МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ

**Національний технічний університет України
«Київський політехнічний інститут»**

ЛІНІЙНА АЛГЕБРА ТА АНАЛІТИЧНА ГЕОМЕТРІЯ

(Частина 1)

Курс лекцій

НАВЧАЛЬНИЙ ПОСІБНИК

для студентів напрямів підготовки
6.040301 «Прикладна математика»
та 6.050103 «Програмна інженерія»

Рекомендовано
Вченою радою ФПМ НТУУ «КПІ»
(___.___.2013 р., протокол № __)

Київ
НТУУ «КПІ»
2013

Лінійна алгебра та аналітична геометрія (Частина 1). Курс лекцій. Навчальний посібник для студентів напрямів підготовки 6.040301 «Прикладна математика» та 6.050103 «Програмна інженерія» [Електронне видання] / В.В. Мальчиков, К.О. Костенко. – К. : НТУУ «КПІ», 2013. – 174 с.

Навчально-методичне видання

ЛІНІЙНА АЛГЕБРА ТА АНАЛІТИЧНА ГЕОМЕТРІЯ (ЧАСТИНА 1).
КУРС ЛЕКЦІЙ
НАВЧАЛЬНИЙ ПОСІБНИК
ДЛЯ СТУДЕНТІВ НАПРЯМІВ ПІДГОТОВКИ 6.040301 «ПРИКЛАДНА МАТЕМАТИКА» ТА 6.050103 «ПРОГРАМНА ІНЖЕНЕРІЯ»

Навчальний посібник розроблено для ознайомлення студентів з теоретичними відомостями з дисципліни «Лінійна алгебра та аналітична геометрія». Навчальне видання призначене для студентів, які навчаються за напрямом 6.040301 «Прикладна математика» та 6.050103 «Програмна інженерія» факультету прикладної математики НТУУ «КПІ».

Автори: *Мальчиков Володимир Вікторович*
Костенко Катерина Олексіївна

Відповідальний за випуск *Сущук-Слюсаренко Вікторія Ігорівна*

Рецензент *Мінарченко Олександр Миколайович, к.ф.-м.н.*

ЗМІСТ

1. МАТРИЦІ ТА ВИЗНАЧНИКИ 4

1.1. Матриці. Типи матриць. Операції над матрицями та їх властивості 4

1.2. Визначник. Властивості визначників. Методи обчислення 16

1.3. Обернені матриці. Матричні рівняння 27

1.4. Лінійна залежність та незалежність. База системи елементів 38

1.5. Ранг матриці. Базисні мінори. Методи обчислення рангу 43

2. СИСТЕМИ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ 50

2.1. Системи лінійних алгебраїчних рівнянь. Основні поняття. 50

2.2. Однорідні та неоднорідні системи. Методи пошуку розв’язків .. 58

3. ЕЛЕМЕНТИ ВЕКТОРНОЇ АЛГЕБРИ 67

3.1. Вектори на площині та в просторі. Операції над векторами 67

3.2. Добутки векторів: скалярний, векторний, подвійний векторний, мішаний 88

4. АНАЛІТИЧНА ГЕОМЕТРІЯ НА ПЛОЩИНІ 104

4.1. Лінія на площині. Способи задання 104

4.2. Пряма на площині. Рівняння прямої 113

4.3. Лінії 2-го порядку. Перетворення координат в прямокутній декартовій системі координат 132

5. ЕЛЕМЕНТИ ТЕОРІЇ МНОГОЧЛЕНІВ 157

5.1. Операції над многочленами. Найбільший спільний дільник. Алгоритм Евкліда 157

5.2. Корені многочленів. Теорема Безу. Основна теорема алгебри. 166

1. МАТРИЦІ ТА ВИЗНАЧНИКИ

1.1. Матриці. Типи матриць. Операції над матрицями та їх властивості

Матрицею називають прямокутну таблицю з чисел, яка містить деяку кількість m рядків і деяку кількість n стовпців.

Іншими словами, матриця – це сукупність чисел, записаних у вигляді таблиці з m рядками і n стовпцями. Натуральні числа m та n задають розмірність матриці, яку позначають як $m \times n$. Тобто, якщо говорять про те, що задано матрицю розмірністю 3×5, то це матриця з трьома рядками і п'ятьма стовпцями. Множину матриць розмірності $m \times n$ позначатимемо $M_{m \times n}$.

Слід зауважити, що перше число завжди позначає кількість рядків.

Якщо в матриці однакова кількість рядків та стовпців, тобто $m = n$, то таку матрицю називають **квадратною** і говорять, що задано матрицю n-го (або m-го) порядку.

Матриці позначаються великими латинськими літерами (*A, B, C, …*), елементи матриці – маленькими літерами a_{ij}, $b_{ij}, \ldots$; $i = 1, 2, \ldots m$; $j = 1, 2, \ldots, n$.

Записи $i = \overline{1..m}$, $j = \overline{1..n}$ вказують на те, в яких межах змінюються індекси рядків та стовпців відповідно.

Елемент a_{ij} знаходиться на перетині i-го рядка та j-го стовпця матриці A.

Це компактне позначення (ім’я) матриці. Можливі також такі варіанти:

1) Запис у вигляді таблиці (елементи в круглих дужках або у подвоєних рисках)

$$\begin{pmatrix} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ .... & .... & .... & .... \\ a_{m1} & a_{m2} & .... & a_{mn} \end{pmatrix} \quad \text{або} \quad \begin{Vmatrix} a_{11} & a_{12} & .... & a_{1n} \\ a_{21} & a_{22} & .... & a_{2n} \\ .... & .... & .... & .... \\ a_{m1} & a_{m2} & .... & a_{mn} \end{Vmatrix}$$

2) $A_{m\times n}$ – ім'я матриці з урахуванням розмірностей

3) $A = \|a_{ij}\| = \left(a_{ij}\right) = \left\{a_{ij}\right\} \quad i = \overline{1..m} \quad j = \overline{1..n}$

У випадку квадратної матриці n-го порядку

$$\begin{pmatrix} a_{11} & .... & a_{1n} \\ .... & .... & .... \\ a_{n1} & .... & a_{nn} \end{pmatrix}$$

можна ввести ще декілька понять:

– **головна діагональ** – діагональ, яка починається з лівого верхнього кута матриці та прямує в правий нижній кут (елементи a_{11}, $a_{22}, \ldots, a_{nn}$).

– **побічна діагональ** – діагональ, яка починається з лівого нижнього кута матриці та прямує в правий верхній кут (елементи a_{n1}, $a_{n-1,2}, \ldots, a_{1n}$).

– **слід матриці** (позначається *Tr A* або *Sp A*) – сума елементів головної діагоналі.

Індекси елементів матриці є парами натуральних чисел (*i, j*). Тобто, вони є елементами декартового добутку множин

$I \times J = \{(1,1),(1,2),...,(i,j),...,(m,n)\}$. Фактично, кожній парі чисел декартового добутку $I \times J$ матриця ставить у відповідність елемент матриці a_{ij} – дійсне чи комплексне число. Таким чином, матриця є **відображенням** декартового добутку $I \times J$ на множину дійсних чисел R (всі елементи матриці є дійсними числами – матриця називається **дійсною**) чи на множину комплексних чисел C (хоча б один з елементів є комплексним числом – матриця називається **комплексною**):

$$A: \; I \times J \to R \; (I \times J \to C)$$
$$(i,j) \to a_{ij}$$

Окремим випадком матриць, який досить часто використовується та розглядається, є випадок коли хоча б один з порядків матриці дорівнює одиниці, при цьому отримуємо або **матрицю-рядок** ($m = 1$) $A_{1\times n}$

$$A = \begin{pmatrix} a_{11}, & a_{12}, & ..., & a_{1n} \end{pmatrix}$$

або матрицю-**стовпець** ($n = 1$) $B_{m\times 1}$

$$B = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{m1} \end{pmatrix}$$

Такі матриці називаються **векторами** (вектор з алгебраїчної точки зору) – **вектор-рядок** та **вектор-стовпець** відповідно.

Матриця 1-го порядку $(m = n = 1)$ – це **звичайне число**.

Розглянемо деякі різновиди матриць:

1) **Нульова матриця (нуль-матриця)** – матриця розмірності $m \times n$, всі елементи якої дорівнюють нулю

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

2) **Діагональна матриця** – квадратна матриця *n*-го порядку, в якій всі елементи окрім (можливо) елементів головної діагоналі дорівнюють нулю

$$D = \begin{pmatrix} a_{11} & 0 & ... & 0 \\ 0 & a_{22} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & a_{nn} \end{pmatrix} = diag\left(a_{11}, \quad a_{22}, \quad ..., \quad a_{nn}\right)$$

Приклад:

$$D = diag\left(2, \quad -1, \quad 0, \quad 0, \quad 3\right) \Rightarrow D = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

3) **Одинична матриця** – це діагональна матриця, в якій всі елементи головної діагоналі дорівнюють одиниці

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

Якщо в діагональній матриці всі елементи головної діагоналі однакові та дорівнюють числу λ ($D = diag(\lambda, \; ..., \; \lambda)$), то її можна записати за допомогою одиничної матриці $D = \lambda \cdot I$

4) **Трикутна матриця** – квадратна матриця *n*-го порядку, в якій всі елементи під/над головною/побічною діагоналлю нульові.

$\begin{pmatrix} a_{11} & a_{12} & .... & a_{1n} \\ 0 & a_{22} & .... & a_{2n} \\ .... & .... & .... & .... \\ 0 & 0 & .... & a_{mn} \end{pmatrix}$ – верхня трикутна відносно головної діагоналі

$\begin{pmatrix} 0 & .... & 0 & a_{1n} \\ 0 & .... & a_{2,n-1} & a_{2n} \\ .... & .... & .... & .... \\ a_{m1} & .... & a_{m,n-1} & a_{mn} \end{pmatrix}$ – нижня трикутна відносно побічної діагоналі

Слід зауважити, що серед перелічених чотирьох типів матриць перша - прямокутна, а решта є квадратними.

Розглянемо основні операції, які можна виконувати над матрицями.

1) **Порівняння матриць**

Дві матриці *A* та *B* вважають рівними, якщо ці матриці мають однакову розмірність та всі їх відповідні елементи співпадають

$A_{m \times n} \quad B_{p \times q}$

а) $m = p \;\&\; n = q$

b) $\forall i = \overline{1..m} \quad \forall j = \overline{1..n} \quad a_{ij} = b_{ij}$

Приклад:

$$A = \begin{pmatrix} 1 & -4 & -3 \\ 0 & 2 & 5 \end{pmatrix},\ B = \begin{pmatrix} 1 & 0 \\ -4 & 2 \\ -3 & 5 \end{pmatrix},\ C = \begin{pmatrix} 1 & 7 & -3 \\ 0 & 2 & -5 \end{pmatrix},\ D = \begin{pmatrix} 1 & -4 & -3 \\ 0 & 2 & 5 \end{pmatrix}$$

$A \neq B$ (різна розмірність)

$A \neq C$ (різні елементи)

$A = D$

2) **Сума матриць**

Сумою двох матриць *A* та *B* однакової розмірності *m* та *n* називають матрицю *C,* елементи c_{ij} якої дорівнюють

$$c_{ij} = a_{ij} + b_{ij} \quad i = \overline{1..m} \quad j = \overline{1..n}$$

Аналогічно визначають суму довільної кількості матриць:

$$A + B + C + D = (A + B) + C + D = K + C + D = (K + C) + D = L + D = P$$

3) **Множення матриці на число**

Добутком матриці *A* розмірності $m \times n$ на дійсне число α називається матриця *B* розмірності $m \times n$, елементи b_{ij} якої дорівнюють

$$b_{ij} = \alpha \cdot a_{ij}$$

При цьому записують $B = \alpha A$

Дві останні операції називаються **лінійними**.

4) **Протилежна матриця**

Протилежною матрицею до матриці $A_{m \times n}$ називають таку матрицю *B*, яку можна отримати з матриці *A* шляхом множення всіх її елементів на «-1».

$$B = -A = (-1) \cdot A$$

$$b_{ij} = (-1) \cdot a_{ij}$$

5) **Різниця матриць**

Різницею матриць *A* та *B* однакової розмірності $m \times n$ називають таку матрицю *C*, яку отримують шляхом додавання матриці *A* та матриці, що протилежна до матриці *B*.

$$C = A - B = A + (-1) \cdot B$$

Очевидно, що $A + (-A) = 0$

Властивості лінійних операцій над матрицями:

a) Комутативність додавання (розподільна властивість)

$$A + B = B + A$$

b) Асоціативність додавання (сполучна властивість)

$$(A + B) + C = A + (B + C)$$

c) Додавання нульової матриці

$$A + 0 = A = 0 + A$$

d) Сполучна властивість відносно числового множника

$$\alpha\left(\beta \cdot A\right) = \left(\alpha\beta\right) A \ , \ \alpha, \beta \in R$$

e) Розподільна властивість відносно суми матриць

$$\alpha(A + B) = \alpha A + \alpha B$$

f) Розподільна властивість відносно суми чисел

$$(\alpha + \beta) A = \alpha A + \beta A$$

6) Транспонування матриці

Транспонуванням матриці називають таку операцію, коли міняють місцями рядки та стовпці матриці із збереженням порядку їх слідування. Матрицю, транспоновану відносно матриці *A,* позначають A^T (або A').

Якщо

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix},$$

то

$$A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \dots & \dots & \dots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

Приклади:

$$A=\begin{pmatrix} 1 & -1 & -4 \\ -2 & 3 & 5 \\ 7 & 8 & -9 \end{pmatrix}, \qquad A^T=\begin{pmatrix} 1 & -2 & 7 \\ -1 & 3 & 8 \\ -4 & 5 & -9 \end{pmatrix}$$

$$B=\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \qquad B^T=\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Властивості операції транспонування:

a) $(A^T)^T = A$

b) $(\alpha A)^T = \alpha \cdot A^T$

c) $(A+B)^T = A^T + B^T$

Матриця A називається **симетричною**, якщо $A = A^T$ та **кососиметричною**, якщо $A = -A^T$.

7) Добуток матриць

Матриці $A_{m\times n}$ та $B_{p\times q}$ називаються узгодженими, якщо кількість стовпців n першої матриці дорівнює кількості рядків p другої матриці:

Слід зазначити, що лише **узгоджені** матриці можна множити.

Результатом множення матриці $A_{m\times n}$ на матрицю $B_{n\times q}$ є матриця C розмірності $m\times q$, елементи якої визначають таким чином:

$$c_{ij} = a_{i1}\cdot b_{1j} + a_{i2}\cdot b_{2j} + \ldots + a_{in}\cdot b_{nj} = \sum_{k=1}^{n} a_{ik}\cdot b_{kj}$$

Тобто, елементи рядка першої матриці множать на відповідні елементи стовпця другої матриці та обчислюється сума всіх цих добутків.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ - & - & - & \rightarrow \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & | & b_{1q} \\ b_{21} & \dots & b_{2j} & | & b_{2q} \\ \dots & \dots & \dots & | & \dots \\ b_{n1} & \dots & b_{nj} & \downarrow & b_{nq} \end{pmatrix} = \begin{pmatrix} \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & c_{ij} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

Іншими словами, елемент c_{ij}, що стоїть на перетині і-го рядка та j-го стовпця матриці $C = A \cdot B$, дорівнює сумі попарних добутків відповідних елементів і-го рядка матриці А та j-го стовпця матриці В.

Приклад:

1) Нехай задано матриці $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ та $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$

Тоді елементи матриці $C = A \cdot B$ обчислюються наступним чином:

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} \\ c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} \\ c_{22} &= a_{21}b_{12} + a_{22}b_{22} \end{aligned}$$

2) Для матриць $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ та $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ їх добутки

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \; BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Властивості добутку матриць:

1) Асоціативність (сполучна властивість)

$$(AB)C = A(BC)$$

2) Розподільна властивість відносно суми матриць

$$(A+B)C = AC + BC$$

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

3) Комутативність

Слід зазначити, що в загальному випадку $AB \neq BA$

4) $A \cdot I = I \cdot A = A$, $A \cdot O = O \cdot A = O$

5) $(AB)^T = B^T A^T$

Цілим додатним степенем $K > 1$, $K \in N$ квадратної матриці A називають добуток K матриць, кожна з яких дорівнює A:

$$A^K = \underbrace{A \cdot A \cdot \ldots \cdot A}_{K \quad раз}$$

Нульовий степінь квадратної матриці – це одинична матриця $A^O = I$.

Перший степінь квадратної матриці – це сама матриця A:

$$A^1 = A$$

Поліномом (многочленом) степеня *k* квадратної матриці A називають вираз

$$P(A)=\alpha_0 I+\alpha_1 A+\alpha_2 A^2+\ldots+\alpha_{K-1}A^{K-1}+\alpha_K A^K,$$

де $\alpha_i \in R \quad i=\overline{0..k}$

Поліном $P(A)$ від матриці можна отримати зі звичайного полінома $P(x)=\alpha_0+\alpha_1 x+\alpha_2 x^2+\ldots+\alpha_{K-1}x^{K-1}+\alpha_K x^K$ шляхом заміни x на A. Якщо $P(A)=0$, то матриця A називається **коренем** цього **многочлена**, а сам многочлен $P(A)$ – **анулюючим** матрицю A.

Контрольні запитання:

1. Що таке матриця?
2. Які існують види матриць?
3. Як позначають матриці?
4. Які операції можна проводити над матрицями?
5. Які матриці називаються узгодженими?
6. Як перемножити дві матриці?

Завдання для самостійної роботи:

1. Довести розглянуті властивості.
2. Визначити, що буде результатом множення двох алгебраїчних векторів.
3. Яка ще властивість чисел не виконуватиметься для матриць?

Література: [1], с.12-17, 30; [2], с.190-198; [3], с.78-84

1.2. Визначник. Властивості визначників. Методи обчислення

Розглянемо довільну квадратну матрицю порядку з порядком n

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

В попередній темі для такої матриці була введена чисельна характеристика – слід матриці. Тепер пов'яжемо з кожною квадратною матрицею ще одну чисельну характеристику, яка називається **визначником** (або **детермінантом**) матриці, при цьому будемо використовувати для нього наступні позначення:

1) Δ
2) $\det A$
3) $\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$

Заздалегідь можна сказати, що **визначник** – це число, яке обчислюють на основі елементів матриці.

Для визначення детермінанта матриці n-го порядку скористаємося **індуктивним методом**. Тобто спочатку визначимо безпосередньо детермінант 1-го порядку, потім припустимо, що вже визначено поняття визначника *(n-1)*-го порядку, і спробуємо виразити визначник n-го порядку через визначник *(n-1)*-го порядку.

Детермінант визначається наступним чином:

1) детермінант матриці першого порядку дорівнює самому елементу матриці

$$\det A = |a_{11}| = a_{11}$$

2) детермінант матриці n-го порядку визначають як суму добутків елементів першого рядка a_{1j} $j = \overline{1..n}$ на відповідні цим елементам алгебраїчні доповнення A_{1j}:

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \sum_{j=1}^{n} a_{1j} A_{1j} \qquad (1.1)$$

Алгебраїчне доповнення A_{ij} елемента a_{ij} матриці А – це мінор M_{ij}, який відповідає даному елементу матриці, помножений на (-1) в степені, який дорівнює сумі індексів елемента a_{ij}:

$$A_{ij} = (-1)^{i+j} M_{ij} \qquad (1.2)$$

Мінор M_{ij}, який відповідає елементу матриці a_{ij}, є визначником матриці, яку можна отримати з початкової матриці A шляхом викреслювання з неї рядка і стовпця, відповідних елементу a_{ij} (тобто і-го рядка і j-го стовпця). Слід зазначити, що порядок цього визначника дорівнює *(n-1)*.

Таким чином, завершено рекурентну послідовність визначень: виражено визначник *n*-го порядку через *n* визначників *(n-1)*-го порядку (поняття останнього вважається вже введеним). Зрозуміло, що аналогічним чином, використовуючи формулу (1.1), можна кожен з визначників *(n-1)*-го порядку представити через *(n-1)* визначників *(n-2)* порядку (тобто отримати всього *n(n-1)* таких детермінантів).

Продовжуючи цей процес потрібну кількість разів, можна дійти до визначників 1-го порядку, які вже визначені вище.

Для прискорення цього процесу отримаємо формули для обчислення детермінантів 2-го і 3-го порядків безпосередньо через елементи визначників, використовуючи формулу (1.1), яку також називають **формулою розвинення визначника за першим рядком**.

Розглянемо матрицю другого порядку:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Скориставшись формулою (1.1), отримаємо:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}A_{11} + a_{12}A_{12} = a_{11} \cdot (-1)^{1+1} \cdot M_{11} + a_{12} \cdot (-1)^{1+2} M_{12} =$$

$$= a_{11} \cdot |a_{22}| - a_{12} \cdot |a_{21}| = a_{11}a_{22} - a_{12}a_{22} \qquad (1.3)$$

Тобто при обчисленні визначників 2-го порядку необхідно з добутку елементів головної діагоналі відняти добуток елементів побічної діагоналі.

Тепер візьмемо матрицю третього порядку

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

Застосуємо формулу (1.1):

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

$$= a_{11} \cdot (-1)^{1+1} M_{11} + a_{12} \cdot (-1)^{1+2} M_{12} + a_{13} \cdot (-1)^{1+3} M_{13} =$$

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

$$= \left[\text{на основі формули (1.3)}\right] = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) +$$

$$+ a_{13}(a_{21}a_{32} - a_{31}a_{22}) = \left[\text{групуючи елементи враховуючи знак}\right] =$$

$$(a_{11}a_{22}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32}) - (a_{11}a_{23}a_{32} + a_{12}a_{21}a_{33} + a_{13}a_{31}a_{22}) \quad (1.4)$$

Формулу (1.4) часто називають **«правилом зірочки»**. Досить легко можна запам'ятати знак доданків в цій формулі та які елементи утворюють доданки (Рис. 1.1а та 1.1б):

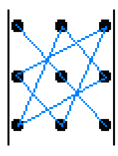

Рис. 1.1а. Добутки зі знаком плюс

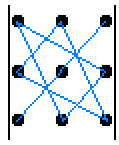

Рис. 1.1б. Добутки зі знаком мінус

Тобто, добуток елементів головної діагоналі та двох трикутників, основи яких паралельні до цієї діагоналі, враховують із знаком «плюс», а добуток елементів побічної діагоналі та двох інших трикутників, основи яких паралельні до цієї діагоналі, зі знаком «мінус».

Таким чином, отримані формули (1.3) і (1.4) дозволяють не доводити обчислення визначника довільного порядку до визначників першого порядку, а зупинитися на 1-2 кроки раніше.

Слід зазначити, що знак, з яким мінор входить в алгебраїчне доповнення, можна визначити ще наступним чином: для елемента a_{11} це знак «+», а далі зсув на один елемент вправо/вліво або вгору/вниз (але не по діагоналі) приводить до чергування знаку.

Формули, аналогічні формулам (1.3) і (1.4), тобто формули, які дозволяють виразити визначник безпосередньо через його елементи, можуть бути отримані для будь-якого порядку детермінанта. У тому числі, можна вивести і загальну формулу для детермінанта довільного n -го порядку.

Проте вони є достатньо громіздкими. Тому, в більшості випадків використовують формулу (1.1) – розвинення за першим рядком. Оскільки всі рядки визначника рівнозначні, то при підрахунку визначників можна скористатися будь-яким рядком (тобто розкласти за i -м рядком).

Оскільки за допомогою операції транспонування стовпці можна перетворити на рядки, то визначник можна обчислювати, розкладаючи його також за будь-яким стовпцем.

Таким чином, для знаходження визначника матриці А n-го порядку можна використовувати або **формулу розвинення визначника за *i*-м рядком**

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} \quad \forall i = \overline{1..n} \tag{1.5}$$

або **формулу розвинення визначника за j-м стовпцем**

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij} \quad \forall j = \overline{1..n} \tag{1.5a}$$

Властивості визначника:

1) Властивість рівноправності рядків та стовпців

При транспонуванні величина визначника зберігається

$$\det A = \det A^T$$

Ця властивість дозволяє всі подальші властивості встановлювати лише для рядків та бути впевненими в їх справедливості і для стовпців.

2) Властивість антисиметрії

При перестановці місцями двох рядків (двох стовпців) визначник зберігає свою абсолютну величину, але змінює знак на протилежний:

$$\begin{vmatrix} \ldots & \alpha_1 & \ldots & \beta_1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \alpha_n & \ldots & \beta_n & \ldots \end{vmatrix} = -\begin{vmatrix} \ldots & \beta_1 & \ldots & \alpha_1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \beta_n & \ldots & \alpha_n & \ldots \end{vmatrix}$$

3) Лінійна властивість

Якщо у визначнику n-го порядку Δ деякий i-й рядок $(a_{i1} \quad \dots \quad a_{in})$ є лінійною комбінацією рядків $(b_{i1} \quad \dots \quad b_{in})$ та $(c_{i1} \quad \dots \quad c_{in})$

$$a_{ij} = \lambda b_{ij} + \mu c_{ij}, \quad \forall j = \overline{1..n},$$

то

$$\Delta = \lambda\Delta_1 + \mu\Delta_2,$$

де Δ_1 - визначник, в якому i-й рядок дорівнює $(b_{i1} \quad \dots \quad b_{in})$, а решта ті самі, що й і у Δ, а Δ_2 – визначник, в якому i-й рядок дорівнює $(c_{i1} \quad \dots \quad c_{in})$, а решта такі самі, що й у Δ.

Слід зазначити, що це справедливо й у випадку більшої кількості рядків. Під лінійною комбінацією елементів $x_1,\dots,x_n$ розуміють вираз вигляду $\alpha_1 x_1 + \dots + \alpha_n x_n$, де $\alpha_1,\dots,\alpha_n$ – довільні дійсні числа.

Ці три властивості є основними властивостями визначника. Наступні п'ять властивостей є логічними наслідками розглянутих трьох основних.

4) Визначник з двома однаковими рядками (стовпцями) дорівнює нулю

$$\begin{vmatrix} a_{11} & \dots & \beta_1 & \dots & \beta_1 & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \beta_n & \dots & \beta_n & \dots & a_{nn} \end{vmatrix} = 0$$

5) Множення всіх елементів деякого рядка (стовпця) визначника на число – це те саме, що й множення визначника на це число α:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ \alpha \cdot a_{i1} & \dots & \alpha \cdot a_{in} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \alpha \cdot \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{i1} & \dots & a_{in} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

6) Якщо всі елементи деякого рядка (стовпця) визначника дорівнюють нулеві, то і сам визначник дорівнює нулеві.

7) Якщо до елементів деякого рядка (стовпця) додати лінійну комбінацію відповідних елементів інших рядків, то значення визначника не зміниться.

8) Теорема анулювання

Сума добутку елементів будь-якого рядка (стовпця) визначника на відповідні алгебраїчні доповнення елементів будь-якого іншого рядка (стовпця) дорівнюють нулеві:

$$a_{k1}A_{i1} + a_{k2}A_{i2} + \ldots + a_{kn}A_{in} = 0, \qquad \forall i,k \quad i \neq k$$

9) Визначник добутку та суми матриць

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

$$\det(A + B) \neq \det A + \det B$$

Серед методів обчислень визначника виділяють:

1) Використання загальної формули, яка виражає значення визначника через значення його елементів.

2) Використання формули розвинення за рядком (стовпцем) (з використанням властивості 7).

3) Зведення визначника до трикутного вигляду (з використанням властивості 7).

Розглянемо визначник верхньої трикутної матриці відносно головної діагоналі:

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}$$

Розкладаючи його за першим стовпцем отримаємо:

$$\Delta_n = \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix} \cdot a_{11}$$

Продовжуючи виконувати аналогічні дії, прийдемо до наступного виразу:

$$\Delta_n = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

Таким чином, визначник верхньої трикутної матриці відносно головної діагоналі дорівнює добутку елементів головної діагоналі.

Очевидно, такий самий результат можна отримати й у випадку нижньої трикутної матриці відносно головної діагоналі.

Приклад:

$$\begin{vmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -2 & -3 & \dots & 0 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & \dots & n \\ 0 & 2 & 6 & \dots & 2n \\ 0 & 0 & 3 & \dots & 2n \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n \end{vmatrix} = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = n!$$

Слід зазначити, що в будь-якому методі можливе використання властивостей на проміжних етапах.

4) Метод рекурентних співвідношень.

Поняття „рекурентність” означає, що об'єкт, який отримують на деякому n-ому кроці, залежить від того, що було отримано на деякій кількості попередніх кроків. Стосовно визначників „рекурентність” означає, що деякий визначник n-го порядку може бути виражений через визначники, які мають такий самий вигляд, але менший порядок (*n-1, n-2*,…).

