В. С. ЧАРІН

Лінійна АЛГЕБРА

В. С. ЧАРІН

Лінійна алгебра

2-ге видання,
стереотипне

Затверджено Міністерством освіти і науки України як підручник для студентів вищих технічних навчальних закладів

Київ
"Техніка"
2005

ББК 22.143я73
Ч20
УДК 512.64(075.8)

Гриф надано Міністерством освіти і науки України, лист №1/11-2103 від 14.05.2004 р.

Видано за рахунок державних коштів. Продаж заборонено.

Рецензенти:
Петунін Ю. І., д-р фіз.-мат. наук, проф. Київського національного університету імені Тараса Шевченка;
Черніков М. С., д-р фіз.-мат. наук, науков. співроб. відділу алгебри Інституту математики НАН України.

У підручнику викладено об'єкти лінійної алгебри трьох типів: лінійні простори, матриці й алгебричні форми. Вони важливі для розв'язування задач геометрії, обчислювальної математики, фізики, математичної економіки. Розглянуто властивості лінійних рівнянь і лінійних перетворень лінійних просторів. Приділено увагу питанням псевдообернення дійсних матриць, що пов'язані з методом найменших квадратів Гаусса. Наведено ефективні методи обчислення псевдооберненої матриці. При розгляданні опуклих множин використано топологічні властивості підмножин евклідового простору. Досліджено діофантові рівняння.

Для студентів вищих технічних навчальних закладів.

ISBN 996-575-053-4

ПЕРЕДМОВА

У лінійній алгебрі вивчаються, головним чином, об'єкти трьох типів: лінійні простори, матриці та алгебричні форми. Теорії цих об'єктів тісно пов'язані між собою. Методи лінійної алгебри застосовуються в теорії кілець і модулів, у теорії зображення груп, у функціональному аналізі. Всі об'єкти, що розглядаються і вивчаються в лінійній алгебрі, важливі для задач геометрії, обчислювальної математики, фізики, математичних методів в економіці тощо.

Теорія алгебричних форм складається з теорії лінійних, квадратичних форм і теорії форм вищого степеня. Алгебра матриць особливо зручна для вивчення лінійних і квадратичних форм, систем лінійних рівнянь та лінійних нерівностей. Об'єднуючим об'єктом усіх частин лінійної алгебри є лінійний (векторний) простір.

Запропонований підручник написано на підставі багаторічного досвіду викладання курсу алгебри в Уральському університеті та на механіко-математичному факультеті й факультеті кібернетики Київського національного університету імені Тараса Шевченка.

Перші два розділи "Поля і многочлени" та "Теорія визначників" є підготовчими для загальної теорії лінійних просторів скінченної розмірності.

У розділі 1 наведено елементи теорії полів і кілець. Доводиться основна теорема вищої алгебри. Заключний параграф розділу присвячено віддільності дійсних коренів многочленів. Ця тема починається з означення звичайного й узагальненого ряду Штурма. Доводяться відповідні теореми Штурма. Розглядаються властивості коренів одного класу многочленів, заданих зворотними співвідношеннями (до цього класу належать многочлени Лежандра, Ерміта та ін.). Теорія визначників викладена так, щоб навички, набуті при їх обчисленні, підготували студентів до сприйняття поняття багатовимірного вектора.

У розділі 2 наведено поняття визначника n-го порядку. Доводиться теорема Крамера для системи n рівнянь з n невідомими.

У розділах 3–7 розглянуто властивості лінійних просторів, систем лінійних рівнянь і лінійних перетворень. Наведено правила й алгоритми обчислення базису та рангу системи векторів, рангу матриці, розв'язання систем лінійних рівнянь, обчислення обернених матриць. При цьому основним засобом розв'язання більшості задач обрано метод елементарних перетворень над векторами. Розглянуто методи обчислення матриць лінійних операторів та їх характеристичних многочленів, а також наведено будову деяких найбільш важливих типів лінійних операторів, зокрема, самоспряжених і ортогональних операторів евклідового простору.

Розділ 8 присвячено питанням псевдоперетворення дійсних матриць. Задача про обчислення псевдооберненої матриці безпосередньо пов'язана з методом найменших квадратів Гаусса. Наведено ефективний метод обчислення псевдооберненої матриці на підставі відомого алгоритму Д. К. Фаддєєва.

У розділі 9 про опуклі множини широко використовуються топологічні властивості підмножин евклідового простору. Доводяться і потім застосовуються теореми віддільності для опуклих множин. Описується будова опуклих багатогранників, а також конусів і опуклих багатогранних множин (тобто множин усіх розв'язків сумісної системи лінійних нерівностей). Матеріал цього розділу є теоретичною основою курсу лінійного і нелінійного програмування, а також спеціального курсу "Лінійні нерівності".

Останній розділ 10 про системи лінійних діофантових рівнянь написаний професором Київського національного університету імені Тараса Шевченка І. В. Протасовим.

У підручнику наведено багато прикладів розв'язання конкретних задач. Окрім того, кожний параграф містить вправи для самостійної роботи.

Автор висловлює щиру подяку кандидату фізико-математичних наук С. С. Шестакову за участь у підготовці підручника до друку.

Розділ 1. **ПОЛЯ І МНОГОЧЛЕНИ**

§ 1.1. Кільця і поля

Для дійсних і раціональних чисел визначено чотири арифметичні операції (дії) – додавання, віднімання, множення і ділення. Результатом цих дій одержуються числа такого ж виду. Для цілих чисел лише перші три арифметичні операції також породжують числа цілі.

Операції, аналогічні арифметичним діям з числами, відповідним чином вводяться і для інших об'єктів. При цьому з'ясовується, що вони мають ті самі властивості, що й дії над числами. Багатовиди таких споріднених об'єктів визначаються аксіоматично за допомогою означень, наведених нижче.

Означення. *Позначимо через K деяку непорожню множину, що складається з елементів* $a, b, c, ..., x, y, z, ...$ *Нехай для кожної пари елементів* a, b *визначено дві операції:*

1. Додавання. Внаслідок застосування цієї операції елементам a і b ставиться у відповідність третій елемент $a + b$, *який належить тій же множині K і називається сумою a і b.*

2. Множення. Цією операцією однозначно ставиться у відповідність елементам a і b третій елемент $a \cdot b$, *який належить тій же множині K і називається добутком a і b.*

Непорожня множина K називається кільцем, якщо ці операції підпорядковані таким основним законам:

I. Закони додавання

I_1. *Асоціативність:*

$$\forall a, b, c \in K \left\{ a + (b + c) = (a + b) + c \right\}.$$

I_2. *Комутативність:*

$$\forall a, b \in K \left\{ a + b = b + a \right\}.$$

I_3. Існування нульового елемента 0:

$$\exists 0 \in K\ \forall a \in K\ \{a + 0 = 0 + a = a\}.$$

I_4. Існування протилежного елемента x, який позначається $-a$:

$$\forall a \in K\ \exists x \in K\ \{a + x = x + a = \theta\}.$$

II. Закони множення

II_1. Асоціативність:

$$\forall a,b,c \in K\ \{a \cdot (b \cdot c) = (a \cdot b) \cdot c\}.$$

III. Розподільні закони

III_1. $\forall a,b,c \in K\{a \cdot (b + c) = a \cdot b + a \cdot c\}.$

III_2. $\forall a,b,c \in K\{(a + b) \cdot c = a \cdot c + b \cdot c\}.$

Кільце K називається комутативним, якщо виконується додатково до основних законів ще один закон – комутативність множення:

II_2. $\forall a,b\ \{a \cdot b = b \cdot a\}.$

Приклади кілець (комутативних):

множина Z всіх цілих чисел відносно дій додавання і множення;

множина всіх цілих чисел, кратних заданому цілому числу m;

множина Q всіх раціональних чисел;

множина всіх раціональних чисел виду $\frac{m}{2^p}$, де m і p набувають відповідно всіх цілих і всіх натуральних чисел.

Доведемо деякі загальні властивості кілець, що випливають з основних законів.

1. У кільці K існує єдиний нульовий елемент 0.

Доведення. Нехай 0 і 0 – його нульові елементи. Тоді $a + \theta = a = \theta + a$ і $a + 0 = a = 0 + a$ для будь-якого $a \in K$. Якщо $a = 0$, тоді $\theta + \theta = 0$ і $0 + \theta = \theta$. Тому $\theta = 0$.

2. Для будь-якого елемента $a \in K$ існує єдиний обернений елемент $x = -a$.

Доведення. Маємо $a+x=0$. Нехай для деякого $y\in K$ має місце $y+a=0$. Тоді $y+(a+x)=y+0=y$. На підставі закону асоціативності (II_1) $y+(a+x)=(y+a)+x=0+x=x$. Тому $y=x$.

3. Нехай $a_i \in K$, $i = 1, 2, \ldots, n$ і $b_j \in K$, $j = 1, 2, \ldots, m$. За розподільним законом маємо

$$(a_1+a_2+\ldots+a_n)\cdot(b_1+b_2+\ldots+b_m)=$$

$$=a_1\cdot\left(\sum_{j=1}^{m} b_j\right)+a_2\cdot\left(\sum_{j=1}^{m} b_j\right)+\ldots+a_n\cdot\left(\sum_{j=1}^{m} b_j\right)=$$

$$=\sum_{j=1}^{m} a_1 b_j+\sum_{j=1}^{m} a_2 b_j+\ldots+\sum_{j=1}^{m} a_n b_j=a_1\cdot b_1+a_1\cdot b_2+\ldots+a_i\cdot b_j+\ldots$$

Отже, щоб помножити суму $a_1+a_2+\ldots+a_n$ на суму $b_1+b_2+\ldots+b_m$, потрібно кожний доданок першої суми помножити на кожний доданок другої суми і одержані добутки додати. Порядок множників a_i і b_j зберігається.

4. У кільці K визначається дія *віднімання*: для будь-яких його елементів a і b визначається різниця $b-a$ таким чином.

Якщо $x=-a$ – обернений елемент для a, тоді

$$b-a=b+(-a).$$

У класі комутативних кілець окремо виділимо поля.

Означення. *Полем називається комутативне кільце P, для якого додатково виконуються закони множення.*

II_3. Існування одиничного елемента e:

$$\exists e \forall a\{a\cdot e=a\}.$$

II_4. Існування оберненого елемента x, який позначатиметься a^{-1}, для $a\neq 0$:

$$\forall a\neq 0\, \exists x\{a\cdot x=e\}.$$

У будь-якому полі P одиничний елемент e і обернений елемент a^{-1} для ненульового елемента a – єдині. Доведення цих тверджень аналогічне доведенням єдиності нульового елемента 0 і оберненого $-a$ відносно дії додавання.

У полі P визначається операція ділення.

Якщо $a, b \in P$ і $a \neq 0$, то часткою $\frac{b}{a}$ називається елемент $c = a^{-1}b = ba^{-1}$ цього поля.

Упорядковані поля і поле дійсних чисел

Означення. *Поле P називається впорядкованим, якщо для його елементів визначено властивість бути додатним і яке характеризується такими законами:*

IV. Закони додатності

IV_1. Для кожного елемента $a \in P$ виконується лише одне із співвідношень: $a = 0$, $a > 0$, $-a > 0$.

IV_2. $a > 0 \ \& \ b > 0 \Rightarrow a + b > 0$.

IV_3. $a > 0 \ \& \ b > 0 \Rightarrow a \cdot b > 0, a^{-1} > 0.$

Елемент a впорядкованого поля P називається від'ємним, якщо $-a > 0$.

У впорядкованому полі P для кожної пари різних елементів a і b можна ввести відношення наступності.

1. Якщо $a, b \in P$, то для них виконується одне із співвідношень, взаємовиключаючих одне одного:

а) $b - a = 0$. Тоді $b = a$;

б) $b - a > 0$. Тоді записуємо $b > a$ (b наступне за a);

в) $-(b - a) > 0$, тобто $a - b > 0.$ Тоді записуємо $a > b$ (a наступне за b).

Отже, для будь-якої пари елементів a, b виконується одне і лише одне з трьох співвідношень:

$$a = b, \ b > a, \ a > b.$$

2. Якщо $a > b$, то $a + c > b + c$ для будь-якого $c \in P$. Це випливає із співвідношень

$$(a + c) - (b + c) = a - b > 0.$$

3. Якщо $a>b$ і $c>0$, тоді $a\cdot c>b\cdot c$. Отже, $a>b\Rightarrow$ $\Rightarrow a-b>0\Rightarrow (a-b)\cdot c>0\Rightarrow ac-bc>0$, тобто $ac>bc$.

4. Якщо $a>b$ і $a>0$, $b>0$, тоді з $a\cdot b\left(b^{-1}-a^{-1}\right)=$ $=a-b>0$ випливає $b^{-1}-a^{-1}>0$ і тому $b^{-1}>a^{-1}$.

Множина R усіх дійсних чисел із загальноприйнятими арифметичними діями над числами і відношенням порядку – важливий приклад упорядкованого поля.

Сучасний виклад теорії дійсних чисел, як правило, починається з аксіоматичної побудови теорії натуральних чисел. Відомий італійський математик Пеано (1891 р.) запропонував побудувати цю теорію, виходячи з такої системи аксіом:

1. 0 – число, яке не є наступним за будь-яким числом.

2. Для кожного числа a є рівно одне число, наступне за a. Воно позначається a'.

3. Із $a'=b'$ випливає, що $a=b$.

4. Якщо число 0 має деяку властивість A і для будь-якого числа a на підставі того, що a теж має властивість A, випливає, що й число a' має властивість A, то і будь-яке число x повинно мати таку саму властивість A.

Остання аксіома називається *принципом математичної індукції*. Мовою математичної логіки вона записується так:

$$A(0)\ \&\ \forall a\{A(a)\Rightarrow A(a')\}\Rightarrow \forall x\ A(x).$$

За допомогою цих аксіом вводяться операції додавання і множення, а потім вивчаються їх властивості та розвивається елементарна теорія чисел.

Поле R усіх дійсних чисел виділяється в класі впорядкованих полів однією фундаментальною властивістю (властивістю неперервності):

V. Принцип Дедекінда

Нехай задано будь-яке розбиття всіх дійсних чисел на два класи E_1 і E_2, що мають такі властивості:

обидва класи непорожні і не перетинаються;

кожне число класу E_2 більше кожного числа класу E_1.

Тоді існує лише одне дійсне число z таке, що кожне число x, менше за z, належить класу E_1, а кожне число y, більше за z, належить класу E_2. При цьому зрозуміло, що число z належить або E_1, або E_2.

Будь-яке поле P, що є підмножиною поля R дійсних чисел, називається дійсним полем. Інакше кажучи, *дійсним полем називається будь-яка підмножина P множини R усіх дійсних чисел, якщо вона має такі властивості:*

$$a, b \in P \Rightarrow a + b \in P,$$

$$a, b \in P \Rightarrow a - b \in P,$$

$$a, b \in P \Rightarrow a \cdot b \in P,$$

$$a, b \in R \;\&\; a \neq 0 \Rightarrow \frac{b}{a} \in P.$$

Вправи

1. Довести, що коли дійсне поле P нетривіальне (тобто містить у крайньому випадку одне число $a \neq 0$), то воно містить поле Q раціональних чисел.

2. Позначимо через $Q\left[\sqrt{2}\right]$ множину всіх чисел виду $x = a + b\sqrt{2}$, де $a, b \in Q$. Довести, що $Q\left[\sqrt{2}\right]$ – поле.

§ 1.2. Поле комплексних чисел

Комплексні числа

Оберемо на площині прямокутну систему координат з початком у точці O та з одиничними координатними векторами $\vec{1} = (1, 0)$ та $\vec{i} = (0, 1)$, спрямованими відповідно вздовж осей Ox і Oy (рис. 1).

Будь-який вектор $\vec{z}$, що міститься в площині Oxy, можна подати у вигляді суми

$$\vec{z} = a \cdot \vec{1} + b \cdot \vec{i},$$

де a і b – проекції вектора відповідно на вісь Ox і Oy.

Як і у випадку операцій з дійсними числами, для векторів, що належать одній площині, введемо операції додавання, віднімання, множення і ділення, а також операції додавання, віднімання та множення на скаляр.

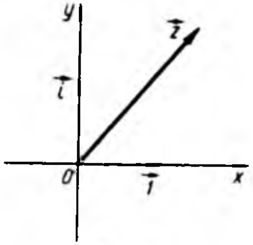


Рис. 1

Операції додавання, віднімання і множення на скаляр. Нехай $\vec{z} = a\cdot\vec{1} + b\cdot\vec{i}$, $\vec{z}' = a'\cdot\vec{1} + b'\cdot\vec{i}$, де a, b, a', b' – довільні дійсні числа (елементи поля R). Тоді

$$\vec{z} + \vec{z}' = (a + a')\vec{1} + (b + b')\vec{i}\ ,$$

$$\vec{z} - \vec{z}' = (a - a')\vec{1} + (b - b')\vec{i}\ ,$$

$$\vec{z}\cdot c = c\cdot\vec{z} = (c\cdot a)\vec{1} + (c\cdot b)\vec{i}$$

для $c \in R$.

Операції множення. Визначимо спочатку правило множення базисних векторів $\vec{1}$ та $\vec{i}$, поклавши $\vec{1}\cdot\vec{1} = \vec{1}$, $\vec{1}\cdot\vec{i} = \vec{i}\cdot\vec{1} = \vec{i}$, $\vec{i}\cdot\vec{i} = -\vec{1}$. Або у вигляді таблиці

	$\vec{1}$	$\vec{i}$
$\vec{1}$	$\vec{1}$	$\vec{i}$
$\vec{i}$	$\vec{i}$	$-\vec{1}$

Добуток $\vec{z}\cdot\vec{z}'$ обчислимо за звичайним правилом множення двочленів $a\cdot\vec{1} + b\cdot\vec{i}$ й $a'\cdot\vec{1} + b'\cdot\vec{i}$ з урахуванням таблиці множення базисних елементів:

$$\vec{z}\cdot\vec{z}' = (a\cdot\vec{1} + b\cdot\vec{i}\,)\cdot(a'\cdot\vec{1} + b'\cdot\vec{i}\,) =$$
$$= aa'\cdot(\vec{1}\cdot\vec{1}\,) + a\cdot b'\cdot(\vec{1}\cdot\vec{i}\,) + b\cdot a'(\vec{i}\cdot\vec{1}) + b\cdot b'(\vec{i}\cdot\vec{i}) =$$
$$= a\cdot a'\cdot\vec{1} + a\cdot b'\cdot\vec{i} + b\cdot a'\cdot\vec{i} + b\cdot b'(-\vec{1}) =$$
$$= (a\cdot a' - b\cdot b')\cdot\vec{1} + (a\cdot b' + b\cdot a')\cdot\vec{i}\ .$$

Отже, дістали вектор $\vec{p} = \vec{z}\cdot\vec{z}'$ з проекціями $a\cdot a' - b\cdot b'$ і $a\cdot b' + b\cdot a'$ на осі координат.

Теорема 1. Якщо на множині C всіх двовимірних векторів увести операції додавання і множення для $\vec{z} = a \cdot \vec{1} + b \cdot \vec{i}$ та $\vec{z}' = a' \cdot \vec{1} + b' \cdot \vec{i}$ за правилами

$$\vec{z} + \vec{z}' = (a + a')\vec{1} + (b + b')\vec{i},$$

$$\vec{z} \cdot \vec{z}' = (a \cdot a' - b \cdot b') \cdot \vec{1} + (a \cdot b' + b \cdot a') \cdot \vec{i},$$

то C – поле.

Доведення. Перевіримо основні закони I, II і III, що характеризують поле.

Виконання законів $I_1 - I_3$ очевидне. При цьому роль нульового елемента відіграватиме нульовий вектор $\vec{0} = 0 \cdot \vec{1} + 0 \cdot \vec{i}$.

Оберненим для $\vec{z} = a\vec{1} + b\vec{i}$ є $-\vec{z} = (-a) \cdot \vec{1} + (-b) \cdot \vec{i}$.

Закон II_2 очевидний, а роль одиничного елемента (закон II_3) виконує вектор $\vec{1} = 1 \cdot \vec{i} + 0 \cdot \vec{i}$.

Перевіримо асоціативність множення (закон II_1).

Нехай $\vec{z} = a \cdot \vec{1} + b \cdot \vec{i}$, $\vec{z}' = a' \cdot \vec{1} + b' \cdot \vec{i}$, $\vec{z}'' = a'' \cdot \vec{1} + b'' \cdot \vec{i}$. Тоді

$$\vec{z} \cdot \vec{z}' = (a \cdot a' - b \cdot b') \cdot \vec{1} + (a \cdot b' + a' \cdot b) \cdot \vec{i},$$

$$(\vec{z} \cdot \vec{z}') \cdot \vec{z}'' = [(a \cdot a' - b \cdot b') \cdot a'' - (a \cdot b' + a' \cdot b) \cdot b''] \cdot \vec{1} +$$
$$+ [(a \cdot a' - b \cdot b') \cdot b'' + (a \cdot b' + a' \cdot b) \cdot a''] \cdot \vec{i} =$$
$$= (a \cdot a' \cdot a'' - b \cdot b' \cdot a'' - a \cdot b' \cdot b'' - a' \cdot b \cdot b'') \cdot \vec{1} +$$
$$+ (a \cdot a' \cdot b'' - b \cdot b' \cdot b'' + a \cdot b' \cdot a'' + b \cdot a' \cdot a'') \cdot \vec{i}.$$

Далі

$$\vec{z}' \cdot \vec{z}'' = (a' \cdot a'' - b' \cdot b'') \cdot \vec{1} + (a' \cdot b'' + b' \cdot a'') \cdot \vec{i}$$

та

$$\vec{z} \cdot (\vec{z}' \cdot \vec{z}'') = (a \cdot \vec{1} + b \cdot \vec{i}) \cdot [(a' \cdot a'' - b' \cdot b'') \cdot \vec{1} + (a' \cdot b'' + b' \cdot a'') \cdot \vec{i}] =$$
$$= [a \cdot (a' \cdot a'' - b' \cdot b'') - b \cdot (a' \cdot b'' + b' \cdot a'')] \cdot \vec{1} +$$
$$+ [a \cdot (a' \cdot b'' + b' \cdot a'') + (a' \cdot a'' - b' \cdot b'') \cdot b] \cdot \vec{i} =$$
$$= (aa'a'' - ab'b'' - ba'b'' - bb'a'') \cdot \vec{1} +$$
$$+ (aa'b'' + ab'a'' + ba'a'' - bb'b'') \cdot \vec{i}.$$

Порівнюючи два останні результати, дістанемо

$$(\vec{z}\cdot\vec{z}')\cdot\vec{z}''=\vec{z}\cdot(\vec{z}'\cdot\vec{z}'').$$

Перевіримо закон II_4.

Нехай $\vec{z}=a\cdot\vec{1}+b\cdot\vec{i}\neq\vec{0}$, тобто $a^2+b^2>0$. Шукаємо вектор $\vec{w}=x\cdot\vec{1}+y\cdot\vec{i}$ такий, щоб $\vec{z}\cdot\vec{w}=\vec{1}$. Маємо

$$\vec{z}\cdot\vec{w}=(a\cdot x-b\cdot y)\cdot\vec{1}+(a\cdot y+b\cdot x)\cdot\vec{i}=1\cdot\vec{1}+0\cdot\vec{i}.$$

Звідси дістанемо рівняння $a\cdot x-b\cdot y=1$, $a\cdot y+b\cdot x=0$, і розв'язок

$$x=\frac{a}{a^2+b^2},\ y=-\frac{b}{a^2+b^2}.$$

Отже, оберненим є вектор

$$\vec{w}=\vec{z}^{-1}=\frac{a}{a^2+b^2}\,\vec{1}-\frac{b}{a^2+b^2}\,\vec{i}.$$

Перевірку розподільних законів залишимо у вигляді вправи.

Теорему доведено.

У побудованому полі C визначається операція *ділення* вектора $\vec{z}'=a'\vec{1}+b'\vec{i}$ на вектор $\vec{z}=a\vec{1}+b\vec{i}\neq\vec{0}$, тобто

$$\frac{\vec{z}'}{\vec{z}}=\vec{z}'\cdot(\vec{z})^{-1}=(a'\vec{1}+b'\vec{i})\cdot\left(\frac{a}{a^2+b^2}\,\vec{1}-\frac{b}{a^2+b^2}\,\vec{i}\right)=$$

$$=\left(a'\cdot\frac{a}{a^2+b^2}-b'\cdot\frac{-b}{a^2+b^2}\right)\cdot\vec{1}+\left(a'\cdot\frac{-b}{a^2+b^2}+b'\cdot\frac{a}{a^2+b^2}\right)\cdot\vec{i}=$$

$$=\frac{a'a+b'b}{a^2+b^2}\,\vec{1}+\frac{b'a-a'b}{a^2+b^2}\,\vec{i}.$$

Елементи поля C називаються *комплексними числами.*

Позначимо через $\overline{R}$ множину всіх комплексних чисел виду $\vec{a}=a\cdot\vec{1}+0\cdot\vec{i}$, де a набуває всіх значень поля R дійсних чисел. Дійсне число a задає комплексне число $\vec{a}=a\cdot\vec{1}$.

Ця підмножина $\overline{R}$ з поля C відносно введених операцій також є полем, тобто $\overline{R}$ – підполе поля C. Можна перевірити, що операція над числами $a\cdot\vec{1}$ точно відповідає арифметичним діям над дійсними числами a, що їх задають. Отже, поля

$\bar{R}$ і R, відрізняючись за формою, однакові по суті. На цій підставі можна вважати поле C розширенням поля R дійсних чисел.

Звідси напрошується ще один висновок. Кожне комплексне число можна записати спрощено у вигляді

$$z = a + bi,$$

де опущено знак вектора ($\rightarrow$), який раніше позначав належність до двовимірних векторів. Доданок a називається дійсною частиною числа z, а bi – його уявною частиною.

Відповідні алгебричні операції над ними зберігаються з урахуванням спрощеного запису. Так, наприклад, таблиця множення базисних чисел має вигляд

	1	i
1	1	i
i	i	-1

Вправа

Уведемо над двовимірними векторами $\vec{z} = a \cdot \vec{1} + b \cdot \vec{i}$, де $\vec{1} = (1,0)$ та $\vec{i} = (0,1)$, операції додавання, аналогічно до того, як це було введено для комплексних чисел, але з новою таблицею множення для базисних векторів $\vec{1}$ та $\vec{i}$:

$$\vec{1} \cdot \vec{1} = \vec{1}\ ,\ \vec{1} \cdot \vec{i} = \vec{i} \cdot \vec{1} = \vec{i}\ ,\ \vec{i} \cdot \vec{i} = -\vec{1}\ .$$

Дістаємо множину K комплексних чисел нової природи.

Чи буде множина K кільцем? полем?

Спряжені комплексні числа

Як і домовились, будемо записувати комплексні числа у спрощеній формі $z = a + bi$, де a, b – дійсні числа. Число i називається уявною одиницею і для неї $i^2 = -1$.

Число $\bar{z} = a - bi$ називається *спряженим* по відношенню до z. Геометрично воно зображується у вигляді вектора, одержаного з вектора z дзеркальним відображенням відносно осі Ox.

Наведемо основні властивості спряжених комплексних чисел. Усі вони легко визначаються.

1. $\bar{\bar{z}} = z$. 2. $\bar{z} \cdot z = a^2 + b^2$.

3. $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$. 4. $\overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2$.

5. $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$. 6. $\overline{\left(\frac{z'}{z}\right)} = \frac{\bar{z}'}{\bar{z}}$.

При обчисленні частки від ділення числа $z' = a' + b'i$ на число $z = a + bi \neq 0$ можна користуватись таким правилом.

Чисельник і знаменник дробу $\frac{z'}{z}$ множаться на спряжене число $\bar{z}$. Тоді обчислення здійснюються так:

$$\frac{z'}{z} = \frac{z' \cdot \bar{z}}{z \cdot \bar{z}} = \frac{(a' + b'i)(a - bi)}{(a + bi)(a - bi)} =$$

$$= \frac{(a'a + b'b) + (b'a - a'b)\,i}{a^2 + b^2} =$$

$$= \frac{a'a + b'b}{a^2 + b^2} + \frac{b'a - a'b}{a^2 + b^2}\,i.$$

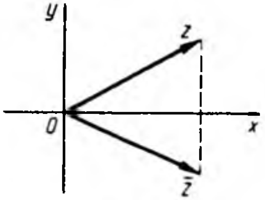


Рис. 2

Це збігається з результатом ділення на підставі загального правила.

Теорема 2. *Якщо внаслідок скінченного числа операцій додавання, віднімання, множення та ділення над комплексними числами* $z_1, z_2, \ldots, z_n$ *одержується число* w, *тоді внаслідок тих самих операцій і у тому ж порядку над спряженими числами* $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_n$ *одержується число* $\bar{w}$, *спряжене з* w.

Дії над комплексними числами в тригонометричній формі

Зобразимо комплексне число $z=a+bi$ у вигляді вектора з проекціями a і b відповідно на осі Ox і Oy. Довжина цього вектора називається *модулем* числа z і позначається через $|z|$.

Безпосередньо з рис. 3 випливає, що

$$|z| = \sqrt{a^2 + b^2}.$$

Кут φ між додатним напрямком осі Ox і вектором z називається *аргументом* числа z і позначається $\arg z$. Аргумент числа визначається неоднозначно, лише з точністю до доданків виду $2\pi k$, $k=0, \pm 1, \dots$

На підставі рис. 3 матимемо

$$\cos\varphi = \frac{a}{|z|}, \ \sin\varphi = \frac{b}{|z|}.$$

Наприклад, якщо $z = 2 - 2i$, то $|z| = \sqrt{2^2 + (-2)^2} = 2\sqrt{2}$ і $\cos\varphi = \frac{\sqrt{2}}{2}$, $\sin\varphi = -\frac{\sqrt{2}}{2}$. Кутом φ буде $-\frac{\pi}{4}$. Або у загальному вигляді $\arg z = -\frac{\pi}{4} + 2\pi k$.

Для числа $z = a + bi$ з модулем $|z|$ і аргументом φ отримаємо

$$a = |z|\cos\varphi, \quad b = |z|\sin\varphi.$$

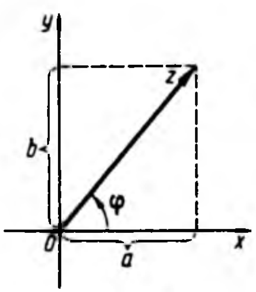


Рис. 3

Тому

$$z = |z| \cdot (\cos\varphi + i\sin\varphi).$$

Позначивши $|z| = r$, матимемо *тригонометричну форму* числа z:

$$z = r \cdot (\cos\varphi + i\sin\varphi).$$

У наведеному вище прикладі для числа $z = 2 - 2i$ дістанемо

$$z = 2\sqrt{2} \cdot \left(\cos\left(-\frac{\pi}{4}\right) + i\sin\left(-\frac{\pi}{4}\right)\right).$$

Множення і піднесення до степеня комплексного числа

Нехай комплексні числа z і z' подано у тригонометричній формі

$$z = r \cdot (\cos\varphi + i\sin\varphi) \text{ і } z' = r' \cdot (\cos\varphi' + i\sin\varphi').$$

Тоді

$$z \cdot z' = r\left(\cos\varphi + i\sin\varphi\right) \cdot r'\left(\cos\varphi' + i\sin\varphi'\right) =$$

$$= rr'\left(\cos\varphi + i\sin\varphi\right) \cdot \left(\cos\varphi' + i\sin\varphi'\right) =$$

$$= r \cdot r'\left[\left(\cos\varphi \cdot \cos\varphi' - \sin\varphi \cdot \sin\varphi'\right) + i\left(\cos\varphi \cdot \sin\varphi' + \sin\varphi \cdot \cos\varphi'\right)\right] =$$

$$= r \cdot r'\left[\cos\left(\varphi + \varphi'\right) + i\sin\left(\varphi + \varphi'\right)\right].$$

Звідси знаходимо, що $r \cdot r' = |z| \cdot |z'|$ – модуль числа $z \cdot z'$, тобто $|z \cdot z'| = |z| \cdot |z'|$, а $\varphi + \varphi'$ – аргумент одержаного числа, тобто $\arg(z \cdot z') = \arg z + \arg z'$.

Правило 1. При множенні комплексних чисел їх модулі перемножуються, а аргументи підсумовуються.

Правило вірне для будь-якої кількості множників: якщо $z_k = r_k(\cos\varphi_k + i\sin\varphi_k)$, $k = \overline{1, m}$, де $r_k = |z_k|$, $\varphi_k = \arg z_k$, тоді

$$z_1 \cdot z_2 \cdots z_m = r_1 \cdot r_2 \cdots r_m [\cos(\varphi_1 + \varphi_2 + \ldots + \varphi_m) +$$
$$+ i\sin(\varphi_1 + \varphi_2 + \ldots + \varphi_m)].$$

Зокрема, при піднесенні до цілого степеня числа $z = r(\cos\varphi + i\sin\varphi)$ дістаємо формулу, яка носить назву *формули Муавра.*

Правило 2 (формула Муавра). При піднесенні до цілого степеня n комплексного числа модуль його підноситься до степеня n, а аргумент множиться на показник степеня n:

$$\left[r\left(\cos\varphi + i\sin\varphi\right)\right]^n = r^n(\cos n\varphi + i\sin n\varphi).$$

Приклад 1. Обчислити $(2 - 2i)^{100}$.

Оскільки $z = 2 - 2i = 2\sqrt{2}\left(\cos\left(-\frac{\pi}{4}\right) + i\sin\left(-\frac{\pi}{4}\right)\right)$, то

$$z^{100} = \left(2\sqrt{2}\right)^{100}\left[\cos\left(-\frac{\pi}{4} \cdot 100\right) + i\sin\left(-\frac{\pi}{4} \cdot 100\right)\right] =$$

$$= 2^{150}[\cos(-25\pi) + i\sin(-25\pi)] = 2^{150}(\cos(-\pi)) = -2^{150}.$$

При $r = 1$ формула Муавра набуває вигляду

$$(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi.$$

Ця формула стає у пригоді при доведенні багатьох тригонометричних тотожностей.

Приклад 2. Подати $\cos 5x$ і $\sin 5x$ через $\cos x$ і $\sin x$.

Застосуємо формулу Муавра до $n=5$. Дістанемо

$$(\cos\varphi + i\sin\varphi)^5 = \cos 5\varphi + i\sin 5\varphi.$$

З другого боку, до лівої частини останньої рівності застосуємо біном Ньютона:

$$(\cos\varphi + i\sin\varphi)^5 = \cos^5\varphi + C_5^1\cos^4\varphi(i\sin\varphi) + C_5^2\cos^3\varphi(i\sin\varphi)^2 +$$
$$+ C_5^3\cos^2\varphi(i\sin\varphi)^3 + C_5^4\cos\varphi(i\sin\varphi)^4 + (i\sin\varphi)^5 =$$
$$= (\cos^5\varphi - 10\cos^3\varphi\sin^2\varphi + 5\cos\varphi\sin^4\varphi) +$$
$$+ i(5\cos^4\varphi\sin\varphi - 10\cos^2\varphi\sin^3\varphi + \sin^5\varphi).$$

Порівнюючи дійсні та уявні частини двох рівностей для $(\cos\varphi + i\sin\varphi)^5$, остаточно доводимо

$$\cos 5\varphi = \cos^5\varphi - 10\cos^3\varphi\sin^2\varphi + 5\cos\varphi\sin^4\varphi,$$
$$\sin 5\varphi = 5\cos^4\varphi\sin\varphi - 10\cos^2\varphi\sin^3\varphi + \sin^5\varphi.$$

Вправи

1. Використовуючи формулу Муавра, довести, що $\cos n\varphi$ можна подати у вигляді

$$\cos n\varphi = a_0\cos^n\varphi + a_1\cos^{n-2}\varphi + a_2\cos^{n-4}\varphi + \ldots,$$

причому $a_0 = 2^{n-1}$, a_1, $a_2, \ldots$ – цілі числа.

2. Нехай

$$P_n = 1 + \cos\varphi + \cos 2\varphi + \ldots + \cos n\varphi,$$
$$Q_n = \sin\varphi + \sin 2\varphi + \ldots + \sin n\varphi.$$

Довести, що

$$P_n = \frac{\sin\frac{n+1}{2}\varphi\cdot\cos\frac{n}{2}\varphi}{\sin\frac{1}{2}\varphi}, \quad Q_n = \frac{\sin\frac{n+1}{2}\varphi\cdot\sin\frac{n}{2}\varphi}{\sin\frac{1}{2}\varphi}.$$

Вказівка. Скласти комплексну суму

$$S_n = P_n + iQ_n = 1 + (\cos\varphi + i\sin\varphi) +$$
$$+ (\cos 2\varphi + i\sin 2\varphi) + \ldots + (\cos n\varphi + i\sin n\varphi).$$

Якщо $z = \cos\varphi + i\sin\varphi$, то за допомогою формули Муавра дістанемо

$$S_n = 1 + z + z^2 + \ldots + z^n.$$

Далі потрібно обчислити останню суму як суму членів геометричної прогресії і в одержаній формулі окремо виділити дійсну частину від уявної.

3. Довести, що

$$\cos\frac{\pi}{11} + \cos\frac{3\pi}{11} + \cos\frac{5\pi}{11} + \cos\frac{7\pi}{11} + \cos\frac{9\pi}{11} = \frac{1}{2}.$$

4. Знайти суми

$$P_n = 1 - C_n^2 + C_n^4 - C_n^6 + \ldots,$$
$$Q_n = C_n^1 - C_n^3 + C_n^5 - \ldots,$$

де $(1+i)^n = P_n + Q_n \cdot i$.

5. Знайти

$$\lim_{n\to\infty}\left(1 + \frac{1}{2}\cos\varphi + \frac{1}{4}\cos 2\varphi + \ldots + \frac{1}{2^n}\cos n\varphi\right).$$

6. Знайти суми

$$P_n = \cos\varphi + 2\cos 2\varphi + \ldots + n\cos n\varphi,$$
$$Q_n = \sin\varphi + 2\sin 2\varphi + \ldots + n\sin n\varphi.$$

7. Довести, що

$$1 + C_n^4 + C_n^8 + \cdots = \frac{1}{2}\left(2^{n-1} + 2^{\frac{n}{2}}\cos\frac{n\pi}{4}\right),$$

$$C_n^1 + C_n^5 + C_n^7 + \cdots = \frac{1}{2}\left(2^{n-1} + 2^{\frac{n}{2}}\sin\frac{n\pi}{4}\right).$$

Ділення комплексних чисел

Для чисел $z = r(\cos\varphi + i\sin\varphi)$ і $z' = r'(\cos\varphi' + i\sin\varphi')$ спробуємо відшукати таке число $w = \rho\,(\cos\psi + i\sin\psi)$, щоб $z' = z \cdot w$ за умови $z \neq 0$. Тоді

$$z' = r\,(\cos\varphi + i\sin\varphi)\cdot\rho(\cos\psi + i\sin\psi) =$$
$$= r\rho\,[\cos(\varphi + \psi) + i\sin\,(\varphi + \psi)].$$

На підставі рівності $r'(\cos\varphi' + i\sin\varphi') = r\rho[\cos(\varphi + \psi) + {}$ ${}+ i\sin(\varphi + \psi)]$ матимемо $r' = \rho \cdot r$, $\varphi' = \varphi + \psi$ і, отже, $\rho = \dfrac{r'}{r}$, $\psi = \varphi' - \varphi$. Остаточно

$$\frac{z'}{z} = \frac{r'}{r}\left[\cos\,(\varphi' - \varphi) + i\sin\,(\varphi' - \varphi)\right].$$

Правило 3. При діленні числа z' *на число* $z \neq 0$ їх модулі діляться, а аргументи віднімаються:

$$\left|\frac{z'}{z}\right| = \frac{|z'|}{|z|},\ \arg\frac{z'}{z} = \arg z' - \arg z\,.$$

Наведемо перелік основних властивостей модулів комплексних чисел:

1. $|z + z'| \leq |z'| + |z|$ (випливає з означення суми векторів).
2. $|z' \cdot z| = |z'| \cdot |z|$ і $|z_1 \cdot z_2 \cdots z_m| = |z_1| \cdot |z_2| \cdots |z_m|$.
3. $|z^n| = |z|^n$.
4. $\left|\dfrac{z'}{z}\right| = \dfrac{|z'|}{|z|},\ z \neq 0$.

Геометричне тлумачення комплексних чисел приводить до такої нерівності:

$$|z + z'| \leq |z| + |z'|.$$

Її тлумачать, як співвідношення між довжинами сторін трикутника. Цю нерівність можна обґрунтувати алгебрично.

Нехай, наприклад, числа z і z' подані у тригонометричній формі

$$z = r(\cos\varphi + i\sin\varphi), \quad z' = \rho(\cos\psi + i\sin\psi),$$

де $r = |z|$, $\rho = |z'|$.

Тоді

$$z + z' = (r\cos\varphi + \rho\cos\psi) + i(r\sin\varphi + \rho\sin\psi)$$

і

$$|z + z'|^2 = (r\cos\varphi + \rho\cos\psi)^2 + (r\sin\varphi + \rho\sin\psi)^2 =$$

$$= r^2 + \rho^2 + 2r\rho(\cos\varphi\cos\psi + \sin\varphi\sin\psi) =$$

$$= r^2 + \rho^2 + 2r\rho\cos(\varphi - \psi).$$

Оскільки $\cos(\varphi - \psi)$ міститься на відрізку $[-1, 1]$, то матимемо нерівності

$$r^2 + \rho^2 - 2r\rho \le |z + z'|^2 \le r^2 + \rho^2 + 2r\rho,$$

тобто

$$(r - \rho)^2 \le |z + z'|^2 \le (\rho + r)^2.$$

Звідси дістанемо нерівності

5. $\big||z| - |z'|\big| \le |z + z'| \le |z| + |z'|$.

Вправи

1. Переконатися, що коли $z + \frac{1}{z} = 2\cos\varphi$, тоді

$$z^m + \frac{1}{z^m} = 2\cos m\varphi.$$

2. Обчислити

$$\frac{\left(\sqrt{3} + i\right)^{23}}{(-1 - i)^{42}}.$$

Добування коренів

Поле C комплексних чисел відрізняється від поля R дійсних чисел двома вимогами.

По-перше. Поле R упорядковане. У полі C неможливо встановити відношення > (більше) для будь-якої пари чисел із збереженням усіх основних властивостей цього відношення.

По-друге. Поле C замкнене відносно ще однієї операції – операції добування коренів будь-якого степеня n: для будь-якого $z \in C$ і будь-якого натурального числа n знайдеться таке $w \in C$, що $w^n = z$.

Розглянемо спочатку добування квадратного кореня з комплексного числа. Нехай $z = a + bi$ – комплексне число. Потрібно знайти таке число $w = x + yi$, щоб $w^2 = z$. У цьому випадку для x і y дістанемо систему рівнянь

$$\begin{cases} x^2 - y^2 = a, \\ 2xy = b. \end{cases}$$

Введемо нові невідомі $t_1 = x^2$, $t_2 = -y^2$. Тоді t_1 і t_2 – корені квадратного рівняння

$$t^2 - at - \frac{1}{4}b^2 = 0 \Rightarrow t_{1,2} = \frac{a \pm \sqrt{a^2 + b^2}}{2} \Rightarrow$$

$$\Rightarrow x = \pm\sqrt{\frac{|z| + a}{2}}, \ y = \pm\sqrt{\frac{|z| - a}{2}},$$

де $|z| = \sqrt{a^2 + b^2}$.

Знаки для x і y потрібно вибирати так, щоб виконувалась рівність

$$2xy = b.$$

Якщо $b > 0$, то x і y мають однакові знаки; якщо $b < 0$ – різні знаки; якщо $b = 0$, тоді $y = 0$, $x = \pm\sqrt{a}$, $a > 0$; $x = 0$, $y = \pm\sqrt{|a|}$, $a < 0$.

За цих умов для будь-якого числа $z \neq 0$ маємо два корені $\sqrt{z}$.

Приклад 1. Обчислити $\sqrt{3 - 4i}$.

Знайдемо корінь у вигляді $w = x + iy$. Дістанемо систему рівнянь

$$\begin{cases} x^2 - y^2 = 3, \\ 2xy = -4. \end{cases}$$

Якщо $t_1 = x^2$, $t_2 = -y^2$, тоді t_1, t_2 – корені рівняння $t^2 - 3t - 4 = 0 \Rightarrow t_1 = 4,\ t_2 = -1 \Rightarrow x^2 = 4,\ y^2 = 1$.

На підставі умови $xy = -2$ знаходимо шукані розв'язки:

$$x = 2,\ y = -1 \text{ і } x = -2,\ y = 1.$$

Остаточно, коренями $\sqrt{3-4i}$ є $w_1 = 2 - i$ і $w_2 = -2 + i$.

Розглянемо загальний випадок. Нехай $z = a + bi \neq 0$. Подамо це число у тригонометричній формі: $z = r(\cos\varphi + i\sin\varphi)$. Спробуємо знайти $\sqrt[n]{z}$ серед комплексних чисел виду $w = \rho(\cos\psi + i\sin\psi)$.

Для знаходження модуля ρ і аргумента ψ використаємо рівність $w^n = z$, тобто

$$\rho^n(\cos n\psi + i\sin n\psi) = r(\cos\varphi + i\sin\varphi).$$

Рівні комплексні числа мають рівні модулі й аргументи, які відрізняються, можливо, на кратні 2π. Звідси

$$\rho^n = r,\ n\psi = \varphi + 2\pi k \quad (k = 0, \pm 1, \pm 2, \ldots).$$

Оскільки корінь w залежить від числа k, то будемо його супроводжувати індексом k

$$w_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right)$$

з будь-яким цілим $k = 0, \pm 1, \pm 2, \ldots$

Виділимо серед цих чисел w_k набір $\langle w_0, w_1, \ldots, w_{n-1}\rangle = W$. Покажемо, що W вичерпує всі попарно різні корені $\sqrt[n]{z}$.

1. Усі числа з W попарно різні, тобто

$$0 \le k < l \le n - 1 \Rightarrow w_k \neq w_l.$$

Доведення проведемо методом від супротивного.

Нехай $0 \le k < l \le n - 1$, але $w_k = w_l$, тобто

$$\sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right) =$$

$$= \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi l}{n} + i\sin\frac{\varphi + 2\pi l}{n}\right).$$

Тоді $\frac{\varphi + 2\pi l}{n} = \frac{\varphi + 2\pi k}{n} + 2\pi m$. Звідси $l = k + mn$, $l - k = mn$; $l - k$ має ділитися на n. За умовою $n - 1 \ge l - k > 0$. Прийшли до суперечності.

2. Покажемо, що будь-який корінь w_k рівний одному з чисел набору W.

Нехай l – остача від ділення числа k на n:

$$k = nh + l,\ 0 \le l \le n - 1.$$

Тоді

$$w_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right) =$$

$$= \sqrt[n]{r}\left[\cos\left(\frac{\varphi + 2\pi l}{n} + 2\pi h\right) + i\sin\left(\frac{\varphi + 2\pi l}{n} + 2\pi h\right)\right] =$$

$$= \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi l}{n} + i\sin\frac{\varphi + 2\pi l}{n}\right) = w_l.$$

Отже, $w_k = w_l$, де $w_l \in W$.

Правило 4. Коренями n-го степеня з числа в тригонометричній формі $z = r(\cos\varphi + i\sin\varphi)$ є числа

$$w_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right),\ k = 0, 1, 2, \ldots, n - 1.$$

При цьому індекс k може набувати будь-яких значень n послідовних натуральних чисел.

Приклад 2. Обчислити корені $\sqrt[4]{-4}$.

Маємо $-4 = 4(\cos\pi + i\sin\pi)$. Тому

$$w_k = \sqrt[4]{-4} = \sqrt[4]{4}\left(\cos\frac{\pi + 2\pi k}{4} + i\sin\frac{\pi + 2\pi k}{4}\right), k = 0, 1, 2, 3;$$

$$w_0 = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) = \sqrt{2}\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) = 1 + i;$$

$$w_1=\sqrt{2}\left(\cos\frac{\pi+2\pi}{4}+i\sin\frac{\pi+2\pi}{4}\right)=\sqrt{2}\left(-\frac{\sqrt{2}}{2}+i\frac{\sqrt{2}}{2}\right)=-1+i;$$

$$w_2=\sqrt{2}\left(\cos\frac{\pi+4\pi}{4}+i\sin\frac{\pi+4\pi}{4}\right)=\sqrt{2}\left(-\frac{\sqrt{2}}{2}-i\frac{\sqrt{2}}{2}\right)=-1-i;$$

$$w_3=\sqrt{2}\left(\cos\frac{\pi+6\pi}{4}+i\sin\frac{\pi+6\pi}{4}\right)=\sqrt{2}\left(\frac{\sqrt{2}}{2}-i\frac{\sqrt{2}}{2}\right)=1-i.$$

Корені з одиниці

Для знаходження коренів з одиниці застосуємо загальне правило. Матимемо $1 = 1(\cos 0 + i\sin 0)$. Тоді шукані корені обчислюються за формулою

$$\varepsilon_k=\sqrt[n]{1}=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n},$$

$$k=0,1,2,\ldots,n-1.$$

На площині кінці векторів, що зображують числа ε_k, ділять одиничне коло з центром на початку координат на n рівних частин (рис. 4).

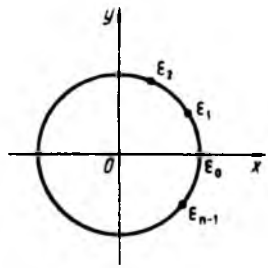


Рис. 4

Серед коренів n-го степеня з одиниці окремо виділяються так звані *первісні* корені, степені яких дають решту коренів даного степеня n. Зокрема, корінь ε_1 серед усіх коренів ε_k є первісним, оскільки

$$\varepsilon_k=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}=\left(\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n}\right)^k=\varepsilon_1^k.$$

Правило відшукання всіх первісних коренів n-го степеня з одиниці наведено в § 1.4.

Вправи

1. Розв'язати рівняння $x^2-(2+i)x+(-1+7i)=0$.

2. Знайти всі первісні корені 6-го степеня з одиниці.

3. Знайти суму всіх різних коренів n-го степеня з одиниці.

4. Знайти суму m-го степеня всіх коренів n-го степеня з одиниці.

Відповідь. Шукана сума дорівнює n, якщо число m ділиться на n, і дорівнює 0, якщо m не ділиться на n.

5. Обчислити суму $1+2\varepsilon+3\varepsilon^2+...+n\varepsilon^{n-1}$, де ε – будь-який корінь n-го степеня з одиниці.

§ 1.3. Комутативні кільця без дільників нуля (кільця цілісності)

Означення. *Комутативне кільце K називається кільцем без дільників нуля, коли для всіх $a, b \in K$ із $a \cdot b = 0$ випливає, що $a = 0$ або $b = 0$. Такі кільця називають також кільцями цілісності.*

Прикладами таких кілець є будь-яке кільце цілих чисел, будь-яке поле тощо.

Нехай K – деяке комутативне кільце. Побудуємо за допомогою нового символу x, який не належить K, такі вирази:

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_mx^m,$$
$$a_j \in K, \ j = 0, 1, 2, \ldots, m$$

зі скінченним числом членів. Вважатимемо, що $ax = xa$ для будь-якого $a \in K$. Такі вирази називаються *многочленами від* x; позначимо їх множину через $K[x]$.

Для будь-яких многочленів

$$f(x) = a_0 + a_1x + \ldots + a_mx^m, \ g(x) = b_0 + b_1x + \ldots + b_nx^n$$

із множини $K[x]$ визначимо операції додавання і множення за правилами:

$$f(x) + g(x) = \sum c_jx^j,$$

де $c_j = a_j + b_j$, $j = 0, 1, 2, ..., N$; $N = \max(m, n)$;

$$f(x) \cdot g(x) = \sum d_jx^j,$$

де $d_j = \sum\limits_{i+k=j} a_ib_k$, $j = 0, 1, 2, ..., M$; $M = n + m$.

Відносно цих операцій множина стає комутативним кільцем. Перехід від кільця K до кільця $K[x]$ називається *приєднанням змінної* x.

Якщо до кільця K послідовно приєднати змінні $x_1, x_2, ..., x_n$, тобто будувати кільця $K_1 = K[x_1]$, $K_2 = K_1[x_2]$ і т. д., тоді дістанемо кільце $K[x_1, x_2, ..., x_n]$, яке складається з виразів виду

$$\sum a_{m_1, m_2, ..., m_n} \cdot x_1^{m_1} \cdot x_2^{m_2} \cdots x_n^{m_n},$$

які називаються *многочленами від змінних* $x_1, ..., x_n$. При цьому можлива будь-яка перестановка множників $x_1^{k_1}, x_2^{k_2}, ..., x_n^{k_n}$ і елементів $a \in K$.

Кільце $K[x_1, x_2, ..., x_n]$ *називається кільцем многочленів від змінних* $x_1, ..., x_n$ *над кільцем* K.

Якщо K – кільце цілісності, то кільце $K[x]$ – також кільце цілісності.

Нехай $f(x), g(x) \in K[x]$ і $f(x) \neq 0$, $g(x) \neq 0$. Позначимо через $a_m \neq 0$ старший коефіцієнт $f(x)$, а через $b_n \neq 0$ – старший коефіцієнт $g(x)$. Тоді $a_m \cdot b_n$ – старший коефіцієнт $f(x)g(x)$ і $a_m \cdot b_n \neq 0$. Отже, $f(x)g(x) \neq 0$.

За допомогою індукції за числом n змінних $x_1, x_2, ..., x_n$ доводиться, що кільце $K[x_1, x_2, ..., x_n]$ не має дільників нуля.

Якщо комутативне кільце K міститься в деякому полі F, то воно не має дільників нуля і з елементів його будуються частки: якщо $a, b \in K$, $b \neq 0$, тоді $\frac{a}{b} = ab^{-1} = b^{-1}a \in F$. Можна переконатися, що для відношень виконуються такі рівності:

$$\frac{a}{b} = \frac{c}{d} \Leftrightarrow a \cdot d = b \cdot c (a, b, c, d \in K),$$

$$\frac{a}{b} + \frac{c}{d} = \frac{da + bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}.$$

Ці частки $\frac{a}{b}$ утворюють поле P, яке міститься в полі F. Воно називається *полем часток комутативного кільця* K.

Теорема. *Кожне кільце цілісності K можна занурити в поле.*

Доведення. Нехай кільце K містить елементи, відмінні від нуля. Розглянемо множину всіх пар (a, b), де $a, b \in K$, $b \neq 0$. Введемо на цій множині відношення еквівалентності, поклавши

$$(a, b) \sim (c, d) \Leftrightarrow ad = bc.$$

Це відношення пар є відношенням еквівалентності.

1. Рефлексивність: $(a, b) \sim (a, b)$;

2. Симетричність: $(a, b) \sim (c, d) \Rightarrow (c, d) \sim (a, b)$;

3. Транзитивність: $(a, b) \sim (c, d), (c, d) \sim (e, f) \Rightarrow (a, b) \sim (e, f)$.

Доведемо транзитивність відношення еквівалентності.

Відношення $(a, b) \sim (c, d)$ означає, що $ad = bc$; відношення $(c, d) \sim (e, f)$ означає, що $cf = de$. Маємо

$$(ad) f = (bc) f = b(cf) = b(de).$$

Звідси $(af)d = (be)d$. Але $d \neq 0$. Тому $af = be$, тобто $(a, b) \sim (e, f)$.

Відношення еквівалентності розбиває всі пари (a, b) на класи еквівалентності: множина всіх еквівалентних пар складає один клас.

Клас, який містить пару (a, b), позначимо через $\frac{a}{b}$. Отже,

$$\frac{a}{b} = \frac{c}{d} \Leftrightarrow (a, b) \sim (c, d) \Leftrightarrow b \neq 0,\ d \neq 0,\ ad = bc.$$

Введемо операції додавання і множення за правилами:

1. $\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$.

2. $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$.

Перш за все треба довести, що введені операції над класами не залежать від вибору представників класів $\frac{a}{b}$ і $\frac{c}{d}$. Нехай

$(a', b') \in \dfrac{a}{b}, (c', d') \in \dfrac{c}{d}$, тобто $a'b = ab'$, $c'd = cd'$. Доводимо збіг класів

$$\frac{ad+bc}{bd}, \quad \frac{a'd'+b'c'}{b'd'}.$$

Для цього слід встановити $(ad+bc, bd) \sim (a'd'+b'c', b'd')$.

Проведемо серію обчислень:

$$(ad+bc)b'd' = adb'd' + bcb'd' = (ab')dd' + (cd')bb' =$$

$$= (a'b)dd' + (c'd)bb' = a'bdd' + c'dbb' = (a'd' + c'b')\, bd \Leftrightarrow$$

$$\Leftrightarrow (ad+bc)b'd' = (a'd' + c'b')\, bd\,.$$

Це і є необхідна нам рівність.

Аналогічно доводиться рівність

$$\frac{ac}{bd} = \frac{a'c'}{b'd'}.$$

З'ясуємо тепер, що введені операції над класами задовольняють усі основні закони поля.

Перевіримо закон асоціативності додавання:

$$\frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) = \frac{a}{b} + \frac{cf+de}{df} = \frac{adf + b(cf+de)}{bdf} = \frac{adf+bcf+bde}{bdf};$$

$$\left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f} = \frac{ad+bc}{bd} + \frac{e}{f} = \frac{adf+bcf+bde}{bdf} = \frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right).$$

Решта законів перевіряється аналогічно.

Остаточно, множина P усіх класів $\dfrac{a}{b}$ є поле. Воно називається *полем часток кільця цілісності K.*

Отже, кільце K міститься у полі P.

З кожним елементом $c \in K$ зіставимо всі частки виду $\dfrac{cb}{b}$, $b \neq 0$. Усі ці частки рівні

$$\frac{cb}{b} = \frac{cb'}{b'}$$

для b, $b' \neq 0$, оскільки $cbb' = cbb'$.

Двом різним елементам $c, c' \in K$ відповідають різні частки $\frac{cb}{b}, \frac{c'b}{b}$. Можна показати, що встановлена відповідність між елементами $c \in K$ і частками $\frac{cb}{b}$ не тільки взаємно однозначна, але й зберігає операції додавання і множення. У цьому плані кільце K вкладено в поле P.

Означення. *Раціональними функціями змінних* $x_1, x_2, ..., x_{n-1}, x_n$ *над полем K називаються частки*

$$\varphi = \frac{f(x_1, x_2, ..., x_n)}{g(x_1, x_2, ..., x_n)}$$

многочленів $f(x_1, x_2, ..., x_n)$ *і* $g(x_1, x_2, ..., x_n) \neq 0$ *з кільця* $K[x_1, x_2, ..., x_n]$.

З доведеної вище теореми випливає, що множина всіх раціональних функцій від змінних $x_1, x_2, ..., x_n$ над полем K є поле.

§ 1.4. Кільце цілих чисел. Найбільший спільний дільник цілих чисел

Почнемо з доведення теореми про подільність цілих чисел.

Теорема 1. *Нехай a і m – цілі числа (припустимо, що* $m \neq 0$*). Для них знайдуться такі цілі числа h і r, причому* $0 \leq r < |m|$*, що* $a = mh + r$.

Число h називається часткою, а r – остачею від ділення a на m. Частка й остача визначаються числами a і m однозначно.

Доведення. Позначимо через X множину всіх цілих чисел, які можна подати у вигляді $x = a - mz$, де z пробігає кільце Z усіх цілих чисел. На множині X містяться додатні та від'ємні числа.

Нехай r – найменше невід'ємне число з X. Тоді $r = a - mh$ для деякого $h \in Z$. Припустимо, що $r \geq |m|$. Якщо $r_1 = r - m$, то $r_1 < r$ і $r_1 \geq 0$. Але $r_1 = a - m(h+1) \in X$. Це суперечить вибору числа r. Аналогічно одержується суперечність і у випадку $m<0$.

Отже, $a = mh + r$ і $0 \le r < |m|$.

Тепер доведемо єдиність. Нехай, крім того, маємо рівність $a = mh' + r'$ і $0 \le r' < |m|$. Вважатимемо, що $r \ge r'$. З двох рівностей відніманням дістанемо $0 = m(h - h') + (r - r')$. Звідси $r - r' = m(h' - h)$. Оскільки $r - r' < |m|$, то остання рівність правильна лише при $r - r' = 0$, тобто при $r = r'$. Тому $m(h - h') = 0 \Rightarrow h - h' = 0 \Rightarrow h' = h$.

Теорему доведено.

Якщо число a ділиться на m, тобто остача $r = 0$, то відношення між числами a і m позначається $m \mid a$.

Означення. *Підмножина I кільця Z цілих чисел називається ідеалом, якщо вона має такі властивості:*

I. $x, y \in I \Rightarrow x + y \in I$;

II. $x \in I,\ z \in Z \Rightarrow x \cdot z \in I$.

Наведемо один із способів побудови ідеалів кільця Z.

Нехай $a_1, a_2, ..., a_m$ – деякі цілі числа, I – множина всіх цілих чисел виду $x = a_1 z_1 + a_2 z_2 + ... + a_m z_m$, де $z_1, z_2, ..., z_m$ пробігає кільце Z.

Легко переконатися, що I – ідеал.

Якщо $x, y \in I$, то $x = a_1 z_1 + ... + a_m z_m$, $y = a_1 z_1' + a_2 z_2' + ... + a_m z_m'$, де $z_k, z_k' \in Z$. Звідси $x + y = a_1(z_1 + z_1') + ... + a_m(z_m + z_m') \in I$. Аналогічно перевіряється і властивість 2, характерна для ідеалу.

Побудований ідеал, породжений числами $a_1, ..., a_m$, позначається через $I = (a_1, a_2, ..., a_m)$.

Ідеал I називається *головним*, якщо він породжується одним цілим числом a:

$$I = (a).$$

Теорема 2. *У кільці цілих чисел будь-який ідеал – головний.*

Доведення. Нехай ідеал I містить числа, відмінні від нуля (в іншому випадку твердження теореми очевидне). Тоді у ньому містяться і додатні числа. Виберемо в ньому найменше додатне число m. Нехай $a \in I$ і r – остача від ділення a на m, тобто $a = mh + r$. Нехай $0 < r < m$. Тоді матимемо $r = a - mh$. Оскільки $a,\ m \in I$, то $r = a - mh \in I$. Припущення, що $0 < r$ суперечить вибору числа m. Тому $r = 0$ і, отже, $a = mh$. Це означає, що число a – елемент головного ідеалу (m), породженого числом m. У результаті довільності вибору числа a з ідеалу I маємо включення $I \subseteq (m)$.

Обернене включення $(m) \subseteq I$ очевидне. Остаточно, $I = (m)$.

Теорему доведено.

Означення. *Найбільшим спільним дільником (НСД) цілих чисел $a_1, a_2, ..., a_p$ називається ціле число m, яке має такі властивості:*

а) m – спільний дільник цих чисел, тобто

$$m \mid a_1, m \mid a_2, ..., m \mid a_p;$$

б) будь-який спільний дільник цих чисел ділить число m, тобто якщо $n \mid a_1, n \mid a_2, ..., n \mid a_p$, тоді $n \mid m$.

Позначення найбільшого спільного дільника має вигляд

$$m = \text{НСД}\left(a_1, a_2, ..., a_p\right).$$

Переконаємось, що будь-які два НСД даних чисел можуть відрізнятися лише знаком.

Нехай також $l = \text{НСД}\left(a_1, a_2, ..., a_p\right)$. З означення випливає: $l \mid m$ і $m \mid l$. Тоді $m = lf$ і $l = mg$, де $f, g \in Z$. Звідси $l = lfg$ і тому $fg = 1$. Остання рівність можлива лише при $f = \pm 1$ і $g = \pm 1$. Отже, $m = \pm 1$.

Зауваження. Доведену вище властивість б) можна ще сформулювати так:

в) дільник m є найбільшим за абсолютною величиною дільником чисел $a_1, a_2, ..., a_p$.

Еквівалентність властивостей а), б) з властивостями а), в) випливає із теореми 3.

Теорема 3. *Для будь-яких цілих чисел* $a_1, a_2, ..., a_p$ *існує найбільший спільний дільник. Якщо* $m = НСД\left(a_1, a_2, ..., a_p\right)$, *то існують такі цілі числа* $z_1, z_2, ..., z_p$, *що*

$$a_1 z_1 + a_2 z_2 + ... + a_p z_p = m.$$

Доведення. Нехай $I = \left(a_1, a_2, ..., a_p\right)$ – ідеал, породжений числами $a_1, a_2, ..., a_p$. Він є головним: $I = (m)$. Тоді $m \mid a_k$, $k = \overline{1, p}$. Існують такі цілі числа $z_1, z_2, ..., z_p$, для яких має місце рівність $m = a_1 z_1 + a_2 z_2 + ... + a_p z_p$. Якщо l – спільний дільник чисел $a_1, a_2, ..., a_p$, то з останньої рівності випливає $l \mid m$. Отже, m – НСД даних чисел.

Теорему доведено.

Алгоритм Евкліда

Для знаходження НСД чисел a і b потрібно провести таку послідовність ділення з остачею:

$$\left.\begin{array}{lll}
(1) & a = bh_1 + r_1, & 0 \le r_1 < |b|; \\
(2) & b = r_1 h_2 + r_2, & 0 \le r_2 < r_1; \\
(3) & r_1 = r_2 h_3 + r_3, & 0 \le r_3 < r_2; \\
\multicolumn{3}{l}{\dotfill} \\
(k) & r_{k-2} = r_{k-1} h_k + r_k, & 0 \le r_k < r_{k-1} \\
(k+1) & r_{k-1} = r_k h_{k+1}. &
\end{array}\right|$$

У цій послідовності на деякому (k-му) кроці одержується додатна остача $\left(r_k > 0\right)$, а на наступному ($k+1$-му) – ділення без остачі.

Доведемо, що r_k = НСД (a, b).

Нехай m = НСД (a, b). Оскільки $I = (a, b) = (m)$, то маємо серію включень у головний ідеал (m):

$$r_1 = a - bh_1 \in (m),$$
$$r_2 = b - r_1 h_2 \in (m),$$
$$r_3 = r_1 - r_2 h_3 \in (m),$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$r_k = r_{k-2} - r_{k-1} h_k \in (m).$$

Тому $m \mid r_k$.

З другого боку, з послідовності рівностей (1), (2), …, (k), $(k+1)$ одержуємо послідовність подільностей:

$$r_k \mid r_{k-1}, \quad r_k \mid r_{k-2}, \quad \dots, \quad r_k \mid r_1, \quad r_k \mid b, \quad r_k \mid a.$$

За означенням НСД (a, b) маємо, що $r_k \mid m$. Вважаючи m додатним, дістаємо рівність $r_k = m$.

Вправи

1. Нехай N – множина натуральних чисел. Довести, що при діленні квадрата n^2 $(n \in N)$ на 3 остача не може дорівнювати 2.

2. Довести, що сума кубів трьох послідовних натуральних чисел ділиться на 9.

3. Довести, що $7 \mid a^2 + b^2 \Rightarrow 7 \mid a,\ 7 \mid b,$ де $a, b \in N$.

4. Знайти НСД $(9n+4,\ 2n+1)$, $n \in N$.

5. Індукцією за $n \in N$ довести, що числа виду $7^{n+2} + 8^{2n+1}$ діляться на 57.

Взаємно прості числа

Означення. *Числа a і b називаються взаємно простими, якщо їх НСД дорівнює ± 1.*

Теорема 4. *Для того щоб числа a і b були взаємно простими, необхідно і достатньо існування таких цілих чисел x, y, для яких $ax + by = 1$.*

Доведення. *Необхідність.* Нехай a і b – взаємно прості числа. Тоді ідеал $(a, b) = (1) = Z$, і тому $ax + by = 1$ для деяких цілих чисел x, y.

Достатність. Нехай виконується рівність $ax+by=1$ для деяких $x, y \in Z$. Якщо $m=$ НСД (a, b), тоді $a=ma'$, $b=mb'$, $a', b' \in Z$ і тому $m(a'x+b'y)=1$, тобто m – дільник 1. Це можливо лише у випадку, якщо $m=\pm 1$, тобто, коли a і b – взаємно прості числа.

Теорему доведено.

Приклад. Переконатися, що числа $a=445$ і $b=192$ – взаємно прості, та знайти такі числа x і y, щоб $445x+192y=1$.

Застосуємо алгоритм Евкліда:

(1) $a=b\cdot 2+r_1$, $r_1=61$;

(2) $b=r_1\cdot 3+r_2$, $r_2=9$;

(3) $r_1=r_2\cdot 6+r_3$, $r_3=7$;

(4) $r_2=r_3\cdot 1+r_4$, $r_4=2$;

(5) $r_3=r_4\cdot 3+r_5$, $r_5=1$.

Одержали НСД $(a, b)=1$. Для знаходження чисел x і y проводимо таку послідовність обчислень:

1. З (5) дістаємо $1=r_5=r_3-3r_4$.
2. З (4) маємо $r_4=r_2-r_3$. Отже, $1=4r_3-3r_2$.
3. З (3) одержуємо $r_3=r_1-6r_2$. Тому $1=4r_1-27r_2$.
4. З (2) дістаємо $r_2=b-3r_1$. Отже, $1=85r_1-27b$.
5. З (1) маємо $r_1=a-2b$. Тоді $1=85a-197b$.

З останньої рівності випливає, що $x=85$, $y=-197$.

Основні властивості взаємно простих чисел

I. Якщо число a взаємно просте з числами b і c, то a взаємно просте з добутком bc.

Доведення. За умовою НСД $(a, b)=1$. Існують такі $x, y \in Z$, що $ax+by=1$. Звідси дістанемо рівність

$$acx+bcy=c.$$

Якщо $m=$ НСД (a, bc), то $a=a'm$, $bc=dm$, де $a', d \in Z$ і, отже, $m(a'x+dy)=c$, тобто $m\,|\,c$. Із відношень $m\,|\,a$ і $m\,|\,c$ випливає відношення $m\,|\,1$, оскільки НСД $(a, c)=1$. Тоді $m=1$.

Число a взаємно просте з добутком bc, що й доводить ця властивість.

II. Якщо число a взаємно просте з числом b і $a \mid bc$, то $a \mid c$.

Доведення. Нехай цілі числа x, y такі, що $ax + by = 1$. Тоді $acx + bcy = c$. Оскільки в останній рівності обидва доданки лівої частини діляться на a, то і права частина ділиться на a, що й доводить властивість II.

III. Якщо числа a і b взаємно прості та є дільниками числа n, то число n ділиться на добуток ab.

Доведення. Нехай x, y – такі цілі числа, що $ax + by = 1$. За умовою $a \mid n$, $b \mid n$, тобто $n = ak$, $n = bl$, де k, l – цілі числа. Тоді

$$n = nax + nby = blax + akby = ab(lx + ky),$$

тобто $ab \mid n$ – властивість III.

Як приклад доведемо теорему про властивість коренів n-го степеня з одиниці.

Теорема 5. *Корінь*

$$w_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}, \quad k = \overline{0, n-1}.$$

n-го степеня з одиниці буде первісним тоді й лише тоді, коли n і k – взаємно прості числа.

Доведення. Нехай d = НСД (n, k). Розглянемо два можливі випадки.

1. *Необхідність*. Нехай $d =$ НСД$(n,k) > 1$. Тоді $k = dl$, $n = dm$, причому $m < n$. Маємо $w_k = \cos\frac{2\pi dl}{n} + i\sin\frac{2\pi dl}{n} = \cos\frac{2\pi l}{m} + i\sin\frac{2\pi l}{m}$ і тому $w_k^m = 1$. Тобто w_k – корінь m-го степеня з одиниці. Степені цього кореня також будуть коренями m-го степеня з одиниці. Остаточно, корінь w_k не є первісним, бо $m < n$.

Отже, для того щоб w_k був первісним коренем, необхідно, щоб НСД $(n, k) = 1$.

2. *Достатність*. Нехай $d = 1$. Знайдемо такі числа x, y, для яких $nx + ky = 1$. Тоді

$$w = w_k^y = \left(\cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}\right)^y = \cos\frac{2\pi ky}{n} + i\sin\frac{2\pi ky}{n} =$$

$$= \cos\frac{2\pi(1-nx)}{n} + i\sin\frac{2\pi(1-nx)}{n} = w_1.$$

Але w_1 – первісний корінь. Тому степені $w_1^l = w_k^{ly}$, $l = 0, 1, ..., n-1$ містять у собі всі корені $\sqrt[n]{1}$. Це й означає, що w_k – корінь первісний.

Теорему доведено.

Вправа

Довести, що найменше спільне кратне попарно взаємно простих чисел $a_1, a_2, ..., a_m$ дорівнює їх добутку.

Прості числа

Означення. *Натуральне число p називається простим, якщо його дільниками є числа ± 1, $\pm p$ і лише вони.*

Наведемо основні властивості простих чисел.

I. Два різних простих числа взаємно прості.

Властивість випливає із означення простого числа.

II. Якщо p – просте, а a – ціле число, то вони взаємно прості, або $p \mid a$.

Доведення. Нехай $m =$ НСД (p, a). Тоді число $m = \pm 1$, або $m = \pm p$, оскільки p має ділитися на m. У першому випадку p і a взаємно прості, у другому – $p \mid a$, що й доводить властивість II.

III. Якщо p – просте число і $p \mid ab$, де a і b – цілі числа, тоді або $p \mid a$, або $p \mid b$.

Властивість випливає з властивості II.

Вправа

Довести основну теорему арифметики:

Будь-яке натуральне число $n > 1$ можна подати у вигляді добутку простих множників: $n = p_1 \cdots p_m$. Таке подання є єдине з точністю до порядку розміщення простих множників.

Класи лишків

Означення. *Нехай m – ціле число більше одиниці. Два числа a і b називаються порівнянними за модулем m, якщо їх різниця $a-b$ ділиться на m. Це співвідношення між числами записується у вигляді*

$$a \equiv b \pmod{m}.$$

Іншими словами, $a \equiv b \pmod{m} \Leftrightarrow a-b \in (m)$, де (m) – головний ідеал, породжений числом m. Для спрощення запишемо

$$a \equiv b(m).$$

Відношення порівнянності має властивості еквівалентності.

1. Рефлексивність: $a \equiv a(m)$;

2. Симетричність: $a \equiv b(m) \Rightarrow b \equiv a(m)$;

3. Транзитивність: $a \equiv b(m)$ і $b \equiv c(m) \Rightarrow a \equiv c(m)$.

Теорема 6. *Якщо $a \equiv a'(m)$ і $b \equiv b'(m)$, тоді $a+b \equiv a'+b'(m)$, $ab \equiv a'b'(m)$.*

Доведення. Порівняння $a \equiv a'(m)$ означає: для деякого цілого числа g маємо $a = a' + mg$. Отже, порівняння $b \equiv b'(m)$ означає, що $b = b' + mf$ для деякого цілого числа f. Але тоді $a+b = a'+b'+m(g+f)$ і $ab = (a'+mg)(b'+mf) = a'b' + {}$ $+m(gb'+fa'+mgf)$. Тому $a+b-a'-b' \in (m)$ і $ab-a'b' \in (m)$, тобто $a+b \equiv a'+b'(m)$, $ab \equiv a'b'(m)$, що й доводить теорему.

Оскільки порівняння чисел за модулем m – відношення еквівалентності, то кільце Z цілих чисел розбивається на класи, у кожному з яких містяться всі попарно порівнянні числа. Різні класи не мають спільних елементів. Вони називаються *класами лишків* за модулем m.

Якщо A – один з цих класів лишків і $a \in A$, тоді A складається з усіх тих чисел, які порівнянні за модулем m. Число a повністю визначає цей клас і називається його представником.

Клас A будемо записувати так:

$$A = \bar{a}.$$

Усі числа даного класу $\bar{a}$ характеризуються такою властивістю: остачі від ділення їх на m збігаються. Тому, якщо $a = mh + r$, $0 \le r < m$, то $\bar{a} = \bar{r}$. Зокрема, клас $\bar{0}$ з представником 0 збігається з ідеалом (m). Отже, усі класи лишків вичерпуються множинами $\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{m-1}$.

Операції додавання і множення над класами лишків

Означення. *Додаванням класів $\bar{a}$ і $\bar{b}$ називається клас з представником $a + b$. Операцію додавання позначаємо знаком плюс:*

$$\bar{a} + \bar{b} = \overline{a + b}.$$

Додавання класів не залежить від вибору представників класів.

Дійсно, нехай $a' \in \bar{a}$, $b' \in \bar{b}$, тобто $a \equiv a'(m)$ і $b \equiv b'(m)$. Маємо рівності $\bar{a} = \overline{a'}$, $\bar{b} = \overline{b'}$, $\bar{a} + \bar{b} = \overline{a + b}$, $\overline{a'} + \overline{b'} = \overline{a' + b'}$. Але $\overline{a+b} = \overline{a' + b'}$, оскільки $a + b \equiv a' + b'(m)$. Тому $\bar{a} + \bar{b} = \overline{a'} + \overline{b'}$, що й доводить твердження.

Означення. *Множенням класів $\bar{a}$ і $\bar{b}$ називається клас з представником ab. Операцію множення позначаємо знаком множення:*

$$\bar{a} \cdot \bar{b} = \overline{ab}.$$

Добуток цих класів не залежить від вибору його представників.

Доведення цієї властивості таке саме, як і для теореми про додавання класів.

Теорема 7. *Множина $K_m = \left(\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{m-1}\right)$ усіх класів лишків за модулем m з введеними операціями додавання і множення є комутативним кільцем з одиницею. Якщо число m – просте, тоді кільце K_m – поле.*

Доведення полягає у перевірці виконання всіх законів, що характеризують комутативні кільця. При цьому роль

нульового елемента на множині K_m відіграє клас $\bar{0}$, оскільки $\bar{a}+\bar{0}=\overline{a+0}=\bar{a}$ для будь-якого $\bar{a}\in K_m$. Роль одиниці виконує клас $\bar{1}$, оскільки $\bar{a}\cdot\bar{1}=\overline{a\cdot 1}=\bar{a}$. Далі, $\bar{a}\left(\bar{b}+\bar{c}\right)=\bar{a}\left(\overline{b+c}\right)=$
$=\overline{a\left(b+c\right)}=\overline{ab+ac}=\overline{ab}+\overline{ac}=\bar{a}\cdot\bar{b}+\bar{a}\cdot\bar{c}$ – розподільний закон (III_1) для кільця.

Нехай тепер m – просте число. Достатньо перевірити закон множення (II_4):

$$\forall\bar{a}\neq\bar{0}\,\exists\bar{x}\left\{\bar{a}\cdot\bar{x}=\bar{x}\cdot\bar{a}=\bar{1}\right\}.$$

Якщо $\bar{a}\neq\bar{0}$, то число a не ділиться на просте число m і, отже, a і m – взаємно прості числа. Тому існують такі цілі числа x, y, що $ax+my=1$. Тоді $ax=1-my$ і $\bar{a}\cdot\bar{x}=\overline{a\cdot x}=\overline{1-mx}=\bar{1}$. Отже, для класу $\bar{a}\neq\bar{0}$ існує клас $\bar{x}$, обернений йому відносно операції множення, що й доводить теорему.

Так перевіряються всі основні закони (аксіоми) для комутативних кілець і полів.

Наведемо приклад поля класів лишків K_3 з таблицями додавання і множення.

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

Вправа

Довести, що скінченне комутативне кільце без дільників нуля є полем.

§ 1.5. Кільце многочленів однієї змінної

Ділення многочленів

Нехай P – довільне поле і x – змінна, яка не належить полю P. Кільце $P[x]$ усіх многочленів з коефіцієнтами із поля P є кільцем цілісності.

Теорема 1 (про подільність). *Якщо $f(x)$, $g(x)$ – многочлени з кільця $P[x]$ і $g(x)\neq 0$, то $f(x)$ можна подати у вигляді*

$$f(x)=g(x)h(x)+r(x), \qquad (1)$$

де $h(x)$ і $r(x)$ – многочлени з того самого кільця і степінь многочлена $r(x)$ менший від степеня $g(x)$. Таке подання єдине, тобто $h(x)$ і $r(x)$ визначаються однозначно. Многочлен $h(x)$ називається часткою, а $r(x)$ – остачею від ділення $f(x)$ на $g(x)$.

Доведення. Відокремимо в $f(x)$ і $g(x)$ їхні старші члени і запишемо їх у вигляді

$$f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n,\ a_0\neq 0;$$

$$g(x)=b_0x^m+b_1x^{m-1}+\ldots+b_{m-1}x+b_m,\ b_0\neq 0.$$

Доведення проведемо індукцією за степенем n многочлена $f(x)$. При цьому можна вважати, що $m\geq 1$.

Якщо $0\leq n<m$, то маємо рівність $f(x)=g(x)\cdot 0+f(x)$ і тому $h(x)=0$, $r(x)=f(x)$, що й вимагалось умовою теореми.

Нехай тепер $n\geq m$ і припустимо, що теорема має місце для многочленів, степені яких менші, ніж n.

Розглянемо многочлен $f(x)$ степеня n. Тоді $\dfrac{a_0x^n}{b_0x^m}=\dfrac{a_0}{b_0}x^{n-m}$ – многочлен. Розглянемо різницю

$$f_1(x)=f(x)-\frac{a_0}{b_0}x^{n-m}g(x)=$$

$$=a_0x^n+a_1x^{n-1}+\ldots-\frac{a_0}{b_0}x^{n-m}\left(b_0x^m+b_1x^{m-1}+\ldots\right)=$$

$$=\left(a_1-\frac{a_0}{b_0}b_1\right)x^{n-1}+\left(a_2-\frac{a_0}{b_0}b_2\right)x^{n-2}+\ldots$$

Степінь многочлена $f_1(x)$ не перевищує $n-1$. Усі його коефіцієнти одержуються з коефіцієнтів $f(x)$ і $g(x)$ за допомогою операцій додавання, віднімання, множення і ділення. Тому $f_1(x)\in P[x]$.

Застосуємо до $f_1(x)$ припущення індукції

$$f_1(x) = g(x)h_1(x) + r(x),$$

де $h_1(x), r(x) \in P[x]$ і степінь $r(x)$ менший, ніж m. Тоді

$$f(x) = g(x)\left[\frac{a_0}{b_0}x^{n-m} + h_1(x)\right] + r(x) = g(x)h(x) + r(x),$$

$$h(x) = \frac{a_0}{b_0}x^{n-m} + h_1(x).$$

І перше твердження теореми доведено.

Доведемо єдиність. Припустимо, що крім співвідношення (1) існує інша рівність:

$$f(x) = g(x)H(x) + R(x)$$

з тими самими властивостями. Матимемо

$$g(x)h(x) + r(x) = g(x)H(x) + R(x) \Rightarrow$$

$$\Rightarrow g(x)[H(x) - h(x)] = r(x) - R(x).$$

Припустимо також, що $H(x) - h(x) \neq 0$. Тоді степінь многочлена $g(x)[H(x) - h(x)]$ буде більшим за m–1. Степінь різниці $r(x) - R(x)$ менший за m. Прийшли до суперечності. Тому $H(x) - h(x) = 0$ і $H(x) = h(x)$, $R(x) = r(x)$.

Теорему доведено.

Доведена вище теорема дає обґрунтування відомого алгоритму ділення многочлена на многочлен (правило “куточка”).

1. Зліва записуємо многочлен $f(x)$, справа – $g(x)$. Ділимо старший член a_0x^n многочлена $f(x)$ на старший член b_0x^m многочлена $g(x)$ і записуємо одержану частку нижче $g(x)$:

$$\begin{array}{r|l} f(x) & g(x) \\ \cline{2-2} & \frac{a_0}{b_0}x^{n-m} \end{array}$$

2. Обчислюємо $\frac{a_0}{b_0}x^{n-m}g(x)$ і записуємо його нижче $f(x)$.

3. Обчислюємо різницю $f_1(x)=f(x)-\frac{a_0}{b_0}x^{n-m}g(x)$ і записуємо її ще нижче.

4. Повторюємо операцію 1 стосовно $f_1(x)$ і т. д.

При діленні многочлена на двочлен виду $x-a$ особливо зручним є алгоритм, відомий як *схема Горнера.* Пояснимо його.

Якщо $f(x)=a_0x^n+a_1x^{n-1}+...+a_n$, тоді $f(x)=(x-a)h(x)+r$, де остача r – стала величина. Знайдемо частку

$$h(x)=b_0x^{n-1}+b_1x^{n-2}+...+b_{n-1}.$$

Матимемо

$$a_0x^n+a_1x^{n-1}+...+a_n=(x-a)(b_0x^{n-1}+b_1x^{n-2}+...+b_{n-1})+r=$$

$$=b_0x^n+(b_1-b_0a)x^{n-1}+(b_2-b_1a)x^{n-2}+...+(r-b_{n-1}a).$$

Звідси

$$b_0=a_0,\quad b_1=a_1+b_0a,\quad b_2=a_2+b_1a,...,$$

$$b_k=a_k+b_{k-1}a,...,\ b_{n-1}=a_{n-1}+b_{n-2}a,\ r=a_n+b_{n-1}a.$$

Процес обчислень задається схемою Горнера.

a_0	a_1	a_2	...	a_{n-1}	a_n
$b_0=a_0$	$b_1=a_1+b_0a$	$b_2=a_2+b_1a$	...	$b_{n-1}=a_{n-1}+b_{n-2}a$	$r=a_n+b_{n-1}a$

Із формули $f(x)=(x-a)h(x)+r$ ділення многочлена на $x-a$ дістаємо ***теорему Безу***:

Остача від ділення довільного многочлена $f(x)$ на $x-a$ дорівнює значенню цього многочлена при $x=a$.

Приклад. Поділити $x^4-8x^3+24x^2-50x+90$ на $x-2$ за наведеною вище схемою.

1	–8	24	–50	90
$b_0=1$	$b_1=-8+2=-6$	$b_2=24-6\cdot2=12$	$b_3=-50+12\cdot2=-26$	$r=90-26\cdot2=38$

Тут частка $h(x)=x^3-6x^2+12x-26$, остача $r=38$.

Означення. *Нехай $f(x)$ – многочлен з коефіцієнтами з поля P, тобто $f(x) \in P[x]$. Елемент $a \in P$ називається його коренем, якщо $f(a) = 0$.*

Оскільки при діленні $f(x)$ на $x-a$ дістаємо $f(x) = (x-a)h(x) + r$, де r – стала величина, тоді $r = f(a) = 0$, якщо a – корінь $f(x)$. Отже, якщо a – корінь $f(x)$, тоді $f(x) = (x-a)h(x)$.

Означення. *Елемент a поля P називається коренем кратності m (m – ціле додатне число), якщо*

$$f(x) = (x-a)^m h(x),\ h(a) \neq 0.$$

Лема. *Нехай P – дійсне поле, $f(x) \in P[x]$ і a – дійсний корінь кратності m многочлена $f(x)$. Тоді a – корінь похідної $f'(x)$ з кратністю $m-1$.*

Доведення. Маємо $f(x) = (x-a)^m h(x),\ h(a) \neq 0$;

$$f'(x) = m(x-a)^{m-1} h(x) + (x-a)^m h'(x) =$$
$$= (x-a)^{m-1}\left[mh(x) + (x-a)h'(x)\right].$$

Якщо $H(x) = mh(x) + (x-a)h'(x)$, тоді $f'(x) = (x-a)^{m-1} H(x)$ і $H(a) = mh(a) \neq 0$.

Вправи

1. Довести, що остача від ділення $x^m - 1$ на $x^n - 1$ дорівнює $x^r - 1$, де r – остача від ділення m на n.

2. Довести, що $x^{3m} + x^{3n+1} + x^{3p+2}$ ділиться на $x^2 + x + 1$.

3. Довести, що $x^{3m} - x^{3n+1} + x^{3p+2}$ ділиться на $x^2 - x + 1$ тоді і лише тоді, коли всі три числа m, n, p однакової парності.

Найбільший спільний дільник многочленів

Наступні означення і доведення властивостей подільності многочленів є повторенням властивостей подільності чисел. При цьому слід лише замінити слова “ціле число” на слово “многочлен”. Тому немає потреби кожен раз наводити детальний виклад відповідних властивостей для многочленів. Лише в окремих місцях з'являться особливості, на яких слід зупинитись.

Введемо поняття *ідеалу* в кільці $P[x]$ многочленів над полем P.

Означення. *Підмножина I кільця $P[x]$ називається ідеалом, якщо вона має такі властивості:*

I. $f(x), g(x) \in I \Rightarrow f(x)+g(x) \in I$.

II. $f(x) \in I$ і $F(x) \in P[x] \Rightarrow f(x)F(x) \in I$.

При цьому ідеал I називається головним, якщо для нього існує такий многочлен $m(x) \in I$, що I вичерпується многочленами виду $m(x)F(x)$, де $F(x)$ пробігає все кільце $P[x]$. Він позначається через $m(x)$.

Теорема 2. *Будь-який ідеал кільця $P[x]$ є головним.*

Доведення цієї теореми аналогічне доведенню подібної теореми про цілі числа.

Означення. *Многочлен $d(x)$ називається найбільшим спільним дільником многочленів (НСД) $f(x)$ і $g(x)$, якщо виконуються властивості:*

I. $d(x) | f(x)$ і $d(x) | g(x)$.

II. $\delta(x) | f(x)$ і $\delta(x) | g(x) \Rightarrow \delta(x) | d(x)$.

Будь-які два НСД $(f(x), g(x))$ асоційовані: вони можуть відрізнятись лише множником $\lambda \neq 0$, що належить основному полю P.

Теорема 3. *Для будь-яких двох (або більшого числа) многочленів існує їхній найбільший спільний дільник. Якщо $d(x)$=НСД (f, g), тоді існують такі многочлени $F(x), G(x) \in P[x]$, для яких*

$$f(x)F(x)+g(x)G(x)=d(x). \tag{1}$$

(Див. § 1.4, доведення теореми 4 про взаємно прості числа).

Алгоритм Евкліда

Як і для чисел, обчислення НСД двох многочленів $f(x)$ і $g(x)$ застосовується схема (Е) послідовних ділень, яка носить назву *алгоритм Евкліда.* Цю схему доцільно подати у вигляді:

$$
\left.\begin{array}{ll}
1. & f(x)=g(x)h_1(x)+\lambda_1 r_1(x),\\
2. & g(x)=r_1(x)h_2(x)+\lambda_2 r_2(x),\\
3. & r_1(x)=r_2(x)h_3(x)+\lambda_3 r_3(x),\\
 & \dots\dots\dots\dots\dots\dots\dots\dots\\
k. & r_{k-2}(x)=r_{k-1}(x)h_k(x)+\lambda_k r_k(x),\\
k+1. & r_{k-1}(x)=r_k(x)h_{k+1}(x).
\end{array}\right| \qquad \text{(E)}
$$

Тут $\lambda_j r_j(x)$ – остача від ділення на j-му кроці. Сталі множники $\lambda_j \neq 0, j=\overline{1,k}$ – довільні і можуть бути вибраними на кожному кроці з міркувань спрощення подальших обчислень. Многочлен $r_k(x)$, який визначено останньою ненульовою остачею, і є шуканий НСД $(f(x), g(x))$.

Наведемо доведення цього твердження.

Спочатку покажемо, що $r_k(x)$ – спільний дільник $f(x)$ і $g(x)$. Многочлен $r_k(x)$ – дільник $r_{k-1}(x)$. Із k-ї рівності схеми (Е) випливає, що $r_{k-2}(x)$ ділиться на $r_k(x)$. Продовжуючи далі таким же чином, дістанемо: $r_{k-1}(x), r_{k-2}(x), \ldots, r_2(x), r_1(x)$ діляться на $r_k(x)$. З 2-ї рівності схеми (Е) одержуємо подільність $g(x)$ на $r_k(x)$, а з 1-ї – подільність $f(x)$ на $r_k(x)$.

Нехай тепер $\delta(x)$ – деякий спільний дільник $f(x)$ і $g(x)$. Із рівності 1 схеми (Е) випливає, що $\delta(x)$ – дільник $r_1(x)$, а з рівності 2 – $\delta(x)$ є дільник $r_2(x)$. Врешті-решт з'ясовуємо остаточно, що $\delta(x)$ – дільник $r_k(x)$.

Нехай $d(x) =$ НСД $(f(x), g(x))$. Тоді на підставі теореми 3 існують многочлени $F(x)$, $G(x)$ такі, для яких виконується рівність

$$F(x)f(x)+G(x)g(x)=d(x).$$

Для їх обчислення використовується або схема (Е) алгоритму Евкліда, або метод невизначених коефіцієнтів. Останній метод є ефективним, коли НСД $d(x)$ наперед відомий.

Проілюструємо застосування алгоритму Евкліда для знаходження НСД многочленів $f(x)=x^5+2x^4-3x^2-3x-1$ і $g(x)=3x^4+7x^3+4x^2-x-1$.

1. Ділимо $f(x)$ на $g(x)$:

$$\begin{array}{l|l} x^5+2x^4+0x^3-3x^2-3x-1 & 3x^4+7x^3+4x^2-x-1 \\ \cline{2-2} -\ \underline{x^5+\frac{7}{3}x^4+\frac{4}{3}x^3-\frac{1}{3}x^2-\frac{1}{3}x} & \frac{1}{3}x-\frac{1}{9} \\ \qquad -\frac{1}{3}x^4-\frac{4}{3}x^3-\frac{8}{3}x^2-\frac{8}{3}x-1 & \\ \quad -\ \underline{-\frac{1}{3}x^4-\frac{7}{9}x^3-\frac{4}{9}x^2+\frac{1}{9}x+\frac{1}{9}} & \\ \qquad\qquad -\frac{5}{9}x^3-\frac{20}{9}x^2-\frac{25}{9}x-\frac{10}{9} & \end{array}$$

Остача

$$R_1(x)=-\frac{5}{9}x^3-\frac{20}{9}x^2-\frac{25}{9}x-\frac{10}{9}=-\frac{5}{9}\left(x^3+4x^2+5x+2\right).$$

Покладемо $h_1(x)=\frac{1}{3}x-\frac{1}{9}$, $r_1(x)=x^3+4x^2+5x+2$.

2. Ділимо $g(x)$ на $r_1(x)$:

$$\begin{array}{l|l} 3x^4+7x^3+4x^2-x-1 & x^3+4x^2+5x+2 \\ \cline{2-2} -\ \underline{3x^4+12x^3+15x^2+6x} & 3x-5 \\ \qquad -5x^3-11x^2-7x-1 & \\ \quad -\ \underline{-5x^3-20x^2-25x-10} & \\ \qquad\qquad 9x^2+18x+9 & \end{array}$$

Остача $R_2(x)=9x^2+18x+9=9\left(x^2+2x+1\right)$.

Покладемо $h_2(x)=3x-5$, $r_2(x)=x^2+2x+1$.

3. Ділимо $r_1(x)$ на $r_2(x)$:

$$\begin{array}{l|l} x^3+4x^2+5x+2 & x^2+2x+1 \\ \cline{2-2} -\ \underline{x^3+2x^2+x} & x+2 \\ \qquad 2x^2+4x+2 & \\ \quad -\ \underline{2x^2+4x+2} & \\ \qquad\qquad 0 & \end{array}$$

Ділення без остачі. Отже, НСД многочленів – $r_2(x) = x^2 + 2x + 1$.

Маємо

$$f(x) = g(x)h_1(x) - \frac{5}{9}r_1(x),$$

$$g(x) = r_1(x)h_2(x) + 9r_2(x),$$

$$r_1(x) = r_2(x)h_3(x).$$

Звідси

$$9r_2(x) = g(x) - r_1(x)h_2(x) = g(x) + \frac{9}{5}[f(x) - g(x)h_1(x)]h_2(x) =$$

$$= \frac{9}{5}h_2(x)f(x) + \left[1 - \frac{9}{5}h_1(x)h_2(x)\right]g(x).$$

Підставивши $h_1(x) = \frac{1}{3}x - \frac{1}{9}$, $h_2(x) = 3x - 5$, знайдемо

$$r_2(x) = f(x)\left(\frac{3}{5}x - 1\right) + g(x)\left(-\frac{1}{5}x^2 + \frac{2}{5}x\right).$$

Остаточно

$$F(x) = \frac{3}{5}x - 1,\ G(x) = -\frac{1}{5}x^2 + \frac{2}{5}x.$$

Застосування методу невизначених коефіцієнтів

Зробимо спочатку таке зауваження. Нехай

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n,\ a_0 \neq 0,$$

$$g(x) = b_0x^m + b_1x^{m-1} + \ldots + b_m,\ b_0 \neq 0.$$

З'ясовується, що серед усіх многочленів $F(x)$, $G(x)$, для яких виконується рівність $f \cdot F + g \cdot G = d$ існують такі, що степінь $F(x)$ менший від m степеня $g(x)$, а степінь $G(x)$ менший від n степеня $f(x)$. Припустимо, що степінь многочлена $F(x)$ не менший за m. Поділимо $F(x)$ на $g(x)$ з деякою остачею $r(x)$:

$$F(x) = g(x) \cdot h(x) + r(x).$$

Тоді

$$d(x)=[g(x)h(x)+r(x)]f(x)+G(x)g(x)=$$
$$=r(x)f(x)+[G(x)+h(x)f(x)]g(x)=r(x)f(x)+S(x)g(x),$$

де $S(x)=G(x)+h(x)f(x)$.

Оскільки степінь $d(x)$ не перевищує $\min\{m,n\}$, а степінь $r(x)$ менший за m, то степінь добутку $S(x)g(x)=d(x)-r(x)f(x)$ менший за $m+n$. Тому степінь $S(x)$ менший від n.

Отже, маємо необхідну рівність $r(x)f(x)+S(x)g(x)=d(x)$, де степінь $r(x)$ менший, ніж m, а степінь $S(x)$ менший, ніж n.

Якщо $d(x)=$НСД $(f(x),g(x))$, тоді в рівності (1) розшукуємо многочлени $F(x)$ і $G(x)$ серед многочленів виду

$$F(x)=A_0x^{m-1}+A_1x^{m-2}+...+A_{m-1},$$
$$G(x)=B_0x^{n-1}+B_1x^{n-2}+...+B_{n-1},$$

де A_i, B_k, $i=\overline{0,m-1}; k=\overline{0,n-1}$ – невідомі коефіцієнти, відшукання яких проілюструємо на конкретному прикладі.

Приклад. Подати у вигляді рівності (1) многочлени $f(x)=x^3$ і $g(x)=x^2-3x+2$.

Розв'язання. НСД $(x^3,\ x^2-3x+2)=1$. Отже, має бути

$$F(x)=A_0x+A_1,\ \ G(x)=B_0x^2+B_1x+B_2;$$
$$(A_0x+A_1)x^3+\left(B_0x^2+B_1x+B_2\right)\left(x^2-3x+2\right)=1.$$

Порівнюючи коефіцієнти при однакових степенях правої та лівої частин цієї рівності, дістаємо систему лінійних рівнянь:

$$\left|\begin{array}{l} A_0+\ B_0=\ 0, \\ A_1-3B_0+\ B_1=\ 0, \\ \qquad 2B_0-3B_1+\ B_2=0, \\ \qquad\qquad 2B_1-3B_2=0, \\ \qquad\qquad\qquad 2B_2=1. \end{array}\right.$$

Після розв'язання цієї системи знаходимо, що

$$A_0 = -\frac{7}{8},\ A_1 = \frac{15}{8},\ B_0 = \frac{7}{8},\ B_1 = \frac{3}{4},\ B_2 = \frac{1}{2},$$

$$F(x) = -\frac{7}{8}x + \frac{15}{8},\ G(x) = \frac{7}{8}x^2 + \frac{3}{4}x + \frac{1}{2}.$$

Взаємно прості многочлени. Незвідні многочлени

Означення. *Многочлени* $f(x)$ і $g(x)$ *називаються взаємно простими, якщо їхній НСД дорівнює одиниці.*

Теорема 4. *Для того щоб многочлени* $f(x)$ *і* $g(x)$ *були взаємно простими, необхідно і достатньо існування таких многочленів* $F(x)$ *і* $G(x)$, *для яких виконується рівність*

$$f(x)F(x) + g(x)G(x) = 1. \qquad (2)$$

Доведення теореми аналогічне доведенню відповідної теореми для взаємно простих чисел.

За аналогією з взаємно простими числами наведемо основні властивості взаємно простих многочленів (без доведення).

I. Якщо многочлен взаємно простий з многочленами $f(x)$ і $g(x)$, тоді він взаємно простий і з їх добутком $f(x)\cdot g(x)$.

II. Якщо многочлен $f(x)$ взаємно простий з $g(x)$ і $f(x) \mid g(x)h(x)$, тоді $f(x) \mid h(x)$.

III. Якщо многочлени $f(x)$ і $g(x)$ взаємно прості і $f(x) \mid h(x)$, $g(x) \mid h(x)$, тоді $f(x)g(x) \mid h(x)$.

У кільці $P(x)$ многочленів аналогом простого числа є незвідний многочлен.

Означення. *Многочлен* $f(x)$ *називається незвідним над полем P, якщо його неможливо подати у вигляді добутку многочленів, коефіцієнти яких належали б тому самому полю P, і степені яких були б менші, ніж степінь* $f(x)$.

Основні властивості незвідних многочленів

I. Якщо многочлени $p(x)$ і $q(x)$ – незвідні (над даним полем P), то вони або взаємно прості, або $q(x)=\lambda p(x)$, де λ – сталий множник, відмінний від нуля.

II. Якщо $p(x)$ – незвідний многочлен і $f(x)$ – довільний многочлен, тоді або $p(x)$ і $f(x)$ взаємно прості, або $p(x) \mid f(x)$.

III. Якщо $p(x)$ – незвідний многочлен і $p(x) \mid f(x)g(x)$, тоді або $p(x) \mid f(x)$, або $p(x) \mid g(x)$.

Незвідність многочлена – властивість відносна. Над одним полем він незвідний, над іншим, більш широким, він може бути звідним. Наприклад, многочлен x^2+1 над полем R дійсних чисел – незвідний, але над полем комплексних чисел C він розкладається на множники $(x+i)(x-i)$.

Теорема 5. *Будь-який многочлен $f(x)$ кільця $P[x]$ подається у вигляді добутку незвідних над полем P множників*

$$f(x)=p_1(x)\cdot p_2(x)\cdots p_r(x). \tag{3}$$

Якщо цей многочлен подано у вигляді добутку незвідних множників двома способами:

$$f(x)=p_1(x)\cdot p_2(x)\cdots p_r(x)=q_1(x)\cdot q_2(x)\cdots q_s(x), \tag{4}$$

тоді $r=s$ і при відповідній нумерації множників маємо, що $q_i(x)=\lambda_i p_i(x)$, $i=\overline{1,r}$, де $\lambda_i \in P$, $\lambda_i \neq 0$, для будь-яких $i=\overline{1,r}$.

Доведення. Нехай n – степінь многочлена $f(x)$. Якщо $f(x)$ – незвідний многочлен, тоді розкладання (3) містить усього один множник. Якщо ж $f(x)$ – звідний, то його можна подати у вигляді добутку многочленів із кільця $P(x)$ з меншими степенями. Якщо серед множників знову знайдеться звідний, то і його подамо у вигляді добутку многочленів менших степенів і т. д. Скінченним числом кроків дістанемо розкладання (3).

Другу частину теореми доведемо індукцією за степенем n многочлена $f(x)$. Якщо $n=1$, тоді твердження теореми очевидне. Припустимо, що воно має місце для многочленів, степені яких менші від n. Єдиність розкладання (3) доведемо для $f(x)$.

