: Four points *A*, *B*, *C*, and *D* lie in the same plane, if the scalar triple product $(\overrightarrow{AB}\times\overrightarrow{AC})\,\overrightarrow{AD}$ is equal to zero.

5.6.1. Properties of the Scalar Triple Product

Consider the scalar triple product $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})$.
1) By the properties of the scalar product, $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = (\boldsymbol{b}\times\boldsymbol{c})\cdot\boldsymbol{a}$.
2) In view of the properties of determinants,

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

Therefore, $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = (\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c}$.

Since the order of the dot and cross symbols is meaningless, the product $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})$ is simply denoted by ***abc***.
Using the properties of determinants it is not difficult to see that

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

Therefore,

$$\mathbf{abc} = \mathbf{cab} = \mathbf{bca}.$$

Likewise,

$$\mathbf{abc} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} = -\begin{vmatrix} b_x & b_y & b_z \\ a_x & a_y & a_z \\ c_x & c_y & c_z \end{vmatrix} = -\begin{vmatrix} a_x & a_y & a_z \\ c_x & c_y & c_z \\ b_x & b_y & b_z \end{vmatrix},$$

and so

$$\mathbf{abc} = -\,\mathbf{bac} = -\,\mathbf{acb}.$$

In view of the theorem of linear dependent vectors, any three linear dependent vectors are coplanar. Hence,

The triple product of non-zero vectors equals zero, if and only if the vectors are linear dependent.

5.6.2. Examples

1) Determine whether the points $A(-1, 2, 2)$, $B(3, 3, 4)$, $C(2, -2, 10)$, and $D(0, 2, 2)$ lie on the same plane.

Solution: Join the point A with the other points to obtain the vectors

$$\mathbf{a} = \overrightarrow{AB} = \{4, 1, 2\}, \qquad \mathbf{b} = \overrightarrow{AC} = \{3, 4, 8\}, \quad \text{and} \qquad \mathbf{c} = \overrightarrow{AD} = \{1, 0, 0\}.$$

Find the scalar triple product:

$$\mathbf{abc} = \begin{vmatrix} 4 & 1 & 2 \\ 3 & 4 & 8 \\ 1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 4 & 8 \end{vmatrix} = 0.$$

Therefore, the vectors lie in a plane, that means the given points lie in the same plane.

2) Find the volume V of the tetrahedron with the vertices at the points $A(1, 0, 2)$, $B(3, -1, 4)$, $C(1, 5, 2)$, and $D(4, 4, 4)$.

Solution: Consider a parallelepiped whose adjacent vertices are at the given points.
The volume V_p of the parallelepiped is equal to the absolute value of the triple scalar product of the vectors $\overrightarrow{AB}$, $\overrightarrow{AC}$, and $\overrightarrow{AD}$.

The volume of the tetrahedron is given by the formula $V = \frac{1}{6} V_p$.

Since

$$\vec{AB} = \{2, -1, 2\}, \quad \vec{AC} = \{0, 5, 0\}, \quad \text{and} \quad \vec{AD} = \{3, 4, 2\},$$

we obtain

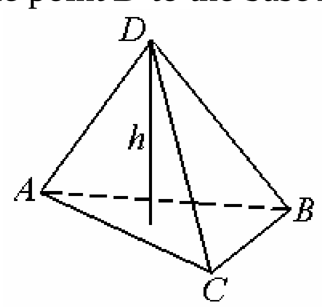
$$\vec{AB}\,\vec{AC}\,\vec{AD} = \begin{vmatrix} 2 & -1 & 2 \\ 0 & 5 & 0 \\ 3 & 4 & 2 \end{vmatrix} = 5\begin{vmatrix} 2 & 2 \\ 3 & 4 \end{vmatrix} = 10.$$

Therefore,

$$V = \frac{10}{6} = \frac{5}{3}.$$

3) The tetrahedron is given by the vertices $A(1, 0, 2)$, $B(3, -1, 4)$, $C(1, 5, 2)$, and $D(4, 4, 4)$.

Find the height from the point *D* to the base *ABC*.



Solution: In view of the formula

$$V = \frac{1}{3} S \cdot h,$$

where *h* is the height from the point *D*, we need to know the volume *V* of the tetrahedron and the area S of the base *ABC* to find *h*.

According to Example 2, the volume of the tetrahedron equals $5/3$.

The area of the triangle *ABC* can be found just in a similar way as in Example 2, section 1.5.2:

$$A = \frac{1}{2} |\vec{AB} \times \vec{AC}|,$$

$$\vec{AB} \times \vec{AC} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & -1 & 2 \\ 0 & 5 & 0 \end{vmatrix} = 10\boldsymbol{i} - 10\boldsymbol{k}, \quad |\vec{AB} \times \vec{AC}| = 10\sqrt{2}.$$

Therefore,

$$h = \frac{3V}{A} = \frac{5}{5\sqrt{2}} = \frac{\sqrt{2}}{2}.$$

5.7. Transformation of Coordinates Under Rotation of a Coordinate System

Consider a rectangular Cartesian coordinate system.
Let $\boldsymbol{e_1} = \{1, 0, 0\}$, $\boldsymbol{e_2} = \{0, 1, 0\}$ and $\boldsymbol{e_3} = \{0, 0, 1\}$ be orthogonal unit vectors of that system:

$$\boldsymbol{e_i} \cdot \boldsymbol{e_j} = \delta_{ij}, \qquad (10)$$

where δ_{ij} is the Kronecker delta.

By rotation of the coordinate system, we obtain a new rectangular coordinate system.
Let $\boldsymbol{e}'_1 = \{1, 0, 0\}$, $\boldsymbol{e}'_2 = \{0, 1, 0\}$ and $\boldsymbol{e}'_3 = \{0, 0, 1\}$ be orthogonal unit vectors of the new coordinate system, that is,

$$\boldsymbol{e}'_i \cdot \boldsymbol{e}'_j = \delta_{ij}. \qquad (11)$$

By the theorem of direction cosines, the coordinates of the vectors $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ and $\boldsymbol{e_3}$ in the basis of vectors $\boldsymbol{e}'_1$, $\boldsymbol{e}'_2$ and $\boldsymbol{e}'_3$ are the direction cosines.
Denote the direction cosines of the vector $\boldsymbol{e}_n$ (with n = 1, 2, 3) by u_{n1}, u_{n2} and u_{n3}, respectively. Then

$$\boldsymbol{e}_1 = u_{11}\boldsymbol{e}'_1 + u_{12}\boldsymbol{e}'_2 + u_{13}\boldsymbol{e}'_3,$$
$$\boldsymbol{e}_2 = u_{21}\boldsymbol{e}'_1 + u_{22}\boldsymbol{e}'_2 + u_{23}\boldsymbol{e}'_3,$$
$$\boldsymbol{e}_3 = u_{31}\boldsymbol{e}'_1 + u_{32}\boldsymbol{e}'_2 + u_{33}\boldsymbol{e}'_3,$$

or a short form

$$\boldsymbol{e}_n = \sum_{k=1}^{3} u_{nk}\,\boldsymbol{e}'_k \qquad (n = 1, 2, 3). \qquad (12)$$

`Equalities (10) and (11) imply

$$\sum_{k=1}^{3} u_{ik}u_{jk} = \sum_{k=1}^{3} u_{ki}u_{kj} = \delta_{ij} \quad (i, j = 1, 2, 3). \qquad (13)$$

Rewrite equalities (12) and (13) in the matrix form.
Let $U = \| u_{ij} \|$ be the matrix of the direction cosines. If we introduce the column matrices $E = \| \boldsymbol{e}_i \|$ and $E' = \| \boldsymbol{e}'_i \|$, then

$$E = U \cdot E', \quad U \cdot U^T = U^T \cdot U = I,$$

where U^T is the transpose of matrix U, and I is the identity matrix.
Note that the transpose of U is the inverse of U.
Therefore, we can easily obtain the formula of the inverse transformation (from the old basis to the new one):

$$E = U \cdot E' \quad \Rightarrow \quad U^{-1}E = U^{-1}U \cdot E' \quad \Rightarrow$$
$$E' = U^{-1}E = U^T E.$$

This matrix equality is equivalent to the system of three vector equalities:

$$\boldsymbol{e}'_n = \sum_{k=1}^{3} u_{kn}\boldsymbol{e}_k \qquad (n = 1, 2, 3).$$

Now consider the transformation of the coordinates of an arbitrary vector $\boldsymbol{a}$. Any vector $\boldsymbol{a}$ can be expressed as the linear combination of basis vectors. If a_1, a_2 and a_3 are the coordinates of $\boldsymbol{a}$ in an the old basis, and a'_1, a'_2 and a'_3 are the coordinates of $\boldsymbol{a}$ in the new basis, then

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_z\boldsymbol{e}_3 = \sum_{i=1}^{3} a_i\boldsymbol{e}_i \ , \tag{14}$$

$$\boldsymbol{a} = a'_1\boldsymbol{e}'_1 + a'_2\boldsymbol{e}'_2 + a'_z\boldsymbol{e}'_3 = \sum_{i=1}^{3} a'_i\boldsymbol{e}'_i \ . \tag{15}$$

Therefore, $\displaystyle\sum_{i=1}^{3} a_i\boldsymbol{e}_i = \sum_{k=1}^{3} a'_k\boldsymbol{e}'_k$.

In view of equality (12) we obtain

$$\sum_{i=1}^{3} a_i \sum_{k=1}^{3} u_{ik}\boldsymbol{e}'_k = \sum_{k=1}^{3} a'_k\boldsymbol{e}'_k \quad \Rightarrow$$

$$\sum_{k=1}^{3}\left(\sum_{i=1}^{3} u_{ik}a_i\right)\boldsymbol{e}'_k = \sum_{k=1}^{3} a'_k\boldsymbol{e}'_k$$

that results in the formulas of transformation of the coordinates:

$$a'_k = \sum_{i=1}^{3} u_{ik}a_i = u_{1k}a_1 + u_{2k}a_2 + u_{3k}a_3 . \tag{16a}$$

Likewise,

$$a_k = \sum_{i=1}^{3} u_{ki}a'_i = u_{k1}a'_1 + u_{k2}a'_2 + u_{k3}a'_3 . \tag{16b}$$

Until now we have interpreted the scalar product of vectors as a scalar quantity without a formal proof, considering this proposition as the self-evident truth. A rigorous justification follows from the below theorem.