Приклад:

$$\Delta_n = \begin{vmatrix} 2 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ 0 & 1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 2 & 1 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ 0 & 1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 1 & \dots & 0 \\ 0 & 1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 \end{vmatrix} =$$

$$= 2\Delta_{n-1} - \Delta_{n-2}$$

Таким чином, рекурентне співвідношення має вигляд:

$$\Delta_n = 2\Delta_{n-1} - \Delta_{n-2}$$

При цьому

$$\Delta_1 = 2 \quad \Delta_2 = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3$$

Із записаного рекурентного співвідношення можна отримати:

$$\Delta_n - \Delta_{n-1} = \Delta_{n-1} - \Delta_{n-2} = \ldots = \Delta_3 - \Delta_2 = \Delta_2 - \Delta_1 = 1$$

Таким чином, маємо, що

$$\Delta_n = \Delta_{n-1} + 1 = n + 1$$

<u>Контрольні запитання:</u>

1. Чим відрізняється алгебраїчне доповнення від мінора?
2. Які властивості має визначник?
3. Які є методи обчислення визначників?

<u>Завдання для самостійної роботи</u>

1. Довести властивості визначника.
2. Вивести формулу для обчислення визначника трикутної матриці відносно побічної діагоналі.
3. Чому дорівнює визначник суми матриць?

<u>Література</u>: [1], с.19-39; [4], с.182-200; [3], с.102-121

1.3. Обернені матриці. Матричні рівняння

В першій темі розглядалося, які операції можна застосовувати до матриць. Проводячи аналогію між матрицями і звичайними числами, можна зазначити, що були застосовані всі операції крім ділення.

Введемо аналог операції ділення чисел в множині матриць.

Слід зазначити, що матриці не ділять одна на одну, тому запис вигляду A/B – некоректний та вважається помилкою.

Розглянемо квадратну матрицю n-го порядку

$$A = \{a_{ij}\}, \quad i, j = \overline{1..n}.$$

Матрицю A будемо називати **невиродженою**, якщо її визначник відмінний від нуля $(\det A \neq 0)$, та **виродженою** - якщо її визначник дорівнює нулю ($\det A = 0$).

Матриця B є **оберненою до квадратної матриці** A (або **оберненою матрицею**), якщо виконується наступне відношення:

$$A \cdot B = B \cdot A = I$$

Обернену матрицю позначатимемо як A^{-1}.

Таким чином, за визначенням

$$A \cdot A^{-1} = A^{-1} \cdot A = I \tag{1.6}$$

Зазначимо, що виходячи з умов узгодженості, обернена матриця A^{-1} також є квадратною матрицею *n*-го порядку.

Наступний крок, виходячи з визначеного вище поняття оберненої матриці, полягатиме в знаходженні умов існування оберненої матриці. Для цього необхідне ще одне поняття.

Матрицю A^* називають **приєднаною до матриці** *A*, якщо кожен її елемент a_{ij}^* є алгебраїчним доповненням до елемента a_{ij}^T транспонованої матриці A^T, або інакше:

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

де A_{ij} – алгебраїчне доповнення елементів початкової матриці.

$$A^* = \begin{pmatrix} A_{11}^T & A_{12}^T & \dots & A_{1n}^T \\ A_{21}^T & A_{22}^T & \dots & A_{2n}^T \\ \dots & \dots & \dots & \dots \\ A_{n1}^T & A_{n2}^T & \dots & A_{nn}^T \end{pmatrix},$$

де A_{ij}^T – алгебраїчне доповнення елементів транспонованої матриці.

Для приєднаної матриці A^* має місце наступна рівність:

$$A \cdot A^* = A^* \cdot A = \det A \cdot I \qquad (1.7)$$

Доведемо її справедливість.

□

Розглянемо добуток $A \cdot A^*$:

$$A \cdot A^* = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} A_{11} & \dots & A_{j1} & \dots & A_{n1} \\ A_{12} & \dots & A_{j2} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ A_{1n} & \dots & A_{jn} & \dots & A_{nn} \end{pmatrix} \equiv T$$

Елемент добутку з індексами *i* та *j* буде дорівнювати:

$$t_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \sum_{k=1}^{n} a_{ik}A_{jk} =$$

$$= \left[\text{згідно теореми анулювання}\right] = \begin{cases} 0, & i \neq j \\ \sum\limits_{k=1}^{n} a_{ik}A_{ik} & i = j \end{cases}$$

Згадавши, що вираз $\sum\limits_{k=1}^{n} a_{ik}A_{ik}$ являється формулою розвинення визначника матриці A за i-м рядком, можна отримати:

$$T = A \cdot A^* = \begin{pmatrix} \det A & 0 & \dots & 0 \\ 0 & \det A & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \det A \end{pmatrix} = diag(\det A, \det A, ..., \det A) = \det A \cdot I$$

Аналогічним чином доводиться друга частина рівності (1.7).

■

Теорема (критерій існування оберненої матриці)

Для того, аби існувала обернена матриця A^{-1} до матриці A, необхідно та достатньо, щоб матриця A була невиродженою, тобто, $\det A \neq 0$

□

Необхідність

Нехай для матриці A існує обернена матриця A^{-1}. При цьому, згідно (1.6),

$$A^{-1} \cdot A = I$$

та

$$\det(A^{-1} \cdot A) = \det I$$

Згідно властивостей визначника

$$\det(A^{-1} \cdot A) = \det A \cdot \det A^{-1}$$

Таким чином,

$$\det A \cdot \det A^{-1} = 1$$

Отже,

$$\det A \neq 0 \quad , \quad \det A^{-1} \neq 0$$

Наслідок

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

Достатність

Нехай матриця A невироджена. Розглянемо матрицю B, елементи якої будуть визначатися наступним чином:

$$b_{ij} = \frac{A_{ji}}{\det A} \quad i, j = \overline{1..n}$$

тобто

$$B = \frac{1}{\det A} A^* \tag{1.8}$$

Тоді добуток матриць A та B дає одиничну матрицю:

$$A \cdot B = A \cdot \frac{1}{\det A} \cdot A^* = \frac{1}{\det A} A \cdot A^* = \left[\text{згідно} \quad (2)\right] = \frac{1}{\det A} \cdot \det A \cdot I = I$$

Аналогічно можна показати, що $B \cdot A = I$. Таким чином, згідно визначення оберненої матриці (1.6) матриця B, яка представлена виразом (1.8), є оберненою матрицею до матриці A:

$$A^{-1} = B = \frac{1}{\det A} \cdot A^* \tag{1.8a}$$

■

Формула (1.8а) дає метод знаходження оберненої матриці, який називають **методом приєднаної матриці.**

Теорема (про єдиність оберненої матриці)

Для невиродженої матриці *A* існує лише одна обернена матриця.

□

Використаємо метод доведення від супротивного. Припустимо, що існують дві обернені матриці B_1 та B_2. За визначенням

$$AB_1 = B_1A = I\ ,\ AB_2 = B_2A = I \qquad (1.9)$$

Розглянемо рівність $AB_1 = I$. Помножимо її зліва на B_2:

$$B_2AB_1 = B_2I = B_2 \qquad (1.10)$$

Згідно (1.9) $B_2A = I$, тому ліву частину (1.10) можна переписати в іншому вигляді:

$$B_2AB_1 = IB_1 = B_1 \qquad (1.11)$$

З (1.10) та (1.11) випливає, що $B_2 = B_1$ та матриця A^{-1} - єдина.

■

Очевидно, що властивість „*бути оберненою матрицею*” є взаємною в тому розумінні, що якщо *B* є оберненою матрицею до *A*, то *A* є оберненою до *B*. Тобто $(A^{-1})^{-1} = A$.

Обернена матриця має також наступні властивості:

1) $(A^{-1})^K = (A^K)^{-1}$

2) $(AB)^{-1} = B^{-1}A^{-1}$

3) $(A^T)^{-1} = (A^{-1})^T$

Приклад:

Знайти обернену матицю до матриці $A = \begin{pmatrix} 5 & -7 \\ 13 & -6 \end{pmatrix}$.

$$\det A = -30+91 = 61$$

$$A^T = \begin{pmatrix} 5 & 13 \\ -7 & -6 \end{pmatrix}$$

$$A_{11}^T = |-6| = -6 \qquad A_{12}^T = -|-7| = 7 \qquad A_{21}^T = -|13| = -13 \qquad A_{22}^T = |5| = 5$$

$$A^* = \begin{pmatrix} -6 & 7 \\ -13 & 5 \end{pmatrix}$$

$$A^{-1} = \frac{1}{61}\begin{pmatrix} -6 & 7 \\ -13 & 5 \end{pmatrix}$$

Як можна побачити, у випадку невиродженої матриці 2-го порядку для пошуку оберненої необхідно елементи головної діагоналі поміняти місцями, а у елементів побічної діагоналі змінити знак на протилежний і кожен з них поділити на визначник матриці A.

Ще один метод знаходження оберненої матриці – **метод елементарних перетворень.**

Під **елементарними перетвореннями** матриці будемо розуміти наступні дії:

1) перестановка рядків (стовпців) місцями;

2) множення рядка (стовпця) на число, відмінне від нуля;

3) додавання до елементів рядка (стовпця) відповідних елементів іншого рядка (стовпця), попередньо помножених на деяке число.

В основі методу елементарних перетворень лежить наступна

Теорема (про елементарні перетворення матриці)

Якщо до одиничної матриці n-го порядку застосувати ті самі перетворення над рядками і в тому самому порядку, за допомогою яких не вироджена матриця A n-го порядку зводиться до одиничної, то отримана при цьому матриця буде оберненою до A:

$$A \xrightarrow{\text{елем. перетвор.}} I \xrightarrow{\text{елем. перетвор.}} B \Rightarrow B = A^{-1}$$

Приймемо цю теорему без доведення.

Приклад:

Знайти методом елементарних перетворень обернену матрицю до матриці A:

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

Застосуємо до неї наступні елементарні перетворення:

- від 1-го рядка віднімемо 2-й
- від 2-го рядка віднімемо 3-й
- …………………………..
- від *n*-го рядка віднімемо (*n-1*)-й

В результаті отримаємо одиничну матрицю

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

Тоді, виконавши аналогічні дії над одиничною матрицею, ми отримаємо обернену

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

Один із аспектів практичного застосування ділення числе полягає в розв'язанні рівняння виду $a \cdot x = b$. Його корінь $x = \dfrac{b}{a}$. Або, переписавши в іншому вигляді:

$$x = \frac{b}{a} = b \cdot \frac{1}{a} = b \cdot a^{-1}$$

Тобто розв’язок рівняння записаний через обернений елемент до множника в лівій частині. Узагальнимо це на випадок матриць, враховуючи, що добуток матриць не комутативний.

Розглянемо різні можливі варіанти:

1) $A \cdot X = B$

Щоб позбавитися від матриці A в лівій частині, помножимо обидві частини рівняння зліва на матрицю A^{-1}:

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

З урахуванням визначення оберненої матриці

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

2) $X \cdot C = B$

Аналогічно до попереднього, але помноживши обидві частини рівняння зправа на матрицю C^{-1}:

$$X = BC^{-1}$$

3) $A \cdot X \cdot C = B$

$$X = A^{-1}BC^{-1}$$

Слід зазначити, що в останніх виразах для розв’язків рівнянь матриці A та C – квадратні, X та B – довільні, які задовольняють умові узгодженості.

Контрольні питання:

1. Що таке обернена матриця?
2. Які властивості оберненої матриці?
3. В чому полягає метод приєднаної матриці?

Завдання для самостійної роботи:

1. Довести властивості оберненої матриці.
2. Чи можна ввести поняття оберненої матриці на випадок виродженої та прямокутної матриць?

Література: [1], с.39-41; [2], с. 198-201; [3], с. 93-96

1.4. Лінійна залежність та незалежність. База системи елементів

Розглянемо множину алгебраїчних векторів $x_1, x_2, ..., x_n$.

Лінійною комбінацією векторів $x_1, x_2, ..., x_n$ називають суму добутків цих векторів на дійсні числа:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n \tag{1.12}$$

Тривіальною лінійною комбінацією називають лінійну комбінацію виду (1.12), де всі коефіцієнти дорівнюють нулеві:

$$\alpha_i = 0, \ \ i = \overline{1..n}$$

Очевидно, що тривіальна лінійна комбінація завжди дорівнює нулеві:

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0,$$

де в правій частині – нуль-вектор.

Якщо в лінійній комбінації виду (1.12) хоча б один з коефіцієнтів відмінний від нуля, тобто $\sum_{i=1}^{n} |\alpha_i| \neq 0$, то це **нетривіальна лінійна комбінація**.

Система векторів $x_1, x_2, ..., x_n$ називається **лінійно залежною**, якщо існує нетривіальна лінійна комбінація цих векторів, яка дорівнює нулеві:

$$\begin{cases} \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0 \\ \qquad\qquad \text{при} \quad \sum_{i=1}^{n} |\alpha_i| \neq 0 \end{cases} \tag{1.13}$$

Лінійно незалежною називають таку систему векторів, для якої не існує нетривіальної нульової лінійної комбінації, тобто:

$$\begin{cases} \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0 \\ \sum_{i=1}^{n} |\alpha_i| = 0 \end{cases} \qquad (1.14)$$

Іншими словами, система векторів буде лінійно незалежною, якщо лише їх тривіальна лінійна комбінація дорівнює нулю. Тоюбто, з того, що лінійна комбінація векторів дорівнює нулеві, випливає, що вона тривіальна:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0 \Rightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$$

<u>*Приклад:*</u>

Нехай задано систему векторів:

$$\begin{aligned} e_1 &= \begin{pmatrix} 1, & 0, & 0, & \ldots, & 0 \end{pmatrix} \\ e_2 &= \begin{pmatrix} 0, & 1, & 0, & \ldots, & 0 \end{pmatrix} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ e_n &= \begin{pmatrix} 0, & 0, & 0, & \ldots, & 1 \end{pmatrix} \end{aligned}$$

Дана система векторів є лінійно незалежною:

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow \begin{matrix} \alpha_1 = 0 \\ \alpha_2 = 0 \\ \ldots \\ \alpha_n = 0 \end{matrix}$$

Очевидно, що один ненульовий вектор $x_1 \neq 0$ завжди утворює лінійно незалежну систему. Властивість системи векторів $x_1, x_2, ..., x_n$ бути лінійно незалежною жодним чином не пов'язана з порядком векторів, оскільки доданки в (1.12) можуть бути переставлені в довільному порядку.

Слід зазначити, що замість векторів система може складатися з елементів довільного лінійного простору.

Зазначимо ряд тверджень стосовно лінійної залежності та незалежності:

1) якщо серед елементів системи $x_1, x_2, ..., x_n$ є нульовий елемент, то така система лінійно залежна;

2) якщо в системі векторів $x_1, x_2, ..., x_n$ є два однакові елементи, то така система лінійно залежна;

3) якщо частина системи векторів $x_1, x_2, ..., x_n$ (підсистема) лінійно залежна, тоді вся система також лінійно залежна;

4) якщо система векторів $x_1, x_2, ..., x_n$ лінійно незалежна, тоді й будь-яка її підсистема також лінійно незалежна;

5) якщо дописати вектор y до лінійно незалежної системи векторів $x_1, x_2, ..., x_n$ і вона стане лінійно залежною, тоді вектор y можна представити у вигляді лінійної комбінації векторів системи.

Розглянемо критерій **лінійної залежності системи векторів**.

Теорема (критерій лінійної залежності)

Для того, щоб вектори $x_1, x_2, ..., x_n$ були лінійно залежні, необхідно і достатньо, щоб хоча б один з них лінійно виражався через інші вектори.

□

Необхідність

Нехай вектори $x_1, x_2, ..., x_n$ лінійно залежні, тобто справедливий вираз (1.13).

При цьому хоча б один з коефіцієнтів α_i відмінний від нуля. Нехай для визначеності це коефіцієнт α_k, $1 \le k \le n$. Тоді, поділивши обидві частини на значення α_K, перенісши решту елементів $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$ в праву частину та ввівши позначення $\beta_i = -\dfrac{\alpha_i}{\alpha_k}$, отримаємо, що:

$$x_k = \beta_1 x_1 + \ldots + \beta_{k-1} x_{k-1} + \beta_{k+1} x_{k+1} + \ldots + \beta_n x_n$$

тобто елемент x_k представлений у вигляді лінійної комбінації інших векторів.

Достатність

Нехай один з елементів лінійно виражається за допомогою інших (для визначеності візьмемо елемент x_j):

$$x_j = \gamma_1 x_1 + \ldots + \gamma_{j-1} x_{j-1} + \gamma_{j+1} x_{j+1} + \ldots + \gamma_n x_n$$

Перепишемо цей вираз наступним чином:

$$\gamma_1 x_1 + \ldots + \gamma_{j-1} x_{j-1} + (-1) x_j + \gamma_{j+1} x_{j+1} + \ldots + \gamma_n x_n = 0$$

В результаті отримано лінійну комбінацію векторів $x_1, x_2, ..., x_n$, яка дорівнює нулеві. Ця комбінація нетривіальна, оскільки серед її коефіцієнтів хоча б один (а саме $\gamma_j = -1$) відмінний від нуля. Отже, система векторів $x_1, x_2, ..., x_n$ лінійно залежна.

■

Нехай в системі векторів $x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n$ перші k векторів лінійно незалежні, а вся система лінійно залежна. Вектори $x_1, x_2, ..., x_k$ будемо називати базою системи векторів $x_1, x_2, ..., x_n$. Тобто **база системи векторів** – це максимальна лінійно незалежна підсистема заданої системи векторів. Кількість векторів в базі називається **рангом** системи.

Контрольні питання:

1. Яка система елементів називається лінійно залежною, а яка лінійно незалежною?
2. Що таке ранг та база системи векторів?

Література: [1], с. 51-53; [3], с. 68-69.

1.5. Ранг матриці. Базисні мінори. Методи обчислення рангу

Розглянемо довільну прямокутну матрицю

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

Мінором k-го порядку матриці A називають визначник k-го порядку, складений з елементів, які розташовані на перетині будь-яких k рядків і будь-яких k стовпців матриці A. Зрозуміло, що k не перевищує мінімального з порядків m та n.

Припустимо, що хоча би один з елементів a_{ij} матриці A відмінний від нуля. Тоді знайдеться таке ціле додатнє число r, що будуть виконуватись наступні дві умови:

1) в матриці A є мінор r-го порядку, який відмінний від нуля;
2) будь-який мінор $(r+1)$-го та вищого порядку (якщо такі існують) дорівнює нулеві.

Число r, що задовольняє цим вимогам, називають **рангом матриці** A та позначають rang A.

Зазначимо, що ранг нульової матриці дорівнює нулеві.

Мінор r-го порядку, який відмінний від нуля, називають **базисним мінором**. Рядки та стовпці, на перетині яких знаходиться базисний мінор, називають відповідно **базисними рядками і базисними стовпцями**.

Слід зазначити, що матриця A може мати декілька ненульових мінорів r-го порядку. Для ненульової матриці базисний мінор завжди існує.

Теорема (про базисний мінор)

Базисні рядки (стовпці) лінійно незалежні. Будь-який рядок (стовпець) матриці A є лінійною комбінацією базисних рядків (стовпців).

□

Всі міркування наводяться для рядків. Рядки матриці можна розглядати як систему векторів $x_1 = \begin{pmatrix} a_{11} & \dots & a_{1n} \end{pmatrix}, \dots, x_m = \begin{pmatrix} a_{m1} & \dots & a_{mn} \end{pmatrix}$. Якби базисні рядки були лінійно залежними, тоді один з цих рядків був би лінійною комбінацією інших базисних рядків. Тоді, можна було б, не змінюючи величину базисного мінору, відняти від цього рядка вказану лінійну комбінацію, отримавши рядок, який складався б з нулів. Але це суперечило б тому, що базисний мінор відмінний від нуля. Тобто базисні рядки лінійно залежні.

Покажемо далі, що будь-який рядок матриці A є лінійною комбінацією базисних рядків. Оскільки при довільних перестановках рядків (або стовпців) визначник дорівнює нулеві, можна вважати, що базисний мінор знаходиться в лівому верхньому куту матриці A, тобто розташований на перших r рядках та перших r стовпцях. Нехай l – будь-яке число від 1 до n, а k – будь-яке число від 1 до m.

Розглянемо визначник ($r+1$)-го порядку

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2l} \\ \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rl} \\ a_{k1} & a_{k2} & \dots & a_{kr} & a_{kl} \end{vmatrix}$$

Цей визначник дорівнює нулеві при будь-яких значеннях k та l. Дійсно, якщо $k \le r$ або $l \le r$, то у визначнику буде два однакових стовпця або два однакових рядка, і за властивостями він буде дорівнювати нулеві. Якщо ж обидва числа k та l перевищують r, тоді він є мінором ($r+1$)-го порядку матриці A, а будь-який такий мінор дорівнює нулеві за визначенням базисного мінору.

Тобто розглянутий визначник дорівнює нулеві при будь-яких значеннях $k = \overline{1..m}$ та $l = \overline{1..n}$. Але тоді, розклавши цей визначник за останнім стовпцем та позначивши незалежні від номера l алгебраїчні доповнення елементів цього стовпця символами $A_{1l} = c_1$ $A_{2l} = c_2 \ldots$ $A_{rl} = c_r$ $A_{k1} = c_{r+1}$, можна отримати, що

$$c_1 a_{1l} + c_2 a_{2l} + \ldots + c_r a_{rl} + c_{r+1} a_{kl} = 0$$

для всіх $l = \overline{1..n}$.

Враховуючи, що алгебраїчне доповнення $A_{kl} = c_{r+1}$ співпадає з відмінним від нуля базовим мінором, то можна поділити на c_{r+1}.

Позначивши

$$\alpha_1 = -\frac{c_1}{c_{r+1}} \quad \alpha_2 = -\frac{c_2}{c_{r+1}} \quad \ldots \quad \alpha_r = -\frac{c_r}{c_{r+1}}$$

отримаємо, що

$$a_{kl} = \alpha_1 a_{1l} + \alpha_2 a_{2l} + \ldots + \alpha_r a_{rl} \qquad l = \overline{1..n}$$

Це й означає, що k-й рядок є лінійною комбінацією перших r (тобто в даному випадку базисних) рядків.

■

Наслідки

1) будь-який не базисний рядок (стовпець) матриці A є лінійною комбінацією всіх рядків (стовпців) матриці;

2) максимальна кількість лінійно незалежних рядків (стовпців) дорівнює рангу матриці.

Тобто ранг матриці можна також визначити як максимальне число лінійно незалежних рядків.

Можна стверджувати, що в будь-якій матриці максимальне число лінійно незалежних рядків збігається з максимальним числом лінійно незалежних стовпців.

Розглянемо далі методи пошуку рангу матриці. Перший метод – це метод **обвідних мінорів**.

Мінор M' називають **обвідним** для мінора M (обидва мінори будуються для матриці A), якщо мінор M отримують з мінора M' за допомогою викреслювання одного рядка і одного стовпця. Іншими словами M' – **обвідний** мінор для мінора M, якщо він отриманий шляхом дописування до мінора M одного рядка і одного стовпця матриці A (які не входять в M).

Важливо, що обвідний мінор $(k+1)$-го порядку містить в собі повністю мінор k-го порядку, для якого він є обвідним.

Приклад:

Збудувати для заданої матриці послідовність обвідних мінорів

$$A = \begin{pmatrix} 1 & 3 & 5 & -1 & 2 \\ 2 & -1 & -3 & 4 & 4 \\ 5 & 1 & -1 & 7 & 10 \\ 7 & 7 & 9 & 1 & 14 \end{pmatrix}$$

Як ненульові мінори першого та другого порядків розглянемо:

$$M_1 = \left|-3\right| \quad , \quad M_2 = \begin{vmatrix} -3 & 4 \\ -1 & 7 \end{vmatrix}$$

До цього мінора рядок можна дописати двома способами, стовпець – трьома. Отже, для цього мінора можна побудувати 6 обвідних мінорів третього порядку:

$$M_3^{/} = \begin{vmatrix} -3 & 4 & 4 \\ -1 & 7 & 10 \\ 9 & 1 & 14 \end{vmatrix} \; M_3^{//} = \begin{vmatrix} 1 & 5 & -1 \\ 2 & -3 & 4 \\ 5 & -1 & 7 \end{vmatrix} \; M_3^{///} = \begin{vmatrix} -1 & -3 & 4 \\ 1 & -1 & 7 \\ 7 & 9 & 1 \end{vmatrix} ..$$

<u>Теорема (про метод обвідних мінорів)</u>

При обчисленні рангу матриці A слід переходити від мінорів менших порядків до мінорів вищих порядків. Якщо для матриці A вже знайдений мінор r-го порядку, який не дорівнює нулеві, то потребують обчислення лише обвідні мінори $(r+1)$-го порядку. Якщо всі вони дорівнюють нулеві, тоді $rangA = r$.

Метод обвідних мінорів дозволяє визначити не тільки ранг, а й базисні рядки та стовпці.

Другий метод обчислення рангу – це метод **елементарних перетворень.**

<u>Теорема (про елементарні перетворення)</u>

Елементарні перетворення не змінюють рангу матриці. Якщо матрицю B отримують з матриці A за допомогою елементарних перетворень, то ранги цих матриць співпадають.

Доведення теореми базується на наступному твердженні.

<u>Твердження</u>

Якщо задана матриця A виду

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & a_{1k+1} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2k} & a_{2k+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{kk} & a_{kk+1} & \dots & a_{kn} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

то тоді $rangA = k$

Контрольні запитання:

1. Що таке ранг матриці?
2. Як будуються обвідні мінори?
3. Що визначає базисний мінор?

Завдання для самостійної роботи

1. Довести твердження про лінійну незалежність та залежність
2. Довести твердження про ранг трапецієподібної матриці

Література: [1], с. 42-44, 62, 87-88; [2], с. 74-76, 125-127.

2. СИСТЕМИ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ

2.1. Системи лінійних алгебраїчних рівнянь. Основні поняття.

У загальному випадку система m лінійних алгебраїчних рівнянь з n невідомими має наступний вигляд:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{2.1}$$

При цьому через $x_1, x_2, ..., x_n$ позначені невідомі, які потрібно знайти. Величини a_{ij}, $i = \overline{1...m}$, $j = \overline{1...n}$ називають коефіцієнтами системи, а величини b_j, $i = \overline{1...m}$ називають вільними членами. Ці величини припускаються відомими.

Досить зручно записувати лінійну систему (2.1) в матричній формі.

Для цього побудуємо:
матрицю з коефіцієнтів при невідомих, яку називають матрицею системи

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \tag{2.2}$$

вектор-стовпець з невідомих

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{2.3}$$

та вектор-стовпець вільних членів

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{2.4}$$

Матриці A та x узгоджені, тому їх можна перемножити:

$$Ax = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \ldots \\ a_{m_1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = \left[\text{згідно}(2.1)\right] = \begin{pmatrix} \mathrm{b}_1 \\ \vdots \\ \mathrm{b}_\mathrm{m} \end{pmatrix} = \left[\text{згідно}(2.4)\right] = \mathrm{b}$$

Тобто матричний запис системи (2.1) можна замінити одним еквівалентним матричним рівнянням:

$$Ax = b \tag{2.5}$$

де матриці A, x, b визначаються відношеннями (2.2), (2.3) та (2.4) відповідно.

Надалі разом з матрицею системи A буде використовуватись також розширена матриця системи $\tilde{A}$, яку отримують шляхом дописування до матриці A вектора правої частини *b*:

$$\tilde{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right) = \left(A \,|\, b \right) \tag{2.6}$$

Якщо всі вільні члени $b_1, b_2, ..., b_m$ дорівнюють нулеві $\left(b \equiv 0 \right)$, то система (2.1) та (2.5) називається **однорідною**. Інакше, якщо хоча б один з вільних членів відмінний від нуля $\left(b \not\equiv 0 \right)$, система називається **неоднорідною**.

У випадку, коли кількість рівнянь системи m дорівнює кількості невідомих n, система (2.1) називається **квадратною**.

Розв'язком системи (2.1) називається така сукупність n чисел $c_1, c_2, ..., c_n$, яка при підстановці в систему на місце невідомих $x_1, x_2, ..., x_n$ перетворює всі рівняння цієї системи в тотожності.

Якщо система записана в матричному вигляді (2.5), то розв'язок системи – це вектор $c = (c_1 ... c_n)$, який перетворює рівність (2.5) в тотожність. Не кожна система має розв'язок.

Приклад:

Розв'язати систему $\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases}$

Наведена система не має жодного розв'язку. Якби розв'язок існував, то виникло протиріччя виду 1=2.

Лінійну систему будемо називати **сумісною**, якщо вона має хоча б один розв'язок, і **несумісною**, якщо в неї не існує жодного розв'язку.

Очевидно, що сумісна система може мати або один розв'язок, або декілька. Два розв'язки сумісної системи

$$c^{(1)} = \begin{pmatrix} c_1^{(1)} \\ c_2^{(1)} \\ \vdots \\ c_n^{(1)} \end{pmatrix} \qquad \text{та} \qquad c^{(2)} = \begin{pmatrix} c_2^{(1)} \\ c_2^{(2)} \\ \vdots \\ c_n^{(2)} \end{pmatrix}$$

називаються різними, якщо $c^{(1)} \not\equiv c^{(2)}$.

