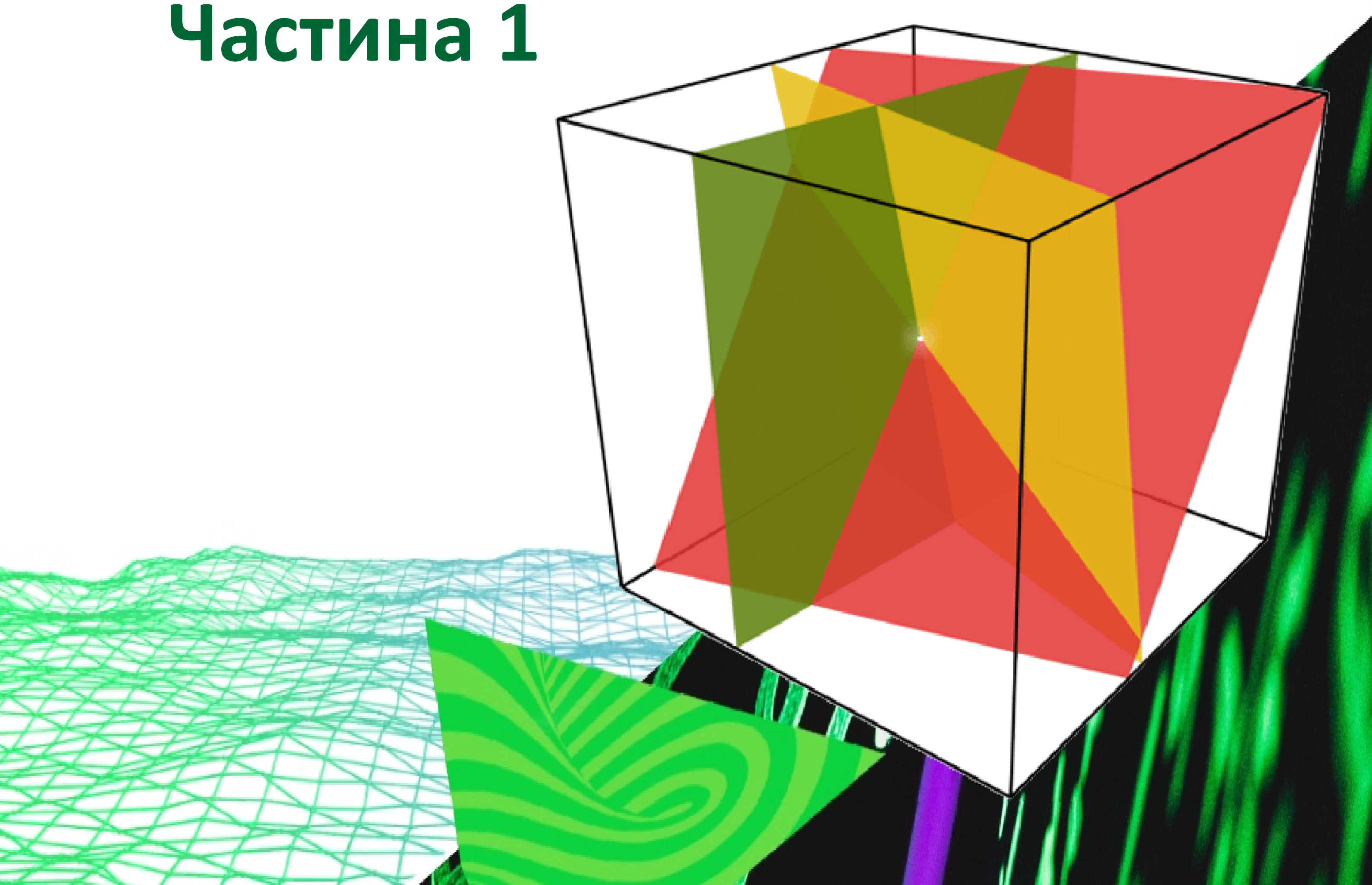
Пащенко З.Д
Турка Т.В.

Лінійна алгебра і аналітична геометрія

Частина 1



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДВНЗ «Донбаський державний педагогічний університет»
кафедра методики навчання математики та методики навчання інформатики

Лінійна алгебра і аналітична геометрія

Частина 1

навчальний посібник
для студентів спеціальності 014 Середня освіта (Інформатика)

Слов'янськ – 2020

УДК 512.64+514.12(075.8)
П22

Затверджено на засіданні Вченої ради
Протокол № 8 від 30.06.2020 р.

Рецензенти:

Загребельний С.Л., кандидат педагогічних наук, доцент, завідувач кафедри інформатики і інженерної графіки Донбаської державної машинобудівної академії

Кадубовський О.А., кандидат фізико-математичних наук, доцент кафедри математики та інформатики ДВНЗ «ДДПУ»

Пащенко З.Д., Турка Т.В.
Лінійна алгебра та аналітична геометрія. Частина 1: навчальний посібник для студентів спеціальності 014 Середня освіта (Інформатика). Слов'янськ: ДВНЗ «ДДПУ», 2020. 169 с.

В посібнику в стислій, доступній формі викладено зміст лекцій першого семестру курсу «Лінійна алгебра та аналітична геометрія», що відповідає освітньо-професійній програмі Середня освіта (Інформатика) фізико-математичного факультету ДВНЗ «Донбаський державний педагогічний університет». Охоплено теми: комплексні числа, матриці, системи лінійних рівнянь, лінійна залежність та лінійна незалежність векторів, лінійні простори, евклідові простори, лінійні оператори, білінійні та квадратичні форми. Наводяться приклади понять та розв'язування задач. Для ґрунтовнішого засвоєння матеріалу в кінці кожного пункту розміщено контрольні питання та завдання.

Рекомендовано для студентів математичних спеціальностей вищих педагогічних навчальних закладів.

ЗМІСТ

Розділ 1. КОМПЛЕКСНІ ЧИСЛА ... 4
1.1. Алгебраїчна форма комплексних чисел ... 4
1.2. Геометричне зображення комплексних чисел. ... 6
1.3. Видобування кореня ... 12
Розділ 2. СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ ... 16
2.1. Матриці. Операції над матрицями ... 16
2.2. Елементарні перетворення матриць ... 22
2.3. Визначники. Властивості визначників ... 25
2.4. Мінори. Визначник добутку ... 34
2.5. Обернена матриця ... 40
2.6. Метод Гаусса ... 43
2.7. Матричний метод розв'язування систем лінійних рівнянь ... 54
2.8. Метод Крамера ... 57
2.9. Вектори. Лінійна залежність та лінійна незалежність векторів ... 59
2.10. Ранг матриці. Базисний мінор ... 64
2.11. Критерії сумісності та визначеності ... 69
2.12. Однорідні системи рівнянь. ФСР ... 71
Розділ 3. ЛІНІЙНІ ПРОСТОРИ ... 83
3.1. Поняття поля ... 83
3.2. Поняття лінійного простору ... 85
3.3. Базис лінійного простору. Однозначність розкладення ... 90
3.4. Розмірність лінійного простору ... 95
3.5. Перетворення базисів ... 97
3.6. Підпростір лінійного простору ... 99
3.7. Перетин і сума підпросторів. Пряма сума підпросторів ... 102
3.8. Евклідовий простір ... 107
3.9. Ортогональні системи ... 111
3.10. Унітарний простір ... 117
Розділ 4. ЛІНІЙНІ ОПЕРАТОРИ ... 120
4.1. Поняття лінійного оператору ... 120
4.2. Простір лінійних операторів ... 121
4.3. Матриці лінійних перетворень ... 124
4.4. Перетворення матриці лінійного оператору при зміні базису ... 128
4.5. Оборотні оператори (перетворення) ... 130
4.6. Образ і ядро ... 132
4.7. Власні значення та власні вектори ... 137
Розділ 5. БІЛІНІЙНІ ТА КВАДРАТИЧНІ ФОРМИ ... 145
5.1. Білінійні форми ... 145
5.2. Симетричні білінійні та квадратичні форми ... 148
5.3. Канонічний вид квадратичної форми ... 150
5.4.Метод Лагранжа ... 152
5.5. Метод Якобі ... 157
5.6. Класифікація квадратичних форм ... 161
СПИСОК ПОЗНАЧЕНЬ ... 166
СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ ... 168

Розділ 1. КОМПЛЕКСНІ ЧИСЛА

1.1. Алгебраїчна форма комплексних чисел

Комплексним числом називається вираз виду $a+bi$, де a і b – дійсні числа, i – ***уявна одиниця***, а саме $i^2=-1$. Числа a і b називаються ***компонентами*** комплексного числа $z=a+bi$.

$\mathbf{C}=\{a+bi \mid a,b\in\mathbf{R}\}$ – множина комплексних чисел.

Будемо вважати, що $a+0i=a\in\mathbf{R}$, тому $\mathbf{R}\subset\mathbf{C}$.

Введемо функції Re та Im, що діють із $\mathbf{C}$ в $\mathbf{R}$: якщо $z=a+bi$, то $\mathrm{Re}\, z=a$ (дійсна частина z), $\mathrm{Im}\, z=b$ (уявна частина z). Число $\overline{z}=a-bi$ називається ***спряженим*** до z. Два комплексні числа ***рівні***, якщо рівні їх дійсні та уявні частини.

Введемо арифметичні дії над комплексними числами.

Нехай $z_1=a+bi$, $z_2=c+di$.

Додавання на множині комплексних чисел $\mathbf{C}$ задається за правилом:

$$\mathbf{z_1+z_2=(a+c)+(b+d)i}.$$

Оскільки різниця двох чисел x і y – це таке число r, що $r+y=x$, то природно, що **різниця** $\mathbf{z_1-z_2=(a-c)+(b-d)i}$, бо $(a-c)+(b-d)i+z_2=z_1$.

Множення на множині комплексних чисел задається за правилом:

$$\mathbf{z_1\cdot z_2=(ac-bd)+(ad+bc)i}.$$

Це правило можна розглядати як добуток двох сум, враховуючи, що $i^2=-1$: $z_1\cdot z_2=(a+bi)(c+di)=ac+adi+bci+bdi^2==(ac-bd)+(ad+bc)i$

Розглянемо дію **ділення**.

Якщо $z_2=m\in\mathbf{R}$, то $z_1:z_2=\dfrac{z_1}{m}=\dfrac{a}{m}+\dfrac{b}{m}i$, тобто ввести поняття ділення комплексного числа на відмінне від нуля дійсне не складає труднощів. Але якщо $z_2\notin\mathbf{R}$, то розглянемо $z_2\cdot\overline{z_2}$, де $\overline{z_2}=c-di$:

$z_2\cdot\overline{z_2}=c^2+d^2+(-cd+dc)i=c^2+d^2\in\mathbf{R}$, тоді

$$\frac{z_1}{z_2}=\frac{z_1\cdot\overline{z_2}}{z_2\cdot\overline{z_2}}=\frac{(ac+bd)+(-ad+bc)i}{c^2+d^2}=\frac{ac+bd}{c^2+d^2}+\frac{-ad+bc}{c^2+d^2}i,$$

і правилом ділення на довільне ненульове комплексне число можна вважати: $\mathbf{\dfrac{z_1}{z_2}=\dfrac{z_1\cdot\overline{z_2}}{z_2\cdot\overline{z_2}}}$ або $\mathbf{\dfrac{z_1}{z_2}=\dfrac{ac+bd}{c^2+d^2}+\dfrac{-ad+bc}{c^2+d^2}i}$.

Ми розглянули ***алгебраїчний вид*** комплексних чисел: $a+bi$, $a,b\in\mathbf{R}$.

Приклад. Дано комплексні числа z_1 і z_2. Знайти z_1+z_2, $z_1\cdot z_2$, z_1-z_2, $z_1:z_2$ та їх дійсні і уявні частини, якщо $z_1=3-2i$, $z_2=-5+3i$.

▼ Виконаємо обчислення за вказаними попереду правилами

1) $z_1+z_2=(3-5)+(-2+3)i=-2+i$, $\operatorname{Re}(z_1+z_2)=-2$, $\operatorname{Im}(z_1+z_2)=1$;

2) $z_1\cdot z_2=3\cdot(-5)+3\cdot 3i+(-2)\cdot(-5)i+(-2)\cdot 3i^2=(-15+6)+(9+10)i=-9+19i$,
$\operatorname{Re}(z_1\cdot z_2)=-9$, $\operatorname{Im}(z_1\cdot z_2)=19$;

3) $z_1-z_2=(3+5)+(-2-3)i=8-5i$, $\operatorname{Re}(z_1-z_2)=8$, $\operatorname{Im}(z_1-z_2)=-5$;

4) $z_1:z_2=\dfrac{z_1\cdot\bar z_2}{z_2\cdot\bar z_2}=\dfrac{(3-2i)(-5-3i)}{(-5+3i)(-5-3i)}=\dfrac{(-15-6)+(-9+10)i}{25+9}=-{}^{21}\!/_{34}+{}^{1}\!/_{34}\,i$,

$\operatorname{Re}(z_1:z_2)=-{}^{21}\!/_{34}$, $\operatorname{Im}(z_1:z_2)={}^{1}\!/_{34}$. ▲

Властивості операцій додавання та множення

1. *Додавання комплексних чисел комутативне.*

▼ Нехай $z_1=a+bi$, $z_2=c+di$. Розглянемо $z_2+z_1=(c+a)+(d+b)i$. Так як додавання в множині дійсних чисел комутативне, то

$$c+a=a+c,\ b+d=d+b,\ \Rightarrow\quad z_1+z_2=z_2+z_1,$$

що доводить комутативність додавання в множині комплексних чисел. ▲

2. *Множення комплексних чисел комутативне.*

▼ Аналогічно $z_1\cdot z_2=z_2\cdot z_1$, бо $z_2\cdot z_1=(ca-db)+i(da+cb)=z_1\cdot z_2$. ▲

Також мають місце наступні властивості додавання і множення комплексних чисел.

3. *Додавання комплексних чисел асоціативне:*

$$\forall\, z_1,z_2,z_3\in\mathbf{C}\ (z_1+z_2)+z_3=z_1+(z_2+z_3).$$

4. *Множення комплексних чисел асоціативне*:

$$\forall\, z_1,z_2,z_3\in\mathbf{C}\ (z_1\cdot z_2)\cdot z_3=z_1\cdot(z_2\cdot z_3).$$

5. $0+z=z$, $\forall\, z\in\mathbf{C}$.

6. $1\cdot z=z$, $\forall\, z\in\mathbf{C}$.

Контрольні питання і завдання

1. Що таке уявна одиниця? Який алгебраїчний вигляд комплексного числа?

2. Що таке дійсна та уявна частини комплексного числа? Знайдіть дійсні та уявні частини комплексних чисел $z_1=-2+i, z_2=-7, z_3=5-4i, z_4=8i.$

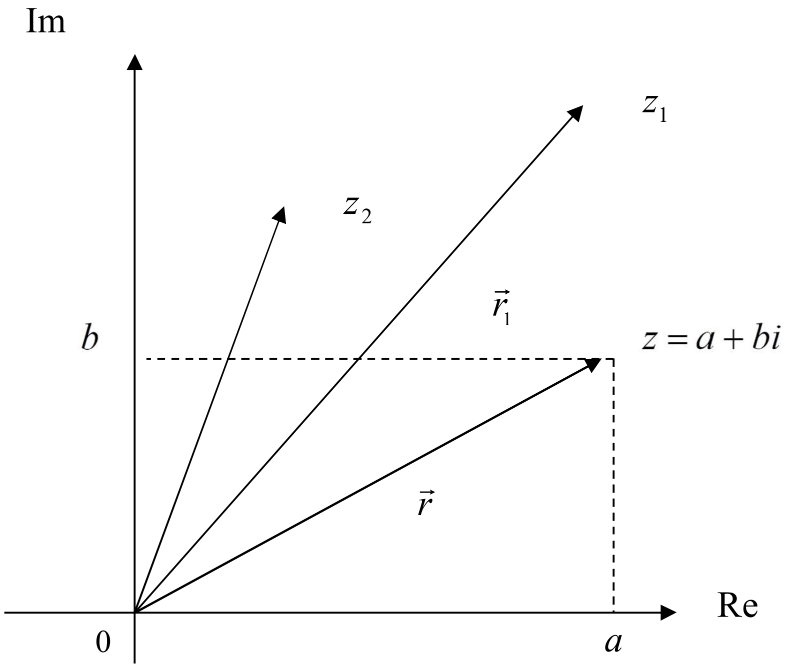
3. Що називається числом, спряженим до комплексного числа z?

4. Знайдіть z_1+z_3, $z_1\cdot z_3$, $z_1:z_3$, де z_1,z_3 – числа із другого завдання.

1.2. Геометричне зображення комплексних чисел.

Та обставина, що комплексне число задається парою дійсних чисел, дає можливість зображати комплексні числа на координатній площині.

Кожному комплексному числу $z = a + bi$ поставимо у відповідність точку (a, b) на координатній площині.



Цю площину називають ***комплексною площиною***. Її вісь абсцис називається дійсною віссю, вісь ординат – уявною віссю. Точка перетину O цих осей має координати (0,0) і відповідає комплексному числу $z = 0 + 0i$. Також довільній точці на комплексній площині з координатами (a_1, b_1) відповідає комплексне число $z_1 = a_1 + b_1 i$.

Кожному комплексному числу $z = a + bi$ відповідає радіус-вектор $\vec{r}$, що утворюється з'єднанням точок O і (a, b). Навпаки, кожному радіус-вектору $\vec{r}_1$ відповідає комплексне число $z_1 = a_1 + b_1 i$, де (a_1, b_1) – координати кінця радіус-вектору $\vec{r}_1$.

Кожний ненульовий радіус-вектор можна задавати його довжиною та кутом між ним і додатнім напрямом дійсної осі (проти годинникової стрілки).

Нехай комплексному числу $z \neq 0$ відповідає радіус-вектор $\vec{r}$.

Довжина вектора $\vec{r}$ називається ***модулем*** числа z і позначається $|z|$. (Якщо $z = 0$, то $|z|$=0). Іншими словами, ***модуль*** комплексного числа z – це відстань від початку координат комплексної площини до точки, яка зображає це число.

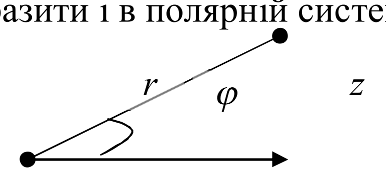
Величина кута, виміряного проти годинникової стрілки, від додатного напряму дійсної осі до ненульового вектору $\vec{r}$, що відповідає числу z, називається ***аргументом*** числа z і позначається $\arg z$. (Аргумент $z = 0$ не визначається).

Наприклад, $\arg(1+i) = \dfrac{\pi}{4}$; $\arg(-1+i) = \dfrac{3\pi}{4}$; $\arg i = \dfrac{\pi}{2}$.

Як можна помітити за означенням, аргумент числа не є однозначним, бо якщо φ_0 – значення аргументу числа z, то аргументом також буде і кут $\varphi = \varphi_0 + 2\pi$, тобто $\arg z = \varphi = \varphi_0 + 2k\pi$, $k \in \mathbf{Z}$. Наприклад, $\arg(1-i) = \dfrac{7\pi}{4} = -\dfrac{\pi}{4}$.

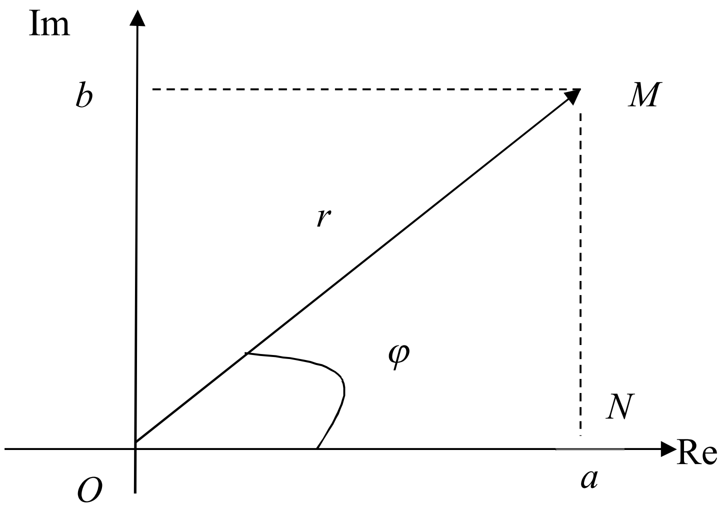
Але ця багатозначність не вносить непорозумінь в теорію і немає необхідності від неї позбавлятися.

Ввівши поняття аргументу і модуля, кожне комплексне число можна зобразити і в полярній системі координат.



Якщо $\arg z = \varphi$, $|z| = r$, то можна писати $z = (r, \varphi)$, оговоривши обрану систему координат.

Між компонентами числа z і його полярними координатами існує зв'язок, який ми зараз встановимо. Розглядаючи одне й те ж число z в різних формах та системах координат, маємо $z = a + bi = M(a,b) = (r,\varphi)$.



1. Розглянемо залежність a та b від r та φ. Оскільки OMN – прямокутний трикутник з катетами a і b та гіпотенузою r, то з геометричних залежностей маємо:

$a = ON = r \cdot \cos\varphi$, $b = MN = r \cdot \sin\varphi$, тоді $z = a + bi = r\cos\varphi + r\sin\varphi \cdot i$, тобто $z = r(\cos\varphi + i\sin\varphi)$ (1.2.1)

Форма (1.2.1) запису комплексного числа називається ***тригонометричною формою комплексного числа***.

2. Одержимо залежність r та φ від a та b, спираючись на той же малюнок з його геометричними залежностями:

$a^2 + b^2 = r^2 \Rightarrow r = \sqrt{a^2 + b^2}$, тобто $|z| = \sqrt{a^2 + b^2}$ – (1.2.2),

формула знаходження модуля комплексного числа $z = a + bi$.

$$\cos\varphi = \frac{a}{r} = \frac{a}{\sqrt{a^2 + b^2}} \quad \Rightarrow \quad \varphi = \pm\arccos\frac{a}{\sqrt{a^2 + b^2}} + 2\pi n, n \in \mathbf{Z}, \text{ (1.2.3)}$$

$$\sin\varphi = \frac{b}{r} = \frac{b}{\sqrt{a^2 + b^2}} \quad \Rightarrow \quad \varphi = (-1)^n \arcsin\frac{b}{\sqrt{a^2 + b^2}} + \pi n, n \in \mathbf{Z}, \text{ (1.2.4)}$$

$$tg\varphi = \frac{b}{a} \quad \Rightarrow \quad \varphi = \mathrm{arctg}\frac{b}{a} + \pi n, n \in \mathbf{Z}. \text{ (1.2.5)}$$

$\arg z$ – це такий кут φ, який задовольняє системі залежностей (1.2.3), (1.2.4), (1.2.5). Зауважимо, що для знаходження аргументу φ достатньо задовольнити дві із вказаних залежностей або, якщо врахувати координатну чверть, в якій знаходиться точка (a, b), то достатньо знайти φ із однієї із цих залежностей. Отже, всі формули (1.2.3), (1.2.4), (1.2.5) є ***формулами для знаходження аргументу комплексного числа*** $z = a + bi$.

<u>Приклад 1</u>. Знайти модуль, аргумент, та тригонометричну форму комплексних чисел $z_1 = -2\sqrt{3} + 2i$ та $z_2 = -2\sqrt{3} - 2i$.

▼ 1) У даного числа z_1 дійсна частина $a_1 = -2\sqrt{3}$, а уявна частина $b_1 = 2$. Знайдемо модуль z_1 за формулою (1.2.2):

$$|z_1| = r_1 = \sqrt{(-2\sqrt{3})^2 + 2^2} = \sqrt{12 + 4} = 4.$$

За формулами (1.2.3) та (1.2.4) знайдемо аргумент z_1:

$$\cos\varphi_1 = \frac{a_1}{r_1} = \frac{-2\sqrt{3}}{4} = -\frac{\sqrt{3}}{2} \quad \Rightarrow \quad \varphi_1 = \pm\arccos\left(-\sqrt{3}/2\right) + 2\pi n, n \in \mathbf{Z},$$

$$\sin\varphi_1 = \frac{b_1}{r_1} = \frac{2}{4} = \frac{1}{2} \quad \Rightarrow \quad \varphi_1 = (-1)^n \arcsin 1/2 + \pi n, n \in \mathbf{Z}.$$

Один із кутів φ_1, що задовольняють попереднім умовам, це кут $\varphi_1 = \frac{5\pi}{6}$. Для поверхневої перевірки вірності знайденого аргументу впевнимося, що і точка, що зображає комплексне число z_1 з координатами $(-2\sqrt{3}, 2)$, і кут $\varphi_1 = \dfrac{5\pi}{6}$ знаходяться в одному й тому ж (другому) координатному куті. Отже остаточно $\arg z_1 = \dfrac{5\pi}{6}$, тому тригонометрична форма (1.2.1) числа z_1 має вигляд: $z_1 = 4(\cos\dfrac{5\pi}{6} + i\sin\dfrac{5\pi}{6})$.

2) Дійсна частина числа z_2 дорівнює $a_2 = -2\sqrt{3}$, а уявна частина $b_2 = -2$. Знайдемо модуль z_2 за формулою (1.2.2):

$$|z_2| = r_2 = \sqrt{(-2\sqrt{3})^2 + (-2)^2} = \sqrt{12+4} = 4.$$

Так як точка з координатами $(-2\sqrt{3}, -2)$, знаходиться в третьому координатному куті, а $\cos\varphi_2 = \dfrac{a_2}{r_2} = \dfrac{-2\sqrt{3}}{4} = -\dfrac{\sqrt{3}}{2}$, то $\varphi_2 = \dfrac{7\pi}{6}$, а тригонометрична форма (1.2.1) числа z_1 має вигляд:

$$z_2 = 4(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}). \ \blacktriangle$$

Наведемо приклади комплексних чисел в алгебраїчній та тригонометричних формах.

$$1 = 1\cdot(\cos 0 + i\sin 0),\ \sqrt{3} + i = 2\cdot(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}),$$

$$-1 = 1\cdot(\cos\pi + i\sin\pi),\ 1 - i = \sqrt{2}\cdot(\cos(-\frac{\pi}{4}) + i\sin(-\frac{\pi}{4})).$$

Нехай числа α_1 і α_2 задано в тригонометричній формі:

$$\alpha_1 = r_1(\cos\varphi_1 + i\sin\varphi_1) \Rightarrow \quad |\alpha_1| = r_1,\ \arg\alpha_1 = \varphi_1,$$

$$\alpha_2 = r_2(\cos\varphi_2 + i\sin\varphi_2) \Rightarrow \quad |\alpha_2| = r_2,\ \arg\alpha_2 = \varphi_2.$$

Обчислимо їх добуток за правилом множення в алгебраїчній формі:

$$\alpha_1\cdot\alpha_2 = r_1\cdot r_2(\cos\varphi_1 + i\sin\varphi_1)\cdot(\cos\varphi_2 + i\sin\varphi_2) =$$

$$= r_1\cdot r_2((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\sin\varphi_1\cos\varphi_2 + \sin\varphi_2\cos\varphi_1)) =$$

$$= r_1\cdot r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2) \quad \Rightarrow$$

$$\alpha_1\cdot\alpha_2 = r_1\cdot r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\ (1.2.6)$$

Формула (1.2.6) – це **правило множення** комплексних чисел **в тригонометричній формі**. З неї маємо:

$|\alpha_1\alpha_2| = r_1\cdot r_2 = |\alpha_1|\cdot|\alpha_2|$ – *модуль добутку дорівнює добутку модулів,*

$\arg(\alpha_1\alpha_2) = \varphi_1 + \varphi_2 = \arg\alpha_1 + \arg\alpha_2$ – *аргумент добутку дорівнює сумі аргументів*.

Це правило множення поширюється на довільну кількість k співмножників $\alpha_1, \alpha_2, ..., \alpha_k$:

$|\alpha_1\alpha_2...\alpha_k| = |\alpha_1| \cdot |\alpha_2| \cdot ... \cdot |\alpha_k|$,

$\arg(\alpha_1...\alpha_k) = \arg\alpha_1 + \arg\alpha_2 + ... + \arg\alpha_k$. (1.2.7)

Обчислимо частку: $\dfrac{\alpha_1}{\alpha_2} = \dfrac{\alpha_1 \cdot \overline{\alpha_2}}{\alpha_2 \cdot \overline{\alpha_2}} = \dfrac{r_1 \cdot r_2(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))}{r_2^2(\cos^2\varphi_2 + \sin^2\varphi_2)} =$

$= \dfrac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$ – (1.2.8) **правило ділення комплексного числа в тригонометричній формі**.

<u>Приклад 2.</u> Знайти добуток і частку чисел $z_1 = -2\sqrt{3} + 2i$ і $z_2 = 1 - i$.

▼ Із попередніх прикладів нам відомо, що тригонометрична форма $z_1 = 4(\cos\dfrac{5\pi}{6} + i\sin\dfrac{5\pi}{6})$, а $z_2 = \sqrt{2} \cdot (\cos(-\dfrac{\pi}{4}) + i\sin(-\dfrac{\pi}{4}))$. Скористаємося формулами множення (1.2.6) і ділення (1.2.8) в тригонометричній формі і одержимо:

$$z_1 \cdot z_2 = 4 \cdot \sqrt{2}(\cos(\frac{5\pi}{6} + (-\frac{\pi}{4})) + i\sin(\frac{5\pi}{6} + (-\frac{\pi}{4})) = 4\sqrt{2}(\cos\frac{7\pi}{12} + i\sin\frac{7\pi}{12}),$$

$$z_1 : z_2 = \frac{4}{\sqrt{2}}(\cos(\frac{5\pi}{6} - (-\frac{\pi}{4})) + i\sin(\frac{5\pi}{6} - (-\frac{\pi}{4})) = 2\sqrt{2}(\cos\frac{13\pi}{12} + i\sin\frac{13\pi}{12}).$$ ▲

У формулі (1.2.7) розглянемо випадок, коли всі співмножники рівні: $\alpha_1 = ... = \alpha_k = \alpha$, тоді $|\alpha^k| = |\alpha|^k$, $\arg(\alpha^k) = k \cdot \arg\alpha$. Тобто, якщо $\alpha = r(\cos\varphi + i\sin\varphi)$, то

$\alpha^k = r^k(\cos k\varphi + i\sin k\varphi)$, $\forall k \in \mathbf{N}$ (1.2.9)

Нехай $k = 0$. Як і в множині дійсних чисел, домовимося, що α^0=1. Тоді $\alpha^0 = 1 = 1(\cos 0 + i\sin 0) = r^0(\cos 0 \cdot \varphi + i\sin 0 \cdot \varphi)$. Очевидно, що $k = 0$ задовольняє формулі (1.2.9).

Нехай $k < 0$ – ціле. Тоді $k = -m$, $m \in \mathbf{N}$. Домовимося, що як і в дійсних числах $\alpha^{-m} = \dfrac{1}{\alpha^m}$. Тоді

$$\alpha^k = \frac{\cos 0 + i\sin 0}{r^m(\cos m\varphi + i\sin m\varphi)} = \frac{(\cos(-m\varphi) + i\sin(-m\varphi))}{r^m} =$$

$$= r^{-m}(\cos(-m)\varphi + i\sin(-m)\varphi) = r^k(\cos k\varphi + i\sin k\varphi).$$

Тобто формула (1.2.9) задовольняється і при $k < 0$. Одержуємо:

$$\alpha^k = r^k(\cos k\varphi + i \sin k\varphi), \ \forall k \in \mathbf{Z} \ (1.2.10)$$

Ця формула (1.2.10) називається ***формулою Муавра***.

<u>Приклад 3.</u> Знайти z^{10}, якщо $z = -2\sqrt{3} + 2i$.

▼ Із попереднього нам відомо, що тригонометрична форма $z = 4(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6})$. Скористаємося формулою Муавра (1.2.10):

$$z^{10} = 4^{10}(\cos(10\cdot\frac{5\pi}{6}) + i\sin(10\cdot\frac{5\pi}{6})) = 4^{10}(\cos\frac{50\pi}{6} + i\sin\frac{50\pi}{6} = 4^{10}(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}).$$

Одержане значення неважко записати в алгебраїчному вигляді, якщо підставити значення $\cos\frac{\pi}{3} = \frac{1}{2}$ і $\sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}$:

$$z^{10} = 4^{10}(\frac{1}{2} + \frac{\sqrt{3}}{2}i) = 2\cdot 4^9 + 2\cdot\sqrt{3}\cdot 4^9 i = 2^{19} + 2^{19}\sqrt{3}i. \ \blacktriangle$$

Контрольні питання і завдання

1. Як зображуються комплексні числа на площині?

2. Що таке модуль комплексного числа (в геометричній інтерпретації)? Знайдіть модулі чисел $2i; \quad 5; \quad -3; \quad -4i$.

3. Що таке аргумент комплексного числа (в геометричній інтерпретації)? Знайдіть аргументи чисел $2i; \quad 5; \quad -3; \quad -4i; \quad -2-2i$.

4. Напишіть формули для знаходження модуля і аргументу комплексного числа в алгебраїчній формі. Знайдіть модуль та аргумент комплексного числа $z = -2 + 2\sqrt{3}i$.

5. Що таке тригонометрична форма комплексного числа? Запишіть число $z = -2 + 2\sqrt{3}i$ в тригонометричній формі.

6. Напишіть формули множення і ділення комплексних чисел в тригонометричній формі. Знайдіть добуток і частку комплексних чисел $z_1 = -2 + 2\sqrt{3}i$ і $z_2 = -2 - 2i$.

7. Напишіть формулу Муавра. Знайдіть z^7, якщо $z = -2 + 2\sqrt{3}i$.

1.3. Видобування кореня

Нехай $n \in \mathbf{N}$. ***Корінь*** n***-го степеня*** з комплексного числа α, це таке комплексне число β, що $\beta^n = \alpha$. Такий корінь може бути не один, але кожне таке число β називається ***коренем*** n***-го степеня від*** α **:** $\beta = \sqrt[n]{\alpha}$.

Зрозуміло, що $\sqrt[n]{0} = 0$. Нехай $\alpha \neq 0$, $\alpha = r(\cos\varphi + i\sin\varphi)$. Знайдемо $\beta = \sqrt[n]{\alpha} = R(\cos\Theta + i\sin\Theta)$. За означенням і формулою (1.2.10) $\alpha = \beta^n = R^n(\cos n\Theta + i\sin n\Theta)$, звідки:

1) $r = R^n \Rightarrow$ (так як $r > 0$, $R > 0$), то $R = \sqrt[n]{r}$ – арифметичний корінь із дійсного числа,

2) $n\Theta = \varphi = \varphi + 2k\pi$, $k \in \mathbf{Z} \Rightarrow \Theta = \dfrac{\varphi + 2k\pi}{n}$, $k \in \mathbf{Z}$.

Таким чином, корені n-го степеня існують і знаходяться по правилу:

$$\sqrt[n]{\alpha} = \sqrt[n]{r}(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}) = \beta_k, \quad k \in \mathbf{Z} \ (1.3.1)$$

Як видно з цієї формули, коренів n-го степеня більше одного (при $n \geq 2$).

Теорема 1.3.1. (Про корені n-го степеня) *Існує рівно n різних значень кореня n-го степеня відмінного від нуля комплексного числа $\alpha = r(\cos\varphi + i\sin\varphi)$. Їх дає формула*

$$\sqrt[n]{\alpha} = \sqrt[n]{r}(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}), \ k = 0,1,2,...,n-1. \ (1.3.2)$$

Доведення. Те, що за формулою (1.3.2) знаходяться корені n-го степеня ми вже довели. Залишилось довести, що різних коренів n (тобто при $k = 0,1,2,...,n-1$ одержуються різні значення β_k, а при інших $k \in \mathbf{Z}$ вони повторюються).

Розглянемо умову, при якій $\beta_{k_1} = \beta_{k_2}$ при $k_1 \neq k_2$.

$$\beta_{k_1} = \sqrt[n]{r}(\cos\frac{\varphi + 2k_1\pi}{n} + i\sin\frac{\varphi + 2k_1\pi}{n}), \ \arg\beta_{k_1} = \frac{\varphi + 2k_1\pi}{n} + 2\pi l, \ l \in \mathbf{Z}$$

$$\beta_{k_2} = \sqrt[n]{r}(\cos\frac{\varphi + 2k_2\pi}{n} + i\sin\frac{\varphi + 2k_2\pi}{n}), \ \arg\beta_{k_2} = \frac{\varphi + 2k_2\pi}{n} + 2\pi s, \ s \in \mathbf{Z}$$

Так як $|\beta_{k_1}| = |\beta_{k_2}| = \sqrt[n]{r}$, то

$$\beta_{k_1} = \beta_{k_2} \Leftrightarrow \arg\beta_{k_1} = \arg\beta_{k_2} \Leftrightarrow \frac{\varphi + 2k_1\pi}{n} + 2\pi l = \frac{\varphi + 2k_2\pi}{n} + 2\pi s \Leftrightarrow$$

$$\Leftrightarrow \frac{\varphi + 2k_1\pi}{n} = \frac{\varphi + 2k_2\pi}{n} + 2\pi t, \text{ де } t = s - l \in \mathbf{Z} \Leftrightarrow$$

$\Leftrightarrow \quad \dfrac{\varphi + 2k_1\pi - \varphi - 2k_2\pi}{n} = 2\pi t \quad \Leftrightarrow \quad \dfrac{k_1 - k_2}{n} = t \quad \Leftrightarrow \quad k_1 - k_2 = tn \quad \Leftrightarrow \quad k_1$ і k_2 – числа, що мають один і той же залишок при діленні на n.

Відношення „мати один і той же залишок при діленні на якесь число n ” на множині $\mathbf{Z}$ – це відношення еквівалентності (перевірте самостійно). Воно множину цілих чисел розбиває на n класів еквівалентності, а числа 0, 1, 2, …, $n-1$ мають різні остачі при діленні на n, отже входять в різні класи еквівалентності.

Так от: $\beta_{k_1} = \beta_{k_2}$ тоді і тільки тоді, коли k_1 і k_2 знаходяться в одному класі еквівалентності відносно вказаного відношення. Тому $\beta_k \neq \beta_m$ тоді і тільки тоді, коли k і m не будуть належати одному класу еквівалентності. Так як різних класів існує n, то і різних коренів буде n. Щоб виписати їх, візьмемо по одному числу з кожного класу, а саме 0, 1, 2,..., $n-1$:

$$\beta_0 = \sqrt[n]{r}(\cos\frac{\varphi}{n} + i\sin\frac{\varphi}{n}),$$

$$\beta_1 = \sqrt[n]{r}(\cos\frac{\varphi + 2\pi}{n} + i\sin\frac{\varphi + 2\pi}{n}),$$

$$\beta_2 = \sqrt[n]{r}(\cos\frac{\varphi + 2\cdot 2\pi}{n} + i\sin\frac{\varphi + 2\cdot 2\pi}{n}),$$

...

$$\beta_{n-1} = \sqrt[n]{r}(\cos\frac{\varphi + 2(n-1)\pi}{n} + i\sin\frac{\varphi + 2(n-1)\pi}{n}).$$

Тоді $\sqrt[n]{\alpha} = \{\beta_0, \beta_1, ..., \beta_{n-1}\}$ або

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

<u>Приклад</u> $\alpha = -1 + \sqrt{3}i$. Знайти $\sqrt[4]{\alpha}$.

▼ Знайдемо модуль α за формулою (2):

$|\alpha| = r = \sqrt{1+3} = 2$. Знайдемо аргумент α із системи умов (1.2.3), (1.2.4): $\cos\varphi = -\dfrac{1}{2}$, $\sin\varphi = \dfrac{\sqrt{3}}{2} \quad \Rightarrow \quad \arg\alpha = \varphi = \dfrac{2\pi}{3}$. Тоді

$\alpha = 2(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3})$ – тригонометрична форма числа α. Формула (1.3.2) в нашому випадку має вигляд:

$$\sqrt[4]{\alpha} = \sqrt[4]{2}(\cos\frac{2\pi/3 + 2k\pi}{4} + i\sin\frac{2\pi/3 + 2k\pi}{4}),\ k = 0, 1, 2, 3.$$

А саме

$$\beta_0 = \sqrt[4]{2}(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}) = \frac{\sqrt[4]{2}\sqrt{3}}{2} + i\frac{\sqrt[4]{2}}{2},$$

$$\beta_1 = \sqrt[4]{2}(\cos\frac{2\pi/3 + 2\pi}{4} + i\sin\frac{2\pi/3 + 2\pi}{n}) = \sqrt[4]{2}(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}) = -\frac{\sqrt[4]{2}}{2} + i\frac{\sqrt[4]{2}\sqrt{3}}{2},$$

$$\beta_2 = \sqrt[4]{2}(\cos\frac{2\pi/3 + 4\pi}{4} + i\sin\frac{2\pi/3 + 4\pi}{n}) = \sqrt[4]{2}(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}) = -\frac{\sqrt[4]{2}\sqrt{3}}{2} - i\frac{\sqrt[4]{2}}{2}$$

$$\beta_3 = \sqrt[4]{2}(\cos\frac{2\pi/3 + 6\pi}{4} + i\sin\frac{2\pi/3 + 6\pi}{n}) = \sqrt[4]{2}(\cos\frac{5\pi}{3} + i\sin\frac{5\pi}{3}) = -\frac{\sqrt[4]{2}}{2} - i\frac{\sqrt[4]{2}\sqrt{3}}{2}.$$

Одержимо відповідь:

$$\sqrt[4]{-1 + \sqrt{3}i} =$$

$$= \left\{\frac{\sqrt[4]{2}\sqrt{3}}{2} + i\frac{\sqrt[4]{2}}{2},\ -\frac{\sqrt[4]{2}}{2} + i\frac{\sqrt[4]{2}\sqrt{3}}{2},\ -\frac{\sqrt[4]{2}\sqrt{3}}{2} - i\frac{\sqrt[4]{2}}{2},\ -\frac{\sqrt[4]{2}}{2} - i\frac{\sqrt[4]{2}\sqrt{3}}{2}\right\}.$$

▲

Оскільки $1 = \cos 0 + i\sin 0$, то корені степеня n із одиниці мають вигляд: $\varepsilon_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}$, $k = 0, 1, 2, ..., n-1$.

Звернемо увагу на те, що

$$\varepsilon_0 = 1,$$

$$\varepsilon_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n},\ |\varepsilon_1| = 1,\ \arg\varepsilon_1 = 2\pi/n,$$

$$\varepsilon_2 = \cos 2\cdot\frac{2\pi}{n} + i\sin 2\cdot\frac{2\pi}{n},\ |\varepsilon_2| = 1,\ \arg\varepsilon_2 = 2\cdot 2\pi/n,\ \ldots\ .$$

З цього аналізу видно, що при зображенні на комплексній площині всі корені n-го степеня із одиниці розташовані на одиничному колі під кутом, кратним $\frac{2\pi}{n}$ (n- на частина кола), який вимірюється від додатної частини дійсної осі. Наприклад, так як $\frac{2\pi}{4} = \frac{\pi}{2}$, то

$$\sqrt[4]{1} = \left\{1, \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}, \cos\pi + i\sin\pi, \cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\right\} = \{1, i, -1, -i\}.$$

Твердження 1.3.2. (Про добуток коренів із одиниці) *Добуток двох коренів із одиниці степеня n є корінь із одиниці степеня n.*

Доведення. Нехай α і β – корені n-го степеня із одиниці. Розглянемо $(\alpha \cdot \beta)^n$. Так як множення комплексних чисел комутативне і асоціативне, то

$$(\alpha \cdot \beta)^2 = (\alpha \cdot \beta) \cdot (\alpha \cdot \beta) = \alpha(\beta(\alpha\beta)) = \alpha((\beta\alpha)\beta) =$$
$$= \alpha((\alpha\beta)\beta) = \alpha(\alpha(\beta\beta)) = (\alpha\alpha)(\beta\beta) = \alpha^2\beta^2$$

і аналогічно $(\alpha \cdot \beta)^n = \alpha^n \beta^n$.

Таким чином $(\alpha \cdot \beta)^n = \alpha^n \beta^n = 1 \cdot 1 = 1 \Rightarrow \alpha\beta = \sqrt[n]{1}$. ■

Твердження 1.3.3. *Всі значення $\sqrt[n]{\alpha}$ $(\alpha \in \mathbf{C},\ \alpha \neq 0)$ утворюються із якогось одного значення $\sqrt[n]{\alpha}$ множенням на всі корені степеня n із одиниці.*

Доведення. 1). Візьмемо деяке $\beta = \sqrt[n]{\alpha}$ і довільне $\varepsilon = \sqrt[n]{1}$. Розглянемо $\gamma = \beta \cdot \varepsilon$.

$\gamma^n = \beta^n \cdot \varepsilon^n = \alpha \cdot 1 = \alpha \Rightarrow \gamma = \sqrt[n]{\alpha}$. Тобто $\beta \cdot \varepsilon = \sqrt[n]{\alpha}$ для довільного $\varepsilon = \sqrt[n]{1}$ і вибраного β.

2). Тепер покажемо, що всі корені δ n-го степеня із α мають вигляд $\beta\varepsilon$:

$$(\delta \cdot \beta^{-1})^n = \left(\frac{\delta}{\beta}\right)^n = \frac{\delta^n}{\beta^n} = \frac{\alpha}{\alpha} = 1 \Rightarrow$$

$\Rightarrow \delta \cdot \beta^{-1} = \varepsilon$ – деякий корінь n-го степеня із 1 $\Rightarrow \delta = \varepsilon\beta$. ■

<u>Приклад</u>. Знайти всі корені 4-го степеня із числа $\alpha = -8 + 8\sqrt{3}i$, якщо відомо, що один з його коренів $\beta = -1 + \sqrt{3}i$.

▼ За твердженням 3.4.3 всі $\sqrt[4]{\alpha}$ можна знайти, помноживши β на всі $\sqrt[4]{1} = \{1, i, -1, -i\}$. Отже,

$$\sqrt[4]{\alpha} = \{\beta, \beta \cdot i, -\beta, -\beta \cdot i\} = \{-1 + \sqrt{3}i, -\sqrt{3} - i, 1 - \sqrt{3}i, \sqrt{3} + i\}. \ ▲$$

Контрольні питання і завдання

1. Дайте означення кореня n-го степеня комплексного числа.
2. Напишіть формулу кореня n-го степеня. Знайдіть $\sqrt[6]{64}$.
3. Напишіть формулу коренів n-го степеня із одиниці. Знайдіть корені 6-го степеня із одиниці.

4. Знайдіть всі корені 12-го степеня із одиниці та їх зображення. Знайдіть зображення всіх $\sqrt[12]{\alpha}$, якщо відомо, що один із цих коренів $\beta = 2 + 3i$.

Розділ 2. СИСТЕМИ ЛІНІЙНИХ РІВНЯНЬ

2.1. Матриці. Операції над матрицями

В цьому розділі ми будемо розглядати системи лінійних рівнянь та матриці з дійсними коефіцієнтами. Позначимо через **R** множину дійсних чисел. ***Матрицею*** розміром m на n ($m \times n$) називається прямокутна таблиця дійсних чисел, що має m рядків і n стовпців. Множина всіх матриць розміром m на n позначається $M_{m\times n}(\mathbf{R})$.

Будемо користуватися такими позначеннями матриці:

$$A = A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} = (a_{ij})_m^n = (a_{ij})_{m\times n} = (a_{ij}).$$

Числа a_{ij} називають ***елементами матриці***, перший індекс i позначає номер рядка, другий індекс j – номер стовпця елементу a_{ij}.

У випадку $m = n$ матриця A називається ***квадратною***, а множина таких матриць позначається $\mathrm{M}_n(\mathbf{R})$. Число n називається ***порядком*** квадратної матриці A. Ряд чисел $a_{11}, a_{22}, \ldots, a_{nn}$ квадратної матриці A називають ***головною діагоналлю***, а ряд $a_{1n}, a_{2(n-1)}, \ldots, a_{n1}$ – ***бічною діагоналлю***.

Дві $m\times n$- матриці $A = (a_{ij})$ і $B = (b_{ij})$ називаються ***рівними*** ($A = B$), якщо їх елементи відповідно рівні: $a_{ij} = b_{ij}$, $i = 1, \ldots, m;$ $j = 1, \ldots, n.$

На множині матриць $M_{m\times n}(\mathbf{R})$ одного розміру ($m \times n$) можна визначити операції **додавання** двох матриць і **множення** матриці **на число**.

Нехай $A, B \in M_{m\times n}(\mathbf{R})$, $A = (a_{ij})$, $B = (b_{ij})$. Операція ***додавання*** задається за таким правилом: $A + B = C \in M_{m\times n}(\mathbf{R})$,

де $C = (c_{ij})$, $c_{ij} = a_{ij} + b_{ij}$ для всіх $1 \le i \le m$, $1 \le j \le n$, тобто додаються елементи, що стоять на однакових місцях і результат записується на таке ж саме місце.

<u>Приклад 1.</u>

$$\begin{pmatrix} 2 & 3 & 7 \\ 5 & -2 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 3 & 8 \\ 7 & 6 & -5 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 15 \\ 12 & 4 & -2 \end{pmatrix}.$$

Нехай $A = (a_{ij}) \in M_{m\times n}(\mathbf{R})$, $\alpha \in \mathbf{R}$. Операція ***множення на число*** α задається за таким правилом: $\alpha \cdot A = C \in M_{m\times n}(\mathbf{R})$,
де $C = (c_{ij})$, $c_{ij} = \alpha \cdot a_{ij}$ для всіх $1 \le i \le m$, $1 \le j \le n$, тобто кожний елемент одержаної матриці є добутком числа α на відповідний елемент даної матриці.

Приклад 2.

$$-4 \cdot \begin{pmatrix} 2 & -3 & 7 \\ 5 & -2 & 3 \end{pmatrix} = \begin{pmatrix} -8 & 12 & -28 \\ -20 & 8 & -12 \end{pmatrix}.$$

Із означень цих операцій випливають

Властивості операцій додавання матриць та множення їх на число.

Для будь-яких $A, B, C \in M_{m\times n}(\mathbf{R})$ та будь-яких $\alpha, \beta \in \mathbf{R}$ операції додавання і множення на число мають такі властивості:

1. $(A + B) + C = A + (B + C)$;
2. $A + B = B + A$;
3. Якщо $O = (0)$ – матриця, що складається з нулів, то $A + O = A$;
4. Для кожної матриці $A = (\alpha_{ij})$ існує протилежна матриця $-A \;\; = (-\alpha_{ij})$, тобто така, що $A + (-A) = O$;
5. $1 \cdot A = A$;
6. $\alpha \cdot (\beta \cdot A) = (\alpha\beta) \cdot A$;
7. $\alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$;
8. $(\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$;
9. $(-1) \cdot A = -A$;
10. $0 \cdot A = O$.

До довільної матриці можна застосувати таке перетворення, яке називається **транспонуванням**. Для того, щоб транспонувати матрицю $A = (\alpha_{ij})$ розміром $m \times n$, розташуємо рядки цієї матриці у вигляді стовпчиків, не змінюючи їх порядку (тобто перший рядок стане першим стовпчиком і т. д.). Одержимо $n \times m$- матрицю:

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

яка називається ***транспонованою*** по відношенню до матриці A і позначається A^T. Якщо позначити $A^T = (\alpha'_{ij})$, то

$\alpha'_{ij} = \alpha_{ji}, \quad i = 1,...,n,\ j = 1,...,m$. (2.1.1)

Для довільних матриць $A, B \in M_{m\times n}(\mathbf{R})$ та довільного числа $\alpha \in \mathbf{R}$ виконуються наступні властивості:

1. $(A+B)^T = A^T + B^T$;
2. $(\alpha A)^T = \alpha A^T$.

Для будь-яких матриць $A = (a_{ij}) \in M_{m\times n}(\mathbf{R})$ та $B = (b_{ij}) \in M_{n\times p}(\mathbf{R})$ існує операція **множення** за правилом:

$$A \cdot B = C \in M_{m\times p}(\mathbf{R}),\text{ де } C = (c_{ij}),\ c_{ij} = a_{i1}b_{1j} + \dots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} \quad (2.1.2)$$

для всіх $1 \le i \le m$, $1 \le j \le p$, причому матриця C називається ***добутком*** матриць A і B.

Зауважимо, що добуток AB визначений тільки для таких матриць, у яких кількість стовпчиків матриці A (першого співмножника) дорівнює кількості рядків матриці B (другого співмножника). При цьому кількість рядків матриці $C = AB$ дорівнює кількості рядків матриці A, а кількість стовпчиків матриці C дорівнює кількості стовпчиків матриці B. Отже, раціонально записувати

$$A_{m\times n} \cdot B_{n\times p} = C_{m\times p}.$$

Звернімо також увагу на те, що для одержання значення елементу c_{ij} за правилом множення беруться елементи із i-го рядка матриці A та j-го стовпчика матриці B, відповідні елементи множаться, а ці добутки додаються. Умовно говорять, що «множиться i-й рядок на j-й стовпчик».

<u>Приклад 3.</u>

$$AB = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & -2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & 7 & 1 \\ 2 & -1 & 5 \\ 3 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 29 & 19 & 31 \\ 2 & 9 & -3 \end{pmatrix} = C.$$

▼ Матриця C знаходиться за правилом (2.1.2):

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} + a_{14}b_{41} = 2\cdot 3 + 3\cdot 2 + 4\cdot 3 + 5\cdot 1 = 29;$$
$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} + a_{14}b_{42} = 2\cdot 7 + 3\cdot(-1) + 4\cdot 2 + 5\cdot 0 = 19;$$
$$c_{13} = a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} + a_{14}b_{43} = 2\cdot 1 + 3\cdot 5 + 4\cdot 1 + 5\cdot 2 = 31;$$
$$c_{21} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} + a_{24}b_{41} = 1\cdot 3 + (-2)\cdot 2 + 0\cdot 3 + 3\cdot 1 = 2 \text{ і т. д. } \blacktriangle$$

Твердження 2.1.1. *Якщо існує добуток* AB, *то* $(AB)^T = B^T A^T$.

Д о в е д е н н я. Нехай $A = A_{m\times n} = (a_{ij})$, $B = B_{n\times p} = (b_{ij})$, $A\cdot B = C_{m\times p} = (c_{ij})$. Тоді $A^T = A^T_{n\times m} = (a'_{st})$, $B^T = B^T_{p\times n} = (b'_{st})$ і, очевидно, що існує добуток $B^T A^T = D_{p\times m} = (d_{st})$, де $d_{st} = \sum_{i=1}^{n} b'_{si} a'_{it}$, і розмір $D_{p\times m}$ співпадає з розміром $(A\cdot B)^T = C^T_{p\times m} = (c'_{st})$.

Порівняємо значення c'_{st} і d_{st}. Оскільки за умовою (2.1.1) $c'_{st} = c_{ts}$, то елемент c'_{st} дорівнює добутку t-го рядка матриці A та s-го стовпчика матриці B: $c'_{st} = c_{ts} = \sum_{i=1}^{n} a_{ti} b_{is}$. Так як за умовою (2.1.1) $a'_{it} = a_{ti}$, $b'_{si} = b_{is}$, то $\sum_{i=1}^{n} a_{ti} b_{is} = \sum_{i=1}^{n} a'_{it} b'_{si} = \sum_{i=1}^{n} b'_{si} a'_{it} = d_{st}$. Отже, $c'_{st} = d_{st}$ для всіх s і t, тому $D_{p\times m} = C^T_{p\times m}$ і $(AB)^T = B^T A^T$. ■

Властивості операцій множення матриць

1. Нехай існує добуток $A(BC)$. Ця умова виконується, якщо $A \in M_{m\times k}(\mathbf{R})$, $B \in M_{k\times r}(\mathbf{R})$, $C \in M_{r\times p}(\mathbf{R})$ при деяких натуральних m, k, r, p. Тоді існують добутки AB і $(AB)C$. І навіть більше того:

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

тобто *множення матриць асоціативне*.