Theorem: The scalar product is invariant under rotation of the coordinate system.

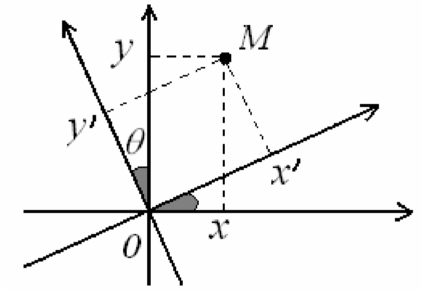
Proof: Really, using formulas (16) we obtain

$$\sum_{k=1}^{3} a'_k b'_k = \sum_{k=1}^{3}\sum_{i=1}^{3} u_{ik}a_i \sum_{j=1}^{3} u_{jk}b_j$$

$$= \sum_{i,j}\left(\sum_{k=1}^{3} u_{ik}u_{jk}\right)a_i b_j = \sum_{i,j}\delta_{ij}a_i a_j = \sum_{i=1}^{3} a_i b_i .$$

5.7.1. Rotation of the *x,y*–Plane Around the *z*-Axis

Consider a particular case of transformation of the rectangular coordinate system by rotation of the *x,y*–plane around the *x*–axis.



Let θ be the angle of the rotation, and $\boldsymbol{r}$ be the radius-vector of a point *M*. Then

$$\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} = x'\,\boldsymbol{i}' + y'\,\boldsymbol{j}'.$$

By the properties of the scalar product,

$$\boldsymbol{i}\cdot\boldsymbol{i} = 1, \quad \boldsymbol{i}\cdot\boldsymbol{j} = 0, \quad \boldsymbol{i}\cdot\boldsymbol{i}' = \cos\theta, \quad \boldsymbol{i}\cdot\boldsymbol{j}' = \cos(90^\circ + \theta) = -\sin\theta,$$

$$\boldsymbol{j}\cdot\boldsymbol{i} = 0, \quad \boldsymbol{j}\cdot\boldsymbol{j} = 1, \quad \boldsymbol{j}\cdot\boldsymbol{i}' = \cos(90^\circ - \theta) = \sin\theta, \quad \boldsymbol{j}\cdot\boldsymbol{j}' = \cos\theta.$$

Therefore, the scalar products $\boldsymbol{i}\cdot\boldsymbol{r}$ and $\boldsymbol{j}\cdot\boldsymbol{r}$ can be expressed, respectively, as

$$x = x'\cos\theta - y'\sin\theta \qquad (17a)$$

and

$$y = x'\sin\theta + y'\cos\theta. \qquad (17b)$$

Likewise, formulas of the inverse transformation follow from the scalar products $\boldsymbol{i}'\cdot\boldsymbol{r}$ and $\boldsymbol{j}'\cdot\boldsymbol{r}$:

$$x' = x\cos\theta + y\sin\theta, \qquad (18a)$$

$$y' = -x\sin\theta + y\cos\theta. \qquad (18b)$$

Formulas (17) – (18) are particular cases of general formulas (16).

Example: Let $\theta = 45^\circ$. Express the quadratic form $x'y'$ in the old coordinate system.

Solution: Apply formulas (18), taking into account that $\sin\theta = \cos\theta = \dfrac{\sqrt{2}}{2}$:

$$x' = \frac{\sqrt{2}}{2}(x + y), \qquad y' = \frac{\sqrt{2}}{2}(-x + y),$$

$$x'y' = \frac{\sqrt{2}}{2}\frac{\sqrt{2}}{2}(y + x)(y - x) = \frac{1}{2}(y^2 - x^2).$$

ANALYTICAL GEOMETRY

6. Straight Lines

6.1. Equations of Lines

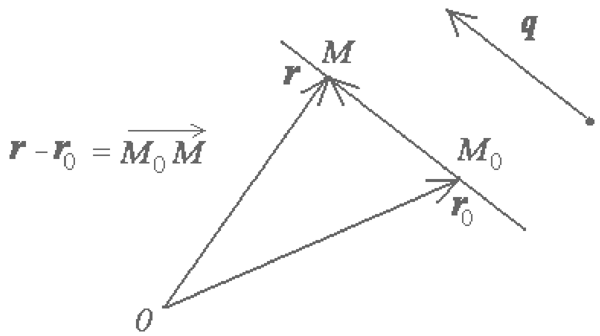
A direction vector of a straight line is a vector parallel to the line.
According to the postulates of geometry, a point M_0 and a direction vector $\boldsymbol{q}$ determine the straight line L.
Let M be an arbitrary point on the line. The difference $\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}_0$ between the radius-vectors of the points M and M_0 is a vector in the line, that is,

$$\boldsymbol{r} - \boldsymbol{r}_0 \parallel \boldsymbol{q}.$$

Two parallel vectors are proportional:

$$\boldsymbol{r} - \boldsymbol{r}_0 = t\boldsymbol{q} \qquad (1)$$

This vector equality is called the **vector equation of the line**. An arbitrary number t is said to be a parameter.



Assume that a rectangular Cartesian coordinate system is chosen. Then equation (1) can be written in the coordinate form as the system of three linear equations

$$\begin{cases} x = x_0 + q_x t \\ y = y_0 + q_y t \\ z = z_0 + q_z t \end{cases} \qquad (2)$$

where x, y and z are running coordinates of a point on the line. Vectors $\boldsymbol{r}$, $\boldsymbol{r}_0$ and $\boldsymbol{q}$ are represented by their coordinates:

$$\boldsymbol{r} - \boldsymbol{r}_0 = \{x - x_0, y - y_0, z - z_0\},$$

$$\boldsymbol{q} = \{q_x, q_y, q_z\}.$$

Equations of a line in coordinate form (2) are called the **parametric equations** of a line.

Solving system (2) by elimination of the parameter *t*, we obtain the **canonical equations of a line**:

$$\frac{x-x_0}{q_x}=\frac{y-y_0}{q_y}=\frac{z-z_0}{q_z}. \qquad (3)$$

If $M_0(x_0, y_0, z_0)$ and $M_1(x_1, y_1, z_1)$ are two given points on a line then the vector

$$\boldsymbol{q}=\{x_1-x_0, y_1-y_0, z_1-z_0\}$$

joining these points serves as a direction vector of the line.

Therefore, we get the following equations of a line passing through two given points:

$$\frac{x-x_0}{x_1-x_0}=\frac{y-y_0}{y_1-y_0}=\frac{z-z_0}{z_1-z_0} \qquad (4)$$

Examples:

1) Let L be a line passing through the points $M_1(1,0,2)$ and $M_2(3,1,-2)$.

   Check whether the point $A(7,3,10)$ lie on the line L.

Solution: Using (4) we get the equations of L:

$$\frac{x-1}{2}=\frac{y}{1}=\frac{z-2}{-4}.$$

The coordinates of the point A satisfy the equation:

$$\frac{7-1}{2}=\frac{3}{1}=\frac{-10-2}{-4},$$

and so A is a point of the line L.

2) Write down the canonical equations of the line passing through the point $A(2,3,4)$ and being parallel to the vector $\boldsymbol{q}=\{5,0,-1\}$.

Solution: By equation (3), we obtain

$$\frac{x-2}{5}=\frac{y-3}{0}=\frac{z-4}{-1}.$$

Note that a symbolical notation $\dfrac{y-3}{0}$ means the equation $y=3$.

6.2. Lines in a Plane

On the x, y–plane, a line is described by the linear equation

$$Ax + By + C = 0\,. \qquad (5)$$

If $M_0(x_0, y_0)$ is a point on the line then

$$Ax_0 + By_0 + C = 0\,. \qquad (6)$$

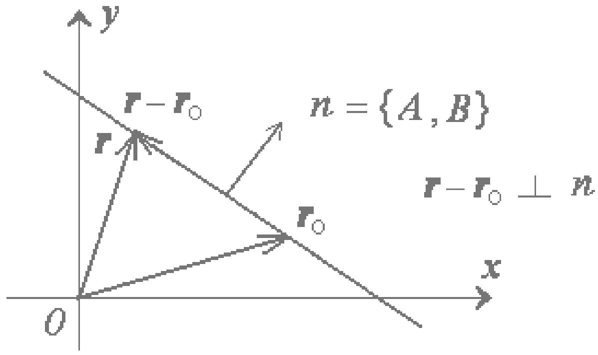
Subtracting identity (6) from equation (5) we obtain the equation of a line passing through the point $M_0(x_0, y_0)$:

$$A(x - x_0) + B(y - y_0) = 0\,. \qquad \text{(6a)}$$

The expression on the left hand side has a form of the scalar product of the vectors $\boldsymbol{n} = \{A, B\}$ and $\boldsymbol{r} - \boldsymbol{r}_0 = \{x - x_0, y - y_0\}$:

$$\boldsymbol{n} \cdot (\boldsymbol{r} - \boldsymbol{r}_0) = 0\,.$$

Therefore, the coefficients A and B can be interpreted geometrically as the coordinates of a vector in the x, y–plane, being perpendicular to the line.



The canonical equation of a line in the x, y–plane has a form

$$\frac{x - x_0}{q_x} = \frac{y - y_0}{q_y}\,,$$

where $\boldsymbol{q} = \{q_x, q_y\}$ is a direction vector of the line.

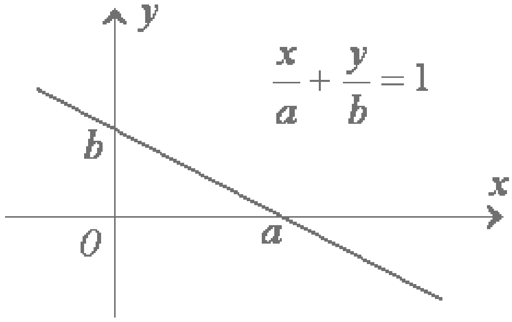
In the x, y–plane, an equation of a line passing through two given points, $M_0(x_0, y_0)$ and $M_1(x_1, y_1)$, is written as follows

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0}\,.$$

Sometimes it is helpful to express a straight-line equation in the x, y–plane as

$$\frac{x}{a} + \frac{y}{b} = 1\,. \qquad (7)$$

In this case, $y = 0$ implies $x = a$, and $x = 0$ implies $y = b$.