Нехай виконується рівність (4). З неї випливає, що $q_1(x)$ – дільник $f(x)$, а тому (на підставі властивості III для незвідних многочленів) на $q_1(x)$ ділиться один із множників

$$p_1(x), p_2(x), ..., p_r(x).$$

Наприклад, $q_1(x)$ – дільник $p_1(x)$. Це можливо лише при виконанні рівності $q_1(x) = \lambda_1 p_1(x)$, де λ_1 – ненульовий елемент поля P. Підставимо це подання $q_1(x)$ у праву частину рівності (4). Після спрощення матимемо нову рівність

$$g(x) = p_2(x) \cdots p_r(x) = \lambda_1 q_2(x) q_3(x) \cdots q_s(x).$$

Степінь многочлена $g(x)$ менший, ніж степінь $f(x)$. Використовуючи припущення індукції, дістанемо: $r - 1 = s - 1$, а тому $r = s$ і

$$\lambda_1 q_2(x) = \lambda_2 p_2(x),\ q_3(x) = \lambda_3 p_3(x),\ .....,\ q_r(x) = \lambda_r p_r(x).$$

Звідси маємо необхідні нам рівності

$$q_1(x) = \lambda_1 p_1(x),\ q_2(x) = \lambda_2 p_2(x),$$

$$q_3(x) = \lambda_3 p_3(x),\ ...,\ q_r(x) = \lambda_r p_r(x).$$

Теорему доведено.

На підставі теореми можна записати однозначне розкладання так:

$$f(x) = a_0 \cdot p_1(x) \cdot p_2(x) \cdots p_r(x), \qquad (5)$$

де $p_1(x), p_2(x), \ldots, p_r(x)$ – незвідні множники, у кожного з яких старший коефіцієнт дорівнює одиниці. Будь-які два розкладання цього подання відрізняються лише порядком розміщення множників.

Множники в розкладанні (5) можуть зустрічатися декілька разів. Якщо $p_1(x),\ p_2(x),\ \ldots,\ p_k(x)$ – всі попарно різні множники цього розкладання, то його можна подати у вигляді

$$f(x) = a_0 \cdot p_1^{m_1}(x) \cdot p_2^{m_2}(x) \cdots p_k^{m_k}(x). \qquad (6)$$

Тут $p_i(x) \neq p_j(x)$, якщо $i \neq j$, а числа m_j – кратності відповідних множників.

Звідси напрошується висновок:

Нехай для многочленів $f(x)$ і $g(x)$ має місце розкладання на незвідні множники типу (6). Тоді НСД цих многочленів дорівнює добутку множників, що одночасно містяться в обох розкладаннях. При цьому кожний незвідний спільний множник має степінь, який дорівнює найменшій з його кратностей у розкладаннях $f(x)$ і $g(x)$.

Вправи

1. Довести, що многочлени $f(x)=x^4-x^3-4x^2+4x+1$ і $g(x)=x^2-x-1$ взаємно прості. Спочатку за допомогою алгоритму Евкліда, а потім методом невизначених коефіцієнтів знайти многочлени $F(x)$ і $G(x)$ такі, щоб мала місце рівність (2).

2. За допомогою алгоритму Евкліда знайти НСД многочленів x^m-1 і x^n-1.

3. Нехай $d(x)=$ НСД $(f(x), g(x))$ і $d(x)=f(x)F(x)+$ $+g(x)G(x)$. Чому дорівнює НСД многочленів $F(x)$ і $G(x)$?

4. Довести, що коли многочлен $f(x)$ з цілими коефіцієнтами звідний над полем раціональних чисел, то його можна подати у вигляді добутку многочленів менших степенів з цілими коефіцієнтами (лема Гаусса).

5. Довести, що многочлен x^n+p, де p – просте число, звідний над полем раціональних чисел для будь-якого $n\geq 1$.

6. Переконатися, що многочлени $1+x+\ldots+x^n$ і $1+2x+$ $+3x^2+\ldots+nx^{n-1}$ взаємно прості.

§ 1.6. Многочлени над полями комплексних і дійсних чисел

Нехай $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$ – многочлен з комплексними коефіцієнтами $a_0, a_1, \ldots, a_n$ $(a_0\neq 0)$ і комплексною змінною $x=y+iz$ (y, z – дійсні змінні).

Виведемо для многочлена $f(x)$ формулу Тейлора:

$$f(x+h)=f(x)+h\cdot f'(x)+\frac{h^2}{2!}f''(x)+\ldots+\frac{h^n}{n!}f^{(n)}(x).$$

Обчислюємо похідні многочлена $f(x)$:

$$f'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + ... + 1 \cdot a_{n-1},$$

$$f''(x) = n(n-1)a_0x^{n-2} + (n-1)(n-2)a_1x^{n-3} + ... + 1 \cdot 2 \cdot a_{n-2},$$

....

$$f^{(n-1)}(x) = n(n-1)\cdots 3 \cdot 2 \cdot a_0x + (n-1)(n-2)\cdots 2 \cdot 1 \cdot a_1,$$

$$f^{(n)}(x) = n!a_0.$$

Для будь-якого приросту h маємо

$$f(x+h) = \sum_{k=0}^{n} a_k(x+h)^k = \sum_{k=0}^{n} a_k \left[\sum_{l=0}^{n-k} C_{n-k}^{l} x^{n-k-l} h^l \right] =$$

$$= \sum_{k=0}^{n} \sum_{l=0}^{n-k} C_{n-k}^{l} a_k x^{n-k-l} h^l.$$

У цій подвійній сумі згрупуємо послідовно всі доданки для $k=0$, $k=1$, $k=2$ і т. д. Дістанемо

$$f(x+h) = \sum_{l=0}^{n} C_n^l a_0 x^{n-l} h^l + \sum_{l=0}^{n-1} C_{n-1}^{l} a_1 x^{n-1-l} h^l +$$

$$+ \sum_{l=0}^{n-2} C_{n-2}^{l} a_2 x^{n-2-l} h^l + ... + a_n x^0 h^0.$$

В одержаній рівності приведемо подібні члени відносно h^k, $k = \overline{0, n}$:

$$f(x+h) = h^0 \left[C_n^0 a_0 x^n + C_{n-1}^{0} a_1 x^{n-1} + ... + a_n \right] +$$

$$+ h \left[C_n^1 a_0 x^{n-1} + C_{n-1}^{1} a_1 x^{n-2} + ... + C_1^1 a_{n-1} \right] +$$

$$+ h^2 \left[C_n^2 a_0 x^{n-2} + C_{n-1}^{2} a_1 x^{n-3} + ... + C_2^2 a_{n-2} \right] + ... + h^n C_n^n a_0.$$

Порівнюючи вирази в квадратних дужках останньої рівності з похідними многочлена $f(x)$, остаточно дістанемо формулу Тейлора.

Для многочлена $f(x)$ з комплексною змінною $x = y + zi$ функція $u = |f(x)|$ – дійсна функція двох дійсних змінних y, z.

Функція $f(x)$ є неперервна функція комплексної змінної x: для будь-якого $\varepsilon > 0$ знайдеться дійсне число $\delta > 0$, що з умови $|x_1 - x_2| < \delta$ випливатиме нерівність $|f(x_1) - f(x_2)| < \varepsilon$. Проте для будь-яких комплексних чисел x_1 і x_2 виконується нерівність

$$\big||f(x_1)| - |f(x_2)|\big| \leq |f(x_1) - f(x_2)|$$

Але тоді $\big||f(x_1)| - |f(x_2)|\big| < \varepsilon$. Це й доводить неперервність функції $u = |f(x)|$ від дійсних змінних y, z: $u = u(y, z)$.

На підставі відомої теореми функція $u(y, z)$, яка є неперервною функцією від y, z, у будь-якій обмеженій області має щонайменше одну точку y_0, z_0 мінімальності зі значенням

$$\mu = u(y_0,\ z_0) = |f(x_0)|,\ x_0 = y_0 + z_0 i.$$

***Лема 1**. Якщо $f(x)$ – многочлен степеня $n \geq 1$, k – будь-яке додатне дійсне число, то для достатньо великих за модулем значень змінної x має місце нерівність*

$$k\left|a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n\right| < \left|a_0 x^n\right|. \qquad (1)$$

Доведення. Нехай $a = \max\left(|a_1|, |a_2|, \ldots, |a_n|\right)$.

Тоді

$$\left|a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n\right| \leq a\left(|x|^{n-1} + |x|^{n-2} + \ldots + 1\right) = a\frac{|x|^n - 1}{|x| - 1}.$$

Якщо

$$|x| > \frac{ka}{|a_0|} + 1, \qquad (2)$$

тоді

$$|x| - 1 > \frac{ka}{|a_0|} \Rightarrow |a_0| > ka\frac{1}{|x| - 1} \Rightarrow$$

$$\Rightarrow |a_0||x|^n > ka\frac{|x|^n}{|x| - 1} \Rightarrow a\frac{|x|^n - 1}{|x| - 1} < \frac{1}{k}|a_0||x|^n.$$

На підставі отриманої вище нерівності

$$\left|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n\right| \le a\frac{|x|^n - 1}{|x| - 1}$$

випливає нерівність

$$\left|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n\right| < \frac{1}{k}|a_0||x|^n,$$

якщо виконується нерівність (2). Отже, виконується і нерівність (1).

Лему доведено.

Наслідок. *Якщо в лемі* 1 $k=2$, *тоді при*

$$|x| > \frac{2a}{|a_0|} + 1 \tag{3}$$

виконується

$$2\left|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n\right| < |a_0|\,|x|^n. \tag{4}$$

Лема 2. *Якщо виконується нерівність (3), тоді для* $f(x)$ *виконується*

$$|f(x)| > \frac{1}{2}|a_0||x|^n. \tag{5}$$

Доведення. На підставі наслідку леми 1 із нерівності (3) випливає нерівність

$$\left|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n\right| < \frac{1}{2}|a_0||x|^n.$$

Тоді

$$|f(x)| = \left|a_0x^n + \left(a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n\right)\right| \ge$$

$$\ge \left|a_0x^n\right| - \left|a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n\right| >$$

$$> \left|a_0x^n\right| - \frac{1}{2}|a_0||x|^n = \frac{1}{2}|a_0||x|^n.$$

Лему доведено.

Теорема (основна теорема вищої алгебри)*. *Будь-яке алгебричне рівняння*

$$f(x) = a_0x^n + a_1x^{n-1} + \ \ldots \ + a_{n-1}x + a_n = 0,$$

коефіцієнти якого – дійсні або комплексні числа, має принаймні один корінь $x_0 = y_0 + z_0 i$, *де* $y_0, z_0 \in R$.

Доведення. Припустимо, що $a_0 \neq 0$. Обираємо додатне число R_0 настільки великим, що

$$R_0 > A = \frac{2a}{|a_0|} + 1 \ \text{ і } \ \frac{1}{2}|a_0| R_0^n > |a_n|, \tag{6}$$

де $a = \max\left(|a_1|, |a_2|, \ldots, |a_n|\right)$.

Тоді на підставі леми 2 в точках x кола $|x| = R_0$ матимемо

$$|f(x)| > |f(0)| = |a_n|.$$

Тому на крузі $K(R_0) = \left\{x : |x| \leq R_0\right\}$ функція $u = \left|f\left(x\right)\right|$ набуває мінімального значення μ у деякій точці $x_0 = y_0 + z_0 i$, яка є внутрішньою точкою цього круга ($\mu \leq |f(0)|$).

Якщо $\mu = 0$, тоді теорему доведено, оскільки $f(x_0) = 0$.

Припустимо, що $\mu > 0$ і $\mu = |f(x_0)|$.

В останньому випадку з'ясуємо, що існує число $h \neq 0$, як завгодно мале за модулем, таке, що $|f(x_0 + h)| < |f(x_0)|$ (твердження відоме як лема Д'Аламбера).

Нехай h – деяке комплексне число. Застосуємо теорему Тейлора до функції f:

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!} f''(x_0) + \ldots + \frac{h^n}{n!} f^{(n)}(x_0) \,. \tag{7}$$

Припустимо, що першим коефіцієнтом, відмінним від нуля, є коефіцієнт при h^m, тобто $f^{(m)}(x_0) \neq 0 \ (m \geq 1)$. Поділивши праву і ліву частини рівності (7) на $f(x_0)$, дістанемо

---

*Строге доведення вперше дав Гаусс (1799 р.). Ця теорема раніше доводилась Д'Аламбером, але з недостатньою повнотою. Наведене доведення належить Коші (1821 р.).

$$\frac{f(x_0+h)}{f(x_0)}=1+\frac{h^m f^{(m)}(x_0)}{m!f(x_0)}+\ldots+\frac{h^n f^{(n)}(x_0)}{n!f(x_0)}.$$

Подамо кожен доданок правої частини цього розкладання в комплексній формі:

$$\frac{f^{(m)}(x_0)}{m!f(x_0)}=R_m(\cos\varphi_m+i\sin\varphi_m),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{f^{(n)}(x_0)}{n!f(x_0)}=R_n(\cos\varphi_n+i\sin\varphi_n),\ h=\rho\left(\cos\varphi+i\sin\varphi\right),$$

де $R_m,\ldots,R_n,\rho$ – модулі відповідних чисел, а $\varphi_m,\ldots,\varphi_n,\varphi$ – їх аргументи.

Тоді

$$\frac{f(x_0+h)}{f(x_0)}=1+R_m\rho^m\left[\cos\left(m\varphi+\varphi_m\right)+i\sin\left(m\varphi+\varphi_m\right)\right]+ \\ +\ldots+R_n\rho^n\left[\cos\left(n\varphi+\varphi_n\right)+i\sin\left(n\varphi+\varphi_n\right)\right]. \quad (8)$$

Оберемо аргумент φ таким, щоб $m\varphi+\varphi_m=\pi$, тобто $\varphi=\frac{\pi-\varphi_m}{m}$, а модуль ρ настільки малим, щоб $R_m\rho^m<1$. Із рівності (8) перші два доданки дають число $1-R_m\rho^m>0$. Тоді на підставі властивостей модуля суми з рівності (8) дістанемо нерівність

$$\left|\frac{f(x_0+h)}{f(x_0)}\right|\le 1-R_m\rho^m+R_{m+1}\rho^{m+1}+\ldots+R_n\rho^n,$$

або

$$\left|\frac{f(x_0+h)}{f(x_0)}\right|\le 1-\rho^m(R_m-R_{m+1}\rho-\ldots-R_n\rho^{n-m}).$$

Зрозуміло, що сума в дужках при достатньо малому ρ буде близькою до числа R_m, а вся права частина – близькою до числа $1-R_m\rho^m$ і, отже, буде додатною. Тому $|f(x_0+h)|<|f(x_0)|$. Можна вважати, що $x_0+h\in K(R_0)$. Прийшли до суперечності. Тим самим теорему доведено.

Наслідки основної теореми

Нехай $f(x)=a_0x^n+a_1x^{n-1}+...+a_n$ – многочлен з коефіцієнтами $a_0,a_1,...,a_n$ з поля C комплексних чисел, $a_0\neq 0, n\geq 1$. За основною теоремою $f(x)$ має корінь α_1, що належить полю C. Тоді

$$f(x)=(x-\alpha_1)f_1(x),$$

де $f_1(x)$ – многочлен степеня $n-1$ з коефіцієнтами з поля C. Якщо степінь $f_1(x)$ більший за 0, тоді на тій же підставі існує корінь α_2 і, отже, $f_1(x)=(x-\alpha_2)f_2(x)$. Тому

$$f(x)=(x-\alpha_1)(x-\alpha_2)f_2(x).$$

Продовжуючи цю процедуру, на n-му кроці дістанемо розкладання

$$f(x)=(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n)b,$$

де $b\in C$. Неважко переконатися, що старший коефіцієнт в останньому розкладанні дорівнює bx^n. Порівнюючи його зі старшим членом a_0x^n у початковому розкладанні, дістанемо $b=a_0$. Тому

$$f(x)=a_0(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_n). \tag{9}$$

Зрозуміло, що всі числа $\alpha_1,\alpha_2,...,\alpha_n$ є коренями $f(x)$. З другого боку, нехай α – корінь $f(x)$. Покладемо $x=\alpha$ у правій і лівій частинах (9). Дістанемо

$$0=a_0(\alpha-\alpha_1)(\alpha-\alpha_2)\cdots(\alpha-\alpha_n).$$

Звідси випливає, що для деякого k множник $\alpha-\alpha_k=0\Rightarrow$ $\Rightarrow\alpha=\alpha_k$, оскільки $a_0\neq 0$.

Наслідок 1. *Будь-який многочлен $f(x)=a_0x^n+a_1x^{n-1}+$ $+...+a_n$, $a_0\neq 0$ над полем комплексних чисел можна подати у вигляді добутку (9), де $\alpha_1,\alpha_2,...,\alpha_n$ – корені цього многочлена. Інших коренів він не має.*

Наслідок 2. *Незвідними многочленами над полем комплексних чисел є лише многочлени степеня 1.*

На підставі теореми 5, § 1.5 розкладання (5) на незвідні множники є єдиним з точністю до порядку розташування множників.

У цьому розкладанні об'єднаємо всі однакові множники. Тоді його можна записати у вигляді

$$f(x) = a_0(x-\alpha_1)^{m_1}(x-\alpha_2)^{m_2} \cdots (x-\alpha_k)^{m_k}, \qquad (10)$$

де $\alpha_1, \alpha_2, ..., \alpha_k$ – усі попарно різні корені многочлена і m_j – кратність кореня α_j, причому $m_1 + m_2 + ... + m_k = n$.

Наслідок 3. *Будь-який многочлен степеня $n \geq 1$ з комплексними (зокрема, дійсними) коефіцієнтами має рівно n коренів, якщо кожний з коренів рахувати стільки разів, скільки його кратність.*

Розглянемо тепер многочлен, старший коефіцієнт якого дорівнює 1, тобто

$$f(x) = x^n + a_1x^{n-1} + ... + a_n.$$

Запишемо для нього розкладання (9):

$$f(x) = (x-\alpha_1)(x-\alpha_2) \cdots (x-\alpha_n),$$

причому кожний корінь взятий стільки разів, скільки його кратність.

Розкривши дужки та звівши подібні члени після порівняння правих частин в обох розкладах $f(x)$, приходимо до заключного наслідку.

Наслідок 4 (формули Вієта). *Між коефіцієнтами і коренями многочлена $f(x)$ зі старшим коефіцієнтом 1 мають місце співвідношення*

$$a_1 = -(\alpha_1 + \alpha_2 + ... + \alpha_n),$$

$$a_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + ... + \alpha_i\alpha_k + ... + \alpha_{n-1}\alpha_n,$$

....

$$a_k = (-1)^k \left(\sum \alpha_{i_1} \cdot \alpha_{i_2} \cdots \alpha_{i_k}\right).$$

....

$$a_n = (-1)^n \alpha_1\alpha_2 \cdots \alpha_n.$$

У формулі для коефіцієнта a_k під знаком суми містяться різні добутки за сполученням по k коренів – усього C_n^k доданків.

Зокрема, для многочленів 3-го степеня $f(x)=x^3+a_1x^2+$ $+a_2x+a_3$ з коренями α_1, α_2, α_3 формули Вієта мають вигляд

$$
\begin{aligned}
a_1 &= -(\alpha_1+\alpha_2+\alpha_3),\\
a_2 &= \alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3,\\
a_3 &= -\alpha_1\cdot\alpha_2\cdot\alpha_3.
\end{aligned}
$$

Приклад 1. Корені многочлена x^n-1 є

$$\zeta_k=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n},\ k=\overline{0,n-1}.$$

На підставі формул Вієта знаходимо:

$$\zeta_0+\zeta_1+\zeta_2+\ldots+\zeta_{n-1}=0,\ \zeta_0\cdot\zeta_1\cdot\zeta_2\cdots\zeta_{n-1}=(-1)^{n-1}.$$

Приклад 2. Знайти НСД многочленів x^m-1 і x^n-1 методом розкладання їх на незвідні множники.

Розв'язання. Подамо x^m-1 у вигляді добутку незвідних множників над полем C:

$$x^m-1=(x-\omega_0)(x-\omega_1)\cdots(x-\omega_{m-1}),$$

де

$$\omega_k=\cos\frac{2\pi k}{m}+i\sin\frac{2\pi k}{m},\ k=\overline{0,m-1}.$$

Аналогічно

$$x^n-1=(x-\zeta_0)(x-\zeta_1)\cdots(x-\zeta_{n-1}),$$

де

$$\zeta_l=\cos\frac{2\pi l}{n}+i\sin\frac{2\pi l}{n},\ l=\overline{0,n-1}.$$

Тоді НСД многочленів з умови дорівнює добутку спільних множників у першому і другому розкладаннях. Задача зводиться до відшукання спільних коренів $\sqrt[m]{1}$ і $\sqrt[n]{1}$.

Нехай $\omega_k = \zeta_l$, тобто $\cos\frac{2\pi k}{m} + i\sin\frac{2\pi k}{m} = \cos\frac{2\pi l}{n} + i\sin\frac{2\pi l}{n}$.

Тоді $\frac{k}{m} = \frac{l}{n} + N$, N – ціле число. Числа $\frac{k}{m}$ і $\frac{l}{n}$ містяться у напівінтервалі $[0, 1)$, а тому $N = 0$.

Нехай $d =$ НСД (m, n). Тоді $m = dm_0$, $n = dn_0$, причому m_0 і n_0 – числа взаємно прості. Маємо рівність $kn_0 = lm_0$, а тому $m_0 \mid k$, $n_0 \mid l$ і, отже, $k = m_0 K$, $l = n_0 L$. Але тоді

$$\omega_k = \cos\frac{2\pi K}{d} + i\sin\frac{2\pi K}{d},\ \zeta_l = \cos\frac{2\pi L}{d} + i\sin\frac{2\pi L}{d}$$

корені многочлена $x^d - 1$. Обернене твердження також має місце: кожний корінь многочлена $x^d - 1$ – спільний корінь многочленів $x^m - 1$ і $x^n - 1$.

Отже, НСД дорівнює $x^d - 1$, де $d =$ НСД (m, n).

Многочлени з дійсними коефіцієнтами

Нехай R – поле дійсних чисел.

Лема 3. *Якщо многочлен* $f(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n \in R[x]$ *і* $\alpha = a + bi$ *– його комплексний корінь, тоді спряжене число* $\overline{\alpha} = a - bi$ *також буде коренем цього многочлена.*

Доведення. Нехай спочатку $\alpha = a + bi$ – довільне комплексне число. Скористаємось властивостями комплексно спряжених чисел і дістанемо таку послідовність рівностей:

$$\overline{f(\alpha)} = \overline{a_0\alpha^n + a_1\alpha^{n-1} + \ldots + a_{n-1}\alpha + a_n} =$$

$$= \overline{a_0\alpha^n} + \overline{a_1\alpha^{n-1}} + \ldots + \overline{a_{n-1}\alpha} + \overline{a_n} = a_0(\overline{\alpha})^n + a_1(\overline{\alpha})^{n-1} + \ldots + a_{n-1}\overline{\alpha} + a_n.$$

Отже, $\overline{f(\alpha)} = f(\overline{\alpha})$. Зокрема, якщо α – корінь $f(x)$, тоді $0 = \overline{f(\alpha)} = f(\overline{\alpha})$. Таким чином, $\overline{\alpha}$ – корінь того самого многочлена.

Лему доведено.

Наслідок 5. *Кожний многочлен з дійсними коефіцієнтами можна подати у вигляді добутку многочленів першого або другого степеня. При цьому множники другого степеня мають від'ємні дискримінанти.*

Доведення. Нехай $f(x)=a_0x^n+a_1x^{n-1}+...+a_n \in R[x]$ і $n \geq 1$, $\alpha_1, \alpha_2, ..., \alpha_n$ – його корені. Тоді $f(x)=a_0(x-\alpha_1)\times$ $\times(x-\alpha_2)\cdots(x-\alpha_n)$. Вважатимемо, що дійсними коренями, якщо вони взагалі тут є, будуть перші m коренів $\alpha_1, \alpha_2, ..., \alpha_m$. Комплексні корені $\alpha_{m+1}, ..., \alpha_n$ розбиваються на пари спряжених чисел. Розмістимо ці пари послідовно: $\alpha_{m+1}=a_1+b_1i$, $\alpha_{m+2}=a_1-b_1i$ і т. д. Якщо l – число таких пар коренів, тоді $n=m+2l$. Маємо

$$f(x)=a_0(x-\alpha_1)\cdots(x-\alpha_m)(x-a_1-b_1i)\times$$
$$\times(x_1-a_1+b_1i)\cdots(x-a_l-b_li)(x-a_l+b_li).$$

Знайдемо $(x-a_1-b_1i)(x-a_1+b_1i)=x^2-2a_1x+a_1^2+b_1^2$ і позначимо $p_1=-2a_1$, $q_1=a_1^2+b_1^2$. Тоді

$$(x-a_1-b_1i)(x-a_1+b_1i)=x^2+p_1x+q_1,\ D_1=p_1^2-4q_1<0,$$

де D_1 – дискримінант тричлена $x^2+p_1x+q_1$. Пробивши таке перетворення і для решти комплексних множників, остаточно дістанемо необхідне нам розкладання:

$$f(x)=a_0(x-\alpha_1)\cdots(x-\alpha_m)(x^2+p_1x+q_1)\cdots(x^2+p_lx+q_l), \quad (11)$$

причому $D_j=p_j^2-4q_j<0\ (j=\overline{1,l})$, $2l+m=n$.

Якщо врахувати присутність однакових множників, тоді дістанемо

$$f(x)=a_0(x-\alpha_1)^{m_1}\cdots(x-\alpha_r)^{m_r}\times$$
$$\times(x^2+p_1x+q_1)^{n_1}\cdots(x^2+p_sx+q_s)^{n_s}, \quad (12)$$

де $x-\alpha_1, ..., x-\alpha_r, x^2+p_1x+q_1, ..., x^2+p_sx+q_s$, – попарно відмінні один від одного дійсні множники. Числа $m_1, ..., m_r$,

$n_1, \ldots, n_s$ називаються кратностями цих множників, серед яких $m_1, \ldots, m_r$ – кратності дійсних коренів.

Наслідок 6. *Над полем R дійсних чисел незвідними многочленами будуть лише многочлени 1-го і 2-го степеня з від'ємними дискримінантами.*

Наслідок 7. *Кожний многочлен $f(x) \in R[x]$ непарного степеня має хоч би один дійсний корінь.*

Нехай n – степінь многочлена $f(x)$, m – число його дійсних коренів. Тоді, як відомо, $n = m + 2l$. Якщо n – число непарне, тоді і m – число непарне.

Вправи

Розкласти над полем дійсних чисел на незвідні множники такі многочлени:

1. $x^6 + 1$.

2. $x^4 - ax^2 + 1, \ -2 < a < 2$.

3. $x^{2n} + x^n + 1$.

4. $T_n(x) = \cos n(\arccos x)$ – многочлен Чебишова.

Вказівка. Подати $\cos n\varphi$ у вигляді степенів $\cos\varphi$.

Відокремлення кратних множників

Нехай P – довільне поле чисел (підполе поля C). Будь-який многочлен $f(x) \in P[x]$ можна подати у вигляді добутку

$$f(x) = a_0 p_1^{m_1}(x) p_2^{m_2}(x) \cdots p_k^{m_k}(x),$$

де $p_1(x), p_2(x), \ldots, p_k(x)$ – попарно відмінні один від одного незвідні множники зі старшим коефіцієнтом 1, а $m_1, m_2, \ldots, m_k$ – кратності цих множників.

Нехай $p(x)$ – один із множників з кратністю m входить у це розкладання. Тоді

$$f(x) = p^m(x) g(x),$$

причому многочлен $g(x)$ взаємно простий з $p(x)$. Зокрема, якщо $p(x) = x - \alpha$, тоді α – корінь $f(x)$ кратності m.

Лема 5. Якщо $p(x)$ – незвідний дільник многочлена $f(x)$ кратності m, тоді $p(x)$ міститься в розкладанні на незвідні множники похідної $f'(x)$ з кратністю $m-1$.

Доведення. Із рівності $f(x)=p^m(x)g(x)$ випливає, що

$$f'(x)=p^{m-1}(x)\left[mp'(x)g(x)+p(x)g'(x)\right].$$

Нехай

$$\varphi(x)=mp'(x)g(x)+p(x)g'(x),$$

причому $d(x)$ = НСД $\left(p(x),\varphi(x)\right)$. Оскільки $p(x)$– незвідний многочлен, то $d(x)=p(x)$ або $d(x)=1$.

Припустимо, що $d(x)=p(x)$. Тоді $p(x)|\varphi(x)$ і, отже, $p(x)|p'(x)g(x)$. Оскільки $p(x)$ і $g(x)$ взаємно прості, то $p(x)|p'(x)$, що неможливо. Отже, $d(x)=1$ і многочлени $p(x)$ і $\varphi(x)$ – взаємно прості.

Лему доведено.

Наведемо тепер метод відшукання кратних множників для многочленів з коефіцієнтами із поля R дійсних чисел. Нехай $f(x)$ – такий многочлен і

$$f(x)=a_0(x-\alpha_1)^{m_1}\cdots(x-\alpha_r)^{m_r}(x^2+p_1x+q_1)^{n_1}\cdots -$$

його розкладання на незвідні множники.

Позначимо через X_1 добуток усіх незвідних множників цього розкладання, що входять з кратністю 1; через X_2 – добуток усіх тих незвідних множників, які входять з кратністю 2; через X_3 – відповідно з кратністю 3 і т. д. Може трапитись, що в розкладанні $f(x)$ відсутні множники якої-небудь кратності. Наприклад, відсутні незвідні множники кратності 1. Тоді покладаємо $X_1=1$.

Врешті-решт дістаємо, наприклад, розкладання у вигляді

$$f(x)=a_0X_1X_2^2X_3^3X_4^4X_5^5.$$

На підставі леми знаходимо

$$D = \text{НСД}\ (f, f') = X_2 X_3^2 X_4^3 X_5^4,$$

$$D_1 = \text{НСД}\ (D, D') = X_3 X_4^2 X_5^2,$$

$$D_2 = \text{НСД}\ (D_1, D_1') = X_4 X_5^2,$$

$$D_3 = \text{НСД}\ (D_2, D_2') = X_5.$$

Звідси

$$E_1 = \frac{f(x)}{D} = a_0 X_1 X_2 X_3 X_4 X_5,\ E_2 = \frac{D}{D_1} = X_2 X_3 X_4 X_5.$$

$$E_3 = \frac{D_1}{D_2} = X_3 X_4 X_5,\ E_4 = \frac{D_2}{D_3} = X_4 X_5,\ E_5 = D_5 = X_5$$

і

$$a_0 X_1 = \frac{E_1}{E_2},\ X_2 = \frac{E_2}{E_3},\ X_3 = \frac{E_3}{E_4},\ X_4 = \frac{E_4}{E_5},\ X_5 = E_5.$$

Такий процес відокремлення кратних множників дає змогу в багатьох випадках подати многочлен у вигляді добутку многочленів більш простого виду.

Приклад. Подати многочлен $f(x) = x^5 - 3x^4 + 4x^3 - 4x^2 + 3x - 1$ у вигляді добутку незвідних многочленів.

Розв'язання. 1) знаходимо НСД $(f(x), f'(x)) = D$ за допомогою алгоритму Евкліда. Знаходимо $D = x^2 - 2x + 1$;

2) знаходимо НСД $(D, D') = D_1 = x - 1$;

3) знаходимо НСД $(D_1, D_1') = 1 = D_2$;

4) обчислюємо E_1, E_2, E_3:

$$E_1 = \frac{f}{D} = x^3 - x^2 + x - 1;\ E_2 = \frac{D}{D_1} = x - 1;\ E_3 = \frac{D_1}{D_2} = x - 1;$$

5) обчислюємо X_1, X_2, X_3:

$$X_1 = \frac{E_1}{E_2} = x^2 + 1;\ X_2 = \frac{E_2}{E_3} = 1;\ X_3 = E_3 = x - 1.$$

Відповідь: $f(x) = (x^2 + 1)(x - 1)^3$.

§ 1.7. Теорема Ролля

Теорема 1. *При достатньо великих абсолютних значеннях дійсної змінної x многочлен $f(x)$ з дійсними коефіцієнтами зберігає знак свого старшого члена.*

Доведення. Нехай $f(x)=a_0x^n+a_1x^{n-1}+...+a_n$, $a_0\neq 0$. Позначимо $\varphi(x)=a_1x^{n-1}+...+a_{n-1}x+a_n$ і $a=\max\{|a_1|,|a_2|,...,|a_n|\}$. Тоді

$$|\varphi(x)|\le a\left(|x|^{n-1}+...+|x|+1\right)=a\frac{|x|^n-1}{|x|-1}.$$

Отже,

$$|a_0x^n|-a\frac{|x|^n-1}{|x|-1}=\frac{|x|^n\{|a_0||x|-|a_0|-a\}+a}{|x|-1}.$$

Якщо має місце нерівність $|a_0||x|-|a_0|-a>0$, тобто $|x|>1+\dfrac{a}{|a_0|}$, то чисельник і знаменник у правій частині останньої рівності додатні. Тому

$$|f(x)|\ge|a_0x^n|-|\varphi(x)|\ge|a_0x^n|-a\frac{|x|^n-1}{|x|-1}>0,$$

якщо $|x|>1+\dfrac{a}{|a_0|}$. Звідси випливає, що при $x>1+\dfrac{a}{|a_0|}$ знак визначається знаком його старшого члена a_0x^n.

Теорему доведено.

Наслідок. *Число $A=1+\dfrac{a}{|a_0|}$ може бути верхньою оцінкою (межею) модулів дійсних коренів многочлена $f(x)$. Усі такі корені містяться на відрізку $[-A,A]$.*

Приклад. Для многочлена $f(x)=3x^4-6x^3+x^2-2x+4$ маємо $a_0=3$, $a=\max\{|-6|,1,|-2|,4\}=6$. Тоді $A=3$ і всі його дійсні корені належать відрізку [–3, 3].

Теорема 2. *При достатньо малих значеннях $|x|$ многочлен $f(x)$ зберігає знак нижчого члена.*

Доведення. Нехай $f(x)=a_0x^m+a_1x^{m+1}+...+a_nx^{m+n}$, $m,n\in N$. Покладемо $x=\frac{1}{y}$. Тоді $F(y)=y^{m+n}f\left(\frac{1}{y}\right)=a_0y^n+a_1y^{n-1}+...+a_n$. Для достатньо великих $N>0$ і за умови $|y|>N$ маємо

$$|a_0y^n|>|a_1y^{n-1}+...+a_{n-1}y+a_n|.$$

Тому

$$|a_px^m|>|a_1x^{m+1}+\ ...\ +a_nx^{m+n}|\ \text{при}\ |x|<\frac{1}{N}.$$

Теорему доведено.

Теорема 3. *Якщо $f'(x_0)=0$ і якщо x, зростаючи, проходить через x_0, тоді многочлен $f(x)$ зростає при $f'(x_0)>0$ і спадає при $f'(x_0)<0$.*

Доведення. Із формули Тейлора для многочлена $f(x)=$ $=a_0x^n+a_1x^{n-1}+...+a_n$, де $a_j\in R$, $j=\overline{0,\ n}$,

$$f(x+h)=f(x)+\frac{h}{1!}f'(x)+\frac{h^2}{2!}f''(x)+...+\frac{h^n}{n!}f^{(n)}(x) \qquad (1)$$

при $x=x_0$ маємо:

$$f(x_0+h)-f(x_0)=hf'(x_0)+\frac{h^2}{2!}f''(x_0)+... \qquad (2)$$

При достатньо малих h знак правої частини визначається знаком її нижчого члена $hf'(x_0)$. Тому при $f'(x_0)>0$ різниця $f(x_0+h)-f(x_0)$ від'ємна при $h<0$ і додатна при $h>0$. Отже, функція $f(x)$ зростає поблизу x_0. Аналогічно, при $f'(x_0)<0$ вона спадає.

Теорему доведено.

Теорема 4. *Якщо змінна x, зростаючи, проходить через корінь многочлена $f(x)$ з дійсними коефіцієнтами, то $\frac{f(x)}{f'(x)}$ проходить при цьому від від'ємних до додатних значень.*

Доведення. Нехай x_0 – корінь многочлена $f(x)$ і m – його кратність. Тоді x_0 – корінь похідної $f'(x)$ і кратність його дорівнює $m-1$. Із формули Тейлора матимемо, що

$$f(x_0+h)=\frac{h^m}{m!}f^{(m)}(x_0)+\frac{h^{m+1}}{(m+1)!}f^{(m+1)}(x_0)+\ldots,$$

$$f'(x_0+h)=\frac{h^{m-1}}{(m-1)!}f^{(m)}(x_0)+\ldots$$

і тому

$$\frac{f(x_0+h)}{f'(x_0+h)}=h\frac{\frac{1}{m!}f^{(m)}(x_0)+\frac{h}{(m+1)!}f^{(m+1)}(x_0)+\ldots}{\frac{1}{(m-1)!}f^{(m)}(x_0)+\frac{h}{m!}f^{(m+1)}(x_0)+\ldots}.$$

При достатньо малих значеннях h дріб, який міститься у правій частині рівності, має знак, що визначається першими доданками – в чисельнику доданком $\frac{1}{m!}f^{(m)}(x_0)$, а в знаменнику $\frac{1}{(m-1)!}f^{(m)}(x_0)$. Але обидва вони одного знака. Тому дріб, який є множником після h, додатний (при малих h). Звідси випливає $\frac{f(x_0+h)}{f'(x_0+h)}<0$ при $h<0$ і $\frac{f(x_0+h)}{f'(x_0+h)}>0$ при $h>0$, що й доводить теорему.

Теорема 5 (Ролля). *Між двома послідовними коренями многочлена міститься у крайньому випадку один корінь похідної.*

Доведення. Нехай a і b – два послідовних корені многочлена $f(x)$, причому $a<b$. Якщо h – достатньо мале додатне число, то з попередньої теореми випливає, що

$$\frac{f(a+h)}{f'(a+h)}>0,\ \frac{f(b-h)}{f'(b-h)}<0.$$

Між $a+h$ і $b-h$ немає коренів многочлена, і тому числа $f(a+h)$ і $f(b-h)$ одного знака. Тому $f'(a+h)$ і $f'(b-h)$ різних знаків і, отже, похідна $f'(x)$ має корені в проміжку від $a+h$ до $b-h$.

Тут використано властивість неперервних функцій, яка стверджує, що коли на кінцях відрізка така функція має *різні* знаки, то на проміжку між кінцями відрізка вона обертається на нуль.

Вправи

1. За допомогою теореми Ролля показати, що всі корені многочлена Лежандра

$$P_n(x) = \frac{1}{n!2^n}\frac{d^n}{dx^n}\left[\left(x^2-1\right)\right]^n$$

дійсні, відмінні один від одного і розміщені в інтервалі $(-1, 1)$.

Вказівка. Розглянути многочлен $f(x)=(x^2-1)^n$ і його похідні $f'(x), f''(x), ..., f^n(x)$, а потім послідовно застосувати до них теорему Ролля.

2. Показати, що всі корені многочлена Лагерра

$$L_n(x) = (-1)^n e^x \frac{d^n}{dx^n}\left(x^n e^{-x}\right)$$

дійсні, відмінні один від одного і містяться в інтервалі $(0, \infty)$.

3. Показати, що всі корені многочлена Ерміта

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right)$$

дійсні і відмінні один від одного.

4. Показати, що многочлен $x^{2m+1}+x^{2n+1}+x^{2p+1}+a$ має лише один дійсний корінь.

§ 1.8. Правило знаків Декарта

Нехай $a_0, a_1, ..., a_n$ – послідовність дійсних чисел. Номер $m\ (0<m\le n)$ називається місцем зміни знака цієї послідовності, якщо при $a_m \ne 0$ і $a_{m-1}\ne 0 \Rightarrow a_{m-1}a_m<0$ і при $a_{m-1}=a_{m-2}=...=a_{m-k+1}=0$, $a_{m-k}\ne 0 \Rightarrow a_{m-k}a_m<0$. В обох випадках пари a_{m-1}, a_m і a_{m-k}, a_m утворюють зміну знака.

При підрахуванні числа змін знака в послідовності нулями можна нехтувати.

Розглянемо деякі властивості числових послідовностей, пов'язаних з числом змін знака.

I. Нехай дано послідовність

$$a_0, a_1, a_2, ..., a_n \tag{1}$$

дійсних чисел і $p_0 > 0, p_1 > 0, ..., p_n > 0$. Тоді послідовність

$$p_0 a_0, p_1 a_1, p_2 a_2, ..., p_n a_n \tag{2}$$

має ті ж місця змін знаків, що й послідовність (1).

II. Нехай у послідовності (1) $a_0 \neq 0$ і $a_n \neq 0$. Тоді число W змін її знака буде парним, якщо a_0 і a_n мають однакові знаки, і непарним, якщо a_0 і a_n мають протилежні знаки.

Уся послідовність (1) розбивається на $W+1$ підпослідовностей: перша з них починається з a_0 і складається з чисел зі знаком $\eta = \operatorname{sign} a_0$ і закінчується числом a_{m_1-1}, де m_1 – перше місце зміни знака; друга послідовність починається з числа a_{m_1} і складається з чисел зі знаком $\operatorname{sign} a_{m_1}$ і закінчується числом a_{m_2-1}, де m_2 – друге місце зміни знака і т. д. Остання серед них з номером $W+1$ складається з чисел зі знаком $\operatorname{sign} a_n$. Цей знак збігається з $\operatorname{sign} a_0$, якщо $W+1$ – непарне, тобто W – парне число. І він збігається з $\operatorname{sign} a_0$, якщо $W+1$ – число парне, тобто W – непарне число.

III. Припустимо, що в послідовності (1) єдиними місцями зміни знака є 1 і n за умови $n > 1$. Тоді число змін знака в послідовності

$$a_1 - a_0, a_2 - a_1, a_3 - a_2, ..., a_n - a_{n-1} \tag{3}$$

непарне.

Оскільки a_0, a_1 утворюють зміну знака, то $\operatorname{sign}(a_1 - a_0) = \operatorname{sign} a_1$. Аналогічно $\operatorname{sign}(a_n - a_{n-1}) = \operatorname{sign} a_n$. Разом з тим, змін знака в послідовності $a_1, a_2, ..., a_{n-1}$ немає. Тому $\operatorname{sign} a_1 = -\operatorname{sign} a_n$

і, отже, $\text{sign}(a_1 - a_0) = -\text{sign}(a_n - a_{n-1})$. На підставі властивості II число змін знака в послідовності (3) – непарне.

IV. Нехай послідовність (1) має W змін знаків. Тоді різницева послідовність (3) матиме не менше $W-1$ змін знаків.

Припустимо, що $m_1, m_2, ..., m_W$ – місця зміни знака в послідовності (1), причому $m_1 < m_2 < ... < m_W$. Розглянемо такі послідовності чисел із послідовності (1):

$$\overline{A_1} : a_{m_1-1}, a_{m_1}, ..., a_{m_2-1}, a_{m_2};$$

$$\overline{A_2} : a_{m_2-1}, a_{m_2}, ..., a_{m_3-1}, a_{m_3};$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\overline{A_{W-1}} : a_{m_{W-1}-1}, a_{m_W}, ..., a_{m_W-1}, a_{m_W}.$$

Кожна з них має рівно дві зміни знака на початку і в кінці. Тому за властивістю II в кожній із послідовностей

$$A_1 : \quad a_{m_1} - a_{m_1-1}, \quad a_{m_1+1} - a_{m_1}, \quad ..., a_{m_2} - a_{m_2-1},$$

$$A_2 : \quad a_{m_2} - a_{m_2-1}, \quad a_{m_2+1} - a_{m_2}, \quad ..., a_{m_3} - a_{m_3-1};$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$A_{W-1} : a_{m_{W-1}} - a_{m_{W-1}-1}, a_{m_{W-1}+1} - a_{m_{W-1}}, ..., a_{m_W} - a_{m_W-1}$$

є у крайньому випадку одна зміна знака. Тому в послідовності (3) є не менше, ніж $W - 1$ змін знаків.

V. Якщо в ненульовій послідовності (1) є W змін знаків, тоді в послідовності

$$a_0, a_1 - a_0, a_2 - a_1, ..., a_n - a_{n-1}, -a_n \tag{4}$$

є не менше, ніж $W+1$ змін знаків.

Окрім послідовностей $A_1, A_2, ..., A_{W-1}$ із послідовності (4) побудуємо ще дві послідовності

$$A_0 : a_0, a_1 - a_0, ..., a_{m_1} - a_{m_1-1};$$

$$A_W : a_{m_W} - a_{m_W-1}, ..., a_n - a_{n-1}, -a_n.$$

Оскільки в послідовності $a_0, a_1, ..., a_{m_1-1}, a_{m_1}$ є лише одна зміна знака (на її кінці) і $\text{sign}\, a_{m_1} = \text{sign}(a_{m_1} - a_{m_1-1})$, то $\text{sign}(a_{m_1} - a_{m_1-1}) =$

$= -\operatorname{sign} a_0$. А тому в послідовності A_1, згідно з властивістю II, є у крайньому випадку одна зміна знака. Аналогічно в послідовності A_W є хоч би одна зміна знака. Звідси випливає, що в послідовності (4) є не менше, ніж $W+1$ змін знаків.

VI. Якщо в послідовності (1) є W змін знака, то послідовність (при $\alpha > 0$)

$$\alpha a_0, \alpha a_1 - a_0, \alpha a_2 - a_1, ..., \alpha a_n - a_{n-1}, \ -a_n \tag{5}$$

має не менше, ніж $W+1$ змін знаків.

Ця властивість – наслідок властивостей I і V. Спочатку будуємо послідовність чисел $a_0, \alpha a_1, \alpha^2 a_2, ..., \alpha^n a_n$, а потім застосуємо до неї властивість V і т. д.

Теорема (правило знаків Декарта). *Нехай W – число змін знака в послідовності $a_0, a_1, a_2, ..., a_n$ коефіцієнтів многочлена $f(x) = a_0 + a_1 x + ... + a_n x^n$ і N – число додатних коренів цього многочлена. Тоді має місце нерівність $W \geq N$.*

Доведення. Позначимо через $\alpha_1, \alpha_2, ..., \alpha_N$ усі додатні корені многочлена $f(x)$. Подамо цей многочлен у вигляді добутку

$$f(x) = g(x)(\alpha_1 - x)(\alpha_2 - x) \cdots (\alpha_N - x),$$

де $g(x) = b_0 + b_1 x + ... + b_m x^m$ – многочлен з дійсними коефіцієнтами степеня $m = n - N$. Нехай W_0 – число змін знака в послідовності $b_0, b_1, ..., b_m$ його коефіцієнтів. Многочлен

$$f_1(x) = g(x)(\alpha_1 - x) = \alpha_1 b_0 + (\alpha_1 b_1 - b_0)x +$$
$$+ ... + (\alpha_1 b_m - b_{m-1})x^m - b_m x^{m+1}$$

має послідовність коефіцієнтів з числом змін знака $W_1 \geq W_0 + 1$. Далі многочлен $f_2(x) = f_1(x)(\alpha_2 - x)$ також має число змін знака в послідовності його коефіцієнтів $W_2 \geq W_1 + 1 \geq W_0 + 2$. Продовжуючи цю процедуру, на N-му кроці дістанемо очікувану нерівність $W \geq N$.

Теорему доведено.

§ 1.9. Відокремлення дійсних коренів

Задача відокремлення дійсних коренів многочлена формулюється так.

Потрібно між межами l і L дійсних коренів вставити ряд зростаючих чисел $l_1, l_2, ..., l_m$ так, щоб у кожному з проміжків між числами $l, l_1, l_2, ..., l_m, L$ існувало не більше одного кореня даного многочлена $f(x)$.

Існує декілька способів відокремлення коренів многочленів. Серед них найбільш теоретично обґрунтований – метод Штурма.

Означення. *Послідовність многочленів з дійсними коефіцієнтами*

$$f_0(x), f_1(x), f_2(x), ..., f_m(x) \qquad \text{(S)}$$

називається рядом Штурма на відрізку $[a, b]$, *якщо ці многочлени мають такі властивості*:

I. Останній многочлен $f_m(x)$ *цього ряду не обертається на нуль на відрізку* $[a, b]$.

II. Жодні два сусідні многочлени цього ряду не мають спільних дійсних коренів.

III. Якщо λ *– дійсний корінь якого-небудь проміжного многочлена* $f_k(x)$ *(тобто* $k \neq 0$ *і* $k \neq m$*), тоді* $f_{k-1}(\lambda)$ *і* $f_{k+1}(\lambda)$ *мають протилежні знаки.*

IV. Якщо x, зростаючи, проходить через дійсний корінь многочлена $f_0(x)$*, тоді добуток* $f_0(x)f_1(x)$ *змінює знак з мінуса на плюс.*

Послідовність (S) називається *узагальненим* рядом Штурма, якщо її многочлени мають лише властивості I–III.

Позначення. Нехай $f_0(x), f_1(x), f_2(x), ..., f_m(x)$ – узагальнений ряд Штурма на відрізку $[a, b]$ і $f_0(a) \neq 0$, $f_0(b) \neq 0$. Нехай λ – корінь многочлена $f_0(x)$, який міститься в $[a, b]$. При зростанні x і переході через λ знак добутку $f_0(x)f_1(x)$ може змінюватися з плюса на мінус, або з мінуса на плюс, або залишатися без зміни. Число всіх коренів першого типу позначимо через Λ^+, а число всіх коренів другого – через Λ^-.

Якщо цей ряд має ще й властивість IV (ряд Штурма), тоді всі корені многочлена $f_0(x)$ – лише другого типу і, отже, число їх дорівнює Λ^-.

Теорема 1 (Штурма). *Нехай*

$$f_0(x), f_1(x), f_2(x), \dots, f_m(x)$$

– узагальнений ряд Штурма на відрізку $[a,b]$ *і* $f_0(a)\neq 0$, $f_0(b)\neq 0$. *Нехай, крім того, для будь-якого дійсного числа* c *число змін знака в послідовності чисел* $f_0(c), f_1(c), f_2(c), \dots, f_m(c)$ *дорівнює* $W(c)$. *Тоді виконується рівність*

$$W(a)-W(b)=\Lambda^- - \Lambda^+. \qquad (1)$$

Якщо дана послідовність многочленів має ще й властивість IV, то різниця $W(a)-W(b)$ невід'ємна і дорівнює числу різних коренів многочлена $f_0(x)$ на відрізку $[a,b]$, тобто Λ^-.

Доведення. Можна припустити, що при $x=a$ і $x=b$ жоден з многочленів ряду не обертається на нуль.

Нехай змінна x зростає, починаючи з $x=a$. Тоді, доки x не проходитиме через дійсні корені многочленів $f_k(x)$ ряду, зміна знаків у ньому не відбувається.

Нехай λ – корінь деякого проміжного многочлена $f_k(x)$ $(0<k<m)$. На підставі властивості III маємо $\operatorname{sign} f_{k-1}(\lambda) = -\operatorname{sign} f_{k+1}(\lambda)$. При достатньо малому $\varepsilon>0$ в інтервалі $(\lambda-\varepsilon, \lambda+\varepsilon)$ знаки $f_{k-1}(x)$ і $f_{k+1}(x)$ зберігаються.

Позначимо через η знаки значень функцій. Тоді при проходженні x через λ на відрізку ряду $f_{k-1}(x), f_k(x), f_{k+1}(x)$ можливі такі комбінації знаків:

	f_{k-1}	f_k	f_{k+1}
$\lambda-\varepsilon$	η	η	$-\eta$
λ	η	0	$-\eta$
$\lambda+\varepsilon$	η	η	$-\eta$

	f_{k-1}	f_k	f_{k+1}
$\lambda-\varepsilon$	η	$-\eta$	$-\eta$
λ	η	0	$-\eta$
$\lambda+\varepsilon$	η	$-\eta$	$-\eta$

	f_{k-1}	f_k	f_{k+1}
$\lambda-\varepsilon$	η	$-\eta$	$-\eta$
λ	η	0	$-\eta$
$\lambda+\varepsilon$	η	η	$-\eta$

	f_{k-1}	f_k	f_{k+1}
$\lambda-\varepsilon$	η	η	$-\eta$
λ	η	0	$-\eta$
$\lambda+\varepsilon$	η	$-\eta$	$-\eta$

Тут η – знак многочлена f_j ($j = k-1, k, k+1$). З цих таблиць випливає, що при переході через λ ряд $f_{k-1}(x), f_k(x), f_{k+1}(x)$ не набуває і не втрачає жодної зміни знака. Тому і для всієї послідовності $f_k(x)$ ($0 \le k \le m$) маємо те ж саме.

Тепер розглянемо випадок, коли λ – корінь многочлена $f_0(x)$. Для достатньо малого $\varepsilon > 0$ в інтервалі $(\lambda - \varepsilon, \lambda + \varepsilon)$ знак $f_1(x)$ зберігається. Зміна кількості чергувань знаків у ряді Штурма при переході x через λ можлива лише за рахунок зміни знака $f_0(x)$. Окремо розглянемо такі випадки:

Випадок 1. Добуток $f_0(x)f_1(x)$ змінює знак з плюса на мінус

	f_0	f_1
$\lambda-\varepsilon$	η	η
λ	0	η
$\lambda+\varepsilon$	$-\eta$	η

Кількість змін знака в ряді f_0, f_1 збільшилась на 1.

Випадок 2. Добуток $f_0(x)f_1(x)$ змінює знак з мінуса на плюс

	f_0	f_1
$\lambda-\varepsilon$	$-\eta$	η
λ	0	η
$\lambda+\varepsilon$	η	η

Кількість змін знака в ряді f_0, f_1 зменшилась на 1.

Випадок 3. Добуток $f_0(x)f_1(x)$ не змінює знак

	f_0	f_1
$\lambda-\varepsilon$	η	η
λ	0	η
$\lambda+\varepsilon$	η	η

	f_0	f_1
$\lambda-\varepsilon$	$-\eta$	η
λ	0	η
$\lambda+\varepsilon$	$-\eta$	η

Кількість змін знака в ряді f_0, f_1 зберігається.

Отже, при зростанні x від a до b зміна кількості чергувань знаків відбувається лише при проходженні через корені $f_0(x)$. При кожному такому проходженні в узагальненому ряді Штурма додається одна зміна знака у першому випадку, коли $f_0(x)f_1(x)$ змінює знак з плюса на мінус, усього кількість таких змін дорівнює Λ^+; у другому випадку, коли $f_0(x)f_1(x)$ змінює знак з мінуса на плюс – втрачається одна зміна знака, всього таких втрат Λ^-; у третьому випадку, коли $f_0(x)f_1(x)$ не змінює знак – зміна в знаках ряду не відбувається.

Тому маємо рівність

$$W(a)+\Lambda^+-\Lambda^-=W(b),$$

або

$$W(a)-W(b)=\Lambda^- -\Lambda^+.$$

Друге твердження теореми випливає як очевидне.

Теорему доведено.

Звернемо увагу на один клас нескінченних послідовностей многочленів, з яких виділяються узагальнені ряди Штурма.

Нехай нескінченна послідовність многочленів

$$\dots, F_n(x), F_{n-1}(x), \dots, F_2(x), F_1(x), F_0(x) \tag{F}$$

така, що $F_0(x)=1, F_1(x)=\alpha_1 x+\beta_1$, $\alpha_1>0$ і $F_n(x)$ при $n\geq 2$ визначається зворотним співвідношенням

$$F_n(x)=\left(\alpha_n x+\beta_n\right)F_{n-1}(x)-\gamma_n F_{n-2}(x). \tag{2}$$

Неважко переконатися, що при виконанні умов $\gamma_n>0$ для всіх $n\geq 2$ скінченний відрізок такої зворотної послідовності

$$F_n(x), F_{n-1}(x), \dots, F_2(x), F_1(x), F_0(x) \tag{F$_n$}$$

є узагальненим рядом Штурма.

Приклади

1. Многочлени Лежандра

$$P_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}\left(x^2-1\right)^n,\ P_0(x)=1,\ P_1(x)=x.$$

Вони визначаються співвідношенням

$$nP_n(x)=(2n-1)xP_{n-1}(x)-(n-1)P_{n-2}(x). \tag{3}$$

2. Многочлени Лагерра

$$L_n(x) = (-1)^n e^x \frac{d^n}{dx^n}\left(x^n e^{-x}\right),\ L_0(x) = 1,\ L_1(x) = x - 1.$$

Вони задовольняють співвідношення

$$L_n(x) = \left(x^2 - 2n + 1\right) L_{n-1}(x) - (n-1)^2 L_{n-2}(x). \qquad (4)$$

3. Многочлени Ерміта

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right),\ H_0(x) = 1,\ H_1(x) = 2x.$$

Вони визначаються співвідношенням

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x). \qquad (5)$$

Перевіримо ці співвідношення для многочленів Ерміта. Застосуємо правило Лейбніца для обчислення похідної (n–1)-го порядку:

$$\frac{d^n}{dx^n}\left(e^{-x^2}\right) = \frac{d^{n-1}}{dx^{n-1}}\left(-2xe^{-x^2}\right) =$$

$$= -2x\frac{d^{n-1}}{dx^{n-1}}\left(e^{-x^2}\right) - 2(n-1)\frac{d^{n-2}}{dx^{n-2}}\left(e^{-x^2}\right).$$

Тоді

$$H_n(x) = 2x\left[(-1)^{n-1} e^{x^2} \frac{d^{n-1}}{dx^{n-1}}\left(e^{-x^2}\right)\right] -$$

$$-2(n-1)\left[(-1)^{n-2} e^{x^2} \frac{d^{n-2}}{dx^{n-2}}\left(e^{-x^2}\right)\right] = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x).$$

Співвідношення (5) доведено.

Наслідок теореми 1. *Кожний многочлен $F_n(x)$ зворотної послідовності (F) має n різних дійсних коренів, якщо виконуються умови $\alpha_n > 0$, $\gamma_{n+1} > 0$ для будь-яких $n \ge 1$.*

Доведення. Індукцією за n можна довести, що старшим коефіцієнтом многочлена $F_n(x)$ є число $\alpha_n\alpha_{n-1}\cdots\alpha_2\alpha_1 > 0$. Для узагальненого ряду Штурма ($\bar{F}_n$) виберемо достатньо велике додатне число N_n, щоб на кінцях відрізка $[-N_n, N_n]$ знаки всіх членів ряду (F_n) визначались їх старшими доданками. Тоді кількість чергувань знаків цього ряду на кінцях відрізка $W(-N_n)=n$, $W(N_n)=0$. Застосовуємо рівність (1). Дістаємо $\Lambda^- - \Lambda^+ = n$. Звідси $\Lambda^- \geq n$. Але число Λ^- дорівнює числу різних коренів многочлена $F_n(x)$, які мають певні властивості. Тому $\Lambda^- = n$ і всі корені $F_n(x)$ дійсні та різні.

Наголосимо тепер на одній загальній схемі побудови ряду Штурма для даного многочлена $f(x)$.

Застосуємо алгоритм Евкліда для обчислення НСД многочлена $f(x)$ і його похідної $f'(x)$ в такій формі:

$$
\begin{aligned}
&f(x) = f'(x)h_1(x) - f_2(x),\\
&f'(x) = f_2(x)h_2(x) - f_3(x),\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&f_{m-2}(x) = f_{m-1}(x)h_m(x) - f_m(x),\\
&f_{m-1}(x) = f_m(x)h_{m+1}(x),
\end{aligned}
\tag{E}
$$

де $-f_2(x)$, $-f_3(x)$, ..., $-f_m(x)$ – остачі, що одержані при діленні.

При цьому $f_m(x) \neq 0$ і $f_m(x) = \text{НСД}(f(x), f'(x))$.

У наведеній схемі послідовних ділень на кожному кроці одержану остачу можна перемножити на будь-яке від'ємне число. Зауважимо, що $f_m(x)$ – дільник усіх многочленів одержаного ряду:

$$f_0(x) = f(x), f_1(x) = f'(x), f_2(x), ..., f_m(x). \tag{G}$$

Цей ряд лише в деяких випадках має всі властивості I–IV рядів Штурма. Так, наприклад, це можливо у випадку, коли $f(x)$ і $f'(x)$ взаємно прості, тобто коли останній многочлен $f_m(x)$ цього ряду – стала функція.

Виходячи з ряду (G), побудуємо нову послідовність многочленів

$$F_0(x) = \frac{f(x)}{f_m(x)},\ F_1(x) = \frac{f'(x)}{f_m(x)},\ F_2(x) = \frac{f_2(x)}{f_m(x)},\ ...,\ F_m(x) = \frac{f_m(x)}{f_m(x)} = 1.$$

По-перше, для цієї послідовності виконуються такі співвідношення:

$$
\begin{aligned}
&F_0(x) = F_1(x)h_1(x) - F_2(x),\\
&F_1(x) = F_2(x)h_2(x) - F_3(x),\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&F_{m-2}(x) = F_{m-1}(x)h_{m-1}(x) - F_m(x).
\end{aligned}
$$

По-друге, дійсні корені многочленів $F_0(x)$ і $f_0(x) = f(x)$ збігаються і будь-який дійсний корінь $F_0(x)$ – простий.

І, по-третє, з'ясовується, що ряд

$$F_0(x),\ F_1(x),\ F_2(x),\ \dots,\ F_{m-1}(x),\ F_m(x) = 1 \qquad \text{(H)}$$

є рядом Штурма на будь-якому відрізку $[a, b]$.

Властивості I–III очевидні. Відношення $\dfrac{F_0(x)}{F_1(x)} = \dfrac{f(x)}{f'(x)}$ при зростанні x і переході через корінь многочлена $f(x)$ змінює знак з мінуса на плюс (див. § 1.7, теорема 4). Тому добуток

$$F_0(x)F_1(x) = \frac{F_0(x)}{F_1(x)} F_1^2(x)$$

при зростанні x і переході через корінь $F_0(x)$ (тобто через корінь $f(x)$) змінює знак з мінуса на плюс.

Сформулюємо і доведемо теорему Штурма у такій формі.

Теорема 2. *На базі многочлена $f(x)$ побудуємо ряд*

$$f_0(x) = f(x), f_1(x) = f'(x), f_2(x),\ \dots,\ f_m(x) \qquad \text{(G)}$$

за схемою (E) *Евкліда. Нехай a, b – довільні дійсні числа, причому $f(a) \neq 0$ і $f(b) \neq 0$, $a < b$ і $f(x)$ не має кратних коренів на відрізку $[a, b]$. Тоді число різних коренів $f(x)$, які містяться на відрізку $[a, b]$, дорівнює $W(a) - W(b)$,* де *$W(a)$* і *$W(b)$ – число чергувань знаків ряду* (G) *відповідно при $x = a$* і *$x = b$.*

Доведення. На основі ряду (G) створимо ряд Штурма

$$F_0(x),\ F_1(x),\ F_2(x),\ \dots,\ F_m(x) = 1, \qquad \text{(H)}$$

де

$$F_0(x)=\frac{f(x)}{f_m(x)},\ F_k(x)=\frac{f_k(x)}{f_m(x)},\ k=\overline{1,m}.$$

За теоремою 1 число дійсних коренів многочлена $F_0(x)$, що містяться на відрізку $[a, b]$, дорівнює різниці $W_0(a)-W_0(b)$, де $W_0(a)$ і $W_0(b)$ – число чергувань знаків ряду (Н) відповідно при $x = a$ і $x = b$.

Оскільки $f_m(a)\neq 0$, то число чергувань знаків у послідовності

$$F_0(a),\ F_1(a),\ F_2(a),\ ...,\ F_m(a)=1$$

дорівнює числу чергувань знаків у послідовності чисел

$$F_0(a)f_m(a),\ F_1(a)f_m(a),\ ...,f_m(a),$$

тобто у послідовності чисел

$$f(a),f'(a),f_2(a),\ ...,f_m(a).$$

Отже, маємо рівність $W_0(a)=W(a)$. Аналогічно дістанемо $W_0(b)=W(b)$.

Оскільки $f_m(x)=\text{НСД}(f(x),f'(x))$, то корені многочленів $f(x)$ і $F_0(x)=\dfrac{f(x)}{f_m(x)}$ на відрізку $[a, b]$ одні й ті самі. Кожний корінь $F_0(x)$ – простий, він же буде і коренем $f(x)$, але не може бути кратним, $f_m(x)$ не змінює знак.

Число різних коренів многочлена $f(x)$, які містяться на відрізку $[a,b]$, дорівнює $W_0(a)-W_0(b)=W(a)-W(b)$.

Теорему доведено.

Приклад 4. Відокремити дійсні корені многочлена

$$f(x)=x^4+2x^3-x^2-2x+1.$$

Обчислюємо похідну $f'(x)=4x^3+6x^2-2x-2$ і замість неї при побудові ряду (F) візьмемо $f_1(x)=2x^3+3x^2-x-1$.

Ділимо $f(x)$ на $f_1(x)$:

$$
\begin{array}{r|l}
x^4+2x^3-x^2-2x+1 & 2x^3+3x^2-x-1 \\
\hline
-\;x^4+\frac{3}{2}x^3-\frac{1}{2}x^2-\frac{1}{2}x & \frac{1}{2}x+\frac{1}{4} \\
\frac{1}{2}x^3-\frac{1}{2}x^2-\frac{3}{2}x+1 & \\
-\;\frac{1}{2}x^3+\frac{3}{4}x^2-\frac{5}{4}x-\frac{1}{4} & \\
-\frac{5}{4}x^2-\frac{5}{4}x+\frac{5}{4} &
\end{array}
$$

Перша остача $R_3(x)=-\frac{5}{4}x^2-\frac{5}{4}x+\frac{5}{4}$. Обираємо $f_2(x)=$ $=x^2+x-1$. Ділимо $f_1(x)$ на $f_2(x)$:

$$
\begin{array}{r|l}
2x^3+3x^2-x-1 & x^2+x-1 \\
\hline
-\;2x^3+2x^2-2x & 2x+1 \\
x^2+x-1 & \\
-\;x^2+x-1 & \\
0 &
\end{array}
$$

Ділення без остачі. Ряд (F) складається з $f(x)$, $f_1(x), f_2(x)$. Наводимо таблицю:

	$f(x)$	$f_1(x)$	$f_2(x)$	Кількість чергувань знаків
$-\infty$	+	–	+	2
$+\infty$	+	+	+	0
–2	+	–	+	2
–1	+	+	–	1
0	+	–	–	1
1	+	+	+	0

Многочлен $f(x)$ має два дійсні корені, які містяться на відрізках [–2,–1] і [0, 1].

Приклад 5. Покажемо, що для многочленів Ерміта

$$H_n(x)=(-1)^n e^{x^2}\frac{d^n}{dx^n}\left(e^{-x^2}\right)$$

послідовність

$$H_n(x),\ H_{n-1}(x),\ \ldots,\ H_2(x), H_1(x)=2x,\ H_0(x)=1 \qquad \text{(H)}$$

є ряд Штурма.

Обчислюємо

$$H_n'(x)=(-1)^n 2xe^{x^2}\frac{d^n}{dx^n}\left(e^{-x^2}\right)+(-1)^n e^{x^2}\frac{d^{n+1}}{dx^{n+1}}\left(e^{-x^2}\right)=$$

$$=2xH_n(x)-H_{n+1}(x).$$

На підставі зворотного співвідношення (5) для многочленів Ерміта маємо

$$H_{n+1}(x)=2xH_n(x)-2nH_{n-1}(x).$$

Звідси дістаємо формулу

$$H_n'(x)=2nH_{n-1}(x). \qquad (6)$$

Виходячи з рівностей (5) і (6), робимо висновок, що послідовність (H) – ряд Штурма.

Вправи

1. Скласти ряд Штурма і відокремити корені многочлена

$$x^5-5x^3-10x^2+2.$$

2. Нехай

$$Q_n(x)=(-1)^n\left(x^2+1\right)^{n+\frac{1}{2}}\frac{d^n}{dx^n}\left[\left(x^2+1\right)^{-\frac{1}{2}}\right],\ n\in N.$$

Використовуючи формулу Лейбніца, довести, що многочлени $Q_n(x)$ задовольняють зворотне співвідношення:

$$Q_0(x)=1,\ Q_1(x)=x,$$

$$Q_n(x)=(2n-1)Q_{n-1}(x)-(n-1)^2(x^2+1)Q_{n-2}(x).$$

Переконатися, що $Q_n'=n^2Q_{n-1}$ і послідовність

$$Q_n(x),\ Q_{n-1}(x),\ Q_{n-2}(x),\ \ldots,\ Q_1(x),\ Q_0(x)$$

є ряд Штурма.

§ 2.1. Визначники 2-го і 3-го порядків

Поняття визначників 2-го і 3-го порядків виникає у зв'язку з обчисленням площ і об'ємів найпростіших фігур і тіл у геометрії та при розв'язанні систем лінійних рівнянь з двома або трьома невідомими.

Нехай на площині у прямокутній системі координат Oxy задано два вектори $\vec{a} = (X, Y)$, $\vec{b} = (X', Y')$ з їх проекціями X, Y і X', Y' на осі координат. Потрібно знайти площу паралелограма, побудованого на цих векторах. Відомо, що шукана площа S дорівнює абсолютному значенню величини $XY' - X'Y$, яка називається визначником 2-го порядку і позначається

$$\begin{vmatrix} X & Y \\ X' & Y' \end{vmatrix}.$$

Отже,

$$\pm S = \begin{vmatrix} X & Y \\ X' & Y' \end{vmatrix}.$$

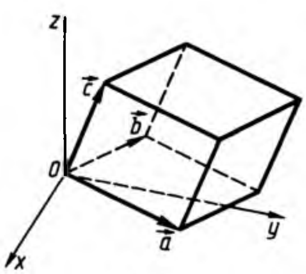


Рис. 5

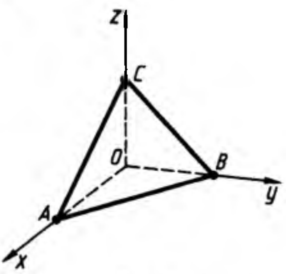


Рис. 6

Аналогічно розв'язується задача про обчислення об'єму паралелепіпеда, який визначається трьома векторами $\vec{a}$, $\vec{b}$, $\vec{c}$ в просторі.

Нехай у прямокутній системі координат ці вектори задано своїми проекціями на осі координат

$$\vec{a} = (X, Y, Z),$$
$$\vec{b} = (X', Y', Z'),$$
$$\vec{c} = (X'', Y'', Z'').$$

Тоді шуканий об'єм V дорівнює абсолютному значенню змішаного векторного добутку $(\vec{a}, \vec{b}, \vec{c}) = ([\vec{a}, \vec{b}]\vec{c})$ цих векторів (векторний добуток $\vec{v} = [\vec{a}, \vec{b}]$ скалярно множиться на вектор $\vec{c}$). Оскільки векторний добуток

$$\vec{v} = [\vec{a}, \vec{b}] = \left(\begin{vmatrix} Y & Z \\ Y' & Z' \end{vmatrix}, -\begin{vmatrix} X & Z \\ X' & Z' \end{vmatrix}, \begin{vmatrix} X & Y \\ X' & Y' \end{vmatrix} \right),$$

то

$$(\vec{a}, \vec{b}, \vec{c}) = \begin{vmatrix} Y & Z \\ Y' & Z' \end{vmatrix} \cdot X'' - \begin{vmatrix} X & Z \\ X' & Z' \end{vmatrix} \cdot Y'' + \begin{vmatrix} X & Y \\ X' & Y' \end{vmatrix} \cdot Z'' =$$

$$= YZ'X'' - ZY'X'' - XZ'Y'' + ZX'Y'' + XY'Z'' - YX'Z'' =$$

$$= XY'Z'' + YZ'X'' + ZX'Y'' - ZY'X'' - XZ'Y'' - YX'Z''.$$

Одержана алгебрична сума називається визначником 3-го порядку і позначається

$$\begin{vmatrix} X & Y & Z \\ X' & Y' & Z' \\ X'' & Y'' & Z'' \end{vmatrix}.$$

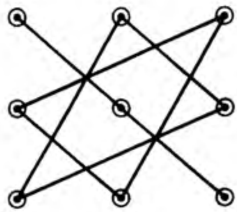
Об'єм паралелепіпеда, побудованого на трьох векторах $\vec{a}$, $\vec{b}$, $\vec{c}$, обчислюється за формулою $V = \left|(\vec{a}, \vec{b}, \vec{c})\right|$.

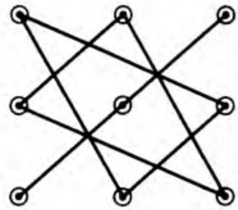
Отже,

$$(\vec{a}, \vec{b}, \vec{c}) = \begin{vmatrix} X & Y & Z \\ X' & Y' & Z' \\ X'' & Y'' & Z'' \end{vmatrix} = \pm V.$$

Для обчислення цього визначника скористаємось таким правилом.

Доданки зі знаком плюс одержуються за схемою, тобто складаються добутки елементів визначника: спочатку тих, які розміщені на головній діагоналі; потім добутки елементів, що розміщені у вершинах двох трикутників, одна зі сторін яких паралельна головній діагоналі.





Доданки зі знаком мінус одержуються за схемою, тобто складаються добутки елементів визначника: спочатку тих елементів, які розміщені на другій діагоналі; потім елементів, що розміщені у вершинах трикутників, одна зі сторін яких паралельна другій діагоналі.

Для встановлення закономірності обчислення визначників 2-го і 3-го порядків, яка має бути основою для переходу до визначників вищих порядків, запровадимо позначення його елементів. Кожному елементу визначника будемо присвоювати два індекси: перший вказуватиме на номер рядка, а другий – на номер стовпця, в якому він розміщений.

Тоді для визначника 2-го порядку матимемо

$$\begin{vmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{vmatrix} = X_{11}X_{22} - X_{12}X_{21}.$$

Визначник 3-го порядку запишеться так:

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

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

$$- X_{13}X_{22}X_{31} - X_{12}X_{21}X_{33} - X_{11}X_{23}X_{32}.$$

Виявляється, що закономірність розміщення знаків “+” і “–” перед добутками залежить від властивостей перестановок других індексів множників, якщо перші індекси множників розміщуються у порядку їх зростання. Запишемо всі ці перестановки других індексів для визначника 3-го порядку у вигляді таблиць

$$\left.\begin{matrix} 1, & 2, & 3 \\ 2, & 3, & 1 \\ 3, & 1, & 2 \end{matrix}\right\} + \qquad \left.\begin{matrix} 3, & 2, & 1 \\ 2, & 1, & 3 \\ 1, & 3, & 2 \end{matrix}\right\} -$$

Усього кількість перестановок з трьох номерів 1, 2, 3 дорівнює 6. У перестановці a, b, c цих номерів пару номерів називають інверсією, якщо той, що стоїть ліворуч, більший того, що стоїть праворуч. Наприклад, у перестановці 3, 2, 1 номери 3 і 2, 3 і 1, 2 і 1 утворюють інверсії. Число інверсій перестановки a, b, c позначимо через $[a, b, c]$. У нашому випадку

$$\begin{matrix} [1, 2, 3] = 0, & [3, 2, 1] = 3, \\ [2, 3, 1] = 2, & [2, 1, 3] = 1, \\ [3, 1, 2] = 2, & [1, 3, 2] = 1. \end{matrix}$$

Отже, перестановки типу “+” мають парні числа інверсій, а типу “–” – непарні.

Встановленою закономірністю для визначників 3-го порядку ми надалі й скористаємось.

§ 2.2. Перестановки

Розглянемо деякий набір із n предметів $A_1, A_2, ..., A_n$. Для вивчення їх взаємного розміщення можна обмежитись розглядом лише номерів 1, 2, …, n. Тому будемо розглядати всі можливі перестановки a, b, c, …, w цих номерів. Число всіх таких перестановок дорівнює $n! = 1 \cdot 2 \cdot 3 \cdots n$.

Означення. *Пара чисел i* та *k із перестановок a, b, c, …, i, …, k, …, w утворює інверсію, якщо попереднє число з цієї пари більше наступного числа k.*

Число всіх інверсій в перестановці $a, b, c, \ldots, w$ позначимо через $[a, b, c, \ldots, w]$. Для підрахунку цього числа застосовується просте правило.

Беремо число a, яке розміщене в перестановці на першому місці, і підраховуємо, скільки інверсій воно утворює зі всіма наступними числами перестановки. Нехай таким чином одержали α інверсій. Потім переходимо до числа b і знаходимо число β інверсій, які утворює це число зі всіма наступними числами перестановки, і т. д. Тоді

$$[a, b, c, \ldots, w] = \alpha + \beta + \ldots$$

Означення. *Перестановка $a, b, c, \ldots, w$ натуральних чисел $1, 2, \ldots, n$ називається парною, якщо число $[a, b, c, \ldots, w]$ її інверсій – парне, і непарною, якщо $[a, b, c, \ldots, w]$ – непарне.*

Приклад. З'ясувати, коли перестановка $n, n-1, n-2, \ldots, 2, 1$ чисел $1, 2, \ldots, n$ буде парною.

Розв'язання. Спочатку підрахуємо $t = [n, n-1, \ldots, 2, 1]$. Число n утворює інверсію зі всіма наступними числами, тобто всього $n-1$ інверсій. Далі, число $n-1$ утворює $n-2$ інверсій зі всіма наступними числами і т. д. Усього дістаємо

$$t = (n-1) + (n-2) + \ldots + 1 = \frac{n(n-1)}{2}$$

інверсій у даній перестановці.

Парність числа t залежить від остачі при діленні n на 4:

а) якщо $n = 4k$, тоді $t = 2k(4k-1)$ – парне число;

б) якщо $n = 4k+1$, тоді $t = (4k+1)2k$ – парне число;

в) якщо $n = 4k+2$, тоді $t = (2k+1)(4k+1)$ – непарне число;

г) якщо $n = 4k+3$, тоді $t = (4k+3)(2k+1)$ – непарне число.

Означення. *Якщо в перестановці $a, b, c, \ldots, w$ натуральних чисел $1, 2, \ldots, n$ поміняти місцями два яких-небудь числа, тоді таке перетворення називається транспозицією.*

Теорема 1. *При одній транспозиції змінюється парність перестановки.*

Доведення. Нехай в перестановці $a, b, \ldots, i, \ldots, k, \ldots, w$ поміняли місцями i та k і дістали нову перестановку $a, b, \ldots, k, \ldots, i, \ldots, w$.

Початкову перестановку запишемо у вигляді $Ai...kB$, де A – "лівий кінець" – група чисел, яка розміщена ліворуч від числа i; B – "правий кінець" – група чисел, яка розміщена праворуч від числа k.

Розглянемо два випадки.

Випадок 1. Елементарна транспозиція. Поміняли місцями два числа i та k, що розміщувались поруч. У результаті з перестановки $AikB$ одержали перестановку $AkiB$. При цьому число інверсій може змінитися лише за рахунок взаємного розміщення i та k, тобто на число 1. Отже,

$$[AkiB] = [AikB] \pm 1 .$$

Випадок 2. Між i та k міститься m чисел ($m \geq 1$). Задана перестановка має вигляд $Aii_1 i_2 ... i_m kB$. Перестановку, одержану з неї транспозицією номерів i та k, можна дістати за допомогою серії елементарних транспозицій:

$$Aii_1 i_2 ... i_m kB \to Ai_1 ii_2 ... i_m kB \to Ai_1 i_2 i ... i_m kB \to \ldots \to$$

$$\to Ai_1 i_2 ... i_m ikB \to Ai_1 i_2 ... i_m kiB \to Ai_1 i_2 ... ki_m iB \to$$

$$\to \ldots \to Aki_1 i_2 ... i_m iB .$$

Кожна з них змінює парність перестановки, а їх число становитиме $2m+1$ – непарне. Тому, якщо перестановка $Aii_1 i_2 \ldots i_m kB$ – парна, то остання перестановка $Aki_1 i_2 \ldots i_m iB$ – непарна, і навпаки. Отже,

$$[Aki_1 i_2 \ldots i_m iB] = [Aii_1 i_2 \ldots i_m kB] + 2m + 1.$$

Теорема 2. *Серед усіх перестановок чисел* $1, 2, \ldots,$ n *при* $n > 1$ *число парних і непарних перестановок збігається.*

Доведення. Нехай p – число парних, а q – число непарних перестановок. У всіх перестановках класу парних виконаємо транспозицію чисел 1 і 2. При цьому ми утворимо непарну перестановку. Звідси випливає співвідношення $p \leq q$. Аналогічно доводиться співвідношення $q \leq p$. Отже,

$$p = q = \frac{n!}{2}.$$

Вправи

1. Яка з перестановок чисел 1, 2, …, n має найбільше число інверсій?

2. Нехай t – число інверсій у перестановці $a_1, a_2, ..., a_n$ чисел 1, 2, …, n. Обчислити $s = [a_n, a_{n-1}, ..., a_2, a_1]$.

3. Переконатися, що від перестановки $a_1, a_2, ..., a_n$ можна перейти до перестановки $b_1, b_2, ..., b_n$ тих же предметів за допомогою не більш як $n-1$ транспозицій.

4. Довести, що для будь-якого цілого числа t, яке міститься в межах від 0 до $n\ (n-1)/2$, існує така перестановка $a_1, a_2, ..., a_n$ чисел 1, 2, …, n, що $t = [a_1, a_2, ..., a_n]$.

§ 2.3. Визначники та їх властивості

Нехай K – довільне комутативне кільце (зокрема – поле), елементи якого позначимо через α, β, γ, … Запишемо квадратну матрицю (таблицю)

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

з елементами α_{ik} кільця K.

Означення. *Визначником* (*детермінантом*) *матриці A над кільцем K називається алгебрична сума всіх можливих добутків виду*

$$(-1)^t \alpha_{1a}\alpha_{2b}\alpha_{3c} \cdots \alpha_{nw},$$

де α_{1a}, α_{2b}, α_{3c}, α_{nw} – відповідно елементи першого, другого, третього, n-го рядків матриці A; a, b, c, …, w – перестановка чисел 1, 2, …, n; t – число інверсій цієї перестановки. Підсумовування здійснюється за всіма можливими перестановками a, b, c, …, w других індексів.

Подамо цю суму у вигляді

$$\Delta = \sum_{(a,b,c,\dots,w)} (-1)^{[a,b,c,\dots,w]} \alpha_{1a}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw} .$$

Число її доданків становить $n!$.

Визначник матриці A позначається так:

$$\det A \equiv |A| = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} .$$

Якщо $n=1$, тоді $\det A = \alpha_{11}$.

Для $n=2$ і $n=3$ означення збігається з раніше введеними поняттями визначників 2-го і 3-го порядків з елементами із поля R дійсних чисел.

Дане означення узагальнює попереднє в двох напрямах. По-перше, число n (порядок визначника) – будь-яке натуральне число, починаючи з одиниці. По-друге, елементами визначника можуть бути елементи будь-якого комутативного кільця або поля, наприклад, з кільця многочленів, з поля раціональних функцій.

Розглянемо найпростіші властивості визначників.

I. Якщо всі елементи деякого рядка визначника дорівнюють нулю, то й сам визначник дорівнює нулю.

Доведення. Властивість I випливає з означення визначника. Дійсно, кожний доданок алгебричної суми містить по одному представнику з кожного рядка і кожного стовпця.

II. Якщо всі елементи деякого фіксованого рядка визначника Δ помножити на елемент λ основного кільця K, то дістанемо новий визначник Δ_1, причому $\Delta_1 = \lambda\Delta$.

Доведення. Нехай, наприклад, у визначнику

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} \qquad (1)$$

таким фіксованим рядком є перший. Тоді

$$\Delta_1 = \begin{vmatrix} \lambda\alpha_{11} & \lambda\alpha_{12} & \dots & \lambda\alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \lambda\Delta .$$

III. Нехай у визначнику (1) елементи деякого фіксованого рядка подаються у вигляді суми двох доданків (наприклад, елементи першого рядка $\alpha_{11} = \alpha'_{11} + \alpha''_{11}$, $\alpha_{12} = \alpha'_{12} + \alpha''_{12}$, ..., $\alpha_{1n} = \alpha'_{1n} + \alpha''_{1n}$). Тоді він дорівнює сумі двох визначників $\Delta = \Delta' + \Delta''$, де

$$\Delta' = \begin{vmatrix} \alpha'_{11} & \alpha'_{12} & \dots & \alpha'_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}, \ \Delta'' = \begin{vmatrix} \alpha''_{11} & \alpha''_{12} & \dots & \alpha''_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}.$$

Аналогічне твердження зберігає силу, якщо елементи деякого рядка подаються у вигляді суми m доданків: визначник Δ дорівнюватиме сумі m визначників $\Delta', \Delta'', \dots, \Delta^{(m)}$.

Як і властивість I, доведення властивостей II і III безпосередньо випливає з означення визначника.

IV. Якщо елементи першого рядка визначника $\Delta = \det A$ такі, що $\alpha_{1a} = 0$ при $a > 1$, тоді

$$\det A = \begin{vmatrix} \alpha_{11} & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{vmatrix}.$$

Доведення. За означенням

$$\det A = \sum_{(a,b,c,\dots,w)} (-1)^t \alpha_{1a}\alpha_{2b}\alpha_{3c} \cdots \alpha_{nw} ,$$

де $t = [a, b, c, \dots, w]$ у кожному з доданків. Оскільки $\alpha_{1a} = 0$ при $a > 1$, тоді

$$\det A = \sum_{(1,b,c,\ldots,w)} (-1)^t \alpha_{11}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw},$$

де $t=[1, b, c, \ldots, w] = [b, c, \ldots, w]$ і $b, c, \ldots, w$ – перестановка чисел 2, 3, …, n. Отже,

$$\det A = \alpha_{11}\Big(\sum_{(b,c,\ldots,w)} (-1)^t \alpha_{2b}\cdots\alpha_{nw}\Big),$$

де $t=[b, c, \ldots, w]$. Сума, яка міститься в дужках, – визначник

$$\begin{vmatrix} \alpha_{22} & \alpha_{23} & \ldots & \alpha_{2n} \\ \alpha_{32} & \alpha_{33} & \ldots & \alpha_{3n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n2} & \alpha_{n3} & \ldots & \alpha_{nn} \end{vmatrix}$$

$(n-1)$-го порядку.

V. Визначник нижнього трикутного виду

$$\begin{vmatrix} \alpha_{11} & 0 & 0 & \ldots & 0 \\ \alpha_{21} & \alpha_{22} & 0 & \ldots & 0 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \ldots & \alpha_{nn} \end{vmatrix}$$

($\alpha_{ik}=0$ за умови, що $i<k$) дорівнює добутку $\alpha_{11}\alpha_{22}\cdots\alpha_{nn}$ його діагональних елементів.

Ця властивість є наслідком властивості IV.

VI. Якщо у визначнику (1) порядку $n \geq 2$ є два однакові рядки (з різними номерами), тоді він дорівнює нулю.

Доведення. Нехай у визначнику (1) є два однакові рядки. Для наочності вважатимемо однаковими два перші рядки: $\alpha_{11}=\alpha_{21}$, $\alpha_{12}=\alpha_{22}$, …, $\alpha_{1n}=\alpha_{2n}$. За означенням

$$\Delta = \sum_{(a,b,c,\ldots,w)} (-1)^t \alpha_{1a}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw},$$

де $t=[a, b, c, \ldots, w]$.

Зафіксуємо в цій сумі який-небудь доданок

$$h=(-1)^t \alpha_{1a}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw},\ t=[a, b, c, \ldots, w].$$

У цій сумі присутній доданок, який однозначно відповідає h і який ми позначимо через h_1:

$$h_1 = (-1)^{t_1}\alpha_{1b}\alpha_{2a}\alpha_{3c}\cdots\alpha_{nw},\ t_1 = [b, a, c, \ldots, w].$$

За умовою $\alpha_{1a} = \alpha_{2a}$, $\alpha_{1b} = \alpha_{2b}$, а тому $\alpha_{1a}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw} = \alpha_{1b}\alpha_{2a}\alpha_{3c}\cdots\alpha_{nw}$, оскільки в цих добутках співмножники, починаючи з третього, збігаються. Перестановка b, a, c,…, w одержується з перестановки a, b, c,…, w транспозицією чисел a і b, а тому вони мають різні парності й, отже, $(-1)^{t_1} = -(-1)^t$. Але тоді $h_1 = -h$ і $h + h_1 = 0$.

Отже, для кожного доданка h суми Δ є однозначно визначений доданок h_1 такий, що $h + h_1 = 0$. Звідси і випливає рівність $\Delta = 0$.

Вправи

1. Довести, що визначник верхнього трикутного виду

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ 0 & 0 & \alpha_{33} & \cdots & \alpha_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \alpha_{nn} \end{vmatrix}$$

дорівнює добутку $\alpha_{11}\alpha_{22}\cdots\alpha_{nn}$ елементів головної діагоналі.

2. Довести, що визначник трикутного виду

$$\begin{vmatrix} 0 & \cdots & 0 & 0 & \alpha_{1n} \\ 0 & \cdots & 0 & \alpha_{2n-1} & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \cdots & \alpha_{nn-2} & \alpha_{nn-1} & \alpha_{nn} \end{vmatrix}$$

дорівнює $(-1)^{\frac{n(n-1)}{2}}\alpha_{1n}\alpha_{2n-1}\ldots\alpha_{n1}$, де α_{1n}, α_{2n-1}, …, α_{n1} – елементи другої (побічної) діагоналі.

Означення. *Нехай*

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

квадратна матриця з елементами α_{ik} *із комутативного кільця К. Визначником матриці (або детермінантом) називається алгебрична сума з n! доданків виду*

$$\Delta' = \sum_{(a,b,c,\cdots,w)} (-1)^{[a,b,c,\dots,w]} \alpha_{a1}\alpha_{b2}\alpha_{c3}\cdots\alpha_{wn},$$

де a, b, c,…, w утворюють різні можливі перестановки чисел 1, 2, …, *n. Позначимо цю суму через* $\det' A$ *або* $|A|'$.

Теорема 1. *Обидва означення детермінанта матриці A еквівалентні*:

$$\det A = \det' A .$$

Доведення. Порівняємо дві суми, породжені елементами однієї і тієї ж матриці A:

$$\Delta = \sum_{(a,b,c,\dots,w)} (-1)^t \alpha_{1a}\alpha_{2b}\alpha_{c3}\cdots\alpha_{nw}, \text{ де } t=[a, b, c, \dots, w];$$

$$\Delta' = \sum_{(a,b,c,\dots,w)} (-1)^{[a,b,c,\dots,w]} \alpha_{a1}\alpha_{b2}\alpha_{c3}\cdots\alpha_{wn}, \text{ де } t=[a, b, c, \dots, w].$$

Оберемо в першій з них довільний доданок $h=(-1)^t \times \times \alpha_{1a}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw}$ і перетворимо його за певним правилом.

Розглянемо спочатку добуток $\alpha_{1a}\alpha_{2b}\alpha_{3c}\cdots\alpha_{nw}$ і знайдемо в ньому множник виду $\alpha_{a'1}$, потім – множник виду $\alpha_{b'2}$ і т. д. Переставимо в добутку множники і подамо його у вигляді $\alpha_{a'1}\alpha_{b'2}\cdots\alpha_{w'n}$. Перестановка $a', b', \dots, w'$ перших індексів визначається перестановкою $a, b, \dots, w$ других індексів однозначно. Перестановка множників може бути здійснена за допомогою деякого числа s транспозицій множників. За допомогою цього числа s транспозицій перестановка a, b, ..., w перетворюється в послідовність 1, 2, ..., n, а перші індекси 1, 2, ..., n – в перестановку $a', b', ..., w'$.

Тому перестановку $a, b, ..., w$ можна перетворити в перестановку $a', b', ..., w'$ за допомогою $2s$ транспозицій. Отже, перестановки $a, b, ..., w$ і $a', b', ..., w'$ мають однакову парність. Дістали рівність $(-1)^t = (-1)^{t_1}$ для $t = [a, b, ..., w]$ і $t' = [a', b', ..., w']$.

Отже,

$$(-1)^t \alpha_{1a}\alpha_{2b} \cdots \alpha_{nw} = (-1)^{t_1} \alpha_{a'1}\alpha_{b'2} \cdots \alpha_{w'n}.$$

Перетворивши всі доданки суми за описаним правилом, дістанемо

$$\Delta = \det A = \det' A = \Delta'.$$

Теорему доведено.

Для даної матриці

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

матриця

$$A^T = \begin{vmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} \end{vmatrix}$$

називається *транспонованою*. Вона одержується з матриці A заміною всіх елементів α_{ik} на α_{ki} (рядків на стовпці, і навпаки).

Теорема 2. *Визначники транспонованих матриць рівні*:

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

Доведення. Запишемо $\det A$ на підставі означення:

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

$$= \sum_{(a,b,c,\dots,w)} (-1)^{[a,b,c,\dots,w]} \alpha_{1a}\alpha_{2b}\cdots\alpha_{nw} .$$

Якщо праворуч і ліворуч цієї тотожності замість α_{ik} всюди поставити α_{ki}, то дістанемо рівність

$$\det A^T = \begin{vmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{n2} \\ \dots & \dots & \dots & \dots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{vmatrix} =$$

$$= \sum_{(a,b,c,\dots,w)} (-1)^{[a,b,c,\dots,w]} \alpha_{a1}\alpha_{b2}\cdots\alpha_{wn} .$$

Права частина цієї рівності збігається з $\det' A$. На підставі теореми 1 маємо $\det A = \det' A$. Тому $\det A^T = \det A$.

Теорему доведено.

Із умови цієї теореми випливає твердження: всі властивості визначника, які доведені для його рядків, зберігають силу і для його стовпців.

Вправи

1. Довести, що визначник n-го порядку, в якого кожний елемент α_{ik} – комплексне число, спряжене з α_{ki}, є дійсне число.

2. Довести, що визначник непарного порядку дорівнює нулю, якщо всі його елементи задовольняють умову $\alpha_{ik} + \alpha_{ki} = 0$.

§ 2.4. Елементарні перетворення визначників

При обчисленні визначника (з елементами з поля дійсних чисел) виникають певні труднощі, якщо при цьому обчисленні використовувати лише означення. Визначник n-го порядку містить $n!$ доданків, а кожен з цих доданків – добуток його n елементів. Якщо всі ці арифметичні дії провести на найсучаснішій ЕОМ, то для обчислення, наприклад, визначника 100-го порядку потрібно було б витратити дуже багато часу. Тому треба знаходити інші методи, що базуються на властивостях визначників.

Означення. *Елементарними перетвореннями визначника* Δ *(з елементами* α_{ik} *з поля P)*

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

називаються такі перетворення:

1. Множення всіх елементів деякого фіксованого рядка визначника на будь-який скаляр $\lambda \neq 0$ *(тобто на елемент* $\lambda \neq 0$ *основного поля P).*

2. Додавання до елементів деякого рядка відповідних елементів іншого рядка, помножених на скаляр λ.

3. Перестановка місцями двох будь-яких рядків.

Аналогічні перетворення 1, 2, 3 для стовпців також називаються елементарними.

Теорема 1*. При елементарних перетвореннях (з рядками або стовпцями) визначника* Δ *одержується новий визначник* Δ_1:

1) при елементарному перетворенні типу 1 $\Delta_1 = \lambda\Delta$;

2) при елементарному перетворенні типу 2 $\Delta_1 = \Delta$;

3) при елементарному перетворенні типу 3 $\Delta_1 = -\Delta$.

Доведення. Перше твердження теореми – це властивість II визначника (див. § 2.3).

При доведенні другого твердження нехай визначник Δ має порядок $n \geq 2$ і визначник Δ_1 одержується з нього додаванням до елементів другого рядка елементів першого, помножених на скаляр λ:

$$\Delta_1 = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} + \lambda\alpha_{11} & \alpha_{22} + \lambda\alpha_{12} & \dots & \alpha_{2n} + \lambda\alpha_{1n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}.$$

На підставі властивості III визначників його можна подати у вигляді суми

$$\Delta_1 = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \lambda\alpha_{11} & \lambda\alpha_{12} & \dots & \lambda\alpha_{1n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}.$$

У другому доданку множник λ виноситься за знак визначника. У доданку, який після нього залишиться, матимемо два однакові рядки, а тому визначник дорівнює нулю (властивість VI). Звідси маємо рівність $\Delta_1 = \Delta$.

При доведенні третього твердження теореми обмежимося розглядом визначника 3-го порядку (міркування зберігають силу і для загального випадку):

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

Переконаємося, що при перестановці двох рядків змінюється лише знак визначника. Застосуємо до нього елементарні перетворення типу 1 і 2:

а) До другого додамо елементи першого рядка. Дістанемо

$$\Delta_1 = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} + \alpha_{11} & \alpha_{22} + \alpha_{12} & \alpha_{23} + \alpha_{13} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

б) Від елементів першого рядка в Δ_1 віднімемо елементи другого рядка. Знайдемо

$$\Delta_2 = \begin{vmatrix} -\alpha_{21} & -\alpha_{22} & -\alpha_{23} \\ \alpha_{21}+\alpha_{11} & \alpha_{22}+\alpha_{12} & \alpha_{23}+\alpha_{13} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

в) До елементів другого рядка в Δ_2 додамо елементи першого рядка. Матимемо

$$\Delta_3 = \begin{vmatrix} -\alpha_{21} & -\alpha_{22} & -\alpha_{23} \\ \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

г) Спільний множник (–1) першого рядка виносимо за знак визначника. Дістанемо

$$\Delta_4 = (-1)\begin{vmatrix} \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

Оскільки $\Delta_1 = \Delta_2 = \Delta_3 = \Delta_4$, тоді

$$\begin{vmatrix} \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = -\Delta.$$

Теорему доведено.

Означення. *Виділимо у визначнику*

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{vmatrix}$$

i-й рядок і k-й стовпець та викреслимо елементи цього рядка і стовпця. Дістанемо визначник $(n-1)$-го порядку

$$\Delta_{ik}=\begin{vmatrix} \alpha_{11} & \dots & \alpha_{1k-1} & \alpha_{1k+1} & \dots & \alpha_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{i-11} & \dots & \alpha_{i-1k-1} & \alpha_{i-1k+1} & \dots & \alpha_{i-1n} \\ \alpha_{i+11} & \dots & \alpha_{i+1k-1} & \alpha_{i+1k+1} & \dots & \alpha_{i+1n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{nk-1} & \alpha_{nk+1} & \dots & \alpha_{nn} \end{vmatrix}.$$

Він називається мінором елемента α_{ik}, *який міститься в* Δ *на перетині i-го рядка і k-го стовпця.*

Добуток $A_{ik}=(-1)^{i+k}\Delta_{ik}$ *називається алгебричним доповненням елемента* α_{ik}.

Теорема 2. *Якщо на деякій лінії визначника* Δ (*i-му рядку або k-му стовпці*) *є єдиний елемент* $\alpha_{ik}\neq 0$, *тоді* $\Delta=\alpha_{ik}A_{ik}$.

Доведення. Нехай лінією, яка фігурує в умові теореми, є i-й рядок визначника Δ: $\alpha_{ik}\neq 0$, $\alpha_{ij}=0$, якщо $j\neq k$.

Якщо $i=k=1$, тоді визначник має вигляд

$$\Delta=\begin{vmatrix} \alpha_{11} & 0 & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{vmatrix}.$$

На підставі властивості IV (див. § 2.3) маємо

$$\Delta=\alpha_{11}\begin{vmatrix} \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ \alpha_{32} & \alpha_{33} & \dots & \alpha_{3n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{vmatrix}=$$

$$=\alpha_{11}\Delta_{11}=\alpha_{11}\cdot(-1)^{1+1}\Delta_{11}=\alpha_{11}A_{11}.$$

У загальному випадку

$$\Delta = \begin{vmatrix} \alpha_{11} & \dots & \alpha_{1k-1} & \alpha_{1k} & \alpha_{1k+1} & \dots & \alpha_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{i-11} & \dots & \alpha_{i-1k-1} & \alpha_{i-1k} & \alpha_{i-1k+1} & \dots & \alpha_{i-1n} \\ 0 & \dots & 0 & \alpha_{ik} & 0 & \dots & 0 \\ \alpha_{i+11} & \dots & \alpha_{i+1k-1} & \alpha_{i+1k} & \alpha_{i+1k+1} & \dots & \alpha_{i+1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{nk-1} & \alpha_{nk} & \alpha_{nk+1} & \dots & \alpha_{nn} \end{vmatrix}.$$

Проведемо серію елементарних перетворень типу 3.

Переставляємо i-й рядок послідовно з тими рядками, що розміщені вище нього, доти, доки не досягнемо першого рядка. Потім в одержаному визначнику переставляємо k-й стовпець доти, доки він не займе місце першого стовпця. Врешті-решт знайдемо

$$\Delta' = \begin{vmatrix} \alpha_{ik} & 0 & \dots & 0 & 0 & \dots & 0 \\ \alpha_{1k} & \alpha_{11} & \dots & \alpha_{1k-1} & \alpha_{1k+1} & \dots & \alpha_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{i-1k} & \alpha_{i-11} & \dots & \alpha_{i-1k-1} & \alpha_{i-1k+1} & \dots & \alpha_{i-1n} \\ \alpha_{i+1k} & \alpha_{i+11} & \dots & \alpha_{i+1k-1} & \alpha_{i+1k+1} & \dots & \alpha_{i+1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{nk} & \alpha_{n1} & \dots & \alpha_{nk-1} & \alpha_{nk+1} & \dots & \alpha_{nn} \end{vmatrix}$$

і

$$\Delta = (-1)^{(i-1)+(k-1)} \cdot \Delta' = (-1)^{i+k} \Delta', \ \Delta' = \alpha_{ik} \Delta_{ik}.$$

Звідси матимемо

$$\Delta = \alpha_{ik} \cdot (-1)^{i+k} \cdot \Delta_{ik} = \alpha_{ik} A_{ik}.$$

Теорему доведено.

Метод елементарних перетворень

Доведені вище теореми можуть бути підставою для обчислення визначників. Розглянемо два варіанти.

Варіант 1. Метод зниження порядку. Розглядаємо визначники, елементи яких належать полю P.

У визначнику

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \dots & \dots & \dots & \dots \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{in} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

виділимо елемент $\alpha_{ik} \neq 0$, який належить i-му рядку. До елементів p-го стовпця ($p \neq k$) визначника додамо елементи k-го стовпця, помноженого на скаляр $\lambda_p = -\dfrac{\alpha_{ip}}{\alpha_{ik}}$. У результаті такого елементарного перетворення дістанемо визначник Δ_1, рівний Δ, в якому на місці (i, p) (на перетині i-го рядка і p-го стовпця) буде нуль. Подібне перетворення застосуємо до всіх стовпців з номерами $p \neq k$. Тоді дістанемо визначник Δ_0, у якого в i-му рядку залишиться єдиний елемент α_{ik}, відмінний від нуля. При цьому $\Delta = \Delta_0$. Але Δ_0 дорівнює добутку α_{ik} на його алгебричне доповнення.

Отже, обчислення визначника звелось до обчислення визначника $(n-1)$-го порядку.

Приклад 1. Обчислити визначник

$$\Delta = \begin{vmatrix} 5 & -2 & -12 & 15 \\ 1 & -1 & -4 & 5 \\ -1 & 3 & 3 & 5 \\ 6 & 2 & -12 & 25 \end{vmatrix}.$$

Фіксуємо рядок з номером $i = 2$ і ведучим елементом $\alpha_{21} = 1$. Проведемо такі елементарні перетворення:

а) додамо до стовпця 2 стовпець 1,

б) додамо до стовпця 3 стовпець 1, помножений на 4,

в) віднімемо від стовпця 4 стовпець 1, помножений на 5.

Матимемо

$$\Delta_0 = \begin{vmatrix} 5 & 3 & 8 & -10 \\ 1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 10 \\ 6 & 8 & 12 & -5 \end{vmatrix}.$$

Задача звелась до обчислення визначника

$$\Delta_0 = 1 \cdot A_{21} = (-1)^{2+1} \cdot \begin{vmatrix} 3 & 8 & -10 \\ 2 & -1 & 10 \\ 8 & 12 & -5 \end{vmatrix} = -\begin{vmatrix} 3 & 8 & -10 \\ 2 & -1 & 10 \\ 8 & 12 & -5 \end{vmatrix}.$$

Винесемо спільний множник з 3-го стовпця одержаного визначника

$$\Delta_0 = -5 \cdot \begin{vmatrix} 3 & 4 & -2 \\ 2 & -1 & 2 \\ 8 & 6 & -1 \end{vmatrix}$$

і дістанемо визначник

$$\Delta_1 = \begin{vmatrix} 3 & 4 & -2 \\ 2 & -1 & 2 \\ 8 & 6 & -1 \end{vmatrix}.$$

Виділимо в ньому елемент $\alpha_{33} = -1$, що міститься в 3-му рядку і 3-му стовпці. Елементарними перетвореннями типу 2 знайдемо рівний йому визначник

$$\Delta_2 = \begin{vmatrix} -13 & -8 & -2 \\ 18 & 11 & 2 \\ 0 & 0 & -1 \end{vmatrix} = (-1)(-1)^{3+3} \begin{vmatrix} -13 & -8 \\ 18 & 11 \end{vmatrix} = -1.$$

Тому

$$\Delta = \Delta_0 = -5 \cdot \Delta_1 = -5 \cdot \Delta_2 = 5.$$

Варіант 2. Метод зведення до трикутного вигляду. Суть цього перетворення полягає в тому, щоб звести визначник до одного з таких, що

$$\begin{vmatrix} \alpha_{11} & 0 & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \alpha_{n3} & \dots & \alpha_{nn} \end{vmatrix} \quad \text{або} \quad \begin{vmatrix} 0 & \dots & 0 & 0 & \alpha_{1n} \\ 0 & \dots & 0 & \alpha_{2n-1} & \alpha_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{nn-2} & \alpha_{nn-1} & \alpha_{nn} \end{vmatrix}.$$

Приклад 2. Обчислити визначник

$$\Delta = \begin{vmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{vmatrix}$$

n-го порядку.