2. Якщо $A \in M_{m\times n}(\mathbf{R})$, $B \in M_{n\times p}(\mathbf{R})$, то існує добуток AB, але не завжди існує добуток BA (він існує тільки при $m = p$).

Якщо ж $A \in M_{m\times n}(\mathbf{R})$, $B \in M_{n\times m}(\mathbf{R})$, тоді існують одночасно AB порядку m і BA порядку n. Рівність AB і BA може виконуватись тільки при $m = n$. Взагалі кажучи, навіть якщо $A, B \in M_n(\mathbf{R})$, то

$$AB \neq BA,$$

тобто *множення матриць не комутативне*.

3. Нехай $A \in M_{m\times n}(\mathbf{R})$, $B, C \in M_{n\times p}(\mathbf{R})$, $D \in M_{p\times k}(\mathbf{R})$. Тоді

$$A(B + C) = AB + AC \text{ і } (B + C)D = BD + CD,$$

тобто виконуються так звані *закони дистрибутивності*.

4. Якщо $\alpha \in \mathbf{R}$, $A \in M_{m\times n}(\mathbf{R})$, $B \in M_{n\times p}(\mathbf{R})$, то. $\alpha(AB) = (\alpha A)B$.

Описані властивості випливають із означень операцій, але потребують доведень. (Доведіть самостійно).

Введемо так звані ***символи Кронекера***:

$\delta_{ij} = \begin{cases} 1, & \text{при} \quad i = j \\ 0, & \text{при} \quad i \neq j \end{cases}$, де i і j – довільні натуральні числа.

Матриця $E = (\delta_{ij})_{n \times n} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ називається ***одиничною*** матрицею n-го порядку. Її також позначають E_n.

Нехай $A \in M_{m \times n}(\mathbf{R})$. Тоді $AE_n = A$, $E_m A = A$. (При $A \in M_n(\mathbf{R})$ виконуються рівності $AE = EA = A$).

Матриця A^{-1} називається ***оберненою*** до A, якщо виконуються наступні рівності: $AA^{-1} = A^{-1}A = E$. Так як E – квадратна матриця, то для того, щоб у матриці A існувала обернена матриця, необхідно, щоб A (і разом з нею A^{-1}) була квадратною матрицею. Далі буде доведено, що не кожна квадратна матриця має обернену, а якщо A^{-1} *існує, то вона єдина для довільної матриці* A.

Блочні матриці.

Матриця A за допомогою горизонтальних та вертикальних прямих може розбиватися на блоки (на матриці менших розмірів), тоді матриця

$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & & & \\ & & & \dots \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix} \cdot = \begin{pmatrix} A_{11} & & \dots & A_{1t} \\ \dots & & & \\ & & & \dots \\ A_{s1} & & & A_{st} \end{pmatrix}$, де A_{ij} – матриці: в кожному рядку стоять матриці з однаковою кількістю рядків, а в кожному стовпчику, з однаковою кількістю стовпчиків.

Цікаво, що блочні матриці додаються за тим же правилом, що і звичайні.

$A + B = (A_{\alpha\beta}) + (B_{\alpha\beta}) = (A_{\alpha\beta} + B_{\alpha\beta})$, якщо $A_{\alpha\beta}$ та $B_{\alpha\beta}$ матриці одного розміру.

Блочні матриці з відповідним розподілом кліток і множаться, як і звичайні: $A \cdot B = (\sum_{\beta} A_{\alpha\beta} B_{\beta\gamma})$.

Прямою сумою двох квадратних матриць A і B порядків m і n відповідно називається квадратна матриця С порядку ($m+n$):

$$C=\begin{pmatrix} A & O \\ O & B \end{pmatrix} (C = A \oplus B).$$

Контрольні питання та завдання

1. Що таке матриця? Чи можуть всі елементи матриці дорівнювати нулю?

2. Які матриці називаються рівними? Чи рівні матриці $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ і $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$?

3. Як додати дві матриці? Чи можна додати дві матриці з розмірами 2×3 і 3×2?

4. Яка матриця називається транспонованою по відношенню до матриці A? Чи може виконуватись рівняння $A^T = A$? Відповідь обґрунтуйте.

5. Як обчислюються елементи добутку матриць?

6. Відомо, що $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot A = \begin{pmatrix} 0 & 1 \end{pmatrix}$. Які розміри матриці A?

7. Якого розміру буде добуток матриць A і B, якщо вони відповідно мають розміри:

   а) (4×7) та (7×5);

   б) (7×5) та (4×7).

8. Що таке одинична матриця? Чи є одиничною матриця:

   а) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$; б) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; в) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$?

9. Яка матриця називається оберненою до даної?

2.2. Елементарні перетворення матриць

Існує три види ***елементарних перетворень*** рядків матриці:

1) перестановка двох рядків місцями;

2) множення рядка на число, відмінне від нуля;

3) додавання до одного рядка іншого, помноженого на деяке число.

Такі ж елементарні перетворення існують і для стовпчиків.

Матриця розміром $m\times n$ виду

$$C=\begin{pmatrix} 0 & \dots & 0 & c_{1s_1} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & c_{1n} \\ 0 & \dots & 0 & 0 & \dots & 0 & c_{2s_2} & \dots & \dots & \dots & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & c_{rs_r} & \dots & c_{rn} \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}, \ (2.2.1)$$

де $c_{is_i} \neq 0$, $i=1,2,...,r$, $r \leq m,n$ називається ***східчастою матрицею***.

Теорема 2.2.1. (Про приведення до східчастої матриці) *За допомогою елементарних перетворень рядків кожну матрицю можна привести до східчастої матриці.*

Д о в е д е н н я. Якщо матриця $A=O$, то вона східчаста. Нехай дано матрицю $A=(a_{ij})_m^n \neq O$.

Доведення проведемо методом математичної індукції по числу рядків m .

Якщо m=1, то довільна матриця A – східчаста.

Нехай теорему доведено для всіх матриць із $\mathrm{M}_{(m-1)\times n}(\mathbf{R})$. Доведемо теорему для матриці $A \in \mathrm{M}_{m\times n}(\mathbf{R})$.

Нехай s_1 – мінімальний номер ненульового стовпчика матриці A. Тоді існує таке i, що елемент $a_{is_1} \neq 0$. Нехай $i=1$ ($a_{1s_1} \neq 0$).

Виконаємо наступні елементарні перетворення рядків третього типу: до k-го рядка матриці A додамо 1-й, помножений на $-\dfrac{a_{ks_1}}{a_{1s_1}}$. Одержимо матрицю B з елементами $b_{1j}=a_{1j}$, $j=1,..,n$;

$$b_{kj}=a_{kj}-\frac{a_{ks_1}}{a_{1s_1}}\cdot a_{1j},\ k=2,...,m,\ j=1,..,n.$$

Оскільки перші стовпчики нульові, то $b_{kj}=0$, $k=2,...,m$, $j=1,..,s_1$. Якщо $s_1=n$, то

$$B=\begin{pmatrix} 0 & \dots & 0 & b_{1n} \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 \end{pmatrix} - (1)$$

східчаста матриця і теорема доведена.

При $s_1 < n$ матриця A приводиться до матриці B виду

$$B = \begin{pmatrix} 0 & \dots & 0 & b_{1s_1} & \dots & \dots & b_{1n} \\ 0 & \dots & 0 & 0 & b_{2(s_1+1)} & \dots & b_{2n} \\ 0 & \dots & 0 & 0 & b_{3(s_1+1)} & \dots & b_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & b_{m(s_1+1)} & \dots & b_{mn} \end{pmatrix}. \ (2)$$

Якщо $i \neq 1$ ($a_{1s_1} = 0$), то спочатку виконаємо елементарне перетворення першого типу: поміняємо місцями 1-й та i-й рядки. Одержимо матрицю $A' = (a_{ij}^{'})$, $a'_{1s_1} \neq 0$, і застосуємо до неї попередні елементарні перетворення, що приведе до матриці виду (1) або (2).

Якщо матриця B виду (2) (не східчаста), то розглянемо матрицю

$$B' = \begin{pmatrix} 0 & \dots & 0 & b_{2(s_1+1)} & \dots & b_{2n} \\ 0 & \dots & 0 & b_{3(s_1+1)} & \dots & b_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & b_{m(s_1+1)} & \dots & b_{mn} \end{pmatrix} \in \mathrm{M}_{(m-1)\times n}(\mathbf{R}),$$ яка відрізняється від B відсутністю першого рядка. За припущенням індукції матриця B' за допомогою елементарних перетворень своїх рядків приводиться до східчастої форми:

$$C' = \begin{pmatrix} 0 & \dots & 0 & c_{2s_2} & \dots & \dots & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & c_{rs_n} & \dots & c_{rn} \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \end{pmatrix}.$$

Оскільки рядки матриці B' є останніми рядками матриці B, то при позначеннях $c_{1j} = b_{1j}$, $j = 1,\dots,n$, матриця B за допомогою елементарних перетворень матриці B' приведеться до східчастої матриці C виду (2.2.1). ■

<u>Приклад 1.</u> Приведемо дану матрицю до східчастої форми:

$$\begin{pmatrix} 1 & -2 & 3 & 2 & 0 & 5 \\ 3 & -6 & 2 & 4 & 5 & 0 \\ 2 & -4 & -1 & 2 & 3 & 4 \\ -2 & 4 & 3 & -1 & -3 & 2 \end{pmatrix} \begin{array}{lll} (-3) & (-2) & 2 \\ \hookleftarrow & & \\ & \hookleftarrow & \\ & & \hookleftarrow \end{array} \approx$$

$$\approx\begin{pmatrix} 1 & -2 & 3 & 2 & 0 & 5 \\ 0 & 0 & -7 & -2 & 5 & -15 \\ 0 & 0 & -7 & -2 & 3 & -6 \\ 0 & 0 & 9 & 3 & -3 & 12 \end{pmatrix}\begin{array}{ll} & \\ (-1) & 9 \\ \hookleftarrow & \\ & 7\hookleftarrow \end{array}\approx$$

$$\approx\begin{pmatrix} 1 & -2 & 3 & 2 & 0 & 5 \\ 0 & 0 & -7 & -2 & 5 & -15 \\ 0 & 0 & 0 & 0 & -2 & 9 \\ 0 & 0 & 0 & 3 & 24 & -51 \end{pmatrix}\begin{array}{l} \\ \\ (-1) \\ :3 \end{array}\approx\begin{pmatrix} 1 & -2 & 3 & 2 & 0 & 5 \\ 0 & 0 & -7 & -2 & 5 & -15 \\ 0 & 0 & 0 & 1 & 8 & -17 \\ 0 & 0 & 0 & 0 & 2 & -9 \end{pmatrix}. \ \blacktriangle$$

Контрольні питання та завдання

1. Яка матриця називається східчастою? Чи східчаста матриця

$$\text{а)}\begin{pmatrix} 0 & 2 & -3 & 0 & 3 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}?\ \text{б)}\begin{pmatrix} 7 & -3 & 2 & 3 & 4 \\ 0 & 2 & 5 & 6 & 1 \\ 0 & 1 & -1 & 0 & 0 \end{pmatrix}?$$

2. Які існують види елементарних перетворень матриць?

2.3. Визначники. Властивості визначників

Перестановкою із n елементів називається будь-яка упорядкована множина n перших натуральних чисел. В загальному вигляді перестановку p із n елементів можна записати

$$p=(\alpha_1,\alpha_2,...,\alpha_n),\ \alpha_i\in\{1,2,...,n\};\ \alpha_i\neq\alpha_j,\text{якщо } i\neq j.$$

Різні перестановки n елементів відрізняються порядком чисел.

(2, 3, 1, 4) $\neq$ (3, 1, 4, 2)

Теорема 2.3.1. (Про число різних перестановок) *Число різних перестановок із n елементів дорівнює $n!$ ($P_n=n!$).*

Доведення проводиться методом математичної індукції (виконайте самостійно).

Якщо в деякій перестановці $(\alpha_1,\alpha_2,...,\alpha_n)$ поміняти місцями два елементи α_i та α_j, а останні залишити на місці, то утвориться нова перестановка, а це перетворення має назву ***транспозиція*** (позначається τ_{ij}).

$\{(3,4,2,1)\to(3,1,2,4)\}$ – транспозиція τ_{24}.

Нехай дана перестановка $p=(\alpha_1,\alpha_2,...,\alpha_i,...,\alpha_j,...,\alpha_n)$. Ясно, що $i<j$.

Якщо у перестановці p має місце нерівність $\alpha_i > \alpha_j$, то говорять, що числа α_i та α_j складають ***інверсію***. Загальна кількість інверсій перестановки p позначається $\nu(p)$.

Приклад. Знайти загальну кількість інверсій:

1. p_1 = (7, 2, 5, 3, 4, 1, 6).

▼2 і 3, які стоять відповідно на місцях 2 і 4, не складають інверсію, бо $2<4$ і $2<3$; 7 і 3, які стоять відповідно на місцях 1 і 4, складають інверсію, бо $1<4$, але $7>3$. Загальна кількість інверсій у цій перестановці $\nu(p_1)=12$:

(7,2), (7,5), (7,3), (7,4), (7,1), (7,6), (2,1), (5,3), (5,4), (5,1),(3,1), (4,1). ▲

2. p_2 = (1, 2, 3, 4, 5, 6, 7) – не має інверсій, $\nu(p_2)=0$.

3. p_3 = (1, 2, 4, 3, 5, 6, 7) – 1 інверсія (4 і 3), $\nu(p_3)=1$.

Перестановка називається ***парною***, якщо число її інверсій парне, і ***непарною*** – навпаки.

$\nu(4{,}5{,}1{,}3{,}6{,}2)=8$ – парна («4» – 3 інверсії, «5» – 3 інверсії, «1» – 0, «3» – 1, «6» – 1, «2» – 0).

$\nu(3{,}2{,}5{,}1{,}4{,}6)=5$ – непарна («3» – 2, «2» – 1, «5» – 2, «4» – 0, «6» – 0).

$\nu(1{,}2{,}3{,}4{,}5)=0$ – парна.

$\nu(1{,}2{,}3{,}...{,}n)=0$ – парна.

Теорема 2.3.2. (Про зміну парності при транспозиції) *Кожна транспозиція змінює парність перестановки.*

Д о в е д е н н я. 1) Нехай дана перестановка $p=(\alpha_1\alpha_2...\alpha_i\alpha_{i+1}...\alpha_n)$.

Будемо робити транспозицію елементів α_i та α_{i+1} (це називається ***транспозицією сусідніх***). Одержимо $p_1=(\alpha_1\alpha_2...\alpha_{i+1}\alpha_i...\alpha_n)$. Від цієї транспозиції не змінюється кількість інверсій елементів α_i та α_{i+1} з останніми.

Якщо $\alpha_i>\alpha_{i+1}$ – (α_i і α_{i+1} складають інверсію в p), то після транспозиції в p_1 цієї інверсії не буде.

Якщо $\alpha_i<\alpha_{i+1}$ – (не складають інверсію), то після транспозиції α_i та α_{i+1} будуть складати інверсію.

В обох випадках загальне число інверсій змінюється на 1 ($\nu(p_1)=\nu(p)\pm 1$), тобто зміниться парність перестановки p_1 відносно p. Одержуємо, що транспозиція сусідніх змінює парність перестановки. ▲

2). Нехай тепер в перестановці $p=(\alpha_1\alpha_2...\alpha_i\alpha_{i+1}\alpha_{i+2}...\alpha_{i+s}\alpha_{i+s+1}...\alpha_i)$ треба зробити транспозицію елементів α_i та α_{i+s+1} (між ними s елементів).

Зробимо цю транспозицію послідовною транспозицією сусідних:

$$\left.\begin{array}{ll|l} 1. & \alpha_i \leftrightarrow \alpha_{i+1} & ...\alpha_{i+1},\alpha_i,\alpha_{i+2},...,\alpha_{i+s+1} \quad ... \\ 2. & \alpha_i \leftrightarrow \alpha_{i+2} & ...\alpha_{i+1},\alpha_{i+2},\alpha_{i,},...,\alpha_{i+s+1} \quad ... \\ & ... & \quad ... \\ s+1. & \alpha_i \leftrightarrow \alpha_{i+s+1} & ...\alpha_{i+1},\alpha_{i+2},...,\alpha_{i+s+1},\alpha_i \quad ... \end{array}\right\}(s+1) \quad - \quad \text{транспозицій}$$

сусідніх.

$$\left.\begin{array}{ll|l} s. & \alpha_{i+s+1} \leftrightarrow \alpha_{i+s} & ...,\alpha_{i+1},\alpha_{i+2},...,\alpha_{i+s+1},\alpha_{i+s}\alpha_i,... \\ s-1. & \alpha_{i+s+1} \leftrightarrow \alpha_{i+s-1} & ...,\alpha_{i+1},...\alpha_{i+s+1},\alpha_{i+s-1},\alpha_{i+s},\alpha_i,... \\ & ... & \quad ... \\ 2. & \alpha_{i+s+1} \leftrightarrow \alpha_{i+2} & ...,\alpha_{i+1},\alpha_{i+s+1},\alpha_{i+2},...,\alpha_{i+s},\alpha_i,... \\ 1. & \alpha_{i+s+1} \leftrightarrow \alpha_{i+1} & ...,\alpha_{i+s+1},\alpha_{i+1},\alpha_{i+2},...,\alpha_{i+s},\alpha_i,... \end{array}\right\} s \quad - \quad \text{транспозицій}$$

сусідніх.

Всього транспозицій сусідніх $2s+1$, а так як кожна транспозиція сусідніх змінює парність, то парна кількість транспозицій сусідніх не змінюють парність, а $(2s+1)$ таких транспозицій змінюють парність, що й треба було довести. ■

Твердження 2.3.3. *При $n \geq 2$ число парних перестановок з n елементів дорівнює числу непарних, тобто* $\frac{1}{2}n!$

Кожне взаємно однозначне відображення множини перших n натуральних чисел на себе називається ***підстановкою*** ***n-го степеня***. Множина всіх підстановок n-го степеня називається ***симетричною групою*** ***n-го порядку*** і позначається S_n.

Як записуються підстановки? Підстановка n-го степеня α може бути записана у наступному вигляді

$\alpha = \begin{pmatrix} 1 & 2 & ... & n \\ \alpha_1 & \alpha_2 & ... & \alpha_n \end{pmatrix}$, де $\alpha_1,\alpha_2,...,\alpha_n$ – різні натуральні числа від 1 до n, якщо $\alpha(i)=\alpha_i$ $(\alpha: \quad i \to \alpha_i)$.

Можна записати $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$, якщо α – підстановка 4-го степеня і $\alpha(1) = 2$, $\alpha(2) = 3$, $\alpha(3) = 1$, $\alpha(4) = 4$. Нехай $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$ – інша підстановки 4-го степеня, а підстановка γ записана у вигляді $\gamma = \begin{pmatrix} 3 & 4 & 1 & 2 \\ 2 & 3 & 1 & 4 \end{pmatrix}$. Чи дорівнює вона α або β?

$\left.\begin{array}{l} \gamma(3) = 2 \\ \alpha(3) = 1 \neq 2 \end{array}\right\}$, тому, за означенням рівності відображень, $\alpha \neq \gamma$.

А от $\beta = \gamma$,бо

$$\gamma(3) = 2 = \beta(3),\ \gamma(4) = 3 = \beta(4),\ \gamma(1) = 1 = \beta(1),\ \gamma(2) = 4 = \beta(2),$$

хоча на перший погляд вони відрізняються.

Якщо верхні рядки двох різних підстановок n-го степеня впорядкувати по зростанню, то вони відрізняються лише нижніми рядками, тобто перестановками n-го порядку. Тому елементів в симетричній групі n-го степеня стільки, скільки різних перестановок n-го порядку, тобто $n!$.

Кількістю інверсій $\nu(\alpha)$ підстановки α називається сума кількостей інверсій її верхнього і нижнього рядків.

$$\alpha = \begin{pmatrix} \alpha_1\alpha_2...\alpha_n \\ \beta_1\beta_2...\beta_n \end{pmatrix},\ \nu(\alpha) = \nu(\alpha_1,\alpha_2,...,\alpha_n) + \nu(\beta_1,\beta_2,...,\beta_n).$$

<u>Приклад</u>. Знайти кількість інверсій підстановки $\alpha = \begin{pmatrix} 2 & 5 & 1 & 3 & 4 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix}$.

▼ $\nu(\alpha) = (1+3+0+0+0) + (2+1+2+0+0) = 9$. ▲

$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 4 & 3 & 7 & 6 & 2 & 1 \end{pmatrix}$ – підстановка 7-го степеня.

Підстановка, яка має парну (непарну) кількість інверсій називається ***парною*** **(*непарною*).** (α – непарна підстановка)

Зауважимо, що при різних записах однієї й тієї ж підстановки кількість інверсій може змінюватись, але не змінюється парність, тобто не змінюється величина $(-1)^{\nu(\alpha)}$ (за рахунок того, що кожна транспозиція виконується в двох рядках – верхньому і нижньому одночасно).

Наприклад, підстановка із попереднього прикладу $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix}$,

$\nu(\alpha) = 0 + (4+2+0+1+0) = 7 \neq 9$, але $(-1)^9 = (-1)^7$.

Таким чином, *кількість парних та непарних підстановок* така ж, як і у відповідних перестановок, а саме $\frac{1}{2}n!$.

Розглянемо квадратну матрицю $A = (a_{ij}) \in M_n$:

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

Із елементів цієї матриці складемо вираз $a_{1\alpha_1} a_{2\alpha_2} ... a_{n\alpha_n}$, який є добутком елементів, взятих по одному і тільки одному із кожного рядка і кожного стовпчика матриці A. Тоді набір номерів стовпчиків $\alpha_1, \alpha_2, ..., \alpha_n$ будуть утворювати перестановку із n елементів, а кожному добутку $a_{1\alpha_1} a_{2\alpha_2} ... a_{n\alpha_n}$ можна поставити у відповідність підстановку індексів його елементів $\alpha = \begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$. Нагадаємо, що число інверсій підстановки дорівнює сумі інверсій її верхнього і нижнього рядків. У даному випадку число інверсій $\nu(\alpha)$ дорівнює числу інверсій перестановки ($\alpha_1, \alpha_2, ..., \alpha_n$).

Визначником квадратної матриці n-го порядку називається алгебраїчна сума $n!$ доданків $(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} ... a_{n\alpha_n}$, які складаються із добутків її елементів, взятих по одному і тільки одному із кожного рядка і кожного стовпчика, причому знак кожного такого доданка визначається множником $(-1)^{\nu}$, де $\nu = \nu(\alpha)$ – число інверсій у підстановці α індексів його елементів.

Визначник матриці $A = (a_{ij}) \in M_n$ позначається символами:

$$\Delta,\ |A|,\ \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n1} & \dots & a_{nn} \end{vmatrix},\ \left\| a_{ij} \right\|,\ \det A.$$

З огляду на означення визначника, можна писати

$$\Delta = |A| = \sum_{\alpha \in S_n} (-1)^{\nu(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} = \sum_{n!} (-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}$$

Також будемо користуватися записом $(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} ... a_{n\alpha_n} \in \Delta$, якщо $(-1)^{\nu} a_{1\alpha_1} a_{2\alpha_2} ... a_{n\alpha_n}$ є одним із доданків визначника Δ $(\nu = \nu(\alpha))$.

Домовимося надалі визначник матриці n-го порядку називати *визначником* n-го порядку, а рядки і стовпчики матриці називати також *рядками і стовпчиками визначника*.

З означення випливає, що

1. Якщо $A=(a_{11})$, то $|A|=a_{11}$.

2. Якщо $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, то $|A|=a_{11}a_{22}-a_{12}a_{21}$.

3. Якщо $A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, то $|A|=(-1)^0 a_{11}a_{22}a_{33}+(-1)^1 a_{11}a_{23}a_{32}+$

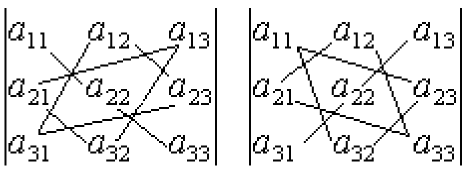
$+(-1)^1 a_{12}a_{21}a_{33}+(-1)^2 a_{12}a_{23}a_{31}+(-1)^3 a_{13}a_{22}a_{31}+(-1)^2 a_{13}a_{21}a_{32}=$

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

<u>Приклад.</u> $\left|-2\right|=-2$. ▲

<u>Приклад.</u> $\begin{vmatrix} 3 & 2 \\ 5 & 7 \end{vmatrix}=3\cdot 7-5\cdot 2=11$. ▲

Для простоти обчислення визначника 3-го порядку існує *правило трикутників*, яке продемонстроване на наступній схемі:



«+» «–»

Всі множники добутків, які входять до визначника із знаком «+» знаходяться на головній діагоналі та трикутниках із стороною, паралельною їй; всі множники добутків, які входять до визначника із знаком «–», знаходяться на бічній діагоналі та трикутниках із стороною, паралельною їй.

<u>Приклад.</u>

$$\begin{vmatrix} 1 & 2 & -3 \\ 7 & 5 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 1\cdot 5\cdot 2+7\cdot 0\cdot(-3)+2\cdot 4\cdot(-1)-(-3)\cdot 5\cdot(-1)-0\cdot 4\cdot 1-7\cdot 2\cdot 2= \quad ▲$$
$$=10+0-8-15-0-28=-41.$$

В л а с т и в о с т і в и з н а ч н и к і в.

При доведенні цих властивостей будемо користуватися позначеннями $A=(a_{ij}),\ B=(b_{ij}),\ A'=(a'_{ij}),\ \ |A|=\sum_{\alpha\in S_n}(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}\ldots a_{n\alpha_n}$, де

$\alpha=\begin{pmatrix}1 & 2 & \ldots & n\\ \alpha_1 & \alpha_2 & \ldots & \alpha_n\end{pmatrix}$.

1. *Визначник при транспонуванні матриці не змінюється.*

Звідси випливає, що *властивості визначників, які доведені для рядків, мають місце і для стовпчиків.*

Д о в е д е н н я.. Нехай $B=A^T$, тоді $b_{ij}=a_{ji}$ і кожен добуток $(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}...a_{n\alpha_n}=(-1)^{\nu(\alpha)}b_{\alpha_1 1}b_{\alpha_2 2}...b_{\alpha_n n},\ \alpha\in S_n$. Так як всі множники такого добутку також знаходяться в різних рядках і різних стовпчиках матриці B, то за означенням визначника виконується рівність

$$|B|=\sum_{\beta\in S_n}(-1)^{\nu(\beta)}b_{\alpha_1 1}b_{\alpha_2 2}\ldots b_{\alpha_n n},\ \text{де}\ \beta=\begin{pmatrix}\alpha_1 & \alpha_2 & \ldots & \alpha_n\\ 1 & 2 & \ldots & n\end{pmatrix}.$$

Враховуючи поняття кількості інверсій підстановки, маємо, що $\nu(\beta)=\nu(\alpha)$ і $\left|A^T\right|=\sum_{\beta\in S_n}(-1)^{\nu(\beta)}b_{\alpha_1 1}b_{\alpha_2 2}\ldots b_{\alpha_n n}=\sum_{\alpha\in S_n}(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}...a_{n\alpha_n}=|A|$. ■

2. *Якщо один з рядків визначника складається з нулів, то визначник дорівнює нулю.*

Д о в е д е н н я. Нехай k-й рядок $|A|$ нульовий. Тоді $a_{kj}=0,\ j=1,2,...,n.$ Так як кожен член визначника містить $a_{k\alpha_k}=0$, то $|A|=\sum_{\alpha\in S_n}(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}...\alpha_{k\alpha_k}...a_{n\alpha_n}=0$. ■

3. *Якщо два рядки визначника поміняти місцями, то знак визначника зміниться на протилежний, а його абсолютна величина не зміниться.* **(Перше елементарне перетворення змінює знак визначника на протилежний.)**

Д о в е д е н н я. Нехай у матриці A переставлені місцями s-й та k-й рядки і одержана матриця B. Тоді $b_{sj}=a_{kj}$, $b_{kj}=a_{sj}$ $b_{ij}=a_{ij},\ i\neq s,k,\ j=1,2,...,n$

$$|B|=\sum_{\beta\in S_n}(-1)^{\nu(\beta)}b_{1\alpha_1}b_{2\alpha_2}\ldots b_{s\alpha_s}\ldots b_{k\alpha_k}\ldots b_{n\alpha_n},$$

$\beta=\begin{pmatrix}1 & \ldots & s & \ldots & k & \ldots & n\\ \alpha_1 & \ldots & \alpha_s & \ldots & \alpha_k & \ldots & \alpha_n\end{pmatrix}$, а значить

$$|B|=\sum_{\beta\in S_n}(-1)^{\nu(\beta)}a_{1\alpha_1}a_{2\alpha_2}\dots a_{k\alpha_s}\dots a_{s\alpha_k}\dots a_{n\alpha_n}\,.\ (1)$$

Якщо $\alpha=\begin{pmatrix}1 & \dots & k & \dots & s & \dots & n\\ \alpha_1 & \dots & \alpha_s & \dots & \alpha_k & \dots & \alpha_n\end{pmatrix}$, то підстановки α і β відрізняються однією транспозицією у верхньому рядку, тоді за теоремою 2.3.2 вони мають різну парність, а саме $(-1)^{\nu(\beta)}=-(-1)^{\nu(\alpha)}$. Підставивши це в (1), маємо $|B|=\sum\limits_{\alpha\in S_n}-(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}\dots a_{k\alpha_s}\dots a_{s\alpha_k}\dots a_{n\alpha_n}=-|A|$. ■

4. *Визначник, що має два однакових рядки, дорівнює нулю.*

Д о в е д е н н я. Нехай у матриці A однакові s-й та k-й рядки, тобто $a_{sj}=a_{kj}$, $j=1,2,...n$. Якщо поміняти місцями s-й та k-й рядки, то, з одного боку, одержана матриця B не буде відрізнятися від матриці A, звідки $|B|=|A|$, а з іншого боку (за попередньою властивістю), $|B|=-|A|$, що може виконуватись тільки при $|A|=0$. ■

5. *Якщо всі елементи довільного рядка визначника помножити на деяке число λ, то визначник помножиться на це число.* **(Друге елементарне перетворення множить визначник на відповідне число).**

Д о в е д е н н я. Нехай у матриці A на λ помножено k-й рядок і одержано матрицю B. Тоді $b_{kj}=\lambda a_{kj}$, $b_{ij}=a_{ij}, i\neq k,\ j=1,2,...,n$. Отже

$$|B|=\sum_{\alpha\in S_n}(-1)^{\nu(\alpha)}b_{1\alpha_1}b_{2\alpha_2}\dots b_{k\alpha_k}\dots b_{n\alpha_n}=\sum_{\alpha\in S_n}(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}\dots\lambda a_{k\alpha_k}\dots a_{n\alpha_n}=$$
$$=\lambda\sum_{\alpha\in S_n}(-1)^{\nu(\alpha)}a_{1\alpha_1}a_{2\alpha_2}\dots a_{k\alpha_k}\dots a_{n\alpha_n}=\lambda|A|.\ ■$$

6. *Визначник, що має два пропорційні рядки, дорівнює нулю.*

Д о в е д е н н я. Нехай у матриці A пропорційні s-й та k-й рядки, тобто $a_{sj}=\lambda a_{kj}$, $j=1,2,...n$. Тоді за попередньою властивістю $|A|=\lambda|A'|$, де $a'_{sj}=a_{kj}$, $a'_{ij}=a_{ij}, i\neq s,\ i,j=1,2,...n$. Отже матриця A' має однакові s-й та k-й рядки, тоді за властивістю 4 $|A|=\lambda|A'|=\lambda 0=0$. ■

7. *Якщо елементи s-го рядка визначника $|A|$ є сумами двох доданків, то цей визначник можна подати як суму двох визначників $|A|=\Delta_1+\Delta_2$, де Δ_1 і Δ_2 є визначниками, які утворено із $|A|$ заміною елементів s-го рядка відповідно першими та другими доданками цих елементів.*

Д о в е д е н н я. Нехай елементи s-го рядка $a_{sj} = b_j + c_j$, $j = 1,2,...,n$. Тоді $|A| = \sum_{\alpha \in S_n} (-1)^{\nu(\alpha)} a_{1\alpha_1} a_{2\alpha_2} ... a_{s\alpha_s} ... a_{n\alpha_n} = \sum_{\alpha \in S_n} (-1)^{\nu(\alpha)} a_{1\alpha_1} a_{2\alpha_2} ... (b_{\alpha_s} + c_{\alpha_s}) ... a_{n\alpha_n} =$
$= \sum_{\alpha \in S_n} (-1)^{\nu(\alpha)} a_{1\alpha_1} a_{2\alpha_2} .. b_{\alpha_s} ... a_{n\alpha_n} + \sum_{\alpha \in S_n} (-1)^{\nu(\alpha)} a_{1\alpha_1} a_{2\alpha_2} .. c_{\alpha_s} ... a_{n\alpha_n} = \Delta_1 + \Delta_2$. ■

Ця властивість поширюється на довільну кількість доданків.

8. *Якщо один з рядків визначника є лінійною комбінацією інших рядків, то визначник дорівнює нулю.*

Поняття лінійної комбінації визначається в п. 2.9, а доведення цієї властивості випливає із властивостей 7 і 6.

9. *Якщо до одного рядка визначника додати інший, помножений на якесь число, то визначник матриці не зміниться.* **(Третє елементарне перетворення не змінює визначника.)**

Д о в е д е н н я. Нехай до s-го рядка матриці A додали k-й рядок, помножений на число λ і одержали матрицю B. Тоді $b_{sj} = a_{sj} + \lambda a_{kj}$, $b_{ij} = a_{ij}, i \neq s$, $j = 1,2,...,n$. За властивістю 7 визначник $|B|$ дорівнює сумі визначника $|A|$ і визначника з пропорційними s-м і k-м рядками, який дорівнює 0 за властивістю 6. Отже $|B| = |A|$. ■

Наслідок. *Якщо до одного рядка визначника додати лінійну комбінацію інших рядків, то визначник не зміниться.*

Ці властивості використовуються при дослідженні визначників та дозволяють обчислювати деякі визначники n-го порядку, не розгортаючи їх у суму $n!$ доданків.

Контрольні питання та завдання

1. Що таке перестановка? Що таке кількість інверсій перестановки? Яка перестановка називається парною?
2. Визначте парність перестановки (2, 5, 8, 3, 4, 7, 1, 6).
3. Що відбувається з парністю перестановки при транспозиції?
4. Що таке підстановка? У якому вигляді записуються підстановки?
5. Яке значення має $\alpha(3)$, якщо $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$?
6. Чи будуть рівними підстановки α і β, якщо $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$, $\beta = \begin{pmatrix} 3 & 1 & 4 & 2 \\ 4 & 2 & 1 & 3 \end{pmatrix}$?

7. Що таке кількість інверсій підстановки?

8. Визначте парність підстановки $\alpha = \begin{pmatrix} 1 & 4 & 3 & 5 & 2 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}$.

9. Із скількох доданків складається визначник

а) 2-го порядку? б) 3-го порядку? в) 6-го порядку?

10. Назвіть елементи головної та бічної діагоналей у визначнику матриці $\left(a_{ij}\right)_4^4$.

11. Відомо, що $|A| = 3$. Чому дорівнює $\left|A^T\right|$?

12. Відомо, що $A \in M_5$ і $|A| = 3$. Чому дорівнює $|2A|$?

13. Чи зміниться визначник, якщо до першого стовпчика визначника додати відповідні елементи всіх останніх стовпчиків?

2.4. Мінори. Визначник добутку

Для обчислення визначників n-го порядку при $n \geq 4$ не існує простого для запам'ятовування правила. Обчислювати їх за означенням стає складно, бо збільшується число доданків в сумі, про яку йде мова в означенні: при $n = 4$ число доданків дорівнює $4! = 24$, при $n = 5$ – 5!=120. Тому в таких випадках застосовують способи обчислення визначників, які використовують поняття мінору та алгебраїчного доповнення.

Мінором k-го порядку $M = M_{i_1, i_2, \dots, i_k}^{j_1, j_2, \dots, j_k}$ визначника Δ називається визначник, утворений елементами, що стоять на перетині $i_1, i_2, \dots, i_k$ рядків і $j_1, j_2, \dots, j_k$ стовпчиків. Цей визначник має порядок $k < n$.(У випадку $k = n$ мінор співпадає з визначником). Визначник матриці, що утворюється після викреслювання $i_1, i_2, \dots, i_k$ рядків і $j_1, j_2, \dots, j_k$ стовпчиків, називається ***доповняльним мінором*** до мінору $M_{i_1, i_2, \dots, i_k}^{j_1, j_2, \dots, j_k}$ і позначається $M' = M'^{j_1, j_2, \dots, j_k}_{i_1, i_2, \dots i_k}$. Він має порядок $(n-k)$. Зауважимо, що M' є мінором $(n-k)$-го порядку, який утворюється елементами, що стоять на перетині тих рядків і стовпчиків, що не дорівнюють $i_1, i_2, \dots, i_k$, $j_1, j_2, \dots, j_k$, а його доповняльний мінор дорівнює M, тобто M і M' взаємно доповняльні мінори.

Позначимо $s_M = (i_1 + i_2 + \dots + i_k) + (j_1 + j_2 + \dots + j_k)$.

Величину $\overline{M} = (-1)^{s_M} M'$ називають ***алгебраїчним доповненням мінору*** M.

Приклад. Нехай дано визначник

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

Тоді мінор $M = M_{1,2,4}^{1,4,5} = \begin{vmatrix} 3 & 2 & 4 \\ 6 & 5 & 2 \\ 0 & 0 & 0 \end{vmatrix} = 0$, його доповняльний мінор $M' = M'^{1,4,5}_{1,2,4} = \begin{vmatrix} 1 & -2 \\ 4 & -9 \end{vmatrix} = -1$, при цьому $S_M = 1+2+4+1+4+5 = 17$, тому його алгебраїчне доповнення $\overline{M} = \overline{M}_{1,2,4}^{1,4,5} = (-1)^{S_M} M'^{1,4,5}_{1,2,4} = (-1)^{17} \begin{vmatrix} 1 & -2 \\ 4 & -9 \end{vmatrix} = 1$. ▲

Теорема 2.4.1. (Про добуток мінору на його алгебраїчне доповнення) *Добуток довільного мінору на його алгебраїчне доповнення у визначнику* Δ *є алгебраїчна сума, доданки якої будуть деякими членами визначника* Δ*, причому їх знаки в цій сумі співпадають з тими знаками, з якими вони входять в склад визначника.*

Д о в е д е н н я. Нехай визначник Δ n-го порядку, а мінор M k-го порядку ($k < n$). Тоді необхідно довести, що $M\overline{M} \in \Delta$.

1. Розглянемо випадок, коли мінор M розташований у верхньому лівому кутку визначника Δ, тобто $M = M_{1,2,\dots,k}^{1,2,\dots,k}$, тоді мінор $M' = M_{(k+1),\dots,n}^{(k+1),\dots,n}$ і кожний його s-й рядок або стовпчик має номер $k+s$ у визначнику Δ,

$$S_M = 1+2+\ldots+k+1+2+\ldots+k = 2(1+2+\ldots+n), \qquad \overline{M} = (-1)^{S_M} M' = M',$$

$$M \cdot \overline{M} = M \cdot M' = \sum_{k!} (-1)^{\nu_1} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k\alpha_k} \cdot \sum_{(n-k)!} (-1)^{\nu_2} a_{(k+1)\beta_1} a_{(k+2)\beta_2} \ldots a_{n\beta_{(n-k)}} =$$

$$= \sum_{k!(n-k)!} (-1)^{\nu_1+\nu_2} a_{1\alpha_1} \cdot \ldots \cdot a_{k\alpha_k} \cdot a_{(k+1)\beta_1} \cdot \ldots \cdot a_{n\beta_{(n-k)}},$$

де $\{\alpha_1, \dots, \alpha_k\} = \{1, \dots, k\}$, $\nu_1 = \nu \begin{pmatrix} 1 & 2 & \dots & k \\ \alpha_1 & \alpha_2 & \dots & \alpha_k \end{pmatrix}$,

$\{\beta_1, \dots, \beta_{n-k}\} = \{k+1, \dots, n\}$, $\nu_2 = \nu \begin{pmatrix} k+1 & k+2 & \dots & n \\ \beta_1 & \beta_2 & \dots & \beta_{n-k} \end{pmatrix}$.

Таким чином, кожен доданок виразу $M \cdot \overline{M}$ буде членом визначника Δ, якщо $(-1)^{\nu_1+\nu_2} = (-1)^{\nu}$, де $\nu = \nu(p)$, $p = \begin{pmatrix} 1 & \dots & k & k+1 & \dots & n \\ \alpha_1 & \dots & \alpha_k & \beta_1 & \dots & \beta_{n-k} \end{pmatrix}$.

Обчислимо кількість інверсій підстановки p:

1) елементи верхнього рядка утворюють нуль інверсій;

2) елементи $\alpha_i, i = 1,\dots,k$ утворюють ν_1 інверсій;

3) елементи $\beta_j, j = 1,\dots,n-k$ утворюють ν_2 інверсій;

4) $\forall i, j : \alpha_i < \beta_j$, причому кожне α_i стоїть на перших k місцях, а кожне β_j стоїть на останніх місцях $k+1, k+2, ..., n$, тому всі α_i та β_j утворюють нуль інверсій.

Отже $\nu = \nu(p) = \nu_1 + \nu_2$ і $(-1)^{\nu} = (-1)^{\nu_1+\nu_2}$,що доводить теорему у випадку 1 $\left(M \cdot \overline{M} \in \Delta\right)$.

2. Розглянемо загальний випадок:

$M = M_{i_1 i_2 \dots i_k}^{j_1 j_2 \dots j_k}$, $\overline{M} = (-1)^{S_M} \cdot M'$, де $S_M = i_1 + i_2 + \dots + i_k + j_1 + j_2 + \dots + j_k$.

Поступовими перестановками рядків і стовпчиків у визначнику Δ перенесемо мінор M у лівий верхній куток (зведемо до випадку 1). Для цього будемо виконувати такі перестановки рядків $i_1, i_2, ..., i_k$:

$$\left.\begin{matrix} i_1 \leftrightarrow i_1 - 1 \\ i_1 \leftrightarrow i_1 - 2 \\ \dots\dots\dots \\ i_1 \leftrightarrow 1 \end{matrix}\right\} \begin{matrix}(i_1 - 1) \\ \text{перестановок,}\end{matrix} \quad \left.\begin{matrix} i_2 \leftrightarrow i_2 - 1 \\ \dots\dots \\ i_2 \leftrightarrow i_1 + 1 \\ i_2 \leftrightarrow i_1 - 1 \\ \dots\dots \\ i_2 \leftrightarrow 1 \end{matrix}\right\} \begin{matrix}(i_2 - 2) \\ \text{перестановок,}\end{matrix} \quad \dots \quad \left.\begin{matrix} i_k \leftrightarrow i_k - 1 \\ \dots\dots\dots \\ i_k \leftrightarrow i_{k-1} + 1 \\ i_k \leftrightarrow i_{k-1} - 1 \\ \dots\dots\dots \\ i_k \leftrightarrow i_1 + 1 \\ i_k \leftrightarrow i_1 - 1 \\ \dots\dots \\ i_k \leftrightarrow 1 \end{matrix}\right\} \begin{matrix}(i_k - k) \\ \text{перестановок.}\end{matrix}$$

Тобто всього виконаємо

$$\left(i_1 - 1\right) + \left(i_2 - 2\right) + \dots + \left(i_k - k\right) = i_1 + i_2 + \dots i_k - (1 + 2 + \dots + k)$$

перестановок рядків.

Аналогічно переставимо стовпчики $j_1, j_2, ..., j_k$ в перші k стовпчиків і одержимо $j_1 + j_2 + \dots + j_k - (1 + 2 + \dots + k)$ перестановок стовпчиків. В результаті цього матимемо бажаний визначник Δ' з мінором M у верхньому лівому куті, а одержаний він буде перестановкою

$$i_1 + i_2 + \dots + i_k + j_1 + j_2 + \dots + j_k - 2(1 + 2 + \dots + k)$$

рядків та стовпчиків. Тоді за доведеним у випадку 1 маємо $M \cdot M' \in \Delta'$. (1)

Так як визначник Δ' одержано поступовою перестановкою $S_M - 2(1+2+...+k)$ рядків та стовпчиків, а за властивістю визначників кожна така перестановка змінює знак визначника на протилежний, то

$\Delta' = (-1)^{S_M - 2(1+2+...+k)} \cdot \Delta = (-1)^{S_M} \cdot \Delta$ і

$\Delta = (-1)^{2S_M} \cdot \Delta = (-1)^{S_M} \cdot \Delta'$ (2)

Врахуємо (1) та (2) і одержимо:

$M \cdot \overline{M} = M \cdot (-1)^{S_M} \cdot M' = (-1)^{S_M} \cdot M \cdot M' \in (-1)^{S_M} \cdot \Delta' = \Delta.$ ■

Зауваження. Всі доданки $M \cdot \overline{M}$ – це *різні* члени визначника Δ (бо визначники M і M' розташовані в різних рядках і різних стовпчиках) і їх *кількість* $k!(n-k)!$, бо кожний мінор $M = M_{i_1 i_2 \dots i_k}^{j_1 j_2 \dots j_k}$ має порядок k і складається з $k!$ доданків, а доповняльний мінор M', а значить і алгебраїчне доповнення $\overline{M}$, із $(n-k)!$ доданків.

Теорема 2.4.2. (Лапласа) *Нехай у визначнику Δ порядку n вибрано k рядків. Тоді сума попарних добутків всіх мінорів k-го порядку, розміщених на цих рядках, на їх алгебраїчні доповнення дорівнює визначнику Δ.*

Д о в е д е н н я. Нехай у визначнику Δ вибрано $i_1, i_2, ..., i_k$ рядки. Тоді треба довести, що $\Delta = \sum_{(j)} M_{i_1, i_2, ..., i_k}^{(j)} \cdot \overline{M}_{i_1, i_2, ..., i_k}^{(j)}$, де (j) – всі можливі набори номерів по k стовпчиків із n. Таких наборів $(j) = (j_1, j_2, ..., j_k)$, де $j_1 < j_2 < ... < j_k$, а значить і мінорів $M = M_{i_1, i_2, ..., i_k}^{(j)}$ буде $C_n^k = \dfrac{n!}{k!(n-k)!}$ штук. Із теореми 2.4.1 $M \cdot \overline{M} \in \Delta$, а із зауваження до цієї теореми кожен добуток $M \cdot \overline{M}$ складається із $k!(n-k)!$ різних доданків, тоді сума $\sum_{(j)} M_{i_1, i_2, ..., i_k}^{(j)} \cdot \overline{M}_{i_1, i_2, ..., i_k}^{(j)}$ усіх можливих таких добутків складається із $C_n^k \cdot k!(n-k)! = n!$ різних доданків визначника Δ, бо кожний набір (j) визначає різний набір стовпчиків в $M_{i_1, i_2, ..., i_k}^{(j)}$ та $M'^{(j)}_{i_1, i_2, ..., i_k}$. Оскільки визначник Δ складається всього із $n!$ доданків, то $\Delta = \sum_{(j)} M_{i_1, i_2, ..., i_k}^{(j)} \cdot \overline{M}_{i_1, i_2, ..., i_k}^{(j)}$. ■

Наслідок 2.4.3. *Нехай матриці A, B, C мають відповідно розмірності $m \times m, k \times m, k \times k$. Тоді* $\begin{vmatrix} A & O \\ B & C \end{vmatrix} = |A| \cdot |C|$.

Д о в е д е н н я. Позначимо $M = M_{1,2,...,m}^{1,2,...,m}$. Тоді $M = |A|$, $M' = |C|$, а $\overline{M} = (-1)^{2(1+2+...+m)} \cdot |C| = |C|$. Так як $M_{1,2,...,m}^{j_1, j_2, ..., j_m} = 0$, якщо $\{j_1, j_2, ..., j_m\} \neq \{1, ..., m\}$ (як мінор що містить нульовий стовпчик), то за теоремою 2.4.2

$$\begin{vmatrix} A & O \\ B & C \end{vmatrix} = \sum_{(j)} M_{1,2,...,m}^{(j)} \cdot \overline{M}_{1,2,...,m}^{(j)} = |A| \cdot |C|. \blacksquare$$

Наслідок 2.4.4.. *Нехай матриці* A, B, C *мають відповідно розмірності* $k \times k, m \times m, k \times m$. *Тоді* $\begin{vmatrix} O & B \\ A & C \end{vmatrix} = (-1)^{mk} |A| \cdot |B|$.

Д о в е д е н н я. Нехай $M = M_{1,2,...,m}^{k+1, k+2, ..., k+m}$, тоді $M = |B|$, $M' = |A|$, $\overline{M} = (-1)^{1+2+...+m+(k+1)+(k+2)+...+(k+m)} \cdot |A| = (-1)^{2(1+2+...+m)+mk} \cdot |A| = (-1)^{mk} \cdot |A|$ Так як $M_{1,2,...,m}^{j_1, j_2, ..., j_m} = 0$, якщо $\{j_1, j_2, ..., j_m\} \neq \{k+1, k+2, ..., k+m\}$ (як мінор що містить нульовий стовпчик), то за теоремою 2.4.2

$$\begin{vmatrix} O & B \\ A & C \end{vmatrix} = \sum_{(j)} M_{1,2,...,m}^{(j)} \cdot \overline{M}_{1,2,...,m}^{(j)} = (-1)^{mk} |A| \cdot |B|. \blacksquare$$

Якщо $|A| = \|a_{ij}\|$, то мінор *першого порядку* $M_s^k = a_{sk}$. Алгебраїчне доповнення такого мінору позначають $A_{sk} = (-1)^{s+k} M_s'^k$ і називають ***алгебраїчним доповненням елементу*** a_{sk}.

Наслідок 2.4.5. (Про розкладення визначника за елементами рядка або стовпця) *Визначник* $\Delta = \|a_{ij}\|$ n*-го порядку дорівнює сумі всіх добутків елементів* i*-го рядка, або* k*-го стовпця на їх алгебраїчні доповнення, тобто*

$\Delta = a_{i1} A_{i1} + a_{i2} A_{i2} + ... + a_{in} A_{in}$, (2.4.3)

або

$\Delta = a_{1k} A_{1k} + a_{2k} A_{2k} + ... + a_{nk} A_{nk}$. (2.4.4)

Формула (2.4.3) називається ***розкладом визначника за елементами*** i***-го рядка***, а формула (2.4.4) – ***розкладом за елементами*** k***-го стовпця***.

Д о в е д е н н я. Якщо теорему 2.4.2 розглянути для $k = 1$ (вибрати один рядок i), то $M_i^j = a_{ij}$, $\overline{M}_i^j = A_{ij}$, а $\Delta = \sum_{j=1}^{n} M_i^j \cdot \overline{M}_i^j = \sum_{j=1}^{n} a_{ij} \cdot A_{ij}$, що доводить формулу (2.4.3).

Наслідок із першої властивості визначників забезпечує доведення формули (2.4.4). $\blacksquare$

Матриця, у якої всі елементи, що стоять з одного боку від головної діагоналі, дорівнюють нулю, називається ***трикутною матрицею***. *Зауважимо*, що послідовне використання розкладення визначника трикутної матриці $\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}$ по першому стовпцю дозволяє зробити висновок, що цей визначник $\Delta = a_{11}a_{22}\dots a_{nn}$:

$\Delta = a_{11}A_{11} + 0 \cdot A_{21} + \dots + 0 \cdot A_{n1} = a_{11}A_{11}$, де A_{11} – визначник трикутного вигляду $(n-1)$-го порядку і т.д.

Оскільки одинична матриця $E \in M_n(R)$ також має трикутний вигляд, то $|E| = 1^n = 1$. Аналогічно $|-E| = (-1)^n$, якщо $-E \in M_n(R)$.

Нехай $\Delta = \|a_{ij}\|$. При $k \neq s$ сума $a_{k1}A_{s1} + a_{k2}A_{s2} + \dots + a_{kn}A_{sn}$ називається ***розкладенням визначника*** Δ ***по чужому рядку***, а сума $a_{1k}A_{1s} + a_{2k}A_{2s} + \dots + a_{nk}A_{ns}$ – ***розкладенням по чужому стовпчику***.

Теорема 2.4.6. *Розкладення визначника по чужому рядку (стовпчику) дорівнює нулю.*

Д о в е д е н н я. Нехай у визначнику $\Delta = \|a_{ij}\|$ однакові k-й та s-й рядки ($k \neq s$), тобто $a_{sj} = a_{kj}$, $j = 1,2,\dots n$. Якщо розкласти Δ за формулою (2.4.3) по s-му рядку, то

$$\Delta = a_{s1}A_{s1} + a_{s2}A_{s2} + \dots + a_{sn}A_{sn} = a_{k1}A_{s1} + a_{k2}A_{s2} + \dots + a_{kn}A_{sn}.$$

З іншого боку, за властивістю 4 визначників $\Delta = 0$. Отже маємо

$$a_{k1}A_{s1} + a_{k2}A_{s2} + \dots + a_{kn}A_{sn} = 0.$$

Для розкладу по чужому стовпчику доведення аналогічне. ■

Використовуючи властивості визначників та мінорів, доводиться важлива

Теорема 2.4.7. *Визначник добутку двох квадратних матриць дорівнює добутку визначників цих матриць:* $|AB| = |A| \cdot |B|, \ \forall A, B \in M_n(\mathbf{R})$.

Квадратна матриця називається ***виродженою (невиродженою)***, якщо її визначник дорівнює (не дорівнює) нулю. Із теореми 2.4.7 випливає

Наслідок 2.4.8. *Добуток невироджених матриць є невироджена матриця.*

Контрольні питання та завдання

1. Поясніть, що таке мінор матриці, доповняльний мінор, алгебраїчне доповнення мінору. Чому дорівнює доповняльний мінор елементу a_{12} матриці $\left(a_{ij}\right)_2^2$? Чому дорівнює алгебраїчне доповнення елементу a_{12} матриці $\left(a_{ij}\right)_2^2$?

2. Як пов'язані між собою доповняльний мінор і алгебраїчне доповнення?

3. Що означає розкласти визначник за елементами i-го рядка?

4. Чому дорівнює визначник трикутної матриці?

5. Яка матриця називається невиродженою? Наведіть приклад невиродженої матриці.

6. Із трьох матриць A, B і C n-го порядку одна вироджена. Чому дорівнює $|ABC|$?

2.5. Обернена матриця

Нагадаємо, що матриця A^{-1} називається ***оберненою*** до A, якщо виконуються наступні рівності: $AA^{-1} = A^{-1}A = E$. Розглянемо детальніше ***в л а с т и в о с т і о б е р н е н о ї м а т р и ц і*** та знайдемо формулу для її обчислення.

1. Оскільки рівність $AA^{-1} = A^{-1}A$ може виконуватись тільки для квадратних матриць, то *обернена матриця A^{-1} може існувати тільки при* $A \in M_n(\mathbf{R})$, причому і сама $A^{-1} \in M_n(\mathbf{R})$.

