О. Б. Панасенко

Лекції з лінійної алгебри

Видання 2-е, доповнене

ТОВ «Нілан-ЛТД»
Вінниця, 2015

УДК 512.64 (075.8)
ББК 22.143
П16

Рекомендовано до друку Вченою радою Вінницького державного педагогічного університету імені Михайла Коцюбинського (протокол №5 від 28.10.2015 р.)

Рецензенти:

Працьовитий М.В., доктор фізико-математичних наук, професор, завідувач кафедри вищої математики Національного педагогічного університету імені М.П. Драгоманова

Величко І.Г., кандидат фізико-математичних наук, доцент кафедри вищої математики та фізики Таврійського державного агротехнологічного університету

Панасенко О. Б.

П16 Панасенко О. Б. Лекції з лінійної алгебри / О. Б. Панасенко. — Вид. 2-е, доповн. — Вінниця : ТОВ «Нілан-ЛТД», 2015. — 222 с.
ISBN 978-966-2770-19-3

Навчальний посібник охоплює теоретичний матеріал з лінійної алгебри за програмами підготовки бакалаврів математики фізико-математичних факультетів педагогічних університетів. Виклад традиційного матеріалу з лінійної алгебри подано у відповідності до сучасних підходів у вивченні цієї дисципліни. Систематизовані теоретичні відомості наведено з доведеннями і супроводжуються прикладами, рисунками, схемами та завданнями для самостійної роботи.

Посібник призначений для студентів педагогічних університетів.

УДК 512.64 (075.8)
ББК 22.143

ISBN 978-966-2770-19-3

Зміст

Передмова

> То, что хотел бы я высказать, высказыванию не подлежит,
> ибо вот то, что я высказать хотел бы, оно таково,
> что, когда его всё же высказать пытаешься, оно бежит,
> а когда не пытаешься, ввек не избавишься от него...
>
> *М. Щербаков*

Лінійна алгебра є розділом математики, що має численні застосування в різних галузях людських знань, а тому займає вагоме місце в підготовці майбутніх учителів математики. Даний навчальний посібник допоможе студентам педагогічних університетів в опануванні теоретичних основ цього розділу математики.

Лінійній алгебрі присвячено чимало підручників, посібників і монографій (перелік рекомендованої літератури подано в кінці цього посібника). Разом з цим, як показав досвід викладання цієї дисципліни, для нинішніх студентів назріли потреби уточнення і узагальнення в доступній формі теоретичних основ лінійної алгебри. Це і було враховано при написанні цього посібника, мета якого — охопити весь обов'язковий теоретичний матеріал з курсу лінійної алгебри, окреслити наукові і практичні рекомендації застосування положень лінійної алгебри при розв'язанні задач. Реалізація мети зумовила необхідність розв'язання таких завдань: вивчення зарубіжного і вітчизняного досвіду викладання лінійної алгебри у вищих навчальних закладах, дослідження структури лінійної алгебри, уточнення її понятійно-термінологічного апарату.

Матеріал посібника структуровано і викладено на двох рівнях: базовому та розширеному. Базовий рівень містить означення, основні теореми з повними доведеннями та коментарями до них, розв'язання типових задач, застосування понять і методів лінійної алгебри. Розширений рівень доповнює базовий додатковими твердженнями, повними доведеннями складних, але важливих теорем. Вони наведені у тексті дрібним шрифтом. Наприкінці кожного розділу наведено додаткові відомості і завдання для самостійної роботи, доведення яких пропонується читачам виконати самостійно з використанням допоміжної літератури.

Структура посібника є такою: 28 параграфів об'єднано у 8 розділів. Параграфи можуть розбиватись на кілька пунктів, які позначаються подвійною нумерацією. На початку посібника є перелік основних умовних позначень лінійної алгебри, які фігурують по тексту, а в кінці — предметний покажчик і перелік рекомендованої літератури. Знаком ■ позначається кінець доведення твердження (теореми, леми) або розв'язання прикладу.

В друге видання посібника додано нові пункти «Ортогональна проекція вектора на підпростір» та «Самоспряжений лінійний оператор».

Автор сподівається, що цей навчальний посібник буде корисним як для студентів очної, заочної форм навчання, що навчаються у педагогічних університетах, так і для студентів вузів іншого спрямування, які вивчають лінійну алгебру.

Основні позначення

> Начинающий $<\ldots>$ не должен смущаться, если $<\ldots>$ он обнаружит, что у него не хватает предварительных знаний даже для чтения предварительных сведений.
>
> *П. Халмош*

$\mathbb{N}$	множина натуральних чисел
$\mathbb{R}$	множина дійсних чисел
$\mathbb{C}$	множина комплексних чисел
$\mathbb{R}^n$	n-вимірний арифметичний векторний простір
V_n	n-вимірний векторний простір
$\dim V$	розмірність векторного простору V
$\vec{0}_V$	нуль-вектор векторного простору V
$M_{m\times n}(\mathbb{R})$, $M_{m\times n}$	множина матриць розмірності $m \times n$ з дійсними коефіцієнтами
$M_n(\mathbb{R})$, M_n	множина квадратних матриць n-го порядку з дійсними коефіцієнтами
$GL_n(\mathbb{R})$	група оборотних квадратних матриць n-го порядку з дійсними коефіцієнтами з операцією матричного множення (повна лінійна група)
$\operatorname{rank}(A), r(A)$	ранг матриці A
A^T	транспонована матриця A
A^{-1}	матриця, обернена до A
$[A]_{ij}$	елемент i-го рядка і j-го стовпця матриці A
$\operatorname{row}(A)$	простір рядків матриці A
$\operatorname{col}(A)$	простір стовпців матриці A
$\operatorname{null}(A)$	нуль-простір матриці A
$M_{ij}(A)$	мінор елемента $[A]_{ij}$ матриці A
$A_{ij}(A)$	алгебраїчне доповнення елемента $[A]_{ij}$ матриці A