Додамо до першого рядка решту рядків (серія перетворень типу 2). Матимемо рівний йому визначник

$$\Delta = \begin{vmatrix} n-1 & n-1 & n-1 & \dots & n-1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{vmatrix}.$$

Винесемо спільний множник $n-1$ з першого рядка за знак визначника. Дістанемо $\Delta = (n-1)\Delta_1$, де

$$\Delta_1 = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 0 \end{vmatrix}.$$

Перший рядок цього визначника віднімаємо від кожного з решти рядків (серія перетворень типу 2). Одержимо рівний йому визначник трикутного виду

$$\Delta_1 = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 \end{vmatrix} = (-1)^{n-1}.$$

Остаточно, $\Delta = (-1)^{n-1} \cdot (n-1)$.

§ 2.5. Розклад визначника за елементами рядка або стовпця

Теорема 1. *Визначник*

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}$$

дорівнює сумі добутків елементів будь-якого виділеного рядка на їх алгебричні доповнення:

$$\Delta = \alpha_{i1} A_{i1} + \alpha_{i2} A_{i2} + \dots + \alpha_{in} A_{in}. \tag{1}$$

Аналогічний розклад має місце і для стовпців:

$$\Delta = \alpha_{1k} A_{1k} + \alpha_{2k} A_{2k} + \dots + \alpha_{nk} A_{nk}. \tag{2}$$

Доведення. Для наочності та простоти викладу розглянемо визначник 3-го порядку:

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}.$$

Виділимо в ньому третій рядок і подамо визначник у вигляді

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

За властивістю III визначника (див. § 2.3) і теоремою 2 з § 2.4 матимемо

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

$$= \alpha_{31} A_{31} + \alpha_{32} A_{32} + \alpha_{33} A_{33}.$$

Теорему доведено.

Теорема 2. *Нехай Δ – визначник n-го порядку. Сума добутків елементів деякого його рядка на відповідні алгебричні доповнення елементів будь-якого іншого рядка дорівнює нулю:*

$$\alpha_{i1} A_{j1} + \alpha_{i2} A_{j2} + \ldots + \alpha_{in} A_{jn} = 0, \ i \neq j. \tag{3}$$

Аналогічна властивість має місце і для стовпців:

$$\alpha_{1k} A_{1l} + \alpha_{2k} A_{2l} + \ldots + \alpha_{nk} A_{nl} = 0, \ k \neq l. \tag{4}$$

Доведення. Відмітимо, що алгебричні доповнення елементів деякого рядка визначника

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix}$$

не залежать від елементів цього рядка. Виділимо, наприклад, другий рядок. У формулах для алгебричних доповнень A_{21}, $A_{22}, \ldots, A_{2n}$ не фігурують елементи $\alpha_{21}, \alpha_{22}, \ldots, \alpha_{2n}$ другого рядка. Якщо замість цих елементів поставити елементи першого рядка, тоді дістанемо новий визначник

$$\Delta' = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{11} & \alpha_{12} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix}$$

тотожно рівний нулю (властивість VI). Алгебричні доповнення елементів другого рядка цього визначника залишаться такими самими, як і у визначнику Δ.

На підставі теореми 1 маємо рівність

$$\alpha_{11}A_{21} + \alpha_{12}A_{22} + \ldots + \alpha_{1n}A_{2n} = 0,$$

тобто сума добутків елементів першого рядка на відповідні алгебричні доповнення елементів другого рядка дорівнює нулю, що й треба було довести.

Вправа

Квадратна матриця A називається клітинно-трикутною, якщо вона має вигляд

$$A = \left(\begin{array}{c|c} B & 0 \\ \hline D & C \end{array}\right),$$

де

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \beta_{n1} & \beta_{n2} & \ldots & \beta_{nn} \end{pmatrix}, \quad C = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \ldots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \ldots & \gamma_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \gamma_{m1} & \gamma_{m2} & \ldots & \gamma_{mm} \end{pmatrix},$$

$$D = \begin{pmatrix} \delta_{11} & \delta_{12} & \ldots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \ldots & \delta_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \delta_{m1} & \delta_{m2} & \ldots & \delta_{mn} \end{pmatrix}$$

(D має m рядків і n стовпців) і

$$0 = \begin{pmatrix} 0 & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & 0 \end{pmatrix} \text{ – нульова матриця.}$$

Довести, що

$$\det A = \det B \det C .$$

Вказівка. Доведення провести індукцією за n і використати теорему 1.

Метод зворотних співвідношень

Тут розглядаються визначники над числовими полями і кільцями многочленів. Для їх обчислення часто використовують метод зворотних співвідношень.

Означення. *Послідовність чисел*

$$u_1, u_2, u_3, \dots, u_n, u_{n+1}, \dots$$

називається зворотною, якщо для неї існує таке натуральне число k і такий набір чисел $a_1, a_2, \dots, a_k$ із основного поля, що має місце рівність

$$u_{n+k} = a_1 u_{n+k-1} + a_2 u_{n+k-2} + \dots + a_k u_n \qquad (5)$$

для будь-яких номерів $n = 1, 2, \dots$ (або починаючи з деякого $n \geq N$).

Співвідношення (5) називається зворотним рівнянням, а натуральне число k зворотної послідовності (5) – її порядком.

Наведемо приклади зворотних послідовностей порядків 1, 2 і 3.

Геометрична прогресія. Послідовність

$$u_1 = a\,,\ u_2 = aq\,, \dots,\ u_n = u_{n-1}q\,, \dots$$

визначається рівнянням

$$u_{n+1} = u_n q \qquad (6)$$

причому $k = 1$.

Арифметична прогресія. Послідовність

$$u_1 = a\,,\ u_2 = u_1 + d\,, \dots,\ u_n = u_{n-1} + d\,, \dots$$

така, що $u_n = a + (n-1)d$. Легко переконатися, що вона визначається зворотним рівнянням

$$u_{n+2} = 2u_{n+1} - u_n\,. \qquad (7)$$

Для неї $k = 2$. Арифметична прогресія – зворотна послідовність 2-го порядку.

Послідовність квадратів:

$$u_1 = 1\,,\ u_2 = 2^2,\ u_3 = 3^2, \dots,\ u_n = n^2, \dots$$

Для неї маємо

$$u_{n+1} = (n+1)^2 = n^2 + 2n + 1 = u_n + 2n + 1\,,$$

$$u_{n+2} = u_{n+1} + 2n + 3\,,\ u_{n+2} - u_{n+1} = u_{n+1} - u_n + 2\,.$$

Тому

$$u_{n+2} = 2u_{n+1} - u_n + 2,$$
$$u_{n+3} = 2u_{n+2} - u_{n+1} + 2$$

і

$$u_{n+3} - u_{n+2} = 2u_{n+2} - 3u_{n+1} + u_n.$$

Звідси дістаємо зворотне рівняння 3-го порядку

$$u_{n+3} = 3u_{n+2} - 3u_{n+1} + u_n. \tag{8}$$

Зупинимося детальніше на зворотних послідовностях 2-го порядку.

Нехай послідовність

$$u_1, u_2, u_3, \ldots, u_n, u_{n+1}, \ldots \tag{U}$$

визначається співвідношенням виду

$$u_n = pu_{n-1} + qu_{n-2} \ (n \geq 3). \tag{9}$$

Рівняння

$$x^2 - px - q = 0 \tag{10}$$

називається характеристичним для послідовності (U).

Слід зазначити спільні властивості послідовностей такого виду.

Якщо дві послідовності

$$x_1, x_2, x_3, \ldots, x_n, x_{n+1}, \ldots \tag{X}$$

$$y_1, y_2, y_3, \ldots, y_n, y_{n+1}, \ldots \tag{Y}$$

задовольняють співвідношення (9), то для будь-яких чисел a і b третя послідовність

$$z_1, z_2, z_3, \ldots, z_n, z_{n+1}, \ldots, \tag{Z}$$

де $z_n = ax_n + by_n$, також задовольняє співвідношення (9). Тепер спробуємо підібрати таке число $\alpha \neq 0$, для якого геометрична прогресія

$$1, \alpha, \alpha^2, \ldots, \alpha^n, \ldots$$

задовольняла б те саме співвідношення (9), тобто

$$\alpha^n = p\alpha^{n-1} + q\alpha^{n-2} \ (n \geq 2).$$

Взявши до уваги рівняння (10), матимемо рівність $\alpha^2 - p\alpha - q = 0$. Нехай многочлен $x^2 - px - q$ має два різні корені α і β, тоді геометрична прогресія

$$1, \beta, \beta^2, ..., \beta^n, ...$$

має також задовольняти співвідношення (9).

Побудуємо нову послідовність

$$\gamma_1, \gamma_2, \gamma_3, ..., \gamma_n, ..., \tag{11}$$

де $\gamma_1 = \alpha + \beta$, $\gamma_2 = a\alpha + b\beta, \ldots$, $\gamma_n = a\alpha^{n-1} + b\beta^{n-1}$, і a, b – неозначені коефіцієнти. Вона задовольняє те саме співвідношення (9). Відшукаємо a і b таким чином, щоб виконувались рівності

$$a + b = u_1, \; a\alpha + b\beta = u_2, \tag{12}$$

де u_1, u_2 – перші два члени початкової послідовності (U).

Оскільки $\alpha \neq \beta$, система (12) з невідомими a, b має єдиний розв'язок. Але послідовність (U) визначається єдиним способом через задання початкових членів u_1, u_2 і рівнянням (9). Тому при знайдених значеннях коефіцієнтів послідовності (U) і (11) збігаються і, отже, u_n подається через корені характеристичного рівняння (10) так:

$$u_n = a\alpha^{n-1} + b\beta^{n-1}.$$

Розглянемо тепер випадок, коли характеристичне рівняння (10) має лише один корінь α (кратності 2). Це може статись тоді, коли його дискримінант $p^2 + 4q = 0$. Тоді $\alpha = \dfrac{p}{2}$.

Складемо послідовність

$$1 \cdot 0, \alpha \cdot 1, \alpha \cdot 2, ..., \alpha^{n-1} \cdot (n-1), \ldots$$

Можна переконатись, що вона задовольняє зворотне співвідношення (9). Тоді послідовність

$$\delta_1, \delta_2, \delta_3, ..., \delta_n, ...,$$

де $\delta_n = a\alpha^{n-1} + b\alpha^{n-1}(n-1)$ з неозначеними коефіцієнтами a, b задовольняє співвідношення (9).

Знайдемо a і b такі, щоб виконувались рівняння системи

$$a + 0 \cdot b = u_1 , \ a\alpha + b\alpha = u_2 .$$

При $\alpha \neq 0$ ця система має єдиний розв'язок. Звідси для членів послідовності (U) випливає формула

$$u_n = \alpha^{n-1}[a + b(n-1)] .$$

Все, про що мовилось вище, залишається в силі і для послідовності многочленів

$$u_1(x), \ u_2(x), u_3(x), \ldots, u_n(x), \ \ldots ,$$

які визначаються зворотними рівняннями

$$u_n(x) = p(x) u_{n-1}(x) + q(x) u_{n-2}(x) \ (n > 2) ,$$

де $u_1(x), u_2(x), p(x), q(x)$ – задані многочлени.

При обчисленні деяких типів визначників з'являються зворотні рівняння. Використання їх і складає метод зворотних співвідношень. Проілюструємо це на прикладах.

Приклад 1. Обчислити визначник n-го порядку:

$$\Delta_n = \begin{vmatrix} 1 & 1 & 0 & \ldots & 0 & 0 \\ -1 & 1 & 1 & \ldots & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -1 & 1 \end{vmatrix} .$$

На його головній діагоналі містяться $\alpha_{ii} = 1$, нижче діагоналі $\alpha_{ii-1} = -1$, а вище – $\alpha_{ii+1} = 1$, на решті місць – нулі. Розкладемо Δ_n за елементами 1-го рядка. Дістанемо співвідношення

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

для послідовності

$$\Delta_1, \Delta_2, \Delta_3, \ldots, \Delta_n, \ldots ,$$

причому $\Delta_1 = 1$, $\Delta_2 = 2$. Ця послідовність називається *послідовністю чисел Фібоначчі.*

Характеристичне рівняння $x^2 - x - 1 = 0$ має корені

$$\alpha = \frac{1+\sqrt{5}}{2},\ \beta = \frac{1-\sqrt{5}}{2}.$$

Тому спільний член Δ_n послідовності можна подати у вигляді

$$\Delta_n = a\alpha^{n-1} + b\beta^{n-1},$$

де коефіцієнти a і b мають задовольняти системі:

$$\begin{cases} a + b = \Delta_1 = 1, \\ a\alpha + b\beta = \Delta_2 = 2. \end{cases}$$

Тоді

$$a = \frac{\sqrt{5}+3}{2\sqrt{5}},\ b = \frac{\sqrt{5}-3}{2\sqrt{5}}$$

і

$$\Delta_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right].$$

Приклад 2. Обчислити визначник n-го порядку

$$\Delta_n = \begin{vmatrix} \alpha+\beta & \alpha\cdot\beta & 0 & \dots & 0 \\ 1 & \alpha+\beta & \alpha\cdot\beta & \dots & 0 \\ 0 & 1 & \alpha+\beta & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha+\beta \end{vmatrix}.$$

Маємо $\Delta_1 = \alpha + \beta$, $\Delta_2 = \alpha^2 + \beta^2 + \alpha\beta$. Розклавши визначник Δ_n за елементами 1-го рядка, дістанемо зворотне рівняння

$$\Delta_n = (\alpha+\beta)\Delta_{n-1} - \alpha\cdot\beta\Delta_{n-2}$$

і відповідне йому характеристичне рівняння $x^2 - (\alpha+\beta)x + {}$ $+\alpha\beta = 0$ з коренями α і β.

Випадок 1. Нехай $\alpha \neq \beta$. Тоді $\Delta_n = a\alpha^{n-1} + b\beta^{n-1}$. Для a і b складаємо систему

$$\begin{cases} a + b = \Delta_1 = \alpha + \beta, \\ a\alpha + b\beta = \Delta_2 = \alpha^2 + \beta^2 + \alpha\beta. \end{cases}$$

Отже,

$$a = \frac{\alpha^2}{\alpha-\beta},\ b = \frac{\beta^2}{\beta-\alpha},\ \Delta_n = \frac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}.$$

Випадок 2. Нехай $\alpha = \beta$. Подамо Δ_n у вигляді $\alpha^{n-1}(a+bn)$ і знаходимо $a = b = \alpha$. Тоді

$$\Delta_n = \alpha^n(n+1).$$

Вправи

1. Обчислити визначник n-го порядку

$$\Delta_n = \begin{vmatrix} 5 & 3 & 0 & 0 & \dots & 0 & 0 \\ 2 & 5 & 3 & 0 & \dots & 0 & 0 \\ 0 & 2 & 5 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 2 & 5 \end{vmatrix}.$$

2. Довести, що визначник n-го порядку

$$\Delta_n = \begin{vmatrix} \cos\varphi & 1 & 0 & \dots & 0 \\ 1 & 2\cos\varphi & 1 & \dots & 0 \\ 0 & 1 & 2\cos\varphi & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2\cos\varphi \end{vmatrix} = \cos n\varphi.$$

Вказівка. Розкласти Δ_n за елементами останнього рядка та знайти зворотне співвідношення.

3. Показати, що

$$\Delta_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \prod_{n\ge i>k\ge 1}(x_i - x_k)$$

(визначник Вандермонда).

4. Індукцією за числом n обчислити визначник

$$\Delta_n = \begin{vmatrix} x+x_1 & x & x & \dots & x \\ x & x+x_2 & x & \dots & x \\ x & x & x+x_3 & \dots & x \\ \dots & \dots & \dots & \dots & \dots \\ x & x & x & \dots & x+x_n \end{vmatrix}.$$

Відповідь: $\Delta_n = x_1 x_2 \cdots x_n \left(\frac{x}{x_1} + \frac{x}{x_2} + \ldots + \frac{x}{x_n} + 1 \right)$.

5. Довести, що

$$\Delta_n = \begin{vmatrix} (x_1+y_1)^{-1} & (x_1+y_2)^{-1} & \dots & (x_1+y_n)^{-1} \\ (x_2+y_1)^{-1} & (x_2+y_2)^{-1} & \dots & (x_2+y_n)^{-1} \\ \dots & \dots & \dots & \dots \\ (x_n+y_1)^{-1} & (x_n+y_2)^{-1} & \dots & (x_n+y_n)^{-1} \end{vmatrix} =$$

$$= \frac{\prod\limits_{i>k} (x_i - x_k)(y_i - y_k)}{\prod\limits_{i,k=1} (x_i + y_k)}$$

(визначник Коші).

§ 2.6. Система лінійних рівнянь. Теорема Крамера

Розглянемо систему лінійних рівнянь з коефіцієнтами із довільного поля P і невідомими $x_1, x_2, \ldots, x_n$:

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = \beta_2, \\ \dots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n = \beta_m. \end{cases} \qquad (1)$$

Матриця

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \hdotsfor{4} \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix},$$

складена з коефіцієнтів α_{ik} при невідомих $x_1, x_2, \dots, x_n$, називається *основною*, а матриця

$$\overline{A} = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2 \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m \end{array}\right)$$

називається *розширеною* матрицею системи (1).

Розв'язком системи (1) називається будь-яка скінченна послідовність $l = (\lambda_1, \lambda_2, \dots, \lambda_n)$ елементів $\lambda_j \ (j = \overline{1, n})$ поля P, яка задовольняє рівняння цієї системи і яка обертає їх у тотожності, тобто

$$\alpha_{i1}\lambda_1 + \alpha_{i2}\lambda_2 + \dots \alpha_{in}\lambda_n \equiv \beta_i, \ i = \overline{1, m}.$$

Система (1), яка має хоч би один розв'язок, називається *сумісною*, в противному разі – *несумісною*.

Сумісна система лінійних рівнянь називається означеною, якщо вона має лише один розв'язок. В іншому випадку вона називається *неозначеною*.

Система лінійних рівнянь (1) називається *однорідною*, якщо праві частини її містять лише нулі, тобто

$$\beta_1 = \beta_2 = \dots = \beta_m = 0.$$

Нехай поряд з системою (1) маємо іншу систему рівнянь з тими самими невідомими $x_1, x_2, \dots, x_n$:

$$\begin{cases} \gamma_{11}x_1 + \gamma_{12}x_2 + \dots + \gamma_{1n}x_n = \delta_1, \\ \gamma_{21}x_1 + \gamma_{22}x_2 + \dots + \gamma_{2n}x_n = \delta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \gamma_{r1}x_1 + \gamma_{r2}x_2 + \dots + \gamma_{rn}x_n = \delta_r. \end{cases} \quad (2)$$

Означення. *Система* (2) *називається наслідком системи* (1), *якщо будь-який розв'язок системи* (1) *є одночасно і розв'язком системи* (2).

Надалі символічно таке відношення систем записуватимемо у вигляді $(1) \Rightarrow (2)$.

Означення. *Системи* (1) *і* (2) *називаються рівносильними, якщо будь-який розв'язок однієї з них є одночасно розв'язком другої, або якщо обидві системи несумісні.*

Символічно таке відношення записуватимемо у вигляді $(1) \Leftrightarrow (2)$.

Зрозуміло, що перетворення системи (1) шляхом застосування дій почленного додавання і віднімання, а також множення рівнянь на скаляри (елементи основного поля P) дає змогу одержати нову систему лінійних рівнянь, яка буде наслідком заданої, але не обов'язково рівносильною їй.

Означення. *Елементарними перетвореннями системи лінійних рівнянь* (1) *називаються перетворення таких трьох типів*:

1. Множення правої та лівої частин будь-якого рівняння системи на елемент $\lambda \neq 0$ *поля* P.

2. Додавання до правої та лівої частин будь-якого рівняння системи відповідно правої та лівої частин другого рівняння цієї системи, помножених на елемент $\lambda \in P$.

3. Перестановка місцями двох будь-яких рівнянь системи.

Означення. *Нехай задано матрицю*

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

з елементами β_{ik} *з поля* P. *Елементарними перетвореннями її називаються перетворення таких трьох типів:*

1. Множення всіх елементів деякого рядка на елемент $\lambda \neq 0$ *основного поля* P.

2. Додавання до елементів деякого рядка матриці відповідних елементів іншого рядка, помножених на елемент λ *поля* P.

3. Перестановка двох будь-яких рядків.

Нехай A – основна матриця системи лінійних рівнянь (1), а

$$\overline{A} = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2 \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m \end{array}\right)$$

– її розширена матриця.

Будь-яке елементарне перетворення системи рівнянь (1) відтворюється на матрицях A і $\overline{A}$ у вигляді їх аналогічного елементарного перетворення. Пряме й обернене твердження: будь-яке елементарне перетворення розширеної матриці $\overline{A}$ знаходить відображення у вигляді аналогічних дій з рівняннями системи. Зокрема, якщо (1) – система однорідних рівнянь, тоді така відповідність перетворень вірна для неї та основної матриці A.

Теорема 1. *Нехай з системи лінійних рівнянь* (1) *за допомогою послідовних елементарних перетворень одержали систему* (2). *Тоді системи* (1) *і* (2) *– рівносильні.*

Доведення. Доведемо теорему для випадку, коли система (2) одержується із системи (1) за допомогою одного елементарного перетворення. Припустимо, що таким елементарним перетворенням є перетворення типу 2. Нехай, наприклад, до лівої та правої частин першого рівняння додали відповідно ліву і праву частини другого, помножені на λ. Тоді система (2) набуде вигляду

$$\begin{cases} (\alpha_{11} + \lambda\alpha_{21})x_1 + (\alpha_{12} + \lambda\alpha_{22})x_2 + \dots + (\alpha_{1n} + \lambda\alpha_{2n})x_n = \beta_1 + \lambda\beta_2, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n = \beta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n = \beta_m. \end{cases}$$

Зрозуміло, що $(1) \Rightarrow (2)$. Але із системи (2) задана система одержується за допомогою подібного перетворення: потрібно до першого рівняння із системи (2) додати друге рівняння, помножене на $-\lambda$. Отже, $(1) \Leftrightarrow (2)$.

Теорему доведено.

Якщо система (1) перетворена шляхом декількох елементарних перетворень, в результаті яких одержали нову систему

(2), то практично таку трансформацію простіше здійснити, якщо відразу залучити для цієї мети розширену матрицю $\overline{A}$. Припустимо, що за допомогою декількох послідовних елементарних перетворень вона трансформована в нову матрицю

$$\overline{C} = \left(\begin{array}{cccc|c} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & \delta_1 \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} & \delta_2 \\ \multicolumn{4}{c|}{\dotfill} & \ldots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mn} & \delta_m \end{array}\right).$$

Позначимо зв'язок між цими матрицями знаком еквівалентності, тобто $\overline{A} \sim \overline{C}$. Тоді система (1) рівносильна системі

$$\begin{cases} \gamma_{11}x_1 + \gamma_{12}x_2 + \ldots + \gamma_{1n}x_n = \delta_1, \\ \gamma_{21}x_1 + \gamma_{22}x_2 + \ldots + \gamma_{2n}x_n = \delta_2, \\ \dotfill \\ \gamma_{m1}x_1 + \gamma_{m2}x_2 + \ldots + \gamma_{mn}x_n = \delta_m \end{cases} \tag{2}$$

з розширеною матрицею $\overline{C}$.

Тимчасово обмежимось вивченням систем лінійних рівнянь, у яких число невідомих n дорівнює числу рівнянь, тобто систем виду

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = \beta_2, \\ \dotfill \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n = \beta_n \end{cases} \tag{3}$$

з коефіцієнтами $\alpha_{ik}, \beta_j \in P,\ i, k, j = \overline{1, n}$. Позначимо через

$$\Delta = \det A = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{vmatrix}$$

основний визначник системи.

Теорема 2 (Крамера). *Якщо система n лінійних рівнянь* (3) *з n невідомими має основний визначник* $\Delta = \det A \neq 0$, *то вона сумісна і визначена.*

Доведення. Перетворимо систему так: праву і ліву частини i-го рівняння помножимо на A_{i1} – алгебричне доповнення елемента α_{i1} визначника Δ. Усі одержані рівняння

$$\alpha_{i1}A_{i1}x_1 + \alpha_{i2}A_{i1}x_2 + \ldots + \alpha_{in}A_{i1}x_n = \beta_i A_{i1}$$

для $i = \overline{1, n}$ підсумуємо:

$$x_1\left(\sum_{i=1}^{n}\alpha_{i1}A_{i1}\right) + x_2\left(\sum_{i=1}^{n}\alpha_{i2}A_{i1}\right) + \ldots + x_n\left(\sum_{i=1}^{n}\alpha_{in}A_{i1}\right) = \sum_{i=1}^{n}\beta_i A_{i1}.$$

Оскільки

$$\sum_{i=1}^{n}\alpha_{i1}A_{i1} = \Delta,\ \sum_{i=1}^{n}\alpha_{i2}A_{i1} = 0,\ \ldots,\ \sum_{i=1}^{n}\alpha_{in}A_{i1} = 0,$$

матимемо

$$\Delta \cdot x_1 = \beta_1 A_{11} + \beta_2 A_{21} + \ldots + \beta_n A_{n1}.$$

Права частина цього рівняння дорівнює визначнику

$$\Delta_1 = \begin{vmatrix} \beta_1 & \alpha_{12} & \ldots & \alpha_{1n} \\ \beta_2 & \alpha_{22} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \beta_n & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix}.$$

Аналогічно із системи дістаємо рівняння

$$\Delta \cdot x_2 = \Delta_2 = \begin{vmatrix} \alpha_{11} & \beta_1 & \alpha_{13} & \ldots & \alpha_{1n} \\ \alpha_{21} & \beta_2 & \alpha_{23} & \ldots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \beta_n & \alpha_{n3} & \ldots & \alpha_{nn} \end{vmatrix}$$

і т. д. Отже, в результаті матимемо нову систему лінійних рівнянь

$$\Delta \cdot x_1 = \Delta_1,\ \Delta \cdot x_2 = \Delta_2,\ \ldots,\ \Delta \cdot x_n = \Delta_n, \tag{4}$$

де $\Delta_i (i = \overline{1,n})$ – визначники, що одержуються з визначника Δ заміною елементів i-го стовпця на відповідні елементи $\beta_k (k = \overline{1,n})$ правих частин заданої системи рівнянь. При цьому (3) $\Rightarrow$ (4). Система (4) має лише єдиний розв'язок

$$\lambda_1 = \frac{\Delta_1}{\Delta},\ \lambda_2 = \frac{\Delta_2}{\Delta},\ \ldots,\ \lambda_n = \frac{\Delta_n}{\Delta}. \tag{5}$$

Тепер доведемо сумісність системи (3). З'ясовується, що одержана послідовність $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ елементів основного поля буде розв'язком цієї системи.

Дійсно, підставивши в перше рівняння елементи λ_k замість невідомих x_k, отримаємо

$$\alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + \ldots + \alpha_{1n}\lambda_n =$$
$$= \alpha_{11} \cdot \frac{\Delta_1}{\Delta} + \alpha_{12} \cdot \frac{\Delta_2}{\Delta} + \ldots + \alpha_{1n} \cdot \frac{\Delta_n}{\Delta} = \frac{1}{\Delta}\left(\sum_{k=1}^{n} \alpha_{ik}\Delta_k\right).$$

Розкладемо Δ_k за елементами k-го стовпця:

$$\Delta_k = \beta_1 A_{1k} + \beta_2 A_{2k} + \ldots + \beta_n A_{nk}.$$

Тому

$$\alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + \ldots + \alpha_{1n}\lambda_n = \frac{1}{\Delta}\sum_{k=1}^{n} \alpha_{1k}\left(\sum_{j=1}^{n} \beta_j A_{jk}\right) =$$
$$= \frac{1}{\Delta}\sum_{j=1}^{n} \beta_j \left(\sum_{k=1}^{n} \alpha_{1k} A_{jk}\right).$$

Але

$$\sum_{k=1}^{n} \alpha_{1k} A_{1k} = \Delta,\ \sum_{k=1}^{n} \alpha_{1k} A_{jk} = 0,\ \text{при}\ j \neq 1.$$

Звідси знаходимо

$$\alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + \ldots + \alpha_{1n}\lambda_n = \beta_1.$$

Так само можна переконатися, що $\lambda_1, \lambda_2, \dots, \lambda_n$ задовольняють решту рівнянь системи (3).

Розв'язки системи (5) носять назву *формул Крамера*. Поряд з системою (3) розглянемо систему однорідних рівнянь

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n = 0, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \dots + \alpha_{nn}x_n = 0 \end{cases} \tag{3°}$$

з тією ж основною квадратною матрицею $A = \left(\alpha_{ik}\right)$. Вона має тривіальний (нульовий) розв'язок (0, 0, …, 0). Виникає запитання: за яких умов вона має нетривіальні розв'язки? При цьому її розв'язок ($\lambda_1, \lambda_2, \dots, \lambda_n$) називається *нетривіальним* (ненульовим), якщо в цій послідовності елементів поля існує хоч би один елемент, відмінний від нуля.

Теорема 3. *Для того щоб система* (3°) *лінійних однорідних рівнянь з **n** невідомими мала ненульовий розв'язок, необхідно і достатньо, щоб визначник цієї системи дорівнював нулю:*

$$\Delta = \det A = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = 0.$$

Доведення. *Необхідність.* Припустимо, що визначник системи (3°) відмінний від нуля. Тоді за теоремою Крамера вона має лише один розв'язок – нульовий. Тому, якщо система має ненульові розв'язки, обов'язково $\Delta = 0$.

Достатність. Нехай $\Delta = 0$. Доведення про існування ненульового розв'язку проведемо індукцією за числом n.

Якщо $n = 1$, то система має вигляд $\alpha_{11}x_1 = 0$ і $\Delta = \alpha_{11} = 0$. Ненульові розв'язки у цьому випадку існують (оскільки основне поле P має елементи, відмінні від нуля).

Припустимо, що твердження теореми має місце для систем лінійних однорідних рівнянь з $n-1$ невідомими і $n-1$ рівняннями. Розглянемо систему (3°) з n невідомими. Випишемо її матрицю

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

Розглянемо два випадки.

Випадок 1. Усі елементи першого стовпця матриці A дорівнюють нулю:

$$\alpha_{11}=\alpha_{21}=\dots=\alpha_{n1}=0.$$

Тоді послідовність (1, 0, 0, ..., 0) – ненульовий розв'язок системи.

Випадок 2. Не всі елементи першого стовпця дорівнюють нулю. Нехай, наприклад, $\alpha_{11}\neq 0$. Декількома елементарними перетвореннями типу 2 над рядками матриці A дістаємо нову матрицю (наприклад, до елементів другого рядка додати відповідні елементи першого, помноженого на $-\dfrac{\alpha_{21}}{\alpha_{11}}$ і т. д.):

$$B=\begin{pmatrix}\alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n}\\ 0 & \beta_{22} & \beta_{23} & \dots & \beta_{2n}\\ \dots & \dots & \dots & \dots & \dots\\ 0 & \beta_{n2} & \beta_{n3} & \dots & \beta_{nn}\end{pmatrix}.$$

Система (3°) рівносильна системі

$$\begin{cases}\alpha_{11}x_1+\alpha_{12}x_2+\alpha_{13}x_3+\dots+\alpha_{1n}x_n=0,\\ \qquad\quad\ \beta_{22}x_2+\beta_{23}x_3+\dots+\beta_{2n}x_n=0,\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ \qquad\quad\ \beta_{n2}x_2+\beta_{n3}x_3+\dots+\beta_{nn}x_n=0\end{cases}\qquad(3')$$

з визначником $\Delta'=\det B$. Цей визначник Δ' одержано з визначника Δ за допомогою декількох елементарних перетворень типу 2 над рядками і тому $\Delta'=\Delta=0$. Оскільки

$$\Delta' = \alpha_{11} \cdot \begin{vmatrix} \beta_{22} & \beta_{23} & \cdots & \beta_{2n} \\ \beta_{32} & \beta_{33} & \cdots & \beta_{3n} \\ \dots & \dots & \dots & \dots \\ \beta_{n2} & \beta_{n3} & \cdots & \beta_{nn} \end{vmatrix} \text{ і } \alpha_{11} \neq 0,$$

то

$$\Delta'' = \begin{vmatrix} \beta_{22} & \beta_{23} & \cdots & \beta_{2n} \\ \beta_{32} & \beta_{33} & \cdots & \beta_{3n} \\ \dots & \dots & \dots & \dots \\ \beta_{n2} & \beta_{n3} & \cdots & \beta_{nn} \end{vmatrix} = 0.$$

Застосуємо до системи

$$\begin{cases} \beta_{22}x_2 + \beta_{23}x_3 + \ldots + \beta_{2n}x_n = 0, \\ \beta_{32}x_2 + \beta_{33}x_3 + \ldots + \beta_{3n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_{n2}x_2 + \beta_{n3}x_3 + \ldots + \beta_{nn}x_n = 0 \end{cases} \qquad (3'')$$

припущення індукції. Вона має ненульовий розв'язок

$$(\lambda_2, \lambda_3, \ldots, \lambda_n).$$

Оскільки $\alpha_{11} \neq 0$, то можна підшукати $\lambda_1 \in P$ таким, щоб виконувалась рівність $\alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + \alpha_{13}\lambda_3 + \ldots + \alpha_{1n}\lambda_n = 0$. Тоді послідовність $(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n)$ – ненульовий розв'язок системи $(3')$ і, отже, системи (3°).

Зауваження. Із доведеної теореми випливає ефективний метод знаходження ненульових розв'язків систем лінійних однорідних рівнянь розглянутого виду (3°). Він полягає в тому, що система з невідомими зводиться до системи меншого числа невідомих.

Наслідок. *Система лінійних однорідних рівнянь*

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

має ненульові розв'язки, якщо число n її невідомих перевищує число m її рівнянь.

Систему такого виду можна розширити до квадратної додаванням нових рівнянь з ненульовими коефіцієнтами і тоді твердження буде очевидним наслідком теореми 3.

Вправа

Розв'язати систему рівнянь

$$\begin{cases} 0 \cdot x_1 + x_2 + x_3 + \dots + x_{n-1} + x_n = 0, \\ x_1 + 0 \cdot x_2 + x_3 + \dots + x_{n-1} + x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_1 + x_2 + x_3 + \dots + x_{n-1} + 0 \cdot x_n = 0. \end{cases}$$

Вказівка. Скористатися формулами Крамера.

§ 3.1. Арифметичні лінійні простори

Непорожня множина Λ елементів $a, b, c, ..., x, y, z, ...$ називається лінійним простором над полем P, якщо:

а) дано правило, за яким довільній парі елементів a, b із множини Λ однозначно ставиться у відповідність новий елемент із Λ, що називається сумою двох перших і позначається $a+b$;

б) дано правило, за яким довільному елементу a із множини Λ і довільному елементу λ з поля P однозначно ставиться у відповідність новий елемент із Λ, який називається добутком a на λ і позначається λa;

в) операції додавання і множення підкоряються таким законам:

1) Додавання комутативне: $a+b=b+a$.

2) Додавання асоціативне: $a+(b+c)=(a+b)+c$.

3) Для будь-яких двох елементів a, b із множини Λ існує один і лише один такий елемент x із Λ, що

$$a+x=b.$$

4) Для одиничного елемента 1 із поля P має місце рівність $1a=a$, $a\in\Lambda$.

5) Множення асоціативне: $(\mu\cdot\lambda)a=\mu(\lambda\cdot a)$.

6) Множення підкоряється 1-му розподільному закону:

$$\lambda(a+b)=\lambda a+\lambda b.$$

7) Множення підкоряється 2-му розподільному закону:

$$(\lambda+\mu)a=\lambda a+\mu a.$$

Елементи лінійного простору Λ називаються *векторами*, а елементи поля P – *скалярами*.

Лінійний простір Λ інакше називають *векторним простором* над полем P.

Лінійний простір над полем C комплексних чисел називається *комплексним лінійним простором*, а над полем R дійсних чисел – *дійсним лінійним простором.*

Із законів 1) і 2) випливає, що сума скінченного числа векторів із простору Λ не залежить ні від порядку розміщення доданків, ні від способу розміщення дужок. Наприклад, $(a+b)+(c+d)=\big((a+b)+d\big)+c$. Тому можна записати $a+b+c+d=$ $=a+b+d+c$.

Серед векторів простору єдиний вектор θ, який має властивість $a+\theta=\theta+a=a$ для будь-якого вектора a. Він називається *нульовим* вектором.

Для будь-якого вектора a знайдеться один і лише один вектор x, для якого $a+x=\theta$. Такий вектор x називається *оберненим* для a, і позначається $-a$.

Із основних законів можна вивести такі правила:

$$-\theta=\theta\,,\quad -(-a)=a\,,\quad -(a+b)=(-a)+(-b)\,.$$

Операція віднімання, визначена формулою

$$a-b=a+(-b)\,,$$

називається *відніманням* вектора b від вектора a.

Із законів 4) і 7) випливають такі тотожності:

1) $0a=\theta$, де 0 – нульовий елемент поля P.

2) $ma=\underbrace{a+a+\ldots+a}_{m}$, де m – натуральне число.

3) $(-\lambda)a=-(\lambda a)$.

4) $\lambda\theta=\theta$.

Наприклад, оскільки $\lambda a+(-\lambda)a=(\lambda-\lambda)a=\theta$, маємо $(-\lambda)a=$ $=-(\lambda a)$.

Вираз виду

$$\lambda_1 a_1+\lambda_2 a_2+\ldots+\lambda_m a_m\,,$$

де $a_1,a_2,\ldots,a_m\in\Lambda$ і $\lambda_1,\lambda_2,\ldots,\lambda_m\in P$, називається лінійною комбінацією векторів $a_1,a_2,\ldots,a_m$.

Наведемо конкретні приклади векторних просторів.

1. Множина всіх напрямлених відрізків тривимірного евклідового простору з відомими правилами додавання і множення на дійсне число (з поля R). Ця множина є векторним простором і позначається R^3.

2. Множина $P[x]$ усіх многочленів змінної x з коефіцієнтами з поля P відносно звичайних операцій додавання і множення на скаляр – лінійний простір. Тут роль векторів відіграють многочлени.

3. Нехай P – довільне поле і n – фіксоване натуральне число. Розглянемо будь-які можливі послідовності $a = (\alpha_1, \alpha_2, ..., \alpha_n)$, складені з n елементів, $\alpha_k \in P$.

Якщо $b = (\beta_1, \beta_2, ..., \beta_n)$ – така ж послідовність елементів $\beta_k \in P$, то будемо вважати їх рівними і записувати $a = b$ тоді і лише тоді, коли $\alpha_k = \beta_k$ для будь-яких $k = \overline{1, n}$.

Введемо для цих послідовностей такі операції:

Додавання. Якщо $a = (\alpha_1, \alpha_2, ..., \alpha_n)$ і $b = (\beta_1, \beta_2, ..., \beta_n)$, тоді

$$a + b = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, ..., \alpha_n + \beta_n).$$

Множення на скаляр. Якщо $a = (\alpha_1, \alpha_2, ..., \alpha_n)$ і $\lambda \in P$, тоді

$$\lambda a = (\lambda\alpha_1, \lambda\alpha_2, ..., \lambda\alpha_n).$$

Відносно таких операцій усі послідовності елементів із поля P з однією довжиною n складають лінійний простір. Він називається *арифметичним векторним простором* і позначається P^n.

Для вектора $a = (\alpha_1, \alpha_2, ..., \alpha_n)$ із P^n елементи α_k називаються його *координатами.* Нульовим елементом у ньому є послідовність $\theta = (0, 0, ..., 0)$ поля P^n.

Лінійна залежність векторів простору P^n

Оберемо у лінійному просторі P^n систему векторів

$$a_1 = (\alpha_{11}, \alpha_{12}, ..., \alpha_{1n}),$$
$$a_2 = (\alpha_{21}, \alpha_{22}, ..., \alpha_{2n}),$$
$$\dots\dots\dots\dots\dots\dots$$
$$a_m = (\alpha_{m1}, \alpha_{m2}, ..., \alpha_{mn}).$$

Вектор $c=(\gamma_1,\gamma_2,...,\gamma_n)$ – лінійна комбінація векторів $a_1,a_2,...,a_m$ (або інакше – лінійно подається через $a_1,a_2,...,a_m$), якщо

$$c=\lambda_1 a_1+\lambda_2 a_2+...+\lambda_m a_m \tag{1}$$

для деяких елементів $\lambda_1,\lambda_2,\ldots,\lambda_m$ із поля P.

Рівність (1) можна подати у вигляді системи рівностей для координат

$$\begin{aligned}
\gamma_1 &=\lambda_1\alpha_{11}+\lambda_2\alpha_{21}+...+\lambda_m\alpha_{m1},\\
\gamma_2 &=\lambda_1\alpha_{12}+\lambda_2\alpha_{22}+...+\lambda_m\alpha_{m2},\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
\gamma_n &=\lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+...+\lambda_m\alpha_{mn}.
\end{aligned}$$

Зрозуміло, що нульовий вектор θ лінійно подається через будь-яку систему векторів

$$\theta=(0,0,...,0)=0\cdot a_1+0\cdot a_2+...+0\cdot a_m\,. \tag{2}$$

Означення. *Система векторів $a_1,a_2,...,a_m$ називається л і н і й н о з а л е ж н о ю, якщо існує така ненульова послідовність скалярів $\lambda_1,\lambda_2,...,\lambda_m$, для якої*

$$\lambda_1 a_1+\lambda_2 a_2+...+\lambda_m a_m=(0,0,...,0)=\theta\,.$$

При цьому домовимось називати послідовність $\lambda_1,\lambda_2,...,\lambda_m$ ненульовою, якщо серед цих скалярів існує хоч би один, відмінний від нуля, і записуватимемо

$$(\lambda_1,\lambda_2,...,\lambda_m)\neq\theta\,.$$

Використовуючи символи математичної логіки, означення викладемо у такій формі.

Система векторів $a_1,a_2,...,a_m$ лінійно залежна, якщо

$$\exists\,\lambda_1,\lambda_2,...,\lambda_m\left\{(\lambda_1,\lambda_2,...,\lambda_m)\neq\theta\ \&\ \sum_{j=1}^{m}\lambda_j a_j=\theta\right\}. \tag{3}$$

Сформулюємо найпростіші властивості лінійно залежних векторів.

I. Система векторів, яка містить нульовий вектор, лінійно залежна.

II. Система векторів лінійно залежна, якщо будь-яка її підсистема лінійно залежна.

III. Система векторів $a_1, a_2, \dots, a_m$, $m \geq 2$ лінійно залежна тоді і лише тоді, коли який-небудь з них лінійно подається через решту векторів.

Перші дві властивості встановлюються порівняно легко. Доведемо властивість III.

Необхідність. Нехай вектори $a_1, a_2, \dots, a_m$ лінійно залежні, тобто для них існує ненульова послідовність $\lambda_1, \lambda_2, \dots, \lambda_m$ скалярів така, що $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_m a_m = \theta$. Припустимо, що $\lambda_1 \neq 0$. Тоді матимемо

$$a_1 = \left(-\frac{\lambda_2}{\lambda_1}\right) a_2 + \dots + \left(-\frac{\lambda_m}{\lambda_1}\right) a_m .$$

Достатність. Нехай деякий вектор, наприклад a_1, лінійно подається у вигляді лінійної комбінації решти векторів:

$$a_1 = \lambda_2 a_2 + \dots + \lambda_m a_m .$$

Звідси дістаємо

$$(-1) a_1 + \lambda_2 a_2 + \dots + \lambda_m a_m = \theta .$$

Остання рівність означає лінійну залежність векторів $a_1, a_2, \dots, a_m$.

Приклад. Довести, що три компланарні вектори $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ тривимірного векторного простору R^3 лінійно залежні.

Опишемо ці вектори їх проекціями на осі прямокутної системи координат:

$$\vec{a}_1 = (\alpha_{11}, \alpha_{12}, \alpha_{13}),$$
$$\vec{a}_2 = (\alpha_{21}, \alpha_{22}, \alpha_{23}),$$
$$\vec{a}_3 = (\alpha_{31}, \alpha_{32}, \alpha_{33}).$$

На підставі компланарності цих векторів змішаний добуток $(\vec{a}_1 \vec{a}_2 \vec{a}_3) = 0$, а тому

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

Складемо систему лінійних рівнянь

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 = 0, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3 = 0, \\ \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3 = 0. \end{cases}$$

Визначник цієї системи обертається на нуль і, отже, вона має ненульовий розв'язок λ_1, λ_2, λ_3 (див. § 2.6, теорема 3). Тому маємо тотожні рівності

$$\alpha_{1i}\lambda_1 + \alpha_{2i}\lambda_2 + \alpha_{3i}\lambda_3 \equiv 0,$$

де $i = 1, 2, 3$.

У векторній формі їх можна подати у вигляді однієї рівності

$$\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \lambda_3\vec{a}_3 = \theta,$$

що й вказує на лінійну залежність векторів $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$.

Означення. *Лінійною оболонкою системи векторів $b_1, b_2, ..., b_p$ називається множина всіх векторів лінійного простору P^n, які лінійно подаються через вектори $b_1, b_2, ..., b_p$. Їх лінійну оболонку домовимось позначати $L\ (b_1, b_2, ..., b_p)$.*

Теорема 1 (головна теорема про лінійну залежність). *Нехай $a_1, a_2, ..., a_m$ і $b_1, b_2, ..., b_p$ – дві системи векторів простору P^n. Якщо $a_1, a_2, ..., a_m \in L\ (b_1, b_2, ..., b_p)$ і $m > p$, то вектори $a_1, a_2, ..., a_m$ лінійно залежні.*

Нижче наводяться два доведення цієї теореми. Перше з них пов'язане з поняттям визначника, а друге – без залучення цього поняття.

Доведення 1. За умовою вектори $a_j\ (j = \overline{1,m})$ лінійно подаються через $b_k\ (k = \overline{1,p})$:

$$\left.\begin{aligned} a_1 &= \alpha_{11}b_1 + \alpha_{21}b_2 + \ldots + \alpha_{p1}b_p, \\ a_2 &= \alpha_{12}b_1 + \alpha_{22}b_2 + \ldots + \alpha_{p2}b_p, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_m &= \alpha_{1m}b_1 + \alpha_{2m}b_2 + \ldots + \alpha_{pm}b_p, \end{aligned}\right\}$$

де $\alpha_{kj} \in P$.

Складемо систему лінійних однорідних рівнянь

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1m}x_m = 0, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2m}x_m = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{p1}x_1 + \alpha_{p2}x_2 + \ldots + \alpha_{pm}x_m = 0. \end{cases} \tag{4}$$

Оскільки $m > p$, то ця система має ненульовий розв'язок $\lambda_1, \lambda_2, \ldots, \lambda_m$ (див. § 2.6, теорема 3 і її наслідок). Якщо в системі (4) замість невідомих $x_1, x_2, \ldots, x_m$ підставити скаляри $\lambda_1, \lambda_2, \ldots, \lambda_m$, тоді одержимо систему тотожностей

$$\begin{cases} \alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + \ldots + \alpha_{1m}\lambda_m \equiv 0, \\ \alpha_{21}\lambda_1 + \alpha_{22}\lambda_2 + \ldots + \alpha_{2m}\lambda_m \equiv 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{p1}\lambda_1 + \alpha_{p2}\lambda_2 + \ldots + \alpha_{pm}\lambda_m \equiv 0. \end{cases} \tag{5}$$

З урахуванням системи (5) маємо

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_m a_m = \lambda_1(\alpha_{11}b_1 + \alpha_{21}b_2 + \ldots + \alpha_{p1}b_p) +$$

$$+\lambda_2(\alpha_{12}b_1 + \alpha_{22}b_2 + \ldots + \alpha_{p2}b_p) + \ldots +$$

$$+\lambda_m(\alpha_{1m}b_1 + \alpha_{2m}b_2 + \ldots + \alpha_{pm}b_p) = (\alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + \ldots + \alpha_{1m}\lambda_m)b_1 +$$

$$+(\alpha_{21}\lambda_1 + \alpha_{22}\lambda_2 + \ldots + \alpha_{2m}\lambda_m)b_2 + \ldots +$$

$$+(\alpha_{p1}\lambda_1 + \alpha_{p2}\lambda_2 + \ldots + \alpha_{pm}\lambda_m)b_p = 0 \cdot b_1 + 0 \cdot b_2 + \ldots + 0 \cdot b_p = \theta,$$

що й доводить лінійну залежність векторів $a_1, a_2, \ldots, a_m$.

Доведення 2. Вважатимемо, що серед векторів $a_1, a_2, \ldots, a_m$ відсутній нульовий. Крім того, теорему достатньо довести для випадку, коли $m = p+1$. Отже, $a_1, a_2, ..., a_{p+1} \in L(b_1, b_2, ..., b_p)$ і вектори $a_1, a_2, ..., a_{p+1}$ ненульові. Доведення лінійної залежності векторів $a_1, a_2, ..., a_{p+1}$ проведемо індукцією за числом p. Якщо $p = 1$, тоді $a_1, a_2 \in L(b_1)$ і $a_1 = \alpha b_1$, $a_2 = \beta b_1$, $\alpha \neq 0$, $\beta \neq 0$. Отже, $\alpha^{-1} a_1 + (-\beta^{-1}) a_2 = \theta$, що й означає лінійну залежність a_1 і a_2.

Припустимо, що теорема має місце, коли лінійна оболонка породжена $q = p-1$ векторами. Доведемо справедливість теореми на випадок, коли лінійна оболонка породжена p векторами.

За умовою

$$\left.\begin{array}{l} a_1 = \alpha_{11} b_1 + \alpha_{21} b_2 + ... + \alpha_{p1} b_p, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_p = \alpha_{1p} b_1 + \alpha_{2p} b_2 + ... + \alpha_{pp} b_p, \\ a_{p+1} = \alpha_{1p+1} b_1 + \alpha_{2p+1} b_2 + ... + \alpha_{pp+1} b_p. \end{array}\right|$$

Якщо при цьому коефіцієнти $\alpha_{p1} = \alpha_{p2} = ... = \alpha_{pp} = \alpha_{pp+1} = 0$, то матимемо включення $a_1, ..., a_p, a_{p+1} \in L(b_1, ..., b_{p-1})$ і за припущенням індукції вектори $a_1, a_2, ..., a_p$ лінійно залежні, а тому лінійно залежні й вектори $a_1, ..., a_p, a_{p+1}$.

Припустимо, що деякий з цих коефіцієнтів при векторі b_p відмінний від нуля, наприклад, $\alpha_{pp+1} \neq 0$. Тоді виключимо вектор b_p з перших p рівностей, віднімаючи від них останню рівність, помножену на деякі скаляри. Дістаємо систему рівностей

$$\begin{cases} a_1 - \lambda_1 a_{p+1} = \beta_{11} b_1 + ... + \beta_{p-1\,1} b_{p-1}, \\ a_2 - \lambda_2 a_{p+1} = \beta_{12} b_1 + ... + \beta_{p-1\,2} b_{p-1}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_p - \lambda_p a_{p+1} = \beta_{1p} b_1 + ... + \beta_{p-1p} b_{p-1}. \end{cases}$$

Це означає, що маємо включення

$$a_1-\lambda_1a_{p+1},\ a_2-\lambda_2a_{p+1},\ \ldots,\ a_p-\lambda_pa_{p+1}\in L(b_1,b_2,\ldots,b_{p-1}).$$

За припущенням індукції вектори $a_1-\lambda_1a_{p+1}, a_2-\lambda_2a_{p+1}, \ldots,$ $a_p-\lambda_pa_{p+1}$ лінійно залежні. Для них існує ненульова послідовність скалярів $\gamma_1,\gamma_2,\ldots,\gamma_p$ таких, що

$$\gamma_1(a_1-\lambda_1a_{p+1})+\gamma_2(a_2-\lambda_2a_{p+1})+\ldots+\gamma_p(a_p-\lambda_pa_{p+1})=\theta.$$

Звідси дістаємо

$$\gamma_1a_1+\gamma_2a_2+\ldots+\gamma_pa_p+\gamma a_{p+1}=\theta,$$

де $\gamma=-\gamma_1\lambda_1-\gamma_2\lambda_2-\ldots-\gamma_p\lambda_p$.

Оскільки послідовність $\gamma_1,\gamma_2,\ldots,\gamma_p,\gamma$ – ненульова, то вектори $a_1,a_2,\ldots,a_p,a_{p+1}$ лінійно залежні.

Теорему доведено.

Теорема 2*. Система векторів* $a_1,a_2,\ldots,a_m$ *лінійного простору* P^n *лінійно залежна, якщо* $m>n$.

Доведення. Виділимо в просторі P^n систему одиничних векторів:

$$\begin{aligned} e_1&=(1,0,0,\ldots,0),\\ e_2&=(0,1,0,\ldots,0),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\\ e_n&=(0,0,0,\ldots,1). \end{aligned} \tag{6}$$

Будь-який вектор $x=(\xi_1,\xi_2,\ldots,\xi_n)$ подається у вигляді лінійної комбінації векторів e_j ($j=\overline{1,n}$):

$$x=\xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n.$$

Тому $a_1,a_2,\ldots,a_m\in L(e_1,e_2,\ldots,e_n)$. За теоремою 1 вектори $a_1,a_2,\ldots,a_m$ лінійно залежні, оскільки за умовою $m>n$.

Теорему доведено.

Вправи

1. Довести, що система векторів e_j ($j=\overline{1,n}$) із (6) – лінійно незалежна.

2. Довести, що вектори

$$a_1 = (1, 1, -1, 1),$$
$$a_2 = (1, -1, 1, -1),$$
$$a_3 = (-1, 1, 1, 1),$$
$$a_4 = (1, 1, 1, 1)$$

простору R^4 лінійно залежні.

Лінійна незалежність векторів простору

Означення. *Вектори $a_1, a_2, ..., a_m$ простору P^n називаються лінійно незалежними, якщо для будь-якої ненульової послідовності $\lambda_1, \lambda_2, ..., \lambda_m$ скалярів виконується нерівність*

$$\lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_m a_m \neq \theta.$$

Це означення еквівалентне такому.

Вектори $a_1, a_2, ..., a_m$ лінійно незалежні, коли з рівності $\lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_m a_m = \theta$ випливає, що $\lambda_1 = \lambda_2 = ... = \lambda_m = 0$.

Наведене означення вважатимемо головним.

Приклад 1. Нехай визначник n-го порядку

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{vmatrix} \neq 0.$$

Розглянемо стовпці цього визначника як вектори

$a_1 = (\alpha_{11}, \alpha_{21}, ..., \alpha_{n1})$, $a_2 = (\alpha_{12}, \alpha_{22}, ..., \alpha_{n2})$, ..., $a_n = (\alpha_{1n}, \alpha_{2n}, ..., \alpha_{nn})$ простору P^n.

Припустимо, що $\lambda_1, \lambda_2, ..., \lambda_n$ такі скаляри, для яких

$$\lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_n a_n = \theta \Rightarrow$$

$$\Rightarrow \begin{cases} \alpha_{11}\lambda_1 + \alpha_{12}\lambda_2 + ... + \alpha_{1n}\lambda_n = 0, \\ \alpha_{21}\lambda_1 + \alpha_{22}\lambda_2 + ... + \alpha_{2n}\lambda_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{n1}\lambda_1 + \alpha_{n2}\lambda_2 + ... + \alpha_{nn}\lambda_n = 0. \end{cases}$$

Оскільки визначник останньої системи відмінний від нуля, то за теоремою Крамера матимемо, що $\lambda_1 = \lambda_2 = ... = \lambda_n = 0$. Отже, стовпці лінійно незалежні. Аналогічно доводиться, що й рядки цього визначника також лінійно незалежні.

Приклад 2. Нехай задано деяку систему векторів з простору P^n:

$$a_i = (\alpha_{i1}, ..., \alpha_{iq}, \alpha_{iq+1}, ..., \alpha_{in}), \quad i = \overline{1,m}.$$

У цій системі відкинемо координати, починаючи з номера $q+1$. Дістанемо вкорочені вектори з простору P^q:

$$\underline{a}_i = (\alpha_{i1}, \alpha_{i2}, ..., \alpha_{iq}), \quad i = \overline{1,m}.$$

Покажемо, що коли ці вектори $\underline{a}_1, \underline{a}_2, ..., \underline{a}_m$ лінійно незалежні, то і первинні вектори $a_1, a_2, ..., a_m$ також лінійно незалежні.

Складемо рівність $\lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_m a_m = \theta$ з деякими коефіцієнтами $\lambda_j \in P$. Цю рівність можна подати у вигляді n рівностей, перейшовши до координатного запису:

$$\begin{cases} \lambda_1\alpha_{11} + \lambda_2\alpha_{21} + ... + \lambda_m\alpha_{m1} = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \lambda_1\alpha_{1q} + \lambda_2\alpha_{2q} + ... + \lambda_m\alpha_{mq} = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \lambda_1\alpha_{1n} + \lambda_2\alpha_{2n} + ... + \lambda_m\alpha_{mn} = 0. \end{cases}$$

Виділимо з них перші q рівностей. У векторній формі вони задають одну рівність: $\lambda_1\underline{a}_1 + \lambda_2\underline{a}_2 + ... + \lambda_m\underline{a}_m = \theta$. Тому $\lambda_1 = \lambda_2 = ... = \lambda_m = 0$.

Окремо відмітимо деякі спільні властивості лінійно незалежних векторів.

I. Якщо система векторів простору P^n лінійно незалежна, то будь-яка її підсистема також лінійно незалежна.

II. Система векторів $a_1, a_2, ..., a_m$, $m \geq 2$ лінійно незалежна тоді і лише тоді, коли кожен з її векторів лінійно не подається через решту векторів цієї системи.

III. Якщо вектори $a_1, a_2, ..., a_m$ лінійно незалежні і $a_1, a_2, ..., a_m \in L(b_1, b_2, ..., b_p)$, тоді $m \le p$.

IV. Якщо $a_1, a_2, ..., a_m$ – лінійно незалежні вектори лінійного простору P^n, тоді $m \le n$.

Ці властивості – безпосередні наслідки доведених вище властивостей лінійно залежних векторів.

Ступеневі матриці

Нехай $A = (\alpha_{ik})$ – прямокутна матриця розмірності $m \times n$ і $\alpha_{ik} \in P$. *Провідним елементом* рядка цієї матриці називається перший (рахуючи зліва направо) ненульовий елемент цього рядка.

Матриця A називається *ступеневою*, якщо її рядки мають такі властивості:

а) нульові рядки (якщо вони є) розміщені нижче всіх ненульових рядків;

б) якщо $\alpha_{1k_1}, \alpha_{2k_2}, ..., \alpha_{rk_r}$ – провідні елементи ненульових рядків, тоді $k_1 < k_2 < ... < k_r$.

Приклади ступеневих матриць.

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & 0 & \alpha_{23} & \alpha_{24} \\ 0 & 0 & 0 & \alpha_{34} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

де $\alpha_{11} \ne 0$, $\alpha_{23} \ne 0$, $\alpha_{34} \ne 0$.

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \alpha_{nn} \end{pmatrix},$$

де $\alpha_{11} \ne 0$, $\alpha_{22} \ne 0$, ..., $\alpha_{nn} \ne 0$.

Матриця A називається *псевдоступеневою*, якщо з неї перестановкою рядків можна дістати ступеневу матрицю.

Вправи

1. Довести, що ненульові рядки ступеневої та псевдоступеневої матриць лінійно незалежні.

2. Переконатися, що коли кожна з систем векторів $a_1, a_2, ..., a_m$ і $b_1, b_2, ..., b_p$ лінійно незалежна і $L\ (a_1, a_2, ..., a_m) = L\ (b_1, b_2, ..., b_p)$, тоді $m = p$.

§ 3.2. Базис і ранг підмножини лінійного простору

Означення. *Дві системи $A = < a_1, a_2, \dots, a_m >$ і $B = < b_1, b_2, \dots, b_p >$ векторів лінійного простору P^n називаються лінійно еквівалентними, якщо кожен вектор із A лінійно подається через вектори системи B і, навпаки, кожен вектор із B лінійно подається через вектори системи A.*

Лінійну еквівалентність систем A і B позначаємо у вигляді відношення $A \sim B$.

Як і раніше, через $L(a_1, a_2, \dots, a_m)$ позначаємо лінійну оболонку системи $A = < a_1, a_2, \dots, a_m >$. Спрощено: $L(A)$.

Лема. *Якщо $a \in L(A)$ і $A \subset L(B)$, тоді $a \in L(B)$.*

Доведення. За умовою

$$a = \sum_{j=1}^{m} \lambda_j a_j\ ,\ a_j = \sum_{k=1}^{p} \mu_{jk} b_k\ ,\ \lambda_j, \mu_{ik} \in P.$$

Тоді

$$a = \sum_{j=1}^{m} \lambda_j \left(\sum_{k=1}^{p} \mu_{jk} b_k \right) = \left(\sum_{j=1}^{m} \lambda_j \mu_{j1} \right) b_1 +$$

$$+ \left(\sum_{j=1}^{m} \lambda_j \mu_{j2} \right) b_2 + \dots + \left(\sum_{j=1}^{m} \lambda_j \mu_{jp} \right) b_p \in L(B).$$

Твердження 1. *Дві системи векторів $A = < a_1, a_2, \dots, a_m >$ і $B = < b_1, b_2, \dots, b_p >$ лінійно еквівалентні тоді і лише тоді, коли збігаються їх лінійні оболонки: $L(A) = L(B)$.*

Доведення. *Необхідність.* Якщо $A \sim B$, тоді $A \subset L(B)$, а тому з леми випливає включення $L(A) \subseteq L(B)$. Вірним є також обернене включення $L(B) \subseteq L(A)$.

Обернене твердження (*достатність*) очевидне.

Відношення лінійної еквівалентності для систем векторів $A = < a_1, a_2, \ \dots \ , a_m >$, $B = < b_1, b_2, \ \dots \ , b_p >$ і $C = < c_1, c_2, \ \dots \ , c_s >$ має властивості еквівалентності.

1. Рефлексивність: $A \sim A$.
2. Симетричність: $A \sim B \Rightarrow B \sim A$.
3. Транзитивність: $A \sim B \ \& \ B \sim C \Rightarrow A \sim C$.

Елементарними перетвореннями системи векторів $A = \langle a_1, a_2, ..., a_m \rangle$ називаються перетворення таких типів:

1. Множення деякого вектора a_j системи на скаляр $\lambda \neq 0$.
2. Додавання до вектора a_j деякої системи іншого вектора цієї ж системи, помноженого на будь-який скаляр λ.
3. Перестановка двох будь-яких векторів системи.

Твердження 2. *Якщо система векторів $B = < b_1, b_2, \ \dots \ , b_m >$ одержана із системи $A = < a_1, a_2, \ \dots \ , a_m >$ за допомогою деякого ланцюга елементарних перетворень, то A і B лінійно еквівалентні.*

Доведення. Достатньо встановити еквівалентність $A \sim B$ лише в тому випадку, коли B одержується з A за допомогою одного елементарного перетворення будь-якого типу. Нехай, наприклад, B одержується з A за допомогою перетворення типу 2:

$$B = < a_1 + \lambda a_2, \ a_2, a_3, \ \dots \ , a_m > .$$

Отже, вектори із B лінійно подаються через систему A. Обернене твердження також зберігає силу: $a_1 = (a_1 + \lambda a_2) + (-\lambda)a_2 + 0 \cdot a_3 + \ \dots$ – лінійна комбінація векторів із B і т. д.

Теорема 1 (про базис). *Нехай Σ – довільна підмножина векторного простору P^n, яка містить ненульові вектори. Існує така скінченна підмножина $B = < b_1, b_2, \ \dots \ , b_p >$ із Σ, яка має такі властивості:*

I. Вектори системи B лінійно незалежні.

II. Кожний вектор $x \in \Sigma$ лінійно подається через вектори системи B: $x = \xi_1 b_1 + \xi_2 b_2 + \ \dots \ + \xi_p b_p$ і таке подання єдине.

Система B з такими властивостями називається базисом множини Σ.

Усі базиси множини Σ містять однакову кількість векторів.

Доведення. Позначимо через Σ' множину різних лінійно незалежних систем векторів, що належать Σ. Нам відомо, що будь-яка з таких систем містить не більше, ніж n векторів (див. § 3.1). У множині Σ' знайдеться система $B = < b_1, b_2, \ \dots \ , b_p >$, яка містить найбільше число p векторів (максимальна система лінійно незалежних векторів). Переконаємось, що вона має властивість II.

Нехай $x \in \Sigma$. Систему B розширимо, додавши до неї вектор x. Одержана система $x, b_1, b_2, \dots, b_p$ лінійно залежна: для неї існує ненульова послідовність скалярів $\lambda, \lambda_1, \lambda_2, \ \dots \ , \lambda_p$ така, що $\lambda x + \lambda_1 b_1 + \lambda_2 b_2 + \ \dots \ + \lambda_p b_p = \theta$. При цьому $\lambda \neq 0$, оскільки у протилежному випадку разом з $\lambda = 0$ випливало б, що $\lambda_1 = \lambda_2 = \ \dots \ = \lambda_p = 0$, а це суперечить припущенню.

Тому маємо

$$x = \xi_1 b_1 + \xi_2 b_2 + \ \dots \ + \xi_p b_p, \text{ де } \xi_j = -\frac{\lambda_j}{\lambda}, \quad j = \overline{1, p}.$$

Залишилось довести єдиність останнього подання вектора x. Припустимо, що $x = \eta_1 b_1 + \eta_2 b_2 + \ \dots \ + \eta_p b_p$. На підставі рівності

$$\xi_1 b_1 + \xi_2 b_2 + \ \dots \ + \xi_p b_p = \eta_1 b_1 + \eta_2 b_2 + \ \dots \ + \eta_p b_p$$

знаходимо $(\xi_1 - \eta_1) b_1 + (\xi_2 - \eta_2) b_2 + \ \dots \ + (\xi_p - \eta_p) b_p = \theta$.

Отже,

$$\xi_1 - \eta_1 = 0, \xi_2 - \eta_2 = 0, \dots, \xi_p - \eta_p = 0$$

і тому

$$\xi_1 = \eta_1, \xi_2 = \eta_2, ..., \xi_p = \eta_p.$$

Нехай $B = < b_1, b_2, \, ..., b_p >$ і $C = < c_1, c_2, ..., c_q >$ – два будь-які базиси множини Σ. Вони лінійно еквівалентні. Із властивостей лінійно незалежних векторів (див. § 3.1) випливає рівність $p = q$.

Означення. *Рангом множини Σ векторів називається число векторів його базису. Якщо множина Σ містить лише нульовий вектор, то ранг її вважається рівним нулю.*

Теорема 2. *Якщо системи векторів Σ і Δ лінійно еквівалентні, то їхні ранги рівні.*

Доведення. Нехай $B = < b_1, b_2, \, ..., b_p >$ – базис множини Σ і $C = < c_1, c_2, \, ..., c_s >$ – базис множини Δ. Оскільки $\Sigma \sim \Delta$, $\Sigma \sim B$, $\Delta \sim C$, то на підставі транзитивності відношення лінійної еквівалентності маємо $B \sim C$. Звідси випливає, що $p = s$.

Теорему доведено.

Зауваження. Наведене вище доведення існування базису множини Σ не є конструктивним. Але для деяких множин векторів існують ефективні методи побудови базисів.

Нехай, наприклад, множина Σ збігається з усім простором P^n. У просторі P^n одиничні вектори

$$\begin{aligned} e_1 &= (1, 0, 0, \, ... \, , 0), \\ e_2 &= (0, 1, 0, \, ... \, , 0), \\ &\dots\dots\dots\dots\dots\dots \\ e_n &= (0, 0, 0, \, ... \, , 1) \end{aligned}$$

складають його базис. Надалі називатимемо його стандартним базисом.

Виділимо будь-яку квадратну матрицю порядку n:

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

Якщо $\det B \neq 0$, то її рядки $b_1 = (\beta_{11}, \beta_{12}, \ldots, \beta_{1n})$, $b_2 = (\beta_{21}, \beta_{22}, \ldots, \beta_{2n})$, ..., $b_n = (\beta_{n1}, \beta_{n2}, \ldots, \beta_{nn})$ лінійно незалежні. Тому вони утворюють базис простору P^n.

Аналогічно стовпці цієї матриці також утворюють базис P^n.

Нехай $Б = < b_1, b_2, \ldots, b_n >$ – базис простору P^n і $x \in P^n$. Коефіцієнти $\xi_1, \xi_2, \ldots, \xi_n$ в поданні

$$x = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n$$

називаються координатами вектора x у цьому базисі. Інакше ці координати будемо записувати у вигляді стовпця. Щоб відмітити їхню залежність від базису використаємо такий запис:

$$x_Б = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = (\xi_1, \xi_2, \ldots, \xi_n)^T.$$

Рядки матриці

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

з елементами з поля P є векторами лінійного простору P^n. Ранг системи цих векторів

$$a_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}), \quad i = \overline{1,m},$$

називається *рангом матриці A за рядками.*

Для обчислень введемо інше означення рангу цієї самої матриці.

Нехай p – натуральне число, що задовольняє нерівності $1 \leq p \leq \min(m,n)$. Зафіксуємо які-небудь p різних рядків цієї матриці і p різних стовпців; складемо визначник порядку p з елементів, що розміщені на перетині цих рядків і стовпців:

$$\begin{vmatrix} \alpha_{i_1 j_1} & \alpha_{i_1 j_2} & \dots & \alpha_{i_1 j_p} \\ \alpha_{i_2 j_1} & \alpha_{i_2 j_2} & \dots & \alpha_{i_2 j_p} \\ \dots & \dots & \dots & \dots \\ \alpha_{i_p j_1} & \alpha_{i_p j_2} & \dots & \alpha_{i_p j_p} \end{vmatrix},$$

причому розташуємо елементи так, щоб $i_1 < i_2 < \dots < i_p$ і $j_1 < j_2 < \dots < j_p$. Такий визначник називається *мінором* p-го *порядку* матриці A, або – p-мінором.

Наприклад, для $p = 2$ такими мінорами будуть

$$\begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{41} & \alpha_{43} \end{vmatrix}, \quad \begin{vmatrix} \alpha_{11} & \alpha_{14} \\ \alpha_{31} & \alpha_{34} \end{vmatrix}, \dots$$

Зауважимо, що коли всі мінори деякого порядку p дорівнюють нулю, то й усі мінори $(p+1)$-го порядку також дорівнюють нулю. А тому всі мінори порядку $> p$ дорівнюють нулю.

Означення. *Нехай A – ненульова матриця розмірності $m \times n$ і r – натуральне число з такими властивостями:*

I. Матриця A має r-мінори, відмінні від нуля.

II. Усі мінори, порядок яких перевищує r (якщо такі наявні), дорівнюють нулю.

Число r називається рангом матриці A за мінорами, а будь-який відмінний від нуля мінор порядку r називається базисним мінором.

Теорема 3 (про ранг матриці). *Ранг матриці за мінорами дорівнює рангу цієї матриці за рядками.*

Доведення. Позначимо через r ранг матриці A за мінорами. Вважатимемо матрицю ненульовою. Оберемо який-небудь її базисний мінор. Для наочності викладу припустимо, що його елементи розміщені у лівому верхньому кутку, тобто

$$\Delta = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2r} \\ \dots & \dots & \dots & \dots \\ \alpha_{r1} & \alpha_{r2} & \dots & \alpha_{rr} \end{vmatrix} \neq 0 .$$

Запишемо матрицю з більшою кількістю елементів. Окремо виділимо рядок a_i з номером $i > r$ і стовпець з довільним номером k:

$$A = \begin{pmatrix} \boxed{\begin{matrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2r} \\ \dots & \dots & \dots & \dots \\ \alpha_{r1} & \alpha_{r2} & \dots & \alpha_{rr} \end{matrix}} & \begin{matrix} \dots & \alpha_{1k} & \dots & \alpha_{1n} \\ \dots & \alpha_{2k} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \alpha_{rk} & \dots & \alpha_{rn} \end{matrix} \\ \dots\dots\dots & \dots\dots\dots \\ \begin{matrix} \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ir} \end{matrix} & \begin{matrix} \dots & \alpha_{ik} & \dots & \alpha_{in} \end{matrix} \\ \dots\dots\dots & \dots\dots\dots \end{pmatrix}.$$

Покажемо, що рядки $a_1, a_2, \ \dots, a_r$ матриці A складають базис для всіх рядків.

Спочатку доведемо лінійну незалежність векторів $a_1, a_2, \ \dots, a_r$. Розглянемо вкорочені вектори

$$\underline{a_1} = (\alpha_{11}, \alpha_{12}, \ \dots, \alpha_{1r}),$$
$$\underline{a_2} = (\alpha_{21}, \alpha_{22}, \ \dots, \alpha_{2r}),$$
$$\dots\dots\dots\dots\dots\dots$$
$$\underline{a_r} = (\alpha_{r1}, \alpha_{r2}, \ \dots, \alpha_{rr}).$$

Вони лінійно незалежні, оскільки $\Delta \neq 0$. Але тоді й вектори $a_1, a_2, \ \dots, a_r$ також лінійно незалежні.

Доведемо, що рядки $a_i \, (i > r)$ лінійно задаються через $a_1, a_2, \ \dots, \ a_r$. Розглянемо визначник $(r+1)$-го порядку з довільним індексом k в межах від 1 до n:

$$\Delta' = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1r} & \alpha_{1k} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2r} & \alpha_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{r1} & \alpha_{r2} & \dots & \alpha_{rr} & \alpha_{rk} \\ \alpha_{i1} & \alpha_{i2} & \dots & \alpha_{ir} & \alpha_{ik} \end{vmatrix}.$$

Для будь-яких значень індексу k в межах від $r+1$ до n визначник Δ' є не що інше, як мінор порядку $r+1$ матриці A і тому він дорівнює нулю. Для всіх значень k в межах від 1 до r він також дорівнює нулю, оскільки в ньому є два однакові стовпці. Отже, $\Delta' = 0$ для будь-яких значень індексу k. Розкладемо Δ' за елементами останнього стовпця. Дістанемо рівність

$$\alpha_{1k} A_{1r+1} + \alpha_{2k} A_{2r+1} + \ \ldots\ + \alpha_{rk} A_{rr+1} + \alpha_{ik}\Delta = 0,$$

де A_{jr+1}, $j = \overline{1,r}$, Δ – алгебричні доповнення елементів останнього стовпця. Вони не залежать від елементів k-го стовпця матриці A. Звідси знаходимо

$$\alpha_{ik} = \left(-\frac{A_{1r+1}}{\Delta}\right)\alpha_{1k} + \left(-\frac{A_{2r+1}}{\Delta}\right)\alpha_{2k} + \ldots + \left(-\frac{A_{rr+1}}{\Delta}\right)\alpha_{rk}.$$

Коефіцієнти $\lambda_1 = -\dfrac{A_{1r+1}}{\Delta}$, ... , $\lambda_r = -\dfrac{A_{rr+1}}{\Delta}$ не залежать від елементів k-го стовпця матриці A. З одержаних рівностей

$$\alpha_{ik} = \lambda_1\alpha_{1k} + \lambda_2\alpha_{2k} + \ldots + \lambda_r\alpha_{rk}, \quad k = \overline{1,n}$$

випливає, що $a_i = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_r a_r$. Отже, перші r рядків (що накривають базисний мінор) лінійно незалежні й усі рядки матриці A лінійно подаються через них. Вони становлять базис рядків матриці A.

Наслідок. *Стовпці матриці*

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

складають систему векторів

$$c_1 = (\alpha_{11}, \alpha_{21}, \ \ldots\ , \alpha_{m1}),$$
$$c_2 = (\alpha_{12}, \alpha_{22}, \ \ldots\ , \alpha_{m2}),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$c_n = (\alpha_{1n}, \alpha_{2n}, \ \ldots\ , \alpha_{mn})$$

простору P^n. *Ранг цієї системи називається рангом матриці* A *за стовпцями.*

Як і для рангу за рядками, доводиться, що *ранг матриці за стовпцями дорівнює її рангу за мінорами.*

Тому *ранг матриці за стовпцями дорівнює рангу матриці за рядками, і навпаки.*

Вправи

1. Довести, що у будь-якій множині Σ векторів простору P^n будь-яку лінійно незалежну систему її векторів можна доповнити до базису цієї множини.

2. Показати, що коли дві системи векторів мають однакові ранги, а вектори однієї з цих систем лінійно подаються через іншу, тоді системи лінійно еквівалентні.

3. Квадратна матриця $A = (\alpha_{ik})$ називається симетричною, якщо $\alpha_{ik} = \alpha_{ki}$. Головними її мінорами називаються такі, елементи яких містяться на перетинах рядків з номерами $i_1, i_2, \ldots, i_p$ і стовпців з тими самими номерами. Довести, що ранг симетричної матриці дорівнює найвищому порядку відмінних від нуля головних мінорів.

4. Довести, що в матриці A рангу r на перетині будь-яких r лінійно незалежних рядків і r лінійно незалежних стовпців містяться елементи ненульового мінора Δ.

Вказівка. Виділити в матриці A підматрицю A', яка складається з r лінійно незалежних рядків матриці A. Оскільки всі стовпці матриці A лінійно задаються через виділені r лінійно незалежних стовпців, то всі стовпці підматриці A' лінійно подаються через стовпці мінора Δ. Тому стовпці мінора Δ лінійно незалежні й, отже, $\Delta \neq 0$.

Метод виключення

Твердження 3. *Матрицю*

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

розмірності $m \times n$ за допомогою деякої послідовності елементарних перетворень рядків можна привести до ступеневого вигляду.

Доведення проведемо індукцією за числом m рядків матриці. Якщо $m = 1$, тоді матриця вже матиме такий вигляд.

Припустимо, що твердження має місце для матриць з $m-1$ рядками. Доведемо твердження для матриці A з m рядками. Розглянемо її перший стовпець. Якщо всі елементи цього стовпця дорівнюють нулю, то перейдемо до другого стовпця і так доти, доки на якомусь кроці знайдеться ненульовий стовпець. Нехай, наприклад, уже в першому стовпці є ненульові елементи. Вважатимемо, що $\alpha_{11} \neq 0$. Інакше цього можна досягти перестановкою рядків. За допомогою декількох елементарних перетворень типу 2 дістанемо матрицю

$$B = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & \beta_{22} & \beta_{23} & \cdots & \beta_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \beta_{m2} & \beta_{m3} & \cdots & \beta_{mn} \end{pmatrix}.$$

Для цього спочатку до другого рядка додамо перший, помножений на скаляр $\lambda = -\frac{\alpha_{21}}{\alpha_{11}}$. Далі аналогічно змінюємо 3-й рядок і т. д. Одержуємо лінійну еквівалентність рядків матриць A і B, яку записуємо у вигляді

$$A \sim B.$$

Матриця

$$\hat{B} = \begin{pmatrix} \beta_{22} & \beta_{23} & \cdots & \beta_{2n} \\ \beta_{32} & \beta_{33} & \cdots & \beta_{3n} \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{m2} & \beta_{m3} & \cdots & \beta_{mn} \end{pmatrix}$$

має $m-1$ рядків і до неї можна застосувати припущення індукції. За допомогою деякої послідовності елементарних перетворень вона зводиться до ступеневої матриці $\hat{C}$. Тоді

$$A \sim C = \left(\begin{array}{c|ccc} \alpha_{11} & \alpha_{12} & \alpha_{13} \;\cdots & \alpha_{1n} \\ \hline 0 & & & \\ \cdots & & \hat{C} & \\ 0 & & & \end{array}\right)$$

і C – матриця ступенева.

Наслідок. *Ненульові рядки $c_1, c_2, ..., c_p$ ступеневої матриці C лінійно незалежні. Вони одержані з рядків матриці A за допомогою декількох елементарних перетворень. Тому рядки матриці A лінійно еквівалентні з рядками матриці C (твердження 2). За теоремою 2 їхні ранги збігаються. Отже, ранг матриці A дорівнює числу p ненульових рядків матриці C.*

На завершення сформулюємо метод виключення.

Для обчислення рангу системи $a_1, a_2, \ ... \ , a_m$ векторів складаємо матрицю A, рядки якої збігаються з цими векторами

$$a_i = (\alpha_{i1}, \alpha_{i2}, \ ... \ , \alpha_{in}), \text{ при } i = \overline{1,m}.$$

За допомогою елементарних перетворень зводимо її до ступеневої матриці C. Якщо $c_1, c_2, \ ... \ , c_p$ – ненульові рядки останньої, то число p – ранг матриці A.

Більше того, збігаються лінійні оболонки

$$L(a_1, a_2, \ ... \ , a_m) = L(c_1, c_2, \ ... \ , c_p)$$

і система $C = \langle c_1, c_2, \ ... \ , c_p \rangle$ – базис цієї оболонки.

Правило. Для обчислення базису рядків матриці A потрібно:

1. Методом виключення обчислити її ранг.

2. Якщо p – ранг цієї матриці, то серед її p-мінорів відшукати який-небудь мінор $\Delta \neq 0$ (базисний мінор).

Тоді рядки матриці A, що покривають мінор Δ, будуть для неї базисними.

Метод напрямленого виключення

Означення. *Головними елементарними перетвореннями рядків (векторів) $a_1 = (\alpha_{11}, \alpha_{12}, \ ... \ , \alpha_{1n})$, $a_2 = (\alpha_{21}, \alpha_{22}, ..., \alpha_{2n})$, ... , $a_m = (\alpha_{m1}, \alpha_{m2}, ..., \alpha_{mn})$ матриці*

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

називаються перетворення таких двох типів:

1°. Множення рядка a_k *на скаляр* $\lambda \neq 0$.

2°. Додавання до рядка $a_k \neq \theta$ *іншого рядка* a_j, *помноженого на скаляр* λ.

Перестановку двох ненульових рядків a_k, a_j можна здійснити за допомогою головних елементарних перетворень. Перестановка з нульовим рядком у цьому випадку неможлива.

Означення. *Індексом* ($\operatorname{ind} a_k$) *ненульового рядка* a_k *називається номер s першого його елемента (координати)* $\alpha_{ks} \neq 0$.

Якщо

$$a_k = (0, \ldots, 0, \alpha_{ks}, \alpha_{k\,s+1}, \ldots, \alpha_{kn}),$$

де $\alpha_{ks} \neq 0$, *тоді* $\operatorname{ind} a_k = s$.

Опишемо один із способів зведення матриці A до псевдоступеневого вигляду, що носить назву *методу напрямленого виключення.*

Нехай число рядків m матриці A більше одиниці.

Фіксуємо в A перший рядок (рахуючи зверху), індекс якого мінімальний. Нехай, наприклад, таким рядком є a_1:

$$\operatorname{ind} a_1 \leq \operatorname{ind} a_j \quad (j = \overline{1,m}).$$

Проведемо першу серію головних елементарних перетворень. Нехай $\operatorname{ind} a_k$ будь-якого рядка a_k, що розташований нижче рядка a_1, $a_k \neq \theta$ ($k > 1$), дорівнює $\operatorname{ind} a_1$. Додамо до рядка a_k рядок a_1, помножений на скаляр $\lambda \neq 0$, такий, що індекс перетвореного рядка $b_k = a_k + \lambda a_1$ збільшився: $\operatorname{ind} b_k > \operatorname{ind} a_1$ (головне елементарне перетворення типу 2°).

Аналогічне перетворення застосуємо до всіх рядків, що розташовані нижче, з індексами, рівними $\operatorname{ind} a_1$. При цьому щоразу (при необхідності) одержані рядки можна множити на ненульові скаляри (перетворення типу 1°). У результаті такої серії перетворень дістанемо матрицю $A_1 \sim A$.

Нехай $\hat{A}_1$ – матриця, одержана з A_1 шляхом викреслення рядка a_1. До неї застосуємо ту саму процедуру: переглянемо всі її рядки і виберемо перший з тих, індекс якого мінімальний і т. д. Одержимо деяку матрицю $\hat{A}_2$. Допишемо до неї вилучений перший рядок a_1. Матимемо нову матрицю $A_2 \sim A_1$.

Продовжимо далі цей процес і на якомусь p-му кроці дістанемо матрицю A_p псевдоступеневого вигляду. Позначимо її через $C = A_p$. Матимемо еквівалентність $A \sim C$.

Усі елементарні перетворення при обчисленні матриці C спрямовані зверху вниз. Проілюструємо процес виключення на матриці

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

Перша серія перетворень. Застосуємо перетворення типу 2° до рядків з номерами 2, 3, 4, 5. Матимемо

$$A \sim \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 3 & -3 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & -1 & 4 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & -1 & 4 & -1 \end{pmatrix} = A_1.$$

Друга серія перетворень. Застосуємо перетворення типу 2° до рядків з номерами 3, 4, 5 і знайдемо

$$A_1 \sim \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 3 & -1 \end{pmatrix} = A_2.$$

Третя серія перетворень. Перетворення типу 2° зводить матрицю A_2 до псевдоступеневого вигляду

$$A_2 \sim \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = A_3 = C.$$

Теорема 4. *Приведемо матрицю*

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

за допомогою методу напрямленого виключення до матриці C псевдоступеневого вигляду. Якщо $c_{j1}, c_{j2}, \ldots, c_{jr}$ *– ненульові рядки матриці C з номерами* $j_1, j_2, \ldots, j_r$ $(1 \le j_1 < j_2 < \ldots < j_r \le m)$, *тоді рядки* $a_{j1}, a_{j2}, \ldots, a_{jr}$ *матриці A з тими самими номерами становлять базис її рядків.*

Доведення проведемо індукцією за числом m рядків матриці.

Якщо $m = 1$, тоді теорема очевидна. Припустимо, що $m > 1$ і теорема має місце для всіх матриць з числом рядків $m - 1$. Позначимо через $\hat{A}$ підматрицю з A, одержану з неї викреслюванням останнього рядка:

$$A = \left(\frac{\hat{A}}{a_m} \right). \tag{1}$$

Процес обчислення матриці C свідчить, що останній її рядок c_m подається у вигляді $c_m = a_m - f$, де f – лінійна комбінація рядків $a_1, a_2, \ldots, a_{m-1}$ матриці $\hat{A}$. Якщо через $\hat{C}$ позначити матрицю, одержану з C викреслюванням її останнього рядка, тобто

$$C = \left(\frac{\hat{C}}{c_m} \right), \tag{2}$$

то $\hat{C}$ – псевдоступенева матриця, одержана з $\hat{A}$ в процесі перетворення матриці A.

За припущенням індукції теорема має місце для $\hat{A}$. Тому, якщо $c_{j1}, c_{j2}, \ldots, c_{jp}$ – ненульові рядки матриці $\hat{C}$ з номерами $j_1, j_2, \ldots, j_p$ $(1 \le j_1 < j_2 < \ldots < j_p \le m)$, то рядки $a_{j_1}, a_{j_2}, \ldots, a_{j_p}$ утворюють базис рядків матриці $\hat{A}$.

Якщо $a_m = \theta$, то теорему доведено.

Нехай $a_m \ne \theta$. Тоді

$$A \sim \left(\frac{\hat{C}}{a_m} \right) = \left(\frac{\hat{C}}{c_m + f} \right). \quad (3)$$

Розглянемо дві можливості:

а) $c_m = \theta$.

У даному випадку $a_m = f$ – лінійна комбінація рядків $a_1, a_2, \ldots, a_{m-1}$ і, отже, a_m лінійно подається через $a_{j1}, a_{j2}, \ldots, a_{jp}$. Рядки $a_{j1}, a_{j2}, \ldots, a_{jp}$ є базисом рядків матриці A.

б) $c_m \ne \theta$.

Покажемо, що у цьому випадку система рядків

$$a_{j1}, a_{j2}, \ldots, a_{jp}, a_m \quad (4)$$

лінійно незалежна і тому є базисом рядків матриці A.

Система (4) лінійно еквівалентна системі

$$a_1, a_2, \ldots, a_{m-1}, a_m \quad (5)$$

рядків матриці A. Остання лінійно еквівалентна системі

$$c_{j1}, c_{j2}, \ldots, c_{jp}, c_m \quad (6)$$

ненульових рядків матриці C (твердження 2). Тому системи (4) і (6) лінійно еквівалентні і за теоремою 2 їхні ранги збігаються. Отже, ранг системи (4) дорівнює $p + 1$. Вектори її лінійно незалежні.

Теорему доведено.

У розглянутому вище прикладі за допомогою методу напрямленого виключення матрицю було зведено до псевдоступеневої матриці C з ненульовими рядками $c_1 = (1, 1, -1, 1)$,

$c_2 = (0, 1, -1, 0)$, $c_4 = (0, 0, 3, -1)$. Тому рядки $a_1 = (1, 1, -1, 1)$, $a_2 = (-1, 2, -2, -1)$, $a_4 = (2, 4, -1, 1)$ початкової матриці A складають базис її рядків.

Метод напрямленого виключення можна застосувати для розв'язання такої обчислювальної задачі. Нехай система $<a_1, a_2, \ \dots, a_m>$ векторів лінійно незалежна. Необхідно виділити з іншої системи $<b_1, b_2, \ \dots, b_s>$ векторів таку максимальну лінійно незалежну підсистему $<b_{j_1}, b_{j_2}, \dots, b_{j_p}>$, щоб система $<a_1, a_2, \dots, a_m, b_{j_1}, b_{j_2}, \dots, b_{j_p}>$ була лінійно незалежна.

Докладне її розв'язання пропонується провести самостійно.

Вправи

1. Обчислити ранг матриці

$$\begin{pmatrix} 3 & 1 & 1 & 4 \\ \lambda & 4 & 10 & 1 \\ 1 & 7 & 17 & 3 \\ 2 & 2 & 4 & 3 \end{pmatrix}$$

залежно від значень параметра λ.

2. Довести, що коли $A = (\alpha_{ik})$ – квадратна матриця і її $\det A \neq 0$, то за допомогою елементарних перетворень рядків її можна звести до одиничної матриці

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

3. Знайти базис системи векторів

$$a_1 = (1, 2, 1, -1),$$
$$a_2 = (2, -1, 2, 0),$$
$$a_3 = (-1, 3, -1, -1),$$
$$a_4 = (3, -1, 3, -1),$$
$$a_5 = (2, 2, 2, -2).$$

4. Нехай A – матриця розмірності $m \times n$, рядки якої $a_1, a_2, \ldots, a_m$ лінійно незалежні, B – матриця розмірності $l \times n$ і $b_1, b_2, \ldots, b_l$ – її рядки. Методом виключення зведемо матрицю A до ступеневої матриці A_1. Застосувавши до розширеної матриці $\left(\frac{A_1}{B}\right)$ метод напрямленого виключення, дістанемо матрицю виду $\left(\frac{A_1}{C}\right)$, де C – псевдоступенева матриця. Показати, що коли $c_{j1}, c_{j2}, \ldots, c_{jp}$ – ненульові рядки матриці C, то система рядків $a_1, a_2, \ldots, a_m$, $b_{j1}, b_{j2}, \ldots, b_{jp}$ – базис рядків розширеної матриці $\left(\frac{A}{B}\right)$.

§ 3.3. *Лінійні простори скінченної розмірності*

Нехай N і M – довільні лінійні простори над одним полем P. *Лінійним перетворенням*

$$F : N \to M$$

простору N у простір M називається закон, за яким кожному вектору $x \in N$ однозначно ставиться у відповідність вектор $y = F(x) \in M$, причому задовольняються такі умови:

1) $F(x + x') = F(x) + F(x')$;

2) $F(\alpha \cdot x) = \alpha \cdot F(x)$,

де $x, x' \in N$, $\alpha \in P$.

Ці умови еквівалентні одній такій

3) $F(\alpha x + \beta x') = \alpha F(x) + \beta F(x')$ *для будь-яких* $x, x' \in N$ *і* $\alpha, \beta \in P$.

Простори N і M називаються *ізоморфними*, якщо існує таке взаємно однозначне відображення простору N на простір M, яке поряд з цим є лінійним перетворенням.

Зауважимо, що для довільного лінійного простору N означення лінійної залежності і лінійної незалежності векторів повністю збігається у тому вигляді, як вони були введені для арифметичного векторного простору P^n.

Можна легко перевірити таке твердження: лінійно залежна система векторів $a_1, a_2, ..., a_m$ простору N при лінійному перетворенні $F: N \to M$ переходить в лінійно залежну систему векторів $F(a_1), F(a_2), ..., F(a_m)$ простору M.

Якщо F – ізоморфне відображення N на M, то й лінійно незалежна система векторів простору N переходить у лінійно незалежну систему векторів із M. При цьому обернене відображення $F^{-1}: M \to N$ буде лінійним перетворенням.

Ізоморфізм просторів N і M означає тотожність їх властивостей, які не залежать від конкретної природи їх елементів.

Означення. *Лінійний простір N над полем P називається* ***скінченновимірним****, якщо він має базис* Б$=<b_1, b_2, ..., b_n>$, *складений із скінченного числа векторів. Число n векторів базису називається* ***розмірністю*** *простору N і позначається* $\dim N$. Якщо N складається з одного нульового вектора, то покладаємо $\dim N = 0$.

Існують лінійні простори, які не мають скінченних базисів. Прикладом такого простору є $P[x]$ – простір усіх многочленів із коефіцієнтами з поля P.

Теорема. *Якщо N – лінійний простір над полем P розмірності n, то він ізоморфний арифметичному векторному простору* P^n.

Доведення. У просторі P^n вибираємо базис, що складається з одиничних векторів (стандартний базис):

$$\begin{aligned} e_1 &= (1, 0, 0, ..., 0), \\ e_2 &= (0, 1, 0, ..., 0), \\ e_3 &= (0, 0, 1, ..., 0), \\ &\dots\dots\dots\dots\dots \\ e_n &= (0, 0, 0, ..., 1). \end{aligned}$$

Нехай $b_1, b_2, ..., b_n$ – базис простору N. Будуємо відображення $F: P^n \to N$ за таким правилом.

Обираємо довільний вектор $x = (\xi_1, \xi_2, ..., \xi_n) = \xi_1 e_1 + \xi_2 e_2 + \\ + ... + \xi_n e_n$ із P^n. Образом цього вектора в N вважаємо

$$y = F(x) = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n .$$

1) Доведемо лінійність відображення F.

Якщо $x = (\xi_1, \xi_2, \ldots, \xi_n)$ і $x' = (\xi_1', \xi_2', \ldots, \xi_n')$, $\alpha, \beta \in P$, тоді

$$y = F(x) = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n ,$$

$$y' = F(x') = \xi_1' b_1 + \xi_2' b_2 + \ldots + \xi_n' b_n ,$$

$$\alpha x + \beta x' = (\alpha\xi_1 + \beta\xi_1', \alpha\xi_2 + \beta\xi_2', \ldots, \alpha\xi_n + \beta\xi_n') .$$

Отже,

$$F(\alpha x + \beta x') = (\alpha\xi_1 + \beta\xi_1')b_1 + (\alpha\xi_2 + \beta\xi_2')b_2 + \ldots + (\alpha\xi_n + \beta\xi_n')b_n =$$

$$= \alpha(\xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n) + \beta(\xi_1' b_1 + \xi_2' b_2 + \ldots + \xi_n' b_n) =$$

$$= \alpha y + \beta y' = \alpha F(x) + \beta F(x') .$$

2) F – відображення на весь простір N.

Обираємо довільний вектор $y = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n$ із N. Тоді для вектора $x = (\xi_1, \xi_2, \ldots, \xi_n)$ із P^n маємо $y = F(x)$.

3) Відображення F взаємно однозначне.

Припустимо, що $F(x) = F(x') = y$ для деяких $x = (\xi_1, \xi_2, \ldots, \xi_n)$ і $x' = (\xi_1', \xi_2', \ldots, \xi_n')$ із P^n. Це означає, що $\xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n = \xi_1' b_1 + \xi_2' b_2 + \ldots + \xi_n' b_n$. Але $b_1, b_2, \ldots, b_n$ – вектори базису. Тому $\xi_1' = \xi_1$, $\xi_2' = \xi_2$, ..., $\xi_n' = \xi_n$ і $x' = x$.

Вправи

1. Нехай R_{n+1} – множина всіх многочленів $f(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n$ степеня $\leq n$ з дійсними коефіцієнтами. Довести, що відносно звичайних операцій додавання і множення на скаляри множина R_{n+1} – лінійний простір розмірності $n+1$.

2. Те саме довести для множини R_{n+1} усіх функцій виду $f(t) = \alpha_0 + \alpha_1 e^t + \ldots + \alpha_n e^{nt}$, де $n \in N$.

3. Нехай N – множина всіх нескінченних послідовностей $(u_1, u_2, \ldots, u_n, u_{n+1}, \ldots)$ дійсних чисел, заданих зворотним рівнянням

$$u_{n+3} = \alpha_0 u_n + \alpha_1 u_{n+1} + \alpha_2 u_{n+2} \ (\alpha_0, \alpha_1, \alpha_2 \neq 0).$$

Показати, що N – лінійний простір відносно операцій додавання і множення цих послідовностей на скаляри. Яка його розмірність?

§ 3.4. Лінійні підпростори

Означення. *Непорожня підмножина L лінійного простору N (над полем P) називається його лінійним підпростором, коли вона має такі властивості:*

I. Якщо $x, y \in L$, то $x + y \in L$.

II. Якщо $x \in L$ і $\lambda \in P$, то $\lambda x \in L$.

Ці дві властивості еквівалентні одній.

III. Якщо $x, y \in L$ і $\lambda, \mu \in P$, то $\lambda x + \mu y \in L$.

Будь-яка лінійна комбінація векторів із L належить L. Підпростір містить нульовий вектор θ.

Якщо Λ – лінійний простір скінченної розмірності і L – його лінійний підпростір, то і L – лінійний простір скінченної розмірності і $\dim L \leq \dim \Lambda$. При цьому $\dim L = \dim \Lambda$ тоді й лише тоді, коли $L = \Lambda$.

Приклад 1. Лінійна оболонка $L = L(a_1, a_2, ..., a_m)$ будь-якого скінченного набору векторів $a_1, a_2, ..., a_m$ з простору Λ є лінійний підпростір скінченної розмірності. Базис Б = $= <a_{i_1}, a_{i_2}, ..., a_{i_r}>$ системи векторів $a_1, a_2, ..., a_m$ можна вибрати як базис L. Тому $\dim L$ дорівнює рангу заданої системи векторів, лінійною оболонкою якої є простір L.

Означення. *Нехай L_1 і L_2 – підмножини лінійного простору Λ. Сумою $L_1 + L_2$ називається множина всіх тих векторів $w \in \Lambda$, які подаються у вигляді суми $w = a_1 + a_2$, де $a_1 \in L_1$, $a_2 \in L_2$, або*

$$L_1 + L_2 = \{w \in \Lambda : w = a_1 + a_2, a_1 \in L_1, a_2 \in L_2\}.$$

Означення певним чином розповсюджується на будь-який скінченний набір підмножин $L_1, L_2, ..., L_k$.

Приклад 2. На площині R^2 з прямокутною системою координат $O\xi\eta$ візьмемо два відрізки L_1 і L_2, паралельні осі $O\eta$. Нехай L_1 – відрізок з кінцями (α, γ_1), (α, δ_1) і $\gamma_1 \le \delta_1$; L_2 – відрізок з кінцями (β, γ_2), (β, δ_2) і $\gamma_2 \le \delta_2$. Їх можна розглядати як множини двовимірних векторів

$$L_1 = \{a = (\alpha, \eta) : \gamma_1 \le \eta \le \delta_1\}, \quad L_2 = \{b = (\beta, \eta) : \gamma_2 \le \eta \le \delta_2\}.$$

Сума $L_1 + L_2$ складається з усіх векторів $w = (\alpha + \beta, \eta_1 + \eta_2)$, де η_1 пробігає відрізок $[\gamma_1, \delta_1]$, а η_2 – відрізок $[\gamma_2, \delta_2]$. Зрозуміло, що $L_1 + L_2 \le L$, де $L = \{c = (\alpha + \beta, \eta) : \gamma_1 + \gamma_2 \le \eta \le \delta_1 + \delta_2\}$.

Покажемо, що має місце й обернене включення $L \subseteq L_1 + L_2$.

Виберемо будь-яке число η на відрізку $[\gamma_1 + \gamma_2, \delta_1 + \delta_2]$. Його можна подати у вигляді $\eta = (1 - \lambda)(\gamma_1 + \gamma_2) + \gamma(\delta_1 + \delta_2)$, де λ – деяке число з відрізка $[0,1]$. Обчислюємо $\eta_1 = (1 - \lambda)\gamma_1 + {}$ $+ \lambda\delta_1$, $\eta_2 = (1 - \lambda)\gamma_2 + \lambda\delta_2$. Тоді $\gamma_1 \le \eta_1 \le \delta_1$ і $\gamma_2 \le \eta_2 \le \delta_2$, $c_1 = (\alpha, \eta_1) \in L_1$ і $c_2 = (\beta, \eta_2) \in L_2$. Звідси матимемо

$$c_1 + c_2 = \left(\alpha + \beta, (1 - \lambda)\gamma_1 + \lambda\delta_1 + (1 - \lambda)\gamma_2 + \lambda\delta_2\right) =$$
$$= (\alpha + \beta, \eta) = c \in L.$$

Будь-який вектор $c = (\alpha + \beta, \eta)$ з множини L подається у вигляді суми $c = c_1 + c_2$, $c_1 \in L_1$, $c_2 \in L_2$. Доведено включення $L \subseteq L_1 + L_2$ і, отже, рівність $L = L_1 + L_2$.

Приклад 3. Нехай у лінійному просторі Λ задано два підпростори $L_1 = L(a_1, a_2, \ldots, a_m)$, $L_2 = L(b_1, b_2, \ldots, b_k)$, що являють собою лінійні оболонки системи векторів $a_1, a_2, \ldots, a_m$ і $b_1, b_2, \ldots, b_k$. Тоді $L = L_1 + L_2 = L(a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_k)$ – лінійний підпростір, породжений об'єднанням цих систем.

Теорема 1. *Якщо L_1 і L_2 – лінійні підпростори простору Λ, то перетин $L_1 \cap L_2$ і сума $L_1 + L_2$ – його лінійні підпростори. Якщо L_1 і L_2 скінченновимірні, тоді*

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

Доведення першої частини теореми досить очевидне. Доведемо її для суми L_1+L_2. Нехай $w_1,w_2\in L_1+L_2$ і $\lambda,\mu\in P$. Із означення матимемо

$$w_1=x_1+y_1,\ w_2=x_2+y_2,$$

де $x_1,x_2\in L_1$, $y_1,y_2\in L_2$. Тоді $w=\lambda w_1+\mu w_2=(\lambda x_1+\mu x_2)+$ $+(\lambda y_1+\mu y_2)$. Але $x=\lambda x_1+\mu x_2\in L_1$, $y=\lambda y_1+\mu y_2\in L_2$. Тому $w=x+y\in L_1+L_2$.