2. Нехай $A \in M_n(\mathbf{R})$, нехай $\exists\, A^{-1} \in M_n(\mathbf{R})$. Оскільки $AA^{-1} = E$, то за теоремою 2.4.7 $\left|A \cdot A^{-1}\right| = |A| \cdot \left|A^{-1}\right|$, а з іншого боку $\left|A \cdot A^{-1}\right| = |E| = 1$, отже $|A| \cdot \left|A^{-1}\right| = 1$, звідки

1) $\left|A^{-1}\right| = 1/|A|$ – *визначник матриці, оберненої до матриці* A, *дорівнює числу, оберненому до визначника матриці* A;

2) *матриця, визначник якої дорівнює нулю, не має оберненої*, бо $0 \cdot \left|A^{-1}\right| \neq 1$

3. Якщо A^{-1} *існує, то вона єдина для довільної матриці* A.

Д о в е д е м о від супротивного. Нехай матриця X обернена до A. Тоді $AX = E$, а також $X = EX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}E = A^{-1}$. Отже

$X = A^{-1}$. ■

4. *Кожна квадратна матриця* $A \in M_n(\mathbf{R})$*, визначник якої не дорівнює нулю (*$|A| \neq 0$*,) має обернену* A^{-1}*, яка обчислюється за формулою:*

$$A^{-1} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & & \dots & \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \ (2.5.1)$$

де A_{ij} *- алгебраїчні доповнення до елементів* a_{ij} *матриці* A*.*

Д о в е д е м о вірність формули (2.5.1). Нехай $A^{-1} \cdot A = C = (c_{ij})_n$. Тоді знайдемо елементи матриці C за правилом множення матриць (2.1.2):

$c_{11} = \frac{1}{|A|}(A_{11}a_{11} + A_{21}a_{21} + \ldots + A_{n1}a_{n1}) = \frac{1}{|A|} \cdot |A| = 1$ – за наслідком 2.4.5 про розкладення визначника по стовпчику,

$c_{1j} = \frac{1}{|A|}(A_{11}a_{1j} + A_{21}a_{2j} + \ldots + A_{n1}a_{nj}) = \frac{1}{|A|} \cdot 0 = 0$ – за теоремою 2.4.6 про розкладення по чужому стовпчику при $j = 2,3,\ldots,n$.

І взагалі, за аналогією $c_{ij} = \delta_{ij}$, $i, j = 1,2,\ldots,n$, тобто $C = E$, отже

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

Так само за наслідком 2.4.5 та теоремою 2.4.6 відносно рядків доводиться рівність $A \cdot A^{-1} = E$, що за означенням оберненої матриці доводить вірність формули (2.5.1). ■

Внаслідок властивостей 2 та 4 можемо стверджувати, що *матриця має обернену тоді і тільки тоді, коли вона невироджена.*

5. *Якщо* A, B *– невироджені матриці одного порядку, то*

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

Д о в е д е н н я. Нехай $(AB)^{-1} = X$. Тоді за означенням матриці, оберненої до AB, маємо $(AB)X = E$. За асоціативною властивістю множення матриць та за означенням матриць, обернених до A і B, маємо:

$[\ A(BX) = E\ ]$ – помножимо рівність на A^{-1} зліва $\Rightarrow$

$\Rightarrow [\ A^{-1}(A(BX)) = A^{-1}E\ ] \Rightarrow [\ (A^{-1}A)(BX) = A^{-1}\ ] \Rightarrow$

$[\ E(BX) = BX = A^{-1}\ ]$ – помножимо рівність $BX = A^{-1}$ на B^{-1} зліва $\Rightarrow$

$\Rightarrow [\ B^{-1}(BX) = (B^{-1}B)X = EX = X = B^{-1}A^{-1}\ ]$. Тобто $(AB)^{-1} = B^{-1}A^{-1}$. ■

6. $(A^T)^{-1} = (A^{-1})^T$ – *матриця, обернена до транспонованої, є матрицею, транспонованою до оберненої.*

Д о в е д е н н я. Якщо $A = (a_{ij}),\ A^T = (b_{ij})$, то $a_{ij} = b_{ji},\ B_{ji} = A_{ij}$ – відповідні алгебраїчні доповнення. Тоді, враховуючи першу властивість визначників, що $|A| = |A^T|$, маємо:

$$(A^T)^{-1} = \frac{1}{|A^T|}\begin{pmatrix} B_{11} & B_{21} & \dots & B_{n1} \\ B_{12} & B_{22} & \dots & B_{n2} \\ \dots & & \dots & \\ B_{1n} & B_{2n} & \dots & B_{nn} \end{pmatrix} = \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & & \dots & \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix} =$$

$$= \frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & & \dots & \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}^T = (A^{-1})^T. \blacksquare$$

<u>Приклад 1</u>. Знайти матрицю, обернену до даної:

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

▼ Так як $|A| = 2 \neq 0$, то матриця A невироджена, тому має обернену. Знайдемо алгебраїчні доповнення A_{ij} елементів a_{ij} матриці A. Для їх обчислення скористаємось формулою $A_{ij} = (-1)^{i+j} M'^{\,j}_{i}$. Маємо

$$A_{11} = (-1)^{1+1}\begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = 2;\quad A_{21} = (-1)^{2+1}\begin{vmatrix} 2 & -3 \\ 0 & 2 \end{vmatrix} = -4;\quad A_{31} = (-1)^{3+1}\begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix} = 7;$$

$$A_{12} = (-1)^{1+2}\begin{vmatrix} 0 & 2 \\ 0 & 2 \end{vmatrix} = 0;\quad A_{22} = (-1)^{2+2}\begin{vmatrix} 1 & -3 \\ 0 & 2 \end{vmatrix} = 2;\quad A_{32} = (-1)^{3+2}\begin{vmatrix} 1 & -3 \\ 0 & 2 \end{vmatrix} = -2;$$

$$A_{13} = (-1)^{1+3}\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} = 0;\quad A_{23} = (-1)^{2+3}\begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} = 0;\quad A_{33} = (-1)^{3+3}\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1.$$

За формулою (2.5.1) одержуємо:

$$A^{-1} = \frac{1}{2}\begin{pmatrix} 2 & -4 & 7 \\ 0 & 2 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3{,}5 \\ 0 & 1 & -1 \\ 0 & 0 & 0{,}5 \end{pmatrix}. \blacktriangle$$

Приклад 2. Знайти обернену матрицю до матриці

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

▼ Знайдемо визначник матриці $|A| = 5 + 12 + 8 - 6 - 8 - 10 = 1 \neq 0$. Матриця A невироджена, тому має обернену. Знайдемо алгебраїчні доповнення A_{ij} елементів a_{ij} матриці A:

$$A_{11} = (-1)^{1+1}\begin{vmatrix} -1 & 2 \\ 4 & -5 \end{vmatrix} = -3;\quad A_{21} = (-1)^{2+1}\begin{vmatrix} 2 & -3 \\ 4 & -5 \end{vmatrix} = -2;\quad A_{31} = (-1)^{3+1}\begin{vmatrix} 2 & -3 \\ -1 & 2 \end{vmatrix} = 1;$$

$$A_{12} = (-1)^{1+2}\begin{vmatrix} -1 & 2 \\ 2 & -5 \end{vmatrix} = -1;\quad A_{22} = (-1)^{2+2}\begin{vmatrix} 1 & -3 \\ 2 & -5 \end{vmatrix} = 1;\quad A_{32} = (-1)^{3+2}\begin{vmatrix} 1 & -3 \\ -1 & 2 \end{vmatrix} = 1;$$

$$A_{13} = (-1)^{1+3}\begin{vmatrix} -1 & -1 \\ 2 & 4 \end{vmatrix} = -2;\quad A_{23} = (-1)^{2+3}\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0;\quad A_{33} = (-1)^{3+3}\begin{vmatrix} 1 & 2 \\ -1 & -1 \end{vmatrix} = 1.$$

За формулою (2.5.1) одержуємо:

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

Контрольні питання та завдання

1. Як пов'язані невиродженість матриці та існування у неї оберненої матриці?

2. Напишіть формулу для обчислення оберненої матриці n-го порядку. Знайдіть обернену матрицю до матриці $\begin{pmatrix} 1 & 2 \\ 6 & -3 \end{pmatrix}$. Перевірте правильність знайденої матриці, виходячи з означення оберненої матриці.

2.6. Метод Гаусса

Система m ***лінійних рівнянь з*** n ***невідомими*** має вигляд

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \quad\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \qquad (2.6.1)$$

де a_{ij}, b_i – дійсні числа, x_j – ***невідомі змінні***, $i = 1,\ldots,m$, $j = 1,\ldots,n$.

Числа a_{ij} утворюють ***матрицю системи*** $A=(a_{ij})$, а числа b_i – ***стовпчик вільних членів*** $B=(b_i)$. Матриця розміру $m\times(n+1)$, яка утворюється приєднанням до матриці A стовпчика вільних членів, називається ***розширеною матрицею системи*** $(A|B)$.

Розв'язком системи (2.6.1) називають такий набір чисел $(l_1,l_2,...,l_n)$, що при відповідній підстановці замість невідомих $x_1,x_2,...,x_n$ кожне рівняння системи перетворюється в числову рівність. Всі такі набори утворюють ***множину розв'язків системи***.

Система, яка має хоч один розв'язок, називається ***сумісною***, в протилежному випадку – ***несумісною***. Сумісна система, яка має єдиний розв'язок, називається ***визначеною***. Сумісна система, що має більше одного розв'язку, називається ***невизначеною***. Система називається ***однорідною***, якщо її стовпчик вільних членів є нульовим: $b_1=b_2=...=b_n=0$, і ***неоднорідною*** – в протилежному випадку.

Дві системи називаються ***еквівалентними***, якщо вони несумісні або їх множини розв'язків співпадають.

Елементарними перетвореннями системи вважають:

1) перестановку рівнянь системи місцями;

2) множення обох частин довільного рівняння системи на одне й те ж саме число, відмінне від нуля;

3) додавання до довільного рівняння системи іншого, помноженого на деяке число.

Твердження 2.6.1. *Система, одержана за допомогою елементарних перетворень даної, еквівалентна їй.*

Д о в е д е н н я розглянемо для системи з двох рівнянь, тотожно не рівних нулю, оскільки кожне з елементарних перетворень торкається не більше двох рівнянь. Отже, нехай дано систему лінійних рівнянь:

$$\begin{cases} a_{11}x_1+a_{12}x_2+...+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+...+a_{2n}x_n=b_2 \end{cases}. \quad (1)$$

Виконаємо перше елементарне перетворення. Система матиме вигляд:

$$\begin{cases} a_{21}x_1+a_{22}x_2+...+a_{2n}x_n=b_2 \\ a_{11}x_1+a_{12}x_2+...+a_{1n}x_n=b_1 \end{cases}. \quad (2)$$

Очевидно, що кожний розв'язок $(l_1,l_2,\ldots,l_n)$ системи (1) буде розв'язком системи (2). І навпаки, кожний розв'язок $(\beta_1,\beta_2,\ldots,\beta_n)$ системи (2) буде розв'язком системи (1).

Виконаємо друге елементарне перетворення: помножимо перше рівняння на число $k \neq 0$. Система набуде вигляду:

$$\begin{cases} a'_{11}x_1 + a'_{12}x_2 + ... + a'_{1n}x_n = b'_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \end{cases}, \text{ де } a'_{1i} = ka_{1i},\ i = 1,\ldots,n,\ b'_1 = kb_1. \ (3)$$

Якщо $(l_1, l_2, \ldots, l_n)$ – розв'язок системи (1), то має місце числова рівність $a_{11}l_1 + a_{12}l_2 + ... + a_{1n}l_n = b_1$, а з неї випливає наступна числова рівність: $ka_{11}l_1 + ka_{12}l_2 + ... + ka_{1n}l_n = kb_1$. Тобто, всі розв'язки системи (1) задовольняють першому рівнянню системи (3). Оскільки друге рівняння системи (3) співпадає з другим рівнянням системи (1), то кожний розв'язок $(l_1, l_2, \ldots, l_n)$ системи (1) буде розв'язком системи (3).

Так як обидва рівняння систем не є нульовими тотожностями, то $\exists a_{1s} \neq 0,\ s \in \{1,\ldots,n\}$,а значить $a'_{1s} \neq 0$ (бо $a'_{1s} = ka_{1s}$). Тоді, помноживши перше рівняння системи (3) на a_{1s}/a'_{1s} ,з одного боку, одержимо перше рівняння системи (1), так як $a_{1s}/a'_{1s} = 1/k$, а з іншого боку, виконаємо елементарне перетворення другого типу в системі (3). Тобто, як тільки що доведено, кожний розв'язок $(\beta_1, \beta_2, \ldots, \beta_n)$ системи (3) буде розв'язком системи (1).

Виконаємо в системі (1) третє елементарне перетворення: до другого рівняння додамо перше, помножене на k. Система набуде вигляду:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a'_{21}x_1 + a'_{22}x_2 + ... + a'_{2n}x_n = b'_2 \end{cases}, \text{ де } a'_{2i} = a_{2i} + ka_{1i},\ i = 1,\ldots,n,\ b'_2 = b_2 + kb_1. \ (4)$$

Якщо $(l_1, l_2, \ldots, l_n)$ – розв'язок системи (1), то мають місце числові рівності $a_{11}l_1 + a_{12}l_2 + ... + a_{1n}l_n = b_1$ і $a_{21}l_1 + a_{22}l_2 + ... + a_{2n}l_n = b_2$, а з них випливає наступна числова рівність:

$$(a_{21} + ka_{11})l_1 + (a_{22} + ka_{12})l_2 + ... + (a_{2n} + ka_{1n})l_n = b_2 + kb_1.$$

Тобто, всі розв'язки системи (1) задовольняють другому рівнянню системи (4). Оскільки перше рівняння системи (4) співпадає з першим рівнянням системи (1), то кожний розв'язок $(l_1, l_2, \ldots, l_n)$ системи (1) буде розв'язком системи (4).

Навпаки, якщо в системі (4) виконаємо елементарне перетворення третього типу: до другого рівняння додамо перше, помножене на $-k$, то одержимо систему (1), оскільки $a'_{2i} - ka_{1i} = a_{2i},\ i = 1,\ldots,n,\ b'_2 - kb_1 = b_2$. Тоді, як тільки що доведено, кожний розв'язок $(\beta_1, \beta_2, \ldots, \beta_n)$ системи (4) буде розв'язком системи (1).

Якщо система (1) несумісна, то кожна із одержаних шляхом елементарних перетворень систем (2), (3), (4) також несумісна. Це легко довести методом від супротивного. Припустивши, що система (2), (3), або (4) має розв'язок $(\beta_1, \beta_2, \ldots, \beta_n)$, одержимо, що система (1) має цей же розв'язок і прийдемо до суперечності.

Отже, при довільному із елементарних перетворень системи одержуємо еквівалентну систему. ■

Метод Гаусса полягає у знаходженні розв'язку системи послідовним виключенням невідомих з деяких рівнянь за допомогою елементарних перетворень. Він складається з двох етапів.

Перший етап полягає в тому, що система (2.6.1) за допомогою елементарних перетворень приводиться до еквівалентної їй, яка має вигляд:

$$\begin{cases} c_{1s_1}x_{s_1} + \ldots + c_{1s_2}x_{s_2} + \ldots\ldots\ldots\ldots + c_{1n}x_n = d_1 & \\ \qquad\qquad c_{2s_2}x_{s_2} + \ldots\ldots\ldots\ldots + c_{2n}x_n = d_2 & (I) \\ \qquad\qquad\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ \qquad\qquad\qquad c_{rs_r}x_{s_r} + \ldots + c_{rn}x_n = d_r & \\ \text{---} & \\ \qquad\qquad\qquad\qquad 0 = d_{r+1} & \\ \qquad\qquad\qquad\qquad \ldots & (II) \\ \qquad\qquad\qquad\qquad 0 = d_m & \end{cases} \qquad (2.6.2)$$

Перша частина отриманої системи (2.6.2, *I*) називається ***розв'язуючою***, друга (2.6.2, *II*) – ***характеристичною***.

Зауважимо, що елементарним перетворенням системи відповідають такі ж перетворення розширеної матриці цієї системи, а матриця системи виду (2.6.2) має східчасту форму. Тому можна вважати, що перший етап методу Гаусса полягає в приведенні розширеної матриці системи до такої, в якій перші n стовпчиків утворюють східчасту матрицю. Можливість такого приведення спирається на теорему 2.2.1 (про приведення до східчастої матриці), яка застосовується до матриці системи. Тоді розширена матриця системи приводиться до вигляду:

$$\left(\begin{array}{cccccccccc|c} 0 & \dots & 0 & c_{1s_1} & \dots & c_{1s_2} & \dots & c_{1s_r} & \dots & c_{1n} & d_1 \\ 0 & \dots & 0 & 0 & \dots & c_{2s_2} & \dots & c_{2s_r} & \dots & c_{2n} & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & c_{rs_r} & \dots & c_{rn} & d_r \\ -- & -- & -- & -- & -- & -- & -- & -- & -- & -- & -- \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & d_{r+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 & d_m \end{array}\right) \begin{array}{c} (I) \\ \\ (II) \end{array} \qquad (2.6.3)$$

На цьому етапі доцільно провести аналіз сумісності системи (2.6.2), (а значить і системи (2.6.1)). Для цього використовують характеристичну частину (2.6.2, II), або (2.6.3, II). Якщо існує $d_t \neq 0$, $t \in \{r+1, r+2, ..., m\}$, то система (2.6.2) несумісна (ніякий набір чисел $(l_1, l_2, ..., l_n)$ не задовольнить рівняння $0 = d_t$).

Якщо ж $d_{r+1} = ... = d_m = 0$, то будь-який набір $(l_1, l_2, ..., l_n)$ задовольнить рівнянням $r+1, ..., m$. Тому розв'язки (2.6.2) слід шукати в розв'язуючій частині (2.6.2, I). Вона характеризується таблицею:

x_1 ... x_{s_1-1} x_{s_1} x_{s_1+1} ... x_{s_2} ….. x_{s_r} x_{s_r+1} ... x_n

0 ... 0 c_{1s_1} c_{1s_1+1} ... c_{1s_2} .... c_{1s_r} … … … c_{1n} d_1

0 ... 0 0 0 ... c_{2s_2} .... c_{2s_r} … … … c_{2n} d_2

………

0 ... 0 0 0 ... 0 ... c_{rs_r} … … c_{rn} d_r

В цій таблиці міститься матриця (2.6.3, I), в кожному i-му стовпчику якої стоять коефіцієнти, що відповідають змінній x_i, $i = 1, \dots, n$, а в $(n+1)$-му стовпчику – вільні члени. Очевидно, що матриця (2.6.3) має вигляд сходів з r сходинками, на яких стоять коефіцієнти відповідних змінних (наприклад, на 1-й сходинці стоять коефіцієнти змінних $x_{s_1}, x_{s_1+1}, \dots, x_{s_2-1}$: будемо говорити «стоять змінні

$x_{s_1}, x_{s_1+1}, \ldots, x_{s_2-1}$", на 2-й сходинці стоять змінні $x_{s_2}, x_{s_2+1}, \ldots, x_{s_3-1}$», на останній сходинці стоять змінні $x_{s_r}, \ldots, x_n$).

Для знаходження розв'язку необхідно обрати ***головні*** і ***незалежні*** змінні (при одержанні відповіді головні змінні виражаються через незалежні). Цей вибір можна здійснити за таким умовним правилом: *„ з кожної сходинки по одній головній змінній"*. Наприклад, з першої сходинки $c_{1s_1}c_{1s_1+1}...c_{1s_2-1}$ головною змінною може бути обрана та змінна x_j, $j = s_1, s_1+1, \ldots, s_2-1$, що $c_{1j} \neq 0$, з другої сходинки $c_{2s_2}c_{2s_2+1}...c_{2s_3-1}$ головною змінною може бути обрана та змінна x_j, $j = s_2, s_2+1, \ldots, s_3-1$, що $c_{2j} \neq 0$. Отже, враховуючи означення східчастої матриці, головними змінними можуть бути обрані r змінних: $x_{s_1}, x_{s_2}, ..., x_{s_r}$, а решту $(n-r)$ – незалежними. На практиці для полегшення обчислень найчастіше обирають головними ті змінні x_j, у яких найбільш зручні (якщо вони є) коефіцієнти $c_{ij} = 1, -1,\ 2,\ -2, ...$, де i – номер сходинки, з якої обирається змінна x_j.

Елементарним перетворенням рядків матриці (2.6.3, I) будуть відповідати аналогічні перетворення рівнянь системи (2.6.2, I). Будемо виконувати такі елементарні перетворення матриці (2.6.3, I), щоб коефіцієнти при головних змінних в тих рядках, з яких вони вибрані, дорівнювали 1, а в решті рядках дорівнювали 0. Для цього помножимо останній рядок розширеної матриці (2.6.3, I) на $1/c_{rs_r}$. Він набуде вигляду:

$$0...0...01\alpha_{s_r+1}...\alpha_n \mid q_r, \text{ де } \alpha_j = c_{rj}/c_{rs_r},\ j = s_r+1, ..., n;\ q_r = d_r/c_{rs_r}.$$

За допомогою множення його на $-c_{is_r}$, $i = 1, 2, ..., r-1$, та додавання до i-го рядка, матриця (2.6.3, I)зміниться на таку:

$$\left(\begin{array}{cccccccccccc|c} 0 & ... & 0 & c_{1s_1} & ... & c_{1s_2} & & ... & 0 & * & ... & * & q_1' \\ 0 & ... & 0 & 0 & ... & c_{2s_2} & & ... & 0 & * & ... & * & q_2' \\ ... & ... & ... & ... & ... & ... & & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & 0 & ... & 0 & & ... & 1 & \alpha_{s_r+1} & ... & \alpha_n & q_r \end{array}\right) \quad (2.6.5)$$

Тут через (*) позначено одержані в результаті перетворень елементи, які можуть не дорівнювати нулю.

Далі помножимо ($r-1$)-й рядок матриці (2.6.5) на $1/c_{(r-1)(s_{r-1})}$. Одержимо новий ($r-1$)-й рядок і виконаємо елементарні перетворення, аналогічні попереднім, в матриці (2.6.5). Продовжуючи цей процес до 1-го рядка, одержимо матрицю вигляду

$$\left(\begin{array}{ccccccccccccccc|c} 0 & ... & 0 & 1 & * & ... & * & 0 & * & ... & * & 0 & * & ... & * & q_1 \\ 0 & ... & 0 & 0 & 0 & ... & 0 & 1 & * & ... & * & 0 & * & ... & * & q_2 \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & 0 & 0 & 0 & ... & 0 & 0 & 0 & ... & 0 & 1 & \alpha_{s_r+1} & ... & \alpha_n & q_r \end{array}\right) \quad (2.6.6)$$

Пам’ятаючи про те, що ця матриця відповідає системі r лінійних рівнянь з n невідомими, можна записати розв’язок системи. Наприклад, останній рядок матриці (2.6.6) відповідає рівнянню

$$x_{s_r} + \alpha_{s_r+1} x_{s_r+1} + ... + \alpha_n x_n = q_r,$$

з якого одержуємо:

$$x_{s_r} = q_r - \alpha_{s_r+1} x_{s_r+1} - ... - \alpha_n x_n = f_r(x_{s_r+1}, ..., x_n).$$

Змінні x_{s_1}, x_{s_2}, ..., $x_{s_{r-1}}$ виражаються аналогічно змінній x_{s_r}. Тобто

$$\begin{cases} x_{s_1} = f_1(x_{s_1+1}, ..., x_n) = f_1, \\ x_{s_2} = f_2(x_{s_2+1}, ..., x_n) = f_2, \\ \qquad \cdots \\ x_{s_r} = f_n(x_{s_r+1}, ..., x_n) = f_r, \end{cases}$$

де f_i, $i = 1, ..., n$ не залежить від змінних x_{s_1}, x_{s_2}, ..., x_{s_r}. Це означає, що кожен розв’язок системи (2.6.2) (а значить і (2.6.1))має вигляд:

$$(x_1; ...; x_{s_1-1}; f_1; x_{s_1+1}; ...; f_2; ...; f_r; x_{s_r+1}; ...; x_n),$$

де всі змінні, крім головних x_{s_1}, x_{s_2}, ..., x_{s_r}, приймають довільні числові значення.

Приклад 1. Розв'язати систему лінійних рівнянь методом Гаусса:
$$\begin{cases} x_1 - 2x_2 + 3x_3 + 2x_4 + 5x_6 = 2 \\ 3x_1 - 6x_2 + 2x_3 + 4x_4 + 5x_5 = 4 \\ 2x_1 - 4x_2 - x_3 + 2x_4 + 3x_5 + 4x_6 = 4 \end{cases}.$$

▼ Складемо розширену матрицю даної системи і приведемо матрицю системи до східчастої форми разом із стовпчиком вільних членів:
$$\left(\begin{array}{cccccc|c} 1 & -2 & 3 & 2 & 0 & 5 & 2 \\ 3 & -6 & 2 & 4 & 5 & 0 & 4 \\ 2 & -4 & -1 & 2 & 3 & 4 & 4 \end{array}\right)\begin{array}{c}(-3)\,(-2) \\ \hookleftarrow \\ \hookleftarrow\end{array} \approx \left(\begin{array}{cccccc|c} 1 & -2 & 3 & 2 & 0 & 5 & 2 \\ 0 & 0 & -7 & -2 & 5 & -15 & -2 \\ 0 & 0 & -7 & -2 & 3 & -6 & 0 \end{array}\right)\begin{array}{c} \\ (-1) \\ \hookleftarrow\end{array} \approx$$
$$\approx \left(\begin{array}{cccccc|c} 1 & -2 & 3 & 2 & 0 & 5 & 2 \\ 0 & 0 & -7 & -2 & 5 & -15 & -2 \\ 0 & 0 & 0 & 0 & -2 & 9 & 2 \end{array}\right).$$

Характеристична частина системи відсутня, тому система сумісна і вся одержана матриця буде відповідати розв'язуючій частині системи. Виберемо головні змінні «з кожної сходинки по одній»: x_1, x_4, x_5, тоді змінні x_2, x_3, x_6 будуть незалежними. Далі виконаємо такі елементарні перетворення рядків останньої розширеної матриці, щоб в усіх рядочках, крім одного, коефіцієнти при головних змінних дорівнювали нулю (помножимо третій рядок на $5/2$ та додамо до другого, а потім одержаний другий рядок додамо до першого):
$$\left(\begin{array}{cccccc|c} 1 & -2 & -4 & 0 & 0 & 25/2 & 5 \\ 0 & 0 & -7 & -2 & 0 & 15/2 & 3 \\ 0 & 0 & 0 & 0 & -2 & 9 & 2 \end{array}\right).$$

Щоб коефіцієнти при всіх головних змінних дорівнювали 1, поділимо другий і третій рядки на (–2):
$$\left(\begin{array}{cccccc|c} 1 & -2 & -4 & 0 & 0 & 25/2 & 5 \\ 0 & 0 & 7/2 & 1 & 0 & -15/4 & -3/2 \\ 0 & 0 & 0 & 0 & 1 & -9/2 & -1 \end{array}\right).$$

Відтворюючи відповідні до одержаних рядків рівняння (наприклад, другому рядку відповідає рівняння $\frac{7}{2}x_3 + x_4 - \frac{15}{4}x_6 = -\frac{3}{2}$), знаходимо значення головних змінних:

$$x_1 = 2x_2 + 4x_3 - \frac{25}{2}x_6 + 5;$$

$$x_4 = -\frac{7}{2}x_3 + \frac{15}{4}x_6 - \frac{3}{2};$$

$$x_5 = \frac{9}{2}x_6 - 1.$$

Відповідь. Система невизначена. Загальний розв'язок має вигляд:

$$F = \left\{\left(2x_2 + 4x_3 - \frac{25}{2}x_6 + 5,\quad x_2,\quad x_3,\quad -\frac{7}{2}x_3 + \frac{15}{4}x_6 - \frac{3}{2},\right.\right.$$
$$\left.\left.\frac{9}{2}x_6 - 1,\quad x_6\right),\ x_2, x_3, x_6 \in \mathbf{R}\right\}.\blacktriangle$$

<u>Приклад 2.</u> Розв'язати систему лінійних рівнянь методом Гауса:

$$\begin{cases} x_1 - 2x_2 + 3x_3 + 2x_4 + 2x_5 + 5x_6 = 2 \\ 3x_1 - 6x_2 + 2x_3 + 4x_4 + 5x_5 + 9x_6 = 4 \\ 2x_1 - 4x_2 - x_3 + 2x_4 + 3x_5 + 4x_6 = 4 \end{cases}.$$

▼ Складемо розширену матрицю даної системи і приведемо матрицю системи до східчастої форми разом із стовпчиком вільних членів:

$$\left(\begin{array}{cccccc|c} 1 & -2 & 3 & 2 & 2 & 5 & 2 \\ 3 & -6 & 2 & 4 & 5 & 9 & 4 \\ 2 & -4 & -1 & 2 & 3 & 4 & 4 \end{array}\right)\begin{array}{cc} (-3) & (-2) \\ \hookleftarrow & \\ & \hookleftarrow \end{array} \approx$$

$$\left(\begin{array}{cccccc|c} 1 & -2 & 3 & 2 & 2 & 5 & 2 \\ 0 & 0 & -7 & -2 & -1 & -6 & -2 \\ 0 & 0 & -7 & -2 & -1 & -6 & 0 \end{array}\right)\begin{array}{c} \\ (-1) \\ \hookleftarrow \end{array} \approx$$

$$\approx \left(\begin{array}{cccccc|c} 1 & -2 & 3 & 2 & 2 & 5 & 2 \\ 0 & 0 & -7 & -2 & 5 & -6 & -2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right)\begin{array}{c} \\ (I) \\ (II) \end{array}.$$

Аналіз характеристичної частини (II) показує, що система несумісна ($2 \neq 0$), тобто не має розв'язків. ▲

<u>Приклад 3.</u> Розв'язати систему лінійних рівнянь методом Гауса:

$$\begin{cases} x_1 - 2x_2 - x_3 = 2 \\ 4x_1 - 9x_2 + 2x_3 = 2 \\ 3x_1 - 4x_2 - x_3 = 4 \\ x_1 - 5x_2 + 3x_3 = -2 \end{cases}.$$

▼ Складемо розширену матрицю даної системи і приведемо матрицю системи до східчастої форми разом із стовпчиком вільних членів:

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 2 \\ 4 & -9 & 2 & 2 \\ 3 & -4 & -1 & 4 \\ 1 & -5 & 3 & -2 \end{array}\right) \begin{array}{l} (-4)(-3)(-1) \\ \hookleftarrow \\ \hookleftarrow \\ \hookleftarrow \end{array} \approx \left(\begin{array}{ccc|c} 1 & -2 & -1 & 2 \\ 0 & -1 & 6 & -6 \\ 0 & 2 & 2 & -2 \\ 0 & -3 & 4 & -4 \end{array}\right) \begin{array}{l} \\ (2)(-3)(-1) \\ \hookleftarrow \\ \hookleftarrow \end{array} \approx$$

$$\approx \left(\begin{array}{ccc|c} 1 & -2 & -1 & 2 \\ 0 & 1 & -6 & 6 \\ 0 & 0 & 14 & -14 \\ 0 & 0 & -14 & 14 \end{array}\right) \begin{array}{l} \\ \\ (1):14 \\ \hookleftarrow \end{array} \approx \left(\begin{array}{ccc|c} 1 & -2 & -1 & 2 \\ 0 & 1 & -6 & 6 \\ 0 & 0 & 1 & -1 \\ \hline 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} (I) \\ \\ (II) \end{array}$$

Аналіз характеристичної частини (II) говорить про сумісність системи, а розв'язок слід шукати в розв'язуючій частині (I). Східчаста матриця цієї частини має три сходинки, тому головних змінних – три (x_1, x_2, x_3), незалежні змінні відсутні. Для одержання розв'язку виконаємо такі елементарні перетворення рядків розширеної матриці (I), щоб в усіх рядочках, крім одного, коефіцієнти при головних змінних дорівнювали нулю (додамо третій рядок до першого та помножений третій рядок на 6 та додамо до другого, а потім одержаний другий рядок помножимо на 2 та додамо до першого):

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

Відтворюючи відповідну до одержаних рядків систему рівнянь, знаходимо значення головних змінних:

$$\begin{cases} x_1 = 1 \\ x_2 = 0 \\ x_3 = -1 \end{cases}.$$

Відповідь. Система визначена, її розв'язок $(1, 0, -1)$. ▲

Зауважимо, якщо розглядати матрицю (2.6.6) з точністю до перестановки змінних, то вона може мати вигляд:

$$\left(\begin{array}{cccccccc|c} 1 & 0 & 0 & ... & 0 & * & ... & * & q_1 \\ 0 & 1 & 0 & ... & 0 & * & ... & * & q_2 \\ 0 & 0 & 1 & ... & 0 & * & ... & * & q_3 \\ ... & ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & * & ... & * & q_r \end{array}\right)$$

Тут легко бачити, що змінні x_1, x_2, ..., x_r – головні, а x_{r+1}, ..., x_n – незалежні. Множина розв'язків буде мати вигляд:

$$F = \{\ (f_1(x_{r+1},...,x_n);\ ...;\ f_r(x_{r+1},...,x_n);\ x_{r+1};\ ...;\ x_n),\ \big|\ x_{r+1},...,x_n \in \mathbf{R}\ \}.$$

Очевидно, що при $r < n$ (кількість головних змінних менша їх загальної кількості) кількість розв'язків нескінченна за рахунок нескінченності множини дійсних чисел, значення яких приймають $(n-r) > 0$ незалежних змінних, тобто, в цьому випадку, система (2.6.1) невизначена.

Розглянемо випадок $r = n$. Тоді система (2.6.2, *I*) буде мати вигляд:

$$\begin{cases} c_{11}x_1 + c_{12}x_2 + ... + c_{1n}x_n = d_1, \\ \qquad\quad c_{22}x_2 + ... + c_{2n}x_n = d_2, \\ \qquad\qquad\qquad\qquad \cdots \\ \qquad\qquad\qquad\quad c_{nn}x_n = d_n, \end{cases}$$

(тобто, матриця цієї системи має трикутний вигляд). А матриця (2.6.6) виглядатиме так:

$$\left(\begin{array}{cccc|c} 1 & 0 & ... & 0 & q_1 \\ 0 & 1 & ... & 0 & q_2 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & q_n \end{array}\right)$$

З неї видно, що ліва частина цієї матриці є одиничною матрицею E, а система має єдиний розв'язок $(q_1; q_2; ...; q_n)$.

Підводячи підсумок проведеного аналізу, одержуємо, що методом Гаусса можна розв'язати довільну систему лінійних рівнянь та виконуються ряд тверджень.

Твердження 2.6.2. *Система лінійних рівнянь (2.6.1) сумісна, якщо її характеристична частина* (2.6.2, *II*) *задовольняє умові*

$$d_{r+1} = d_{r+2} = ... = d_m = 0.$$

Твердження 2.6.3. *Якщо матриця розв'язуючої частини сумісної системи має трикутний вигляд, то така система визначена (має єдиний розв'язок).*

Твердження 2.6.4. *Якщо кількість головних змінних сумісної системи співпадає з кількістю невідомих, то система визначена. В протилежному випадку система невизначена, причому має безліч розв'язків.*

Розглянемо достатні умови для визначеності системи лінійних рівнянь. Нехай $(l_1, l_2, ..., l_n)$ – єдиний розв'язок системи (2.6.1). Тоді очевидно, що система

$$\begin{cases} x_1 = l_1 \\ x_2 = l_2 \\ \quad ... \\ x_n = l_n \end{cases} \qquad (2.6.7)$$

еквівалентна системі (2.6.1) і кількість їх головних змінних дорівнює n, так як матриця системи (2.6 7) одинична, а значить має n сходинок. Одержуємо

Твердження 2.6.5. *Сумісна система лінійних рівнянь визначена тоді і тільки тоді, коли кількість її головних змінних співпадає з кількістю невідомих.*

І як наслідок даного твердження одержуємо

Твердження 2.6.6. *Сумісна система лінійних рівнянь невизначена тоді і тільки тоді, коли кількість її головних змінних менша кількості невідомих.*

Контрольні питання та завдання

1. Яка матриця називається матрицею системи лінійних рівнянь, а яка – розширеною матрицею системи?
2. Складіть розширену матрицю системи двох рівнянь з трьома невідомими x, y, z: $x = 1, \quad y - z = 0$.
3. Що називається розв'язком системи лінійних рівнянь?
4. Яка система лінійних рівнянь називається сумісною, а яка визначеною?
5. Чи сумісна система рівнянь $x + y = 1, \quad 2x + 2y = 4$?
6. Які системи називаються еквівалентними?

2.7. Матричний метод розв'язування систем лінійних рівнянь

Розглянемо матричне рівняння

$$AX = B, \; (3.7.1)$$

де $A=(a_{ij})_m^n$, $B=(bij)_m^p$ –дані матриці, а $X=(x_{ij})_n^p$ – невідома матриця.

Якщо матриця A квадратна і $|A|\neq 0$, то існує обернена матриця A^{-1}, а тоді обидві частини рівняння (2.7.1) можна помножити зліва на A^{-1}, в результаті чого одержимо $A^{-1}(AX)=A^{-1}B$, а так як $A^{-1}(AX)=(A^{-1}A)X=$ $=EX=X$, то

$X=A^{-1}B$ – розв'язок рівняння (2.7.1).

Зауважимо, що при відповідних матрицях A і B матричне рівняння $XA=B$ має розв'язок $X=BA^{-1}$.

Нехай A – матриця системи (2.6.1), X – стовпчик невідомих, B – стовпчик вільних членів цієї системи ($A\in M_{m\times n}, X\in M_{n\times 1}, B\in M_{m\times 1}$). Тоді, враховуючи рівності системи (2.6.1),

$$AX=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}=\begin{pmatrix} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \\ \dots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}=B.$$

Тобто можна говорити, що систему (2.6.1) можна розглядати у матричному вигляді $AX=B$. Тоді, якщо матриця A – квадратна і $|A|\neq 0$, розв'язок системи (2.6.1) можна знайти як

$$X=A^{-1}B. \qquad (2.7.2)$$

Знаходження розв'язку системи (2.6.1) за формулою (2.7.2) називається ***матричним методом*** розв'язування системи.

Приклад. Розв'язати систему матричним методом:

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

▼ У матричному вигляді ця система має вигляд

$$AX=\begin{pmatrix} 1 & 2 & -3 \\ -1 & -1 & 2 \\ 2 & 4 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} -1 \\ -2 \\ -2 \end{pmatrix}=B.$$

Для знаходження розв'язку X за формулою (2.7.2) необхідно знайти матрицю A^{-1} і помножити її на стовпчик вільних членів B. Обернена матриця до матриці цієї системи знайдена в прикладі 2 п.2.5 цього розділу і $A^{-1}=\begin{pmatrix}-3 & -2 & 1\\ -1 & 1 & 1\\ -2 & 0 & 1\end{pmatrix}$. Отже $X=A^{-1}B=\begin{pmatrix}-3 & -2 & 1\\ -1 & 1 & 1\\ -2 & 0 & 1\end{pmatrix}\begin{pmatrix}-1\\ -2\\ -2\end{pmatrix}=\begin{pmatrix}5\\ -3\\ 0\end{pmatrix}$. ▲

Зустрічаються задачі, пов'язані з розв'язуванням матричних рівнянь виду $AX=B$ відносно X, причому матриця A може бути не квадратною або $|A|=0$. Наведемо один з методів розв'язування таких рівнянь.

1. Підбираємо елементарні перетворення рядків, які приводять матрицю A до східчастої матриці A'. Застосувавши ці ж самі перетворення до матриці B, одержимо матрицю B'. Процес такого перетворення далі будемо демонструвати за допомогою наступної схеми:

$(A|B)\to(A'|B')$.

Ясно, що рівняння $AX=B$ та $A'X=B'$ є еквівалентними.

Далі проводимо аналіз існування розв'язку рівняння, який є аналогічним до аналізу сумісності системи лінійних рівнянь по характеристичній частині. Після цього за допомогою елементарних перетворень рядків приведемо матрицю A' до матриці A'', що має вигляд лівої частини (2.6.6). При цих елементарних перетвореннях рядків матриця B' перетвориться в B''. Тоді рівняння $A''X=B''$ еквівалентне рівнянням $A'X=B'$ та $AX=B$:

$$(A|B)\to(A'|B')\to(A''|B'').$$

2. Матричне рівняння $A''X=B''$ можна представити як сукупність систем рівнянь, матриці яких співпадають з матрицею A'', стовпчиками їх невідомих є стовпчики матриці X, а стовпчики вільних членів є відповідними стовпчиками матриці B''. Стовпчики розв'язків цих систем утворюють шукану матрицю X.

Якщо виявляється, що $A''=E$, тобто $(A|B)\to(E|B'')$, то одержуємо рівняння $EX=B''$, значить $X=B''$ – єдиний розв'язок рівняння $AX=B$.

В окремому випадку, коли матричне рівняння має вигляд $AX=E$, де A – квадратна матриця, його розв'язок, якщо він існує, задовольняє означенню оберненої матриці, тобто $X=A^{-1}$ і цей розв'язок єдиний. Тоді даний метод розв'язування матричних рівнянь дає нам нове **правило знаходження оберненої матриці за допомогою елементарних перетворень рядків:**

щоб знайти матрицю, обернену до A*, треба елементарними перетвореннями рядків привести її до одиничної і цими ж елементарними перетвореннями подіяти на одиничну матрицю. Якщо матриця* A *приводиться до одиничної, то оберненою до неї буде та, що утворилася з одиничної, інакше – матриця* A *не має оберненої:*

$$(A|E) \to (E|A^{-1}).$$

Зауваження. Для розв'язування рівнянь $XA = B$ використовують аналогічні перетворення стовпчиків, що можна продемонструвати схемою:

$$\left(\frac{A}{B}\right) \to \left(\frac{A'}{B'}\right) \to \left(\frac{A''}{B''}\right).$$

Контрольні питання та завдання

1. Які системи лінійних рівнянь можна розв'язувати матричним методом?
2. Напишіть формулу для розв'язування системи лінійних рівнянь матричним методом.

2.8. Метод Крамера

Метод Крамера застосовується для розв'язування систем лінійних рівнянь $AX = B$, що мають такі ж обмеження, як і в попередньому пункті: визначник матриці системи існує і не дорівнює нулю. Обґрунтуємо цей метод.

Введемо такі позначення:

$$\Delta = |A|; \Delta_i = \begin{vmatrix} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots & a_{1n} \\ a_{21} & \dots & a_{2(i-1)} & b_2 & a_{2(i+1)} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots & a_{nn} \end{vmatrix}, i = 1,2,\dots,n.$$

Як видно, кожен визначник Δ_i утворюється із визначника Δ заміною i-го стовпчика на стовпчик вільних членів B, тому алгебраїчні доповнення елементу b_j будуть дорівнювати A_{ji} і розклавши визначник по i-му стовпчику одержимо

$$\Delta_i = b_1 A_{1i} + b_2 A_{2i} + \dots + b_n A_{ni}. \qquad (2.8.1)$$

Будемо розглядати систему у вигляді (2.6.1). При кожному фіксованому j $(j \in \{1,2,\dots,n\})$ помножимо кожне i-е рівняння системи

(2.6.1) $(i=1,2,...,n)$ на алгебраїчне доповнення A_{ij}, складемо ліві і праві частини цих рівнянь, виділимо коефіцієнти при змінних $x_1, x_2,..., x_n$.

При $j=1$ одержимо

$$x_1(a_{11}A_{11}+a_{21}A_{21}+...+a_{n1}A_{n1})+x_2(a_{12}A_{11}+a_{22}A_{21}+...+a_{n2}A_{n1})+...+$$
$$+x_n(a_{1n}A_{11}+a_{2n}A_{21}+...+a_{nn}A_{n1})=b_1A_{11}+b_2A_{21}+...+b_nA_{n1}. \quad (2.8.2)$$

В лівій частині коефіцієнт при x_1 – це розкладення визначника Δ по першому стовпчику (наслідок 2.4.5)), а коефіцієнти при останніх змінних – це розкладення Δ по чужому стовпчику, яке дорівнює нулю (твердження 2.4.6)). Права частина цього рівняння за (2.8.1) дорівнює Δ_1. Отже з рівняння (2.8.2) маємо $\Delta \cdot x_1 = \Delta_1$. Враховуючи, що $\Delta \neq 0$, одержуємо

$$x_1 = \frac{\Delta_1}{\Delta} \quad (2.8.3)$$

При довільних інших $j=2,3,...,n$ результат буде аналогічним. Тоді розв'язок системи знаходиться за наступними *формулами Крамера:*

$$x_j = \frac{\Delta_j}{\Delta}, j=1,2,...,n. \quad (2.8.4)$$

Знаходження розв'язку системи лінійних рівнянь за формулами (2.8.4) називається ***методом Крамера.***

Приклад. Розв'язати систему методом Крамера:

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

▼ Ми розглядаємо ту ж саму систему, що і в п.2.7. Тому нам відомо значення визначника матриці системи $\Delta=1$. Знайдемо значення визначників $\Delta_i, i=1,2,3$, які одержуються із визначника Δ заміною відповідного стовпчика на стовпчик вільних членів:

$$\Delta_1 = \begin{vmatrix} -1 & 2 & -3 \\ -2 & -1 & 2 \\ -2 & 4 & -5 \end{vmatrix} = 5; \; \Delta_2 = \begin{vmatrix} 1 & -1 & -3 \\ -1 & -2 & 2 \\ 2 & -2 & -5 \end{vmatrix} = -3; \; \Delta_3 = \begin{vmatrix} 1 & 2 & -1 \\ -1 & -1 & -2 \\ 2 & 4 & -2 \end{vmatrix} = 0.$$

За формулами Крамера (2.8.4) одержуємо:

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{5}{1} = 5; \quad x_2 = \frac{\Delta_2}{\Delta} = -3; \quad x_3 = \frac{\Delta_3}{\Delta} = 0.$$

Відповідь: (5; –3; 0). ▲

Контрольні питання та завдання

1. Які системи лінійних рівнянь можна розв'язувати методом Крамера?

2. Напишіть формули для розв'язування системи лінійних рівнянь методом Крамера.

2.9. Вектори. Лінійна залежність та лінійна незалежність векторів

Вектором розміру n називається упорядкований набір з n чисел $a = (\alpha_1, \alpha_2, ..., \alpha_n)$, $\alpha_i \in \mathbf{R}$, $i = 1,2,...,n$, числа α_i називаються ***компонентами*** вектора a. Множина всіх n- мірних векторів позначається $\mathbf{R}^n$ або V_n і називається ***арифметичним*** n***- мірним векторним простором***. На множині векторів задаються операції *додавання* і *множення* на число: $\forall\, a = (\alpha_1, \alpha_2, ..., \alpha_n), b = (b_1, b_2, ..., b_n) \in \mathbf{R}^n$, $\forall \lambda \in \mathbf{R}$

$$a + b = (\alpha_1 + \beta_1, ..., \alpha_n + \beta_n),\ \lambda a = (\lambda \alpha_1, ..., \lambda \alpha_n).$$

Таким чином матриці розміру $(1 \times n)$ можна вважати n- мірними векторами, а кожну матрицю A розміру $(m \times n)$ можна розглядати як упорядкований набір m векторів-рядків розміру n.

Узагальнюючи цей підхід, вектором-стовпчиком можна назвати матрицю розміру $(m \times 1)$ (множина всіх стовпчиків з m елементами позначається T_m), а матрицю A розглядати як упорядкований набір n векторів-стовпчиків із T_m.

Два вектори називаються ***рівними***, якщо їх відповідні компоненти рівні: $a = (\alpha_1, \alpha_2, ..., \alpha_n), b = (b_1, b_2, ..., b_n) \in \mathbf{R}^n$, $(\, a = b\,) \Leftrightarrow (\alpha_i = \beta_i, i = 1,...,n\,)$

Вектор, всі компоненти якого нульові, називається ***нульовим*** і позначається $\theta, \vec{0}, \overline{0}$ тощо. Вектор $-a$ називається ***протилежним*** до вектору a, якщо $a + (-a) = \theta$. Маємо

В л а с т и в о с т і в е к т о р і в

$\forall a, b, c \in \mathbf{R}^n, \forall \lambda, \mu \in \mathbf{R}$

1. $(a + b) + c = a + (b + c)$;

2. $a + b = b + a$;

3. $a + \theta = a$;

4. Якщо $a = (\alpha_1, \alpha_2, ..., \alpha_n)$, то $\exists -a = (-\alpha_1, -\alpha_2, ..., -\alpha_n) \in \mathbf{R}^n$;

5. $1 \cdot a = a$;

6. $(\lambda\mu)\cdot a = \lambda\cdot(\mu\cdot a)$;

7. $(\lambda+\mu)\cdot a = \lambda\cdot a + \mu\cdot a$;

8. $\lambda\cdot(a+b) = \lambda\cdot a + \lambda\cdot b$;

9. $0\cdot a = \theta$;

10. $(-1)\cdot a = -a$;

11. $\lambda\cdot\theta = \theta$.

Дані властивості пропонується довести самостійно.

Лінійною комбінацією векторів $a_1, a_2, ..., a_s \in \mathbf{R}^n$ з дійсними коефіцієнтами $k_1, k_2, ...k_s$ називається вектор $y \in \mathbf{R}^n$, що дорівнює сумі

$$y = k_1 a_1 + k_2 a_2 + ... + k_s a_s = \sum_{i=1}^{s} k_i a_i.$$

На основі властивостей 9 і 3 лінійна комбінація довільних векторів з нульовими коефіцієнтами дорівнює θ. Вектори $a_1, a_2, ..., a_s$ називаються ***лінійно залежними (ЛЗ),*** якщо для них існують не всі рівні нулю коефіцієнти $k_1, k_2, ..., k_s$ такі, що лінійна комбінація $\sum_{i=1}^{s} k_i a_i = \theta$. В протилежному випадку, якщо лінійна комбінація векторів дорівнює нулю тільки при нульових коефіцієнтах, ці вектори називаються ***лінійно незалежними (ЛНЗ).***

Приклад. Дано вектори:

$a_1 = (1, 2, -3, 4)$;

$a_2 = (2, 1, 5, 7)$;

$a_3 = (-1, -2, 3, -2)$;

$a_4 = (-1, -5, 14, -5)$.

Довести, що 1) вектори a_1, a_2, a_3 лінійно незалежні, а 2) вектори a_1, a_2, a_3, a_4 лінійно залежні.

▼ 1) Припустимо, що a_1, a_2, a_3 лінійно залежні. Тоді існує їх лінійна комбінація, рівна нульовому вектору:

$k_1 a_1 + k_2 a_2 + k_3\, a_3 = \theta$, (1)

причому хоча б один із коефіцієнтів k_1, k_2, k_3 не дорівнює нулю. Підставивши в рівність (1) значення векторів a_1, a_2, a_3, одержимо рівність

$$(k_1 + 2k_2 - k_3,\ 2k_1 + k_2 - 2k_3,\ -3k_1 + 5k_2 + 3k_3,\ 4k_1 + 7k_2 - 2k_3) = (0, 0, 0, 0),$$

якій еквівалентна наступна однорідна система лінійних рівнянь з невідомими k_1, k_2, k_3:

$$\begin{cases} k_1 + 2k_2 - k_3 = 0 \\ 2k_1 + k_2 - 2k_3 = 0 \\ -3k_1 + 5k_2 + 3k_3 = 0 \\ 4k_1 + 7k_2 - 2k_3 = 0. \end{cases}$$

Приведемо матрицю цієї системи до східчастої форми (зауважимо, що матриця системи складається із векторів a_1, a_2, a_3, записаних в стовпчик):

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & -2 \\ -3 & 5 & 3 \\ 4 & 7 & -2 \end{pmatrix} \approx \begin{pmatrix} 1 & 2 & -1 \\ 0 & -3 & 0 \\ 0 & 11 & 0 \\ 0 & -1 & 2 \end{pmatrix} \approx \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

звідки видно, що система має тільки нульовий розв'язок $k_1 = k_2 = k_3 = 0$. Одержали протиріччя з тим, що хоча б один із коефіцієнтів k_1, k_2, k_3 не дорівнює нулю. Значить дані вектори лінійно незалежні. □

2) Покажемо, що a_1, a_2, a_3, a_4 лінійно залежні. Розглянемо їх лінійну комбінацію, рівну нульовому вектору:

$k_1 a_1 + k_2 a_2 + k_3\ a_3 + k_4 a_4 = \theta$. (2)

Визначимо, чи може хоча б один із коефіцієнтів k_1, k_2, k_3, k_4 цієї рівності не дорівнювати нулю. Підставивши в рівність (2) значення векторів a_1, a_2, a_3, a_4, одержимо рівність

$(k_1 + 2k_2 - k_3 - k_4, 2k_1 + k_2 - 2k_3 - 5k_4, -3k_1 + 5k_2 + 3k_3 + 14k_4, 4k_1 + 7k_2 - 2k_3 - 5k_4) =$
$= (0, 0, 0, 0)$,

якій еквівалентна наступна однорідна система лінійних рівнянь з невідомими k_1, k_2, k_3, k_4:

$$\begin{cases} k_1 + 2k_2 - k_3 - k_4 = 0 \\ 2k_1 + k_2 - 2k_3 - 5k_4 = 0 \\ -3k_1 + 5k_2 + 3k_3 + 14k_4 = 0 \\ 4k_1 + 7k_2 - 2k_3 - 5k_4 = 0. \end{cases}$$

Приведемо матрицю цієї системи до східчастої форми (матриця системи складається із векторів a_1, a_2, a_3, a_4, записаних в стовпчик):

$$\begin{pmatrix} 1 & 2 & -1 & -1 \\ 2 & 1 & -2 & -5 \\ -3 & 5 & 3 & 14 \\ 4 & 7 & -2 & -5 \end{pmatrix} \approx \begin{pmatrix} 1 & 2 & -1 & -1 \\ 0 & -3 & 0 & -3 \\ 0 & 11 & 0 & 11 \\ 0 & -1 & 2 & -9 \end{pmatrix} \approx \begin{pmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

звідки видно, що система має незалежну змінну k_4, тому існує ненульовий розв'язок (наприклад, $k_4 = 1$, $k_3 = -4$, $k_2 = -1$, $k_1 = -1$). А це означає, що система векторів a_1, a_2, a_3, a_4 лінійно залежна. ▲

Зауважимо, що словосполучення „система векторів” використовується в розумінні „сукупність векторів” і це поняття не має аналогії з поняттям „система рівнянь”.