Сумісна лінійна система називається **визначеною**, якщо вона має єдиний розв'язок, і **невизначеною**, якщо існують принаймні два різні розв'язки.

Дві системи **еквівалентні (рівносильні)**, якщо будь-який розв'язок однієї з них є розв'язком іншої, та навпаки, тобто множини їх розв'язків співпадають.

Основними питаннями при розгляді лінійних систем є:

1) встановлення сумісності системи;

2) встановлення визначеності системи;

3) пошук єдиного розв'язку системи (у випадку її визначеності) та пошуку всіх її розв'язків (у випадку її невизначеності).

Почнемо з встановлення сумісності лінійної системи.

Розглянемо спочатку однорідну систему виду (2.1):

$$\begin{cases} a_{11}x_1 + ... + a_{1n}x_n = 0 \\ \qquad \dots \\ a_{m_1}x_1 + ... + a_{mn}x_n = 0 \end{cases} \tag{2.7}$$

або в матричному вигляді

$$Ax = 0 \tag{2.7a}$$

Зазначимо те, що однорідна система завжди має тривіальний (нульовий) розв'язок $x_1 = x_2 = ... = x_n = 0 \quad (x = 0)$, тобто однорідна система завжди сумісна.

Постає питання про те, за яких умов однорідна система має, крім тривіального розв'язку, ще й інші розв'язки (тобто є «нетривіально сумісною»).

Відповідь на це полягає в наступному. Існування нетривіального розв'язку системи (2.7) еквівалентне лінійній залежності стовпців матриці (2.2) (оскільки з лінійної залежності стовпців виходить, що існують такі числа $x_1 ... x_n$, які не всі дорівнюють нулю й такі, що справедлива рівність (2.7))

$$A = (\vec{a}_n^{(1)}, \vec{a}_n^{(2)}, ..., a_a^{(n)}) \qquad a^{(j)} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad j = \overline{1...n}$$

Формула (2.7) тоді еквівалентна співвідношенню

$$x_1 a^{(1)} + x_2 a^{(2)} + ... + x_n a^{(n)} = 0 .$$

Але тоді, згідно теореми про базисний мінор, лінійна залежність стовпців матриці A матиме місце тоді і лише тоді, коли не всі стовпці цієї матриці є базисними, тобто тоді і лише тоді, коли порядок її базисного мінору менший за кількість стовпців n.

Теорема (про нетривіальні розв'язки однорідної системи)

Однорідна система має нетривіальні розв'язки тоді і лише тоді, коли ранг її матриці менший за кількість її стовпців:

$$r = rangA < n$$

Наслідок

Квадратна однорідна система має нетривіальні розв'язки тоді і лише тоді, коли визначник її матриці дорівнює нулеві:

$$\det A = 0$$

Лінійною оболонкою векторів $f_1, f_2, ..., f_k$ називатимемо всі можливі лінійні комбінації цих векторів:

$$\Lambda\left(f_1, f_2, ..., f_k\right) = \left\{ f \,\middle|\, f = \sum_{i=1}^{k} \alpha_i f_i , \quad \forall \alpha_i \in R \right\}$$

У загальному випадку необхідну і достатню умову сумісності лінійної системи визначає наступна теорема.

Теорема Крокенера-Капеллі

Для того, щоб лінійна система була сумісною, необхідно і достатньо, щоб ранг розширеної матриці цієї системи дорівнював рангу її основної матриці:

$$rang \quad A = rang \quad \tilde{A}$$

□

Необхідність

Нехай система сумісна, тобто існують такі числа $c_1,...,c_n$, що виконуються рівності

$$\begin{cases} a_{11}c_1 + ... + a_{1n}c_n = b_1 \\ \qquad \dots \\ a_{m1}c_1 + ... + a_{mn}c_n = b_\mathrm{m} \end{cases} \tag{2.8}$$

чи в компактній формі запису

$$c_1 \cdot a^{(1)} + c_2 \cdot a^{(2)} + ... + c_n \cdot a^{(n)} = b \tag{2.8a}$$

Нехай r – ранг основної матриці системи $(r = rangA)$. Розглянемо лінійну оболонку r базисних стовпців матриці. Згідно теореми про базисний мінор будь-який стовпець матриці A належить цій лінійній оболонці. Іншими словами, будь-який стовпець розширеної матриці $\tilde{A}$ також належить розширеній лінійній оболонці (оскільки він, згідно (8), лінійно виражається через всі стовпці основної матриці, а, відповідно, і через базисні).

Тобто, всі стовпці розширеної матриці належать даній лінійній оболонці. Ранг цієї оболонки дорівнює r . Це означає, що будь-які $(r+1)$ стовпці розширеної матриці лінійно залежні, тобто її ранг $rang\tilde{A}$ (рівний максимальному числу лінійно незалежних стовпців матриці) також дорівнює r, тобто $rangA = r = rang\tilde{A}$.

Достатність

Нехай ранги матриць співпадають. Тоді r базисних стовпців матриці A є базисними стовпцями $\tilde{A}$. Згідно теореми про базисний мінор останній стовпець розширеної матриці є лінійною комбінацією r базисних стовбців, а, отже, і всіх стовпців основної матриці. Тобто існують числа $c_1, c_2, ..., c_n$, такі, що справедливі (2.8) і (2.8a). З цього випливає, що числа $c_1, ..., c_n$ є розв'язком лінійної системи.

■

<u>Контрольні питання:</u>

1. Що таке розв'язок лінійної системи?
2. Яка система називається сумісною та несумісною?
3. Коли система матиме розв'язок?

<u>Література</u>: [1], с. 70-75; [2], с. 67-68, 156-160; [3], с. 20-28, 123-125.

2.2. Однорідні та неоднорідні системи. Методи пошуку розв'язків

Теорема Кронекера-Капеллі встановлює необхідну і достатню умову сумісності лінійної системи, проте не дає способу знаходження розв'язків цієї системи.

Розглянемо найбільш використовувані методи пошуку розв'язку лінійної системи

$$Ax = b$$

або в розгорнутому вигляді

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases} \qquad (2.9)$$

Почнемо з окремого випадку, а саме розглянемо квадратну систему лінійних рівнянь

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n \end{cases} \qquad (2.10)$$

з відмінним від нуля визначником основної матриці ($\Delta = \det A \neq 0$).

Покажемо, що дана система має розв'язок (причому єдиний) і знайдемо його.

За теоремою Кронекера-Капеллі достатньо довести, що ранг розширеної матриці $\tilde{A}$ дорівнює рангу основної матриці A. Згідно з умовою $\Delta \neq 0$ ранг основної матриці дорівнює n, а ранг матриці $\tilde{A}$, яка містить n рядків, не може бути більший за число n, тому

$$rang \quad A = rang \quad \tilde{A}$$

тобто система сумісна (має розв'язок). Покажемо тепер, що система може мати лише один розв'язок.

Нехай n чисел $x_1 ... x_n$ такі, що при підстановці в систему (2.9) вони перетворюють всі рівняння в тотожності.

Домножаючи ці тотожності на алгебраїчні доповнення елементів $A_{1j}, A_{2j} ... A_{nj}$ j-го стовпця визначника матриці A і додаючи їх, отримаємо:

$$\sum_{i=1}^{n} x_i \left(a_{1i} A_{1j} + a_{2i} A_{2j} + ... + a_{ni} A_{nj} \right) = b_1 A_{1j} + b_2 A_{2j} + ... + b_n A_{nj}$$

Згідно властивостей визначника (теореми анулювання і формули розкладання за рядком), з останньої рівності отримаємо:

$$x_j \cdot \Delta = b_1 A_{1j} + b_2 A_{2j} + ... + b_n A_{nj}$$

Введемо позначення Δ_j для визначника, який отримують з визначника Δ основної матриці A заміною його j-го стовпця стовпцем з вільних членів із збереженням без зміни решти стовпців визначника Δ .

Тоді останню рівність можна записати у вигляді:

$$x_j \cdot \Delta = \Delta_j \qquad j = \overline{1 ... n}$$

Оскільки $\Delta \neq 0$, то:

$$x_j = \frac{\Delta_j}{\Delta} \qquad j = \overline{1...n} \tag{2.10}$$

Таким чином, розв'язок системи (2.9) з ненульовим визначником основної матриці однозначно визначається формулами (2.10), які називаються **формулами Крамера**.

Доведене твердження про існування і єдиність розв'язку квадратної системи з визначником основної матриці, відмінним від нуля, ще простіше встановлюється матричним способом. Запишемо систему

$$Ax = b \tag{2.11}$$

Оскільки її визначник не дорівнює нулеві ($\Delta = \det A \neq 0$), то A – невироджена, і для неї існує обернена матриця A^{-1}. Припустимо, що існує розв'язок цієї системи, тобто існує стовпець x, що перетворює в тотожність матричне рівняння (2.11).

Його розв'язок має вигляд:

$$x = A^{-1} \cdot b \tag{2.12}$$

Розгорнувши останній вираз і врахувавши вид оберненої матриці, для елементів стовпця x можна отримати формули Крамера.

Легко перевірити, що стовпець x, який визначається рівністю (2.12), є розв'язком (2.11). Тобто, якщо визначник матриці A відмінний від нуля, то існує єдиний розв’язок рівняння (2.11), який визначається співвідношенням (2.12), еквівалентним формулам Крамера (2.10).

Основне призначення формул Крамера – дати явне представлення для розв’язків квадратної системи лінійних алгебраїчних рівнянь через її коефіцієнти та вільні члени. Практичне їх використання пов'язане з досить громіздкими обчисленнями (для розв’язку системи n-го порядку необхідно обчислити $(n+1)$ визначник n-го порядку).

Розглянемо загальну лінійну систему m рівнянь та n невідомих

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ \cdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{cases} \tag{2.13}$$

Нехай дана система сумісна, причому $rangA = rang\tilde{A} = r$.

Не обмежуючи загальності, можна припустити, що базисний мінор матриці A знаходиться в лівому верхньому куті.

Тоді перші r рядків як основної, так і розширеної матриці є базисними рядками. Згідно теореми про базисний мінор, кожен з рядків розширеної матриці, починаючи з $(r+1)$-го, являється лінійною комбінацією перших r рядків.

В термінах системи (2.13) це означає, що кожне з її рівнянь, починаючи з $(r+1)$-го, є лінійною комбінацією перших r рівнянь, тобто будь-який розв’язок перших r рівнянь перетворює в тотожність і всі наступні рівняння цієї системи. Отже, достатньо знайти всі розв’язки перших r рівнянь. Запишемо їх та перенесемо невідомі $x_{r+1}...x_n$ (тобто ті, що не відносяться до базисного мінору) в праву частину:

$$\begin{cases} a_{11}x_1 + ... + a_{1r}x_r = b_1 - a_{1r+1}x_{r+1} - ... - a_{1r}x_r \\ \qquad\qquad\qquad\qquad \cdots \\ a_{1r}x_1 + ... + a_{rr}x_r = b_r - a_{rr+1}x_{r+1} - ... - a_{rn}x_n \end{cases} \tag{2.14}$$

Якщо надати невідомим $x_{r+1}...x_n$ абсолютно довільні значення $c_{r+1}...c_n$, то система (2.14) перетвориться на квадратну r-го порядку відносно невідомих $x_1...x_r$ з ненульовим визначником матриці системи, який збігається з базисним мінором.

Тоді існує єдиний розв'язок цієї системи, який можна отримати скориставшись формулами Крамера.

Перш, ніж записати розв'язок, необхідно ввести позначення. Нехай $M_j(b)$ – визначник, який отримують з базисного мінору M матриці A шляхом заміни j-го стовпця на вектор-стовпець b із збереженням без зміни решти стовпців.

Тоді, використовуючи формули Крамера і лінійну властивість визначника, можна записати розв'язок системи (2.14) в наступному вигляді:

$$x_j = \frac{1}{M}\left[M_j(b) - C_{r+1}M_j(a^{(r+1)}) - ... - C_r M_j(a^{(n)})\right]$$

або, згадуючи, що $c_{r+1}...c_n$ було замінено на $x_{r+1}...x_n$,

$$x_j = \frac{1}{M}\left[M_j(b) - x_{r+1}M_j(a^{(r+1)}) - ... - x_r M_j(a^{(n)})\right] \quad j = \overline{1...r} \qquad (2.15)$$

Тобто значення базисних змінних виражено через вільні змінні (або параметри). Тепер для отримання будь-якого розв’язку досить задати значення вільних змінних і за формулами (2.15) знайти значення базисних. Усі можливі комбінації значень параметрів визначають всю множину розв’язків початкової системи (2.14). Формули (2.15) визначають розв’язок системи (2.14) з теоретичної точки зору. На практиці вони застосовуються трохи інакше.

Приклад:

Знайти загальний розв’язок системи

$$\begin{cases} x_1 - x_2 + x_3 - x_4 = 4 \\ x_1 + x_2 + 2x_3 + 3x_4 = 8 \\ 2x_1 + 4x_2 + 5x_3 + 10x_4 = 20 \\ 2x_1 - 4x_2 + x_3 - 6x_4 = 4 \end{cases}$$

Система сумісна, причому $rangA = rang\tilde{A} = 2$

В якості базисний мінору обирається мінор

$$\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2 \neq 0$$

Відкидаючи два останні рівняння і переносячи змінні x_3, x_4 в праву частину, отримаємо:

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

Виражаємо змінні x_1, x_2:

$$x_1 = 6 - \tfrac{3}{2} x_3 - x_4$$
$$x_2 = 2 - \tfrac{1}{2} x_3 - 2x_4$$

Тобто, розв'язок – вектор-стовпець

$$\left(6 - \tfrac{3}{2} x_3 - x_4, 2 - \tfrac{1}{2} x_3 - 2x_4, x_3, x_4\right)^T$$

У випадку однорідної системи всі вільні члени нульові, і формули (2.15) можна переписати наступним чином:

$$x_j = \frac{-1}{M}\left[x_{r+1} M_j(a^{(r+1)}) + \ldots + x_r M_j(a^{(n)})\right] \tag{2.16}$$

Розв'язок неоднорідної системи (2.14) записують у вигляді

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_r \end{pmatrix} = x_{r+1} f^{(r+1)} + \ldots + x_n f^{(n)} + f^{(0)} \tag{2.17}$$

де

$$f^{(0)}=\left(\frac{M_1(b)}{M},...,\frac{M_r(b)}{M},0,0,...,0\right)^T$$

$$f^{(r+1)}=\left(-\frac{M_1(a^{(r+1)})}{M},...,-\frac{M_r(a^{(r+1)})}{M},1,0,...,0\right)^T$$

....

$$f^{(n)}=\left(-\frac{M_1(a^{(n)})}{M},...,-\frac{M_r(a^{(n)})}{M},0,0,...,1\right)^T$$

Очевидно, що вектори $f^{(r+1)},...,f^{(n)}$ лінійно незалежні.

Формула (2.17) є загальним розв'язком неоднорідної системи. У випадку однорідної системи в цій формулі відсутній останній доданок.

Теорема (про фундаментальну систему розв'язків)

Для однорідної системи $Ax=0$, ранг якої менше за кількість невідомих

$$rangA=r<n$$

існує $(n-r)$ лінійно незалежних розв'язків $e_1,...,e_{n-r}$, і будь-який інший розв'язок представляється у вигляді лінійної комбінації цих розв'язків. Така система називається **фундаментальною системою розв'язків** (або **базисом простору розв'язків**).

Фундаментальна система розв'язків, яка отримана шляхом задання значень параметрів (1,0,…,0) ……(0,0,…,1), називається **нормальною (нормованою**) фундаментальною системою.

Твердження:

1) сума будь-якого розв’язку неоднорідної системи з будь-яким розв’язком відповідної однорідної системи є розв’язком неоднорідної системи;

2) різниця двох довільних розв’язків неоднорідної системи є розв’язком відповідної однорідної системи;

3) сума частинного розв’язку неоднорідної системи і загального розв’язку відповідної однорідної дає загальний розв’язок неоднорідної.

Контрольні запитання:

1. Який вигляд мають формули Крамера і що вони дозволяють визначити?
2. Що таке базисні та вільні змінні?
3. Як у загальному випадку знайти розв’язок неоднорідної системи лінійних алгебраїчних рівнянь?
4. Що таке фундаментальна система розв’язків?

Завдання для самостійної роботи

1. Довести вище сформульовані твердження.
2. Опрацювати методи Гауса та Жордана-Гауса.
3. Переконатися в еквівалентності співвідношень (3) та $(3')$.
4. З’ясувати, що буде у випадку, коли у формулах Крамера $\det A \equiv 0$.

Література: [1], с. 75-87; [2], с. 68-70, 156-160; [3], с. 20-28, 123-125; [4], с. 68-76, 44-47.

3. ЕЛЕМЕНТИ ВЕКТОРНОЇ АЛГЕБРИ

3.1. Вектори на площині та в просторі. Операції над векторами

У першій темі вже розглядалося поняття вектора при введенні матриць, один з порядків яких дорівнював одиниці. Це були алгебраїчні вектори, тобто набори чисел.

У векторній алгебрі розглядаються **вектори** з іншої точки зору. Це **величини**, які, крім свого **чисельного значення**, характеризуються ще й **напрямом**.

Геометричним вектором (далі скорочено просто **вектором**) називають напрямлений відрізок, тобто відрізок, який має початок та кінець. Іншими словами – це **впорядкована пара точок** простору А та В (рис. 3.1).

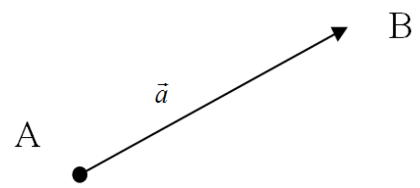

Рис. 3.1. Геометричний вектор

Точка A – початок вектора, точка B – кінець вектора. Початок вектора також називають ще точкою прикладання.

Позначають вектори або за допомогою його початкової та кінцевої точки $\left(\overrightarrow{AB}\right)$, або за допомогою малих латинських літер $\left(\vec{a}\right)$. **Довжина вектора** – (**модуль** вектора) – це довжина відрізка АВ. Довжина позначається як $\left|\overrightarrow{AB}\right|$ або $\left|\vec{a}\right|$.

Вектор, довжина якого дорівнює одиниці, називається **одиничним вектором**, або **ортом**. Вектор, початок та кінець якого співпадають, називають **нульовим вектором**.

Нульовий вектор не має певного напряму, проте має довжину, яка дорівнює нулеві. Це дозволяє при записі ототожнювати нульовий вектор з дійсним числом нуль, тобто не обов'язково його виділяти стрілкою.

Вектори називаються **колінеарними**, якщо вони лежать або на одній прямій, або на паралельних прямих, при цьому вони можуть бути однакового чи різного напряму (рис. 3.2):

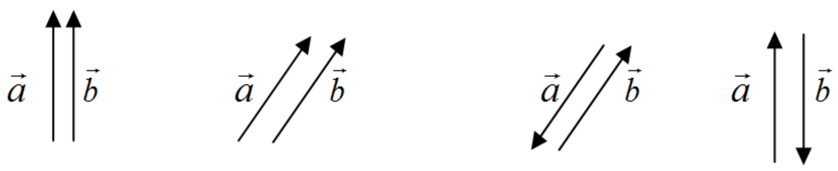

Рис. 3.2. Приклади колінеарних векторів

Нульовий вектор колінеарний до будь-якого вектора.

Два вектори $\vec{a}$ та $\vec{b}$ називаються **рівними**, якщо вони колінеарні, мають однакову довжину і однаковий напрям (рис. 3.3). Всі нульові вектори вважаються рівними.

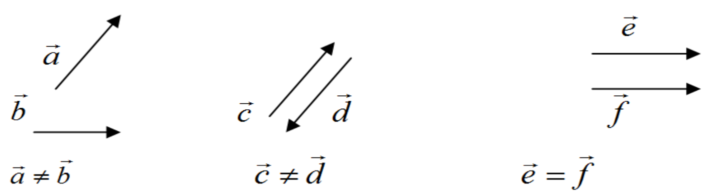

Рис. 3.3. Приклади рівних та нерівних векторів

З визначення рівності векторів безпосередньо випливає, що які б не були вектор $\vec{a}$ і точка P, існує, і притому єдиний, вектор $\overrightarrow{PQ}$ з початком в точці P, рівний вектору $\vec{a}$, тобто точка прикладання даного вектора $\vec{a}$ може бути вибрана довільно. Це дозволяє в подальшому не розрізняти два рівні вектори, що мають різні точки прикладання, і які можна отримати один з одного шляхом паралельного переносу.

Відповідно до того, що було сказано вище, вектори, які вивчаються в геометрії, називають **вільними**.

Вектором, **протилежним** до вектора $\vec{a}$, називається вектор $(-\vec{a})$, що має таку саму довжину, що й $\vec{a}$, та протилежний напрям:

$$\begin{cases} |-\vec{a}| = |\vec{a}| \\ (-\vec{a}) \uparrow\downarrow \vec{a} \end{cases}$$

Визначимо далі лінійні операції над векторами:

1) **Операція додавання векторів**

Сумою $\bar{c} = \bar{a} + \bar{b}$ двох векторів $\vec{a}$ та $\vec{b}$ називають вектор, який проведено від початку вектора $\vec{a}$ в кінець вектора $\vec{b}$ за умови, що вектор $\vec{b}$ прикладений до кінця вектора $\vec{a}$. Це правило додавання векторів називається **правилом трикутника** (рис. 3.4).

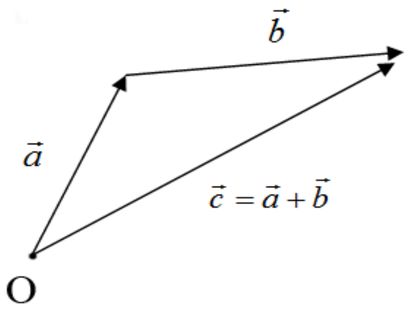

Рис. 3.4. Додавання векторів за правилом трикутника

Властивості додавання векторів такі самі, що й властивості додавання дійсних чисел:

a) $\overline{a} + \overline{b} = \overline{b} + \overline{0}$ (комутативність, або переставна властивість)

b) $(\overline{a} + \overline{b}) + \overline{c} = \overline{a} + (\overline{b} + \overline{c})$ (асоціативність, або сполучна властивість)

c) існує нульовий вектор $\vec{0}$ такий, що $\vec{a} + \vec{0} = \vec{a}$

d) для кожного вектора $\vec{a}$ існує протилежний вектор $(-\vec{a})$ такий, що $\vec{a} + (-\vec{a}) = 0$.

Доведемо першу властивість. Прикладемо два довільні вектори $\vec{a}$ та $\vec{b}$ до загального початку – точки *O*. Позначимо буквами *A* та *B* кінці векторів $\vec{a}$ та $\vec{b}$ і розглянемо паралелограм *OBCA* (рис. 3.5). З нього випливає, що $\overrightarrow{BC} = \overrightarrow{OA} = \vec{a}$, $\overrightarrow{AC} = \overrightarrow{OB} = \vec{b}$

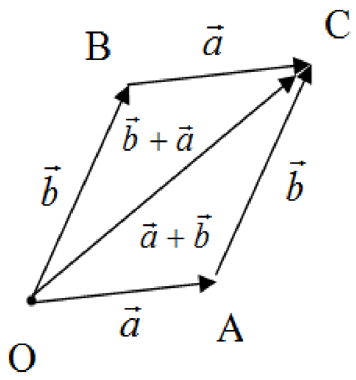

Рис. 3.5. Додавання векторів за правилом паралелограма

Якщо розглянути трикутник *OBC*, то з визначення суми векторів отримаємо $\overrightarrow{OC}=\vec{b}+\vec{a}$. Аналогічно з розгляду трикутника *OAC* отримаємо $\overrightarrow{OC}=\vec{a}+\vec{b}$ тобто $\overrightarrow{OC}=\vec{b}+\vec{a}=\vec{a}+\vec{b}$.

При доведенні цієї властивості було обґрунтувано ще одне правило додавання векторів – **правило паралелограма**: якщо вектори $\vec{a}$ та $\vec{b}$ неколінеарні, прикладені до загального початку і на них побудовано паралелограм, то сума цих векторів є діагоналлю паралелограма, що виходить із загального початку векторів.

Ці властивості дозволяють оперувати з сумою векторів так само, як і з сумою дійсних чисел. Вони дозволяють розповсюдити правило додавання для будь-якого скінченого числа векторів (при цьому немає необхідності виконувати додавання послідовно, фіксуючи кожен проміжний результат). Сума будь-якого числа векторів може бути побудована за допомогою наступного правила:

Якщо прикласти вектор $\vec{a}_2$ до кінця вектора $\vec{a}_1$, вектор $\vec{a}_3$ до кінця вектора $\vec{a}_2$, …, вектор $\vec{a}_n$ до кінця вектора $\vec{a}_{n-1}$, то сума $\vec{c}=\vec{a}_1+\vec{a}_2+...+\vec{a}_n$ буде представляти собою вектор, що виходить з початку вектора $\vec{a}>\vec{b}$ в кінець вектора $\vec{a}_n$ – тобто правило замикання ламаної до многокутника (рис. 3.6).

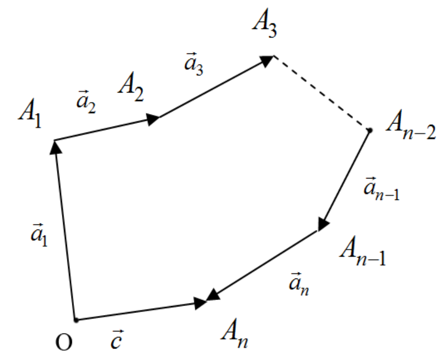

Рис. 3.6. Правило замикання ламаної до многокутника

2) **Операція віднімання векторів**

Різницею $\vec{a}-\vec{b}$ векторів $\vec{a}$ та $\vec{b}$ називається такий вектор $\vec{c}$, який в сумі з вектором $\vec{b}$ дає вектор $\vec{a}$.

Очевидно, що існує, причому єдиний, вектор $\vec{c}$, який є різницею $\vec{a}-\vec{b}$, та дорівнює $\vec{c}=\vec{a}+(-\vec{b})$.

Правило побудови різниці випливає з визначення і правила трикутника. Різниця $\vec{a}-\vec{b}$ приведених до загального початку векторів $\vec{a}$ і $\vec{b}$ є вектором, що йде з кінця вектора $\vec{b}$ в кінець вектора $\vec{a}$ (рис. 3.7).

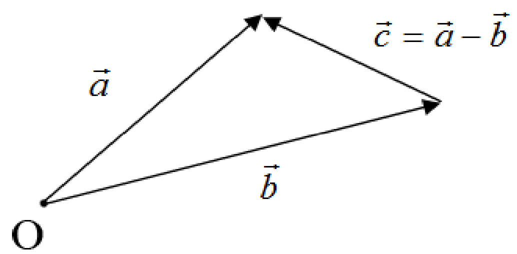

Рис. 3.7. Віднімання векторів

3) **Множення вектора на число**

Добутком $\alpha\vec{a}$ вектора $\vec{a}$ на число α називається вектор $\vec{b}$, колінеарний вектору $\vec{a}$, що має довжину, яка дорівнює $|\alpha|\cdot|\overline{a}|$ та напрям, що співпадає з напрямом вектора $\vec{a}$ при $\alpha > 0$ або протилежний у випадку $\alpha < 0$.

З цього визначення випливає твердження про те що, якщо вектор $\vec{b}$ колінеарний до вектора $\vec{a}$, то існує дійсне таке число λ, що $\overline{b} = \lambda \cdot \overline{a}$ та навпаки.

Властивості операції множення на число:

a) $\alpha\left(\overline{a}+\overline{b}\right)=\alpha\overline{a}+\alpha\overline{b}$ (розподільна властивість);

b) $\left(\alpha+\beta\right)\overline{a}=\alpha\overline{a}+\beta\overline{a}$ (розподільна властивість);

c) $\alpha(\beta\overline{a})=(\alpha\beta)\overline{a}$ (сполучна властивість);

d) $1\cdot\overline{a}=\overline{a}$.

Вектори $\vec{a},\overline{b},\vec{c}$ називають **компланарними**, якщо вони лежать в одній площині.

Сформулюємо ряд важливих властивостей.

<u>Теорема 1</u>

Два вектори лінійно залежні тоді і тільки тоді, коли вони колінеарні.

<u>Наслідок №1</u>

Якщо $\vec{a}$ і $\vec{b}$ не колінеарні, то вони лінійно незалежні.

<u>Наслідок №2</u>

Серед двох не колінеарних векторів не може бути нульового.

Теорема 2

Три вектори $\vec{a}$, $\vec{b}$ і $\vec{c}$ лінійно залежні тоді і тільки тоді, коли вони компланарні.