Нехай $\dim L_1=m$, $\dim L_2=l$, $\dim(L_1\cap L_2)=p$. Оберемо базис $c_1,c_2,\ldots,c_p$ перетину $L_1\cap L_2$. Доповнимо його до базису $c_1,c_2,\ldots,c_p,a_{p+1},\ldots,a_m$ підпростору L_1. Аналогічно доповнимо його до базису $c_1,c_2,\ldots,c_p,b_{p+1},\ldots,b_l$ підпростору L_2. З'ясовується, що система векторів усіх трьох типів $c_1,c_2,\ldots,c_p$, $a_{p+1},\ldots,a_m$, $b_{p+1},\ldots,b_l$ – базис суми L_1+L_2.

1) Доведемо лінійну незалежність цих векторів. Нехай $\gamma_1,\gamma_2,\ldots,\gamma_p$, $\alpha_{p+1},\ldots,\alpha_m$, $\beta_{p+1},\ldots,\beta_l$ – такі скаляри, для яких має місце рівність

$$\gamma_1c_1+\ldots+\gamma_pc_p+\alpha_{p+1}a_{p+1}+\ldots+\alpha_ma_m+$$
$$+\beta_{p+1}b_{p+1}+\ldots+\beta_lb_l=\theta.$$

Тоді

$$\gamma_1c_1+\ldots+\gamma_pc_p+\alpha_{p+1}a_{p+1}+\ldots+\alpha_ma_m=$$
$$=-(\beta_{p+1}b_{p+1}+\ldots+\beta_lb_l).$$

Ліва частина цієї рівності – вектор із L_1, права – вектор із L_2, а тому $-\beta_{p+1}b_{p+1}-\ldots-\beta_lb_l\in L_1\cap L_2$, тобто $-\beta_{p+1}b_{p+1}-\ldots-\beta_lb_l=$ $=\gamma'_1c_1+\ldots+\gamma'_pc_p$. Звідси дістаємо рівність $\gamma'_1c_1+\ldots+\gamma'_pc_p+$ $+\beta_{p+1}b_{p+1}+\ldots+\beta_lb_l=\theta$. З неї випливає, що $\gamma'_1=\ldots=\gamma'_p=$ $=\beta_{p+1}=\ldots=\beta_l=0$. Остаточно, $\gamma_1=\ldots=\gamma_p=\alpha_{p+1}=\ldots=\alpha_m=0$.

2) Якщо $w\in L_1+L_2$, тоді $w=x+y$, $x\in L_1$, $y\in L_2$. Подамо x у базисі підпростору L_1:

$$x=\gamma_1 c_1+\ldots+\gamma_p c_p+\alpha_{p+1}a_{p+1}+\ldots+\alpha_m a_m.$$

Також $y=\gamma_1' c_1+\ldots+\gamma_p' c_p+\beta_{p+1}b_{p+1}+\ldots+\beta_l b_l$. Тоді

$$w=(\gamma_1+\gamma_1')c_1+\ldots+(\gamma_p+\gamma_p')c_p+$$
$$+\alpha_{p+1}a_{p+1}+\ldots+\alpha_m a_m+\beta_{p+1}b_{p+1}+\ldots+\beta_l b_l.$$

Вектори $c_1,c_2,\ldots,c_p$, $a_{p+1},\ldots,a_m$, $b_{p+1},\ldots,b_l$ становлять базис простору L_1+L_2. Тому

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

Означення. *Сума* $L=L_1+L_2+\ldots+L_k$ *– лінійних підпросторів* L_j *векторного простору* Λ *називається прямою, якщо* $L_j\cap(L_1+\ldots+L_{j-1}+L_{j+1}+\ldots+L_k)=\{\theta\}$ *для всіх* $j=\overline{1,k}$. Зокрема, сума двох підпросторів L_1+L_2 – пряма, якщо $L_1\cap L_2=\{\theta\}$.

Позначення прямої суми підпросторів $L_1,\ldots,L_k$:

$$L=L_1\oplus L_2\oplus\ldots\oplus L_k.$$

Теорема 2. *Сума* $L=L_1+L_2+\ldots+L_k$ *лінійних підпросторів* $L_1,L_2,\ldots,L_k$ *є прямою тоді і лише тоді, коли будь-який вектор* $x\in L$ *можна подати у вигляді*

$$x=x_1+x_2+\ldots+x_k,\ x_j\in L_j$$

єдиним способом, тобто якщо $x=y_1+y_2+\ldots+y_k$, $y_j\in L_j$, *тоді* $x_j=y_j$ *для всіх* $j=\overline{1,k}$.

Доведення. Нехай $L=L_1\oplus L_2\oplus\ldots\oplus L_k$, $x\in L$ і

$$x=x_1+x_2+\ldots+x_k=y_1+y_2+\ldots+y_k,\text{ де } x_j,y_j\in L_j.$$

Маємо рівність $(x_1-y_1)+(x_2-y_2)+\ldots+(x_k-y_k)=\theta$, а тому

$$y_1-x_1=(x_2-y_2)+\ldots+(x_k-y_k).$$

Але $y_1-x_1\in L_1$ і $(x_2-y_2)+\ldots+(x_k-y_k)\in L_2+\ldots+L_k$.

За умовою перетин $L_1 \cap (L_2 + ... + L_k) = \{\theta\}$. Отже, $y_1 - x_1 = \theta \Rightarrow y_1 = x_1$.

Таким же чином доводяться рівності $y_j = x_j$ для всіх $j = \overline{1,k}$.

Висунемо обернене припущення: нехай будь-який вектор x із суми

$$L = L_1 + L_2 + ... + L_k$$

подається у вигляді

$$x = x_1 + x_2 + ... + x_k, \; x_j \in L_j$$

єдиним способом.

Оберемо $x \in L_1 \cap (L_2 + ... + L_k)$. Тоді $x = x_2 + ... + x_k$, $x_j \in L_j$, $j \geq 2$. Звідси на підставі єдиності подання нульового вектора отримаємо

$$-x = \theta, \; x_2 = \theta, ..., x_k = \{\theta\}$$

і, отже, $L_1 \cap (L_2 + ... + L_k) = \theta$.

Аналогічно для будь-якого j матимемо

$$L_j \cap (L_1 + ... + L_{j-1} + L_{j+1} + ... + L_k) = \{\theta\}.$$

Теорему доведено.

Теорема 3. *Якщо* $L = L_1 \oplus L_2 \oplus ... \oplus L_k$, *де* $L_1, L_2, ..., L_k$ – *лінійні підпростори простору* Λ, *тоді об'єднання базисів усіх доданків* L_j *становить базис підпростору* L *і*

$$\dim L = \dim L_1 + \dim L_2 + ... + \dim L_k.$$

Доведення випливає з теореми 1. При цьому достатньо використати метод математичної індукції за числом k доданків.

Вправи

1. Перелічити всі типи лінійних підпросторів дійсного векторного простору R^3.

2. У просторі R^3 задано підпростір $L = L(a,b)$, де $a = (1,1,1)$, $b = (2,-1,1)$. Знайти всі такі підпростори M, що $R^3 = L \oplus M$.

3. Нехай L і M – лінійні підпростори векторного простору Λ. У якому випадку теоретико-множинне об'єднання $L \cup M$ буде підпростором?

4. Нехай L і M – лінійні підпростори векторного простору N. Довести, що сума $L+M$ буде прямою сумою, якщо будь-який один вектор $z \in L+M$ однозначно подається у вигляді суми $z=x+y$, де $x \in L, y \in M$.

Відносний базис

Нехай L – лінійний підпростір векторного простору P^n над полем P. Для будь-яких двох векторів a, b з P^n введемо відношення $a \equiv b$, якщо $a-b \in L$. У цьому випадку будемо писати $a \equiv b (\operatorname{mod} L)$ і вважати, що a і b *конгруентні за модулем L*.

Це відношення між векторами є еквівалентністю, оскільки має властивості:

I. $a \equiv a (\operatorname{mod} L)$.

II. $a \equiv b (\operatorname{mod} L) \Rightarrow b \equiv a (\operatorname{mod} L)$.

III. $a \equiv b (\operatorname{mod} L)\ \&\ b \equiv c (\operatorname{mod} L) \Rightarrow a \equiv c (\operatorname{mod} L)$.

Означення. *Система векторів* $a_1, a_2, \ldots, a_k$ *називається лінійно залежною відносно підпростору L (лінійно залежною* $\operatorname{mod} L$), *коли існує така ненульова послідовність* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *скалярів, що*

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_k a_k \equiv \theta (\operatorname{mod} L).$$

Система векторів $a_1, a_2, \ldots, a_k$ *називається лінійно незалежною відносно L (лінійно незалежною* $\operatorname{mod} L$), *якщо має місце висновок*

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_k a_k \equiv \theta (\operatorname{mod} L) \Rightarrow$$
$$\Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

Твердження 1. *Система векторів* $A = \langle a_1, a_2, \ldots, a_k \rangle$ *лінійно незалежна відносно лінійного підпростору L, якщо для будь-якого базису* $Б = \langle b_1, b_2, \ldots, b_l \rangle$ *підпростору L об'єднання*

$$Б \cup A = \langle b_1, b_2, ..., b_l, a_1, a_2, ..., a_k \rangle$$

– лінійно незалежна система векторів.

Якщо Б *– будь-який з базисів підпростору* L і $Б \cup A$ *– лінійно незалежна система векторів, то система* A *лінійно незалежна* $\mathrm{mod}\, L$.

Означення. *Нехай L і M – лінійні підпростори. Система векторів* $a_1, a_2, ..., a_k$ *підпростору M називається його базисом відносно L (базисом* $\mathrm{mod}\, L$*), якщо*

1) $a_1, a_2, ..., a_k$ *лінійно незалежні* $\mathrm{mod}\, L$;

2) *для будь-якого вектора* $x \in M$ *існують такі скаляри* $\lambda_1, \lambda_2, ..., \lambda_k$, що $x \equiv \lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_k a_k \ (\mathrm{mod}\, L)$.

Із лінійної незалежності векторів $a_1, a_2, ..., a_k$ відносно L випливає, що коефіцієнти $\lambda_1, \lambda_2, ..., \lambda_k$ в останньому порівнянні визначаються однозначно для будь-якого вектора $x \in M$.

Твердження 2. *Нехай* $Б = \langle b_1, b_2, ..., b_l \rangle$ *– базис лінійного підпростору L. Для того щоб система* $A = \langle a_1, a_2, ..., a_k \rangle$ *векторів лінійного підпростору M була базисом* $\mathrm{mod}\, L$, *необхідно і достатньо, щоб об'єднання* $Б \cup A = \langle b_1, b_2, ..., b_l, a_1, a_2, ..., a_k \rangle$ *було базисом суми* $L + M$.

Твердження 1 і 2 доводяться на підставі означень.

Звернемо увагу на один ефективний спосіб обчислення базису $\mathrm{mod}\, L$. Нехай $Б = \langle b_1, b_2, ..., b_l \rangle$ – базис підпростору L, $A = \langle a_1, a_2, ..., a_m \rangle$ – базис підпростору M.

Підпростір $L + M$ – лінійна оболонка системи векторів $\Sigma = Б \cup A = \langle b_1, b_2, ..., b_l, a_1, a_2, ..., a_m \rangle$. Для вилучення потрібного базису з Σ утворимо розширену матрицю

$$\overline{B} = \left(\frac{B}{A} \right),$$

де B – матриця розмірності $l \times n$, рядки якої збігаються з векторами базису Б; A – матриця розмірності $m \times n$, рядки якої збігаються з векторами базису A.

Застосуємо до матриці $\overline{B}$ метод напрямленого виключення.

Оскільки рядки матриці B лінійно незалежні, то за результатами обчислень дістанемо базис суми $L+M$ з векторів

$$b_1, b_2, \dots, b_l, a_{j_1}, a_{j_2}, \dots, a_{j_k},$$

де $\Gamma = <a_{j_1}, a_{j_2}, \dots, a_{j_k}>$ – підмножина із A. Вона є шуканим відносним базисом.

Вправи

1. Нехай $\text{Б} = <b_1, b_2, \dots, b_l>$ – базис лінійного підпростору L, $\text{А} = <a_1, a_2, \dots, a_m>$ – базис лінійного підпростору M, $\Gamma = <a_1, a_2, \dots, a_k>$ – базис $\mod L$. Виразимо $p = m - k$ векторів $a_{k+1}, a_{k+2}, \dots, a_m$ через базис $\text{Б} \cup \Gamma$ суми $L+M$:

$$a_{k+1} = \beta_{11} b_1 + \beta_{12} b_2 + \dots + \beta_{1l} b_l + \alpha_{11} a_1 + \alpha_{12} a_2 + \dots + \alpha_{1k} a_k,$$

$$a_{k+2} = \beta_{21} b_1 + \beta_{22} b_2 + \dots + \beta_{2l} b_l + \alpha_{21} a_1 + \alpha_{22} a_2 + \dots + \alpha_{2k} a_k,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_m = \beta_{p1} b_1 + \beta_{p2} b_2 + \dots + \beta_{pl} b_l + \alpha_{p1} a_1 + \alpha_{p2} a_2 + \dots + \alpha_{pk} a_k.$$

Довести, що вектори

$$c_i = \beta_{i1} b_1 + \beta_{i2} b_2 + \dots + \beta_{il} b_l, \quad i = \overline{1, p}$$

становлять базис перетину $L \cap M$.

2. Знайти базис суми і перетину лінійних підпросторів

$$L = L(b_1, b_2, b_3), \ \ M = L(a_1, a_2, a_3),$$

де

$$b_1 = (1, 1, 0, 0), \ \ b_2 = (0, 1, 1, 0), \ \ b_3 = (0, 0, 1, 1),$$

$$a_1 = (1, 0, 1, 0), \ \ a_2 = (0, 2, 1, 1), \ \ a_3 = (1, 2, 1, 2).$$

Розділ 4. СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ (ЗАГАЛЬНА ТЕОРІЯ)

§ 4.1. Метод виключення невідомих (метод Гаусса). Умови сумісності

Нехай задано довільну систему лінійних рівнянь з невідомими $x_1, x_2, ..., x_n$ і коефіцієнтами α_{ij} $(i = \overline{1,m},\ j = \overline{1,n})$ із поля P:

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + ... + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + ... + \alpha_{2n}x_n = \beta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + ... + \alpha_{mn}x_n = \beta_m. \end{cases} \tag{1}$$

Основна матриця цієї системи

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

і розширена

$$\overline{A} = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2 \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m \end{array}\right).$$

Якщо в системі є рівняння виду

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

то його можна відкинути. Одержана система буде рівносильна заданій.

Якщо в системі є рівняння виду

$$0 \cdot x_1 + 0 \cdot x_2 + ... + 0 \cdot x_n = \beta, \quad \beta \neq 0,$$

то система не має розв'язків.

Обґрунтуємо метод виключення невідомих (метод Гаусса) для розв'язання системи (1).

За допомогою елементарних перетворень над рядками розширеної матриці $\overline{A}$ зведемо її до ступеневого вигляду

$$\overline{C}=\left(C\left|\begin{matrix}\delta_1\\ \delta_2\\ \delta_3\\ \vdots\end{matrix}\right.\right)=\left(\begin{matrix}\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p_2} & \cdots & \gamma_{1p_3} & \cdots & \gamma_{1n}\\ 0 & 0 & \cdots & \gamma_{2p_2} & \cdots & \gamma_{2p_3} & \cdots & \gamma_{2n}\\ 0 & 0 & \cdots & 0 & \cdots & \gamma_{3p_3} & \cdots & \gamma_{3n}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0\end{matrix}\left|\begin{matrix}\delta_1\\ \delta_2\\ \delta_3\\ \vdots\\ \end{matrix}\right.\right).$$

Вважаємо, що перші r рядків одержаної матриці $c_1, c_2, ..., c_r$ – ненульові, а решта – нульові. Тут $\gamma_{11}, \gamma_{2p_2}, \gamma_{3p_3}, ...$, ($1 < p_2 < p_3 < ...$) – провідні елементи, одержані після декількох елементарних перетворень рядків матриці A. Позначимо одержану матрицю C, а її розширену – $\overline{C}$. Як було доведено в § 2.6, система рівнянь (1) рівносильна системі

$$\begin{cases}\gamma_{11}x_1+\gamma_{12}x_2+...+\gamma_{1p_2}x_{p_2}+...+\gamma_{1p_3}x_{p_3}+...+\gamma_{1n}x_n=\delta_1,\\ \qquad\qquad\qquad +\gamma_{2p_2}x_{p_2}+...+\gamma_{2p_3}x_{p_3}+...+\gamma_{2n}x_n=\delta_2,\\ \qquad\qquad\qquad\qquad\qquad \gamma_{3p_3}x_{p_3}+...+\gamma_{3n}x_n=\delta_3.\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\end{cases} \qquad (2)$$

Система (2) містить r рівнянь (решту рівнянь виду $0=0$ відкинуто).

Для розширеної матриці $\overline{C}$ цієї системи є дві можливості, що виключають одна одну. Розглянемо їх окремо.

Можливість 1. Останній її ненульовий рядок (з номером r) має вигляд

$$c_r=(0,0,...,\gamma_{rp_r},...,\gamma_{rn}\mid\delta_r),$$

де $\gamma_{rp_r}\neq 0$ і $r\leq n$.

Тому матриці C і $\overline{C}$ ступеневого вигляду мають однакове число r ненульових рядків і, отже, ранги їх збігаються:

$$r=\text{рангу } C=\text{рангу } \overline{C}.$$

Але ранг A = рангу C і ранг $\overline{A}$ = рангу $\overline{C}$. Тому

$$r = \text{рангу } A = \text{рангу } \overline{A}.$$

Разом з тим, система (2) сумісна. Невідомі $x_1, x_{p_2}, ..., x_{p_r}$, які містяться при провідних коефіцієнтах $\gamma_{11}, \gamma_{2p_2}, ..., \gamma_{rp_r}$ – називатимемо їх *основними*, – можна подати через решту невідомих. Останні невідомі називаються *вільними* і число їх $n-r$.

Для того щоб подати основні невідомі через вільні, потрібно розпочинати з останнього рівняння системи (2):

$$\gamma_{rp_r} \cdot x_{p_r} + \gamma_{rp_r+1} \cdot x_{p_r+1} + ... + \gamma_{rn} \cdot x_n = \delta_r$$

і подати x_{p_r} через невідомі $x_{p_r+1}, ..., x_n$, які увійдуть у сукупність вільних невідомих. Далі слід перейти до рівняння, яке знаходиться вище щойно розглянутого, і т. д.

Дістаємо загальний розв'язок системи (2) і, разом з тим, загальний розв'язок системи (1).

Зокрема, якщо ранг $r = n$, тобто збігається з кількістю невідомих, тоді система (2) матиме вигляд

$$\begin{cases} \gamma_{11}x_1 + \gamma_{12}x_2 + ... + \gamma_{1n-1}x_{n-1} + \gamma_{1n}x_n = \delta_1, \\ \gamma_{22}x_2 + ... + \gamma_{2n-1}x_{n-1} + \gamma_{2n}x_n = \delta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \gamma_{n-1n-1}x_{n-1} + \gamma_{n-1n}x_n = \delta_{n-1}, \\ \gamma_{nn}x_n = \delta_n, \end{cases}$$

де $\gamma_{kk} \neq 0, k = \overline{1,n}$.

Тоді система має єдиний розв'язок

$$x_n = \frac{\delta_n}{\gamma_{nn}}, \quad x_{n-1} = \frac{\delta_{n-1}}{\gamma_{n-1n-1}} - \frac{\gamma_{n-1n}}{\gamma_{n-1n-1}} x_n, ...$$

Можливість 2. Останній ненульовий рядок матриці $\overline{C}$ має вигляд

$$c_r = (0, 0, ..., 0 \mid \delta_r), \quad \delta_r \neq 0.$$

Тоді ранг $C=r-1$, ранг $\overline{C}=r$, а тому ранг $A=r-1$, ранг $\overline{A}=r$. Система (2) у цьому випадку не матиме розв'язків, оскільки містить рівняння

$$0\cdot x_1+0\cdot x_2+\ldots+0\cdot x_n=\delta_r,\quad \delta_r\neq 0$$

типу $0=1$. На підставі цього задана система не матиме розв'язків.

Теорема 1 (Кронекера–Капеллі). *Для того щоб система лінійних рівнянь була сумісною, необхідно і достатньо, щоб збігалися ранги її основної та розширеної матриць.*

Доведення випливає з наведених вище міркувань.

Теорема 2 (про визначеність системи лінійних рівнянь). *Для того щоб сумісна система лінійних рівнянь мала лише один розв'язок, необхідно і достатньо, щоб ранг її матриці збігався з числом її невідомих.*

Вказівка щодо застосування методу виключення невідомих. При застосуванні елементарних перетворень з рядками матриці для зведення її до ступеневого (або навіть до псевдоступеневого) вигляду іноді доцільно на якомусь кроці переставити стовпці (за винятком останнього). При такій перестановці слід поміняти місцями і відповідні невідомі. Проілюструємо це.

Приклад. Дослідити систему

$$\begin{cases} -x_1+(1+\lambda)x_2+(2-\lambda)x_3+\lambda x_4=3,\\ \lambda x_1-x_2+(2-\lambda)x_3+\lambda x_4=2,\\ \lambda x_1+\lambda x_2+(2-\lambda)x_3+\lambda x_4=2,\\ \lambda x_1+\lambda x_2+(2-\lambda)x_3-x_4=2 \end{cases}$$

та знайти її розв'язки залежно від параметра λ.

Запишемо розширену матрицю системи

$$\overline{A}=\left(\begin{array}{cccc|c} -1 & 1+\lambda & 2-\lambda & \lambda & 3\\ \lambda & -1 & 2-\lambda & \lambda & 2\\ \lambda & \lambda & 2-\lambda & \lambda & 2\\ \lambda & \lambda & 2-\lambda & -1 & 2 \end{array}\right).$$

Застосуємо до матриці $\overline{A}$ серію елементарних перетворень: віднімемо від останнього рядка 3-й рядок, потім від 3-го віднімемо 2-й, і, нарешті, до 2-го рядка додамо 1-й, помножений на λ. Дістанемо

$$\overline{B}=\left(\begin{array}{cccc|c} -1 & 1+\lambda & 2-\lambda & \lambda & 3 \\ 0 & \lambda^2+\lambda-1 & -\lambda^2+\lambda+2 & \lambda^2+\lambda & 2+3\lambda \\ 0 & \lambda+1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda-1 & 0 \end{array}\right).$$

Переставимо 2-й і 3-й стовпці. Дістанемо ступеневу матрицю

$$\overline{C}=\left(\begin{array}{cccc|c} -1 & 2-\lambda & 1+\lambda & \lambda & 3 \\ 0 & -\lambda^2+\lambda+2 & \lambda^2+\lambda-1 & \lambda^2+\lambda & 2+3\lambda \\ 0 & 0 & 1+\lambda & 0 & 0 \\ 0 & 0 & 0 & -\lambda-1 & 0 \end{array}\right).$$

Отже, якщо $-\lambda^2+\lambda+2=(\lambda+1)(2-\lambda)\neq 0 \Rightarrow \lambda\neq -1$ і $\lambda\neq 2$, то матриця $\overline{C}$ – ступенева. У відповідній системі рівнянь слід поміняти місцями x_2 і x_3. Рівносильну систему рівнянь у цьому випадку можна подати у вигляді

$$\begin{cases} -x_1+(2-\lambda)x_3+(1+\lambda)x_2+\lambda x_4=3, \\ (\lambda+1)(2-\lambda)x_3+(\lambda^2+\lambda-1)x_2+\lambda(\lambda+1)x_4=2+3\lambda, \\ +(\lambda+1)x_2=0, \\ -(\lambda+1)x_4=0. \end{cases}$$

Її розв'язок

$$x_1=\frac{-1}{\lambda+1},\quad x_2=0,\quad x_3=\frac{2+3\lambda}{(\lambda+1)(2-\lambda)},\quad x_4=0$$

при $\lambda\neq -1$ і $\lambda\neq 2$.

При $\lambda=-1$ безпосередньо із системи дістанемо

$$\overline{C}=\left(\begin{array}{cccc|c} -1 & 3 & 0 & -1 & 3 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Рівносильна система матиме вигляд

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

Загальний розв'язок: $x_1 = -3 + 3x_3 - x_4$, $x_2 = 1$, x_3, x_4, де x_3, x_4 – вільні невідомі.

При $\lambda = 2$ дістанемо

$$\overline{C} = \left(\begin{array}{cccc|c} -1 & 0 & 3 & 2 & 3 \\ 0 & 0 & 5 & 6 & 8 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \end{array}\right) \sim \left(\begin{array}{cccc|c} -1 & 0 & 3 & 2 & 3 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Рівносильна система має рівняння виду $0 = 1$.
У цьому випадку система не має розв'язків.

Вправа

Дослідити систему лінійних рівнянь і знайти її розв'язки залежно від параметра λ:

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

Відокремлення базисних розв'язків

Нехай система лінійних рівнянь (1) з розширеною матрицею

$$\overline{A} = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2 \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m \end{array}\right)$$

сумісна, тобто ранг A = рангу $\overline{A} = r$. Тоді матриця A має базисний мінор Δ порядку r, який одночасно буде і базисним мінором матриці $\overline{A}$.

Для простоти викладу вважатимемо, що

$$\Delta = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2r} \\ \dots & \dots & \dots & \dots \\ \alpha_{r1} & \alpha_{r2} & \dots & \alpha_{rr} \end{pmatrix} \neq 0 .$$

На підставі теореми про ранг матриці рядки матриці $\overline{A}$, які заповнюють базисний мінор, одночасно будуть її базисними рядками. За нашим припущенням перші r її рядків $a_1, a_2, \dots, a_r$ будуть базисними.

За допомогою елементарних перетворень рядків матрицю $\overline{A}$ можна подати у вигляді

$$\overline{B} = \left(B \left| \begin{matrix} \beta_1 \\ \vdots \\ \beta_r \\ 0 \\ \vdots \end{matrix} \right. \right) = \left(\begin{matrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1r} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2r} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_{r1} & \alpha_{r2} & \dots & \alpha_{rr} & \dots & \alpha_{rn} \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{matrix} \left| \begin{matrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_r \\ 0 \\ \dots \\ 0 \end{matrix} \right. \right).$$

Вона одержується з $\overline{A}$ заміною всіх рядків з номерами $i > r$ нульовими рядками.

Наведемо ту серію елементарних перетворень (усі вони типу 2), яка перетворює матрицю $\overline{A}$ в $\overline{B}$.

Рядок a_i з номером $i > r$ матриці $\overline{A}$ лінійно подається через рядки $a_1, a_2, \dots, a_r$:

$$a_i = \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_r a_r .$$

До рядка a_i додамо рядок a_1, помножений на $-\lambda_1$, потім до нього додамо рядок a_2, помножений на $-\lambda_2$, і т. д. У результаті i-й рядок заповниться нулями. Продовжуючи далі ці перетворення, аналогічно дістанемо врешті-решт матрицю $\overline{B}$.

Задана система (1) рівносильна новій системі рівнянь

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1r}x_r + \alpha_{1r+1}x_{r+1} + \ldots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \ldots + \alpha_{2r}x_r + \alpha_{2r+1}x_{r+1} + \ldots + \alpha_{2n}x_n = \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{r1}x_1 + \ldots + \alpha_{rr}x_r + \alpha_{rr+1}x_{r+1} + \ldots + \alpha_{rn}x_n = \beta_r. \end{cases} \qquad (3)$$

Рівняння цієї системи називаються *базисними*.

Перетворимо її до вигляду

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1r}x_r = \beta_1 - \alpha_{1r+1}x_{r+1} - \ldots - \alpha_{1n}x_n, \\ \alpha_{21}x_1 + \ldots + \alpha_{2r}x_r = \beta_2 - \alpha_{2r+1}x_{r+1} - \ldots - \alpha_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{r1}x_1 + \ldots + \alpha_{rr}x_r = \beta_r - \alpha_{rr+1}x_{r+1} - \ldots - \alpha_{rn}x_n. \end{cases} \qquad (4)$$

Якщо $r < n$, то невідомі $x_{r+1}, x_{r+2}, \ldots, x_n$ ми назвемо *вільними* і подамо невідомі $x_1, x_2, \ldots, x_r$ через ці вільні невідомі, застосувавши будь-який з методів, зокрема, метод Крамера.

Якщо $r = n$, то система (3) набуде вигляду

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots + \alpha_{nn}x_n = \beta_n. \end{cases}$$

Оскільки її визначник $\Delta \neq 0$, система має лише єдиний розв'язок. Цей розв'язок буде єдиним і для заданої системи (1).

Вправи

1. Знайти необхідні й достатні умови для того, щоб n прямих площини

$$a_i x + b_i y + c_i = 0\,, \quad i = \overline{1,n}$$

проходили через одну точку.

2. Знайти умови необхідні й достатні для того, щоб чотири точки площини (x_i, y_i), $i = \overline{1,4}$, що не належать одній прямій, належали б одному колу.

3. Довести, що коли система лінійних рівнянь має два різні розв'язки a і b, то їх лінійна комбінація виду $\lambda a+(1-\lambda)b$ буде також розв'язком цієї системи при будь-яких значеннях параметра λ.

4. Встановити необхідну і достатню умову для того, щоб лінійна комбінація $\lambda_1 a_1+\lambda_2 a_2+\ldots+\lambda_k a_k$ розв'язків $a_1, a_2, \ldots, a_k$ неоднорідної системи лінійних рівнянь, була розв'язком цієї системи.

5. Встановити необхідну і достатню умову для того, щоб сума $a+b$ двох розв'язків a і b системи рівнянь (1) була б розв'язком цієї ж системи. Те саме для λa, $\lambda \neq 1$.

6. За якої умови n точок (x_i, y_i, z_i), $i=\overline{1,n}$ належать одній площині та містяться на одній прямій?

§ 4.2. Системи лінійних однорідних рівнянь

Система лінійних рівнянь

$$\begin{cases} \alpha_{11}x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n=0, \\ \alpha_{21}x_1+\alpha_{22}x_2+\ldots+\alpha_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1+\alpha_{m2}x_2+\ldots+\alpha_{mn}x_n=0 \end{cases} \qquad (1)$$

з коефіцієнтами α_{ik} $(i=\overline{1,m},\ k=\overline{1,n})$ із поля P називається *системою лінійних однорідних рівнянь* з невідомими $x_1, x_2, \ldots, x_n$.

Система (1) завжди сумісна, оскільки нульова послідовність (0, 0, ..., 0) скалярів – її розв'язок.

Теорема 1. *Множина L усіх розв'язків системи лінійних однорідних рівнянь є лінійний підпростір простору P^n. Розмірність його $\dim L=n-r$, де n – число невідомих і r – ранг матриці $A=(\alpha_{ik})$, $i=\overline{1,m}$, $k=\overline{1,n}$ цієї системи.*

Доведення. Нехай $x, y \in L$ і $\lambda, \mu \in P$. Якщо $x=(\xi_1, \xi_2, \ldots, \xi_n)$ і $y=(\eta_1, \eta_2, \ldots, \eta_n)$, то $\sum_{k=1}^{n}\alpha_{ik}\xi_k=0$ і $\sum_{k=1}^{n}\alpha_{ik}\eta_k=0$ для будь-яких $i=\overline{1,m}$.

Координатами лінійної комбінації $w = \lambda x + \mu y$ будуть скаляри $\omega_k = \lambda\xi_k + \mu\eta_k$, $k = \overline{1,n}$. Тоді

$$\sum_{k=1}^{n}\alpha_{ik}\omega_k = \sum_{k=1}^{n}\alpha_{ik}(\lambda\xi_k + \mu\eta_k) = \lambda\cdot\sum_{k=1}^{n}\alpha_{ik}\xi_k + \mu\cdot\sum_{k=1}^{n}\alpha_{ik}\eta_k = 0,$$

$$w \in L.$$

Множина L – лінійний підпростір.

Нехай r – ранг матриці

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

системи (1). Згідно з загальною теорією за умови $r = n$ система має лише один розв'язок – нульовий (тривіальний). У цьому випадку $\dim L = 0$.

Розглянемо випадок $r < n$.

За допомогою серії елементарних перетворень зведемо матрицю A до ступеневого вигляду

$$C = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1r} & \cdots & \gamma_{1n} \\ 0 & \gamma_{22} & \cdots & \gamma_{2r} & \cdots & \gamma_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \gamma_{rr} & \cdots & \gamma_{rn} \\ 0 & 0 & \ldots & 0 & \ldots & 0 \end{pmatrix}.$$

Тут заради наочності викладу вважатимемо, що провідними елементами ненульових рядків $c_1, c_2, \ldots, c_r$ є $\gamma_{11}, \gamma_{22}, \ldots, \gamma_{rr}$. Тоді система рівнянь, рівносильна заданій (1), матиме вигляд

$$\begin{cases} \gamma_{11}x_1 + \gamma_{12}x_2 + \ldots + \gamma_{1r}x_r + \ldots + \gamma_{1n}x_n = 0, \\ \gamma_{22}x_2 + \ldots + \gamma_{2r}x_r + \ldots + \gamma_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \gamma_{rr}x_r + \ldots + \gamma_{rn}x_n = 0. \end{cases} \tag{2}$$

Вільними невідомими будуть $x_{r+1}, x_{r+2}, ..., x_n$. Отже,

$$x_1 = \beta_{1r+1}x_{r+1} + \ldots + \beta_{1n}x_n,$$
$$x_2 = \beta_{2r+1}x_{r+1} + \ldots + \beta_{2n}x_n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_r = \beta_{rr+1}x_{r+1} + \ldots + \beta_{rn}x_n.$$

Загальний розв'язок подається у векторному запису

$$X = (\beta_{1r+1}x_{r+1} + \ldots + \beta_{1n}x_n, \ldots, \beta_{rr+1}x_{r+1} + \ldots + \beta_{rn}x_n, x_{r+1}, \ldots, x_n).$$

Виділимо $n-r$ частинних розв'язків, надаючи вільним невідомим послідовно скалярні значення, наведені в рядках квадратної матриці порядку $d=n-r$:

x_{r+1}	x_{r+2}	...	x_n
δ_{11}	δ_{12}	...	δ_{1d}
δ_{21}	δ_{22}	...	δ_{2d}
...	...	...	...
δ_{d1}	δ_{d2}	...	δ_{dd}

за умови, що

$$\det D = \begin{vmatrix} \delta_{11} & \delta_{12} & \ldots & \delta_{1d} \\ \delta_{21} & \delta_{22} & \ldots & \delta_{2d} \\ \ldots & \ldots & \ldots & \ldots \\ \delta_{d1} & \delta_{d2} & \ldots & \delta_{dd} \end{vmatrix} \neq 0.$$

Залежно від їх значень обчислюються $x_1, x_2, ..., x_r$.

Часто такою матрицею обирається одинична матриця

x_{r+1}	x_{r+2}	...	x_n
1	0	...	0
0	1	...	0
...	...	...	...
0	0	...	1

При такому заданні матриці частинними розв'язками будуть вектори

$$b_1 = (\beta_{1\,r+1}, \beta_{2\,r+1}, ..., \beta_{r\,r+1}, 1, 0, 0, ..., 0),$$

$$b_2 = (\beta_{1\,r+2}, \beta_{2\,r+2}, ..., \beta_{r\,r+2}, 0, 1, 0, ..., 0),$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$b_{n-r} = (\beta_{1n}, \beta_{2n}, ..., \beta_{rn}, 0, 0, 0, ..., 1).$$

Ця система векторів $b_1, b_2, ..., b_{n-r}$ – лінійно незалежна. Тому будь-який вектор X можна подати у вигляді

$$X = x_{r+1}b_1 + x_{r+2}b_2 + ... + x_n b_{n-r}. \qquad (3)$$

Отже, будь-який розв'язок системи рівнянь (2) і системи рівнянь (1) лінійно виражається через вектори $b_1, b_2, ..., b_{n-r}$. Вони становлять базис підпростору L з $\dim L = n - r$.

Базис підпростору L називається *фундаментальною системою розв'язків* системи лінійних рівнянь (1).

З доведеної вище теореми випливає правило обчислення фундаментальної системи розв'язків системи рівнянь:

1. Методом виключення невідомих знаходимо ранг r і загальний розв'язок системи рівнянь (1).

2. Якщо $x_{r+1}, x_{r+2}, ..., x_n$ – вільні невідомі, тоді складаємо таблицю,

	x_1	x_2	…	x_r	x_{r+1}	x_{r+2}	…	x_n
b_1	$\beta_{1\,r+1}$	$\beta_{2\,r+1}$	…	$\beta_{r\,r+1}$	1	0	…	0
b_2	$\beta_{1\,r+2}$	$\beta_{2\,r+2}$	…	$\beta_{r\,r+2}$	0	1	…	0
…	…	…	…	…	…	…	…	…
b_{n-r}	β_{1n}	β_{2n}	…	β_{rn}	0	0	…	1

заповнюючи послідовно її рядки. У 1-му рядку b_1 надаємо вільним змінним значення $x_{r+1} = 1$, $x_{r+2} = 0, ...,$ $x_n = 0$ і, залежно від них, обчислюємо $x_1, x_2, ..., x_r$. Потім у 2-му рядку b_2 надаємо вільним змінним значення $x_{r+1} = 0$, $x_{r+2} = 1, ...,$ $x_n = 0$ і, залежно від них, обчислюємо значення $x_1, x_2, ..., x_r$ і т. д. Одержані рядки $b_1, b_2, ..., b_{n-r}$ і є фундаментальна система розв'язків.

3. Замість одиничної матриці в останніх $n-r$ стовпцях можна вибрати будь-яку квадратну матрицю, визначник якої відмінний від нуля.

Приклад. Знайти фундаментальну систему розв'язків однорідної системи

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

За допомогою елементарних перетворень матрицю

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

зведемо до вигляду

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

Тут у матриці A до 2-го рядка додали 1-й, потім до 3-го додали 1-й, помножений на –2, і, нарешті, до 4-го рядка додали 1-й, помножений на –3.

Застосуємо перетворення: до 3-го рядка додамо 2-й, а до 4-го додамо 2-й, помножений на 2. Дістанемо

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

Отже, система, рівносильна заданій, матиме вигляд

$$\begin{cases} x_1 + 2x_2 - x_3 + 2x_4 + x_5 = 0, \\ x_2 + 3x_3 + 3x_4 - x_5 = 0. \end{cases}$$

Звідси

$$x_1 = 7x_3 + 4x_4 - 3x_5, \quad x_2 = -3x_3 - 3x_4 + x_5 .$$

Загальний розв'язок матиме вигляд

$$X = (7x_3 + 4x_4 - 3x_5, \ -3x_3 - 3x_4 + x_5, \ x_3, x_4, x_5).$$

Складаємо таблицю

	x_1	x_2	x_3	x_4	x_5
b_1	7	–3	1	0	0
b_2	4	–3	0	1	0
b_3	–3	1	0	0	1

Рядки цієї таблиці

$$b_1 = (7, -3, 1, 0, 0),$$
$$b_2 = (4, -3, 0, 1, 0),$$
$$b_3 = (-3, 1, 0, 0, 1)$$

є шуканою фундаментальною системою розв'язків. Неважко переконатися, що вектори b_1, b_2, b_3 – лінійно незалежні, і будь-який вектор із R^5, поданий у вигляді лінійної комбінації цих векторів b_1, b_2, b_3, також буде розв'язком заданої системи.

Теорема 2. *Рівняння*

$$\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n = 0 \qquad (4)$$

є наслідком системи лінійних рівнянь (1) *тоді і лише тоді, коли вектор* $b = (\beta_1, \beta_2, ..., \beta_n)$ *лінійно подається через вектори* $a_1, a_2, ..., a_m$ *–рядки матриці системи* (1).

Доведення. *Необхідність.* Якщо вектор b подається у вигляді лінійної комбінації $b = \lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_m a_m$, тоді рівняння (3) одержується із системи (1) за допомогою елементарних перетворень. При цьому до рівняння виду $0 \cdot x_1 + 0 \cdot x_2 + \ldots + +0 \cdot x_n = 0$ додаємо ліву і праву частини першого рівняння системи, помножені на λ_1, потім додаємо ліву і праву частини другого рівняння, помножені на λ_2 і т. д. В результаті одержуємо рівняння (4). Тому воно – наслідок системи (1).

Достатність. Нехай рівняння (4) – наслідок системи (1). Розглянемо систему лінійних рівнянь, одержаних об'єднанням рівнянь систем (1) і (4):

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = 0, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n = 0, \\ \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n = 0 \end{cases} \tag{5}$$

з розширеною матрицею

$$\overline{A} = \left(\frac{A}{\beta_1 \beta_2 \ldots \beta_n} \right) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}.$$

Позначимо через $c_1, c_2, \ldots, c_n$ стовпці матриці A, а $\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_n$ – стовпці розширеної матриці $\overline{A}$. Нехай r – ранг матриці $\overline{A}$. Вважатимемо, що $\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_r$ – базис, складений із стовпців. Будь-який стовпець $\overline{c}_j$ лінійно подається через ці базисні стовпці. Тоді й довільний стовпець c_j матриці A лінійно подається через стовпці $c_1, c_2, \ldots, c_r$.

Доведемо лінійну незалежність векторів $c_1, c_2, \ldots, c_r$.

Нехай $\lambda_1, \lambda_2, \ldots, \lambda_r$ – така послідовність скалярів, для якої виконується рівність $\lambda_1 c_1 + \lambda_2 c_2 + \ldots + \lambda_r c_r = \theta$. Тоді $\lambda_1 c_1 + \lambda_2 c_2 + \ldots + \lambda_r c_r + 0 \cdot c_{r+1} + \ldots + 0 \cdot c_n = \theta$ і, отже, послідовність $\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0$ – розв'язок системи рівнянь (1). За припущенням рівняння (4) – наслідок системи (1). Тому послідовність $\lambda_1, \lambda_2, ..., \lambda_r, 0, ..., 0$ – розв'язок системи рівнянь (5), тобто

$$\lambda_1\overline{c_1} + \lambda_2\overline{c_2} + ... + \lambda_r\overline{c_r} + 0\cdot\overline{c}_{r+1} + ... + 0\cdot\overline{c}_n = \theta$$

і, отже, $\lambda_1\overline{c_1} + \lambda_2\overline{c_2} + ... + \lambda_r\overline{c_r} = \theta$. Звідси одержуємо $\lambda_1 = = \lambda_2 = ... = \lambda_r = 0$. Останнє і означає, що вектори-стовпці $c_1, c_2, ..., c_r$– лінійно незалежні.

Ранг матриці A дорівнює рангу r розширеної матриці $\overline{A}$.

Із теореми про ранг матриці випливає, що ранги матриць A і $\overline{A}$ за рядками також дорівнюють r. Тому останній рядок у матриці лінійно подається через рядки $a_1, a_2, ..., a_r$.

Теорему доведено.

Вправи

1. Знайти базис і розмірність лінійного підпростору L простору P^n, якщо

а) L задано рівнянням $x_1 + x_2 + ... + x_n = 0$;

б) L – множина всіх векторів, у яких координати з парними номерами збігаються;

в) L – множина всіх векторів, у яких рівні всі координати з непарними і парними номерами.

2. Встановити розмірність підпростору розв'язків системи залежно від значень параметра λ

$$\begin{cases} (8-\lambda)x_1 + 2x_2 + 3x_3 + \lambda x_4 = 0, \\ x_1 + (9-\lambda)x_2 + 4x_3 + \lambda x_4 = 0, \\ x_1 + 2x_2 + (10-\lambda)x_3 + \lambda x_4 = 0, \\ x_1 + 2x_2 + 3x_3 + \lambda x_4 = 0. \end{cases}$$

3. У просторі многочленів степеня ≤ 5 знайти базис лінійного підпростору многочленів $f(t)$, для яких $f(0) = f(1) = f(2) = = f(3) = 0$.

§ 4.3. Спряжені лінійні підпростори

Нехай у просторі P^n задано дві системи векторів

$$\begin{aligned} a_1 &= (\alpha_{11}, \alpha_{12}, ..., \alpha_{1n}), \\ a_2 &= (\alpha_{21}, \alpha_{22}, ..., \alpha_{2n}), \\ &\dots\dots\dots\dots\dots\dots \\ a_m &= (\alpha_{m1}, \alpha_{m2}, ..., \alpha_{mn}). \end{aligned} \tag{A}$$

$$\begin{aligned} b_1 &= (\beta_{11}, \beta_{12}, ..., \beta_{1n}), \\ b_2 &= (\beta_{21}, \beta_{22}, ..., \beta_{2n}), \\ &\dots\dots\dots\dots\dots\dots \\ b_l &= (\beta_{l1}, \beta_{l2}, ..., \beta_{ln}). \end{aligned} \tag{Б}$$

Якщо вони лінійно еквівалентні, то їхні лінійні оболонки збігаються

$$L(a_1, a_2, ..., a_m) = L(b_1, b_2, ..., b_l) = L.$$

Система рівнянь

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

рівносильна системі

$$\begin{cases} \beta_{11}x_1 + \beta_{12}x_2 + ... + \beta_{1n}x_n = 0, \\ \beta_{21}x_1 + \beta_{22}x_2 + ... + \beta_{2n}x_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta_{l1}x_1 + \beta_{l2}x_2 + ... + \beta_{ln}x_n = 0, \end{cases} \tag{1'}$$

якщо системи векторів (А) і (Б) лінійно еквівалентні.

Позначимо через L^* лінійний простір розв'язків системи (1) і, отже, рівносильної системи (1′). Він не залежить від вибору систем векторів, що породжують підпростір L. Такими системами векторів, зокрема, можна вибрати будь-який базис підпростору L.

Тому доцільно ввести таке означення.

Означення. *Спряженим підпростором L^* для лінійного простору $L = L(a_1, a_2, ..., a_m)$ називається множина всіх розв'язків системи* (1) *лінійних однорідних рівнянь. Рядками матриці $A = (\alpha_{ik})$ цієї системи будуть вектори $a_1, a_2, ..., a_m$.*

Із означення випливає

$$\dim L^* = n - r ,$$

де r – ранг системи векторів (A).

Теорема 1 (про двоїстість). *Для лінійного підпростору L векторного простору P^n виконується рівність*

$$\left(L^*\right)^* = L .$$

Доведення. Подамо підпростір L у вигляді лінійної оболонки деякої системи векторів

$$a_i = (\alpha_{i1}, \alpha_{i2}, ..., \alpha_{in}),\ i = \overline{1, m}.$$

Тоді L^* – множина всіх розв'язків системи рівнянь (1).

Нехай $c_1, c_2, ..., c_p$ – фундаментальна система її розв'язків, $p = n - r$ і $c_i = (\gamma_{j1}, \gamma_{j2}, ..., \gamma_{jn}), j = \overline{1, p}$.

Простір $M = \left(L^*\right)^*$ збігається з множиною розв'язків системи рівнянь

$$\begin{cases} \gamma_{11}y_1 + \gamma_{12}y_2 + ... + \gamma_{1n}y_n = 0, \\ \gamma_{21}y_1 + \gamma_{22}y_2 + ... + \gamma_{2n}y_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \gamma_{p1}y_1 + \gamma_{p2}y_2 + ... + \gamma_{pn}y_n = 0. \end{cases} \tag{2}$$

Оскільки c_j – розв'язок системи (1), то має виконуватись рівність

$$\alpha_{i1}\gamma_{j1} + \alpha_{i2}\gamma_{j2} + ... + \alpha_{in}\gamma_{jn} = 0 \tag{3}$$

для будь-яких $i=\overline{1,m}; j=\overline{1,p}$.

Із рівності (3) випливає, що

$$a_1, a_2, ..., a_m \subseteq M = \left(L^*\right)^*.$$

Тому $L = L(a_1, a_2, ..., a_m) \subseteq M$.

Доведемо обернене включення $M \subseteq L$.

Оберемо довільний вектор $b = (\beta_1, \beta_2, ..., \beta_n) \in M$. Для нього виконується рівність

$$\gamma_{j1}\beta_1 + \gamma_{j2}\beta_2 + ... + \gamma_{jn}\beta_n = 0 \tag{4}$$

для всіх $j=\overline{1,p}$. Оскільки будь-який вектор $c = (\gamma_1, \gamma_2, ..., \gamma_n)$ із підпростору L^* – лінійна комбінація векторів $c_1, c_2, ..., c_p$, то c – розв'язок рівняння

$$\beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n = 0. \tag{5}$$

Отже, будь-який розв'язок системи (1) є одночасно і розв'язком рівняння (5). За теоремою про рівняння – наслідок (теорема 2, § 4.2) вектор b лінійно подається через $a_1, a_2, ..., a_m$, тобто $b \in L$.

Включення $M \subseteq L$ доведено. Поряд з цим доведено рівність

$$\left(L^*\right)^* = L^{**} = L.$$

Зауваження. Доведення може бути спрощене, коли врахувати, що $\dim L = \dim\left(L^*\right)^*$.

Способи задання лінійних підпросторів

Із ефективних способів задання лінійних підпросторів особливо відзначимо два:

I. Підпростір $L = L(a_1, a_2, ..., a_m)$ – лінійна оболонка системи векторів $a_1, a_2, ..., a_m$.

II. Підпростір L – множина розв'язків заданої системи лінійних однорідних рівнянь.

При розв'язанні деяких задач потрібно вміти переходити від одного з цих способів задання підпростору до іншого.

Якщо лінійний підпростір L задано системою лінійних рівнянь, тоді перехід до першого способу полягає в розв'язанні цієї системи рівнянь.

Із теореми двоїстості випливає такий спосіб переходу від першого способу задання лінійного підпростору до другого.

Нехай $L = L(a_1, a_2, ..., a_m)$, де

$$\begin{gathered}
a_1 = (\alpha_{11}, \alpha_{12}, ..., \alpha_{1n}), \\
a_2 = (\alpha_{21}, \alpha_{22}, ..., \alpha_{2n}), \\
\dots\dots\dots\dots\dots\dots\dots\dots \\
a_m = (\alpha_{m1}, \alpha_{m2}, ..., \alpha_{mn}).
\end{gathered}$$

Складемо систему рівнянь

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

та обчислимо фундаментальну систему її розв'язків $c_1, c_2, ..., c_p$. Якщо $c_j = (\gamma_{j1}, \gamma_{j2}, ..., \gamma_{jn}), j = \overline{1, p}$, то система лінійних рівнянь

$$\begin{cases}
\gamma_{11}y_1 + \gamma_{12}y_2 + ... + \gamma_{1n}y_n = 0, \\
\gamma_{21}y_1 + \gamma_{22}y_2 + ... + \gamma_{2n}y_n = 0, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\gamma_{p1}y_1 + \gamma_{p2}y_2 + ... + \gamma_{pn}y_n = 0
\end{cases}$$

визначає підпростір L, і множина всіх розв'язків цієї системи збігається з L.

Методичні вказівки. Нехай $L = L_1 + L_2$, де L_1, L_2 – лінійні підпростори. Для знаходження базису підпростору L слід подати L_1 і L_2 у вигляді лінійних оболонок $L_1 = L(a_1, a_2, ..., a_m)$ і $L_2 = L(b_1, b_2, ..., b_p)$. Тоді $L = L(a_1, a_2, ..., a_m, b_1, b_2, ..., b_p)$.

Нехай $L = L_1 \cap L_2$. Для знаходження базису підпростору L слід подати L_1 і L_2 лінійними рівняннями. Якщо L_1 подається системою S_1 лінійних однорідних рівнянь, а L_2 – системою S_2 рівнянь, то перетин $L_1 \cap L_2$ подається системою $S_1 \cup S_2$ – об'єднанням усіх рівнянь систем S_1 і S_2.

Приклад 1. Нехай $L_1 = L(a_1, a_2, a_3)$ і $L_2 = L(b_1, b_2, b_3)$, де

$$\begin{aligned} a_1 &= (1, 2, 1), & b_1 &= (2, 3, -1), \\ a_2 &= (1, 1, -1), & b_2 &= (1, 2, 2), \\ a_3 &= (1, 3, 3), & b_3 &= (1, 1, -3). \end{aligned}$$

Знайти базис підпростору $L_1 \cap L_2$.

Складемо систему рівнянь

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

що визначає L_1^*. Її матриця

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

Рівносильною системою буде

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ \quad\quad x_2 + 2x_3 = 0. \end{cases}$$

Загальний розв'язок її $X = (3x_3, -2x_3, x_3)$.

Фундаментальна система розв'язків містить лише один вектор $c_1 = (3, -2, 1)$ і $L_1^* = L(c_1)$.

Підпростір L_1 збігається з множиною розв'язків рівняння

$$3y_1 - 2y_2 + y_3 = 0 . \qquad (S_1)$$

Система рівнянь

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

визначає L_2^*. Її матриця

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

Рівносильною системою буде

$$\begin{cases} x_1 + x_2 - 3x_3 = 0, \\ \qquad x_2 + 5x_3 = 0. \end{cases}$$

Загальний її розв'язок $X = (8x_3, -5x_3, x_3)$. Фундаментальна система розв'язків містить лише один вектор $c_2 = (8, -5, 1)$ і $L_2^* = L(c_2)$.

Простір L_2 визначається рівнянням

$$8y_1 - 5y_2 + y_3 = 0 . \qquad (S_2)$$

Перетин $L_1 \cap L_2$ визначається системою рівнянь

$$\begin{cases} 3y_1 - 2y_2 + y_3 = 0, \\ 8y_1 - 5y_2 + y_3 = 0. \end{cases}$$

Загальний її розв'язок $Y = \left(y_1, \frac{5}{3}y_1, \frac{1}{3}y_1 \right)$.

Вектор $c = (3, 5, 1)$ – базис перетину $L_1 \cap L_2$.

Теорема 2. *Якщо L – лінійний підпростір векторного простору P^n з дійсним основним полем P (P – підполе поля R дійсних чисел), тоді*

$$P^n = L \oplus L^*.$$

Доведення. Подамо підпростір у вигляді лінійної оболонки деякої системи векторів $a_1, a_2, ..., a_m$: $L = L(a_1, a_2, ..., a_m)$. Нехай $M = L + L^*$. Нам відомо, що коли $\dim L = r$, тоді $\dim L^* = n - r$.

Покажемо, що $L \cap L^* = \{\theta\}$.

Нехай $x = (\xi_1, \xi_2, ..., \xi_n) \in L \cap L^*$. Система векторів $a_1, a_2, ..., a_m, x$ лінійно еквівалентна системі $a_1, a_2, ..., a_m$, оскільки $x \in L$. Якщо $a_i = (\alpha_{i1}, \alpha_{i2}, ..., \alpha_{in}), i = \overline{1, n}$, тоді підпростір L^* – множина всіх розв'язків системи рівнянь

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n = 0, \\ \xi_1 x_1 + \xi_2 x_2 + \ldots + \xi_n x_n = 0. \end{cases}$$

Оскільки $x \in L^*$, то координати його $\xi_1, \xi_2, ..., \xi_n$ задовольняють цю систему і, зокрема, задовольняють останнє рівняння. Це означає, що

$$\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 = 0.$$

Для дійсних чисел $\xi_1, \xi_2, ..., \xi_n$ ця рівність правильна лише при $\xi_1 = \xi_2 = \ldots = \xi_n = 0$. Отже, $x = \theta$. Тому $L \cap L^* = \{\theta\}$ і $M = L \oplus L^*$. Оскільки $\dim M = \dim L + \dim L^* = n$, то $L + L^* = P^n$.

Доведена теорема не розповсюджується на простори над довільним полем P. Вона втрачає силу вже для комплексного векторного простору C^n. Проілюструємо це на прикладі простору C^3.

Приклад 2. Нехай $a=(1, i, 0)\in C^3$ і $L=L(a)$. Знайдемо базис спряженого простору L^*. Для цього складемо відповідну систему рівнянь. Вона містить лише одне рівняння

$$1\cdot x_1 + i\cdot x_2 + 0\cdot x_3 = 0.$$

Загальний розв'язок системи $X=(-ix_2, x_2, x_3)$.

За відомою схемою знаходимо фундаментальну систему розв'язків. Вона складається з двох векторів $c_1=a=(1, i, 0)$ і $c_2=(0, 0, 1)$. При цьому $L^*=L(a, c_2)$, $a\in L\cap L^*$ і $a\neq\theta$; $C^3\neq L\oplus L^*$.

Вправи

1. Знайти базис і розмірність суми та перетину підпросторів $L_1=L(a_1, a_2, a_3)$ і $L_2=L(b_1, b_2)$, де

$a_1=(1, 2, -1, -2)$, $b_1=(2, 5, -6, -5)$,
$a_2=(3, 1, 1, 1)$, $b_2=(-1, 2, -7, -3)$.
$a_3=(-1, 0, 1, -1)$,

2. Нехай L_1, L_2 – лінійні підпростори векторного простору P^n над довільним полем P. Довести:

а) $(L_1+L_2)^*=L_1^*\cap L_2^*$;

б) $(L_1\cap L_2)^*=L_1^*+L_2^*$.

3. Нехай $P=\{\overline{0}, \overline{1}\}$ – поле з двох елементів. Навести приклади підпросторів із P^3, для яких їхній перетин із спряженим підпростором відмінний від тривіального підпростору.

§ 4.4. Лінійні багатовиди

Означення. *Нехай L – довільний лінійний підпростір векторного простору P^n і $a\in P^n$. Множина $M=a+L=\{w\in P^n\colon w=a+x, x\in L\}$ називається лінійним багатовидом. Його одержано паралельним зсувом підпростору L.*

Прикладами лінійних багатовидів у просторі R^3 є прямі лінії. Вони одержуються паралельними зсувами одновимірних лінійних підпросторів – прямих ліній, що проходять через початок координат.

Наведемо деякі загальні властивості лінійних багатовидів.

I. Лінійний багатовид $M = a + L$ з базисним підпростором L буде лінійним підпростором тоді і лише тоді, коли $M = L$ (тобто, коли $a \in L$).

Доведення. *Необхідність.* Припустимо, що M – лінійний підпростір. Тоді $\theta \in M$ і тому $\theta = a + x_0$, де $x_0 \in L$. Звідси $a = -x_0 \in L$ і тому $M = a + L = L$.

Достатність очевидна.

II. Базисний підпростір L лінійного багатовиду єдиний: якщо $a + L = a' + L'$, де L і L' – підпростори, тоді $L = L'$.

Доведення. Якщо $M = a + L = a' + L'$, тоді $(a - a') + L = L'$ і, отже, $a - a' \in L$. Тоді $(a - a') + L = L$ і $L = L'$.

Означення. *Розмірністю лінійного багатовиду M з базисним підпростором L називається число, що дорівнює* $\dim L$. *Воно позначається* $\dim M$.

Теорема 1. *Множина всіх розв'язків сумісної системи лінійних рівнянь є лінійний багатовид.*

Доведення. Нехай

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n = \beta_m \end{cases} \tag{1}$$

є сумісна система лінійних рівнянь і $l = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ – який-небудь її частинний розв'язок. Поряд з системою (1) розглянемо систему

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1n}x_n = 0, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \ldots + \alpha_{mn}x_n = 0 \end{cases} \tag{1'}$$

однорідних рівнянь. Її будемо називати асоційованою по відношенню до заданої системи (1).

Позначимо через L підпростір усіх розв'язків системи (1′) і нехай $M = l + L$. Якщо $w \in M$, тоді $w = l + x$, де $x = (\xi_1, \xi_2, ..., \xi_n) \in L$ і $w = (w_1, w_2, ..., w_n) = (\lambda_1 + \xi_1, \lambda_2 + \xi_2, ..., \lambda_n + \xi_n)$. Отже, $\sum_{k=1}^{n} \alpha_{ik} w_k = \sum_{k=1}^{n} \alpha_{ik} (\lambda_k + \xi_k) = \sum_{k=1}^{n} \alpha_{ik} \lambda_k + \sum_{k=1}^{n} \alpha_{ik} \xi_k = \beta_i$, оскільки $\sum_{k=1}^{n} \alpha_{ik} \xi_k = 0$. Одержана рівність правильна для всіх $i = \overline{1, m}$.

Таким чином, усі вектори w із багатовиду $M = L + l$ будуть розв'язками системи (1).

Доведемо обернене твердження: будь-який розв'язок системи (1) належить багатовиду M.

Нехай $y = (\eta_1, \eta_2, ..., \eta_n)$ – розв'язок системи (1), тобто $\sum_{k=1}^{n} \alpha_{ik} \eta_k = \beta_i$ для всіх $i = \overline{1, m}$. Вектор $x = y - l = (\xi_1, \xi_2, ..., \xi_n) \Rightarrow \xi_1 = \eta_1 - \lambda_1, \xi_2 = \eta_2 - \lambda_2, ..., \xi_n = \eta_n - \lambda_n$, і тоді

$$\sum_{k=1}^{n} \alpha_{ik} \xi_k = \sum_{k=1}^{n} \alpha_{ik} (\eta_k - \lambda_k) = \sum_{k=1}^{n} \alpha_{ik} \eta_k - \sum_{k=1}^{n} \alpha_{ik} \lambda_k = \beta_i - \beta_i = 0 .$$

Вектор x належить L і, отже, $y = l + x \in M$.

Множина всіх розв'язків системи (1) збігається з багатовидом $M = l + L$, базисний підпростір L якого визначається системою (1′), асоційованою з (1).

Теорема 2. *Будь-який лінійний багатовид M збігається з множиною всіх розв'язків деякої сумісної системи лінійних рівнянь. Якщо $r = \dim M$, то таку систему лінійних рівнянь можна вибрати таким чином, щоб вона мала $n - r$ рівнянь.*

Доведення. Нехай $M = l + L$, де $l = (\lambda_1, \lambda_2, ..., \lambda_n)$ і L – базисний підпростір багатовиду. Якщо $r = \dim M = \dim L$, то виберемо який-небудь базис $a_1, a_2, ..., a_r$ підпростору L. Тоді $L = L(a_1, a_2, ..., a_r)$. Із попереднього параграфа нам відомо, що лінійний підпростір L можна визначити як множину розв'язків системи лінійних однорідних рівнянь

$$\begin{cases} \beta_{11}x_1 + \beta_{12}x_2 + \ldots + \beta_{1n}x_n = 0, \\ \beta_{21}x_1 + \beta_{22}x_2 + \ldots + \beta_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_{p1}x_1 + \beta_{p2}x_2 + \ldots + \beta_{pn}x_n = 0, \end{cases}$$

причому можна вважати, що число рівнянь $p=n-r$ ($p=n-r$ – розмірність спряженого підпростору L^*).

Обчислимо $\beta_i = \sum_{k=1}^{n} \beta_{ik}\lambda_k$ для всіх $i=\overline{1,p}$ і утворимо систему лінійних рівнянь

$$\begin{cases} \beta_{11}x_1 + \beta_{12}x_2 + \ldots + \beta_{1n}x_n = \beta_1, \\ \beta_{21}x_1 + \beta_{22}x_2 + \ldots + \beta_{2n}x_n = \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_{p1}x_1 + \beta_{p2}x_2 + \ldots + \beta_{pn}x_n = \beta_p. \end{cases}$$

Вона сумісна: її частинним розв'язком буде вектор $l=(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

За теоремою 1 множина розв'язків цієї системи збігається з багатовидом $M = l + L$.

Доведені теореми дають змогу сформулювати інше означення лінійного багатовиду, еквівалентне початковому.

Означення. *Лінійним багатовидом називається множина всіх векторів векторного простору P^n, які є розв'язками сумісної системи лінійних рівнянь з коефіцієнтами із поля P.*

Теорема 3. *Будь-який лінійний багатовид M векторного простору P^n над полем P з розмірністю $\dim M < n$ належить множині розв'язків деякого лінійного рівняння*

$$\beta_1x_1 + \beta_2x_2 + \ldots + \beta_nx_n = \beta, \tag{2}$$

у якого $b=(\beta_1,\beta_2,\ldots,\beta_n) \neq \theta$.

Якщо $\dim M = n-1$, тоді M збігається з множиною всіх розв'язків рівняння виду (2).

Ця теорема – наслідок теореми 2.

Вправи

1. Довести, що коли M – лінійний багатовид простору P^n, то для будь-яких $x, y \in M$ і $\lambda, \mu \in P$, для яких $\lambda + \mu = 1$, вектор $w = \lambda x + \mu y \in M$.

2. Нехай M – підмножина простору P^n з такими властивостями: якщо $x, y \in M$ і $\lambda, \mu \in P$, для яких $\lambda + \mu = 1$, тоді $\lambda x + \mu y \in M$. Довести, що M – лінійний багатовид.

§ 5.1. Лінійні перетворення матриці

Нехай задано матрицю

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

розмірності $m \times n$ з елементами із поля P.

Лінійним перетворенням змінних $x_1, x_2, ..., x_m$ через нові змінні $y_1, y_2, ..., y_n$, визначеним матрицею A (умовний запис $x_1, x_2, ..., x_m \xrightarrow{A} y_1, y_2, ..., y_n$), називається перетворення

$$\begin{cases} x_1 = \alpha_{11} y_1 + \alpha_{12} y_2 + ... + \alpha_{1n} y_n, \\ x_2 = \alpha_{21} y_1 + \alpha_{22} y_2 + ... + \alpha_{2n} y_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ x_m = \alpha_{m1} y_1 + \alpha_{m2} y_2 + ... + \alpha_{mn} y_n. \end{cases} \tag{1}$$

Аналогічно визначимо друге лінійне перетворення

$$\begin{cases} y_1 = \beta_{11} z_1 + \beta_{12} z_2 + ... + \beta_{1p} z_p, \\ y_2 = \beta_{21} z_1 + \beta_{22} z_2 + ... + \beta_{2p} z_p, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ y_{n1} = \beta_{n1} z_1 + \beta_{n2} z_2 + ... + \beta_{np} z_p \end{cases} \tag{2}$$

за допомогою матриці

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2p} \\ \dots & \dots & \dots & \dots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{np} \end{pmatrix}$$

розмірності $n \times p$ з елементами із того самого поля P (умовний запис $y_1, y_2, ..., y_n \xrightarrow{B} z_1, z_2, ..., z_p$).

У результаті послідовного виконання двох перетворень – спочатку перетворення (1), а потім перетворення (2) – змінні $x_1, x_2, ..., x_m$ будуть подані через змінні $z_1, z_2, ..., z_p$. Це третє перетворення надалі називатимемо *добутком* лінійних перетворень (1) і (2).

Доведемо, що добуток двох лінійних перетворень змінних також є лінійним перетворенням.

Встановимо зв'язок змінних x_1 і $z_1, z_2, ..., z_p$:

$$x_1 = \alpha_{11} y_1 + \alpha_{12} y_2 + ... + \alpha_{1n} y_n =$$
$$= \alpha_{11}\left(\beta_{11} z_1 + \beta_{12} z_2 + ... + \beta_{1p} z_p\right) + \alpha_{12}\left(\beta_{21} z_1 + \beta_{22} z_2 + ... + \beta_{2p} z_p\right) +$$
$$+ ... + \alpha_{1n}\left(\beta_{n1} z_1 + \beta_{n2} z_2 + ... + \beta_{np} z_p\right) =$$
$$= \left(\alpha_{11}\beta_{11} + \alpha_{12}\beta_{21} + ... + \alpha_{1n}\beta_{n1}\right) z_1 + ... +$$
$$+ \left(\alpha_{11}\beta_{1p} + \alpha_{12}\beta_{2p} + ... + \alpha_{1n}\beta_{np}\right) z_p.$$

Позначимо коефіцієнти при $z_1, z_2, ..., z_p$ через

$$\gamma_{1k} = \alpha_{11}\beta_{1k} + \alpha_{12}\beta_{2k} + ... + \alpha_{1n}\beta_{nk},\ k = \overline{1, p}.$$

Тоді $x_1 = \gamma_{11} z_1 + \gamma_{12} z_2 + ... + \gamma_{1p} z_p$ і т. д. для $x_2, x_3, ...$ Добуток перетворень (1) і (2) подається у вигляді

$$\begin{cases} x_1 = \gamma_{11} z_1 + \gamma_{12} z_2 + ... + \gamma_{1p} z_p, \\ x_2 = \gamma_{21} z_1 + \gamma_{22} z_2 + ... + \gamma_{2p} z_p, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_m = \gamma_{m1} z_1 + \gamma_{m2} z_2 + ... + \gamma_{mp} z_p. \end{cases} \qquad (3)$$

Це – лінійне перетворення з матрицею

$$C = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\ \dots & \dots & \dots & \dots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mp} \end{pmatrix}$$

розмірності $m \times p$ і з елементами

$$\gamma_{ik} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \ldots + \alpha_{in}\beta_{nk} \quad \left(i = \overline{1,m};\ k = \overline{1,p}\right).$$

Вони обчислюються за таким правилом.

Позначимо через a_i рядок матриці A з номером i, тобто $a_i = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$, а через $b^{(k)}$ – стовпець матриці B з номером k, тобто

$$b^{(k)} = \begin{pmatrix} \beta_{1k} \\ \beta_{2k} \\ \vdots \\ \beta_{nk} \end{pmatrix}.$$

Їхній добуток визначимо за формулою

$$a_i \cdot b^{(k)} = (\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}) \begin{pmatrix} \beta_{1k} \\ \beta_{2k} \\ \vdots \\ \beta_{nk} \end{pmatrix} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \ldots + \alpha_{in}\beta_{nk}.$$

Означення. *Добутком $A \cdot B$ матриці A розмірності $m \times n$ на матрицю B розмірності $n \times p$ називається матриця C розмірності $m \times p$, елементи γ_{ik} якої обчислюються за формулою*

$$\gamma_{ik} = a_i \cdot b^{(k)} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \ldots + \alpha_{in}\beta_{nk},$$

тобто γ_{ik} дорівнює сумі добутків елементів i-го рядка лівого множника A на відповідні елементи k-го стовпця правого множника B.

Наведені вище обчислення показали, що добуток двох лінійних перетворень є лінійним перетворенням, і матриця його дорівнює добутку матриць даних перетворень, причому порядок розміщення перетворень і порядок матриць одержаного добутку збігаються.

Означення. *Сумою матриць $A = (\alpha_{ik})$ і $B = (\beta_{ik})$ однакової розмірності $m \times n$ називається матриця $C = (\gamma_{ik})$ такої самої* *розмірності і з елементами* $\gamma_{ik} = \alpha_{ik} + \beta_{ik}$. *Сума позначається* $A + B$. *Отже,*

$$A + B = (\alpha_{ik} + \beta_{ik}).$$

Добутком матриці $A = (\alpha_{ik})$ на скаляр λ називається матриця $(\lambda \cdot \alpha_{ik})$. Позначення: $\lambda \cdot A = (\lambda \cdot \alpha_{ik})$.

Теорема 1. *Операції з матрицями мають такі основні властивості (розподілимо їх на групи):*

I. Властивості додавання

I_1. Асоціативність: $A + (B + C) = (A + B) + C$.

I_2. Існування єдиної матриці O, для якої $A + O = O + A = A$ (*O – нульова матриця*).

I_3. Для будь-якої матриці A існує єдина матриця X, для якої $A + X = O$.

I_4. $A + B = B + A$.

II. Властивості множення

II_1. Асоціативність: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

II_2. $E \cdot A = A \cdot E' = A$ (E і E' *– ліва і права одиничні матриці порядків m і n для матриці A розмірності* $m \times n$).

II_3. $A \cdot (B + C) = A \cdot B + A \cdot C$.

II_4. $(A + B) \cdot C = A \cdot C + B \cdot C$.

III. Властивості множення на скаляр

III_1. $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$.

III_2. $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$.

III_3. $(\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A)$.

III_4. $1 \cdot A = A$.

III_5. $\alpha \cdot (A \cdot B) = (\alpha \cdot A) \cdot B = A \cdot (\alpha \cdot B)$.

Тут $\alpha, \beta \in P$ *і* 1 *– одиничний елемент поля P.*

Відповідні операції допустимі. Так, наприклад, властивістю II_1 (асоціативність множення) передбачається: матриця A має розмірність $m \times n$, матриця B – розмірність $n \times p$, матриця C – розмірність $p \times q$.

Позначимо через M_n множину всіх квадратних матриць порядку n (тобто розмірності $n \times n$) з елементами з поля P.

Для них операції додавання, множення і множення на скаляр допустимі без обмежень. Ця множина називається *алгеброю матриць*.

При $n > 1$ в алгебрі M_n матриць не виконується закон комутативності множення. Наприклад, якщо в алгебрі M_2 взяти

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \text{ і } B = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \text{ тоді } A \cdot B = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix} \text{ і } B \cdot A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

Приклад 1. Нехай $A = (\alpha_{ik}) \in M_n$. Матриця

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

називається *приєднаною* для A, якщо її елементи є алгебричними доповненнями елементів визначника

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

Обчислимо добуток

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

За правилом множення матриць маємо

$$\gamma_{11} = \alpha_{11}A_{11} + \alpha_{12}A_{12} + \ldots + \alpha_{1n}A_{1n} = \Delta,$$

$$\gamma_{12} = \alpha_{11}A_{21} + \alpha_{12}A_{22} + \ldots + \alpha_{1n}A_{2n} = 0, \ldots$$

Для довільного γ_{ik} :

$$\gamma_{ik} = \begin{cases} \Delta, \text{ якщо } & i = k; \\ 0, \text{ якщо } & i \neq k. \end{cases}$$

Отже,

$$A \cdot \tilde{A} = \begin{pmatrix} \Delta & 0 & \dots & 0 \\ 0 & \Delta & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \Delta \end{pmatrix} = \Delta \cdot E,$$

де E – одинична матриця розмірності $n \times n$:

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

Приклад 2. Знайти добуток матриць

$$A = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} \text{ і } B = \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix}.$$

$$A \cdot B = \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & \beta \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

З наведеного прикладу напрошується висновок: добуток ненульових матриць може бути нульовою матрицею.

Вправи

1. Довести сформульовану вище теорему про операції з матрицями, звернувши особливу увагу на доведення властивості II_1 .

2. Нехай матриця B одержується з матриці $A = (\alpha_{ik})$ розмірності $m \times n$ за правилом: перший рядок матриці A записуємо першим стовпцем матриці B, другий рядок матриці A записуємо другим стовпцем матриці B і т. д. Одержана матриця називається *транспонованою* для A і позначається A^T . Очевидно, що $\left(A^T\right)^T = A$. Довести рівність $(A \cdot C)^T = C^T \cdot A^T$.

3. Довести, що коли $r = r(A)$ – ранг матриці A, $r_1 = r(B)$ – ранг матриці B, $r_2 = r(C)$ – ранг матриці C і $A = B \cdot C$, тоді

$$r \le \min(r_1, r_2).$$

4. Знайти загальний вигляд квадратної матриці X (або розв'язати матричне рівняння $AX - XA = 0$), яка переставна з матрицею

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

Множення визначників

Теорема 2. *Якщо A і B – квадратні матриці одного порядку, тоді добуток їх визначників дорівнює визначнику добутку $A \cdot B$ цих матриць, тобто* $\det(A \cdot B) = \det A \cdot \det B$.

Доведення. Нехай

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix},$$

$$C = A \cdot B = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix},$$

де $\gamma_{ik} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \ldots + \alpha_{in}\beta_{nk}$. Подамо елементи першого стовпця матриці C у вигляді $\gamma_{i1} = \alpha_{i1}\beta_{11} + \alpha_{i2}\beta_{21} + \ldots + \alpha_{in}\beta_{n1}$.

Тоді

$$\det C = \begin{vmatrix} \alpha_{11}\beta_{11} + \alpha_{12}\beta_{21} + \ldots + \alpha_{1n}\beta_{n1} & \gamma_{12} & \cdots & \gamma_{1n} \\ \alpha_{21}\beta_{11} + \alpha_{22}\beta_{21} + \ldots + \alpha_{2n}\beta_{n1} & \gamma_{22} & \cdots & \gamma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1}\beta_{11} + \alpha_{n2}\beta_{21} + \ldots + \alpha_{nn}\beta_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{vmatrix} =$$

$$
= \sum_{i=1}^{n} \beta_{i1} \begin{vmatrix} \alpha_{1i} & \gamma_{12} & \cdots & \gamma_{12} \\ \alpha_{2i} & \gamma_{22} & \cdots & \gamma_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{ni} & \gamma_{n2} & \cdots & \gamma_{nn} \end{vmatrix}.
$$

У кожному визначнику під знаком суми елементи другого стовпця подамо у вигляді суми $\gamma_{i2} = a_i \cdot b^{(2)} = \sum_{j=1}^{n} \alpha_{ij}\beta_{j2}$ та розкладемо одержаний визначник на суму n визначників і т. д.

Урешті-решт дістанемо

$$
\det C = \sum_{i,j,\dots,w=1}^{n} \beta_{i1}\beta_{j2}\dots\beta_{wn} \cdot \begin{vmatrix} \alpha_{1i} & \alpha_{1j} & \dots & \alpha_{1w} \\ \alpha_{2i} & \alpha_{2j} & \dots & \alpha_{2w} \\ \dots & \dots & \dots & \dots \\ \alpha_{ni} & \alpha_{nj} & \dots & \alpha_{nw} \end{vmatrix},
$$

де $i, j, \dots, w$ незалежно набувають значень від одиниці до n (усього n^n доданків). Але у цій сумі присутні доданки, які дорівнюють нулю: якщо два яких-небудь індекси в послідовності $i, j, \dots, w$ рівні, то відповідний визначник під знаком суми має однакові стовпці, а тому він обертається на нуль. Звільнившись від таких доданків, дістанемо

$$
\det C = \sum_{(i,j,\dots,w)} \beta_{i1}\beta_{j2}\dots\beta_{wn} \cdot \begin{vmatrix} \alpha_{1i} & \alpha_{1j} & \dots & \alpha_{1w} \\ \alpha_{2i} & \alpha_{2j} & \dots & \alpha_{2w} \\ \dots & \dots & \dots & \dots \\ \alpha_{ni} & \alpha_{nj} & \dots & \alpha_{nw} \end{vmatrix},
$$

де підсумування проводиться за всіма можливими перестановками $(i, j, \dots, w)$ номерів 1, 2, ..., n. Для деякої фіксованої перестановки $(i, j, \dots, w)$ розглянемо визначник

$$
\Delta = \begin{vmatrix} \alpha_{1i} & \alpha_{1j} & \dots & \alpha_{1w} \\ \alpha_{2i} & \alpha_{2j} & \dots & \alpha_{2w} \\ \dots & \dots & \dots & \dots \\ \alpha_{ni} & \alpha_{nj} & \dots & \alpha_{nw} \end{vmatrix}.
$$

Він одержується із $\det A$ перестановкою стовпців:

$$a^{(1)}, a^{(2)}, ..., a^{(n)} \to a^{(i)}, a^{(j)}, ..., a^{(w)}.$$

Якщо для здійснення цього перетворення потрібно буде s транспозицій, то $\Delta = (-1)^s \det A$. Число інверсій $t = [i, j, ..., w]$ має однакову парність з числом s, а тому $(-1)^s = (-1)^t$. Тоді

$$\det C = \sum_{(i,j,...,w)} (-1)^t \beta_{i1}\beta_{j2}...\beta_{wn} \cdot \det A =$$

$$= \det A \cdot \left(\sum_{(i,j,...,w)} (-1)^t \beta_{i1}\beta_{j2}...\beta_{wn} \right) = \det A \cdot \det B,$$

що й потрібно було довести.

Цією теоремою користуються при обчисленні визначників деяких спеціальних типів.

Приклад 3. Обчислити визначник типу

$$\Delta = \begin{vmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{vmatrix}.$$

Такі визначники носять назву *циклічних*.

Помножимо його на визначник Вандермонда.

$$D = \begin{vmatrix} 1 & 1 & \dots & 1 \\ \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_n \\ \dots & \dots & \dots & \dots \\ \varepsilon_1^{n-1} & \varepsilon_2^{n-2} & \dots & \varepsilon_n^{n-1} \end{vmatrix} = \prod_{k>j} \left(\varepsilon_k - \varepsilon_j\right),$$

де $\varepsilon_k = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}$ – корені рівняння $x^n - 1 = 0$, $k = \overline{1,n}$.

Позначимо $f(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1}$. Користуючись правилом множення відповідних матриць, дістанемо

$$\Delta\cdot D=\begin{pmatrix} f(\varepsilon_1) & f(\varepsilon_2) & \dots & f(\varepsilon_n)\\ \varepsilon_1 f(\varepsilon_1) & \varepsilon_2 f(\varepsilon_2) & \dots & \varepsilon_n f(\varepsilon_n)\\ \dots & \dots & \dots & \dots\\ \varepsilon_1^{n-1} f(\varepsilon_1) & \varepsilon_2^{n-1} f(\varepsilon_2) & \dots & \varepsilon_n^{n-1} f(\varepsilon_n)\end{pmatrix}=D\cdot\prod_{k=1}^{n} f(\varepsilon_k).$$

Але $D\neq 0$. Тому $\Delta=f(\varepsilon_1)\cdot f(\varepsilon_2)\cdots f(\varepsilon_n)$.

Приклад 4. Обчислити визначник циклічного типу

$$\Delta=\begin{pmatrix} 1 & 2^2 & 3^2 & \dots & n^2\\ n^2 & 1 & 2^2 & \dots & (n-1)^2\\ \dots & \dots & \dots & \dots & \dots\\ 2^2 & 3^2 & 4^2 & \dots & 1\end{pmatrix}.$$

Для нього $f(x)=1+2^2\cdot x+\ldots+n^2\cdot x^{n-1}$. Якщо $g(x)=$ $=x+x^2+\ldots+x^n=\dfrac{x^{n+1}-x}{x-1}$, то $g'(x)=1+2x+3x^2+\ldots+nx^{n-1}$.

Нехай $h(x)=g'(x)x=x+2x^2+\ldots+nx^n$. Тоді

$$f(x)=[g'(x)\cdot x]' \Rightarrow f(x)=$$

$$=\frac{x^n\left[n^2x^3-\left(3n^2+2n-1\right)x^2+\left(3n^2+4n\right)x-(n+1)^2\right]-x^2+1}{(x-1)^4}.$$

Для коренів ε_k, $k=\overline{1,n-1}$, відмінних від 1, маємо

$$f(\varepsilon_k)=\frac{n^2(\varepsilon_k-1)^3-2n(\varepsilon_k-1)^2}{(\varepsilon_k-1)^4}=-\frac{n(n+2)-n^2\varepsilon_k}{(1-\varepsilon_k)^2}.$$

Для $\varepsilon_n=1$ знаходимо відому суму

$$f(\varepsilon_n)=1+2^2+3^2+\ldots+n^2=\frac{n(n+1)(2n+1)}{6}.$$

Таким чином,

$$\Delta = (-1)^{n-1} \cdot \frac{n(n+1)(2n+1)}{6} \cdot \frac{\prod_{k=1}^{n-1}\left\{n(n+2) - n^2\varepsilon_k\right\}}{\left\{\prod_{k=1}^{n-1}(1-\varepsilon_k)\right\}^2}.$$

Оскільки

$$\prod_{k=1}^{n-1}(x-\varepsilon_k) = \frac{x^n - 1}{x-1} = 1 + x + \ldots + x^{n-1},$$

то $\prod_{k=1}^{n-1}(1-\varepsilon_k) = n.$ Аналогічно знаходимо, що

$$\prod_{k=1}^{n-1}\left\{n(n+2) - n^2\varepsilon_k\right\} = n^{2n-2} \cdot \prod_{k=1}^{n-1}\left\{\frac{n+2}{n} - \varepsilon_k\right\} =$$

$$= \frac{1}{2} n^{n-1}\left[(n+2)^n - n^n\right].$$

Остаточно,

$$\Delta = (-1)^{n-1} \cdot \frac{(n+1)(2n+1)}{12} \cdot n^{n-2}\left[(n+2)^n - n^n\right].$$

Вправи

1. Піднесенням до квадрата обчислити визначник

$$\Delta = \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix}.$$

2. Використовуючи циклічність, довести, що

$$\Delta = \begin{vmatrix} x & y & 0 & \ldots & 0 \\ 0 & x & y & \ldots & 0 \\ 0 & 0 & x & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ y & 0 & 0 & \ldots & x \end{vmatrix} = x^n + (-1)^{n-1} \cdot y^n.$$

3. Обчислити

$$\Delta = \begin{vmatrix} 1 & 2 & 3 & \dots & n \\ n & 1 & 2 & \dots & n-1 \\ n-1 & n & 1 & \dots & n-2 \\ \dots & \dots & \dots & \dots & \dots \\ 2 & 3 & 4 & \dots & 1 \end{vmatrix}.$$

4. Методом математичної індукції за натуральним числом n довести рівність

$$1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

§ 5.2. Оборотні лінійні перетворення

Лінійне перетворення

$$\left.\begin{array}{llll} x_1 = \alpha_{11}y_1 & +\alpha_{12}y_2 & +\dots+ & \alpha_{1n}y_n, \\ x_2 = \alpha_{21}y_1 & +\alpha_{22}y_2 & +\dots+ & \alpha_{2n}y_n, \\ \dots & \dots & \dots & \dots \\ x_n = \alpha_{n1}y_1 & +\alpha_{n2}y_2 & +\dots+ & \alpha_{nn}y_n. \end{array}\right\} \qquad (1)$$

з квадратною матрицею $A = (\alpha_{ik})$ порядку n можна записати іншим способом, використовуючи правило матричного множення. Для цього змінні $x_1, x_2, \dots, x_n$ і $y_1, y_2, \dots, y_n$ подамо стовпцями

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Тоді система рівностей (1) набуде вигляду

$$X = A \cdot Y. \qquad (1')$$

Означення. *Лінійне перетворення* (1′) *називається оборотним, якщо знайдеться таке лінійне перетворення*

$$Y = V \cdot X \tag{2}$$

з квадратною матрицею V, що подає змінні $y_1, y_2, \dots, y_n$ *через змінні* $x_1, x_2, \dots, x_n$. *Воно називається оберненим по відношенню до* (1′).

Теорема 1. *Якщо визначник* $\Delta = \det A$ *матриці A лінійного перетворення* (1) *відмінний від нуля, тоді воно оборотне.*

Доведення. Подамо змінні $y_1, y_2, \dots, y_n$ через змінні $x_1, x_2, \dots, x_n$ за допомогою формул Крамера. Дістанемо

$$y_1 = \frac{1}{\Delta}\begin{vmatrix} x_1 & \alpha_{12} & \dots & \alpha_{1n} \\ x_2 & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ x_n & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = \frac{1}{\Delta}\left(x_1 A_{11} + x_2 A_{21} + \dots + x_n A_{n1}\right) =$$

$$= \frac{A_{11}}{\Delta} x_1 + \frac{A_{21}}{\Delta} x_2 + \dots + \frac{A_{n1}}{\Delta} x_n.$$

Аналогічно подамо y_2 через $x_1, x_2, \dots, x_n$ і т. д. Оберненим перетворенням буде система рівностей

$$y_k = \frac{A_{1k}}{\Delta} x_1 + \frac{A_{2k}}{\Delta} x_2 + \dots + \frac{A_{nk}}{\Delta} x_n, \quad k = \overline{1, n}, \tag{2′}$$

з матрицею $V = \frac{1}{\Delta}\tilde{A}$, де $\Delta = \det A$ і $\tilde{A}$ – приєднана матриця для A.

Означення. *Квадратна матриця* $A = (\alpha_{ik})$ *називається оборотною, якщо для неї знайдеться така матриця* $B = (\beta_{ik})$, *для якої*

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

– одинична матриця. Матриця B називається оберненою для A.

Теорема 2. *Квадратна матриця A оборотна тоді і лише тоді, коли її визначник відмінний від нуля. Для оборотної матриці обернена матриця єдина.*

Доведення. *Необхідність.* Нехай матриця A – оборотна. Для неї існує така матриця B, що $A \cdot B = B \cdot A = E$. За теоремою про множення визначників $\det(A \cdot B) = \det A \cdot \det B = 1$. Звідси випливає, що $\Delta = \det A \neq 0$.

Достатність. Нехай $\Delta = \det A \neq 0$. Позначимо $V = \frac{1}{\Delta} \cdot \tilde{A}$, де $\tilde{A}$ – приєднана матриця. Тоді $A \cdot V = \frac{1}{\Delta}(A \cdot \tilde{A}) = \frac{1}{\Delta} \cdot \Delta \cdot E = E$. Отже, $V \cdot A = E$, тобто V – обернена матриця.

Доведемо єдиність. Нехай B – також обернена матриця для A:

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

Тоді $(B \cdot A) \cdot V = E \cdot V = V$, $(B \cdot A) \cdot V = B \cdot (A \cdot V) = B \cdot E = B$. Звідси $B = V$.

Матриця, обернена до A, позначається A^{-1}. Отже, $V = A^{-1}$.

Методи обчислення оберненої матриці

Метод I. Якщо A – оборотна матриця порядку n, то

$$A^{-1} = \frac{1}{\Delta} \cdot \tilde{A},$$

де $\Delta = \det A \neq 0$; $\tilde{A}$ – приєднана матриця.

Застосування цієї формули вимагає обчислення $n^2 + 1$ визначників порядків $n - 1$ і n. Якщо, наприклад, $n > 4$, застосування цього методу вимагає великих обчислень. Навпаки, при $2 \leq n \leq 4$ цією формулою можна користуватись.

Вправи

1. Розв'язати матричне рівняння $AXB = C$, де

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

2. Обчислити

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}^n$$

для всіх цілих показників n .

Метод II (виключення невідомих). Для оборотної матриці $A = (\alpha_{ik})$ складемо систему лінійних рівнянь

$$\begin{cases} \alpha_{11}x_1 & +\alpha_{12}x_2 & +\ldots+ & \alpha_{1n}x_n & = y_1, \\ \alpha_{21}x_1 & +\alpha_{22}x_2 & +\ldots+ & \alpha_{2n}x_n & = y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{n1}x_1 & +\alpha_{n2}x_2 & +\ldots+ & \alpha_{nn}x_n & = y_n \end{cases}$$

з розширеною матрицею

$$\overline{A} = \left(\begin{array}{cccc|c} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & y_1 \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & y_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} & y_n \end{array}\right).$$

Вважаючи змінні $y_1, y_2, \ldots, y_n$ параметрами, застосуємо метод виключення невідомих $x_1, x_2, \ldots, x_n$. На підставі умови $\det A \neq 0$ змінні $x_1, x_2, \ldots, x_n$ подаються через $y_1, y_2, \ldots, y_n$. Дістанемо рівність

$$X = B \cdot Y$$

для стовпців X і Y з невідомими $x_1, x_2, \ldots, x_n$ і $y_1, y_2, \ldots, y_n$. Тоді $B = A^{-1}$.

Приклад 1. Знайти обернену матрицю для

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

Складемо систему рівнянь

$$\begin{cases} 0\cdot x_1 & +x_2 & +x_3 & +x_4 & = y_1, \\ x_1 & +0\cdot x_2 & +x_3 & +x_4 & = y_2, \\ x_1 & +x_2 & +0\cdot x_3 & +x_4 & = y_3, \\ x_1 & +x_2 & +x_3 & +0\cdot x_4 & = y_4 \end{cases}$$

з розширеною матрицею

$$\overline{A} = \left(\begin{array}{cccc|c} 0 & 1 & 1 & 1 & y_1 \\ 1 & 0 & 1 & 1 & y_2 \\ 1 & 1 & 0 & 1 & y_3 \\ 1 & 1 & 1 & 0 & y_4 \end{array}\right).$$

Зробимо такі послідовні елементарні перетворення.

1. Поміняємо місцями перший і останній рядки:

$$\overline{A}_1 = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & y_4 \\ 1 & 0 & 1 & 1 & y_2 \\ 1 & 1 & 0 & 1 & y_3 \\ 0 & 1 & 1 & 1 & y_1 \end{array}\right).$$

2. Віднімемо з другого і третього перший рядок:

$$\overline{A}_2 = \left(\begin{array}{cccc|l} 1 & 1 & 1 & 0 & y_4 \\ 0 & -1 & 0 & 1 & y_2 - y_4 \\ 0 & 0 & -1 & 1 & y_3 - y_4 \\ 0 & 1 & 1 & 1 & y_1 \end{array}\right).$$

3. Додамо до четвертого другий рядок:

$$\overline{A}_3 = \left(\begin{array}{cccc|l} 1 & 1 & 1 & 0 & y_4 \\ 0 & -1 & 0 & 1 & y_2 - y_4 \\ 0 & 0 & -1 & 1 & y_3 - y_4 \\ 0 & 0 & 1 & 2 & y_1 + y_2 - y_4 \end{array}\right).$$

4. Додамо до четвертого третій рядок:

$$\overline{A}_4 = \left(\begin{array}{cccc|l} 1 & 1 & 1 & 0 & y_4 \\ 0 & -1 & 0 & 1 & y_2 - y_4 \\ 0 & 0 & -1 & 1 & y_3 - y_4 \\ 0 & 0 & 0 & 3 & y_1 + y_2 + y_3 - 2y_4 \end{array}\right).$$

Рівносильна система матиме вигляд

$$\begin{cases} x_1 + x_2 + x_3 = y_4, \\ -x_2 + x_4 = y_2 - y_4, \\ -x_3 + x_4 = y_3 - y_4, \\ 3x_4 = y_1 + y_2 + y_3 - 2y_4. \end{cases}$$

Розв'язавши цю порівняно просту систему, знайдемо

$$\begin{aligned} x_1 &= -\frac{2}{3}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 + \frac{1}{3}y_4, \\ x_2 &= \frac{1}{3}y_1 - \frac{2}{3}y_2 + \frac{1}{3}y_3 + \frac{1}{3}y_4, \\ x_3 &= \frac{1}{3}y_1 + \frac{1}{3}y_2 - \frac{2}{3}y_3 + \frac{1}{3}y_4, \\ x_4 &= \frac{1}{3}y_1 + \frac{1}{3}y_2 + \frac{1}{3}y_3 - \frac{2}{3}y_4. \end{aligned}$$

Тому

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

Вправа

Знайти обернену матрицю для матриці n-го порядку:

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

Метод III (Гаусса – Жордана). Нехай $A = (\alpha_{ik})$ – квадратна оборотна матриця n-го порядку і $B = (\beta_{ik})$ – прямокутна матриця розмірності $n \times p$. Наведемо метод обчислення добутку $A^{-1} \cdot B$. Для цього складемо систему рівнянь з невідомими $x_1, x_2, \dots, x_n$ і параметрами $y_1, y_2, \dots, y_p$:

$$\begin{cases} \alpha_{11}x_1+\alpha_{12}x_2+\ \ldots\ +\alpha_{1n}x_n=\beta_{11}y_1+\beta_{12}y_2+\ \ldots\ +\beta_{1p}y_p, \\ \alpha_{21}x_1+\alpha_{22}x_2+\ \ldots\ +\alpha_{2n}x_n=\beta_{21}y_1+\beta_{22}y_2+\ \ldots\ +\beta_{2p}y_p, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{n1}x_1+\alpha_{n2}x_2+\ \ldots\ +\alpha_{nn}x_n=\beta_{n1}y_1+\beta_{n2}y_2+\ \ldots\ +\beta_{np}y_p. \end{cases} \quad (1')$$

Якщо

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},\quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix},$$

тоді систему $(1')$ можна подати у вигляді рівності $A\cdot X=B\cdot Y$. Оскільки матриця A – оборотна, то

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

Для знаходження $A^{-1}\cdot B$ застосуємо метод виключення невідомих. Більш зручна процедура обчислень наведена нижче.

Запишемо розширену матрицю системи $(1')$ у вигляді

$$\overline{A}=\left(\begin{array}{cccc|cccc} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} & \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} & \beta_{21} & \beta_{22} & \cdots & \beta_{2p} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} & \beta_{n1} & \beta_{n2} & \cdots & \beta_{np} \end{array}\right).$$

Будь-яке перетворення системи еквівалентне необхідним перетворенням рядків матриці $\overline{A}$, і навпаки. Оскільки за умовою матриця A – оборотна, то за допомогою серії елементарних перетворень вона зводиться до одиничної матриці E. Застосуємо до $\overline{A}$ декілька елементарних перетворень таких, щоб підматриця A перетворилася в матрицю E. В результаті $\overline{A}$ зводиться до вигляду

$$\overline{C}=(E|C)=\left(\begin{array}{cccc|cccc} 1 & 0 & \ldots & 0 & \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\ 0 & 1 & \ldots & 0 & \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{np} \end{array}\right).$$

Тоді система (1′) рівносильна системі

$$\left.\begin{array}{l} x_1 = \gamma_{11}y_1 + \gamma_{12}y_2 + \ldots + \gamma_{1p}y_p, \\ x_2 = \gamma_{21}y_1 + \gamma_{22}y_2 + \ldots + \gamma_{2p}y_p, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = \gamma_{n1}y_1 + \gamma_{n2}y_2 + \ldots + \gamma_{np}y_p, \end{array}\right\} \qquad (2')$$

тобто

$$X = C \cdot Y .$$

Одержано рівність $A^{-1} \cdot B = C$, де

$$C = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{np} \end{pmatrix}.$$

Зокрема, якщо $B = E$, то $C = A^{-1}$. Отже, для знаходження оберненої матриці для A потрібно використати таке правило. Розширену матрицю

$$\overline{A} = \left(A|E\right) = \left(\begin{array}{cccc|cccc} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} & 1 & 0 & \ldots & 0 \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} & 0 & 0 & \ldots & 1 \end{array}\right)$$

за допомогою елементарних перетворень рядків зведемо до вигляду

$$\overline{C} = \left(E|C\right) = \left(\begin{array}{cccc|cccc} 1 & 0 & \ldots & 0 & \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ 0 & 1 & \ldots & 0 & \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{array}\right).$$

Тоді $C = A^{-1}$.

Приклад 2. Обчислити обернену матрицю для

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

Розширена матриця

$$\overline{A} = \left(\begin{array}{cccc|cccc} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Проробимо такі елементарні перетворення з рядками.

1. Поміняємо місцями перший і четвертий рядки:

$$\overline{A}_1 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right).$$

2. Віднімемо з другого і третього рядків перший:

$$\overline{A}_2 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right).$$

3. Додамо до четвертого другий рядок:

$$\overline{A}_3 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 2 & 1 & 1 & 0 & -1 \end{array}\right).$$

4. Додамо до четвертого третій рядок:

$$\overline{A}_4 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 3 & 1 & 1 & 1 & -2 \end{array}\right).$$

5. Помножимо останній рядок на $\frac{1}{3}$:

$$\overline{A}_5 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{array}\right).$$

6. Віднімемо від другого і третього рядків четвертий:

$$\overline{A}_6 = \left(\begin{array}{cccc|cccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & -1 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{array}\right).$$

7. Додамо до першого другий і третій рядки:

$$\overline{A}_7 = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & -1 & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & -1 & 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{array}\right).$$

8. Помножимо другий і третій рядки на -1:

$$\overline{A}_8 = \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 & 0 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} \end{array}\right).$$

Остаточно дістанемо обернену матрицю

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

Вправи

1. Довести, що коли $A = B \cdot C$, де C – квадратна і невироджена матриця ($\det C \neq 0$), тоді ранги матриць A і B збігаються:

$$r(A) = r(B).$$

2. Обчислити обернену матрицю для

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

§ 5.3. Квадратичні форми

Означення. *Нехай $A = (\alpha_{ik})$ – квадратна матриця порядку n з елементами α_{ik} із поля P. Функція F від $2n$ змінних $x_1, x_2, \dots, x_n$, $y_1, y_2, \dots, y_n$, яка визначається формулою*

$$F = F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = \sum_{i,k=1}^{n} \alpha_{ik} x_i y_k,$$

називається білінійною формою, заданою матрицею A.

Означення. *Квадратна матриця $A = (\alpha_{ik})$ називається симетричною, якщо $A^T = A$, тобто вона збігається з транспонованою. Це означає, що для всіх її елементів виконується рівність $\alpha_{ik} = \alpha_{ki}$.*

Симетричній матриці A відповідає *квадратична форма*, тобто функція f від змінних $x_1, x_2, \ldots, x_n$, що визначається формулою

$$f = f(x_1, x_2, \ldots, x_n) = \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k .$$

Враховуючи, що $\alpha_{ik} = \alpha_{ki}$ для коефіцієнтів цієї функції, помічаємо, що доданки $\alpha_{ik} x_i x_k$ та $\alpha_{ki} x_k x_i$ для цієї форми збігаються.

Тому форму f можна подати у вигляді

$$f = \sum_{i=1}^{n} \alpha_{ii} x_i^2 + 2 \cdot \sum_{i<k} \alpha_{ik} x_i x_k =$$

$$= \alpha_{11} x_1^2 + \alpha_{22} x_2^2 + \ \ldots\ + \alpha_{nn} x_n^2 + 2\alpha_{12} x_1 x_2 + 2\alpha_{13} x_1 x_3 + \ \ldots \qquad (1)$$

Навпаки, кожній однорідній функції від змінних $x_1, x_2, \ldots, x_n$ степеня 2 можна поставити у відповідність симетричну матрицю, яка і визначить її як квадратичну форму.

Наприклад, якщо задано однорідну функцію

$$f = x_1^2 + x_2^2 - 2x_3^2 + 6x_1 x_2 + 12 x_1 x_3 - 20 x_2 x_3 ,$$

то у відповідність цій формі ставиться симетрична матриця

$$A = \begin{pmatrix} 1 & 3 & 6 \\ 3 & 1 & -10 \\ 6 & -10 & -2 \end{pmatrix}.$$

Подамо форму (1) іншим способом, використовуючи правило матричного множення. Для цього запишемо змінні $x_1, x_2, \ldots, x_n$ у вигляді стовпця

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Тоді

$$X^T \cdot A \cdot X = (x_1, x_2, \ \dots \ , x_n) \cdot \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= (x_1, x_2, \ \dots \ , x_n) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \ y_i = \sum_{k=1}^{n} \alpha_{ik} \cdot x_k \ .$$

Звідси

$$X^T \cdot A \cdot X = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} x_i \left(\sum_{k=1}^{n} \alpha_{ik} x_k \right) =$$

$$= \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k = f(x_1, x_2, \ \dots \ , x_n)$$

і тому

$$f(x_1, x_2, \ \dots \ , x_n) = f(X) = X^T \cdot A \cdot X \ . \qquad (2)$$

Означення. *Полярою квадратичної форми*

$$f = \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k$$

називається білінійна форма

$$\hat{f} = \hat{f}(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = \sum_{i,k=1}^{n} x_i \cdot \alpha_{ik} \cdot y_k$$

змінних $x_1, x_2, \ \dots \ , x_n, y_1, y_2, \ \dots \ , y_n$.

Скорочено подамо її у вигляді

$$\hat{f} = \hat{f}(X, Y) = X^T \cdot A \cdot Y, \qquad (3)$$

де X, Y – стовпці змінних $x_1, x_2, \ \dots \ , x_n$ і $y_1, y_2, \ \dots \ , y_n$.

Теорема 1. *Поляра* $\hat{f}$ *квадратичної форми* f *є симетрична білінійна форма*

$$\hat{f}(X, Y) = \hat{f}(Y, X) \, .$$

Доведення.

$$\hat{f}(X,Y)=\sum_{i,k=1}^{n} x_i\alpha_{ik}y_k=\sum_{i,k=1}^{n} x_i\alpha_{ki}y_k=\sum_{i,k=1}^{n} y_k\alpha_{ki}x_i=$$
$$=Y^T\cdot A\cdot X=\hat{f}(Y,X).$$

У квадратичній формі $f=X^T\cdot A\cdot X$ подамо змінні $x_1,x_2,\ \dots,x_n$ через нові змінні $y_1,y_2,\ \dots,y_n$ за допомогою оборотного лінійного перетворення

$$\begin{cases} x_1= & \gamma_{11}y_1+\gamma_{12}y_2 & +\ \dots & +\gamma_{1n}y_n, \\ x_2= & \gamma_{21}y_1+\gamma_{22}y_2 & +\ \dots & +\gamma_{2n}y_n, \\ \dots\dots\dots\dots & & & \\ x_n= & \gamma_{n1}y_1+\gamma_{n2}y_2 & +\ \dots & +\gamma_{nn}y_n. \end{cases}$$

Матриця $C=(\gamma_{ik})$ його оборотна ($\det C\neq 0$). Це можна подати шляхом компактного запису

$$X=C\cdot Y,\quad Y=(y_1,y_2,\ \dots,y_n)^T.$$

Тоді форма f у нових змінних $y_1,y_2,\ \dots,y_n$ набуде вигляду

$$f=(C\cdot Y)^T\cdot A\cdot(C\cdot Y)=Y^T\cdot(C^T\cdot A\cdot C)\cdot Y,$$

тобто

$$f=Y^T\cdot B\cdot Y\ ,$$

де
$$B=C^T\cdot A\cdot C\ .$$

При цьому матриця B також симетрична:

$$B^T=(C^T\cdot A\cdot C)^T=C^T\cdot A\cdot(C^T)^T=C^T\cdot A\cdot C=B\ .$$

Означення. *Матриця B називається конгруентною матриці A ($B\equiv A$), якщо існує така оборотна матриця C, для якої $B=C^T\cdot A\cdot C$ (вважатимемо, що A, B, C – квадратні матриці однієї і тієї самої розмірності).*

Відношення конгруентності матриць – відношення еквівалентності:

рефлексивність $A \equiv A$,

симетричність $B \equiv A \Rightarrow A \equiv B$,

транзитивність $B \equiv A$, $D \equiv B \Rightarrow D \equiv A$.