Therefore, the quantities *a* and *b* are, respectively, the *x*-intercept and the *y*-intercept of a graph of the line. Equation (7) is called an equation of a line in the intercept form.

A line on the *x*,*y*–plane may be also given by the equation in the slope-intercept form

$$y = kx + b,$$

where *b* is the *y*-intercept of a graph of the line, and *k* is the slope of the line.

If $M_0(x_0, y_0)$ is a point on the line, that is, $y_0 = kx_0 + b$, then the point–slope equation of the line is

$$y - y_0 = k(x - x_0).$$

Examples:

1) A line on the *x, y*–plane is given by the equation

$$2x - 3y + 24 = 0.$$

Find: (i) any two points on the line; (ii) the slope of the line; (iii) the *x*– and *y*–intercepts.

Solution:

(i) Setting *x* = 0 we obtain *y* = 8.

If *x* = 3 then *y* = 10.

Therefore, the points *P*(0, 8) and *Q*(3, 10) lie on the line.

(ii) $2x - 3y + 24 = 0 \quad \Rightarrow \quad y = \frac{2}{3}x + 8,$

Therefore, the slope of the line is $k = 2/3$.

(iii) The *y*–intercept equals 8. The *x*–intercept is the solution of the equation $y = 0$, that is, *x* = –12.

2) In the *x*, *y*–plane, find the equation of the line passing through the point $M_1(5, 3)$ and being perpendicular to the vector $\boldsymbol{N} = \{2, -1\}$.

Solution: Using equation (6a) we obtain

$$2(x - 5) - (y - 3) = 0 \qquad \Rightarrow \qquad y = 2x - 7.$$

3) Let $M_1(-2, 4)$ and $M_2(1, 6)$ be the points on a line.
Which of the following points, $A(-3, 1)$, $B(0, 3)$ and $C(3, -6)$, are the points on the line?

Solution: In view of the equation of a line passing through two given points, we have

$$\frac{x+2}{1+2} = \frac{y-4}{6-4} \quad \Rightarrow \quad \frac{x+2}{3} = \frac{y-4}{2} \quad \Rightarrow$$
$$2x - 3y + 12 = 0.$$

Substituting the coordinates of the points we obtain that

$A(-3, 1)$ is not a point on the line, since
$$2 \cdot (-3) - 3 \cdot 1 + 12 = 3 \neq 0;$$
$B(0, 3)$ is not a point on the line, since
$$2 \cdot 0 + 3 \cdot 3 + 12 = 21 \neq 0;$$
$C(3, -6)$ is a point on the line, since
$$2 \cdot 3 - 3 \cdot 6 + 12 = 0 \equiv 0.$$

6.3. Angle Between Two Lines

The angle between two lines is the angle between direction vectors of the lines.

If $\boldsymbol{p} = \{p_x, p_y, p_z\}$ and $\boldsymbol{q} = \{q_x, q_y, q_z\}$ are direction vectors of lines, then the cosine of the angle between the lines is given by the following formula:

$$\cos\theta = \frac{\boldsymbol{p} \cdot \boldsymbol{q}}{|\boldsymbol{p}| \cdot |\boldsymbol{q}|} = \frac{p_x q_x + p_y q_y + p_z q_z}{\sqrt{p_x^2 + p_y^2 + p_z^2}\sqrt{q_x^2 + q_y^2 + q_z^2}}.$$

If two lines are perpendicular to each other then their direction vectors are also perpendicular. This means that the scalar product of the direction vectors is equal to zero:

$$\boldsymbol{p} \cdot \boldsymbol{q} = p_x q_x + p_y q_y + p_z q_z = 0.$$

If two lines are parallel then their direction vectors are proportional:

$$\boldsymbol{p} = c\boldsymbol{q},$$

where c is a number.

In the coordinate form, this condition looks like

$$\frac{p_x}{q_x} = \frac{p_y}{q_y} = \frac{p_z}{q_z}.$$

We need direction vectors of lines to find the angle between the lines. Consider a few particular cases.

1) Let a line be given by two points $M_1(x_1, y_1, z_1)$ and $M_2(x_2, y_2, z_2)$. Then
$$\boldsymbol{p} = \{x_2 - x_1,\, y_2 - y_1,\, z_2 - z_1\}$$
is a direction vector of the line.

2) If a line in the *x*, *y*–plane is given by the equation
$$Ax + By + C = 0,$$
then we can easily find two points on the line. For instance, $M_1(0, -C/B)$ and $M_2(-C/A, 0)$ are two points on the line.
If two lines in the *x*, *y*–plane are given by the equations
$$A_1x + B_1y + C_1 = 0 \quad \text{and} \quad A_2x + B_2y + C_2 = 0$$
then the angle between the lines is equal to the angle between perpendicular vectors $\boldsymbol{n_1} = \{A_1, B_1\}$ and $\boldsymbol{n_2} = \{A_2, B_2\}$ to the lines:
$$\cos\theta = \frac{\boldsymbol{n_1} \cdot \boldsymbol{n_2}}{|\boldsymbol{n_1}| \cdot |\boldsymbol{n_2}|}.$$
Note that a perpendicular vector to a line is also called a **normal vector** to the line.

3) If a line in the *x*, *y*–plane is given by the equation
$$\frac{x}{a} + \frac{y}{b} = 1,$$
then $M_1(0, b)$ and $M_2(a, 0)$ are two points on the line, and so $\boldsymbol{p} = \{a, -b\}$ is a direction vector of the line.

4) If two lines in the *x*, *y*–plane are given by the equations in the slope-intercept form
$$y = k_1x + b_1 \quad \text{and} \quad y = k_2x + b_2,$$
and θ is the angle between the lines, then
$$\tan\theta = \frac{k_2 - k_1}{1 + k_1k_2}.$$
The lines are parallel, if
$$k_1 = k_2.$$
The lines are perpendicular, if
$$k_1k_2 = -1.$$

Examples:

1) Find the angle θ between two lines in the *x*, *y*–plane, if they are given by the following equations:
$$3x-4y+1=0 \qquad \text{and} \qquad 2x+y-5=0\,.$$

Solution: Normal vectors to the lines are, respectively, $\boldsymbol{n_1}=\{3,-4\}$ and $\boldsymbol{n_2}=\{2,1\}$. Therefore,
$$\cos\theta=\frac{\boldsymbol{n_1}\cdot\boldsymbol{n_2}}{|\boldsymbol{n_1}|\cdot|\boldsymbol{n_2}|}$$
$$=\frac{3\cdot4+(-4)\cdot1}{\sqrt{3^2+(-4)^2}\sqrt{2^2+1^2}}=\frac{8}{5\sqrt{5}}=\frac{8}{25}\sqrt{5}.$$

2) Find the angle θ between two lines in the *x*, *y*–plane, if they are given by the equations in the slope-intercept form:
$$y=-\sqrt{3}x+1 \quad \text{and} \quad y=\frac{\sqrt{3}}{3}x+5\,.$$

Solution: We have $k_1=-\sqrt{3}$ and $k_2=\sqrt{3}/3$.
Since
$$k_1k_2=-\sqrt{3}\sqrt{3}/3=-1\,,$$
the lines are orthogonal: $\theta=\dfrac{\pi}{2}$.

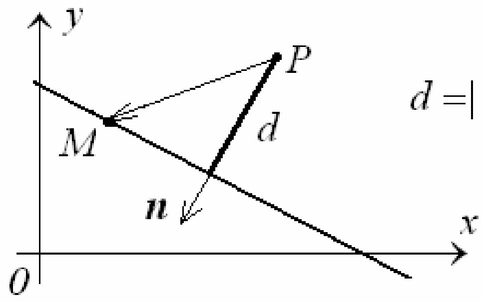
3) Let *A* = {2, –1}, *B* = {4, 4} and *C* = {9, 7} be the vertices of a triangle. Find the equation of the altitude from the vertex *A*, and write down the equation in the intercept form.

Solution: If *D* = {*x*, *y*} is an arbitrary point on the altitude from the vertex *A*, then the vectors $\overrightarrow{AD}=\{x-2,\,y+1\}$ and $\overrightarrow{BC}=\{5,3\}$ are orthogonal. Therefore, the scalar product of $\overrightarrow{AD}$ and $\overrightarrow{BC}$ is equal to zero, and we obtain the desired equation:
$$\overrightarrow{AD}\cdot\overrightarrow{BC}=5(x-2)+3(y+1)=0 \qquad \Rightarrow$$
$$5x+3y-7=0 \qquad \Rightarrow$$
$$\frac{x}{7/5}+\frac{y}{7/3}=1\,.$$

6.4. Distance From a Point to a Line

Consider a line in the *x*, *y*–plane.

Let $\boldsymbol{n}$ be a normal vector to the line and $M(x_0, y_0)$ be any point on the line. Then the distance *d* from a point *P* not on the line is equal to the absolute value of the projection of $\overrightarrow{PM}$ on $\boldsymbol{n}$:



$$d = |\mathrm{Proj}_{\boldsymbol{n}} \overrightarrow{PM}| = \left| \frac{\overrightarrow{PM} \cdot \boldsymbol{n}}{|\boldsymbol{n}|} \right|$$

In particular, if the line is given by the equation

$$Ax + By + C = 0,$$

and the coordinates of the point *P* are x_1 and y_1, that is,

$$\boldsymbol{n} = \{A, B\} \quad \text{and} \quad \overrightarrow{PM} = \{x_1 - x_0,\, y_1 - y_0\},$$

then the distance from the point $P(x_1, y_1)$ to the line is calculated according to the following formula:

$$d = \frac{|A(x_1 - x_0) + B(y_1 - y_0)|}{\sqrt{A^2 + B^2}}.$$

Since $M(x_0, y_0)$ is a point on the line,

$$Ax_0 + By_0 + C = 0.$$

Therefore, we obtain

$$d = \frac{|Ax_1 + By_1 + C|}{\sqrt{A^2 + B^2}}.$$

Example: Let *ABC* be a triangle in *x*, *y*–plane with the vertices at the points $A = \{2, -1\}$, $B = \{4, 4\}$ and $C = \{9, 7\}$.