Наслідок №1

Які б не були два не колінеарні вектори $\vec{a}$ і $\vec{b}$, для будь-якого вектора $\vec{c}$, який лежить в площині векторів $\vec{a}$ і $\vec{b}$ знайдуться дійсні числа α і β такі, що вектор $\vec{c}$ можна представити лінійною комбінацією векторів:

$$\vec{c} = \alpha\vec{a} + \beta\vec{b} \tag{3.1}$$

Наслідок №2

Якщо вектори $\vec{a}, \vec{b}, \vec{c}$ не компланарні, то вони лінійно незалежні.

Наслідок №3

Серед трьох не компланарних векторів не може бути нульового вектора, а також не може бути двох колінеарних векторів.

Теорема 3

Будь-які чотири вектори в просторі лінійно залежні.

□

Неважко зрозуміти що, якщо серед чотирьох векторів будь-які три компланарні, то твердження теореми виконується (оскільки компланарні вектори лінійно залежні, а якщо частина системи лінійно залежна, то і вся система лінійно залежна).

Тому розглянемо випадок, коли серед чотирьох векторів $\vec{a},\vec{b},\vec{c},\vec{d}$ жодна трійка не є компланарною.

Приведемо всі чотири вектори до спільного початку. і проведемо через кінець вектора $\vec{d}$ три площини, які визначаються парами векторів $\{\vec{a},\vec{b}\},\{\vec{b},\vec{c}\},\{\vec{a},\vec{c}\}$ (рис. 3.8). Це можна зробити, оскільки вектори, що входять до складу цих пар, не колінеарні (на основі наслідку 3 з теореми 2).

Точки перетину цих площин з векторами $\vec{a},\vec{b},\vec{c}$ позначимо A, B, C відповідно.

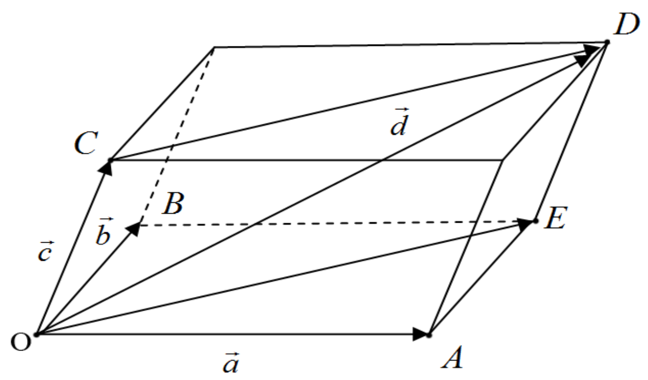

Рис. 3.8. Лінійна залежність чотирьох векторів у просторі

За правилом паралелограма на основі паралелограма OCDE можна записати:

$$\vec{d} = \overrightarrow{OD} = \overrightarrow{OC} + \overrightarrow{OE},$$

а з паралелограма OBEA:

$$\overrightarrow{OE} = \overrightarrow{OA} + \overrightarrow{OB},$$

тобто

$$\vec{d} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} \qquad (3.2)$$

Оскільки вектори $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ колінеарні векторам $\vec{a}, \vec{b}, \vec{c}$ відповідно, то:

$$\overrightarrow{OA} = \alpha\vec{a} \quad \overrightarrow{OB} = \beta\vec{b} \quad \overrightarrow{OC} = \gamma\vec{c} \qquad (3.3)$$

Підставивши (3.3) у (3.2), можна отримати представлення вектора $\vec{d}$ у вигляді лінійної комбінації векторів $\vec{a}, \vec{b}, \vec{c}$:

$$\vec{d} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c} \qquad (3.4)$$

що й свідчить про лінійну залежність векторів $\vec{a}, \vec{b}, \vec{c}, \vec{d}$.

■

<u>Наслідок</u>

Які б не були три некомпланарні вектори $\vec{a}, \vec{b}, \vec{c}$, то для будь-якого довільного вектору $\vec{d}$ існують такі дійсні числа α, β, γ, що справедлива рівність (3.4).

Доведена теорема 3 дає можливість сформулювати ключове поняття базису.

Три лінійно незалежних вектори $\vec{a}, \vec{b}, \vec{c}$ утворюють в просторі **базис**, якщо будь-який вектор $\vec{d}$ цього простору може бути представлений у вигляді їх лінійної комбінації, тобто якщо знайдуться такі числа α, β, γ, що буде справедлива рівність (3.4).

Твердження 1

Будь-яка трійка некомпланарних векторів $\vec{a}, \vec{b}, \vec{c}$ утворює базис в просторі.

Два лінійно незалежних вектори $\vec{a}$ та $\vec{b}$, що лежать у деякій площині, утворюють на цій площині базис, якщо будь-який з векторів $\vec{c}$ цієї площини може бути представлений у вигляді їх лінійної комбінації, тобто знайдуться такі числа α та β, що буде справедлива рівність (3.1).

Твердження 2

Будь-яка пара неколінеарних векторів $\vec{a}$ та $\vec{b}$, що лежать у деякій площині, утворюють базис на цій площині.

Для визначеності далі будемо розглядати базис простору. Все, що буде сформульоване, залишиться справедливим також і для площини.

Нехай $\vec{a}, \vec{b}, \vec{c}$ довільна трійка некомпланарних векторів, тобто базис в просторі. Рівність (3.4) в цьому випадку називається розкладенням вектора $\vec{d}$ за базисом із векторів $\vec{a}, \vec{b}, \vec{c}$. Коефіцієнти цієї комбінації α, β, γ називають **координатами вектора** $\vec{d}$ в базисі векторів $\vec{a}, \vec{b}, \vec{c}$.

Теорема 4

Кожен вектор $\vec{d}$ може бути єдиним чином розкладений за базисом $\vec{a},\vec{b},\vec{c}$. Іншими словами, координати кожного вектора $\vec{d}$ відносно базиса $\vec{a},\vec{b},\vec{c}$ визначаються однозначно.

□

Підемо методом від супротивного. Припустимо, що для деякого вектора $\vec{d}$ разом із розкладенням (3.4) справедливе ще одне розкладення за цим базисом:

$$\vec{d} = \alpha_1\vec{a} + \beta_1\vec{b} + \gamma_1\vec{c} \qquad (3.5)$$

Віднімаючи (3.5) з (3.4), можна отримати:

$$(\alpha_1 - \alpha)\vec{a} + (\beta_1 - \beta)\vec{b} + (\gamma_1 - \gamma)\vec{c} = \vec{0}$$

що, внаслідок лінійної незалежності векторів $\vec{a},\vec{b},\vec{c}$, дає:

$$\begin{cases} \alpha - \alpha_1 = 0 \\ \beta - \beta_1 = 0 \\ \gamma - \gamma_1 = 0 \end{cases}, \text{ або } \begin{cases} \alpha = \alpha_1 \\ \beta = \beta_1 \\ \gamma = \gamma_1 \end{cases}$$

■

Основне значення базиса полягає в тому, що у випадку його задання лінійні операції над векторами перетворюються на звичайні лінійні операції над числами – координатами векторів.

Твердження 3

Для векторів $\vec{d}_1 = (\alpha_1, \beta_1, \gamma_1)$ і $\vec{d}_2 = (\alpha_2, \beta_2, \gamma_2)$, які задані своїми координатами в базисі $\vec{a}, \vec{b}, \vec{c}$, мають місце співвідношення:

$$\vec{d}_1 + \vec{d}_2 = (\alpha_1 + \alpha_2, \beta_1 + \beta_2, \gamma_1 + \gamma_2)$$
$$\lambda \vec{d}_1 = (\lambda\alpha_1, \lambda\beta_1, \lambda\gamma_1)$$

На основі поняття базису можна визначити в просторі систему координат.

Афінна система координат в просторі визначається заданням точки *O*, яка має назву „**початок системи координат**” та базисними векторами $\vec{a}, \vec{b}, \vec{c}$, які прикладені до цієї точки. Прямі, що проходять через базисні вектори, називаються **осями координат**. У будь-якій координатній системі можна визначити координати вектора.

Нехай *M* – довільна точка в просторі. Вектор $\overrightarrow{OM}$ називається радіус-вектором точки *M*. Координати радіус-вектора $\overrightarrow{OM} = (\alpha, \beta, \gamma)$ називаються афінними координатами точки *M* (рис. 3.9):

$$M(\alpha; \beta; \gamma) \qquad \overrightarrow{OM} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}$$

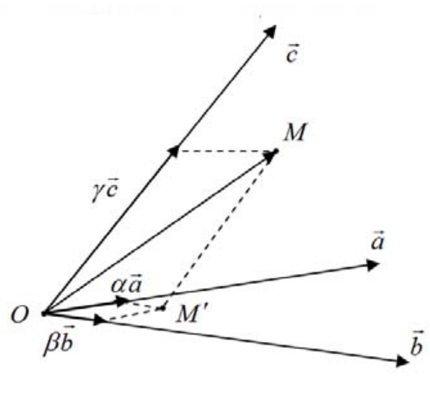

Рис. 3.9. Розклад вектора за базисом

Оскільки розкладення вектора за базисом завжди існує, і притому єдине, то кожній точці *M* однозначно відповідає трійка афінних координат.

Розглянемо в просторі вектор $\overrightarrow{AB}$ та деяку вісь $\vec{u}$. **Ортогональною проекцією** вектора $\overrightarrow{AB}$ на вісь $\vec{u}$ (напрямок $\vec{u}$) називають число, яке дорівнює довжині вектору $\overrightarrow{A'B'}$, яка береться зі знаком «+», якщо він напрямлений в той самий бік, що й вектор $\vec{u}$, або зі знаком «-» – в іншому випадку (рис. 3.10):

$$pr_{\vec{u}}\overrightarrow{AB} = \begin{cases} \left|\overrightarrow{A'B'}\right|, & \overrightarrow{A'B'} \uparrow\uparrow \vec{u} \\ -\left|\overrightarrow{A'B'}\right|, & \overrightarrow{A'B'} \uparrow\downarrow \vec{u} \end{cases} \tag{3.6}$$

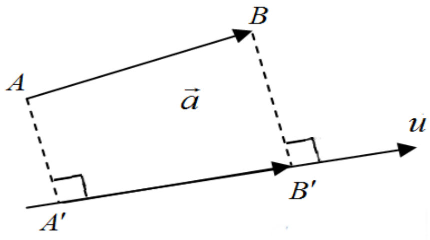

Рис. 3.10. Проекція вектора на вектор

Якщо визначити кут φ нахилу вектора $\vec{a}$ до осі $\vec{u}$ як кут між двома променями, які виходять з довільної точки *M*, причому напрям одного променя співпадає з напрямом вектора $\vec{a}$, а іншого співпадає з напрямом осі $\vec{u}$, то проекція вектора $\overrightarrow{AB}$ на вісь $\vec{u}$ буде дорівнювати довжині вектора $\overrightarrow{AB}$, помноженої на косинус кута φ нахилу вектора $\overrightarrow{AB}$ до осі $\vec{u}$:

$$pr_{\vec{u}}\overrightarrow{AB} = \overrightarrow{AB}\cos\varphi \tag{3.6a}$$

Основні властивості проекції вектора на вісь полягають в тому, що лінійні операції над векторами приводять до відповідних лінійних операцій над проекціями цих векторів, а саме:

$$pr_{\vec{u}}(\vec{a}+\vec{b}) = pr_{\vec{u}}\vec{a} + pr_{\vec{u}}\vec{b}$$
$$pr_{\vec{u}}(\alpha\vec{a}) = \alpha pr_{\vec{u}}\vec{a}$$

Систем координат в просторі можна ввести багато. Найчастіше використовується **прямокутна декартова система координат** (ПДСК).

Упорядковану трійку ненульових векторів $\vec{a},\vec{b},\vec{c}$ називають **правою трійкою**, якщо з вершини вектора $\vec{c}$ поворот від вектора $\vec{a}$ до $\vec{b}$ виконується проти годинникової стрілки, і **лівою трійкою** – якщо за годинниковою. Іншими словами, трійка некомпланарних векторів буде **правою**, якщо поворот від першого вектора до другого й від другого до третього виконується проти годинникової стрілки.

На рис. 3.11 вектори $\vec{a},\vec{b},\vec{c}$ – це права трійка, а вектори $\vec{a},\vec{b},\vec{c}'$ – ліва трійка.

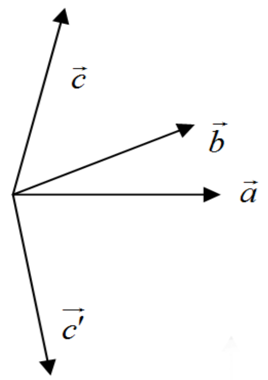

Рис. 3.11 Права та ліва трійки векторів

Прямокутна декартова система координат у просторі – це права трійка попарно ортогональних одиничних векторів, які позначаються $\vec{i}, \vec{j}, \vec{k}$ (рис. 3.12).

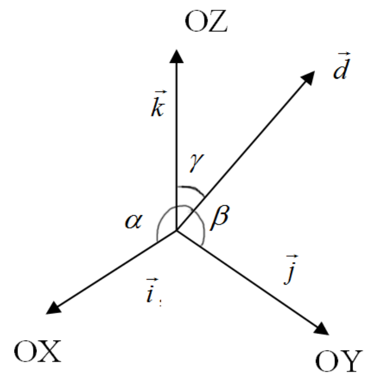

Рис. 3.12 Вектор в прямокутній декартовій системі координат

У представлені вектора $\vec{d}$

$$\vec{d} = x \cdot \vec{i} + y \cdot \vec{j} + z \cdot \vec{k} = (x, y, z)$$

числа x, y, z називаються **декартовими координатами** вектора $\vec{d}$.

З геометричної точки зору декартові координати вектора – це проекції цього вектора на осі координат OX, OY, OZ відповідно.

Позначимо через α, β, γ кути нахилу вектора $\vec{d}$ до осі OX, OY, OZ відповідно. Числа $\cos\alpha, \cos\beta, \cos\gamma$ називаються **напрямними косинусами** вектора $\vec{d}$.

Виходячи з геометричного змісту декартових координат та формули (3.6а), можна записати:

$$x = \left|\vec{d}\right| \cos\alpha \quad y = \left|\vec{d}\right| \cos\beta \quad z = \left|\vec{d}\right| \cos\gamma \tag{3.7}$$

Оскільки квадрат діагоналі прямокутного паралелепіпеда дорівнює сумі квадратів його сторін, то можна записати наступний вираз для визначення довжини вектора через його координати

$$\left|\vec{d}\right| = \sqrt{x^2 + y^2 + z^2} \tag{3.8}$$

З формул (3.7) та (3.8) випливає:

$$\cos\alpha = \frac{x}{\sqrt{x^2 + y^2 + z^2}} \quad \cos\beta = \frac{y}{\sqrt{x^2 + y^2 + z^2}} \quad \cos\gamma = \frac{z}{\sqrt{x^2 + y^2 + z^2}} \tag{3.9}$$

Підносячи ці вирази до квадрата та додаючи, отримаємо:

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1 \tag{3.10}$$

тобто, сума квадратів напрямних косинусів будь-якого вектора дорівнює одиниці. Таким чином, довільний вектор $\vec{d}$ в просторі однозначно визначається заданням довжини та трьох напрямних косинусів.

Орт-вектором $\vec{d}_0$ вектора $\vec{d}$ називають вектор, який має той самий напрям, що й $\vec{d}$, та одиничну довжину.

Очевидно, що

$$\vec{d}_0 = \frac{\vec{d}}{\left|\vec{d}\right|} \tag{3.11}$$

З формул (3.11) та (3.7) випливає, що

$$\vec{d}_0 = (\cos\alpha; \cos\beta; \cos\gamma) \tag{3.12}$$

тобто в ПДСК координатами орт-вектора $\vec{d}_0$ є напрямні косинуси вектора $\vec{d}$.

Наостанок розглянемо дві прості (проте дуже важливі) задачі аналітичної геометрії.

Перша з них пов'язана із **знаходження довжини відрізка між двома точками**

Розглянемо в просторі ПДСК дві точки $M_1(x_1, y_1, z_1)$ та $M_2(x_2, y_2, z_2)$. Очевидно, що відстань $\rho(M_1, M_2)$ між точками M_1 та M_2 дорівнює довжині вектора $\overrightarrow{M_1M_2}$ (рис. 3.13).

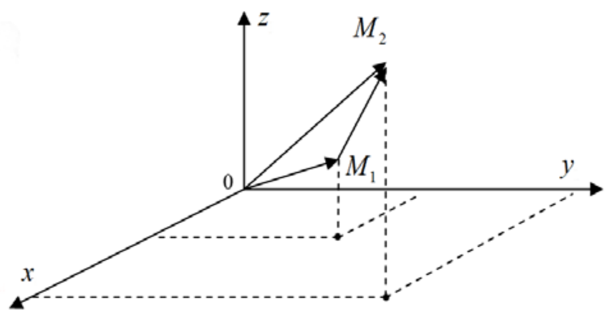

Рис. 3.13 Обчислення відстані між точками в просторі

Маємо:

$$\overrightarrow{M_1M_2} = ((\overrightarrow{M_1M_2})_x, (\overrightarrow{M_1M_2})_y, (\overrightarrow{M_1M_2})_z)$$

$$\overrightarrow{OM_1} + \overrightarrow{M_1M_2} = \overrightarrow{OM_2}$$

$$\overrightarrow{M_1M_2} = (\overrightarrow{OM_2} - \overrightarrow{OM_1})$$

$$\overrightarrow{M_1M_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$$

$$\rho(M_1, M_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

Друга задача пов'язана з **діленням відрізка у заданому відношенні.**

Розглянемо в просторі дві точки A та B, і пряму, яка проходить через ці точки (рис. 3.14):

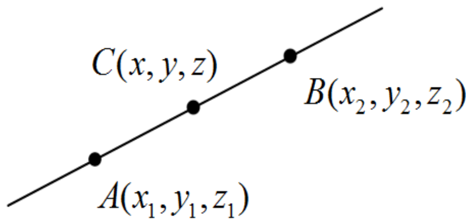

Рис. 3.14 Ділення відрізка у заданому відношенні

Число $\lambda = \dfrac{\overrightarrow{AC}}{\overrightarrow{CB}}$ будемо називати відношенням, в якому точка C ділить відрізок AB, тобто будь-яка точка C, яка відмінна від точки B, ділить цей відрізок в деякому відношенні.

Координати точки C обчислюють наступним чином:

$$\overrightarrow{AC} = (x - x_1, y - y_1, z - z_1)$$

$$\overrightarrow{CB} = (x_2 - x, y_2 - y, z_2 - z)$$

$$\overrightarrow{AC} = \lambda \cdot \overrightarrow{CB}$$

Тобто отримуємо настіпні співвідношення:

$$\begin{array}{l} x - x_1 = \lambda(x_2 - x) \\ y - y_1 = \lambda(y_2 - y) \\ z - z_1 = \lambda(z_2 - z) \end{array} \Rightarrow \begin{cases} x = \dfrac{x_1 + \lambda x_2}{1 + \lambda} \\ y = \dfrac{y_1 + \lambda y_2}{1 + \lambda}, \ \lambda \neq -1 \\ z = \dfrac{z_1 + \lambda z_2}{1 + \lambda} \end{cases} \qquad (3.13)$$

Формули (3.13) називають **формулами ділення відрізка в заданому відношенні λ.**

Зауваження 1

Значення $\lambda = -1$ не підходить з геометричних міркувань, оскільки в цьому випадку має бути $\overrightarrow{AC} = -\overrightarrow{CB}$.

Зауваження 2

У випадку, коли $\lambda = 1$, точка C є серединою відрізка AB. Формули (3.13) тоді називають формулами ділення відрізка навпіл.

Зауваження 3

Коли $\lambda > 0$, точка C знаходиться всередині відрізка AB. Якщо $\lambda < 0$, то точка C заходиться поза межами відрізка AB.

Контрольні запитання:

1. Дайте визначення вектора та його довжини
2. Як додають вектори?
3. Що таке компланарні та колінеарні вектори?
4. Дайте визначення базису та системи координат

Завдання для самостійної роботи

1. Довести всі властивості операції додавання векторів
2. Довести властивості операції множення вектора на число
3. Довести сформульовані твердження щодо базисів

Література: [5], с. 44-63, 17-21; [6], с. 9-31.

3.2. Добутки векторів: скалярний, векторний, подвійний векторний, мішаний

Скалярним добутком двох векторів $\vec{a}$ та $\vec{b}$ називають число, яке дорівнює добутку довжин цих векторів на косинус кута між ними:

$$\left(\vec{a},\vec{b}\right)=|\vec{a}|\left|\vec{b}\right|\cos\varphi \tag{3.14}$$

де φ – кут між $\vec{a}$ та $\vec{b}$.

Добуток позначають $\left(\vec{a},\vec{b}\right)$ або $\vec{a}\cdot\vec{b}$.

Сформулюємо інше визначення скалярного добутку, яке еквівалентне вже визначеному, та ґрунтується на понятті проекції вектора на вектор.

Виходячи з визначення проекції:

$$pr_{\bar{a}}\vec{b}=\left|\vec{b}\right|\cos\varphi \tag{3.15}$$

$$pr_{\bar{b}}\vec{a}=|\vec{a}|\cos\varphi \tag{3.15а}$$

Тоді твердження (3.14) з урахуванням формул (3.15) та (3.15а) матиме наступний вигляд:

$$\left(\vec{a},\vec{b}\right)=|\vec{a}|\,pr_{\bar{a}}\vec{b}=\left|\vec{b}\right|pr_{\bar{b}}\vec{a} \tag{3.16}$$

Тобто під скалярним добутком векторів $\vec{a}$ та $\vec{b}$ розуміють число, яке дорівнює добутку довжини одного з них на проекцію іншого вектора на вісь, визначену першим вектором.

З фізичної точки зору робота сили, яка визначається вектором $\vec{F}$, при переміщенні точок, до якої вона прикладена, з початку в кінець вектора $\vec{S}$, дорівнює скалярному добутку цих векторів:

$$A = \left(\vec{F}, \vec{S}\right)$$

Алгебраїчні властивості скалярного добутку:

1) $\left(\vec{a}, \vec{b}\right) = \left(\vec{b}, \vec{a}\right)$ – комутативність;

2) $\left(\alpha\vec{a}, \vec{b}\right) = \alpha\left(\vec{a}, \vec{b}\right)$ – асоціативність відносно числового множника;

3) $\left(\vec{a} + \vec{b}, \vec{c}\right) = \left(\vec{a}, \vec{c}\right) + \left(\vec{b}, \vec{c}\right)$ – дистрибутивність;

4) $\left(\vec{a}, \vec{a}\right) \geq 0,\ \left(\vec{a}, \vec{a}\right) = 0 \ \Leftrightarrow\ \vec{a} = 0$.

Властивості 2) і 3) характеризують лінійність скалярного добутку.

Доведемо справедливість цих властивостей.

Властивість 1) безпосередньо випливає з визначення. Властивості 2) і 3) легко доводяться, якщо використати визначення скалярного добутку в формі (3.16). Оскільки проекція вектора на вісь має властивості лінійності, то

$$\left(\alpha\vec{a}, \vec{b}\right) = \left|\vec{b}\right| \cdot pr_{\vec{b}}\left(\alpha\vec{a}\right) = \alpha \cdot \left|\vec{b}\right| \cdot pr_{\vec{b}} \cdot \vec{a} = \alpha\left(\vec{a}, \vec{b}\right)$$

та

$$\left(\vec{a} + \vec{b}, \vec{c}\right) = \left|\vec{c}\right| \cdot pr_{\vec{c}}\left(\vec{a} + \vec{b}\right) = \left|\vec{c}\right| \cdot pr_{\vec{c}}\left(pr_{\vec{c}}\vec{a} + pr_{\vec{c}}\vec{b}\right) = \left(\vec{a}, \vec{c}\right) + \left(\vec{b}, \vec{c}\right).$$

У випадку властивості 4) зазначимо спочатку, що з визначення (3.14) буде випливати, що:

$$\left(\vec{a},\vec{a}\right)=\left|\vec{a}\right|^2 \tag{3.17}$$

тобто скалярний добуток вектора самого на себе дорівнює квадрату його довжини, звідки й випливає справедливість останньої властивості.

Доведені властивості дозволяють надалі при скалярному добутку лінійних комбінацій векторів виконувати необхідні дії почленно, не піклуючись при цьому про порядок векторів і сполучення числових множників.

Розглянемо далі **геометричні властивості**, які сформульовані у вигляді теорем.

Теорема (про ортогональність векторів)

Два вектори $\vec{a}$ та $\vec{b}$ ортогональні тоді і тільки тоді, коли їх скалярний добуток дорівнює нулеві

$$\vec{a}\perp\vec{b} \quad \Leftrightarrow \quad \left(\vec{a},\vec{b}\right)=0$$

□

Необхідність

Нехай вектори $\vec{a}$ та $\vec{b}$ ортогональні, φ – кут між ними. Тоді $\cos\varphi=0$ та за визначенням (3.14) $\left(\vec{a},\vec{b}\right)=0$.

Достатність

Нехай скалярний добуток векторів дорівнює нулю $\left(\vec{a},\vec{b}\right)=0$. Якщо хоча б один з векторів нульовий, то, оскільки нульовий вектор має невизначений напрям, його можна вважати ортогональним до будь-якого вектора, і вектори $\vec{a}$ та $\vec{b}$ ортогональні.

Нехай тепер обидва вектори ненульові. Отже, $\left|\vec{a}\right|>0$ та $\left|\vec{b}\right|>0$. Але тоді зі співвідношення $\left(\vec{a},\vec{b}\right)=0=\left|\vec{a}\right|\cdot\left|\vec{b}\right|\cdot\cos\varphi$ випливає, що $\cos\varphi=0$ і вектори $\vec{a}$ та $\vec{b}$ ортогональні.

■

Уточнимо поняття кута φ між векторами $\vec{a}$ та $\vec{b}$, що було сформульовано в попередній темі.

Приведемо вектори $\vec{a}$ та $\vec{b}$ до загального початку – деякої точки *O* (рис. 3.15):

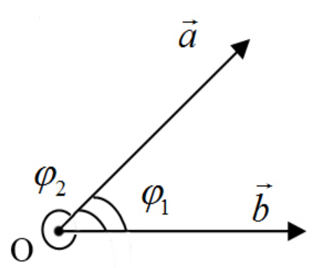

Рис. 3.15 Кут між векторами

Тоді кутом між векторами можна вважати як кут φ_1, так і кут φ_2. Насправді, їх сума дорівнює 2π, а отже $\cos\varphi_1=\cos\varphi_2$ (а у визначення (3.14) входить лише косинус). Тому в подальших міркуваннях під кутом між векторами будемо розуміти той кут, який не перевищує числа π (хоча б один з них обов'язково не перевищує π).

Теорема (про кут між векторами)

Два ненульові вектори $\vec{a}$ та $\vec{b}$ утворюють гострий (тупий) кут тоді і лише тоді, коли їх скалярний добуток додатній (від'ємний).

Розглянемо далі представлення скалярного добутку через координати векторів в ПДСК.

У просторі визначено базис з трьох взаємоперпендикулярних векторів $\vec{i}, \vec{j}, \vec{k}$, які утворюють праву трійку. Розкладання векторів $\vec{a}$ та $\vec{b}$ за цим базисом має вигляд:

$$\vec{a} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k} \qquad b = b_1\vec{i} + b_2\vec{j} + b_3\vec{k}$$

Тоді

$(\vec{a}, \vec{b}) = \left(a_1\vec{i} + a_2\vec{j} + a_3\vec{k}, b_1\vec{i} + b_2\vec{j} + b_3\vec{k}\right)$=[з урахуванням лінійних властивостей]= $a_1b_1\left(\vec{i}, \vec{i}\right) + a_1b_2\left(\vec{i}, \vec{j}\right) + a_1b_3\left(\vec{i}, \vec{k}\right) + a_2b_1\left(\vec{j}, \vec{i}\right) +$

$+ a_2b_2\left(\vec{j}, \vec{j}\right) + a_2b_3\left(\vec{j}, \vec{k}\right) + a_3b_1\left(\vec{k}, \vec{i}\right) + a_3b_2\left(\vec{k}, \vec{j}\right) + a_3b_3\left(\vec{k}, \vec{k}\right) =$

=[ортогональність векторів та формула (3.17)]=$a_1b_1\left|\vec{i}\right|^2 + a_2b_2\left|\vec{j}\right|^2 + a_3b_3\left|\vec{k}\right|^2$,

тобто

$$\left(\vec{a}, \vec{b}\right) = a_1b_1 + a_2b_2 + a_3b_3 \qquad (3.18)$$

Також з формул (3.14), (3.17) та (3.18) випливає, що

$$\cos\varphi = \frac{\left(\vec{a}, \vec{b}\right)}{|\vec{a}|\,|\vec{b}|} = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}\,\sqrt{b_1^2 + b_2^2 + b_3^2}} \qquad (3.19)$$

Векторним добутком вектора $\vec{a}$ на вектор $\vec{b}$ називають вектор $\vec{c}$ (позначають $\vec{c} = \left[\vec{a}, \vec{b}\right]$ або $\vec{c} = \vec{a} \times \vec{b}$), який задовольняє наступні три умови:

1. Довжина вектора $\vec{c}$ дорівнює добутку довжин векторів $\vec{a}$ та $\vec{b}$ на синус кута між ними

$$|\vec{c}| = \left|\left[\vec{a}, \vec{b}\right]\right| = |\vec{a}| \cdot |\vec{b}| \cdot \sin\varphi \qquad (3.20)$$

2. Вектор $\vec{c}$ ортогональний до кожного з векторів $\vec{a}$ та $\vec{b}$

3. Вектори $\vec{a}$, $\vec{b}$, $\vec{c}$ утворюють праву трійку (рис. 3.16).