Властивостісистемивекторів

1.*Система, що складається з одного нульового вектора, лінійно залежна*:

▼ при $a = \theta$ за властивістю 11, наприклад, $1 \cdot a = \theta$, що задовольняє означенню лінійної залежності. ▲

2.*Система, що складається з одного ненульового вектора, лінійно незалежна:*

▼ $a = (\alpha_1, \alpha_2, ..., \alpha_n) \neq \theta \Leftrightarrow \exists \alpha_i \neq 0$. Тоді $k \cdot a = \theta \Leftrightarrow k\,\alpha_j = 0$, $j = 1,2,...,n$, в тому числі $k\alpha_i = 0$, а це можливо тільки при $k = 0$, що задовольняє означенню лінійної незалежності. ▲

3.*Якщо вектори a,b пропорційні* ($a = \lambda b$), *то система векторів a,b – лінійно залежна*:

▼ $1 \cdot a + (-\lambda) \cdot b = \theta$, причому коефіцієнт при a не дорівнює нулю. ▲

4. *Кожна підсистема лінійно незалежної системи векторів лінійно незалежна.*

▼ Від супротивного: нехай $a_1, a_2, ..., a_s$ – лінійно незалежна система, а $a_1, a_2, ..., a_r$, $r < s$ – її лінійно залежна підсистема. Тоді за означенням лінійної залежності $\exists k_1, k_2, ..., k_r$, не всі рівні нулю, що $\sum_{i=1}^{r} k_i a_i = \theta$, а значить $\sum_{i=1}^{r} k_i a_i + 0a_{i+1} + ... + 0a_s = \theta$, тобто вектори $a_1, a_2, ..., a_s$ задовольняють означенню лінійної залежності, що суперечить припущенню. ▲

5.*Якщо підсистема векторів лінійно залежна, то вся система лінійно залежна.*

▼ Доведення цієї властивості міститься в попередньому доведенні. ▲

Теорема 2.9.1. (Критерій лінійної залежності.) *Система векторів лінійно залежна тоді і тільки тоді, коли один з векторів можна виразити через лінійну комбінацію решти векторів системи.*

Н е о б х і д н і с т ь. Нехай система векторів $a_1, a_2, ..., a_t$ лінійно залежна, тоді за означенням існує набір коефіцієнтів $k_1, k_2, ..., k_t$, не всі з яких рівні нулю (наприклад $k_1 \neq 0$), а $\sum_{i=1}^{t} k_i a_i = \theta$. Звідси

$$a_1 = \left(\frac{-k_2}{k_1}\right) a_2 + \left(\frac{-k_3}{k_1}\right) a_3 + ... + \left(\frac{-k_t}{k_1}\right) a_t. \blacktriangle$$

Д о с т а т н і с т ь. Нехай $a_1 = k_2 a_2 + k_3 a_3 + ... + k_t a_t$, тоді

$$(-1)a_1 + k_2 a_2 + k_3 a_3 + ... + k_t a_t = \theta,$$

причому коефіцієнт при a_1 дорівнює $-1 \neq 0$, що доводить лінійну залежність системи векторів $a_1, a_2, ..., a_t$. ■

Твердження 2.9.2. *Якщо система векторів $a_1, a_2, ..., a_t$ лінійно незалежна, а система векторів $a_1, a_2, ..., a_t, a_{t+1}$ лінійно залежна, то вектор a_{t+1} лінійно виражається через $a_1, a_2, ..., a_t$.*

Д о в е д е н н я. Якщо система векторів $a_1, a_2, ..., a_t, a_{t+1}$ лінійно залежна, то за означенням $\exists \lambda_i \in \mathbf{R}$, $i = 1,2,...,t+1$, не всі рівні нулю такі, що

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_t a_t + \lambda_{t+1} a_{t+1} = \theta. \quad (2.9.1)$$

Припустимо, що $\lambda_{t+1} = 0$. Тоді $\sum_{i=1}^{t+1} \lambda_i a_i = \sum_{i=1}^{t} \lambda_i a_i = \theta$, де не всі коефіцієнти рівні нулю, що суперечить умові лінійної незалежності $a_1, a_2, ..., a_t$. Значить припущення було не вірним. Отже, $\lambda_{t+1} \neq 0$ і тому із співвідношення (2.9.1) маємо

$$a_{t+1} = \left(\frac{-\lambda_1}{\lambda_{t+1}}\right) a_2 + \left(\frac{-\lambda_2}{\lambda_{t+1}}\right) a_3 + ... + \left(\frac{-\lambda_t}{\lambda_{t+1}}\right) a_t. \blacksquare$$

Контрольні питання та завдання

1. Дайте означення лінійної залежності та лінійної незалежності векторів.
2. Сформулюйте критерій лінійної залежності.
3. Чи є лінійно залежними вектори арифметичного простору V_2: а) x *і* $2x$; б) $x_1, x_2, \ldots, x_n, \theta$; в) (1,0) і (0,1)? Відповіді обґрунтуйте.

2.10. Ранг матриці. Базисний мінор

Рядковим(стовпчиковим) рангом матриці називається максимальна кількість лінійно незалежних векторів-рядків (векторів-стовпчиків) цієї матриці.

Мають місце наступні твердження, якими ми скористаємося без доведення.

Твердження 2.10.1. *Стовпчиковий ранг матриці не змінюється при довільних елементарних перетвореннях рядків. Рядковий ранг матриці не змінюється при довільних елементарних перетвореннях стовпчиків.*

Твердження 2.10.2. *Рядковий ранг дорівнює стовпчиковому рангу матриці.*

Тому говорять просто, ***ранг матриці*** – це максимальна кількість лінійно незалежних рядків або стовпчиків. Якщо ранг матриці A дорівнює r, то позначають $rang\ A = r$. Із попередніх тверджень також випливає

Твердження 2.10.3. *Елементарні перетворення рядків і стовпчиків не змінюють рангу матриці.*

Теорема 2.10.4. *Ранг східчастої матриці дорівнює кількості її ненульових рядків.*

Д о в е д е н н я. Позначимо кожний i-й рядок матриці (2.2.1) через $\overline{c}_i$. Доведемо, що система рядків $\overline{c}_1, \overline{c}_2, ..., \overline{c}_r$ лінійно незалежна, а саме, що $\sum_{i=1}^{r} k_i \overline{c}_i = \theta$ тільки при $k_i = 0,\ i = 1,2,...,r$.

Розглянемо покоординатно рівність, яка нас цікавить:

1. $\sum_{i=1}^{r} k_i \overline{c}_i = (0,...,0,k_1 c_{1s_1},...) = \theta \quad \Rightarrow \quad k_1 c_{1s_1} = 0$. Оскільки $c_{1s_1} \neq 0$ за означенням

східчастої матриці, то $k_1 = 0$. Тоді

2. $\sum_{i=1}^{r} k_i \overline{c}_i = \sum_{i=2}^{r} k_i \overline{c}_i = (0,...,0,k_2 c_{2s_2},...) = \theta \quad \Rightarrow \quad k_2 c_{2s_2} = 0$, звідки, аналогічно до попереднього кроку, одержуємо $k_2 = 0$.

Продовжуючи цей процес до r-го кроку, одержимо $k_i = 0,\ i = 1,2,...,r$, що доводить лінійну незалежність $\overline{c}_1, \overline{c}_2, ..., \overline{c}_r$.

Довільна система із $(r+1)$ або більше рядків матриці (2.2.1) містить нульовий рядок $\overline{c}_t,\ t \in \{r+1, r+2, ..., m\}$, тоді за першою та п’ятою властивостями лінійної залежності така система лінійно залежна. Значить ранг східчастої матриці (2.2.1) дорівнює r, як і кількість її ненульових рядків. ■

За теоремою 2.2.1, твердженням 2.10.3 і теоремою2.10.4 ранг матриці можна знайти, привівши її елементарними перетвореннями до східчастої матриці. Ранг буде дорівнювати кількості ненульових рядків східчастої матриці, яка знаходиться звичайним підрахунком.

<u>Приклад 1.</u> Знайти ранг даної матриці:

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

▼ Приведемо дану матрицю елементарними перетвореннями рядків до східчастої матриці:

$$\begin{pmatrix} 1 & -2 & 3 & 2 & 1 & 3 \\ 3 & -6 & 2 & 4 & 5 & 7 \\ 2 & -4 & -1 & 2 & 4 & 4 \\ -2 & 4 & 3 & -1 & -3 & 2 \end{pmatrix}\begin{array}{lll} (-3) & (-2) & 2 \\ \hookleftarrow & & \\ & \hookleftarrow & \\ & & \hookleftarrow \end{array} \approx$$

$$\approx\begin{pmatrix} 1 & -2 & 3 & 2 & 1 & 3 \\ 0 & 0 & -7 & -2 & 2 & -2 \\ 0 & 0 & -7 & -2 & 2 & -2 \\ 0 & 0 & 9 & 3 & -3 & 12 \end{pmatrix}\begin{array}{ll} & \\ (-1) & 9 \\ \hookleftarrow & \\ & 7\hookleftarrow \end{array} \approx$$

$$\approx\begin{pmatrix} 1 & -2 & 3 & 2 & 1 & 3 \\ 0 & 0 & -7 & -2 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & -3 & 66 \end{pmatrix}\approx\begin{pmatrix} 1 & -2 & 3 & 2 & 1 & 3 \\ 0 & 0 & -7 & -2 & 2 & -2 \\ 0 & 0 & 0 & 1 & -1 & 22 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Одержана східчаста матриця має три ненульових рядочки, тому $rang\ A = 3$. ▲

Поняття мінору існує не тільки для визначника, а і для матриці довільного розміру як визначник, утворений елементами, що стоять на перетині $i_1, i_2, \ldots, i_k$ рядків і $j_1, j_2, \ldots, j_k$ стовпчиків. Кожний мінор максимального порядку довільної матриці A, відмінний від нуля, називається ***базисним мінором***. Рядки (стовпчики), на яких він розташований – ***базисними рядками (стовпчиками)***; кількість цих рядків і стовпчиків однакова, як у довільного мінору, і називається ***порядком***

базисного мінору. Зверніть увагу, як випливає з означення, всі базисні мінори матриці мають однаковий порядок.

Теорема 2.10.5. (Про базисний мінор).

1.Базисні рядки (стовпчики) матриці лінійно незалежні.

2.Кожний рядок (стовпчик) матриці є лінійною комбінацією базисних рядків (стовпчиків) цієї матриці.

Д о в е д е н н я.

1. Нехай $\overline{a}_1, \overline{a}_2, ..., \overline{a}_r$ – базисні рядки матриці $A = (a_{ij})_m^n$. Припустимо, що вони лінійно залежні і $\overline{a}_r = \sum_{i=1}^{r-1} \lambda_i \overline{a}_i$ (за критерієм 2.9.1). Нехай базисний мінор M вибрано у стовпчиках $j_1, j_2, ..., j_r$. Тоді його r-й рядок дорівнює $(\sum_1^{r-1} \lambda_i a_{ij_1}, \sum_1^{r-1} \lambda_i a_{ij_2}, ..., \sum_1^{r-1} \lambda_i a_{ij_r})$, тобто є лінійною комбінацією останніх рядків, звідки за 8-ю властивістю визначників випливає, що $M = 0$, а це суперечить означенню базисного мінору і доводить першу частину теореми.

2. Не обмежуючи загальності доведення, можна вважати, що базисний мінор M знаходиться в лівому верхньому куту матриці A, тобто $M = M_{1,2,...,r}^{1,2,...,r} \neq 0$. Розглянемо можливі випадки для значень визначника M_1 при різних k і j, де

$$M_1 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2r} & a_{2j} \\ \dots & \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} & a_{rj} \\ a_{k1} & a_{k2} & \dots & a_{kr} & a_{kj} \end{vmatrix}, \ k \leq m, \ j \leq n:$$

1) якщо $k \leq r$, то $M_1 = 0$ за 4-ю властивістю визначників, бо $(r+1)$-й його рядок дорівнює k-му рядку;

2) якщо $j \leq r$, то $M_1 = 0$ аналогічно, бо $(r+1)$-й його стовпчик дорівнює j-му стовпчику;

3) якщо $r < k \leq m, r < j \leq n$, то $M_1 = M_{1,2,...,r,k}^{1,2,...,r,j}$ – мінор матриці A порядку $(r+1)$, тобто $M_1 = 0$ із означення базисного мінору.

Отже визначник $M_1 = 0$ при всіх можливих значеннях k і j. Розкладемо його по $(r+1)$-му стовпчику:

$$M_1 = a_{1j} A_{1(r+1)} + a_{2j} A_{2(r+1)} + ... + a_{rj} A_{r(r+1)} + a_{kj} A_{(r+1)(r+1)} = 0. \ (1)$$

В цій рівності $A_{i(r+1)}$ – алгебраїчні доповнення до елементів $(r+1)$-го стовпчика визначника M_1. Позначимо $c_i(k) = A_{i(r+1)}$, $i = 1,2,...,r+1$. Зауважимо, що $c_i(k)$ не залежить від j $\forall i \in \{1,2,...,r\}$, а $c_{r+1}(k)$ не залежить і від k, а саме $c_{r+1}(k) = M \neq 0$. Отже із рівності (1) маємо представлення кожного елементу k-го рядка через елементи відповідного стовпчика базисних рядків

$$\forall j \in \{1,2,...,n\} \ a_{kj} = \left(\frac{-c_1(k)}{M}\right)a_{1j} + \left(\frac{-c_2(k)}{M}\right)a_{2j} + ... + \left(\frac{-c_r(k)}{M}\right)a_{rj},$$

де $\left(\frac{-c_i(k)}{M}\right)$ не залежить від j. Значить кожний k-й рядок матриці A має представлення $\overline{a}_k = \left(\frac{-c_1(k)}{M}\right)\overline{a}_1 + \left(\frac{-c_2(k)}{M}\right)\overline{a}_2 + ... + \left(\frac{-c_r(k)}{M}\right)\overline{a}_r$. ■

Теорема 2.10.6. (Про зв'язок порядку базисного мінору і рангу матриці) *Порядок базисного мінору матриці дорівнює рангу цієї матриці.*

Д о в е д е н н я. Нехай порядок базисного мінору матриці A дорівнює r. Якщо $r = 0$, то не існує мінору матриці A, відмінного від 0. Значить за теоремою 2.10.5(1) не існує лінійно незалежних рядків, а за означенням рангу $rang\ A = 0$ і в цьому випадку теорема доведена. (Очевидно, що цей випадок можливий тільки для нульової матриці).

Якщо $r \neq 0$, то існують базисні рядки матриці A (нехай це будуть рядки $\overline{a}_1, \overline{a}_2, ..., \overline{a}_r$), які за теоремою 2.10.5 (1) лінійно незалежні, звідки $rang\ A \geq r$.

Доведемо, що $rang\ A = r$, а саме, що довільні рядки $\overline{s}_1, \overline{s}_2, ..., \overline{s}_p$ $(p > r)$ матриці A лінійно залежні.

Складемо матрицю B із рядків $\overline{s}_1, \overline{s}_2, ..., \overline{s}_p$. Кожний мінор матриці B з точністю до порядку розташування рядків $\overline{s}_1, \overline{s}_2, ..., \overline{s}_p$ в матриці A є мінором матриці A, тому всі ненульові мінори матриці B знаходяться серед ненульових мінорів матриці A, звідки порядок r_1 базисного мінору матриці B задовольняє умові $r_1 \leq r$, тобто $r_1 < p$. Значить за теоремою 2.10.5(2) існує рядок матриці B, який не є її базисним рядком, але лінійно виражається через базисні рядки. Отже, $\forall p > r$ рядки $\overline{s}_1, \overline{s}_2, ..., \overline{s}_p$ лінійно залежні, значить $rang\ A = r$ і теорему доведено. ■

Із теорем 2.10.5 і 2.10.6 випливає

Наслідок 2.10.7. *Для того, щоб рядки матриці були лінійно залежні (лінійно незалежні), необхідно і достатньо, щоб її ранг був менший числа (дорівнював числу) рядків.*

Також мають місце наслідки, які пропонуються для самостійного доведення.

Наслідок 2.10.8. *Визначник матриці n-го порядку не дорівнює нулю (квадратна матриця n-го порядку невироджена) тоді і тільки тоді, коли її ранг дорівнює n.*

Наслідок 2.10.9. ***М**атриця n-го порядку вироджена тоді і тільки тоді, коли її ранг менший n.*

Наслідок 2.10.10 ***М**атриця невироджена тоді і тільки тоді, коли її рядки лінійно незалежні.*

Наслідок 2.10.11. ***М**атриця вироджена тоді і тільки тоді, коли її рядки лінійно залежні.*

Всі наслідки, сформульовані для рядків, мають місце і для стовпчиків.

Із теореми 2.10.6 випливає, що для знаходження рангу матриці достатньо знайти її базисний мінор. Його порядок буде рангом матриці.

Для знаходження базисного мінору і рангу матриці використовують *метод обмежуючих мінорів*, який полягає у послідовному переході від мінорів менших порядків матриці до мінорів більших порядків. Мінор D_1 називається ***обмежуючим*** для $D = M_{i_1 i_2 \ldots i_k}^{j_1 j_2 \ldots j_k}$, якщо $D_1 = M_{i'_1 i'_2 \ldots i'_{k+1}}^{j'_1 j'_2 \ldots j'_{k+1}}$, де $\{i_1, i_2, \ldots, i_k\} \subset \{i'_1, i'_2, \ldots, i'_{k+1}\}$ і $\{j_1, j_2, \ldots, j_k\} \subset \{j'_1, j'_2, \ldots, j'_{k+1}\}$. Якщо вже знайдений мінор k-го порядку Δ_k, відмінний від нуля, то шукаємо мінор $(k+1)$-го порядку Δ_{k+1} такий, що є обмежуючим мінором для Δ_k і відмінним від нуля. Якщо такого мінору $(k+1)$-го порядку не існує, то ранг матриці дорівнює k, а мінор Δ_k є базисним мінором. В протилежному випадку процес продовжується. Його скінченність забезпечується скінченною кількістю рядочків та стовпчиків матриці.

<u>Приклад 2</u>. Знайти ранг матриці методом обмежуючих мінорів:

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

▼ Почнемо обчислення із знаходження кутових мінорів:

$$\Delta_1 = M_1^1 = |1| = 1 \neq 0; \quad \Delta_2 = M_{1,2}^{1,2} = \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = 1 \neq 0;$$

$$\Delta_3 = M_{1,2,3}^{1,2,3} = \begin{vmatrix} 1 & 2 & -3 \\ 2 & 5 & 4 \\ 3 & 7 & 1 \end{vmatrix} = 0; \quad \Delta_3' = M_{1,2,3}^{1,2,4} = \begin{vmatrix} 1 & 2 & 6 \\ 2 & 5 & -3 \\ 3 & 7 & 3 \end{vmatrix} = 0.$$

Отже, обмежуючого мінору третього порядку не існує, решту мінорів третього порядку перевіряти немає потреби. Базисним мінором може бути мінор другого порядку $M_{1,2}^{1,2}$, а за теоремою 2.10.6 $rang\ A = 2$. ▲

Контрольні питання та завдання

1. Дайте означення рангу матриці? Які операції над рядками та стовпчиками матриці не змінюють її рангу?

2. Чому дорівнює ранг транспонованої матриці A^T, якщо $rang\ A = r$?

3. Дайте означення базисного мінору, базисних рядків і базисних стовпчиків.

4. Знайдіть ранг, всі базисні мінори і відповідні їм базисні рядки і стовпчики матриці $\begin{pmatrix} 0 & 1 & 1 & -1 & 3 \\ 0 & 2 & 1 & -2 & 3 \end{pmatrix}$.

5. Сформулюйте теорему про зв’язок порядку базисного мінору і рангу матриці.

6. Що можна сказати про порядок базисного мінору матриці A n-го порядку, якщо A – вироджена матриця?

7. Чи може ранг матриці з розмірами 5×6 дорівнювати: 3; 5; 6; 7?

8. Нехай ранг $(m \times n)$- матриці A дорівнює r. Чи є рядки матриці A лінійно залежні, якщо: а) $r < n$; б) $r = n$?

9. Чи можуть бути лінійно незалежними рядки з трьома елементами, якщо число рядків дорівнює: 2; 3; 4; 5?

10. Чи може ранг матриці дорівнювати r, якщо:

а) деякі r її рядків лінійно залежні;

б) довільні r рядків матриці лінійно залежні;

в) деякі $r+1$ її рядків лінійно незалежні? Відповіді обґрунтуйте.

2.11. Критерії сумісності та визначеності

Розглянемо систему (2.6.1). Якщо вектори-стовпчики матриці A цієї системи позначити $\overline{a}^1, \overline{a}^2, ..., \overline{a}^n$, то систему можна записати у векторному вигляді:

$$\overline{a}^1 x_1 + \overline{a}^2 x_2 + \ldots + \overline{a}^n x_n = B, \text{ (2.11.1)}$$

а існування розв'язку $(l_1, l_2, \ldots, l_n)$ цієї системи означає, що стовпчик вільних членів B лінійно виражається через стовпчики $\overline{a}^1, \overline{a}^2, \ldots, \overline{a}^n$ матриці системи:

$$B = \overline{a}^1 l_1 + \overline{a}^2 l_2 + \ldots + \overline{a}^n l_n.$$

Теорема 2.11.1. (Критерій сумісності[1]*). *Система лінійних рівнянь сумісна тоді і тільки тоді, коли ранг розширеної матриці дорівнює рангу матриці системи.*

Н е о б х і д н і с т ь. Нехай система (2.6.1) – сумісна. За попереднім зауваженням це означає, що система векторів-стовпчиків розширеної матриці цієї системи $\overline{a}^1, \overline{a}^2, \ldots, \overline{a}^n, B$ лінійно залежна, $rang(A|B) \le n$ і максимальна лінійно незалежна система стовпчиків розширеної матриці може бути вибрана серед стовпчиків $\overline{a}^1, \overline{a}^2, \ldots, \overline{a}^n$, тому $rang\ (A|B) = rang\ A$. ▲

Д о с т а т н і с т ь. Нехай $rang\ (A|B) = rang\ A = r$ і стовпчики $\overline{a}^1, \overline{a}^2, \ldots, \overline{a}^r$ лінійно незалежні. Тоді $\overline{a}^1, \overline{a}^2, \ldots, \overline{a}^r, B$ – лінійно залежна система, звідки за твердженням 2.9.2 $\exists\, \lambda_1, \lambda_2, \ldots, \lambda_r \in \mathbf{R}$, що

$$B = \lambda_1 \overline{a}^1 + \lambda_2 \overline{a}^2 + \ldots + \lambda_r \overline{a}^r,$$

а значить $B = \lambda_1 \overline{a}^1 + \lambda_2 \overline{a}^2 + \ldots + \lambda_r \overline{a}^r + 0\overline{a}^{r+1} + \ldots + 0\overline{a}^n$, тобто $(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$ – розв'язок рівняння (2.11.1) та системи (2.6.1). ■

Зауважимо, що розширена матриця $(A|B)$ системи лінійних рівнянь (2.6.1) може бути приведена до матриці $(C|D)$ виду (2.6.3), де матриця C східчаста і $rang\ A = rang\ C$. Так як ранг східчастої матриці дорівнює кількості ненульових рядочків, тобто кількості сходинок, а головні змінні у випадку сумісної системи вибираються „з кожної сходинки по одній", то $rang\ C =$ „кількості головних змінних". Врахувавши одержане зауваження, твердження 2.6.5 та теорему 2.11.1 маємо

[1] Цю теорему 2.11.1 часто називають теоремою Кронекера-Капеллі.

Твердження 2.11.2. (Критерій визначеності). *Система лінійних рівнянь визначена тоді і тільки тоді, коли ранг матриці системи дорівнює рангу розширеної матриці і кількості невідомих.*

Зауважимо, що ранг матриці системи взагалі може бути рівним або меншим на одиницю від рангу розширеної матриці системи і не перевищувати кількість невідомих. У випадку, коли матриця системи лінійних рівнянь квадратна, її порядок дорівнює кількості невідомих n та існує її визначник. Якщо ця матриця невироджена, то за наслідком 2.10.8 її ранг дорівнює n. Розширена матриця цієї системи має n рядочків, значить її ранг не може перевищувати n, а, з огляду на попереднє зауваження, її ранг дорівнює n, як і ранг матриці системи. Таким чином, система сумісна. З врахуванням попереднього твердження 2.11.2, одержуємо інший критерій визначеності.

Твердження 2.11.3. *Квадратна система лінійних рівнянь визначена тоді і тільки тоді, коли матриця системи невироджена.*

Спираючись на даний критерій, тепер можна стверджувати, що системи, які розв'язуються методом Крамера або матричним методом, мають єдиний розв'язок.

Контрольні питання та завдання

1. Сформулюйте критерій сумісності системи лінійних рівнянь.
2. Чи сумісна система рівнянь $x+y=1, \quad 3x+3y=4$?
3. Сформулюйте критерій визначеності системи лінійних рівнянь.
4. Як пов'язані визначеність квадратної системи і невиродженість її матриці?

2.12. Однорідні системи рівнянь. ФСР

Система лінійних рівнянь називається ***однорідною***, якщо її стовпчик вільних членів складається з нулів.

В матричній формі така система має вигляд $AX=O$. Вона завжди сумісна, бо $rang(A)=rang(A\,|\,O)$. Нульовий розв'язок завжди є розв'язком цієї системи. Тому при розв'язуванні однорідної системи питання полягає в тому, чи має система ненульові розв'язки. Також можна зауважити, якщо однорідна система має хоч один ненульовий розв'язок, то вона невизначена, а значить має безліч розв'язків. Неважко помітити, що розв'язки системи з n невідомими можна розглядати як вектори арифметичного n- мірного векторного простору з існуючими там операціями додавання векторів і множення їх на число. В цьому розумінні розв'язки задовольняють наступним твердженням.

Теорема 2.12.1. (Про розв'язки однорідної системи) ***1.*** *Сума двох довільних розв'язків однорідної системи є розв'язком цієї системи. 2.Довільний розв'язок однорідної системи, помножений на число, є розв'язком цієї системи.*

Д о в е д е н н я. 1. Нехай α_1, α_2 – розв'язки системи $AX = O$. Тоді $A\alpha_1 = O$ і $A\alpha_2 = O$. Звідки $A\alpha_1 + A\,\alpha_2 = A(\alpha_1 + \alpha_2) = O$, тобто $\alpha_1 + \alpha_2$ – розв'язок системи $AX = O$.

2. Нехай α – розв'язок системи $AX = O$. Тоді $A\alpha = O$ і $\forall\,\lambda \in R$ маємо $A(\lambda\alpha) = \lambda(A\alpha) = \lambda O = O$, тобто $\lambda\alpha$ – розв'язок системи $AX = O$. ■

Розповсюдження цієї теореми на довільну кількість доданків приводить до наступного наслідку.

Наслідок 2.12.2. *Кожна лінійна комбінація розв'язків однорідної системи є розв'язком цієї системи.*

Нехай $AX = O$ – однорідна система лінійних рівнянь з n невідомими і $rang\ A = r \quad (r \le n)$. Дана система може бути приведена до еквівалентної їй виду $CX = O$, де C – східчаста матриця, причому, як зауважено в попередньому пункті, $rang\ A = rang\ C =$ „кількості головних змінних” $= r$. За твердженням 2.11.2 зрозуміло, що при $r = n$ система $AX = O$ має тільки нульовий розв'язок (визначена), а також маємо

Твердження 2.12.3. *Однорідна система невизначена тоді і тільки тоді, коли ранг її матриці менший кількості невідомих.*

Максимальна лінійно незалежна сукупність розв'язків однорідної системи лінійних рівнянь називається ***фундаментальною сукупністю розв'язків (ФСР)*** цієї системи. Зрозуміло, якщо ранг матриці однорідної системи дорівнює кількості невідомих, то фундаментальної сукупності розв'язків не існує.

Лема 2.12.4. *Якщо* $\alpha_1, \alpha_2, ..., \alpha_p$ *– лінійно незалежна сукупність векторів, а кожний вектор сукупності* $\beta_1, \beta_2, ..., \beta_m\ (m > p)$ *є лінійною* *комбінацією сукупності векторів $\alpha_1, \alpha_2, ..., \alpha_p$, то сукупність $\beta_1, \beta_2, ..., \beta_m$ – лінійно залежна.*

Д о в е д е н н я. Доведемо лінійну залежність $\beta_1, \beta_2, ..., \beta_m$ $(m > p)$, а саме, що векторне рівняння

$$\sum_{i=1}^{m} y_i \beta_i = \theta (1)$$

має ненульові розв'язки.

За умовою $\exists \lambda_{ji} \in \mathbf{R}$, $j = 1,2,...,p$, що $\beta_i = \sum_{j=1}^{p} \lambda_{ji} \alpha_j$. Тоді рівняння (1) має вигляд $\sum_{i=1}^{m} y_i \sum_{j=1}^{p} \lambda_{ji} \alpha_j = \theta$,

що відповідає рівнянню з лінійною комбінацією $\alpha_1, \alpha_2, ..., \alpha_p$:

$$\sum_{j=1}^{p} (\sum_{i=1}^{m} \lambda_{ji} y_i) \alpha_j = \theta . \qquad (2)$$

За означенням $\alpha_1, \alpha_2, ..., \alpha_p$ лінійно незалежні, тому рівняння (2) може виконуватись тільки при нульових коефіцієнтах, тобто повинна виконуватись система рівнянь:

$$\sum_{i=1}^{m} \lambda_{ji} y_i = 0 , \; j = 1,2,...,p . \qquad (3)$$

Матриця цієї однорідної системи з m невідомими має розміри $(p \times m)$, тому її ранг не перевищує $\min(p, m)$, а так як $m > p$, то ранг системи (3) менший кількості невідомих і за твердженням 2.12.3 ця система, а значить і рівняння (1), мають ненульові розв'язки, що доводить лінійну залежність системи $\beta_1, \beta_2, ..., \beta_m$. ■

Теорема 2.12.5. (Про фундаментальну сукупність розв'язків) *Нехай r – ранг матриці однорідної системи лінійних рівнянь з n невідомими. Якщо $r < n$, то кожна фундаментальна сукупність розв'язків однорідної системи складається з $(n - r)$ розв'язків $\alpha_1, \alpha_2, ..., \alpha_{n-r}$ однорідної системи.*

Д о в е д е н н я. Нехай система $AX = O$ приводиться до еквівалентної їй системи $CX = O$, де C – східчаста матриця, яка, як зауважено вище, має r ненульових рядків. Для зручності, нехай система має вид:

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1(r-1)}x_{r-1} + a_{1r}x_r + a_{1(r+1)}x_{r+1} + \ldots a_{1n}x_n = 0 \\
a_{22}x_2 + \ldots + a_{2(r-1)}x_{r-1} + a_{2r}x_r + a_{2(r+1)}x_{r+1} + \ldots a_{2n}x_n = 0 \\
\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
a_{(r-1)(r-1)}x_{r-1} + a_{(r-1)r}x_r + a_{(r-1)(r+1)}x_{r+1} + \ldots a_{(r-1)n}x_n = 0 \\
a_{rr}x_r + a_{r(r+1)}x_{r+1} + \ldots a_{rn}x_n = 0
\end{cases} \qquad (2.12.1)
$$

де $x_1, x_2, ..., x_r$ – головні змінні, $x_{r+1}, x_{r+2}, ..., x_n$ – незалежні.

Задамо довільний визначник $\Delta \neq 0$ порядку $(n-r)$:

$$
\Delta = \begin{vmatrix} c_{1(r+1)} & c_{1(r+2)} & \cdots & c_{1n} \\ c_{2(r+1)} & c_{2(r+2)} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{(n-r)(r+1)} & c_{(n-r)(r+2)} & \cdots & c_{(n-r)n} \end{vmatrix}.
$$

1. Замість незалежних змінних візьмемо відповідні значення першого рядка визначника Δ: $x_{r+1} = c_{1(r+1)}$, $x_{r+2} = c_{1(r+2)}$, ... , $x_n = c_{1n}$.

Із системи (2.12.1) знайдемо відповідні значення головних змінних та присвоїмо їм позначення:

$$
x_r = -\frac{a_{r(r+1)}}{a_{rr}} c_{1(r+1)} - ... - \frac{a_{rn}}{a_{rr}} c_{1n} = c_{1r},
$$

$$
x_{r-1} = -\frac{a_{(r-1)r}}{a_{(r-1)(r-1)}} c_{1r} - \frac{a_{(r-1)(r+1)}}{a_{(r-1)(r-1)}} c_{1(r+1)} - ... - \frac{a_{(r-1)n}}{a_{(r-1)(r-1)}} c_{1n} = c_{1(r-1)},
$$

. . . .

$$
x_1 = -\frac{a_{12}}{a_{11}} c_{12} - \frac{a_{13}}{a_{11}} c_{13} - ... - \frac{a_{1n}}{a_{11}} c_{1n} = c_{11}.
$$

Тоді $\alpha_1 = (c_{11}, c_{12}, ..., c_{1r}, c_{1(r+1)}, ..., c_{1n})$ – розв'язок систем (2.12.1) і $AX = O$.

2. Проведемо аналогічний процес:

1) візьмемо значення 2-го рядка визначника Δ:

$$
x_{r+1} = c_{2(r+1)},\ x_{r+2} = c_{2(r+2)},\ ....\ ,\ x_n = c_{2n}.
$$

2) знайдемо відповідні значення головних змінних:

$$
x_r = c_{2r},\ x_{r-1} = c_{2(r-1)},\ ....\ ,\ x_2 = c_{22},\ x_1 = c_{21}
$$

3) одержимо розв'язок систем (2.12.1) і $AX = O$.

$$
\alpha_2 = (c_{21}, c_{22}, ..., c_{2r}, c_{2(r+1)}, ..., c_{2n})
$$

Аналогічно одержимо розв'язки систем (2.12.1) і $AX = O$:

$$
\alpha_3, \alpha_4, ..., \alpha_{n-r}.
$$

Доведемо, що одержані розв’язки $\alpha_1,\alpha_2,...,\alpha_{n-r}$ утворюють лінійно незалежну сукупність. Для цього складемо з них матрицю T розв’язків-рядків:

$$T=\begin{pmatrix} c_{11} & \dots & c_{1r} & c_{1(r+1)} & \dots & c_{1n} \\ c_{21} & \dots & c_{2r} & c_{2(r+1)} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{(n-r)1} & \dots & c_{(n-r)r} & c_{(n-r)(r+1)} & \dots & c_{(n-r)n} \end{pmatrix}.$$

Мінор $M=M_{12...(n-r)}^{(r+1)...n}$ матриці T дорівнює $\Delta\neq 0$, має порядок $(n-r)$, а мінору матриці T більшого порядку не існує. Тому M – базисний мінор матриці T, $rang\,T=n-r$, а за наслідком 2.10.7 рядки $\alpha_1,\alpha_2,...,\alpha_{n-r}$ утворюють лінійно незалежну сукупність.

Нехай $\beta=(b_1,b_2,...,b_r,b_{r+1},...,b_n)$ – довільний розв’язок системи $AX=O$. Покажемо, що сукупність $(n-r+1)$ розв’язків $\alpha_1,\alpha_2,...,\alpha_{n-r},\beta$ системи $AX=O$ лінійно залежна.

Позначимо $\alpha_1'=(c_{1(r+1)},c_{1(r+2)},...,c_{1n})$,

$\alpha_2'=(c_{2(r+1)},c_{2(r+2)},...,c_{2n})$,

....

$\alpha_{n-r}'=(c_{(n-r)(r+1)},c_{(n-r)(r+2)},...,c_{(n-r)n})$,

$\beta'=(b_{r+1},b_{r+2},...,b_n)$.

Складемо матрицю X з цих векторів-рядочків:

$$X=\begin{pmatrix} c_{1(r+1)} & c_{1(r+2)} & \dots & c_{1n} \\ c_{2(r+1)} & c_{2(r+2)} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{(n-r)(r+1)} & c_{(n-r)(r+2)} & \dots & c_{(n-r)n} \\ b_{r+1} & b_{r+2} & \dots & b_n \end{pmatrix}.$$

Мінор $M_1=M_{12...(n-r)}^{1,2,...,(n-r)}$ матриці X дорівнює $\Delta\neq 0$ і має максимально можливий порядок $(n-r)$. Тому M_1 – базисний мінор матриці X, $rang\,X=n-r$ і рядки $\alpha_1',\alpha_2',...,\alpha_{n-r}'$ – лінійно незалежні, а рядки $\alpha_1',\alpha_2',...,\alpha_{n-r}',\beta'$ – лінійно залежні. За твердженням 2.9.2 $\exists\lambda_i\in\mathbf{R}$, $i=1,2,...,n-r$, що

$\beta'=\lambda_1\alpha_1'+\lambda_2\alpha_2'+...+\lambda_{n-r}\alpha_{n-r}'\Rightarrow$

$\lambda_1\alpha_1'+\lambda_2\alpha_2'+...+\lambda_{n-r}\alpha_{n-r}'+(-1)\beta'=\theta$. (2.12.2)

Рівність (2.12.2) у векторно-стовпчиковій формі має вигляд

$$\lambda_1\begin{pmatrix}c_{1(r+1)}\\ \dots\\ c_{1n}\end{pmatrix}+\lambda_2\begin{pmatrix}c_{2(r+1)}\\ \dots\\ c_{2n}\end{pmatrix}+\dots+\lambda_{n-r}\begin{pmatrix}c_{(n-r)(r+1)}\\ \dots\\ c_{(n-r)n}\end{pmatrix}+(-1)\begin{pmatrix}b_{r+1}\\ \dots\\ b_n\end{pmatrix}=\begin{pmatrix}0\\ \dots\\ 0\end{pmatrix} \quad (2.12.3)$$

Розглянемо лінійну комбінацію δ розв'язків $\alpha_1,\alpha_2,...,\alpha_{n-r},\beta$ зі знайденими коефіцієнтами $\lambda_1,\lambda_2,...,\lambda_{n-r}$:

$$\delta=\lambda_1\alpha_1+\lambda_2\alpha_2+...+\lambda_{n-r}\alpha_{n-r}+(-1)\beta=(d_1,d_2,...,d_r,d_{r+1},...,d_n). \quad (2.12.4)$$

Рівність (2.12.4) перепишемо у векторно-стовпчиковій формі з врахуванням покоординатних рівностей (2.12.3):

$$\delta=\lambda_1\begin{pmatrix}c_{11}\\ \dots\\ c_{1r}\\ c_{1(r+1)}\\ \dots\\ c_{1n}\end{pmatrix}+\lambda_2\begin{pmatrix}c_{21}\\ \dots\\ c_{2r}\\ c_{2(r+1)}\\ \dots\\ c_{2n}\end{pmatrix}+\dots+\lambda_{n-r}\begin{pmatrix}c_{(n-r)1}\\ \dots\\ c_{(n-r)r}\\ c_{(n-r)(r+1)}\\ \dots\\ c_{(n-r)n}\end{pmatrix}+(-1)\begin{pmatrix}b_1\\ \dots\\ b_r\\ b_{r+1}\\ \dots\\ b_n\end{pmatrix}=\begin{pmatrix}d_1\\ \dots\\ d_r\\ d_{r+1}\\ \dots\\ d_n\end{pmatrix}=\begin{pmatrix}d_1\\ \dots\\ d_r\\ 0\\ \dots\\ 0\end{pmatrix}.$$

За наслідком 2.12.2 випливає, що δ – розв'язок системи $AX=O$ і еквівалентної їй (2.12.1), а числа $d_1,d_2,...,d_r$ – значення головних змінних $x_1,x_2,...,x_r$ при незалежних змінних $x_{r+1}=d_{r+1}=0$, ... , $x_n=d_n=0$. Підставляючи ці значення незалежних змінних в r-е рівняння системи (2.12.1), одержимо $x_r=d_r=0$, в $(r-1)$-е рівняння – одержимо $d_{r-1}=0$, а також $d_{r-2}=0$,..., $d_2=0$, $d_1=0$, звідки $\delta=(0,...,0,0,...,0)=\theta$.

Підставивши одержане в (2.12.4), будемо мати

$$\beta=\lambda_1\alpha_1+\lambda_2\alpha_2+...+\lambda_{n-r}\alpha_{n-r},$$

тобто одержуємо висновок, що *довільний розв'язок однорідної системи* $AX=O$ *лінійно виражається через* $\alpha_1,\alpha_2,...,\alpha_{n-r}$.

Тоді за лемою 2.12.4 довільна множина $\beta_1,\beta_2,...,\beta_m$ $(m>n-r)$ розв'язків системи $AX=O$ буде лінійно залежною, отже, $\alpha_1,\alpha_2,...,\alpha_{n-r}$ – ФСР (фундаментальна сукупність розв'язків). ■

В процесі доведення попередньої теореми, ми одержали

Твердження 2.12.6. Д*овільний розв'язок однорідної системи лінійно виражається через фундаментальну сукупність розв'язків.*

В силу цього твердження та наслідку 2.12.2 очевидно, що для знаходження загального розв'язку (множини всіх розв'язків) однорідної системи достатньо знайти ФСР $\alpha_1,\alpha_2,...,\alpha_{n-r}$. Тоді загальний розв'язок записується у вигляді довільних лінійних комбінацій ФСР:

$$F=\{\ c_1\alpha_1+c_2\alpha_2+...+c_{n-r}\alpha_{n-r},\quad c_i\in\mathbf{R},\ i=1,2,...,n-r\ \}. \quad (2.12.5)$$

Алгоритм знаходження фундаментальної сукупності розв'язків

1. Матриця системи $AX=O$ приводиться до східчастої форми A', яка має r ненульових рядків $(rang\ A=rang\ A'=r)$. Отримаємо систему $A'X=O$.

2. За допомогою нескладного аналізу матриці A' вибираємо „з кожної сходинки по одній" r головних та решту $n-r$ незалежних змінних. Нехай $x_1,...,x_r$ – головні, а $x_{r+1},...,x_n$ – незалежні змінні системи.

3. Складається таблиця (2.12.6) з $n-r$ рядків:

x	x_1	x_2	...	x_r	x_{r+1}	x_{r+2}	...	x_n
α_1			...		1	0	...	0
α_2			...		0	1	...	0
...	...	...	...	...	...	...	...	...
α_{n-r}					0	0	...	1

(2.12.6)

де під $n-r$ незалежними змінними вписується одинична матриця.

4. Значення незалежних змінних $x_{r+1},x_{r+2},...,x_n$, що відповідають першому рядку таблиці (2.12.6) – $1,0,0,...0$, підставляємо в систему $A'X=O$. З неї, починаючи з останнього рівняння, почергово знаходяться і вносяться в перший рядок таблиці (2.12.6) значення головних змінних

$$x_r=x_r^1,...,x_2=x_2^1,x_1=x_1^1.$$

Вектор $\alpha_1=\left(x_1^1,x_2^1,...,x_r^1,1,0,0,...,0\right)$ буде першим розв'язком ФСР.

Потім замість незалежних змінних в систему $A'X=O$ підставляються відповідні значення другого рядка таблиці (2.12.6): $0,1,0,...,0$ і аналогічно знаходимо і вносимо в таблицю другий розв'язок ФСР:

$$\alpha_2=\left(x_1^2,x_2^2,...,x_r^2,0,1,0,...,0\right).$$

Так продовжуємо до $(n-r)$-го розв'язку α_{n-r}.

Таблиця буде мати вигляд:

x	x_1	x_2	...	x_r	x_{r+1}	x_{r+2}	...	x_n
α_1	x_1^1	x_2^1	...	x_r^1	1	0	...	0
α_2	x_1^2	x_2^2	...	x_r^2	0	1	...	0
...	...	...	...	...	...	...	...	...
α_{n-r}	x_1^{n-r}	x_2^{n-r}	...	x_r^{n-r}	0	0	...	1

(2.12.7)

Одержані вектори $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ утворюють ФСР однорідної системи $AX = O$.

5. Якщо необхідно знайти загальний розв'язок системи $AX = O$ через ФСР, то його можна записати у вигляді (2.12.5):

$$F = \sum_{i=1}^{n-r} c_i \alpha_i = \{\ (c_1 x_1^1 + c_2 x_1^2 + \ldots + c_{n-r} x_1^{n-r},\ c_1 x_2^1 + c_2 x_2^2 + \ldots + c_{n-r} x_2^{n-r}, \ldots,$$

$$c_1 x_r^1 + c_2 x_r^2 + \ldots + c_{n-r} x_r^{n-r}, c_1, c_2, \ldots, c_{n-r}), \quad c_i \in \mathbf{R}, i = 1,2,\ldots, n-r\ \}.$$

Зауваження. *У таблиці (2.12.6) під незалежними змінними можна вписувати довільну матрицю, визначник якої не дорівнює нулю. Наприклад, з метою уникнення дробових значень змінних, замість одиниць можна вписувати довільне число, відмінне від нуля.*

Приклад. Знайти ФСР і загальний розв'язок системи рівнянь

$$\begin{cases} 2x_1 - x_2 + 3x_3 - 2x_4 + 4x_5 = 0, \\ 4x_1 - 2x_2 + 5x_3 + x_4 + 7x_5 = 0, \\ 2x_1 - x_2 + x_3 + 8x_4 + 2x_5 = 0. \end{cases}$$

▼ Скористаємося вказаним алгоритмом знаходження ФСР.

1. Матрицю системи приведемо до східчастої форми:

$$A = \begin{pmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{pmatrix} \begin{matrix} (-2)(-1) \\ \hookleftarrow \quad \\ \quad \hookleftarrow \end{matrix} \approx \begin{pmatrix} 2 & -1 & 3 & -2 & 4 \\ 0 & 0 & -1 & 5 & -1 \\ 0 & 0 & -2 & 10 & -2 \end{pmatrix} \begin{matrix} \\ (2) \\ \hookleftarrow \end{matrix} \approx$$

$$\approx \begin{pmatrix} 2 & -1 & 3 & -2 & 4 \\ 0 & 0 & -1 & 5 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A',\ r = rang\ A = rang\ A' = 2.$$

2. З кожної сходинки матриці A' вибираємо дві головних та 5 – 2 = 3 незалежних змінних. Нехай x_2, x_3 – головні (так як на першій і другій сходинці коефіцієнти при них дорівнюють –1), а x_1, x_4, x_5 – незалежні змінні системи.

3. Складемо таблицю з трьох рядків, де під незалежними змінними вписується одинична матриця:

x	x_1	x_2	x_3	x_4	x_5
α_1	1			0	0
α_2	0			1	0
α_3	0			0	1

(1)

4. Випишемо систему $A'X = O$:

$$\begin{cases} 2x_1 & -x_2 & +3x_3 & -2x_4 & +4x_5 & = & 0, \\ & & -x_3 & +5x_4 & -x_5 & = & 0. \end{cases}$$

Значення незалежних змінних x_1, x_4, x_5, що відповідають першому рядку таблиці (1) – $1,0,0$, підставляємо в цю систему. З неї, починаючи з останнього рівняння, почергово знаходяться значення головних змінних x_3, x_2 і вносяться в перший рядок таблиці (1):

$$x_3 = 5x_4 - x_5 = 5 \cdot 0 - 0 = 0,$$

$$x_2 = 2x_1 + 3x_3 - 2x_4 + 4x_5 = 2 \cdot 1 + 3 \cdot 0 - 2 \cdot 0 + 4 \cdot 0 = 2.$$

Вектор $\alpha_1 = (1, 2, 0, 0, 0)$ буде першим розв'язком фундаментальної сукупності.

Потім замість незалежних змінних x_1, x_4, x_5 в систему $A'X = O$ підставляються відповідні значення другого рядка таблиці (1) $-0, 1, 0$ і аналогічно знаходимо і вносимо в таблицю другий розв'язок фундаментальної сукупності:

$$x_3 = 5x_4 - x_5 = 5 \cdot 1 - 0 = 5,$$

$$x_2 = 2x_1 + 3x_3 - 2x_4 + 4x_5 = 2 \cdot 0 + 3 \cdot 5 - 2 \cdot 1 + 4 \cdot 0 = 13.$$

Вектор $\alpha_2 = (0, 13, 5, 1, 0)$ буде другим розв'язком фундаментальної сукупності.

Третій розв'язок фундаментальної сукупності знаходиться аналогічно, підставляючи замість незалежних змінних x_1, x_4, x_5 в систему $A'X = O$ значення третього рядка таблиці (1) $-0, 0, 1$:

$$x_3 = 5x_4 - x_5 = 5 \cdot 0 - 1 = -1,$$

$$x_2 = 2x_1 + 3x_3 - 2x_4 + 4x_5 = 2 \cdot 0 + 3 \cdot (-1) - 2 \cdot 0 + 4 \cdot 1 = 1.$$

І в третьому рядку таблиці (1) буде записано розв'язок фундаментальної сукупності $\alpha_3 = (0, 1, -1, 0, 1)$. Таблиця (1) буде мати вигляд:

x	x_1	x_2	x_3	x_4	x_5
α_1	1	2	0	0	0
α_2	0	13	5	1	0
α_3	0	1	–1	0	1

Отже, ФСР, яка складається із розв'язків $\alpha_1, \alpha_2, \alpha_3$, побудована.

5. Запишемо тепер загальний розв'язок даної однорідної системи:

$$F=\left\{ c_1\alpha_1+c_2\alpha_2+c_3\alpha_3 \;\middle|\; c_1,c_2,c_3\in\mathbf{R}\right\}=$$
$$=\left\{(c_1,\,2c_1+13c_2+c_3,\,5c_2-c_3,\,c_2,\,c_3),\quad c_1,c_2,c_3\in\mathbf{R}\right\}.\ \blacktriangle$$

Неоднорідну систему лінійних рівнянь з n невідомими $AX=B$ можна розв'язати, використовуючи фундаментальну систему відповідної однорідної системи $AX=O$, що ми і розглянемо далі.

Теорема 2.12.7. *Нехай β – частинний розв'язок системи $AX=B$. Для довільного розв'язку α відповідної однорідної системи $AX=O$ сума $\beta+\alpha$ також є розв'язком системи $AX=B$.*

Д о в е д е н н я. Так як виконуються рівності $A\beta=B$ і $A\alpha=O$, то сума їх лівих і правих частин перетвориться на істинну рівність $A(\beta+\alpha)=B$, звідки $\beta+\alpha$ – розв'язок $AX=B$. ■

Теорема 2.12.8. *Нехай β – частинний розв'язок системи $AX=B$. Для довільного розв'язку γ системи $AX=B$ існує такий розв'язок α відповідної однорідної системи $AX=O$, що $\gamma=\beta+\alpha$.*

Д о в е д е н н я. Так як виконуються рівності $A\gamma=B$ і $A\beta=B$, то різниця їх правих і лівих частин приведе до істинного рівняння $A(\gamma-\beta)=O$, тобто різниця $\alpha=\gamma-\beta$ є деяким розв'язком однорідної системи $AX=O$, звідки і одержуємо бажане: $\gamma=\beta+(\gamma-\beta)=\beta+\alpha$. ■

З цих теорем випливає, що умовно можна говорити, що загальний розв'язок системи $AX=B$ є сумою деякого її частинного розв'язку β і загального розв'язку F відповідної однорідної системи $AX=O$:

$\Phi=\beta+F$. (2.12.8)

<u>Приклад.</u> Знайти загальний розв'язок системи рівнянь

$$\begin{cases} 2x_1 - x_2 + 3x_3 - 2x_4 + 4x_5 = 1, \\ 4x_1 - 2x_2 + 5x_3 + x_4 + 7x_5 = 1, \\ 2x_1 - x_2 + x_3 + 8x_4 + 2x_5 = -1. \end{cases}$$

▼ 1. Розширену матрицю даної системи приведемо до східчастої матриці:

$$
(A|B) = \left(\begin{array}{ccccc|c} 2 & -1 & 3 & -2 & 4 & 1 \\ 4 & -2 & 5 & 1 & 7 & 1 \\ 2 & -1 & 1 & 8 & 2 & -1 \end{array}\right)
\begin{array}{l} (-2)\,(-1) \\ \hookleftarrow \\ \qquad\hookleftarrow \end{array} \approx
$$

$$
\left(\begin{array}{ccccc|c} 2 & -1 & 3 & -2 & 4 & 1 \\ 0 & 0 & -1 & 5 & -1 & -1 \\ 0 & 0 & -2 & 10 & -2 & -2 \end{array}\right)
\begin{array}{l} \\ (2) \approx \\ \hookleftarrow \end{array}
\qquad \approx \left(\begin{array}{ccccc|c} 2 & -1 & 3 & -2 & 4 & 1 \\ 0 & 0 & -1 & 5 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) = (A'|B'),
$$

$r = rang\ A = rang\ A' = 2$

2. Так як $rang\ (A|B) = rang\ (A'|B') = 2$ і $rang\ A = 2$, то система сумісна. Як і в попередньому прикладі, x_2, x_3 – головні змінні, а x_1, x_4, x_5 – незалежні змінні системи.

3. Складаємо і заповнюємо таблицю з $n - r + 1 = 5 - 2 + 1 = 4$ рядків. Перші $n - r = 3$ рядки заповнюються значеннями ФСР відповідної однорідної системи як і в попередньому прикладі.

x	x_1	x_2	x_3	x_4	x_5	Зауваження
α_1	1	2	0	0	0	
α_2	0	13	5	1	0	$A'X = O$
α_3	0	1	–1	0	1	
β	0	2	1	0	0	$A'X = B'$

Щоб знайти частинний розв’язок β та заповнити останній рядок таблиці, замість незалежних змінних x_1, x_4, x_5 можна взяти довільні значення і записати їх у стовпчиках під відповідними змінними, а потім із системи $A'X = B'$, починаючи з останнього рівняння, знайти x_3, x_2. Якщо покласти $x_1 = x_4 = x_5 = 0$, то $x_3 = 1 + 5x_4 - x_5 = 1 + 5 \cdot 0 - 0 = 1$,

$$x_2 = -1 + 2x_1 + 3x_3 - 2x_4 + 4x_5 = -1 + 2 \cdot 0 + 3 \cdot 1 - 2 \cdot 0 + 4 \cdot 0 = 2,$$

а частинний розв’язок $\beta = (0, 2, 1, 0, 0)$ і саме він вноситься в таблицю.

4. Знаходимо загальний розв’язок системи $AX = B$ за формулою (2.12.8):

$$\Phi = \beta + F = \{\ (c_1, 2 + 2c_1 + 13c_2 + c_3, 1 + 5c_2 - c_3, c_2, c_3),\quad c_1, c_2, c_3 \in \mathbf{R}\ \}. \ \blacktriangle$$

Контрольні питання та завдання

1. Чи може однорідна система рівнянь бути несумісною?

2. У вигляді елементів якої множини записуються розв'язки однорідної системи лінійних рівнянь з n невідомими?

3. Що називається фундаментальною сукупністю розв'язків однорідної системи лінійних рівнянь?

4. Скільки ФСР має однорідна система лінійних рівнянь?

5. Напишіть формулу, яка описує загальний розв'язок однорідної системи через ФСР.

6. Чи може неоднорідна система лінійних рівнянь бути несумісною?

7. Чи буде сума розв'язків неоднорідної системи лінійних рівнянь розв'язком цієї системи?

Розділ 3. ЛІНІЙНІ ПРОСТОРИ

3.1. Поняття поля

Нехай M – довільна множина елементів a, b, c, …. ***Бінарною алгебраїчною операцією*** на множині M називається закон, по якому кожним двом елементам a і b цієї множини, що беруться у певному порядку, ставиться у відповідність однозначно визначений елемент c цієї множини.

Бінарними алгебраїчними операціями на множині цілих чисел будуть, наприклад, операції додавання, віднімання, множення. А от на множині натуральних чисел операція віднімання не буде алгебраїчною, оскільки закон віднімання парі чисел 5 і 7 ставить у відповідність число –2 (5 – 7 = –2), яке не належить множині натуральних чисел. Цьому висновку не суперечить те, що, наприклад, парі (7, 5) закон віднімання ставить у відповідність натуральне число 2 (7 – 5 = 2), оскільки, за означенням, *кожній парі ставиться у відповідність елемент цієї ж множини.* Виділена остання умова називається ***замкненістю бінарної алгебраїчної операції.*** Надалі будемо розглядати бінарні алгебраїчні операції і називати просто операцією.

Операція $*$ на множині M називається ***комутативною***, якщо $\forall a,b \in M \;\big|\; a*b = b*a$.

Операція $*$ на множині M називається ***асоціативною***, якщо $\forall a,b,c \in M \;\big|\; (a*b)*c = a*(b*c)$.

Множина M з визначеною на ній асоціативною операцією називається ***напівгрупою***.

Елемент $e \in M$ називається ***нейтральним*** відносно операції $*$, якщо $\forall a \in M$ виконуються рівності $a*e = a$ і $e*a = a$.

Елемент $a' \in M$ називається ***симетричним*** для $a \in M$, якщо $a*a' = a'*a = e$.