Find the altitude from the vertex *A*.

Solution: The altitude from the vertex *A* equals the distance *d* from the point *A* to the line passing through the points *B* and *C*.

Find the equation of the line *BC*:

$$\frac{x-2}{2} = \frac{y+1}{5} \qquad \Rightarrow \qquad 5x - 2y - 12 = 0.$$

Therefore, a normal vector to the line *BC* is $\boldsymbol{n} = \{5, -2\}$.

Since $\overrightarrow{AC} = \{7, 8\}$, we finally obtain

$$d = \left| \frac{\overrightarrow{AC} \cdot \boldsymbol{n}}{|\boldsymbol{n}|} \right| = \frac{7 \cdot 5 + 8 \cdot (-2)}{\sqrt{5^2 + (-2)^2}} = \frac{19}{\sqrt{29}} = \frac{19}{29}\sqrt{29}.$$

6.5. Relative Position of Lines

Let two lines, L_1 and L_2, be given by their equations, e.g., in the canonical form:

$$L_1: \qquad \frac{x - x_1}{p_x} = \frac{y - y_1}{p_y} = \frac{z - z_1}{p_z},$$

$$L_2: \qquad \frac{x - x_2}{q_x} = \frac{y - y_2}{q_y} = \frac{z - z_2}{q_z},$$

where $\{p_x, p_y, p_z\} = \boldsymbol{p}$ and $\{q_x, q_y, q_z\} = \boldsymbol{q}$ are direction vectors of the lines.

In order to determine the relative position of the lines, it is necessary to consider the equations of both lines as a system of linear equations. Each lines is described by two linear equations, and so we have the following system of four linear equations with three unknowns *x*, *y* and *z*:

$$\begin{cases} (x - x_1)/p_x = (y - y_1)/p_y \\ (x - x_1)/p_x = (z - z_1)/p_z \\ (x - x_2)/q_x = (y - y_2)/q_y \\ (x - x_2)/q_x = (z - z_2)/q_z \end{cases} \qquad (1)$$

Let us analyze all possible cases.

1) Assume that system (1) is inconsistent. Then the lines are either parallel or skew. If the coordinates of the direction vectors $\boldsymbol{p}$ and $\boldsymbol{q}$ are proportional, that is,

$$\frac{p_x}{q_x} = \frac{p_y}{q_y} = \frac{p_z}{q_z}$$

then the lines are parallel; otherwise, they are skew.

2) Suppose that system (1) is consistent, and the rank of the coefficient matrix equals 3. Then L_1 and L_2 are intersecting lines, that is, they have exactly one point of intersection.

3) If system (1) is consistent, and the rank of the coefficient matrix equals 2, then the lines coincide with each other.

7. Planes

7.1. General Equation of a Plane

A **normal vector** to a plane is a perpendicular vector to the plane.
According to geometrical postulates,

- A point and a vector determine a plane.
- Three points determine a plane.

The general equation of a plane in a rectangular Cartesian coordinate system has the following form:

$$Ax + By + Cz + D = 0, \qquad (1)$$

where *x*, *y* and *z* are running coordinates of a point in the plane.
Let $M_1(x_1, y_1, z_1)$ be a point in the plane, that is,

$$Ax_1 + By_1 + Cz_1 + D = 0. \qquad (2)$$

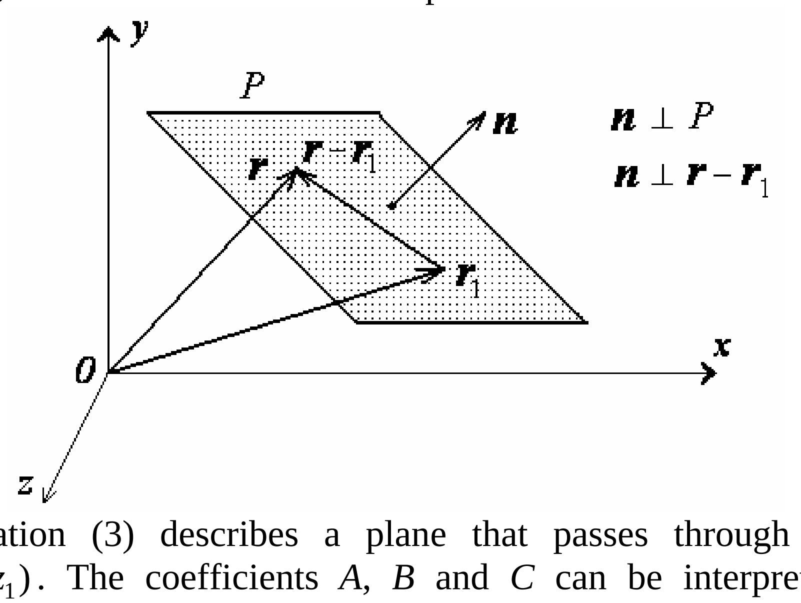
Subtracting identity (2) from equation (1) we obtain another form of the general equation of a plane:

$$A(x - x_1) + B(y - y_1) + C(z - z_1) = 0. \qquad (3)$$

Assume that *A*, *B* and *C* are the coordinates of some vector $\boldsymbol{n}$.
Then the left hand side of equation (3) is the scalar product of the vectors $\boldsymbol{n}$ and $\boldsymbol{r} - \boldsymbol{r}_1 = \{x - x_1, y - y_1, z - z_1\}$:

$$(\boldsymbol{r} - \boldsymbol{r}_1) \cdot \boldsymbol{n} = 0. \qquad (3a)$$

By the properties of the scalar product this equality implies that the vector $\boldsymbol{n}$ is perpendicular to the vector $\boldsymbol{r} - \boldsymbol{r}_1$. Since $\boldsymbol{r} - \boldsymbol{r}_1$ is an arbitrary vector in the plane *P*, $\boldsymbol{n}$ is a normal vector to the plane *P*.



Thus, equation (3) describes a plane that passes through the point $M_1(x_1, y_1, z_1)$. The coefficients *A*, *B* and *C* can be interpreted as the coordinates of a normal vector to the plane.
Consider a few particular cases of equation (1).

1) If $D = 0$ then the plane
$$Ax + By + Cz = 0$$
passes through the origin.
2) If $C = 0$ then the plane
$$Ax + By + D = 0$$
is parallel to the z–axis, that is, it extends along the x-axis.
3) If $B = 0$ then the plane
$$Ax + Cz + D = 0$$
is parallel to the y–axis.
4) If $A = 0$ then the plane
$$By + Cz + D = 0$$
is parallel to the x–axis.
5) If $A = B = 0$ then the plane
$$Cz + D = 0$$
is parallel to the x, y–plane, that is, the plane is perpendicular to the z–axis.

Examples:

1) Let $M_1(1, -2, 3)$ be a point in a plane, and $\boldsymbol{n} = \{4, 5, -6\}$ be a normal vector to the plane. Then the plane is described by the following equation
$$4(x-1) + 5(y+2) - 6(z-3) = 0 \quad \Rightarrow \quad 4x + 5y - 6z + 24 = 0 .$$

2) A plane is given by the equation
$$x - 2y + 3z - 6 = 0 .$$
Find a unit normal vector $\boldsymbol{u}$ to the plane and find any two points in the plane.

Solution: Since $\boldsymbol{n}$ = {1, –2, 3} and $|\boldsymbol{n}| = \sqrt{1^2 + (-2)^2 + 3^2} = \sqrt{14}$, then
$$\boldsymbol{u} = \frac{\boldsymbol{n}}{|\boldsymbol{n}|} = \frac{1}{\sqrt{14}}(\boldsymbol{i} - 2\boldsymbol{j} + 3\boldsymbol{k}) .$$
Setting $x = y = 0$, we obtain $z = 2$.
Likewise, if $x = z = 0$, then $y = -3$.
Therefore, $M_1(0, 0, 2)$ and $M_2(0, -3, 0)$ are the points in the given plane.

7.2. Equation of a Plane Passing Through Three Points

Let $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ and $M_3(x_3, y_3, z_3)$ be three given points in a plane *P*, and $M(x, y, z)$ be an arbitrary point in *P*.
Consider three vectors,

$$\overrightarrow{M_1M} = \boldsymbol{r} - \boldsymbol{r}_1 = \{x - x_1, y - y_1, z - z_1\},$$

$$\overrightarrow{M_1M_2} = \boldsymbol{r}_2 - \boldsymbol{r}_1 = \{x_2 - x_1, y_2 - y_1, z_2 - z_1\}$$

and

$$\overrightarrow{M_1M_3} = \boldsymbol{r}_3 - \boldsymbol{r}_1 = \{x_3 - x_1, y_3 - y_1, z_3 - z_1\}.$$

They all lie in the plane *P*, and so their scalar triple product is equal to zero:

$$(\boldsymbol{r} - \boldsymbol{r}_1)(\boldsymbol{r}_2 - \boldsymbol{r}_1)(\boldsymbol{r}_3 - \boldsymbol{r}_1) = 0 \qquad \Rightarrow$$

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \qquad (4)$$

Equation (4) describes a plane passing through three given points.

Example: Let $M_1(2, 5, -1)$, $M_2(2, -3, 3)$ and $M_3(4, 5, 0)$ be points in a plane.
Find an equation of that plane.

Solution: By equation (4), we have

$$\begin{vmatrix} x-2 & y-5 & z+1 \\ 0 & -8 & 4 \\ 2 & 0 & 1 \end{vmatrix} = 0 \quad \Rightarrow$$

$$\begin{vmatrix} x-2 & y-5 & z+1 \\ 0 & 2 & -1 \\ 2 & 0 & 1 \end{vmatrix} = 0 \quad \Rightarrow$$

$$2(x-2) - 2(y-5) - 4(z+1) = 0 \qquad \Rightarrow$$

$$2x - 2y - 4z + 2 = 0 \qquad \Rightarrow$$

$$x - y - 2z + 1 = 0.$$

7.3. Other Forms of Equations of a Plane

1) Let $\boldsymbol{p} = \{p_x, p_y, p_z\}$ and $\boldsymbol{q} = \{q_x, q_y, q_z\}$ be two vectors that are parallel to a plane *P*, and $M_1(x_1, y_1, z_1)$ be a point in *P*.