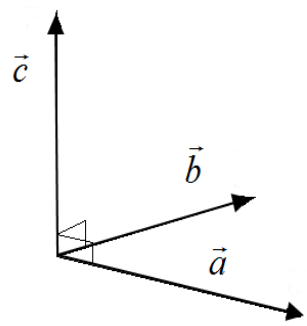

Рис. 3.16 Векторний добуток векторів

З точки зору фізики, якщо вектор $\vec{b}$ відповідає прикладеній в деякій точці *M* силі, а вектор $\vec{a}$ з'єднує точки *O* та *M*, то вектор $\vec{c} = \left[\vec{a}, \vec{b}\right]$ є моментом сили $\vec{b}$ відносно точки *O*.

Алгебраїчні властивості векторного добутку:

1) $\left[\vec{a}, \vec{b}\right] = -\left[\vec{b}, \vec{a}\right]$ – антикомутативність;

2) $\left[\alpha\vec{a}, \vec{b}\right] = \alpha\left[\vec{a}, \vec{b}\right]$ – асоціативність відносно числового множника;

3) $\left[\vec{a} + \vec{b}, \vec{c}\right] = \left[\vec{a}, \vec{c}\right] + \left[\vec{b}, \vec{c}\right]$ – дистрибутивність відносно суми;

4) $\left[\vec{a}, \vec{a}\right] = 0$ для будь-якого вектора $\vec{a}$.

Геометричні властивості векторного добутку полягають в наступному:

1) вектори $\vec{a}$ та $\vec{b}$ колінеарні тоді і лише тоді, коли їх векторний добуток дорівнює нулю

$$\vec{a} \parallel \vec{b} \Leftrightarrow \left[\vec{a}, \vec{b}\right] = 0$$

2) модуль векторного добутку вектора $\vec{a}$ на вектор $\vec{b}$ дорівнює площі паралелограма, побудованого на приведених до загального початку векторах.

Покажемо справедливість цих властивостей.

Друга геометрична властивість безпосередньо випливає з визначення (3.20) і формули площі паралелограма.

Розглянемо першу геометричну властивість. Необхідність випливає з формули (3.20), оскільки у випадку колінеарних векторів $\varphi = 0$ або $\varphi = \pi$. Тоді $\sin\varphi = 0$, і, відповідно, $\left[\vec{a}, \vec{b}\right] = 0$.

Достатність можна показати наступним чином.

Нехай векторний добуток нульовий $\left[\vec{a}, \vec{b}\right] = 0$. Якщо хоча б один з векторів $\vec{a}$ та $\vec{b}$ нульовий, то вони, очевидно, колінеарні (оскільки нульовий вектор є колінеарним до будь-якого вектора). Якщо ж обидва вектори $\vec{a}$ та $\vec{b}$ не нульові, то $\left|\vec{a}\right| > 0$ і $\left|\vec{b}\right| > 0$, тоді отримаємо, що $\sin\varphi = 0$, тобто $\varphi = 0$ або $\varphi = \pi$, і вектори $\vec{a}$ та $\vec{b}$ колінеарні, що і потрібно було довести.

Тепер повернемось до доведення алгебраїчних властивостей.

Для доведення першого з них покладемо $\vec{c} = \left[\vec{a}, \vec{b}\right]$, $\vec{d} = \left[\vec{b}, \vec{a}\right]$. Якщо вектори $\vec{a}$ та $\vec{b}$ колінеарні, то $\vec{c} = \vec{d} = 0$. Якщо ж колінеарність відсутня, то, по-перше, вектори $\vec{c}$ і $\vec{d}$ однакової довжини (на основі (3.20)), а по-друге, колінеарні (оскільки вони обидва ортогональні до площини векторів $\vec{a}$ та $\vec{b}$). Але тоді $\vec{c} = \vec{d}$ або $\vec{c} = -\vec{d}$. Перший варіант неможливий, оскільки тоді трійки $\vec{a}, \vec{b}, \vec{c}$ і $\vec{b}, \vec{a}, \vec{c}$ обидві були б правими, таким чином $\vec{c} = -\vec{d}$, тобто $\left[\vec{a}, \vec{b}\right] = -\left[\vec{b}, \vec{a}\right]$.

Четверта властивість автоматично випливає з вже доведеної першої геометричної. Друга та третя властивість (лінійні властивості векторного добутку) приймемо без доведення.

Лінійні властивості сформульовані відносно першого аргумента векторного добутку. Неважко показати (самостійно), що вони справедливі також і відносно другого аргумента.

Як і у випадку скалярного добутку, лінійні властивості дозволяють при векторному множенні лінійних комбінацій векторів виконувати дії почленно і робити сполучення числових множників. Проте, на відміну від скалярного добутку, при цьому необхідно або зберігати порядок векторних множників, або в разі його зміни міняти знак на протилежний.

Розглянемо векторний добуток векторів в просторі із заданою ПДСК.

Нехай вектори задані своїми координатами $\vec{a} = \left(a_1, a_2, a_3\right)$ та $\vec{b} = \left(b_1, b_2, b_3\right)$. Векторний добуток цих векторів має наступний вигляд:

$$\left[\vec{a}, \vec{b}\right] = \left[a_1\vec{i} + a_2\vec{j} + a_3\vec{k},\ b_1\vec{i} + b_2\vec{j} + b_3\vec{k}\right] = a_1b_1\left[\vec{i}, \vec{i}\right] + a_1b_2\left[\vec{i}, \vec{j}\right] + a_1b_3\left[\vec{i}, \vec{k}\right] +$$
$$+ a_2b_1\left[\vec{j}, \vec{i}\right] + a_2b_2\left[\vec{j}, \vec{j}\right] + a_2b_3\left[\vec{j}, \vec{k}\right] + a_3b_1\left[\vec{k}, \vec{i}\right] + a_3b_2\left[\vec{k}, \vec{j}\right] + a_3b_3\left[\vec{k}, \vec{k}\right]$$

враховуючи, що

$$\left[\vec{i}, \vec{i}\right] = \left[\vec{j}, \vec{j}\right] = \left[\vec{k}, \vec{k}\right] = 0$$
$$\left[\vec{i}, \vec{j}\right] = \vec{k} \qquad \left[\vec{i}, \vec{k}\right] = -\vec{j} \qquad \left[\vec{j}, \vec{k}\right] = \vec{i}$$

отримаємо

$$\left[\vec{a}, \vec{b}\right] = a_1b_2\vec{k} - a_1b_3\vec{j} - a_2b_1\vec{k} + a_2b_3\vec{i} + a_3b_1\vec{j} - a_3b_2\vec{i}$$

чи, перегрупувавши елементи

$$\left[\vec{a},\vec{b}\right]=\left(a_2b_3-a_3b_2\right)\vec{i}-\left(a_1b_3-a_3b_1\right)\vec{j}+\left(a_1b_2-a_2b_1\right)\vec{k} \qquad (3.21)$$

Беручи до уваги теорію визначників (обчислення визначників другого порядку і формулу розкладання за рядком), вираз (3.21) можна переписати у наступному вигляді:

$$\left[\vec{a},\vec{b}\right]=\vec{i}\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}-\vec{j}\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}+\vec{k}\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

або

$$\left[\vec{a},\vec{b}\right]=\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \qquad (3.22)$$

З формули (3.21) випливає, що векторний добуток дорівнює нулю, якщо одночасно виконується

$$a_2b_3=a_3b_2 \qquad a_1b_3=a_3b_1 \qquad a_1b_2=a_2b_1$$

Останні вирази еквівалентні пропорціям

$$\frac{a_1}{b_1}=\frac{a_2}{b_2}=\frac{a_3}{b_3} \qquad (3.23)$$

Оскільки в знаменниках (3.23) можуть бути нулі, то будь-яку пропорцію $\frac{a}{b}=\frac{c}{d}$ слід розуміти як рівність $ad=bc$.

Таким чином, формулами (3.23) підтверджено вже доведений вище факт, що якщо вектори $\vec{a}=(a_1,a_2,a_3)$ та $\vec{b}=(b_1,b_2,b_3)$ колінеарні, то їх координати пропорційні.

Нехай тепер задано три довільні вектори $\vec{a}$, $\vec{b}$ та $\vec{c}$. Якщо вектор $\vec{b}$ векторно множать на вектор $\vec{c}$, а вектор $\vec{a}$ так само векторно множать на векторний добуток $\left[\vec{b},\vec{c}\right]$, то отримуваний вектор

$$\vec{d}=\left[\vec{a},\left[\vec{b},\vec{c}\right]\right] \tag{3.24}$$

називають **подвійним векторним добутком**. Слід зауважити, що вектор $\vec{d}$, який визначається формулою (3.24), лежить в площині векторів $\vec{b}$ та $\vec{c}$.

Для обчислення подвійного векторного добутку можна також скористатися формулою

$$\left[\vec{a},\left[\vec{b},\vec{c}\right]\right]=\vec{b}\cdot(\vec{a},\vec{c})-\vec{c}\cdot\left(\vec{a},\vec{b}\right)$$

Для запам’ятовування цієї формули можна скористатися мнемонічним правилом: права частина читається як “бац-мінус-цаб”.

Мішаним добутком векторів $\vec{a}$, $\vec{b}$ та $\vec{c}$ називають число, яке дорівнює скалярному добутку вектора, що в свою чергу дорівнює векторному добутку векторів $\vec{a}$ та $\vec{b}$ на вектор $\vec{c}$:

$$\vec{a}\cdot\vec{b}\cdot\vec{c}=\left(\left[\vec{a},\vec{b}\right],c\right)$$

Геометричний зміст мішаного добутку виражає наступна теорема.

<u>Теорема (про геометричний зміст мішаного добутку)</u>

Мішаний добуток векторів $\vec{a}$, $\vec{b}$ та $\vec{c}$, дорівнює об'єму паралелепіпеда, побудованого на приведених до загального початку векторах $\vec{a}$, $\vec{b}$ та $\vec{c}$, взятому із знаком «+», якщо вони утворюють праву трійку, та зі знаком «-», якщо ліву. У випадку компланарності векторів їх мішаний добуток дорівнює нулю.

□

Якщо вектори $\vec{a}$ та $\vec{b}$ колінеарні, то їх векторний добуток нульовий, і, відповідно, мішаний добуток також дорівнює нулеві.

Якщо вектори $\vec{a}$, $\vec{b}$ та $\vec{c}$ компланарні, то, оскільки вектор $\vec{f}=\left[\vec{a},\vec{b}\right]$ перпендикулярний площині векторів $\vec{a}$ та $\vec{b}$, він також буде перпендикулярний до вектора $\vec{c}$, а отже, $\left(\vec{f},\vec{c}\right)=0$, тобто мішаний добуток також нульовий.

Нехай тепер вектори $\vec{a}$, $\vec{b}$ та $\vec{c}$ не компланарні. Приведемо їх до спільного початку та побудуємо на них паралелепіпед (рис. 3.18).

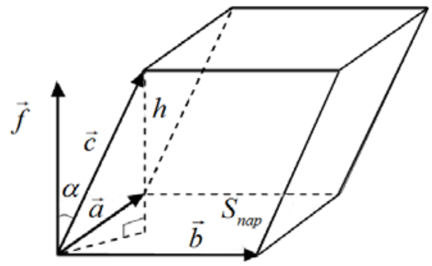

Рис. 3.18 Геометричний зміст мішаного добутку

Нехай спочатку трійка векторів $\vec{a}$, $\vec{b}$, $\vec{c}$ права.

Об'єм побудованого паралелепіпеда $V_{пар}$ дорівнює добутку висоти h на площу паралелограма, побудованого на векторах $\vec{a}$ та $\vec{b}$:

$$V_{пар} = S_{пар} \cdot |h|$$

З властивості векторного добутку маємо

$$S_{пар} = \left| \left[\vec{a}, \vec{b} \right] \right|$$

Довжина висоти h дорівнює проекції вектора $\vec{c}$ на вектор $\vec{f} = \left[\vec{a}, \vec{b} \right]$,тобто

$$|h| = pr_{\vec{f}}\, \vec{c} = |\vec{c}| \cdot \cos\alpha, \;\; \alpha = \left(\vec{c} \wedge \vec{f} \right) \in \left(0; \pi/2 \right)$$

Оскільки трійка векторів $\vec{a}$, $\vec{b}$, $\vec{c}$ права, то вектор $\vec{f}$ напрямлений в один бік з вектором $\vec{c}$, до того ж проекція береться із знаком «мінус».

Таким чином, маємо

$$V_{пар} = \left|\left[\vec{a}, \vec{b}\right]\right| \cdot |\vec{c}| \cdot \cos\alpha = \left(\left[\vec{a}, \vec{b}\right], \vec{c}\right).$$

У випадку, якщо вектори утворюють ліву трійку, то кут α тупий, отже, його можна записати у вигляді $\alpha = \pi - \beta$, де β – гострий кут між $\vec{c}$ і прямою, що містить $\vec{f}$

$$V_{пар} = \left|\left[\vec{a}, \vec{b}\right]\right| \cdot |\vec{c}| \cdot \cos\alpha = -\left|\left[\vec{a}, \vec{b}\right]\right| \cdot |\vec{c}| \cdot \cos\beta = -\left(\left[\vec{a}, \vec{b}\right], \vec{c}\right)$$

■

<u>Наслідок 1</u>

Справедлива рівність $\left(\left[\vec{a}, \vec{b}\right], \vec{c}\right) = \left(\vec{a}, \left[\vec{b}, \vec{c}\right]\right)$.

Останнє співвідношення дозволяє записувати мішаний добуток не вказуючи які саме два вектори (перші або останні) перемножуються векторно.

<u>Наслідок 2</u>

Необхідною і достатньою умовою компланарності трьох векторів є рівність нулю їх мішаного добутку.

<u>Наслідок 3</u>

Якщо серед трьох векторів два співпадають, то їх мішаний добуток дорівнює нулю.

Алгебраїчні властивості:

1) $\vec{a}\;\vec{b}\;\vec{c}=\vec{c}\;\vec{a}\;\vec{b}=\vec{b}\;\vec{c}\;\vec{a}=-\vec{b}\;\vec{a}\;\vec{c}=-\vec{c}\;\vec{b}\;\vec{a}=-\vec{a}\;\vec{c}\;\vec{b}$;

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

3) $\left(\vec{a}+\vec{b}\right)\vec{c}\;\vec{d}=\vec{a}\,\vec{c}\;\vec{d}+\vec{b}\;\vec{c}\;\vec{d}$.

В ПДСК мішаний добуток має наступний вигляд. Згідно (3.21)

$$\left[\vec{a},\vec{b}\right]=\left(a_2b_3-a_3b_2,\,-a_1b_3+a_3b_1,\,a_1b_2-a_2b_1\right)$$

Тоді

$$\left(\left[\vec{a},\vec{b}\right],\,\vec{c}\right)=\left(a_2b_3-a_3b_2\right)c_1+\left(a_3b_1-a_1b_3\right)c_2+\left(a_1b_2-a_2b_1\right)c_3=$$
$$=a_2b_3c_1-a_3b_2c_1+a_3b_1c_2-a_1b_3c_2+a_1b_2c_3-a_2b_1c_3=$$
$$=\left(a_1b_2c_3+a_2b_3c_1+a_3b_1c_2\right)-\left(a_3b_2c_1+a_1b_3c_2+a_2b_1c_3\right)=$$
$$=[\text{правило зірочки}]=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Тобто

$$\vec{a}\cdot\vec{b}\cdot\vec{c}=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \qquad (3.25)$$

З (3.25) випливає, що критерієм компланарності векторів $\vec{a}$, $\vec{b}$, $\vec{c}$, які задані своїми координатами в ПДСК, є рівність нулеві визначника, рядками якого є координати цих векторів:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0 .$$

Таким чином, оскільки базис в просторі E^3 утворює будь-яка трійка некомпланарних векторів, то останнє співвідношення можна розглядати як метод визначення того, чи будуть вектори утворювати базис.

Контрольні запитання:

1. Як визначають скалярний добуток і якими є його властивості?
2. Як визначають векторний добуток і якими є його властивості?
3. Як визначається мішаний добуток і якими є його властивості?

Завдання для самостійної роботи

1. Довести теорему про кут між векторами
2. Показати справедливість лінійності векторного добутку відносно другого аргумента

Література: [5], с. 63-80; [6], с. 31-36.

4. АНАЛІТИЧНА ГЕОМЕТРІЯ НА ПЛОЩИНІ

4.1. Лінія на площині. Способи задання

Нехай задано деяку площину, на якій визначено ПДСК. Нехай у цій площині задано деяку лінію L.

Рівняння виду

$$F(x, y) = 0 \tag{4.1}$$

визначає **пласку лінію** L, якщо це рівняння задовольняють координати x та y будь-якої точки кривої L, та не задовольняють координати x та y жодної точки, що не належить лінії L.

Сама лінія L являється геометричним місцем точок, координати яких задовольняють рівняння (4.1) (звичайно ж, у заданій системі координат).

Якщо рівняння виду (4.1) у заданій системі координат є рівнянням лінії L, то це рівняння визначає лінію L.

Можливий варіант, коли рівняння (4.1) або визначає образ, відмінний від того, що звичайно розуміють під терміном "лінія", (наприклад, $x^2 + y^2 = 0$), або взагалі не визначає жодного геометричного образу (наприклад, $x^2 + y^2 + 13 = 0$).

Пласкі лінії поділяють на дві групи: алгебраїчні і трансцендентні.

Пласку лінію називають **алгебраїчною**, якщо функція $F(x, y)$ в лівій частині рівняння (4.1) є алгебричним поліномом, тобто її можна представити у вигляді:

$$\Phi(x,y)=\sum_{k,l} a_{kl}\, x^k y^l \tag{4.2}$$

Будь-яку не алгебраїчну лінію називають **трансцендентною**.

Алгебраїчну лінію будемо називатимемо лінією n*-го* порядку, якщо в деякій ПДСК функція $F(x,y)$ є алгебраїчним поліномом n*-го* степеня, тобто $\max\limits_{k,l}(k+l)=n$

Іншими словами, алгебраїчна лінія n*-го* порядку – це лінія, яка визначається в деякій ПДСК алгебраїчним рівнянням n*-го* степеня з двома невідомими.

Приклад:

$5x^5 - tg\, x + 3y = 0$ – трансцендентна лінія;

$\sin(x^2 + 2x - y^3) = 0$ – трансцендентна лінія;

$3x^5 + 13x^3y^4 - 2y^3x + 2 = 0$ – алгебраїчна лінія 7-го порядку.

Розглянемо далі способи задання лінії на площині. Перший з них вже записаний на основі рівняння (4.1). У цьому випадку говорять, що лінію задано за допомогою неявної функції, або неявно.

Другий спосіб – явне задання лінії. Можливо лише в тому випадку, коли рівняння (4.1) однозначно розв'язується відносно однієї змінної x або y. У цьому випадку рівняння (4.1) перетворюється в

$$y = f_1(x) \tag{4.3a}$$

або

$$x = f_2(y) \qquad (4.3б)$$

Для аналітичного представлення лінії L часто буває зручно виражати змінні координати x та y точок цієї лінії за допомогою третьої допоміжної змінної (параметра) t:

$$\begin{cases} x = \varphi(t) \\ y = \psi(t) \end{cases} \qquad t \in T \qquad (4.4)$$

де функції $\varphi(t)$ та $\psi(t)$ неперервні.

Це третій спосіб задання лінії – параметричний.

Параметричне представлення лінії на площині природно виникає, якщо розглядати лінію як шлях, пройдений матеріальною точкою, яка неперервно рухається за певним законом.

<u>*Приклади*</u>:

1) Розглянемо **коло** $x^2 + y^2 = R^2$ (рис. 4.1)

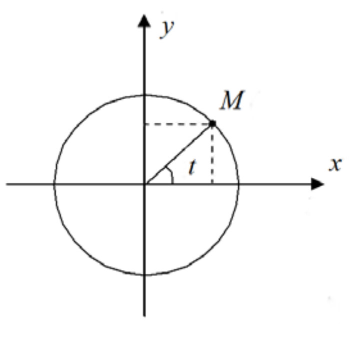

Рис. 4.1 Коло радіусу R

Нехай $M(x,y)$ – його довільна точка, а t – кут між її радіус-вектором та віссю OX, який відлічують проти годинникової стрілки. Очевидно, що

$$\begin{cases} x = R\cos t \\ y = R\sin t \end{cases} \qquad t \in [0;\, 2\pi)$$

2) **Астроїда** (рис. 4.2)

$$\begin{cases} x = a\cos^3 t \\ y = a\sin^3 t \end{cases} \quad \Leftrightarrow \quad x^{2/3} + y^{2/3} = 1 \quad t \in [0;\, 2\pi)$$

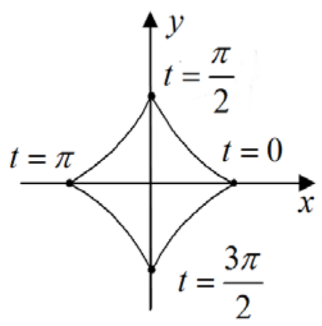

Рис. 4.2 Астроїда

3) **Лист Декарта** (рис. 4.3)

$$\begin{cases} x = \dfrac{3at}{1+t^3} \\ y = \dfrac{3at^2}{1+t^3} \end{cases} \quad \Leftrightarrow \quad x^3 + y^3 - 3axy = 0 \quad t \in \mathbb{R} \setminus \{-1\}$$

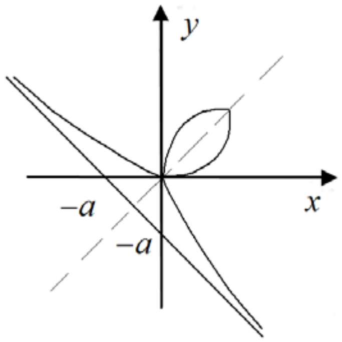

Рис. 4.3 Лист Декарта

4) **Циклоїда**

Це, фактично, шлях, який описується однією з точок кола M, що котиться без ковзання по нерухомій прямій (рис. 4.4).

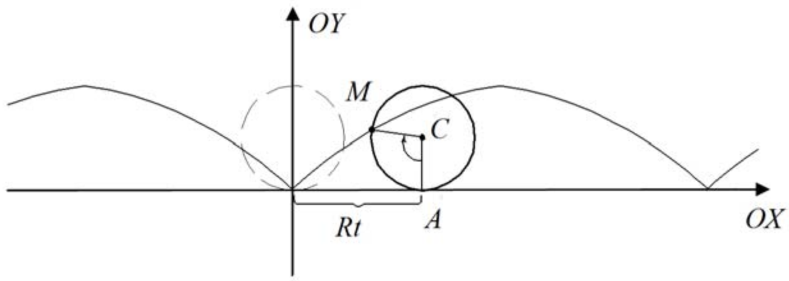

Рис. 4.4 Циклоїда

Введемо ПДСК наступним чином. Вісь OX – пряма, по якій котиться коло, точка O – одна з точок, в яких точка M виходить на вказану пряму, вісь OY спрямована так, щоб її додатня піввісь лежала по той бік OX, що і коло, яке котиться.

Фіксуємо довільне положення кола. Позначимо літерою C його центр, а літерою A – точку дотику з віссю OX. Нехай параметр t – кут, на який повернулося коло при переміщенні з положення з точкою дотику на початку координат в те положення, що розглядається. Оскільки ковзання немає, то $OA = Rt$, де R – радіус кола.

На основі визначення прямокутних декартових координат x, y та лінійних властивостей проекції отримуємо:

$$x = pr_{OX}\overrightarrow{OM} = pr_{OX}\overrightarrow{OA} + pr_{OX}\overrightarrow{AC} + pr_{OX}\overrightarrow{CM}$$
$$y = pr_{OY}\overrightarrow{OM} = pr_{OY}\overrightarrow{OA} + pr_{OY}\overrightarrow{AC} + pr_{OY}\overrightarrow{CM}$$

Враховуючи, що $\angle ACM = t + 2\pi k,\ \ k \in \mathbb{Z}$, маємо

$$pr_{OX}\overrightarrow{OA} = Rt \qquad pr_{OX}\overrightarrow{AC} = 0 \qquad pr_{OX}\overrightarrow{CM} = -R\sin t$$
$$pr_{OY}\overrightarrow{OA} = 0 \qquad pr_{OY}\overrightarrow{AC} = R \qquad pr_{OY}\overrightarrow{CM} = -R\cos t.$$

Підставивши ці результати в попередній запис, отримаємо параметричне рівняння циклоїди:

$$\begin{cases} x = R(t - \sin t) \\ y = R(1 - \cos t) \end{cases} \qquad t \in \mathbb{R}$$

Окрім ПДСК, на площині можуть також бути введені й інші системи координат. Найчастіше використовують **полярну систему координат** (ПСК). Для її задання необхідно на площині обрати деяку точку O, яку називають полюсом, і провести з неї промінь.

Тоді будь-яка точка M площини визначатиметься полярними координатами (рис. 4.5):

- **Полярний радіус** $\rho \in (0;\ +\infty)$ – довжина радіус-вектора $\overrightarrow{OM}$.
- **Полярний кут** $\varphi \in [0;\ 2\pi)$ – кут повороту радіус-вектора $\overrightarrow{OM}$ від заданого кута проти годинникової стрілки.

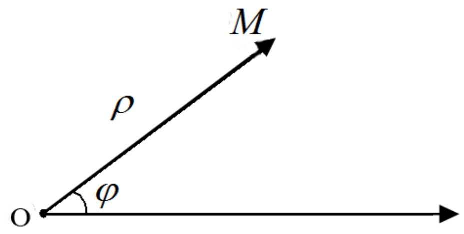

Рис. 4.5 Полярна система координат

Використання недекартових систем координат, зокрема, полярної, викликане тим, що при цьому рівняння лінії мають простіший вигляд.

При заданні ПДСК і ПСК на одній площині зазвичай початки цих систем координат збігаються, а нульовий промінь ПСК збігається з додатним напрямком осі ПДСК. Координати точки в цих системах зв'язані наступними очевидними співвідношеннями:

$$\begin{cases} x = \rho\cos\varphi \\ y = \rho\sin\varphi \end{cases} \Leftrightarrow \begin{cases} \rho = \sqrt{x^2 + y^2} \\ \varphi = arctg\, {}^{y}\!/_{x} \end{cases}$$

Відмітимо, що при визначенні значення кута φ дуже важливо враховувати знак координат x та y.

Таким чином, останній спосіб задання лінії на площині – задання її в полярній системі координат:

$$\rho = \rho(\varphi) \tag{4.5}$$

Для побудови лінії в ПСК будемо використовувати наступний метод. З інтервалу зміни полярного кута φ візьмемо точки, у яких досить легко обчислити значення функції $\rho(\varphi)$, а потім з'єднаємо ці точки плавною лінією.

Приклади:

1) $\rho = R$ (рис. 4.6)

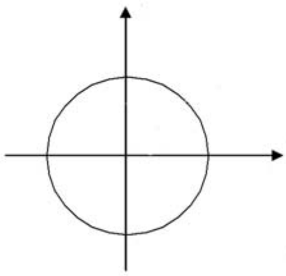

Рис. 4.6 Коло в полярній системі координат

2) $\rho = (1 - \cos\varphi)$ (рис. 4.7)

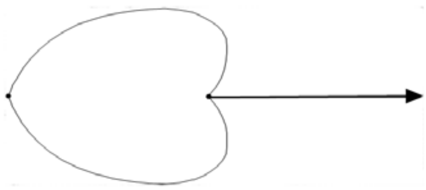

Рис. 4.7 Кардіоїда

3) **Лемніската Бернуллі** (рис. 4.8)

$$\rho = a\sqrt{\cos 2\varphi}$$

$$\cos 2\varphi \geq 0 \;\Rightarrow\; \varphi \in \left[-\pi/4;\, \pi/4\right],\; \varphi \in \left[3\pi/4;\, 5\pi/4\right]$$

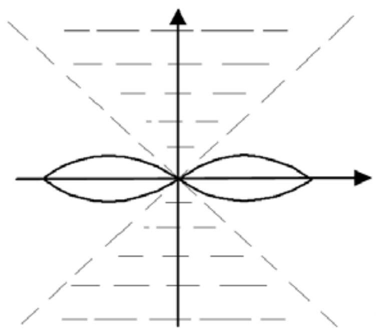

Рис. 4.8 Лемніската Бернулі

Контрольні запитання:

1. Що таке пласка лінія?
2. Чим відрізняється алгебраїчна лінія від трансцендентної?
3. Які існують способи задання лінії на площині?