$\det A$, $\lvert A\rvert$	детермінант матриці A
I_n	одинична матриця n-го порядку
I	одинична матриця, розмірність якої зрозуміла з контексту
O	нульова матриця
$\big(\vec{a}_1, \ldots, \vec{a}_n\big)$	впорядкована система векторів
$\{\vec{a}_1, \ldots, \vec{a}_n\}$	невпорядкована система векторів
$[\vec{x}]_{\mathcal{B}}$	координати вектора $\vec{x}$ в базисі $\mathcal{B}$
$P_{\mathcal{C}\leftarrow\mathcal{B}}$	матриця переходу від базису $\mathcal{B}$ до базису $\mathcal{C}$
$\mathcal{P}_n, \mathbb{R}_n[x]$	векторний простір многочленів від змінної x степеня не вище n
$\mathcal{C}_{[a,b]}$	векторний простір неперервних на відрізку $[a,b]$ функцій
$[T]$	стандартна матриця лінійного відображення T
$[T]_{\mathcal{C}\leftarrow\mathcal{B}}$	матриця лінійного відображення T відносно базисів $\mathcal{B}$ та $\mathcal{C}$
$\operatorname{Ker} T$	ядро лінійного відображення T
$\operatorname{Im} T$	образ лінійного відображення T
$L\,(\vec{a}_1, \ldots, \vec{a}_n)$	лінійна оболонка системи векторів $\big(\vec{a}_1, \ldots, \vec{a}_n\big)$
$\operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$	діагональна матриця $[a_{ij}]$ з $a_{ii} = \lambda_i$
$\big(\vec{x}, \vec{y}\big)$	скалярний добуток векторів $\vec{x}$ та $\vec{y}$
T^*	оператор, спряжений до оператора T
$\lVert\vec{a}\rVert$	норма (довжина) вектора $\vec{a}$
$U^\perp$	ортогональне доповнення до підпростору U
$\vec{a} \perp \vec{b}$	ортогональність векторів $\vec{a}$ та $\vec{b}$
$\operatorname{pr}_U \vec{x}$	ортогональна проекція вектора $\vec{x}$ на підпростір U
$\operatorname{ort}_U \vec{x}$	ортогональна складова вектора $\vec{x}$ відносно підпростору U

Розділ 1

Системи лінійних рівнянь. Арифметичний векторний простір.

— ... Вот об этом я бы поговорил с удовольствием, но вы ничего не поймёте.
— Почему же это я ничего не пойму? Я всё-таки инженер...
— Потому что я сам не понимаю, — сказал Валентин. — У меня есть системы уравнений, но как их истолковать, я представления не имею...

Аркадий и Борис Стругацкие, «Пикник на обочине»

§ 1. Системи лінійних рівнянь та методи їх розв'язування.

Системи лінійних рівнянь лежать в основі лінійної алгебри. В цьому параграфі ми розглянемо основний метод їх розв'язування — метод Гауса.

1.1. Поняття про системи лінійних рівнянь.

Означення 1.1. *Лінійним рівнянням з* n *невідомими* $x_1, x_2, \ldots, x_n$ *над полем* P називається рівняння, яке може бути записане у вигляді

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b, \tag{1.1}$$

де $a_1, a_2, \ldots, a_n, b$ — елементи поля[1] P (зазвичай відомі).

Приклад 1.1. Рівняння $4x_1 - 5x_2 + 2 = x_1$ і $x_2 = 2(\sqrt{6} - x_1) + x_3$ є лінійними, хоча і не записані у вигляді (1.1), а рівняння $x_1x_2 = 6 + x_1 + x_2$ не є лінійним.

[1] Передбачається, що читачі знайомі з поняттями група, кільце, поле (див., наприклад [4]).

Означення 1.2. *Системою лінійних рівнянь* називається набір з одного чи більше лінійних рівнянь відносно однієї і тієї ж множини змінних (наприклад, $x_1, x_2, \ldots, x_n$).

В загальному вигляді система лінійних рівнянь, що містить m рівнянь з n невідомими, має вигляд

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{1n}x_n & = & b_2, \\ \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m. \end{cases} \tag{1.2}$$

Означення 1.3. *Розв'язком* системи лінійних рівнянь (1.2) називається впорядкована множина чисел $(s_1, s_2, \ldots, s_n)$, яка при підстановці в дану систему відповідно замість $x_1, x_2, \ldots, x_n$ перетворює кожне рівняння системи на правильну числову рівність.

Наприклад, пара чисел $(3, -1)$ є розв'язком системи

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

Якщо існує хоча б один розв'язок системи (1.2), то система називається *сумісною*, в іншому випадку — *несумісною*. Сумісна система називається *визначеною*, якщо вона має єдиний розв'язок; система, що має більше одного розв'язків називається *невизначеною*.

Розв'язати систему — означає знайти множину її розв'язків.

Система лінійних рівнянь

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0, \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{1n}x_n & = & 0, \\ \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & 0 \end{cases}$$

називається *однорідною*. Однорідна система завжди сумісна (оскільки $(0, 0, \ldots, 0)$ є її розв'язком). Разом з цим вона може бути як визначеною, так і невизначеною. Якщо $(0, 0, \ldots, 0)$ не є розв'язком системи лінійних рівнянь, то вона називається *неоднорідною*.

Розв'язування систем лінійних рівнянь з двома невідомими вивчаються в шкільному курсі алгебри (методи підстановки і додавання). Для дослідження питання сумісності системи можна використати графічний метод.