If $\boldsymbol{r} = \{x, y, z\}$ is the radius-vector of an arbitrary point in the plane *P*, then three vectors, $\boldsymbol{r} - \boldsymbol{r}_1 = \{x - x_1, y - y_1, z - z_1\}$, $\boldsymbol{p}$ and $\boldsymbol{q}$, are coplanar, and so the scalar triple product is equal to zero:

$$(\boldsymbol{r} - \boldsymbol{r}_1)\boldsymbol{p}\boldsymbol{q} = 0.$$

This equality expresses an equation of a plane in the vector form. It can also be written in the coordinate form as follows:

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ p_x & p_y & p_z \\ q_x & q_y & q_z \end{vmatrix} = 0. \qquad (5)$$

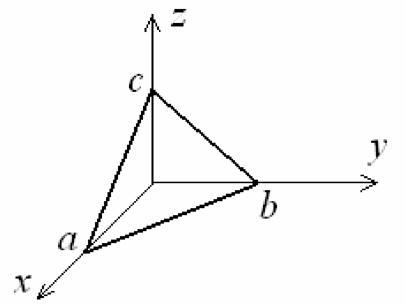
2) Assume that the general equation of a plane is expressed in the form of the following equality:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \qquad (6)$$

Then

$$\begin{aligned} y = z = 0 \quad &\Rightarrow \quad x = a, \\ x = z = 0 \quad &\Rightarrow \quad y = b, \\ x = y = 0 \quad &\Rightarrow \quad z = c. \end{aligned}$$

Therefore, the quantities *a*, *b* and *c* are, respectively, the *x*-intercept, *y*-intercept and *z*-intercept of the plane.



Equation (6) is called the equation of a plane in the intercept form.

For instance, the equation

$$\frac{x}{2} + \frac{y}{-5} + \frac{z}{4} = 1$$

describes the plane with the *x*–, *y*–, *z*–intercepts equal 2, –5 and 4, respectively.

7.4. Angle Between Two Planes

The angle θ between two planes equals the angle between their normal vectors $\boldsymbol{n}$ and $\boldsymbol{m}$:

$$\cos\theta = \frac{\boldsymbol{n}\cdot\boldsymbol{m}}{|\boldsymbol{n}|\cdot|\boldsymbol{m}|}.$$

If the planes are given by equations in the general form

$$A_1x + B_1y + C_1 + D_1 = 0,$$

$$A_2x + B_2y + C_2 + D_2 = 0,$$

then

$$\cos\theta = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2}\sqrt{A_2^2 + B_2^2 + C_2^2}}. \qquad (7)$$

If two planes are perpendicular to each other then their normal vectors are also perpendicular:

$$\boldsymbol{n}\cdot\boldsymbol{m} = A_1A_2 + B_1B_2 + C_1C_2 = 0.$$

If two planes are parallel then the normal vectors are proportional:

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

Note that the vector product of two non-parallel vectors in a plane gives a normal vector to the plane. In particular, if a plane is given by three points $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ and $M_2(x_3, y_3, z_3)$, then a normal vector to the plane is

$$\boldsymbol{n} = \overrightarrow{M_1M_2}\times\overrightarrow{M_1M_3} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix}. \qquad (8)$$

Example: Find the angle between two planes P_1 and P_2, if P_1 passes through the points $M_1(-2, 2, 2)$, $M_2(0, 5, 3)$ and $M_3(-2, 3, 4)$, and P_2 is given by the equation

$$3x - 4y + z + 5 = 0,$$

Solution: A normal vector to the plane P_2 is determined by

$$\boldsymbol{m} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 5\boldsymbol{i} - 4\boldsymbol{j} + 2\boldsymbol{k}.$$

A normal vector to the plane P_1 is $\boldsymbol{n} = \{3, -4, 1\}$.

Therefore, the cosine of the angle between the given planes is

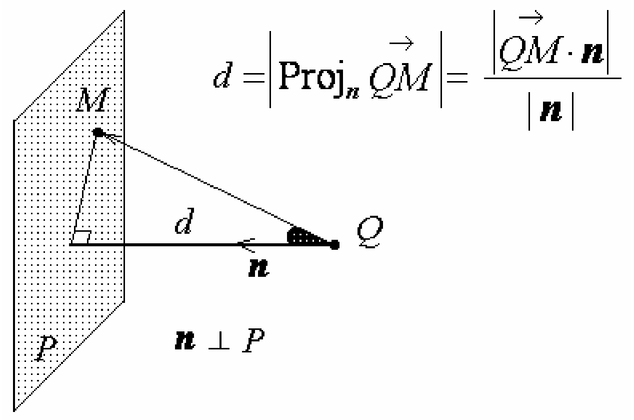
$$\cos\theta = \frac{3\cdot 5+(-4)\cdot(-4)+1\cdot 2}{\sqrt{3^2+(-4)^2+1^2}\sqrt{5^2+(-4)^2+2^2}} = \frac{11}{\sqrt{130}}.$$

7.5. Distance From a Point To a Plane

Assume that a plane *P* is determined by the equation in the general form:

$$Ax+By+Cz+D=0. \qquad (9)$$

Let $Q(x_1, y_1, z_1)$ be a given point not in the plane, and $M(x,y,z)$ be an arbitrary point in *P*. Then the distance *d* between the point *Q* and the plane *P* is equal to the absolute value of the projection of $\overrightarrow{QM}$ on $\boldsymbol{n}=\{A,B,C\}$.



Therefore,

$$d = \left|\frac{A(x-x_1)+B(y-y_1)+C(z-z_1)}{\sqrt{A^2+B^2+C^2}}\right|.$$

By equality (9),

$$Ax+By+Cz=-D,$$

and so the distance between point $Q(x_1, y_1, z_1)$ and plane (9) is given by the following formula:

$$d = \left|\frac{Ax_1+By_1+Cz_1+D}{\sqrt{A^2+B^2+C^2}}\right|. \qquad (9)$$

Example: Let the plane be given by the equation

$$2x+3y-4z+5=0.$$

The distance from the point $Q(8,-7,1)$ to the plane is

$$d = \left|\frac{2\cdot 8+3\cdot(-7)-4\cdot 1+5}{\sqrt{2^2+3^2+(-4)^2}}\right| = \frac{4}{\sqrt{29}} = \frac{4}{29}\sqrt{29}.$$

7.6. Relative Position of Planes

Let two planes, P_1 and P_2, be given by their general equations

$$P_1: \qquad A_1x + B_1y + C_1z + D_1 = 0,$$

$$P_1: \qquad A_2x + B_2y + C_2z + D_2 = 0.$$

Consider the system of two linear equations

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \tag{10}$$

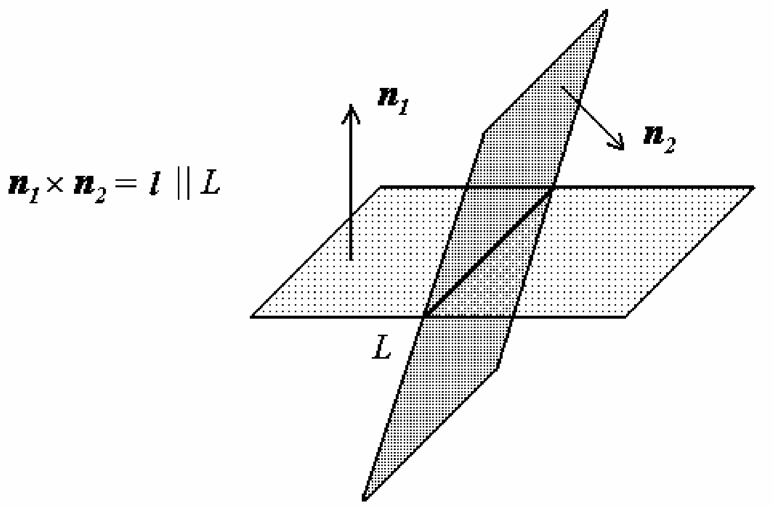
1) If system (10) is inconsistent, then the planes are parallel, and so the coordinates of the normal vectors $\boldsymbol{n_1} = \{A_1, B_1, C_1\}$ and $\boldsymbol{n_2} = \{A_2, B_2, C_2\}$ are proportional:

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

2) If system (10) is consistent and the equations are proportional to each other, then P_1 is just the same plane as P_2:

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

3) If system (10) is consistent, and the rank of the coefficient matrix equals 2, then P_1 and P_2 are intersecting planes. The locus of these distinct intersecting planes is exactly one line L. The vector product of normal vectors to the planes P_1 and P_2 is the vector, which is perpendicular to the normal vectors, and so it lies in both planes. Therefore, $\boldsymbol{n_1} \times \boldsymbol{n_2}$ is a direction vector $\boldsymbol{l}$ of the intersection line L:



In a similar way we can consider the relative position of any number of planes. The only difference is the number of possible cases.

7.7. Relative Position of a Plane and a Line

Let a plane P be given by the equation in the general form

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

and a line L be determined by the system of two linear equations

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

To investigate the relative positions of the line and the plane, consider the integrated system of equations:

$$\begin{cases} Ax + By + Cz + D = 0, \\ A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \qquad (11)$$

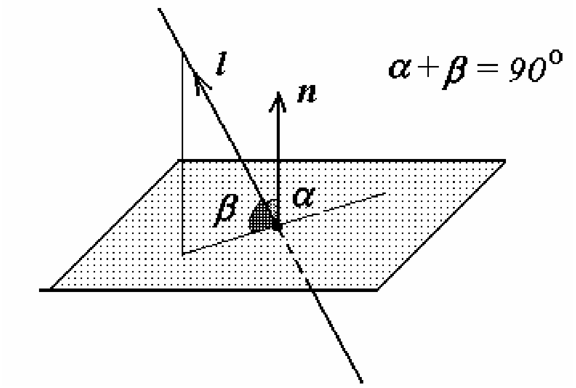
There are three possible cases.

1) If the rank of the coefficient matrix equals 3, then the system is consistent and has a unique solution $\{x_0, y_0, z_0\}$. It means that $M_0(x_0, y_0, z_0)$ is the point of intersection of the plane and the line.
2) If system (11) is consistent, and the rank of the coefficient matrix equals 2, then the line L lies in the plane P.
3) If system (11) is inconsistent then the line L is parallel to the plane P.