Множина M з визначеною на ній (замкненою) операцією $(*)$ називається ***групою***, якщо операція задовольняє умовам:

1) $(*)$ – асоціативна,

2) існує нейтральний елемент $e \in M$ відносно $(*)$: $\forall\, a \in M$ $e*a = a*e = a$,

3) $\forall\, a \in M$ існує симетричний елемент $a' \in M$: $a*a' = a'*a = e$.

Група з комутативною операцією називається ***абелевою*** або ***комутативною***.

Наприклад, на множині натуральних чисел з нулем операція додавання буде алгебраїчною, оскільки сума довільних натуральних чисел є число натуральне, асоціативною та комутативною, оскільки відомі закони: $\forall m,n,k \in \mathbf{N_0}$ $(m+n)+k = m+(n+k)$ та $m+n = n+m$. В даній множині міститься нейтральний елемент 0, оскільки $\forall m \in \mathbf{N_0}$ $m+0 = m$. Симетричних елементів ні для якого натурального числа не існує, бо розв'язок рівняння $m+x=0$ не належить множині натуральних чисел $\forall m \neq 0$. Отже, ($\mathbf{N_0}, +$) не є групою.

На відміну від множини натуральних чисел з нулем множина цілих чисел відносно додавання утворює абелеву групу ($\mathbf{Z}$, +):

0) операція (+) замкнена на множині $\mathbf{Z}$;

1) операція (+) асоціативна на множині $\mathbf{Z}$;

2) нейтральний елемент $0 \in \mathbf{Z}$: $\forall a \in \mathbf{Z}$ $a+0=0+a=a$;

3) $\forall a \in \mathbf{Z}$ існує симетричний елемент $a' = -a \in \mathbf{Z}$: $a+a' = a+(-a) = 0$;

4) операція (+) комутативна на множині $\mathbf{Z}$.

Аналогічно перевіряється, що абелеву групу відносно додавання утворюють множини $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, $M_{m\times n}$. Також неважко перевірити, що ($\mathbf{N}, \cdot$), ($\mathbf{Z}, \cdot$), ($M_n, \cdot$) не утворюють груп, бо, хоча у них і виконуються умови 0) –3), вони мають нейтральний елемент e відносно ($\cdot$) ($e=1$: $a\cdot 1 = a$ $\forall a \in \mathbf{N}, \mathbf{Z}$ $\left[e = E : AE = A \ \ \forall A \in M_n\right]$), але не мають симетричного ($\forall a \neq 1 \left[|A| \neq 0\right]$ рівняння $a \cdot x = 1 \left[AX = E\right]$ не має розв'язку ні в $\mathbf{N}$, ні в $\mathbf{Z}$, ні в M_n). А, на відміну від них, ($\mathbf{Q}_{\mathbf{-0}}, \cdot$), ($\mathbf{R}_{\mathbf{-0}}, \cdot$), ($\mathbf{C}_{\mathbf{-0}}, \cdot$) утворюють комутативні групи.

Існують традиційні назви та позначення нейтральних та симетричних елементів відносно операцій додавання та множення, а саме відносно додавання (+) нейтральний (e) називають ***нульовим*** і позначають 0, а симетричний (a') називають ***протилежним*** і позначають $-a$, відносно множення ($\cdot$) нейтральний (e) називають ***одиничним*** і позначають 1, а симетричний (a') називають ***оберненим*** і позначають a^{-1}.

Не порожня множина P, на якій задано операції (+) і ($\cdot$) називається ***полем***, якщо виконуються наступні умови:

I $(P,+)$ комутативна група:

1) (+) – асоціативне,

2) (+) – комутативне,

3) нульовий $0 \in P$,

4) $\forall a \in P\ \exists -a \in P$.

II $(P_{-0},\cdot)$ – комутативна група:

5) ($\cdot$) – асоціативне,

6) одиничний $1 \in P$,

7) $\forall a \neq 0 \in P,\ \exists a^{-1} \in P$,

8) ($\cdot$) – комутативне.

III Виконується дистрибутивний закон:

9) $\forall a,b,c \in P\ \ (a+b)c = ac+bc$.

Виходячи з наведених прикладів груп, можемо стверджувати, що числові множини $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ утворюють поля, оскільки для них виконуються умови 1) – 9).

Контрольні питання і завдання

1. Дайте означення групи. Які числові множини утворюють групу відносно операції: а) додавання; б) множення?

2. Що таке поле? Чи утворює поле множина раціональних чисел? А множина дійсних квадратних матриць? Чому?

3. Доведіть, що $\forall a \in (P,+,\cdot)$, де P – поле, виконуються умови: $0 \cdot a = 0,\ \ (-1)\cdot a = -a$, де 0 – нульовий елемент поля, –1 – протилежний до одиничного, $-a$ – протилежний до a.

3.2. Поняття лінійного простору

Лінійним простором L над полем P називається множина з бінарною алгебраїчною операцією (+) $(\forall l_1, l_2 \in L \mid (l_1 + l_2) \in L)$ і зовнішньою бінарною операцією $(\cdot)$ на елемент поля P $(a \cdot l \in L \mid \forall a \in P, \forall l \in L)$, які задовольняють наступним умовам:

I. $(L,+)$ – комутативна група:

1) операція (+) – асоціативна;

2) операція (+) – комутативна;

3) $\exists \theta \in L$ – нульовий елемент: $\forall l \in L \;\; l+\theta = l$;

4) $\forall l \in L, \exists l' \in L$ – протилежний елемент: $l+(l') = \theta$;

II. операція $(\cdot)$ унітарна і асоціативна:

5) $1 \cdot l = l \mid \forall l \in L$, де 1 – одиниця поля P;

6) $(\cdot)\, a(bl) = (ab)l \quad \mid \forall a, b \in P, \forall l \in L$;

III. (+) і $(\cdot)$ пов'язані дистрибутивними законами:

7) $(a+b)\cdot l = al + bl \quad \mid \forall a, b \in P, \forall l \in L$;

8) $a(l_1 + l_2) = al_1 + al_2 \quad \mid \forall a \in P, \forall l_1, l_2 \in L$.

В цьому означенні P – довільне поле. У випадку, коли P – числове поле ($\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$), то простір називається *числовим*. В загальному випадку лінійний простір L над полем P позначається L_P.

<u>Приклад 1</u>. L – множина геометричних векторів трьохмірного простору ($L = V_3$, $P = \mathbf{R}$). Покажемо, що L – дійсний лінійний простір.

▼ Нехай $\overline{a} = (a_1, a_2, a_3)$, $\overline{b} = (b_1, b_2, b_3) \in V_3$, $\lambda \in \mathbf{R}$. Покажемо замкненість операцій додавання і множення на число на множині V_3:

0_1) $\overline{a} + \overline{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3) \in V_3$.

0_2) $\lambda \overline{a} = (\lambda a_1, \lambda a_2, \lambda a_3) \in V_3$

Перевіримо виконання аксіом лінійного простору:

I. 1) операція (+) – асоціативна: $\forall \overline{a}, \overline{b}, \overline{c} \in V_3 \mid (\overline{a} + \overline{b}) + \overline{c} = \overline{a} + (\overline{b} + \overline{c})$;

2) операція (+) – комутативна: $\forall \overline{a}, \overline{b} \in V_3 \mid \overline{a} + \overline{b} = \overline{b} + \overline{a}$;

3) $\overline{0} = (0,0,0) \in V_3$ – нейтральний елемент: $\overline{a} + \overline{0} = \overline{a} \;\; |\forall \overline{a} \in V_3$;

4) $\forall \overline{a} \in V_3 \;\; \exists -\overline{a} = (-a_1, -a_2, -a_3) \in V_3$ – протилежний до $\overline{a}$: $\overline{a} + (-\overline{a}) = \overline{0}$.

II. 5) операція множення унітарна: $1 \cdot \overline{a} = \overline{a} \;\; |\forall \overline{a} \in V_3$;

6) множення асоціативне: $(\alpha\beta)\overline{a} = \alpha(\beta \overline{a}) \;\; |\forall \alpha, \beta \in \mathbf{R}, \forall \overline{a} \in V_3$;

III. Виконуються дистрибутивні закони:

7) $(\alpha + \beta)\overline{a} = ((\alpha + \beta)a_1, (\alpha + \beta)a_2, (\alpha + \beta)a_3) =$

$= (\alpha a_1, \alpha a_2, \alpha a_3) + (\beta a_1, \beta a_2, \beta a_3) = \alpha \overline{a} + \beta \overline{a} \;\; |\forall \alpha, \beta \in \mathbf{R}, \forall \overline{a} \in V_3$.

8) $\alpha(\overline{a} + \overline{b}) = (\alpha(a_1 + b_1), \alpha(a_2 + b_2), \alpha(a_3 + b_3)) = \alpha \overline{a} + \alpha \overline{b} \quad | \forall \alpha \in \mathbf{R}, \; \forall \overline{a}, \overline{b} \in V_3$.

Із 1–8 одержуємо, що L_R – лінійний простір. ▲ (Перевірте самостійно, що L_Q також лінійний простір, а L_C не є лінійним простором).

<u>Приклад 2</u>. $L = \mathbf{R}_+$, $P = \mathbf{R}$. Доведемо, що L_P – лінійний простір відносно операцій $\oplus$ і $\otimes$: $x \oplus y := x \cdot y \in L$; $\lambda \otimes x = x^\lambda \in L \,\Big|\, \forall x, y \in L, \forall \lambda \in P$.

▼ I. $(L, \oplus) = (L, \cdot)$ – абелева група (доведіть самостійно).

II. Операція ($\otimes$) унітарна і асоціативна:

а) $1 \otimes x = x^1 = x \;\Big|\; \forall x \in L$;

б) $(\lambda\mu) \otimes x = x^{\lambda\mu} = (x^\mu)^\lambda = \lambda \otimes (x^\mu) = \lambda \otimes (\mu \otimes x) \;\Big|\; \forall \lambda, \mu \in P, \forall x \in L$;

III. Виконуються дистрибутивні закони: $\forall x, y \in L, \forall \lambda, \mu \in P$

а) $(\lambda + \mu) \otimes x = x^{\lambda+\mu} = x^\lambda \cdot x^\mu = (\lambda \otimes x) \oplus (\mu \otimes x)$;

б) $\lambda \otimes (x \oplus y) = \lambda \otimes (xy) = (xy)^\lambda = x^\lambda \oplus y^\lambda = (\lambda \otimes x) \oplus (\lambda \otimes y)$.

Із I – III випливає, що L_P – дійсний лінійний простір. ▲

<u>Приклад 3</u>. $L = \{0\}$, K – довільне поле, $1 \in K$. Доведемо, що L_K – лінійний простір.

▼ Зрозуміло, що додавання і множення на такій множині повинні бути заданими тільки так: $0 + 0 = 0,\quad \lambda \cdot 0 = 0 \;\; \forall \lambda \in K$. Тоді $\forall \lambda, \mu \in K$

I. $(L, +)$ – комутативна група;

II. а) $1 \cdot 0 = 0$;

б) $(\lambda\mu) \cdot 0 = 0$, $\lambda \cdot (\mu \cdot 0) = \lambda \cdot 0 = 0$;

III. а) $(\lambda + \mu) \cdot 0 = 0 = \lambda \cdot 0 + \mu \cdot 0$;

б) $\lambda \cdot (0 + 0) = 0 = \lambda \cdot 0 + \lambda \cdot 0$.

Із означення L_K – лінійний простір (***нульовий*** лінійний простір).

В л а с т и в о с т і л і н і й н о г о п р о с т о р у

Нехай L_P – довільний лінійний простір над полем P. Нагадаємо, що довільне поле містить 0 і 1 (нейтральні відносно додавання і множення), а також -1 (протилежний до 1).

1. *Нульовий елемент* $\theta \in L$ *у лінійному просторі єдиний.*

▼ Нехай існує інший нульовий елемент $\theta_1 \in L$. Тоді $\theta_1 = \theta_1 + \theta = \theta$, бо і θ, і θ_1 може виступати в ролі нейтрального. ■

2. *Нульовий елемент* $\theta \in L$ *дорівнює добутку довільного елементу* $x \in L$ *на* $0 \in P$ ($\theta = 0 \cdot x$).

▼ $(L,+)$ – група, тому $\forall x \in L, \quad \exists x' \in L$ такий, що $x + x' = \theta$. Тоді

$$\underset{(\in L)}{0 \cdot x} \overset{3)}{=} 0 \cdot x + \theta \overset{4),5),1)}{=} (0 \cdot x + 1 \cdot x) + x' \overset{7)}{=} (0+1)x + x' = 1 \cdot x + x' \overset{5)}{=} x + x' = \theta. \blacksquare$$

3. $\forall x \in L$ *протилежний елемент* x' *єдиний.*

▼ Нехай x'' – інший протилежний до x. Тоді $x'' = x'' + \theta =$ $= x'' + (x + x') = (x'' + x) + x' = \theta + x' = x'$. $\blacksquare$

4. $\forall x \in L$ *протилежний елемент* x' *дорівнює добутку* x *на* -1 ($x' = (-1) \cdot x$):

▼ Нехай $y = (-1) \cdot x$. Тоді $x + y \overset{5)}{=} 1 \cdot x + (-1) \cdot x \overset{7)}{=} (1 + (-1))x = 0 \cdot x = \theta$, тобто $x' = y = (-1)x$ і позначають $x' = -x$. $\blacksquare$

Різницею елементів x і y називається такий елемент z, що $y + z = x$. Позначається: $z = x - y$.

5.*Різниця елементів* x *і* y *дорівнює сумі* x і $-y$: $x - y = x + (-y)$.

▼ $y + (x + (-y)) = y + (-y) + x = \theta + x = x \Rightarrow x - y = x + (-y)$. $\blacksquare$

6. *При множенні нульового елементу отримується нульовий елемент* ($\forall a \in P: \quad a \cdot \theta = \theta$).

▼ $a\theta = a(x + (-1)x) \overset{8)}{=} ax + a((-1)x) \overset{6)}{=} ax + (-a)x \overset{7)}{=}$ $(a + (-a))x = 0x = \theta$. $\blacksquare$

7. $\forall a \in P \; \forall x \in L_P$ *із рівності* $ax = \theta$ *випливає, що або* $a = 0$, *або* $x = \theta$.

▼ а) Якщо $a = 0$, то $\forall x \in L \;\; ax = \theta$ (із 1).

б) Якщо $ax = \theta, a \neq 0,$ то $\exists a^{-1} \in P$.Тоді за властивостями лінійного простору та за його означенням $\theta = a^{-1}\theta = a^{-1}(ax) = (a^{-1}a)x = 1 \cdot x = x$. $\blacksquare$

Лінійною комбінацією елементів $x, y, ..., z \in L_P$ називається сума добутків цих елементів на довільні елементи $a, b, ..., c$ поля P:

$$ax + by + ... + cz$$

Елементи $x, y, ..., z \in L_P$ називаються ***лінійно залежними (ЛЗ)***, якщо існує їх лінійна комбінація рівна нулю, в якій не всі коефіцієнти рівні нулю, тобто знайдуться такі $a, b, ..., c \in P$, не всі рівні $0 \in P$, що $ax + by + ... + cz = \theta$.

Елементи $x, y, ..., z \in L_P$ називаються ***лінійно незалежними*** (***ЛНЗ***), якщо лінійна комбінація $ax + by + ... + cz = \theta$ тоді і тільки тоді, коли $a = b = ... = c = 0$

Всі властивості ЛЗ та ЛНЗ елементів лінійного простору такі ж, як і розглянуті раніше (п. 2.9.) властивості ЛЗ та ЛНЗ векторів, мають аналогічні доведення і випливають із критерію ЛЗ.

Теорема 3.2.1 (Критерій лінійної залежності). *Для того, щоб елементи $x, y, ..., z \in L$ були лінійно залежними, необхідно і достатньо, щоб один з цих елементів був лінійною комбінацією останніх.*

Твердження 3.2.2 *Якщо в лінійному просторі над полем P елементи $a_1, a_2, ..., a_t$ – лінійно незалежні, та елементи $a_1, a_2, ..., a_t, b$ – лінійно залежні, то b лінійно виражається через лінійну комбінацію останніх, тобто $\exists \lambda_i \in P, i = 1,2,...,t$ такі, що $b = \lambda_1 a_1 + \lambda_2 a_2 + ... + \lambda_t a_t$.*

Контрольні питання і завдання

1. Сформулюйте означення лінійного простору і наведіть приклади лінійних просторів.

2. L – множина розв'язків однорідної системи лінійних рівнянь. Доведіть, що L – лінійний простір відносно додавання розв'язків і множення їх на число.

3. Доведіть, що $L = \{a_n x^n + ... + a_1 x + a_0 \mid a_i \in \mathbf{R}\} = P_n[x]$ – лінійний простір многочленів степеню не вище n.

4. Чим відрізняється раціональний лінійний простір від дійсного? Покажіть, що R_Q – раціональний лінійний простір, а R_R – дійсний лінійний простір.

5. S – довільна множина. $L = \{\ f \mid f : S \to \mathbf{R}$ – функція $\}$. Доведіть, що L – дійсний лінійний простір з операціями $\oplus$ і $\otimes$:

$$f \oplus g: \quad (f \oplus g)(s) = f(s) + g(s) \quad \Big|\ \forall s \in S, \forall f, g \in L$$

$$a \otimes f : (af)(s) = a \cdot (f(s)) \ \Big|\forall s \in S, \forall f \in L, \forall a \in \mathbf{R}.$$

Чи будуть лінійними просторами L_Q і L_C?

6. Чи можуть в лінійному просторі існувати :

   а) два нульових елементи;

   б) два протилежних елементи для деякого елементу x?

7. Як виражається через x елемент, протилежний до x?

8. Чи справедлива рівність $\theta = -\theta$?

9. Що таке різниця елементів x і y?

10. Чи є лінійно залежними елементи лінійного простору L:

а) x і $2x$; б) $x_1, x_2, ..., x_n, \theta$; в) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ і $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, якщо $L = T_2$;

г) функції $\sin x$ і $\cos x$, якщо $L = C_{[0,\, \pi/2]}$;

д) функції $f_1 = 2\sin^2 x$, $f_2 = 1$, $f_3 = 3\cos^2 x$, якщо $L = C_{[0,\, \pi/2]}$?

Відповіді обґрунтуйте.

11. Відомо, що деякий елемент можна представити у вигляді двох різних лінійних комбінацій одних і тих самих елементів $e_1, e_2, \ldots, e_n$. Чи будуть елементи $e_1, e_2, \ldots, e_n$ лінійно незалежними?

3.3. Базис лінійного простору. Однозначність розкладення

Далі будемо розглядати дійсний лінійний простір L_R.

Сукупність ЛНЗ елементів $e_1, e_2, ..., e_n$ простору L_R називається ***базисом*** цього простору, якщо $\forall x \in L$ знайдуться такі дійсні числа $x_1, x_2, ..., x_n$, що виконується рівність

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n = \sum_{i=1}^{n} x_i e_i \qquad (3.3.1)$$

При цьому (1.3.1) називається ***розкладенням елементу*** x ***по базису*** $e_1, e_2, ..., e_n$, а числа $x_1, x_2, ..., x_n$ – ***координатами елементу*** x відносно базису $e_1, e_2, ..., e_n$. У цьому випадку позначають $x = (x_1, x_2, ..., x_n)_e$, $x_e = (x_1, x_2, ..., x_n)$, де e – базис. Зауважимо також, що рівність (1.3.1) може мати матричний вигляд

$$x = (e_1, e_2, ..., e_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (e_1, e_2, ..., e_n) \cdot x_e^T = e \cdot x_e^T, \qquad (3.3.2)$$

де $e = (e_1, e_2, ..., e_n)$ – рядок, складений із елементів базису, x_e^T – стовпчик із координат $x_1, x_2, ..., x_n$ елементу x в базисі e (транспонований рядок x_e).

Твердження 1.3.1 (Про однозначність розкладення по базису) *Довільний елемент лінійного простору однозначно розкладається по базису.*

Д о в е д е н н я. Нехай існує два розкладення довільного елементу $x \in L$ по базису $e_1, e_2, ..., e_n$:

$$\overline{x} = x_1\overline{e}_1 + x_2\overline{e}_2 + ... + x_n\overline{e}_n \qquad (*)$$

$$\overline{x} = x_1'\overline{e}_1 + x_2'\overline{e}_2 + ... + x_n'\overline{e}_n \qquad (**)$$

Тоді із аксіом та властивостей лінійного простору різниця рівнянь (*) та (**) має вид

$$\theta = \overline{x} - \overline{x} = (x_1 - x_1')\overline{e}_1 + ... + (x_n - x_n')\overline{e}_n \quad (***)$$

Так як $e_1, e_2, ..., e_n$ – ЛНЗ, то із (***) випливає:

$$\left.\begin{array}{c} x_1 - x_1' = 0 \Rightarrow x_1 = x_1' \\ .... \\ x_n - x_n' = 0 \Rightarrow x_n = x_n' \end{array}\right\} \Rightarrow (*) = (**) .\blacksquare$$

Розглянемо приклади базисів лінійних просторів та координат елементів.

<u>Приклад 1</u>. $V_n = \{(a_1, ..., a_n) \mid a_i \in R\}$. Нехай

$$\overline{e}_1 = (1,0,...,0)$$
$$\overline{e}_2 = (0,1,...,0)$$
$$...$$
$$\overline{e}_n = (0,0,...,1)$$

Покажемо, що ці вектори утворюють базис V_n.

▼ 1. Очевидно, що $\overline{e}_1, \overline{e}_2, ..., \overline{e}_n$ – ЛНЗ.

2. $\forall \overline{x} = (x_1, x_2, ..., x_n) \in V_n$ маємо, що $\overline{x} = x_1\overline{e}_1 + x_2\overline{e}_2 + ... + x_n\overline{e}_n$.

Отже одержали, що $\overline{e}_1, \overline{e}_2, ..., \overline{e}_n$ – базис V_n і $x_1, x_2, \ldots, x_n$ – координати вектора $\overline{x}$ відносно $\overline{e}_1, \overline{e}_2, ..., \overline{e}_n$. ▲

Зауважимо, що *базис у лінійному просторі не є єдиним*. Наприклад у V_3 поряд з традиційним базисом e наведемо приклад іншого базису e' і координат вектору $\overline{x} = (3,-2,1)$ в цих двох базисах:

$e_1 = (1,0,0)$	$e_1' = (2,0,0)$
$e_2 = (0,1,0)$	$e_2' = (0,1,1)$
$e_3 = (0,0,1)$	$e_3' = (0,1,0)$
$\overline{x} = (3,-2,1)_e$	$\overline{x} = \sum x_i'e_i' = (2x_1', x_2' + x_3', x_2')$
	$\overline{x} = (x_1', x_2', x_3')_{e'} = (3/2, 1, -3)_{e'}$

Координати вектору $\overline{x}=(3,-2,1)$ в базисі e' знаходяться із рівності

$$\overline{x}=(2x_1',x_2'+x_3',x_2')=(3,-2,1),$$

яка перетворюється в систему трьох лінійних рівнянь з трьома невідомими

$$\begin{cases} 2x_1'=3; \\ x_2'+x_3'=-2; \\ x_2'=1. \end{cases}$$

Зазначимо, що матриця цієї системи

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

складається із стовпчиків компонент векторів базису e_1',e_2',e_3'.

<u>Приклад 2</u>. $P_n[x]=\{a_nx^n+...+a_1x+a_0 \mid a_i\in\mathbf{R},\ \ n\in\mathbf{N_0}\ \}$**.** Знайти базис $P_n[x]$**.** Знайти координати многочлена $g(x)=(x-2)^2$ в цьому базисі.

▼ 1) Нульовим елементом простору $P_n[x]$ є многочлен $\mathrm{O}(x)=0x^n+\ldots+0x+0$. Розглянемо сукупність елементів $e:1,x,x^2,...,x^n$. Вони лінійно незалежні, бо $a+bx+cx^2+...+dx^n=\mathrm{O}(x)\Leftrightarrow$ $\Leftrightarrow a=b=c=...=d=0$.

2) $\forall f(x)\in P_n[x]\ \exists a_0,a_1,...,a_n\in\mathbf{R}$, що $f(x)=a_nx^n+...+a_1x+a_0$.

Отже, $e:1,\ x,\ x^2,...,\ x^n$ – базис $P_n[x]$. Кількість елементів цього базису дорівнює $n+1$.

Координатами многочлена $g(x)=(x-2)^2$ в цьому базисі будуть 4,–4,1,0,...,0, бо

$$g(x)=x^2-4x+4=4\cdot1+(-4)\cdot x+1\cdot x^2+0\cdot x^3+...+0\cdot x^n=(4,-4,1,0,...,0)_e.$$

Зауважимо, що базисом простору всіх многочленів $P[x]$ буде сукупність елементів $1,x,x^2,...,x^n,...$, тобто базис складається з *нескінченної* кількості елементів. ▲

Теорема 3.3.2. (про додавання і множення елементів лінійного простору в базисі). *При додаванні довільних елементів лінійного простору їх координати додаються. При множенні довільного елементу на число всі координати множаться на це число.*

Д о в е д е н н я. Нехай $e_1, e_2, ..., e_n$ – базис простору L, елементи $x, y \in L$ і $x = x_1e_1 + x_2e_2 + ... + x_ne_n$, $y = y_1e_1 + y_2e_2 + ... + y_ne_n$.

Тоді $x + y = x_1e_1 + x_2e_2 + ... + x_ne_n + y_1e_1 + y_2e_2 + ... + y_ne_n \overset{I,III}{=}$

$= (x_1 + y_1)e_1 + (x_2 + y_2)e_2 + ... + (x_n + y_n)e_n$;

$\lambda x = \lambda(x_1e_1 + x_2e_2 + ... + x_ne_n) \overset{II,III}{=} (\lambda x_1)e_1 + (\lambda x_2)e_2 + ... + (\lambda x_n)e_n$

Тобто, якщо $x = (x_1, x_2, ..., x_n)_e$, $y = (y_1, y_2, ..., y_n)_e$, то $x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)_e$, а $\lambda x = (\lambda x_1, \lambda x_2, ..., \lambda x_n)_e$. ■

Зауваження: $x' = (-1)x = (-x_1, -x_2, ..., -x_n)_e$

Два простори L_P та F_P називаються ***ізоморфними (*** $L \cong F$ ***)***, якщо існує взаємно однозначне відображення $\varphi : L \to F$ таке, що виконуються наступні умови

1) $\forall x, y \in L \quad | \quad \varphi(x + y) = \varphi(x) + \varphi(y)$;

2) $\forall x \in L, \forall \lambda \in P \quad | \quad \varphi(\lambda x) = \lambda \varphi(x)$.

Відповідність $f : X \to Y$ називається ***взаємно однозначним відображенням***, якщо вона задовольняє умовам:

$$\begin{bmatrix} 1)\forall\ x \in X \quad \exists!\ y \in Y (f(x) = y) \\ 2)\forall\ y \in Y \quad \exists\ \ x \in X \quad (y = f(x)) \\ 3)\forall\ x_1, x_2 \in X \quad (x_1 \neq x_2) \Rightarrow (f(x_1) \neq f(x_2)) \end{bmatrix}$$

Відповідність $f : X \to Y$, що задовольняє умову 1), називається ***відображенням***, що задовольняє умови 1), 2) – ***сюр'єктивним*** відображенням, що задовольняє умови 1), 3) – ***ін'єктивним*** відображенням.

При цьому відображення φ називається ***ізоморфізмом*** просторів L_P та F_P.

Розглянемо окремі

В л а с т и в о с т і і з о м о р ф і з м і в

Нехай $\varphi : L \to F$ – ізоморфізм простору L_P.

1. *Ізоморфний образ нульового елементу і тільки його дорівнює нульовому елементу.*

▼ Нехай θ, θ' – нульові елементи L і F відповідно. Тоді $\forall x \in L$

1) $\varphi(\theta) = \varphi(x + (-1)x) = \varphi(x) + (-1)\varphi(x) = \theta'$;

2) якщо $\varphi(x)=\theta'$ то за умовою ін'єктивності φ $(\varphi(x_1)=\varphi(x_2)\Rightarrow x_1=x_2)$ маємо, що $\varphi(x)=\varphi(0)\Rightarrow x=\theta$. ■

2. *Елементи* $a_1,a_2,...,a_t\in L$ *лінійно залежні тоді і тільки тоді, коли лінійно залежні їх ізоморфні образи.*

▼ Нехай $a_1,a_2,...,a_t\in L$ – лінійно залежні. Це означає, що існують не всі рівні нулю коефіцієнти $\lambda_i\in R$ такі, що виконується умова $\sum\lambda_i a_i=\theta$. Тоді за попередньою властивістю $\theta'=\varphi(\theta)=\varphi(\sum_{i=1}^{t}\lambda_i a_i)=\sum_{i=1}^{t}\lambda_i\varphi(a_i)$, що означає лінійну залежність $\varphi(a_1),\varphi(a_2),...,\varphi(a_t)$. ■

3. *Елементи* $a_1,a_2,...,a_t\in L$ *лінійно незалежні тоді і тільки тоді, коли* $\varphi(a_1),\varphi(a_2),...,\varphi(a_t)\in F$ *лінійно незалежні.*

Ця властивість випливає із попередньої властивості.

4. *Якщо* $\varphi: L\to F$ *– ізоморфізм, то* $\varphi^{-1}: F\to L$ *також ізоморфізм.*

5. *Якщо* $L\cong F$, $F\cong W$, *то* $L\cong W$.

Наслідок 3.3.3. *Кожний лінійний простір* L_R, *базис якого складається із* n *елементів, ізоморфний векторному простору* V_n.

Д о в е д е н н я. Виберемо базис $e_1,e_2,...,e_n$ простору L. Задамо відображення $f: L_R\to V_n$ наступним чином:

$\forall\, x\in L$, якщо $x=(x_1,x_2,...,x_n)_e$, то $f(x)=(x_1,x_2,...,x_n)\in V_n$.

Легко перевірити, що дане відображення взаємно однозначне. А також, за теоремою 1.3.2, якщо $y=(y_1,y_2,...,y_n)_e$, то

1) $f(x+y)=f((x_1+y_1,x_2+y_2,...,x_n+y_n)_e)=$

$=(x_1+y_1,x_2+y_2,...,x_n+y_n)=(x_1,x_2,...,x_n)+(y_1,y_2,...,y_n)=f(x)+f(y)$;

2) $\forall\lambda\in R$ $f(\lambda x)=f((\lambda x_1,\lambda x_2,...,\lambda x_n)_e)=(\lambda x_1,\lambda x_2,...,\lambda x_n)=$

$=\lambda(x_1,x_2,...,x_n)=\lambda f(x)$, тобто виконуються всі умови ізоморфізму . ■

Контрольні питання і завдання

1. Дайте означення базису лінійного простору.
2. Чи може базис містити нульовий елемент? Відповідь обґрунтуйте.
3. Що таке координати елементу лінійного простору в даному базисі?

4. Елементи x_1 і x_2 мають в деякому базисі координати $-3,5,7$ і $10,0,1$. Чому рівні координати елементів $x_1 - x_2$, $7x_1 + x_2$, $2x_1 + 3x_2$ в цьому ж базисі?

5. Чи утворюють функції $p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = x^2$ базис:

а) в просторі $P_2[x]$;

б) в просторі $C_{[a,b]}$ неперервних на відрізку $[a,b]$ функцій?

6. Вкажіть базис в просторі M_3 такий, що елементами базисних матриць є тільки 0 і 1. Якою є кількість базисних матриць? Знайдіть в цьому базисі координати матриці $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$.

7. Які лінійні простори називаються ізоморфними?

8. Чи може ізоморфним образом нульового елементу бути ненульовий елемент?

9. Чи ізоморфні лінійний простір T_3 стовпчиків з трьома елементами і лінійний простір $P_3[x]$ многочленів степеня не вище 3? Відповідь обґрунтуйте.

10. Наведіть приклади трьох лінійних просторів, ізоморфних лінійному простору $P_3[x]$ многочленів степеня не вище 3?

3.4. Розмірність лінійного простору

Лінійний простір L називається n-***мірним***, якщо в ньому існує n ЛНЗ елементів, а довільні $(n+1)$ елементів ЛЗ. Число n при цьому називається ***розмірністю*** простору L. Іншими словами говорять, що ***розмірність*** простору – це максимальна кількість ЛНЗ елементів простору. Позначається $\dim L = n$.

Знайдемо розмірності деяких лінійних просторів в наступних прикладах.

<u>Приклад 1.</u> $M_2(R)$ – дійсні матриці 2-го порядку.

Елементи $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ – ЛНЗ, а довільні 5 елементів ЛЗ (перевірте самостійно), тому $\dim M_2(R) = 4$.

<u>Приклад 2</u>. Множина F розв'язків однорідної системи лінійних рівнянь з n невідомими має фундаментальну систему розв'язків, яка

складається з $(n-r)$ розв'язків, де r – ранг матриці системи (теорема 2.12.5), тому (за означенням фундаментальної системи розв'язків)

$$\dim F = n-r.$$

<u>Приклад 3.</u> $P[x] = \{a_n x^n + ... + a_1 x + a_0 \mid a_i \in R\}$

$1, x, x^2, 2x^2+3$ – ЛЗ, але це не означає, що $\dim P[x] < 4$;

$1, x, x^2, x^3$ – ЛНЗ, але це не означає, що $\dim P[x] = 4$;

$1, x, x^2, x^3, ..., x^n, ...$ – ЛНЗ, тому $\dim P[x] = \infty$

В прикладах 1, 2 лінійні простори ***скінченномірні***, в прикладі 3 лінійний простір ***нескінченномірний***.

Теорема 3.4.1. (Про зв'язок розмірності і базису) *Якщо лінійний простір L має базис із n елементів, то* $\dim L = n$.

Д о в е д е н н я. Нехай $e_1, e_2, ..., e_n$ – базис L. Тоді $e_1, e_2, ..., e_n$ – ЛНЗ.

Візьмемо довільні $(n+1)$ елементи $x_1, x_2, ..., x_{n+1} \in L$. За означенням базису $\exists a_{ij}, i = \overline{1, n+1}, j = \overline{1, n}$, що

$$x_1 = a_{11}e_1 + a_{12}e_2 + ... + a_{1n}e_n = (a_{11}, a_{12}, ..., a_{1n})_e,$$
$$x_2 = a_{21}e_1 + a_{22}e_2 + ... + a_{2n}e_n = (a_{21}, a_{22}, ..., a_{2n})_e,$$
$$.....$$
$$x_{n+1} = a_{(n+1)1}e_1 + a_{(n+1)2}e_2 + ... + a_{(n+1)n}e_n = (a_{(n+1)1}, a_{(n+1)2}, ..., a_{(n+1)n})_e$$

Так як із наслідку 1.3.3 простір $L \cong V_n$, то розглянемо вектори

$$(a_{11}, a_{12}, ..., a_{1n}),$$
$$(a_{21}, a_{22}, ..., a_{2n}), (*)$$
$$.....$$
$$(a_{(n+1)1}, a_{(n+1)2}, ..., a_{(n+1)n})$$

на предмет їх лінійної залежності або незалежності. Ці вектори утворюють матрицю A:

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{(n+1)1} & a_{(n+1)2} & ... & a_{(n+1)n} \end{pmatrix}$$

Вона має $(n+1)$ рядів і n стовпчиків, тому $rg\,A \le n < n+1$. Звідси $(n+1)$ векторів-рядків (*), а з ними за другою властивістю ізоморфізмів і елементи $x_1,...,x_{n+1}$ – ЛЗ. Отже, за означенням, $\dim L = n$. ■

Теорема 3.4.2. *(*обернена*) Якщо* $\dim L = n$*, то довільні* n *ЛНЗ елементів цього простору утворюють його базис.*

Д о в е д е н н я. Нехай $e_1, e_2, ..., e_n$ – ЛНЗ (вони існують із означення розмірності простору). Треба показати, що довільний елемент простору L можна виразити через $e_1, e_2, ..., e_n$, тоді це буде означати, що $e_1, e_2, ..., e_n$ – базис L.

Візьмемо довільний елемент $x \in L$. За означенням $\dim L$ елементи $x, e_1, ..., e_n$ – ЛЗ (бо їх $(n+1)$). За твердженням 1.3.1 елемент x виражається через елементи $e_1, e_2, ..., e_n$. Враховуючи їх ЛНЗ, одержуємо, що $e_1, e_2, ..., e_n$ – базис L. ■

Наслідок 3.4.3. *Якщо відомо один базис* $e_1, e_2, ..., e_n$ *простору, то довільні* n *ЛНЗ елементів* $f_1, f_2, ..., f_n$ *цього простору теж утворюють його базис.*

Наслідок 3.4.4. *Лінійні простори однакової розмірності ізоморфні.*

Отже, єдиною істотною відмінністю структури скінченномірних лінійних просторів є їх розмірність.

Контрольні питання і завдання

1. Дайте означення розмірності лінійного простору.
2. Як пов’язані між собою розмірність простору і число елементів базису?
3. Відомо, що елементи $e_1, e_2, \ldots, e_n$ утворюють базис лінійного простору. Чи буде сукупність елементів $c_1e_1, c_2e_2, \ldots, c_ne_n$, де c_p – числа, не рівні нулю, базисом цього простору? Відповідь обґрунтуйте.
4. Скільки базисів має n- мірний лінійний простір?

3.5. Перетворення базисів

Можна помітити той факт, що у кожному просторі базис не єдиний. Охарактеризуємо зв’язок між різними базисами та координатами одного й того ж елементу в різних базисах одного й того ж простору.

Нехай $e: e_1, e_2, ..., e_n$ – 1-й базис простору L.

$e' : e_1', e_2', ..., e_n'$ – 2-й базис простору L

Так як $e_1', e_2', ..., e_n'$ – елементи L, то $\exists a_{ij} \in R, i, j = 1,2,...,n,$ що

$$e_1' = a_{11}e_1 + a_{12}e_2 + ... + a_{1n}e_n = (a_{11}, a_{12}, ..., a_{1n})_e,$$
$$e_2' = a_{21}e_1 + a_{22}e_2 + ... + a_{2n}e_n = (a_{21}, a_{22}, ..., a_{2n})_e,$$
$$\text{...} \qquad (3.5.1)$$
$$e_n' = a_{n1}e_1 + a_{n2}e_2 + ... + a_{nn}e_n = (a_{n1}, a_{n2}, ..., a_{nn})_e.$$

Говорять, що другий базис e' виражено через перший e, а перехід від базису e до e' задається матрицею U:

$$U = \begin{pmatrix} a_{11} & a_{21} & ... & a_{n1} \\ a_{12} & a_{22} & ... & a_{n2} \\ ... & ... & ... & ... \\ a_{1n} & a_{1n} & ... & a_{nn} \end{pmatrix}.$$

Матрицею переходу від базису e ***до базису*** e' називається матриця U, яка складається із стовпчиків a^i координат елементів базису e' в базисі e. Позначають $e \xrightarrow{U} e'$.

Розглянемо з в о р о т н е п е р е т в о р е н н я б а з и с і в: $e' \to e$. Із матричного рівняння (1.5.3) одержуємо еквівалентне матричне рівняння $e'U^{-1} = e$, яке визначає матрицю переходу від базису e' до базису e, а саме

Твердження 3.5.1. *Якщо перехід від базису* e *до* базису e' *простору* L *задається матрицею* U, *то перехід від* e' *до* e *задається оберненою матрицею* U^{-1}:

$$e' \xrightarrow{U^{-1}} e.$$

Розглянемо з м і н и в к о о р д и н а т а х елементів при переході від базису e до базису e'. Нехай $x \in L$, $x_e = (x_1, x_2, ..., x_n)$, $x_{e'} = (x_1', x_2', ..., x_n')$. Відомо, що в матричному вигляді (1.2.2) $x = ex_e^T$ та $x = e'x_{e'}^T$, тобто $ex_e^T = e'x_{e'}^T$ Якщо $e \xrightarrow{U} e'$, то $e' = eU$ і одержуємо рівняння $ex_e^T = eUx_{e'}^T$, обидві частини якого є розклад елементу x по базису e. З однозначності цього розкладу (теорема 1.2.3) випливає матричне рівняння

$$x_e^T = Ux_{e'}^T, \qquad (3.5.4)$$

яке еквівалентне рівнянням

$$x_{e'}^T = U^{-1} x_e^T \text{ та } x_{e'} = x_e (U^{-1})^T. \qquad (3.5.5)$$

Рівняння (1.5.5) одержуються із (1.6.4) з урахуванням властивостей транспонування матриць і добутку матриць ($(AB)^T = B^T A^T$ і $(A^T)^T = A$) Вигляд другого рівняння (1.5.5) приводить до

Твердження 3.5.2. *Якщо перехід від базису e до* базису *e' простору L задається матрицею U, то перехід від координат в базисі e до координат в базисі e' задається матрицею $(U^{-1})^T$, транспонованою до оберненої.*

Контрольні питання і завдання

1. Як скласти матрицю переходу від одного базису до іншого?
2. Чи може матриця переходу від одного базису до іншого бути виродженою? Відповідь обґрунтуйте.
3. Складіть матрицю переходу від базису e_1, e_2, e_3 до базису f_1, f_2, f_3, якщо $f_1 = e_1 + e_2 + e_3$, $f_2 = -e_1 + e_2 + e_3$, $f_3 = 3e_1 - e_2 + 4e_3$. Знайдіть матрицю зворотного переходу.
4. Запишіть формули перетворення координат елементу при перетворенні базису?

3.6. Підпростір лінійного простору

Нехай $L \subset V$, V_P – лінійний простір над полем P. Множина L називається ***підпростором*** простору V, якщо L_P – лінійний простір.

Теорема 3.6.1. (критерій підпростору) *Кожна підмножина L лінійного простору V_P утворює підпростір тоді і тільки тоді, якщо вона задовольняє наступним умовам:*

1°. $\forall x, y \in L, \quad x + y \in L$;

2°. $\forall x \in L, \quad \forall \lambda \in P, \ \lambda x \in L.$

Н е о б х і д н і с т ь. Якщо L – підпростір, то умови $1^\circ, 2^\circ$, які відповідають умовам замкненості операцій додавання і множення на число на множині L, виконуються за означенням лінійного простору L. ▲

Д о с т а т н і с т ь. Нехай $L \subset V$ і виконуються умови $1^\circ, 2^\circ$. Тоді залишилось довести, що L задовольняє умовам 1) – 8) означення лінійного простору.

Зрозуміло, що умови 1), 2), 5) – 8) виконуються для довільної підмножини простору V. Тому залишилось перевірити виконання умов 3), 4), а саме належність множині L нейтрального та протилежних для всіх елементів із L, що вже існують у множині V.

Так як за властивістю 2 лінійних просторів $0 \cdot x = \theta \quad \forall x \in L \subset V$, а за властивістю 4 лінійних просторів $-x = (-1) \cdot x$, та із умови 2° випливає, що $0 \cdot x \in L$ і $(-1) \cdot x \in L$, то $\theta \in L$ і $-x \in L \quad \forall x \in L$, а значить виконуються бажані умови 3) і 4). ■

Зауважимо, що даний критерій можна розглядати в такому вигляді:

Теорема 3.6.1а. *Кожна підмножина L лінійного простору V_P утворює підпростір тоді і тільки тоді, якщо вона задовольняє наступній умові:*

$$\forall x, y \in L, \forall \lambda, \mu \in P \; \middle| \lambda x + \mu y \in L.$$

Найважливішим прикладом лінійних підпросторів є лінійна оболонка.

Нехай V – дійсний лінійний простір

Лінійною оболонкою L елементів $x, y, ..., z \in V$ називається сукупність всіх лінійних комбінацій цих елементів:

$$L = \langle x, y, ..., z \rangle = \{\alpha x + \beta y + ... + \gamma z \mid \alpha, \beta, ..., \gamma \in \mathbf{R}\}$$

Так як всі лінійні комбінації елементів простору належать даному простору, то $\langle x, y, ..., z \rangle \subset V$ і доведемо, що $L = \langle x, y, ..., z \rangle$ *утворює підпростір простору* V, користуючись критерієм підпростору 1.6.1:

1°. $(\alpha_1 x + \beta_1 y + ... + \gamma_1 z) + (\alpha_2 x + \beta_2 y + ... + \gamma_2 z) =$

$$= (\alpha_1 + \alpha_2)x + (\beta_1 + \beta_2)y + ... + (\gamma_1 + \gamma_2)z \in L;$$

2°. $\lambda(\alpha x + \beta y + ... + \gamma z) = (\lambda\alpha)x + (\lambda\beta)y + ... + (\lambda\gamma)z \in L$. ■

Зауваження 3.6.2. *Нехай $x_1, x_2, ..., x_t \in V$. Лінійна оболонка $\langle x_1, x_2, ..., x_t \rangle$ є найменшим підпростором простору V, який включає елементи $x_1, x_2, ..., x_t$.*

Д о в е д е н н я. Нехай W – саме такий найменший підпростір. Тоді

1. $x_i = 1 \cdot x_i \in \langle x_1, x_2, ..., x_t \rangle, \quad i = 1,2,...,t.$

2. Для виконання умов критерію підпростору $1^\circ, 2^\circ$ всі лінійні комбінації $\sum_1^t \alpha_i x_i$ повинні належати W, отже $\langle x_1, x_2, ..., x_t \rangle \subset W$.

Так як із попереднього доведення лінійна оболонка утворює лінійний підпростір, то із умови мінімальності підпростору W і одержаних умов 1,2 маємо, що $\langle x_1, x_2, ..., x_t \rangle = W$. ■

Теорема 3.6.3. (про розмірність лінійної оболонки) *Розмірність лінійної оболонки* $\langle x, y, ..., z \rangle$ *дорівнює максимальному числу ЛНЗ елементів в системі елементів* $x, y, ..., z$, *а самі ці ЛНЗ елементи утворюють базис.*

Зауваження 3.6.4. *Нехай* $e_1, e_2, ..., e_n$ *– базис простору* V. *Тоді*

$$V = \langle e_1, e_2, ..., e_n \rangle.$$

Розглянемо питання про зв'язок базису простору V і базису підпростору $L \subset V$. Не викликає сумніву, що розмірність довільного підпростору L n- мірного лінійного простору V не перевищує n.

Точніше, підпростір $L = V$, тоді і тільки тоді, коли $\dim L = \dim V$:

1) $L = V \Rightarrow \dim L = \dim V$ – очевидно,

2) якщо $\dim L = \dim V$, то базис $e_1, e_2, ... e_t$ підпростору L за теоремами 1.4.1 і 1.4.2 буде базисом простору V і за попереднім зауваженням 1.6.4 маємо

$$L = \langle e_1, e_2, ..., e_t \rangle = V.$$

В протилежному випадку: $L \neq V$ тоді і тільки тоді, коли $\dim L < \dim V$.

Теорема 3.6.5. (про доповнення базису). *Якщо елементи* $e_1, e_2, ..., e_k$ *складають базис* k- *мірного підпростору* n- *мірного лінійного простору* V, *то цей базис можна доповнити елементами* $e_{k+1}, e_{k+2}, ..., e_n \in V$ *так, що сукупність елементів* $e_1, e_2, ..., e_n$ *буде утворювати базис всього простору* V.

Зауважимо, що елементи $e_{k+1}, e_{k+2}, ..., e_n$ називаються ***доповняльним базисом*** простору L до простору V.

Звернемо увагу також на те, що з довільного базису простору V не завжди можна виділити базис підпростору L.

Наприклад. $\bar{e}_1 = (1,0,0)$, $\bar{e}_2 = (0,1,0)$, $\bar{e}_3 = (0,0,1)$ – базис V_3. Якщо $L = \langle \bar{x}, \bar{y} \rangle$, де $\bar{x} = (1,2,0)$, $\bar{y} = (0,1,1)$, то $L \subseteq V_3$ і $\bar{x}, \bar{y}$ – *ЛНЗ* $\Rightarrow \dim L = 2$.

Спробуємо з базису e_1, e_2, e_3 вибрати два ЛНЗ елементи простору L.
$L = \langle \bar{x}, \bar{y} \rangle = \{\alpha\bar{x} + \beta\bar{y} \mid \alpha, \beta \in \mathbf{R}\} = \{(\alpha, 2\alpha + \beta, \beta) \mid \alpha, \beta \in \mathbf{R}\}$.

$e_1 = (1,0,0) \in L$, якщо

$$\begin{cases} \alpha & = 1 \\ 2\alpha + \beta & = 0 \\ \beta & = 0 \end{cases} \Rightarrow \begin{cases} \alpha = 1 \\ \beta = -2 \\ \beta = 0 \end{cases} \text{– система не сумісна} \Rightarrow e_1 \notin L.$$

Аналогічно $e_2, e_3 \notin L$, тому базис L не можна вибрати з базису V.

Контрольні питання і завдання

1. Що таке підпростір лінійного простору?
2. Сформулюйте критерій підпростору лінійного простору.
3. Доведіть твердження 1.6.4.
4. Чи може розмірність лінійного простору бути меншою (рівною, більшою) розмірності його підпростору?
5. Чи можуть лінійний простір і його підпростір, який не співпадає з усім простором, бути ізоморфними?
6. Дайте означення лінійної оболонки.
7. Який лінійний простір утворює лінійна оболонка, породжена системою функцій $f_0(x) = 1, \quad f_1(x) = x, \ \dots, \ f_n(x) = x^n$?
8. Доведіть, що $\langle\ x, 3x\ \rangle = \langle\ x\ \rangle$.

3.7. Перетин і сума підпросторів. Пряма сума підпросторів

Нехай L_1 і L_2 – два різних підпростори простору V.

Твердження 3.7.1. (про перетин підпросторів) *Перетин підпросторів L_1 і L_2 утворює підпростір.*

Д о в е д е м о, що $L = L_1 \bigcap L_2 = \{x \in V \mid x \in L_1 \ i \ x \in L_2\}$ – підпростір в V. Так як $L \subset V$, то можемо користуватись критерієм підпростору 3.6.1:

1° $\forall x, y \in L \Rightarrow x, y \in L_1$, і $x, y \in L_2 \Rightarrow x + y \in L_1$ і $x + y \in L_2 \Rightarrow x + y \in L$.

2°. $\forall x, y \in L \Rightarrow x \in L_1$ і $x \in L_2 \Rightarrow \forall \lambda \in \mathbf{R} \big| \ \lambda x \in L_1 \ i \ \lambda x \in L_2 \Rightarrow \lambda x \in L$.

Отже $L = L_1 \bigcap L_2$ – підпростір. ■

Сумою підпросторів L_1 *і* L_2 називається множина

$$L = \left\{\ x + y \ \middle|\ x \in L_1, y \in L_2\ \right\}.$$

Позначається $L = L_1 + L_2$.

Твердження 3.7.2. (про суму підпросторів) *Сума підпросторів L_1 і L_2 утворює підпростір в V.*

Д о в е д е н н я. Оскільки довільний елемент l множини $L = L_1 + L_2$ має вигляд $l = x + y$, де $x \in L_1, y \in L_2$, а підпростори $L_1, L_2 \subset V$, то за рахунок замкненості додавання в V елемент $l \in V$, а $L \subset V$. Отже, скористаємось критерієм підпростору 3.6.1:

1°. $\forall l_1, l_2 \in L \big| l_1 + l_2 = (x_1 + y_1) + (x_2 + y_2) = (\underset{\in L_1}{x_1 + x_2}) + (\underset{\in L_2}{y_1 + y_2}) \in L$.

2°. $\forall l \in L, \forall \lambda \in \mathbf{R} \big| \lambda l = \lambda x + \lambda y \in L_1 + L_2 = L$.

Отже $L = L_1 + L_2$ – підпростір. ■

Приклад. Нехай $\vec{a}, \vec{b}$ – неколінеарні вектори декартового трьохмірного простору. Тоді $L_1 = \langle \vec{a} \rangle = \{\lambda \vec{a} | \lambda \in \mathbf{R}\}$ та $L_2 = \langle \vec{b} \rangle = \{\mu \vec{b} | \mu \in \mathbf{R}\}$ зображають прямі l_1 та l_2, паралельні відповідно векторам $\vec{a}$ і $\vec{b}$, що проходять через початок координат, $L_1 \cap L_2 = \{\vec{0}\}$, а $L_1 + L_2 = \{\lambda \vec{a} + \mu \vec{b} | \lambda, \mu \in \mathbf{R}\}$ – описує площину, яка визначається прямими l_1 та l_2.

Теорема 3.7.3. (про суму розмірностей) *Сума розмірностей довільних підпросторів скінченномірного лінійного простору дорівнює сумі розмірностей перетину і суми цих підпросторів.*

$$\dim L_1 + \dim L_2 = \dim(L_1 \cap L_2) + \dim(L_1 + L_2) \qquad (3.7.1)$$

Д о в е д е н н я. Нехай $L_1 + L_2 = L$, $L_1 \cap L_2 = L_0$,

$$e_1, e_2, ..., e_k \text{ – базис } L_0. \qquad (1)$$

Так як $L_0 \subset L_1$ і $L_0 \subset L_2$, то за теоремою 3.6.5 про доповнення базису підпростору, доповнимо базис (1) до базисів L_1 і L_2:

$$e_1, e_2, ..., e_k, g_1, ..., g_s \text{ – базис } L_1, \qquad (2)$$

$$e_1, e_2, ..., e_k, f_1, ..., f_m \text{ – базис } L_2. \qquad (3)$$

Тоді $\dim L_0 = k$, $\dim L_1 = k + s$, $\dim L_2 = k + m$. При цих позначеннях формула (1.7.1) буде мати вигляд: $(k + s) + (k + m) = k + \dim L$. Тобто треба довести, що $\dim L = k + s + m$.

Розглянемо систему із $k + s + m$ елементів простору L:

$$e_1, e_2, ..., e_k, g_1, ..., g_s, f_1, ..., f_m. \qquad (4)$$

1. З'ясуємо питання, чи буде система (4) ЛНЗ ?

Розглянемо рівність

$$\sum_1^k \alpha_i e_i + \sum_1^s \beta_j g_j + \sum_1^m \gamma_t f_t = \theta \qquad (5).$$

$$\sum_{1}^{k}\alpha_i e_i+\sum_{1}^{s}\beta_j g_j+\sum_{1}^{m}\gamma_t f_t=\theta \Leftrightarrow \underbrace{\sum_{1}^{k}\alpha_i e_i+\sum_{1}^{s}\beta_j g_j}_{\in L_1}=\underbrace{\sum_{1}^{m}(-\gamma_t)f_t}_{\in L_2}\Rightarrow,$$

$\Rightarrow \sum_{1}^{k}\alpha_i e_i+\sum_{1}^{s}\beta_j g_j$, $\sum_{1}^{m}(-\gamma_t)f_t \in L_1 \bigcap L_2 = L_0$. Тобто, $\exists\, \lambda_i \in \mathbf{R},\ i=1,...,k$ такі, що $\sum_{1}^{m}(-\gamma_t)f_t=\sum_{1}^{k}\lambda_i e_i$. Значить $\sum_{1}^{k}\lambda_i e_i+\sum_{1}^{m}\gamma_t f_t=\theta$. і $\lambda_i=0,\ i=1,..k$, $\gamma_t=0,\ t=1,...,m$, так як базис $e_1,e_2,...,e_k,f_1,...,f_m$ – ЛНЗ.

Підставимо $\gamma_t=0$ в (5): Тоді $\sum_{1}^{k}\alpha_i e_i+\sum_{1}^{s}\beta_j g_j=\theta$ і знову $\alpha_i=0,\ i=1,...,k,\ \beta_j=0,\ j=1,...,s$, бо базис $e_1,e_2,...,e_k,g_1,...,g_s$ –ЛНЗ.