Приклад 1.2. Дослідити на сумісність систему
$$\begin{cases} x - 2y = -1, \\ -x + 3y = 3. \end{cases}$$

Розглянемо графіки обох рівнянь системи в прямокутній декартовій системі координат. Добре відомо, що це прямі, які перетинаються (рис. 1.1). Отже, система сумісна, причому $(3, 2)$ — розв'язок системи.

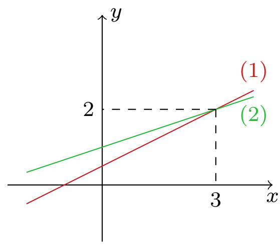


Рис. 1.1. Прямі, задані рівняннями $x - 2y = -1$ та $-x + 3y = 3$.

Означення 1.4. Дві системи називаються *еквівалентними (рівносильними)*, якщо множини їхніх розв'язків співпадають, тобто кожний розв'язок першої системи є розв'язком другої і, навпаки, кожний розв'язок другої системи є розв'язком першої.

Основна ідея при розв'язуванні систем лінійних рівнянь — заміна даної системи на еквівалентну до даної, яку розв'язати з якихось міркувань простіше.

Наступні перетворення системи лінійних рівнянь називаються *елементарними*:

1) Перестановка двох рівнянь у системі.
2) Множення будь-якого рівняння системи на відмінний від нуля елемент поля P.
3) Додавання до будь-якого рівняння системи будь-якого іншого рівняння, помноженого на довільний елемент поля P.
4) Вилучення із системи рівняння виду $0 = 0$ або приписування до системи такого рівняння.

Теорема 1.1. *Елементарне перетворення кожної системи лінійних рівнянь переводить її у еквівалентну систему.*

Доведення. Нехай система S_2 одержується з системи S_1 шляхом певного елементарного перетворення. Покажемо, що кожний розв'язок системи S_1 є розв'язком системи S_2 і навпаки. Для перетворень 1, 2, 4 це очевидно. Доведемо це для третього елементарного перетворення.

Але спочатку відмітимо, що елементарні перетворення систем лінійних рівнянь оборотні. Це означає, що якщо ми з деякої системи S_1 шляхом елементарного перетворення отримали систему S_2, то існує елементарне перетворення, що переводить систему S_2 назад в систему S_1.

Нехай тепер S_1 має вигляд (1.2), а S_2 відрізняється від S_1 лише i-м рівнянням, що має вигляд

$$(a_{i1} + ca_{j1})x_1 + (a_{i2} + ca_{j2})x_2 + \ldots + (a_{in} + ca_{jn})x_n = b_i + cb_j.$$

Тоді якщо $(s_1, \ldots, s_n)$ — розв'язок системи S_1, тобто $a_{k1}s_1 + a_{k2}s_2 + \ldots + a_{kn}s_n = b_k$ $(k = \overline{1,m})$, то $(s_1, \ldots, s_n)$ задовольняє і останню рівність, а значить є розв'язком системи S_2. Оскільки елементарні перетворення оборотні, то і, навпаки, кожний розв'язок системи S_2 є розв'язком системи S_1, тобто системи S_1 і S_2 еквівалентні. ∎

Нехай маємо систему лінійних рівнянь (1.2). Випишемо коефіцієнти цієї системи в таку таблицю, яку називатимемо *матрицею коефіцієнтів* цієї системи (*головною матрицею*):

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}.$$

Розширена матриця системи відрізняється від головної тим, що додатково виписаний також і стовпець вільних членів. Стовпець вільних членів прийнято відокремлювати від головної матриці вертикальною рискою:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

Приклад 1.3. Розв'язати систему:

$$\begin{cases} x - y - z = 2, \\ \quad y + 3z = 5, \\ \qquad 5z = 10. \end{cases}$$

Розв'язання. З останнього рівняння однозначно знаходимо, що $z = 2$. Тоді з другого знаходимо, що $y = 5 - 3z = -1$, а з першого $x = 2 + y + z = 3$.

Відповідь. $\{(3, -1, 2)\}$. ■

Приклад 1.4. Розв'язати систему:

$$\begin{cases} x - y - z = 2, \\ 3x - 3y + 2z = 16, \\ 2x - y + z = 9. \end{cases}$$

Розв'язання. Спробуємо за допомогою елементарних перетворень звести дану систему до так званого «трикутного» виду, тобто такого виду, як в попередньому прикладі.

$$\begin{cases} x - y - z = 2, \\ 3x - 3y + 2z = 16, \\ 2x - y + z = 9. \end{cases} \xrightarrow[\text{IIIр.}-2\cdot\text{Iр.}]{\text{IIр.}-3\cdot\text{Iр.}} \begin{cases} x - y - z = 2, \\ \qquad 5z = 10, \\ \quad y + 3z = 5. \end{cases} \xrightarrow{\text{IIр.}\leftrightarrow\text{IIIр.}}$$

$$\begin{cases} x - y - z = 2, \\ \quad y + 3z = 5, \\ \qquad 5z = 10, \end{cases}$$

тобто дістали систему лінійних рівнянь з попереднього прикладу.

Відповідь. $\{(3, -1, 2)\}$. ■

1.2. Східчаста форма матриці. Метод Гауса розв'язування систем лінійних рівнянь.

Означення 1.5. Говорять, що матриця має *східчасту форму* (або *рядково східчасту форму*), якщо вона володіє такими властивостями:

1) усі ненульові рядки знаходяться вище всіх нульових рядків;
2) в кожному ненульовому рядку перший ненульовий елемент (він називається *ведучим* елементом рядка) знаходиться в стовпці справа від ведучого елемента у рядку над ним.