7.8. The Angle Between a Plane and a Line

Let α be the angle between a normal vectors $\boldsymbol{n}$ to a plane and a direction vector $\boldsymbol{l}$ of a line, and β be the angle between the plane and the line.

Then α and β are complementary angles shown in the figure below.



Therefore,

$$\sin\beta = \cos\alpha = \frac{\boldsymbol{n} \cdot \boldsymbol{l}}{|\boldsymbol{n}| \cdot |\boldsymbol{l}|}.$$

8. Quadratic Curves

8.1. Circles

A **circle** is a set of points in a plane that are equidistant from a fixed point. The fixed point is called the **center**. A line segment that joins the center with any point of the circle is called the **radius**.

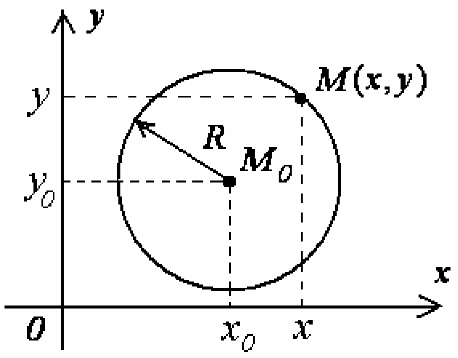
In the *x*,*y*–plane, the distance between two points $M(x, y)$ and $M_0(x_0, y_0)$ equals

$$\sqrt{(x-x_0)^2+(y-y_0)^2},$$

and so the circle is described by the equation

$$(x-x_0)^2+(y-y_0)^2=R^2, \qquad (1)$$

where x_0 and y_0 are the coordinates of the center, and R is the radius.



Equation of a circle centered at the origin

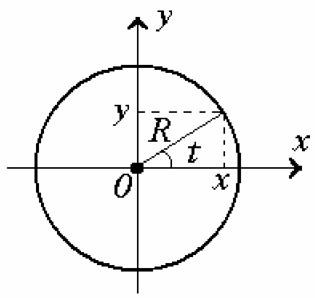
$$x^2+y^2=R^2 \qquad (2)$$

is known as the canonical equation of the circle.

If t is a real parameter, then

$$\begin{cases} x = R\cos t \\ y = R\sin t \end{cases}$$

are the parametric equations of the circle centered at the origin with radius R.



By elimination of the parameter t, we return to canonical equation (2):

$$\begin{cases} x^2 = R^2\cos^2 t \\ y^2 = R^2\sin^2 t \end{cases} \quad \Rightarrow \quad x^2 + y^2 = R^2.$$

Likewise,

$$\begin{cases} x = x_0 + R\cos t \\ y = y_0 + R\sin t \end{cases}$$

are the parametric equations of the circle centered at the point $M_0(x_0, y_0)$ with radius *R*.

Examples:

1) The circle is given by the equation

$$x^2 - 4x + y^2 + 6y = 12.$$

Find the radius and the coordinates of the center.

Solution: Transform the quadratic polynomial on the left-hand side of the equation by adding and subtracting the corresponding constants to complete the perfect squares:

$$x^2 - 4x = (x^2 - 4x + 4) - 4 = (x-2)^2 - 4$$

$$y^2 + 6y = (y^2 + 6y + 9) - 9 = (y+3)^2 - 9.$$

Then the given equation is reduced to the form

$$(x-2)^2 + (y+3)^2 = 5^2,$$

which describes the circle centered at the point $M_0(2,-3)$ with radius 5.

2) Let

$$x^2 + 2x + y^2 - 8y + 17 = 0.$$

Find the canonical equation of the circle.

Solution:

$$x^2 + 2x + y^2 - 8y + 17 = 0 \quad \Rightarrow \quad (x^2 + 2x + 1) + (y^2 - 8y + 16) = 0 \quad \Rightarrow$$

$$(x+1)^2 + (y-4)^2 = 0.$$

The radius of the circle equals zero, that means the given equation corresponds to a null point circle.

3) The equation

$$x^2 + 2x + y^2 + 5 = 0$$

can be reduced to the form

$$(x+1)^2 + y^2 = -4,$$

which has no solutions. In this case they say that the equation describes an imaginary circle.

8.2. Ellipses

An **ellipse** is a plane curve, which is represented by the equation

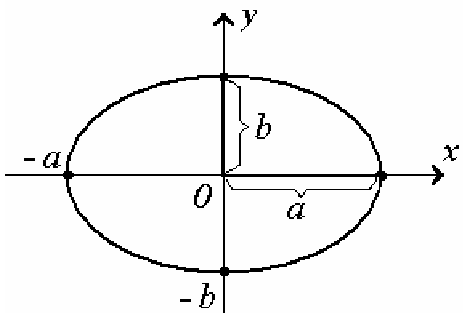
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \qquad (3)$$

in some Cartesian coordinate system.

Equation (3) is called the **canonical equation of an ellipse**, or the equation of an ellipse in the canonical system of coordinates. The positive quantities $2a$ and $2b$ are called the axes of the ellipse. One of them is said to be the major axis, while the other is the minor axis.

In the canonical system, the coordinate axes are the axes of symmetry, that means if a point (x, y) belongs to the ellipse, then the points $(-x, y)$, $(x,-y)$ and $(-x,-y)$ also belong to the ellipse.

The intersection points of the ellipse with the axes of symmetry are called the **vertices** of the ellipse. Hence, the points $(\pm a, 0)$ and $(0, \pm b)$ are the vertices of ellipse (3).



If $a = b = R$ then equation (3) is reduced to equation (2) of a circle. Thus, one can consider a circle as a specific ellipse.

The parametric equations of the ellipse have the following form:

$$\begin{cases} x = a\cos t \\ y = b\sin t \end{cases}$$

One can easily eliminate the parameter t to obtain the canonical equation of the ellipse:

$$\begin{cases} \dfrac{x^2}{a^2} = \cos^2 t \\ \dfrac{y^2}{b^2} = \sin^2 t \end{cases} \quad \Rightarrow \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=1.$$

The equation

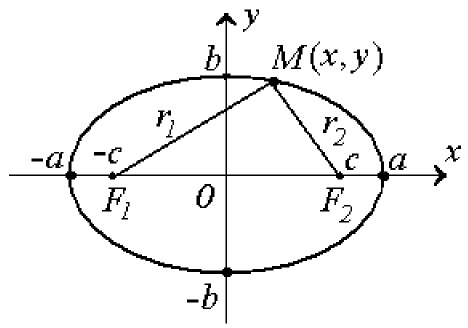
$$\frac{(x-x_0)^2}{a^2}+\frac{(y-y_2)^2}{b^2}=1$$

corresponds to the ellipse with the center at the point $M_0(x_0, y_0)$. The axes of symmetry of this ellipse pass through M_0, being parallel to the coordinate axes.

8.2.1. Properties of Ellipses

Consider an ellipse, which is given by equation (3) with the major axis $2a$. Two fixed points, $F_1(-c,0)$ and $F_2(c,0)$, are called the **focuses** of the ellipse, if equality $c^2 = a^2 - b^2$ is satisfied.
Correspondingly, the distances r_1 and r_2 from any point $M(x, y)$ of the ellipse to the points F_1 and F_2 are called the **focal distances**.

The ratio $\dfrac{c}{a} = \varepsilon$ is called the **eccentricity** of ellipse.
Note that $0 < \varepsilon < 1$.



1) Let x be the abscissa of a point of ellipse (3). Then the focal distances of the point can be expressed as follows:

$$r_1 = a + x\varepsilon, \qquad (4a)$$

$$r_2 = a - x\varepsilon. \qquad (4b)$$

Proof: By the definition, the distance between two points, $M(x, y)$ and $F_1(-c,0)$, is

$$r_1 = \sqrt{(x+c)^2 + y^2}.$$

Consider the expression under the sign of the radical.
By substituting

$$y^2 = (a^2 - x^2)\frac{b^2}{a^2},$$

$$c = a\varepsilon \quad \text{and} \quad b^2 = a^2 - c^2 = a^2(1-\varepsilon^2)$$

in r_1^2, we obtain

$$r_1^2 = (x+c)^2 + y^2 = x^2 + 2cx + c^2 + y^2$$
$$= x^2 + 2ax\varepsilon + a^2\varepsilon^2 + (a^2 - x^2)(1-\varepsilon^2),$$

which results in

$$r_1^2 = a^2 + 2a\,x\,\varepsilon + x^2\varepsilon^{\,2} = (a + x\,\varepsilon)^2 .$$

Likewise,

$$r_2 = \sqrt{(x-c)^2 + y^2} \quad \Rightarrow \quad r_2^2 = a^2 - 2a\,x\,\varepsilon + x^2\varepsilon^{\,2} = (a - x\,\varepsilon)^2 .$$

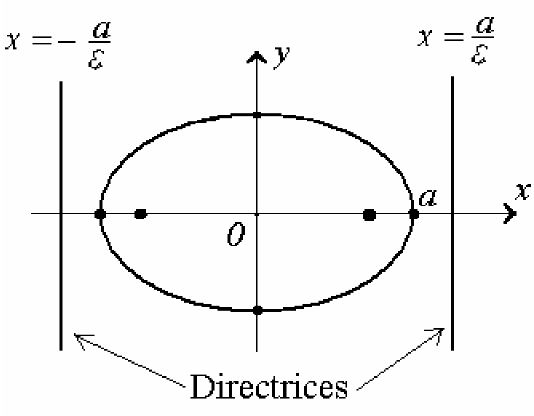
Since $a \pm x\,\varepsilon > 0$, the above formulas give the desired statement.

2) For any point of ellipse (3), the sum of the focal distances is the constant quantity $2a$:

$$r_1 + r_2 = 2a . \qquad (5)$$

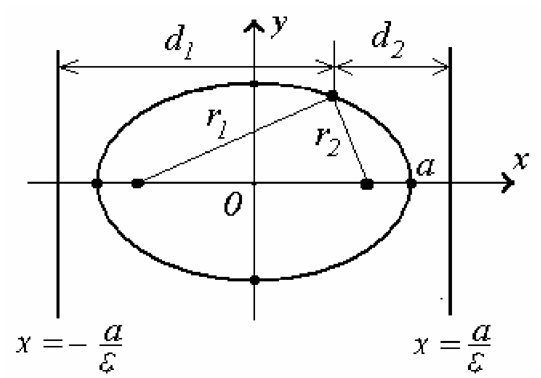
This property follows from formulas (4a) and (4b).