Пропонуємо самостійно довести ці властивості.

Нами доведена така теорема.

Теорема 2. *При оборотному лінійному перетворенні змінних матриця квадратичної форми перетворюється в конгруентну матрицю.*

Означення. *Квадратична форма*

$$f = \beta_1 y_1^2 + \beta_2 y_2^2 + \ \dots \ + \beta_n y_n^2$$

називається канонічною, якщо її матриця діагональна.

Теорема 3 (Лагранжа). *Квадратична форма*

$$f = f(X) = \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k$$

з коефіцієнтами з будь-якого поля P за допомогою оборотного лінійного перетворення змінних може бути зведена до канонічного вигляду.

Доведення проведемо індукцією за числом n змінних. Якщо $n=1$, тоді $f = \alpha_{11} x_1^2$ – канонічна форма. Припустимо, що $n > 1$ і теорема має місце для форм з числом змінних $< n$.

Розглянемо форму f з числом змінних n. Подамо її у вигляді

$$f = \alpha_{11} x_1^2 + \alpha_{22} x_2^2 + \ \dots \ + \alpha_{nn} x_n^2 + 2\sum_{i<k} \alpha_{ik} x_i x_k \, .$$

Окремо розглянемо дві можливості.

1. Серед коефіцієнтів $\alpha_{11}, \alpha_{22}, \ \dots \ , \alpha_{nn}$ існує хоч би один, відмінний від нуля.

2. Усі вони дорівнюють нулю:

$$\alpha_{11} = \alpha_{22} = \dots = \alpha_{nn} = 0 \, .$$

У першому випадку можна вважати, що $\alpha_{11} \neq 0$. Виділимо у формі f усі доданки, які містять x_1:

$$f=\alpha_{11}x_1^2+2\alpha_{12}x_1x_2+2\alpha_{13}x_1x_3+\ldots+2\alpha_{1n}x_1x_n+g\,,$$

де g – форма зі змінними $x_2, x_3, \ldots, x_n$.

Перетворимо суму перших n доданків

$$\alpha_{11}x_1^2+2\alpha_{12}x_1x_2+\ldots+2\alpha_{1n}x_1x_n=$$

$$=\alpha_{11}\left(x_1^2+2\frac{\alpha_{12}}{\alpha_{11}}x_1x_2+\ldots+2\frac{\alpha_{1n}}{\alpha_{11}}x_1x_n\right)=$$

$$=\alpha_{11}\left(x_1+\frac{\alpha_{12}}{\alpha_{11}}x_2+\ldots+\frac{\alpha_{1n}}{\alpha_{11}}x_n\right)^2-\frac{\alpha_{12}^2}{\alpha_{11}}x_2^2-\ldots-\frac{\alpha_{1n}^2}{\alpha_{11}}x_n^2-$$

$$-2\frac{\alpha_{12}\cdot\alpha_{13}}{\alpha_{11}}x_2x_3-\ldots=\alpha_{11}\left(x_1+\frac{\alpha_{12}}{\alpha_{11}}x_2+\ldots+\frac{\alpha_{1n}}{\alpha_{11}}x_n\right)^2+g_1,$$

де g_1 – форма від змінних $x_2, x_3, \ldots, x_n$. Отже, форму f можна подати у вигляді

$$f=\alpha_{11}\left(x_1+\frac{\alpha_{12}}{\alpha_{11}}x_2+\ \ldots\ +\frac{\alpha_{1n}}{\alpha_{11}}x_n\right)^2+\bar{f}(x_2, x_3, \ldots, x_n)\,,$$

де $\bar{f}(x_2, x_3, \ldots, x_n)$ – форма, яка не містить x_1.

Введемо нові змінні

$$x_1+\frac{\alpha_{12}}{\alpha_{11}}x_2+\ \ldots\ +\frac{\alpha_{1n}}{\alpha_{11}}x_n=y_1\,,\ x_2=y_2,\ \ldots,\ x_n=y_n\,.$$

Тоді

$$f=\alpha_{11}y_1^2+\bar{f}(y_2, y_3, \ldots, y_n)\,.$$

Подамо старі змінні через нові

$$x_1=y_1-\frac{\alpha_{12}}{\alpha_{11}}y_2-\ldots-\frac{\alpha_{1n}}{\alpha_{11}}y_n\,,$$

$$x_2=y_2,\ \ldots,\ x_n=y_n\,.$$

Матриця цього лінійного перетворення

$$C_1 = \begin{pmatrix} 1 & -\dfrac{\alpha_{12}}{\alpha_{11}} & \dots & -\dfrac{\alpha_{1n}}{\alpha_{11}} \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

оборотна, оскільки $\det C_1 = 1$.

До форми $\bar{f} = \bar{f}(y_2, y_3, \dots, y_n)$ застосуємо припущення індукції.

Існує оборотне лінійне перетворення

$$\begin{aligned} y_2 &= \beta_{22}z_2 + \dots + \beta_{2n}z_n, \\ y_3 &= \beta_{32}z_2 + \dots + \beta_{3n}z_n, \\ &\dots\dots\dots\dots \\ y_n &= \beta_{n2}z_2 + \dots + \beta_{nn}z_n \end{aligned}$$

таке, що форма $\bar{f}$ у змінних $z_2, z_3, \dots, z_n$ має канонічний вигляд

$$\bar{f} = \beta_2 z_2^2 + \beta_3 z_3^2 + \dots + \beta_n z_n^2 .$$

Якщо покласти, що $y_1 = z_1$, то матрицею переходу від $y_1, y_2, \dots, y_n$ до $z_1, z_2, \dots, z_n$ є

$$C_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \beta_{22} & \dots & \beta_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & \beta_{n2} & \dots & \beta_{nn} \end{pmatrix}$$

і $\det C_2 \neq 0$.

У змінних $z_1, z_2, \dots, z_n$ форма f має канонічний вигляд

$$f = \alpha_{11} z_1^2 + \beta_2 z_2^2 + \beta_3 z_3^2 + \dots + \beta_n z_n^2 .$$

Матрицею переходу від початкових змінних $x_1, x_2, \dots, x_n$ до змінних $z_1, z_2, \dots, z_n$ є матриця $C = C_1 \cdot C_2$ і $\det C = \det C_1 \cdot \det C_2 \neq 0$.

Перейдемо до розгляду другого випадку, коли $\alpha_{11} = \alpha_{22} = \ldots = \alpha_{nn} = 0$ і форма

$$f = 2\alpha_{12}x_1x_2 + 2\alpha_{13}x_1x_3 + \ldots$$

Серед її коефіцієнтів є ненульові. Наприклад, нехай $\alpha_{12} \neq 0$. Введемо нові змінні $x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, $x_3 = y_3, \ldots, x_n = y_n$.

Матриця цього лінійного перетворення

$$C_1 = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 1 & -1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

оборотна ($\det C_1 = -2$). У нових змінних форма має вигляд

$$f = 2\alpha_{12}y_1^2 - 2\alpha_{12}y_2^2 + (2\alpha_{13} + 2\alpha_{23})y_1y_3 + \ldots$$

Існує лінійне перетворення змінних $y_1, y_2, \ldots, y_n$ через змінні $z_1, z_2, \ldots, z_n$ з оборотною матрицею C_2, яке зводить форму f до канонічного вигляду.

Остаточне лінійне перетворення змінних $x_1, x_2, \ldots, x_n$ у змінні $z_1, z_2, \ldots, z_n$ дає оборотну матрицю $C = C_1 \cdot C_2$.

Приклад. Звести квадратичну форму

$$f = x_1^2 + x_2^2 - 2x_3^2 + 6x_1x_2 + 12x_1x_3 - 20x_2x_3$$

до канонічного вигляду.

Матриця цієї квадратичної форми має вигляд

$$A = \begin{pmatrix} 1 & 3 & 6 \\ 3 & 1 & -10 \\ 6 & -10 & -2 \end{pmatrix}.$$

Відокремимо в цій формі квадрат лінійного виразу, який містить усі члени зі змінною x_1.

$$f = x_1^2 + 6x_1x_2 + 12x_1x_3 + x_2^2 - 2x_3^2 - 20x_2x_3 =$$
$$= (x_1 + 3x_2 + 6x_3)^2 - 8x_2^2 - 38x_3^2 - 56x_2x_3.$$

Позначимо $x_1 + 3x_2 + 6x_3 = y_1$, $x_2 = y_2$, $x_3 = y_3$. Тоді

$$x_1 = y_1 - 3y_2 - 6y_3,$$
$$x_2 = y_2,$$
$$x_3 = y_3.$$

Матриця цього перетворення

$$C_1 = \begin{pmatrix} 1 & -3 & -6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Квадратична форма у нових змінних матиме вигляд:

$$f = y_1^2 - 8y_2^2 - 56y_2y_3 - 38y_3^2.$$

Виділимо квадрат лінійного виразу, що містить усі члени зі змінною y_2:

$$f = y_1^2 - 8(y_2^2 + 7y_2y_3) - 38y_3^2 = y_1^2 - 8\left(y_2 + \frac{7}{2}y_3\right)^2 + 60y_3^2.$$

Знову-таки введемо нові змінні

$$y_1 = z_1, \quad y_2 = z_2 - \frac{7}{2}z_3, \quad y_3 = z_3.$$

Матриця цього перетворення

$$C_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{7}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

У нових змінних квадратична форма матиме канонічний вигляд

$$f = z_1^2 - 8z_2^2 + 60z_3^2.$$

Матриця перетворення, що подає x_1, x_2, x_3 через z_1, z_2, z_3, дорівнює

$$C = C_1 \cdot C_2 = \begin{pmatrix} 1 & -3 & -6 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{7}{2} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -3 & \frac{9}{2} \\ 0 & 1 & -\frac{7}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

Після очевидних обчислень

$$B = C^T \cdot A \cdot C =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ \frac{9}{2} & -\frac{7}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 6 \\ 3 & 1 & -10 \\ 6 & -10 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 & \frac{9}{2} \\ 0 & 1 & -\frac{7}{2} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & 60 \end{pmatrix}.$$

Це і є матриця канонічної форми f.

§ 5.4. Додатно визначені квадратичні форми

У цьому параграфі вивчаються матриці та квадратичні форми над полем R дійсних чисел.

Якщо у дійсній квадратичній формі

$$f(X) = f(x_1, x_2, ..., x_n) = \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k$$

замість змінних $x_1, x_2, ..., x_n$ покладемо будь-які числові значення $\xi_1, \xi_2, ..., \xi_n \in R$, тоді будемо записувати $x = (\xi_1, \xi_2, ..., \xi_n) \in R^n$ і $f(x) = f(\xi_1, \xi_2, ..., \xi_n)$ – значення квадратичної форми f зі змінними $(x_1, x_2, ..., x_n)^T = X$.

Нехай над змінними $x_1, x_2, ..., x_n$ здійснено оборотне лінійне перетворення

$$x_i = \sum_{k=1}^{n} \gamma_{ik} y_k,\ i = \overline{1,n}, \tag{1}$$

в результаті якого форма набуває вигляду

$$g(Y) = g(y_1, y_2, ..., y_n) = f\left(\sum_{k=1}^{n} \gamma_{1k} y_k, \sum_{k=1}^{n} \gamma_{2k} y_k, ..., \sum_{k=1}^{n} \gamma_{nk} y_k \right).$$

На підставі оборотності матриці $C = (\gamma_{ik})$ для довільних значень $\xi_1, \xi_2, ..., \xi_n$ змінних $x_1, x_2, ..., x_n$ існує єдиний набір $\eta_1, \eta_2, ..., \eta_n$ значень змінних $y_1, y_2, ..., y_n$ такий, що

$$\xi_i = \sum_{k=1}^{n} \gamma_{ik} \eta_k,\ i = \overline{1,n}.$$

При цьому має місце рівність

$$f(\xi_1, \xi_2, ..., \xi_n) = g(\eta_1, \eta_2, ..., \eta_n). \tag{2}$$

Означення. *Дійсна квадратична форма*

$$f(x_1, x_2, ..., x_n) = f(X)$$

називається додатно визначеною, якщо вона набуває додатних значень при будь-яких значеннях змінних, відмінних від нуля одночасно:

$$(\xi_1, \xi_2, ..., \xi_n) \neq \theta \Rightarrow f(\xi_1, \xi_2, ..., \xi_n) > 0.$$

Теорема 1. *Для того щоб дійсна квадратична форма $f(X)$ була додатно визначеною, необхідно і достатньо, щоб її канонічна форма $g(Y) = \beta_1 y_1^2 + \beta_2 y_2^2 + ... + \beta_n y_n^2$ при змінних $y_1, y_2, ..., y_n$ мала коефіцієнти $\beta_1 > 0, \beta_2 > 0, ..., \beta_n > 0$.*

Доведення. *Достатність.* Припустимо, що форму $f(X)$ лінійним перетворенням (1) зведено до вигляду

$$g(Y) = \beta_1 y_1^2 + \beta_2 y_2^2 + ... + \beta_n y_n^2.$$

Нехай $\xi_1, \xi_2, ..., \xi_n$ – деяка послідовність значень змінних $x_1, x_2, ..., x_n$, а $\eta_1, \eta_2, ..., \eta_n$ – відповідна послідовність значень нових змінних $y_1, y_2, ..., y_n$, тоді матимемо рівність (2):

$$f(\xi_1,\xi_2,...,\xi_n)=\beta_1\eta_1^2+\beta_2\eta_2^2+...+\beta_n\eta_n^2.$$

Припустимо, що $\beta_1>0$, $\beta_2>0$, ..., $\beta_n>0$. Якщо $(\xi_1,\xi_2,...,\xi_n)\neq\theta$, тоді $(\eta_1,\eta_2,...,\eta_n)\neq\theta$ і $\beta_1\eta_1^2+...+\beta_n\eta_n^2>0$, а тому $f(\xi_1,\xi_2,...,\xi_n)>0$.

Форма $f(X)$ – додатно визначена.

Необхідність. Припустимо, що форма $f(X)$ – додатно визначена. Покладемо $\eta_1=1, \eta_2=0, ..., \eta_n=0$. Тоді відповідними значеннями початкових змінних є числа $\xi_1=\gamma_{11}$, $\xi_2=\gamma_{21}$, ..., $\xi_n=\gamma_{n1}$ (із формули (1)). Вони не всі дорівнюють нулю, оскільки стовпці матриці $C=(\gamma_{ik})$ – ненульові. Тому $f(\xi_1,\xi_2,...,\xi_n)>0$. Але

$$f(\xi_1,\xi_2,...,\xi_n)=\beta_1\cdot 1+\beta_2\cdot 0+...+\beta_n\cdot 0.$$

Звідси $\beta_1>0$, а разом з тим і $\beta_2>0,...,\beta_n>0$.

Наслідок. *Квадратична форма f від n змінних $x_1,x_2,...,x_n$ є додатно визначеною тоді і лише тоді, коли вона за допомогою оборотного лінійного перетворення старих змінних може бути зведена до суми квадратів*

$$f=z_1^2+z_2^2+...+z_n^2$$

нових змінних $z_1,z_2,...,z_n$.

Доведення. *Необхідність.* Нехай форма $f(X)$ додатно визначена. За допомогою лінійного перетворення змінних зведемо її до канонічного вигляду

$$g(Y)=\beta_1y_1^2+\beta_2y_2^2+...+\beta_ny_n^2.$$

За теоремою 1 коефіцієнти $\beta_1,\beta_2,...,\beta_n$ додатні. Здійснимо ще одне лінійне перетворення:

$$y_1=\frac{1}{\sqrt{\beta_1}}z_1,\ y_2=\frac{1}{\sqrt{\beta_2}}z_2,\ ...,\ y_n=\frac{1}{\sqrt{\beta_n}}z_n.$$

У нових змінних

$$f = z_1^2 + z_2^2 + \ldots + z_n^2.$$

Достатність очевидна.

Теорема 2 (критерій Сильвестра). *Для того щоб квадратична форма* $f = \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k$ *була додатно визначеною, необхідно і достатньо, щоб головні мінори*

$$\Delta_1 = \alpha_{11},\ \Delta_2 = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}, \ldots,$$

$$\Delta_p = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pp} \end{vmatrix}, \ldots, \Delta_n = \det A$$

її матриці $A = (\alpha_{ik})$ *були числами додатними.*

Доведення. *Необхідність.* Нехай $f(x) = \sum_{i,k=1}^{n} \alpha_{ik} x_i x_k$ – додатно визначена квадратична форма. Зведемо її до канонічного вигляду

$$g(Y) = \beta_1 y_1^2 + \beta_2 y_2^2 + \ldots + \beta_n y_n^2$$

за допомогою лінійного перетворення $X = C \cdot Y$ з оборотною матрицею C. При цьому $\beta_1 > 0$, $\beta_2 > 0, \ldots, \beta_n > 0$.

Матриця

$$B = \begin{pmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \beta_n \end{pmatrix}$$

конгруентна матриці A: $B = C^T \cdot A \cdot C$. Тому $\det B = \beta_1 \cdot \beta_2 \cdots \beta_n = \det\left(C^T \cdot A \cdot C\right) = \det C^T \cdot \det A \cdot \det C = (\det C)^2 \cdot \det A.$

Оскільки $\beta_1 > 0,\ \beta_2 > 0,\ ...,\ \beta_n > 0$ і $(\det C)^2 > 0$, то $\det A = \Delta_n > 0$.

Розглянемо форму

$$f_p = f(x_1,...,x_p,0,...,0) = \sum_{i,k=1}^{p} \alpha_{ik} x_i x_k$$

з матрицею

$$A_p = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p} \\ \dots & \dots & \dots & \dots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pp} \end{pmatrix}.$$

Вона також додатно визначена і тому

$$\Delta_p = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2p} \\ \dots & \dots & \dots & \dots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pp} \end{vmatrix} > 0$$

для будь-яких $p < n$.

Достатність. Нехай головні мінори матриці A додатні: $\Delta_1 > 0, \Delta_2 > 0, ..., \Delta_n > 0$. Додатну визначеність форми доводимо індукцією за числом n змінних $x_1, x_2, ..., x_n$.

Якщо $n = 1$, тоді $f = \alpha_{11} x_1^2$ і $\Delta_1 = \alpha_{11} > 0$.

Форма f – додатно визначена.

Припустимо, що твердження теореми має місце для форм з $n-1$ змінними.

Якщо позначити

$$f_{n-1} = f(x_1, x_2, ..., x_{n-1}, 0) = \sum_{i,k=1}^{n-1} \alpha_{ik} x_i x_k ,$$

тоді

$$f = f_{n-1} + \alpha_{nn} x_n^2 + 2x_n \cdot \left(\sum_{i=1}^{n-1} \alpha_{in} x_i \right).$$

За припущенням індукції форма f_{n-1} додатно визначена. Тому в деяких змінних $y_1, y_2, ..., y_{n-1}$ вона подається у вигляді суми квадратів $f_{n-1} = y_1^2 + y_2^2 + ... + y_{n-1}^2$ і тоді

$$f = \sum_{i=1}^{n-1} y_i^2 + \alpha_{nn} x_n^2 + 2\sum_{i=1}^{n-1} \beta_{in} y_i x_n$$

з деякими новими коефіцієнтами β_{in}. Для однозначності покладемо, що $x_n = y_n$. Тоді

$$f = \sum_{i=1}^{n-1} y_i^2 + \alpha_{nn} y_n^2 + 2\sum_{i=1}^{n-1} \beta_{in} y_i y_n .$$

Перетворимо цей вираз, використовуючи рівності

$$y_i^2 + 2\beta_{in} y_i y_n = (y_i + \beta_{in} y_n)^2 - \beta_{in}^2 y_n^2 .$$

Дістанемо

$$f = \sum_{i=1}^{n-1} (y_i + \beta_{in} y_n)^2 + \beta y_n^2 ,$$

де $\beta = \alpha_{nn} - \sum_{i=1}^{n-1} \beta_{in}^2$.

Введемо нові змінні $y_i + \beta_{in} y_n = z_i$, $i = \overline{1, n-1}$, $y_n = z_n$.

Подамо y_i через z_j:

$$y_1 = z_1 - \beta_{1n} z_n,\ y_2 = z_2 - \beta_{22} z_n,\ ...,\ y_{n-1} = z_{n-1} - \beta_{n-1n} z_n,\ y_n = z_n.$$

У змінних $z_1, z_2, ..., z_n$ форма f набуде вигляду

$$g(z_1, z_2, ..., z_n) = z_1^2 + z_2^2 + ... + z_{n-1}^2 + \beta z_n^2$$

з діагональною матрицею

$$B = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \beta \end{pmatrix}.$$

Якщо C – матриця перетворення, що подає $x_1, x_2, ..., x_n$ через $z_1, z_2, ..., z_n$, тоді $B = C^T \cdot A \cdot C$ і тому $\beta = \det B = \det C^T \times$ $\times \det A \cdot \det C = (\det C)^2 \cdot \det A$. Але $(\det C)^2 > 0$ і $\det A = \Delta_n > 0$ (за умовою). Звідси $\beta > 0$.

Форма f – додатно визначена.

Приклад. З'ясувати, за яких значень параметра λ форма $f = x_1^2 + 2x_2^2 + 2x_3^2 - 2\lambda x_1 x_2 + 4x_1 x_3 - 6x_2 x_3$ додатно визначена.

Матриця форми

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

Головні мінори

$$\Delta_1 = 1, \quad \Delta_2 = \begin{vmatrix} 1 & -\lambda \\ -\lambda & 2 \end{vmatrix} = 2 - \lambda^2,$$

$$\Delta_3 = \begin{vmatrix} 1 & -\lambda & 2 \\ -\lambda & 2 & -3 \\ 2 & -3 & 2 \end{vmatrix} = -2\lambda^2 + 12\lambda - 13.$$

Для додатної визначеності повинні мати місце нерівності

$$\Delta_2 = 2 - \lambda^2 > 0, \quad \Delta_3 = -2\lambda^2 + 12\lambda - 13 > 0.$$

З першої нерівності знаходимо, що $\lambda \in \left(-\sqrt{2}; \sqrt{2}\right)$, а з другої – $\lambda \in \left(3 - \sqrt{\frac{5}{2}}; 3 + \sqrt{\frac{5}{2}}\right)$. З'ясуємо, чи перетинаються інтервали $\left(-\sqrt{2}; \sqrt{2}\right)$ і $\left(3 - \sqrt{\frac{5}{2}}; 3 + \sqrt{\frac{5}{2}}\right)$. Для цього порівняємо числа $\sqrt{2}$ і $3 - \sqrt{\frac{5}{2}}$. Маємо $\sqrt{2} > 3 - \sqrt{\frac{5}{2}} \Leftrightarrow \sqrt{2} + \sqrt{\frac{5}{2}} > 3 \Leftrightarrow$

$$\Leftrightarrow\left(\sqrt{2}+\sqrt{\frac{5}{2}}\right)^2>3^2\Leftrightarrow\frac{9}{2}+2\sqrt{5}>9\Leftrightarrow 2\sqrt{5}>\frac{9}{2}\Leftrightarrow 4\sqrt{5}>9\Leftrightarrow$$

$\Leftrightarrow 80>81$. Але $80<81$. Тому $\sqrt{2}<3-\sqrt{\frac{5}{2}}$. Інтервали не перетинаються. Отже, не існує таких значень λ, для яких форма f додатно визначена.

Вправи

1. Знайти всі значення λ, для яких квадратична форма

$$f=2x_1^2+x_2^2+3x_3^2+2\lambda x_1x_2+2x_1x_3$$

додатно визначена.

2. Звести форму $f=x_1^2+x_2^2+\ldots+x_n^2+\sum\limits_{i<k}x_ix_k$ до канонічного вигляду.

3. *Дискримінантом* квадратичної форми f називається визначник $D_f=\det A$ її матриці A. Довести, що коли до додатно визначеної форми $f=f(x_1,x_2,\ldots,x_n)$ приєднати квадрат ненульової лінійної форми $\gamma_1x_1+\gamma_2x_2+\ldots+\gamma_nx_n$, тоді дискримінант форми збільшиться.

4. Нехай $f=f(x_1,x_2,\ldots,x_n)$ – додатно визначена квадратична форма і $\overline{f}=f(0,x_2,\ldots,x_n)$. Тоді для дискримінантів цих форм виконується нерівність $D_f\le\alpha_{11}\cdot D_{\overline{f}}$.

Розділ 6. ЛІНІЙНІ ПЕРЕТВОРЕННЯ ВЕКТОРНИХ ПРОСТОРІВ

§ 6.1. Базиси векторних просторів

Базисом векторного простору P^n над полем P є будь-який набір із n його лінійно незалежних векторів.

Нехай $Б = < b_1, b_2, ..., b_n >$ – який-небудь базис простору P^n.

Якщо $x \in P^n$ і $x = \xi_1 b_1 + \xi_2 b_2 + ... + \xi_n b_n$, тоді координати $\xi_1, \xi_2, ..., \xi_n$ записуватимемо у вигляді стовпця

$$x_Б = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

або $x_Б = (\xi_1, \xi_2, ..., \xi_n)^T$.

Оберемо інший базис $Г = < c_1, c_2, ..., c_n >$ того ж простору. Всі вектори цього нового базису лінійно задаються через вектори старого так:

$$\begin{aligned} c_1 &= \delta_{11} b_1 + \delta_{21} b_2 + \ ... \ + \delta_{n1} b_n, \\ c_2 &= \delta_{12} b_1 + \delta_{22} b_2 + \ ... \ + \delta_{n2} b_n, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ c_n &= \delta_{1n} b_1 + \delta_{2n} b_2 + \ ... \ + \delta_{nn} b_n. \end{aligned}$$

Матрицею переходу від базису Б до базису Г називається матриця виду

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

стовпці якої складені з координат векторів $c_1, c_2, ..., c_n$ в базисі Б.

Теорема. *Матриця переходу D від базису до базису невироджена, тобто* $\det D \neq 0$.

Доведення проведемо методом від супротивного. Припустимо, що $\det D = 0$. При цьому припущенні складемо систему лінійних однорідних рівнянь

$$\begin{cases} \delta_{11}x_1 + \delta_{12}x_2 + \ldots + \delta_{1n}x_n = 0, \\ \delta_{21}x_1 + \delta_{22}x_2 + \ldots + \delta_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \delta_{n1}x_1 + \delta_{n2}x_2 + \ldots + \delta_{nn}x_n = 0. \end{cases}$$

Оскільки $\det D = 0$, то система має ненульовий розв'язок $\lambda_1, \lambda_2, \ldots, \lambda_n$. Якщо замість невідомих підставити цей розв'язок, дістанемо систему рівностей

$$\begin{cases} \delta_{11}\lambda_1 + \delta_{12}\lambda_2 + \ldots + \delta_{1n}\lambda_n = 0, \\ \delta_{21}\lambda_1 + \delta_{22}\lambda_2 + \ldots + \delta_{2n}\lambda_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \delta_{n1}\lambda_1 + \delta_{n2}\lambda_2 + \ldots + \delta_{nn}\lambda_n = 0. \end{cases}$$

У векторній формі ця система записується у вигляді

$$\lambda_1 c_1 + \lambda_2 c_2 + \ldots + \lambda_n c_n = 0,$$

що означає лінійну залежність векторів. Дістали суперечність. Тому $\det D \neq 0$.

Тепер знайдемо співвідношення між координатами будь-якого вектора $x \in P^n$ в різних базисах Б і Г.

Нехай $x_Б = (\xi_1, \xi_2, \ldots, \xi_n)^T$ і $x_Г = (\eta_1, \eta_2, \ldots, \eta_n)^T$, тобто матимемо рівності:

а) $x = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n$;

б) $x = \eta_1 c_1 + \eta_2 c_2 + \ldots + \eta_n c_n$.

Виразимо вектори c_j через базис Б і підставимо в рівність б). Дістанемо

$$x=\eta_1\left(\sum_{i=1}^{n}\delta_{i1}b_i\right)+\eta_2\left(\sum_{i=1}^{n}\delta_{i2}b_i\right)+\ldots+\eta_n\left(\sum_{i=1}^{n}\delta_{in}b_i\right)=$$

$$=\left(\sum_{k=1}^{n}\delta_{1k}\eta_k\right)b_1+\left(\sum_{k=1}^{n}\delta_{2k}\eta_k\right)b_2+\ldots+\left(\sum_{k=1}^{n}\delta_{nk}\eta_k\right)b_n.$$

Порівнюючи цю рівність з рівністю а), остаточно знаходимо співвідношення

$$\begin{cases}\xi_1=\delta_{11}\eta_1+\delta_{12}\eta_2+\ldots+\delta_{1n}\eta_n,\\ \xi_2=\delta_{21}\eta_1+\delta_{22}\eta_2+\ldots+\delta_{2n}\eta_n,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \xi_n=\delta_{n1}\eta_1+\delta_{n2}\eta_2+\ldots+\delta_{nn}\eta_n,\end{cases}$$

або у формі матричного рівняння

$$x_Б=D\cdot x_Г.$$

Отже, *стовпець координат вектора x у старому базисі дорівнює добутку матриці переходу на стовпець координат у новому базисі.*

Приклад. Переконатися, що вектори

$$b_1=(2,2,-1),\ \ b_2=(2,-1,2),\ \ b_3=(-1,2,2)$$

становлять базис простору R^3 та знайти координати вектора $x=(1,1,1)$ у цьому базисі.

Розв'язання. На підставі означення встановлюємо, що $Б=<b_1,b_2,b_3>$ – базис простору R^3. Початковим базисом вважатимемо $Е=<e_1,e_2,e_3>$, де $e_1=(1,\ 0,\ 0)$, $e_2=(0,\ 1,\ 0)$, $e_3=(0,\ 0,\ 1)$ – одиничні координатні вектори (вектори b_1,b_2,b_3 задані у цьому базисі).

Маємо $x_Е=(1,\ 1,\ 1)^T$. Матриця переходу від базису Е до базису Б має вигляд

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

Якщо $x_Б = (\eta_1, \eta_2, \eta_3)^T$, тоді

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}. \qquad (1)$$

Система рівнянь

$$\begin{cases} 2\eta_1 + 2\eta_2 - \eta_3 = 1, \\ 2\eta_1 - \eta_2 + 2\eta_3 = 1, \\ -\eta_1 + 2\eta_2 + 2\eta_3 = 1 \end{cases}$$

має розв'язок $\eta_1 = \eta_2 = \eta_3 = \frac{1}{3}$.

Координати η_1, η_2, η_3 можна знайти іншим способом. Дійсно, подамо (1) у вигляді матричного рівняння

$$x_E = D \cdot x_Б.$$

З цього рівняння маємо

$$x_Б = D^{-1} \cdot x_E.$$

Враховуючи, що

$$D^{-1} = \frac{1}{9} \cdot \begin{pmatrix} 2 & 2 & -1 \\ 2 & -1 & 2 \\ -1 & 2 & 2 \end{pmatrix},$$

знаходимо

$$x_Б = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} = D^{-1} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}.$$

Метод обчислення матриці переходу

Нехай $Б = \langle b_1, b_2, \ldots, b_n \rangle$ – базис простору P^n. Побудуємо матрицю B, стовпцями якої будуть вектори цього базису: 1-й стовпець складається з координат 1-го вектора

$b_1 = (\beta_{11}, \beta_{21}, ..., \beta_{n1})$, 2-й – з координат 2-го вектора $b_2 = (\beta_{12}, \beta_{22}, ..., \beta_{n2})$ і т. д.

Нехай $\Gamma =< c_1, c_2, ..., c_n >$ – інший (новий) базис. Для нього побудуємо матрицю C, стовпцями якої будуть вектори цього базису: 1-й стовпець складено з координат 1-го вектора $c_1 = (\gamma_{11}, \gamma_{21}, ..., \gamma_{n1})$, 2-й – із координат 2-го вектора $c_2 = (\gamma_{12}, \gamma_{22}, ..., \gamma_{n2})$ і т. д.

Запишемо матрицю переходу від базису Б до базису Г :

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

За її означенням

$$c_i = \delta_{1i} b_1 + \delta_{2i} b_2 + \dots + \delta_{ni} b_n .$$

Звідси i-й стовпець матриці C подається у вигляді лінійної комбінації стовпців матриці B з коефіцієнтами $\delta_{1i}, \delta_{2i}, ..., \delta_{ni}$, що заповнюють i-й стовпець матриці D. Тоді мають місце рівності

$$\begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \dots & \dots & \dots & \dots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix} \cdot \begin{pmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1n} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2n} \\ \dots & \dots & \dots & \dots \\ \delta_{n1} & \delta_{n2} & \cdots & \delta_{nn} \end{pmatrix} =$$

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

тобто $B \cdot D = C$. Звідси знаходимо

$$D = B^{-1} \cdot C .$$

§ 6.2. Лінійні перетворення і матриці

Лінійним перетворенням векторного простору P^n над полем P у векторний простір P^m називається таке однозначне відображення

$$\mathcal{A}: P^n \to P^m,$$

яке має властивості:

1. $\mathcal{A}(x_1 + x_2) = \mathcal{A}(x_1) + \mathcal{A}(x_2)$ *для будь-яких* $x_1, x_2 \in P^n$ (*адитивність*).

2. $\mathcal{A}(\lambda x) = \lambda\, \mathcal{A}(x)$ *для всіх* $x \in P^n$, $\lambda \in P$ *(однорідність)*.

Ці дві властивості за сукупністю еквівалентні одній властивості:

3. $\mathcal{A}(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \mathcal{A}(x_1) + \lambda_2 \mathcal{A}(x_2)$ (*лінійність*), $x_1, x_2 \in P^n$ і $\lambda_1, \lambda_2 \in P$.

Зокрема, *лінійне відображення*

$$\mathcal{A}: P^n \to P^n$$

називається лінійним оператором простору P^n.

Із означення лінійного перетворення $\mathcal{A}$ випливає, що коли $x_1, x_2, \ldots, x_m \in P^n$ і $\lambda_1, \lambda_2, \ldots, \lambda_m \in P$, тоді

$$\mathcal{A}(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_m x_m) =$$
$$= \lambda_1 \mathcal{A}(x_1) + \lambda_2 \mathcal{A}(x_2) + \ldots + \lambda_m \mathcal{A}(x_m).$$

Наведена нижче теорема дає загальний спосіб побудови лінійних перетворень.

Теорема 1. *Нехай* Б$=<b_1, b_2, \ldots, b_n>$ *– базис векторного простору* P^n *і* $f_1, f_2, \ldots, f_n$ *– будь-який набір, що містить у собі n векторів простору* P^m*. Існує єдине лінійне перетворення* $\mathcal{A}$ *простору* P^n *у простір* P^m*, для якого*

$$\mathcal{A}(b_i) = f_i, \quad i = \overline{1, n}. \qquad (1)$$

Доведення. Визначимо відображення $\mathcal{A}: P^n \to P^m$ за правилом: якщо $x = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n$ – вектор із P^n, тоді

Його образом у просторі P^m покладемо вектор $y = \mathcal{A}(x) = \xi_1 f_1 + +\xi_2 f_2 + \ldots + \xi_n f_n$.

Спочатку доведемо лінійність побудованого відображення. Нехай

$$x_1 = \sum_{i=1}^{n} \xi_i b_i, \; x_2 = \sum_{i=1}^{n} \eta_i b_i, \; \lambda_1, \lambda_2 \in P^n.$$

За означенням $\mathcal{A}(x_1) = \sum_{i=1}^{n} \xi_i f_i, \; \mathcal{A}(x_2) = \sum_{i=1}^{n} \eta_i f_i$.

Оскільки

$$\lambda_1 x_1 + \lambda_2 x_2 = \lambda_1 \left(\sum_{i=1}^{n} \xi_i b_i \right) + \lambda_2 \left(\sum_{i=1}^{n} \eta_i b_i \right) = \sum_{i=1}^{n} (\lambda_1 \xi_i + \lambda_2 \eta_i) b_i,$$

тоді

$$\mathcal{A}(\lambda_1 x_1 + \lambda_2 x_2) = \sum_{i=1}^{n} (\lambda_1 \xi_i + \lambda_2 \eta_i) f_i =$$

$$= \lambda_1 \left(\sum_{i=1}^{n} \xi_i f_i \right) + \lambda_2 \left(\sum_{i=1}^{n} \eta_i f_i \right) = \lambda_1 \mathcal{A}(x_1) + \lambda_2 \mathcal{A}(x_2).$$

Отже, $\mathcal{A}(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \mathcal{A}(x_1) + \lambda_2 \mathcal{A}(x_2)$.

Відображення $\mathcal{A}$ має необхідну властивість: воно відображає вектор b_i у вектор f_i для будь-яких $i = \overline{1, n}$.

Доведемо єдиність перетворення $\mathcal{A}$. Нехай $\mathcal{B}$ – лінійне перетворення простору P^n в P^m, яке перетворює вектор b_i у вектор f_i, $i = \overline{1, n}$. Якщо $x = \sum_{i=1}^{n} \xi_i b_i$, тоді $\mathcal{B}(x) = \mathcal{B}\left(\sum_{i=1}^{n} \xi_i b_i \right) =$ $= \sum_{i=1}^{n} \xi_i \, \mathcal{B}(b_i) = \sum_{i=1}^{n} \xi_i f_i$. Маємо рівність $\mathcal{B}(x) = \mathcal{A}(x)$ для будь-яких $x \in P^n$. Це означає, що перетворення $\mathcal{B}$ і $\mathcal{A}$ збігаються.

Позначимо через L множину всіх лінійних відображень векторного простору P^n у простір P^m, а через M – множину всіх матриць розмірності $m \times n$ з елементами з поля P.

У просторі P^n оберемо який-небудь базис $Б = < b_1, b_2, ..., b_n >$ і в просторі P^m оберемо базис $Г = < c_1, c_2, ..., c_m >$.

Теорема 2 (про задання лінійних перетворень матрицями). *Існує взаємно однозначне відображення (яке залежить від базисів* Б *і* Г):

$$\varphi : L \to M$$

множини L лінійних перетворень на множину M матриць таке, що коли $\mathcal{A} \in L$, $A = \varphi(\mathcal{A})$, $x \in P^n$ *і* $y = \mathcal{A}(x)$, *тоді*

$$y_Г = A \cdot x_Б, \tag{2}$$

де $x_Б$ *– стовпець координат вектора x в базисі* Б, *а* $y_Г$ *– стовпець координат вектора y* в *базисі* Г.

Доведення. Побудова відображення $\varphi : L \to M$. Для будь-якого перетворення $\mathcal{A} \in L$ поставимо у відповідність матрицю $A = (\alpha_{ik})$ розмірності $m \times n$ за таким правилом (правило φ):

Нехай

$$\begin{aligned}
\mathcal{A}(b_1) &= \alpha_{11}c_1 + \alpha_{21}c_2 + ... + \alpha_{m1}c_m, \\
\mathcal{A}(b_2) &= \alpha_{12}c_1 + \alpha_{22}c_2 + ... + \alpha_{m2}c_m, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\mathcal{A}(b_n) &= \alpha_{1n}c_1 + \alpha_{2n}c_2 + ... + \alpha_{mn}c_m.
\end{aligned}$$

Перший стовпець матриці A заповнюється коефіцієнтами $\alpha_{11}, \alpha_{21}, ..., \alpha_{m1}$, тобто координатами вектора $\mathcal{A}(b_1)$ у базисі Г; другий стовпець – коефіцієнтами $\alpha_{12}, \alpha_{22}, ..., \alpha_{m2}$, тобто координатами вектора $\mathcal{A}(b_2)$ в базисі Г і т. д. Тоді

$$\varphi(\mathcal{A}) = A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}.$$

Побудоване відображення φ являє собою відображення на всю множину M матриць. Нехай A – матриця розмірності $m \times n$ з елементами з поля P. Розглянемо систему векторів із простору P^m:

$$f_k = \alpha_{1k}c_1 + \alpha_{2k}c_2 + \ldots + \alpha_{mk}c_m, \ k = \overline{1, n},$$

де $\alpha_{1k}, \alpha_{2k}, \ldots, \alpha_{mk}$ – елементи k-го стовпця матриці A.

За теоремою 1 існує, причому єдиним способом, лінійне перетворення $\mathcal{A}: P^n \to P^m$ таке, що виконуються умови

$$\mathcal{A}(b_k) = f_k, \ k = \overline{1, n}. \tag{3}$$

За правилом φ матрицею перетворення $\mathcal{A}$ буде дана матриця A, тобто $\varphi(\mathcal{A}) = A$.

Відображення φ – взаємно однозначне.

Нехай $\mathcal{A}, \mathcal{B} \in L$ і $\varphi(\mathcal{A}) = \varphi(\mathcal{B}) = A$.

Тоді $\mathcal{A}(b_k) = \sum_{i=1}^{m} \alpha_{ik}c_i = \mathcal{B}(b_k)$ і для довільного вектора $x = \sum_{i=1}^{n} \xi_i b_i$ маємо

$$\mathcal{A}(x) = \mathcal{A}\left(\sum_{i=1}^{n} \xi_i b_i\right) = \sum_{i=1}^{n} \xi_i \, \mathcal{A}(b_i) = \sum_{i=1}^{n} \xi_i \, \mathcal{B}(b_i) = \mathcal{B}(x).$$

Це означає, що $\mathcal{A} = \mathcal{B}$.

Нехай $x \in P^n$, $y = \mathcal{A}(x) \in P^m$ і

$$x_{\text{Б}} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad y_{\Gamma} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{pmatrix},$$

тобто $x = \xi_1 b_1 + \xi_2 b_2 + \ldots + \xi_n b_n$, $y = \eta_1 c_1 + \eta_2 c_2 + \ldots + \eta_m c_m$. Тоді

$$y = \mathcal{A}(x) = \mathcal{A}\left(\sum_{i=1}^{n} \xi_i b_i\right) = \sum_{i=1}^{n} \xi_i \mathcal{A}(b_i) =$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{m} \xi_i \alpha_{ki} c_k = \sum_{k=1}^{m} \left(\sum_{i=1}^{n} \alpha_{ki} \xi_i\right) c_k.$$

Звідси $\eta_k = \sum_{i=1}^{n} \alpha_{ki} \xi_i$, $k = \overline{1, m}$, тобто $y_{\Gamma} = A \cdot x_{\text{Б}}$.

Ядро лінійного перетворення

Ядром лінійного перетворення $\mathcal{A}: P^n \to P^m$ *називається множина всіх тих векторів* $x \in P^n$, *для яких* $\mathcal{A}(x) = \theta$. Ядро позначається

$$\operatorname{Ker}\mathcal{A} = \{x \in P^n : \mathcal{A}(x) = \theta\}.$$

Теорема 3. *Ядро лінійного перетворення – лінійний підпростір векторного простору* P^n. *Розмірність його дорівнює* $n - r$, *де* r *– ранг матриці цього перетворення.*

Доведення. Нехай $x_1, x_2 \in \operatorname{Ker}\mathcal{A}$; $\lambda_1, \lambda_2 \in P$. Тоді маємо

$$\mathcal{A}(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \mathcal{A}(x_1) + \lambda_2 \mathcal{A}(x_2) = \theta,$$

тобто $\lambda_1 x_1 + \lambda_2 x_2 \in \operatorname{Ker}\mathcal{A}$. Множина $\operatorname{Ker}\mathcal{A}$ – лінійний підпростір.

Нехай відносно базисів Б і Г просторів P^n і P^m матриця $A = (\alpha_{ik})$ відповідає перетворенню $\mathcal{A}$. Тоді для будь-якого $x \in \operatorname{Ker}\mathcal{A}$ виконується рівність

$$A \cdot x_Б = \theta, \text{ де } x_Б = (\xi_1, \xi_2, ..., \xi_n)^T.$$

Ядро $\operatorname{Ker}\mathcal{A}$ визначається підпростором усіх розв'язків системи однорідних рівнянь

$$\begin{gathered} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + ... + \alpha_{1n}\xi_n = 0, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + ... + \alpha_{2n}\xi_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + ... + \alpha_{mn}\xi_n = 0. \end{gathered}$$

Останній підпростір матиме розмірність $n - r$, де r – ранг матриці $A = (\alpha_{ik})$. Таку ж розмірність матиме підпростір $\operatorname{Ker}\mathcal{A}$.

Образ лінійного перетворення

Образом лінійного перетворення $\mathcal{A}: P^n \to P^m$ *називається множина всіх тих векторів* $y \in P^m$, *для яких існують вектори* $x \in P^n$ *такі, що* $y = \mathcal{A}(x)$. Образ позначається

$$\operatorname{Im}\mathcal{A} = \{y \in P^m : y = \mathcal{A}(x),\ x \in P^n\}.$$

Теорема 4. *Образ лінійного перетворення – лінійний підпростір простору P^m. Розмірність його дорівнює рангу матриці цього перетворення.*

Доведення. Нехай $y_1, y_2 \in \operatorname{Im}\mathcal{A}$; $\lambda_1, \lambda_2 \in P$. Існують такі вектори $x_1, x_2 \in P^n$, що $y_1 = \mathcal{A}(x_1)$ і $y_2 = \mathcal{A}(x_2)$. Тоді

$$\mathcal{A}(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 y_1 + \lambda_2 y_2, \ \lambda_1 y_1 + \lambda_2 y_2 \in \operatorname{Im}\mathcal{A},$$

де $\operatorname{Im}\mathcal{A}$ – лінійний підпростір.

Нехай відносно базисів Б і Г просторів P^n і P^m матриця $\mathcal{A} = (\alpha_{ik})$ відповідає перетворенню $\mathcal{A}$, тобто її коефіцієнти задовольняють співвідношення

$$\begin{aligned} f_1 &= \mathcal{A}(b_1) = \alpha_{11}c_1 + \alpha_{21}c_2 + \ldots + \alpha_{m1}c_m, \\ f_2 &= \mathcal{A}(b_2) = \alpha_{12}c_1 + \alpha_{22}c_2 + \ldots + \alpha_{m2}c_m, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n &= \mathcal{A}(b_n) = \alpha_{1n}c_1 + \alpha_{2n}c_2 + \ldots + \alpha_{mn}c_m. \end{aligned}$$

Підпростір $\operatorname{Im}\mathcal{A}$ збігається з лінійною оболонкою $L(f_1, f_2, \ldots, f_n)$. Разом з тим ранг системи векторів $f_1, f_2, \ldots, f_n$ дорівнює рангу r матриці A. Тому розмірність $\operatorname{Im}\mathcal{A}$ дорівнює r.

§ 6.3. Лінійні оператори

Лінійне перетворення $\mathcal{A}: P^n \to P^n$ векторного простору P^n на себе називається його лінійним оператором.

Теорему про подання лінійних операторів сформулюємо так:

Нехай Б $= \langle b_1, b_2, \ldots, b_n \rangle$ – базис простору P^n. Існує взаємно однозначне відображення (яке залежить від вибору базису Б)

$$\varphi_Б : L_n \to M_n$$

множини L_n усіх лінійних операторів простору P^n на множину M_n усіх квадратних матриць порядку n над полем P таке, що коли $\mathcal{A} \in L_n$ і $y = \mathcal{A}(x)$, $x \in P^n$, тоді

$$y_Б = A \cdot x_Б, \tag{1}$$

де $A = \varphi_Б(\mathscr{A})$; $x_Б$ *– стовпець координат вектора x у базисі* Б, $y_Б$ *– стовпець координат вектора y в тому самому базисі.*

Надалі записуватимемо матрицю A, яка відповідає оператору $\mathscr{A}$ при відображенні $\varphi_Б$, спрощено у вигляді

$$A = \mathscr{A}_Б .$$

Отже, якщо

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}, \quad x_Б = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad y_Б = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix},$$

тоді співвідношення (1) набуде вигляду

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}.$$

При цьому правило побудови матриці A (правило φ) спрощується, оскільки використовується лише один базис Б. Якщо

$$\mathscr{A}(b_k) = \alpha_{1k}b_1 + \alpha_{2k}b_2 + \ldots + \alpha_{nk}b_n, \; k = \overline{1, n}, \tag{2}$$

тоді k-й стовпець матриці A заповнюється координатами $\alpha_{1k}, \alpha_{2k}, \ldots, \alpha_{nk}$ вектора $\mathscr{A}(b_k)$ для всіх k, $k = \overline{1, n}$.

Серед усіх лінійних операторів особливо виділимо два:

1. Оператор $\mathscr{E}$ – тотожне перетворення простору. Він перетворює будь-який вектор x у той же вектор $x : \mathscr{E}(x) = x$.

У будь-якому базисі цьому оператору відповідає одинична матриця E.

2. Оператор $\mathscr{O}$ перетворює будь-який вектор x простору в нульовий вектор θ: $\mathscr{O}(x) = \theta$.

У будь-якому базисі цьому оператору відповідає нульова матриця O.

Приклад 1. Задано деякий базис Б$=<b_1, b_2, ..., b_n>$ і набір векторів $f_1, f_2, ..., f_n$ векторного простору P^n. Припустимо, що координати всіх векторів b_k і f_k подано в базисі Е$=<e_1, e_2, ..., e_n>$, де

$$\begin{aligned} e_1 &= (1, 0, 0, ..., 0), \\ e_2 &= (0, 1, 0, ..., 0), \\ &\dots\dots\dots\dots\dots\dots \\ e_n &= (0, 0, 0, ..., 1) \end{aligned}$$

– стандартний базис. Нехай $\mathcal{A}$ – лінійний оператор такий, що

$$\mathcal{A}(b_k) = f_k, \; k = \overline{1, n}.$$

Знайти матрицю оператора $\mathcal{A}$ в базисі Е.

Розв'язання. Запишемо матрицю

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \dots & \dots & \dots & \dots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix},$$

в якої перший стовпець містить координати вектора $b_1 = (\beta_{11}, \beta_{21}, \ldots, \beta_{n1})$ (в базисі Е), другий – координати вектора $b_2 = (\beta_{12}, \beta_{22}, \ldots, \beta_{n2})$ і т. д.

Аналогічно складемо матрицю

$$F = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2n} \\ \dots & \dots & \dots & \dots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{pmatrix},$$

перший стовпець якої заповнено координатами вектора $f_1 = (\varphi_{11}, \varphi_{21}, \ldots, \varphi_{n1})$, другий – координатами вектора $f_2 = (\varphi_{12}, \varphi_{22}, \ldots, \varphi_{n2})$ і т. д.

Якщо $A = \mathcal{A}_E$, то за теоремою про подання лінійного оператора матимемо систему рівностей

$$(f_1)_E = A \cdot (b_1)_E,$$
$$(f_2)_E = A \cdot (b_2)_E,$$
$$\dots\dots\dots\dots\dots\dots$$
$$(f_n)_E = A \cdot (b_n)_E.$$

Тут $(b_1)_E$, $(b_2)_E$, ..., $(b_n)_E$ – стовпці матриці B; $(f_1)_E$, $(f_2)_E$, ..., $(f_n)_E$ – стовпці матриці F.

У матричній формі наведені вище рівності можна подати у вигляді

$$F = A \cdot B.$$

Звідси $A = F \cdot B^{-1}$.

Приклад 2. Нехай $b_1 = (2, 1, -1)$, $b_2 = (2, -1, 2)$, $b_3 = (3, 0, 1)$ і $f_1 = (1, 2, -1)$, $f_2 = (1, 0, 0)$, $f_3 = (0, 1, 1)$. Обчислити матрицю лінійного оператора $\mathcal{A}$, який перетворює b_i в f_i ($i = \overline{1,3}$) в тому базисі, в якому задано координати векторів.

Розв'язання. Спочатку переконаємося, що вектори b_1, b_2, b_3 – лінійно незалежні. Потім запишемо матриці

$$B = \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}, F = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

Якщо $A = \mathcal{A}_E$, тоді $A = F \cdot B^{-1}$. Обчислюємо

$$B^{-1} = \begin{pmatrix} 1 & -4 & -3 \\ 1 & -5 & -3 \\ -1 & 6 & 4 \end{pmatrix}.$$

Тоді

$$A = \mathcal{A}_E = F \cdot B^{-1} = \begin{pmatrix} 2 & -9 & -6 \\ 1 & -2 & -2 \\ -2 & 10 & 7 \end{pmatrix}.$$

Приклад 3. Нехай $Б = <b_1, b_2, \dots, b_n>$ і $\Gamma = <c_1, c_2, \dots, c_n>$ – базиси векторного простору P^n. Припустимо, що відома матриця A лінійного оператора $\mathcal{A}$ відносно базису Б: $A=\mathcal{A}_{\text{Б}}$. Обчислити матрицю A' того самого оператора $\mathcal{A}$ відносно базису Г, тобто знайти $A'=\mathcal{A}_{\Gamma}$.

Розв'язання. Нехай $y=\mathcal{A}(x)$ і $x_{\text{Б}}=(\xi_1,\xi_2,\ \ldots,\ \xi_n)^T$, $y_{\text{Б}}=(\eta_1,\eta_2,\ \ldots,\ \eta_n)^T$, $x_{\Gamma}=(\xi_1',\xi_2',\ \ldots,\ \xi_n')^T$, $y_{\Gamma}=(\eta_1',$ $\eta_2',\ \ldots,\ \eta_n')^T$. На підставі теореми про подання лінійних операторів маємо співвідношення

$$y_{\text{Б}}=A\cdot x_{\text{Б}},\ y_{\Gamma}=A'\cdot x_{\Gamma}.$$

Позначимо через D матрицю переходу від базису Б до базису Г. Залежність між координатами одного й того самого вектора в різних базисах (див. § 5.4) подається рівністю $x_{\text{Б}}=D\cdot x_{\Gamma}$ для вектора x і рівністю $y_{\text{Б}}=D\cdot y_{\Gamma}$ для вектора y.

Отже,

$$D\cdot y_{\Gamma}=AD\cdot x_{\Gamma}\Rightarrow y_{\Gamma}=$$
$$=D^{-1}\cdot AD\cdot x_{\Gamma}\Rightarrow A'\cdot x_{\Gamma}=D^{-1}\cdot AD\cdot x_{\Gamma}.$$

Якщо в останній рівності покласти $x_{\Gamma}=(1,\ 0,\ \ldots,\ 0)^T$, то дістанемо

$$A'\cdot\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}=D^{-1}\cdot AD\cdot\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}.$$

Це означає, що перші стовпці матриць A' і $D^{-1}\cdot AD$ збігаються. Аналогічно встановлюється збіг інших стовпців матриць A' і $D^{-1}\cdot AD$. Отже,

$$A'=D^{-1}\cdot AD.$$

Матриця A' одержується з матриці A множенням її справа на невироджену матрицю D і зліва – на обернену D^{-1}. Таке перетворення матриці A називається *перетворенням подібності*.

Вправи

1. Показати, що в просторі R^3 оператор

$$\mathcal{A}(x) = [a, x]$$

– лінійний для заданого вектора $a = (\alpha, \beta, \gamma)$ та знайти його матрицю в базисі $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ (базис E).

2. Довести лінійність оператора

$$\mathcal{A}(x) = (x, a)\, a$$

у просторі R^3 при заданому векторі $a = (\alpha, \beta, \gamma)$ та знайти його матрицю в базисі E.

3. Нехай у просторі R^4 задано системи векторів

$$b_1 = (1, 1, 1, 1), \quad b_2 = (1, 1, -1, -1),$$
$$b_3 = (1, -1, 1, -1), \quad b_4 = (1, -1, -1, 1);$$
$$f_1 = (0, 1, -1, 0), \quad f_2 = (0, -1, 1, 0),$$
$$f_3 = (0, 2, -2, 0), \quad f_4 = (0, 0, 0, 1).$$

Знайти матриці лінійного оператора $\mathcal{A}$, що переводять b_k в f_k ($k = \overline{1, 4}$):

а) в тому самому базисі, що й задано координати векторів b_k і f_k;

б) в базисі $c_1 = (1, 0, 0, 0)$, $c_2 = (1, 1, 0, 0)$, $c_3 = (1, 1, 1, 0)$, $c_4 = (1, 1, 1, 1)$.

Ядро і образ лінійного оператора

Ядро $\operatorname{Ker}\mathcal{A}$ лінійного оператора $\mathcal{A}$ за означенням

$$\operatorname{Ker}\mathcal{A} = \{x \in P^n : \mathcal{A}(x) = \theta\}.$$

Розмірність ядра називається дефектом оператора $\mathcal{A}$.

Образ цього оператора

$$\operatorname{Im}\mathcal{A} = \{y \in P^n : y = \mathcal{A}(x), x \in P^n\}.$$

Розмірність образа називається рангом оператора $\mathcal{A}$.

Для обчислення $\operatorname{Ker}\mathcal{A}$ і $\operatorname{Im}\mathcal{A}$ потрібно подати оператор за допомогою матриці у деякому базисі, наприклад, у базисі E $< e_1, e_2, \ldots, e_n >$ одиничних векторів (стандартному базисі).

Якщо $A = \mathcal{A}_\text{E}$, тоді задача зводиться до обчислення ядра $\operatorname{Ker}\mathcal{A}$ і образа $\operatorname{Im}\mathcal{A}$ цієї матриці:

$$\operatorname{Ker}\mathcal{A} = \{x \in P^n : A \cdot x_\text{E} = \theta\},$$

$$\operatorname{Im}\mathcal{A} = \{y \in P^n : y_\text{E} = A \cdot x_\text{E},\, x \in P^n\},$$

де x_E, y_E – стовпці, складені з координат векторів x і y.

Приклад 4. Лінійний оператор $\mathcal{A}$ перетворює вектори $b_1 = (1,-1,\ 1)$, $b_2 = (1,\ 2,-1)$, $b_3 = (2,\ 2,-1)$ у відповідні вектори $f_1 = (0,\ 1,\ 1)$, $f_2 = (0,\ 2,\ 0)$, $f_3 = (0,-1,\ 1)$.

Знайти ядро і образ цього оператора в базисі E.

Розв'язання. На підставі методу розв'язання задачі 1 цього параграфа випишемо матрицю B зі стовпцями b_1, b_2, b_3 і матрицю F зі стовпцями f_1, f_2, f_3:

$$B = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 2 \\ 1 & -1 & -1 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Тоді матриця A лінійного оператора $\mathcal{A}$ в базисі E обчислюється з рівності $A \cdot B = F$, тобто $A = F \cdot B^{-1}$.

Здійснимо необхідні обчислення:

$$B^{-1} = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix};$$

$$A = F \cdot B^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 2 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ -3 & 9 & 13 \\ 1 & -1 & -1 \end{pmatrix}.$$

Обчислимо $\operatorname{Ker}\mathcal{A}$.

Якщо $x = (\xi_1, \xi_2,\ \xi_3) \in \operatorname{Ker}\mathcal{A}$, тоді $x_\text{E} = (\xi_1, \xi_2,\ \xi_3)^T$ і $A \cdot x_\text{E} = \theta$. Для невідомих $\xi_1, \xi_2,\ \xi_3$ дістанемо систему

$$\begin{cases} -3\xi_1 + 9\xi_2 + 13\xi_3 = 0; \\ \xi_1 - \xi_2 - \xi_3 = 0. \end{cases}$$

Її матриця

$$\begin{pmatrix} -3 & 9 & 13 \\ 1 & -1 & -1 \end{pmatrix}$$

за допомогою елементарних перетворень з рядками зводиться до вигляду

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

Еквівалентна система лінійних рівнянь

$$\begin{cases} \xi_1 & -\xi_2 & -\xi_3 & = 0; \\ & 3\xi_2 & +5\xi_3 & = 0. \end{cases}$$

Її загальний розв'язок $X = \left(-\frac{2}{3}\xi_3,\ -\frac{5}{3}\xi_3,\ \xi_3\right)$.

Фундаментальна система розв'язків складається лише з одного розв'язку: $a = (2,\ 5,\ -3)$. Отже, $\operatorname{Ker}\mathcal{A} = L(a)$ і дефект оператора $\mathcal{A}$ дорівнює одиниці.

Обчислимо $\operatorname{Im}\mathcal{A}$, який збігається з лінійною оболонкою $L(a_1, a_2, a_3)$ стовпців $a_1 = (0,\ -3,\ 1)^T$, $a_2 = (0,\ 9,\ -1)^T$, $a_3 = (0,\ 13,\ -1)^T$ матриці A.

Запишемо матрицю

$$A^T = \begin{pmatrix} 0 & -3 & 1 \\ 0 & 9 & -1 \\ 0 & 13 & -1 \end{pmatrix}.$$

За допомогою елементарних перетворень стосовно її рядків зводимо її до вигляду

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

Вектори $b_2 = (0,\ 1,\ 0)$ і $b_3 = (0,\ 0,\ 1)$ становлять базис підпростору $\operatorname{Im}\mathcal{A}$.

Вправи

1. Знайти ядро та образ лінійного оператора $\mathcal{A}(x)=[a,x]$ простору R^3.

2. Навести приклад лінійного оператора $\mathcal{A}$ простору R^3, для якого $\operatorname{Ker}\mathcal{A}\neq\theta$ і $\operatorname{Ker}\mathcal{A}\subseteq\operatorname{Im}\mathcal{A}$.

§ 6.4. Алгебра лінійних операторів

Означення. *Нехай $\mathcal{A},\mathcal{B}\in L_n$. Сумою лінійних операторів $\mathcal{A}+\mathcal{B}$ називається оператор $\mathcal{F}$ такий, що $\mathcal{F}(x)=\mathcal{A}(x)+\mathcal{B}(x)$ для будь-яких $x\in P^n$. Сума $\mathcal{F}=\mathcal{A}+\mathcal{B}$ лінійних операторів $\mathcal{A}$ і $\mathcal{B}$ – лінійний оператор.*

Нехай x_1, $x_2\in P^n$, λ_1, $\lambda_2\in P$. Тоді

$$\mathcal{F}(\lambda_1x_1+\lambda_2x_2)=\mathcal{A}(\lambda_1x_1+\lambda_2x_2)+\mathcal{B}(\lambda_1x_1+\lambda_2x_2)=$$
$$=\lambda_1\mathcal{A}(x_1)+\lambda_2\mathcal{A}(x_2)+\lambda_1\mathcal{B}(x_1)+\lambda_2\mathcal{B}(x_2)=$$
$$=[\lambda_1\mathcal{A}(x_1)+\lambda_1\mathcal{B}(x_1)]+[\lambda_2\mathcal{A}(x_2)+\lambda_2\mathcal{B}(x_2)]=$$
$$=\lambda_1[\mathcal{A}(x_1)+\mathcal{B}(x_1)]+\lambda_2[\mathcal{A}(x_2)+\mathcal{B}(x_2)]=$$
$$=\lambda_1\mathcal{F}(x_1)+\lambda_2\mathcal{F}(x_2).$$

Основні властивості суми операторів:

I_1. $\mathcal{A}+(\mathcal{B}+\mathcal{C})=(\mathcal{A}+\mathcal{B})+\mathcal{C}$,

I_2. $\exists\,\mathcal{O}\,\forall\,\mathcal{A}\,\{\mathcal{A}+\mathcal{O}=\mathcal{O}+\mathcal{A}=\mathcal{A}\}$,

I_3. $\forall\,\mathcal{A}\,\exists\,\mathcal{X}\,\{\mathcal{A}+\mathcal{X}=\mathcal{X}+\mathcal{A}=\mathcal{O}\}$,

I_4. $\mathcal{A}+\mathcal{B}=\mathcal{B}+\mathcal{A}$.

Тут $\mathcal{O}$ – нульовий оператор, $\mathcal{X}=-\mathcal{A}$ – протилежний оператор для $\mathcal{A}$.

Операцію віднімання введемо так:

$$\mathcal{A}-\mathcal{B}=\mathcal{A}+(-\mathcal{B}).$$

Означення. *Оператор $\mathcal{P}$ називається добутком лінійних операторів $\mathcal{A}$ і $\mathcal{B}$, якщо для будь-якого $x\in P^n$ має місце рівність $\mathcal{P}(x)=\mathcal{A}(\mathcal{B}(x))$. Позначення добутку: $\mathcal{P}=\mathcal{A}\cdot\mathcal{B}$.*

Ця операція добутку полягає в послідовній дії спочатку оператором $\mathcal{B}$ на вектор x, а потім оператором $\mathcal{A}$ на вектор $y=\mathcal{B}(x)$: $z=\mathcal{A}(y)$, тоді $z=\mathcal{P}(x)$.

Добуток лінійних операторів – лінійний оператор.

Нехай $x_1, x_2 \in P^n$ і $\lambda_1, \lambda_2 \in P$. Тоді $(\mathcal{AB})(\lambda_1 x_1 + \lambda_2 x_2) =$
$= \mathcal{A}(\mathcal{B}(\lambda_1 x_1 + \lambda_2 x_2)) = \mathcal{A}(\lambda_1 \mathcal{B}(x_1) + \lambda_2 \mathcal{B}(x_2)) = \lambda_1 \mathcal{A}(\mathcal{B}(x_1)) +$
$+ \lambda_2 \mathcal{A}(\mathcal{B}(x_2)) = \lambda_1 (\mathcal{AB})(x_1) + \lambda_2 (\mathcal{AB})(x_2)$.

Основні властивості добутку операторів:

II$_1$. $\mathcal{A}\cdot(\mathcal{B}\cdot\mathcal{C}) = (\mathcal{A}\cdot\mathcal{B})\cdot\mathcal{C}$.

II$_2$. $\mathcal{A}\cdot\mathcal{E} = \mathcal{E}\cdot\mathcal{A} = \mathcal{A}$.

Тут $\mathcal{E}$ – одиничний оператор (тотожне перетворення).

Розподільні закони:

II$_3$. $\mathcal{A}\cdot(\mathcal{B} + \mathcal{C}) = \mathcal{A}\cdot\mathcal{B} + \mathcal{A}\cdot\mathcal{C}$.

II$_4$. $(\mathcal{A} + \mathcal{B})\cdot \mathcal{C} = \mathcal{A}\cdot\mathcal{C} + \mathcal{B}\cdot\mathcal{C}$.

З урахуванням наведених вище властивостей, множина L_n лінійних операторів простору P^n є кільцем.

Означення. *Добутком лінійного оператора $\mathcal{A}$ на скаляр λ називається такий оператор $\mathcal{H}$, для якого $\mathcal{H}(x) = \lambda\cdot\mathcal{A}(x)$, $x \in P^n$. Позначення $\mathcal{H} = \lambda\mathcal{A}$.*

Оператор $\lambda\mathcal{A}$ – лінійний.

Основні властивості добутку оператора на скаляр:

III$_1$. $1\cdot \mathcal{A} = \mathcal{A}$, $0\cdot \mathcal{A} = \mathcal{O}$, де 1, 0 – одиниця і нуль із P.

III$_2$. $(\lambda\mu)\mathcal{A} = \lambda(\mu\mathcal{A})$.

III$_3$. $(\lambda + \mu)\mathcal{A} = \lambda\mathcal{A} + \mu\mathcal{A}$.

III$_4$. $\lambda(\mathcal{A} + \mathcal{B}) = \lambda\mathcal{A} + \lambda\mathcal{B}$.

III$_5$. $\lambda(\mathcal{A}\cdot\mathcal{B}) = (\lambda\mathcal{A})\cdot\mathcal{B} = \mathcal{A}\cdot(\lambda\mathcal{B})$.

Алгебрична система з наведеними вище властивостями називається *лінійною алгеброю* над полем P.

Множина L_n усіх лінійних операторів векторного простору P^n – один з конкретних прикладів лінійної алгебри.

Вище нами розглядалася алгебра M_n матриць розмірності $n \times n$. Це ще один з конкретних прикладів лінійної алгебри над полем P.

Вправа. Довести наведені вище властивості лінійних операторів.

Теорема 1 (про ізоморфізм). *Нехай* Б $= \langle b_1, b_2, \ldots, b_n \rangle$ – *базис векторного простору P^n і*

$$\varphi_Б : L_n \to M_n$$

– взаємно однозначне відображення алгебри L_n лінійних операторів простору P^n на алгебру M_n матриць, побудоване в теоремі про подання лінійних операторів. Воно має такі властивості:

I. $\varphi_Б(\mathcal{A}+\mathcal{B})_Б = \varphi_Б(\mathcal{A}) + \varphi_Б(\mathcal{B})$.

II. $\varphi_Б(\mathcal{A}\cdot\mathcal{B}) = \varphi_Б(\mathcal{A}) \cdot \varphi_Б(\mathcal{B})$.

III. $\varphi_Б(\lambda\mathcal{A}) = \lambda\varphi_Б(\mathcal{A})$.

Доведення. Раніше ми умовились щодо позначення

$$\varphi_Б(\mathcal{A}) = \mathcal{A}_Б .$$

Тоді сформульовані в теоремі властивості можна записати так:

I. $(\mathcal{A}+\mathcal{B})_Б = \mathcal{A}_Б + \mathcal{B}_Б$.

II. $(\mathcal{A}\cdot\mathcal{B})_Б = \mathcal{A}_Б \cdot \mathcal{B}_Б$.

III. $(\lambda\mathcal{A})_Б = \lambda\cdot\mathcal{A}_Б$.

Стовпець з номером k матриці $\mathcal{A}_Б = (\alpha_{ik})$ містить коефіцієнти $\alpha_{1k}, \alpha_{2k}, \ldots, \alpha_{nk}$, що визначаються рівністю

$$\mathcal{A}(b_k) = \alpha_{1k}b_1 + \alpha_{2k}b_2 + \ldots + \alpha_{nk}b_n,\ k = \overline{1,n} .$$

Стовпець з номером k матриці $\mathcal{B}_Б = (\beta_{ik})$ містить коефіцієнти $\beta_{1k}, \beta_{2k}, \ldots, \beta_{nk}$, що визначаються рівністю

$$\mathcal{B}(b_k) = \beta_{1k}b_1 + \beta_{2k}b_2 + \ldots + \beta_{nk}b_n,\ k = \overline{1,n} .$$

Обчислення матриці $(\mathcal{A}+\mathcal{B})_Б$.

$$(\mathcal{A}+\mathcal{B})(b_k) = \mathcal{A}(b_k) + \mathcal{B}(b_k) =$$
$$= \sum_{i=1}^{n}\alpha_{ik}b_i + \sum_{i=1}^{n}\beta_{ik}b_i = \sum_{i=1}^{n}(\alpha_{ik} + \beta_{ik})b_i .$$

Це означає, що k-й стовпець матриці $(\mathcal{A}+\mathcal{B})_Б$ складається з сум $\alpha_{1k} + \beta_{1k}$, $\alpha_{2k} + \beta_{2k}$, ..., $\alpha_{nk} + \beta_{nk}$. Оскільки це має місце для всіх номерів k, то

$$(\mathcal{A}+\mathcal{B})_Б = \mathcal{A}_Б + \mathcal{B}_Б .$$

Обчислення матриці $(\mathcal{A}\cdot\mathcal{B})_Б$:

$$(\mathcal{A}\cdot\mathcal{B})(b_k)=\mathcal{A}(\mathcal{B}(b_k))=\mathcal{A}(\beta_{1k}b_1+\beta_{2k}b_2+\ldots+\beta_{nk}b_n)=$$

$$=\beta_{1k}\mathcal{A}(b_1)+\beta_{2k}\mathcal{A}(b_2)+\ldots+\beta_{nk}\mathcal{A}(b_n)=$$

$$=\beta_{1k}\left(\sum_{j=1}^{n}\alpha_{j1}b_j\right)+\beta_{2k}\left(\sum_{j=1}^{n}\alpha_{j2}b_j\right)+\ldots+\beta_{nk}\left(\sum_{j=1}^{n}\alpha_{jn}b_j\right)=$$

$$=\left(\sum_{j=1}^{n}\alpha_{1j}\beta_{jk}\right)b_1+\left(\sum_{j=1}^{n}\alpha_{2j}\beta_{jk}\right)b_2+\ldots+\left(\sum_{j=1}^{n}\alpha_{nj}\beta_{jk}\right)b_n=$$

$$=\gamma_{1k}b_1+\gamma_{2k}b_2+\ldots+\gamma_{nk}b_n,$$

де

$$\gamma_{ik}=\alpha_{i1}\beta_{1k}+\alpha_{i2}\beta_{2k}+\ldots+\alpha_{in}\beta_{nk}$$

– елементи k-го стовпця матриці $C=\mathcal{A}_Б\cdot\mathcal{B}_Б$. Останнє означає, що

$$(\mathcal{A}\cdot\mathcal{B})_Б=\mathcal{A}_Б\cdot\mathcal{B}_Б.$$

Третя властивість очевидна.

Оборотні лінійні оператори

Означення. *Лінійний оператор $\mathcal{A}$ простору P^n називається оборотним, коли для нього існує такий лінійний оператор $\mathcal{X}$, що $\mathcal{A}\mathcal{X}=\mathcal{X}\mathcal{A}=\mathcal{E}$ – тотожне(одиничне) перетворення.*

Теорема 2. *Для лінійного оператора $\mathcal{A}$ еквівалентні такі умови:*

I. *Дефект оператора $\mathcal{A}$ дорівнює нулю.*

II. *Ранг оператора $\mathcal{A}$ дорівнює n.*

III. *Матриця оператора $\mathcal{A}$ у будь-якому базисі оборотна.*

IV. *Оператор $\mathcal{A}$ оборотний.*

Доведення. 1) I $\Rightarrow$ II

Нехай дефект оператора $\mathcal{A}$ дорівнює нулю, тобто $\operatorname{Ker}\mathcal{A}=\theta$. Покажемо, що будь-яка лінійно незалежна система векторів $b_1, b_2, \ldots, b_p$ оператором $\mathcal{A}$ перетворюється в систему лінійно незалежних векторів. Нехай $c_1=\mathcal{A}(b_1)$, $c_2=\mathcal{A}(b_2)$, ..., $c_p=$ $=\mathcal{A}(b_p)$. Припустимо, що $\lambda_1c_1+\lambda_2c_2+\ldots+\lambda_pc_p=\theta$ або

$$\lambda_1 \mathcal{A}(b_1) + \lambda_2 \mathcal{A}(b_2) + \ldots + \lambda_p \mathcal{A}(b_p) = \theta.$$

Тому $\mathcal{A}(\lambda_1 b_1 + \lambda_2 b_2 + \ldots + \lambda_p b_p) = \theta$. Тоді

$$\lambda_1 b_1 + \lambda_2 b_2 + \ldots + \lambda_p b_p = \theta.$$

На підставі лінійної незалежності векторів $b_1, b_2, \ldots, b_p$ має бути $\lambda_1 = \lambda_2 = \ldots = \lambda_p = 0$, що й доводить лінійну незалежність векторів $c_1, c_2, \ldots, c_p$. Якщо $Б = \langle b_1, b_2, \ldots, b_n \rangle$ – базис простору P^n, тоді сукупність $C = \langle c_1, c_2, \ldots, c_n \rangle$ векторів $c_1 = \mathcal{A}(b_1)$, $c_2 = \mathcal{A}(b_2)$, ..., $c_n = \mathcal{A}(b_n)$ також буде базисом простору. Ранг оператора $\mathcal{A}$ дорівнює n.

2) $\mathrm{II} \Rightarrow \mathrm{III}$

Нехай ранг оператора $\mathcal{A}$ дорівнює n. Вектори базису $Б = \langle b_1, b_2, \ldots, b_n \rangle$ простору перетворюються у вектори $c_1 = \mathcal{A}(b_1)$, $c_2 = \mathcal{A}(b_2)$, ..., $c_n = \mathcal{A}(b_n)$. Покажемо, що система векторів $c_1, c_2, \ldots, c_n$ – базис простору. За умовою їх лінійна оболонка $L(c_1, c_2, \ldots, c_n)$ збігається з простором P^n, а тому ранг системи Г дорівнює n і, отже, вектори $c_1, c_2, \ldots, c_n$ лінійно незалежні.

Разом з тим матриця $A = \mathcal{A}_Б$ лінійного оператора $\mathcal{A}$ є матрицею переходу від базису Б до базису Г. Тому вона невироджена, а отже, оборотна.

3) $\mathrm{III} \Rightarrow \mathrm{IV}$

Нехай Б – базис простору P^n і $A = \mathcal{A}_Б$ – оборотна матриця. Існує така матриця X, для якої $AX = XA = E$. Поряд з цим, існує лінійний оператор $\mathcal{X}$, для якого $X = \mathcal{X}_Б$. Тоді $(\mathcal{A}\mathcal{X})_Б = \mathcal{A}_Б \mathcal{X}_Б = A \cdot X = E$. Звідки випливає, що $\mathcal{A}\mathcal{X} = \mathcal{E}$. Аналогічно доводиться $\mathcal{X}\mathcal{A} = \mathcal{E}$. Отже, оператор $\mathcal{A}$ оборотний.

4) $\mathrm{IV} \Rightarrow \mathrm{I}$

Припустимо, що оператор $\mathcal{A}$ оборотний. Для нього знайдеться такий оператор $\mathcal{X}$, для якого $\mathcal{A}\mathcal{X} = \mathcal{E}$.

Якщо $x \in \operatorname{Ker} \mathcal{A}$, тоді $x = \mathcal{E}(x) = \mathcal{X}\mathcal{A}(x) = \mathcal{X}(\theta) = \theta$, тобто $\operatorname{Ker} \mathcal{A} = \theta$.

Якщо оператор $\mathcal{A}$ оборотний і оператор $\mathcal{X}$ такий, що $\mathcal{A}\mathcal{X} = \mathcal{X}\mathcal{A} = \mathcal{E}$, тоді він називається *оберненим* для $\mathcal{A}$. Він єдиний і позначається

$$\mathcal{X} = \mathcal{A}^{-1}.$$

Вправи

1. Довести, що лінійні оператори $\mathcal{A}$ і $\mathcal{B}$ оборотні тоді й лише тоді, коли оборотні оператори $\mathcal{A}\mathcal{B}$ і $\mathcal{B}\mathcal{A}$.

2. Довести, що коли $\mathcal{A}$, $\mathcal{B}$ – лінійні оператори і $\mathcal{A}$ оборотний, тоді ранг $\mathcal{A}\mathcal{B}$ = ранг $\mathcal{B}\mathcal{A}$ = ранг $\mathcal{B}$.

§ 6.5. Характеристичний многочлен

Означення. *Матриця B називається подібною до квадратної матриці $A = (\alpha_{ik})$ n-го порядку над полем P, якщо існує така невироджена матриця D, для якої*

$$B = D^{-1}AD.$$

Позначимо відношення подібності матриці B до матриці A через $B \simeq A$. Воно водночас є відношенням еквівалентності, тобто має властивості:

I. $A \simeq A$.

II. $B \simeq A \Rightarrow A \simeq B$.

III. $B \simeq A$ і $C \simeq B \Rightarrow C \simeq A$.

Відомо, що матриця лінійного оператора при переході від одного базису до іншого перетворюється подібно.

Означення. *Нехай $A = (\alpha_{ik})$ – квадратна матриця n-го порядку над полем P, t – змінна величина, яка не належить полю P. Тоді матриця $tE - A$ називається характеристичною матрицею для A, а многочлен*

$$\mathcal{X}_A(t) = \det(tE - A) = \begin{vmatrix} t-\alpha_{11} & -\alpha_{12} & \dots & -\alpha_{1n} \\ -\alpha_{21} & t-\alpha_{22} & \dots & -\alpha_{2n} \\ \dots & \dots & \dots & \dots \\ -\alpha_{n1} & -\alpha_{n2} & \dots & t-\alpha_{nn} \end{vmatrix} \qquad (1)$$

називається характеристичним многочленом матриці A.

Якщо на підставі означення записати цей визначник у вигляді $n!$ доданків, то можемо переконатися, що один з цих доданків збігається з добутком $(t-\alpha_{11})(t-\alpha_{22})\cdots(t-\alpha_{nn})$. Решта доданків – добутки n множників і серед цих множників не більше $n-2$ множників виду $t-\alpha_{kk}$. Тому після зведення всієї суми за степенями змінної t дістанемо

$$\mathcal{X}_A(t)=t^n-\beta_1 t^{n-1}-\beta_2 t^{n-2}-\ldots-\beta_n,$$

де $\beta_1=\alpha_{11}+\alpha_{22}+\ldots+\alpha_{nn}$ – сума діагональних елементів матриці A. Ця сума називається слідом матриці A і позначається:

$$trA=\alpha_{11}+\alpha_{22}+\ldots+\alpha_{nn}.$$

Решта коефіцієнтів β_k пов'язана з коефіцієнтами матриці A більш складними залежностями. Зауважимо, що

$$-\beta_n=\mathcal{X}_A(0)=\begin{vmatrix} -\alpha_{11} & -\alpha_{12} & \ldots & -\alpha_{1n} \\ -\alpha_{21} & -\alpha_{22} & \ldots & -\alpha_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -\alpha_{n1} & -\alpha_{n2} & \ldots & -\alpha_{nn} \end{vmatrix}=(-1)^n\,|A|. \quad (1°)$$

Теорема 1. *Характеристичні многочлени подібних матриць збігаються.*

Доведення. Нехай $B=D^{-1}AD$, де D – оборотна матриця. Тоді

$$\mathcal{X}_B(t)=|tE-B|=\left|tE-D^{-1}AD\right|=\left|tD^{-1}D-D^{-1}AD\right|=$$
$$=\left|D^{-1}(tE-A)D\right|=\left|D^{-1}\right|\cdot|tE-A|\cdot|D|=\left|D^{-1}\right|\cdot|D|\cdot|tE-A|=$$
$$=\left|D^{-1}D\right|\cdot|tE-A|=|E|\cdot|tE-A|=\mathcal{X}_A(t).$$

Означення. *Характеристичним многочленом лінійного оператора $\mathcal{A}$ простору P^n називається характеристичний многочлен матриці $A=\mathcal{A}_Б$ цього оператора у деякому базисі* Б *простору.*

Оскільки матриці оператора у різних базисах подібні, то їх характеристичний многочлен – інваріант.

Поліноміальною матрицею називається матриця, елементами якої є многочлени від деякої змінної t. Наприклад, характеристична матриця (1) – поліноміальна.

Матриця

$$\begin{pmatrix} t^2+1 & t-2 \\ 2t-1 & t^2+2 \end{pmatrix}$$

також поліноміальна і її можна подати у вигляді

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} t^2 + \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} t + \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}.$$

Будь-яка поліноміальна матриця $B(t)$, елементами якої є многочлени степеня $\le m$, може бути подана у вигляді суми

$$B(t) = B_0 t^m + B_1 t^{m-1} + \ldots + B_{m-1} t + B_m,$$

де коефіцієнти B_0, B_1, ..., B_{m-1}, B_m – скалярні матриці.

Теорема 2 (Гамільтона – Келі). *Якщо $\mathcal{X}(t)$ – характеристичний многочлен квадратної матриці A, то*

$$\mathcal{X}(A) = 0,$$

де 0 – нульова матриця.

Доведення. Для характеристичної матриці $X = tE - A$ і її приєднаної матриці $\tilde{X}$ виконується співвідношення

$$X \cdot \tilde{X} = \tilde{X} \cdot X = |X| \cdot E = \mathcal{X}(t)E. \tag{2}$$

Матриця $\tilde{X}$ – поліноміальна. Її елементи – алгебричні доповнення елементів X і, отже, многочлени від t степеня $\le n-1$. Тому

$$\tilde{X} = B_0 t^{n-1} + B_1 t^{n-2} + \ldots + B_{n-2} t + B_{n-1}.$$

Звідси

$$X \cdot \tilde{X} = (tE - A)\left(B_0 t^{n-1} + B_1 t^{n-2} + \ldots + B_{n-2} t + B_{n-1}\right) =$$
$$= B_0 t^n + \left(B_1 - AB_0\right) t^{n-1} + \left(B_2 - AB_1\right) t^{n-2} + \ldots +$$
$$+ \left(B_{n-1} - AB_{n-2}\right) t - AB_{n-1}.$$

Характеристичний многочлен записуємо у вигляді

$$\mathcal{X}(t) = t^n - \beta_1 t^{n-1} - \beta_2 t^{n-2} - \ldots - \beta_n$$

і тому

$$\mathcal{X}(t)E = Et^n - \beta_1 E t^{n-1} - \ldots - \beta_{n-1} t E - \beta_n E.$$

На підставі (2) матимемо систему рівностей

$$\begin{aligned} B_0 &= E, \\ B_1 - AB_0 &= -\beta_1 E, \\ B_2 - AB_1 &= -\beta_2 E, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ B_{n-1} - AB_{n-2} &= -\beta_{n-1} E, \\ -AB_{n-1} &= -\beta_n E. \end{aligned} \tag{Г–К}$$

Помноживши першу рівність на A^n, другу – на A^{n-1}, ..., передостанню – на A_1 (кожен раз зліва), дістанемо систему

$$\begin{cases} A^n B_0 = A^n, \\ A^{n-1} B_1 - A^n B_0 = -\beta_1 A^{n-1}, \\ A^{n-2} B_2 - A^{n-1} B_1 = -\beta_2 A^{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -AB_{n-1} = -\beta_n E. \end{cases}$$

Підсумуємо ліві та праві частини цих рівностей. Одержимо

$$A^n - \beta_1 A^{n-1} - \beta_2 A^{n-2} - \ldots - \beta_{n-1} A - \beta_n E = 0,$$

тобто $\mathcal{X}(A) = 0$.

Вправи

1. Обчислити характеристичний многочлен лінійного оператора $\mathcal{A}(x) = [a, x]$ простору R^3.

2. Знайти характеристичний многочлен матриці

$$\begin{pmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \ldots & \ldots & \ldots & \ldots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_n \end{pmatrix}.$$

Відповідь: $\mathcal{X}(t) = t^n - (x_1 y_1 + x_2 y_2 + \ldots + x_n y_n) t^{n-1}$.

3. Довести, що для квадратних матриць A і B одного й того самого порядку характеристичні многочлени матриць AB і BA збігаються.

Характеристичні числа

Нехай P – числове поле (підполе поля C комплексних чисел). *Корені характеристичного многочлена* $\mathcal{X}(t)=|tE-A|$ *матриці* $A=(\alpha_{ik})$ *з елементами* α_{ik} *із поля* P *називаються характеристичними числами.* Характеристичне число λ матриці може і не належати основному полю P.

Для характеристичного числа λ матриці $A=(\alpha_{ik})$ складемо систему лінійних однорідних рівнянь

$$\begin{cases}(\lambda-\alpha_{11})x_1-\alpha_{12}x_2-\ldots-\alpha_{1n}x_n=0,\\ -\alpha_{21}x_1+(\lambda-\alpha_{22})x_2-\ldots-\alpha_{2n}x_n=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ -\alpha_{n1}x_1-\alpha_{n2}x_2-\ldots+(\lambda-\alpha_{nn})x_n=0.\end{cases} \tag{3}$$

Оскільки визначник цієї системи $\mathcal{X}(\lambda)=|\lambda E-A|=0$, то вона має ненульові розв'язки $x=(\xi_1, \xi_2, \ldots, \xi_n)$ з координатами з поля C комплексних чисел. Якщо x^T – стовпець координат цього вектора, то має виконуватись рівність

$$(\lambda E-A)\cdot x^T=\theta, \tag{4}$$

або

$$A\cdot x^T=\lambda x^T. \tag{5}$$

Теорема 3. *Всі характеристичні числа дійсної симетричної матриці – числа дійсні.*

Доведення. Нехай $A=(\alpha_{ik})$ – симетрична матриця, тобто $A=A^T$, і λ – її характеристичне число. Складемо для цієї матриці систему (3) і знайдемо який-небудь нульовий розв'язок $x=(\xi_1, \xi_2, \ldots, \xi_n)$ цієї системи. Для нього повинна мати місце рівність (5), тобто

$$\alpha_{i1}\xi_1+\alpha_{i2}\xi_2+\ldots+\alpha_{in}\xi_n=\lambda\xi_i \quad (i=\overline{1,\,n}).$$

Помноживши праву і ліву частини останньої рівності на спряжене число $\overline{\xi}_i$ та підсумувавши одержані рівності за всіма індексами i, дістанемо

$$\sum_{i,k=1}^{n}\alpha_{ik}\xi_k\overline{\xi}_i=\lambda\left(\sum_{i=1}^{n}\xi_i\overline{\xi}_i\right).$$

Отже,

$$\lambda=\frac{w}{\rho},$$

де $w=\sum\limits_{i,k=1}^{n}\alpha_{ik}\xi_k\overline{\xi}_i$, $\rho=\sum\limits_{i=1}^{n}\xi_i\overline{\xi}_i$ – число додатне.

Крім того,

$$\overline{w}=\overline{\sum_{i,k=1}^{n}\alpha_{ik}\xi_k\overline{\xi}_i}=\sum_{i,k=1}^{n}\overline{\alpha}_{ik}\overline{\xi}_k\overline{\overline{\xi}}_i=\sum_{i=1}^{n}\alpha_{ik}\overline{\xi}_k\xi_i .$$

Оскільки $\alpha_{ik}=\alpha_{ki}$, то $\overline{w}=\sum\limits_{i,k=1}^{n}\alpha_{ki}\overline{\xi}_k\xi_i=\sum\limits_{i,k=1}^{n}\alpha_{ki}\xi_i\overline{\xi}_k=w$.

Число w – дійсне, а тому і λ – дійсне.

Теорема 4. *Якщо λ_1, λ_2, ..., λ_n – усі характеристичні числа матриці A, $g(x)$ – многочлен, тоді $g(\lambda_1)$, $g(\lambda_2),\ldots,$ $g(\lambda_n)$ – усі характеристичні числа матриці $g(A)$. При цьому, якщо $g(x)=a_0x^m+a_1x^{m-1}+\ldots+a_m$, то $g(A)=a_0A^m+a_1A^{m-1}+\ldots+$ $+a_mE$.*

Доведення. Нехай $\mathcal{X}(t)=|tE-A|$ – характеристичний многочлен матриці A і $\mathcal{X}(t)=(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$.

Розкладемо також многочлен $g(x)$ на лінійні множники

$$g(x)=a_0(x-\mu_1)(x-\mu_2)\cdots(x-\mu_m).$$

В обидві частини останньої рівності замість змінної x поставимо матрицю A:

$$g(A)=a_0E(A-\mu_1E)(A-\mu_2E)\cdots(A-\mu_mE).$$

Тоді

$$|g(A)| = a_0^n |A - \mu_1 E| \cdot |A - \mu_2 E| \cdots |A - \mu_m E| =$$

$$= a_0^n (-1)^n |\mu_1 E - A| \cdot (-1)^n |\mu_2 E - A| \cdots (-1)^n |\mu_m E - A| =$$

$$= (-1)^{mn} a_0^n \mathcal{X}(\mu_1)\mathcal{X}(\mu_2)\cdots\mathcal{X}(\mu_m) = (-1)^{mn} a_0^n \prod_{i=1}^{m}\prod_{k=1}^{n}(\mu_i - \lambda_k) =$$

$$= a_0^n \prod_{i=1}^{m}\prod_{k=1}^{n}(\lambda_k - \mu_i) = \prod_{k=1}^{n}\left\{a_0 \prod_{i=1}^{m}(\lambda_k - \mu_i)\right\} =$$

$$= g(\lambda_1)\cdot g(\lambda_2)\cdots g(\lambda_n).$$

В отриманій рівності

$$|g(A)| = g(\lambda_1)\cdot g(\lambda_2)\cdots g(\lambda_n)$$

замість многочлена $g(x)$ поставимо $t - g(x)$, де t – параметр.

Дістанемо

$$|tE - g(A)| = (t - g(\lambda_1))(t - g(\lambda_2))\cdots(t - g(\lambda_n)).$$

Тут $|tE - g(A)| = \chi_{g(A)}(t)$ – характеристичний многочлен матриці $g(A)$.

Теорему доведено.

Покажемо, що добуток усіх характеристичних чисел матриці A дорівнює визначнику цієї матриці.

Нехай $\mathcal{X}(t)$ – характеристичний многочлен (1). Якщо $\lambda_1, \lambda_2, ..., \lambda_n$ його корені, тоді $\mathcal{X}(0) = (-1)^n \lambda_1 \cdots \lambda_n$ і на підставі (1) матимемо $\lambda_1 \cdot \lambda_2 \cdots \lambda_n = |A|$.

Приклад 1. Знайти характеристичні числа матриці n-го порядку

$$\begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

(в ній $\alpha_{kk-1} = 1$, $\alpha_{k-1k} = -1$, і на решті місць – нулі).

Характеристичним многочленом для цієї матриці буде

$$\mathcal{X}_n(t)=\begin{vmatrix} t & 1 & 0 & \dots & 0 \\ -1 & t & 1 & \dots & 0 \\ 0 & -1 & t & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & t \end{vmatrix}.$$

Нескінченна послідовність $\mathcal{X}_1(t),\ \mathcal{X}_2(t), \dots,\ \mathcal{X}_n(t)\dots$ підкоряється зворотному співвідношенню $\mathcal{X}_n(t)=t\mathcal{X}_{n-1}(t)+$ $+\mathcal{X}_{n-2}(t)$. Згідно з загальною методикою обчислення членів зворотної послідовності (див. § 2.5) характеристичне рівняння для $\chi_n(t)$ матиме вигляд $x^2=t\cdot x+1$. Корені цього рівняння

$$x_{1,2}=\frac{t\pm\sqrt{t^2+4}}{2}.$$

Отже, $\mathcal{X}_n(t)=Ax_1^{n-1}+Bx_2^{n-1}$. Коефіцієнти A і B задовольняють систему рівнянь

$$\begin{cases} A+B=\mathcal{X}_1(t)=t; \\ Ax_1+Bx_2=\mathcal{X}_2(t)=t^2+1. \end{cases}$$

Звідси матимемо

$$A=\frac{t^2+1-tx_2}{x_1-x_2},\quad B=\frac{t^2+1-tx_1}{x_2-x_1}.$$

Враховуючи, що $x_1+x_2=t$ і $x_1x_2=-1$, після очевидних перетворень знаходимо

$$\mathcal{X}_n(t)=\frac{x_1^{n+1}-x_2^{n+1}}{x_1-x_2}.$$

Коренями многочлена $\mathcal{X}_n(t)$ є ті значення змінної t, для яких $x_1^{n+1}-x_2^{n+1}=0$ і $x_1-x_2\neq 0$. Тому співвідношення $\dfrac{x_1}{x_2}$ має дорівнювати кореням $(n+1)$-го степеня з одиниці, але відмінні від одиниці:

$$\varepsilon_k = \cos\frac{2\pi k}{n+1} + i\sin\frac{2\pi k}{n+1},\ \ k = \overline{1,\ n}\,.$$

Корені многочлена $\mathcal{X}_n(t)$ обчислюються з рівності

$$\frac{t+\sqrt{t^2+4}}{t-\sqrt{t^2+4}} = \varepsilon_k,\ \ k = \overline{1,\ n}\,.$$

Звідси

$$\frac{1+\sqrt{1+\dfrac{4}{t^2}}}{1-\sqrt{1+\dfrac{4}{t^2}}} = \varepsilon_k,\ k = \overline{1,\ n},\ t = \pm 2i\cos\frac{\pi k}{n+1}\,.$$

Можна вибрати $t = -2i\cos\dfrac{\pi k}{n+1}$. Тоді

$$\frac{t+\sqrt{t^2+4}}{t-\sqrt{t^2+4}} = \frac{-2i\cos\dfrac{\pi k}{n+1} + 2\sin\dfrac{\pi k}{n+1}}{-2i\cos\dfrac{\pi k}{n+1} - 2\sin\dfrac{\pi k}{n+1}} =$$

$$= \left(\cos\frac{\pi k}{n+1} + i\sin\frac{\pi k}{n+1}\right)^2 =$$

$$= \cos\frac{2\pi k}{n+1} + i\sin\frac{2\pi k}{n+1} = \varepsilon_k \quad (k = 1,\ 2,\ \ldots,\ n).$$

Приклад 2. Характеристичним многочленом матриці n-го порядку

$$A_n = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 1 & -1 & \ldots & 0 \\ 0 & 1 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

є $\varphi_n(t) = \mathcal{X}_n(t-1)$, де $\mathcal{X}_n(t)$ – характеристичний многочлен матриці з попереднього прикладу.

Метод Д. К. Фаддєєва

Нехай $A=(\alpha_{ik})$ – квадратна матриця порядку n з коефіцієнтами із числового поля і $X=tE-A$ – її характеристична матриця.

Викладемо метод обчислення характеристичного многочлена

$$\mathcal{X}(t)=|tE-A|=t^n-\beta_1 t^{n-1}-\beta_2 t^{n-2}-\ldots-\beta_n$$

і приєднаної матриці

$$\tilde{X}=B_0 t^{n-1}-B_1 t^{n-2}+\ldots+B_{n-2}t+B_{n-1},$$

запропонований Д. К. Фаддєєвим [21].

Якщо λ_1, λ_2, ..., λ_n – характеристичні числа матриці A, тоді $\beta_1=\lambda_1+\lambda_2+\ldots+\lambda_n$ і, разом з тим, $\beta_1=trA=\alpha_{11}+$ $+\alpha_{22}+\ldots+\alpha_{nn}$.

Характеристичними числами степеня A^k $(k=1, 2, \ldots)$ будуть степені λ_1^k, λ_2^k, ..., λ_n^k і

$$S_k=\lambda_1^k+\lambda_2^k+\ldots+\lambda_n^k=trA^k.$$

Суми S_k k-х степенів коренів λ_1, λ_2, ..., λ_n пов'язані з коефіцієнтами многочлена $\chi(t)$ такими співвідношеннями (формули Ньютона):

$$k\cdot\beta_k=S_k-\beta_1 S_{k-1}-\beta_2 S_{k-2}-\ldots-\beta_{k-1}S_1,\ k=1, 2, 3, \ldots, n.$$

Для матриці A обчислюються послідовності матриць A_1, A_2, ..., A_n, $B_0=E$, B_1, B_2, ..., B_n і чисел γ_1, γ_2, ..., γ_n за правилом:

$$\begin{cases} A_1=A, & \gamma_1=trA_1, & B_1=A_1-\gamma_1 E; \\ A_2=AB_1, & \gamma_2=\dfrac{1}{2}trA_2, & B_2=A_2-\gamma_2 E; \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ A_{n-1}=AB_{n-2}, & \gamma_{n-1}=\dfrac{1}{n-1}trA_{n-1}, & B_{n-1}=A_{n-1}-\gamma_{n-1}E; \\ A_n=AB_{n-1}, & \gamma_n=\dfrac{1}{n}trA_n, & B_n=A_n-\gamma_n E. \end{cases} \qquad (\Phi)$$

Доведемо, що $\mathcal{X}(t)=t^n-\gamma_1 t^{n-1}-\gamma_2 t^{n-2}-...-\gamma_n$, $B_n=A_n-$ $-\gamma_n E=0$. Маємо $\gamma_1=\beta_1$. Далі застосуємо індукцію, припустивши, що $\gamma_1=\beta_1$, $\gamma_2=\beta_2$, …, $\gamma_{k-1}=\beta_{k-1}$ – рівність коефіцієнтів $\chi(t)$ при степенях t^{n-1}, t^{n-2}, ..., t^{n-k+1}.

Із схеми (Ф) матимемо

$$A_k=AB_{k-1}=A(A_{k-1}-\gamma_{k-1}E)=AA_{k-1}-\gamma_{k-1}A.$$

Продовживши далі для A_{k-1}, дістанемо

$$A_k=A^k-\sum_{j=1}^{k-1}\gamma_j A^{k-j}. \tag{7}$$

Звідси випливає рівність

$$trA_k=trA^k-\sum_{j=1}^{k-1}\gamma_j trA^{k-j}.$$

Отже,

$$k\gamma_k=S_k-\sum_{j=1}^{k-1}\gamma_j S_{k-j}. \tag{8}$$

За припущенням індукції $\gamma_j=\beta_j$, $j=\overline{1,k-1}$. Тоді

$$k\gamma_k=S_k-\sum_{j=1}^{k-1}\beta_j S_{k-j}.$$

На підставі формули Ньютона матимемо рівність $k\gamma_k=k\beta_k$ і, отже, $\gamma_k=\beta_k$. Індукцію завершено. Коефіцієнти β_1, β_2, ..., β_n многочлена $\mathcal{X}(t)$ збігаються з числами γ_1, γ_2, ..., γ_n. Якщо в рівності (7) покласти $k=n-1$ і використати одержані рівності $\gamma_j=\beta_j$ для будь-яких $j=\overline{1,n}$, дістанемо

$$B_n=A_n-\gamma_n E=A^n-\beta_1 A^{n-1}-\beta_2 A^{n-2}-...-\beta_n E.$$

Права частина цієї рівності дорівнює нульовій матриці на підставі теореми 2 (теорема Гамільтона – Келі). Коефіцієнти B_0, B_1, B_2, ..., B_{n-1} із системи рівностей (Г–К) збігаються з матрицями B_k, одержаними в процесі обчислень за схемою (Ф) Фаддєєва.

Приклад 3. Обчислити коефіцієнти характеристичного многочлена матриці

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

Обчислення.

1. $\gamma_1 = trA = 4, \quad B_1 = A_1 - \gamma_1 E = \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & -1 & -1 \\ 1 & -1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{pmatrix}.$

2. $A_2 = A \cdot B_1 = \begin{pmatrix} 0 & -4 & -4 & -4 \\ -4 & 0 & 4 & 4 \\ -4 & 4 & 0 & 4 \\ -4 & 4 & 4 & 0 \end{pmatrix}, \; \gamma_2 = \frac{1}{2} trA_2 = 0,$

$$B_2 = 4\begin{pmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

3. $A_3 = A \cdot B_2 = 4\begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & -1 & -1 \\ 1 & -1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{pmatrix}, \; \gamma_3 = \frac{1}{3} trA_3 = -16,$

$$B_3 = A_3 - \gamma_3 E = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 4 & 4 & -4 & -4 \\ 4 & -4 & 4 & -4 \\ 4 & -4 & -4 & 4 \end{pmatrix} = 4 \cdot A.$$

4. Контрольний крок

$$A_4 = A \cdot B_3 = \begin{pmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix} = 16 \cdot E,$$

$$\gamma_4 = 16, \quad B_4 = A_4 - \gamma_4 E = 0.$$

Остаточно

$$\mathcal{X}(t) = t^4 - 4t^3 + 0 \cdot t^2 + 16t - 16 = (t^2 - 4)(t - 2)^2.$$

§ 6.6. Власні значення лінійного оператора

Нехай $\mathcal{A}$ – лінійний оператор векторного простору P^n над полем P.

Означення. *Елемент λ поля P називається регулярним значенням лінійного оператора $\mathcal{A}$, якщо оператор $\lambda\mathcal{E} - \mathcal{A}$ оборотний.*

Елемент λ поля P називається власним значенням оператора $\mathcal{A}$, якщо оператор $\lambda\mathcal{E} - \mathcal{A}$ необоротний.

Теорема. *Власні значення лінійного оператора $\mathcal{A}$ – корені характеристичного многочлена цього оператора. Будь-який корінь характеристичного многочлена лінійного оператора $\mathcal{A}$ з поля P, буде його власним значенням.*

Доведення. Нехай $\mathcal{A}$ – лінійний оператор простору P^n, Б – деякий базис цього простору, $A = \mathcal{A}_Б$ – матриця оператора $\mathcal{A}$. Якщо λ – власне значення оператора, тоді матриця $\lambda E - A = (\lambda\mathcal{E} - \mathcal{A})_Б$ необоротна. Тому визначник $|\lambda E - A| = 0$. Характеристичним многочленом оператора $\mathcal{A}$ є $\mathcal{X}(t) = |tE - A|$. Це означає, що власне значення λ – корінь многочлена $\mathcal{X}(t)$.

Доведення оберненого твердження досить просте.