Приклад 1.5. Наступні матриці мають східчасту форму:

$$\begin{bmatrix} \boxed{2} & 4 & 1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \boxed{1} & 4 & 1 & 1 \\ 0 & 0 & \boxed{-1} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \boxed{\star} & \star & \star & \star & \star & \star \\ 0 & 0 & 0 & \boxed{\star} & \star & \star & \star \\ 0 & 0 & 0 & 0 & \boxed{\star} & \star & \star \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{\star} \end{bmatrix}.$$

(тут ведучі елементи кожного рядка обведені в рамку; $\star$ — довільний елемент, $\boxed{\star}$ — довільний ненульовий елемент).

Елементарними перетвореннями рядків матриці називатимемо такі:

1) Перестановка двох рядків місцями.
2) Множення рядка матриці на ненульовий елемент поля P.
3) Додавання до одного рядка іншого, помноженого на певне число.
4) Відкидання від матриці нульового рядка (якщо він є) або приписування до матриці нульовий рядок.

Якщо над матрицею виконується елементарне перетворення, то це буде позначатися відповідним чином справа від матриці. Наприклад, запис

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{array}{l} \mid \cdot 7 \leftarrow \\ \leftarrow \quad \cdot 5 \\ \leftarrow \quad + \end{array}$$

означає, що перший рядок множиться на 7, переставляється місцями з другим рядком, в той час як до третього рядка додаються 5 других.

Дві матриці називаються *рядково еквівалентними*, якщо існує послідовність елементарних рядкових операцій, які переводять одну матрицю в іншу.

Кожну матрицю шляхом елементарних перетворень можна звести до східчастої форми (можливо, не єдиним способом).

Приклад 1.6. Звести до східчастої форми матрицю

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix}$$

Алгоритм наступний. Працюємо зверху вниз, зліва направо. Спочатку помічаємо, що елемент першого рядка і першого стовпця матриці дорівнює 1, що є зручним для того, щоб в першому стовпці з допомогою елементарного перетворення «заміщення» (третє елементарне перетворення) утворити нулі в інших позиціях. Отже, першим кроком обираємо «опорний елемент» — той, який стане ведучим у східчастій формі матриці і утворюємо нулі під цим опорним елементом.

$$\begin{bmatrix} \boxed{1} & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix} \begin{array}{lll} \cdot(-2) & \cdot(-2) & \\ + & & \\ & + & \\ & & + \end{array} \sim$$

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 3 & -1 & 2 & 10 \end{bmatrix} \begin{array}{l} \\ \mid : 8 \\ \\ \\ \end{array}$$

Тепер помічаємо, що, по-перше, усі елементи другого рядка діляться на 8, і, по-друге, в цьому рядку на другій позиції знаходиться 0, тоді як в третьому і четвертому рядках — відмінні від нуля елементи. Тому далі виконаємо ділення елементів другого рядка на 8 і поміняємо його місцем з третім рядком, ведучим елементом якого є -1. Тепер обираємо опорний елемент (-1 в другому рядку) і утворюємо нулі в усіх позиціях під ним.

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & \boxed{-1} & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 3 & -1 & 2 & 10 \end{bmatrix} \begin{array}{l} \\ \cdot 3 \\ \\ + \end{array} \sim \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 29 & 29 & -5 \end{bmatrix}$$

Переходимо до третього рядка. Опорний елемент дорівнює 1. Віднімемо від четвертого рядка 29 третіх і остаточно отримаємо:

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}.$$

Бачимо, що одержана матриця у східчастій формі і вона рядково еквівалентна до вихідної матриці.

Метод Гауса[1] розв'язування систем лінійних рівнянь (метод послідовного виключення невідомих) полягає в наступному:

1. Виписати розширену матрицю системи лінійних рівнянь.
2. За допомогою елементарних перетворень рядків матриці звести її до східчастої форми.
3. Повернутись від одержаної матриці до системи та розв'язати її.

Зауважимо, що звести матрицю до східчастої форми можна різними способами. Відзначимо деякі поради того, як цього можна досягнути:

- Обрати крайній зліва стовпець, не всі елементи якого дорівнюють нулю. Створити ведучий елемент першого рядка (він повинен бути ненульовим, але бажано, якщо ведучий елемент дорівнює ± 1), використовуючи елементарне перетворення «перестановка рядків».
- Використовуючи цей ведучий елемент, створити в стовпці під ним нулі. Цього завжди можна досягти, додаючи до кожного рядка перший рядок, що помножений на відповідний коефіцієнт.
- Повторити описані вище дії для підматриці, яка одержується з даної після викреслення першого рядка і першого (ненульового) стовпця.

Приклад 1.7. Розв'язати систему:

$$\begin{cases} \phantom{2x_1+{}} 2x_2 + 3x_3 = 8, \\ 2x_1 + 3x_2 + x_3 = 5, \\ x_1 - x_2 - 2x_3 = -5. \end{cases}$$

[1] Johann Carl Friedrich Gauss (1777–1855) — видатний німецький математик.

Розв'язання. Розширена матриця системи має вигляд

$$\left[\begin{array}{ccc|c} 0 & 2 & 3 & 8 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right].$$

Зведемо її до східчастої форми. Перший ненульовий стовпець — крайній зліва. Переставивши місцями перший і третій рядки матриці, ми досягнемо того, що ведучий елемент першого рядка буде рівним 1, що є зручним для подальших обчислень.

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -5 \\ 2 & 3 & 1 & 5 \\ 0 & 2 & 3 & 8 \end{array}\right].$$

Далі утворюємо нулі під ведучим елементом. Для цього до другого рядка додаємо перший, що помножений на -2:

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -5 \\ 0 & 5 & 5 & 15 \\ 0 & 2 & 3 & 8 \end{array}\right].$$

Тепер «забуваємо» про перший рядок і повторюємо описані дії для одержаної підматриці. Елементи другого рядка помножимо на $\frac{1}{5}$.