Two symmetric lines passing at the distance $\frac{a}{\varepsilon}$ from the center of an ellipse and being perpendicular to the major axis are called the **directrices**.



3) For any point of ellipse (3), the ratio of the focal distance to the distance from the corresponding directrix is equal to the eccentricity of the ellipse:

$$\frac{r_1}{d_1} = \frac{r_2}{d_2} = \varepsilon \qquad (6)$$

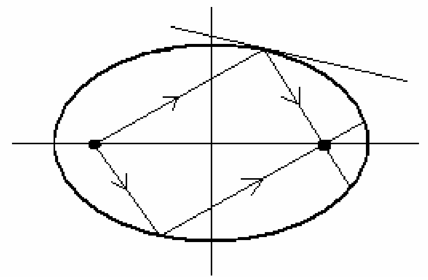


Proof: By Property 1 and in view of the fact that

$$d_1 = \frac{a}{\varepsilon} + x \quad \text{and} \quad d_1 = \frac{a}{\varepsilon} - x,$$

we obtain the desired results.

4) Assume that the curve of an ellipse has the mirror reflection property. If a point light source is located at a focus of the ellipse, then rays of light meet at the other focus after being reflected.



In other words, at any point of an ellipse, the tangent line forms equal angles with the focal radiuses.

5) The orbital path of a planet around the sun is an ellipse such that the sun is located at a focus.

Example: Reduce the equation

$$2x^2 + 4x + 3y^2 - 12y = 1$$

to the canonical form. Give the detailed description of the curve.

Solution: Complete the perfect squares.

$$2x^2 + 4x + 3y^2 - 12y = 1 \quad \Rightarrow$$

$$2(x^2 + 2x + 1) + 3(y^2 - 4y + 4) = 15 \quad \Rightarrow$$

$$2(x+1)^2 + 3(y-2)^2 = 15 \quad \Rightarrow$$

$$\frac{(x+1)^2}{15/2} + \frac{(y-2)^2}{5} = 1.$$

Thus, the given equation describes the ellipse with the center at the point (–1, 2).

The major semi-axis equals $\sqrt{15/2}$, and the minor semi-axis is $\sqrt{5}$. The focuses are located on the horizontal line $y = 2$. The distance between each focus and the center is

$$c = \sqrt{a^2 - b^2} = \sqrt{\frac{15}{2} - 5} = \sqrt{\frac{5}{2}} = \frac{\sqrt{10}}{2}.$$

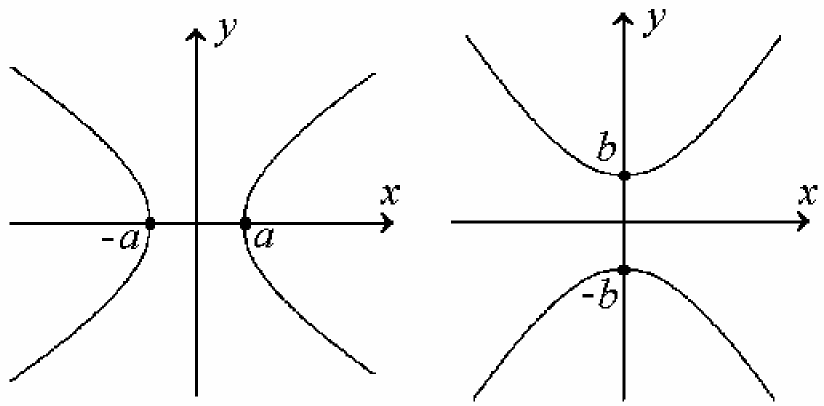
The eccentricity equals

$$\varepsilon = \frac{c}{a} = \frac{\sqrt{5}}{\sqrt{15}} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}.$$

8.3. Hyperbolas

A **hyperbola** is a plane curve, which can be represented in some Cartesian coordinate system by one of the below equations

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \qquad \frac{x^2}{a^2}-\frac{y^2}{b^2}=-1 \qquad (7)$$



Equations (7) are called the **canonical equations of a hyperbola**. The corresponding coordinate system is said to be the canonical system. In this coordinate system, the coordinate axes are axes of symmetry, that is, if a point (x, y) belongs to the hyperbola then the points $(-x, y)$, $(x, -y)$ and $(-x, -y)$ also belong to the hyperbola.

The intersection points of the hyperbola with the axis of symmetry are called the **vertices** of the hyperbola. Any hyperbola has two vertices.

If $a = b$ then the hyperbola is called an equilateral hyperbola.

The equations

$$\frac{(x-x_0)^2}{a^2}-\frac{(y-y_2)^2}{b^2}=\pm 1$$

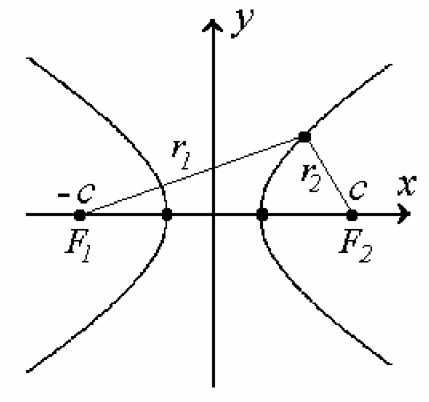
describe hyperbolas with the center at the point $M_0(x_0, y_0)$. The axes of symmetry of the hyperbolas pass through M_0, being parallel to the coordinate axes.

Consider a hyperbola, which is given by the equation

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \qquad (8)$$

Two fixed points, $F_1(-c, 0)$ and $F_2(c, 0)$, are called the **focuses** of the hyperbola, if the equality $c^2 = a^2 + b^2$ is satisfied.

Correspondingly, the distances r_1 and r_2 from any point $M(x, y)$ of the hyperbola to the points F_1 and F_2 are called the **focal distances**.



The ratio $\frac{c}{a} = \varepsilon$ is called the **eccentricity** of hyperbola.
Note that $\varepsilon > 1$.

8.3.1. Properties of Hyperbolas

1) Let x be the abscissa of a point of hyperbola (8). Then the focal distances of the point are the following:

$$r_1 = \pm(x\varepsilon + a), \qquad (9a)$$

$$r_2 = \pm(x\varepsilon - a). \qquad (9b)$$

In the above formulas we have to apply the sign '+' for the points on the right half-hyperbola, while the sign '–' is used for the points on the left half-hyperbola.

This property is similar to the corresponding one of ellipses.
The distance between two points $M(x, y)$ and $F_1(-c, 0)$ is

$$r_1 = \sqrt{(x+c)^2 + y^2},$$

where

$$y^2 = (x^2 - a^2)\frac{b^2}{a^2},$$

$$c = a\varepsilon \quad \text{and} \quad b^2 = c^2 - a^2 = a^2(\varepsilon^2 - 1).$$

Therefore,

$$r_1^2 = (x+c)^2 + y^2 = x^2 + 2cx + c^2 + y^2$$
$$= x^2 + 2ax\varepsilon + a^2\varepsilon^2 + (x^2 - a^2)(\varepsilon^2 - 1)$$
$$= x^2\varepsilon^2 + 2ax\varepsilon + a^2 = (x\varepsilon + a)^2.$$

Likewise,

$$r_2 = \sqrt{(x-c)^2 + y^2} \quad \Rightarrow \quad r_2^2 = (x\varepsilon - a)^2.$$

Since $x\varepsilon \pm a > 0$ for the points of the right half-hyperbola, and $x\varepsilon \pm a < 0$ for points of the left half-hyperbola, we have got the desired results.

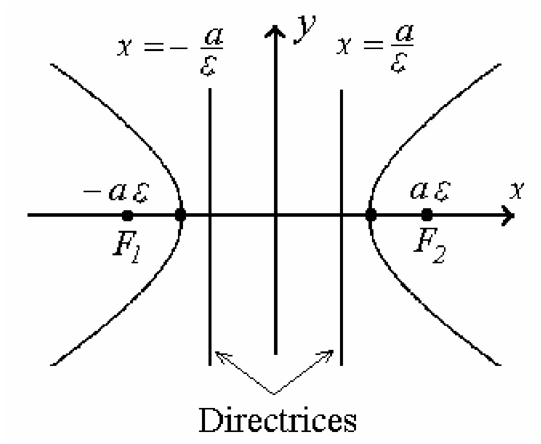
2) For any point of hyperbola (8), the difference between the focal distances is the constant quantity $(\pm 2a)$:

$$r_1 - r_2 = \pm 2a\,. \tag{10}$$

The sign depends on whether the point lies on the right or left half-hyperbola.

The proof is straightforward. We only need to apply Property 1.

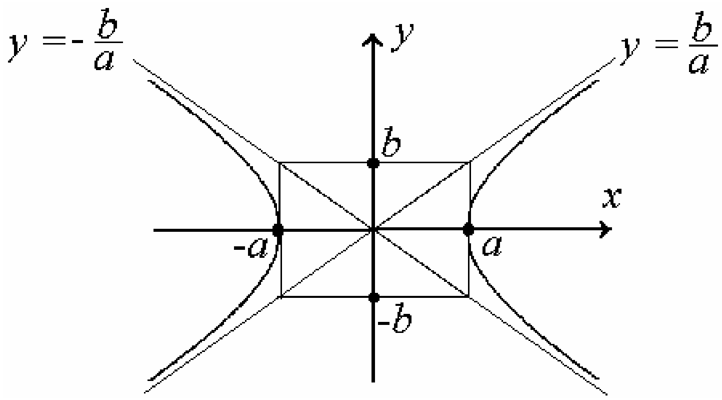
The **directrices** of hyperbola (8) are two vertical lines $x = \pm \frac{a}{\varepsilon}$.



3) For any point of hyperbola (8) the ratio of the focal distance to the distance from the corresponding directrix is equal to the eccentricity of the hyperbola.

$$\frac{r_1}{d_1} = \frac{r_2}{d_2} = \varepsilon \tag{11}$$

4) Two straight lines $y = \pm \frac{b}{a}$ are the asymptotes of hyperbola (8).