Тобто рівність (5) виконується тоді і тільки тоді, коли всі коефіцієнти лінійної комбінації дорівнюють нулю, значить система (4) – ЛНЗ.

2. Розглянемо тепер, чи можна кожен елемент $x \in L$ лінійно виразити через систему (4).

$\forall x \in L \Rightarrow \exists\, x_1 \in L_1, x_2 \in L_2$, що $x = x_1 + x_2$. Крім того $\exists \lambda_i, \mu_j, \alpha_i, \beta_t \in \mathbf{R}$ такі, що

$$x_1=\lambda_1 e_1+...+\lambda_k e_k+\mu_1 g_1+...+\mu_l g_s,$$
$$x_2=\alpha_1 e_1+...+\alpha_k e_k+\beta_1 e_1+...+\beta_m f_m.$$

Тоді $x=x_1+x_2=\sum_{1}^{k}(\lambda_i+\alpha_i)e_i+\sum_{1}^{s}\mu_j g_j+\sum_{1}^{m}\beta_t f_t$, тобто x – лінійна комбінація системи елементів ряду (4).

Із 1. і 2. випливає, що (4) – базис L. Так як система (4) має $k+s+m$ елементів, то $\dim L = k+s+m$. ■

Попередній приклад так ілюструє дану теорему:

$$\dim L_1=\dim L_2=1,\ \dim L_1 \bigcap L_2=0,\ \dim L_1+L_2=2 \text{ і}$$
$$1 + 1 = 0 + 2.$$

Нехай L_1 і L_2 – підпростори n - мірного простору V. Простір L називається ***прямою сумою*** підпросторів L_1 і L_2, якщо кожний елемент x простору L може бути однозначно утворений сумою $x_1+x_2=x$, де $x_1 \in L_1, x_2 \in L_2$. Позначається $L = L_1 \oplus L_2$.

Говорять також, що $L = L_1 \oplus L_2$ – ***розкладення*** простору L ***в пряму суму*** підпросторів L_1 і L_2.

Теорема 3.7.4. (Критерій прямої суми) *Для того, щоб простір L представляв собою пряму суму підпросторів L_1 і L_2, необхідно і достатньо, щоб виконувались наступні умови*

1) $L_1 \bigcap L_2 = \{\theta\}$,

2) $\dim L = \dim L_1 + \dim L_2$.

Н е о б х і д н і с т ь. Якщо $L = L_1 \oplus L_2$, то $L = L_1 + L_2$, де $\forall x\, \exists!\, x_1 \in L_1, x_2 \in L_2$ такі, що $x = x_1 + x_2$. З теореми 1.7.3 про суму розмірностей підпросторів маємо:

$$\dim L = \dim L_1 + \dim L_2 - \dim L_1 \bigcap L_2. \tag{1}$$

Розглянемо підпростір $L_1 \bigcap L_2$. Із означення лінійного простору $\forall\, y \in L_1 \bigcap L_2, \exists -y \in L_1 \bigcap L_2$, тобто $y, -y \in L_1, L_2$. Тоді

$$x = x_1 + x_2 = x_1 + x_2 + (y - y) = \underset{\in L_1}{(x_1 + y)} + \underset{\in L_2}{(x_2 - y)}.$$

Якщо $y \neq \theta$, то $x_1 + y \neq x_1$, тобто $x = x_1 + x_2$ має інше розкладення $x = x_1' + x_2'$ (де $x_1' = x_1 + y \in L_1$, $x_2' = x_2 - y \in L_2$), що суперечить означенню прямої суми. Отже $y = \theta$ і виконується перша умова теореми: $L_1 \bigcap L_2 = \{\theta\}$. Тоді $\dim L_1 \bigcap L_2 = 0$ і за формулою (1) $\dim L = \dim L_1 + \dim L_2 - 0$, тобто виконується друга умова теореми. ▲

Д о с т а т н і с т ь. Нехай простір L містить в собі підпростори L_1 і L_2 та виконуються умови 1) $L_1 \bigcap L_2 = \{0\}$ і 2) $\dim L = \dim L_1 + \dim L_2$. Нехай

$e_1, e_2, ..., e_k$ – базис L_1,

$g_1, g_2, ..., g_m$ – базис L_2.

Тоді умова 2) буде мати вигляд: $\dim L = k + m$.

Елементи $e_1, e_2, ..., e_k, g_1, g_2, ..., g_m$ належать простору L. Розглянемо систему елементів $e_1, e_2, ..., e_k, g_1, g_2, ..., g_m$. Якщо вона ЛНЗ, то за теоремою 1.4.2 ця система буде базисом простору L. Перевіримо це:

$$\sum \alpha_i e_i + \sum \beta_j g_j = \theta \;\Rightarrow\; \underset{\in L_1}{\sum \alpha_i e_i} = -\underset{\in L_2}{\sum \beta_j g_j} \Rightarrow$$

$\sum \alpha_i e_i, \sum \beta_j g_j \in L_1 \bigcap L_2$. Так як $L_1 \bigcap L_2 = \{\theta\}$ за умовою 1), то $\sum \alpha_i e_i = \theta$ і $\sum \beta_j g_j = \theta$. Оскільки

$$e_1, e_2, ..., e_k \text{ – базис } L_1,\ g_1, g_2, ..., g_m \text{ – базис } L_2,$$

то $\alpha_i = 0, i = 1,...,k$ і $\beta_j = 0, j = 1,...,m$. Значить система $e_1,...,e_k, g_1,...,g_m$ лінійно незалежна і утворює базис L.

Покажемо, що простір L утворює суму підпросторів. Так як система $e_1,...,e_k, g_1,...,g_m$ утворює базис L, то $\forall x \in L \quad \exists \lambda_i, \mu_j \in \mathbf{R},\ i = 1,...,k,$ $j = 1,2,...,m$ такі, що $x = \sum_{i=1}^{k} \lambda_i e_i + \sum_{j=1}^{m} \mu_j g_j = x_1 + x_2$, де $x_1 = \sum_{i=1}^{k} \lambda_i e_i \in L_1$, а $x_2 = \sum_{j=1}^{m} \mu_j g_j \in L_2$. Значить, $L = L_1 + L_2$. Те, що ця сума пряма, доведемо від супротивного.

Нехай L не утворює пряму суму, тоді існує два розклади деякого елементу $x \in L$:
$x = x_1 + x_2 = x_1' + x_2'$, де $x_1, x_1' \in L_1$, $x_2, x_2' \in L_2$, причому $x_1 \neq x_1'$, або $x_2 \neq x_2'$. Тоді

$$-\begin{array}{l} x = x_1 + x_2 \\ \underline{x = x_1' + x'} \\ \theta = (x_1 - x_1') + (x_2 - x_2') \end{array} \qquad \Rightarrow \underbrace{x_1 - x_1'}_{\in L_1} = \underbrace{x_2' - x_2}_{\in L_2} \in L_1 \cap L_2 = \{\theta\}.$$

Отже $x_1 - x_1' = \theta$ і $x_2 - x_2' = \theta$, значить $x_1 = x_1'$ і $x_2 = x_2'$, що суперечить припущенню, тобто x єдиним способом розкладається в суму $x_1 + x_2$. Тому

$$L = L_1 \oplus L_2. \ \blacksquare$$

Нехай $L \subset V$ і $L \oplus L_1 = V$, тоді говорять, що L_1 – ***пряме доповнення*** простору L до простору V.

Зауваження 1. Цей критерій розповсюджується на довільну кількість прямих доданків. *Для того, щоб простір* L *представляв собою пряму суму підпросторів* L_1, L_2, … L_s, *необхідно і достатньо, щоб виконувались наступні умови*

1) $L_i \cap L_j = \{\theta\}, \quad i \neq j, \quad i, j = 1, \dots, s,$

2) $\dim L = \dim L_1 + \dim L_2 + \dots + \dim L_s.$

Зауваження 2. Із доведення останньої частини критерію 1.7.4 та теореми про суму розмірностей 1.7.3 маємо критерій прямої суми у такому вигляді. *Сума* $L_1 + L_2$ *пряма тоді і тільки тоді, коли* $L_1 \cap L_2 = \{\theta\}$. Його розповсюдження на s доданків виглядає наступним чином: *сума* $L_1 + L_2 + \dots + L_s$ *пряма тоді і тільки тоді, коли* $L_i \cap L_j = \{\theta\}, \quad i \neq j, \quad i, j = 1, \dots, s$.

Нехай $L \subset V$ і $L \oplus L_1 = V$, тоді говорять, що L_1 – ***пряме доповнення*** простору L до простору V.

Твердження 3.7.5. *Для довільного підпростору L лінійного простору V існує пряме доповнення.*

Д о в е д е н н я. Нехай $e_1, e_2, ..., e_k$ – базис L, $\dim V = n$. Для даних просторів L і V, $L \subset V$ знайдемо пряме доповнення L до V.

За теоремою 1.6.5 про доповнення базису $\exists\, e_{k+1}, \ldots, e_n \in V$ такі, що $e_1, e_2, ..., e_k, e_{k+1}, ..., e_n$ – базис V. Позначимо $L_1 = \langle\, e_{k+1}, ..., e_n \,\rangle$. Так як $e_{k+1}, \ldots, e_n$ – ЛНЗ, то $\dim L = k$, $\dim L_1 = n - k$. Отже, $n = \dim L + \dim L_1$, що задовольняє умову 2) критерію 1.7.4.

Розглянемо $L \bigcap L_1$. Нехай $x \in L \bigcap L_1$. Тоді $\exists\, \alpha_i, \beta_j \in \mathbf{R}$ такі, що $x = \sum\limits_{i=1}^{k} \alpha_i e_i = \sum\limits_{j=k+1}^{n} \beta_j e_j$. Тоді $\alpha_1 e_1 + ... + \alpha_k e_k - \beta_{k+1} e_{k+1} - ... - \beta_n e_n = \theta$. Так як $e_1, e_2, ..., e_k, e_{k+1}, ..., e_n$ – ЛНЗ, то $\alpha_i = 0, i = 1, ..., k, \beta_j = 0, j = k+1, ..., n \Rightarrow x = \theta$. Тобто $L \bigcap L_1 = \{\theta\}$ і виконується умова 1) критерію 1.7.4. Отже $V = L \oplus L_1$, а прямим доповненням L до V буде лінійна оболонка елементів доповняльного базису: $L_1 = \langle\, e_{k+1}, ..., e_n \,\rangle$.

Контрольні питання і завдання

1. Що таке сума підпросторів?
2. Знайдіть суму і перетин підпросторів L_1 і L_2, якщо $L_1 = \{(t_1 + t_2, t_1 - t_2, t_2, 0) | t_1, t_2 \in \mathbf{R}\}$, $L_2 = \{(2p_1 - p_2, 2p_1 - 3p_2, 0, p_2) | p_1, p_2 \in \mathbf{R}\}$. Знайдіть розмірності $L_1 + L_2$ і $L_1 \cap L_2$.
3. Чи буде $L_1 + L_2$ розкладенням V_4 в пряму суму, якщо L_1 і L_2 простори із попереднього завдання?

3.8. Евклідовий простір

Дійсний лінійний простір V називається ***евклідовим простором***, якщо:

І. Існує правило, за яким кожним двом елементам $x, y \in V$ ставиться у відповідність *дійсне число*, яке називається ***скалярним добутком*** цих елементів і позначається (x, y).

II. Це правило задовольняє наступним аксіомам:

1^0. $(x,y)=(y,x)$

2^0. $(x_1+x_2,y)=(x_1,y)+(x_2,y)$

3^0. $(\lambda x,y)=\lambda(x,y)$

4^0. $(x,x)>0\ \forall x\neq\theta$ і $(x,x)=0 \Leftrightarrow x=\theta$

$\forall x,y,x_1,x_2\in V,\ \forall\lambda\in R$.

Приклади евклідових просторів.

Приклад 1. Нехай V – лінійний простір векторів площини і скалярний добуток задано по правилу:

$$\forall \bar{x},\bar{y}\in V\big|\ (\bar{x},\bar{y})=|\bar{x}|\cdot|\bar{y}|\cdot\cos(\bar{x}\hat{\cdot}\bar{y}).$$

▼ Як відомо із шкільного курсу математики, умови 1^0. – 4^0. виконуються, тому V – евклідовий простір. ▲

Приклад 2. $V=V_n$. Нехай $x=(x_1,...,x_n)$, $y=(y_1,...,y_n)$, задано скалярний добуток

$$(x,y)=x_1y_1+x_2y_2+...+x_ny_n=\sum_{i=1}^{n}x_iy_i\ .$$

▼ Умови 1^0 – 4^0 виконуються (перевірте самостійно), отже V_n з таким скалярним добутком утворює евклідовий простір. ▲

Властивості евклідового простору

Зручно евклідовий простір позначати E.

5. $\forall x,y,z\in E\ \big|\ (x,y+z)=(x,y)+(x,z)$:

$$(x,y+z)\overset{1^0}{=}(y+z,x)\overset{2^0}{=}(y,x)+(z,x)\overset{1^0}{=}(x,y)+(x,z).\ \blacksquare$$

6. $\forall x,y\in E,\ \forall\lambda\in R\ \big|\ (x,\lambda y)=\lambda(x,y)$:

$(x,\lambda y)=(\lambda y,x)=\lambda(y,x)=\lambda(x,y)\blacksquare$.

7. $\forall x\in E\ \big|\ (\theta,x)=0$:

$(\theta,x)=(0\cdot y,x)=0(y,x)=0\blacksquare$.

8. $\forall x,y\in E$ *виконується нерівність Коші - Буняковського:*

$$(x,y)^2\leq(x,x)(y,y). \tag{3.8.1}$$

Д о в е д е н н я.

1. Нехай $y=\theta \overset{7,4^\circ}{\Rightarrow} (x,y)=0, (y,y)=0 \Rightarrow (x,y)^2=(x,x)(y,y)$, бо $0^2=(x,x)\cdot 0$.

2. Якщо $y\neq\theta$, то $(y,y)>0$ і

$\forall y\neq\theta, x\in E, \forall\lambda,\mu\in R \mid (\lambda x+\mu y, \lambda x+\mu y)\geq 0 \Rightarrow$.

$\Rightarrow \lambda^2(x,x)+2\lambda\mu(x,y)+\mu^2(y,y)\geq 0$. (1)

Нехай $\lambda=(y,y)$, $\mu=-(x,y)$. Підставимо ці значення в (1). Одержимо

$(y,y)^2(x,x)-2(y,y)(x,y)^2+(x,y)^2(y,y)==(y,y)[(y,y)(x,x)-(x,y)^2]\geq 0$ (2)

Так як $(y,y)>0$, то обидві частини нерівності (2) можна поділити на (y,y) і знак нерівності при цьому не зміниться. Одержимо

$(x,y)^2\leq(x,x)(y,y)$. ■

Лінійний простір V називається ***нормованим***, якщо виконуються наступні умови:

I. Існує правило, яке кожному елементу простору ставить у відповідність *дійсне число*, яке називається ***нормою*** даного елементу і позначається $\|x\|$.

II. Це правило задовольняє аксіомам:

1^*. $\|x\|>0, \forall x\neq\theta$ і $\|x\|=0 \Leftrightarrow x=\theta$.

2^*. $\|\lambda x\|=|\lambda|\cdot\|x\| \quad \forall x\in V, \forall\lambda\in R$.

3^*. $\forall x,y\in V \quad | \ \|x+y\|\leq\|x\|+\|y\|$ – нерівність трикутника.

9. **Твердження 3.8.1**. *Довільний евклідовий простір E є нормованим.*

Д о в е д е н н я. Задамо на евклідовому просторі E правило, яке $\forall x\in E$ визначається рівнянням $\|x\|=\sqrt{(x,x)}$. Покажемо, що воно задовольняє аксіомам норми.

Перевіримо 1^*. Із аксіоми 4^0 евклідового простору

$$\|x\|=\sqrt{(x,x)}=\begin{cases}0, & x=\theta\\ >0, & x\neq\theta\end{cases}.$$

Перевіримо 2^*. Із аксіоми 3^0 і властивості 6 евклідового простору

$$\|\lambda x\| =\sqrt{(\lambda x,\lambda x)}=\sqrt{\lambda^2(x,x)}=|\lambda|\cdot\|x\| \quad | \quad \forall x\in E, \forall\lambda\in R.$$

<u>Зауваження</u>. Нерівність (3.8.1) Коші - Буняковського при визначеному таким чином понятті $\|x\|$ може бути записана у вигляді

$(x,y)^2 \le \|x\|^2 \cdot \|y\|^2$, або $|(x,y)| \le \|x\| \cdot \|y\|$, або

$$-\|x\| \cdot \|y\| \le (x,y) \le \|x\| \cdot \|y\|, \qquad (3.8.2)$$

а також маємо

$$(x,y) \le \|x\| \cdot \|y\|. \qquad (3.8.3)$$

Перевіримо 3^*. Із аксіоми 2^0 і властивості 5 евклідового простору $\|x+y\| = \sqrt{(x+y,x+y)} = \sqrt{(x,x)+2(x,y)+(y,y)}$. Врахувавши нерівність (1.10.3), одержимо, що

$$\|x+y\| = \sqrt{(x,x)+2(x,y)+(y,y)} \le \sqrt{\|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2} = \|x\| + \|y\|.$$

Отже $\|x\| = \sqrt{(x,x)}$ задовольняє аксіомам норми і E – нормований простір. ■

Кутом між елементами $x, y \ne \theta$ евклідового простору E називається дійсне число φ, яке задовольняє умові $\cos\varphi = \dfrac{(x,y)}{\|x\| \cdot \|y\|}, 0 \le \varphi \le \pi$.

Зауважимо, що із нерівності (1.8.2) випливає $-1 \le \dfrac{(x,y)}{\|x\| \cdot \|y\|} \le 1$, тобто кут між x та y завжди однозначно визначений. Елементи x та y називаються ***ортогональними***, якщо $(x,y) = 0$. За означенням кут між такими елементами дорівнює 90^0.

10. Твердження 3.8.2. *В довільному евклідовому просторі справедлива теорема Піфагора: квадрат гіпотенузи дорівнює сумі квадратів катетів* (x і y називаються ***катетами***, якщо вони ортогональні, а $x+y$ у цьому випадку називається ***гіпотенузою***), тобто $\|x+y\|^2 = \|x\|^2 + \|y\|^2$, де $(x,y) = 0$.

Д о в е д е н н я випливає із властивостей евклідового простору та означення ортогональних елементів :

$$\|x+y\|^2 = (x+y,x+y) = (x,x) + 2(x,y) + (y,y) = \|x\|^2 + \|y\|^2. ■$$

Наслідок 3.8.3. *Якщо* $x_1, x_2, \ldots, x_k$ *– попарно ортогональні, а* $z = x_1 + x_2 + \ldots + x_k$ *(z представляє діагональ* k*- мірного прямокутного паралелепіпеда), то*

$$\|z\|^2 = \|x_1\|^2 + \|x_2\|^2 + \ldots + \|x_k\|^2.$$

Розглянемо залежність скалярного добутку елементів $x, y \in E$ від їх координат в базисі $e_1, e_2, ..., e_n$. Нехай $x_e = (x_1, x_2, ...x_n)$, $y_e = (y_1, y_2, ..., y_n)$. Тоді, враховуючи властивості $2^0, 3^0, 5, 6$ евклідового простору, одержуємо

$$(x, y) = (\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j (e_i, e_j) - (3.8.4)$$

скалярний добуток двох елементів дорівнює сумі попарних добутків координат цих елементів, помножених на скалярні добутки відповідних елементів базису.

Контрольні питання і завдання

1. Сформулюйте визначення дійсного евклідового простору і наведіть приклад такого простору.
2. Яким чином можна ввести скалярне множення елементів: у просторі T_n; у просторі $C[a, b]$?
3. Як виражається скалярний добуток елементів евклідового простору через координати елементів у довільному базисі?
4. Як визначаються норма (довжина) елементу й кут між ненульовими елементами евклідового простору?
5. Яка нерівність називається нерівністю Коші-Буняковського?
6. Запишіть значення норми елемента і нерівність Коші-Буняковського в просторах T_n, $C[a, b]$.

3.9. Ортогональні системи

Теорема 3.9.1. (про лінійну незалежність ортогональних елементів). *Якщо ненульові елементи $u_1, u_2, ..., u_m$ евклідового простору попарно ортогональні, то вони лінійно незалежні.*

Д о в е д е н н я. Нехай

$$\alpha_1 u_1 + ... + \alpha_m u_m = \theta. \qquad (1)$$

Помножимо скалярно на u_i обидві частини рівності (1) $\forall i = 1, 2, ..., m$. Враховуючи властивості 2^0, 3^0, 7 евклідового простору, одержимо

$$\alpha_1(u_1, u_i) + \alpha_2(u_2, u_i) + ... + \alpha_i(u_i, u_i) + ... + \alpha_m(u_m, u_i) = (\theta, u_i) = 0.$$

Враховуючи, що $(u_k, u_i) = 0$, при $k \neq i$, то

$$\alpha_i (u_i, u_i) = 0. \qquad (2)$$

А так як $u_i \neq \theta$ і $(u_i, u_i) \neq 0$, то одержуємо, що із рівності (2) випливає $\alpha_i = 0 \forall i = 1, \ldots, m$. ■

Базис евклідового простору, що утворює ортогональну систему, називається ***ортогональним базисом***. З теорем 3.8.1 та 3.4.2 довільна ортогональна система ненульових елементів в кількості, рівній розмірності простору E, утворює ортогональний базис E.

Теорема 3.9.2. (про побудову ортогональної системи). *Якщо ненульові елементи евклідового простору $u_1, u_2, \ldots, u_s$ утворюють ортогональну систему, і елемент v не виражається лінійно через них, то при*

$$w = v + \lambda_1 u_1 + \ldots + \lambda_s u_s, \text{ де } \lambda_i = -\frac{(v, u_i)}{(u_i, u_i)} \quad (i = 1, \ldots, s)$$

система $u_1, u_2, \ldots, u_s, w$ буде ортогональною. При цьому $w \neq \theta$.

Д о в е д е н н я. Нехай $w = v + \lambda_1 u_1 + \ldots + \lambda_s u_s$, $\lambda_i = -\dfrac{(v, u_i)}{(u_i, u_i)} \quad (i = 1, \ldots, s)$.

Враховуючи, що $(u_i, u_j) = 0, \quad i \neq j$, розглянемо

$$(w, u_i) = (v, u_i) + \lambda_1 (u_1, u_i) + \ldots + \lambda_i (u_i, u_i) + \ldots + \lambda_s (u_s, u_i) = (v, u_i) + \lambda_i (u_i, u_i) =$$

$$= (v, u_i) + \left[-\frac{(v, u_i)}{(u_i, u_i)} \right] (u_i, u_i) = 0.$$

Отже система $u_1, u_2, \ldots, u_s, w$ – ортогональна.

Нехай $w = \theta$. Тоді із умови теореми $v = -\lambda_1 u_1 - \ldots - \lambda_s u_s$, а це суперечить умові, що елемент v не виражається лінійно через $u_1, u_2, \ldots, u_s$.Значить зроблене припущення не вірне і $w \neq \theta$. ■

Наслідок 3.9.3. *Нехай в системі лінійно незалежних елементів евклідового простору $u_1, u_2, \ldots, u_m$ елементи $u_1, u_2, \ldots, u_s$ $(s \leq m)$ утворюють ортогональну систему. Тоді існує ортогональна система з ненульових елементів $v_1, v_2, \ldots, v_m$, в якій виконуються умови*

(1) $v_i = u_i$ при $i = 1, \ldots, s$*

(2) $\forall k = 1, 2, \ldots, m$ елемент v_k лінійно виражається через $u_1, u_2, \ldots, u_k$.*

Д о в е д е н н я проведемо методом математичної індукції по m.

Якщо $s = m$, то все очевидно $\forall m = 1,2,...$. В тому числі, при $m = 1$ твердження істинне.

Нехай $s < m$. Припустимо, що твердження істинне для $m = t$, ($t > s$), тобто $v_1, v_2, ..., v_t$ – ортогональна система і виконуються умови (1*) і (2*). Доведемо твердження для $m = t+1$. Розглянемо елемент

$$v_{t+1} = u_{t+1} + \lambda_1 v_1 + ... + \lambda_t v_t,\ \lambda_j = -\frac{(u_t, v_j)}{(v_j, v_j)},\ j = 1,2,...,t.$$

За теоремою 1.9.2 він взаємно ортогональний з усіма $v_1, v_2, ..., v_t$ а також лінійно виражається через $u_1, u_2, ..., u_{t+1}$, бо за припущенням $v_1, v_2, ..., v_t$ лінійно виражаються через $u_1, u_2, ..., u_t$ за умовою (2*). ■

На основі теореми 3.9.2. та наслідку 3.9.3 має місце **процес ортогоналізації:** *побудова ортогональної системи елементів* $v_1, v_2, ..., v_m$ *евклідового простору із даної лінійно незалежної системи* $u_1, u_2, ..., u_m$.

1. Вибираємо всі попарно ортогональні елементи $u_1, u_2, ..., u_s$ (в усякому випадку можемо взяти $s = 1$, бо $u_1 \neq \theta$ є ортогональною системою).

2. Будуємо шукану систему $v_1, v_2, ..., v_m$:

1) $v_i = u_i,\ i = 1,...,s$.

2) $v_t = u_t + \lambda_1^{(t)} v_1 + ... + \lambda_{t-1}^{(t)} v_{t-1},\ t = s+1,...,m$ при

$$\lambda_j^{(t)} = -\frac{(u_t, v_j)}{(v_j, v_j)},\ j = 1,...,t-1.$$

Застосування такого процесу до довільного базису евклідового простору дає можливість побудувати ортогональний базис і має місце

Твердження 3.9.4. (про ортогональні базиси) *Довільний скінченномірний евклідовий простір* E *має ортогональні базиси.*

Також довільну ортогональну систему ненульових елементів із E можна доповнити до ортогонального базису простору. Для цього дану систему доповнюють до деякого базису, а потім до нього застосовують процес ортогоналізації.

Розглянемо залежність с к а л я р н о г о д о б у т к у елементів $x, y \in E$ від їх координат в о р т о г о н а л ь н о м у базисі $e_1, e_2, ..., e_n$. Нехай

$x_e = (x_1, x_2, ... x_n)$, $y_e = (y_1, y_2, ..., y_n)$. Тоді, враховуючи залежність (2.10.4) та умову ортогональності базису ($(e_i, e_j) = 0, i \neq j$), одержуємо

$$(x, y) = \sum_{i=1}^{n} x_i y_i (e_i, e_i) \; - \tag{3.9.1}$$

скалярний добуток двох елементів дорівнює сумі добутків відповідних координат цих елементів в ортогональному базисі, помножених на скалярні квадрати відповідних елементів базису.

Елемент $e \in E$ називається ***нормованим***, якщо його скалярний квадрат дорівнює 1: $(e, e) = 1$.

Твердження 3.9.5. *$\forall u \in E, u \neq \theta$ елементи λu при $\lambda = \pm \frac{1}{\sqrt{(u,u)}}$ і тільки при таких λ є нормованими.*

Н е о б х і д н і с т ь. $(\lambda u, \lambda u) = \lambda^2 (u, u) = \frac{1}{(u,u)}(u, u) = 1$. ▲

Д о с т а т н і с т ь. Нехай αu є нормованим елементом. Тоді $1 = (\alpha u, \alpha u) = \alpha^2 (u, u) \Rightarrow \alpha = \pm \frac{1}{\sqrt{(u,u)}}$. ■

Перехід від u до $\frac{1}{\sqrt{(u,u)}} u$ називають ***нормуванням*** u. Якщо всі елементи ортогонального базису нормовані, то такий базис називається ***ортонормованим***. Із довільного ортогонального базису можна утворити ортонормований, якщо нормувати кожний його елемент, так як множення векторів на число не впливає на їх ортогональність, отже приходимо до

Твердження 3.9.6. *Довільний скінченномірний евклідовий простір E має ортонормовані базиси.*

З теорем 3.9.1 та 2.4.2 довільна ортонормована система ненульових елементів в кількості, рівній розмірності простору E, утворює ортонормований базис $e_1, e_2, ..., e_n$ простору E, який задовольняє умовам: $(e_i, e_j) = \delta_{ij}$, де $\delta_{ij} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$ – символи Кронекера.

Розглянемо залежність с к а л я р н о г о д о б у т к у елементів $x, y \in E$ від їх координат в о р т о н о р м о в а н о м у базисі $e_1, e_2, ..., e_n$. Нехай

$x_e=(x_1,x_2,...x_n)$, $y_e=(y_1,y_2,...,y_n)$. Тоді, враховуючи залежність (2.8.1) та умову ортонормованості базису ($(e_i,e_j)=\delta_{ij}$), одержуємо

$$(x,y)=\sum_{i=1}^{n}x_i y_i \; - \qquad (3.9.2)$$

скалярний добуток двох елементів дорівнює сумі добутків відповідних координат цих елементів в ортонормованому базисі.

Звернемо увагу, що справедливе і обернене твердження: *якщо скалярний добуток задано по правилу (1.9.2), то базис ортонормований.* В цьому легко впевнитися, якщо розглянути координати елементів базису в цьому ж базисі:

$$e_1=(1,0,0,...,0)_e,$$
$$e_2=(0,1,0,...,0)_e,$$
$$.....,$$
$$e_n=(0,0,0,...,1)_e$$

і попарно перемножити їх за правилом (3.9.2): $(e_i,e_j)=\delta_{ij}$, $i,j=1,2,...,n$.

Отже справедлива

Теорема 3.9.7. *Скалярний добуток елементів простору E дорівнює сумі добутків відповідних координат в деякому базисі тоді і тільки тоді, коли цей базис ортонормований.*

Зауваження 3.9.8. *Координати елементу евклідового простору в ортонормованому базисі дорівнюють скалярному добутку цього елементу на відповідні елементи базису.*

Д о в е д е н н я. Нехай $e_1,e_2,...,e_n$ – ортонормований базис простору E, $x\in E$, $x_e=(x_1,x_2,...x_n)$. Тоді $(x,e_k)=\sum_{i=1}^{n}x_i(e_i,e_k)=x_k, k=1,2,...,n$. ■

Зауваження 3.9.9. *Для довільних додатних чисел a_{i1}, a_{i2} справедлива нерівність Шварца:*

$$\left(\sum_{i=1}^{n}\frac{a_{i1}}{a_{i2}}\right)\cdot\left(\sum_{i=1}^{n}a_{i1}\cdot a_{i2}\right)\geq\left(\sum_{i=1}^{n}a_{i1}\right)^2.$$

Д о в е д е н н я. Нехай $e_1,e_2,...,e_n$ – ортонормований базис простору E, $x,y\in E,\quad x_e=(x_1,\ldots,x_n), y_e=(y_1,\ldots,y_n)$. Тоді

$$(x,y)=\sum_{i=1}^{n} x_i y_i, \quad (x,x)=\sum_{i=1}^{n} x_i^2, \quad (y,y)=\sum_{i=1}^{n} y_i^2 .$$

Якщо взяти $x_i = \sqrt{\dfrac{a_{i1}}{a_{i2}}},\ y_i = \sqrt{a_{i1} \cdot a_{i2}}$, то

$$(x,y)=\sum_{i=1}^{n}\sqrt{\frac{a_{i1}}{a_{i2}}}\sqrt{a_{i1}a_{i2}}=\sum_{i=1}^{n} a_{i1},\ (x,x)=\sum_{i=1}^{n}\frac{a_{i1}}{a_{i2}},\ (y,y)=\sum_{i=1}^{n} a_{i1}\cdot a_{i2}$$

Нерівність Коші - Буняковського має вид $(x,y)^2 \le (x,x)(y,y)$. Підставляючи в неї одержані співвідношення, отримуємо нерівність Шварца:

$$\left(\sum_{i=1}^{n}\frac{a_{i1}}{a_{i2}}\right)\cdot\left(\sum_{i=1}^{n} a_{i1}\cdot a_{i2}\right)\ge\left(\sum_{i=1}^{n} a_{i1}\right)^2 . \blacksquare$$

Контрольні питання і завдання

1. Які елементи евклідового простору називаються ортогональними?

2. У просторі $C[0,1]$ скалярний добуток функцій $f(x)$ і $g(x)$ визначено формулою $(f,g)=\int_0^1 f(x)g(x)dx.$

3. Чи є ортогональними функції

$f_1(x)=1$ і $f_2(x)=x-\dfrac{1}{2}$; $f_1(x)$ і $f_3(x)=\cos\pi x$; $f_2(x)$ і $f_3(x)$?

4. Доведіть, що якщо елементи x_1 і x_2 ортогональні, то елементи c_1x_1 і c_2x_2 також ортогональні при будь-яких числах c_1 і c_2.

5. Довести, що якщо ненульові елементи x і y евклідового простору ортогональні, то вони лінійно незалежні. Чи вірне обернене твердження?

6. Що таке ортогональний базис? Що таке ортонормований базис? Наведіть приклади ортонормованих базисів.

7. Опишіть процедуру побудови ортонормованого базису на основі довільного базису (процесу ортогоналізації).

8. Скільки ортонормованих базисів існує в даному n- мірному евклідовому просторі?

9. Напишіть формулу для координат довільного елементу в ортонормованому базисі.

10. Напишіть формулу скалярного добутку елементів x і y через їхні координати в ортонормованому базисі евклідового простору?

11. Як виражається норма довільного елементу через його координати в ортонормованому базисі евклідового простору?

3.10. Унітарний простір

Нехай $V = V_{\mathbf{C}}$ – комплексний лінійний простір.

Комплексний лінійний простір V називається ***унітарним (комплексним евклідовим) простором***, якщо:

I. Існує правило, за яким кожним двом елементам $x, y \in V$ ставиться у відповідність *комплексне число*, яке називається ***скалярним добутком*** цих елементів і позначається (x, y).

II. Це правило задовольняє наступним аксіомам:

1^0. $(x, y) = \overline{(y, x)}$ ($\bar{z}$ – спряжене до z);

2^0. $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$;

3^0. $(\lambda x, y) = \lambda(x, y)$;

4^0. $(x, x) > 0\ \forall x \neq \theta$ і $(x, x) = 0 \Leftrightarrow x = \theta$.

$\forall x, y, x_1, x_2 \in V, \forall \lambda \in \mathbf{C}$.

Властивості унітарного простору

Зручно унітарний простір позначати U.

Із аксіом 1^0 – 4^0 випливають властивості унітарного простору, які є аналогічними до властивостей дійсного евклідового простору, але мають певні відмінності по суті і відрізняються доведеннями в тій мірі, в якій відрізняються множини дійсних і комплексних чисел: $\forall x, y, y_1, y_2 \in U, \forall \lambda \in \mathbf{C}$

5. $(x, \lambda y) = \bar{\lambda}(x, y)$: ▼

$(x, \lambda y) = \bar{\lambda}(x, y)(x, \lambda y) = \overline{(\lambda y, x)} = \overline{\lambda \cdot (y, x)} = \bar{\lambda} \cdot \overline{(y, x)} = \bar{\lambda}(x, y)$ ▲

6. $(x_1, y_1 + y_2) = (x_1, y_1) + (x_1, y_2)$ (довести самостійно).

7. $\forall x \in U \mid (\theta, x) = 0$: ▼ $(\theta, x) = (0 \cdot y, x) = 0(y, x) = 0$. ▲

8. *Для унітарного простору виконується нерівність Коші-Буняковського* $|(x, y)|^2 \leq (x, x)(y, y)$.

9. *Довільний унітарний простір нормований*, якщо $\|x\| = \sqrt{(x,x)}$ (а значить *в унітарному просторі виконується нерівність трикутника* $\|x+y\| \le \|x\| + \|y\|$).

<u>Зауваження</u>. Поняття кута φ між елементами x, y унітарного простору втрачає своє значення, так як $(x, y) \in \mathbf{C}$, а $\cos\varphi \notin \mathbf{C}$. Але елементи $x, y \in U$ називаються ***ортогональними***, якщо $(x, y) = 0$.

10. *В довільному унітарному просторі справедлива теорема Піфагора.*

Для унітарних просторів справедливими залишаються всі властивості ортогональних систем евклідового простору, не змінюються поняття норми і нормування елементу, тому справедливі також твердження 3.9.4, 3.9.6.

Розглянемо скалярний добуток елементів $x, y \in U$ в ортонормованому базисі $e_1, e_2, ..., e_n$. Нехай $x_e = (x_1, x_2, ..., x_n)$, $y_e = (y_1, y_2, ..., y_n)$.Враховуючи властивості унітарного простору і те, що $(e_i, e_j) = \delta_{ij}$, маємо

$$(x, y) = (\sum x_i e_i, \sum y_j e_j) = \sum_{i,j} x_i \overline{y_j} (e_i, e_j) = \sum_i x_i \overline{y_i}.$$

Скалярний добуток двох елементів унітарного простору в ортонормованому базисі дорівнює сумі добутків координат одного елементу на число, спряжене до відповідної координати другого елементу.

Контрольні питання і завдання

1. Сформулюйте визначення унітарного простору і наведіть приклад такого простору.

2. Яким чином можна ввести скалярне множення елементів:

а) у просторі T_n^* стовпчиків з n комплексними елементами;

б) у просторі $C^*[a,b]$ неперервних на $[a,b]$ комплекснозначних функцій?

3. Доведіть властивість скалярного добутку в унітарному просторі, що $(x, y+z) = (x, y) + (x, z)$.

4. Як виражається скалярний добуток елементів унітарного простору через координати елементів у довільному базисі?

5. Запишіть значення норми елемента і нерівність Коші-Буняковського в просторах T_n^*, $C^*[a,b]$.

6. Напишіть формулу скалярного добутку елементів x і y через їхні координати в ортонормованому базисі унітарного простору?

7. Як виражається норма довільного елементу через його координати в ортонормованому базисі унітарного простору?

Розділ 4. ЛІНІЙНІ ОПЕРАТОРИ

4.1. Поняття лінійного оператору

Нехай V і W – числові лінійні простори.

Оператором називається відображення $\widetilde{A}$, що діє із V в W $\left(\widetilde{A}: V \to W\right)$ та ставить у відповідність кожному елементу $x \in V$ деякий елемент $y \in W$. Елемент y називається ***образом*** елементу x при дії оператора $\widetilde{A}$, а елемент x – ***прообразом*** елементу y. При цьому позначаємо

$y = \widetilde{A}(x)$ або $y = \widetilde{A}x$. (4.1.1)

Оператор $\widetilde{A}$, що діє із V в W називається ***лінійним***, якщо $\forall\, x, y \in V$ і для довільного числа λ виконуються наступні співвідношення.

1^0. $\widetilde{A}(x+y) = \widetilde{A}(x) + \widetilde{A}(y)$;

2^0. $\widetilde{A}(\lambda x) = \lambda \widetilde{A} x$.

Якщо $W = R$ або $W = C$ то $\widetilde{A}: V \to W$ називається ***лінійною формою*** або ***лінійним функціоналом***.

Якщо $W = V$, то $\widetilde{A}: V \to V$ називають ***лінійним перетворенням***.

Перевіримо на прикладах лінійність окремих операторів.

Приклад 1. $\widetilde{A}: R^3 \to R^4$; $\widetilde{A}(x_1, x_2, x_3) = (x_1 - x_3, x_3, x_1, x_3 + x_2)$.

▼ Нехай $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$. Тоді

$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$, $\widetilde{A}(x) = (x_1 - x_3, x_3, x_1, x_3 + x_2)$,

$\widetilde{A}(y) = \widetilde{A}(y_1, y_2, y_3) = (y_1 - y_3, y_3, y_1, y_3 + y_2)$,

$\widetilde{A}(x+y) = (x_1 + y_1 - (x_3 + y_3), x_3 + y_3, x_1 + y_1, x_3 + y_3 + x_2 + y_2)$. Так як

$\widetilde{A}(x) + \widetilde{A}(y) = (x_1 - x_3 + y_1 - y_3, x_3 + y_3, x_1 + y_1, x_3 + x_2 + y_3 + y_2) = \widetilde{A}(x+y)$, то виконується співвідношення 1^0 із означення лінійного оператору.

Так як $\lambda x = (\lambda x_1, \lambda x_2, \lambda x_3)$, то $\widetilde{A}(\lambda x) = (\lambda x_1 - \lambda x_3, \lambda x_3, \lambda x_1, \lambda x_3 + \lambda x_2) =$
$\lambda(x_1 - x_3, x_3, x_1, x_3 + x_2) = \lambda \widetilde{A}(x)$. Значить виконується співвідношення 2^0 із означення лінійного оператору, тобто $\widetilde{A}(x_1, x_2, x_3) = (x_1 - x_3, x_3, x_1, x_3 + x_2)$ – лінійний оператор. ▲

Приклад 2. $\widetilde{A}:V \to V$; $\widetilde{A}x = 2x$.

▼ Так як $\widetilde{A}y = 2y$, $\widetilde{A}(x+y) = 2(x+y)$, $\widetilde{A}x + \widetilde{A}y = 2x + 2y = \widetilde{A}(x+y)$, то виконується співвідношення 1^0 із означення лінійного оператору.

Так як $\widetilde{A}(\lambda x) = 2(\lambda x) = \lambda(2x) = \lambda\widetilde{A}x$, то виконується співвідношення 2^0 із означення лінійного оператору, тобто $\widetilde{A}x = 2x$ – лінійний оператор. ▲

Приклад 3. $\widetilde{A}:R^2 \to R^2$; $\widetilde{A}x = (2,2) + x$.

▼ Так як $\widetilde{A}y = (2,2) + y$, $\widetilde{A}(x+y) = (2,2) + x + y$,

$\widetilde{A}x + \widetilde{A}y = (2,2) + x + (2,2) + y = (4,4) + x + y \neq \widetilde{A}(x+y)$, то не виконується співвідношення 1^0 із означення лінійного оператору, а значить $\widetilde{A}x = (2,2) + x$ не є лінійним оператором. ▲

Приклад 4. $\widetilde{A}:R^3 \to R^2$; $\widetilde{A}(x_1, x_2, x_3) = (x_1 \cdot x_2, x_3)$.

▼ Нехай $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$. Тоді

$$x + y = (x_1 + y_1, x_2 + y_2, x_3 + y_3),\ \widetilde{A}(x) = (x_1 \cdot x_2, x_3),\ \widetilde{A}(y) = (y_1 \cdot y_2, y_3),$$

$$\widetilde{A}(x+y) = ((x_1 + y_1)\cdot(x_2 + y_2), x_3 + y_3) = (x_1x_2 + x_1y_2 + y_1x_2 + y_1y_2, x_3 + y_3).$$

Так як $\widetilde{A}(x) + \widetilde{A}(y) = (x_1x_2 + y_1y_2, x_3 + y_3) \neq \widetilde{A}(x+y)$, то не виконується співвідношення 1^0 із означення лінійного оператору, а значить $\widetilde{A}(x_1, x_2, x_3) = (x_1 \cdot x_2, x_3)$ не є лінійним оператором. ▲

Зауваження 4.1.1. *Лінійний оператор* $\widetilde{A}$ *нейтральний елемент відображає у нейтральний* ($\widetilde{A}\theta = \theta'$):

Враховуючи властивість 2 лінійного простору (п. 3.2), одержуємо $\forall x \in V$, $\widetilde{A}\theta = \widetilde{A}(0x) \overset{2^0}{=} 0\widetilde{A}(x) = \theta'$. ■

Зауваження 4.1.2. *Умови* $1^0, 2^0$ *лінійного оператора можуть бути замінені однією умовою:* $\widetilde{A}(\lambda x_1 + \mu x_2) = \lambda\widetilde{A}x_1 + \mu\widetilde{A}x_2$ $\forall\, x_1, x_2 \in V$, $\forall\, \lambda, \mu \in R$.

4.2. Простір лінійних операторів

Розглянемо множину всіх лінійних операторів $\widetilde{A}:V \to W$, яку позначимо $L(V,W)$. На цій множині задамо внутрішню операцію додавання та зовнішню множення на число. Нехай $\widetilde{A}, \widetilde{B} \in L(V,W)$.

$\widetilde{A}+\widetilde{B}:\quad (\widetilde{A}+\widetilde{B})x=\widetilde{A}x+\widetilde{B}x\quad \lambda A:\quad (\lambda\widetilde{A})x=\lambda\widetilde{A}x$.

У множині $L(V,W)$ існує оператор $\widetilde{O}:V\to W$ $(\forall x\in V,\ \widetilde{O}x=\theta,\ \theta\in W)$. Він називається ***нульовим,*** задовольняє умовам лінійності, виконує роль нейтрального на множині $L(V,W)$ відносно додавання ($\widetilde{A}+\widetilde{O}=\widetilde{O}+\widetilde{A}=\widetilde{A}$).

$\forall\widetilde{A}\in L(V,W)$ існує оператор $-\widetilde{A}:V\to W$ $(\forall x\in V,\ (-\widetilde{A})x=-1(\widetilde{A}x)$, тобто $-\widetilde{A}=(-1)\widetilde{A}$). Цей оператор буде протилежним до $\widetilde{A}$ ($\widetilde{A}+(-\widetilde{A})=\widetilde{O}=-\widetilde{A}+\widetilde{A}$, і він задовольняє умовам лінійності.

Існування двох операцій на $L(V,W)$ (внутрішньої та зовнішньої) дає нам можливість перевіряти виконання умов лінійного простору. $\forall\widetilde{A},\widetilde{B},\widetilde{C}\in L(V,W),\quad \forall\alpha,\beta\in R$

1. $\widetilde{A}+\widetilde{B}=\widetilde{B}+\widetilde{A}$;
2. $(\widetilde{A}+\widetilde{B})+\widetilde{C}=\widetilde{A}+(\widetilde{B}+\widetilde{C})$;
3. $\exists\widetilde{O}\in L(V,W):\widetilde{O}+\widetilde{A}=\widetilde{A}$;
4. $\forall\widetilde{A}\in L(V,W)\,\exists-\widetilde{A}\in L(V,W):\widetilde{A}+(-\widetilde{A})=O$;
5. $1\cdot\widetilde{A}=\widetilde{A}$;
6. $(\alpha\beta)\widetilde{A}=\alpha(\beta\widetilde{A})$;
7. $(\alpha+\beta)\widetilde{A}=\alpha\widetilde{A}+\beta\widetilde{A}$;
8. $\alpha(\widetilde{A}+\widetilde{B})=\alpha\widetilde{A}+\alpha\widetilde{B}$. (перевірте).

Всі ці умови виконуються, отже маємо

Твердження 4.2.1. *Множина $L(V,W)$ всіх лінійних операторів із простору V в простір W із заданими операціями додавання та множення на число утворює лінійний простір.*

В просторі лінійних перетворень $L(V,V)$***,*** крім нульового перетворення $\widetilde{O}x=\theta$, існує *тотожне* перетворення $\widetilde{I}x=x,\ \forall x\in V$ (покажіть, що $\widetilde{I}\in L(V,V)$).

Добутком операторів $\widetilde{A}$ і $\widetilde{B}\in L(V,W)$ називається оператор $\widetilde{A}\widetilde{B}$, що діє за правилом: $(\widetilde{A}\widetilde{B})x=\widetilde{A}(\widetilde{B}x)$. (Покажіть, що $\widetilde{A},\widetilde{B},\widetilde{A}\widetilde{B}\in L(V,V)$.)

Властивості добутку в просторі лінійних перетворень

$$\forall \widetilde{A}, \widetilde{B}, \widetilde{C} \in L(V,V),\ \forall\ \lambda \in R$$

1^0 $\lambda(\widetilde{A}\widetilde{B}) = (\lambda\widetilde{A})\widetilde{B}$;

2^0 $\widetilde{A}(\widetilde{B}\widetilde{C}) = (\widetilde{A}\widetilde{B})\widetilde{C}$;

Властивість 2^0 дає змогу визначити добуток $\widetilde{A}\widetilde{B}...\widetilde{C}$ довільного скінченного числа співмножників із $L(V,V)$ та, зокрема, $\widetilde{A}^n = \underbrace{\widetilde{A}\widetilde{A}...\widetilde{A}}_{n}$.

Очевидно, $\widetilde{A}^m\widetilde{A}^n = \widetilde{A}^{m+n} = \widetilde{A}^{n+m} = \widetilde{A}^n\widetilde{A}^m$.

3^0 Взагалі кажучи, $\widetilde{A}\widetilde{B} \neq \widetilde{B}\widetilde{A}$.

4^0 Тотожній оператор $\widetilde{I} \in L(V,V)$ виконує роль нейтрального (одиничного) відносно множення на $L(V,V)$:

1) $(\widetilde{I}\widetilde{A})x = \widetilde{I}(\widetilde{A}x) = \widetilde{A}x \Rightarrow \widetilde{I}\widetilde{A} = \widetilde{A}$;

2) $(\widetilde{A}\widetilde{I})x = \widetilde{A}(\widetilde{I}x) = \widetilde{A}x \Rightarrow \widetilde{A}\widetilde{I} = \widetilde{A}$.

5^0 $(\widetilde{A} + \widetilde{B})\widetilde{C} = \widetilde{A}\widetilde{C} + \widetilde{B}\widetilde{C}$;

6^0 $\widetilde{A}(\widetilde{B} + \widetilde{C}) = \widetilde{A}\widetilde{B} + \widetilde{A}\widetilde{C}$.

Контрольні питання і завдання

1. Що таке оператор, який діє в лінійному просторі?

2. Який оператор називається лінійним? Наведіть приклад лінійного оператору.

3. Чи буде лінійним перетворенням оператор подібності $\widetilde{A}x = \mu x$?

4. Як визначається сума двох операторів?

5. Який оператор одержується в результаті додавання двох операторів подібності з коефіцієнтами подібності μ_1 і μ_2? Чи буде він лінійним? Доведіть.

6. Як визначається добуток оператора на число?

7. Нехай $\widetilde{A}$ – оператор повороту на кут φ в просторі V_2 векторів на площині. Як діє оператор: а) 2 $\widetilde{A}$; б) (-1) $\widetilde{A}$?

4.3. Матриці лінійних перетворень

Надалі в нашому курсі ми будемо розглядати лише лінійні перетворення. Дозволимо собі називати їх операторами. Нехай $\widetilde{A} \in L(V,V)$, $e: e_1, e_2, ..., e_n$ – фіксований базис простору V. Подіє').

Подіємо оператором $\widetilde{A}$ на базисні елементи і розкладемо образи базисних елементів $\widetilde{A}e_1, \widetilde{A}e_2, \dots, \widetilde{A}e_n$ по цьому ж базису:

$$\begin{aligned} \widetilde{A}e_1 &= a_{11}e_1 + a_{12}e_2 + ... + a_{1n}e_n = \sum_{j=1}^{n} a_{1j}e_j; \\ \widetilde{A}e_2 &= a_{21}e_1 + a_{22}e_2 + ... + a_{2n}e_n = \sum_{j=1}^{n} a_{2j}e_j; \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \widetilde{A}e_n &= a_{n1}e_1 + a_{n2}e_2 + ... + a_{nn}e_n = \sum_{j=1}^{n} a_{nj}e_j. \end{aligned} \tag{4.3.1}$$

Квадратна матриця n-го порядку

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

називається ***матрицею лінійного перетворення*** $\widetilde{A}$ в базисі $e_1, e_2, ..., e_n$.

Зауважимо, що i-й стовпчик матриці A_e складається із коефіцієнтів a_{ij} розкладу елементу $\widetilde{A}e_i$ по базису $e_1, e_2, ..., e_n$. Тобто можна говорити, що ***матрицею лінійного перетворення*** базисі e буде матриця, стовпчики якої є координатами образів базисних елементів в базисі e. Систему рівнянь (4.3.1) можна записати в матричній формі:

$$\widetilde{A}e = eA_e, \tag{4.3.2}$$

де, як звичайно, позначається $e = (e_1, e_2, \dots, e_n)$ – рядок, складений із елементів базису, $\widetilde{A}e = (\widetilde{A}e_1, \widetilde{A}e_2, \dots, \widetilde{A}e_n)$ – рядок із образів базисних елементів, а добуток eA_e одержується за правилом множення матриць: $(1 \times n)$ - матриця e множиться на $(n \times n)$- матрицю A_e.

Зауваження 4.3.1. *Кожному лінійному оператору* $\widetilde{A}$ *в заданому базисі* e *відповідає єдина матриця* A_e (єдиність забезпечується однозначністю образу довільного елементу та однозначністю розкладу по базису).

Часто номери елементів матриць позначають дещо інакше, а саме: номер рядка позначають верхнім, а номер стовпчика – нижнім індексами (a_j^i – елемент i-го рядка і j-го стовпчика) і так само координати елемента позначають верхнім індексом. Такі позначення відіграють важливу роль при описанні тензорних величин і широко використовуються в фізиці. Тоді рівняння (2.3.1) будуть мати вигляд: $\widetilde{A}e_j = \sum_{i=1}^{n} a_j^i e_i$, $j = 1,2,\ldots,n$, розклад елементу по базису – $x = \sum_{i=1}^{n} x^i e_i$, що при таких позначеннях також записується у вигляді $\widetilde{A}e_j = a_j^i e_i$ $j = 1,2,\ldots,n$ та $x = x^i e_i$

Нехай $\widetilde{A}x = y$, $x = x_1e_1 + \ldots + x_ne_n = ex_e^T$, $y = y_1e_1 + \ldots + y_ne_n = ey_e^T$, де x_e^T і y_e^T – стовпчики координат елементів x і y в даному базисі e. Тоді із рівності (4.1.1), (4.3.2) та означення лінійного оператора одержуємо

$$ey_e^T = y \overset{(2.1.1)}{=} \widetilde{A}x = \widetilde{A}(x_1e_1 + x_2e_2 + \ldots + x_ne_n) =$$
$$= x_1\widetilde{A}e_1 + x_2\widetilde{A}e_2 + \ldots + x_n\widetilde{A}e_n = \widetilde{A}ex_e^T \overset{(2.3.2)}{=} eA_ex_e^T,$$

Зліва і справа в цій рівності маємо розклад одного й того ж елементу по базису e. Так як цей розклад однозначний (твердження 1.4.3), то

$$y_e^T = A_ex_e^T. \tag{4.3.3}$$

Формула (4.3.3) дозволяє знайти координати образа y через координати прообразу x в даному базисі, якщо відома матриця A_e оператора $\widetilde{A}$ в цьому базисі.

<u>Приклад 1.</u> Лінійне перетворення $\widetilde{A}$ діє в просторі R^3 за правилом $\widetilde{A}(x_1,x_2,x_3) = (x_1 + x_3, x_3, x_2 - x_1)$. Знайти матрицю цього перетворення в базисі

а) e: $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$;

б) e': $e_1' = (1,1,0)$, $e_2' = (0,1,0)$, $e_3' = (0,0,1)$.

Знайти образ елементу $x = (2,3,5)$ за допомогою матриці перетворення.

▼ а) Знайдемо образи базисних елементів та розкладемо їх по базису e:

$\widetilde{A}e_1 = (1,0,-1) = 1\cdot e_1 + 0\cdot e_2 - 1\cdot e_3$;

$\widetilde{A}e_2 = (1,0,-1) = 0\cdot e_1 + 0\cdot e_2 + 1\cdot e_3$;

$\widetilde{A}e_3 = (1,1,0) = 1\cdot e_1 + 1\cdot e_2 + 0\cdot e_3$.