Означення. *Нехай λ – власне значення лінійного оператора $\mathcal{A}$. Тоді підпростір $V_\lambda = \mathrm{Ker}(\lambda\mathcal{E} - \mathcal{A})$ має додатну розмір-*

ність. Будь-який ненульовий вектор із V_λ називається власним вектором, відповідним власному значенню λ (λ – власним вектором).

Інакше кажучи, вектор $x \neq \theta - \lambda$ – власний вектор, якщо виконується рівність

$$(\lambda\mathcal{E} - \mathcal{A})(x) = \theta. \tag{1}$$

Його можна подати в еквівалентній формі

$$\mathcal{A}(x) = \lambda x. \tag{2}$$

Із викладеного вище випливає такий спосіб обчислення власних значень і власних векторів лінійного оператора $\mathcal{A}$.

Нехай у стандартному базисі простору P^n $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, ..., 0)$, ..., $e_n = (0, 0, ..., 1)$ (як і в будь-якому іншому базисі) оператору $\mathcal{A}$ відповідає матриця

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

Обчислюємо характеристичний многочлен

$$\mathcal{X}(t) = \begin{vmatrix} t - \alpha_{11} & -\alpha_{12} & \dots & -\alpha_{1n} \\ -\alpha_{21} & t - \alpha_{22} & \dots & -\alpha_{2n} \\ \dots & \dots & \dots & \dots \\ -\alpha_{n1} & -\alpha_{n2} & \dots & t - \alpha_{nn} \end{vmatrix},$$

та знаходимо його корені, що належать полю P.

Нехай λ – корінь $\mathcal{X}(t)$ і $\lambda \in P$. Усі λ – власні вектори $x = (\xi_1, \xi_2, ..., \xi_n) \neq 0$ задовольняють рівняння

$$(\lambda E - A)x^T = \theta, \tag{3}$$

де x^T – стовпець координат вектора x.

Підпростір $V_\lambda = \operatorname{Ker}(\lambda\mathcal{E} - \mathcal{A})$ збігається з множиною всіх розв'язків системи лінійних рівнянь

$$\begin{cases} (\lambda-\alpha_{11})x_1-\alpha_{12}x_2-\ldots-\alpha_{1n}x_n=0,\\ -\alpha_{21}x_1+(\lambda-\alpha_{22})x_2-\ldots-\alpha_{2n}x_n=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ -\alpha_{n1}x_1-\alpha_{n2}x_2-\ldots+(\lambda-\alpha_{nn})x_n=0. \end{cases} \qquad (V_\lambda)$$

Усі ненульові розв'язки цієї системи – власні вектори.

Наслідок 1. *Якщо $\mathcal{A}$ – лінійний оператор комплексного векторного простору (основне поле – поле C комплексних чисел), тоді всі корені його характеристичного многочлена – власні значення оператора.*

Будь-який оператор цього простору має власні вектори.

Наслідок 2. *Якщо $\mathcal{A}$ – лінійний оператор дійсного векторного простору (основне поле – поле R дійсних чисел), тоді його власними значеннями будуть усі дійсні корені характеристичного многочлена.*

Не всякий лінійний оператор простору R^n має власні значення. Нехай, наприклад, у просторі R^2 через $\mathcal{A}_\alpha$ позначимо оператор обертання навколо початку координат на кут $\alpha \neq k\pi$ $(k=0,\ \pm 1,\ \pm 2, \ldots)$.

У базисі $e_1=(1, 0)$, $e_2=(0, 1)$ йому відповідає матриця

$$A=\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Характеристичними числами цієї матриці будуть комплексні числа $\lambda_1=\cos\alpha+i\sin\alpha$, $\lambda_2=\cos\alpha-i\sin\alpha$, оскільки $\alpha\neq k\pi$. Оператор $\mathcal{A}_\alpha$ не має власних значень, а тому і власних векторів.

Відсутність власних векторів оператора $\mathcal{A}_\alpha$ обертання на кут $\alpha\neq 0, \pi, 2\pi, \ldots$ геометрично очевидна.

Приклад. Знайти власні вектори матриці

$$A=\begin{pmatrix} -7 & 12 & -6 \\ -3 & 5 & -3 \\ 3 & -6 & 2 \end{pmatrix}.$$

Розв'язання. Обчислимо характеристичний многочлен

$$\mathcal{X}(t)=\begin{vmatrix} t+7 & -12 & 6 \\ 3 & t-5 & 3 \\ -3 & 6 & t-2 \end{vmatrix}=(t+1)^2(t-2).$$

Власні значення $\lambda_1=\lambda_2=-1,\ \lambda_3=2$.

Для $\lambda_1=-1$ складаємо систему рівнянь

$$\begin{cases} 6x_1-12x_2+6x_3=0, \\ 3x_1-6x_2+3x_3=0, \\ -3x_1+6x_2-3x_3=0. \end{cases}$$

Вона еквівалентна одному рівнянню $x_1-2x_2+x_3=0$. Загальний розв'язок $X=(2x_2-x_3,x_2,x_3)$. Фундаментальна система розв'язків складається з $a_1=(2,\ 1,\ 0)$ і $a_2=(-1,\ 0,\ 1)$. Усі ненульові лінійні комбінації

$$x=\alpha a_1+\beta a_2$$

– власні вектори з власним значенням -1.

Для $\lambda=2$ складаємо систему рівнянь

$$\begin{cases} 9x_1-12x_2+6x_3=0, \\ 3x_1-3x_2+3x_3=0, \\ -3x_1+6x_2=0. \end{cases}$$

Загальний розв'язок $X=(-2x_3,\ -x_3,x_3)$.

Фундаментальна система розв'язків містить лише один вектор $b=(2,\ 1,\ -1)$. Усі ненульові вектори виду $x=\gamma\cdot b$ – власні вектори з власним значенням 2.

Вправи

1. Знайти власні значення і власні вектори матриці порядку n:

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

2. Знайти власні значення і власні вектори матриці порядку n:

$$A = \begin{pmatrix} a & b & b & \dots & b \\ b & a & b & \dots & b \\ \dots & \dots & \dots & \dots & \dots \\ b & b & b & \dots & a \end{pmatrix}.$$

3. Довести, що всі відмінні від нуля вектори простору P^n тоді і лише тоді будуть власними векторами оператора $\mathcal{A}$, коли $\mathcal{A}$ – скалярний оператор, тобто $\mathcal{A} = \alpha\,\mathcal{E}$.

4. Довести, що коли оператор $\mathcal{A}$ оборотний, то $\mathcal{A}$ і $\mathcal{A}^{-1}$ мають одні й ті самі власні вектори. Знайти залежність між їх власними значеннями.

§ 6.7. Інваріантні підпростори

Означення. *Лінійний підпростір L векторного простору P^n називається інваріантним відносно лінійного оператора $\mathcal{A}$, якщо для будь-якого вектора $x \in L$ вектор $y = \mathcal{A}(x) \in L$.*

Тривіальними інваріантними підпросторами довільного лінійного оператора $\mathcal{A}$ є весь простір P^n і нульовий підпростір $\{\theta\}$.

Якщо оператор $\mathcal{A}$ має власне значення λ, тоді λ – власний вектор x породжує одновимірний інваріантний підпростір $L = L(x)$.

Дійсно, оскільки $\mathcal{A}(x) = \lambda x$, то для будь-якого вектора $z = \zeta x$ із L маємо

$$\mathcal{A}(z) = \mathcal{A}(\zeta x) = \zeta\,\mathcal{A}(x) = \zeta\lambda x = \lambda \cdot (\zeta x) = \lambda \cdot z.$$

Усі ненульові вектори із L також будуть власними векторами. І навпаки, якщо L – одновимірний підпростір, інваріантний для оператора $\mathcal{A}$, то всі ненульові вектори із L будуть власними векторами цього оператора з єдиним власним значенням.

Існують лінійні оператори дійсного простору R^2, що не мають інваріантних підпросторів розмірності 1. Таким оператором, наприклад, є оператор $\mathcal{A}_\alpha$ повороту на кут $\alpha \neq k\pi$ $(k = 0, \pm 1, \pm 2, \ldots)$.

Лема. *Власні вектори* $x_1, x_2, \ldots, x_m$ *лінійного оператора* $\mathcal{A}$ *з попарно відмінними власними значеннями* $\lambda_1, \lambda_2, \ldots, \lambda_m$ – *лінійно незалежні.*

Доведення. Застосуємо індукцію за числом m векторів. Якщо $m = 1$, тоді лема виконується. Припустимо, що лема має місце для $m-1$ власних векторів. Доведемо її для $x_1, x_2, \ldots, x_m$.

Нехай виконується рівність (а) $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m = \theta$. Оскільки $\mathcal{A}(x_k) = \lambda_k x_k$ для будь-яких k, то матимемо (б) $\mathcal{A}\left(\sum_{k=1}^{m} \alpha_k x_k\right) = \alpha_1\lambda_1 x_1 + \alpha_2\lambda_2 x_2 + \ldots + \alpha_m\lambda_m x_m = \theta$. З другого боку, помноживши рівність (а) на скаляр λ_m, дістанемо (в) $\alpha_1\lambda_m x_1 + \alpha_2\lambda_m x_2 + \ldots + \alpha_m\lambda_m x_m = \theta$.

Віднявши з рівності (б) рівність (в), знайдемо

$$\alpha_1(\lambda_1 - \lambda_m)x_1 + \alpha_2(\lambda_2 - \lambda_m)x_2 + \ldots +$$
$$+\alpha_{m-1}(\lambda_{m-1} - \lambda_m)x_{m-1} = \theta.$$

За припущенням індукції вектори $x_1, x_2, \ldots, x_{m-1}$ лінійно незалежні. Тому

$$\alpha_1(\lambda_1 - \lambda_m) = 0,$$
$$\alpha_2(\lambda_2 - \lambda_m) = 0, \ldots, \alpha_{m-1}(\lambda_{m-1} - \lambda_m) = 0.$$

Але $\lambda_k - \lambda_m \neq 0$ при $k < m$. Звідси випливає, що $\alpha_1 = \alpha_2 = \ldots = \alpha_{m-1} = 0$. Тоді з рівності (а) знаходимо, що $\alpha_m = 0$. А це й доводить лему.

Теорема 1. *Нехай* $\lambda_1, \lambda_2, \ldots, \lambda_m$ $(m \geq 2)$ – *попарно різні власні значення лінійного оператора* $\mathcal{A}$. *Підпростори* $V_{\lambda_k} = \operatorname{Ker}(\lambda_k\mathcal{E} - \mathcal{A})$ *є інваріантними відносно* $\mathcal{A}$ *і їх сума* $V = V_{\lambda_1} + V_{\lambda_2} + \ldots + V_{\lambda_m}$ є прямою сумою $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_m}$.

Доведення. Нехай λ – будь-яке власне значення оператора і $x \in V_\lambda = \operatorname{Ker}(\lambda\mathcal{E} - \mathcal{A})$. Тоді $\mathcal{A}(x) = \lambda x$. Якщо $y =$

$= \mathcal{A}(x)$, то $\mathcal{A}(y) = \mathcal{A}(\mathcal{A}(x)) = \mathcal{A}(\lambda x) = \lambda\mathcal{A}(x) = \lambda y$. Це означатиме, що $y \in V_\lambda$ і V_λ інваріантний підпростір для $\mathcal{A}$.

Нехай $x \in V_{\lambda_1} \cap \left(V_{\lambda_2} + \ldots + V_{\lambda_m}\right)$. Тоді $x \in V_{\lambda_1}$ і $x = x_2 + \ldots + x_m$, $x_k \in V_{\lambda_k}$. Якщо $x \neq \theta$, то остання рівність суперечить лемі. Тому $x = \theta$ і $V_{\lambda_1} \cap \left(V_{\lambda_2} + \ldots + V_{\lambda_m}\right) = \{\theta\}$. Аналогічно $V_{\lambda_k} \cap \left(V_{\lambda_1} + \ldots + V_{\lambda_{k-1}} + V_{\lambda_{k+1}} + \ldots + V_{\lambda_m}\right) = \{\theta\}$, якщо $k > 1$.

Остаточно $V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_m}$.

Теорему доведено.

Наслідок. *Нехай лінійний оператор $\mathcal{A}$ має власні вектори і $\lambda_1, \lambda_2, \ldots, \lambda_r$ – усі попарно різні його власні значення. Тоді найменшим інваріантним підпростором, який містить усі власні вектори оператора $\mathcal{A}$, є $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_r}$. Цей підпростір назвемо лінійною оболонкою всіх власних векторів (ВВ-оболонкою) і позначимо $V(\mathcal{A})$.*

Лінійний оператор $\mathcal{A}$ на інваріантному підпросторі $M \subset P^n$ індукує лінійне перетворення, яке називається *обмеженням* $\mathcal{A}$ на M. Воно позначається як $\mathcal{A}\big|_M$.

Теорема 2. *Нехай $\mathcal{X}(t)$ – характеристичний многочлен лінійного оператора $\mathcal{A}$ і M – інваріантний підпростір. Тоді характеристичний многочлен $\mathcal{X}_1(t)$ оператора $\mathcal{A}\big|_M$ – дільник многочлена $\mathcal{X}(t)$.*

Доведення. Вважатимемо, що $m = \dim M$ і $0 < m < n$. Оберемо який-небудь базис $\mathrm{C} = \langle b_1, b_2, \ldots, b_m\rangle$ підпростору M і доповнимо його до базису $\mathrm{Б} = \langle b_1, \ldots, b_m, b_{m+1}, \ldots, b_n\rangle$ простору P^n. Побудуємо матрицю A оператора $\mathcal{A}$ у цьому базисі.

Врахувавши, що $\mathcal{A}(b_1), \ldots, \mathcal{A}(b_m) \in M$, матимемо рівності

$$\mathcal{A}(b_1) = \alpha_{11}b_1 + \alpha_{21}b_2 + \ldots + \alpha_{m1}b_m + 0\cdot b_{m+1} + \ldots + 0\cdot b_n,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\mathcal{A}(b_m) = \alpha_{1m}b_1 + \alpha_{2m}b_2 + \ldots + \alpha_{mm}b_m + 0\cdot b_{m+1} + \ldots + 0\cdot b_n,$$

$$\mathcal{A}(b_{m+1}) = \alpha_{1m+1}b_1 + \alpha_{2m+1}b_2 + \ldots + \alpha_{mm+1}b_m +$$
$$+\alpha_{m+1m+1} \cdot b_{m+1} + \ldots + \alpha_{nm+1} \cdot b_n,$$

....

$$\mathcal{A}(b_n) = \alpha_{1n}b_1 + \alpha_{2n}b_2 + \ldots + \alpha_{mn}b_m +$$
$$+\alpha_{m+1n} \cdot b_{m+1} + \ldots + \alpha_{nn} \cdot b_n.$$

Матриця A матиме клітинно-трикутний вигляд

$$A = \left(\begin{array}{c|c} A_1 & B \\ \hline 0 & A_2 \end{array}\right),$$

де матриця A_1 – розмірності $m \times m$, матриця A_2 – розмірності $(n-m)\times(n-m)$ і матриця B – розмірності $m\times(n-m)$. При цьому A_1 – матриця оператора $\mathcal{A}|_M$ у базисі $C = \langle b_1, b_2, \ldots, b_m \rangle$.

Характеристичний многочлен оператора $\mathcal{A}|_M$ дорівнює $\mathcal{X}_1(t) = |tE_m - A_1|$, а характеристичний многочлен оператора $\mathcal{A}$:

$$\mathcal{X}(t) = |tE - A| = \det\left(\begin{array}{c|c} tE_m - A_1 & -B \\ \hline 0 & tE_{n-m} - A_2 \end{array}\right) =$$
$$= |tE_m - A_1| \cdot |tE_{n-m} - A_2| = \mathcal{X}_1(t) \cdot \mathcal{X}_2(t),$$

де $\mathcal{X}_2(t)$ – характеристичний многочлен матриці A_2.

Наслідок. *Нехай l – кратність характеристичного числа λ лінійного оператора $\mathcal{A}$ і $\lambda \in P$. Тоді дефект оператора $\lambda\mathcal{E} - \mathcal{A}$ не перевищує кратності l.*

Доведення. Оскільки λ – власне значення оператора, то дефект m оператора $\lambda\mathcal{E} - \mathcal{A}$ (розмірність підпростору $V_\lambda = \mathrm{Ker}(\lambda\mathcal{E} - \mathcal{A})$) додатний. Матрицею A_1 індукованого оператора $\mathcal{A}|_{V_\lambda}$ є діагональна матриця

$$\begin{pmatrix} \lambda & & & 0 \\ & \lambda & & \\ & & \ddots & \\ 0 & & & \lambda \end{pmatrix}$$

порядку m. Його характеристичний многочлен – $\mathcal{X}_1(t) = (t-\lambda)^m$. Оскільки $\mathcal{X}_1(t)$ – дільник характеристичного многочлена оператора $\mathcal{A}$, то $m \le l$.

Означення. *Лінійний оператор $\mathcal{A}$ називається звідним, якщо в просторі P^n існує нетривіальний підпростір M ($0 < \dim M < n$), інваріантний для $\mathcal{A}$.*

Лінійний оператор $\mathcal{A}$ називається цілком звідним, якщо для нього існують такі інваріантні підпростори M і L, що $P^n = M \oplus L$ і $0 < \dim M < n$.

Для цілком звідного оператора $\mathcal{A}$ побудуємо базис в P^n: оберемо який-небудь базис Г в M і який-небудь базис Д в L. Тоді їх об'єднання Б = Г $\cup$ Д – базис простору P^n.

У базисі Б матриця оператора має клітинно-діагональний вигляд:

$$A = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right).$$

При цьому A_1 і A_2 – відповідно матриці операторів $\mathcal{A}|_M$ і $\mathcal{A}|_L$.

Матриця звідного лінійного оператора перетворюється подібно до клітинно-трикутного вигляду.

Матриця цілком звідного оператора перетворюється подібно до клітинно-діагонального вигляду.

Теорема 3. *У дійсному векторному просторі R^n для будь-якого лінійного оператора $\mathcal{A}$ існує інваріантний підпростір розмірності 1 або* 2.

Доведення. Нехай $A = (\alpha_{ik})$ – матриця оператора $\mathcal{A}$ у деякому базисі, λ – характеристичне число оператора (матриці).

Якщо λ – число дійсне, то воно є власним значенням матриці і, отже, відповідний власний вектор визначить інваріантний підпростір розмірності 1.

Нехай $\lambda = \alpha + \beta i$ – комплексне число ($\beta \ne 0$). Розглянемо рівняння

$$(\lambda E - A) \cdot x^T = \theta, \tag{1}$$

де $x = (\xi_1, \xi_2, ..., \xi_n)$ – вектор комплексного простору C^n. Подамо його у вигляді системи рівнянь

$$
\begin{cases}
(\lambda-\alpha_{11})\xi_1-\alpha_{12}\xi_2-\ldots-\alpha_{1n}\xi_n=0,\\
-\alpha_{21}\xi_1+(\lambda-\alpha_{22})\xi_2-\ldots-\alpha_{2n}\xi_n=0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
-\alpha_{n1}\xi_1-\alpha_{n2}\xi_2-\ldots+(\lambda-\alpha_{nn})\xi_n=0.
\end{cases}
\tag{2}
$$

Вона має ненульовий розв'язок

$$
\xi_1^0=\eta_1+\zeta_1 i,\ \ \xi_2^0=\eta_2+\zeta_2 i,\ \ldots,\ \xi_n^0=\eta_n+\zeta_n i\,.
$$

Підставивши цей розв'язок у систему рівнянь (2), дістанемо

$$
\begin{cases}
\alpha_{11}\xi_1^0+\alpha_{12}\xi_2^0+\ldots+\alpha_{1n}\xi_n^0=\lambda\xi_1^0,\\
\alpha_{21}\xi_1^0+\alpha_{22}\xi_2^0+\ldots+\alpha_{2n}\xi_n^0=\lambda\xi_2^0,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\alpha_{n1}\xi_1^0+\alpha_{n2}\xi_2^0+\ldots+\alpha_{nn}\xi_n^0=\lambda\xi_n^0.
\end{cases}
\tag{3}
$$

Виділивши окремо дійсні та уявні частини, з системи (3) матимемо

$$
\begin{cases}
\alpha_{11}\eta_1+\alpha_{12}\eta_2+\ldots+\alpha_{1n}\eta_n=\alpha\eta_1-\beta\zeta_1,\\
\alpha_{21}\eta_1+\alpha_{22}\eta_2+\ldots+\alpha_{2n}\eta_n=\alpha\eta_2-\beta\zeta_2,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\alpha_{n1}\eta_1+\alpha_{n2}\eta_2+\ldots+\alpha_{nn}\eta_n=\alpha\eta_n-\beta\zeta_n.
\end{cases}
$$

$$
\begin{cases}
\alpha_{11}\zeta_1+\alpha_{12}\zeta_2+\ldots+\alpha_{1n}\zeta_n=\beta\eta_1+\alpha\zeta_1,\\
\alpha_{21}\zeta_1+\alpha_{22}\zeta_2+\ldots+\alpha_{2n}\zeta_n=\beta\eta_2+\alpha\zeta_2,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\alpha_{n1}\zeta_1+\alpha_{n2}\zeta_2+\ldots+\alpha_{nn}\zeta_n=\beta\eta_n+\alpha\zeta_n.
\end{cases}
$$

Для векторів $y=(\eta_1,\eta_2,\ldots,\eta_n)$ і $z=(\zeta_1,\zeta_2,\ldots,\zeta_n)$ із простору R^n останні дві системи означають:

$$
A\cdot y^T=\alpha y^T-\beta z^T\text{, тобто } \mathcal{A}(y)=\alpha y-\beta z\,;
$$

$$
A\cdot z^T=\beta y^T+\alpha z^T\text{, тобто } \mathcal{A}(z)=\beta y+\alpha z.
$$

Лінійна оболонка $L = L(y, z)$ інваріантна відносно оператора $\mathcal{A}$.

Переконаємось, що вектори y і z лінійно незалежні. Зрозуміло, що один з них обов'язково відмінний від нульового вектора.

Припустимо, що $z \neq 0$ і $y = \delta z$, де δ – дійсне число. Тоді

$$\mathcal{A}(y) = \alpha(\delta \cdot z) - \beta z,$$

$$\mathcal{A}(y) = \delta \cdot \mathcal{A}(z) = \delta(\beta\delta z + \alpha z) = \beta\delta^2 z + \delta\alpha z.$$

Звідси $\alpha\delta z - \beta z = \beta\delta^2 z + \alpha\delta z$ і тому $\beta\left(1 + \delta^2\right) z = \theta$.

Оскільки $\delta^2 \geq 0$ і $z \neq \theta$, то $\beta = 0$. Останнє суперечить припущенню, що $\beta \neq 0$.

Отже, вектори y і z – лінійно незалежні і $\dim L = 2$.

Вправи

1. Знайти всі інваріантні підпростори оператора $\mathcal{A}(x) = [a, x]$ у просторі R^3.

2. Нехай лінійні оператори $\mathcal{A}$ і $\mathcal{B}$ переставні. Довести, що $\operatorname{Ker} \mathcal{B}$ і $\operatorname{Im} \mathcal{B}$ інваріантні відносно $\mathcal{A}$.

3. Знайти всі інваріантні підпростори оператора диференціювання в просторі M многочленів степеня $\leq n$ з дійсними коефіцієнтами.

4. Довести, що коли лінійний оператор $\mathcal{A}$ оборотний, тоді $\mathcal{A}$ і $\mathcal{A}^{-1}$ мають одні й ті самі інваріантні підпростори.

Лінійні оператори простої структури

Теорема 4. *Такі властивості лінійного оператора $\mathcal{A}$ векторного простору P^n еквівалентні:*

I. *У просторі P^n існує базис, складений із власних векторів оператора $\mathcal{A}$.*

II. *Усі корені характеристичного многочлена $\mathcal{X}(t)$ оператора $\mathcal{A}$ належать полю P і*

$$P^n = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_r},$$

де $V_{\lambda_k} = \mathrm{Ker}(\lambda_k \mathcal{E} - \mathcal{A})$; $\lambda_1, \lambda_2, \ldots, \lambda_r$ *– усі попарно різні корені* $\mathcal{X}(t)$.

III. *Усі корені характеристичного многочлена* $\mathcal{X}(t)$ *належать полю* P *і кратність будь-якого кореня* λ *дорівнює дефекту оператора* $\lambda\mathcal{E} - \mathcal{A}$.

Доведення. Нехай $\Phi = \langle f_1, f_2, \ldots, f_n \rangle$ – базис простору P^n, який містить власні вектори оператора $\mathcal{A}$. Тоді у цьому базисі його матриця $A = \mathcal{A}_\Phi$ має діагональний вигляд

$$A = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

і $\mathcal{X}(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n)$. Усі корені λ_k належать полю P. Нехай серед них $\lambda_1, \lambda_2, \ldots, \lambda_r$ – усі різні корені. Тоді ВВ-оболонка оператора $V(\mathcal{A}) = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}$ (у позначеннях теореми 1) містить базис Φ і, отже, збігається з P^n:

$$P^n = V(\mathcal{A}) = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}.$$

Цим доведено, що I $\Rightarrow$ II.

Доведемо, що II $\Rightarrow$ I. На підставі властивості II обираємо базис простору P^n таким чином. Нехай $Б_k$ – базис підпростору $V_{\lambda_k} = \mathrm{Ker}(\lambda_k \mathcal{E} - \mathcal{A})$ $(k = \overline{1, r})$. Тоді об'єднання $Б = Б_1 \cup Б_2 \cup \ldots \cup Б_r$ – базис простору P^n, який складається з власних векторів оператора $\mathcal{A}$.

Доведемо еквівалентність властивостей II і III. Нехай виконується властивість II. Якщо m_k – дефект оператора $\lambda_k \mathcal{E} - \mathcal{A}$ і l_k – кратність кореня λ_k многочлена $\mathcal{X}(t)$ $(k = \overline{1, r})$, то $m_k \le l_k$. Оскільки з властивості II випливає рівність $m_1 + m_2 + \ldots + m_r = n$ і завжди виконується рівність $l_1 + l_2 + \ldots + l_r = n$, то дістаємо $m_k = l_k$ для будь-яких $k = \overline{1, r}$.

Цим доведено, що II $\Rightarrow$ III. Обернене твердження III $\Rightarrow$ II – очевидне.

Теорему доведено.

Лінійний оператор $\mathcal{A}$ простору P^n називається *оператором простої структури*, коли в просторі існує базис, що складається з його власних векторів.

Квадратна матриця A з елементами з поля P називається *матрицею простої структури*, якщо вона подібна до діагональної матриці.

Якщо $\mathcal{A}$ – лінійний оператор простої структури, тоді його матриця $A = \mathcal{A}_Б$ у довільному базисі Б простору P^n має просту структуру.

Сформулюємо достатні умови простоти структури лінійного оператора (матриці).

Якщо лінійний оператор $\mathcal{A}$ простору P^n має n різних власних значень $\lambda_1, \lambda_2, ..., \lambda_n$, тоді структура його проста.

Оберемо деякий власний вектор l_k з власним значенням λ_k для кожного номера $k = \overline{1, n}$. За доведеною вище лемою вектори $f_1, f_2, ..., f_n$ лінійно незалежні й, отже, становлять базис простору P^n.

Приклад 1. Квадратна дійсна матриця $A_n = (\alpha_{ik})$ порядку n називається *якобієвою (тридіагональною)*, якщо її елементи $\alpha_{ik} = 0$ за умови $|i-k| > 1$.

Розв'язання. Введемо позначення

$$\alpha_i = \alpha_{ii} \ (i = \overline{1, n});\ \beta_i = -\alpha_{i\,i+1};\ \gamma_i = -\alpha_{i+1\,i} \ (i = \overline{1, n-1}).$$

Тоді цю матрицю можна записати так:

$$A_n = \begin{pmatrix} \alpha_1 & -\beta_1 & \dots & 0 & 0 \\ -\gamma_1 & \alpha_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \alpha_{n-1} & -\beta_{n-1} \\ 0 & 0 & \dots & -\gamma_{n-1} & \alpha_n \end{pmatrix}.$$

Її характеристичний многочлен

$$\mathcal{X}_n(t)=\begin{vmatrix} t-\alpha_1 & \beta_1 & \dots & 0 \\ \gamma_1 & t-\alpha_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & t-\alpha_n \end{vmatrix}.$$

Розклавши цей визначник за елементами останнього рядка, дістанемо зворотне співвідношення для многочленів $\mathcal{X}_n(t)$:

$$\mathcal{X}_n(t)=(t-\alpha_n)\mathcal{X}_{n-1}(t)-\beta_{n-1}\gamma_{n-1}\mathcal{X}_{n-2}(t).$$

Розглянемо послідовність

$$\mathcal{X}_n(t), \mathcal{X}_{n-1}(t), \dots, \mathcal{X}_2(t), \mathcal{X}_1(t), \mathcal{X}_0(t), \qquad (\mathcal{X}_n)$$

де покладемо $\mathcal{X}_0(t)=1$.

За умови $\beta_k\gamma_k>0$ $(k=\overline{1, n-1})$ ця послідовність має такі властивості:

I. $\mathcal{X}_0(t)=1$ – величина стала.

II. Два многочлени, що містяться поруч, $\mathcal{X}_{k-1}(t)$ і $\mathcal{X}_k(t)$ $(1\le k\le n)$ не обертаються на нуль при одному й тому ж значенні змінної t.

III. Якщо λ – корінь многочлена $\mathcal{X}_k(t)$, $0<k<n$, тоді $\mathcal{X}_{k+1}(t)$ і $\mathcal{X}_{k-1}(t)$ мають протилежні знаки при $t=\lambda$.

Послідовність $(\mathcal{X}_n)$ – узагальнений ряд Штурма. Раніше було доведено (див. § 1.9): многочлен $\mathcal{X}_n(t)$ має n попарно різних дійсних коренів.

Тридіагональна матриця A_n за умови $\beta_k\gamma_k>0$ $(1\le k\le n-1)$ має просту структуру.

Приклад 2. Лінійний оператор у базисі $e_1=(1,\ 0,\ 0,\ 0)$, $e_2=(0,\ 1,\ 0,\ 0)$, $e_3=(0,\ 0,\ 1,\ 0)$ і $e_4=(0,\ 0,\ 0,\ 1)$ задано матрицею

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

Показати, що структура цієї матриці проста, і звести її подібним перетворенням до діагонального вигляду.

Розв'язання. Обчислюємо характеристичний многочлен матриці A:

$$\mathcal{X}(t)=\begin{vmatrix} t+3 & -3 & 2 & -2 \\ 4 & t-4 & 2 & -2 \\ 4 & -3 & t+1 & -2 \\ 4 & -3 & 2 & t-3 \end{vmatrix}=t(t-1)^3 .$$

Корені многочлена: $\lambda_1=0$, $\lambda_2=\lambda_3=\lambda_4=1$.

Для власного значення $\lambda_1=0$ ядро матриці $\lambda_1 E-A=-A$ складається з усіх векторів $x=(\xi_1,\xi_2,\xi_3,\xi_4)$, для яких $(-A)x^T=\theta$.

Перепишемо цю рівність у вигляді системи рівнянь:

$$\begin{cases} 3\xi_1-3\xi_2+2\xi_3-2\xi_4=0, \\ 4\xi_1-4\xi_2+2\xi_3-2\xi_4=0, \\ 4\xi_1-3\xi_2+\xi_3-2\xi_4=0, \\ 4\xi_1-3\xi_2+2\xi_3-3\xi_4=0. \end{cases}$$

Ранг системи дорівнює 3. Фундаментальна система розв'язків має лише один розв'язок $f_1=(1, 1, 1, 1)$.

Для власного значення $\lambda=1$ матриця $\lambda E-A$ має вигляд

$$E-A=\begin{pmatrix} 4 & -3 & 2 & -2 \\ 4 & -3 & 2 & -2 \\ 4 & -3 & 2 & -2 \\ 4 & -3 & 2 & -2 \end{pmatrix}.$$

Її ранг дорівнює 1. Ядро її складається з векторів $x=(\xi_1,\xi_2,\xi_3,\xi_4)$, для яких

$$(E-A)x^T=\theta .$$

У координатній формі це рівняння еквівалентне одному рівнянню

$$4\xi_1-3\xi_2+2\xi_3-2\xi_4=0 .$$

Загальний розв'язок $x=\left(\frac{3}{4}\xi_2-\frac{1}{2}\xi_3+\frac{1}{2}\xi_4,\ \xi_2,\ \xi_3,\ \xi_4\right)$. Фундаментальна система розв'язків складається з векторів

$$f_2 = (3, 4, 0, 0),\ f_3 = (-1, 0, 2, 0) \text{ і } f_4 = (1, 0, 0, 2).$$

Новим базисом простору обираємо $f=\left(f_1, f_2, f_3, f_4\right)$. Матрицею переходу від старого базису до базису f буде

$$D=\begin{pmatrix} 1 & 3 & -1 & 1 \\ 1 & 4 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix},\ D^{-1}=\frac{1}{2}\begin{pmatrix} 8 & -6 & 4 & -4 \\ -2 & 2 & -1 & 1 \\ -4 & 3 & -1 & 2 \\ -4 & 3 & -2 & 3 \end{pmatrix}.$$

Тоді

$$D^{-1}AD=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Вправи

1. Довести, що оператор $\mathcal{A}(x)=(x,a)a$ простору R^3, де $a=(1, 2, 3)$, має просту структуру.

2. Показати, що матриця n-го порядку

$$\begin{pmatrix} 1 & 1 & 1 & \ldots & \ldots & \ldots & 1 \\ 1 & 1 & 1 & \ldots & \ldots & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & 1 & \ldots & \ldots & \ldots & 1 \end{pmatrix}$$

простої структури. Подібним перетворенням звести її до діагонального вигляду.

3. Довести, що коли оператор $\mathcal{A}$ простору P^n має n різних власних значень, то оператор $\mathcal{B}$, переставний з $\mathcal{A}$, має просту структуру. При цьому всі власні вектори оператора $\mathcal{A}$ будуть також власними векторами оператора $\mathcal{B}$.

4. Показати, що коли $\mathcal{A}$ – оператор простої структури, то для будь-якого многочлена $f(t)$ оператор $f(\mathcal{A})$ також простої структури.

§ 6.8. Нормальна форма Жордана

Позначимо через $I_a(\lambda)$ квадратну матрицю виду

$$\begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

порядку a , де на головній діагоналі містяться елементи λ з основного поля P, нижче цієї діагоналі – одиниці, решта елементів – нулі. Така матриця називається *клітиною Жордана.*

Матриця виду

$$\begin{pmatrix} I_{a_1}(\lambda_1) & & & 0 \\ & I_{a_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & I_{a_k}(\lambda_k) \end{pmatrix},$$

в якій уздовж головної діагоналі містяться клітини Жордана $I_{a_j}(\lambda_j)$, $j = \overline{1, k}$, *а решта елементів матриці – нулі, називається матрицею, яка має нормальну форму Жордана.*

Наприклад, матриця

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

має нормальну форму Жордана.

Виявляється, що будь-яка матриця над полем комплексних чисел подібна до матриці в нормальній формі Жордана.

Нільпотентні оператори і матриці

Лінійний оператор $\mathcal{A}$ векторного простору P^n називається нільпотентним, якщо деякий його степінь $\mathcal{A}^k = \mathcal{O}$ ($k \in N,\ k \geq 1$).

Квадратна матриця A називається нільпотентною, якщо деякий її степінь $A^k = 0$ ($k \in N,\ k \geq 1$).

Лінійний оператор $\mathcal{A}$ (матриця A) буде нільпотентним, якщо його характеристичний многочлен $\mathcal{X}(t)$ не має ненульових коренів, тобто $\mathcal{X}(t) = t^n$. Тоді за теоремою Гамільтона – Келі маємо $A^n = 0$.

Жорданова клітина $I_a(0)$ є нільпотентною матрицею. Тому матриця жорданового виду

$$\begin{pmatrix} I_{a_1}(0) & & & 0 \\ & I_{a_2}(0) & & \\ & & \ddots & \\ 0 & & & I_{a_k}(0) \end{pmatrix}$$

(з нулями на головній діагоналі) – нільпотентна.

Виявляється, що *будь-яка нільпотентна матриця подібна до матриці такого виду.* Щоб обґрунтувати це, розглянемо будову нільпотентних операторів.

Для нільпотентного оператора $\mathcal{A}$ знайдеться таке натуральне число $m \geq 1$, для якого $\mathcal{A}^m = 0$, але $\mathcal{A}^q \neq \mathcal{O}$ для всіх цілих чисел q, що задовольняють умову $0 < q < m$. Таке число m називається *порядком нільпотентності* оператора.

Розмістимо всі степені $\mathcal{A}^q$, $0 \leq q \leq m$ оператора у такому порядку

$$\mathcal{E} = \mathcal{A}^0, \mathcal{A}^1, \mathcal{A}^2, \ldots, \mathcal{A}^{m-1}, \mathcal{A}^m = \mathcal{O}$$

(де $\mathcal{A}^0$ – тотожне перетворення $\mathcal{E}$).

Позначимо через K_q ядро оператора $\mathcal{A}^q$: $K_q = \mathrm{Ker}\,\mathcal{A}^q$. При цьому $K_0 = \{\theta\}$, $K_m = P^n$. Далі розглянемо декілька окремих фактів.

1. Підпростори K_q утворюють зростаючу послідовність за включенням:

$$\theta = K_0 \subset K_1 \subset \ldots \subset K_q \subset K_{q+1} \subset \ldots \subset K_{m-1} \subset K_m = P^n.$$

Покажемо, що $K_q \subset K_{q+1}$. Якщо $x \in K_q$, тобто $\mathcal{A}^q(x) = \theta$, то $\mathcal{A}^{q+1}(x) = \mathcal{A}(\mathcal{A}^q x) = \theta$. Це означає, що $x \in K_{q+1} = \operatorname{Ker} \mathcal{A}^{q+1}$.

2. Доведемо інваріантність K_q відносно оператора $\mathcal{A}$ для всіх $0 \le q \le m$. Якщо $x \in K_q$, тобто $\mathcal{A}^q x = \theta$ і $y = \mathcal{A}(x)$, то $\mathcal{A}^{q-1}(y) = \theta$. Це означає, що $y \in K_{q-1}$. Оператор $\mathcal{A}$ відображає K_q в K_{q-1}. Оскільки $K_{q-1} \subset K_q$, то K_q інваріантний відносно $\mathcal{A}$.

3. Оберемо в просторі $P^n = K_m$ який-небудь базис $Б_1 = \langle e_1, e_2, \ldots, e_s \rangle$ відносно підпростору K_{m-1} (базис mod K_{m-1}). Покажемо, що вектори $f_1 = \mathcal{A}(e_1)$, $f_2 = \mathcal{A}(e_2)$, ..., $f_s = \mathcal{A}(e_s)$ належать K_{m-1} і лінійно незалежні відносно K_{m-2}. Їх належність підпростору K_{m-1} доведено в пункті 2. Покажемо їх лінійну незалежність mod K_{m-2}. Нехай

$$\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_s f_s = h \in K_{m-2},$$

тобто

$$\alpha_1 \mathcal{A}(e_1) + \alpha_2 \mathcal{A}(e_2) + \ldots + \alpha_s \mathcal{A}(e_s) \in K_{m-2}.$$

Тоді

$$\mathcal{A}(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_s e_s) \in K_{m-2}.$$

Це означає, що $\mathcal{A}^{m-1}(\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_s e_s) = \theta$. Отже, $\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_s e_s$ належить K_{m-1}. Але вектори $e_1, e_2, \ldots, e_s$ становлять базис простору K_m відносно підпростору K_{m-1}. Тому $\alpha_1 = \alpha_2 = \ldots = \alpha_s = 0$.

4. Набір векторів $f_i = \mathcal{A}(e_i)$, $i = \overline{1, s}$, доповнимо (за необхідністю) до базису підпростору K_{m-1} відносно K_{m-2}. Нехай цей базис mod K_{m-2} складається з системи

$$Б_2 = \langle \mathcal{A}(e_1), \mathcal{A}(e_2), \ldots, \mathcal{A}(e_s), e_{s+1}, \ldots, e_r \rangle.$$

Аналогічно доводиться, що вектори

$$\mathcal{A}^2(e_1), \mathcal{A}^2(e_2), \ldots, \mathcal{A}^2(e_s), \mathcal{A}(e_{s+1}), \ldots, \mathcal{A}(e_r)$$

належать підпростору K_{m-2} і лінійно незалежні mod K_{m-3}.

Набір векторів $\mathcal{A}^2(e_1)$, ..., $\mathcal{A}^2(e_s)$, $\mathcal{A}(e_{s+1})$, ..., $\mathcal{A}(e_r)$ може бути доповненим (за необхідністю) до базису mod K_{m-3} . Нехай цим відносним базисом буде сукупність векторів

$$Б_3 = <\mathcal{A}^2(e_1), ..., \mathcal{A}^2(e_s), \mathcal{A}(e_{s+1}), ..., \mathcal{A}(e_r), e_{r+1}, ..., e_p>.$$

Описаний процес продовжується і закінчується побудовою базису останнього, відмінного від $\{\theta\}$, підпростору K_1 .

Об'єднання всіх побудованих відносних базисів $Б_1$, $Б_2$, $Б_3$, ... складає базис усього простору

$$Б = \begin{cases} \hfill e_1, e_2, ..., e_s; \\ \hfill \mathcal{A}(e_1), \mathcal{A}(e_2), ..., \mathcal{A}(e_s), e_{s+1}, ..., e_r; \\ \mathcal{A}^2(e_1), ..., \mathcal{A}^2(e_s), \mathcal{A}(e_{s+1}), ..., \mathcal{A}(e_r)\ e_{r+1}, ..., e_p; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{cases}$$

Перегрупуємо базисні вектори на такі підсистеми

$$C_1 = <e_1, \mathcal{A}(e_1), ..., \mathcal{A}^{a-1}(e_1)>, \mathcal{A}^a(e_1) = \theta;$$
$$C_2 = <e_2, \mathcal{A}(e_2), ..., \mathcal{A}^{b-1}(e_2)>, \mathcal{A}^b(e_2) = \theta;$$
$$C_3 = <e_3, \mathcal{A}(e_3), ..., \mathcal{A}^{c-1}(e_3)>, \mathcal{A}^c(e_3) = \theta;$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

Тоді оператору $\mathcal{A}$ в цьому базисі $C = C_1 \cup C_2 \cup ...$ відповідає матриця

$$A = \begin{pmatrix} I_a(0) & & & 0 \\ & I_b(0) & & \\ & & I_c(0) & \\ 0 & & & \ddots \end{pmatrix},$$

де вздовж головної діагоналі містяться нільпотентні клітини Жордана.

Побудований базис $C = C_1 \cup C_2 \cup C_3 \cup ...$ називається *канонічним базисом* для нільпотентного оператора.

Більш детально проілюструємо викладене вище на конкретному прикладі. Нехай лінійний оператор задано матрицею

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

Її характеристичний многочлен

$$\mathcal{X}(t) = |tE - A| = \begin{vmatrix} t & -1 & -1 & 0 \\ 1 & t-2 & 0 & -1 \\ 1 & 0 & t+2 & -1 \\ 0 & 1 & 1 & t \end{vmatrix} = t^4 .$$

За теоремою Гамільтона – Келі $\mathcal{X}(A) = A^4 = 0$. У дійсному векторному просторі R^4 і в деякому базисі (наприклад, у стандартному базисі e_1, e_2, e_3, e_4, складеному з одиничних векторів) лінійний оператор $\mathcal{A}$ з матрицею A нільпотентний. Необхідно для нього побудувати канонічний базис Жордана.

Обчислення дають

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

Порядок нільпотентності дорівнює 3. Тоді дістанемо:

1. K_1 = Ker A = $L(b_1, b_2)$, де $b_1 = (-2, -1, 1, 0)^T$, $b_2 = (1, 0, 0, 1)^T$ – базисні вектори.

2. K_2 = Ker $A^2 = L(c_1, c_2, c_3)$, де $c_1 = (1, 1, 0, 0)^T$, $c_2 = (-1, 0, 1, 0)^T$, $c_3 = (1, 0, 0, 1)^T$ – базисні вектори.

3. $K_3 = R^4$.

Базисом простору R^4відносно K_2 обираємо один вектор $f = (1, 0, 0, 0)^T$. При цьому $A \cdot f = (0, -1, -1, 0)^T \in K_2$ і $A \cdot f$ становить базис підпростору K_2 відносно K_1. Далі, $A^2 \cdot f = (-2, -2, 2, 2)^T \in K_1$. Базис підпростору K_1 складається з двох векторів: вектора $A^2 \cdot f$ і вектора $b_2 = (1, 0, 0, 1)^T$. Канонічний базис С складається з векторів f, $A \cdot f$, $A^2 \cdot f$, b_2 .

У цьому базисі матрицею оператора $\mathcal{A}$ є матриця

$$I = \mathcal{A}_C = \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right).$$

Розщеплення оператора

Нехай $\mathcal{A}$ – лінійний оператор з характеристичним многочленом $\mathcal{X}(t)$. Розкладемо цей многочлен на попарно взаємно прості множники

$$\mathcal{X}(t) = \varphi_1(t) \cdot \varphi_2(t) \cdots \varphi_k(t)$$

над основним полем P лінійного простору P^n. Тоді виявляється, що простір P^n розкладається в пряму суму $P^n = L_1 \oplus L_2 \oplus \ldots \oplus L_k$ лінійних підпросторів L_j ($j = \overline{1, k}$), які мають дві властивості:

1) кожен L_j інваріантний відносно оператора $\mathcal{A}$;

2) оператор $\mathcal{A}$ індукує на L_j такий лінійний оператор (позначаємо його також через $\mathcal{A}$), що $\varphi_j(\mathcal{A}) = \mathcal{O}$.

Для простоти викладу розглянемо випадок, коли $\mathcal{X}(t)$ розкладається лише на два взаємно прості множники, тобто $\mathcal{X}(t) = \varphi_1(t) \cdot \varphi_2(t)$. Загальний випадок зводиться до цього.

Позначимо $L_1 = \operatorname{Ker} \varphi_1(\mathcal{A})$, $L_2 = \operatorname{Ker} \varphi_2(\mathcal{A})$ і покажемо, що

а) L_1 і L_2 інваріантні відносно $\mathcal{A}$;

б) $P^n = L_1 \oplus L_2$.

Доведення. Нехай $x \in L_j$, тобто $\varphi_j(\mathcal{A})(x) = \theta$. Якщо $y = \mathcal{A}(x)$, тоді $\varphi_j(\mathcal{A})(y) = \varphi_j(\mathcal{A})(\mathcal{A}(x)) = \mathcal{A}\varphi_j(\mathcal{A})(x) = \mathcal{A}(\theta) = \theta$. Отже, $y \in L_j$, і інваріантність L_j доведено ($j = 1, 2$).

Оскільки множники $\varphi_1(t)$ і $\varphi_2(t)$ взаємно прості, то можна відшукати такі многочлени $f_1(t)$ і $f_2(t)$, що $f_1(t)\varphi_1(t) + f_2(t)\varphi_2(t) = 1$. Тоді $\mathcal{E} = f_1(\mathcal{A})\varphi_1(\mathcal{A}) + f_2(\mathcal{A})\varphi_2(\mathcal{A})$ і для будь-якого $x \in P^n$ матимемо

$$x = f_2(\mathcal{A})\varphi_2(\mathcal{A})(x) + f_1(\mathcal{A})\varphi_1(\mathcal{A})(x) = x_1 + x_2,$$

де $x_1 = f_2(\mathcal{A})\varphi_2(\mathcal{A})(x)$; $x_2 = f_1(\mathcal{A})\varphi_1(\mathcal{A})(x)$.

При цьому

$$\varphi_1(\mathcal{A})(x_1) = \varphi_1(\mathcal{A})f_2(\mathcal{A})\varphi_2(\mathcal{A})(x) = f_2(\mathcal{A})\varphi_1(\mathcal{A})\varphi_2(\mathcal{A})(x) = \\ = f_2(\mathcal{A})\chi(\mathcal{A})(x) = \theta,$$

тобто $x_1 \in \operatorname{Ker} \varphi_1(\mathcal{A})$.

Тут використано теорему Гамільтона – Келі, згідно з якою $\mathcal{X}(\mathcal{A}) = 0$. Так само $x_2 \in \operatorname{Ker}\varphi_2(\mathcal{A}) = L_2$. Отже, $x = x_1 + x_2$, $x_j \in L_j$. Це означає, що $P^n = L_1 + L_2$. Залишається встановити: $L_1 \cap L_2 = \{\theta\}$. Якщо $x \in L_1 \cap L_2$, то $\varphi_1(\mathcal{A})(x) = \theta$ і $\varphi_2(\mathcal{A})(x) = \theta$. Тоді $x = f_1(\mathcal{A})\varphi_1(\mathcal{A})(x) + f_2(\mathcal{A})\varphi_2(\mathcal{A})(x) = \theta$.

Зрозуміло, що $\mathcal{A}$ на L_j індукує такий оператор, що $\varphi_j(\mathcal{A}) = \mathbf{0}$.

Теорема 1. *Якщо характеристичний многочлен $\mathcal{X}(t)$ лінійного оператора $\mathcal{A}$ розкладається над основним полем P простору P^n у вигляді добутку $\mathcal{X}(t) = (t-\lambda_1)^{n_1} \times \times (t-\lambda_2)^{n_2} \dots (t-\lambda_k)^{n_k}$, де $\lambda_1, \lambda_2, \dots, \lambda_k$ – його попарно різні корені, тоді у відповідному базисі матриця оператора $\mathcal{A}$ має нормальну форму Жордана.*

Доведення. Спочатку розглянемо випадок, коли $\mathcal{X}(t) = (t-\lambda)^n$. Якщо $\lambda = 0$, тоді $\mathcal{X}(t) = t^n$ і оператор $\mathcal{A}$ нільпотентний згідно з доведеною вище теоремою. Якщо $\mathcal{X}(t) = (t-\lambda)^n$, $\lambda \neq 0$, то оператор $\mathcal{B} = \mathcal{A} - \lambda\mathcal{E}$ нільпотентний. У відповідному базисі С матимемо

$$\mathcal{B} \underset{C}{\leftrightarrow} B = \begin{pmatrix} I_a(0) & & & 0 \\ & I_b(0) & & \\ & & I_c(0) & \\ 0 & & & \ddots \end{pmatrix}.$$

Тоді

$$\mathcal{A} = \mathcal{B} + \lambda\mathcal{E} \underset{C}{\leftrightarrow} B + \lambda E = \begin{pmatrix} I_a(\lambda) & & & 0 \\ & I_b(\lambda) & & \\ & & I_c(\lambda) & \\ 0 & & & \ddots \end{pmatrix}.$$

Матриця оператора $\mathcal{A}$ має потрібну форму в тому ж базисі.

Нехай тепер число k попарно різних коренів многочлена $\mathcal{X}(t)$ більше одиниці. Многочлени $\varphi_1(t) = (t - \lambda_1)^{n_1}$, $\varphi_2(t) = (t - \lambda_2)^{n_2}, \ldots, \varphi_k(t) = (t - \lambda_k)^{n_k}$ попарно взаємно прості. Тому простір розпадається в пряму суму $P^n = L_1 \oplus L_2 \oplus \ldots \oplus L_k$, в якій доданки L_j мають властивості 1) і 2).

Оберемо в лінійному підпросторі L_j такий базис C_j, для якого

$$\mathcal{A} \underset{C_j}{\leftrightarrow} B_j = \begin{pmatrix} I_a(\lambda_j) & & & 0 \\ & I_b(\lambda_j) & & \\ & & \ddots & \\ 0 & & & \ddots \end{pmatrix}.$$

Об'єднання C всіх обраних базисів $C_1, C_2, \ldots, C_k$ відповідних підпросторів $L_1, L_2, \ldots, L_k$ становить базис усього простору P^n і відносно його

$$\mathcal{A} \underset{C}{\leftrightarrow} B_j = \begin{pmatrix} B_1 & & & 0 \\ & B_2 & & \\ & & \ddots & \\ 0 & & & B_k \end{pmatrix}.$$

Матриця B має нормальну форму Жордана.

Теорема 2. *Нехай елементи квадратної матриці A належать полю P. Якщо всі її характеристичні числа належать тому самому полю, то A подібна матриці в нормальній формі Жордана, причому перетворення подібності здійснюється матрицею F з елементами з поля P. Зокрема, матриця з елементами з поля комплексних чисел подібна матриці в нормальній формі Жордана.*

Доведення. Якщо n – порядок матриці A, тоді в просторі P^n оберемо який-небудь базис Б. У цьому базисі матриця A визначає деякий лінійний оператор $\mathcal{A}$. За теоремою 1 існує новий базис C, в якому оператору $\mathcal{A}$ відповідає матриця B, яка має форму Жордана. Якщо F – матриця переходу від Б до C, тоді $F^{-1} \cdot A \cdot F = B$.

Означення. *Базис векторного простору, в якому матриця лінійного оператора $\mathcal{A}$ має жорданову форму, називається канонічним базисом Жордана цього оператора.*

Вправи

1. Побудувати канонічний базис і знайти жорданову форму таких матриць:

а) $A=\begin{pmatrix} 3 & 0 & -1 \\ -2 & 1 & 1 \\ 3 & -1 & -1 \end{pmatrix}$. Відповідь: $B=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

б) $A=\begin{pmatrix} -3 & 4 & 3 & 15 \\ -1 & 1 & 0 & 5 \\ 0 & 0 & -3 & -3 \\ 0 & 0 & 2 & 2 \end{pmatrix}$. Відповідь: $B=\left(\begin{array}{ccc|c} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$.

2. Знайти жорданову форму оператора $\mathcal{A}^2$, якщо відома жорданова форма лінійного оператора $\mathcal{A}$.

Вказівка. Жорданову форму оператора $\mathcal{A}^2$ можна дістати із жорданової форми оператора $\mathcal{A}$ так: у кожній клітині, яка має відношення до $\lambda \neq 0$, поміняти λ на λ^2; кожну клітину порядку $k \geq 1$, яка має відношення до $\lambda = 0$, замінити двома клітинами порядку l, якщо $k = 2l$, і двома клітинами порядків відповідно l і $l + 1$, якщо $k = 2l + 1$.

Розділ 7. ЕВКЛІДОВІ ПРОСТОРИ

§ 7.1. Основні означення

Означення. *Нехай V – лінійний простір над полем R дійсних чисел. Припустимо, що кожному вектору $x \in V$ поставлено у відповідність дійсне число $\|x\|$ з такими властивостями:*

1. $\|x\| \geq 0$ і $\|x\| = 0 \Leftrightarrow x = \theta$.
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$ *для будь-якого* $\lambda \in R$.
3. $\|x + y\| \leq \|x\| + \|y\|$; $x, y \in V$ (*нерівність трикутника*).

Число з такими властивостями називається нормою вектора.

Дійсний лінійний простір V, для векторів якого за будь-яким правилом або законом визначена норма, називається *нормованим*. Зазначимо, що аналогічно вводиться поняття норми і для комплексних лінійних просторів

Наведемо два прості приклади нормованих просторів.

1. В арифметичному дійсному лінійному просторі R^n для вектора $x = (\xi_1, \xi_2, ..., \xi_n)$ покладемо

$$\|x\| = |\xi_1| + |\xi_2| + \ldots + |\xi_n|.$$

Легко переконатися, що цей закон визначає норму вектора $x \in R^n$, і R^n стає нормованим простором.

2. У лінійному просторі $C_{[a,b]}$ дійсних неперервних функцій $f(t)$, заданих на відрізку $[a, b]$, покладемо

$$\|f\| = \max_{a \leq t \leq b} |f(t)|.$$

І в цьому просторі функціонал $\|f\|$ має всі властивості норми.

Означення. *Нехай у дійсному лінійному просторі V кожній парі векторів x, y однозначно поставлено у відповідність дійсне число $\zeta = (x, y)$. Така вектор-функція визначає скалярний добуток векторів x і y, якщо вона має такі властивості:*

1. $(\lambda x+\mu y,z)=\lambda(x,z)+\mu(y,z)$; $x,y,z\in V$; $\lambda,\mu\in R$.

2. $(x,y)=(y,x)$.

3. $x\neq\theta \Rightarrow (x,x)>0$ i $(\theta,\theta)=0$.

Із властивостей 1 і 2 випливає, що $(\theta,y)=(y,\theta)=0$ для будь-якого $y\in V$.

Лінійний простір V, в якому визначено скалярний добуток векторів, називається *евклідовим простором.*

Приклади евклідових просторів.

1. В арифметичному просторі R^n для векторів $x=(\xi_1,\xi_2,...,\xi_n)$ і $y=(\eta_1,\eta_2,...,\eta_n)$ покладемо

$$(x,y)=\xi_1\eta_1+\xi_2\eta_2+...+\xi_n\eta_n.$$

Визначена за таким правилом вектор-функція (x, y) є скалярним добутком. А з введеним таким чином скалярним добутком простір R^n стає евклідовим.

2. Розглянемо квадратичну форму

$$f(x,x)=\sum_{i,k=1}^{n}\alpha_{ik}\xi_i\xi_k,$$

визначену симетричною матрицею $A=(\alpha_{ik})$ порядку n. Тут $x=(\xi_1,\xi_2,...,\xi_n)$ – довільний вектор арифметичного простору R^n.

Припустимо, що квадратична форма $f(x)$ додатно визначена (умови додатної визначеності наводяться в § 5.4, теорема Сильвестра).

Складемо поляру

$$\hat{f}(x,y)=\sum_{i,k=1}^{n}\alpha_{ik}\xi_i\eta_k,$$

де $y=(\eta_1,\eta_2,...,\eta_n)$ – також змінний вектор із R^n.

Вектор-функція $\hat{f}(x,y)$ має всі властивості скалярного добутку. Отже, на лінійному просторі R^n структура евклідового простору може бути введена багатьма способами – залежно від вибору матриці $A=(\alpha_{ik})$.

3. У лінійному просторі $V = C_{[a,b]}$ неперервних функцій для $f(t), g(t) \in V$ покладемо

$$(f,g) = \int_a^b f(t)g(t)dt .$$

Властивості 1, 2, 3, характерні для скалярного добутку, для функціонала (f,g) є простими наслідками відомих властивостей означеного інтеграла від неперервних функцій.

***Теорема 1* (*нерівність Коші – Буняковського*).** *В евклідовому просторі V для скалярного добутку* (x, y) *векторів x, y виконується нерівність*

$$|(x,y)| \le \sqrt{(x,x)} \cdot \sqrt{(y,y)} ,$$

причому рівність досягається тоді і лише тоді, коли вектори x та y лінійно залежні.

Доведення. Якщо $x = \theta$ або $y = \theta$, твердження теореми очевидне. Припустимо, що $x \ne \theta$ і $y \ne \theta$. При цьому $(x, x) > 0$ і $(y, y) > 0$. Розглянемо функцію

$$F(\lambda) = (x + \lambda y, x + \lambda y) .$$

Із властивостей скалярного добутку випливає

$$F(\lambda) = (x,x) + 2\lambda(x,y) + \lambda^2(y,y) \ge 0$$

для будь-яких λ. Тому дискримінант тричлена $F(\lambda)$ не більше нуля: $D = (x,y)^2 - (x,x)(y,y) \le 0$ і отже,

$$|(x,y)| \le \sqrt{(x,x)} \cdot \sqrt{(y,y)} .$$

$F(\lambda)$ обертається на нуль при деякому дійсному значенні λ_0 тоді і лише тоді, коли $x + \lambda_0 y = \theta$, тобто коли x і y лінійно залежні. Таке число λ_0 для $F(\lambda)$ існує і єдине тоді й лише тоді, коли $D = (x,y)^2 - (x,x)(y,y) = 0$, тобто коли $|(x,y)| = \sqrt{(x,x)} \cdot \sqrt{(y,y)}$.

Теорема 2. *Евклідовий простір V є нормований, якщо для будь-якого його вектора x функцію $\|x\|$ визначити за правилом*

$$\|x\| = \sqrt{(x,x)}\,.$$

Доведення. Переконаємось, що функція $\|x\|$ має всі властивості норми. Виконання властивості 1 для $\|x\|$ випливає з властивості 3 скалярного добутку.

Доведемо властивість 2. Маємо

$$\|\lambda x\| = \sqrt{(\lambda x, \lambda x)} = \sqrt{\lambda^2 (x,x)} = \sqrt{|\lambda|^2 (x,x)} =$$
$$= |\lambda| \cdot \sqrt{(x,x)} = |\lambda| \|x\|.$$

Доведемо властивість 3. Маємо систему рівностей

$$\|x+y\|^2 = (x+y, x+y) = (x,x) + 2(x,y) + (y,y) =$$
$$= \|x\|^2 + 2(x,y) + \|y\|^2\,.$$

На підставі нерівності Коші – Буняковського $|(x,y)| \le \|x\|\|y\|$ дістаємо

$$\|x+y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 =$$
$$= (\|x\| + \|y\|)^2 \Rightarrow \|x+y\| \le \|x\| + \|y\|.$$

Наведена норма в просторі V називається *евклідовою* нормою вектора x (*довжиною* вектора x).

Зауваження. У нерівності трикутника $\|x+y\| \le \|x\| + \|y\|$ для евклідової норми рівність досягається тоді й лише тоді, коли вектори x і y лінійно залежні. Це випливає з відповідного зауваження щодо нерівності Коші – Буняковського в теоремі 1.

Для ненульових векторів x і y на підставі нерівності Коші – Буняковського матимемо

$$-1 \le \frac{(x,y)}{\|x\| \cdot \|y\|} \le 1\,.$$

Тому існує такий кут φ, $0 \le \varphi \le \pi$, для якого

$$\cos\varphi = \frac{(x, y)}{\|x\| \cdot \|y\|}.$$

Його називають кутом між цими векторами. Вектори x, y називаються *ортогональними*, якщо їх скалярний добуток $(x, y) = 0$, тобто кут φ між ними має дорівнювати $\pi/2$ за умови $x \ne \theta$ і $y \ne \theta$.

Означення. *Система векторів $x_1, x_2, ..., x_m$ евклідового простору називається ортогональною, якщо кожна пара різних її векторів ортогональна: $(x_i, x_k) = 0$, якщо $i \ne k$.*

Теорема 3. *Ортогональна система ненульових векторів лінійно незалежна.*

Доведення. Нехай $x_1, x_2, ..., x_m$ – ортогональна система ненульових векторів. Припустимо, що має місце рівність

$$\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m = \theta$$

зі скалярними коефіцієнтами $\lambda_1, \lambda_2, ..., \lambda_m$. Складемо скалярний добуток правої та лівої частин останньої рівності на x_1. Дістанемо

$$0 = (\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m, x_1) =$$
$$= \lambda_1 (x_1, x_1) + \lambda_2 (x_2, x_1) + ... + \lambda_m (x_m, x_1).$$

Оскільки $(x_2, x_1) = 0$, $(x_3, x_1) = 0, ..., (x_m, x_1) = 0$, то $\lambda_1 (x_1, x_1) = 0$. За умовою $(x_1, x_1) \ne 0$. Тому $\lambda_1 = 0$. Аналогічно дістаємо $\lambda_2 = 0$, $\lambda_3 = 0$, $..., \lambda_m = 0$. Це й доводить лінійну незалежність векторів $x_1, x_2, ..., x_m$.

Означення. *Нехай M – непорожня підмножина евклідового простору V. Вектор y називається ортогональним до підмножини M, якщо він ортогональний до кожного вектора x із M*

Позначимо це відношення y до M так: $y \perp M$. Ортогональним доповненням $M^{\perp}$ підмножини M називається множина всіх тих векторів $y \in V$, які ортогональні M:

$$M^{\perp} = \{y \in V : y \perp M\} =$$
$$= \{y \in V : \forall x[(y, x) = 0,\ x \in M]\}.$$

Теорема 4. *Для будь-якої непорожньої підмножини M евклідового простору ортогональне доповнення його $M^{\perp}$ – лінійний підпростір. Якщо для двох підмножин M, L маємо $M \subseteq L$, то $L^{\perp} \subseteq M^{\perp}$.*

Доведемо перше твердження. Якщо $y, z \in M^{\perp}$ і $\lambda, \mu \in R$, тоді для довільного вектора $x \in M$ маємо $(\lambda y + \mu z, x) = = \lambda(y, x) + \mu(z, x) = 0$. Це означає, що $\lambda y + \mu z \in M^{\perp}$.

Доведення другого твердження – очевидне.

Вправи

1. Довести правило паралелограма:

$$\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

2. Довести, що коли система векторів $x_1, x_2, ..., x_m$ ортогональна, тоді

$$\|x_1 + x_2 + ... + x_m\|^2 = \|x_1\|^2 + \|x_2\|^2 + ... + \|x_m\|^2$$

(узагальнена теорема Піфагора).

3. Використавши узагальнену теорему Піфагора, довести лінійну незалежність ортогональної системи ненульових векторів.

4. Показати, що в лінійному просторі R_∞ усіх многочленів з дійсними коефіцієнтами (як і в просторі R_{n+1} многочленів степеня $\leq n$, де n – фіксоване натуральне число) скалярний добуток многочленів $f(t) = \alpha_0 + \alpha_1 t + ... + \alpha_p t^p$ і $g(t) = \beta_0 + \beta_1 t + ... + \beta_p t^p$ можна визначити за правилом

$$(f, g) = \alpha_0\beta_0 + \alpha_1\beta_1 + ... + \alpha_p\beta_p.$$

§ 7.2. Ортогональне проектування

Починаючи з цього параграфа, розглянемо скінченновимірні евклідові простори. Оскільки кожний дійсний лінійний простір розмірності n ізоморфний арифметичному простору R^n, то й обмежимось вивченням евклідового простору R^n.

Нехай $x_1, x_2, ..., x_m$ $(m \le n)$ – ряд векторів евклідового простору. Визначник

$$\Gamma(x_1, x_2, ..., x_m) = \begin{vmatrix} (x_1, x_1) & (x_1, x_2) & \dots & (x_1, x_m) \\ (x_2, x_1) & (x_2, x_2) & \dots & (x_2, x_m) \\ \dots & \dots & \dots & \dots \\ (x_m, x_1) & (x_m, x_2) & \dots & (x_m, x_m) \end{vmatrix}$$

називається *визначником Грама* для цього ряда векторів.

Теорема 1. *Вектори $x_1, x_2, ..., x_m$ лінійно залежні тоді й лише тоді, коли визначник Грама $\Gamma(x_1, x_2, ..., x_m)$ дорівнює нулю.*

Доведення. *Необхідність.* Нехай вектори $x_1, x_2, ..., x_m$ лінійно залежні. Для них існує ненульова послідовність чисел $\lambda_1, \lambda_2, ..., \lambda_m$ така, що $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m = \theta$. Тоді

$$(x_i, \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m) =$$
$$= \lambda_1 (x_i, x_1) + \lambda_2 (x_i, x_2) + ... + \lambda_m (x_i, x_m) = 0, \; i = \overline{1, m}.$$

Це означає, що система лінійних однорідних рівнянь

$$\begin{cases} (x_1, x_1)\xi_1 + (x_1, x_2)\xi_2 + ... + (x_1, x_m)\xi_m = 0, \\ (x_2, x_1)\xi_1 + (x_2, x_2)\xi_2 + ... + (x_2, x_m)\xi_m = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ (x_m, x_1)\xi_1 + (x_m, x_2)\xi_2 + ... + (x_m, x_m)\xi_m = 0. \end{cases} \quad (1)$$

має ненульовий розв'язок $\lambda_1, \lambda_2, ..., \lambda_m$. Тому визначник $\Gamma(x_1, x_2, ..., x_m)$ цієї системи дорівнює нулю.

Достатність. Нехай $\Gamma(x_1, x_2, ..., x_m) = 0$. Тоді система рівнянь (1) має ненульові розв'язки. Оберемо деякий ненульовий розв'язок $\lambda_1, \lambda_2, ..., \lambda_m$ і розглянемо вектор

$$x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m .$$

Оскільки мають місце рівності

$$(x_i, x_1)\lambda_1 + (x_i, x_2)\lambda_2 + ... + (x_i, x_m)\lambda_m = 0,$$

то $(x_i, x) = 0,\ i = \overline{1,m}.$

Помножимо i-ту рівність на λ_i, а потім просумуємо по i. Дістанемо

$$\lambda_1 (x_1, x) + \lambda_2 (x_2, x) + ... + \lambda_m (x_m, x) = (x, x) = 0.$$

Звідси $x = \theta$, тобто $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m = \theta$. Отже, вектори $x_1, x_2, ..., x_m$ – лінійно залежні.

Наслідок. *Вектори $x_1, x_2, ..., x_m$ лінійно незалежні тоді й лише тоді, коли визначник Грама $\Gamma(x_1, x_2, ..., x_m) \neq 0$.*

Теорема 2 (про ортогональне проектування). *Нехай M – лінійний підпростір евклідового простору R^n. Довільний вектор $x \in R^n$ подається у вигляді суми*

$$x = \hat{x} + \overset{\perp}{x},$$

де доданок $\hat{x}$ належить M, а доданок $\overset{\perp}{x}$ – ортогональний M Таке подання вектора єдине. Доданок $\hat{x}$ називається ортогональною проекцією x на підпростір M, а доданок $\overset{\perp}{x}$ – ортогональною складовою.

Доведення. Нехай $m = \dim M$ і $x_1, x_2, ..., x_m$ – базис підпростору *M*. Знаходимо $\hat{x}$ у вигляді лінійної комбінації $\hat{x} = \xi_1 x_1 + ... + \xi_m x_m$. Оскільки $\overset{\perp}{x} = x - \hat{x}$ ортогональний *M*, тоді $(x_i, x - \hat{x}) = 0$ для будь-яких $i = \overline{1,m}$. Дістаємо систему рівнянь

$$(x_i, x) - \xi_1 (x_i, x_1) - \xi_2 (x_i, x_2) - \ldots - \xi_m (x_i, x_m) = 0,\ i = \overline{1,m},$$

або у вигляді системи

$$\begin{cases} (x_1,x_1)\xi_1+(x_1,x_2)\xi_2+ & \dots & +(x_1,x_m)\xi_m=(x_1,x), \\ (x_2,x_1)\xi_1+(x_2,x_2)\xi_2+ & \dots & +(x_2,x_m)\xi_m=(x_2,x), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & \\ (x_m,x_1)\xi_1+(x_m,x_2)\xi_2+ & \dots & +(x_m,x_m)\xi_m=(x_m,x). \end{cases} \tag{2}$$

Визначник цієї системи $\Gamma(x_1,x_2,...,x_m)\neq 0$, отже система (2) має лише один розв'язок $\xi_1,\xi_2,...,\xi_m$. Цей розв'язок і визначає доданок $\hat{x}$.

Наслідок. *Для будь-якого підпростору M евклідового простору R^n виконується рівність*

$$R^n=M\oplus M^{\perp}. \tag{3}$$

Приклад 1. У просторі R^4 скалярний добуток векторів $x=(\xi_1,\xi_2,\xi_3,\xi_4)$ і $y=(\eta_1,\eta_2,\eta_3,\eta_4)$ задано правилом $(x,y)=\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3+\xi_4\eta_4$. Нехай M – лінійна оболонка векторів $a_1=(1,\ 1,\ 1,\ 1)$, $a_2=(1,\ 2,\ 2,-1)$, $a_3=(1,\ 0,\ 0,\ 3)$. Знайти ортогональну проекцію $\hat{x}$ і ортогональну складову $\overset{\perp}{x}$ вектора $x=(4,\ -1,\ -3,\ 4)$ на підпростір M.

Розв'язання. Ранг системи векторів a_1,a_2,a_3 дорівнює 2; вектори a_1,a_2 лінійно незалежні і становлять базис M.

Знаходимо проекцію $\hat{x}$ у вигляді $\hat{x}=\xi_1a_1+\xi_2a_2$. Для невідомих ξ_1,ξ_2 матимемо систему

$$\begin{cases} (a_1,a_1)\xi_1+(a_1,a_2)\xi_2=(a_1,x), \\ (a_2,a_1)\xi_1+(a_2,a_2)\xi_2=(a_2,x). \end{cases}$$

Обчислюємо коефіцієнти $(a_1,a_1)=4$, $(a_1,a_2)=(a_2,a_1)=4$, $(a_2,a_2)=10$, $(a_1,x)=4$, $(a_2,x)=-8$. Система

$$\begin{cases} 4\xi_1+4\xi_2=4, \\ 4\xi_1+10\xi_2=-8 \end{cases}$$

має розв'язки $\xi_1=3,\xi_2=-2$. Тоді $\hat{x}=3a_1-2a_2=(1,\ -1,\ -1,\ 5)$, $\overset{\perp}{x}=x-\hat{x}=(3,\ 0,\ -2,\ -1)$. Перевіркою переконаємось, що одержаний вектор $\overset{\perp}{x}$ перпендикулярний до кожного із заданих векторів a_1, a_2, a_3.

Доведемо формули для знаходження розв'язку системи рівнянь (2) у загальному вигляді.

Для зручності позначимо стовпці розширеної матриці цієї системи так:

$$X_1 = \begin{pmatrix} (x_1, x_1) \\ (x_2, x_1) \\ \vdots \\ (x_m, x_1) \end{pmatrix}, \; X_2 = \begin{pmatrix} (x_1, x_2) \\ (x_2, x_2) \\ \vdots \\ (x_m, x_2) \end{pmatrix}, \ldots,$$

$$X_m = \begin{pmatrix} (x_1, x_m) \\ (x_2, x_m) \\ \vdots \\ (x_m, x_m) \end{pmatrix}, \; X = \begin{pmatrix} (x_1, x) \\ (x_2, x) \\ \vdots \\ (x_m, x) \end{pmatrix}.$$

Тоді стосовно до системи (2) формули Крамера подамо у вигляді

$$\xi_1 = \frac{1}{\Gamma_m} \det\left(X, X_2, X_3, \ldots, X_m\right),$$

$$\xi_2 = \frac{1}{\Gamma_m} \det\left(X_1, X, X_3, \ldots, X_m\right),$$

....

$$\xi_m = \frac{1}{\Gamma_m} \det\,\left(X_1, X_2, \ldots, X_{m-1}, X\right),$$

де $\Gamma_m = \Gamma\left(x_1, x_2, \ldots, x_m\right)$.

Перестановкою стовпців визначників перетворимо ці формули:

$$\xi_1 = \frac{1}{\Gamma_m} \det\left(X, X_2, X_3, \ldots, X_m\right),$$

$$\xi_2 = -\frac{1}{\Gamma_m} \det\,\left(X, X_1, X_3, \ldots, X_m\right),$$

$$\xi_3 = \frac{1}{\Gamma_m}\det\left(X, X_1, X_2, X_4, \ldots, X_m\right),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\xi_m = \frac{(-1)^{m-1}}{\Gamma_m}\det\left(X, X_1, X_2, \ldots, X_{m-1}\right).$$

Нехай $h = \left\| \overset{\perp}{x} \right\|$. Тоді

$$h^2 = \left(\overset{\perp}{x}, \overset{\perp}{x}\right) = \left(\overset{\perp}{x}, x\right) = \left(x - \hat{x}, x\right) = (x, x) - (\hat{x}, x) = (x, x) -$$

$$-(x_1, x)\xi_1 - (x_2, x)\xi_2 - \ldots - (x_m, x)\xi_m =$$

$$= \frac{1}{\Gamma_m}\left\{\begin{array}{l} (x,x)\Gamma_m - (x_1,x)\det\left(X, X_2, X_3, \ldots, X_m\right) + \\ +(x_2,x)\det\left(X, X_1, X_3, \ldots, X_m\right) - \\ -(x_3,x)\det\left(X, X_1, X_2, X_4, \ldots, X_m\right) + \ldots + \\ +(-1)^{m-1}(x_m,x)\det\left(X, X_1, X_2, \ldots, X_{m-1}\right) \end{array}\right\}.$$

Якщо визначник

$$\Delta_{m+1} = \begin{vmatrix} (x,x) & (x_1,x) & (x_2,x) & \ldots & (x_m,x) \\ (x_1,x) & (x_1,x_1) & (x_1,x_2) & \ldots & (x_1,x_m) \\ (x_2,x) & (x_2,x_1) & (x_2,x_2) & \ldots & (x_2,x_m) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (x_m,x) & (x_m,x_1) & (x_m,x_2) & \ldots & (x_m,x_m) \end{vmatrix}$$

розкласти за елементами першого рядка, дістанемо суму, що міститься в фігурних дужках попередньої рівності. Отже, дістали формулу

$$h^2 = \left(\overset{\perp}{x}, \overset{\perp}{x}\right) = \frac{\Delta_{m+1}}{\Gamma_m}. \qquad (4)$$

Перетворимо визначник Δ_{m+1}. Перший його рядок переставляємо послідовно з кожним із рядків доти, доки він не займе місце останнього рядка. Перший стовпець переставляємо послідовно з кожним стовпцем доти, доки він не займе місце останнього стовпця. Після такого перетворення визначник не зміниться, хоча й набуде вигляду

$$\Delta_{m+1}=\begin{vmatrix}(x_1,x_1) & (x_1,x_2) & \dots & (x_1,x_m) & (x_1,x)\\ (x_2,x_1) & (x_2,x_2) & \dots & (x_2,x_m) & (x_2,x)\\ \dots & \dots & \dots & \dots & \dots\\ (x_m,x_1) & (x_m,x_2) & \dots & (x_m,x_m) & (x_m,x)\\ (x,x_1) & (x,x_2) & \dots & (x,x_m) & (x,x)\end{vmatrix}=$$

$$=\Gamma(x_1,x_2,\dots,x_m,x).$$

Остаточно для $h=\left\|\overset{\perp}{x}\right\|$ дістали

$$h^2=\left(\overset{\perp}{x},\overset{\perp}{x}\right)=\frac{\Gamma(x_1,x_2,\dots,x_m,x)}{\Gamma(x_1,x_2,\dots,x_m)}. \tag{5}$$

Теорема 3. *Якщо вектори* $x_1,x_2,\dots,x_m$ *лінійно незалежні, тоді*

$$\Gamma(x_1,x_2,\dots,x_m)>0\,.$$

Доведення здійснимо за допомогою індукції за числом m векторів. Якщо $m=1$, тоді $\Gamma(x_1)=(x_1,x_1)>0$. Припустимо, що твердження теореми має місце для $m-1$ лінійно незалежних векторів.

Якщо $L=L(x_1,\dots,x_{m-1})$ – лінійна оболонка векторів $x_1,x_2,\dots,x_{m-1}$, то $x_m\notin L$. Тоді подамо x_m у вигляді суми $x_m=\hat{x}_m+\overset{\perp}{x}_m$, де $\hat{x}_m$ – ортогональна проекція на L і $\overset{\perp}{x}_m$ – ортогональна складова до L. Оскільки $\overset{\perp}{x}_m\neq\theta$, тоді $h_m=\left\|\overset{\perp}{x}_m\right\|>0$.

Застосувавши формулу (5), матимемо

$$h_m^2=\left(\overset{\perp}{x}_m,\overset{\perp}{x}_m\right)=\frac{\Gamma(x_1,x_2,\dots,x_{m-1},x_m)}{\Gamma(x_1,x_2,\dots,x_{m-1})}>0\,. \tag{6}$$

За припущенням індукції $\Gamma(x_1,x_2,\dots,x_{m-1})>0$. Тому $\Gamma(x_1,\dots,x_{m-1},x_m)>0$.

Геометрична інтерпретація визначника Грама

Якщо $x_1, x_2, ..., x_m$ *– лінійно незалежні вектори, то m-вимірним паралелепіпедом* Π_m *простору* R^n, *побудованим на цих векторах, називається множина всіх точок* $x \in R^n$ *таких, для яких має місце лінійна комбінація*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m,\ 0 \le \lambda_j \le 1,\ j = \overline{1, m}.$$

Об'єм V_m паралелепіпеда Π_m визначається індукцією за m. Для будь-якого k $(1 \le k \le m)$ позначимо через $L_k = L(x_1, x_2, ..., x_k)$ – лінійну оболонку векторів $x_1, x_2, ..., x_k : \Gamma_k = \Gamma(x_1, x_2, ..., x_k)$. На підставі формули (6) матимемо $\frac{\Gamma_k}{\Gamma_{k-1}} = h_k^2$, де $h_k = \left\| \overset{\perp}{x}_k \right\|$ і $\overset{\perp}{x}_k$ – ортогональна складова x_k відносно підпростору L_{k-1} ($k > 1$). Тоді $V_1 = \sqrt{(x_1, x_1)} = \sqrt{\Gamma_1}$ – довжина вектора x_1; $V_2 = V_1 \cdot h_2$, де $h_2 = \left\| \overset{\perp}{x}_2 \right\|$ – висота паралелограма Π_2, побудованого на x_1, x_2. Дістаємо $V_2 = \sqrt{\Gamma_2}$; $V_3 = V_2 \cdot h_3$, де $h_3 = \left\| \overset{\perp}{x}_3 \right\|$ – висота паралелепіпеда з основою Π_2. Отже, $V_3 = \sqrt{\Gamma_3}$.

Продовжуючи наші міркування, нарешті знайдемо $V_m = V_{m-1} \cdot h_m$, де $h_m = \left\| \overset{\perp}{x}_m \right\|$ – висота паралелепіпеда Π_m з основою Π_{m-1}. Тому $V_m = \sqrt{\Gamma_m}$.

Приклад 2. Нехай у просторі R^n скалярний добуток векторів $x = (\xi_1, \xi_2, ..., \xi_n)$ і $y = (\eta_1, \eta_2, ..., \eta_n)$ визначено правилом

$$(x, y) = \sum_{k=1}^{n} \xi_k \eta_k .$$

Для системи лінійно незалежних векторів

$$x_k = (\xi_{1k}, \xi_{2k}, ..., \xi_{nk}),\ k = \overline{1, n}$$

складемо матрицю

$$X = \begin{pmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n} \\ \hdotsfor{4} \\ \xi_{n1} & \xi_{n2} & \cdots & \xi_{nn} \end{pmatrix}.$$

Тоді $|X^T \cdot X| = |X^T| \cdot |X| = |X|^2$. Разом з тим $|X^T \cdot X| = \Gamma(x_1, x_2, ..., x_n)$ – визначник Грама векторів $x_1, x_2, ..., x_n$.

Об'єм паралелепіпеда, утвореного цими векторами, дорівнює $V_n = \sqrt{\Gamma(x_1, x_2, ..., x_n)}$ – модулю визначника матриці.

Отже,

$$V_n = \operatorname{mod} \begin{vmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n} \\ \hdotsfor{4} \\ \xi_{n1} & \xi_{n2} & \cdots & \xi_{nn} \end{vmatrix}.$$

Одержана формула – узагальнення відомих формул для обчислення площі паралелограма та об'єму паралелепіпеда відповідно в просторах R^2 і R^3.

Вправи

1. Підпростір M задано системою лінійних рівнянь

$$\begin{cases} 2\xi_1 + \xi_2 + \xi_3 + 3\xi_4 = 0, \\ 3\xi_1 + 2\xi_2 + 2\xi_3 + \xi_4 = 0, \\ \xi_1 + 2\xi_2 + 2\xi_3 - 9\xi_4 = 0. \end{cases}$$

Знайти ортогональну проекцію та ортогональну складову вектора $X = (7, -4, -1, 2)$ відносно M.

2. Нехай M – лінійний підпростір в R^n. *Відстанню від точки, заданої вектором x, до M* називається мінімум довжин векторів $\|x - u\|$, де u пробігає M. Довести, що відстань d дорівнює довжині ортогональної складової $\overset{\perp}{x}$ вектора x відносно M.

3. Нехай $L = a + M$ – лінійний багатовид з базисним підпростором M. Відстанню від точки, заданої вектором x, до L називається мінімум довжин векторів $\|x - u\|$, де u пробігає L.

Довести, що відстань d дорівнює довжині ортогональної складової вектора $y = x - a$ відносно підпростору M.

4. Знайти відстань точки, заданої вектором $x = (4, 2, -5, 1)$ до лінійного багатовиду, що визначається системою рівнянь

$$\begin{cases} 2x_1 - 2x_2 + x_3 + 2x_4 = 9, \\ 2x_1 - 4x_2 + 2x_3 + 3x_4 = 12. \end{cases}$$

5. Нехай в евклідовому просторі задано лінійний підпростір M і вектор x. *Кутом між вектором x і підпростором M* називається найменший з кутів, який утворює x з векторами із M. Довести, що кут між x і M дорівнює куту x і його ортогональній проекції $\hat{x}$ на M.

Довести, що серед векторів підпростору M той самий кут з вектором x утворюють вектори виду $\lambda \cdot \hat{x}$, $\lambda > 0$, і лише вони.

6. Знайти кут між вектором $x = (-3, 15, 1, -5)$ і підпростором $M = L(a_1, a_2, a_3)$, де $a_1 = (2, 3, -4, -6)$, $a_2 = (1, 8, -2, -16)$, $a_3 = (1, -5, -2, 10)$.

7. Нехай L і M – лінійні підпростори. Довести:

а) $\left(M^{\perp}\right)^{\perp} = M$ (теорема двоїстості);

б) $(L + M)^{\perp} = L^{\perp} \cap M^{\perp}$;

в) $(L \cap M)^{\perp} = L^{\perp} + M^{\perp}$.

8. Довести нерівність $\Gamma(x_1, x_2, ..., x_m) \le \Gamma(x_1)\Gamma(x_2) \cdots \Gamma(x_m)$ (*нерівність Адамара*).

Вказівка. Довести для лінійно незалежних векторів, використовуючи формулу (6).

§ 7.3. Ортогоналізація системи векторів. Ортонормований базис

Означення. *Упорядковані системи векторів*

$$\text{А} = \langle a_1, a_2, ..., a_m \rangle; \quad \text{Б} = \langle b_1, b_2, ..., b_m \rangle,$$

складені з однакового числа m векторів, називаємо еквівалентними (*позначаємо* А ≈ Б), *якщо для кожного* $p = \overline{1, m}$

підсистеми $\langle a_1, a_2, ..., a_p \rangle$ і $\langle b_1, b_2, ..., b_p \rangle$ *лінійно еквівалентні, тобто їх лінійні оболонки збігаються*:

$$L(a_1, a_2, ..., a_p) = L(b_1, b_2, ..., b_p).$$

Означення. *Ортогоналізацією упорядкованої лінійно-незалежної системи* $\text{А} = \langle a_1, a_2, ..., a_m \rangle$ *векторів називається заміна її на нову упорядковану систему* $\text{Б} = \langle b_1, b_2, ..., b_m \rangle$ *векторів з такими властивостями:*

1) $\text{А} \approx \text{Б}$;

2) *система* Б – *ортогональна.*

Теорема 1 (про ортогоналізацію). *Будь-яку упорядковану систему* $\text{А} = \langle a_1, a_2, ..., a_m \rangle$ *можна ортогоналізувати. Будь-які способи ортогоналізації лінійно незалежної системи векторів зводять її до єдиної нової системи ненульових векторів з точністю до скалярних множників.*

Доведення. Будуємо нову систему $\text{Б} = \langle b_1, b_2, ..., b_m \rangle$ векторів з потрібними властивостями послідовно, починаючи з вектора b_1.

Покладемо $b_1 = a_1$. Припустимо, що $p > 1$ і вже побудовано ортогональну систему $\langle b_1, b_2, ..., b_{p-1} \rangle$ векторів, еквівалентну підсистемі $\langle a_1, a_2, ..., a_{p-1} \rangle$ із $\text{А} = \langle a_1, a_2, ..., a_m \rangle$. При цьому рівні їх лінійні оболонки:

$$L(a_1, a_2, ..., a_{p-1}) = L(b_1, b_2, ..., b_{p-1}) = M.$$

Побудуємо вектор b_p. Для цього розкладемо вектор a_p на суму $a_p = \hat{a}_p + \overset{\perp}{a}_p$, де $\hat{a}_p$ – його ортогональна проекція на M і $\overset{\perp}{a}_p$ – ортогональна складова $\left(\overset{\perp}{a}_p \perp M\right)$. Покладемо $b_p = \overset{\perp}{a}_p = -\hat{a}_p + a_p$. Тут $-\hat{a}_p \in M$ і, отже, $-\hat{a}_p$ лінійно задається векторами $b_1, b_2, ..., b_{p-1}$; $-\hat{a}_p = \beta_1 b_1 + \beta_2 b_2 + ... + \beta_{p-1} b_{p-1}$. Тому b_p подається у вигляді суми

$$b_p = \beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_{p-1} b_{p-1} + a_p .$$

За побудовою система векторів $b_1, b_2, \ldots, b_{p-1}, b_p$ ортогональна. Залишається перевірити рівність $L(b_1, b_2, \ldots, b_p) = L(a_1, a_2, \ldots, a_p)$. Для цього достатньо перевірити включення $b_p \in L(a_1, a_2, \ldots, a_{p-1}, a_p)$ і $a_p \in L(b_1, b_2, \ldots, b_{p-1}, b_p)$. Перше включення: оскільки $b_p = a_p - \hat{a}_p$ і $-\hat{a}_p \in M = L(a_1, a_2, \ldots, a_{p-1})$, тоді $b_p \in L(a_1, \ldots, a_{p-1}, a_p)$. Друге включення: оскільки $a_p = \hat{a}_p + \overset{\perp}{a}_p = \hat{a}_p + b_p$ і $\hat{a}_p \in M = L(b_1, b_2, \ldots, b_{p-1})$, тоді $a_p \in L(b_1, \ldots, b_{p-1}, b_p)$.

Продовжуючи далі таким же чином, на m-му кроці дістанемо потрібну систему векторів.

Припустимо тепер, що система векторів $\text{А} = \langle a_1, a_2, \ldots, a_m \rangle$ лінійно незалежна. Покажемо, що всі вектори одержаної ортогональної системи $\text{Б} = \langle b_1, b_2, \ldots, b_m \rangle$ відмінні від нульового вектора.

Припустимо, що вектор $b_p = \theta$ (зрозуміло, що $p > 1$). Оскільки $b_p = \overset{\perp}{a}_p = -\hat{a}_p + a_p = \theta$, то $a_p = \hat{a}_p \in M$. Вектор a_p лінійно залежить від $a_1, a_2, \ldots, a_{p-1}$. Дістали суперечність.

Отже, $b_1 \neq \theta,\ b_2 \neq \theta, \ldots, b_m \neq \theta$.

Доведення єдиності. Нехай $\text{С} = \langle c_1, c_2, \ldots, c_n \rangle$ – третя система векторів, яка має ті самі властивості, що й система Б:

1) $\text{А} \approx \text{С}$;

2) С – ортогональна система.

Для будь-якого p, $1 \le p \le m$, маємо рівності

$$L(a_1, \ldots, a_{p-1}, a_p) = L(b_1, \ldots, b_{p-1}, b_p) = L(c_1, \ldots, c_{p-1}, c_p).$$

Тоді $b_p = x + \lambda \cdot c_p$, де $x \in M = L(b_1, b_2, \ldots, b_{p-1}) = L(c_1, c_2, \ldots, c_{p-1})$. Тому b_p подається у вигляді лінійної комбінації $b_p=\lambda_1 b_1+\ldots+\lambda_{p-1}b_{p-1}+\lambda c_p$. Вектори b_p і c_p ортогональні до всіх векторів $b_1, b_2, \ldots, b_{p-1}$. Звідси дістаємо

$$0=(b_p,b_1)=(\lambda_1 b_1+\ldots+\lambda_{p-1}b_{p-1}+\lambda c_p, b_1)=\lambda_1(b_1,b_1).$$

Оскільки $(b_1,b_1)\neq 0$, то $\lambda_1=0$. Аналогічно встановлюються рівності $\lambda_2=0, \ldots, \lambda_{p-1}=0$. Остаточно матимемо рівність $b_p=\lambda\cdot c_p$, $\lambda\neq 0$. Вона правильна для будь-яких $p=\overline{1,m}$.

Викладений при доведенні теореми метод ортогоналізації упорядкованої системи векторів $\mathrm{A}=\langle a_1, a_2, \ldots, a_m\rangle$ називається *процесом ортогоналізації*. Його можна описати у вигляді правила (алгоритма):

А) Обираємо $b_1=a_1$;

Б) Якщо вектори $b_1, b_2, \ldots, b_{p-1}$ побудовано, тоді вектор b_p $(p>1)$ матимемо у вигляді

$$b_p=\beta_1 b_1+\beta_2 b_2+\ldots+\beta_{p-1}b_{p-1}+a_p.$$

Коефіцієнти $\beta_1, \beta_2, \ldots, \beta_{p-1}$ обчислюються за умов

$$(b_p,b_1)=0,\ (b_p,b_2)=0, \ldots, (b_p,b_{p-1})=0.$$

Правило можна де в чому поміняти:

А) Обираємо $b_1=\beta a_1$ $(\beta\neq 0)$;

Б) Якщо вектори $b_1, b_2, \ldots, b_{p-1}$ побудовано, тоді вектор b_p $(p>1)$ знаходимо у вигляді

$$b_p=\beta_1 b_1+\beta_2 b_2+\ldots+\beta_{p-1}b_{p-1}+\beta_p a_p\ (\beta_p\neq 0).$$

Коефіцієнти $\beta_1, \beta_2, \ldots, \beta_{p-1}$ обчислюються за умов

$$(b_p,b_1)=0,\ (b_p,b_2)=0, \ldots, (b_p,b_{p-1})=0.$$

Приклад. У лінійному просторі R^4 скалярний добуток векторів $x=(\xi_1,\xi_2,\xi_3,\xi_4)$, $y=(\eta_1,\eta_2,\eta_3,\eta_4)$ визначається законом

$$(x,y)=\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3+\xi_4\eta_4.$$

Ортогоналізувати систему векторів

$$a_1=(1,1,1,1),\ a_2=(3,3,-1,-1),\ a_3=(-2,0,6,8).$$

Розв'язання.

1. Покладемо $b_1=a_1=(1,1,1,1)$.

2. Відшукаємо $b_2=\beta b_1+a_2$ за умови $(b_2,b_1)=0$. Для неозначеного коефіцієнта β матимемо рівняння $\beta(b_1,b_1)+$ $+(a_2,b_1)=0$. Оскільки $(b_1,b_1)=4$, $(a_2,b_1)=4$, то $\beta=-1$. Отже, $b_2=(-1)b_1+a_2=(2,2,-2,-2)$.

3. Відшукаємо $b_3=\beta_1 b_1+\beta_2 b_2+a_3$ за умов $(b_3,b_1)=0$, $(b_3,b_2)=0$. Дістанемо рівняння для коефіцієнтів β_1 і β_2:

$$\beta_1(b_1,b_1)+(a_3,b_1)=0,\ \beta_2(b_2,b_2)+(a_3,b_2)=0.$$

Оскільки $(b_1,b_1)=4$, $(a_3,b_1)=12$, $(b_2,b_2)=16$, $(a_3,b_2)=-32$, то $\beta_1=-3$, $\beta_2=2$. Знаходимо

$$b_3=-3b_1+2b_2+a_3=(-1,1,-1,1).$$

Отже, векторами нової системи будуть

$$b_1=(1,1,1,1),\ b_2=(2,2,-2,-2),\ b_3=(-1,1,-1,1).$$

Зауваження. В ході доведення теореми 1 встановлено, що коли в процесі ортогоналізації системи векторів a_1, $a_2,\ldots,a_m$ на деякому p-му кроці одержується нульовий вектор $(b_p=\theta)$, то це означає, що вектор a_p лінійно виражається через попередні вектори $a_1,a_2,\ldots,a_{p-1}$, а система векторів $a_1,a_2,\ldots,a_m$ лінійно залежна.

Означення. *Базис* $E=\langle e_1,e_2,\ldots,e_n\rangle$ *евклідового простору* R^n *називається ортонормованим, якщо він ортогональний і норми всіх його векторів дорівнюють* 1: $(e_i,e_k)=0$, *якщо* $i\neq k$, *і* $\|e_i\|=1$ *для будь-яких* $i=\overline{1,n}$, *або*

$$(e_i, e_k) = \begin{cases} 1, \textit{ якщо } i = k; \\ 0, \textit{ якщо } i \neq k. \end{cases}$$

Із будь-якого базису $\mathrm{A} = \langle a_1, a_2, ..., a_n \rangle$ простору R^n можна побудувати ортонормований базис таким чином. За допомогою процесу ортогоналізації з базису А отримаємо ортогональний базис $\mathrm{Б} = \langle b_1, b_2, ..., b_n \rangle$. Із останнього базису дістанемо ортонормований базис $\mathrm{E} = \langle e_1, e_2, ..., e_n \rangle$, якщо покладати

$$e_i = \frac{1}{\|b_i\|} \cdot b_i, \; i = \overline{1, n}.$$

Теорема 2. *Нехай* $\mathrm{E} = \langle e_1, e_2, ..., e_n \rangle$ – *ортонормований базис евклідового простору* R^n. *Якщо* $\xi_1, \xi_2, ..., \xi_n$ – *координати вектора* x, *а* $\eta_1, \eta_2, ..., \eta_n$ – *координати вектора* y *в цьому базисі, то їх скалярний добуток* (x, y) *обчислюється за формулою*

$$(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + ... + \xi_n\eta_n. \tag{1}$$

Доведення. Оскільки

$$x = \xi_1 e_1 + \xi_2 e_2 + ... + \xi_n e_n \text{ і } y = \eta_1 e_1 + \eta_2 e_2 + ... + \eta_n e_n,$$

то $(x, y) = \sum_{i,k=1}^{n} \xi_i \eta_k (e_i, e_k)$. Але $(e_i, e_k) = 0$, якщо $i \neq k$, і $(e_i, e_k) = 1$, якщо $i = k$. Тому

$$(x, y) = \sum_{i=1}^{n} \xi_i \eta_i = \xi_1\eta_1 + \xi_2\eta_2 + ... + \xi_n\eta_n.$$

Якщо в ортонормованому базисі Е вектори x і y подати у вигляді стовпців їх координат

$$x_{\mathrm{E}} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \; y_{\mathrm{E}} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix},$$

тоді скалярний добуток записують так:

$$(x, y) = x^T \cdot y. \tag{2}$$

У евклідовому просторі R^n зручно проводити операції з векторами відносно якого-небудь виділеного ортонормованого базису, котрий відігравав би роль стандартного базису.

Теорема 3 (про ізоморфізм). *Нехай* V_1 і V_2 *– евклідові простори, структура яких визначена на базі арифметичного лінійного простору* R^n *скалярними добутками* $(x, y)_1$ і $(x, y)_2$ *векторів x, y відповідно. Існує оборотне лінійне перетворення* $\mathcal{A}$ *простору* R^n*, яке зберігає скалярний добуток між векторами:*

$$(x, y)_1 = (\mathcal{A}(x), \mathcal{A}(y))_2 .$$

Доведення. Оберемо в лінійному просторі R^n який-небудь ортонормований базис $\mathrm{E} = \langle e_1, e_2, ..., e_n \rangle$ відносно скалярного добутку $(x, y)_1$, тобто в евклідовому просторі V_1. Якщо $x = \xi_1 e_1 + \xi_2 e_2 + ... + \xi_n e_n$ і $y = \eta_1 e_1 + \eta_2 e_2 + ... + \eta_n e_n$, то $(x, y)_1 =$
$= \xi_1\eta_1 + \xi_2\eta_2 + ... + \xi_n\eta_n$.

Аналогічно оберемо базис $\Phi = \langle f_1, f_2, ..., f_n \rangle$, ортонормований відносно скалярного добутку $(x, y)_2$, тобто у просторі V_2.

На основі теорії лінійних операторів існує лінійне перетворення $\mathcal{A}$ арифметичного векторного простору R^n таке, для якого

$$\mathcal{A}(e_k) = f_k,\ k = \overline{1, n}.$$

Воно єдине і оборотне. При цьому

$$x_1 = \mathcal{A}(x) = \sum_{k=1}^{n} \xi_k f_k ,$$

$$y_1 = \mathcal{A}(y) = \sum_{k=1}^{n} \eta_k f_k .$$

Скалярний добуток $(x_1, y_1)_2 = (\mathcal{A}(x), \ \mathcal{A}(y))_2 = \sum_{k=1}^{n} \xi_k \eta_k$.

Дістали рівність $(x, y)_1 = (\mathcal{A}(x), \mathcal{A}(y))_2$.

Теорему доведено.

Зауваження. Раніше в загальній теорії лінійних рівнянь було введено поняття спряженого підпростору для заданого лінійного підпростору векторного простору P^n над довільним полем P. Якщо підпростір M буде лінійною оболонкою $L(a_1, a_2, ..., a_m)$ системи векторів

$$\begin{aligned} a_1 &= (\alpha_{11}, \alpha_{12}, ..., \alpha_{1n}), \\ a_2 &= (\alpha_{21}, \alpha_{22}, ..., \alpha_{2n}), \\ &\dots\dots\dots\dots\dots\dots \\ a_m &= (\alpha_{m1}, \alpha_{m2}, ..., \alpha_{mn}), \end{aligned}$$

то спряжений підпростір M^* – множина всіх розв'язків системи однорідних рівнянь

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + ... + \alpha_{1n}\xi_n = 0, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + ... + \alpha_{2n}\xi_n = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + ... + \alpha_{mn}\xi_n = 0. \end{cases}$$

Нехай основним полем буде поле дійсних чисел R. Тоді в евклідовому просторі R^n спряжений підпростір M^* для $M = L(a_1, a_2, ..., a_m)$ збігається з множиною всіх векторів $x \in R^n$, ортогональних векторам системи $a_1, a_2, ..., a_m$:

$$\begin{cases} (a_1, x) = 0, \\ (a_2, x) = 0, \\ \dots\dots\dots\dots \\ (a_m, x) = 0. \end{cases}$$

Інакше кажучи, M^* збігається з ортогональним доповненням $M^{\perp}$.

Вправи

1. Застосувавши процес ортогоналізації, знайти ортогональний базис підпростору $L(a_1, a_2, a_3, a_4)$, де $a_1 = (2, 3, -4, -6)$, $a_2 = (1, 8, -2, -16)$, $a_3 = (12, 5, -14, 5)$, $a_4 = (3, 11, 4, -7)$.

2. Доповнити систему векторів $a_1 = (1/2, -1/2, 1/2, -1/2)$, $a_2 = (-1/2, 1/2, 1/2, -1/2)$ до ортонормованого базису.

3. Знайти систему лінійних рівнянь, яка визначатиме лінійний підпростір M, натягнутий на вектори $a_1 = (1, 0, 0, -1)$, $a_2 = (2, 1, 1, 0)$, $a_3 = (1, 1, 1, 1)$, $a_4 = (1, 2, 3, 4)$, $a_5 = (0, 1, 2, 3)$.

4. Знайти базиси суми та перерізу лінійних підпросторів $M_1 = L(a_1, a_2, a_3)$ і $M_2 = L(b_1, b_2, b_3)$, де $a_1 = (1, 2, 1, -2)$, $a_2 = (2, 3, 1, 0)$, $a_3 = (1, 2, 2, -3)$; $b_1 = (1, 1, 1, 1)$, $b_2 = (1, 0, 1, -1)$, $b_3 = (1, 3, 0, -4)$.

5. Підмножину K_n евклідового простору R^n, складену з усіх точок $x = (\xi_1, \xi_2, ..., \xi_n)$, для яких $0 \le \xi_i \le 1$, $i = \overline{1, n}$, називатимемо *n-вимірним кубом*. Точки $a = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$, де $\varepsilon_i = 0, 1$, назвемо його *вершинами*, причому протилежні вершини визначимо як $a = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ і $\overline{a} = (\overline{\varepsilon}_1, \overline{\varepsilon}_2, ..., \overline{\varepsilon}_n)$, де $\overline{\varepsilon}_i = 1$, якщо $\varepsilon_i = 0$, і $\overline{\varepsilon}_i = 0$, якщо $\varepsilon_i = 1$.

Для протилежних вершин a і $\overline{a}$ множину всіх точок виду

$$x = \lambda a + \mu \overline{a},\ \lambda \ge 0,\ \mu \ge 0,\ \lambda + \mu = 1$$

назвемо *діагоналлю з кінцями a і $\overline{a}$*.