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 2 & 3 & 8 \end{array}\right].$$

Утворимо нуль на другій позиції третього рядка (додаємо до третього рядка другий, помножений на -2):

$$\left[\begin{array}{ccc|c} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right].$$

Одержали матрицю у східчастій формі. Тепер повертаємось до системи

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

і розв'язуємо її «знизу вгору»: $x_3 = 2$, $x_2 = 3 - x_3 = 1$, $x_1 = -5 + x_2 + 2x_3 = 0$.

Відповідь. $\{(0, 1, 2)\}$. ■

Приклад 1.8. Розв'язати систему:

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

Розв'язання. Випишемо розширену матрицю системи і зведемо її до східчастої форми.

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right] \sim$$

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Система, що відповідає останній матриці, має вигляд

$$\begin{cases} x_1 - x_2 - x_3 + 2x_4 = 1, \\ \qquad\qquad x_3 - \ x_4 = 1. \end{cases}$$

Ця система має безліч розв'язків. Існує декілька способів введення параметрів для опису загального розв'язку цієї системи. Ми будемо використовувати такий: вираження змінних, що відповідають ведучим елементам рядків східчастої форми матриці (ці змінні називаються *основними*) через інші змінні (вони називаються *вільними*). В нашому прикладі основними є змінні x_1 та x_3, а вільними — x_2 та x_4.

Маємо: $x_3 = x_4 + 1$, $x_1 = x_2 + x_3 - 2x_4 + 1 = x_2 - x_4 + 2$. Для запису відповіді введемо параметри $a = x_2, b = x_4$.

Відповідь. $\{(a - b + 2, a, b + 1, b) \colon a, b \in \mathbb{R}\}$. ■

Відзначимо також, що якщо останній ненульовий рядок східчастої форми матриці містить лише один відмінний від нуля елемент (останній), то система є несумісною, а в іншому випадку — сумісною. Якщо кількість рядків в східчастій формі матриці співпадає з кількістю невідомих сумісної системи, то вона має єдиний розв'язок; якщо ж їх менше — то безліч.

1.3. Зведена східчаста форма матриці і розв'язування систем лінійних рівнянь методом Йордана–Гауса.

Означення 1.6. Говорять, що матриця має *зведену східчасту форму*, якщо виконується умови:

1) вона має східчасту форму;
2) ведучі елементи кожного ненульового рядка дорівнюють 1;
3) кожний стовпець, що містить ведучу одиницю якогось рядка, містить в усіх інших позиціях нулі.

Приклад 1.9. Наступні матриці мають зведену східчасту форму:

$$\begin{bmatrix} \boxed{1} & 2 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}, \quad \begin{bmatrix} 0 & \boxed{1} & \star & 0 & 0 & \star & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 & \star & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

(тут $\star$ — довільні числа, $\boxed{1}$ — ведучий елемент відповідного рядка).

Виявляється, що кожна матриця рядково еквівалентна лише одній матриці у зведеній східчастій формі.

Метод Йордана[1]–Гауса розв'язування систем лінійних рівнянь полягає в наступному:

1. Записуємо розширену матрицю системи лінійних рівнянь.
2. Зводимо матрицю до східчастої форми і встановлюємо, чи сумісна система. Якщо сумісна, то переходимо до наступного пункту.
3. Утворюючі нулі в стовпцях з ведучими елементами рядків матриці, приводимо її до зведеної східчастої форми.
4. Повертаємось від матриці до системи. Ті змінні, що відповідають ведучим одиницям, назвемо *основними*, усі інші (якщо вони є) — *вільними*. Виражаємо основні змінні через вільні і записуємо відповідь.

Розв'яжемо систему з прикладу 1.8 методом Йордана–Гауса. Як було показано вище, східчастою формою матриці цієї системи є

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Відзначаємо, що система сумісна, оскільки ми не отримали рядка виду $\left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 1 \end{array}\right]$. Приведемо її до зведеної східчастої форми. Для цього

[1] Wilhelm Jordan (1842–1899) — німецький геодезіст.

слід утворити нулі в першій і третій позиціях третього стовпця. Цього можна досягти, додавши до першого рядка другий:

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Одержали матрицю у зведеній східчастій формі. Система, що їй відповідає має вигляд:

$$\begin{cases} x_1 - x_2 \quad + x_4 = 2, \\ \qquad\quad x_3 - x_4 = 1. \end{cases}$$

Змінні x_1 та x_3 є основними, а x_2 та x_4 — вільними. Виражаємо основні змінні через вільні:

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

і виписуємо остаточну відповідь: $\{(a - b + 2, a, b + 1, b) \colon a, b \in \mathbb{R}\}$.

§ 2. n-ВИМІРНИЙ АРИФМЕТИЧНИЙ ВЕКТОРНИЙ ПРОСТІР

При розв'язуванні систем лінійних рівнянь виникає поняття впорядкованого набору з n чисел — розв'язок системи рівнянь з n невідомими. Якщо на множині всіх таких впорядкованих наборів чисел ввести операції додавання і множення на число, то дістанемо алгебраїчну структуру, яка називається n-вимірним арифметичним векторним простором.

2.1. Поняття n-вимірного арифметичного векторного простору.

Нехай n — натуральне.

Означення 1.7. *n-вимірним вектором* над полем P назвемо впорядкований набір з n елементів цього поля.

Зазвичай впорядкований набір чисел $x_1, x_2, \ldots, x_n$ позначається одним з двох наступних способів:

$$(x_1, x_2, \ldots, x_n) \quad \text{або} \quad \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

n-вимірний вектор позначають малими латинськими літерами зі стрілкою зверху або грубим шрифтом: $\vec{x}$ або $\boldsymbol{x}$. Ми будемо використовувати перше позначення.

Зауважимо, що часто для спрощення обґрунтування окремих фактів використовується тлумачення n-вимірного вектора як матриці, що містить n рядків та один стовпець.