За означенням матриці лінійного оператору

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

Тоді $A_e \cdot x_e^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \\ 1 \end{pmatrix}$, тобто $\widetilde{A}x = (7,5,1)_e$. ▲

▼б) Знайдемо образи базисних елементів та розкладемо їх по базису e':

$$\begin{matrix} \widetilde{A}e_1' = (1,0,0) = 1\cdot e_1' - 1\cdot e_2' + 0\cdot e_3' \\ \widetilde{A}e_2' = (0,0,1) = 0\cdot e_1' + 0\cdot e_2' + 1\cdot e_3' \\ \widetilde{A}e_3' = (1,1,0) = 1\cdot e_1' + 0\cdot e_2' + 0\cdot e_3' \end{matrix} \Rightarrow A_{e'} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Тоді для знаходження образу елементу x розкладемо його по базису e': $x = 2e_1' + e_2' + 5e_3'$(перевірте), знайдемо $A_e \cdot x_e^T = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 7 \\ -2 \\ 1 \end{pmatrix}$, тобто $\widetilde{A}x = (7,-2,1)_{e'}$. ▲

Зауваження 4.3.2. *В різних базисах матриці лінійного оператору мають різний вигляд.*

Зауваження 4.3.3. *Якщо $\widetilde{A} = \widetilde{O}$, то матриця цього оператору нульова в довільному базисі.*

Зауваження 4.3.4. *Якщо $\widetilde{A} = \widetilde{I}$, то матриця цього оператору одинична в довільному базисі.*

Теорема 4.3.5. *Нехай в лінійному просторі V задано базис $e: e_1, e_2, ..., e_n$ і нехай $A = (a_{ij})$ – квадратна матриця $n \times n$. Існує єдиний лінійний оператор $\widetilde{A} \in L(V,V)$, матриця A_e якого в заданому базисі буде дорівнювати A.*

Д о в е д е н н я.

І. Задамо $\widetilde{A}$ на базисних елементах:

$$\begin{array}{l}\widetilde{A}e_1 = a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n \\ \widetilde{A}e_2 = a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \widetilde{A}e_n = a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n\end{array},$$

тобто $\widetilde{A}e_i = \sum_{j=1}^{n} a_{ji}e_j$, $i = 1,2,\ldots,n$. Враховуючи, що $\forall x \in V \quad \exists!\, x_i, i = 1,\ldots,n,$ що $x = \sum_{1}^{n} x_i e_i = e x_e^T$, задамо образ довільного елементу x наступним чином:

$$\widetilde{A}x = \sum_{i=1}^{n} x_i \widetilde{A}e_i = \sum_i x_i (\sum_j a_{ji}e_j) = \sum_j (\sum_i a_{ji}x_i)e_j = eAx_e^T. \qquad (4.3.4)$$

II. Перевіримо, чи такий оператор $\widetilde{A}$ лінійний. За теоремою 3.4.4 про додавання та множення елементів лінійного простору в базисі маємо, що $\forall u, v \in V, \forall \alpha, \beta \in R \quad (\alpha u + \beta v)_e = \alpha u_e + \beta v_e$. Тоді за умовою (4.3.4) $\widetilde{A}(\alpha u + \beta v) = eA(\alpha u + \beta v)_e^T = eA(\alpha u_e^T + \beta v_e^T)$. Останнє співвідношення – це множення відповідних матриць, тому із властивостей добутку матриць (Частина 1) маємо лінійність утвореного оператора $\widetilde{A}$:

$$\widetilde{A}(\alpha u + \beta v) = \alpha eAu_e^T + \beta eAv_e^T = \alpha\widetilde{A}u + \beta\widetilde{A}v.$$

III. Доведемо, що він єдиний (*від супротивного*):

Нехай $\exists \widetilde{B} \in L(V,V)$, $\widetilde{B} \neq \widetilde{A}$, матрицею якого теж буде матриця A. Тоді $\forall x \in V$ маємо $\widetilde{A}x = eAx_e^T$ і $\widetilde{B}x = eAx_e^T$. Тому $\widetilde{A}x - \widetilde{B}x = eAx_e^T - eAx_e^T = \theta \quad \Rightarrow$

$\Rightarrow \quad (\widetilde{A} - \widetilde{B})x = \theta \quad \Rightarrow (\widetilde{A} - \widetilde{B})x = \theta \Rightarrow \widetilde{A} - \widetilde{B} = \widetilde{O} \Rightarrow \widetilde{A} = \widetilde{B}$ – протиріччя. ■

Зауважимо, що матрицею C_e суми лінійних операторів $\widetilde{A}, \widetilde{B} \in L(V,V)$ буде сума матриць цих операторів:

$(\widetilde{A} + \widetilde{B})x = \widetilde{A}x + \widetilde{B}x = eA_e x_e^T + eB_e x_e^T = e(A_e + B_e)x_e^T \Rightarrow \quad C_e = A_e + B_e$; матрицею оператора $\widetilde{D} = \lambda\widetilde{A}$ буде матриця $D_e = \lambda A_e$ (доведення аналогічне); а матрицею F_e добутку операторів $\widetilde{A}$ і $\widetilde{B}$ буде добуток матриць A_e і B_e:

$(\widetilde{A}\widetilde{B})x = \widetilde{A}(\widetilde{B}x) = \widetilde{A}(eB_e x_e^T) = (\widetilde{A}e)(B_e x_e^T) \overset{(2.3.2)}{=} eA_e B_e x_e^T \overset{(2.3.4)}{\Rightarrow} F_e = A_e B_e$.

Із попередніх зауважень та зауваження 4.3.4 випливає

Зауваження 4.3.6. *Матрицею оператора* $\widetilde{A}-\lambda\widetilde{B}$ *в базисі* e *буде матриця* $A_e-\lambda B_e$. *Зокрема, матрицею оператора* $\widetilde{A}-\lambda\widetilde{I}$ *в базисі* e *буде матриця* $A_e-\lambda E$ *(*$\widetilde{I}$ *– тотожній оператор,* E *– одинична матриця).*

Контрольні питання і завдання

1. Як визначається матриця лінійного оператору в заданому базисі?
2. Який вигляд в довільному базисі має матриця: а) нульового оператору; б) тотожного оператору;

в)оператору подібності з коефіцієнтом подібності μ?

3. Лінійне перетворення $\widetilde{A}$ діє в просторі R^3 за правилом : кожному елементу x з координатами x_1,x_2,x_3 в деякому базисі ставить у відповідність елемент $\widetilde{A}x$ з координатами x_1+x_3,x_3,x_2-x_1 в цьому ж базисі. Знайти матрицю цього перетворення в даному базисі.
4. Чому дорівнює матриця суми лінійних операторів в заданому базисі?
5. Яка розмірність лінійного простору лінійних перетворень, які діють в просторі розмірності n?

4.4. Перетворення матриці лінійного оператору при зміні базису

Нехай $\widetilde{A}\in L(V,V)$, e і e' – два базиси V, U – матриця переходу від e до e'.

Теорема 4.4.1. *Матриці* A_e і $A_{e'}$ *лінійного оператору* $\widetilde{A}$ *в базисах* e *та* e' *відповідно пов'язані співвідношенням* $A_{e'}=U^{-1}A_eU$.

Д о в е д е н н я. Виберемо довільний елемент $x\in V$. Нагадаємо позначення координат цього елементу та його образу в базисах e та e': $x=x_e$, $x=x_{e'}$, $y=\widetilde{A}x$, $y=y_e$, $y=y_{e'}$. Тоді

$$y_e^T=A_ex_e^T, \tag{1}$$

$$y_{e'}^T=A_{e'}x_{e'}^T. \tag{2}$$

Якщо U – матриця переходу від e до e',то

$$x_{e'}^T=U^{-1}x_e^T, \tag{3}$$

$$y_{e'}^T = U^{-1} y_e^T. \quad (4)$$

Підставимо (3) і (4) в (2). Одержимо

$$U^{-1} y_e^T = A_{e'} (U^{-1} x_e^T). \quad (5)$$

Підставимо (1) в (5) і врахуємо асоціативність множення матриць:

$$(U^{-1} A_e) x_e^T = (A_{e'} U^{-1}) x_e^T. \quad (6)$$

Оскільки (6) виконується $\forall x \in V$, то $U^{-1} A_e = A_{e'} U^{-1}$, що еквівалентно співвідношенню $A_{e'} = U^{-1} A_e U$. ■

Якщо A, B, C – квадратні матриці одного порядку і $B = C^{-1} \cdot A \cdot C$, то говорять, що A і B – ***подібні***. З урахуванням цього попередня теорема може бути сформульована наступним чином.

Теорема 4.4.1(а). Матриці лінійного оператору в різних базисах подібні.

Зауваження 4.4.2. Якщо A_e і B_e – матриці лінійних операторів $\widetilde{A}$ і $\widetilde{B}$ в базисі e, $A_{e'}$ і $B_{e'}$ – матриці цих лінійних операторів в базисі e', то *матриці* $A_e + \lambda B_e$ *та* $A_{e'} + \lambda B_{e'}$ *подібні* як матриці лінійного оператору $\widetilde{A} + \lambda \widetilde{B}$:

Нехай U – матриця переходу від базису e до базису e'. Тоді

$U^{-1}(A_e + \lambda B_e) U = U^{-1} A_e U + U^{-1} \lambda B_e U = A_{e'} + \lambda B_{e'}$. ■

Зокрема, якщо $\widetilde{B} = \widetilde{I}$, то, за зауваженням: 2.3.4, $B_e = B_{e'} = E$, а подібними матрицями лінійного оператору $\widetilde{A} - \lambda \widetilde{I}$ в базисах e і e' будуть $A_e - \lambda E$ і $A_{e'} - \lambda E$.

Наслідок 4.4.3. *Визначники матриць лінійного оператору в різних базисах рівні.*

Д о в е д е н н я.

$|A_{e'}| = |U^{-1} A_e U| = |U^{-1}| \cdot |A_e| \cdot |U| = |U^{-1}| \cdot |U| \cdot |A_e| = |U^{-1} \cdot U| \cdot |A_e| = 1 \cdot |A_e| = |A_e|$. ■

Тобто визначник матриці лінійного оператору не залежить від вибору базису, що дає змогу ввести поняття: ***детермінантом оператору*** $\widetilde{A}$ називається визначник матриці цього оператору в деякому базисі ($\det \widetilde{A} = |A_e|$).

Величини, пов'язані з лінійним оператором, які не залежать від вибору базису будемо називати ***інваріантними*** величинами цього оператору. Отже, із зауваження 2.4.2 та наслідку 2.4.3 маємо, що $|A_e - \lambda E| = |A_{e'} - \lambda E|$, а це приводить до

Зауваження 4.4.4. *Інваріантною величиною оператору* $\widetilde{A}$ *є не тільки* $\det\widetilde{A}$*, а й* $\det(\widetilde{A} - \lambda\widetilde{I}) = |A_e - \lambda E|$.

Контрольні питання і завдання

1. Як змінюється матриця лінійного оператору при переході від одного базису до іншого?

2. Чому дорівнює $\det\lambda\widetilde{A}$, якщо оператор $\widetilde{A}$ діє в просторі розмірності n ?

4.5. Оборотні оператори (перетворення)

Серед операторів особливу роль відіграють ті, для яких існують обернені оператори у відомому нам загальному понятті оберненого відображення.

Означення 4.5.1. Відображення $f^{-1}: Y \to X$ називається ***оберненим*** до відображення $f: X \to Y$, якщо $\forall a' \in Y$ $f^{-1}(a') = a \in X$, причому $fa = a'$.

Нагадаємо критерій оборотності відображення: обернене відображення f^{-1} існує тоді і тільки тоді, коли f – взаємно однозначне.

Існує поняття оберненого перетворення (не обов’язково лінійного оператору), яке визначається через поняття добутку перетворень.

Означення 4.5.2. Для перетворення f множини M ***оберненим*** перетворенням називається таке перетворення f^{-1} множини M, що:

$$f \cdot f^{-1} = f^{-1} \cdot f = I,$$

де I – тотожне перетворення множини M.

Перетворення, для якого існує обернене перетворення, називається ***оборотним***. Сукупність всіх оборотних перетворень множини M позначають через G_M.

Ці означення еквівалентні.

Звернемо увагу на найпростіші властивості оборотних перетворень.

1. *Оборотне перетворення має єдине обернене перетворення.*

Якщо φ' і φ'' обернені до f, то $\varphi' = \varphi' I = \varphi'(f\varphi'') = (\varphi' f)\varphi'' = I\varphi'' = \varphi''$. ■

2. *Тотожне перетворення оборотне і* $I^{-1} = I$: $I \cdot I = I$. ■

3. *Якщо* f *оборотне, то* f^{-1} *теж оборотне і* $(f^{-1})^{-1} = f$:

$f\cdot f^{-1}=f^{-1}\cdot f=I\quad\Rightarrow\quad$ – обернене до f^{-1}, тобто $f=(f^{-1})^{-1}$. ■

4. *Якщо два перетворення f і θ оборотні, то $f\cdot\theta$ теж оборотне перетворення, причому*

$$(f\cdot\theta)^{-1}=\theta^{-1}\cdot f^{-1}.$$

$$(f\cdot\theta)(f\cdot\theta)^{-1}=(f\cdot\theta)(\theta^{-1}\cdot f^{-1})=f(\theta\cdot\theta^{-1})f^{-1}=f\cdot I\cdot f^{-1}=f\cdot f^{-1}=I,$$

$$(f\cdot\theta)^{-1}(f\cdot\theta)=(\theta^{-1}\cdot f^{-1})(f\cdot\theta)=f(\theta\cdot\theta^{-1})f^{-1}=\theta^{-1}\cdot I\cdot\theta=\theta^{-1}\cdot\theta=I.\ ■$$

5. *Для довільної множини M сукупність всіх оборотних перетворень G_M утворює групу відносно множення.*

Лема 4.5.3. *Ін'єктивний лінійний оператор простору тільки нульовий елемент відображає в нульовий.*

Д о в е д е н н я. Із ін'єктивності оператора $\tilde{A}$ випливає $\forall\, x_1\neq x_2\Rightarrow\tilde{A}x_1\neq\tilde{A}x_2$, тоді, враховуючи зауваження 2.1.1 , $\forall\, x_1\neq\theta\ \Rightarrow\ \tilde{A}x_1\neq\tilde{A}\theta\ \Rightarrow\ \tilde{A}x_1\neq\theta$.

Повернемося до простору лінійних перетворень $L(V,V)$.

Твердження 4.5.4. *Ін'єктивне лінійне перетворення $\tilde{A}\in L(V,V)$ є сюр'єктивним.*

Д о в е д е н н я. Нехай $\dim V=n$. Покажемо, що n ЛНЗ елементів $x_1,x_2,...,x_n\in V$ ін'єктивним оператором $\tilde{A}$ відображаються в n ЛНЗ елементів

$$\tilde{A}x_1,\tilde{A}x_2,...,\tilde{A}x_n\in V:$$

$$\sum_i\alpha_i\tilde{A}x_i=\theta\overset{\text{озн.Л.О.}}{\Rightarrow}\tilde{A}(\sum_i\alpha_i x_i)=\theta\overset{\text{Лема 2.5.2.}}{\Rightarrow}\sum_i\alpha_i x_i=\theta\Rightarrow\alpha_i=0,i=1,...,n\quad\Rightarrow\quad\tilde{A}x_1,\tilde{A}x_2,...,\tilde{A}x_n-$$
ЛНЗ

Враховуючи теорему про те, що довільні n ЛНЗ елементів простору V при $\dim V=n$ утворюють базис, маємо, що $\tilde{A}x_1,\tilde{A}x_2,...,\tilde{A}x_n$ утворюють базис і довільний елемент $y\in V$ можна представити через їх лінійну комбінацію. Звідси $\forall\ y\in V\ \exists\lambda_i\in R$, $i=1,2,...,n$, що

$$y=\sum\lambda_i(\tilde{A}x_i)\overset{A-Л.О.}{=}\tilde{A}(\sum\lambda_i x_i)=\tilde{A}x,\ \text{де}\ x=\sum\lambda_i x_i\in V.$$

Ця рівність показує метод знаходження прообразу елементу $y\in V$: вибрати базис $\mathrm{I}:x_1,...,x_n\in V$, одержати базис $\mathrm{II}:\tilde{A}x_1,...,\tilde{A}x_n$, знайти координати λ_i елементу y в базисі II та його прообраз $x=\sum\lambda_i x_i$, що забезпечує виконання умови сюр'єктивності ($\forall\,y\in V\quad\exists x\in V$, що $\tilde{A}x=y$).

Враховуючи теорему 4.5.1. та твердження 4.5.3 одержимо

Зауваження 4.5.5. *Ін'єктивні лінійні перетворення є взаємно однозначними.*

Контрольні питання і завдання

1. Доведіть, що для довільної множини M сукупність всіх оборотних перетворень G_M утворює групу відносно множення.

2. Доведіть, якщо $e_1, e_2, ..., e_n \in V$ – базис простору V, $\widetilde{A} \in L(V,V)$ – ін'єктивний, то елементи $\widetilde{A}x_1, \widetilde{A}x_2, ..., \widetilde{A}x_n$ утворюють базис простору V.

4.6. Образ і ядро

Ядром лінійного оператору $\widetilde{A}$ називається множина всіх тих елементів $x \in V$, для яких $\widetilde{A}x = \theta$. Позначається $Ker\widetilde{A}$:

$$Ker\widetilde{A} = \{x \in V \mid \widetilde{A}x = \theta\}.$$

Образом лінійного оператора $\widetilde{A} \in L(V,W)$ називається множина всіх елементів $y \in W$, які можуть бути представлені у вигляді $y = \widetilde{A}x$. Позначається $\operatorname{Im}\widetilde{A}$:

$$\operatorname{Im}\widetilde{A} = \{y \in W \mid \exists x \in V, y = \widetilde{A}x\} \text{ або } \operatorname{Im}\widetilde{A} = \{\widetilde{A}x \mid x \in V\}.$$

Зауважимо, якщо $e_1, e_2, ..., e_n$ – базис простору V, то за зауваженням 3.6.4 $V = \langle e_1, e_2, ..., e_n \rangle$. Так як $\forall y \in \operatorname{Im}\widetilde{A}$ $\exists x \in V$, що $y = \widetilde{A}x$, то $\forall y \in \operatorname{Im}\widetilde{A}$ $\exists \alpha_1, \ldots, \alpha_n \in \mathrm{R}$, що $y = \widetilde{A}x = \widetilde{A}\sum_1^n \alpha_i x_i = \sum_1^n \alpha_i \widetilde{A}x_i$, а значить $\operatorname{Im}\widetilde{A} = \langle \widetilde{A}e_1, \widetilde{A}e_2, ..., \widetilde{A}e_n \rangle$.

Найпростіші властивості ядра та образу.

1. Ядро лінійного оператора $\widetilde{A} \in L(V,V)$ нульове тоді і тільки тоді, якщо $\widetilde{A}$ – взаємно однозначне лінійне перетворення.

Н е о б х і д н і с т ь. (*Від супротивного*) Нехай $Ker\widetilde{A} = \{\theta\}$ і $\widetilde{A}$ *не ін'єктивне. Тоді* $\exists x_1 \neq x_2$, $\widetilde{A}x_1 = \widetilde{A}x_2 \quad \Rightarrow$

$$\begin{gathered} \widetilde{A}x_1 - \widetilde{A}x_2 = \theta \quad \Rightarrow \quad \widetilde{A}(x_1 - x_2) = \theta \quad \Rightarrow \quad x_1 - x_2 \in Ker\,\widetilde{A} = \{\theta\} \Rightarrow \\ \Rightarrow x_1 - x_2 = \theta \Rightarrow \quad x_1 = x_2 \end{gathered}.$$

Одержане протиріччя доводить ін'єктивність лінійного перетворення, а значить, за зауваженням 245.4, і його взаємну однозначність. ▲

Д о с т а т н і с т ь випливає безпосередньо із леми 4.5.2. ■

2. Ядро лінійного оператора $\widetilde{A} \in L(V,W)$ утворює лінійний підпростір в просторі V.

Д о в е д е н н я. Скористаємось критерієм підпростору 1.6.1а.:

$\forall x_1, x_2 \in Ker\,\widetilde{A}, \forall \lambda, \mu \in R,$

$\widetilde{A}(\lambda x_1 + \mu x_2) = \lambda\widetilde{A}x_1 + \mu\widetilde{A}x_2 = \theta + \theta = \theta \quad \Rightarrow \quad \lambda x_1 + \mu x_2 \in Ker\,\widetilde{A}$. ■

3 .Для лінійного перетворення $\widetilde{A} \in L(V,V)$ виконується умова

$$\operatorname{Im}\widetilde{A} = V \Leftrightarrow Ker\widetilde{A} = \theta.$$

.Д о в е д е н н я. Нехай $\operatorname{Im}\widetilde{A} = V$.

1) Нехай $x_1, x_2 ... x_n$ – базис $V = \operatorname{Im}\widetilde{A}$. Тоді $V = \langle x_1, x_2, ..., x_n \rangle$, $\dim V = n$, а за попереднім зауваженням $\operatorname{Im}\widetilde{A} = \left\langle \widetilde{A}x_1, ..., \widetilde{A}x_n \right\rangle$. Так як $\dim\,(\operatorname{Im}\widetilde{A}) = \dim V = n$, то, враховуючи теорему1.6.3, одержуємо, що $\widetilde{A}x_1, ..., \widetilde{A}x_n$ – базис $\operatorname{Im}\widetilde{A}$.

2) Покажемо, що $\widetilde{A}$ – ін'єктивне (*від супротивного*). Нехай $x' \neq x,\ x, x' \in V,$ $x = \sum \alpha_i x_i$, $x' = \sum_1^n \beta_i x_i$ і $\widetilde{A}x' = \widetilde{A}x$. Тоді $\widetilde{A}(x' - x) = \theta \Rightarrow \sum_1^n (\beta_i - \alpha_i)\widetilde{A}x_i = \theta$. Із лінійної незалежності $\widetilde{A}x_1, ..., \widetilde{A}x_n$ випливає, що $\beta_i - \alpha_i = 0 \Rightarrow \alpha_i = \beta_i \Rightarrow x = x'$, а це означає суперечність і доводить ін'єктивність $\widetilde{A}$.

З урахуванням зауваження 4.5.4, це рівнозначно умові, що $\widetilde{A}$ – взаємно однозначне перетворення, а за властивістю 1) рівнозначно тому, що $Ker\widetilde{A} = \{\theta\}$. ■

4. Образ лінійного оператора $\widetilde{A} \in L(V,W)$ утворює лінійний підпростір в просторі W.

Д о в е д е н н я. Так як $\forall y_1, y_2 \in \operatorname{Im}\widetilde{A}\ \exists\, x_1, x_2 \in V$, що $\widetilde{A}x_1 = y_1,\ \widetilde{A}x_2 = y_2$ то скористаємось критерієм підпростору 3.6.1а.:

$\forall y_1, y_2 \in \operatorname{Im}\widetilde{A}, \forall \lambda, \mu \in R$ сума $\lambda y_1 + \mu y_2 = \lambda\widetilde{A}x_1 + \mu\widetilde{A}x_2 = \widetilde{A}(\lambda x_1 + \mu x_2)$. Отже $\lambda y_1 + \mu y_2 \in \operatorname{Im}\widetilde{A}$. ■

Теорема 4.6.1. (про зв'язок розмірності образу і ядра). *Нехай* $\dim V = n$, $\widetilde{A} \in L(V,V)$. *Тоді* $\dim(Ker\widetilde{A}) + \dim(\operatorname{Im}\widetilde{A}) = n$.

Д о в е д е н н я. Нехай $Ker\widetilde{A} = V_1$, $\operatorname{Im}\widetilde{A} = V_2$. Так як V_1 підпростір V, то за теоремою 3.7.5 існує пряме доповнення V' простору V ($V = V_1 \oplus V'$: $V_1 \cap V' = \theta$, $\dim V = \dim V_1 + \dim V'$).

Отже, для доведення теореми достатньо довести, що $\dim V' = \dim V_2$.

Нехай $\dim V' = p$, $\dim V_2 = q$, $y_1, y_2, ..., y_q$ – базис V_2.

Із означення образу лінійного оператору маємо:

$$\forall y \in V_2 \quad \exists x \in V \text{ такий, що } \widetilde{A}x = y.$$

З іншого боку, за означенням прямої суми $V = V_1 \oplus V'$, $\forall x \in V \quad \exists! x' \in V_1, x'' \in V'$ такі, що $x = x' + x''$. Тоді, за означенням лінійного оператору, $y = \widetilde{A}x = \widetilde{A}(x' + x'') = \widetilde{A}x' + \widetilde{A}x''$. За означенням ядра, $\widetilde{A}x' = \theta$, тому $y = \theta + \widetilde{A}x'' = \widetilde{A}x''$. Таким чином ми одержали, що

$$\forall y \in V_2 \quad \exists x'' \in V', \text{ що } \widetilde{A}x'' = y.$$

Отже, для базису $y_1, y_2 ... y_q$ $\exists x_1, x_2, ..., x_q \in V'$, що $\widetilde{A}x_i = y_i$.

Доведемо, що ці $x_1, x_2, ..., x_q$ – ЛНЗ. Розглянемо лінійну комбінацію $\sum\limits_1^q \alpha_i x_i = \theta$. Звідси $\theta = \widetilde{A}\theta = \widetilde{A}(\sum \alpha_i x_i) = \sum \alpha_i \widetilde{A}x_i = \sum \alpha_i y_i$. Так як $y_1, y_2 ... y_q$ – ЛНЗ, то $\alpha_i = 0, i = 1, ..., q$, а це означає, що $x_1, x_2 ..., x_q$ – ЛНЗ. Значить $\dim V' \geq \dim V_2$, тобто $p \geq q$.

Нехай $p > q$. Доповнимо ЛНЗ систему $x_1, ..., x_q$ до базису V' елементами $x_{q+1}, ..., x_p$ так, що $x_1, ..., x_p$ – базис V'. Тоді $\widetilde{A}x_1, ..., \widetilde{A}x_p \in V_2$, $p > \dim V_2 \Rightarrow \widetilde{A}x_1, ..., \widetilde{A}x_p$ – ЛЗ. Отже існують не всі рівні нулю числа λ_i, $i = 1, ..., p$, що $\sum\limits_1^p \lambda_i \widetilde{A}x_i = \theta \quad \Rightarrow \quad \widetilde{A}(\sum\limits_1^p \lambda_i x_i) = \theta$. Остання рівність означає, що, з одного боку, $\sum\limits_1^p \lambda_i x_i \in Ker\widetilde{A} = V_1$, а, з іншого боку, $\sum\limits_1^p \lambda_i x_i \in V'$, бо $x_i \in V'$, отже $\sum\limits_1^p \lambda_i x_i \in V_1 \cap V' = \theta \Rightarrow \sum\limits_1^p \lambda_i x_i = \theta \quad \Rightarrow \quad x_1, ..., x_p$ – ЛЗ і одержуємо протиріччя. Значить умова $p > q$ не може виконуватись, отже $p = q \Rightarrow \dim V' = \dim V_2$, що й треба було довести. ■

Теорема 4.6.2 (обернена теорема про зв'язок розмірностей ядра і образа). *Нехай V_1 і V_2 такі підпростори n- мірного простору V, що*

$\dim V_1 + \dim V_2 = n = \dim V$. *Тоді існує лінійний оператор* $\widetilde{A} \in L(V,V)$ *такий, що* $V_1 = \operatorname{Im}\widetilde{A}$, $V_2 = Ker\widetilde{A}$.

Д о в е д е н н я. Введемо позначення: $\dim V_1 = p$, $\dim V_2 = q$ $(p + q = n)$.

$g_1, g_2, ..., g_p$ – базис V_1,

$e_{p+1}, ..., e_{p+q} = e_n$ – базис V_2,

$e_1, e_2, ..., e_p$ – доповнення базису V_2 до базису V.

Задамо лінійний оператор $\widetilde{A}$ на базисних елементах простору V ($\widetilde{A}e_1 = g_1, ..., \widetilde{A}e_p = g_p, \widetilde{A}e_{p+1} = \theta, ..., \widetilde{A}e_n = \theta$) і $\forall x \in V$ (якщо $x = \sum\limits_1^n x_i e_i$, то $\widetilde{A}x = \sum\limits_1^n x_i \widetilde{A}e_i$). Покажемо, що це шуканий лінійний оператор.

По-перше, $V_1 = \operatorname{Im}\widetilde{A}$: так як $\forall x \in V$ $x = \sum\limits_1^n x_i e_i = \sum\limits_{i=1}^p x_i e_i + \sum\limits_{i=p+1}^{n=p+q} x_i e_i$, то

а) $[\forall x \in V \quad \widetilde{A}x = \sum\limits_1^n x_i \widetilde{A}e_i = x_1 g_1 + ... + x_p g_p + \theta + ... + \theta \in V_1] \Rightarrow \operatorname{Im}\widetilde{A} \subseteq V_1$;

б) $[\forall y \in V_1 \Rightarrow y = \sum\limits_1^p y_i g_i = \sum\limits_1^p y_i \widetilde{A}e_i = \widetilde{A}(\sum\limits_1^p y_i e_i) \in \operatorname{Im}\widetilde{A}] \Rightarrow V_1 \subseteq \operatorname{Im}\widetilde{A}$.

Тоді із а), б) випливає $V_1 = \operatorname{Im}\widetilde{A}$.

По-друге, $V_2 = Ker\widetilde{A}$:

а) $\forall x \in V_2 \Rightarrow x = \sum\limits_{i=1}^q x_i e_{p+i}$, значить

$$[\forall x \in V_2 \quad \widetilde{A}x = \sum_{i=1}^q x_i \widetilde{A}e_{p+i} = \sum_{i=1}^q x_i \cdot \theta = \theta \Rightarrow x \in Ker\widetilde{A}] \Rightarrow V_2 \subseteq Ker\widetilde{A},$$

б) $[\forall x \in Ker\widetilde{A} \Rightarrow \widetilde{A}x = \theta \Rightarrow \widetilde{A}(\sum\limits_1^n x_i e_i) = \theta \Rightarrow \sum\limits_1^n x_i \widetilde{A}e_i = \theta \Rightarrow$

$\Rightarrow \sum\limits_1^p x_i \widetilde{A}e_i + \sum\limits_{p+1}^n x_i \cdot \theta = \sum\limits_1^p x_i \widetilde{A}e_i = \theta \Rightarrow$

$\Rightarrow x = \sum\limits_{p+1}^n x_i e_i \in V_2] \Rightarrow Ker\widetilde{A} \subseteq V_2$

$\Rightarrow \sum\limits_1^p x_i g_i = \theta \Rightarrow x_i = 0,\ i = 1, ..., p] \Rightarrow [\forall x \in Ker\widetilde{A} \Rightarrow$. Із а, б) випливає

$$V_2 = Ker\widetilde{A}.$$

По-третє, даний оператор є лінійним. ■

Дефектом лінійного оператору $\widetilde{A} \in L(V,V)$ називається $\dim(Ker\widetilde{A})$.

$$def\widetilde{A} = \dim(Ker\widetilde{A})$$

Рангом $\widetilde{A} \in L(V,V)$ називається $\dim(\mathrm{Im}\widetilde{A})$.

$$rang\widetilde{A} = \dim(\mathrm{Im}\widetilde{A})$$

Зауваження 4.6.3. $def\widetilde{A} + rang\widetilde{A} = n = \dim V$. (Теорема 4.6.1 в нових позначеннях).

Теорема 4.6.4. (про ранг лінійного оператора і його матриці). *Якщо* $\widetilde{A} \in L(V,V)$*, то* $rang\widetilde{A} = rgA_e$.

Д о в е д е н н я. Так як $\mathrm{Im}\widetilde{A} = \langle \widetilde{A}e_1, \widetilde{A}e_2, ..., \widetilde{A}e_n \rangle$, то за теоремою 3.6.3 $\dim(\mathrm{Im}\widetilde{A})$, а значить і $rang\,\widetilde{A}$ дорівнює максимальній кількості ЛНЗ в множині $\{\widetilde{A}e_i, i = 1,...,n\}$. Також за означенням rgA_e дорівнює максимальній кількості ЛНЗ стовпчиків матриці A_e, які є координатами елементів $\widetilde{A}e_i$ $, i = 1,...,n$, тому $rang\,\widetilde{A} = rg\,A_e$. ■

Підмножина M' множини M називається ***інваріантною*** відносно перетворення $f : M \to M$, якщо $\forall a \in M' : fa \in M'$.

Можемо навести приклади вже відомих нам підпросторів, які будуть інваріантними відносно оператора $\widetilde{A}$:

1. $Ker\widetilde{A} : \forall x \in Ker\widetilde{A} \quad \widetilde{A}x = \theta \in Ker\widetilde{A}$;
2. $\mathrm{Im}\widetilde{A} : \forall y \in \mathrm{Im}\widetilde{A} \Rightarrow y \in V \Rightarrow \widetilde{A}y \in \mathrm{Im}\widetilde{A}$.

Особливу роль в теорії лінійних перетворень відіграють інваріантні підпростори розмірності 1. З ними тісно пов'язане поняття власного значення та власного вектора.

Теорема 4.6.5. (про ранг добутку). Мають місце наступні співвідношення: $\forall \mathrm{A}, \mathrm{B} \in L(V,V)$:

$$rang\mathrm{AB} \le rang\mathrm{A}$$

$$rang\mathrm{AB} \le rang\mathrm{B}$$

Доведення. 1) $\mathrm{Im}\,\mathrm{AB} = \{(\mathrm{AB})x \mid x \in V\}$

$$\forall y = (\mathrm{AB})x \in \mathrm{Im}\,\mathrm{AB} \Rightarrow y = (\mathrm{AB})x = \frac{\mathrm{A}(\mathrm{B}x)}{\mathrm{B}x \in V} \Rightarrow \in \mathrm{Im}\,\mathrm{A} \Rightarrow \mathrm{Im}\,\mathrm{AB} \subseteq \mathrm{Im}\,\mathrm{A} \Rightarrow rang\mathrm{AB} \le rang\mathrm{A}.$$

2) ($\mathrm{AB} \ne \mathrm{BA}$, тому доводимо не аналогічно) $Ker\mathrm{B} = \{x \in V \mid \mathrm{B}x = 0\}$

$$\{[\forall x \in Ker\mathrm{B}\ \ (\mathrm{AB})x = \mathrm{A}(\mathrm{B}x) = \mathrm{A}0 = 0] \Rightarrow x \in Ker\mathrm{AB}\} \Rightarrow Ker\mathrm{B} \subseteq Ker\mathrm{AB} \Rightarrow def\mathrm{B} \le def\mathrm{AB}\ (*)$$

Контрольні питання і завдання

1. Який зв'язок між дефектом та рангом лінійного оператору.

2. Чи є ранг матриці лінійного оператору інваріантною величиною цього оператору?

3. Доведіть, що лінійний оператор $\mathrm{A} \in L(V,V)$ оборотний тоді і тільки тоді, коли $rang\mathrm{A} = n = \dim V$.

4.7. Власні значення та власні вектори

Число λ називається ***власним значенням*** оператору $\widetilde{A} \in L(V,V)$, якщо існує ненульовий елемент $x \in V$ такий, що

$$\widetilde{A}x = \lambda x. \qquad (4.7.1).$$

При цьому вектор x називають ***власним вектором*** оператору $\widetilde{A}$, що відповідає власному значенню λ.

Звернемо увагу, що умова (4.7.1) при $\widetilde{I} \in L(V,V)$ еквівалентна умовам $\widetilde{A}x = (\lambda\widetilde{I})x$ і $(\widetilde{A} - \lambda\widetilde{I})x = \theta$.

Тобто число λ буде ***власним значенням*** лінійного оператору $\widetilde{A}$, якщо $Ker(\widetilde{A} - \lambda\widetilde{I}) \neq \{\theta\}$.

Нехай $V_{\widetilde{A}}^{\lambda}$ – множина власних векторів оператору $\widetilde{A}$, що відповідає власному значенню λ:

$V_{\widetilde{A}}^{\lambda} = \left\{x \in V \middle| \widetilde{A}x = \lambda x\right\}$.

1. $\forall x_1, x_2 \in V_{\widetilde{A}}^{\lambda}, \forall \alpha, \beta \in R \quad \Rightarrow$

$\widetilde{A}(\alpha x_1 + \beta x_2) = \alpha\lambda x_1 + \beta\lambda x_2 = \lambda(\alpha x_1 + \beta x_2) \quad \Rightarrow \quad \alpha x_1 + \beta x_2 \in V_{\widetilde{A}}^{\lambda}$. За критерієм 3.7.1а одержуємо, що $V_{\widetilde{A}}^{\lambda}$ – підпростір.

2. Нехай $x \in V_{\widetilde{A}}^{\lambda}$. Тоді $\widetilde{A}x = \lambda x \in V_{\widetilde{A}}^{\lambda} \quad \Rightarrow \quad \widetilde{A}\,V_{\widetilde{A}}^{\lambda} \subseteq V_{\widetilde{A}}^{\lambda} \quad \Rightarrow \quad V_{\widetilde{A}}^{\lambda}$ – інваріантна підмножина відносно оператора $\widetilde{A}$.

Одержали, що $V_{\widetilde{A}}^{\lambda}$ – лінійний підпростір простору V, інваріантний відносно оператора $\widetilde{A}$.

Многочлен $\det(\widetilde{A} - \lambda\widetilde{I})$ відносно λ називається ***характеристичним многочленом*** оператору $\widetilde{A}$.

Якщо $A=(a_{ij})$ – матриця лінійного оператору $\widetilde{A}$ в деякому базисі, то за зауваженням 2.3.6. оператору $\widetilde{A}-\lambda\widetilde{I}$ відповідає матриця $A-\lambda E$ в цьому ж базисі і

$$\det(\widetilde{A}-\lambda\widetilde{I})=\begin{vmatrix} a_{11}-\lambda & \cdots & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & & & \vdots \\ \vdots & & a_{kk}-\lambda & & \vdots \\ \vdots & & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & \cdots & a_{nn}-\lambda \end{vmatrix}=d_0+d_1\lambda+\ldots+d_n\lambda^n$$

Зауваження 4.7.1. *Коефіцієнти характеристичного многочлену оператору* $\widetilde{A}$ $d_0,d_1,\ldots,d_n$ *не залежать від вибору базису, тобто є інваріантами оператору* $\widetilde{A}$, так як за зауваженням 2.4.4 $\det(\widetilde{A}-\lambda\widetilde{I})=|\ A-\lambda E\ |$ є інваріантною величиною оператора $\widetilde{A}$ для довільного числа λ.

Рівняння $\det(\widetilde{A}-\lambda\widetilde{I})=0$ називається ***характеристичним рівнянням*** оператору $\widetilde{A}$.

Звернемо увагу, що *корені характеристичного рівняння теж є інваріантами оператора* $\widetilde{A}$.

Теорема 4.7.2. (критерій власного значення) *Для того, щоб число* λ *було власним значенням оператору* $\widetilde{A}$, *необхідно і достатньо, щоб це число було коренем характеристичного рівняння оператору* $\widetilde{A}$.

Д о в е д е н н я. Нехай λ – власне значення оператору $\widetilde{A}\in L(V,V)$, $\dim V=n$. За означенням, $\exists x\neq\theta$ такий, що $Ker(\widetilde{A}-\lambda\widetilde{I})\neq\{\theta\}$ $\Leftrightarrow$ $def(\widetilde{A}-\lambda\widetilde{I})\geq 1$. Із зауваження 4.6.3 це рівнозначно тому, що $rang(\widetilde{A}-\lambda\widetilde{I})\leq n-1 \overset{(2.6.4)}{\Leftrightarrow} rg(A-\lambda E)\leq n-1<n$. За наслідком 3.10.9 (частина 1) це рівнозначно умові $|A-\lambda E|=0$, що означає: λ – корінь характеристичного рівняння оператору $\widetilde{A}$.

Оскільки в доведенні використовувались рівнозначні наслідки, то доведено необхідну і достатню умови. ■

Приклад 4.7.1. В деякому базисі e дана матриця $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 5 & 3 & -2 \end{pmatrix}$ лінійного оператору дійсного лінійного простору. Знайти власні вектори цього оператору.

Розв'язання. Характеристичне рівняння даного оператору

$$|A - \lambda E| = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 5 & 3 & -2-\lambda \end{vmatrix} = 0 \text{ або } (1-\lambda)^2(-2-\lambda) = 0.$$

Корені цього рівняння $\lambda_1 = -2$, $\lambda_2 = \lambda_3 = 1$.

Так як всі характеристичні корені дійсні, то власними значеннями даного лінійного оператору є: $\lambda_1 = -2$, $\lambda_2 = \lambda_3 = 1$.

Для знаходження власних векторів X лінійного оператору $\mathbf{A}$, що відповідають власному значенню λ, скористаємося матричним рівнянням $(A - \lambda E)X = \Theta$, де X – стовпчик невідомих $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Дане рівняння відповідає однорідній системі лінійних рівнянь з невідомими x_1, x_2, x_3. Для розв'язування цієї системи будемо приводити матрицю $(A - \lambda E)$ до східчастої форми та знаходити ФСР системи.

Знайдемо власні вектори X, що відповідають власному значенню $\lambda_1 = -2$: $(A - \lambda E) = \begin{pmatrix} 1+2 & 1 & 0 \\ 0 & 1+2 & 0 \\ 5 & 3 & -2+2 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 5 & 3 & 0 \end{pmatrix}$.

Матричне рівняння $(A - \lambda E)X = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 5 & 3 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$, що відповідає системі $\begin{cases} 3x_1 + x_2 = 0 \\ 3x_2 = 0 \\ 5x_1 + 3x_2 = 0 \end{cases}$.

Розв'язуючи цю систему, маємо: $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 5 & 3 & 0 \end{pmatrix} \approx \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, x_3 – незалежна змінна,

x_1	x_2	x_3
0	0	1

Загальний розв'язок $F = \{(0,0,t), t \in \mathbf{R}\}$.

Власними векторами даного оператору, що відповідають власному значенню –2 є $(0,0,t)_e$, де $t \in \mathbf{R}, t \neq 0$.

Аналогічно знайдемо власні вектори, що відповідають власному значенню $\lambda = 1$.

Маємо: $F = \{(3s,0,5s), s \in \mathbf{R}\}$. Значить, власними векторами даного оператору, що відповідають власному числу 1 є $(3s,0,5s)_e$, де $s \in \mathbf{R}, s \neq 0$ ..▲

Зауважимо, що існують наступні *інваріантні величини лінійного оператору* $\widetilde{A}$: $rang\widetilde{A}$, $def\widetilde{A}$, $\det \widetilde{A}$, $\det(\widetilde{A} - \lambda\widetilde{I})$ з його коефіцієнтами та коренями, власні значення, власні вектори.

Виникає потреба виділити ті інваріанти, які давали б змогу по зовнішньому вигляду матриць лінійних операторів відрізнити матриці різних операторів в різних базисах і з'ясувати в яких саме базисах. Таким зручним виглядом матриці лінійного оператору є ***діагональна матриця***, у якої всі елементи, що не стоять на головній діагоналі є нульовими. Базис, в якому матриця лінійного оператору має діагональний вид, називається ***канонічним базисом.***

Теорема 4.7.3. (критерій діагоналізовності матриці лінійного оператору) *Для того, щоб матриця A лінійного оператору $\widetilde{A}$ в даному базисі $e_1, e_2, ..., e_n$ була діагоналізовною, необхідно і достатньо, щоб базисні вектори $e_1, e_2, ..., e_n$ були власними векторами цього оператору.*

Н е о б х і д н і с т ь. Нехай $A = \begin{pmatrix} d_1 & 0 & ... & 0 \\ 0 & d_2 & ... & 0 \\ ... & ... & \ddots & ... \\ 0 & 0 & ... & d_n \end{pmatrix}$ – матриця лінійного оператору $\widetilde{A}$ в базисі $e_1, e_2, ..., e_n$. Тоді за означенням матриці лінійного оператору $\widetilde{A}e_1 = d_1e_1 + 0e_2 + \ldots + 0e_n = d_1e_1 \quad \Rightarrow \quad e_1$ – власний вектор, що відповідає власному значенню d_1. Аналогічно, $\widetilde{A}e_k = d_ke_k \quad \Rightarrow \quad e_k$ – власний вектор, що відповідає власному значенню d_k, $k = 1,2,\ldots,n$. ▲

Д о с т а т н і с т ь. Нехай $e_1, e_2, ..., e_n$ – базис, у якому кожний елемент e_k – власний вектор, що відповідає деякому власному значенню λ_k. Тоді

$$\left.\begin{array}{l} \mathrm{A}e_1 = \lambda_1 e_1 = \lambda_1 e_1 + 0\cdot e_2 + ... + 0\cdot e_n \\ \mathrm{A}e_2 = \lambda_2 e_2 = 0\cdot e_1 + \lambda_2 e_2 + ... + 0\cdot e_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathrm{A}e_n = \lambda_n e_n = 0\cdot e_1 + 0\cdot e_2 + ... + \lambda_n \cdot e_n \end{array}\right\} \Rightarrow A_e = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} -$$

діагональна. ■

Теорема 4.7.4. *Нехай власні значення $\lambda_1, \lambda_2, ..., \lambda_p$ – різні. Тоді відповідні їм власні вектори $e_1, e_2, ..., e_p$ – ЛНЗ.*

Д о в е д е н н я. (*методом математичної індукції*).

I. $p = 1 \Rightarrow \exists e_1$ – власний вектор. Так як $e_1 \neq \theta$, то e_1 – ЛНЗ.

II. Нехай теорема доведена для $p = m$ ($\lambda_1, \lambda_2, ..., \lambda_m$ – різні $\Rightarrow$ $e_1, e_2, ..., e_m$ – ЛНЗ). Доведемо теорему для $p = m+1$. Нехай власні значення $\lambda_1, \lambda_2, ..., \lambda_{m+1}$ – різні і власному значенню λ_{m+1} відповідає власний вектор e_{m+1}. Треба довести, що $e_1, e_2, ..., e_m, e_{m+1}$ – ЛНЗ.

Розглянемо рівність $\sum\limits_{i=1}^{m+1} \alpha_i e_i = \theta$. $\qquad$ (1)

На обидві частини цієї рівності подіємо оператором $\widetilde{A}$ та множенням на число λ_{m+1}. Одержимо відповідно

$$\widetilde{A}\sum_{i=1}^{m+1} \alpha_i e_i = \widetilde{A}\theta = \theta \text{ та } \lambda_{m+1}\sum_{i=1}^{m+1} \alpha_i e_i = \lambda_{m+1}\theta = \theta,$$

$$\text{або } \sum_{i=1}^{m+1} \alpha_i \widetilde{A} e_i = \sum_{i=1}^{m+1} \alpha_i \lambda_i e_i = \theta \qquad (2) \text{ та } \sum_{i=1}^{m+1} \alpha_i \lambda_{m+1} e_i = \theta \qquad (3)$$

Останні доданки лівих частин (2) і (3) співпадають, тому різниця рівностей (3) і (2) має вигляд $\sum\limits_{i=1}^{m} \alpha_i (\lambda_{m+1} - \lambda_i) e_i = \theta$.

Так як $e_1, e_2, ..., e_m$ – ЛНЗ, то $\alpha_i(\lambda_{m+1} - \lambda_i) = 0, i = 1,...,m$. Ці умови виконуються, коли один із співмножників нульовий. А так як . $\lambda_{m+1} \neq \lambda_i \quad \Rightarrow \quad \lambda_{m+1} - \lambda_i \neq 0$, то $\alpha_i = 0, i = 1,...,m$. Підставимо цей результат в (1), одержимо $\alpha_{m+1}e_{m+1} = \theta (e_{m+1} \neq \theta) \Rightarrow \ \Rightarrow \alpha_{m+1} = 0 \Rightarrow \alpha_i = 0, i = 1,...,m+1 \Rightarrow e_1,...,e_{m+1}$ – ЛНЗ. ■

Наслідок 4.7.5. *Якщо характеристичний многочлен оператору $\tilde{A} \in L(V,V)$ має n різних коренів ($n = \dim V$), то в деякому базисі матриця A лінійного оператору $\tilde{A}$ має діагональний вид.*

Д о в е д е н н я. Якщо характеристичний многочлен оператору $\tilde{A} \in L(V,V)$ має n різних коренів, то за теоремою 4.7.2 він має n різних власних значень, а за теоремою 4.7.4 відповідні власні вектори $e_1, e_2, ..., e_n$ – ЛНЗ. Так як $n = \dim V$, то за теоремою 3.5.2 базис V може складатися з власних векторів $e_1, e_2, ..., e_n$, а за теоремою 3.7.3 в цьому базисі A – діагональна матриця. ■

Зауваження. Як бачимо, хоча б з прикладу 4.7.1, що власні значення і власні вектори оператору знаходяться за допомогою матриці цього оператору в довільному базисі. З цієї причини про власні значення і власні вектори матриці говорять так само, як і про власні значення і власні вектори оператору: ***власними значеннями матриці*** A називаються корені характеристичного рівняння $|A - \lambda E| = 0$, ***власними векторами матриці*** A, що відповідають власному значенню λ, називаються ненульові розв'язки системи $(A - \lambda E)X = O$.

Приклад 4.7.2. Знайдіть всі власні числа та власні вектори матриці $A \in M_8$, де $A = (a_{ij})$, для якої $a_{ij} = 0$, якщо $i + j$ непарне, $a_{ij} = 1$, якщо $i + j$ парне.

Розв'язання. Шукати характеристичний многочлен у цьому випадку не дуже вже й хочеться. По-перше, це неабияк важко — обчислювати визначник восьмого порядку. По-друге, навіть обчисливши його, ми не маємо впевненості, що зможемо знайти власні числа, бо формул для обчислення коренів рівняння восьмого порядку не існує. Отже, розв'язувати задачу треба якось інакше. Найкраще чи не спробувати просто вгадати власні вектори.

Подивимось спочатку на ранг матриці. Окремо парні і окремо непарні рядки нашої матриці однакові, а перший та другий рядки — лінійно незалежні. Звідси легко отримуємо, що ранг A дорівнює 2, тобто ядро відповідного лінійного оператора має розмірність 8 – 2 = 6. Тобто, ми можемо одразу знайти цілих 6 лінійно незалежних власних векторів із власним значенням 0. Знайдемо їх. Для цього нам треба знайти Ф.С.Р. однорідної С.Л.Р. з матрицею

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

Виберемо за вільні змінні $x_3, x_4, x_5, x_6, x_7, x_8$ та виразимо решту через них:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
α_1	-1	0	1	0	0	0	0	0
α_2	0	-1	0	1	0	0	0	0
α_3	-1	0	0	0	1	0	0	0
α_4	0	-1	0	0	0	1	0	0
α_5	-1	0	0	0	0	0	1	0
α_6	0	-1	0	0	0	0	0	1

Звідси знаходимо такий базис V_A^0:

$\alpha_1 = (-1,0,1,0,0,0,0,0)$; $\alpha_2 = (0,-1,0,1,0,0,0,0)$; $\alpha_3 = (-1,0,0,0,1,0,0,0)$;

$\alpha_4 = (0, -1,0,0,0,1,0,0)$; $\alpha_5 = (-1,0,0,0,0,0,1,0)$; $\alpha_6 = (0, -1,0,0,0,0,0,1)$.

Ми вже знайшли шість власних векторів (лінійно незалежних). Якщо нам пощастить, і ми зможемо вгадати ще два, ми стверджуватимемо, що більше їх бути і не може. Але ті два вектори, що залишились, вгадати дуже легко: це можуть будуть вектори

$\alpha_7 = (1,1,1,1,1,1,1,1)$ та $\alpha_8 = (1,0,1,0,1,0,1, 0)$,

бо $A\alpha_7^T = 4\alpha_7^T$ і $A\alpha_8^T = 4\alpha_8^T$, тобто вони є власними векторами матриці A з власним числом 4. Їх лінійна незалежність очевидна, а лінійна незалежність з попередніми випливає з теореми 2.7.4.. Останнє і завершує розв'язання задачі: $V_A^0 = \langle \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \rangle$, $V_A^4 = \langle \alpha_7, \alpha_8 \rangle$. ▲

Контрольні питання і завдання

1. Який елемент лінійного простору називається власним вектором лінійного оператору?

2. Що таке власне значення лінійного оператору?

3. Доведіть істинність рівності $\widetilde{A}\theta = 5\theta$, де $\widetilde{A}$ – довільний лінійний оператор, θ – нульовий елемент. Чи випливає з цього рівняння, що θ– власний вектор оператора, а число 5 – власне значення цього оператора?

4. Чи має власні значення і власні вектори оператор подібності з коефіцієнтом подібності μ?

5. Чи має оператор повороту на кут φ в просторі V_2 векторів на площині власні значення і власні вектори, якщо: а) $\varphi = \pi/4$; б) $\varphi = \pi$?

6. Яке найбільше число різних власних значень лінійного оператору, який діє в лінійному просторі розмірності n ?

7. Чи довільний лінійний оператор, який діє у дійсному лінійному просторі, має власні значення?

8. Яке рівняння називається характеристичним рівнянням лінійного оператору?

9. Чи завжди корінь характеристичного рівняння є власним значенням лінійного оператору?

Розділ 5. БІЛІНІЙНІ ТА КВАДРАТИЧНІ ФОРМИ

5.1. Білінійні форми

Числова функція $A(x, y)$, аргументами якої є довільні вектори x і y дійсного лінійного простору L, називається ***білінійною формою***, якщо $\forall x, y, z \in L, \forall \lambda \in R$ виконуються рівності:

$$1.\ A(x+z, y) = A(x, y) + A(z, y)$$

$$2.\ A(x, y+z) = A(x, y) + A(x, y)$$

$$3.\ A(\lambda x, y) = \lambda A(x, y)$$

$$4.\ A(x.\lambda y) = \lambda A(x, y).$$

(Тобто числова функція $A(x, y)$ лінійна по кожному аргументу).

Нехай в n- мірному лінійному просторі L задано білінійну форму $A(x, y)$ та базис $e: e_1, e_2, ..., e_n$.

Теорема 5.1.1. *Білінійна форма $A(x, y)$ в лінійному просторі L з базисом e: $e_1, e_2, \ldots, e_n$ може бути однозначно представлена у вигляді:*

$$A(x, y) = \sum_{i,j=1}^{n} a_{ij} \xi_i \eta_j \tag{5.1.1}$$

де $a_{ij} = A(e_i, e_j)$, а ξ_i, η_j – координати елементів x і y в базисі e.

Д о в е д е н н я**.** Так як $x = \sum_{i=1}^{n} \xi_i e_i$, $y = \sum_{j=1}^{n} \eta_j e_j$, а $A(x, y)$ – білінійна форма, то із означення білінійної форми

$$A(x, y) = A(\sum_i \xi_i e_i, \sum_j \eta_j e_j) = \sum_{i,j=1}^{n} A(e_i, e_j) \xi_i \eta_j = \sum_{i,j=1}^{n} a_{ij} \xi_i \eta_j.$$

Доведемо однозначність цього представлення від супротивного. Нехай існує інше представлення $A(x, y) = \sum_{i,j=1}^{n} a'_{ij} \xi_i \eta_j$ цієї білінійної форми. Так як координати кожного елементу e_k в базисі e мають вигляд $e_k = (0,0,...,1,...,0)_e$, де 1 стоїть на k- му місці, то із цього представлення $A(e_k, e_t)$ $= 0 + \ldots + 0 + a'_{kt} \cdot 1 \cdot 1 + 0 + \ldots + 0 = a'_{kt}$. Оскільки $A(x, y)$ – функція, то існує єдине значення $A(e_k, e_t)$ $\forall k, t = 1,2, \ldots, n$, значить $a_{kt} = a'_{kt}$, $\forall k, t = 1,2, \ldots, n$, тобто $A(x, y)$ однозначно представляється у вигляді (5.1.1). ■

Рівність (5.1.1) називають ***загальним видом білінійної форми*** в n- мірному лінійному просторі.