А) Знайти число діагоналей n-вимірного куба K_n, ортогональних до даної діагоналі;

Б) Знайти довжину діагоналі куба та її границю при $n \to \infty$;

В) Довести, що ортогональні проекції вершин куба K_n на будь-яку діагональ ділять її на n рівних частин.

6. Нехай R_{n+1} – евклідовий простір усіх многочленів від змінної t степеня $\le n$, де скалярний добуток многочленів $f(t)$ і $g(t)$ визначається законом

$$(f, g) = \int_{-1}^{1} f(t) g(t) dt .$$

Довести, що многочлени Лежандра

$$P_0(t) = 1,\ P_k(t) = \frac{1}{2^k\, k!}\,\frac{d^k}{dt^k}(t^2 - 1)^k,\ k = \overline{1, n}$$

становлять ортогональний базис цього простору. Показати, що коли до базису $1, t, t^2, \ldots, t^n$ простору R_{n+1} застосувати процес ортогоналізації, то дістанемо многочлени $f_0(t), f_1(t), \ldots, f_n(t)$, які збігатимуться з послідовністю многочленів Лежандра з точністю до сталих множників.

Ортогональні матриці

Означення. *Ортогональною називається матриця переходу від одного ортонормованого базису евклідового простору до іншого ортонормованого базису.*

Нехай $\mathrm{E} = \langle e_1, e_2, \ldots, e_n \rangle$ – ортонормований базис простору R^n і $\Phi = \langle f_1, f_2, \ldots, f_n \rangle$ – інший (новий) ортонормований базис. Подаємо вектори із базису Ф у вигляді лінійних комбінацій векторів базису Е:

$$f_k = \chi_{1k} e_1 + \chi_{2k} e_2 + \ldots + \chi_{nk} e_n,\ k = \overline{1, n}\,.$$

Тоді матрицею переходу буде

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

Знайдемо фундаментальні співвідношення між елементами χ_{ik} цієї матриці. Скалярний добуток (f_i, f_k), обчислений в базисі Е, дорівнює сумі $\chi_{1i}\chi_{1k} + \chi_{2i}\chi_{2k} + \ldots + \chi_{ni}\chi_{nk}$. Але для векторів із базису Ф виконуються рівності

$$(f_i, f_k) = \begin{cases} 0, \text{ якщо } i \neq k; \\ 1, \text{ якщо } i = k. \end{cases}$$

Звідси дістаємо систему рівностей

$$\chi_{1i}\chi_{1k}+\chi_{2i}\chi_{2k}+\ldots+\chi_{ni}\chi_{nk}=\begin{cases}0, \text{ якщо } i\neq k;\\ 1, \text{ якщо } i=k.\end{cases} \tag{3}$$

Сума добутків відповідних елементів різних стовпців матриці дорівнює нулю, а сума квадратів елементів одного стовпця дорівнює одиниці.

Співвідношення (3) коротко можна подати у вигляді матричної рівності

$$X^T\cdot X=\begin{pmatrix}\chi_{11} & \chi_{21} & \cdots & \chi_{n1}\\ \chi_{12} & \chi_{22} & \cdots & \chi_{n2}\\ \ldots & \ldots & \ldots & \ldots\\ \chi_{1n} & \chi_{2n} & \cdots & \chi_{nn}\end{pmatrix}\cdot\begin{pmatrix}\chi_{11} & \chi_{12} & \cdots & \chi_{1n}\\ \chi_{21} & \chi_{22} & \cdots & \chi_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \chi_{n1} & \chi_{n2} & \cdots & \chi_{nn}\end{pmatrix}=$$
$$=\begin{pmatrix}1 & 0 & \ldots & 0\\ 0 & 1 & \ldots & 0\\ \ldots & \ldots & \ldots & \ldots\\ 0 & 0 & \ldots & 1\end{pmatrix}=E. \tag{4}$$

Твердження 1. *Визначник довільної ортогональної матриці дорівнює 1, або –1.*

Доведення. Нехай X – ортогональна матриця. Із рівності $X^T\cdot X=E$ випливає, що $\det(X^T\cdot X)=\det X^T\cdot\det X=\det E=1$. Оскільки $\det X^T=\det X$, то $(\det X)^2=1$. Тому $\det X=\pm1$.

Твердження 2. *Ортогональна матриця X оборотна і*

$$X^{-1}=X^T. \tag{5}$$

Доведення. Оскільки $\det X=\pm1$, то X має обернену матрицю X^{-1}. Із рівності (4) знаходимо $(X^T\cdot X)\cdot X^{-1}=X^{-1}$, з іншого боку $X^T\cdot(X\cdot X^{-1})=X^T$, а тому $X^T=X^{-1}$.

Вправи

1. Показати, що коли дійсна квадратна матриця X підкоряється співвідношенню (4) $X^T\cdot X=E$, то вона ортогональна.

2. Довести, що коли X – ортогональна матриця порядку n, то для будь-яких векторів $x=(\xi_1,\xi_2,\ldots,\xi_n)^T$ і $y=(\eta_1,\eta_2,\ldots,\eta_n)^T$, заданих в ортонормованому базисі, виконується рівність

$$(x,y)=(Xx,\ Xy)$$

(закон збереження скалярного добутку).

3. Нехай на деякій множині G задано операцію – функція двох змінних, що подається у вигляді $a \cdot b = c$, і така, що ставить у відповідність кожній впорядкованій парі (a, b) деякий елемент c:

$$a \cdot b = c,$$

причому так, що виконуються такі вимоги:

I. *Властивість асоціативності*

$$(a \cdot b) \cdot d = a \cdot (b \cdot d).$$

II. Існує такий елемент e, що для будь-якого елемента a

$$a \cdot e = e \cdot a = a$$

(e – *одиничний*, або *нейтральний елемент*).

III. Для кожного $a \in G$ існує елемент $x \in G$ такий, що

$$a \cdot x = x \cdot a = e.$$

Елемент x називається *оберненим* для a і позначається a^{-1}.

Множина G у цьому випадку називається *групою*.

Крім того, може виконуватися вимога IV. *Властивість комутативності:* $a \cdot b = b \cdot a$. Тоді група G називається *комутативною*, або *абелевою* (на честь норвезького математика Н. Абеля).

Показати, що множина всіх ортогональних матриць порядку n становить групу відносно операції множення. Вона називається *ортогональною групою* і позначається O_n.

Нехай F – множина всіх матриць виду

$$X_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},\ 0 \le \varphi \le 2\pi.$$

Показати, що F – комутативна група, яка належить групі O_2 (тобто є її підгрупою). Порівняти її з мультиплікативною групою K всіх комплексних чисел z, для яких $|z| = 1$.

§ 7.4. Лінійні перетворення евклідових просторів

Надалі скористаємося загальноприйнятим у математиці позначенням для евклідових просторів. Нехай лінійний підпростір L розкладається у пряму суму

$$L = L_1 \oplus L_2 \oplus \ldots \oplus L_m$$

попарно ортогональних підпросторів $L_1, L_2, \ldots, L_m$.Тоді

$$L = L_1 \perp L_2 \perp \ldots \perp L_m .$$

Зокрема, евклідовий простір R^n розкладається у пряму суму лінійного підпростору M і його ортогонального доповнення:

$$R^n = M \perp M^{\perp} .$$

Дійсна матриця

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

розмірності $m \times n$ визначає лінійне перетворення

$$\mathcal{A} : R^n \to R^m$$

евклідового простору R^n в евклідовий простір R^m. Для його точного опису оберемо в R^n і, відповідно, в R^m які-небудь ортонормовані базиси і зафіксуємо їх. Тоді лінійне перетворення задамо законом

$$y = A \cdot x ,$$

де x – стовпець координат вектора $x \in R^n$, а y – стовпець координат перетвореного вектора $y = A \cdot x \in R^m$ (у відповідно обраних базисах).

Нагадаємо означення ядра та образу лінійного перетворення.

Ядро $\operatorname{Ker} A = \{x \in R^n : A \cdot x = \theta\}$.

Образ $\operatorname{Im} A = \{y \in R^m:\ y = A \cdot x,\ x \in R^n\}$.

Лема. *Для будь-яких векторів* $x \in R^n$, $y \in R^m$ *виконується рівність*

$$(A \cdot x, y) = \left(x, A^T y\right).$$

Доведення.

$$(Ax, y) = (A \cdot x)^T \cdot y = x^T \cdot A^T \cdot y =$$
$$= x^T \cdot \left(A^T \cdot y\right) = \left(x, A^T y\right).$$

Теорема. *Для будь-якої дійсної матриці* $A = (\alpha_{ik})$ *розмірності* $m \times n$ *виконуються рівності*

$$\operatorname{Ker} A = \left(\operatorname{Im} A^T\right)^{\perp}, \tag{А}$$

$$R^n = \operatorname{Ker} A \perp \operatorname{Im} A^T. \tag{Б}$$

Зокрема, якщо матриця A симетрична, то

$$R^n = \operatorname{Ker} A \perp \operatorname{Im} A, \tag{Б$_0$}$$

$$R^m = \operatorname{Im} A \perp \operatorname{Ker} A^T, \tag{В}$$

$$\operatorname{Ker} A = \operatorname{Ker}\left(A^T \cdot A\right), \tag{Г}$$

$$\operatorname{Im} A^T = \operatorname{Im}\left(A^T \cdot A\right). \tag{Д}$$

Доведення рівності (А).

Нехай $x \in \operatorname{Ker} A$ і $z \in \operatorname{Im} A^T$. Тоді $z = A^T \cdot y$, де $y \in R^m$. Маємо $(x, z) = \left(x, A^T y\right) = (Ax, y) = (\theta, y) = 0$. Оскільки x – довільний вектор із $\operatorname{Ker} A$, тоді $z \in (\operatorname{Ker} A)^{\perp}$. Це означає, що $\operatorname{Im} A^T \subseteq (\operatorname{Ker} A)^{\perp}$. Звідси дістаємо включення

$$\operatorname{Ker} A \subseteq \left(\operatorname{Im} A^T\right)^{\perp}.$$

Для доведення оберненого включення обираємо $z \in \left(\mathrm{Im}A^T\right)^{\perp}$. Це означає, що вектор z ортогональний будь-якому вектору x із $\mathrm{Im}A^T$. Якщо $x = A^T \cdot y, y \in R^m$, то

$$\left(z, x\right) = \left(z, A^T y\right) = \left(Az, y\right) = 0$$

для будь-якого $y \in R^m$. Це можливо лише у тому випадку, коли $Az = \theta$, тобто коли $z \in \mathrm{Ker}A$. Дістали обернене включення

$$\left(\mathrm{Im}A^T\right)^{\perp} \subseteq \mathrm{Ker}A,$$

а звідси і потрібну рівність

$$\mathrm{Ker}A = \left(\mathrm{Im}A^T\right)^{\perp}.$$

Із рівності (А) випливає рівність (Б):

$$R^n = \mathrm{Ker}A \perp \mathrm{Im}A^T.$$

Із рівності (Б) заміною A на A^T отримаємо рівність (В):

$$R^m = \mathrm{Im}A \perp \mathrm{Ker}A^T.$$

А тепер доведемо рівність (Г). Одностороннє включення $\mathrm{Ker}\,A \subseteq \mathrm{Ker}\left(A^T \cdot A\right)$ очевидне.

Нехай тепер $x \in \mathrm{Ker}\left(A^T \cdot A\right)$, тобто $A^T \cdot Ax = \theta$. Тоді $\left(x, A^T \cdot Ax\right) = 0$, а тому $\left(Ax, Ax\right) = 0$. Це можливо лише при $A \cdot x = \theta$, тобто $x \in \mathrm{Ker}A$. Доведено обернене включення $\mathrm{Ker}\left(A^T \cdot A\right) \subseteq \mathrm{Ker}\,A$.

Тепер доведемо рівність (Д).

На підставі рівностей (А) і (Г) одержується така послідовність рівностей:

$$\mathrm{Im}\,A^T = \left(\mathrm{Ker}\,A\right)^{\perp} = \left[\left(\mathrm{Ker}\,A^T \cdot A\right)\right]^{\perp} =$$

$$= \mathrm{Im}\left(A^T \cdot A\right)^T = \mathrm{Im}\left(A^T \cdot A\right).$$

Теорему доведено.

§ 7.5. Спряжені оператори

Означення. *Нехай $\mathcal{A}$ – лінійний оператор евклідового простору R^n. Лінійний оператор $\mathcal{A}^*$ називається спряженим з ним, якщо $(\mathcal{A}(x), y) = (x, \mathcal{A}^*(y))$ для будь-яких $x, y \in R^n$.*

Теорема 1. *Для будь-якого лінійного оператора $\mathcal{A}$ існує (і при цьому єдиний) лінійний оператор, спряжений з ним.*

Якщо $\mathcal{A}^*$– оператор, спряжений з $\mathcal{A}$, тоді у будь-якому ортонормованому базисі їх матриці одержуються одна від одної транспонуванням.

Доведення. Доведемо спочатку друге твердження теореми.

Нехай $\mathrm{E} = \langle e_1, e_2, ..., e_n \rangle$ – ортонормований базис евклідового простору і $A = (\alpha_{ik}) = \mathcal{A}_{\mathrm{E}}$ – матриця оператора $\mathcal{A}$; $B = (\beta_{ik}) = \mathcal{A}^*_{\mathrm{E}}$ – матриця оператора $\mathcal{A}^*$ в базисі E. Тоді

$$\mathcal{A}(e_k) = \sum_{i=1}^{n} \alpha_{ik} e_i \, , \mathcal{A}^*(e_k) = \sum_{i=1}^{n} \beta_{ik} e_i \ \left(k = \overline{1, n}\right).$$

Маємо такі рівності:

$$(\mathcal{A}(e_k), e_j) = \left(\sum_{i=1}^{n} \alpha_{ik} e_i, e_j\right) = \sum_{i=1}^{n} \alpha_{ik} (e_i, e_j) = \alpha_{jk};$$

$$(e_k, \mathcal{A}^*(e_j)) = \left(e_k, \sum_{i=1}^{n} \beta_{ij} e_i\right) = \sum_{i=1}^{n} \beta_{ij} (e_k, e_i) = \beta_{kj}.$$

Оскільки $(\mathcal{A}(e_k), e_j) = (e_k, \mathcal{A}^*(e_j))$, то дістанемо $\beta_{kj} = \alpha_{jk}$. Це означає, що $B = A^T$.

Існування спряженого оператора випливає з таких міркувань. Нехай в ортонормованому базисі $\mathrm{E} = \langle e_1, e_2, ..., e_n \rangle$ маємо

$$\mathcal{A}_{\mathrm{E}} = A = (\alpha_{ik}).$$

Для транспонованої матриці A^T існує такий лінійний оператор $\mathcal{B}$, що $\mathcal{B}_E = A^T$. Для нього виконуються рівності

$$\mathcal{B}(e_k) = \alpha_{k1}e_1 + \alpha_{k2}e_2 + \ldots + \alpha_{kn}e_n \quad \left(k = \overline{1,n}\right).$$

Нехай x, y – довільні вектори простору, $z = \mathcal{A}(x)$, $w = \mathcal{B}(y)$ і

$$x_E = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, y_E = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}, z_E = \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix}, w_E = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix}.$$

Тоді $z_E = A \cdot x_E$, $w_E = A^T \cdot y_E$, тобто $\zeta_i = \sum_{k=1}^{n} \alpha_{ik}\xi_k$, $\omega_i = \sum_{k=1}^{n} \alpha_{ki}\eta_k$.

Обчислюємо

$$(\mathcal{A}(x), y) = (z, y) = \sum_{i=1}^{n} \zeta_i\eta_i = \sum_{i=1}^{n}\left(\sum_{k=1}^{n} \alpha_{ik}\xi_k\right)\eta_i = \sum_{i,k=1}^{n} \alpha_{ik}\xi_k\eta_i.$$

$$\left(x, \mathcal{B}(y)\right) = (x, w) = \sum_{i=1}^{n} \xi_i\omega_i = \sum_{i=1}^{n} \xi_i\left(\sum_{k=1}^{n} \alpha_{ki}\eta_k\right) =$$

$$= \sum_{i,k=1}^{n} \alpha_{ki}\xi_i\eta_k = \sum_{i,k=1}^{n} \alpha_{ik}\xi_k\eta_i.$$

Порівнюючи ці дві рівності, можемо зробити висновок, що $(\mathcal{A}(x), y) = (x, \mathcal{B}(y))$ для будь-яких векторів x, y. Тому $\mathcal{B} = \mathcal{A}^*$.

Єдиність спряженого оператора $\mathcal{A}^*$ для $\mathcal{A}$ випливає з першого пункту цього доведення.

Теорему доведено.

Наслідок. *Операція спряження лінійних операторів має такі властивості:*

I. $(\mathcal{A}^*)^* = \mathcal{A}$.

II. $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$.

III. $(\mathcal{A} \cdot \mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$.

IV. $(\lambda \cdot \mathcal{A})^* = \lambda\, \mathcal{A}^*$.

Ці властивості випливають із відповідних властивостей операції транспонування матриць.

Теорема 2. *Якщо лінійний підпростір M інваріантний відносно лінійного оператора $\mathcal{A}$, то ортогональне доповнення $M^{\perp}$ інваріантне відносно спряженого оператора $\mathcal{A}^*$.*

Доведення. Нехай x – довільний вектор із M і y – вектор із $M^{\perp}$. Оскільки $\mathcal{A}(x) \in M$, то $(\mathcal{A}(x), y) = 0$. Із рівності $(\mathcal{A}(x), y) = (x, \mathcal{A}^*(y))$ випливає, що $(x, \mathcal{A}^*(y)) = 0$.

Це означає, що $\mathcal{A}^*(y) \in M^{\perp}$.

Означення. *Лінійний оператор $\mathcal{A}$ простору R^n називається самоспряженим (симетричним), якщо він збігається зі своїм спряженим оператором $\mathcal{A}^*$, тобто для будь-яких x, y виконується рівність $(\mathcal{A}(x), y) = (x, \mathcal{A}(y))$.*

Наведемо такі властивості самоспряжених операторів:

1. В ортонормованому базисі самоспряженому оператору відповідає симетрична матриця. Має місце й обернене твердження.

2. Усі характеристичні числа самоспряженого оператора дійсні й, отже, є його власними значеннями.

Самоспряженому оператору $\mathcal{A}$ в ортонормованому базисі відповідає симетрична матриця A. Але всі характеристичні числа матриці A дійсні (див. § 6.4, теорема 3).

3. Якщо x, y – власні вектори самоспряженого оператора $\mathcal{A}$ з різними власними значеннями λ, μ, то x і y – ортогональні.

Оскільки $\mathcal{A}(x) = \lambda x$, $\mathcal{A}(y) = \mu y$, то з рівності $(\mathcal{A}(x), y) = (x, \mathcal{A}(y))$ дістаємо $(\lambda x, y) = (x, \mu y) \Rightarrow (\lambda - \mu)(x, y) = 0$. Із $\lambda - \mu \neq 0$ випливає $(x, y) = 0$.

4. Якщо лінійний підпростір M інваріантний відносно самоспряженого оператора $\mathcal{A}$, тоді й ортогональне доповнення $M^{\perp}$ інваріантне відносно $\mathcal{A}$.

Ця властивість випливає з теореми 2.

Теорема 3. *Якщо $\mathcal{A}$ – самоспряжений лінійний оператор простору R^n, тоді*

$$R^n = V_{\lambda_1} \perp V_{\lambda_2} \perp \ldots \perp V_{\lambda_r},$$

де $\lambda_1, \lambda_2, \ldots, \lambda_r$ *– усі відмінні одне від одного власні значення оператора* $\mathcal{A}$ *і підпростори* $V_{\lambda_k} = \mathrm{Ker}(\lambda_k \mathcal{E} - \mathcal{A})$, $k = \overline{1, r}$, *ортогональні.*

Доведення. Мінімальним інваріантним підпростором, який містить усі власні вектори оператора $\mathcal{A}$, є $V = V(\mathcal{A}) = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r}$ (див. § 6.7, теорема 4 і наслідок). Доданки цієї прямої суми ортогональні, оскільки складаються вони з власних векторів з відмінними один від одного власними значеннями. Тоді

$$R^n = V \perp V^{\perp},$$

де ортогональне доповнення $V^{\perp}$ також інваріантне відносно $\mathcal{A}$.

Покажемо, що $V^{\perp} = \{\theta\}$ – тривіальний підпростір. Припустимо, що $V^{\perp} \neq \{\theta\}, l = \dim V^{\perp} > 0$.

Характеристичний многочлен $\chi_1(t)$ оператора $\mathcal{A}_{|V^{\perp}}$ – обмеження $\mathcal{A}$ на $V^{\perp}$ – має степінь $l > 0$. Він – дільник характеристичного многочлена оператора $\mathcal{A}$ (див. § 6.7, теорема 2). Тому будь-який корінь λ многочлена $\chi_1(t)$ збігається з одним із чисел $\lambda_1, \lambda_2, \ldots, \lambda_r$. Знайдемо який-небудь власний вектор $w \in V^{\perp}$ оператора $\mathcal{A}_{|V^{\perp}}$ з власним значенням λ. Він, разом з тим, є власним вектором оператора $\mathcal{A}$.

Таким чином, $w \in V$. Але $w \neq \theta$. Це суперечить розкладу $R^n = V \perp V^{\perp}$.

Отже, $V^{\perp} = \{\theta\}$ і $R^n = V_{\lambda_1} \perp V_{\lambda_2} \perp \ldots \perp V_{\lambda_r}$.

Структура самоспряженого оператора проста.

Теорема 4. *Для дійсної симетричної матриці A існує така ортогональна матриця X, що*

$$X^T \cdot A \cdot X = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ldots & \\ 0 & & & \lambda_n \end{pmatrix}$$

– *діагональна матриця, на головній діагоналі якої містяться власні значення матриці A.*

Доведення. Нехай в ортонормованому базисі $\mathrm{E} = \langle e_1, e_2, \ldots, e_n \rangle$ деякий самоспряжений оператор $\mathcal{A}$ подається симетричною матрицею A. Оберемо новий ортонормований базис $\Phi = \langle f_1, f_2, \ldots, f_n \rangle$, складений із власних векторів оператора $\mathcal{A}$ (матриці A) таким чином. У підпросторі $V_{\lambda_1} = \mathrm{Ker}(\lambda_1 \mathcal{E} - \mathcal{A})$ оберемо ортонормований базис. Далі те саме проробимо у підпросторі V_{λ_2} і т. д. Об'єднання їх усіх становитиме ортонормований базис R^n.

Якщо X – матриця переходу від Е до Ф, то матриця

$$X^{-1}AX = X^T \cdot AX$$

матиме діагональний вигляд.

Правило знаходження ортогональної матриці. Для того щоб знайти ортогональну матрицю X, яка перетворенням конгруентності X^TAX приводить симетричну матрицю $A = (\alpha_{ik})$ до діагонального виду, потрібно:

I. Скласти характеристичний многочлен матриці A та знайти всі його корені. Нехай $\lambda_1, \lambda_2, \ldots, \lambda_r$ – усі попарно різні корені.

II. Для кореня λ_1 складемо систему лінійних рівнянь

$$\begin{cases} (\lambda_1 - \alpha_{11})\xi_1 - \alpha_{12}\xi_2 - \ldots - \alpha_{1n}\xi_n = 0, \\ -\alpha_{21}\xi_1 + (\lambda_1 - \alpha_{22})\xi_2 - \ldots - \alpha_{2n}\xi_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -\alpha_{n1}\xi_1 - \alpha_{n2}\xi_2 - \ldots + (\lambda_1 - \alpha_{nn})\xi_n = 0 \end{cases} \quad \left(V_{\lambda_1}\right)$$

та знайдемо фундаментальну систему її розв'язків. Нехай вона складається із d_1 розв'язків. Число d_1 дорівнює дефекту матриці цієї системи рівнянь.

До одержаної системи розв'язків застосуємо спочатку процес ортогоналізації, а потім – нормування. Дістанемо вектори $f_1, f_2, \ldots, f_{d_1}$, які становитимуть ортонормований базис підпростору V_{λ_1}. Тоді перший стовпець матриці X буде заповнений вектором f_1, другий – вектором f_2 і т. д. Отже, будуть заповнені d_1 стовпців матриці X.

III. Аналогічно складається система $\left(V_{\lambda_2}\right)$ лінійних рівнянь для кореня λ_2 і повторюється процедура II. Знову-таки дістанемо d_2 ортонормованих векторів $f_{d_1+1}, \ldots, f_{d_1+d_2}$, що становлять базис підпростору V_{λ_2}. Координати цих векторів заповнюють d_2 стовпців матриці X.

Цю процедуру продовжуємо доти, доки не буде остаточно побудована матриця X.

Приклад 1. Знайти ортогональну матрицю, яка конгруентно перетворює матрицю

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

у діагональну.

Розв'язання. Спочатку обчислюємо характеристичний многочлен

$$\mathcal{X}(t) = \begin{vmatrix} t & 0 & 0 & -1 \\ 0 & t & -1 & 0 \\ 0 & -1 & t & 0 \\ -1 & 0 & 0 & t \end{vmatrix} = \left(t^2 - 1\right)^2.$$

Із $\mathcal{X}(t) = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = -1$.

Потім складаємо систему $\left(V_{\lambda_1}\right)$ рівнянь. Вона матиме вигляд

$$\begin{cases} \xi_1 - \xi_4 = 0, \\ \xi_2 - \xi_3 = 0. \end{cases}$$

Вектори $a_1 = (1, 0, 0, 1)$ і $a_2 = (0, 1, 1, 0)$ утворюють фундаментальну систему розв'язків. Вони ортогональні. Після їх нормування знайдемо

$$f_1 = \left(\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}\right), f_2 = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right).$$

Система $\left(V_{\lambda_2}\right)$ для $\lambda_2 = -1$ зводиться до вигляду

$$\begin{cases} \xi_1 + \xi_4 = 0, \\ \xi_2 + \xi_3 = 0. \end{cases}$$

Для f_3 і f_4 матимемо:

$$f_3 = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), f_4 = \left(-\frac{1}{\sqrt{2}}, 0, 0, \frac{1}{\sqrt{2}}\right).$$

Остаточно

$$X = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad X^T AX = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Приклад 2. Звести квадратичну форму

$$f = 6x_1^2 + 5x_2^2 + 7x_3^2 - 4x_1x_2 + 4x_1x_3$$

до діагонального виду за допомогою ортогонального перетворення змінних x_1, x_2, x_3.

Розв'язання. Матриця форми

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

Спочатку обчислимо характеристичний многочлен

$$\chi(t) = \begin{vmatrix} t-6 & 2 & -2 \\ 2 & t-5 & 0 \\ -2 & 0 & t-7 \end{vmatrix} = (t-6)(t-9)(t-3) \Rightarrow$$

$$\Rightarrow \lambda_1 = 6, \lambda_2 = 9, \lambda_3 = 3.$$

Складемо систему рівнянь для $\lambda_1 = 6$. Матимемо

$$\begin{cases} 2x_2 - 2x_3 = 0, \\ 2x_1 + x_2 = 0, \\ -2x_1 - x_3 = 0. \end{cases}$$

Загальним її розв'язком буде $X = \left(-\frac{1}{2}x_3, x_3, x_3 \right)$.

Фундаментальна система розв'язків складається лише з одного розв'язку $a_1 = (-1, 2, 2)$, або після нормування

$$f_1 = \left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right).$$

Для $\lambda_2 = 9$ із системи рівнянь

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

знайдемо $a_2 = (2, -1, 2)$. Або після нормування

$$f_2 = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3} \right).$$

Аналогічно для $\lambda_3 = 3$ знаходимо

$$f_3 = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3} \right).$$

Шукана ортогональна матриця

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

і відповідне лінійне перетворення змінних

$$x_1 = -\frac{1}{3}y_1 + \frac{2}{3}y_2 + \frac{2}{3}y_3,$$

$$x_2 = \frac{2}{3}y_1 - \frac{1}{3}y_2 + \frac{2}{3}y_3,$$

$$x_3 = \frac{2}{3}y_1 + \frac{2}{3}y_2 - \frac{1}{3}y_3 .$$

У нових змінних y_1, y_2, y_3 форма набуде вигляду

$$f = 6y_1^2 + 9y_2^2 + 3y_3^2 .$$

Вправи

1. За допомогою ортогонального перетворення звести до канонічного вигляду квадратичні форми:

а) $f = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 - 2x_1x_4 - 2x_2x_3 + 2x_3x_4$;

б) $f = 8x_1x_3 + 2x_1x_4 + 2x_2x_3 + 8x_2x_4$;

в) $f = \sum\limits_{i \neq k} x_i x_k \quad (1 \leq i,\ k \leq n)$.

2. Знайти оператор, спряжений з оператором $\mathcal{A}(x) = [x, a]$ евклідового простору R^3 .

3. Нехай $Б = \langle b_1, b_2, ..., b_n \rangle$ – ортогональний (але не ортонормований) базис простору R^n . Встановити зв'язок між матрицями оператора $\mathcal{A}$ і спряженого оператора $\mathcal{A}^*$ у цьому базисі.

4. У просторі R^2 многочленів $f(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2$ введено скалярний добуток

$$(f, g) = \int_{-1}^{1} f(t) g(t) dt .$$

Знайти матрицю спряженого оператора по відношенню до оператора диференціювання в базисі $1, t, t^2$.

§ 7.6. Ортогональні оператори

Лінійний оператор $\mathcal{X}$ називається *ортогональним*, якщо він зберігає скалярний добуток векторів: $(x, y) = (\mathcal{X}(x), \mathcal{X}(y))$ для будь-яких $x, y \in R^n$.

Наведемо деякі загальні властивості ортогональних операторів.

I. Для того щоб лінійний оператор $\mathcal{X}$ був ортогональним, необхідно і достатньо, щоб він будь-який ортонормований базис переводив у ортонормований базис. Звідси випливає, що у будь-якому ортонормованому базисі йому відповідає ортогональна матриця X.

Оскільки спряженому оператору $\mathcal{X}^*$ відповідає *транспонована* матриця X^T, то з рівності $X^T \cdot X = E$ випливає рівність $\mathcal{X}^* \cdot \mathcal{X} = \mathcal{E}$. Ортогональний оператор $\mathcal{X}$ оборотний і $\mathcal{X}^{-1} = \mathcal{X}^*$.

II. Усі числа, відмінні від ± 1, не можуть бути власними значеннями ортогонального оператора.

Нехай ортогональний оператор $\mathcal{X}$ має власний вектор x з відповідним власним значенням λ. Тоді $(x, x) = (\mathcal{X}(x), \mathcal{X}(x)) =$
$= (\lambda x, \lambda x) = \lambda^2 (x, x)$. Оскільки $(x, x) > 0$, то $\lambda^2 = 1 \Rightarrow \lambda = \pm 1$.

III. Якщо підпростір M інваріантний відносно ортогонального оператора $\mathcal{X}$, тоді й ортогональне доповнення $M^\perp$ інваріантне відносно $\mathcal{X}$.

Нехай $m = \dim M$ і $\mathrm{E} = \langle e_1, e_2, \ldots, e_m \rangle$ – базис підпростору M. За допомогою перетворення $\mathcal{X}$ базис Е перетворюється в базис, складений із векторів $f_1 = \mathcal{X}(e_1)$, $f_2 = \mathcal{X}(e_2)$, ..., $f_m = \mathcal{X}(e_m)$. Якщо $x \in M$ і $x = \sum_{k=1}^{m} \xi_k f_k$, то $\mathcal{X}^{1}(x)$ =

$$= \mathcal{X}^{-1} \left(\sum_{k=1}^{m} \xi_k f_k \right) = \sum_{k=1}^{m} \xi_k \mathcal{X}^{-1}(f_k) = \sum_{k=1}^{m} \xi_k e_k \in M.$$

Це означає, що підпростір M інваріантний відносно $\mathcal{X}^{-1} = \mathcal{X}^*$. Але тоді ортогональне доповнення $M^\perp$ інваріантне відносно спряженого оператора $(\mathcal{X})^{**} = \mathcal{X}$.

Ортогональні оператори простору розмірності 1

Евклідовий простір R^1 має лише один базисний вектор e. Ортогональний оператор $\mathcal{X}$ перетворює його у вектор виду $\mathcal{X}(e) = \lambda e$. Число λ – власне значення оператора $\mathcal{X}$. Тому $\lambda = \pm 1$. Якщо $\lambda = 1$, то для будь-якого вектора $x = \xi e$

матимемо $\mathcal{X}(x) = x$, тобто $\mathcal{X}$ – тотожне перетворення. Якщо $\lambda = -1$, то $\mathcal{X}(x) = -x$ для будь-якого вектора $x \in R^1$.

Висновок. Ортогональний оператор одновимірного евклідового простору є або тотожним перетворенням, або дзеркальним відображенням відносно початку координат.

Ортогональні оператори простору розмірності 2

Нехай в евклідовому просторі R^2 обрано ортонормований базис e_1, e_2. Ортогональному оператору $\mathcal{X}$ у цьому базисі відповідає ортогональна матриця

$$X = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

Спочатку розглянемо випадок, коли $|X| = \alpha\delta - \beta\gamma = 1$. Обчислимо обернену матрицю за допомогою відомої формули

$$X^{-1} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

Але $X^{-1} = X^T$. Тому $\alpha = \delta$, $\gamma = -\beta$, тобто $X = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, причому $\alpha^2 + \beta^2 = 1$. Покладемо $\alpha = \cos\varphi, \beta = -\sin\varphi$. Тоді

$$X = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Це – матриця повороту на кут φ. Якщо $\varphi \neq k\pi$, $k = 0, \pm 1, \ldots$, тоді оператор $\mathcal{X}$ повороту на кут φ не матиме нетривіальних інваріантних підпросторів в евклідовому просторі R^2.

Якщо $\varphi = k\pi$, то $R^2 = L_1 \oplus L_2$, де L_1 і L_2 підпростори відповідно з базисними векторами e_1 і e_2, інваріантні відносно $\mathcal{X}$.

Нехай тепер $|X| = \alpha\delta - \beta\gamma = -1$. Характеристичний многочлен

$$\chi(t) = \begin{vmatrix} t-\alpha & -\beta \\ -\gamma & t-\delta \end{vmatrix} = t^2 - (\alpha+\delta)t - 1.$$

Дискримінант його $\Delta = (\alpha + \delta)^2 + 4$. Характеристичні числа λ_1, λ_2 дійсні і є власними значеннями оператора. Вони можуть бути рівними лише ± 1. Оскільки $\lambda_1 \cdot \lambda_2 = -1$, то одне з них дорівнює 1, а інше – 1. Нехай $\lambda_1 = 1, \lambda_2 = -1$. Їм відповідають власні вектори f_1 і f_2: $\mathcal{X}(f_1) = f_1$, $\mathcal{X}(f_2) = -f_2$. Вони ортогональні: $(f_1, f_2) = 0$. Якщо L_1 – лінійна оболонка вектора f_1, L_2 – вектора f_2, то $R^2 = L_1 \oplus L_2$, і оператор $\mathcal{X}$ індукує на L_1 тотожне перетворення, а на L_2 – дзеркальне відображення. Підпростори L_1 і L_2 – ортогональні.

Висновок. Ортогональний оператор $\mathcal{X}$ простору R^2 діє на ньому або як обертання на кут $\varphi \neq k\pi$, або простір розкладається в пряму суму $R^2 = L_1 \oplus L_2$ одновимірних інваріантних підпросторів L_1, L_2, ортогональних між собою і на яких індукує тотожне перетворення або ж дзеркальне відображення відносно початку координат.

Теорема (про структуру ортогональних операторів). *Для будь-якого ортогонального оператора $\mathcal{X}$ простір R^n розкладається в пряму суму*

$$R^n = L_1 \perp L_2 \perp \ldots \perp L_r$$

попарно ортогональних інваріантних підпросторів L_i, розмірності яких не більше ніж 2. На доданках з розмірністю 1 оператор індукує або тотожне перетворення, або дзеркальне відображення. На доданках з розмірністю 2 він індукує обертання.

Доведення проведемо індукцією за розмірністю n простору. Якщо $n \leq 2$, то справедливість теореми випливає з наведених вище міркувань.

Припустимо, що $n > 2$ і твердження теореми зберігає силу для просторів з розмірністю $< n$. Як і будь-який лінійний оператор дійсного векторного простору, перетворення $\mathcal{X}$ матиме в R^n інваріантний підпростір L_1 розмірності 1 або 2 (див. § 6.7, теорема 3).

Розкладемо простір у пряму суму $R^n = L_1 \oplus L_1^{\perp}$. Ортогональне доповнення $L_1^{\perp}$ інваріантне відносно $\mathcal{X}$ і його розмірність менша за n. За припущенням індукції $L_1^{\perp}$ розкладається у пряму суму

$$L_1^{\perp} = L_2 \perp L_3 \perp \ldots \perp L_r,$$

усі доданки якої $L_j, j = \overline{2,r}$ задовольняють вимоги теореми. Звідси дістаємо потрібний нам розклад

$$R^n = L_1 \perp L_2 \perp \ldots \perp L_r.$$

Наслідок*. Для кожної ортогональної матриці X існує така ортогональна матриця T, для якої*

$$T^{-1}XT = \begin{pmatrix} y_1 & & & 0 \\ & y_2 & & \\ & & \ldots & \\ 0 & & & y_m \end{pmatrix},$$

де $y_k = \pm 1$, *або*

$$y_k = \begin{pmatrix} \cos\varphi_k & -\sin\varphi_k \\ \sin\varphi_k & \cos\varphi_k \end{pmatrix}.$$

Вправи

1. Базиси Е і Ф евклідового простору R^3 називаються *однаково орієнтованими*, якщо визначник матриці переходу від Е до Ф додатний.

Довести теорему Ейлера: якщо ортогональний оператор $\mathcal{X} \neq \mathcal{E}$ тривимірного простору R^3 не змінює орієнтації (тобто його матриця X має $\det X = 1$), то $\mathcal{X}$ є обертанням навколо деякої осі.

Вказівка. Переконатися, що характеристичний многочлен оператора $\mathcal{X}$ має корінь, рівний 1, підпростір L власних векторів, відповідних власному значенню 1, є віссю обертання і $\mathcal{X}_{|L^{\perp}}$ – обмеження на $L^{\perp}$ – обертання в площині $L^{\perp}$.

2. Довести, що лінійне перетворення $\mathcal{A}(x) = [x, a]$ простору R^3 не є ортогональним.

3. Довести, що коли два вектори x, y простору R^3 мають однакову довжину $(\|x\| = \|y\|)$, то існує такий ортогональний оператор $\mathcal{X}$, для якого $\mathcal{X}(x) = y$.

4. Довести, що коли дві пари векторів x_1, x_2 і y_1, y_2 такі, що $\|x_1\| = \|y_1\|$, $\|x_2\| = \|y_2\|$ і кут між x_1, x_2 дорівнює куту між y_1, y_2, то існує такий ортогональний оператор $\mathcal{X}$, для якого $\mathcal{X}(x_1) = y_1$, $\mathcal{X}(x_2) = y_2$.

5. Довести, що коли лінійний оператор $\mathcal{X}$ евклідового простору зберігає довжину всіх векторів, то $\mathcal{X}$ – ортогональний оператор.

6. Довести, що коли лінійний оператор $\mathcal{X}$ зберігає ортогональність будь-яких двох векторів, то він відрізняється від деякого ортогонального оператора лише числовим множником.

7. Довести, що модулі всіх характеристичних чисел ортогональної матриці дорівнюють одиниці.

Вказівка. Це твердження одержується з наслідка структурної теореми про ортогональні оператори. Але його можна довести за допомогою простіших міркувань. Дійсно, нехай X – ортогональна матриця порядку n і λ – її характеристичне число. Число λ – власне значення нашої матриці у комплексному лінійному просторі C^n. Нехай x – відповідний власний вектор і $x = u + iv$, де $u, v \in R^n$. Якщо $\lambda = \alpha + i\beta$, тоді на підставі рівності $X \cdot x = \lambda x$ матимемо

$$Xu = \alpha u - \beta v,\ Xv = \beta u + \alpha v .$$

Оскільки

$$\|u\|^2 = (u, u) = (Xu, Xu) = \alpha^2 \|u\|^2 + \beta^2 \|v\|^2 - 2\alpha\beta(u, v),$$

$$\|v\|^2 = (v, v) = (Xv, Xv) = \alpha^2 \|v\|^2 + \beta^2 \|u\|^2 + 2\alpha\beta(u, v),$$

то

$$\|u\|^2 + \|v\|^2 = (\alpha^2 + \beta^2)(\|v\|^2 + \|u\|^2).$$

Звідси і випливає, що $\alpha^2 + \beta^2 = 1$, $|\lambda| = 1$.

§ 7.7. Метричний простір

Метричним простором називається множина V *елементів деякої природи, для кожної пари* x, y *з яких поставлено у відповідність дійсне число* $\rho(x,y)$ *з такими властивостями*:

1) $\rho(x,y) \geq 0$ і $\rho(x,y)=0 \Leftrightarrow x=y$;

2) $\rho(x,y)=\rho(y,x)$ (*симетричність*);

3) $\rho(x,y) \leq \rho(x,z)+\rho(z,y)$ (*нерівність трикутника*).

Число $\rho(x,y)$ *називається відстанню між елементами* x, y.

Елементи метричного простору часто називаються *точками.*

Кожний евклідовий простір V буде метричним простором, якщо покласти $\rho(x,y)=\|x-y\|$.

Властивості 1) –3) визначеної таким чином функції $\rho(x,y)$ випливають із властивостей норми векторів. Наприклад, нерівність трикутника:

$$\rho(x,y)=\|x-y\|=\|(x-z)+(z-y)\| \leq \|x-z\|+$$
$$+\|z-y\|=\rho(x,z)+\rho(z,y).$$

Наведемо деякі конкретні приклади метричних просторів.

1. R – множина всіх дійсних чисел. Відстань між x, $y \in R$ – це $\rho(x,y)=|x-y|$.

2. R^n – евклідовий простір зі скалярним добутком векторів $x=(\xi_1,\xi_2,...,\xi_n)$ і $y=(\eta_1,\eta_2,...,\eta_n)$, що обчислюються за формулою $(x,y)=\xi_1\eta_1+\xi_2\eta_2+...+\xi_n\eta_n$. Норма вектора

$$\|x\|=\sqrt{\xi_1^2+\xi_2^2+...+\xi_n^2},$$

а відстань між точками, поданими векторами x і y, є

$$\rho(x,y)=\sqrt{(\xi_1-\eta_1)^2+(\xi_2-\eta_2)^2+...+(\xi_n-\eta_n)^2}.$$

3. Простір $C_{[a,b]}$ неперервних функцій на відрізку $[a, b]$. Скалярний добуток $(f,g)=\int_a^b f(t)g(t)dt$, норма функції $\|f\|=\sqrt{\int_a^b f^2(t)dt}$ і відстань між функціями

$$\rho(f,g)=\sqrt{\int_a^b (f(t)-g(t))^2\,dt}\,.$$

Щоб пов'язати введені вище поняття з загальноприйнятою термінологією, нагадаємо деякі поняття топологічного характеру.

Відкритою кулею $S(a,\varepsilon)$ метричного простору V з центром a і радіусом $\varepsilon>0$ називається множина всіх точок $x\in V$ таких, що $\rho(x,a)<\varepsilon$. Відкриту кулю $S(a,\varepsilon)$ називають ε-*околом* точки a.

Нехай M – деяка непорожня множина точок метричного простору V. Точка $a\in V$ називається *граничною* для множини M, якщо будь-який її ε-окіл містить хоч би одну точку M, відмінну від a.

Множина всіх граничних точок називається *похідною множиною* і позначається M'.

Об'єднання $M\cup M'$ називається *замкненням* множини M і позначається $\overline{M}$.

Множина M називається *замкненою*, якщо $\overline{M}=M$. Отже, множина замкнена, якщо вона містить усі свої граничні точки.

Точка a множини M називається *внутрішньою*, якщо існує таке число $\varepsilon>0$, що $S(a,\varepsilon)\subset M$. Множина всіх внутрішніх точок множини M називається її *внутрішньою областю* і позначається M°.

Множина M називається *відкритою*, якщо $M=M^\circ$, тобто, коли всі її точки – внутрішні.

Повний метричний простір

Нескінченна послідовність $x_1,x_2,\dots,x_n,\dots$ точок метричного простору V називається *збіжною до точки* x, якщо кожний її ε-окіл $S(x,\varepsilon)$ містить усі точки x_m послідовності, починаючи з деякого номера $m\ge N_\varepsilon$ (N_ε залежить від ε).

Послідовність $x_1,x_2,\dots,x_m,\dots$ точок простору V називається *фундаментальною*, якщо для будь-якого $\varepsilon>0$ знайдеться таке число N_ε, що при $m,k \geq N_\varepsilon$ виконується нерівність

$$\rho(x_m, x_k) < \varepsilon.$$

Метричний простір V називається *повним*, якщо будь-яка фундаментальна послідовність його точок збігається.

Твердження. *Простір R^n з евклідовою метрикою $\rho(x,y) = \|x-y\|$ є повний.*

Доведення. Розглянемо довільну фундаментальну послідовність $x_1, x_2, ..., x_m, ...$ точок цього простору. Якщо $x_m = (\xi_{m1}, \xi_{m2}, ..., \xi_{mn})$, тоді

$$\|x_m - x_k\| = \sqrt{(\xi_{m1} - \xi_{k1})^2 + (\xi_{m2} - \xi_{k2})^2 + ... + (\xi_{mn} - \xi_{kn})^2}.$$

Для будь-якого числа $\varepsilon > 0$ існує таке число N_ε, що $\|x_m - x_k\| < \varepsilon$ для всіх $m,k \geq N_\varepsilon$. Тому $|\xi_{mi} - \xi_{ki}| < \varepsilon$ для всіх $m,k \geq N_\varepsilon$ ($i = \overline{1,n}$).

Нескінченна послідовність $\xi_{1i}, \xi_{2i}, ..., \xi_{mi}, ...$ – фундаментальна для кожного номера $1 \leq i \leq n$. Але множина R дійсних чисел – повний метричний простір з метрикою $\rho(x,y) = |x-y|$. Отже, послідовність $\xi_{1i}, \xi_{2i}, ..., \xi_{mi}, ...$ збігається до деякого числа ξ_i. Але тоді й послідовність $x_1, x_2, ..., x_m, ...$ збігається до точки $x = (\xi_1, \xi_2, ..., \xi_n)$.

Зауважимо, що метричний простір $C_{[a,b]}$ з відстанню

$$\rho(f,g) = \sqrt{\int_a^b (f(t) - g(t))\,dt}$$

між функціями $f(t)$, $g(t)$ не є повним. Цей факт встановлюється в курсі математичного аналізу.

Вправи

1. *Гіперплощиною* евклідового простору R^n називається множина $\Pi = \{x \in R^n : (a,x) = \beta\}$, $a \neq \theta$. Вона ділить увесь простір на півпростори $P = \{x \in R^n : (a,x) \leq \beta\}$ і $\Sigma = \{x \in R^n : (a,x) \geq \beta\}$.

Довести, що P, Σ, Π – замкнені множини. При цьому внутрішніми областями множин P і Σ є $P^\circ = \{x \in R^n: (a,x) < \beta\}$ і $\Sigma^\circ = \{x \in R^n: (a,x) > \beta\}$.

2. Нехай система лінійних рівнянь

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n = \beta_1, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \ldots + \alpha_{2n}\xi_n = \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n = \beta_m \end{cases}$$

з дійсними коефіцієнтами сумісна. Тоді множина L усіх її розв'язків $x = (\xi_1, \xi_2, \ldots, \xi_n) \in R^n$ (тобто лінійний багатовид L) – множина замкнена.

Довести це. Зокрема, довести, що будь-який лінійний підпростір – множина замкнена.

Те саме для сумісної системи лінійних нерівностей:

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n \le \beta_1, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \ldots + \alpha_{2n}\xi_n \le \beta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n \le \beta_m. \end{cases}$$

3. Довести, що множина $C_{[a,b]}$ всіх дійсних неперервних функцій на відрізку $[a, b]$ з відстанню

$$\rho(f,g) = \max_{a \le t \le b} |f(t) - g(t)|$$

між функціями $f(t)$ і $g(t)$ утворює повний метричний простір.

§ 7.8. Унітарний простір

Оскільки зміст цього параграфу безпосередньо випливає з теорії евклідових просторів, обмежимось розглядом лише окремих прикладів.

Нехай V – векторний простір над полем C комплексних чисел. Припустимо, що на V визначена комплекснозначна вектор-функція $\zeta=(x,y), x,y\in V; \zeta\in C$, яка має такі властивості:

1) $(\lambda x+\mu y, z)=\lambda\cdot(x,z)+\mu\cdot(y,z); x,y,z\in V; \lambda,\mu\in C$;

2) $(x,y)=\overline{(y,x)}$;

3) $x\neq\theta \Rightarrow (x,x)>0$ і $(\theta,\theta)=0$.

Тоді простір V називається *унітарним*, а функція двох аргументів (x, y), яка визначається на елементах простору V, називається *скалярним добутком*.

В унітарному просторі V *норма* вектора x визначається так само, як і в евклідовому просторі:

$$\|x\|=\sqrt{(x,x)}.$$

Основні властивості норми.

I. $x\neq\theta \Rightarrow \|x\|>0$ і $\|\theta\|=0$;

II. $\|\lambda x\|=|\lambda|\cdot\|x\|, \lambda\in C$;

III. $|(x,y)|\leq\|x\|\cdot\|y\|$ (нерівність Коші – Буняковського);

IV. $\|x+y\|\leq\|x\|+\|y\|$ (нерівність трикутника).

Скінченновимірний унітарний простір є арифметичним комплексним векторним простором C^n зі скалярним добутком векторів $x=(\xi_1,\xi_2,...,\xi_n)$ і $y=(\eta_1,\eta_2,...,\eta_n)$, обчисленим за правилом

$$(x,y)=\xi_1\overline{\eta}_1+\xi_2\overline{\eta}_2+...+\xi_n\overline{\eta}_n.$$

У ньому норма векторів знаходиться за формулою

$$\|x\|=\sqrt{|\xi_1|^2+|\xi_2|^2+...+|\xi_n|^2},$$

а нерівність Коші – Буняковського набуває вигляду

$$|\xi_1\overline{\eta}_1+\xi_2\overline{\eta}_2+...+\xi_n\overline{\eta}_n|^2\leq\left(|\xi_1|^2+|\xi_2|^2+...+|\xi_n|^2\right)\times$$
$$\times\left(|\eta_1|^2+|\eta_2|^2+...+|\eta_n|^2\right).$$

Надалі основну увагу приділятимемо унітарному простору C^n. При цьому означення ортогонального і ортонормованого базисів унітарного простору такі самі, як і для евклідового простору.

Для матриці

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

з елементами із поля C комплексних чисел спряженою називається матриця

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

Матриця переходу від ортонормованого базису до іншого ортонормованого базису називається *унітарною*. Унітарна матриця U характеризується властивостями

$$U^{*}\cdot U=U\cdot U^{*}=E\ .$$

Лінійний оператор $\mathcal{B}$ унітарного простору C^{n} називається *спряженим* з лінійним оператором $\mathcal{A}$, якщо виконується рівність

$$(\mathcal{A}(x), y) = (x, \mathcal{B}(y))$$

для будь-яких $x,y\in C^{n}$. Він позначається як $\mathcal{A}^{*}$.

Пропонується самостійно довести такі властивості лінійних операторів:

1. Для будь-якого лінійного оператора $\mathcal{A}$ існує єдиний лінійний оператор $\mathcal{A}^{*}$, спряжений з ним.

Якщо в ортонормованому базисі лінійному оператору відповідає матриця A, тоді спряженому оператору у тому ж базисі відповідає спряжена матриця $\mathcal{A}^{*}$.

2. Для будь-якого лінійного оператора $\mathcal{A}$ унітарного простору існує ортонормований базис, в якому матриця цього оператора має трикутну форму (теорема Шура).

3. Лінійний оператор $\mathcal{U}$ називається унітарним, якщо він зберігає скалярний добуток векторів:

$$(x, y) = (\mathcal{U}(x), \mathcal{U}(y)).$$

В ортонормованому базисі такому оператору відповідає унітарна матриця.

4. Лінійний оператор $\mathcal{A}$ називається *нормальним*, якщо він переставний зі своїм спряженим оператором, тобто

$$\mathcal{A} \cdot \mathcal{A}^* = \mathcal{A}^* \cdot \mathcal{A}.$$

Встановити послідовно:

а) якщо x – власний вектор нормального оператора $\mathcal{A}$ з власним значенням λ, тоді той самий вектор буде власним вектором оператора $\mathcal{A}^*$ з власним значенням $\bar{\lambda}$;

б) власні вектори нормального оператора з різними власними значеннями ортогональні;

в) для нормального оператора існує ортонормований базис простору, складений із власних векторів цього оператора (структура його проста).

Розділ 8. СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ І ПСЕВДООБЕРНЕНІ МАТРИЦІ

§ 8.1. Псевдорозв'язки систем лінійних рівнянь

Розглянемо систему лінійних рівнянь з дійсними коефіцієнтами

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n = \zeta_1, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \ldots + \alpha_{2n}\xi_n = \zeta_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n = \zeta_m \end{cases}$$

із заданим вектором $z = (\zeta_1, \zeta_2, \ldots, \zeta_m)^T$ і змінними векторами $x = (\xi_1, \xi_2, \ldots, \xi_n)^T$. Якщо $H = (\alpha_{ik})$ – матриця цієї системи (розмірності $m \times n$), то запишемо її (систему) в матричній формі

$$Hx = z. \tag{1}$$

Для матриці H використаємо відомі (див. § 7.4) розклади евклідових просторів R^n і R^m в ортогональну пряму суму

$$R^n = \mathrm{Ker}H \perp \mathrm{Im}H^T, \quad R^m = \mathrm{Im}H \perp \mathrm{Ker}H^T,$$

а також рівності

$$\mathrm{Ker}\, H = \mathrm{Ker}\left(H^T \cdot H\right), \ \ \mathrm{Im}H^T = \mathrm{Im}\left(H^T \cdot H\right).$$

Теорема 1 (Фредгольма). *Для того щоб система лінійних рівнянь (1) була сумісною, необхідно і достатньо, щоб вектор правої частини системи був ортогональним будь-якому розв'язку* $y = (\eta_1, \eta_2, \ldots, \eta_m)^T$ *системи лінійних однорідних рівнянь*

$$H^T y = \theta. \tag{2}$$

Доведення. *Необхідність.* Нехай система (1) сумісна і x_0 – який-небудь її розв'язок, а y – довільний розв'язок системи (2). Тоді

$$(z,y)=(Hx_0,y)=\left(x_0,H^T y\right)=(x_0,\theta)=0\,.$$

Достатність. Маємо розклад

$$R^m=\mathrm{Im}H\perp\mathrm{Ker}H^T\,.$$

Нехай $(z,y)=0$ для будь-якого розв'язку y системи (2), тобто для будь-якого вектора $y\in\mathrm{Ker}H^T$. Тоді маємо включення $z\in\mathrm{Im}H$. Це означає, що існує вектор $x_0\in R^n$, для якого $z=Hx_0$. Отже, система (1) сумісна.

Означення. *Якщо* $x\in R^n$, *то для заданого вектора* z *різниця* $y=z-Hx$ *називається відхилом системи* (1). Зокрема, якщо x – розв'язок системи, то для нього відхил нульовий.

Якщо ця система несумісна, то доцільно відшукати такі вектори x^*, для яких відхил $y^*=z-Hx^*$ має мінімальну норму $\|z-Hx^*\|$. У цьому й полягає метод найменших квадратів; обчислюються вектори, що мінімізують форму

$$F(x)=\|z-Hx\|^2\,.$$

Такі вектори x^* називаються *псевдорозв'язками* даної системи лінійних рівнянь.

Теорема 2. *Множина всіх псевдорозв'язків системи лінійних рівнянь*

$$Hx=z$$

непорожня і збігається з багатовидом усіх розв'язків системи

$$Hx=\hat{z}\,, \tag{3}$$

де $\hat{z}$ – *ортогональна проекція вектора* z *на* $\mathrm{Im}\,H$.

Доведення. Знову скористаємось розкладом на ортогональні складові

$$R^m=\mathrm{Im}H\perp\mathrm{Ker}H^T\,.$$

Тоді $z=\hat{z}+\overset{\perp}{z}$, де $\hat{z}$ – ортогональна проекція z на $\mathrm{Im}H$ і $\overset{\perp}{z}\in\mathrm{Ker}H^T$. Для будь-якого $x\in R^n$ виконується рівність

$$F(x) = \|z - Hx\|^2 = \left\|\left(\hat{z} - Hx\right) + \overset{\perp}{z}\right\|^2 = \|\hat{z} - Hx\|^2 + \left\|\overset{\perp}{z}\right\|^2,$$

оскільки $\hat{z} - Hx \in \mathrm{Im}H$. Форма $F(x)$ досягає свого мінімуму $\left\|\overset{\perp}{z}\right\|^2$ лише для тих векторів x, для яких $\|\hat{z} - Hx\| = 0$, тобто для яких виконується рівність $Hx = \hat{z}$. Оскільки $\hat{z} \in \mathrm{Im}H$, то такі вектори неодмінно існують.

§ 8.2. Псевдообернені матриці Мура – Пенроуза

Нехай L – лінійний підпростір евклідового простору R^n. Квадратна матриця A називається *проекційною* на підпростір L, якщо для будь-якого вектора $z \in R^n$ вектор Az є ортогональною проекцією вектора z на L.

Твердження 1. *Для того щоб дійсна квадратна матриця A порядку n була проекційною для деякого лінійного підпростору, достатньо, щоб вона мала такі дві властивості:*

I. *Симетричність:* $A^T = A$.

II. *Ідемпотентність:* $A^2 = A$.

Доведення. Нехай $A^T = A$ і $A^2 = A$. Скористаємось відомою властивістю симетричної матриці:

$$R^n = \mathrm{Im}A \perp \mathrm{Ker}A$$

(теорема § 7.4). Тому кожний вектор $z \in R^n$ подається у вигляді суми $z = \hat{z} + \overset{\perp}{z}$, $\hat{z} \in \mathrm{Im}A$, $\overset{\perp}{z} \in \mathrm{Ker}A$; $\hat{z}$ – ортогональна проекція z на $\mathrm{Im}A$. Звідси $Az = A\hat{z}$ і $\hat{z} = Aw$ для деякого вектора w. Тому

$$Az = A(Aw) = A^2 w = Aw = \hat{z}.$$

Теорема 1. *Якщо A – симетрична матриця порядку n, то існує границя*

$$\lim_{\delta \to 0} \left(A + \delta E\right)^{-1} A = \hat{A}.$$

Матриця $\hat{A}$ – проекційна на підпростір $L = \mathrm{Im}\, A$.

Доведення. Для симетричної матриці A існує така ортогональна матриця X, для якої $A = XДX^T$, де $Д$ – діагональна матриця: $Д = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$; $\lambda_1, \lambda_2, ..., \lambda_n$ – її власні значення (див. § 7.5). Позначимо через δ_0 додатне число, менше абсолютних значень $|\lambda_k|$ для всіх $\lambda_k \neq 0$ (ненульових власних значень). Тоді для будь-яких чисел δ, для яких $0 < |\delta| < \delta_0$, матриця

$$A + \delta E = X \cdot \mathrm{diag}(\lambda_1 + \delta, \lambda_2 + \delta, ..., \lambda_n + \delta) \cdot X^T$$

оборотна. Оскільки $R^n = \mathrm{Im}A \perp \mathrm{Ker}A$, то довільний вектор $z \in R^n$ подається у вигляді $z = \hat{z} + \overset{\perp}{z}$, де $\hat{z} \in \mathrm{Im}A$, $\overset{\perp}{z} \in \mathrm{Ker}A$. Для деякого вектора w матимемо рівність $\hat{z} = Aw$. Тоді

$$(A + \delta E)^{-1} Az = (A + \delta E)^{-1} A\hat{z} = (A + \delta E)^{-1} A^2 w =$$
$$= X(Д + \delta E)^{-1} X^T XД^2 X^T w = X\left[(Д + \delta E)^{-1} Д^2\right] X^T w.$$

Оскільки

$$(Д + \delta E)^{-1} Д^2 = \mathrm{diag}\left(\frac{\lambda_1^2}{\lambda_1 + \delta}, \frac{\lambda_2^2}{\lambda_2 + \delta}, ..., \frac{\lambda_n^2}{\lambda_n + \delta}\right),$$

то

$$\lim_{\delta \to 0} (Д + \delta E)^{-1} Д^2 = Д.$$

Тому існує границя

$$\lim_{\delta \to 0} (A + \delta E)^{-1} Az = XДX^T w = Aw = \hat{z}.$$

У результаті матимемо

$$\lim_{\delta \to 0} (A + \delta E)^{-1} Az = \hat{z},$$

де $\hat{z}$ – ортогональна проекція вектора z на лінійний підпростір $\mathrm{Im}A$. Якщо замість вектора z в останній рівності послідовно підставляти вектори базису $e_1 = (1, 0, ..., 0)^T$, $e_2 = (0, 1, ..., 0)^T$, ..., $e_n = (0, 0, ..., 1)^T$, то праворуч дістанемо вектори $\hat{e}_1, \hat{e}_2, ..., \hat{e}_n$.

Побудуємо матрицю $\hat{A} = (\hat{e}_1, \hat{e}_2, ..., \hat{e}_n)$, стовпці якої збігаються з цими векторами. Тоді

$$\lim_{\delta \to 0} (A + \delta E)^{-1} A = \hat{A} \quad \text{і} \quad \hat{A} z = \hat{z},$$

де $\hat{z}$ – ортогональна проекція z на підпростір $\mathrm{Im} A$.

Матриця $\hat{A}$, яка є границею симетричних матриць, сама також симетрична. Властивість ідемпотентності $\hat{A}^2 = \hat{A}$ очевидна.

Висновок. Якщо A – симетрична невироджена матриця, то $\hat{A} = \lim_{\delta \to 0} (A + \delta E)^{-1} A = E$ – одинична матриця.

Лема. *Для будь-якої дійсної матриці H розмірності $m \times n$ і досить малого числа $\delta \neq 0$ матриця*

$$H^T H + \delta^2 E$$

оборотна.

Доведення леми наведено в теоремі 1 для симетричної матриці $A = H^T H$.

Теорема 2. *Для матриці H розмірності $m \times n$ існує*

$$\lim_{\delta \to 0} \left(H^T H + \delta^2 E \right)^{-1} H^T = H^+ . \quad (1)$$

Розмірність матриці H^+ дорівнює $n \times m$. Для будь-якого вектора $z \in R^m$ вектор

$$v = H^+ z \in \mathrm{Im} H^T \quad (2)$$

і задовольняє рівність

$$H v = \hat{z}, \quad (3)$$

де $\hat{z}$ – ортогональна проекція z на $\mathrm{Im}\, H$.

Доведення. Спочатку доведемо існування границі

$$\lim_{\delta \to 0} \left(H^T H + \delta^2 E \right)^{-1} H^T z$$

для довільного вектора $z \in R^m$.

Використаємо відомий розклад

$$R^m = \mathrm{Im} H \perp \mathrm{Ker} H^T .$$

Тоді $z = \hat{z} + \overset{\perp}{z}, \hat{z} \in \operatorname{Im} H, \overset{\perp}{z} \in \operatorname{Ker} H^T$. Оскільки для деякого вектора w маємо $Hw = \hat{z}$, то

$$H^T z = H^T \left(\hat{z} + \overset{\perp}{z}\right) = H^T \hat{z} = H^T Hw$$

і

$$\left(H^T H + \delta^2 E\right)^{-1} H^T z = \left(H^T H + \delta^2 E\right)^{-1} H^T Hw. \qquad (4)$$

Для симетричної матриці $A = H^T H$ за теоремою 1 існує границя

$$\lim_{\delta \to 0} \left(H^T H + \delta^2 E\right)^{-1} H^T H = \hat{A},$$

причому $\hat{A}$ – проекційна матриця на підпростір $\operatorname{Im} A$ = $= \operatorname{Im}\left(H^T H\right)$. На підставі рівності $\operatorname{Im}\left(H^T H\right) = \operatorname{Im} H^T$ вона буде проекційною матрицею на $\operatorname{Im} H^T$. Тому

$$\lim_{\delta \to 0} \left(H^T H + \delta^2 E\right)^{-1} H^T Hw = \hat{A}w = \hat{w},$$

де $\hat{w}$ – ортогональна проекція вектора w на $\operatorname{Im} H^T$. Із рівності (4) дістаємо

$$\lim_{\delta \to 0} \left(H^T H + \delta^2 E\right)^{-1} H^T z = \hat{w}. \qquad (5)$$

Із останньої рівності випливає, що вектор $v = \hat{w}$ не залежить від вибору w в рівнянні $Hw = \hat{z}$.

Якщо в рівності (5) вектору z послідовно надавати значень базисних векторів $e_1 = (1, 0, ..., 0)^T$, $e_2 = (0, 1, ..., 0)^T$, ..., $e_m = (0, 0, ..., 1)^T$, то дістанемо систему векторів

$$v_k = \lim_{\delta \to 0} \left(H^T H + \delta^2 E\right)^{-1} H^T e_k \ \left(k = \overline{1, m}\right), \qquad (6)$$

яка належить $\operatorname{Im} H^T$. Матриця $H^+ = \left(v_1, v_2, ..., v_m\right)$ зі стовпцями $v_1, v_2, ..., v_m$ має розмірність $n \times m$ і

$$\lim_{\delta \to 0} \left(H^T H + \delta^2 E\right)^{-1} H^T = H^+, \ v \equiv \hat{w} = H^+ z \in \operatorname{Im} H^T.$$

Із розкладу

$$R^n = \mathrm{Im}H^T \perp \mathrm{Ker}H$$

дістаємо $w = \hat{w} + \overset{\perp}{w}$, $\hat{w} \in \mathrm{Im}H^T$, $\overset{\perp}{w} \in \mathrm{Ker}H$. Тому

$$\hat{z} = Hw = H\hat{w}$$

і для вектора $v \equiv \hat{w}$ випливає вимога теореми – рівність (3).

Теорему доведено.

Означення. *Якщо* $H = (\alpha_{ik})$ – *матриця розмірності* $m \times n$ *з дійсними коефіцієнтами, то матриця* $H^+ = \lim\limits_{\delta \to 0}\left(H^T H + \delta^2 E\right)^{-1} H^T$ *розмірності* $n \times m$ *називається псевдооберненою для матриці H.*

Приклад 1. Знайти H^+ для матриці

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

Розв'язання. Для матриці H^+ маємо

$$H^T H + \delta^2 E = \begin{pmatrix} 1+\delta^2 & 1 \\ 1 & 1+\delta^2 \end{pmatrix},$$

$$\left(H^T H + \delta^2 E\right)^{-1} = \frac{1}{\delta^2(2+\delta^2)}\begin{pmatrix} 1+\delta^2 & -1 \\ -1 & 1+\delta^2 \end{pmatrix},$$

$$\left(H^T H + \delta^2 E\right)^{-1} H^T = \frac{1}{2+\delta^2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

Тому

$$H^+ = \lim_{\delta \to 0}(H^T H + \delta^2 E)^{-1} H^T = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2} H^T .$$

Приклад 2. Знайти $Д^+$ для квадратної діагональної матриці $Д = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$.

Розв'язання. Маємо

$$Д^{+} = \lim_{\delta\to 0} (Д^2 + \delta^2 E)^{-1} Д =$$

$$= \lim_{\delta\to 0} \operatorname{diag}\left(\frac{\lambda_1}{\lambda_1^2 + \delta^2}, \frac{\lambda_2}{\lambda_2^2 + \delta^2}, \ldots, \frac{\lambda_n}{\lambda_n^2 + \delta^2} \right) = \operatorname{diag}\left(\lambda_1^{+}, \lambda_2^{+}, \ldots, \lambda_n^{+}\right),$$

де

$$\lambda_k^{+} = \begin{cases} \lambda_k^{-1}, \text{ якщо } \lambda_k \neq 0, \\ 0, \quad \text{якщо } \lambda_k = 0. \end{cases}$$

Приклад 3. Обчислити псевдообернену матрицю A^{+} для симетричної матриці A порядку n.

Розв'язання. Для матриці A знайдемо таку ортогональну матрицю X, яка перетворенням подібності зводить її до діагонального виду. Тоді її можна подати у вигляді добутку $A = XДX^{T}$, де $Д = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $\lambda_1, \lambda_2, \ldots, \lambda_n$ – її власні значення. Тому

$$A^{+} = \lim_{\delta\to 0} (A^2 + \delta^2 E)^{-1} A = \lim_{\delta\to 0} X(Д^2 + \delta^2 E)^{-1} ДX^{T} =$$

$$= X\left[\lim_{\delta\to 0} (Д^2 + \delta^2 E)^{-1} Д \right] X^{T} = XД^{+}X^{T}.$$

Будь-яка симетрична матриця A переставна з A^{+}: $AA^{+} = A^{+}A$. Ця рівність легко встановлюється, якщо скористатися наведеними вище поданнями матриць A і A^{+} у вигляді діагональних матриць.

Зауважимо, що не кожна матриця H переставна зі своєю псевдоoберненою матрицею H^{+}.

У прикладі 1 для $H = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ мали $H^{+} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ і $HH^{+} \neq H^{+}H$.

Псевдообернену H^{+} для матриці H можна відшукати як границю

$$\lim_{\delta\to 0} H^T(HH^T+\delta^2 E)^{-1},$$

де E – одинична матриця відповідного порядку.

Щоб у цьому переконатися, розглянемо рівність

$$H^T(HH^T+\delta^2 E)=(H^T H+\delta^2 E)H^T.$$

Звідси знаходимо

$$(H^T H+\delta^2 E)^{-1}H^T=H^T(HH^T+\delta^2 E)^{-1}.$$

Тому

$$H^+=\lim_{\delta\to 0} H^T(HH^T+\delta^2 E)^{-1}.$$

Твердження 2. *Якщо H^+ – псевдообернена матриця для H, то H^+H – проекційна матриця на лінійний підпростір* $\mathrm{Im}(H^T)$, а HH^+ *– проекційна матриця на* $\mathrm{Im}\, H$.

Доведення. Із означення

$$H^+=\lim_{\delta\to 0}(HH^T+\delta^2 E)^{-1}H^T$$

дістанемо

$$H^+H=\lim_{\delta\to 0}(H^T H+\delta^2 E)^{-1}H^T H.$$

За теоремою 1 ця границя для симетричної матриці $A=H^T H$ є проекційною матрицею $\hat{A}$ на лінійний підпростір $\mathrm{Im}A = \mathrm{Im}(H^T H)$. Але відомо, що $\mathrm{Im}(H^T H)=\mathrm{Im}H^T$. Аналогічне твердження зберігає силу і для HH^+.

Твердження 3. *Якщо H – квадратна невироджена матриця, то $H^+=H^{-1}$.*

Доведення. Для невиродженої матриці H порядку n маємо $\mathrm{Im}\, H=\mathrm{Im}\, H^T=R^n$. Тому ортогональна проекція $\hat{z}=H^+Hz$ вектора z на R^n збігається із z. Рівність $H^+Hz=z$ для довільного вектора z означає, що $H^+H=E$, де E – одинична матриця, тобто $H^+=H^{-1}$ (див. теорему 1).

Твердження 4. $\mathrm{Ker}H=\mathrm{Ker}(H^+H)$.

Доведення. Включення $\mathrm{Ker}H\subseteq\mathrm{Ker}(H^+H)$ очевидне. Для встановлення оберненого включення оберемо $x\in\mathrm{Ker}(H^+H)$;

$H^+Hx = \theta$. На підставі рівності $R^n = \mathrm{Im}H^T \perp \mathrm{Ker}H$ маємо розклад $x = \hat{x} + \overset{\perp}{x}$, де $\hat{x} \in \mathrm{Im}H^T$, $\overset{\perp}{x} \in \mathrm{Ker}H$. Але H^+H – проекційна матриця на $\mathrm{Im}H^T$. Тому $\hat{x} = H^+Hx = \theta$ і, отже,

$$x = \overset{\perp}{x} \in \mathrm{Ker}\, H.$$

Остаточно $\mathrm{Ker}(H^+H) \subseteq \mathrm{Ker}H$.

Твердження 5. *Для матриці H мають місце рівності*:

I. $H^+ = (H^TH)^+H^T$.

II. $(H^T)^+ = (H^+)^T$.

III. $H^+ = H^T(HH^T)^+$.

Доведення рівності I. Скористаємось рівністю (4):

$$H^+z = \lim_{\delta\to 0}(H^TH + \delta^2E)^{-1}H^THw,$$

де z – довільний вектор із R^m; w – вектор із R^n, що задовольняє рівність $\hat{z} = Hw$; $\hat{z}$ – ортогональна проекція z на лінійний підпростір $\mathrm{Im}\, H$.

Симетричну матрицю $A = H^TH$ можна подати у вигляді добутку $A = XДX^T$, де $Д$ = diag $(\lambda_1, \lambda_2, ..., \lambda_n)$ і X – вдало підібрана ортогональна матриця.

Дістанемо рівність

$$H^+z = X\left[\lim_{\delta\to 0}(Д + \delta^2E)^{-1}Д\right]X^Tw. \tag{7}$$

Оскільки

$$(H^TH)^+ = \lim_{\delta\to 0}\left[(H^TH)^2 + \delta^2E\right]^{-1}H^TH,$$

матимемо рівність

$$(H^TH)^+H^Tz = \lim_{\delta\to 0}\left[(H^TH)^2 + \delta^2E\right]^{-1}H^THH^Tz.$$

Знову використаємо розклад у пряму ортогональну суму:

$$R^m = \mathrm{Im}H \perp \mathrm{Ker}H^T.$$

Тоді $z = \hat{z} + \overset{\perp}{z}$, де $\hat{z} \in \operatorname{Im} H$, $\overset{\perp}{z} \in \operatorname{Ker} H^T$; $H^T z = H^T \hat{z}$, $\hat{z} = Hw$. Тому

$$(H^T H)^+ H^T z = \left\{ \lim_{\delta \to 0} \left[(H^T H)^2 + \delta^2 E \right]^{-1} \left(H^T H \right)^2 \right\} w =$$

$$= X \left\{ \lim_{\delta \to 0} \left[Д^2 + \delta^2 E \right]^{-1} Д^2 \right\} X^T w .$$

На підставі рівності $\lim\limits_{\delta \to 0} \left(Д^2 + \delta^2 E \right)^{-1} Д^2 =$ $= \lim\limits_{\delta \to 0} \left(Д + \delta^2 E \right)^{-1} Д$ випливає

$$(H^T H)^+ H^T z = X \left[\lim_{\delta \to 0} (Д + \delta^2 E)^{-1} Д \right] X^T w . \qquad (8)$$

Порівнявши рівності (7) і (8), знайдемо

$$H^+ z = (H^T H)^+ H^T z . \qquad (9)$$

Оскільки остання рівність справедлива для будь-якого вектора z, то

$$H^+ = (H^T H)^+ H^T .$$

На доведеннях рівностей II і III не зупиняємось.

Твердження 6. *Виконуються рівності*

IV. $HH^+ H = H$.

V. $H^+ HH^+ = H^+$.

Доведення. Оскільки HH^+ – проекційна матриця на лінійний підпростір ImH, то для вектора $z = Hx$ виконується рівність $HH^+ z = z$. Тому $HH^+ Hx = Hx$ для будь-якого вектора $x \in R^n$. Це означає, що $HH^+ H = H$.

Для доведення рівності V скористаємося співвідношенням $H^+ = (H^T H)^+ H^T$. Оскільки $H = HH^+ H$, то $H^+ = (H^T H)^+ \times$ $\times (HH^+ H)^T = (H^T H)^+ H^T (HH^+)^T$. Але HH^+ – симетрична матриця. Тому

$$H^+ = (H^T H)^+ H^T HH^+ = [(H^T H)^+ H^T] HH^+ = H^+ HH^+ .$$

Теорема Пенроуза *(без доведення). Нехай H – дійсна матриця розмірності $m \times n$. Для того щоб матриця F розмірності $n \times m$ збігалася з псевдооберненою H^+, необхідно і достатньо, щоб вона мала такі властивості:*

а) HF і FH *– матриці симетричні;*

б) $HFH = H$;

в) $FHF = F$.

Твердження 7. *Для будь-якої матриці H виконуються рівності:*

VI. $H^+ H H^T = H^T$.

VII. $H^T H H^+ = H^T$.

Доведення. Із рівності $H = HH^+ H$ випливає, що

$$H^T = (HH^+ H)^T = (H^+ H)^T H^T = H^+ H H^T;$$

$$H^T = (HH^+ H)^T = H^T (HH^+)^T = H^T H H^+.$$

Вправи

1. Показати, що для симетричної матриці A виконуються такі рівності:

а) $AA^+ = A^+ A$; б) $(A^k)^+ = (A^+)^k$ $(k \ge 1)$;

в) $(A^k)^+ A^k = (A^+ A)^k = A^+ A$ $(k \ge 1)$;

г) $A^k A^+ = A^{k-1}$ $(k \ge 2)$; д) $(A^+)^k A = (A^+)^{k-1}$ $(k \ge 2)$.

2. Довести, що коли X – ортогональна матриця, то

$$(HX)^+ = X^+ H^+ \text{ і } (XH)^+ = H^+ X^+.$$

3. Довести

$$(H^T H)^+ = H^+ (H^T)^+, (HH^T)^+ = (H^T)^+ H^+.$$

4. Довести

$$\operatorname{Ker} H^+ = \operatorname{Ker} H^T, \operatorname{Im} H^+ = \operatorname{Im} H^T.$$

5. Довести

$$(H^+)^+ = H.$$

§ 8.3. Застосування псевдооберненої матриці для розв'язання матричних рівнянь

Знову розглянемо систему лінійних рівнянь

$$Hx = z. \tag{1}$$

Тут H – матриця розмірності $m \times n$, $x = (\xi_1, \xi_2, ..., \xi_n)^T$ – стовпець невідомих, $z = (\zeta_1, \zeta_2, ..., \zeta_m)^T$.

Поряд із системою (1), складемо нову систему лінійних рівнянь

$$H^T Hx = H^T z, \tag{2}$$

яку будемо називати *нормальною* системою для даної системи (1). Виявляється, що нормальна система рівнянь завжди сумісна. Для системи (2) запишемо однорідну систему $(H^T H)^T y = \theta$, тобто

$$H^T Hy = \theta, \tag{2°}$$

їй відповідну на підставі умови теореми Фредгольма. За цією теоремою система (2) сумісна, якщо будь-який розв'язок однорідної системи (2°) ортогональний вектору $H^T z$. Переконаємось у цьому. Нехай y – довільний розв'язок системи (2°). Тоді

$$(Hy, Hy) = (y, H^T Hy) = (y, \theta) = 0$$

і тому $Hy = \theta$. Звідси $(y, H^T z) = (Hy, z) = (\theta, z) = 0$. Отже, система (2) – сумісна.

Твердження 1. *Множина всіх псевдорозв'язків системи лінійних рівнянь* (1) *є рішенням відповідної нормальної системи* (2).

Доведення. Відомо, що множина псевдорозв'язків системи (1) збігається з множиною розв'язків сумісної системи

$$Hx = \hat{z},$$

де $\hat{z}$ – ортогональна проекція вектора z на підпростір ImH. Нехай x^* – псевдорозв'язок системи (1), тобто $Hx^* = \hat{z}$. Оскільки HH^+ – проекційна матриця на лінійний підпростір

$\operatorname{Im} H$ (§ 8.2, твердження 2), то $\hat{z} = HH^+ z$. Із рівності $Hx^* = \hat{z} = HH^+ z$ дістаємо $H^T Hx^* = H^T HH^+ z$. Але $H^T HH^+ = H^T$ (§ 8.2, твердження 7, властивість VII). Тому $H^T Hx^* = H^T z$: псевдорозв'язок x^* задовольняє нормальне рівняння (2).

Псевдообернені матриці використовуються при знаходженні розв'язків матричних рівнянь виду

$$AXB = C, \tag{4}$$

де A, B, C – матриці відповідних розмірностей; X – шукана матриця.

Твердження 2. *Лінійне матричне рівняння* (4) *має розв'язок тоді й лише тоді*, *коли виконується рівність*

$$AA^+CB^+B = C. \tag{5}$$

Доведення. *Необхідність.* Нехай X_0 – деякий розв'язок матричного рівняння (4):

$$AX_0B = C.$$

Тоді

$$AA^+CB^+B = AA^+(AX_0B)B^+B = (AA^+A)X_0(BB^+B) = AX_0B = C.$$

Достатність. Нехай має місце рівність (5). Тоді $X_0 = A^+CB^+$ – розв'язок рівняння (4).

Вправи

1. Яким чином твердження 2 пов'язане з твердженням: матричне рівняння (4) має розв'язок тоді й лише тоді, коли ранг матриці дорівнює рангу розширеної матриці A (теорема Кронекера – Капеллі)?

2. Відповісти на питання 1 для рівняння $XB = C$.

3. Довести, що загальним розв'язком однорідного матричного рівняння $AYB = 0$ буде $Y = M - A^+AMBB^+$, де M – довільна матриця, розмірність якої збігається з розмірністю матриці Y.

4. Довести сумісність нормального рівняння $H^T Hx = H^T z$ за допомогою твердження 2.

§ 8.4. Метод обчислення псевдооберненої матриці

Наведемо один із методів обчислення псевдооберненої матриці, побудований на основі теореми Гамільтона–Келі.

Нехай A – симетрична матриця і

$$\mathcal{X}(t) = \left|tE - A\right| = t^n - \beta_1 t^{n-1} - \ldots - \beta_{n-1} t - \beta_n -$$

її характеристичний многочлен. Подамо його у вигляді

$$\mathcal{X}(t) = \alpha t^m (t\varphi(t) - 1),$$

де $\alpha \neq 0$, $m \geq 0$, $\varphi(t)$ – однозначно визначений многочлен.

На основі теореми Гамільтона–Келі

$$\mathcal{X}(A) = 0 \Rightarrow \alpha A^m (A\varphi(A) - E) = 0 .$$

Якщо матриця A – оборотна, то $m = 0$ і останнє співвідношення набуде вигляду

$$\alpha(A\varphi(A) - E) = 0 . \tag{1}$$

Отже, $A^+ = A^{-1} = \varphi(A)$.

Надалі скористаємось властивостями симетричних матриць (§ 8.2, вправа 1):

а) $AA^+ = A^+ A$;

б) $\left(A^k\right)^+ = \left(A^+\right)^k$ $(k \geq 1)$;

в) $\left(A^k\right)^+ A^k = \left(A^+ A\right)^k = A^+ A$ $(k \geq 1)$;

г) $A^k A^+ = A^{k-1}$, $\left(A^+\right)^k A = \left(A^+\right)^{k-1}$ $(k \geq 2)$.

Коли $m \geq 1$, виконується рівність

$$A^m = A^{m+1}\varphi(A) \tag{2}$$

і тому $(A^+)^{m+1} A^m = (A^+)^{m+1} A^{m+1}\varphi(A)$. Маємо $(A^+)^{m+1} A^{m+1} = (A^+ A)^{m+1} = A^+ A$ – проекційна матриця. Але $(A^+)^{m+1} A^m = A^+ (A^+)^m A^m = A^+ A^+ A = A^+$.

Дістаємо рівність

$$A^+ = AA^+\varphi(A) \tag{3}$$

Оскільки $AA^+\varphi(A) = AA^+[\varphi(A) - \varphi(0)E] + \varphi(0)AA^+$ і $\varphi(A) - \varphi(0)E$ містить доданки лише з додатними степенями матриці A, то з рівності $A^+A^{k+1} = A^k$ $(k \geq 1)$ випливає

$$AA^+\left[\varphi(A) - \varphi(0)E\right] = \varphi(A) - \varphi(0)E .$$

Тому з рівності (3) знаходимо

$$A^+ = \varphi(A) - \varphi(0)E + \varphi(0)AA^+ . \tag{4}$$

Звідси

$$AA^+ = A\varphi(A) - \varphi(0)A + \varphi(0)A^2A^+ .$$

Оскільки $A^2A^+ = A$, то $AA^+ = A\varphi(A)$. Підставимо останній вираз AA^+ в рівність (4). Дістанемо формулу

$$A^+ = \varphi(A) + \varphi(0)[A\varphi(A) - E] . \tag{5}$$

Висновок. У всіх випадках для симетричної матриці псевдообернена обчислюється за формулою

$$A^+ = \varphi(A) + \varphi(0)[A\varphi(A) - E] .$$

Якщо ж A – оборотна, то $A^+ = \varphi(A)$.

Нехай тепер H – довільна матриця розмірності $m \times n$. Обчислимо для симетричної матриці $A = H^TH$ характеристичний многочлен $\mathcal{X}(t)$ і подамо його у вигляді

$$\mathcal{X}(t) = \alpha \cdot t^m (t\varphi(t) - 1) .$$

Тоді

$$A^+ = (H^TH)^+ = \varphi(H^TH) + \varphi(0)[H^TH\varphi(H^TH) - E] .$$

Із співвідношення $AA^+ = A^+A = A\varphi(A)$ одержимо рівність

$$H^TH\varphi(H^TH) = (H^TH)^+H^TH = [(H^TH)^+H^T]H .$$

Скористаємось тепер відомою нам рівністю $(H^T H)^+ H^T = H^+$ (§ 8.2, твердження 5), тоді

$$(H^T H)^+ = \varphi(H^T H) + \varphi(0)[H^+ H - E]$$

і, отже,

$$H^+ = (H^T H)^+ H^T = \varphi(H^T H)H^T + \varphi(0)[H^+ H H^T - H^T].$$

Скориставшись $H^+ H H^T = H^T$ (§ 8.2, твердження 7), дістанемо коротку формулу

$$H^+ = \varphi(A)H^T = \varphi(H^T H)H^T. \tag{6}$$

Якщо r – ранг матриці H, тоді ранг симетричної матриці $A = H^T H$ також дорівнює r. Це, наприклад, випливає з рівності $\mathrm{Im}H^T = \mathrm{Im}(H^T H)$, оскільки ранги матриць H і $A = H^T H$ збігаються з розмірністю цього лінійного простору. Можна послатися також на рівність $\mathrm{Ker}H = \mathrm{Ker}(H^T H)$.

Кількість нульових коренів характеристичного многочлена $\mathcal{X}(t) = t^n - \beta_1 t^{n-1} - \dots - b_n$ матриці $A = H^T H$ дорівнює $n-r$. Тому останнім ненульовим коефіцієнтом його буде $-\beta_r$:

$$\mathcal{X}(t) = t^n - \beta_1 t^{n-1} - \dots - \beta_r t^{n-r}.$$

Отже,

$$\mathcal{X}(t) = \beta_r t^{n-r}(t\varphi(t) - 1),$$

де

$$\varphi(t) = \frac{1}{\beta_r}\left(t^{r-1} - \sum_{k=1}^{r-1} \beta_k t^{r-k-1}\right).$$

Приклад. Обчислити псевдообернену для матриці

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

Розв'язання. Застосуємо формулу (6):

$$A = H^T H = \begin{pmatrix} 5 & -3 & -1 \\ -3 & 3 & 3 \\ -1 & 3 & 5 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \mathcal{X}(t) = \begin{vmatrix} t-5 & 3 & 1 \\ 3 & t-3 & -3 \\ 1 & -3 & t-5 \end{vmatrix} = t(t-4)(t-9) = -36t(t\varphi(t)-1),$$

де $\varphi(t) = -\dfrac{1}{36}(t-13)$. Тоді

$$\varphi(A) = -\frac{1}{36}(A-13E) = \frac{1}{36}\begin{pmatrix} 8 & 3 & 1 \\ 3 & 10 & -3 \\ 1 & -1 & 8 \end{pmatrix},$$

$$H^+ = \varphi(A)H^T = \frac{1}{36}\begin{pmatrix} 13 & -4 & 5 \\ -4 & 4 & 4 \\ 5 & 4 & 13 \end{pmatrix}.$$

Розділ 9. ОПУКЛІ МНОЖИНИ

§ 9.1. Загальні властивості та способи побудови опуклих множин

Нехай x_0, x_1 – дві різні точки евклідового простору R^n. Множина всіх точок $x = (1-\lambda)x_0 + \lambda x_1$, де параметр λ набуває всіх значень із R, називається *прямою лінією*, що проходить через x_0 і x_1.

Нехай x_0, x_1 – будь-які дві точки евклідового простору R^n. *Відрізком* $[x_0, x_1]$ з кінцями x_0 і x_1 називається множина всіх точок $x = (1-\lambda)x_0 + \lambda x_1$, де параметр λ пробігає всі дійсні числа відрізка $[0, 1]$.

Означення. *Непорожня підмножина V евклідового простору R^n називається опуклою, якщо з $x_0, x_1 \in V$ випливає, що $[x_0, x_1] \subset V$.*

Приклад 1. Нехай L – лінійний багатовид простору R^n, тобто множина всіх розв'язків сумісної системи лінійних рівнянь

$$\begin{cases} (a_1, x) = \beta_1, \\ (a_2, x) = \beta_2, \\ \dots\dots\dots\dots \\ (a_m, x) = \beta_m. \end{cases}$$

Якщо x_0, x_1 – два розв'язки цієї системи, то $x = (1-\lambda)x_0 + \lambda x_1$ при будь-яких значеннях параметра λ також буде розв'язком цієї системи. Якщо x_0 і x_1 відмінні один від одного, то пряма лінія, що з'єднує ці точки, належить багатовиду L. Звичайно, багатовид L – опукла множина.

Приклад 2. Нехай $\Pi = \{x \in R^n \colon (a, x) = \beta\}$ – гіперплощина ($a \neq \theta$). Півпростори

$$P=\{x\in R^n:(a,x)\le\beta\},\ Q=\{x\in R^n:(a,x)\ge\beta\},$$
$$P^0=\{x\in R^n:(a,x)<\beta\},\ Q^0=\{x\in R^n:(a,x)>\beta\},$$

які визначаються гіперплощиною, – опуклі множини.

Перевіримо твердження для півпростору P. Якщо $x_0,x_1\in P$, то $(a,x_0)\le\beta$, $(a,x_1)\le\beta$. Нехай $x=(1-\lambda)x_0+\lambda x_1$, $0\le\lambda\le 1$. Тоді

$$(a,x)=(a,(1-\lambda)x_0+\lambda x_1)=$$
$$=(1-\lambda)(a,x_0)+\lambda(a,x_1)\le(1-\lambda)\beta+\lambda\beta=\beta,$$

що й означає $x\in P$. Відрізок $[x_0,x_1]\subset P$.

Приклад 3. Відкрита куля $S=S(a,\varepsilon)=\{x\in R^n:\|x-a\|<\varepsilon\}$ з центром a і радіусом $\varepsilon>0$ – опукла множина.

Нехай $x_0,x_1\in S$, тобто $\|x_0-a\|<\varepsilon$ і $\|x_1-a\|<\varepsilon$. Якщо $x=(1-\lambda)x_0+\lambda x_1$, $0\le\lambda\le 1$, то $\|x-a\|=\|(1-\lambda)x_0+\lambda x_1-a\|=$

$$=\|(1-\lambda)(x_0-a)+\lambda(x_1-a)\|\le\|(1-\lambda)(x_0-a)\|+\|\lambda(x_1-a)\|=$$
$$=(1-\lambda)\cdot\|x_0-a\|+\lambda\|x_1-a\|<(1-\lambda)\varepsilon+\lambda\varepsilon=\varepsilon,\ \|x-a\|<\varepsilon.$$

Отже, $x\in S$ і $[x_0,x_1]\subset S$.

Вправи

1. Нехай K_n – множина всіх точок $x=(\xi_1,\xi_2,\ldots,\xi_n)$ простору R^n таких, що $|\xi_k|\le 1, k=\overline{1,n}$. Ця множина називається n-вимірним кубом. Показати, що K_n – опукла множина.

2. Показати, що множина всіх точок (ξ,η) евклідової площини R^2, обмежена еліпсом $\frac{\xi^2}{\alpha^2}+\frac{\eta^2}{\beta^2}=1$ ($\alpha>0$, $\beta>0$), тобто множина всіх точок (ξ,η), координати яких задовольняють нерівність $\frac{\xi^2}{\alpha^2}+\frac{\eta^2}{\beta^2}\le 1$, є опуклою.

3. Показати, що множина всіх точок (ξ,η) площини R^2, що задовольняють нерівність $\alpha\xi^2+\beta\xi+\gamma\le\eta$ ($\alpha>0$), тобто множина всіх точок, обмежених параболою з рівнянням $\eta = \alpha\xi^2 + \beta\xi + \gamma$, є опуклою.

4. Показати, що множина всіх точок $x = (\xi, \eta, \zeta)$ тривимірного евклідового простору R^3, які задовольняють нерівність $\alpha\xi^2 + \beta\eta^2 \le \zeta$ $(\alpha > 0, \beta > 0)$, є опуклою.

5. Нехай V– множина всіх точок $x = (\xi_1, \xi_2, ..., \xi_n)$ простору R^n, що задовольняють умову $\frac{\xi_1^2}{\alpha_1^2} + \frac{\xi_2^2}{\alpha_2^2} + ... + \frac{\xi_n^2}{\alpha_n^2} \le 1$. Довести опуклість цієї множини.

6. Показати, що коли $\mathcal{A}$ – лінійний оператор простору R^n, то образ $\mathcal{A}(V)$ опуклої множини V – опукла множина.

7. Нехай V – підмножина евклідового простору R^n, яка має такі властивості: якщо $x_0, x_1 \in V$, то $x = \frac{1}{2}x_0 + \frac{1}{2}x_1 \in V$. Довести, що коли V– замкнена множина, то вона – опукла.

Навести приклад, який свідчив би, що такий висновок, взагалі кажучи, невірний.

Операції над опуклими множинами

I. Сума опуклих множин. Нехай V_1 і V_2 – опуклі множини. Їх сума $V = V_1 + V_2$ – опукла множина.

Дійсно, оберемо $v, w \in V_1 + V_2$. Тоді $v = x_1 + x_2$, де $x_1 \in V_1$, $x_2 \in V_2$ і $w = y_1 + y_2$, де $y_1 \in V_1, y_2 \in V_2$. Утворимо лінійну комбінацію $u = (1-\lambda)v + \lambda w$, $0 \le \lambda \le 1$. Тоді

$$u = (1-\lambda)(x_1 + x_2) + \lambda(y_1 + y_2) = \big((1-\lambda)x_1 + \lambda y_1\big) + \\ + \big((1-\lambda)x_2 + \lambda y_2\big).$$

За умовою $(1-\lambda)x_1 + \lambda y_1 \in V_1$ і $(1-\lambda)x_2 + \lambda y_2 \in V_2$. Тому $u \in V_1 + V_2$ і, отже, $V_1 + V_2$ – множина опукла.

II. Пряма сума опуклих множин. Сума $V = V_1 + V_2$ опуклих множин V_1 і V_2 називається *прямою*, якщо із подання точки $v \in V_1 + V_2$ у вигляді $v = x_1 + x_2 = y_1 + y_2$, де $x_1, y_1 \in V_1$, $x_2, y_2 \in V_2$ випливає $x_1 = y_1, x_2 = y_2$.

III. Множення на скаляр. Якщо V – опукла множина і λ – дійсне число, то за означенням $\lambda V = \{y \in R^n : y = \lambda x, x \in V\}$ – опукла множина.

IV. Перетин сім'ї опуклих множин. Нехай дано сім'ю $V_\omega, \omega \in \Omega$ опуклих множин. Їх перетин $V = \bigcap_{\omega \in \Omega} V_\omega$ – або порожня, або опукла множина.

Припустимо, що $V \neq \varnothing$ і $x_0, x_1 \in V$. Тоді $x_0, x_1 \in V_\omega$ для будь-якого $\omega \in \Omega$. На підставі опуклості V_ω відрізок $[x_0, x_1] \subset V_\omega$. А тому $[x_0, x_1] \subset V$. Отже, множина V – опукла.

Зокрема, переріз скінченної сім'ї півпросторів $P_k = \{x \in R^n: (a_k, x) \le \beta_k\}$, $k = \overline{1,m}$, тобто множина всіх розв'язків системи лінійних нерівностей

$$(a_k, x) \le \beta_k, \ k = \overline{1,m}$$

– множина опукла (якщо система сумісна).

Така множина $P = P_1 \cap P_2 \cap \ldots \cap P_m$ називається *багатогранною* множиною.

Багатогранні множини можуть бути обмеженими і необмеженими. Наприклад, у просторі R^3 множина розв'язків системи нерівностей

$$\begin{cases} x + y + z \le 1, \\ -x \le 0, \\ -y \le 0, \\ -z \le 0 \end{cases}$$

обмежена багатогранна множина. Це – тетраедр $OABC$ (рис. 6), обмежений площинами: 1) $x + y + z = 1$, 2) $x = 0$, 3) $y = 0$, 4) $z = 0$, де A (1, 0, 0), B (0, 1, 0), C (0, 0, 1). Обмежену багатогранну множину називатимемо *опуклим багатогранником.*

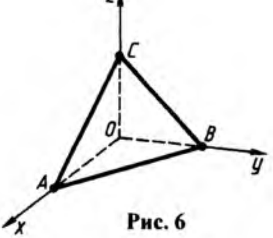


Рис. 6

V. Теоретико-множинне замикання. Замикання $\overline{V}$ опуклої множини V – множина опукла.

Замикання $\bar{V}$ одержується з множини V приєднанням усіх її граничних точок. Якщо $x_0, x_1 \in \bar{V}$ і $x_0, x_1 \notin V$ (або хоч би одна з точок x_0, x_1 не належить V), то існують нескінченні послідовності точок $\{a_m\}$ і $\{b_m\}$ із V, що збігаються відповідно до x_0 і x_1:

$$a_m \to x_0 \text{ при } m \to \infty,$$
$$b_m \to x_1 \text{ при } m \to \infty.$$

Якщо $x = (1-\lambda)x_0 + \lambda x_1$, $0 \le \lambda \le 1$, тоді розглянемо послідовність точок $c_m = (1-\lambda)a_m + \lambda b_m \in V$, $c_m \to x = (1-\lambda)x_0 + \lambda x_1$ при $m \to \infty$. Точка x – гранична для послідовності $\{c_m\}$ і, отже, $x \in \bar{V}$. Множина $\bar{V}$ – опукла.

VI. Якщо V – опукла множина і її внутрішня область V° непорожня, то V° – опукла множина.

Якщо $x_0, x_1 \in V^\circ$, тоді існують відкриті кулі

$$S_1 = S(x_0, \varepsilon) = \{x \in R^n : \|x - x_0\| < \varepsilon\},$$

$$S_2 = S(x_1, \varepsilon) = \{x \in R^n : \|x - x_1\| < \varepsilon\}$$

достатньо малого радіуса ε такі, що $S_1 \subset V$, $S_2 \subset V$. Нехай $x = (1-\lambda)x_0 + \lambda x_1$, $0 \le \lambda \le 1$ і $S = S(x, \varepsilon)$ – ε-окіл точки x. Оберемо довільну точку $w \in S$. Для неї $\|w - x\| < \varepsilon$.

Точка $w_0 = x_0 + (w - x)$ належить S_1, оскільки $\|w_0 - x_0\| = \|w - x\| < \varepsilon$; точка $w_1 = x_1 + (w - x)$ належить S_2, оскільки $\|w_1 - x_1\| = \|w - x\| < \varepsilon$. При цьому $(1-\lambda)w_0 + \lambda w_1 = (1-\lambda)x_0 + \lambda x_1 + (w - x) = x + (w - x) = w$.

На підставі опуклості множини V маємо включення $w \in V$.

Отже, довільна точка із $w \in S$ належить V, тобто $S \subset V$.

Таким чином, $x \in V^\circ \Rightarrow [x_0, x_1] \subset V^\circ$. V° – опукла множина.

§ 9.2. Конус

Означення. *Непорожня підмножина K евклідового простору називається* к о н у с о м, *якщо для неї виконуються умови*:

1. $x, y \in K \Rightarrow x + y \in K$.
2. $x \in K, \lambda \geq 0 \Rightarrow \lambda x \in K$.

Ці умови еквівалентні одній умові

3. $x, y \in K \quad \lambda \geq 0, \mu \geq 0 \Rightarrow \lambda x + \mu y \in K$.

Звідси випливає, що конус K – опукла множина. Її називаємо *опуклим конусом*.

Якщо $x_1, x_2, ..., x_m$ належать конусу K і $\lambda_1 \geq 0, \lambda_2 \geq 0, ..., \lambda_m \geq 0$, то $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m \in K$: будь-яка *невід'ємна лінійна комбінація* векторів із цього конуса належить цьому ж конусу. Конус містить нульовий вектор θ.

Приклади опуклих конусів

1. Множина $R_+^n = \{x = (\xi_1, \xi_2, ..., \xi_n) \in R^n : \xi_1 \geq 0, \ \xi_2 \geq 0, ..., \xi_n \geq 0\}$ називається *невід'ємним ортантом* простору R^n. Це – опуклий конус.

2. Будь-який лінійний підпростір простору R^n є конус.

3. Нехай $\Pi = \{x : (a, x) = 0\}$ – деяка гіперплощина, що проходить через нульову точку θ. Півпростори $P = \{x : (a, x) \leq 0\}$, $Q = \{x : (a, x) \geq 0\}$, визначені цією гіперплощиною, є опуклі конуси.

4. Розглянемо систему лінійних однорідних нерівностей

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + ... + \alpha_{1n}\xi_n \leq 0, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + ... + \alpha_{2n}\xi_n \leq 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + ... + \alpha_{mn}\xi_n \leq 0, \end{cases}$$

або у векторному запису $(a_1, x) \leq 0$, $(a_2, x) \leq 0$, ..., $(a_m, x) \leq 0$, де $a_k = (\alpha_{k1}, \alpha_{k2}, ..., \alpha_{kn})$, $x = (\xi_1, \xi_2, ..., \xi_n)$, $k = \overline{1, m}$.

Множина K всіх її розв'язків збігається з перетином

$$P_1 \cap P_2 \cap ... \cap P_m$$

півпросторів

$$P_k = \{x : (a_k, x) \le 0\},\ k = \overline{1,m}.$$

Ця множина є опуклий конус.

У просторі R^n виділимо скінченну систему векторів $a_1, a_2, ..., a_m$ і через $L^+(a_1, a_2, ..., a_m)$ позначимо множину всіх тих векторів, які можна подати у вигляді невід'ємних лінійних комбінацій цих векторів:

$$L^+(a_1, a_2, ..., a_m) =$$

$$= \{x = \eta_1 a_1 + \eta_2 a_2 + ... + \eta_m a_m : \eta_1 \ge 0, \eta_2 \ge 0, ..., \eta_m \ge 0\}.$$

Якщо через $A = (a_1 | a_2 | ... | a_m)$ позначити матрицю, стовпці якої складені з координат заданих векторів a_k, а $y = (\eta_1, \eta_2, ..., \eta_m)^T \ge \theta$ – невід'ємний вектор-стовпець розмірності m ($\eta_1 \ge 0, \eta_2 \ge 0, ..., \eta_m \ge 0$), то цю множину можна записати так:

$$L^+(a_1, a_2, ..., a_m) = \{x \in R^m : x = Ay, y \ge 0\}.$$

Покажемо, що ця множина є конус. Нехай $x, x_1 \in L^+(a_1, ..., a_m)$ і $\lambda, \mu \ge 0$. Тоді $x = Ay$ і $x_1 = Ay_1$, де $y = (\eta_1, \eta_2, ..., \eta_m)^T \ge \theta$, $y_1 = (\gamma_1, \gamma_2, ..., \gamma_m)^T \ge \theta$ – деякі невід'ємні вектор-стовпці. Невід'ємна лінійна комбінація $w = \lambda x + \mu x_1$ дорівнює $A(\lambda y + \mu y_1)$, де $\lambda y + \mu y_1$ також невід'ємний вектор-стовпець. Тому $w = \lambda x + \mu x_1$ належить множині $L^+(a_1, a_2, ..., a_m)$.

Конус $L^+(a_1, a_2, ..., a_m)$ називається *багатогранним конусом, породженим векторами* $a_1, a_2, ..., a_m$.