Два n-вимірних вектори $\vec{x}$ та $\vec{y}$ називаються *рівними* (позначається $\vec{x} = \vec{y}$), якщо їх відповідні компоненти однакові: нехай $\vec{x} = (x_1, \ldots, x_n)$, $\vec{y} = (y_1, \ldots, y_n)$, тоді $\vec{x} = \vec{y}$ тоді і тільки тоді, коли $x_1 = y_1 \wedge \ldots \wedge x_n = y_n$.

Вектор $(0, 0, \ldots, 0)$ називається *нульовим* і позначається $\vec{0}$.

Множину всіх n-вимірних векторів заданих над полем P позначають P^n. Якщо в тексті не зроблено спеціальних зауважень, ми надалі розглядатимемо випадок, коли $P = \mathbb{R}$.

На множині $\mathbb{R}^n$ всіх n-вимірних векторів, елементами яких є дійсні числа, введемо дві бінарні алгебраїчні операції: внутрішню «додавання» і зовнішню «множення на число».

Означення 1.8. *Сумою* двох n-вимірних векторів $\vec{x} = (x_1, \ldots, x_n)$ та $\vec{y} = (y_1, \ldots, y_n)$ називається вектор $\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n)$.

Означення 1.9. *Добутком* вектора $\vec{x} = (x_1, x_2, \ldots, x_n)$ на число $c \in \mathbb{R}$ називається вектор $c\vec{x} = (cx_1, cx_2, \ldots, cx_n)$.

Серед властивостей вказаних операцій додавання та множення на число відмітимо такі:

1) $\forall \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$: $\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$ (асоціативність);
2) $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$: $\vec{x} + \vec{y} = \vec{y} + \vec{x}$ (комутативність);
3) $\forall \vec{x} \in \mathbb{R}^n$: $\vec{x} + \vec{0} = \vec{x}$;
4) $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$, $\forall c \in \mathbb{R}$: $c(\vec{x} + \vec{y}) = c\vec{x} + c\vec{y}$ (дистрибутивність);
5) $\forall \vec{x} \in \mathbb{R}^n$, $\forall c, d \in \mathbb{R}$: $(c + d)\vec{x} = c\vec{x} + d\vec{x}$ (дистрибутивність);
6) $\forall \vec{x} \in \mathbb{R}^n$, $\forall c, d \in \mathbb{R}$: $(cd)\vec{x} = c(d\vec{x})$;
7) $\forall \vec{x} \in \mathbb{R}^n$: $1 \cdot \vec{x} = \vec{x}$;
8) $\forall \vec{x} \in \mathbb{R}^n$: $0 \cdot \vec{x} = \vec{0}$.

Означення 1.10. Множина усіх n-вимірних векторів з дійсними компонентами, що розглядається з визначеними в ній операціями додавання і множення на число, називається *n-вимірним арифметичним векторним простором.*

Геометричною інтерпретацією просторів $\mathbb{R}^2$ ($\mathbb{R}^3$) є множина всіх векторів площини (простору) з початком в точці O з операціями додавання і множення на число, означеними в стандартний спосіб, який відомий ще зі шкільної геометрії.

2.2. Лінійна комбінація і лінійна оболонка системи векторів.

Далі ми розглядатимемо системи векторів — це скінченна сукупність векторів, в якій, взагалі кажучи, чітко визначено місце кожного вектора.

Означення 1.11. Нехай $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right)$ — система векторів простору $\mathbb{R}^n$. Говорять, що вектор $\vec{b}$ є *лінійною комбінацією* системи векторів $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right)$ (вектор $\vec{b}$ *лінійно виражається* через $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$), якщо існують дійсні $\alpha_1, \alpha_2, \ldots, \alpha_k$ такі, що виконується рівність:

$$\vec{b} = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_k\vec{a}_k.$$

Позначатимемо це так: $\vec{b} \xrightarrow{\text{л.в.}} \vec{a}_1, \ldots, \vec{a}_k$.

Означення 1.12. Множина всіх можливих лінійних комбінацій векторів $\vec{a}_1, \ldots, \vec{a}_k$ простору $\mathbb{R}^n$ називається *лінійною оболонкою* цих векторів і позначається $L\left(\vec{a}_1, \ldots, \vec{a}_k\right)$:

$$L\left(\vec{a}_1, \ldots, \vec{a}_k\right) = \left\{\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k \colon \alpha_i \in \mathbb{R}, i = \overline{1, k}\right\}.$$

Приклад 1.10. Розглянемо вектори $\vec{u} = \begin{bmatrix}-1\\2\end{bmatrix}$ і $\vec{v} = \begin{bmatrix}2\\3\end{bmatrix}$. Ці два вектори можна взяти в основі нової координатної сітки на площині (див. рис. 1.2). При цьому кожний вектор $\vec{w} \in \mathbb{R}^2$ може бути поданий у вигляді лінійної комбінації векторів $\vec{u}$ та $\vec{v}$. Наприклад, як видно з рисунку, вектор $\begin{bmatrix}5\\-3\end{bmatrix}$ є сумою векторів $-3\vec{u}$ та $\vec{v}$. З іншого боку, це підтверджується безпосереднім обчисленням:

$$-3\vec{u} + \vec{v} = -3\begin{bmatrix}-1\\2\end{bmatrix} + \begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}5\\-3\end{bmatrix}.$$

Перехід від однієї системи координат до іншої є зручним інструментом при розв'язуванні багатьох задач.