Завдання для самостійної роботи

1. Записати рівняння лемніскати Бернуллі в ПДСК.
2. Побудувати наступні лінії, які задано в ПСК:

$$\rho = a\sin 3\varphi$$

$$\rho = a\cos\varphi + b\sin\varphi$$

$$\rho = \frac{2}{\cos\varphi - \sin\varphi}$$

Література: [5], с. 101-109, 23-24; [6], с. 37-39.

4.2. Пряма на площині. Рівняння прямої

Пряма лінія на площині із заданою ПДСК – це алгебраїчна лінія першого порядку.

Задамо на площині ПДСК. Нехай пряма L проходить через точку $M_0(x_0, y_0)$ (рис. 4.9):

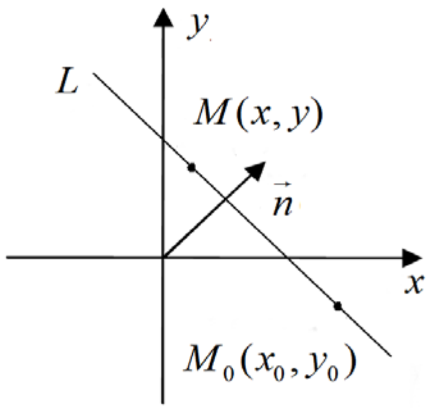

Рис. 4.9 Побудова рівняння прямої на площині

Нехай ненульовий вектор $\vec{n} = (A, B)$ перпендикулярний до проведеної прямої. Такий вектор називають **нормаллю** (або нормальним вектором) прямої L. Довільна точка площини $M(x, y)$ належатиме прямій L в тому і лише в тому випадку, коли вектори $\vec{n}$ та $\overrightarrow{M_0M}$ перпендикулярні:

$$\vec{n} \perp \overrightarrow{M_0M} \quad \Leftrightarrow \quad \left(\vec{n}\,, \overrightarrow{M_0M}\right) = 0$$

Записавши скалярний добуток в координатах, отримаємо

$$A\left(x - x_0\right) + B\left(y - y_0\right) = 0 \qquad (4.6)$$

рівняння прямої, що проходить через задану точку перпендикулярно заданому вектору (або скорочено – **рівняння прямої через нормаль і точку**).

Розкривши в останньому виразі дужки, звівши подібні члени й позначивши $C \equiv -Ax_0 - By_0$, можна отримати **загальне рівняння прямої**

$$Ax + By + C = 0 \tag{4.7}$$

яке дійсно є алгебраїчним рівнянням першого порядку.

Загальне рівняння прямої (4.7) називають **повним**, якщо всі його коефіцієнти відмінні від нуля, і **неповним**, якщо хоча б один з них дорівнює нулю.

Розглянемо, які прямі визначаються за допомогою неповних рівнянь:

а) $C = 0$

рівняння $Ax + By = 0$ визначає пряму, яка проходить через початок координат;

б) $B = 0$

рівняння $Ax + C = 0$ визначає пряму, паралельну осі OY;

в) $A = 0$

рівняння $By + C = 0$ визначає пряму, паралельну осі OX;

г) $B = 0$ та $C = 0$

рівняння $Ax = 0$ визначає вісь OY;

д) $A = 0$ та $C = 0$

рівняння $By = 0$ визначає вісь OX.

Називатимемо **напрямним вектором** прямої L будь-який ненульовий вектор, який паралельний цій прямій.

Розглянемо в ПДСК пряму L, яка проходить через точку $M_0(x_0, y_0)$, та деякий вектор $\vec{s} = (p;q) \neq 0$, паралельний цій прямій (рис. 4.10)

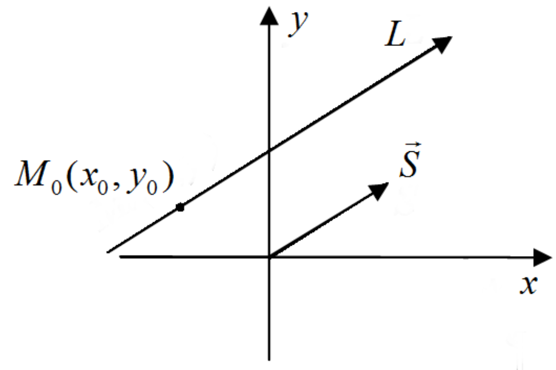

Рис. 4.10 Побудова канонічного рівняння прямої

Зрозуміло, що довільна точка площини $M(x, y)$ належатиме заданій прямій L тоді і лише тоді, коли вектори $\vec{s}$ та $\overrightarrow{M_0M}$ колінеарні, тобто якщо

$$\vec{s} \quad || \quad \overrightarrow{M_0M}$$

або, в координатах

$$\frac{x - x_0}{p} = \frac{y - y_0}{q} \tag{4.8}$$

оскільки для колінеарності векторів необхідно і достатньо пропорційності їх координат.

Формула (4.8) визначає **канонічне рівняння прямої** (або **рівняння прямої через точку та напрямний вектор**).

У канонічному рівнянні (4.8) один із знаменників (p або q) може бути рівним нулеві (обидва не можуть, оскільки вектор $\vec{s}$ ненульовий). Але, через означення пропорції, буде нульовим і відповідний чисельник.

Окремим випадком рівняння (4.8) є **рівняння прямої, що проходить через дві точки** $M_1(x_1, y_1)$ та $M_2(x_2, y_2)$ (рис. 4.11):

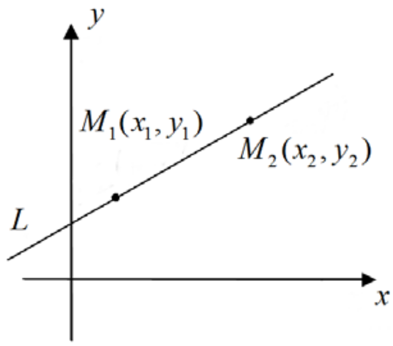

Рис. 4.11 Побудова рівняння прямої через дві точки

Очевидно, в цьому випадку як напрямний вектор $\vec{s}$ можна взяти вектор $\overrightarrow{M_1M_2} = \left(x_2 - x_1;\, y_2 - y_1\right)$. Тепер рівняння (4.8) можна записати у вигляді

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} \tag{4.9}$$

Канонічне рівняння (4.8) – це рівність двох відношень. Можна вважати, що кожне з цих відношень дорівнює деякому параметру t. Тоді (4.8) можна переписати у вигляді

$$\begin{cases} x = x_0 + pt \\ y = y_0 + qt \end{cases} \tag{4.10}$$

де параметр t набуває всіх дійсних значень. Рівняння (4.10) визначають **параметричне рівняння прямої L**.

Розглянемо повне рівняння прямої $Ax + By + C = 0$.

Оскільки всі коефіцієнти відмінні від нуля, то його можна перетворити таким чином:

$$Ax + By = -C$$

$$\left(-\frac{A}{C}\right)x + \left(-\frac{B}{C}\right)y = 1$$

$$-\frac{x}{C/A} - \frac{y}{C/B} = 1$$

Увівши позначення $a \equiv -{C}/{A}$ та $b \equiv -{C}/{B}$, отримаємо **рівняння прямої L у відрізках**:

$$\frac{x}{a} + \frac{y}{b} = 1 \qquad (4.11)$$

Слід зазначити, що в цьому рівнянні числа a та b мають простий геометричний зміст – вони дорівнюють величині відрізків, що відсікаються прямою L на координатних осях (рис. 4.12).

При цьому знак чисел a та b вказує на те, на якому напрямі осі відсікається відповідний відрізок.

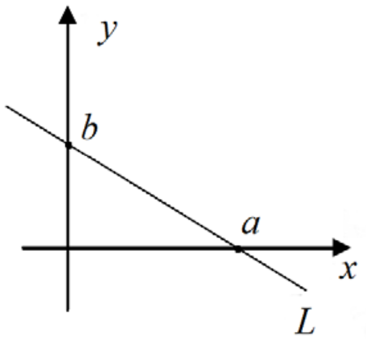

Рис. 4.12 Побудова рівняння прямої у відрізках

Розглянемо в заданій ПДСК довільну пряму L, яка не паралельна осі OY. Позначимо точку перетину прямої та вісі OX літерою A. Візьмемо на вісі OX точку M, що лежить правіше від точки A, а на прямій візьмемо точку N, яка лежить по той бік точки A, куди напрямлена вісь OY (рис. 4.13).

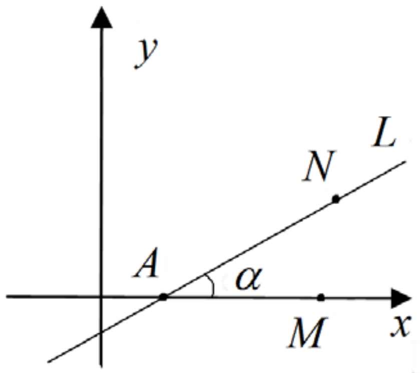

Рис. 4.13 Побудова рівняння прямої через кутовий коефіцієнт

Тоді кут $\alpha = \angle NAM$ будемо називати **кутом нахилу** даної прямої L до осі OX. У випадку, коли $L \| OX$ або співпадає з нею, кут α вважатимемо нульовим. Тангенс кута нахилу прямої до осі OX називатимемо **кутовим коефіцієнтом прямої** $k = tg\,\alpha$.

Для прямої, паралельній осі OX, $k = 0$, для прямої, паралельній осі OY, кутовий коефіцієнт k не існує (при цьому формально кажуть, що він перетворюється в нескінченність).

Нехай задано пряму L, яка проходить через точку $M_0(x_0, y_0)$, а також відомий її напрямний вектор $\vec{S} = (p, q) \neq 0$ (рис. 4.14).

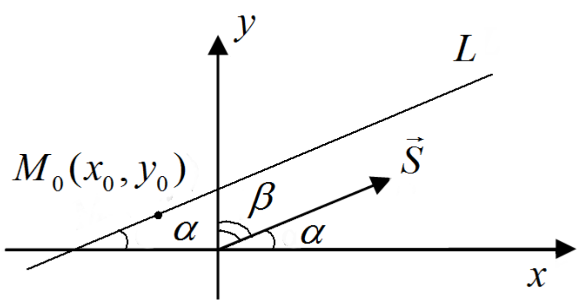

Рис. 4.14 Рівняння прямої через кутовий коефіцієнт

Можна знайти орт-вектор вектора $\vec{S}$, який також буде напрямним вектором для прямої L:

$$\overrightarrow{S_0} = (\cos\alpha; \cos\beta) = \left[\beta = \pi/2 - \alpha\right] = (\cos\alpha; \sin\alpha)$$

Тоді канонічне рівняння (4.8) можна переписати у вигляді:

$$\frac{x - x_0}{\cos\alpha} = \frac{y - y_0}{\sin\alpha}$$

За допомогою послідовних перетворень отримаємо:

$$y - y_0 = \frac{\sin\alpha}{\cos\alpha}(x - x_0) = tg\,\alpha(x - x_0),$$

або

$$y = y_0 + tg\alpha \cdot x - tg\alpha \cdot x_0 .$$

Позначивши $y_0 - x_0\, tg\alpha = b$, отримаємо **рівняння прямої через кутовий коефіцієнт**

$$y = kx + b \qquad (4.12)$$

Залишається отримати останню форму запису рівняння прямої – в **нормальному** (нормованому) вигляді. Для цього розглянемо довільну пряму L. Проведемо з початку координат до прямої L вектор нормалі $\vec{n}$ до перетину з прямою L (рис. 4.15).

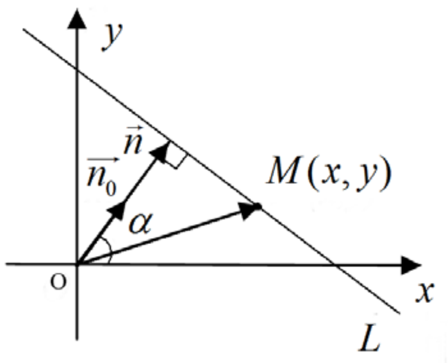

Рис. 4.15 Побудова нормального рівняння прямої

Нехай його довжина p, а кут між ним та віссю OX дорівнює α:

$$\left|\vec{n}\right| = p \geq 0 \qquad \left(\vec{n}\,\widehat{\ }\,OX\right) = \alpha$$

Вектор $\overrightarrow{n_0}$ – орт-вектор вектора нормалі $\vec{n}$:

$$\overrightarrow{n_0} = (\cos\alpha; \sin\alpha)$$

Очевидно, що довільна точка площини $M(x, y)$ лежить на даній прямій L тоді і лише тоді, коли проекція вектора $\overrightarrow{OM}$ на вісь вектора $\overrightarrow{n_0}$ дорівнює довжині вектора $\vec{n}$, тобто $pr_{\overrightarrow{n_0}}\overrightarrow{OM} = p$.

Згідно з визначенням скалярного добутку $pr_{\overrightarrow{n_0}}\overrightarrow{OM} = \left(\overrightarrow{n_0}, \overrightarrow{OM}\right)$.

Розписавши скалярний добуток в координатах, отримаємо:

$$x\cos\alpha + y\sin\alpha - p = 0 \tag{4.13}$$

Формула (4.13) визначає рівняння прямої в нормальному вигляді.

Для зведення загального рівняння прямої (4.7) до нормального вигляду (4.13) потрібно загальне рівняння помножити на нормуючий множник μ, який визначається формулою

$$\mu = \frac{sign\ C}{\sqrt{A^2 + B^2}} \tag{4.14}$$

Відстанню від точки $M_1(x_1, y_1)$ до прямої L називатимемо довжину перпендикуляра, який опущений з точки M на пряму L.

Теорема (про відстань від точки до прямої)

Нехай пряма L задана в ПДСК своїм рівнянням в нормальному вигляді (4.13). Тоді відстань $\rho\left(M_1, L\right)$ від довільної точки $M_1(x_1, y_1)$ до прямої L дорівнює числу

$$\rho\left(M_1, L\right) = \left|x_1 \cos\alpha + y_1 \sin\alpha - p\right| \tag{4.15}$$

□

Проведемо через точку M_1 пряму L_1 паралельно до заданої прямої L. Через початок координат проведемо пряму, перпендикулярну до прямих L та L_1, яка перетинає їх в точках N та N_1 відповідно.

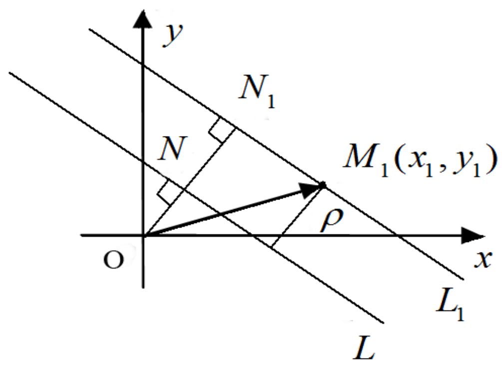

Рис. 4.16 Відстань від точки до прямої

Тоді шукана відстань $\rho\left(M_1, L\right)$ чисельно дорівнює довжині вектора $\overrightarrow{NN_1}$

$$\rho\left(M_1, L\right) = \left|\overrightarrow{NN_1}\right|$$

Вектор $\overline{NN_1}$ можна представити у вигляді $\overrightarrow{NN_1} = \overrightarrow{ON_1} - \overrightarrow{ON}$. Тоді $\left|\overrightarrow{NN_1}\right| = \left\|\overrightarrow{ON_1}\right| - \left|\overrightarrow{ON}\right\|$.

Враховуючи, що

$$\left|\overrightarrow{ON_1}\right| = x_1 \cos\alpha + y_1 \sin\alpha, \quad \left|\overrightarrow{ON}\right| = p$$

отримаємо:

$$\rho\left(M_1, L\right) = \left|x_1 \cos\alpha + y_1 \sin\alpha - p\right|.$$

■

Відхиленням точки $M_1(x_1, y_1)$ від прямої L, заданої рівнянням в нормальному вигляді (4.13), назвемо число, яке дорівнює відстані від точки до прямої, якщо точка M_1 і початок координат лежать по різні боки відносно прямої, та взятій зі знаком "-" відстані, якщо вони лежать по один бік:

$$\rho\left(M_1, L\right) = \pm\rho\left(M_1, L\right)$$

або

$$\rho\left(M_1, L\right) = \pm(x_1 \cos\alpha + y_1 \sin\alpha - p) \qquad (4.16)$$

Підкреслимо, що для обчислення і відхилення точки, і відстані рівняння прямої спочатку необхідно звести до вигляду (4.13).

Розглянемо взаємне розташування двох прямих на площині. Нехай вони задані своїми загальними рівняннями $A_1x + B_1y + C_1 = 0$ та $A_2x + B_2y + C_2 = 0$ відповідно.

Кут φ між прямими визначають як кут між їх векторами нормалей. Тоді з визначення скалярного добутку і його представлення в ПДСК можна отримати:

$$\cos\varphi = \frac{A_1A_2 + B_1B_2}{\sqrt{A_1^2 + A_2^2} \cdot \sqrt{B_1^2 + B_2^2}} \qquad (4.17)$$

Умова паралельності прямих еквівалентна умові колінеарності векторів нормалей:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \qquad (4.18)$$

Якщо при цьому ще й виконується рівність $\frac{A_1}{A_2} = \frac{C_1}{C_2}$, то прямі L_1 та L_2 співпадають.

Прямі L_1 та L_2 перпендикулярні, якщо $\cos\varphi = 0$. Тому умову перпендикулярності запишемо в наступному вигляді (виходячи з (4.17)):

$$A_1A_2 + B_1B_2 = 0 \qquad (4.19)$$

Якщо ж прямі L_1 та L_2 задані канонічними рівняннями $\frac{x-x_1}{p_1}=\frac{y-y_1}{q_1}$ та $\frac{x-x_2}{p_2}=\frac{y-y_2}{q_2}$, то формули (4.17), (4.18), (4.19) можна переписати у вигляді:

$$\cos\varphi=\frac{p_1p_2+q_1q_2}{\sqrt{p_1^2+q_1^2}\cdot\sqrt{p_2^2+q_2^2}} \tag{4.17а}$$

$$\frac{p_1}{p_2}=\frac{q_1}{q_2} \tag{4.18а}$$

$$p_1p_2+q_1q_2=0 \tag{4.19а}$$

У випадку задання прямих рівняннями з кутовими коефіцієнтами $y=k_1x+b_1$ та $y=k_2x+b_2$ отримаємо наступні вирази. Якщо α_1 та α_2 – кути нахилу прямих L_1 та L_2 до осі OX, то кут між цими прямими φ можна отримати як

$$\varphi=\left|\alpha_2-\alpha_1\right|$$

тобто,

$$tg\,\varphi=\left|tg\left(\alpha_2-\alpha_1\right)\right|=\frac{\left|tg\,\alpha_2-tg\,\alpha_1\right|}{1+tg\,\alpha_1\ tg\,\alpha_2}$$

Тоді формула обчислення кута φ матиме вигляд:

$$tg\,\varphi=\frac{\left|k_2-k_1\right|}{1+k_1k_2} \tag{4.17б}$$

Тому прямі паралельні в тому випадку, коли тангенс кута між ними дорівнює нулю, тобто

$$k_1 = k_2 \tag{4.18б}$$

Перпендикулярними прямі будуть в тому випадку, коли $tg\varphi$ не існує, тобто

$$1 + k_1 k_2 = 0 \Leftrightarrow k_1 = -\frac{1}{k_2} \tag{4.19б}$$

Сукупність прямих, що проходять через деяку точку M, називають **пучком прямих** з центром в M.

Теорема (про пучок прямих)

Якщо $A_1 x + B_1 y + C_1 = 0$ та $A_2 x + B_2 y + C_2 = 0$ – рівняння двох різних прямих, які перетинаються в точці $M_0(x_0, y_0)$, а α та β – довільні дійсні числа, які одночасно не дорівнюють нулеві, то рівняння

$$\alpha\left(A_1 x + B_1 y + C_1\right) + \beta\left(A_2 x + B_2 y + C_2\right) = 0 \tag{4.20}$$

визначає пряму, що проходить через точку M. Більше того, яка б не була заздалегідь задана пряма, що проходитиме через точку M, вона визначається рівнянням (4.20) при деяких значеннях α і β.

□

Покажемо спочатку, що при будь-яких значеннях α та β, одночасно ненульових, рівність (4.20) є рівнянням першого порядку. Розкривши в ній дужки та звівши подібні, отримаємо:

$$\left(\alpha A_1 + \beta A_2\right)x + \left(\alpha B_1 + \beta B_2\right)y + \left(\alpha C_1 + \beta C_2\right) = 0$$

Якби мали місце рівності $\alpha A_1 + \beta A_2 = 0$ та $\alpha B_1 + \beta B_2 = 0$, то з них можна було б отримати, що

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = -\frac{\beta}{\alpha}$$

Але цього бути не може, оскільки останній вираз є умовою паралельності прямих, що суперечить припущенню теореми. Тобто, (4.20) – рівняння першого порядку, тому воно визначає деяку пряму. Більше того, ця пряма проходить через точку $M_0(x_0, y_0)$, оскільки

$$A_1x_0 + B_1y_0 + C_1 = 0 \text{ і } A_2x_0 + B_2y_0 + C_2 = 0,$$

а тоді

$$\alpha\left(A_1x_0 + B_1y_0 + C_1\right) + \beta\left(A_2x_0 + B_2y_0 + C_2\right) = 0$$

Покажемо тепер, що будь-яка пряма, що проходить через точку M_0, визначається рівнянням (4.20) при деяких значеннях α і β. Наперед задана пряма, що проходить через точку M_0, однозначно визначається заданням ще однієї своєї точки $M_1(x_1, y_1)$, відмінної від M_0. Тобто, потрібно довести, що можна вибрати не рівні одночасно нулю числа α і β так, що координати наперед заданої точки M_1 будуть задовольняти (4.20).

Підставимо в (4.20) координати точки M_1:

$$\alpha\left(A_1x_1 + B_1y_1 + C_1\right) + \beta\left(A_2x_1 + B_2y_1 + C_2\right) = 0 \qquad (4.21)$$

Останній вираз являється рівнянням відносно α та β, оскільки обидва вирази в круглих дужках одночасно перетворитись в нуль не можуть (інакше прямі L_1 та L_2 проходили б через точку M_1, що неможливо). Тобто, значення хоча б в одній з цих дужок відмінне від нуля. Нехай це значення в перших дужках. Тоді, задавши довільно $\beta \neq 0$, з (4.21) знайдемо α:

$$\alpha = -\beta\frac{A_2x_1 + B_2y_1 + C_2}{A_1x_1 + B_1y_1 + C_1}.$$

Тобто, при вказаних α та β пряма (4.20) проходитиме через точку M_1.

Аналогічно у випадку з виразом у других дужках.

■

Рівняння пучка прямих дає можливість записати рівняння будь-якої прямої, що проходить через точку перетину двох даних прямих, не обчислюючи координат цієї точки перетину.

Приклад:

1) Записати рівняння бісектрис кутів, утворених прямими:

$$L_1: \ 3x-7y+5=0 \quad \text{та} \quad L_2: \ 7x-3y-1=0$$

Для точки $M(x_0, y_0)$, яка лежить на бісектрисі, виконується умова (рис. 4.17):

$$\rho\left(M, L_1\right)=\rho\left(M, L_2\right)$$

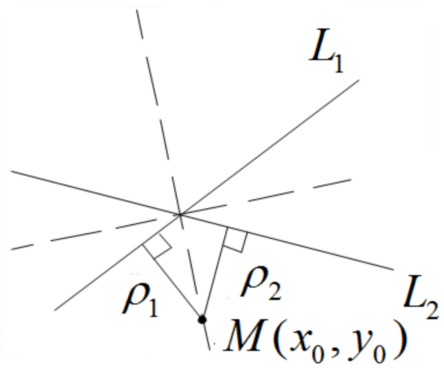

Рис. 4.17 Задача про побудову бісектрис кутів

Запишемо рівняння прямих в нормальному вигляді:

$$L_{1H}: \ -\frac{3}{\sqrt{58}}x+\frac{7}{\sqrt{58}}y-\frac{5}{\sqrt{58}}=0 \ \text{та} \ L_{2H}: \ \frac{7}{\sqrt{58}}x-\frac{3}{\sqrt{58}}y-\frac{1}{\sqrt{58}}=0$$

Тоді отримаємо:

$$\frac{\left|-3x_0+7y_0-5\right|}{\sqrt{58}}=\frac{\left|7x_0-3y_0-1\right|}{\sqrt{58}}$$

Оскільки точка $M(x_0, y_0)$ – будь-яка точка бісектрис, то розкриваючи модулі, отримаємо, що бісектриси визначаються рівніннями

$$10x - 10y + 4 = 0 \text{ і } 4x + 4y + 6 = 0,$$

або

$$5x - 5y + 2 = 0 \text{ і } 2x + 2y + 3 = 0.$$

2) Визначити, у якому з кутів, утворених прямими $L_1:\ 3x - 5y - 4 = 0$ і $L_2:\ x + 2y + 3 = 0$, гострому чи тупому, лежить початок координат.

Проведемо з початку координат O вектори нормалей $\overrightarrow{n_1}$ і $\overrightarrow{n_2}$ до прямих L_1 і L_2 відповідно. Розглянемо чотирикутник $OABC$ (рис. 4.18):

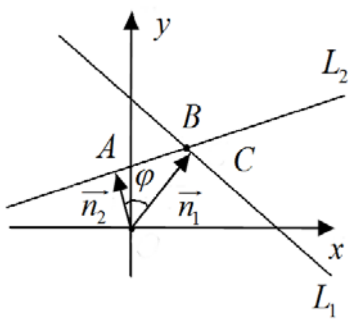

Рис. 4.18 Задача про визначення кута із початком координат

Очевидно, якщо кут $\angle AOC$ (між векторами нормалей) гострий, то кут $\angle ABC$ (між прямими, де лежить початок координат) тупий, і навпаки.

Тобто, необхідно знайти кут між векторами $\overrightarrow{n_1}$ і $\overrightarrow{n_2}$ заданого напрямку. Зведемо рівняння прямих до нормального вигляду:

$$L_{1H}: \ \frac{3}{\sqrt{34}}x-\frac{5}{\sqrt{34}}y-\frac{4}{\sqrt{34}}=0 \text{ та } L_{2H}: \ -\frac{1}{\sqrt{5}}x-\frac{2}{\sqrt{5}}y-\frac{3}{\sqrt{5}}=0$$

Тоді

$$\overrightarrow{n_1}=\left({}^{3}\!\big/\!{}_{\sqrt{34}};-{}^{5}\!\big/\!{}_{\sqrt{34}} \right) \text{ та } \overrightarrow{n_2}=\left(-{}^{1}\!\big/\!{}_{\sqrt{5}};-{}^{2}\!\big/\!{}_{\sqrt{5}} \right)$$

$$\cos\varphi=\left(\overline{n_1},\overline{n_2}\right)=\frac{1}{\sqrt{170}}\left(-3+10\right)=\frac{7}{\sqrt{170}}>0$$

тобто φ гострий, і початок координат лежить в тупому куті.

<u>Контрольні запитання:</u>

1. Якими способами можна записати рівняння прямої?
2. Що таке відхилення та відстань точки від прямої?
3. Що таке пучок прямих?

<u>Література</u>: [5], с. 118-136; [6], с. 41-55.

4.3. Лінії 2-го порядку. Перетворення координат в прямокутній декартовій системі координат

Загальний вигляд рівняння алгебраїчної лінії другого порядку визначається рівнянням:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{4.22}$$

у якому хоча б один із коефіцієнтів A, B, C відмінний від нуля (інакше (4.22) перетворюється на рівняння лінії 1-го порядку). Рівняння (4.22) визначає вироджену криву, якщо вона є лінією першого або нульового порядку, і невироджену в іншому випадку.

Невиродженими лініями другого порядку є еліпс, гіпербола та парабола. Дамо визначення кожній з цих ліній, отримаємо їх канонічні рівняння, побудуємо ці лінії та розглянемо основні характеристики.

Еліпсом називають геометричне місце точок площини, для яких сума відстаней до двох фіксованих точок F_1 та F_2 цієї площини є величина стала. Точки F_1 та F_2 при цьому називають **фокусами** еліпса, а пряму, що проходить через них – **фокальною віссю**. Фокуси еліпса можуть збігтися, і тоді отримують коло.

Задамо на площині ПДСК наступним чином. За початок відліку візьмемо середину відрізка F_1F_2, ось OY проведемо через цю точку перпендикулярно відрізку, вісь OX буде збігатися з прямою F_1F_2, додатний напрям визначимо від F_1 до F_2 (рис. 4.19)

.