Proof: Express the variable *y* from equality (8) in the explicit form.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \Rightarrow \quad y^2 = \frac{b^2}{a^2}(x^2 - a^2) \quad \Rightarrow$$

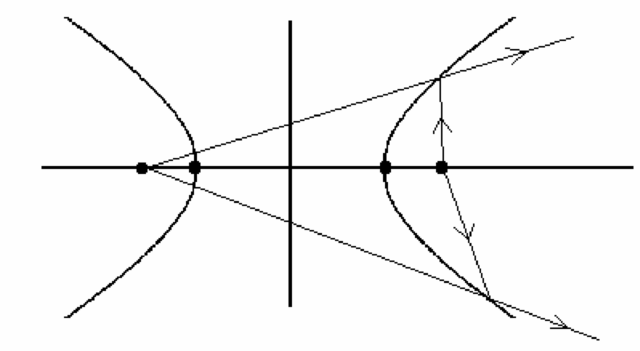
$$y = \pm\frac{b}{a}\sqrt{x^2 - a^2}\ .$$

If x approaches infinity, then constant a^2 is a negligible quantity, that is,

$$y \to \pm\frac{b}{a}x\,.$$

Hence, the property.

5) Assume that the curve of a hyperbola has the mirror reflection property. If a point light source is located at a focus of the hyperbola, then the other focus is the image source of rays that being reflected.



The drawing illustrates that reflected rays form a divergent beam.

Example: Reduce the equation

$$x^2 - 6x + 2y^2 + 8y = 0$$

to the canonical form. Give the detailed description of the curve.

Solution: Complete the perfect squares.

$$x^2 - 6x - 2y^2 - 8y = 7 \qquad \Rightarrow$$

$$(x^2 - 6x + 9) - 2(y^2 + 4y + 4) = 8 \qquad \Rightarrow$$

$$(x-3)^2 - 2(y+2)^2 = 8\,.$$

Dividing both sides by 8 we obtain the equation

$$\frac{(x-3)^2}{8} - \frac{(y+2)^2}{4} = 1\,,$$

which describes the hyperbola with the center at the point (3, –2). The focuses are located on the horizontal line y = –2. The distance between each focus and the center of the hyperbola is

$$c = \sqrt{a^2 + b^2} = \sqrt{8+4} = \sqrt{12} = 3\sqrt{2}\ .$$

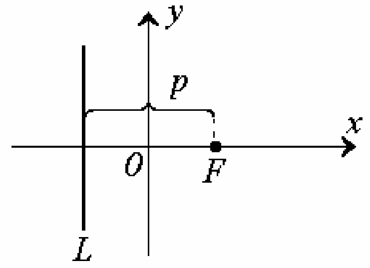
The eccentricity of the hyperbola equals

$$\varepsilon = \frac{c}{a} = \frac{3\sqrt{2}}{\sqrt{8}} = \frac{3}{2}.$$

8.4. Parabolas

A **parabola** is the locus of points, which are equidistant from a given point F and line L. The point F is called the **focus**. The line L is called the **directrix** of the parabola.

Let the focus be on the x-axis and the directrix be parallel to the y-axis at the distance p from the focus as it is shown in the figure below.



Then the focal distance of a point $M(x, y)$ is

$$r = \sqrt{(x - p/2)^2 + y^2}$$

and the distance from M to the directrix is

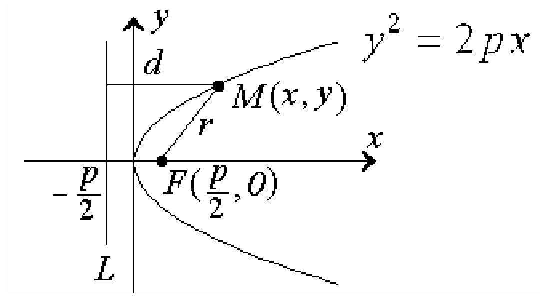
$$d = x + p/2.$$

Therefore, due to the transformations

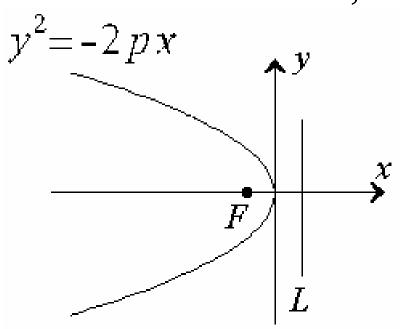
$$r = d \quad \Rightarrow \quad \sqrt{(x - p/2)^2 + y^2} = x + p/2 \quad \Rightarrow$$

$$(x - p/2)^2 + y^2 = (x + p/2)^2,$$

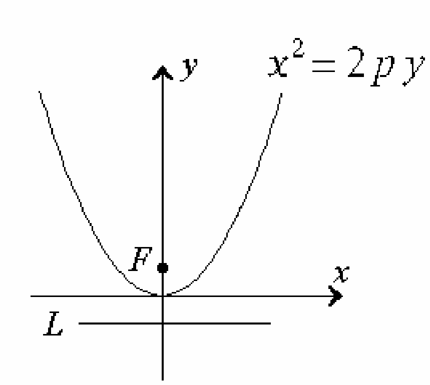
we obtain the following equation of a parabola: $y^2 = 2px$.

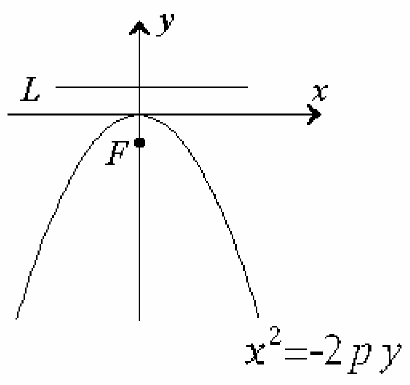


If the focus is located on the left of the directrix, then we obtain



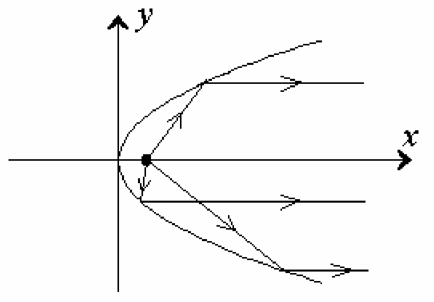
Some more cases are shown in the drawings below.





Parabola properties:

1) Any parabola has the axis of symmetry, which passes through the vertex of the parabola, being perpendicular to the directrix.
2) Let an ellipse be the mirror reflection curve. If a point light source is located at a focus of the ellipse, then rays of light are parallel after being reflected.



The equations

$$(y - y_0)^2 = \pm 2p(x - x_0),$$

$$(x - x_0)^2 = \pm 2p(y - y_0)$$

describe parabolas with the vertex at the point $M_0(x_0, y_0)$.

Example: Reduce the equation

$$x^2 + 4x - 3y = -5.$$

to the canonical form. Give the detailed description of the curve.

Solution:

$$x^2 + 4x - 3y = -5 \quad \Rightarrow$$

$$(x + 2)^2 = 3(y - \frac{5}{3}).$$

This equation describes the parabola with the vertex at the point $M_0(-2, 5/3)$. The axis of symmetry is a line $x = -2$ which is parallel to the y–axis.

8.5. Summary

Let F be a point (focus) and L be a line (directrix) of a quadric curve. Consider the locus of points such that the ratio of the distances to the focus and to the directrix is a constant quantity (eccentricity),

$$\frac{r}{d} = \varepsilon . \qquad (12)$$

If $0 < \varepsilon < 1$ then equation (12) describes an ellipse.
If $\varepsilon = 1$ then equation (12) describes a parabola.
If $\varepsilon > 1$ then (12) is the equation of a hyperbola.
Thus, the curves of the second order can be classified by the value of the eccentricity.
From the algebraic point of view, the equation

$$a_1 x^2 + b_1 x + a_2 y^2 + b_2 y + a_3 xy + c = 0$$

describes a curve of the second order in the x,y–plane, provided that at least one of the leading coefficients is non-zero.
The presence of the term xy means that the axes of symmetry of the curve are rotated with respect to the coordinate axes.
The linear term x (or y) means that the center (or vertex) of the curve is shifted along the corresponding axis.

Examples:

1) The equation

$$xy = \text{const}$$

describes a hyperbola, whose axes of symmetry are rotated on the angle $45°$ with respect to the coordinate axes.

2) If $c \to 0$, then a hyperbola

$$x^2 - y^2 = c$$

collapses to the pair of the lines $y = \pm x$.

3) The equation

$$x^2 + 2y^2 = -1$$

has no solutions and corresponds to an imaginary ellipse.

References

1. K.P. Arefiev, O.V. Boev, A.I. Nagornova, G.P. Stoljarova, A.N. Harlova. Higher Mathematics, Part 1. Textbook. Tomsk, TPU Press, 1998, 97p.
2. Murray H. Plotter, Charles B. Morrey. Intermediate Calculus, Springer, 1985.
3. V.V. Konev. Mathematics: Preparatory Course. Textbook. Tomsk. TPU Press, 2009, 104p.
4. V.V. Konev, The Elements of Mathematics. Textbook. Tomsk. TPU Press, 2009, 140p.

---

5. V.V. Konev, Higher Mathematics, Part 2. Textbook. Tomsk. TPU Press, 2009. 114p.
6. V.V. Konev. Mathematics: Algebra, Workbook. TPU Press, 2009, 60p.
7. V.V. Konev, Mathematics: Geometry and Trigonometry, Workbook. Tomsk. TPU Press, 2009, 34p.
8. V.V. Konev, The Elements of Mathematics. Workbook, Part 1. Tomsk. TPU Press, 2009, 54p.
9. V.V. Konev, The Elements of Mathematics. Workbook, Part 2. Tomsk. TPU Press, 2009, 40p.

Valery V. Konev, Associate Professor of the Higher Mathematics Department, TPU, Ph.D.

Linear Algebra, Vector Algebra and Analytical Geometry

Textbook

Editor: V.A. Kilin, Professor, of the Higher Mathematics Department, TPU, D.Sc.