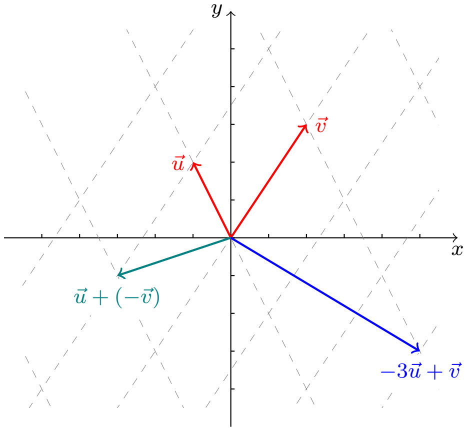


Рис. 1.2. Нова координатна сітка, що породжується векторами $\vec{u}=(-1,2)$ та $\vec{v}=(2,3)$ в просторі $\mathbb{R}^2$.

Приклад 1.11. Нехай $\vec{x}$ — ненульовий вектор простору $\mathbb{R}^3$. Тоді з геометричної точки зору $L(\vec{x})$ — множина векторів на прямій, що містить вектор $\vec{x}$:

$$L(\vec{x})=\{\alpha\vec{x}\colon \alpha\in\mathbb{R}\}.$$

Приклад 1.12. Нехай $\vec{x},\vec{y}$ — ненульові неколінеарні вектори простору $\mathbb{R}^3$. Тоді з геометричної точки зору $L(\vec{x},\vec{y})$ — множина векторів площини, що містить вектори $\vec{x}$ та $\vec{y}$ і проходить через початок координат:

$$L(\vec{x},\vec{y})=\{\alpha_1\vec{x}+\alpha\vec{y}\colon \alpha_1,\alpha_2\in\mathbb{R}\}.$$

Приклад 1.13. Нехай $\vec{a}_1=\begin{bmatrix}1\\-2\\-5\end{bmatrix}$, $\vec{a}_2=\begin{bmatrix}2\\5\\6\end{bmatrix}$, $\vec{b}=\begin{bmatrix}7\\4\\-3\end{bmatrix}$. Довести, що вектор $\vec{b}$ є лінійною комбінацією векторів $\vec{a}_1$ і $\vec{a}_2$ і знайти відповідне подання.

Розв'язання. Припустимо, що вектор $\vec{b}$ є лінійною комбінацією векторів $\vec{a}_1$ і $\vec{a}_2$. Тоді повинні існувати такі коефіцієнти x та y, щоб виконувалась рівність

$$\vec{b}=x\vec{a}_1+y\vec{a}_2.$$

Маємо:

$$\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} = x \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix},$$

звідки з означень добутку вектора на число, суми векторів та їх рівності знаходимо:

$$\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} = \begin{bmatrix} x + 2y \\ -2x + 5y \\ -5x + 6y \end{bmatrix},$$

що рівносильно системі

$$\begin{cases} x+2y = 7, \\ -2x+5y = 4, \\ -5x+6y = -3. \end{cases}$$

Розв'язуємо цю систему методом Гауса:

$$\left[\begin{array}{cc|c} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{array}\right] \begin{array}{l} \cdot 2 \quad \cdot 5 \\ + \\ + \end{array} \sim \left[\begin{array}{cc|c} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{array}\right] \begin{array}{l} \\ \mid \cdot \frac{1}{9} \\ \mid \cdot \frac{1}{16} \end{array} \sim$$

$$\left[\begin{array}{cc|c} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{array}\right].$$

Система, що відповідає останній матриці, вочевидь, сумісна і визначена, а її розв'язком є пара чисел $(3, 2)$. Таким чином, $\vec{b} = 3\vec{a}_1 + \vec{a}_2$. ■

Приклад 1.14. Довести, що $L\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = \mathbb{R}^2$.

Розв'язання. Треба довести, що кожний вектор з простору $\mathbb{R}^2$ можна подати у вигляді лінійної комбінації векторів $\vec{a}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Нехай $\begin{bmatrix} a \\ b \end{bmatrix}$ — довільний вектор з $\mathbb{R}^2$. Спробуємо подати цей вектор у вигляді лінійної комбінації векторів $\vec{a}_1$, $\vec{a}_2$:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Остання рівність рівносильна системі:

$$\begin{cases} 2x + y = a, \\ -x + 3y = b. \end{cases}$$

Додавши до першого рівняння два других, знаходимо $y = \frac{a+2b}{7}$ і $x = \frac{3a-b}{7}$. Отже, якими б не були a і b, справджується рівність

$$\frac{3a-b}{7}\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \frac{a+2b}{7}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

■

Означення 1.13. Система векторів $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ називається *лінійно залежною*, якщо існують такі числа $\alpha_1, \ldots, \alpha_k$, серед яких хоча б одне відмінне від нуля, що виконується рівність

$$\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k = \vec{0}. \qquad (\star)$$

Якщо ж рівність $(\star)$ виконується лише при $\alpha_1 = \ldots = \alpha_k = 0$, то система векторів називається *лінійно незалежною*.

Приклад 1.15. Дослідити, лінійно залежною чи лінійно незалежною є система векторів:

а) $\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{a}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$;

б) $\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{a}_3 = \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix}$.

Розв'язання. а) Складемо рівність $(\star)$ для векторів $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$.

$$\alpha_1\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \alpha_2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Остання рівність приводить до системи

$$\begin{cases} \alpha_1 + \alpha_2 = 0, \\ \alpha_3 = 0, \\ 2\alpha_1 = 0, \end{cases}$$

яка, очевидно, має лише нульовий розв'язок. Таким чином, система векторів $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ є лінійно незалежною.

б) Легко перевірити, що має місце рівність

$$3\vec{a}_1 - \vec{a}_2 - \vec{a}_3 = \vec{0}.$$

Таким чином, система векторів $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$ є лінійно залежною.