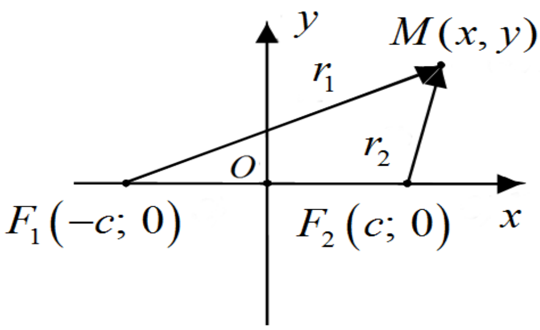

Рис. 4.19 Задання системи координат для еліпсу

Нехай довжина відрізка F_1F_2 дорівнює $2c$. Тоді координати точок F_1 та F_2 у введеній ПДСК дорівнюють $(-c;0)$ та $(c;0)$ відповідно. Позначимо згадану у визначенні еліпса постійну величину через $2a$. Очевидно, що $2a > 2c$ (на підставі того, що сума довжин будь-яких двох сторін трикутника більша за довжину третьої сторони, що залишилася). Нехай $M(x,y)$ – довільна точка площини, а r_1 та r_2 – відстані від неї до точок F_1 та F_2 відповідно. З визначення еліпса випливає, що необхідною і достатньою умовою розташування точки M на даному еліпсі є рівність

$$r_1 + r_2 = 2a \qquad (4.23)$$

Використовуючи формулу відстані між двома точками, отримаємо:

$$r_1 = \sqrt{(x+c)^2 + y^2} \qquad r_2 = \sqrt{(x-c)^2 + y^2} \qquad (*)$$

Тоді

$$r_1 + r_2 = \sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a$$

Позбавимось радикалів:

$$\left(x+c\right)^2+y^2+\left(x-c\right)^2+y^2+2\sqrt{\left(\left(x+c\right)^2+y^2\right)\left(\left(x-c\right)^2+y^2\right)}=4a^2$$

$$x^2+c^2+y^2-2a^2=-\sqrt{\left(\left(x+c\right)^2+y^2\right)\left(\left(x-c\right)^2+y^2\right)}$$

$$x^4+c^4+y^4+4a^4+2x^2c^2+2x^2y^2-4x^2a^2+2c^2y^2-4c^2a^2-4y^2a^2=$$

$$=\left(x^2+y^2+c^2+2cx\right)\left(x^2+y^2+c^2-2cx\right)$$

$$x^4+c^4+y^4+4a^4+2x^2c^2+2x^2y^2-4x^2a^2+2c^2y^2-4c^2a^2-4y^2a^2=$$

$$=x^4+y^4+c^4+2x^2c^2+2y^2c^2+2x^2y^2-4c^2x^2$$

$$a^2c^2+a^2x^2+a^2y^2-c^2x^2-a^4=0$$

Перегрупуємо елементи:

$$x^2\left(a^2-c^2\right)+a^2y^2=a^2\left(a^2-c^2\right)$$

Оскільки $a>c$, то $a^2-c^2>0$, і, ввівши позначення $a^2-c^2=b^2$, отримаємо:

$$b^2x^2+a^2y^2=a^2b^2$$

або, розділивши обидві частини на a^2b^2:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,\quad a>b \tag{4.24}$$

Тобто, координати будь-якої точки M еліпса задовольняють рівняння (4.24).

Оскільки в процесі алгебраїчних перетворень могли з'явитися так звані "зайві корені", то потрібно переконатися в тому, що будь-яка точка M, координати якої задовольняють рівняння (4.24), розташована на даному еліпсі.

Для цього достатньо довести, що величини r_1 та r_2 для такої точки задовольняють співвідношення (4.23). Нехай координати деякої точки M площини задовольняють рівняння (4.24), тобто для цієї точки:

$$y^2 = b^2\left(1 - \frac{x^2}{a^2}\right) = \left(a^2 - c^2\right)\left(1 - x^2\!/a^2\right)$$

Тоді

$$r_1 = \sqrt{\left(x+c\right)^2 + y^2} = \sqrt{\left(x+c\right)^2 + \left(a^2 - c^2\right)\left(1 - x^2\!/a^2\right)} =$$

$$= \sqrt{x^2 + c^2 + 2xc + a^2 - c^2 - x^2 + x^2\frac{c^2}{a^2}} = \sqrt{\left(a + x\frac{c}{a}\right)^2} = a + x\frac{c}{a}$$

Аналогічно,

$$r_2 = a - x\frac{c}{a} \qquad (**)$$

Тоді виконується умова (4.23), тобто точка M лежить на заданому еліпсі. Тобто рівняння (4.24) визначає еліпс, і називається канонічним рівнянням еліпса. Величини a та b називають великою та малою півосями еліпса. Очевидно, що в випадку $a = b$ еліпс перетворюється на коло.

Розглянемо геометричні властивості еліпса, що випливають з канонічного рівняння.

Оскільки в рівнянні (4.24) величини x та y присутні в парних степенях, то тоді, якщо це рівняння задовольняють координати точки $M(x, y)$, йому також задовольняють і координати точок $M_1(x, -y)$, $M_2(-x, y)$, $M_3(-x, -y)$:

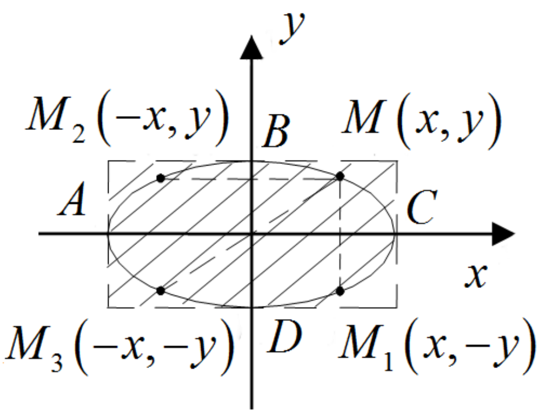

Рис. 4.20 Геометричні властивості еліпса

Тобто еліпс має дві взаємно перпендикулярні осі симетрії (головні осі еліпса) і центр симетрії (центр еліпса). У випадку з канонічним рівнянням головні осі збігаються з осями координат, а центр – з початком координат.

Точки перетину еліпса з головними осями називають **вершинами еліпса**. У випадку канонічного задання вершини еліпса мають координати

$$A(-a, 0), B(0, b), C(a, 0), D(0, -b)$$

Розглянемо принципи побудови еліпса. Оскільки $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, то, очевидно, $|x|\le a$ та $|y|\le b$, тобто весь еліпс міститься всередині прямокутника $|x|\le a$, $|y|\le b$, який називають основним прямокутником.

Крім того, еліпс можна побудувати за допомогою рівномірного стискування кола, радіус якого $R=a$.

Гіперболою називається геометричне місце точок площини, для яких абсолютна величина різниці відстаней до двох фіксованих точок F_1 та F_2 цієї площини є величина стала. Ці фіксовані точки називають **фокусами гіперболи**, а пряму, що проходить через них – **фокальною віссю**. Фокуси гіперболи завжди різні.

Задамо на площині ПДСК аналогічно до випадку з еліпсом. Нехай довжина відрізка F_1F_2 дорівнює $2c$, тоді $(-c,0)$ і $(c,0)$ – координати точок F_1 і F_2 в заданій ПДСК. Позначимо сталу величину, згадану у визначенні гіперболи, через $2a$. Очевидно, що $2a<2c$, тобто $a<c$ (оскільки модуль різниці довжин двох сторін трикутника менший за довжину третьої сторони). Нехай M – точка площини з координатами (x,y), r_1 і r_2 – відстані від неї до фокусів F_1 і F_2 відповідно.

Умову розташування точки $M(x,y)$ на гіперболі (необхідну та достатню), виходячи з визначення, запишемо у вигляді

$$|r_1-r_2|=2a \tag{4.25}$$

З урахуванням виразу (*) для r_1 та r_2 отримаємо:

$$\left|\sqrt{(x+c)^2+y^2}-\sqrt{(x-c)^2+y^2}\right|=2a$$

або, позбавившись радикалів

$$\left[(x+c)^2+y^2\right]+\left[(x-c)^2+y^2\right]=4a^2+2\sqrt{\left[(x+c)^2+y^2\right]\left[(x-c)^2+y^2\right]}$$

$$x^2+c^2+y^2-2a^2=\sqrt{\left(x^2+c^2+y^2\right)^2-4c^2x^2}$$

$$x^4+c^4+y^4-4a^4+2x^2c^2+2x^2y^2-4x^2a^2+2c^2y^2-4c^2a^2-4y^2a^2=$$

$$x^4+y^4+c^4+2x^2c^2+2x^2y^2+2c^2y^2-4c^2x^2$$

Згрупуємо

$$x^2\left(c^2-a^2\right)-y^2a^2=a^2\left(c^2-a^2\right)$$

і введемо позначення $b^2=c^2-a^2$. Отриманий вираз поділимо на a^2b^2:

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \tag{4.26}$$

Тобто, будь-яка точка гіперболи задовольняє рівняння (4.26).

Так само, як і у випадку з еліпсом, потрібно переконатися в тому, що не додалися "зайві корені", тобто потрібно показати, що будь-яка точка M, координати якої задовольняють співвідношення (4.26), лежить на заданій гіперболі, і для неї має місце співвідношення (4.25). Проводячи аналогічні перетворення, як й у випадку виведення формули (**), отримаємо:

$$r_1 = \begin{cases} a + \dfrac{c}{a}x, x > 0 \\ -a - \dfrac{c}{a}x, x < 0 \end{cases} \qquad r_2 = \begin{cases} -a + \dfrac{c}{a}x, x > 0 \\ a - \dfrac{c}{a}x, x < 0 \end{cases}$$

Тобто виконується умова (4.25), і точка M належить заданій гіперболі. Таким чином, рівняння (4.26) визначає гіперболу, і його називають **канонічним рівнянням гіперболи**. Величини a та b, що входять до рівняння, називають відповідно **дійсною** та **уявною півосями** гіперболи.

Гіпербола має дві осі симетрії (головні осі гіперболи) і центр симетрії (центр гіперболи). При цьому одна з осей перетинається з гіперболою в двох точках, які називають вершинами, а сама ця вісь називається дійсною віссю. Інша вісь, яка є серединним перпендикуляром до відрізка, що містить вершини гіперболи, не має з гіперболою спільних точок; тому її називають уявною віссю гіперболи. Тобто, уявна вісь гіперболи розділяє площину на праву й ліву півплощини, у яких розташовані симетричні відносно цієї осі права та ліва гілки гіперболи (рис. 4.21).

Це випливає з того, що в рівнянні (4.26) величини x і y беруть участь з парними степенями, а отже, якщо точка $M(x,y)$ задовольняє це рівняння, то його також задовольняють і точки $M_1(x,-y), M_2(-x,y), M_3(-x,-y)$.

У випадку задання гіперболи канонічним рівнянням центр гіперболи збігається з початком координат, вісь OX є дійсною, вісь OY – уявною віссю гіперболи, вершини гіперболи знаходяться в точках $A(-a,0)$ і $B(a,0)$. Фокуси гіперболи завжди знаходяться на дійсній осі.

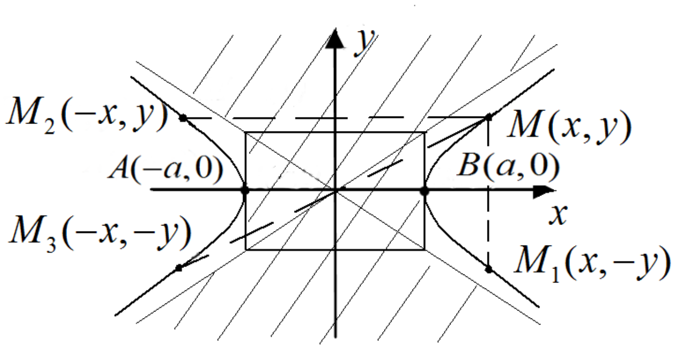

Рис. 4.21 Геометричні властивості гіперболи

Розглянемо далі область G, яку отримують шляхом об'єднання прямокутника $D=\{(x,y)\,|\,|x|<a\,\&\,|y|<b\}$ (його називають основним прямокутником), і тих двох кутів, утворених діагоналями цього прямокутника, в яких лежить уявна вісь гіперболи. В цій області G немає жодної точки гіперболи (4.26).

Для доведення цього факту необхідно розбити область G на дві частини: G_1 – смуга $|x|<a$ та G_2 – решта.

Очевидно, що в смузі G_1 точки гіперболи відсутні, оскільки з рівняння (4.26) випливає, що $|x|\geq a$. Що стосується області G_2, то будь-яка її точка лежить або на діагоналі прямокутника D, або за його діагоналлю. Діагоналі прямокутника D визначаються рівняннями

$$y=\pm\frac{b}{a}x \tag{4.27}$$

Отже, координати точок області G_2 задовольняють нерівність

$$\frac{b}{a} \le \frac{|y|}{|x|}$$

Але тоді

$$\frac{|x|}{a} \le \frac{|y|}{b} \Leftrightarrow \frac{x^2}{a^2} \le \frac{y^2}{b^2} \Leftrightarrow \frac{x^2}{a^2} \le \frac{x^2}{a^2} - 1 \Leftrightarrow 0 \le -1$$

Отримана нерівність підтверджує той факт, що в області G_2 немає точок, що задовольняють співвідношення (4.26).

Наступна властивість полягає в наявності біля гіперболи асимптот, а саме: гілки гіперболи наближаються до діагоналей прямокутника D. Через симетричність гіперболи достатньо встановити істинність цього твердження в першій чверті.

Відповідна гілка гіперболи задається виразом:

$$y = b\sqrt{\frac{x^2}{a^2} - 1}$$

Асимптота є прямою, а її рівняння:

$$y_a = k_a x + b_a,$$

де

$$k_a = \lim_{x\to\infty} \frac{b\sqrt{\dfrac{x^2}{a^2}-1}}{x} = \lim_{x\to\infty} \frac{b}{a}\sqrt{\frac{x^2-a^2}{x^2}} = \frac{b}{a}$$

$$b_a = \lim_{x\to\infty}\left(y(x)-k_a x\right) = \lim_{x\to\infty}\left(b\sqrt{\frac{x^2}{a^2}-1}-\frac{b}{a}x\right) =$$

$$= b\lim_{x\to\infty} \frac{\dfrac{x^2}{a^2}-1-\dfrac{x^2}{a^2}}{\sqrt{\dfrac{x^2}{a^2}-1}+\dfrac{x}{a}} = 0,$$

тобто

$$y_a = \frac{b}{a}x.$$

Таким чином, прямі (4.27) справді є асимптотами гіперболи.

Фактично проведене вище дослідження дає метод побудови гіперболи. Для цього потрібно послідовно побудувати наступні елементи:

1) центр гіперболи,
2) вісі гіперболи,
3) основний прямокутник,
4) асимптоти,
5) вершини,
6) гіпербола.

Разом з гіперболою (4.26) часто розглядають так звану спряжену до неї гіперболу

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=-1 \qquad (4.26а)$$

Парабола – це геометричне місце точок площини, для яких відстань до деякої фіксованої точки F цієї площини дорівнює відстані до деякої фіксованої прямої, також розташованої в цій площині. Точку F називають **фокусом параболи**, а пряму – **директрисою параболи**.

Задамо ПДСК на площині наступним чином. Оберемо початок відліку як середину відрізка FD, що є перпендикуляром, опущеним з фокуса на директрису. Вісь OX проведемо через цей перпендикуляр, а вісь OY – паралельно директрисі (рис. 4.22).

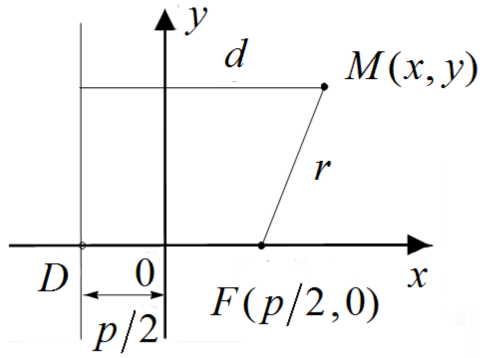

Рис. 4.22 Задання системи координат для параболи

Нехай довжина відрізка FD дорівнює p. Тоді в обраній системі координат точка F має координати $\left(p/2,0\right)$. Нехай M – довільна точка площини з координатами $\left(x,y\right)$. Позначимо через r відстань від M до F, а через d – відстань від M до директриси. За визначенням параболи, рівність

$$r = d \tag{4.28}$$

є необхідною та достатньою умовою розташування точки M на заданій параболі.

З урахуванням того, що

$$r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2}\,,\ d = \frac{p}{2} + x \qquad (***)$$

відношення (4.28) можна переписати у вигляді:

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = x + \frac{p}{2},$$

чи, позбавившись радикала,

$$\left(x - \frac{p}{2}\right)^2 + y^2 = \left(x + \frac{p}{2}\right)^2$$

$$x^2 + \frac{p^2}{4} - px + y^2 = x^2 + \frac{p^2}{4} + px$$

$$y^2 = 2px,\ x \geq 0 \qquad (4.29)$$

Тобто будь-яка точка параболи задовольняє рівняння (4.29).

Зазначимо, що умову $x \geq 0$ можна було накласти ще при записі відношення (***),оскільки легко помітити, що у випадку $x < 0$ – $r > d$, і такі точки одразу виключаються з розгляду.

Перевіримо, що рівняння (4.29) не набуло нових коренів. Для цього доведемо, що для кожної точки $M(x, y)$, яка задовольняє (4.29), величини r та d однакові, тобто виконується відношення (4.28).

Оскільки в (4.29) $x \ge 0$, то відстань до директриси $d = x + \frac{p}{2}$. Тоді, підставивши вираз для y^2 з (4.29) у вираз для r із (***), отримаємо:

$$r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \sqrt{\left(x - \frac{p}{2}\right)^2 + 2px} = \sqrt{x^2 + \frac{p^2}{4} - px + 2px} =$$

$$= \sqrt{x^2 + \frac{p^2}{4} + px} = \sqrt{\left(x + \frac{p}{2}\right)^2} = x + \frac{p}{2}, \text{ бо } x \ge 0$$

Тобто $r = d$, і точка M лежить на параболі.

Рівняння (4.29) називають канонічним рівнянням параболи, а величину p – параметром параболи.

Оскільки в рівнянні (4.29) величина y фігурує в парному степені, то якщо точка $M(x, y)$ задовольняє це рівняння, точка $M'(x, -y)$ також його задовольняє (рис. 4.23).

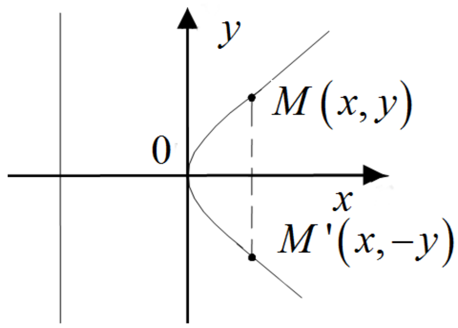

Рис. 4.23 Властивості параболи

Таким чином, парабола має вісь симетрії (вісь параболи). Точку перетину параболи з віссю називають **вершиною параболи**.

Якщо парабола задана своїм канонічним рівнянням (4.29), то ось параболи – це ось OX, а вершина параболи співпадає з початком координат. При цьому вся парабола розташована в правій півплощині площини OXY, а рівняння директриси параболи має вигляд:

$$x = -\frac{p}{2}$$

Зауважимо, що будь-які дві параболи подібні одна одній.

Якщо уважно проаналізувати визначення еліпса, гіперболи та параболи, можна помітити, що способи визначення кривих різняться. Якщо перші дві криві визначались через відстань до двох фіксованих точок, то парабола – через відстань від точки до прямої. Більше того, визначення параболи можна переформулювати наступним чином.

Парабола – геометричне місце точок площини, для яких відношення відстані до фокуса та відстані до відповідної цьому фокусу директриси є величина стала, і вона дорівнює одиниці.

Виявляється, що аналогічним чином можна переформулювати визначення гіперболи й еліпса, відмінного від кола.

Ексцентриситетом еліпса (гіперболи) називають величину ε, яка дорівнює:

$$\varepsilon = \frac{c}{a} \qquad (4.30)$$

Уведемо параметр ε для еліпса $\varepsilon = \sqrt{1 - \frac{b^2}{a^2}} < 1$ та для гіперболи $\varepsilon = \sqrt{1 + \frac{b^2}{a^2}} > 1$.

Ексцентриситет кола дорівнює нулю.

Геометричний зміст ексцентриситету:

1) Для еліпса він вказує міру "сплющеності": чим більше ексцентриситет, тим менше відношення малої півосі еліпса b до його великої півосі a.

2) У випадку гіперболи ексцентриситет можна розглядати як числову характеристику величини кута між її асимптотами, оскільки $\frac{b}{a} = tg\frac{\varphi}{2}$, φ – кут між асимптотами.

Розглянемо еліпс, заданий рівнянням (4.24). **Директрисою** D_i, що відповідає фокусу F_i, називають пряму, перпендикулярну великій осі еліпса, що проходить на відстані $\frac{a}{\varepsilon}$ від його центра.

Рівняння директрис мають вигляд

$$D_1 : x = -\frac{a}{\varepsilon} \quad D_2 : x = \frac{a}{\varepsilon} \tag{4.31}$$

Директриси гіперболи визначають аналогічно, і вони задаються такими самими рівняннями.

Очевидно, що директриси еліпса розташовані поза еліпсом, а директриси гіперболи розташовані в області G, яка не містить точок гіперболи.

Точки еліпса і його центр розташовані по один бік від кожної з його директрис.

Точки лівої (правої) гілки гіперболи та її центр розташовані по різні боки директриси $D_1\left(D_2\right)$, а точки правої (лівої) гілки гіперболи та її центр O розташовані по один бік від директриси $D_1\left(D_2\right)$.

<u>Теорема (про ексцентриситет)</u>

Відношення відстані r_i від точки M еліпса (гіперболи) до фокуса F_i і відстані d_i від цієї точки, що відповідає цьому фокусу директриси D_i, дорівнює ексцентриситету ε цього еліпса (гіперболи).

Приймемо теорему без доведення.

На підставі цієї теореми і визначення параболи можна дати наступне визначення відмінного від кола еліпса, гіперболи і параболи.

Геометричне місце точок площини, для яких відношення ε відстані r (до фокуса F) та відстані d від цієї точки до прямої D, є величина стала та є або еліпсом (при $o<\varepsilon<1$), або параболою (при $\varepsilon=1$), або гіперболою (при $\varepsilon>1$). Точку F при цьому називають **фокусом**, пряма D – **директрисою**, а ε – **ексцентриситетом** геометричного місця точок.

Розглянемо перетворення прямокутних декартових координат на площині. Нехай на площині задано дві праві довільні ПДСК: перша визначається початком O й базисними ортами $\bar{i}$ та $\bar{j}$, друга визначається початком O' і базовими векторами $\bar{i}'$ та $\bar{j}'$, при цьому кути між ортами $\bar{i}$ і $\bar{i}'$ та $\bar{j}$ і $\bar{j}'$ однакові.

Потрібно встановити зв'язок між координатами довільної точки M площини в системі $\left(x,y\right)$ і в системі $\left(x',y'\right)$.

Зазначимо, що координати x і y співпадають з координатами вектора $\overrightarrow{OM}$ в базисі $\{\bar{i}, \bar{j}\}$, а координати x' і y' співпадають з координатами вектора $\overrightarrow{O'M}$ в базисі $\{\bar{i}', \bar{j}'\}$:

$$\overrightarrow{OM} = x\vec{i} + y\vec{j} \tag{4.32}$$

$$\overrightarrow{O'M} = x'\vec{i}' + y'\vec{j}' \tag{4.33}$$

Якщо покласти, що точка O' має в першій системі координати (x_0, y_0), то тоді вектор $\overrightarrow{OO'}$ можна розкласти за базисом $\{\vec{i}, \vec{j}\}$ наступним чином:

$$\overrightarrow{OO'} = x_0\vec{i} + y_0\vec{j} \tag{4.34}$$

Будь-який вектор на площині можна розкласти за базисом:

$$\vec{i}' = \alpha\vec{i} + \beta\vec{j}, \quad \vec{j}' = \gamma\vec{i} + s\vec{j} \tag{4.35}$$

За правилом трикутника додавання векторів маємо:

$$\overrightarrow{OM} = \overrightarrow{OO'} + \overrightarrow{O'M} \tag{4.36}$$

Підставляючи в (4.36) вирази для векторів $\overrightarrow{OM}, \overrightarrow{O'M}$ і $\overrightarrow{OO'}$ з (4.32), (4.33), (4.34), і замінюючи входження векторів $\bar{i}'$ і $\bar{j}'$ іншого базиса на їх вирази (4.35), отримаємо:

$$x\vec{i} + y\vec{j} = x_0\vec{i} + y_0\vec{j} + x'\left(\alpha\vec{i} + \beta\vec{j}\right) + y'\left(\gamma\vec{i} + s\vec{j}\right),$$

або, після групування в правій частині:

$$x\vec{i} + y\vec{j} = \left(x_0 + \alpha x' + \gamma y'\right)\vec{i} + \left(y_0 + \beta x' + s y'\right)\vec{j},$$

звідки, внаслідок єдиності розкладання вектора за базисом, отримаємо:

$$\begin{cases} x = x_0 + \alpha x' + \gamma y' \\ y = y_0 + \beta x' + s y' \end{cases} \tag{4.33}$$

Таким чином, з (4.33) видно, що, які б не були дві довільні ДСК, координати будь-якої точки відносно першої системи є лінійними функціями координат цієї самої точки відносно другої системи.

Позначивши через φ кут між векторами $\bar{i}$ і $\bar{i}'$, і домножаючи скалярно на $\vec{i}$ і $\vec{j}$ по черзі, перепишемо формули (4.33) у вигляді

$$\begin{cases} x = x_0 + x'\cos\varphi - y'\sin\varphi \\ y = y_0 + x'\sin\varphi + y'\cos\varphi \end{cases} \tag{4.34}$$

Таким чином, які б не були дві праві ДСК OXY та $O'X'Y'$, першу з них можна поєднати з другою за допомогою паралельного переносу вздовж вектора $\overrightarrow{OO'}$ і подальшого повороту наввколо початку координат на деякий кут φ. Загальне перетворення координат (4.34) можна розбити на суму двох перетворень, перше з яких відповідає лише паралельному переносу системи, а інше – лише повороту системи на кут φ.

Дійсно, вважаючи в (4.34) кут повороту φ нульовим, отримаємо формули перетворення координат при паралельному переносі системи з точки $O(0,0)$ в точку $O'(x_0, y_0)$:

$$\begin{cases} x = x' + x_0 \\ y = y' + y_0 \end{cases} \tag{4.34а}$$

Поклавши $x_0 = y_0 = 0$, отримаємо формули перетворення координат при обертанні системи навколо початку координат на кут φ:

$$\begin{cases} x = x'\cos\varphi - y'\sin\varphi \\ y = x'\sin\varphi + y'\cos\varphi \end{cases} \tag{4.34б}$$

Зазначимо, що поворот здійснюється проти годинникової стрілки.

Раніше було зазначено, що лініями другого порядку є:

еліпс $\quad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$;

гіпербола $\quad \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$;

парабола $\quad y^2 = 2px$.

Крім цього, рівняння лінії другого порядку також може визначати:

точку $\quad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$ (вироджений еліпс);

уявний еліпс $\quad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$;

прямі, що перетинаються $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$;

паралельні прямі $x^2 = a^2$;

одну пряму $x^2 = 0$;

пару уявних прямих $x^2 = -a^2$.

Визначення типу лінії є досить простим лише в тому випадку, коли воно записано в канонічному вигляді (тобто за допомогою одного з перелічених вище).

У випадку, коли лінію 2-го порядку задано своїм загальним рівнянням (4.22), то його можна звести до одного з перелічених канонічних рівнянь, використовуючи перетворення координат (паралельний перенос й поворот).

Досить легко можна показати, що перетворення паралельного переносу не змінює квадратичну частину рівняння (4.22), а перетворення повороту змінює всі коефіцієнти, крім вільного члена.

При цьому будь-яке рівняння

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \qquad (4.22)$$

лінії другого порядку шляхом спеціального повороту системи координат можна звести до рівняння, у якому не міститиметься доданок $x'y'$, тобто такого рівняння, в якому $B' = 0$. Таке спрощення називатимемо **стандартним**. Очевидно, що при цьому передбачається відмінність від нуля коефіцієнта B вихідного рівняння, оскільки при $B = 0$ це перетворення непотрібне.