Зауважимо, що будь-який лінійний підпростір M простору R^n – багатогранний конус. Подамо цей підпростір у вигляді лінійної оболонки деякого набору векторів: $M = L(a_1, a_2, ..., a_m)$. Тоді

$$M = L^+(a_1, a_2, ..., a_m, -a_1, -a_2, ..., -a_m).$$

Нехай S – довільна підмножина із простору R^n. Позначимо через $L^+(S)$ множину всіх векторів $x \in R^n$, які можна подати у вигляді невід'ємних лінійних комбінацій різних скінченних підмножин із S:

$$L^+(S) = \bigcup_{\langle a_1, a_2, ..., a_m \rangle \subset S} L^+(a_1, a_2, ..., a_m).$$

Ця множина – опуклий конус.

Означення. *Конус K називається загостреним, якщо він не містить лінійних підпросторів, відмінних від нульового підпростору* $\{\theta\}$.

Загострений конус K характеризується властивістю: якщо $a, -a \in K$, тоді $a = \theta$.

Прикладом загостреного конуса у просторі R^n є невід'ємний ортант R^n_+.

Теорема 1. *Будь-який конус K* простору R^n розкладається у пряму суму

$$K = M \oplus V, \tag{1}$$

де M – лінійний підпростір (лінійна частина конуса); V – загострений конус (загострена частина).

Підпростір M визначається конусом K однозначно. Якщо K – багатогранний конус, тоді його загострена частина V – також багатогранний конус.

Доведення. Позначимо через M лінійний підпростір, який міститься у конусі K і має найбільшу розмірність. Якщо M' – також лінійний підпростір, що міститься в K, то $M + M'$ – підпростір, що міститься в K і $M + M' \supseteq M$. Але розмірність M – найбільша. Тому $M + M' = M$ і, отже, $M' \subseteq M$.

Таким чином, M – об'єднання всіх лінійних підпросторів, що містяться в конусі K. Воно визначається однозначно.

Якщо $M^\perp$ – ортогональне доповнення M, то $R^n = M \oplus M^\perp$. Нехай $V = K \cap M^\perp$ (замість $M^\perp$ можна взяти будь-який інший лінійний підпростір L, для якого мала місце рівність $R^n = M \oplus L$).

Множина V – перетин двох конусів і, отже, є конус. Він загострений, оскільки не містить нетривіальних лінійних підпросторів. Сума $M+V$ – пряма і $K \supseteq M \oplus V$.

Якщо $w \in K$, то $w = x + y$, де $x \in M$, $y \in M^{\perp}$. Але $y = w + (-1)x \in K$ і тому $y \in M^{\perp} \cap K = V$. Звідси дістаємо $K \subseteq M \oplus V$.

Отже, $K = M \oplus V$.

Припустимо тепер, що K – багатогранний конус: $K = L^{+}(a_1, a_2, ..., a_m)$. Тоді $a_k = x_k + y_k, k = \overline{1,m}$, де $x_k \in M$, $y_k \in V$. Якщо $v \in V$, тоді

$$v = \sum_{k=1}^{m} \lambda_k a_k = \sum_{k=1}^{m} \lambda_k (x_k + y_k) = \sum_{k=1}^{m} \lambda_k x_k + \sum_{k=1}^{m} \lambda_k y_k ,$$

де всі $\lambda_k \geq 0$. При цьому $x = \sum_{k=1}^{m} \lambda_k x_k \in M$, $y = \sum_{k=1}^{m} \lambda_k y_k \in V$. Маємо рівності $v = \theta + v = x + y$. Але сума $M \oplus V$ – пряма. Тому $\theta = x$, $v = y = \sum_{k=1}^{m} \lambda_k y_k$. Це означає, що $V = L^{+}(y_1, y_2, ..., y_m)$.

Лема. *Якщо $x_1, x_2, ..., x_m$ – ненульові вектори загостреного конуса K, то з рівності $\sum_{k=1}^{m} \lambda_k x_k = \theta$, де $\lambda_1 \geq 0, \lambda_2 \geq 0, ..., \lambda_m \geq 0$, випливає $\lambda_1 = \lambda_2 = ... = \lambda_m = 0$.*

Доведення. Якщо $m = 1$, то лема очевидна. Нехай $m > 1$. Якщо $\lambda_1 > 0$, то $x_1 = -\left(\frac{\lambda_2}{\lambda_1} x_2 + ... + \frac{\lambda_m}{\lambda_1} x_m \right) = -w$, $w \neq \theta$, $w \in K$ і $-w \in K$. Це неможливо для загостреного конуса. Отже, $\lambda_1 = 0$. Таким же чином $\lambda_2 = ... = \lambda_m = 0$.

Теорема 2. *Нехай K – багатогранний загострений конус, породжений ненульовими векторами $x_1, x_2, ..., x_m$, тобто $K = L^{+}(x_1, x_2, ..., x_m)$. Тоді у ньому знайдуться вектори $y_1, y_2, ..., y_p$ з такими властивостями:*

а) $K = L^{+}(y_1, y_2, ..., y_p)$;

б) $y_k \notin L^{+}(y_1, ..., y_{k-1}, y_{k+1}, ..., y_p)$ *для всіх* $k = \overline{1,p}$.

Підмножина векторів з властивостями а) *і* б) *єдина з точністю до порядку розміщення цих векторів і додатних скалярних множників.*

Набір векторів із конуса K з властивостями а) *і* б) *називається кістяком конуса. Півпрямі (промені), що виходять із нульової точки* θ *і які містять вектори кістяка, називаються його ребрами.*

Доведення. Якщо жоден з векторів $x_1, x_2, ..., x_m$ не може бути поданим у вигляді невід'ємної лінійної комбінації решти, тоді векторами y_k з умови теореми будуть вектори $x_1, x_2, ..., x_m$, причому $p = m$. Припустимо, що який-небудь з них, наприклад x_1, – невід'ємна лінійна комбінація решти, тобто $x_1 \in L^{+}(x_2, ..., x_m)$. Тоді $L^{+}(x_2, ..., x_m) = K$. Якщо жоден з векторів $x_2, ..., x_m$ не подається у вигляді невід'ємної лінійної комбінації, то векторами y_k обираємо $x_2, ..., x_m$ і тоді $p = m - 1$. Продовжуючи цю процедуру, дістанемо потрібну нам систему векторів $y_1, y_2, ..., y_p$, яка належатиме заданій системі $x_1, x_2, ..., x_m$ векторів.

А тепер доведемо єдиність. Нехай інший набір векторів $w_1, w_2, ..., w_q$ із $L^{+}(x_1, ..., x_m)$ має властивості а) і б). Тоді

$$y_1 = \sum_{j=1}^{q} \alpha_j w_j,\ \alpha_j \ge 0\ (j = \overline{1,q});\ w_j = \sum_{t=1}^{p} \beta_{jt} y_t,\ \beta_{jt} \ge 0\ (j = \overline{1,q}).$$

Тому

$$y_1 = \sum_{t=1}^{p} \left(\sum_{j=1}^{q} \alpha_j \beta_{jt} \right) y_t = \sum_{t=1}^{p} \gamma_t y_t\ ,\ \gamma_t = \sum_{j=1}^{q} \alpha_j \beta_{jt} \ge 0\,.$$

Тут $\gamma_1 > 0$, оскільки $y_1 \notin L^{+}(y_2, ..., y_p)$.

Якщо $\gamma_1 < 1$, то $y_1 = \frac{\gamma_2}{1-\gamma_1} y_2 + ... + \frac{\gamma_p}{1-\gamma_1} y_p$, що неможливо з того ж приводу. Якщо $\gamma_1 > 1$, то дістанемо рівність

$(\gamma_1 - 1)y_1 + \gamma_2 y_2 + ... + \gamma_p y_p = \theta$, що неможливо на підставі леми.

Залишається можливість $\gamma_1 = 1$. Тоді $\gamma_2 y_2 + ... + \gamma_p y_p = \theta$. Із леми випливає $\gamma_2 = \gamma_3 = ... = \gamma_p = 0$.

Отже, одержано рівності

$$\gamma_1 = \sum_{j=1}^{q} \alpha_j \beta_{j1}, \ \gamma_t = \sum_{j=1}^{q} \alpha_j \beta_{jt} = 0 \text{ при } t > 1.$$

Із останніх рівностей випливає, що

$$\alpha_j \beta_{j2} = 0, \alpha_j \beta_{j3} = 0, ..., \alpha_j \beta_{jp} = 0,$$

причому серед α_j є відмінні від нуля. Нехай $\alpha_k \neq 0$. Тоді $\beta_{k2} = ... = \beta_{kp} = 0$. Це означає, що $w_k = \beta_{k1} y_1$, $\beta_{k1} > 0$. Звідси $y_1 = \delta_1 w_k$, $\delta_1 > 0$. Таким же чином $y_2 = \delta_2 w_l$ і т. д.

Врешті-решт дістаємо нерівність $p \leq q$. З тих самих міркувань встановлюємо обернену нерівність $q \leq p$.

Теорему доведено.

Вправа

Нехай K – опуклий конус. Довести:

а) множина $K - K = \{w = x - y, x \in K, y \in K\}$ – мінімальний лінійний підпростір, що містить конус K;

б) множина $K \cap (-K)$ – максимальний лінійний підпростір, що міститься в конусі K.

§ 9.3. Опукла оболонка

Означення. *Точка x евклідового простору називається опукло залежною від точок $x_1, x_2, ..., x_m$, якщо вона може бути подана у вигляді суми*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_m x_m,$$

де $\lambda_1 \geq 0,\ \lambda_2 \geq 0, ...,\ \lambda_m \geq 0,\ \lambda_1 + \lambda_2 + ... + \lambda_m = 1$.

Мова йтиме про те, що x – *опукла лінійна комбінація* точок $x_1, x_2, ..., x_m$.

Точка x називається *опукло залежною від множини* S точок простору, якщо існує такий набір точок $x_1, x_2, ..., x_m \in S$, що x опукло залежить від $x_1, x_2, ..., x_m$.

Лема. *Якщо точка x опукло залежить від множини S, а кожна точка із S опукло залежить від множини T, то x опукло залежить від T.*

Доведення. Нехай $x = \sum_{k=1}^{m} \lambda_k x_k$, $\lambda_k \geq 0$, $\sum_{k=1}^{m} \lambda_k = 1$ і $x_1, x_2, ..., x_m \in S$. За умовою $x_k = \sum_{j=1}^{p} \mu_{kj} y_j$, де $\mu_{kj} \geq 0$, $\sum_{j=1}^{p} \mu_{kj} = 1$ і $y_1, y_2, ..., y_p \in T$. Тоді $x = \sum_{k=1}^{m} \lambda_k \left(\sum_{j=1}^{p} \mu_{kj} y_j \right) = \sum_{j=1}^{p} \gamma_j y_j$, де $\gamma_j = \sum_{k=1}^{m} \lambda_k \mu_{kj}$. При цьому $\gamma_j \geq 0$ і $\sum_{j=1}^{p} \gamma_j = \sum_{j=1}^{p} \left(\sum_{k=1}^{m} \lambda_k \mu_{kj} \right) =$
$$= \sum_{k=1}^{m} \lambda_k \left(\sum_{j=1}^{p} \mu_{kj} \right) = \sum_{k=1}^{m} \lambda_k = 1.$$

Теорема 1. *Якщо V – опукла множина і точка $x \in R^n$ опукло залежить від V, тоді x належить цій множині V.*

Доведення. Нехай $x = \sum_{k=1}^{m} \lambda_k x_k$, $\lambda_k \geq 0$, $\sum_{k=1}^{m} \lambda_k = 1$ і $x_1, x_2, ..., x_m \in V$.

Якщо $m = 1$ і $m = 2$, тоді твердження теореми вірне. Припустимо, що має місце теорема, коли точка x опукло залежить від $m-1$ точок із V. Доведемо теорему для точки x, яка опукло залежить від точок $x_1, x_2, ..., x_m$ із V.

Якщо $\lambda_m = 0$, то $x = \lambda_1 x_1 + ... + \lambda_{m-1} x_{m-1}$ і за припущенням індукції $x \in V$. Якщо $\lambda_m = 1$, то $x = x_m \in V$.

Залишається розглянути випадок, коли $0 < \lambda_m < 1$. Позначимо $\lambda = \lambda_1 + ... + \lambda_{m-1}$. Тоді $\lambda + \lambda_m = 1$, $0 < \lambda < 1$. Матимемо рівність

$$x = \lambda_1 x_1 + ... + \lambda_{m-1} x_{m-1} + \lambda_m x_m =$$

$$= \lambda \left(\frac{\lambda_1}{\lambda} x_1 + ... + \frac{\lambda_{m-1}}{\lambda} x_{m-1} \right) + \lambda_m x_m.$$

Точка $y = \frac{\lambda_1}{\lambda} x_1 + ... + \frac{\lambda_{m-1}}{\lambda} x_{m-1}$ опукло залежить від $x_1, ..., x_{m-1}$ і за припущенням індукції належить V. Тоді $x = \lambda y + \lambda_m x_m \in V$ на підставі опуклості V.

Означення. *Нехай S – довільна множина точок простору R^n. Перетин класу Ω всіх опуклих множин, що містять S, називається його опуклою оболонкою.*

Позначення опуклої оболонки – $\operatorname{conv}(S) = \bigcap_{V \in \Omega} V$. Інакше кажучи, $\operatorname{conv}(S)$ – найменша опукла множина точок простору, що містить множину S.

Теорема 2. *Нехай S – множина точок, що належать простору R^n. Його опукла оболонка збігається з множиною всіх тих точок простору, які опукло залежать від S.*

Доведення. Позначимо через V множину всіх тих точок $x \in R^n$, які опукло залежать від S. Оскільки $\operatorname{conv}(S)$ – множина опукла, то за теоремою 1 маємо включення $V \subseteq \operatorname{conv}(S)$.

Тепер доведемо обернене включення $\operatorname{conv}(S) \subseteq V$. Для цього переконаємось, що V – опукла множина.

Нехай $x, y \in V$. Це означає, що існує такий набір точок $x_1, x_2, ..., x_m \in S$, для яких

$$x = \sum_{k=1}^{m} \lambda_k x_k,\ \lambda_k \geq 0,\ \sum_{k=1}^{m} \lambda_k = 1;$$

$$y = \sum_{k=1}^{m} \mu_k x_k,\ \mu_k \geq 0,\ \sum_{k=1}^{m} \mu_k = 1.$$

Якщо $w=(1-\lambda)x+\lambda y$, $0\le\lambda\le 1$, то $w=(1-\lambda)\left(\sum_{k=1}^{m}\lambda_k x_k\right)+$ $+\lambda\left(\sum_{k=1}^{m}\mu_k x_k\right)=\sum_{k=1}^{m}[(1-\lambda)\lambda_k+\lambda\mu_k]x_k$. При цьому $(1-\lambda)\lambda_k+$ $+\lambda\mu_k\ge 0$ і $\sum_{k=1}^{m}[(1-\lambda)\lambda_k+\lambda\mu_k]=(1-\lambda)\left(\sum_{k=1}^{m}\lambda_k\right)+\lambda\left(\sum_{k=1}^{m}\mu_k\right)=$ $=(1-\lambda)+\lambda=1$. Це означає, що $w\in V$, а тому $[x,y]\subset V$. Множина V опукла і $V=\mathrm{conv}(S)$.

Наведемо деякі властивості опуклих оболонок.

I. $\mathrm{conv}(\mathrm{conv}(S))=\mathrm{conv}(S)$.

II. Якщо $x\in\mathrm{conv}(S)$, тоді $\mathrm{conv}(x,S)=\mathrm{conv}(S)$.

Ці властивості випливають із леми.

III. Якщо S – відкрита множина, тоді $\mathrm{conv}(S)$ також відкрита множина.

Доведення. Оскільки $\mathrm{conv}(S)\supset S$, то $\mathrm{conv}(S)\supseteq$ $\supseteq(\mathrm{conv}(S))^\circ\supseteq S$ на підставі того, що S – відкрита множина. Але $(\mathrm{conv}(S))^\circ$ – опукла множина. Тому з означення conv матимемо включення $(\mathrm{conv}(S))^\circ\supseteq\mathrm{conv}(S)$. Отже, $\mathrm{conv}(S)=$ $=(\mathrm{conv}(S))^\circ$. Останнє й доводить, що $\mathrm{conv}(S)$ – відкрита множина.

Подібне висловлювання для замкнених множин не вірне: існують замкнені множини, опуклі оболонки яких незамкнені.

Приклад 1. Нехай L – пряма лінія на площині R^2 і p – точка на тій же площині, яка не належить прямій L (рис. 7). Тоді опукла оболонка замкненої множини $L\cup\{p\}$ складається з усіх точок, що належать усім відрізкам $[p, q]$, де точка q пробігає пряму L. Ця множина $V=\bigcup_{q\in L}[p,q]$ не замкнена. Замикання $\overline{V}$ одержується приєднанням точок прямої L', паралельної до L і яка проходить через точку p.

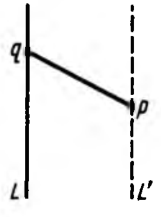


Рис. 7

Зауважимо, що побудова опуклої оболонки даної множини S – задача досить складна. Найчастіше вирішують її в два етапи.

1. Обирають опуклу множину V, що містить S, і яка може бути допустимою у ролі $\operatorname{conv}(S)$.

2. Перевіряють, чи буде довільна точка $x \in V$ опукло залежною від S.

Приклад 2. Нехай S_n – множина всіх точок виду $a = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$, $\varepsilon_k = \pm 1$ простору R^n. Покажемо, що $\operatorname{conv}(S_n) = K_n$ – n-вимірний куб (множина всіх точок $x = (\xi_1, \xi_2, ..., \xi_n)$ таких, що $|\xi_k| \le 1$, $k = \overline{1, n}$).

Розв'язання. Почнемо з випадку $n = 2$. Тоді S_2 складається з 4 точок:

$$a_1 = (1, 1),\ a_2 = (1, -1),\ a_3 = (-1, 1),\ a_4 = (-1, -1).$$

K_2 – квадрат з вершинами a_1, a_2, a_3, a_4 – множина опукла і містить ці точки. Для встановлення рівності $K_2 = \operatorname{conv}(a_1, a_2, a_3, a_4)$ потрібно показати, що будь-яка точка $x^0 = (\xi_1^0, \xi_2^0) \in K_2$ – опукла лінійна комбінація точок a_1, a_2, a_3, a_4.

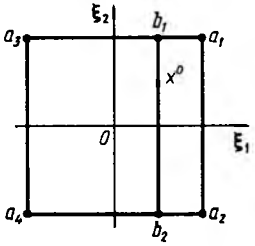


Рис. 8

У прямокутній системі координат $O\xi_1\xi_2$ (рис. 8) перетнемо квадрат K_2 прямою, що проходить через точку x^0 паралельно осі $O\xi_2$. Матимемо $b_1 = (\xi_1^0, 1)$, $b_2 = (\xi_1^0, -1)$. Для деякого $\lambda \in [0, 1]$ маємо $b_1 = (1-\lambda)a_1 + \lambda a_3$ і $b_2 = (1-\lambda)a_2 + \lambda a_4$. Крім того, подамо x^0 у вигляді $x^0 = (1-\mu)b_1 + \mu b_2$. При цьому $\mu \in [0, 1]$, $\lambda = \dfrac{1-\xi_1^0}{2}$, $\mu = \dfrac{1-\xi_2^0}{2}$. Тоді $x^0 = (1-\mu)[(1-\lambda)a_1 + \lambda a_3] + \mu[(1-\lambda)a_2 + \lambda a_4] = (1-\mu)(1-\lambda)a_1 + (1-\mu)\lambda a_3 + \mu(1-\lambda)a_2 + \mu\lambda a_4$ і $(1-\mu)(1-\lambda) + (1-\mu)\lambda + \mu(1-\lambda) + \mu\lambda = 1$.

Далі використаємо індукцію за розмірністю простору. Припустимо, що має місце рівність $\operatorname{conv}(S_{n-1}) = K_{n-1}$.

Куб K_n – опукла множина, що містить усі точки виду

$$a = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n),\ \varepsilon_k = \pm 1,\ k = 1, 2, ..., 2^n.$$

Потрібно показати, що будь-яка точка $x = (\xi_1^0, \xi_2^0, ..., \xi_n^0) \in K_n$ опукло залежить від $a_1, a_2, ..., a_{2^n}$.

Запишемо будь-яку точку $x = (\xi_1, \xi_2, ..., \xi_n)$ у вигляді $x = (\underline{x}, \xi_n)$, де $\underline{x} = (\xi_1, \xi_2, ..., \xi_{n-1})$. Тоді $x^0 = (\underline{x}^0, \xi_n^0)$, де $\underline{x}^0 \in K_{n-1}$, $a_k = (\underline{a}_k, \varepsilon_n)$, де $\underline{a}_k \in S_{n-1}$. Перетнемо K_n гіперплощиною $\xi_n = \xi_n^0$. Перерізом буде множина всіх точок виду $x = (\underline{x}, \xi_n^0)$, де $\underline{x} \in K_{n-1}$.

За припущенням індукції $\underline{x}^0 = (\xi_1^0, \xi_2^0, ..., \xi_{n-1}^0)$ опукло залежить від S_{n-1}:

$$\underline{x}^0 = \lambda_1 \underline{a}_1 + \lambda_2 \underline{a}_2 + ...,\ \lambda_k \geq 0,\ \lambda_1 + \lambda_2 + ... = 1.$$

Тому $(\underline{x}^0, 1) = \lambda_1(\underline{a}_1, 1) + \lambda_2(\underline{a}_2, 1) + ...$, $(\underline{x}^0, -1) = \lambda_1(\underline{a}_1, -1) + \lambda_2(\underline{a}_2, -1) + ...$; $(\underline{x}^0, 1)$ і $(\underline{x}^0, -1)$ опукло залежать від S_n. Але $x^0 = (1-\lambda)(\underline{x}^0, 1) + \lambda(\underline{x}^0, -1)$, де $\lambda = \dfrac{1-\xi_n^0}{2} \in [0, 1]$. На підставі леми x^0 опукло залежить від S_n.

Вправи

Знайти $\operatorname{conv}(S)$ множини S:

1. $S = \{x_0, x_1, x_2, ..., x_n, ...\} \subset R^2$, де $x_0 = (1, 0)$, $x_n = \left(0, \dfrac{1}{n}\right)$, $n = 1, 2, ...$

2. S – множина всіх точок кола $\xi^2 + \eta^2 = \gamma$ ($\gamma > 0$) площини R^2.

3. S – множина всіх точок параболи з рівнянням $\eta^2 = 2p\xi$ ($p > 0$) площини R^2.

4. S – множина всіх точок параболоїда $\zeta = \alpha\xi^2 + \beta\eta^2$ $(\alpha > 0, \beta > 0)$ простору R^3.

5. $S = \{x = (\xi_1, \xi_2, ..., \xi_n) \in R^n : \xi_1^2 + \xi_2^2 + ... + \xi_n^2 = 1\}$ – сфера з радіусом 1 у просторі R^n.

6. V, W – опуклі множини простору R^n. Знайти $\operatorname{conv}(V \cup W)$.

Опуклі багатогранники

Нехай $x_1, x_2, ..., x_m$ – скінченна підмножина точок простору R^n. Опукла оболонка $M = \operatorname{conv}(x_1, x_2, ..., x_m)$ називається *опуклим багатогранником*, натягнутим на ці точки.

Теорема 3. *Існує єдина множина точок* $y_1, y_2, ..., y_p$ *опуклого багатогранника* $\operatorname{conv}(x_1, x_2, ..., x_m)$ *з такими властивостями*:

а) $\operatorname{conv}(y_1, y_2, ..., y_p) = \operatorname{conv}(x_1, x_2, ..., x_m)$;

б) $y_k \notin \operatorname{conv}(y_1, ..., y_{k-1}, y_{k+1} ..., y_p)$ *для всіх* $k = \overline{1, p}$.

Такими точками $y_1, y_2, ..., y_p$ будуть деякі точки множини $x_1, x_2, ..., x_m$. Вони називаються *вершинами* багатогранника.

Доведення. Якщо жодна з точок $x_1, x_2, ..., x_m$ не являє собою опуклу лінійну комбінацію решти точок, тоді точками $y_1, y_2, ..., y_p$ з властивостями а) і б) обираємо задані точки $x_1, x_2, ..., x_m$. Припустимо, що деяка з точок $x_1, x_2, ..., x_m$ опукло залежить від решти точок, наприклад точка x_1, тобто $x_1 \in \operatorname{conv}(x_2, ..., x_m)$. За властивістю II опуклих оболонок матимемо рівність $M = \operatorname{conv}(x_2, ..., x_m)$. Якщо жодна з точок $x_2, x_3, ..., x_m$ не є опуклою лінійною комбінацією решти, то шуканими точками $y_1, y_2, ..., y_p$ з властивостями а) і б) обираємо точки $x_2, x_3, ..., x_m$ і тоді $p = m - 1$. Продовжуючи далі ці міркування, ми, нарешті, знайдемо деякий набір точок $y_1 = x_{k_1}, y_2 = x_{k_2}, ..., y_p = x_{k_p}$ з потрібними властивостями.

А тепер доведемо єдиність. Нехай точки $w_1, w_2, ..., w_q \in M$ мають властивості а) і б). Тоді

$$y_1 = \sum_{j=1}^{q} \alpha_j w_j,\ \alpha_j \geq 0,\ \sum_{j=1}^{q} \alpha_j = 1,$$

$$w_j = \sum_{t=1}^{p} \beta_{jt} y_t,\ \beta_{jt} \geq 0,\ \sum_{t=1}^{p} \beta_{jt} = 1.$$

Тому

$$y_1 = \sum_{j=1}^{q} \alpha_j \left(\sum_{t=1}^{p} \beta_{jt} y_t \right) = \sum_{t=1}^{p} \left(\sum_{j=1}^{q} \alpha_j \beta_{jt} \right) y_t$$

– опукла лінійна комбінація точок $y_1, y_2, ..., y_p$:

$$y_1 = \sum_{t=1}^{p} \gamma_t y_t \text{ і } \sum_{t=1}^{p} \gamma_t = 1, \text{ де } \gamma_t = \sum_{j=1}^{q} \alpha_j \beta_{jt}.$$

Припустимо, що $\gamma_1 < 1$. Тоді $y_1 = \frac{\gamma_2}{1-\gamma_1} y_2 + ... + \frac{\gamma_p}{1-\gamma_1} y_p$ – також опукла лінійна комбінація, тобто $y_1 \in \operatorname{conv}(y_2, ..., y_p)$, що неможливо. Тому $\gamma_1 = 1, \gamma_2 = \gamma_3 = ... = \gamma_p = 0$. Одержано рівності $\sum_{j=1}^{q} \alpha_j \beta_{j1} = 1,\ \sum_{j=1}^{q} \alpha_j \beta_{jt} = 0$ при $t \geq 2$. Звідси $1 = \sum_{j=1}^{q} \alpha_j \beta_{j1} \leq \max \beta_{j1} \leq 1$. Тому $\max \beta_{j1} = 1$, тобто $\beta_{k1} = 1$ для деякого k, $1 \leq k \leq q$ і $\beta_{kt} = 0$ для всіх $t \neq 1$. Отже, $y_1 = w_k$ для деякого номера $k, 1 \leq k \leq q$.

Аналогічно доводиться $y_2 = w_l$ для деякого номера l, $1 \leq l \leq q$ і т. д. Таким чином, множина $y_1, y_2, ..., y_p$ виявилась підмножиною набору точок $w_1, w_2, ..., w_q$.

Якщо в наведених вище міркуваннях поміняти ролями набір точок $y_1, y_2, ..., y_p$ на $w_1, w_2, ..., w_q$, дістанемо: $w_1, w_2, ..., w_q$ –

підмножина набору точок $y_1, y_2, ..., y_p$. Отже, обидва набори точок збігаються.

Теорему доведено.

Означення. *Опуклий багатогранник* $M = \text{conv}(x_0, x_1, ..., x_p)$ *називається p-симплексом, якщо вектори* $x_1 - x_0, x_2 - x_0, ...,$ $x_p - x_0$ *лінійно незалежні.*

У цьому означенні роль точки x_0 може виконувати будь-яка інша точка з $x_0, x_1, ..., x_p$, наприклад x_1. Доведемо лінійну незалежність векторів $x_0 - x_1, x_2 - x_1, ..., x_p - x_1$. Нехай

$$\alpha_1(x_0 - x_1) + \alpha_2(x_2 - x_1) + ... + \alpha_p(x_p - x_1) = \theta .$$

Тоді з цієї рівності матимемо

$$(\alpha_1 + \alpha_2 + ... + \alpha_p)(x_0 - x_1) +$$
$$+\alpha_2(x_2 - x_0) + ... + \alpha_p(x_p - x_0) = \theta .$$

Звідси $\alpha_1 + \alpha_2 + ... + \alpha_p = 0,\ \ \alpha_2 = 0,\ \ \alpha_3 = 0, ..., \alpha_p = 0 \Rightarrow$ $\alpha_1 = 0$.

Теорема 4. *Всякий p-симплекс* $M = \text{conv}(x_0, x_1, ..., x_p)$ *– опуклий багатогранник з вершинами* $x_0, x_1, ..., x_p$. *Будь-яка його точка єдиним чином подається у вигляді опуклої лінійної комбінації його вершин.*

Доведення. Доведемо, що в симплексі M точки x_0, $x_1, ..., x_p$ будуть вершинами, тобто $x_k \notin \text{conv}(x_0, x_1, ..., x_{k-1},$ $..., x_p)$. Припустимо

$$x_k = \sum_{j \neq k} \lambda_j x_j,\ \sum_{j \neq k} \lambda_j = 1,\ \lambda_j \geq 0 .$$

Звідси дістаємо рівність $\sum\limits_{j \neq k} \lambda_j (x_j - x_k) = \theta$. Але вектори $x_0 - x_k, ..., x_{k-1} - x_k, x_{k+1} - x_k, ..., x_p - x_k$ лінійно незалежні. Прийшли до суперечності.

Нехай $x \in M$ і $x = \sum_{k=0}^{p} \lambda_k x_k = \sum_{k=0}^{p} \mu_k x_k$, $\lambda_k \geq 0$, $\mu_k \geq 0$, $\sum_{k=0}^{p} \lambda_k = \sum_{k=0}^{p} \mu_k = 1$. Тоді $\sum_{k=0}^{p} (\lambda_k - \mu_k) x_k = \theta$. Оскільки $\sum_{k=0}^{p} (\lambda_k - \mu_k) x_0 = \theta$, то дістаємо рівність $\sum_{k=0}^{p} (\lambda_k - \mu_k)(x_k - x_0) = \theta$. Але вектори $x_1 - x_0, x_2 - x_0, ..., x_p - x_0$ лінійно незалежні. Маємо рівності $\lambda_k = \mu_k$, $k = \overline{1, p}$, а тому $\lambda_0 = \mu_0$.

Теорема 5. *Усякий n-симплекс евклідового простору R^n має внутрішні точки.*

Доведення. Нехай M – n-симплекс з вершинами x_0, $x_1, ..., x_n$. Визначимо відображення f: $R^n \to R^n$ за таким правилом:

$$w = (\lambda_1, \lambda_2, ..., \lambda_n) \to \left(1 - \sum_{k=1}^{n} \lambda_k\right) x_0 + \lambda_1 x_1 + ... + \lambda_n x_n .$$

Позначимо через L множину всіх точок $w = (\lambda_1, \lambda_2, ..., \lambda_n)$ таких, що $\lambda_1 \geq 0$, $\lambda_2 \geq 0, ..., \lambda_n \geq 0$, $\lambda_1 + \lambda_2 + ... + \lambda_n \leq 1$. Тоді відображення f буде гомеоморфізмом L на M: воно взаємно однозначне і взаємно неперервне в обох напрямках.

Покажемо, що точка $w_0 = \left(\frac{1}{n+1}, \frac{1}{n+1}, ..., \frac{1}{n+1}\right)$ – внутрішня в множині L.

Нехай K – множина всіх точок $w = (\lambda_1, \lambda_2, ..., \lambda_n)$, координати яких задовольняють нерівності $\left|\lambda_1 - \frac{1}{n+1}\right| < \varepsilon$, $\left|\lambda_2 - \frac{1}{n+1}\right| < \varepsilon, ...,$ $\left|\lambda_n - \frac{1}{n+1}\right| < \varepsilon$, де $\varepsilon = \frac{1}{n(n+1)}$. Це – відкритий куб, який містить точку w_0, тобто K – окіл цієї точки.

Координати $\lambda_1, \lambda_2, ..., \lambda_n$ точок $w \in K$ додатні і $\lambda_1 + \lambda_2 + ... + \lambda_n < \frac{n}{n+1} + n\varepsilon = 1$. Це означає, що куб K повністю міститься у множині L. Образ $f(K)$ міститься у множині M і є околом точки $x_0 = f(w_0)$. Отже, $x_0 \in M^0$.

Теорему доведено.

Приклад 3. Довести, що вершинами n-вимірного куба K_n, породженого всіма точками виду $a_k = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$, де $\varepsilon_i = \pm 1, i = \overline{1, n}$, будуть усі ці 2^n точок.

Розв'язання. Для доведення використаємо індукцію за розмірністю n простору. Для $n = 1, 2$ твердження геометрично очевидне.

Припустимо, що воно має місце для куба K_{n-1} розмірності $n-1$. Для K_n доводимо методом від супротивного. Припустимо, що будь-яка з точок $a_1, a_2, ..., a_{2^n}$ виду $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$, $\varepsilon_i = \pm 1$ опукло залежить від решти. Нехай, наприклад, $a_1 = (1, 1, ..., 1) = \sum_{k=2}^{2^n} \lambda_k a_k, \lambda_k \geq 0, \sum_{k=2}^{2^n} \lambda_k = 1$. Якщо a_k подати у вигляді $(\underline{a}_k, \varepsilon_n)$, де $\underline{a}_k$ – вершина куба K_{n-1}, то матимемо рівність $\underline{a}_1 = \sum_{k=2}^{2^n} \lambda_k \underline{a}_k$.

Праворуч у рівності вершина $\underline{a}_1$ зустрічається лише 1 раз (вважатимемо, що $\underline{a}_1$ в сумі займає перше місце), а решта вершин $\underline{a}_k$ входить у суму двічі (вважатимемо, що вони розміщені поруч). Тоді $\underline{a}_1 = \Lambda_1 \underline{a}_1 + \Lambda_2 \underline{a}_2 + ... + \Lambda_{2^{n-1}} \underline{a}_{2^{n-1}}$, де $\Lambda_1 = \lambda_2 < 1$, а решта коефіцієнтів $\Lambda_2 = \lambda_3 + \lambda_4$, $\Lambda_3 = \lambda_5 + \lambda_6$ і т. д. При цьому $\sum_{k=1}^{2^{n-1}} \Lambda_k = 1$. Звідси знаходимо

$$\underline{a}_1 = \sum_{k \geq 2} \frac{\Lambda_k}{1 - \Lambda_1} \underline{a}_k, \ \frac{\Lambda_k}{1 - \Lambda_1} \geq 0, \ \sum_{k \geq 2} \frac{\Lambda_k}{1 - \Lambda_1} = 1.$$

Отже, вершина $\underline{a}_1$ куба K_{n-1} опукло залежить від решти вершин. Прийшли до суперечності.

Теорему доведено.

Вправи

1. Довести, що опуклий багатогранник – компактна множина (тобто множина обмежена і замкнена).

2. Довести, що образ опуклого багатогранника при лінійному перетворенні $\mathcal{A}$ простору R^n є опуклий багатогранник, і що, коли лінійне перетворення $\mathcal{A}$ оборотне, вершини багатогранника при цьому перетворенні переходять у вершини образу.

§ 9.4. Внутрішні та граничні точки опуклої множини

Лема. *Нехай a – внутрішня точка опуклої множини $\overline{V}$, точка b належить замиканню $\overline{V}$ цієї множини. Тоді кожна точка відрізка $[a,\ b]$, за виключенням, можливо, точки b, є внутрішня для множини V.*

Доведення. Для достатньо малого числа $\varepsilon > 0$ відкрита куля $S(a, \varepsilon)$ міститься у множині V. Оберемо будь-яку точку $c_\lambda = (1-\lambda)a + \lambda b$ за умови $0 \le \lambda < 1$. Покажемо, що вона є внутрішньою для V.

Розглянемо кулю $W = S\left(c_\lambda,\ (1-\lambda)\varepsilon\right)$ і доведемо, що $W \subset V$. Оберемо будь-яку точку $w \in W$. Для неї $\|w - c_\lambda\| < (1-\lambda)\varepsilon$. Для будь-якої точки $v \in V$ маємо $\left\|w - [(1-\lambda)a + \lambda v]\right\| = \left\|(w - c_\lambda) + \lambda(b - v)\right\| \le \|w - c_\lambda\| + \lambda\|b - v\|$. Відшукаємо точку v із множини V настільки близьку до точки b, щоб виконувалась нерівність

$$\|w - c_\lambda\| + \lambda\|b - v\| < (1-\lambda)\varepsilon .$$

Це можливо, оскільки різниця $(1-\lambda)\varepsilon - \|w - c_\lambda\|$ додатна. Отже, можна підібрати таку точку $v \in V$, для якої

$$\left\|w - [(1-\lambda)a + \lambda v]\right\| < (1-\lambda)\varepsilon .$$

Звідси дістанемо нерівність

$$\left\|\frac{1}{1-\lambda}(w - \lambda v) - a\right\| < \varepsilon .$$

Отже, точка $x = \frac{1}{1-\lambda}(w - \lambda v) \in S(a, \varepsilon)$; $w = (1-\lambda)x + \lambda v \in V$.

Але точку w обрано в W довільно. Тому $W = S(c_\lambda, (1-\lambda)\varepsilon)$ міститься в V. Точка c_λ – внутрішня точка множини V.

Нагадаємо означення граничної точки множини V простору R^n.

Точка $g \in R^n$ називається *граничною* для V, якщо будь-який її ε-окіл $S(g,\varepsilon)$ містить точки із V, а також точки з доповнення $R^n \setminus V$.

Твердження 1. *Нехай V – опукла відкрита множина простору R^n. Кожна точка $g \in \overline{V} \setminus V$ є гранична для замикання $\overline{V}$ множини V.*

Доведення проведемо методом від супротивного. Припустимо, що існує ε-окіл $S(g,\varepsilon)$ $(\varepsilon > 0)$, що міститься в $\overline{V}$. Оскільки $S(g,\varepsilon) \cap V \neq \varnothing$, то оберемо деяку точку $a \in S(g,\varepsilon) \cap V$. Тоді $b = g + (g - a) \in S(g,\varepsilon)$. Отже, $g = \frac{1}{2}a + \frac{1}{2}b$, де $a \in V = V^0$ і $b \in \overline{V}$. Із леми випливає, що точка g належить множині $V = V^0$. Це суперечить умові твердження.

Твердження 2. *Будь-яка гранична точка опуклої множини V є граничною і для її замикання $\overline{V}$.*

Доведення. Нехай g – гранична точка множини V. Припустимо, що існує ε-окіл $S(g,\varepsilon)$ $(\varepsilon > 0)$, який не містить точок доповнення $R^n \setminus \overline{V}$, тобто містився б у $\overline{V}$. Це означає, що множина $V \cap S(g,\varepsilon)$ всюди щільна в $S(g,\varepsilon)$. Тому існує $n+1$ таких точок $x_0, x_1, ..., x_n \in V \cap S(g,\varepsilon)$, що їх опукла оболонка $\mathrm{conv}(x_0, x_1, ..., x_n)$ буде n-симплексом. Він міститься у множині V і має внутрішні точки. Отже, внутрішня область V^0 множини V непорожня.

Як і при доведенні твердження 1 існують такі точки $a \in V^0 \cap S(g,\varepsilon)$ і $b \in S(g,\varepsilon) \subset \overline{V}$, що $g = \frac{1}{2}a + \frac{1}{2}b$. На підставі леми встановлюємо $g \in V^0$. Це суперечить умові твердження.

Отже, будь-який ε-окіл $S(g,\varepsilon)$ точки g містить точки множини $R^n \setminus \overline{V}$. Вона – гранична для множини $\overline{V}$.

Твердження 3. *Якщо внутрішня область V^0 опуклої множини V непорожня, то $\overline{V} = \overline{V^0}$.*

Доведення. Із включення $V^0 \subseteq V$ випливає $\overline{V^0} \subseteq \overline{V}$. Нехай $a \in V^0$ і $b \in \overline{V}$. За лемою всі точки відрізка [a, b], за винятком, можливо, точки b, є внутрішні для множини V. Тому в будь-якому околі точки b розміщуватимуться точки внутрішньої області V^0. Отже, $b \in \overline{V^0}$. Доведено обернене включення $\overline{V} \subseteq \overline{V^0}$, а разом з тим і рівність $\overline{V} = \overline{V^0}$.

Твердження 4. *На кожному промені, що виходить з внутрішньої точки опуклої множини (якщо такі існують), є не більше однієї граничної точки цієї множини.*

Доведення. Нехай V – опукла множина, $V^0 \neq \varnothing$ і $a \in V^0$. Припустимо, що промінь l, який виходить із a, має дві різні граничні точки b, c множини V. Нехай при цьому $c \in [a, b]$. За лемою точка c – внутрішня для V. Дістали суперечність.

Твердження 5. *Якщо опукла множина V не містить внутрішніх точок, то вона міститься в деякій гіперплощині*

$$\Pi = \left\{ x \in R^n : (a, x) = \beta \right\},\ a \neq \theta .$$

Доведення. Оберемо деяку точку $x_0 \in V$ і розглянемо опуклу множину $W = -x_0 + V$. Вона містить нульову точку θ і її лінійна оболонка $M = L(W)$ є лінійним підпростором простору R^n.

Нехай розмірність підпростору M дорівнює p і y_1, $y_2, ..., y_p$ – деякий його базис. Доведемо, що $p < n$. Для цього припустимо, що $p = n$. Тоді опукла оболонка conv(θ, y_1, $y_2, ..., y_n$) є n-симплексом з вершинами θ, $y_1, y_2, ..., y_n$, що містяться у множині W. Тому точки x_0, $x_0 + y_k = x_k$ $(k = \overline{1, n})$ породжують n-симплекс conv$(x_0, x_1, x_2, ..., x_n)$, що міститься в опуклій множині V. Але оскільки n-симплекси містять внутрішні точки, то й множина V матиме внутрішні точки. Це суперечить припущенню про V.

Із нерівності $p < n$ згідно з теорією про лінійні багатовиди випливає, що багатовид $L = x_0 + M$ складається з усіх розв'язків відповідної системи з $n - p$ лінійних рівнянь (§ 4.4).

Стосовно нашого випадку це означає, що опукла множина V належить перетину $n - p$ гіперплощин $(n - p \geq 1)$.

§ 9.5. Віддільність опуклих множин

Означення. *Нехай V – опукла замкнена множина точок простору R^n і $g \in R^n$. Точка $v \in V$ називається проекцією g на множину, якщо*

$$\| g - v \| = \inf_{x \in V} \| g - x \|.$$

Інакше кажучи, проекція v точки g – це точка найменшого відхилення точок $x \in V$ від точки g.

Теорема 1 (про оптимізацію). *Опукла замкнена множина W простору R^n містить єдину точку з мінімальною нормою.*

Доведення. Нехай $\delta = \inf_{x \in W} \| x \|$. Оберемо таку послідовність $\{x_m\}$ точок із W, що $\lim_{m \to \infty} \| x_m \| = \delta$.

Скористаємось законом паралелограма:

$$\| x + y \|^2 + \| x - y \|^2 = 2\left(\| x \|^2 + \| y \|^2 \right).$$

Поклавши $x = \frac{1}{2} x_k$ і $y = \frac{1}{2} x_m$, матимемо

$$\left\| \frac{1}{2}(x_k - x_m) \right\|^2 = \frac{1}{2}\left(\| x_k \|^2 + \| x_m \|^2 \right) - \left\| \frac{1}{2} x_k + \frac{1}{2} x_m \right\|^2.$$

Оскільки $\frac{1}{2} x_k + \frac{1}{2} x_m \in W$, то $\left\| \frac{1}{2} x_k + \frac{1}{2} x_m \right\|^2 \geq \delta^2$ і $\left\| \frac{1}{2}(x_k - x_m) \right\|^2 \leq \frac{1}{2}\left(\| x_k \|^2 + \| x_m \|^2 \right) - \delta^2$. При $k, m \to \infty$ права частина нерівності прямує до нуля. Отже, $\|x_k - x_m\| \to 0$ при $k, m \to \infty$. Послідовність $\{x_m\}$ фундаментальна. За теоремою Коші вона збігається до деякої точки w. За умовою множина W – замкнена, а тому $w \in W$. Маємо $\|w\| \le \|w - x_m\| + \|x_m\|$. Перейшовши до границі, дістанемо $\|w\| \le \delta$, а звідси $\|w\| = \delta$. Отже, точка w має мінімальну норму.

Тепер доведемо єдиність існування точки з мінімальною нормою. Нехай $w, w' \in W$ і $\|w\| = \delta, \|w'\| = \delta$. За законом паралелограма маємо

$$\left\|\frac{1}{2}(w - w')\right\|^2 = \frac{1}{2}\left(\|w\|^2 + \|w'\|^2\right) - \left\|\frac{1}{2}w + \frac{1}{2}w'\right\|^2 \le 0.$$

Тому $\|w - w'\| = 0$, і, отже, $w = w'$.

Теорема 2. *Якщо V – опукла замкнена множина і $g \in R^n$, то у множині V знайдеться проекція v точки g, причому ця проекція єдина.*

Доведення. Розглянемо опуклу множину $W = -g + V$. Вона замкнена. За теоремою 1 знайдеться єдиний елемент $w = -g + v \in W$ з мінімальною нормою. Точка v – проекція g на V.

Теорема 3. *Нехай V – опукла замкнена множина точок простору R^n. Для того щоб точка $v \in V$ була проекцією точки $g \in R^n$, необхідно і достатньо, щоб виконувалась нерівність*

$$(g - v, x - v) \le 0 \tag{1}$$

для всіх точок $x \in V$.

Доведення. *Необхідність.* Нехай v – проекція точки g на множину V. Якщо $x \in V$, то $z(\lambda) = (1-\lambda)v + \lambda x \in V$ для будь-яких $\lambda \in [0, 1]$. Розглянемо функцію $F(\lambda) = \|z(\lambda) - g\|^2 = A\lambda^2 + 2B\lambda + C$. Її коефіцієнти $A = \|x - v\|^2$, $B = (v - g, x - v)$, $C = \|v - g\|^2$. Оскільки $\|z(\lambda) - g\|^2 \ge \|v - g\|^2$, то $A\lambda^2 + 2B\lambda \ge 0$ і при $\lambda > 0$ матимемо $A\lambda + 2B \ge 0$. Ця нерівність правильна для достатньо малих значень $\lambda > 0$. Тому $B \geq 0$, $(v-g,\ x-v) \geq 0$ і, отже, $(g-v,\ x-v) \leq 0$.

Достатність. Нехай нерівність (1) виконується для будь-яких $x \in V$. Тоді $\| x-g \|^2 = \| (x-v)+(v-g) \|^2 = \| x-v \|^2 + 2(x-v,\ v-g) + \| v-g \|^2 \geq \| v-g \|^2$. Точка v – проекція g на V.

Означення. *Гіперплощина* Π *у просторі* R^n *відокремлює непорожні множини A і B точок, якщо A міститься в одному із замкнених півпросторів, що визначаються гіперплощиною* Π, *а B – в іншому півпросторі.*

Гіперплощина П строго відокремлює A і B, якщо A міститься в одному з відкритих півпросторів, визначених гіперплощиною П, а B – в іншому.

Означення. *Нехай A – деяка множина точок із* R^n. *Гіперплощина* $\Pi = \left\{ x \in R^n : (a,\ x) = \beta \right\}$, $a \neq \theta$, *називається опорною для A, якщо виконуються умови:*

а) множина A міститься в одному із замкнених півпросторів, що визначаються гіперплощиною Π;

б) гіперплощина Π *має хоч би одну спільну точку з множиною A.*

Теорема 4 (теорема віддільності). *Нехай V – опукла і замкнена множина. Якщо точка* $g \in R^n$ *не належить V, то існує така опорна гіперплощина* $\Pi = \left\{ x \in R^n : (a,\ x) = \beta \right\}$ *для V, що* $(a,\ x) \leq \beta$ *для будь-яких* $x \in V$ *і* $(a,\ g) > \beta$.

Доведення. Нехай v – проекція точки g на множину V. Тоді $(g-v,\ x-v) \leq 0$ для всіх $x \in V$. Вектор $a = g - v \neq \theta$. Гіперплощина $\Pi = \left\{ x \in R^n : (a,\ x-v) = 0 \right\}$ є опорна для V.

Позначимо $\beta = (a, v)$. Тоді $\Pi = \left\{ x \in R^n : (a,\ x) = \beta \right\}$ і $(a,\ x) \leq \beta$ для всіх $x \in V$, $(a,\ g) > (a,\ v) = \beta$.

Наслідок. *Для опуклої та замкненої множини V і точки* $g \overline{\in} V$ *існує гіперплощина* $\Pi' = \left\{ x \in R^n : (a,\ x) = \gamma \right\}$, *що проходить через g, і така, що* $(a,\ x) < \gamma$ *для будь-яких* $x \in V$.

Такою гіперплощиною є $\Pi' = \{x : (a, x - g) = 0\}$, де $a = g - v$. Якщо $x \in V$, матимемо $(a, x - g) = (a, x - v + (v - g)) = (a, x - v) + (a, v - g) = (a, x - v) - (a, a) < 0$. Позначимо $\gamma = (a, g)$. Тоді $\Pi' = \{x : (a, x) = \gamma\}$ і $g \in \Pi'$, $(a, x) < \gamma$ для будь-яких $x \in V$.

Теорема 5 (про опорну гіперплощину). *Через граничну точку $v \in V$ опуклої множини V проходить опорна гіперплощина цієї множини: існує вектор $a \neq \theta$ і скаляр γ такі, що гіперплощина $\Pi = \left\{x \in R^n : (a, x) = \gamma\right\}$ має властивості $(a, v) = \gamma$ і $(a, x) \leq \gamma$, $x \in V$.*

Доведення. Гранична точка v опуклої множини V є одночасно і граничною точкою для замикання $\bar{V}$ (див. § 9.3, теорема 1). Нехай $\{v_m\}$ – послідовність точок, що не належить $\bar{V}$ і збігається до точки v. За попередньою теоремою та її наслідком для кожної точки v_m послідовності існує гіперплощина

$$\Pi_m = \left\{x \in R^n : (a_m, x) = \gamma_m\right\}$$

з властивостями: $\gamma_m = (a_m, v_m)$ і $(a_m, x) \leq \gamma_m$ для будь-яких $x \in V$.

Можна вважати, що всі вектори a_m нормовані: $\|a_m\| = 1$. Вони подаються точками сфери одиничного радіуса. Сфера – компактна множина. Тому із послідовності $\{a_m\}$ можна виділити збіжну підпослідовність. Вважатимемо, що сама послідовність $\{a_m\}$ збігається: $a_m \to a$ при $m \to \infty$ і $\|a\| = 1$.

Якщо в рівняннях гіперплощин

$$(a_m, x) = \gamma_m$$

перейдемо до границі, то дістанемо рівняння гіперплощини Π:

$$(a, x) = \gamma,$$

де

$$\gamma = \lim_{m \to \infty} \gamma_m = \lim_{m \to \infty} \left(a_m, v_m\right) = (a, v).$$

При цьому $(a, x) \leq \gamma$ для всіх $x \in V$.

Теорема 6. *Якщо V – опукла відкрита множина і $g \overline{\in} V$, то існує гіперплощина*

$$\Pi = \left\{x \in R^n : (a, x) = \gamma\right\},$$

що проходить через точку g, і така, що $(a, x) < \gamma$ для всіх $x \in V$.

Доведення. Якщо $g \overline{\in} \overline{V}$, тоді твердження теореми випливає з наслідку теореми 4. Якщо $g \in \overline{V} \setminus V$, тоді на підставі теореми 5 існує гіперплощина $\Pi = \left\{x \in R^n : (a, x) = \gamma\right\}$, яка проходить через точку g, і така, що $(a, x) \le \gamma$ для будь-яких $x \in \overline{V}$. Але V – відкрита множина. Тому $(a, x) < \gamma$ для будь-яких $x \in V$.

Теорема 7 (про відокремлюючу гіперплощину). *Нехай U і V – опуклі множини, внутрішня область U^0 непорожня і $U^0 \cap V = \varnothing$. Існує гіперплощина, яка поділяє U і V.*

Доведення. Алгебрична різниця $W = U^0 - V = U^0 + (-1)V$ – відкрита опукла множина, яка не містить нульову точку θ. Тоді існує гіперплощина $\Pi = \{x : (a, x) = 0\}$, що проходить через точку θ, і така, що $(a, x) < 0$ для всіх точок $x \in U^0 - V$. Звідси знаходимо, що для будь-яких $u \in U^0$ і $v \in V$ виконується нерівність $(a, u) < (a, v)$. Тому $\sup\limits_{u \in U^0} (a, u) \le \inf\limits_{v \in V} (a, v)$.

Оберемо число β таким, щоб виконувались нерівності

$$\sup_{u \in U^0} (a, u) \le \beta \le \inf_{v \in V} (a, v).$$

Гіперплощина $\Pi = \left\{x \in R^n : (a, x) = \beta\right\}$ відділяє множини U^0 і V. Оскільки $\overline{U^0} \supset U$, то

$$(a, x) \le \beta \text{ для всіх } x \in U,$$
$$(a, x) \ge \beta \text{ для всіх } x \in V.$$

Гіперплощина Π відділяє множини U і V.

Теореми віддільності опуклих множин застосовуються при доведенні важливих тверджень у лінійній алгебрі, математичному програмуванні, теорії опуклих множин тощо.

Теорема Фаркаша про нерівності-наслідки

Означення. *Нехай задано систему лінійних однорідних нерівностей*

$$\begin{cases} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n \le 0, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \ldots + \alpha_{2n}\xi_n \le 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n \le 0. \end{cases} \tag{1}$$

Нерівність

$$\beta_1\xi_1 + \beta_2\xi_2 + \ldots + \beta_n\xi_n \le 0 \tag{2}$$

називається наслідком системи (1), якщо будь-який розв'язок цієї системи задовольняє нерівність (2).

Це означення дамо у більш стислій формі. Позначимо через $A = (\alpha_{ik})$ матрицю системи (1). Нехай $x = (\xi_1, \xi_2, \ldots, \xi_n)^T$, $b = (\beta_1, \beta_2, \ldots, \beta_n)^T$. Тоді систему (1) можна подати як

$$Ax \le \theta,$$

а нерівність –

$$(b, x) = b^T x \le 0.$$

Останнє є наслідком системи (1), якщо

$$Ax \le \theta \Rightarrow (b, x) \le 0, \quad b \ne \theta.$$

Теорема Фаркаша. *Для того щоб*

$$Ax \le \theta \Rightarrow (b, x) \le 0, \quad b \ne \theta.$$

необхідно і достатньо, щоб система лінійних рівнянь

$$A^T y = b \tag{3}$$

мала невід'ємний розв'язок y (тобто існував би вектор $y \in R^m$ і $y \ge \theta$).

Доведення. *Необхідність.* Нехай із $Ax \leq \theta$ випливає $(b, x) \leq 0$. Множина

$$K = \left\{ x \in R^n : x = A^T y, \ y \geq \theta \right\}$$

– багатогранний конус, породжений рядками матриці A.

Якщо вектор $b \in R^n$ міститься в конусі K, тоді твердження теореми вірне.

Припустимо, що $b \notin K$. Конус K опуклий і замкнений. Застосуємо до точки b і конуса K теорему віддільності (теорему 4). Тоді існує гіперплощина $\Pi = \{x : (a, x) = \beta\}$ така, що $(a, b) > \beta$ і $(a, x) \leq \beta$ для всіх точок $x \in K$.

Інакше кажучи, існує такий вектор $a \neq \theta$, що $(b, a) > \beta \geq (a, x)$ для всіх $x \in K$.

Оскільки конус K містить нульовий вектор, то

$$(b, a) > 0. \tag{4}$$

Якщо замість вектора x із K покласти λx, де $\lambda > 0$, то матимемо нерівність $(b, a) > (a, \lambda x)$ і, отже,

$$\frac{1}{\lambda}(b, a) > (a, x).$$

При $\lambda \to \infty$ дістаємо

$$(a, x) \leq 0 \text{ для всіх } x \in K. \tag{5}$$

З другого боку, маємо $(a, x) = a^T \cdot x = a^T \cdot A^T y = (Aa)^T \cdot y = (Aa, y) \leq 0$, де $y \geq \theta$. Звідси

$$Aa \leq \theta. \tag{6}$$

Із останньої нерівності за умовою має бути $(b, a) \leq 0$, що суперечить нерівності (4). Інша можливість виключена.

Достатність. Нехай існує такий вектор $y \geq \theta$, що $b = A^T y$. Якщо $x \in R^n$ і $Ax \leq \theta$, то $(b, x) = b^T \cdot x = (A^T y)^T \cdot x = y^T \cdot Ax = (y, Ax) \leq \theta$.

Отже, $Ax \leq \theta \Rightarrow (b, x) < 0$.

Крайні точки

Означення. *Точка v опуклої множини V називається крайньою (кутовою), якщо вона не є внутрішньою точкою відрізка з кінцями із множини V.*

Це означення можна подати у такому вигляді: точка v множини V – крайня, якщо із $v=(1-\lambda)x+\lambda y$, $0<\lambda<1$, $x, y \in V$, випливає $x=y=v$.

Приклад. Показати, що крайніми точками опуклого багатогранника будуть його вершини і що інших крайніх точок він не має.

Розв'язання. Нехай вершинами опуклого багатогранника M будуть точки $x_1, x_2, ..., x_m$. Тоді $M=\operatorname{conv}(x_1, x_2, ..., x_m)$ і $x_k \notin \operatorname{conv}(x_1, ..., x_{k-1}, x_{k+1}, ..., x_m)$, $k=\overline{1, m}$.

Спочатку покажемо, що кожна вершина, наприклад, x_1 – крайня точка багатогранника M. Припустимо, що $x_1=(1-\lambda)x+$ $+\lambda y$, $0<\lambda<1$, $x, y \in M$. Тоді $x=\sum_{k=1}^{m}\lambda_k x_k$, $\lambda_k \geq 0$, $\sum_{k=1}^{m}\lambda_k=1$; $y=\sum_{k=1}^{m}\mu_k x_k$, $\mu_k \geq 0$, $\sum_{k=1}^{m}\mu_k=1$. Тому $x_1=\gamma_1 x_1+\gamma_2 x_2+...+$ $+\gamma_m x_m$, де $\gamma_k=(1-\lambda)\lambda_k+\lambda\mu_k$, $\sum_{k=1}^{m}\gamma_k=1$.

Припустимо, що $\gamma_1<1$. Тоді $x_1=\frac{\gamma_2}{1-\gamma_1}x_2+...+\frac{\gamma_m}{1-\gamma_1}x_m$ – опукла лінійна комбінація точок $x_2, x_3, ..., x_m$. Для вершин багатогранника M це неможливо. Звідси випливає, що $\gamma_1=1$, $\gamma_2=...=\gamma_m=0$. Із рівності $(1-\lambda)\lambda_k+\lambda\mu_k=0$ дістанемо $\lambda_k=\mu_k=0$, $k=\overline{2, m}$. Тому $\lambda_1=\mu_1=1$ і, отже, $x=y=x_1$.

Нехай v – крайня точка багатогранника M і $v=\sum_{k=1}^{m}\lambda_k x_k$, $\lambda_k \geq 0$, $\sum_{k=1}^{m}\lambda_k=1$. Припустимо, що $0<\lambda_1<1$. Тоді

$$\lambda = \lambda_2 + \ldots + \lambda_m > 0,\ \lambda_1 + \lambda = 1,\ v = \lambda_1 x_1 + \lambda x,$$

де $x = \frac{\lambda_2}{\lambda} x_2 + \ldots + \frac{\lambda_m}{\lambda} x_m \in M$.

Це неможливо для крайньої точки v, оскільки $x_1 \neq x$. Тому або $\lambda_1 = 0$, або $\lambda_1 = 1$. Такий самий висновок є вірним для будь-якого коефіцієнта λ_k. Це означає, що точка v збігається з однією з вершин $x_1, x_2, \ldots, x_m$.

Теорема 8. *Опукла компактна (тобто обмежена і замкнена) множина V простору R^n має крайні точки. Вона збігається з опуклою оболонкою всіх своїх крайніх точок.*

Доведення проведемо методом математичної індукції за розмірністю n евклідового простору R^n.

Якщо $n = 1$, тоді множина V – відрізок $[a, b]$ і теорема очевидна.

Припустимо, що теорема має місце для опуклих компактних множин простору R^{n-1}. Нехай множина $V \subseteq R^n$. Якщо множина V не містить внутрішніх точок, то вона буде підмножиною деякої гіперплощини

$$\Pi = \left\{x \in R^n : (a, x) = \beta\right\},\ a \neq \theta.$$

Але будь-яка гіперплощина буде евклідовим простором розмірності $n-1$. Тому твердження теореми матиме силу за припущенням індукції.

Нехай множина V має внутрішні точки. Оберемо довільну точку $x_0 \in V$. Окремо розглянемо два випадки.

Випадок 1. Обрана точка x_0 – гранична точка множини V. Існує для V опорна гіперплощина

$$\Pi = \left\{x \in R^n : (a, x) = \beta\right\},$$

що проходить через x_0 (теорема 5). При цьому $(a, x_0) = \beta$ і $(a, x) \leq \beta$ для будь-яких $x \in V$.

Множина $V_0 = V \cap \Pi$ – опукла, компактна і належить евклідовому простору Π розмірності $n-1$. За припущенням індукції множина V_0 має крайні точки і збігається з опуклою оболонкою всіх її крайніх точок. Зокрема, $x_0 = \sum_{k=1}^{m} \lambda_k x_k$, $\lambda_k \geq 0$, $\sum_{k=1}^{m} \lambda_k = 1$ і $x_1, x_2, ..., x_m$ – крайні точки цієї множини.

Покажемо, що будь-яка крайня точка v множини V_0 одночасно є і крайньою точкою множини V.

Нехай $v = (1-\lambda)x + \lambda y$, де $0 < \lambda < 1$, $x, y \in V$. Маємо

$$\beta = (a, v) = (a, (1-\lambda)x + \lambda y) =$$
$$= (1-\lambda)(a, x) + \lambda(a, y) \leq (1-\lambda)\beta + \lambda\beta = \beta,$$

оскільки $(a, x) \leq \beta$ і $(a, y) \leq \beta$. Тому має місце рівність

$$(1-\lambda)(a, x) + \lambda(a, y) = \beta.$$

Це можливо лише у випадку, коли $(a, x) = \beta$, $(a, y) = \beta$, тобто коли $x, y \in V \cap \Pi = V_0$. Але v – крайня точка множини V_0, тому $x = y = v$. Отже, точка x_0 опукло залежить від крайніх точок $x_1, x_2, ..., x_m$ множини V.

Випадок 2. Обрана точка x_0 – внутрішня точка для V. Через цю точку проведемо пряму лінію. Вона перетне границю множини V у двох точках y_1 і y_2. При цьому $x_0 = (1-\alpha)y_1 + \alpha y_2$, $0 < \alpha < 1$. Але точки y_1, y_2 опукло залежать від крайніх точок множини V. Тому і x_0 опукло залежить від крайніх точок множини V.

Вправи

1. Показати, що множина всіх крайніх точок кулі

$$\overline{S}(\theta, r) = \left\{ x \in R^n : \| x \| \leq r \right\}$$

збігається зі сферою

$$C = \left\{ x \in R^n : \| x \| = r \right\}.$$

2. Знайти множину всіх крайніх точок опуклої множини

$$V = \left\{ x = (\xi_1, \xi_2, \ldots, \xi_n) : \sum_{k=1}^{n} \frac{\xi_k^2}{\alpha_k^2} \le 1 \right\}.$$

§ 9.6. Замкнені конуси

Опуклому конусу K поставимо у відповідність множину $K^* = \left\{ x \in R^n : (w, x) \le 0 \text{ для всіх } w \in K \right\}$.

Ця множина K^* є конус. Якщо $x,\ y \in K^*$, $\lambda \ge 0$, $\mu \ge 0$, то $(w, \lambda x + \mu y) = \lambda(w, x) + \mu(w, y) \le 0$ для будь-якого $w \in K$. Це означає, що $\lambda x + \mu y \in K^*$, тобто K^* – конус.

Означення. *Конус K^* називається спряженим для конуса K.*

Спряжений конус K^* – замкнена множина, оскільки він збігається з перетином усіх півпросторів $P_w = \{ x : (w, x) \le 0 \}$, де w пробігає конус K.

Покажемо, що у випадку, коли конус K є лінійним підпростором, спряжений конус K^* дорівнює ортогональному доповненню $K^\perp$.

Нехай $y \in K^*$. Тоді $(x, y) \le 0$ для всіх $x \in K$. Але $-x \in K$ і тому $(-x, y) = -(x, y) \le 0$. Звідси дістанемо $(x, y) = 0$ і, отже, y належить ортогональному доповненню $K^\perp$ лінійного підпростору K. Отже, $K^* \subseteq K^\perp$. Із означення множини K^* випливає включення $K^\perp \subseteq K^*$. Таким чином, $K^* = K^\perp$.

Теорема 1. *Для будь-якого конуса K виконується включення $K^{**} \supseteq K$. Конус K замкнений тоді і лише тоді, коли $K^{**} = K$ (відношення двоїстості).*

Доведення. Справедливість включення $K^{**} \supseteq K$ випливає з означення спряженого конуса. Якщо $K^{**} = K$, то K – замкнена множина, оскільки $K^{**} = \left(K^* \right)^*$ – конус замкнений.

Припустимо, що K – замкнений конус. Потрібно встановити включення $K^{**} \subseteq K$, тобто довести, що $g \in K^{**} \Rightarrow g \in K$. За законом логіки це зводиться до доведення висновку $g \overline{\in} K \Rightarrow g \overline{\in} K^{**}$.

За теоремою віддільності (§ 9.5, теорема 5) для точки g та опуклої, замкненої множини K існує гіперплощина $\Pi = \{x : (a, x) = \beta\}$ така, що $(a, g) > \beta$ і $(a, y) \leq \beta$ для всіх $y \in K$. Для $y \in K$ і $\lambda > 0$ маємо $\lambda y \in K$, а тому $(a, \lambda y) \leq \beta$. Звідси $(a, y) \leq \frac{\beta}{\lambda}$. Збільшуємо λ необмежено і дістаємо $(a, y) \leq 0$. Оскільки це вірно для всіх $y \in K$, то $a \in K^*$, а з того, що $\theta \in K$, знаходимо $0 \leq \beta$. Тому $(a, g) > \beta \geq 0$ і, отже, $g \overline{\in} K^{**}$.

Теорема 2. *Для того щоб замкнений конус K був загостреним, необхідно і достатньо, щоб спряжений йому конус K^* мав внутрішні точки.*

Доведення. *Необхідність.* Припустимо, що конус K загострений, але K^* не має внутрішніх точок. Тоді K^* міститься у деякій гіперплощині $\Pi = \{x : (a, x) = 0\}$. Оскільки $(a, x) = (-a, x) = 0$ для всіх $x \in K^*$, то $a, -a \in K^{**} = K$. Для загостреного конуса таке включення неможливе.

Достатність. Припустимо, що K^* має внутрішні точки і, разом з тим, існує $a \neq \theta$ і $a, -a \in K$. Якщо $x \in K^*$, тоді $(a, x) \leq 0$ і $(-a, x) \leq 0$. Отже, $(a, x) = 0$ для всіх $x \in K^*$, тобто K^* міститься в гіперплощині з рівнянням $(a, x) = 0$. Знову виникає суперечність.

Теорема 3. *Якщо конус K замкнений і загострений, то для нього існує вектор $a \neq \theta$ такий, що $(a, x) < 0$ для будь-яких $x \in K$, $x \neq \theta$.*

Доведення. Оберемо яку-небудь внутрішню точку a із спряженого конуса K^*. При достатньо малому $\varepsilon > 0$ відкрита куля $S(a, \varepsilon)$ міститься в K^*. Якщо $x \in K$ і $x \neq \theta$, то

$y = a + \dfrac{\varepsilon}{2\|x\|}x \in S(a,\varepsilon)$. Отже, $(y,x) \le 0$, тобто $(a,x) + \dfrac{\varepsilon}{2}\|x\| \le 0$ і $(a,x) \le -\dfrac{\varepsilon}{2}\|x\| < 0$.

Наслідок. *Для замкненого загостреного конуса K існує $a \ne \theta$, що $(a,x) > 0$ для всіх $x \in K$, $x \ne \theta$.*

Означення. *Промінь $l \subset K$, що виходить із нульової точки θ, називається крайнім для конуса K, якщо має місце включення: із $w \in l$ і $w = (1-\lambda)x + \lambda y$, де $0 < \lambda < 1$, $x, y \in K$, випливає, що $x, y \in l$. Інакше кажучи, промінь $l \subset K$ – крайній, якщо будь-який відрізок з кінцями із K, і який має внутрішню точку, що належить l, цілком належить l.*

Теорема 4. *Нехай K – замкнений загострений конус, що містить ненульові точки. Тоді він матиме крайні промені. Оберемо на кожному крайньому промені один який-небудь його напрямний вектор і через S позначимо множину всіх таких векторів. Конус, породжений цією множиною, є $K = L^+(S)$.*

Доведення. Для конуса K оберемо такий вектор $a \ne \theta$, що $(a,x) > 0$ для всіх $x \in K$, $x \ne \theta$.

Розглянемо гіперплощину

$$\Pi = \left\{x \in R^n : (a,x) = 1\right\}.$$

Кожний промінь l конуса K (тобто промінь, що виходить із θ і міститься в K) перетинає П лише в одній точці. Дійсно, якщо $w \in l$ і $w \ne \theta$, то $(a,w) > 0$. Тоді $x = \dfrac{1}{(a,w)}w$ належить промeню l. Разом з тим $(a,x) = 1$, тобто $x \in \Pi$. Отже, $x \in l \cap \Pi$.

Якщо $y \in l \cap \Pi$, то $y = \alpha w$ і

$$(a,y) = (a,\alpha w) = \alpha(a,w) = 1,\ \ \alpha = \frac{1}{(a,w)},$$

тобто

$$y = \frac{1}{(a,w)} w = x .$$

Множина $V = K \cap \Pi$ опукла і замкнена. Нами встановлено взаємно однозначну відповідність між всіма променями l конуса K і точками множини $V: l \leftrightarrow x = l \cap \Pi$. При цьому $x = \frac{1}{(a,w)} w$, якщо $w \in l$ і $w \neq \theta$.

На тій же підставі встановлюється взаємно однозначна відповідність між крайніми точками множини V і крайніми променями конуса K.

Нехай x – крайня точка множини V і l – промінь конуса K, який відповідає цій точці, тобто такий, що $x = l \cap \Pi$. Покажемо, що l – крайній промінь конуса K. Оберемо спочатку на l точку x і припустимо, що $x = (1-\lambda)u + \lambda v$, де $0 < \lambda < 1$; $u, v \in K$. Якщо $u = \theta$, то зрозуміло, що $u, v \in l$. Те ж саме, якщо $v = \theta$.

Нехай $u \neq \theta$ і $v \neq \theta$. Тоді $y = \frac{1}{(a,u)} u$ і $z = \frac{1}{(a,v)} v$ – точки перетину відповідних променів з напрямними векторами u і v з гіперплощиною Π. Маємо рівності

$$x = (1-\lambda)(a,u)\frac{1}{(a,u)}u + \lambda(a,v)\frac{1}{(a,v)}v =$$
$$= (1-\lambda)(a,u)y + \lambda(a,v)z,$$

де $y, z \in V$.

При цьому

$$(1-\lambda)(a,u) + \lambda(a,v) = (a,(1-\lambda)u + \lambda v) = (a,x) = 1,$$

$$(1-\lambda)(a,u) > 0, \lambda(a,v) > 0 .$$

Оскільки x – крайня точка множини V, то $y = z = x$. Отже, $x = \frac{1}{(a,u)} u = \frac{1}{(a,v)} v$, а тому $u, v \in l$.

Нехай тепер $w \in l$, $w \neq x$ і $w = (1-\lambda)u + \lambda v$, $0 < \lambda < 1$, $u, v \in K \Rightarrow x = \frac{1}{(a,w)} w = (1-\lambda)\left[\frac{1}{(a,w)} u\right] + \lambda\left[\frac{1}{(a,w)} v\right]$,

де

$$u' = \frac{1}{(a,w)} u \in K, \quad v' = \frac{1}{(a,w)} v \in K .$$

Звідси робимо висновок, що $u', v' \in l$, а тому $u, v \in l$.

Припустимо, що l – крайній промінь конуса K і $x \in l \cap \Pi$. Якщо $x = (1-\lambda)u + \lambda v$, $u, v \in V$, $0 < \lambda < 1$, то $u, v \in l$; $u, v \in \Pi$. Тому $x = u = v$, тобто x – крайня точка множини V.

Для завершення доведення теореми достатньо показати, що множина $V = K \cap \Pi$ компактна, а потім зіслатися на теорему про крайні точки.

Нехай Q – перетин конуса K зі сферою одиничного радіуса з центром у нульовій точці θ: $Q = \left\{ x \in K : \|x\| = 1 \right\}$. Множина Q компактна в просторі R^n.

Побудуємо відображення f: $Q \to V$ за законом

$$f(x) = \frac{1}{(a,x)} x \quad \text{для} \quad x \in Q .$$

Це – взаємно однозначне відображення на всю множину V. Воно неперервне. Множина V, яка є неперервним образом компактної множини – компактна [14].

Як множину S, що породжує конус K і відповідає умовам теореми, можна обрати множину всіх крайніх точок множини

$$V = K \cap \Pi .$$

Теорема 5. *Ребра загостреного багатогранного конуса K є його крайніми променями. Інших крайніх променів конус K не має.*

Ця теорема є наслідком попередньої теореми і властивостей опуклих багатогранників, якщо врахувати, що багатогранний конус – замкнена множина. Однак, доведемо цю теорему, виходячи з інших міркувань.

Нехай $y_1, y_2, ..., y_p$ – кістяк багатогранного конуса K такий, що

$$K = L^+\left(y_1, y_2, ..., y_p\right), \tag{а}$$

$$y_k \overline{\in} L^+\left(y_1, ..., y_{k-1}, y_{k+1}, ..., y_p\right),\ \ k = \overline{1, p} \tag{б}$$

(див. § 9.2). Позначимо через l_k ребро конуса з напрямним вектором y_k. Доведемо, що кожне з них – крайній промінь. Наприклад, здійснимо це для ребра l_1. Оберемо на l_1 який-небудь вектор, зокрема, вектор y_1. Нехай $y_1 = (1-\lambda)x + \lambda y$, $0 < \lambda < 1$, $x, y \in K$. Якщо $x = \alpha_1 y_1 + \alpha_2 y_2 + ... + \alpha_p y_p$, $\alpha_k \geq 0$, $y = \beta_1 y_1 + \beta_2 y_2 + ... + \beta_p y_p$, $\beta_k \geq 0$, тоді $y_1 = \gamma_1 y_1 + \gamma_2 y_2 + ... + \gamma_p y_p$, де $\gamma_k = (1-\lambda)\alpha_k + \lambda\beta_k$, $k = \overline{1, p}$. Якщо $\gamma_1 < 1$, тоді дістанемо $y_1 = \dfrac{\gamma_2}{1-\gamma_1} y_2 + ... + \dfrac{\gamma_p}{1-\gamma_1} y_p$, що неможливо. Якщо ж $\gamma_1 > 1$, то $\left(\gamma_1 - 1\right) y_1 + \gamma_2 y_2 + ... + \gamma_p y_p = \theta$, що неможливо для загострених конусів. Отже, $\gamma_1 = 1$ і $\gamma_2 y_2 + ... + \gamma_p y_p = \theta$. Але тоді $\gamma_2 = \gamma_3 = ... = \gamma_p = 0$ на підставі леми про загострені конуси (див. § 9.2). Із рівності $\gamma_k = (1-\lambda)\alpha_k + \lambda\beta_k = 0$ і умови $0 < \lambda < 1$ випливає, що $\alpha_2 = \beta_2 = ... = \alpha_p = \beta_p = 0$. Тому $x = \alpha_1 y_1$, $y = \beta_1 y_1$ і $x, y \in l_1$. Отже, l_1 – крайній промінь конуса.

Доведемо обернене твердження. Для цього покажемо, що будь-який крайній промінь l конуса збігається з одним із ребер конуса.

Нехай $x \in l$, $x \neq \theta$ і $x = \lambda_1 y_1 + \lambda_2 y_2 + ... + \lambda_p y_p$, $\lambda_k \geq 0$. Припустимо, що в останній рівності $\lambda_1 \neq 0$. Тоді матимемо

$$x = \frac{1}{2}\left(2\lambda_1 y_1\right) + \frac{1}{2}\left(2\lambda_2 y_2 + ... + 2\lambda_p y_p\right).$$

Точка x – опукла лінійна комбінація точок із конуса K. Оскільки l – крайній промінь цього конуса, то $2\lambda_1 y_1 \in l$, а тому $y_1 \in l$. Це означає, що $l = l_1$.

Теорема 6. *Замкнений конус K може бути поданий у вигляді прямої суми*

$$K = M \oplus V$$

лінійного підпростору M і замкненого загостреного конуса V. Доданок M визначається конусом K однозначно і включає всі лінійні підпростори, які містяться в K.

Теорема безпосередньо випливає з теореми 1, доведеної в § 9.2. Доданок V може бути одержаним у вигляді перетину $V = K \cap L$, де L – будь-який лінійний підпростір, для якого $R^n = M \oplus L$. Зокрема, можна покласти $L = M^{\perp}$. Тому V – замкнена множина.

Розділ 10. СИСТЕМИ ЛІНІЙНИХ ДІОФАНТОВИХ РІВНЯНЬ*

§ 10.1. Умови сумісності. Алгоритм Ерміта

Системою лінійних діофантових рівнянь називається система лінійних рівнянь з цілочисловими коефіцієнтами і вільними членами. Як розв'язок цієї системи розуміють цілочисловий.

Теорема 1. *Лінійне діофантове рівняння*

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = \beta, \tag{1}$$

де коефіцієнти $\alpha_1, \alpha_2, \ldots, \alpha_n$ *одночасно не обертаються на нуль, має розв'язок тоді і лише тоді, коли НСД* $(\alpha_1, \alpha_2, \ldots, \alpha_n) \mid \beta$.

Доведення. *Необхідність.* Нехай $\delta = \text{НСД}(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Тоді $\delta \mid \alpha_1 x_1, \delta \mid \alpha_2 x_2, \ldots, \delta \mid \alpha_n x_n$ для будь-яких $x_j \in Z$, $j = \overline{1, n}$, де Z – кільце всіх цілих чисел. Отже, $\delta \mid \sum\limits_{j=1}^{n} \alpha_j x_j$. Якщо $\left(x_1^*, x_2^*, \ldots, x_n^*\right)$ – розв'язок рівняння, то $\sum\limits_{j=1}^{n} \alpha_j x_j^* = \beta$, $\delta \mid \sum\limits_{j=1}^{n} \alpha_j x_j^*$ і, отже, $\delta \mid \beta$.

Достатність. Припустимо, що $\delta \mid \beta$. Покладемо

$$J = \left\{ x \in Z : x = \sum_{j=1}^{n} \alpha_j x_j ;\ x_j \in Z,\ j = \overline{1, n} \right\}.$$

Ця множина має такі властивості:

1. $x, y \in J \Rightarrow x \pm y \in J$.
2. $x \in J,\ z \in Z \Rightarrow x \cdot z \in J$.

Будь-яка підмножина з кільця Z з цими двома властивостями називається *ідеалом*. У нашому випадку множина J є

---

* Цей розділ написаний І. В. Протасовим – проф. Київського національного університету імені Тараса Шевченка.

ідеал, породжений числами $\alpha_1, \alpha_2, ..., \alpha_n$, і позначається через $J = (\alpha_1, \alpha_2, ..., \alpha_n)$.

Кожний ідеал кільця Z є головним, оскільки він породжений одним цілим числом. Тому $J = (\alpha_1, \alpha_2, ..., \alpha_n) = (\delta)$ і $\delta =$ НСД $(\alpha_1, \alpha_2, ..., \alpha_n)$ (§ 1.4, теорема 3).

Існують такі цілі числа $y_1, y_2, ..., y_n$, що

$$\sum_{j=1}^{n} \alpha_j y_j = \delta .$$

За умовою $\delta \mid \beta$ і, отже, $\beta = \delta \cdot m$, де $m \in Z$. Якщо $x_1^* = my_1$, $x_2^* = my_2$, ..., $x_n^* = my_n$, тоді $\sum_{j=1}^{n} \alpha_j x_j^* = \beta$, що й доводить розв'язність рівняння (1).

Теорему доведено.

Тепер зупинимось на знаходженні розв'язків систем діофантових рівнянь. Якщо така система сумісна, то для знаходження її розв'язків достатньо залишити в цій системі лише базисні рівняння. У новій системі число рівнянь дорівнюватиме рангу матриці цієї системи (див. § 4.1). Тому вважатимемо, що ранг матриці наведеної нижче системи дорівнює числу її рівнянь.

Теорема 2 (Штейніца). *Система лінійних діофантових рівнянь*

$$\begin{cases} \alpha_{11}x_1 + \alpha_{12}x_2 + ... + \alpha_{1n}x_n = \beta_1, \\ \alpha_{21}x_1 + \alpha_{22}x_2 + ... + \alpha_{2n}x_n = \beta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + ... + \alpha_{mn}x_n = \beta_m, \end{cases}$$

де ранг матриці $A = (\alpha_{ik})$ цієї системи дорівнює числу m рівнянь, розв'язна тоді й лише тоді, коли НСД мінорів матриці A m-го порядку збігається з НСД мінорів m-го порядку розширеної матриці цієї системи $(A \mid b)$.

Перш ніж доводити теорему, наведемо декілька необхідних тверджень. Зокрема, будемо називати елементарними перетвореннями стовпців (рядків) цілочислової матриці такі:

І. Множення всіх елементів стовпців на -1.

II. Додавання до стовпця (рядка) іншого, помноженого на ціле число.

III. Перестановка стовпців (рядків).

Лема 1. *Кожне елементарне перетворення* τ *цілочислової матриці* $A = (\alpha_{ik})$ *розмірності* $m \times n$ *рівносильне множенню цієї матриці з правого боку на матрицю* $T = (\gamma_{ik})$ *розмірності* $n \times n$, *яка одержується з одиничної матриці порядку* n *тим самим елементарним перетворенням* τ.

Доведення. Доведемо лему для елементарних перетворень II. Нехай елементарне перетворення τ полягає у додаванні до k-го стовпця матриці i-го стовпця, помноженого на ціле число γ. Застосуємо елементарне перетворення τ відповідно до матриці A та одиничної матриці порядку n. Дістанемо матриці

$$B = \begin{pmatrix} \alpha_{11}\alpha_{12} & \dots \alpha_{1i} & \dots \alpha_{1k} + \gamma\alpha_{1i} & \dots \alpha_{1n} \\ \alpha_{21}\alpha_{22} & \dots \alpha_{2i} & \dots \alpha_{2k} + \gamma\alpha_{2i} & \dots \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{i1}\alpha_{i2} & \dots \alpha_{ii} & \dots \alpha_{ik} + \gamma\alpha_{ii} & \dots \alpha_{in} \\ \dots & \dots & \dots & \dots \\ \alpha_{m1}\alpha_{m2} & \dots \alpha_{mi} & \dots \alpha_{mk} + \gamma\alpha_{mi} & \dots \alpha_{mn} \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i, \\ \\ \\ \end{matrix}$$

$$T = \begin{pmatrix} 1\ 0 \dots 0 \dots 0 \dots 0 \\ 0\ 1 \dots 0 \dots 0 \dots 0 \\ \dots\dots\dots\dots\dots\dots \\ 0\ 0 \dots 1 \dots \gamma \dots 0 \\ \dots\dots\dots\dots\dots\dots \\ 0\ 0 \dots 0 \dots 0 \dots 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i. \\ \\ \\ \end{matrix}$$

З іншого боку,

$$AT = \begin{pmatrix} \alpha_{11} \dots \alpha_{1i} \dots \alpha_{1k} \dots \alpha_{1n} \\ \alpha_{21} \dots \alpha_{2i} \dots \alpha_{2k} \dots \alpha_{2n} \\ \dots\dots\dots\dots\dots\dots \\ \alpha_{i1} \dots \alpha_{ii} \dots \alpha_{ik} \dots \alpha_{in} \\ \dots\dots\dots\dots\dots\dots \\ \alpha_{m1} \dots \alpha_{mi} \dots \alpha_{mk} \dots \alpha_{mn} \end{pmatrix} \begin{pmatrix} 1 \dots 0 \dots 0 \dots 0 \\ 0 \dots 0 \dots 0 \dots 0 \\ \dots\dots\dots\dots \\ 0 \dots 1 \dots \gamma \dots 0 \\ \dots\dots\dots\dots \\ 0 \dots 0 \dots 0 \dots 1 \end{pmatrix} =$$

$$=\begin{pmatrix} \alpha_{11} \dots \alpha_{1l} \dots \alpha_{1k} + \gamma\alpha_{1l} \dots \alpha_{1n} \\ \alpha_{21} \dots \alpha_{2l} \dots \alpha_{2k} + \gamma\alpha_{2l} \dots \alpha_{2n} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{i1} \dots \alpha_{il} \dots \alpha_{ik} + \gamma\alpha_{il} \dots \alpha_{in} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_{m1} \dots \alpha_{ml} \dots \alpha_{mk} + \gamma\alpha_{ml} \dots \alpha_{mn} \end{pmatrix}.$$

Отже, $AT = B$ і лему доведено.

Наведемо тепер алгоритм Ерміта, який дає змогу перетворити довільну цілочислову матрицю $A = (\alpha_{ik})$ розмірності $m \times n$ і рангу m за допомогою деякої послідовності елементарних перетворень до матриці виду

$$H = \begin{pmatrix} \overline{\alpha}_{11} & 0 & 0 & \dots & 0 & \dots & 0 \\ \overline{\alpha}_{21} & \overline{\alpha}_{22} & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \overline{\alpha}_{m1} & \overline{\alpha}_{m2} & \overline{\alpha}_{m3} & \dots & \overline{\alpha}_{mm} & \dots & 0 \end{pmatrix},$$

де $\overline{\alpha}_{ii} > 0,\ i = \overline{1, m}$.

Матриця H називається *ермітовою формою* матриці A.

Оскільки ранг матриці A дорівнює числу рядків, то серед елементів $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1n}$ існують відмінні від нуля. Перестановкою стовпців досягаємо того, щоб елемент α_{11} був відмінним від нуля і найменшим за модулем у першому рядку. Можна вважати, що $\alpha_{11} > 0$ (при необхідності множимо перший стовпець на -1). Далі кожний елемент α_{1k} подаємо у вигляді $\alpha_{1k} = \alpha_{11} h_k + r_k$, де $h_k, r_k \in Z$, $0 \le r_k < \alpha_{11}$. Віднімемо від k-го $(k \ge 2)$ стовпця елементи першого стовпця, помножені на число h_k. У результаті всіх таких операцій елементи першого рядка будуть цілими числами від нуля до α_{11}.

Повторивши цю процедуру скінченне число разів, матимемо матрицю

$$A_1 = \begin{pmatrix} \overline{\alpha}_{11} & 0 & 0 & \dots & 0 \\ \overline{\alpha}_{21} & b_{22} & b_{23} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \overline{\alpha}_{m1} & b_{m2} & b_{m3} & \dots & b_{mn} \end{pmatrix},$$

де $\overline{\alpha}_{11} > 0$.

Оскільки елементарні перетворення не змінюють ранг матриці, то серед елементів $b_{22}, b_{23}, \dots, b_{2n}$ будуть такі, що відмінні від нуля.

Застосовуємо аналогічні обчислення до матриці

$$\begin{pmatrix} b_{22} & b_{23} & \dots & b_{2n} \\ b_{32} & b_{33} & \dots & b_{3n} \\ \dots & \dots & \dots & \dots \\ b_{m2} & b_{m3} & \dots & b_{mn} \end{pmatrix}.$$

У результаті побудовано матрицю H потрібного вигляду.

Приклад 1. Знайти матрицю H для матриці

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

Розв'язання.

$$A = \begin{pmatrix} 4 & 1 & 2 & 3 \\ 0 & -1 & -1 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & 2 & 3 \\ -1 & 0 & -1 & 5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 4 & 1 & 8 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 4 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix} = H .$$

Квадратна цілочислова матриця V називається *унімодулярною*, якщо її визначник $\det V = \pm 1$.

Лема 2. *Нехай $A = (\alpha_{ik})$ – цілочислова матриця розмірності $m \times n$ і рангу m. Для неї існує така унімодулярна матриця $V = (v_{ik})$ порядку n, що матриця $A \cdot V$ матиме вигляд*

$$\begin{pmatrix} \overline{\alpha}_{11} & 0 & 0 & \dots & 0 & \dots & 0 \\ \overline{\alpha}_{21} & \overline{\alpha}_{22} & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \overline{\alpha}_{m1} & \overline{\alpha}_{m2} & \overline{\alpha}_{m3} & \dots & \overline{\alpha}_{mm} & \dots & 0 \end{pmatrix},$$

де $\overline{\alpha}_{ii} > 0,\ i = \overline{1,m}$.

Доведення. За допомогою деякої послідовності елементарних перетворень $\tau_1, \tau_2, ..., \tau_s$ (алгоритм Ерміта) матриця A набуває ермітової форми H на підставі леми 1. Кожне елементарне перетворення τ_i можна замінити множенням з правого боку на матрицю T_i, яка одержується з одиничної матриці порядку E таким самим елементарним перетворенням τ_i над відповідними стовпцями.

Тому $H = AT_1T_2...T_s$. Покладемо $V = T_1T_2...T_s$. Оскільки $\det T_i = \pm 1$, то $\det V = \pm 1$.

Лему доведено.

Доведення теореми Штейніца. Систему лінійних діофантових рівнянь запишемо у матричному вигляді

$$A \cdot X = b,\ \ X = (x_1, x_2, ..., x_n)^T, \qquad (2)$$

де b – стовпець чисел, складений із правих частин системи.

Покажемо, що ця система рівнянь розв'язна тоді й лише тоді, коли розв'язна система лінійних діофантових рівнянь

$$H \cdot Y = b,\ \ Y = (y_1, y_2, ..., y_n)^T, \qquad (3)$$

де H – ермітова форма матриці A.

На підставі леми 2 існує така унімодулярна матриця V, що $H = A \cdot V$. Нехай X^* – розв'язок системи (2). Тоді $A\left(VV^{-1}\right)X^* = b$ і тому $H\left(V^{-1}X^*\right) = b$. Оскільки V – матриця цілочислова й унімодулярна, то матриця V^{-1} також цілочислова. Отже, вектор $Y^* = V^{-1}X^*$ – цілочисловий розв'язок системи (3).

Нехай Y^* – розв'язок системи (3). Тоді $A\left(VY^*\right)=b$ і $X^*=VY^*$ – розв'язок системи (2).

На підставі елементарних перетворень легко можна довести, що ці перетворення не змінюють НСД мінорів будь-якого порядку. Тому теорему достатньо довести для системи $HY=b$. Отже, слід довести розв'язність такої системи діофантових рівнянь:

$$\begin{cases} \overline{\alpha}_{11}y_1 & = \beta_1, \\ \overline{\alpha}_{21}y_1+\overline{\alpha}_{22}y_2 & = \beta_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \overline{\alpha}_{m1}y_1+\overline{\alpha}_{m2}y_2+\ldots+\overline{\alpha}_{mm}y_m & = \beta_m. \end{cases} \quad (4)$$

Скористаємось формулами Крамера

$$y_i=\frac{\Delta_i}{\Delta}\ (i=\overline{1,m}),$$

де Δ – визначник системи, а Δ_i – визначники, в яких i-й стовпець замінено стовпцем $\left(\beta_1,\beta_2,\ldots,\beta_m\right)^T$.

Отже, система лінійних діофантових рівнянь розв'язна тоді й лише тоді, коли $\frac{\Delta_i}{\Delta}$ $(i=\overline{1,m})$ – цілі числа, тобто коли $\Delta\mid\Delta_1,\Delta\mid\Delta_2,\ldots,\Delta\mid\Delta_m$.

Необхідно зауважити, що Δ – НСД мінорів m-го порядку матриці H, а $\Delta,\Delta_1,\Delta_2,\ldots,\Delta_m$ – мінори m-го порядку розширеної матриці $(H\mid b)$, причому решта мінорів m-го порядку цієї матриці дорівнює нулю.

Теорему доведено.

Загальний розв'язок системи лінійних діофантових рівнянь

Розглянемо систему лінійних діофантових рівнянь (2), знайдемо ермітову форму H матриці A і перейдемо до системи (4). З першого рівняння системи знайдемо $y_1^*=\frac{\beta_1}{\overline{\alpha}_{11}}$, з другого – y_2^* і т. д. Якщо серед y_j^* $(j=\overline{1,m})$ знайдеться хоч би одне не ціле число, то система (3), а з нею і (2), не має необхідних розв'язків. Якщо всі ці числа цілі, то послідовність $(y_1^*, y_2^*, \ldots, y_m^*, y_{m+1}, \ldots, y_n)$, де $y_{m+1}, \ldots, y_n$ – довільні цілі числа, буде загальним розв'язком системи (3). Враховуючи, що співвідношення $X = V \cdot Y$ визначає загальний розв'язок початкової системи рівнянь, загальний розв'язок $X = (x_1, x_2, \ldots, x_n)^T$ можна подати у вигляді

$$x_i = \sum_{k=1}^{m} v_{ik} y_k^* + \sum_{k=m+1}^{n} v_{ik} y_k, \quad i = \overline{1,n},$$

де v_{ik} – елементи матриці V; $y_{m+1}, \ldots, y_n$ – довільні цілі числа.

Приклад 2. Знайти загальний розв'язок системи діофантових рівнянь

$$4x_1 - 4x_2 + 3x_3 = 3,$$
$$5x_1 - 7x_2 + 6x_3 = 4.$$

Розв'язання. Із леми 2 випливає, що послідовність елементарних перетворень стовпців, яка зводить матрицю A до ермітової H, одночасно зводить одиничну матрицю E (тут – порядку 3) в матрицю V. Тому обчислювати матриці H і V доцільно одночасно, застосовуючи серію елементарних перетворень стовпців розширеної матриці

$$\left(\frac{A}{E}\right) = \left(\begin{array}{ccc} 4 & -4 & 3 \\ 5 & -7 & 6 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right).$$

Матимемо

$$\left(\frac{A}{E}\right) \sim \left(\begin{array}{ccc} 3 & -4 & 4 \\ 6 & -7 & 5 \\ \hline 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc} 3 & -1 & 1 \\ 6 & -1 & -1 \\ \hline 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{array}\right) \sim \left(\begin{array}{ccc} 1 & -1 & 3 \\ -1 & -1 & 6 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{array}\right) \sim$$

$$\sim\begin{pmatrix}1 & 0 & 0\\ -1 & -2 & 9\\ \hline 1 & 1 & -3\\ 0 & 1 & 0\\ -1 & 0 & 4\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 0\\ -1 & -2 & 1\\ \hline 1 & 1 & 1\\ 0 & 1 & 4\\ -1 & 0 & 4\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ \hline 1 & 1 & 3\\ 0 & 4 & 9\\ -1 & 4 & 8\end{pmatrix}.$$

Отже,

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

Розглянемо систему $HY=b$:

$\begin{cases}y_1+0\cdot y_2+0\cdot y_3=3,\\ -y_1+y_2+0\cdot y_3=4\end{cases}\Rightarrow y_1^*=3,\ \ y_2^*=7,\ \ y_3-$ довільне ціле число.

Загальний розв'язок системи:

$$\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=V\begin{pmatrix}y_1^*\\ y_2^*\\ y_3\end{pmatrix}=\begin{pmatrix}1 & 1 & 3\\ 0 & 4 & 9\\ -1 & 4 & 8\end{pmatrix}\begin{pmatrix}3\\ 7\\ y_3\end{pmatrix}=\begin{pmatrix}10+3y_3\\ 28+9y_3\\ 25+8y_3\end{pmatrix}.$$

§ 10.2. Системи лінійних однорідних діофантових рівнянь

Розглянемо систему лінійних однорідних діофантових рівнянь

$$A\cdot X=\theta,\ \ X=\left(x_1,x_2,\ldots,x_n\right)^T,$$

де $A=\left(\alpha_{ik}\right)$ – цілочислова матриця розмірності $m\times n$ і рангу m; θ – нульовий вектор.

Така система розв'язна, оскільки її задовольняє цілочисловий вектор $\theta=(0,0,\ldots,0)$. Якщо $m=n$, тоді на підставі теореми Крамера ця система не матиме нетривіальні розв'язки.

Якщо $m < n$, то, як випливає із попереднього параграфа, система матиме нетривіальні розв'язки.

Теорема Зігеля. *Якщо* $m < n$ *і* $|\alpha_{ik}| \le C$, $C = \text{const}$, *тоді однорідна система лінійних діофантових рівнянь має нетривіальні цілочислові розв'язки* $x_1, x_2, \ldots, x_n$ *такі, що*

$$x_i \le 2(nC)^{\frac{m}{n-m}}, \quad i = \overline{1,n}.$$

Доведення. Визначимо лінійне перетворення $A: R^n \to R^m$ дійсного векторного простору R^n у векторний простір R^m (R – поле дійсних чисел) за допомогою матриці A даної системи:

$$y = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} x,$$

де $x = (\xi_1, \xi_2, \ldots, \xi_n)^T \in R^n$; $y = (\eta_1, \eta_2, \ldots, \eta_m)^T \in R^m$.

Позначимо Z^n і Z^m – множини всіх точок з цілочисловими координатами відповідно у просторах R^n і R^m. Далі позначимо для додатних цілих чисел s і t множини

$$N(s) = \left\{ x \in Z^n : |\xi_i| \le s,\ i = \overline{1,n} \right\},$$

$$M(t) = \left\{ y \in Z^m : |\eta_k| \le t,\ k = \overline{1,m} \right\}.$$

Зауважимо, якщо $x \in N(s)$, то $y = Ax \in M(nCs)$. Число елементів множини $N(s)$ дорівнює $(2s+1)^n$, а число елементів множини $M(nCs)$ дорівнює $(2nCs+1)^m$.

З'ясуємо, за яких умов виконуватиметься нерівність

$$(2s+1)^n > (2nCs+1)^m.$$

Зрозуміло, що

$$(2nCs+1)^m = \left(2s + \frac{1}{nC}\right)^m (nC)^m < (2s+1)^m (nC)^m.$$

Оскільки $(2s+1)^n = (2s+1)^m (2s+1)^{n-m}$, то для потрібної нерівності ми повинні мати

$$(2s+1)^{n-m} \geq (nC)^m .$$

Остання нерівність виконується, коли

$$s = \left[(nC)^{\frac{m}{n-m}} \right].$$

Отже, при такому виборі s знайдуться такі $x, x' \in N\left(\left[(nC)^{\frac{m}{n-m}}\right]\right)$, що $x \neq x'$ і $Ax = Ax'$. Тоді вектор $x - x'$ – шуканий цілочисловий розв'язок системи рівнянь.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Алберт А.* Регрессия, псевдоинверсия и рекуррентное оценивание. – М.: Наука, 1977. – 223 с.
2. *Архангельский А. В.* Конечномерные векторные пространства. – М.: Изд-во Моск. ун-та, 1982. – 350 с.
3. *Беклемишев Д. В.* Дополнительные главы алгебры. – М.: Наука, 1983. – 335 с.
4. *Ван дер Варден.* Алгебра. – М.: Наука, 1979. – 623 с.
5. *Воеводин В. В.* Вычислительные основы линейной алгебры. – М.: Наука, 1977. – 303 с.
6. *Воеводин В. В.* Линейная алгебра. – М.: Наука, 1980. – 400 с.
7. *Гантмахер Ф. Р.* Теория матриц. – М.: Наука, 1988. – 547 с.
8. *Гельфанд И. М.* Лекции по линейной алгебре. – М.: Наука, 1971. – 280 с.
9. *Граве Д. А.* Элементы высшей алгебры. – К., 1914. – 698 с.
10. *Завало С. Т.* Курс алгебри. – К.: Вища шк., 1985. – 500 с.
11. *Икрамов Х. Д.* Задачник по линейной алгебре. – М.: Наука, 1975. – 319 с.
12. *Ильин В. А., Позняк Э. Г.* Линейная алгебра. – М.: Наука, 1984. – 294 с.
13. *Карманов В. Г.* Математическое программирование. – М.: Наука, 1980. – 253 с.
14. *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. – М.: Наука, 1968. – 234 с.
15. *Курош А. Г.* Курс высшей алгебры. – М.: Наука, 1968. – 431 с.
16. *Ланкастер П.* Теория матриц. – М.: Наука, 1978. – 280 с.
17. *Мальцев А. И.* Алгебраические системы. – М.: Наука, 1970. – 392 с.
18. *Маркушевич А. И.* Возвратные последовательности. – М.; Л., 1951. – 47 с.
19. *Проскуряков И. В.* Сборник задач по линейной алгебре. – М.: Наука, 1974. – 384 с.
20. *Фаддеев Д. К., Соминский И. С.* Сборник задач по высшей алгебре. – М.: Наука, 1977. – 307 с.
21. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры. – М.: Физматгиз, 1963. – 734 с.
22. *Чарин В. С.* Линейные преобразования и выпуклые множества. – К.: Вища шк., 1978. – 192 с.
23. *Черников С. Н.* Линейные неравенства. – М.: Наука, 1968. – 488 с.

ЗМІСТ

Навчальне видання

Чарін Віктор Сильвестрович

ЛІНІЙНА АЛГЕБРА

2-ге видання, стереотипне

Затверджено Міністерством освіти
і науки України як підручник для студентів
вищих технічних навчальних закладів

Видано за рахунок державних коштів.
Продаж заборонено

Редактор *Л. Ю. Степанчук*
Оформлення художника *В. О. Гурлєва*
Художній редактор *С. В. Анненков*
Комп'ютерна верстка *Т. В. Антіпової, Л. О. Ємець*
Коректори *І. В. Іванюсь, О. О. Куценко, Н. О. Сазонова*

Підписано до друку 04.08.2005. Формат 84×108 $^{1}/_{32}$. Папір офсетний. Друк офсетний. Гарнітура Times New Roman.
Умов. друк. арк. 21,84. Обл.-вид. арк. 20,96. Тираж 5500 прим.

Зам. № 5356

Видавництво "Техніка". 04053 Київ, вул. Обсерваторна, 25.
Свідоцтво про внесення до Державного реєстру України суб'єктів видавничої справи ДК № 357 від 12.03.2001 р.

Віддруковано на Білоцерківській книжковій фабриці.
09117 Біла Церква, вул. Леся Курбаса, 4.
Свідоцтво про внесення до Державного реєстру України суб'єктів видавничої справи ДК № 567 від 14.08.2001 р.

Чарін В. С.
Ч20 Лінійна алгебра. – 2-ге вид., стер. – К.: Техніка, 2005. – 416 с.
ISBN 996-575-090-9

У підручнику викладено об'єкти лінійної алгебри трьох типів: лінійні простори, матриці й алгебричні форми. Вони важливі для розв'язування задач геометрії, обчислювальної математики, фізики, математичної економіки. Розглянуто властивості лінійних рівнянь і лінійних перетворень лінійних просторів. Приділено увагу питанням псевдообернення дійсних матриць, що пов'язані з методом найменших квадратів Гаусса. Наведено ефективні методи обчислення псевдооберненої матриці. При розгляданні опуклих множин використано топологічні властивості підмножин евклідового простору. Досліджено діофантові рівняння.

Для студентів вищих технічних навчальних закладів.

ББК 22.143я73