Наведемо приклад загального виду білінійної форми при $L = R^3$:

$$A(x,y) = 2\xi_1\eta_1 + 3\xi_1\eta_2 + 0\xi_1\eta_3 - \xi_2\eta_1 - \xi_2\eta_2 + \frac{1}{2}\xi_2\eta_3 + 10\xi_3\eta_1 - 12\xi_3\eta_2 + 7\xi_3\eta_3.$$

Матриця

$$A_e = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

де $a_{ij} = A(e_i, e_j)$, називається ***матрицею білінійної форми*** $A(x,y)$ в базисі e. Очевидно, що вона визначається однозначно. Матрицею білінійної форми попереднього прикладу буде матриця

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & -1 & \frac{1}{2} \\ 10 & -12 & 7 \end{pmatrix}.$$

Зауваження 5.1.2. *В заданому базисі* e *простору* L *довільна квадратна матриця* (a_{ij}) *є матрицею деякої білінійної форми* $A(x,y)$*:*

$$A(x,y) = \sum_{i,j=1}^{n} a_{ij}\xi_i\eta_j$$

Зауваження 5.1.3. *Дія білінійної форми* $A(x,y)$ *з матрицею* A *в базисі* e *на елементи* x *і* y *дорівнює добутку* $x_e A y_e^T$, *де* x_e *– вектор-рядок координат елементу* x *в базисі* e, y^T *– вектор-стовпчик координат елементу* y *в базисі* e.

Д о в е д е н н я. Нехай $x = (\xi_1, \xi_2, ..., \xi_n)_e, y = (\eta_1, \eta_2, ..., \eta_n)_e$. Тоді

$$x_e A y_e^T = (\xi_1, ..., \xi_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} y_e^T = \left(\sum_{i=1}^{n}\xi_i a_{i1}, ..., \sum \xi_i a_{in}\right)\begin{pmatrix} \eta_1 \\ \cdots \\ \eta_n \end{pmatrix} =$$

$$= \sum_{j=1}^{n}\eta_j \sum_{i=1}^{n}\xi_i a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}\xi_i\eta_j = A(x,y). \blacksquare$$

Теорема 5.1.4. *Матриці* A_e *і* A_f *білінійної форми* $A(x,y)$ *в базисах* $e: e_1, e_2, ..., e_n$ *і* $f: f_1, f_2, ..., f_n$ *пов`язані співвідношенням*

$$A_f = U^T A_e U, \tag{5.1.2}$$

де U – матриця переходу від базису e *до базису* f, *а* U^T *– транспонована матриця до U.*

Д о в е д е н н я. Нагадаємо властивості транспонованих матриць:

$$(AB)^T = B^T A^T, \ \left(A^T\right)^T = A, \ \left(A^{-1}\right)^T = \left(A^T\right)^{-1}.$$

Як і раніше, в розділі 1, через x_e, x_f, y_e, y_f будемо позначати вектори-рядки координат елементів x і y в базисах e і f відповідно. Зв'язок координат елементів в різних базисах виражається рівняннями (4.6.5)

$$x_f = x_e\left(U^{-1}\right)^T, \ y_f^T = U^{-1}y_e^T.$$

Із зауваження 5.1.2 $\forall x, y \in L \quad A(x,y) = x_e A_e y_e^T$. Крім того, враховуючи асоціативність множення матриць, одержуємо

$$A(x,y) = x_f A_f y_f^T = \left(x_e\left(U^{-1}\right)^T\right)A_f\left(U^{-1}y_e^T\right) = x_e\left(\left(U^{-1}\right)^T A_f U^{-1}\right)y_e^T.$$

Так як матриця білінійної форми в заданому базисі визначається однозначно, то $A_e = \left(U^{-1}\right)^T A_f U^{-1} \Rightarrow \quad A_e U = \left(U^{-1}\right)^T A_f = \left(U^T\right)^{-1} A_f \quad \Rightarrow$

$U^T A_e U = A_f$. ■

Наслідок 5.1.5. *Ранг матриці* A_f *дорівнює рангу матриці* A_e.

Це зразу випливає із співвідношення (5.1.4), бо U і U^T невироджені матриці, а множення на невироджену матрицю не змінює рангу.

Цей наслідок дозволяє ввести важливий числовий інваріант білінійної форми – ранг білінійної форми.

Рангом білінійної форми, заданої в скінченномірному лінійному просторі L, називається ранг матриці цієї форми в довільному базисі простору L.

Білінійна форма, задана в скінченномірному просторі L, називається ***невиродженою (виродженою),*** якщо її ранг дорівнює (менше) розмірності простору L.

Контрольні питання і завдання

1. Сформулюйте означення білінійної форми.
2. Напишіть загальний вид білінійної форми в 3-мірному просторі.
3. Що таке матриця білінійної форми в заданому базисі?
4. Дано білінійну форму в загальному вигляді в деякому базисі e:

$$B(x,y) = 3x_1y_1 - x_1y_2 + 2x_2y_1 + 4x_1y_3 - 2x_3y_1 + x_2y_3 + x_3y_2 - x_2y_2 + 8x_3y_3$$

Знайдіть матрицю цієї білінійної форми в цьому базисі та її ранг.

5.2. Симетричні білінійні та квадратичні форми

Білінійна форма $A(x,y)$ називається ***симетричною***, якщо $\forall x,y\in L$, виконується співвідношення

$$A(x,y)=A(y,x).$$

Зауваження 5.2.1 *Білінійна форма* $A(x,y)$ *симетрична тоді і тільки тоді, коли її матриця A в довільному базисі* e *симетрична.*

Д о в е д е н н я. Нехай $x=(x_{1,}x_2,...,x_n)_e, y=(y_1,y_2,...,y_n)_e$. Тоді $A(x,y)=\sum\limits_{i,j=1}^{n}A(e_i,e_j)x_iy_j$, $A(y,x)=\sum\limits_{i,j=1}^{n}A(e_j,e_i)y_jx_i$, тобто умова симетричності білінійної форми $A(x,y)$ виконується тоді і тільки тоді, коли $A(e_i,e_j)=A(e_j,e_i)\Leftrightarrow a_{ij}=a_{ji}$, тобто матриця A_e симетрична. ■

Таким чином, за даним критерієм, білінійна форма

$$A(x,y)=2\xi_1\eta_1+3\xi_1\eta_2+0\xi_1\eta_3-\xi_2\eta_1-\xi_2\eta_2+\frac{1}{2}\xi_2\eta_3+10\xi_3\eta_1-12\xi_3\eta_2+7\xi_3\eta_3$$

не є симетричною білінійною формою, а білінійні форми $B(x,y)=x_1y_1-x_1y_2-x_2y_1+2x_2y_2$ і $C(x,y)=3x_1y_1+2x_1y_2+2x_2y_1-x_2y_2+5x_2y_3+5x_3y_2$ – симетричні, так як

$$A=\begin{pmatrix}2&3&0\\-1&-1&\frac{1}{2}\\10&-12&7\end{pmatrix}\text{ не симетрична,}$$

$$B=\begin{pmatrix}1&-1\\-1&2\end{pmatrix},\ C=\begin{pmatrix}3&2&0\\2&-1&5\\0&5&0\end{pmatrix}\text{ – симетричні.}$$

Квадратичною формою називається числова функція $A(x,x)$ одного аргументу $x\in L$, яка одержується із симетричної форми $A(x,y)$ при $x=y$.

Симетрична форма $A(x,y)$ називається ***полярною*** до квадратичної форми $A(x,x)$.

Очевидно, якщо $A(x,y)=\sum\limits_{i,j=1}^{n}a_{ij}x_iy_j$ – загальний вид симетричної білінійної форми в базисі e, то $A(x,x)=\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j$ – загальний вид квадратичної форми в базисі e простору L розмірності n. В такому вигляді *квадратичну форму* можна розглядати як *числову функцію* n *змінних*:

$$A(x,x) = Q(x_1, x_2, ..., x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$, де $a_{ij} = a_{ji}$ – дійсні числа.

Приклади квадратичних форм:

а) Розглянемо симетричну білінійну форму $B(x,y)$ із попереднього прикладу. Тоді $B(x,x) = x_1^2 - x_1 x_2 - x_2 x_1 + 2x_2^2 = x_1^2 - 2x_1 x_2 + 2x_2^2$ – квадратична форма, а $B(x,y)$ – полярна до неї.

б) Аналогічно $C(x,x) = 3x_1^2 + 4x_1 x_2 - x_2^2 + 10x_2 x_3$, а $C(x,y)$ – полярна до неї.

в) Очевидно, що $A(x,x) = x_1^2 + 3x_1 x_2 - 5x_3^2$ – квадратична форма, а $A(x,y) = x_1 y_1 + \frac{3}{2} x_1 y_2 + \frac{3}{2} x_2 y_1 - 5x_3 y_3$ – полярна до неї.

Зауваження 5.2.2. *Полярна білінійна форма* $A(x,y)$ *і квадратична форма* $A(x,x)$ *пов`язані співвідношенням*

$$A(x,y) = \frac{1}{2}\left[A(x+y, x+y) - A(x,x) - A(y,y)\right]:$$

$$A(x+y, x+y) = A(x, x+y) + A(y, x+y) = A(x,x) + A(x,y) + A(y,x) + A(y,y). \blacksquare$$

Матрицею квадратичної форми $A(x,x)$ ***в базисі*** e називається матриця відповідної полярної форми.

Наприклад $A(x,x) = x_1^2 + 3x_1 x_2 - 5x_2^2$, тоді

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

Зауваження 5.2.3. *Всі твердження і поняття п.3.1 справедливі і для квадратичних форм.*

Квадратична форма $A(x,x)$ називається

1. ***додатно визначеною***, якщо $\forall x \neq \theta$ виконується нерівність $A(x,x)$>0.
2. ***від`ємно визначеною***, якщо $\forall x \neq \theta$ виконується нерівність $A(x,x)$<0
3. ***знакозмінною***, якщо $\exists\, x, y \in L$ такі, що $A(x,x)$>0 і $A(y,y)$<0.
4. ***квазідодатно визначеною,*** якщо $\forall x \in L \quad A(x,x) \geq 0$ і існує $x \neq \theta$, що $A(x,x) = 0$.

5. ***квазівід`ємно визначеною,*** якщо $\forall x \in L \quad A(x,x) \le 0$ і існує $x \ne \theta$, що $A(x,x) = 0$.

Надалі ми познайомимось з критеріями, які дозволяють класифікувати квадратичні форми за належністю до одного з цих типів.

Твердження 5.2.4. *Якщо $A(x,y)$ є білінійна форма, полярна до додатно визначеної квадратичної форми $A(x,x)$, то $A(x,y)$ задовольняє всім аксіомами скалярного добутку в евклідовому просторі:*

1^0. $A(x,y) = A(y,x)$

2^0. $A(x_1 + x_2, y) = A(x_1, y) + A(x_2, y)$

3^0. $A(\lambda x, y) = \lambda A(x,y)$

4^0. $A(x,x) > 0 \; \forall x \ne \theta$ і $A(x,x) = 0 \Leftrightarrow x = \theta \;\; \forall x, y, x_1, x_2 \in L, \forall \lambda \in R$. ■

Контрольні питання і завдання

1. Дайте два означення квадратичної форми.

2. Що таке полярна симетрична форма до квадратичної форми $A(x,x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$?

3. Що таке матриця квадратичної форми?

4. Сформулюйте означення знаковизначених, квазізнаковизначених та знакозмінної квадратичних форм.

5.3. Канонічний вид квадратичної форми

Говорять, що квадратична форма має ***канонічний вид*** в базисі f, якщо

$$A(x,x) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

де $(y_1, y_2, ..., y_n) = x_f$ – координати $x \in L$ в базисі f.

Коефіцієнти $\lambda_1, \lambda_2, ..., \lambda_n$ називаються ***канонічними коефіцієнтами квадратичної форми*** , базис $f : f_1, f_2, ..., f_n$ називається ***канонічним базисом квадратичної форми*** .

В канонічному виді квадратичної форми $A(x,x) = \lambda_1 y_1^2 + ... + \lambda_n y_n^2$, взагалі кажучи, не всі коефіцієнти λ_i відмінні від нуля.

Залишивши тільки ненульові коефіцієнти і зробивши відповідну перенумерацію, одержимо: $A(x,x)=\lambda_1 z_1^2+\lambda_2 z_2^2+...+\lambda_k z_k^2$.Ясно, що $k \le n$. Матриця цієї квадратичної форми в деякому базисі

$$A=\begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_k & \\ 0 & & & 0 \end{pmatrix}, \text{ очевидно, } rgA=k\,.$$

Ранг квадратичної форми, за означенням, дорівнює рангу матриці в довільному базисі, тому ранг квадратичної форми дорівнює k, а саме:

Зауваження 5.3.1. *Ранг квадратичної форми дорівнює числу відмінних від нуля канонічних коефіцієнтів.*

Наступне питання полягає в тому, як привести квадратичну форму до канонічного виду. Які існують шляхи для приведення до канонічного виду? Що взагалі змінює загальний вид квадратичної форми?

По-перше, очевидно, що п е р е х і д до н о в о г о б а з и с у здійснює зміну загального виду (теорема 4.1.4, формула (4.1.1), зауваження 4.2.3). Крім того, маємо

Зауваження 5.3.2. *Кожному лінійному перетворенню базису відповідає невироджене перетворення координат. І навпаки, кожному невиродженому перетворенню координат відповідає лінійне перетворення базису.*

Д о в е д е н н я. Нагадаємо, що перетворення невироджене, коли матриця, що йому відповідає, невироджена.

Нехай $x_e \xrightarrow{C} x_f$, U – матриця переходу від базису e простору L до базису f. За твердженням 1.6.2 $C=(U^{-1})^T$. Тоді $|U|\ne 0 \Leftrightarrow |C|\ne 0$. ■

Отже, зміну загального виду здійснює також перетворення координат.

Нехай $A(x,x)=\sum_{i,j=1}^{n} a_{ij}x_i x_j$ – загальний вид квадратичної форми $A(x,x)$, заданий в базисі e ($A_e=(a_{ij})$). Зрозуміло, що загальний вид квадратичної форми $A(x,x)$ зміниться, якщо перейти до іншого базису f, бо $A_f=U^T A_e U=(b_{ij})$:

$$A(x,x)=\begin{cases} x_e A_e x_e^T=\sum a_{ij}x_i x_j & (1) \\ x_f A_f x_f^T=\sum b_{ij}y_i y_j & (2) \end{cases},$$

тобто зміною базису вид (1) квадратичної форми приведено до виду (2).

Із зауваження 5.3.1 випливає також ,що загальний вид квадратичної форми змінює н е в и р о д ж е н е п е р е т в о р е н н я к о о р д и н а т: якщо C – матриця переходу до нових координат, то $x_f = x_e C$, $x_e = x_f C^{-1}$, $x_e^T = (C^{-1})^T x_f^T$. Тоді $A(x,x) = x_e A_e x_e^T = x_f C^{-1} A_e (C^{-1})^T x_f^T$.

Зауваження 5.3.3. *Матриця переходу від початкових координат* $x_1, x_2, ..., x_n$ *до координат канонічного виду дорівнює добутку матриць переходу, одержаних в процесі приведення до квадратичної форми*, а саме:

1) якщо процес складається з двох кроків, то нехай
$x \xrightarrow{C_1} y \xrightarrow{C_2} z$, тоді $y = xC_1, z = yC_2$, звідки $z = xC_1C_2$, тобто
$x \xrightarrow{C} z, C = C_1C_2$;

2) аналогічно, якщо $C_1, C_2, \ldots, C_k$ – матриці переходу до нових координат в результаті всіх k кроків, то матриця переходу до координат канонічного виду $C = C_1 \cdot C_2 \cdot \ldots \cdot C_k$

Контрольні питання і завдання

1. Що таке канонічний вид квадратичної форми, її канонічний базис та канонічні коефіцієнти?

2. Що приводить до зміни загального виду однієї й тієї ж квадратичної форми?

3. Як знайти канонічний вигляд квадратичної форми, якщо відомі матриця цієї квадратичної форми в деякому базисі e та координати канонічного базису в базисі e:

$$A_e = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix}, \quad \begin{aligned} f_1 &= e_1; \\ f_2 &= -2e_1 + 3e_2 \;? \end{aligned}$$

4. Який вигляд має матриця квадратичної форми в канонічному базисі? Як пов’язані ранг квадратичної форми і канонічні коефіцієнти?

5.4.Метод Лагранжа

Теорема 5.4.1. *Довільна квадратична форма* $\mathrm{A}(x,x)$, *яка задана в n-мірному лінійному просторі L, за допомогою невиродженого перетворення координат може бути приведена до канонічного виду.*

Д о в е д е н н я даної теореми є описання методу Лагранжа в загальному вигляді. Він полягає в послідовному доповненні частини доданків загального виду квадратичної форми до повного квадрату і переході до нових координат.

Нехай $A(x,x)$ – ненульова квадратична форма, яка в базисі $e: e_1, e_2 ....e_n$ має загальний вигляд

$$A(x,x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \qquad (5.4.1)$$

Крок 1.

1. Нехай $a_{11} \neq 0$ (якщо $a_{11} = 0$, то знайдемо $a_{ii} \neq 0$, якщо він існує, і проведемо пере нумерацію координат). У виразі (3.4.1) виберемо всі доданки з x_1, доповнимо їх до повного квадрату з одержанням рівнозначного виразу:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n = a_{11}(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n)^2 -$$

$$- \frac{a_{12}^2}{a_{11}}x_2^2 - \ldots - \frac{a_{1n}^2}{a_{11}}x_n^2 - 2\frac{a_{12}a_{13}}{a_{11}}x_2x_3 - \ldots - 2\frac{a_{1(n-1)}a_{1n}}{a_{11}}x_{(n-1)}x_n,$$

так як квадрат суми дорівнює сумі квадратів доданків плюс сума подвоєних попарних добутків всіх доданків.

Підставимо одержаний вираз в (5.4.1):

$$A(x,x) = \underbrace{a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n}_{(*)} + \underbrace{\sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}x_ix_j}_{(**)} =$$

$$= a_{11}(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots. + \frac{a_{1n}}{a_{11}}x_n)^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}^* x_i x_j, \qquad (5.4.2)$$

де $a_{ij}^* = a_{ij} - \frac{a_{1i}a_{1j}}{a_{11}}$, $i,j = 2,3,\ldots,n$, що одержуємо після зведення подібних доданків.

Розглянемо наступне перетворення координат:

$$\begin{cases} y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n \\ y_2 = x_2 \\ \ldots \\ y_n = x_n \end{cases}$$

Матриця цього перетворення

$$C=\begin{pmatrix} 1 & 0 & ... & 0 \\ \dfrac{a_{12}}{a_{11}} & 1 & ... & 0 \\ ... & ... & ... & ... \\ \dfrac{a_{1n}}{a_{11}} & 0 & ... & 1 \end{pmatrix}, \quad |C|=1 \quad \Rightarrow \quad |C|\neq 0,$$

тобто перетворення координат невироджене. Після цього перетворення загальний вид (3.4.2) матиме вигляд

$$A(x,x)=a_{11}y_1^2+\sum_{i=2}^{n}\sum_{j=2}^{n}a_{ij}^{*}y_iy_j \ . \ (3.4.3)$$

2. Якщо в (5.4.1) всі $a_{ii}=0, i=1,...,n$, то нехай $a_{12}\neq 0$. Розглянемо невироджене перетворення координат $x\xrightarrow{C}x'$:

$$\begin{cases} x'=x_1-x_2 \\ x_2'=x_1+x_2 \\ x_3'=x_3 \\ ...... \\ x_n'=x_n \end{cases}, \qquad C=\begin{pmatrix} 1 & 1 & 0 & .... & 0 \\ -1 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix}$$

$$|C|=\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix}\cdot\begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 1 \end{vmatrix}=2\cdot 1=2\neq 0$$

$$x_1=\frac{1}{2}(x_1'+x_2') . \ x_2=\frac{1}{2}(x_2'-x_1')$$

$$2a_{12}x_1x_2=2a_{12}\frac{1}{4}(x_2'^2-x_1'^2)=\frac{a_{12}}{2}x_1'^2-\frac{a_{12}}{2}x_2'^2$$

Після цього перетворення (5.4.1) матиме вигляд:

$$A(x,x)=\sum_{i,j=1}^{n}a_{ij}'x_i'x_j' ,$$

де $a'_{11} = \frac{a_{12}}{2} \neq 0$ і ми можемо використовувати далі процес, описаний в попередньому випадку. і одержати вид (5.4.3).

Крок 2.

Розглянемо у рівності (3.4.3) квадратичну форму $\sum_{i=2}^{n}\sum_{j=2}^{n} a_{ij}^{*} y_i y_j$:

1) якщо вона нульова, то $A(x,x)$ приведена до канонічної форми:

$$A(x,x) = a_{11}y_1^2 + 0y_2^2 + \ldots + 0y_n^2;$$

2) якщо вона ненульова, то звернімо увагу, що в ній відсутня перша координата. Тоді виконаємо процес, аналогічний випадку1 (або 2-1.): виберемо всі доданки з y_2, доповнимо їх до повного квадрату, застосуємо заміну координат, яка змінює другу координату. Очевидно, що ці перетворення координат $y_1, y_2, \ldots, y_n$ будуть невиродженими. В результаті виділяться квадрати перших двох змінних та квадратична форма останніх змінних.

Так як ми маємо скінченну кількість змінних, то за скінченну кількість таких кроків, застосованих до квадратичної форми, що залишається після виділення повних квадратів, квадратична форма $A(x,x)$ приведеться до канонічного виду. ■

Зауваження 5.4.2. *Канонічний базис квадратичної форми, взагалі кажучи, визначається неоднозначно.* Хоча в результаті застосування методу Лагранжа можна знайти канонічний базис $f = e((C_1C_2 \ldots C_k)^{-1})^T$, де $C_1, C_2, \ldots, C_k$ – матриці переходу до нових координат в результаті всіх k кроків.

Приклад. Нехай квадратична форма

$$A(x,x) = x_1^2 + 2x_1x_2 + 4x_1x_3 - 3x_1x_4 + 2x_2^2 + 2x_2x_3 - 3x_2x_4 + 5x_3^2 - 6x_3x_4 + 3x_4^2.$$

Знайти її канонічний вигляд і ранг. Знайти матрицю переходу до канонічного базису.

Δ Крок 1. Виділимо доданки з x_1 і повний квадрат з цієї суми:

$$x_1^2 + 2x_1x_2 + 4x_1x_3 - 3x_1x_4 = (x_1 + x_2 + 2x_3 - \frac{3}{2}x_4)^2 - x_2^2 - 4x_3^2 - \frac{9}{4}x_4^2 - 4x_2x_3 + 3x_2x_4 + 6x_3x_4$$

Замість перших чотирьох доданків $A(x,x)$ підставимо праву частину одержаної рівності, зведемо подібні доданки, зробимо заміну координат:

$$\left[\begin{array}{l} y_1 = x_1 + x_2 + 2x_3 - \frac{3}{2}x_4^2 \\ y_2 = x_2 \\ y_3 = x_3 \\ y_4 = x_4 \end{array}\right], \; A(x,x) = y_1^2 + y_2^2 - 2y_2y_3 + y_3^2 + \frac{3}{4}y_4^2.$$

Крок 2. Виділимо доданки з y_2, повний квадрат з цієї суми, підставимо праву частину одержаної рівності в $A(x,x)$, зведемо подібні доданки, зробимо заміну координат:

$$y_2^2 - 2y_2y_3 = (y_2 - y_3)^2 - y_3^2;$$

$$\left[\begin{array}{l} z_1 = y_1 \\ z_2 = y_2 - y_3 \\ z_3 = y_3 \\ z_4 = y_4 \end{array}\right], \; A(x,x) = z_1^2 + z_2^2 + 0z_3^2 + \frac{3}{4}z_4^2.$$

Одержали канонічний вид квадратичної форми $A(x,x)$. Вона має три ненульових коефіцієнти, тому $rgA(x,x) = 3$.

Матрицями переходу до нових координат кожного кроку будуть матриці $C_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ -\frac{3}{2} & 0 & 0 & 1 \end{pmatrix}$ і $C_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ відповідно. Тоді

$x \xrightarrow{C} z, C = C_1C_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -\frac{3}{2} & 0 & 0 & 1 \end{pmatrix}$, а матрицею переходу до канонічного базису буде матриця

$$U = (C^T)^{-1} = \begin{pmatrix} 1 & 1 & 2 & -\frac{3}{2} \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & -3 & \frac{3}{2} \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Якщо позначити $e: e_1, e_2, e_3, e_4$ – базис, в якому задана квадратична форма $A(x,x)$, а $f: f_1, f_2, f_3, f_4$ – її канонічний базис, то із матриці U одержимо:

$$f_1 = e_1;$$
$$f_2 = -e_1 + e_2;$$
$$f_3 = -3e_1 + e_2 + e_3; \blacksquare$$
$$f_4 = \frac{3}{2}e_1 + e_4.$$

Контрольні питання і завдання

1. Чи існує квадратична форма, яку не можна методом Лагранжа привести до канонічного вигляду?

2. Поясність, в чому полягає доповнення до повного квадрату в методі Лагранжа?

3. Як знайти канонічний базис при використанні методу Лагранжа.

5.5. Метод Якобі

Нехай, як і раніше, квадратична форма $A(x,x)$ задана в n- мірному лінійному просторі L. При деяких додаткових обмеженнях на квадратичну форму $A(x,x)$ можна вказати явні формули переходу від даного базису e до канонічного базису f і формули для канонічних коефіцієнтів (чого не вистачає в методі Лагранжа).

Попередньо введемо поняття трикутного перетворення базисних векторів.

Перетворення базису $e_1, e_2, \ldots, e_n$ називається ***трикутним***, якщо воно має наступний вид:

$$\begin{aligned} f_1 &= e_1 \\ f_2 &= \lambda_{21}e_1 + e_2 \\ f_3 &= \lambda_{31}e_1 + \lambda_{32}e_2 + e_3 \\ &\ldots \\ f_n &= \lambda_{n1}e_1 + \lambda_{n2}e_2 + \ldots + \lambda_{n(n-1)}e_{n-1} + e_n \end{aligned} \tag{5.5.1}$$

Зауваження 5.5.1. *Довільне трикутне перетворення невироджене, а значить* $f_1, f_2, \ldots, f_n$ *– базис* **.**

Обчислимо визначник $|U|$ матриці трикутного перетворення (5.5.1):

$$e \xrightarrow{U} f \qquad U = \begin{pmatrix} 1 & \lambda_{21} & \lambda_{31} & \dots & \lambda_{n1} \\ 0 & 1 & \lambda_{32} & \dots & \lambda_{n2} \\ 0 & 0 & 1 & \dots & \lambda_{n3} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \quad |U| = 1 \neq 0. \blacksquare$$

Визначимо ***кутові мінори*** матриці $A_e = (a_{ij})$ квадратичної форми *A(x,x)* в базисі *e*:

$$\Delta_1 = |a_{11}|, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \Delta_i = \begin{vmatrix} a_{11} & \dots & a_{1i} \\ \dots & \dots & \dots \\ a_{i1} & \dots & a_{ii} \end{vmatrix}.$$

Теорема 5.5.2. (Якобі) *Нехай мінори* $\Delta_1, \Delta_2, ..., \Delta_{n-1}$ *матриці* A_e *квадратичної форми A(x, x) відмінні від нуля. Тоді існує єдине трикутне перетворення базисних векторів* $e_1, e_2, ..., e_n$, *за допомогою якого форму A(x, x) можна привести до канонічного виду.*

Д о в е д е н н я**.** Нехай $f : f_1, f_2, ..., f_n$ – інший базис L, $A_f = (b_{ij})$, тоді $b_{ij} = A(f_i, f_j)$.

Якщо форма *A(x, x)* в базисі f має канонічний вигляд, то $b_{ij} = 0$ при $i \neq j$. Тобто за допомогою трикутного перетворення (3.5.1) треба побудувати такий базис f, щоб $A(f_i, f_j) = 0$ при $i \neq j$. Так як *A(x, x)* – симетрична, то достатньо, щоб $A(f_i, f_j) = 0$ при $i < j$. Тобто треба знайти коефіцієнти трикутного перетворення (5.5.1) $\lambda_{ji}, j = 2,...,n, i = 1,..., j-1$, які задовольняють умовам $A(f_i, f_j) = 0$.

Зафіксуємо деяке j. З одного боку:

$$i = 1 \qquad \left.\begin{matrix} A(f_1, f_j) \overset{(3.5.1)}{=} A(e_1, f_j) \\ A(f_1, f_j) = 0, j \neq 1 \end{matrix}\right] \Rightarrow A(e_1, f_j) = 0; \qquad (1)$$

$$i = 2 \quad \left.\begin{matrix} A(f_2, f_j) = \lambda_{21} A(e_1, f_j) + A(e_2, f_j) \\ A(f_2, f_j) = 0, j \neq 2 \end{matrix}\right\} \overset{(1)}{\Rightarrow} A(e_2, f_j) = 0; \qquad (2)$$

…

$i = j-1 \ldots \overset{(j-2)}{\Rightarrow} A(e_{j-1}, f_j) = 0$. $(j-1)$

З іншого боку, так як $f_j = \lambda_{j1}e_1 + \lambda_{j2}e_2 + \ldots + \lambda_{j(j-1)}e_{j-1} + e_j$, то

$A(e_i, f_j) = \lambda_{j1}A(e_i, e_1) + \lambda_{j2}A(e_i, e_2) + \ldots + \lambda_{j(j-1)}A(e_i, e_{j-1}) + A(e_i, e_j)$, $i < j$

і одержимо наступну систему умов:

$$\begin{cases} 0 = A(e_1, f_j) = \alpha_{j1}a_{11} + \alpha_{j2}a_{12} + \ldots + \alpha_{j(j-1)}a_{1(j-1)} + a_{1j} \\ 0 = A(e_2, f_j) = \alpha_{j1}a_{21} + \alpha_{j2}a_{22} + \ldots + \alpha_{j(j-1)})a_{2(j-1)} + a_{2j} \\ \qquad\qquad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ 0 = A(e_{j-1}, f_j) = \alpha_{j1}a_{(j-1)1} + \alpha_{j2}a_{(j-1)2} + \ldots + \alpha_{j(j-1)}a_{(j-1)(j-1)} + a_{(j-1)j} \end{cases},$$

яка в стандартному виді виглядає як система $j-1$ лінійних рівнянь з $j-1$ невідомими $\lambda_{j1}, \lambda_{j2}, \ldots, \lambda_{j(j-1)}$:

$$\begin{cases} a_{11}\lambda_{j1} + a_{12}\lambda_{j2} + \ldots + a_{1(j-1)}\lambda_{j(j-1)} = -a_{1j} \\ a_{21}\lambda_{j1} + a_{22}\lambda_{j2} + \ldots + a_{2(j-1)}\lambda_{j(j-1)} = -a_{2j} \\ \qquad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ a_{(j-1)1}\lambda_{j1} + a_{j2_{(j-1)2}} + \ldots + a_{(j-1)(j-1)}\lambda_{j(j-1)} = -a_{(j-1)j} \end{cases}. \qquad (5.5.2)$$

Визначник матриці цієї системи $\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1(j-1)} \\ a_{21} & a_{22} & \ldots & a_{2(j-1)} \\ \ldots & \ldots & \ldots & \ldots \\ a_{(j-1)1} & a_{(j-1)2} & \ldots & a_{(j-1)(j-1)} \end{vmatrix} = \Delta_{j-1}$.

За умовою $\Delta_{j-1} \neq 0 \quad \forall j = 2, \ldots, n$. Система (5.5.2 j) має єдиний розв'язок $(\lambda_{j1}, \lambda_{j2}, \ldots, \lambda_{j(j-1)})$.

Розглядаючи $j = 2, \ldots, n$, знайдемо всі λ_{ji} $(i < j)$, які визначені однозначно, значить трикутне перетворення, яке приводить до канонічного базису, єдине. Розв'язувати ці системи будемо методом Крамера.

У кожній системі (5.5.2) стовпчик вільних членів $B = \begin{pmatrix} -a_{1j} \\ -a_{2j} \\ \ldots\ldots. \\ -a_{(j-1)j} \end{pmatrix}$

складається з елементів матриці A_e, взятих з перших $(j-1)$ рядків та з j-го стовпчика з протилежним знаком, тому за формулами Крамера $\lambda_{ji} = \frac{\Delta^i_{j-1}}{\Delta_{j-1}}, \quad i = 1,2,...,j-1$, де

$$\Delta^i_{j-1} = \begin{vmatrix} a_{11} & \dots & a_{1(i-1)} & -a_{1j} & a_{1(i+1)} & \dots & a_{1(j-1)} \\ a_{21} & \dots & a_{2(i-1)} & -a_{2j} & a_{2(i+1)} & \dots & a_{2(j-1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{(j-1)1} & \dots & a_{(j-1)(i-1)} & -a_{(j-1)j} & a_{(j-1)(i+1)} & \dots & a_{(j-1)(j-1)} \end{vmatrix} -$$

визначник, що одержується з матриці системи (5.5.2) заміною i-го стовпчика на стовпчик вільних членів B.

Позначимо мінор матриці A_e, складений з перших ($j-1$) рядочків та перших j стовпчиків без i-го стовпчика

$$M^{12...(i-1)(i+1)...j}_{12...(j-1)} = \begin{vmatrix} a_{11} & \dots & a_{1(i-1)} & a_{1(i+1)} & \dots & a_{1j} \\ a_{21} & \dots & a_{2(i-1)} & a_{2(i+1)} & \dots & a_{2j} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{(j-1)1} & \dots & a_{(j-1)(i-1)} & a_{(j-1)(i+1)} & \dots & a_{(j-1)j} \end{vmatrix} = \Delta_{j-1,i}.$$

Для того, щоб із Δ^i_{j-1} одержати $\Delta_{j-1,i}$, треба стовпчик вільних членів B переставити $(j-1)-i$ разів і помножити його на (–1). Тоді справедлива рівність

$$\Delta_{j-1,i} = (-1)^{j-1-i+1}\Delta^i_{j-1} = (-1)^{j-i}\Delta^i_{j-1} \Rightarrow \Delta^i_{j-1} = \frac{1}{(-1)^{j-i}}\Delta_{j-1,i} = \frac{(-1)^{2j}}{(-1)^{j-i}}\Delta_{j-1,i} = (-1)^{j+i}\Delta_{j-1,i}$$

і формули Крамера будуть мати вигляд

$$\lambda_{ji} = \frac{\Delta^i_{j-1}}{\Delta_{j-1}} = (-1)^{j+i}\frac{\Delta_{j-1,i}}{\Delta_{j-1}}. \quad (5.5.3)$$

Таким чином, формули (5.5.3) – ***формули коефіцієнтів трикутного перетворення.***

Знайдемо канонічні коефіцієнти λ_i, $i = 1,2,...,n$ з умови, що $\lambda_i = A(f_i, f_i)$:

$$\lambda_1 = A(f_1, f_1) = A(e_1, e_1) = a_{11} = \Delta_1;$$

$$\lambda_2 = A(f_2, f_2) = \lambda_{21}A(e_1, f_2) + A(e_2, f_2) \overset{(1)}{=} A(e_2, f_2) = \lambda_{21}A(e_2, e_1) + A(e_2, e_2) =$$

$$= (-1)^3\frac{\Delta_{1,1}}{\Delta_1}a_{21} + a_{22} = \frac{(-1)^3\Delta_{1,1}a_{21} + a_{22}\Delta_1}{\Delta_1} = \frac{-a_{21}a_{12} + a_{22}a_{11}}{\Delta_1} = \frac{\Delta_2}{\Delta_1};$$

Отже, $\forall i = 2,...,n$

$$\lambda_i = A(f_i, f_i) = \lambda_{i1}A(e_1, f_i), + \ldots + \lambda_{ii-1}A(e_{i-1}, f_i) + A(e_i, f_i) \overset{(1),\ldots,(i-1)}{===} A(e_i, f_i) =$$
$$= \lambda_{i1}A(e_i, e_1) + \ldots + \lambda_{ii-1}A(e_{i-1}, e_i) + A(e_i, e_i) = \lambda_{i1}a_{i1} + \ldots + \lambda_{ii-1}a_{ii-1} + a_{ii} =$$
$$= \frac{(-1)^{i+1}\Delta_{i-1,1}a_{i1} + \ldots + (-1)^{i+i-1}\Delta_{i-1i-1}a_{ii-1} + \Delta_{i-1}a_{ii}}{\Delta_{i-1}} = \frac{\Delta_i}{\Delta_{i-1}},$$

бо чисельник передостаннього дробу є розкладом визначника Δ_i по i-му рядку.

Одержуємо ***формули для знаходження канонічних коефіцієнтів***:

$$\lambda_1 = \Delta_1, \lambda_2 = \frac{\Delta_2}{\Delta_1}, \lambda_3 = \frac{\Delta_3}{\Delta_2}, \ldots, \lambda_n = \frac{\Delta_n}{\Delta_{n-1}}. \qquad (5.5.4)$$

Знаходження канонічного вигляду та канонічного базису квадратичної форми за допомогою формул (3.5.3) та (3.5.4) називається ***методом Якобі***.

Контрольні питання і завдання

1. За яких умов квадратичну форму можна привести до канонічного вигляду методом Якобі.

2. Запишіть формули канонічних коефіцієнтів та коефіцієнтів канонічного базису методу Якобі.

5.6. Класифікація квадратичних форм

В зауваженні 5.3.1. відмічалось, що число відмінних від нуля канонічних коефіцієнтів квадратичної форми дорівнює її рангу, а ця величина не залежить від базису та невиродженого перетворення. В дійсності, навіть *кількість від'ємних та додатних коефіцієнтів не залежить від способу приведення до канонічної форми.* Ця властивість називається ***законом інерції*** квадратичної форми.

Індексом інерції квадратичної форми називається кількість відмінних від нуля канонічних коефіцієнтів (k) (тобто рангу), ***додатним індексом інерції*** – кількість додатних канонічних *коефіцієнтів* (p), ***від'ємним індексом інерції*** – число від'ємних канонічних коефіцієнтів (q).

Ясно, що $k = p + q$.

$$A(x, x) = z_1^2 + z_2^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_k^2.$$

Теорема 5.6.1. (закон інерції квадратичної форми). *Кількість доданків з додатними (від'ємними) коефіцієнтами в нормальному виді квадратичної форми не залежить від способу приведення форми до його канонічного виду.*

Д о в е д е н н я. Нехай в базисі e квадратична форма $A(x,x)=\sum_{i,j=1}^{n}a_{ij}x_ix_j$ за допомогою одного невиродженого перетворення приводиться до виду

(1) $$A(x,x)=\lambda_1y_1^2+\ldots+\lambda_qy_q^2+\lambda_{q+1}y_{q+1}^2+\ldots+\lambda_ky_k^2,\quad k\le n,$$ причому

$$\left.\begin{array}{l}\lambda_1>0,\lambda_2>0,\ldots,\lambda_q>0\\ \lambda_{q+1}<0,\ldots,\lambda_k<0\end{array}\right\}\ \lambda_i\ne 0,\quad i=1,\ldots,k.$$

а за допомогою іншого – до виду

(2)

$$A(x,x)=\mu_1t_1^2+\ldots+\mu_pt_p^2+\mu_{p+1}t_{p+1}^2+\ldots+\mu_st_s^2,\quad s\le n,$$ де

$$\left.\begin{array}{l}\mu_1>0,\mu_2>0,\ldots,\mu_p>0\\ \quad\mu_{p+1}<0,\ldots,\mu_s<0\end{array}\right\}.$$

Зрозуміло, що $k=s$, бо кількість ненульових коефіцієнтів дорівнює рангу, який не залежить від базису. Тоді очевидно, що для доведення теореми достатньо показати, що $q=p$(бо тоді від'ємних коефіцієнтів також буде однакова кількість: $k-q=k-p$). Нехай $p>q$, тоді $q-p<0$.

$$x_0=(0,0,\ldots,0,z_{q+1},\ldots,z_n)_f,\ (z_1=z_2=\ldots=z_q=0);$$
$$x_0=(t_1,\ldots,t_p,0,\ldots,0)_g,\qquad \underbrace{(t_{p+1}=t_{p+2}=\ldots=t_n=0)}_{n-p}.$$

Таким чином, на вектор x_0 накладається $q+(n-p)=n+(q-p)<n$ умов, які утворюють однорідну систему $(n+q-p)$ лінійних рівнянь з n невідомими:

$$\begin{cases}z_1=\ u_{11}x_1+\ldots+u_{n1}x_n=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ z_q=\ u_{1p}x_1+\ldots..+u_{np}x_n=0\\ t_{p+1}=\ v_{1p+1}x_1+\ldots+v_{np+1}x_n=0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ t_n=\ v_{1n}x_1+\ldots+v_{nn}x_n=0\end{cases}.\tag{5.6.1}$$

Кількість рівнянь системи (5.6.1), а значить і її ранг, менший кількості невідомих, значить ця однорідна система має ненульовий розв'язок $(x_1^0,x_2^0,\ldots,x_n^0)$, тоді $x_0=(x_1^0,x_2^0,\ldots,x_n^0)_e$ – шуканий вектор.

Таким чином,

$$\left.\begin{array}{l}(1) \Rightarrow A(x_0, x_0) = -z_{q+1}^2 - \ldots - z_n^2 \le 0 \\ (2) \Rightarrow A(x_0, x_0) = t_1^2 + t_2^2 + \ldots + t_p^2 \ge 0\end{array}\right\} \Rightarrow z_{q+1} = \ldots = z_k = t_1 = \ldots = t_p = 0 \Rightarrow$$

$$\Rightarrow x_0 = (0,0\ldots,0)_f = 0 \cdot f_1 + 0 \cdot f_2 + \ldots + 0 \cdot f_n = \theta.$$

Одержане протиріччя виникає від невірного припущення, що $p > q$, тобто $p \le q$. Аналогічно виникає протиріччя, якщо припустити, що $p < q$. Звідки $p = q$ і теорема доведена. ■

Квадратичні форми класифікуються за п'ятьма типами: додатновизначені, від'ємновизначені, знакозмінні, квазідодатновизначені та квазівід'ємновизначені. Ця класифікація спрощується за допомогою поняття індексу інерції.

Теорема 5.6.3. (Критерій знаковизначеності квадратичної форми) *Квадратична форма $A(x,x)$, задана в n- мірному лінійному просторі L, є додатно (від'ємно) визначена тоді і тільки тоді, коли додатний індекс інерції p (від'ємний індекс інерції q) дорівнює n: $p = n \quad (q = n)$.*

Теорема 5.6.4. (Критерій квазівизначеності) *Квадратична форма $A(x,x)$ квазівизначена тоді і тільки тоді, коли або $p < n, \quad q = 0,$ або $p = 0, \quad q < n$.*

Теорема 5.6.5. (Критерій знакозмінності квадратичної форми) *Квадратична форма $A(x,x)$ знакозмінна тоді і тільки тоді, коли і додатний, і від'ємний індекси інерції цієї форми відмінні від нуля: $p > 0 \quad і \quad q > 0$.*

Існує ще один критерій знаковизначеності квадратичної форми, який відрізняється від попередніх тим, що для визначення класу, до якого належить квадратична форма не потрібно приводити до канонічного вигляду.

Нехай $A(x,x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ в базисі $e: e_1, \ldots, e_n$ з матрицею $A_e = (a_{ij})$.

Нехай

$$\Delta_1 = a_{11}, \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ldots, \Delta_n = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ & \ldots & \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$ – кутові мінори матриці A_e.

Теорема 5.6.6. (Критерій Сильвестра)

а) *Для того, щоб квадратична форма $A(x,x)$ була додатно визначена, необхідно і достатньо, щоб виконувались нерівності:* $\Delta_1 > 0, \Delta_2 > 0, \ldots, \Delta_n > 0.$

б) *Квадратична форма від'ємно визначена тоді і тільки тоді, коли знаки кутових мінорів чередуються так:* $\Delta_1 < 0,\ \Delta_2 > 0,\ \Delta_3 < 0,\ \Delta_4 > 0, ...$

Д о в е д е н н я**.** Спочатку розглянемо загальні властивості знаковизначеної квадратичної форми $A(x,x)$. Припустимо, що $\Delta_k = 0$, розглянемо систему лінійних рівнянь

$$(*) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1k}x_k = 0 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2k}x_k = 0 \\ ... \\ a_{k1}x_1 + a_{k2}x_2 + ... + a_{kk}x_k = 0 \end{cases} \left| \begin{array}{c} \cdot x_1 \\ \cdot x_2 \\ ... \\ \cdot x_k \end{array} \right.$$

Це однорідна система з квадратною матрицею системи і її визначником $\Delta_k = 0$, звідки система невизначена, отже існує ненульовий розв'язок $(x_1^0,...,x_k^0)$ системи (*).

Помножимо кожне i- те рівняння системи (*) на x_i і складемо. Одержимо

$$\sum_{i,j=1}^{k} a_{ij}x_i x_j = 0 \ (**)$$

Рівність (**) виконується для всіх розв'язків системи (*).

Візьмемо $x' = (x_1^0,..., x_k^0,0,...,0)_e \neq \theta$.Знайдемо

$$A(x',x') = \sum_{i,j=1}^{n} a_{ij}x_i'x_j' = \sum_{i,j=1}^{k} a_{ij}x_i^0 x_j^0 \overset{(**)}{=} 0.$$

Одержуємо протиріччя із знаковизначеністю $A(x,x)$. Значить $\Delta_i \neq 0, i = 1,...,n.$

Для доведення теореми можемо використати формули методу Якобі для знаходження канонічних коефіцієнтів через кутові мінори.

а) Нехай $A(x,x)$ – додатно визначена квадратична форма. За критерієм 5.6.3 додатний індекс інерції цієї форми $p = n$, значить $\lambda_1 > 0,...,\lambda_n > 0$. Так як $\Delta_i \neq 0, i = 1,...,n$, то розглянемо ці умови, враховуючи формули методу Якобі для знаходження канонічних коефіцієнтів: $\lambda_1 = \Delta_1 > 0,\ \lambda_2 = \dfrac{\Delta_2}{\Delta_1} > 0, ...,\lambda_n = \dfrac{\Delta_n}{\Delta_{n-1}} > 0 \ \Leftrightarrow\ \Delta_1 > 0, \Delta_2 > 0, ..., \Delta_n > 0$.

Всі слідування цього доведення еквівалентні, тому доведено необхідну і достатню умови.

б) Нехай $A(x,x)$ – від'ємно визначена квадратична форма. За критерієм 3.6.3 від'ємний індекс інерції цієї форми $q=n$, значить $\lambda_1<0,...,\lambda_n<0$. Так як $\Delta_i\neq 0, i=1,...,n$, то розглянемо ці умови, враховуючи формули методу Якобі для знаходження канонічних коефіцієнтів: $\left(\lambda_1=\Delta_1<0, \lambda_2=\dfrac{\Delta_2}{\Delta_1}<0, \lambda_3=\dfrac{\Delta_3}{\Delta_2}<0, ...\right) \Leftrightarrow \Delta_1<0, \Delta_2>0, \Delta_3<0$.

Всі слідування цього доведення еквівалентні, тому доведено необхідну і достатню умови. ■

Дві квадратичні форми f і g одного і того ж лінійного простору називаються ***еквівалентними***, якщо існує не вироджене перетворення координат T, при якому загальний вид квадратичної форми f співпадає з загальним видом квадратичної форми g. Позначається $f\sim g$.

Твердження 5.6.7. *Квадратичні форми еквівалентні тоді і тільки тоді, коли їх додатні і від'ємні індекси інерції співпадають.*

Контрольні питання і завдання

1. Дайте означення нормального виду квадратичної форми.
2. Сформулюйте закон інерції квадратичних форм.
3. Сформулюйте критерії знаковизначеності, квазізнаковизначеності та знакозмінності квадратичних форм.
4. Сформулюйте критерій Сильвестра.
5. Дайте означення еквівалентних квадратичних форм.

СПИСОК ПОЗНАЧЕНЬ

$\mathbf{N_0}$ – множина натуральних чисел з числом нуль;

$\mathbf{Z}$ – множина цілих чисел;

$\mathbf{Q}$ – множина раціональних чисел;

$\mathbf{R}$ – множина дійсних чисел;

$\mathbf{R}_+$ – множина додатних дійсних чисел;

$\mathbf{C}$ – множина комплексних чисел;

L_P – лінійний простір L над полем P;

e – рядок, складений із елементів базису $e_1, e_2, ..., e_n$;

x_e – рядок із координат елементу x в базисі e;

x_e^T – стовпчик із координат елементу x в базисі e;

$\dim L$ – розмірність простору L;

$e \xrightarrow{U} e'$ – матрицею переходу від базису e до базису e' є матриця U;

$L_1 \oplus L_2$ – пряма сума підпросторів L_1 і L_2;

$P_n[x]$ – множина многочленів з дійсними коефіцієнтами степеня не більшого n;

V_n – множина дійсних векторів-рядків порядку n;

T_n – множина дійсних векторів-стовпчиків порядку n;

$M_{m\times n}(\mathbf{R})$, $M_{m\times n}$ – множина матриць розміру $m \times n$ з дійсними коефіцієнтами;

$M_n(\mathbf{R})$, M_n – множина квадратних матриць n-го порядку з дійсними коефіцієнтами;

E – одинична матриця;

$C_{[a,b]}$ – множина дійсних функцій неперервних на відрізку $[a,b]$;

$C^*_{[a,b]}$ – множина комплексних функцій, неперервних на відрізку $[a,b]$;

T_n^* – множина комплексних векторів-стовпчиків порядку n;

(x, y) – скалярний добуток векторів x і y;

$M^\perp$ – ортогональне доповнення множини M;

$\widetilde{A}: V \to W$ – лінійне відображення із простору V в простір W;

$L(V,W)$ – простір лінійних відображень із простору V в простір W;

$\widetilde{I}$ – тотожне перетворення;

A_e – матриця лінійного перетворення $\widetilde{A}$ в базисі e;

rgA – ранг матриці A;

$Ker\widetilde{A}$ – ядро лінійного оператору $\widetilde{A}$;

$\operatorname{Im}\widetilde{A}$ – образ лінійного оператору $\widetilde{A}$;

$def\widetilde{A}$ – дефект лінійного оператору $\widetilde{A}$;

$rang\widetilde{A}$ – ранг лінійного оператору $\widetilde{A}$;

$V_{\widetilde{A}}^{\lambda}$ – множина власних векторів перетворення $\widetilde{A}$, що відповідають власному значенню λ.

$\det\widetilde{A}$ – детермінант оператору $\widetilde{A}$;

$\det(\widetilde{A}-\lambda\widetilde{I})$ – характеристичний многочлен оператору $\widetilde{A}$.

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

Основна

1. Андрійчук В.І. Лінійна алгебра: навчальний посібник / В. І. Андрійчук, Б. В. Забавський. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2008. – 238с.

2. Безущак О.О, Ганюшкін О.Г., Кочубінська Є.А. Навчальний посібник з лінійної алгебри для студентів механіко-математичного факультету. – К. : ВПЦ «Київський університет», 2019. – 224 с.

3. Бутузов В. Ф. Линейная алгебра в вопросах и задачах : учеб. пособие / [В. Ф. Бутузов, Н. И. Крутицкая, А. А. Шишкин]; под ред В. Ф. Бутузова. – М. : ФИЗМАТЛИТ, 2010.–248 с.

4. Борозенець Н. С., Пугач В.І. Вища математика. Лінійна та векторна алгебра. Аналітична геометрія. Практикум для студентів 1 курсу інженерно-технологічних спеціальностей денної і заочної форм навчання / Суми: СНАУ, 2017 р.

5. Винберг Э.Б. Курс алгебры / Э.Б. Винберг. – М.: Факториал Пресс, 2011. – 544 с

6. Волошина Т.В. Лінійна алгебра: навч. посібник / Т.В. Волошина. – Луцьк: Вежа-Друк, 2020. – 308 с.

7. Гантмахер Ф.Р. Теория матриц, - М.: Физматгиз, 2010. -353 с.

8. Дрозденко В.О. Вища математика: необхідний теоретичний мінімум: навч. посіб. В.О. Дрозденко, О.Л. Дрозденко Б.: Пшонківський О.В., 2020. 264 с.

9. Довгай Б.В., Шестаков С.С. Комплексні числа та многочлени: посібник до розв'язання задач. – 2017.- 46 с

10. Завало С. Т. Алгебра і теорія чисел : курс лекцій: в 2 ч. / С. Т. Завало, В. Н. Костарчук, Б. І. Хацет. – К. : Вища шк., 2010. – Ч.1. – 398 с.

11. Завало С. Т. Алгебра і теорія чисел : практикум: в 2 ч. / [С. Т. Завало, С. С. Левіщенко, В. В. Пилаєв, І. О. Рокицький] – К. : Вища шк. Головне вид-во, 2010. – Ч.1. – 232 с.)

12. Зайцев О.П. Вища математика: лінійна та векторна алгебра, аналітична геометрія, вступ до матаналізу. – навч. посібник. – К.: Алерта, 2017. – 574 с.

13. Ильин В. А. Линейная алгебра : учеб. пособие / В. А. Ильин, Э. Г. Поздняк. – 2-е изд., стереотипное, серия "Курс высшей математики …" – М. : Наука, 2010. – 304 с.

14. Лінійна алгебра та аналітична геометрія : Навч. посібник / В. В. Булдигін, І. В. Алєксєєва, В. О. Гайдей, О. О. Диховничий, Н. Р. Коновалова, Л. Б. Федорова; за ред. проф. В. В. Булдигіна. — К. : ТВіМС, 2009. — 224 с.

15. Ляпин Е.С., Евсеев А.Е. Алгебра и теория чисел. Ч. II. Линейная алгебра и полиномы. Учебное пособие для студентов физ.-мат. факультетов педагогических институтов. – М.: "Просвещение" – 1978 – 448 с. (не перевидавався)

16. Кострикин А.И. Введение в алгебру. Часть II. Линейная алгебра. Учебник для вузов / А.И.Кострикин. – М.: Физико-математическая литература, 2012, 368 с.

17. Куликов Л.Я. Алгебра и теория чисел: Учебное пособие для пед. инст. / Л. Я. Куликов. – М. : Высш. шк., 2010. – 559 с.

18. Методичні вказівки до практичних занять з курсу «Лінійна алгебра» (1 семестр) / Пащенко З.Д., Турка Т.В. – Слов'янськ: ДВНЗ «ДДПУ», 2016, – 80 с.

19. Методичні вказівки до практичних занять з курсу «Лінійна алгебра» (2 семестр)/ Пащенко З.Д., Турка Т.В. – Слов'янськ: ДВНЗ «ДДПУ», 2017, – 109 с.

20.

21. Проскуряков И.В. Сборник задач по линейной алгебре : учеб. пособие / И. В. Проскуряков. – М., 1962. – 332 с .

22. Рокіцький І.О. Застосування лінійної алгебри / І. О. Рокіцький, О. Б. Панасенко. — Вінниця : Вид. Главацька Р. В., 2012. — 240 с.

23. Сборник задач по линейной алгебре: учеб. пособие / [Р.Ф. Апатенок и др.]. – Минск, 1980. – 192 с.

Підписано до друку 30.06.2020 р.
Формат 60х84 1/16. Ум. др. арк. 10,75.
Наклад 50 прим. Зам. № 1869.

Видавництво Б.І. Маторіна
84116, м. Слов'янськ, вул. Г. Батюка, 19.
Тел.: +38 06262 3-20-99; +38 050 518 88 99. E-mail: matorinb@ukr.net

Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавців, виготівників і розповсюджувачів видавничої продукції ДК №3141, видане Державним комітетом телебачення та радіомовлення України від 24.03.2008 р.