Нехай φ – кут повороту системи координат. Тоді, підставляючи в рівняння (4.22) вирази для координат x і y через координати x' і y' нової системи координат з формул (4.34б), отримаємо:

$$A\left(x'\cos\varphi - y'\sin\varphi\right)^2 + B\left(x'\cos\varphi - y'\sin\varphi\right)\left(x'\sin\varphi + y'\cos\varphi\right) +$$
$$+C\left(x'\sin\varphi + y'\cos\varphi\right)^2 + D\left(x'\cos\varphi - y'\sin\varphi\right) + E\left(x'\sin\varphi + y'\cos\varphi\right) + F = 0$$

або, розкриваючи дужки та перетворюючи:

$$A\left(x'^2\cos^2\varphi + y'^2\sin^2\varphi - x'y'\sin 2\varphi\right) +$$
$$+B\left(x'^2\cos\varphi\sin\varphi + x'y'\cos^2\varphi - x'y'\sin^2\varphi - y'^2\sin\varphi\cos\varphi\right) +$$
$$+C\left(x'^2\sin^2\varphi + y'^2\cos^2\varphi + x'y'\sin 2\varphi\right) + D\left(x'\cos\varphi - y'\sin\varphi\right) +$$
$$+E\left(x'\sin\varphi + y'\cos\varphi\right) + F = 0$$

або, групуючи:

$$x'^2\left(A\cos^2\varphi + B\cos\varphi\sin\varphi + C\sin^2\varphi\right) + x'y'\left(-A\sin 2\varphi + B\cos 2\varphi + C\sin 2\varphi\right) +$$
$$+y'^2\left(A\sin^2\varphi - B\sin\varphi\cos\varphi + C\cos^2\varphi\right) + x'\left(D\cos\varphi + E\sin\varphi\right)$$
$$+y'\left(-D\sin\varphi + E\cos\varphi\right) + F = 0$$

Очевидно, прирівнюючи до нуля коефіцієнт при добутку $x'y'$, отримаємо умову для знаходження потрібного кута повороту:

$$-A\sin 2\varphi + B\cos 2\varphi + C\sin 2\varphi = 0,$$

або

$$tg2\varphi = \frac{B}{A-C} \tag{4.35}$$

Таким чином, поворот системи координат на кут, який визначається співвідношенням (4.35), зводить рівняння лінії 2-го порядку (4.33) до вигляду

$$A'x'^2 + C'y'^2 + D'x' + E'y' + F = 0 \tag{4.36}$$

Для зведення до канонічного вигляду тепер необхідно виділити повні квадрати за змінними x' та y' і здійснити паралельний перенос.

Геометричний поворот системи координат на кут φ означає зміщення напряму осей лінії 2-го порядку та осей координат, а паралельне перенесення означає "суміщення" центрів системи координат і ліній.

<u>*Приклад*</u>:

Привести рівняння лінії другого порядку $9x^2 - 4xy + 6y^2 + 16x - 8y - 2 = 0$ до канонічного вигляду

$$A = 9 \quad B = -4 \quad C = 6$$

$$tg2\varphi = \frac{-4}{9-6} = -\frac{4}{3}$$

$$tg2\varphi < 0 \Rightarrow 2\varphi \in 2\text{-га чверть} \Rightarrow \varphi \in 1\text{-а чверть}$$

$$1 + tg^2 2\varphi = \frac{1}{\cos^2 2\varphi} \Rightarrow \cos 2\varphi = \frac{-1}{\sqrt{1 + tg^2 2\varphi}} \quad \text{(2-а чверть)}$$

$$\cos 2\varphi = \frac{-1}{\sqrt{1+\frac{16}{9}}} = -\frac{3}{5}$$

$$1+\cos 2\varphi = 2\cos^2\varphi \Rightarrow \cos\varphi = \sqrt{\frac{1+\cos 2\varphi}{2}} = \frac{1}{\sqrt{5}}$$

$$1-\cos 2\varphi = 2\sin^2\varphi \Rightarrow \sin\varphi = \sqrt{\frac{1-\cos 2\varphi}{2}} = \frac{2}{\sqrt{5}}$$

Перетворення обертання має вигляд

$$\begin{cases} x = \frac{1}{\sqrt{5}}\left(x'-2y'\right) \\ y = \frac{1}{\sqrt{5}}\left(2x'+y'\right) \end{cases}$$

$$\frac{9}{5}\left(x'-2y'\right)^2 - \frac{4}{5}\left(x'-2y'\right)\left(2x'+y'\right) + \frac{6}{5}\left(2x'+y'\right)^2 + \frac{16}{\sqrt{5}}\left(x'-2y'\right) -$$
$$-\frac{8}{\sqrt{5}}\left(2x'+y'\right) - 2 = \frac{9}{5}x'^2 + \frac{36}{5}y'^2 - \frac{36}{5}x'y' - \frac{8}{5}x'^2 - \frac{4}{5}x'y' +$$
$$+\frac{16}{5}x'y' + \frac{8}{5}y'^2 + \frac{24}{5}x'^2 + \frac{24}{5}x'y' + \frac{6}{5}y'^2 + \frac{16}{\sqrt{5}}x' - \frac{32}{\sqrt{5}}y' -$$
$$-\frac{16}{\sqrt{5}}x' - \frac{8}{\sqrt{5}}y' - 2 = 5x'^2 + 10y'^2 - 8\sqrt{5}y' - 2 = 0$$

$$5x'^2 + 10\left(y'^2 - 2\frac{2}{\sqrt{5}}y' + \frac{4}{5}\right) - 8 - 2 = 0$$

$$5x'^2 + 10\left(y' - \frac{2}{\sqrt{5}}\right)^2 = 10$$

$$\begin{cases} x'' = x' \\ y'' = y' - \frac{2}{\sqrt{5}} \end{cases}$$

$$\left(x''\right)^2 + 2\left(y''\right)^2 = 2\text{, або } \frac{\left(x''\right)^2}{2} + \frac{\left(y''\right)^2}{1} = 1$$

Таким чином, ми отримали канонічне рівняння еліпса.

Контрольні запитання:

1. Дати визначення понять еліпс, гіпербола, парабола.
2. Що таке фокус, директриса, ексцентриситет?
3. Які є способи перетворення системи координат?

Література: [5], с. 155-198, 81-84; [6], с. 69-85, 116-124, 141-145.

5. ЕЛЕМЕНТИ ТЕОРІЇ МНОГОЧЛЕНІВ

5.1. Операції над многочленами. Найбільший спільний дільник. Алгоритм Евкліда

М**ногочленом** називатимемо вираз виду:

$$f(x) = a_0 + a_1 x + a_2 x + ... + a_n x^n \tag{5.1}$$

Многочлен виду (5.1) називають **многочленом степеня *n***, якщо коефіцієнт a_n відмінний від нуля (при цьому коефіцієнти з великими номерами можуть дорівнювати нулю):

$$\deg f(x) = n$$

Коефіцієнт a_n називають **старшим коефіцієнтом**. Многочлени зі старшим коефіцієнтом a_n, що дорівнює 1, часто називають **нормалізованими** (або нормованими).

Єдиним многочленом, що не має визначеного степеня, є многочлен, у якого всі коефіцієнти дорівнюють нулеві. Такий многочлен називається **нульовим многочленом**. Многочлени степеня 1, 2, 3 називають відповідно **лінійними**, **квадратичними** (або квадратними) і **кубічними многочленами**.

Два многочлени називають **рівними**, якщо вони мають однаковий степінь, і рівні всі їх коефіцієнти при однакових степенях аргумента.

Нехай задано два многочлени – $f(x)$ степеня *n* та $g(x)$ степеня *s*:

$$f(x) = a_0 + a_1 x + ... + a_n x^n, \ g(x) = b_0 + b_1 x + ... + b_s x^s \qquad (5.2)$$

причому для визначеності припустімо, що $n \geq s$.

Сумою многочленів $f(x)$ та $g(x)$ назвають многочлен

$$f(x) + g(x) = c_0 + c_1 x + ... + c_n x^n,$$

де

$$c_i = a_i + b_{\mathrm{i}} \quad i = \overline{0...s}$$

$$c_i = a_i \qquad i = \overline{s...n}.$$

Степінь суми многочленів дорівнює *n* у випадку, коли $n>s$; при $n=s$ він може бути меншим за *n*, якщо

$$b_i = -a_i \quad i = \overline{\mathrm{k}...n} \qquad 0 \leq k \leq n.$$

Добутком многочленів $f(x)$ та $g(x)$ називають многочлен

$$f(x) \cdot g(x) = d_0 + d_1 x + ... + d_{n+s} x^{n+s},$$

де

$$d_i = \sum_{k+l=i} a_k \cdot b_{\mathrm{l}} \qquad i = \overline{0...n+s},$$

тобто коефіцієнт d_i дорівнює сумі добутків тих коефіцієнтів многочленів $f(x)$ и $g(x)$, сума індексів яких дорівнює i.

$$d_0 = a_0 b_0 \quad d_2 = a_0 b_2 + a_1 b_1 + a_2 b_0 \quad d_{\mathrm{n+s}} = a_n \cdot b_s .$$

З останньої рівності випливає, що $d_{\mathrm{n+s}} \neq 0$, і степінь добутку ненульових многочленів дорівнює сумі степенів співмножників:

$$\deg(f(x) \cdot g(x)) = \deg(f(x)) + \deg(g(x))$$

тобто добуток ненульових многочленів є ненульовий многочлен.

Окремим випадком добутку многочленів є добуток многочлена $f(x)$ на число α, оскільки ненульове число можна розглядати як многочлен нульового степеня. Для многочлена $f(x)$ його добуток на число α матиме вигляд:

$$\alpha \cdot f(x) = \sum_{i=0}^{n} (\alpha \cdot a_i) x^i$$

Теорема (про ділення многочленів)

Для будь-якого многочлена $f(x)$ і ненульового многочлена $g(x)$ знайдуться такі многочлени $q(x)$ і $r(x)$, що

$$f(x) = q(x) \cdot g(x) + r(x), \tag{5.3}$$

де степінь $r(x)$ менший за степінь $g(x)$ (або $r(x)=0$), причому таке представлення єдине.

□

Нехай степінь многочлена $f(x)$ дорівнює n ($\deg f(x)=n$), а многочлена $g(x)$ – s ($\deg g(x)=s$). Очевидно, що коли $n<s$, в розкладанні (5.3) можна покласти $q(x)=0,\ r(x)=f(x)$. Тому розглянемо далі лише випадок, коли $n\ge s$.

Визначимо новий многочлен f_1(x) наступним чином:

$$f_1(x)=f(x)-\frac{a_n}{b_s}x^{n-s}g(x) \tag{5.4}$$

Позначимо n_1 – степінь многочленна f_1, $a_{n_1}^{(1)}$ – його старший коефіцієнт. Очевидно, що виходячи зі способу побудови цього многочленна, $n_1<n$. Якщо $n_1\ge s$, то побудуємо многочлен $f_2(x)$:

$$f_2(x)=f_1(x)-\frac{a_{n_1}^{(1)}}{b_s}x^{n_1-s}g(x) \tag{5.5}$$

Позначимо $n_2=\deg f_2(x)$, $a_{n_2}^{(2)}$ – його старший коефіцієнт. Якщо $n_2\ge s$, то продовжуємо далі аналогічно:

$$f_3(x)=f_2(x)-\frac{a_{n_2}^{(2)}}{b_s}x^{n_2-s}g(x) \tag{5.6}$$

Зрозуміло, що степені многочленів $f_1(x)$, $f_2(x)$, $f_3(x)$, … спадають. Тому після деякої скінченної кількості кроків можна прийти до многочлена

$$f_k(x) = f_{k-1}(x) - \frac{a_{n_{k-1}}^{(k-1)}}{b_s} x^{n_{k-1}-s} g(x), \tag{5.7}$$

причому цей многочлен $f_k(x)$ або буде нульовим, або його степінь n_k буде меншим за s. Тоді, додаючи всі рівності від (5.4) до (5.7), отримаємо:

$$f(x) = (\frac{a_n}{b_s} x^{n-s} + \frac{a_{n_1}^{(1)}}{b_s} x^{n_1-s} + \ldots + \frac{a_{n_{k-1}}^{(k-1)}}{b_s} x^{n_{k-1}-s}) g(x) + f_k(x)$$

Це означає, що многочлени

$$q(x) = \frac{a_n}{b_s} x^{n-s} + \frac{a_{n_1}^{(1)}}{b_s} x^{n_1-s} + \ldots + \frac{a_{n_{k-1}}^{(k-1)}}{b_s} x^{n_{k-1}-s} \quad \text{і} \quad r(x) = f_k(x)$$

задовольняють розкладанню (5.3), причому $r(x)$ – або нульовий многочлен, або його степінь менший за степінь многочлена $g(x)$.

Покажемо тепер, що таке розкладаня єдине.

Нехай існують ще многочлени $\tilde{q}(x)$ і $\tilde{r}(x)$, такі, що

$$f(x) = \tilde{q}(x) \cdot g(x) + \tilde{r}(x) \tag{5.8}$$

причому $\tilde{r}(x)$ або нульовий, або $\deg \tilde{r}(x) < \deg g(x)$.

Віднімаючи (5.8) з (5.3), отримаємо $g(x)\cdot(q(x)-\tilde{q}(x))=r(x)-\tilde{r}(x)$.

Многочлен в правій частині рівності є нульовим, або його степінь менший за степінь $g(x)$ (за способом побудови).

Многочлен у лівій частині рівності при $q(x)-\tilde{q}(x)\neq 0$ має степінь не менший за степінь $g(x)$. Тому рівність можлива лише у тому випадку, коли $q(x)=\tilde{q}(x)$ та $r(x)=\tilde{r}(x)$.

■

Многочлен $f(x)$ називається **діленим**, многочлен $q(x)$ – **часткою** від ділення $f(x)$ на $g(x)$, а многочлен $r(x)$ – **остачею** від ділення $f(x)$ на $g(x)$.

Якщо остача дорівнює нулю, то кажуть, що $f(x)$ **ділиться** на $g(x)$, а многочлен $g(x)$ називають **дільником** многочлена $f(x)$.

Нехай задано два многочлени $f(x)$ і $g(x)$ виду (5.2). Многочлен $d(x)$ називають **найбільшим спільним дільником** (НСД) многочленів $f(x)$ і $g(x)$ (записують $d(x)=НСД(f(x),g(x))$), якщо:

1) многочлен $d(x)$ є дільником многочленів $f(x)$ і $g(x)$;

2) будь-який інший дільник $\tilde{d}(x)$ многочленів $f(x)$ і $g(x)$ є у свою чергу дільником многочлена $d(x)$.

Многочлени $f(x)$ і $g(x)$ називають взаємно простими, якщо $НСД(f(x),g(x))=1$.

Для знаходження найбільшого спільного дільника многочленів $f(x)$ і $g(x)$ використовують **алгоритм Евкліда** (**алгоритм послідовного ділення**). Його ідея полягає в наступному.

Згідно теорії ділення многочленів, має місце розкладання:

$$f(x) = q_1(x) \cdot g(x) + r_1(x)$$

Аналогічно

$$g(x) = q_2(x) \cdot r_1(x) + r_2(x)$$

Застосовуючи далі скінченну кількість разів теорему про ділення многочленів до многочленів $r_i(x)$, отримаємо:

$$\begin{gathered} r_1(x) = q_3(x) \cdot r_2(x) + r_3(x) \\ \ldots \\ r_{k-1}(x) = q_{k+1}(x) \cdot r_k(x) + r_{k+1}(x) \\ r_k(x) = q_{k+2}(x) \cdot r_{k+1}(x) \end{gathered}$$

тобто останнє ділення відбувається без остачі. Тоді остання відмінна від нуля остача $r_{k+1}(x)$ (тобто та, ділення на яку відбулося націло) і буде найбільшим спільним дільником многочленів $f(x)$ та $g(x)$.

Доведемо останнє твердження.

□

Оскільки $r_k(x)$ ділиться на $r_{k+1}(x)$ без остачі, то, підставивши вираз для $r_k(x)$ у вираз для $r_{k-1}(x)$, отримаємо:

$$r_{k-1}(x) = q_{k+1}(x) \cdot q_{k+2}(x) \cdot r_{k+1}(x) + r_{k+1}(x) = (q_{k+1}(x) \cdot q_{k+2}(x) + 1) \cdot r_{k+1}(x)$$

тобто $r_{k-1}(x)$ ділиться на $r_{k+1}(x)$ без остачі.

Продовжуючи аналогічні міркування, отримаємо, що многочлен $r_{k+1}(x)$ є дільником усіх поліномів $r_i(x)$, а також і початкових многочленів $f(x)$ і $g(x)$.

Нехай $\tilde{d}(x)$ – іще один спільний дільник многочленів $f(x)$ і $g(x)$. Але тоді многочлен $r_1(x) = f(x) - \mathrm{q}_1(x) \cdot g(x)$ також ділиться на $\tilde{d}(x)$, многочлен $r_2(x) = g(x) - \mathrm{q}_2(x) \cdot r_1(x)$ ділиться на $\tilde{d}(x)$, і, продовжуючи цей логічний ланцюжок, отримаємо, що многочлен $\tilde{d}(x)$ є дільником і многочлена $r_{k+1}(x)$.

А тоді, згідно з вище викладеним визначенням, многочлен $r_{k+1}(x)$ є найбільшим спільним дільником многочленів $f(x)$ і $g(x)$, що й потрібно було довести.

■

Слід зазначити, що найбільший спільний дільник визначається однозначно з точністю до постійного множника, тобто якщо

$$d(x) = НСД(f(x), g(x)),$$

то

$$\alpha d(x) = НСД(\beta f(x), \gamma g(x)).$$

Це означає, що під час ділення многочлени можна множити на число.

Приклад

Знайти найбільший спільний дільник многочленів

$$f(x) = x^4 + 3x^3 - x^2 - 4x - 3 \text{ і } g(x) = 3x^3 + 10x^2 + 2x - 3.$$

$$q_1(x) = x - \frac{1}{3} \qquad r_1(x) = -\frac{5}{3}x^2 - \frac{25}{3}x - 10$$

$$3f(x) = q_1(x) \cdot g(x) + r_1(x)$$

$$q_2(x) = 3x - 5 \qquad r_2(x) = 9x + 27$$

$$g(x) = -\frac{3}{5}q_2(x) \cdot r_1(x) + r_2(x)$$

$$q_3(x) = x + 2 \qquad r_3(x) = 0$$

$$-\frac{3}{5}r_1(x) = q_3(x) \cdot \frac{1}{9}r_2(x)$$

Тобто, $НСД(f(x), g(x)) = \frac{1}{3}r_2(x) = x + 3$.

Контрольні запитання:

1. Які операції виконуються над многочленами?
2. Яким методом шукають найбільший спільний дільник?

Література: [2], с. 212-214; [3], с. 180-183, 187-188, 192-195.

5.2. Корені многочленів. Теорема Безу. Основна теорема алгебри

Розглянемо многочлен $f(x)$ степеня *n*:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1} + a_n x^n \qquad (5.9)$$

Число x_0 називають **коренем** (або **розв'язком**) многочлена $f(x)$, якщо

$$f(x_0) = 0$$

Розглянемо теорему, яка фактично є наслідком теореми про ділення многочленів.

<u>Теорема (Безу)</u>

Остача від ділення многочлена $f(x)$ на двочлен (лінійний многочлен) x-x_0 дорівнює значенню многочлена $f(x)$ в точці x_0.

□

Розглянемо ділення многочлена $f(x)$ на лінійний многочлен $x - x_0$. Згідно з теорією ділення многочленів, отримаємо:

$$f(x) = (x - x_0) \cdot q(x) + r(x)$$

Оскільки степінь $r(x)$ має бути меншим за степінь многочленна x-x_0, то остача $r(x)$ є многочленом нульового степеня, тобто константою. Для її знаходження підставимо в обидві частини замість x значення x_0. Отримаємо:

$$r(x) = const \equiv r(x_0) = f(x_0).$$

■

<u>Наслідок</u>

Число x_0 є коренем многочлена $f(x)$ тоді і лише тоді, коли лінійний многочлен x-x_0 є дільником многочлена $f(x)$.

На основі теореми Безу та наслідка з неї можна дати наступне, більш загальне, визначення кореня многочлена.

Число x_0 називається **коренем многочлена** $f(x)$ кратності k, якщо многочлен $f(x)$ ділиться на $(x - x_0)^k$ і не ділиться на $(x - x_0)^{k+1}$.

Іншими словами, число x_0 – корінь кратності k многочлена $f(x)$ тоді і лише тоді, коли

$$f(x) = (x - x_0)^k \cdot g(x) \tag{5.10}$$

причому

$$\mathrm{g}(\mathrm{x}_0) \neq 0$$

Корінь, кратність якого дорівнює 1, називається **простим коренем**.

Твердження (про кратні корені)

Якщо число x_0 – корінь кратності k для многочлена $f(x)$, то це саме число x_0 є коренем кратності $(k-1)$ для многочлена $f'(x)$.

□

Нехай x_0 – корінь кратності k многочлена $f(x)$. Тоді має місце представлення цього многочлена у вигляді (5.10). Продиференціювавши, отримаємо:

$$f'(x) = k(x-x_0)^{k-1} \cdot g(x) + (x-x_0)^k \cdot g'(x) =$$

$$= (x-x_0)^{k-1} \cdot (kg(x) + (x-x_0) \cdot g'(x)) =$$

$$= (x-x_0)^{k-1} \cdot \phi(x), \text{ де } \phi(x) = kg(x) + (x-x_0) \cdot g'(x)$$

Оскільки $\phi(x_0) = kg(x_0) \neq 0$ внаслідок того, що $g(x_0) \neq 0$, то, за визначенням, число x_0 є коренем многочлена $f'(x)$ кратності $k-1$.

■

Далі сформулюємо одне з найважливіших тверджень лінійної алгебри – теорему про алгебраїчну замкненість поля комплексних чисел. Вона гарантує існування розв'язку многочлена в комплексних числах. Її важливість випливає з того, що багато задач математики, зрештою, зводяться до обчислення окремих коренів комплексних многочленів, або до якісного опису їх сукупності.

Зокрема, на цій теоремі ґрунтується теорія спектру лінійного оператора (яка розглядається пізніше в курсі лінійної алгебри) та методи знаходження невизначених інтегралів (які розглядаються в курсі математичного аналізу).

За традицією цю теорему називають "основною теоремою алгебри", оскільки в часи, коли вона була сформульовані і доведена (XVIII століття), розв'язування алгебраїчних рівнянь було одним з головних завдань алгебраїстів. У наш час ця теорема є лише одним з найважливіших тверджень.

Теорема (основна теорема алгебри)

Довільний многочлен $f(x)$ степеня $n \geq 1$ з комплексними коефіцієнтами має принаймні один корінь, в загальному випадку комплексний.

Теорему приймемо без доведення.

Наслідок 1

Довільний многочлен $f(x)$ степеня $n \geq 1$ з комплексними коефіцієнтами має рівно n коренів (у загальному випадку комплексних), з урахуванням їхньої кратності.

□

Розглянемо многочлен $f(x)$ степеня n.

Згідно основної теоремою алгебри, він має принаймні один корінь, який позначимо x_1. Тоді за теоремою Безу і наслідком із неї випливає, що:

$$f(x) = (x - x_1) \cdot g_1(x) \qquad \deg g_1(x) = n - 1$$

Аналогічно, застосовуючи до многочлена $g_1(x)$ основну теорему алгебри та теорему Безу, отримаємо $g_1(x) = (x - x_2) \cdot g_2(x)$, $\deg g_2(x) = n - 2$, x_2 – корінь $g_1(x)$.

Продовжуючи аналогічно, отримаємо, що початковий многочлен $f(x)$ можна представити у вигляді:

$$f(x) = a_n(x - x_1) \cdot (x - x_2)...(x - x_n) = a_n \prod_{i=1}^{n} (x - x_i) \tag{5.11}$$

■

Наслідок 2

Довільний многочлен степеня $n \geq 1$ з комплексними коефіцієнтами можна представити у вигляді:

$$f(x) = a_n(x - x_1)^{S_1} \cdot (x - x_2)^{S_2}...(x - x_r)^{S_r} \tag{5.12}$$

де $x_1,...,x_r$ – різні корені многочлена, $S_1,...,S_r$ – кратності коренів, $S_1 + ... + S_r = n$.

□

Очевидно, що розкладання (5.12) випливає з розкладання (5.11) шляхом об'єднання однакових дужок.

■

Розкладання (5.12) називають **канонічним розкладанням** многочлена $f(x)$ на множники. Воно є єдиним з точністю до порядку слідування співмножників.

Наслідок 3

Якщо значення многочленів $f(x)$ і $g(x)$ степеня n $(n \geq 1)$, збігаються принаймні в $(n+1)$ точках, то ці многочлени рівні.

□

Розглянемо новий многочлен $F(x)$, який визначимо наступним чином

$$\mathrm{F}(x) = f(x) - g(x)$$

Очевидно, що степінь многочлена $F(x)$ не більший за n ($\deg F(x) \leq n$).

За умовою, многочлен $\mathrm{F}(x)$ має як мінімум $(n+1)$ коренів, а за основною теоремою алгебри кількість його розв'язків не може бути більшою за n. Цього протиріччя можна уникнути лише тоді, коли многочлени $f(x)$ і $g(x)$ будуть рівними.

■

Зауваженя до наслідка 3

Многочлен $f(x)$ степеня n повністю визначається своїми значеннями при будь-яких $(n+1)$ різних значеннях аргумента. Це дозволяє відновлювати многочлен за його значеннями. Докладніше про це можна дізнатися в курсі чисельних методів (при розгляді інтерполяції).

Наслідок 4

Якщо комплексне число $z_0 = \alpha + i\beta$ є коренем многочлена $f(x)$ з дійсними коефіцієнтами, то коренем многочлена також є число, спряжене до числа z_0:

$$\overline{z}_0 = \alpha - i\beta$$

□

Розглянемо многочлен $f(x) = a_0 + a_1 x + ... + a_n x^n$, де всі a_i – дійсні числа. За умовою наслідку $f(z_0) = 0$. Розглянемо $f(\overline{z}_0)$:

$$\begin{aligned}
& f(\overline{z}_0) = a_0 + a_1\overline{z}_0 + a_2(\overline{z}_0)^2 + ... + a_n(\overline{z}_0)^n = \\
& = \left[\overline{z}_1 \cdot \overline{z}_2 = \overline{z_1 \cdot z_2}\right] = a_0 + a_1\overline{z}_0 + a_2\overline{z_0^2} + ... + a_n\overline{z_0^n} = \\
& = a_0 + \overline{a_1 z_0} + \overline{a_2 z_0^{\ 2}} + ... + \overline{a_n z_0^{\ n}} = \left[\overline{z}_1 + \overline{z}_2 = \overline{z_1 + z_2}\right] = \\
& = \overline{a_0 + a_1 z_0 + a_2 z_0^{\ 2} + ... + a_n z_0^{\ n}} = \overline{f(z_0)} = \overline{0} = 0,
\end{aligned}$$

тобто $\overline{z}_0 = \alpha - i\beta$ – корінь многочлена $f(x)$.

■

Зауваження до наслідку 4

Многочлен $f(x)$ з дійсними коефіцієнтами, що має комплексні корені, ділиться на квадратний многочлен з дійсними коефіцієнтами.

□

Многочлен $f(x)$ ділиться на $(x - z_0)$ та $(x - \overline{z}_0)$, оскільки z_0 та $\overline{z}_0$ – корені $f(x)$, ($z_0 = \alpha + i\beta$). З цього випливає, що $f(x)$ ділиться також і на многочлен $g(x) = (x - z_0) \cdot (x - \overline{z}_0)$, але

$$g(x) = (x - z_0)\cdot(x - \overline{z}_0) = x^2 - (z_0 + \overline{z}_0)\cdot x + z_0 \cdot \overline{z}_0 = \\ = x^2 - 2\alpha\cdot x + (\alpha^2 + \beta^2), \quad \alpha, \beta \in \mathbb{R}.$$

■

Наслідок 5

Будь-який многочлен з дійсними коефіцієнтами представляється єдиним чином з точністю до перестановки множників у вигляді добутку многочленів степеня не вищого за другий з дійсними коефіцієнтами:

$$f(x) = a_0(x - x_1)^{S_1}\ldots(x - x_k)^{S_k} = (x^2 + c_1 + d_1)^{p_1}\ldots(x^2 + c_e + d_e)^{p_e}$$

$$c_i, d_i \in \mathbb{R} \quad s_1 + \ldots + s_k = 2(p_1 + \ldots + p_e) = n = \deg f(x)$$

Контрольні запитання:

1. Що таке корінь многочлена? Що таке кратний корінь многочлена?
2. Який зміст теореми Безу?
3. Які наслідки випливають з основної теореми алгебри?

Література: [2], с. 216-225; [3], с. 208-210, 232-242.

Література:

1. Ильин В.А., Позняк Э.Г. Линейная алгебра. – М.: Физматлит – 2001 г. – 320 с.
2. Воеводин В.В. Линейная алгебра. – М.: Наука – 1980 г. – 400 с.
3. Кострикин А.И. Введение в алгебру. Часть 1. – М.: Физматлит – 2001 г. – 272 с.
4. Стренг Г. Линейная алгебра и ее применения. – М.: Мир – 1980 г. – 446 с.
5. Ильин В.А., Позняк Э.Г. Аналитическая геометрия. – М.: Физматлит – 2001 г. – 224 с.
6. Александров П.С. Курс аналитической геометрии и линейной алгебры. – М.: Наука – 1979 г. – 512 с.