Зауважимо, що цю задачу можна розв'язати, складаючи рівність $(\star)$ (аналогічно до того, як це було зроблено в п. а)). В цьому випадку одержана система була б невизначеною (мала б безліч розв'язків), з чого робимо висновок про лінійну залежність досліджуваної системи векторів. ■

Теорема 1.2. *Для того, щоб система векторів* $\big(\vec{a}_1, \ldots, \vec{a}_k\big)$ *була лінійно залежною, необхідно і достатньо, щоб хоча б один вектор цієї системи лінійно виражався через інші.*

Доведення. *Необхідність.* Нехай $\big(\vec{a}_1, \ldots, \vec{a}_k\big)$ — деяка лінійно залежна система векторів. Це означає, що рівність $(\star)$, складена для цих векторів, виконується при деякому $c_i \neq 0$:

$$c_1\vec{a}_1 + \ldots + c_i\vec{a}_i + \ldots + c_k\vec{a}_k = \vec{0}, \quad c_i \neq 0.$$

Розділивши обидві частини останньої рівності на c_i, дістанемо:

$$\vec{a}_i = -\frac{c_1}{c_i}\vec{a}_1 - \ldots - \frac{c_{i-1}}{c_i}\vec{a}_{i-1} - \frac{c_{i+1}}{c_i}\vec{a}_{i+1} - \ldots - \frac{c_k}{c_i}\vec{a}_k,$$

тобто вектор $\vec{a}_i$ є лінійною комбінацією інших векторів.

Достатність. Припустимо, що деякий вектор системи векторів є лінійною комбінацією інших векторів: $\vec{a}_i = c_{i_1}\vec{a}_{i_1} + \ldots + c_{i_l}\vec{a}_{i_l}$. Тоді, записавши різницю між лівою і правою частинами останньої рівності, і доповнивши її доданками виду $0 \cdot \vec{a}_j$, складемо правильну рівність $(\star)$ для системи векторів $\big(\vec{a}_1, \ldots, \vec{a}_k\big)$, яка виконується принаймні при одному коефіцієнту відмінному від нуля (коефіцієнт при $\vec{a}_i$ дорівнює 1). ■

Очевидно, що система, яка складається з одного вектора, є лінійно залежною тоді і тільки тоді, коли цей вектор — $\vec{0}$.

Теорема 1.3. *Якщо система векторів* $\big(\vec{a}_1, \ldots, \vec{a}_k\big)$ *містить лінійно залежну підсистему* $\big(\vec{a}_{i_1}, \ldots, \vec{a}_{i_l}\big)$, *то вихідна система лінійно залежна.*

Доведення пропонуємо читачам провести самостійно.

Наслідок. *Система векторів, що містить нуль-вектор, є лінійно залежною.*

Теорема 1.4 (Основна теорема про лінійну залежність)**.** *Якщо кожний вектор системи з більшою кількістю векторів лінійно виражається через вектори системи з меншою кількістю векторів, то перша система лінійна залежна.*

Доведення. Нехай $\mathcal{A} = \big(\vec{a}_1, \ldots, \vec{a}_k\big)$ і $\mathcal{B} = \big(\vec{b}_1, \ldots, \vec{b}_l\big)$ — дві системи векторів, причому $k > l$. Нехай також кожний вектор системи $\mathcal{A}$ є лінійною комбінацією векторів системи $\mathcal{B}$. Якщо серед векторів системи $\mathcal{A}$ є нуль-вектор, то вона є лінійно залежною, а тому далі розглядатимемо випадок, коли усі вектори системи $\mathcal{A}$ ненульові.

1. Оскільки $\vec{a}_1 \xrightarrow{\text{л.в.}} \vec{b}_1, \ldots, \vec{b}_l$, то рівність $c_0\vec{a}_1 + c_1\vec{b}_1 + \ldots + c_l\vec{b}_l$ виконується хоча б при одному $c_i \neq 0$, $i \geqslant 1$ (якщо припустити, що $c_i = 0$ для всіх $i \geqslant 1$, то тоді $c_0 \neq 0$, а значить $\vec{a}_1 = \vec{0}$, що суперечить припущенню про ненульові вектори системи $\mathcal{A}$). Не порушуючи загальності вважатимемо, що $c_1 \neq 0$. Тоді

$$\vec{b}_1 = -\frac{c_0}{c_1}\vec{a}_1 - \frac{c_2}{c_1}\vec{b}_2 - \ldots - \frac{c_l}{c_1}\vec{b}_l,$$

тобто

$$\vec{b}_1 \xrightarrow{\text{л.в.}} \vec{a}_1, \vec{b}_2, \ldots, \vec{b}_l.$$

Тоді очевидно, що і кожний з векторів $\vec{a}_2, \ldots, \vec{a}_k$ також є лінійною комбінацією векторів $\vec{a}_1, \vec{b}_2, \ldots, \vec{b}_l$.

2. Складемо рівність $(\star)$ для векторів $\vec{a}_2, \vec{a}_1, \vec{b}_2, \ldots, \vec{b}_l$. Якщо усі коефіцієнти при векторах $\vec{b}_i$ дорівнюють нулю, то вектори $\vec{a}_1, \vec{a}_2$ утворюють лінійно залежну систему, а тому за теоремою 1.3 і уся система $\mathcal{A}$ буде лінійно залежною. Тому припустимо, що якийсь коефіцієнт при $\vec{b}_i$ відмінний від нуля. Не порушуючи загальності вважатимемо, що це коефіцієнт при $\vec{b}_2$. Тоді знаходимо, що

$$\vec{b}_2 \xrightarrow{\text{л.в.}} \vec{a}_1, \vec{a}_2, \vec{b}_3, \ldots, \vec{b}_l.$$

Крім цього, $\vec{a}_i \xrightarrow{\text{л.в.}} \vec{a}_1, \vec{a}_2, \vec{b}_3, \ldots, \vec{b}_l$, $i = \overline{3, k}$.

3. Продовжимо такі ж міркування далі. Враховуючи, що $k > l$ зрештою одержимо, що для деякого s: $\vec{a}_s \xrightarrow{\text{л.в.}} \vec{a}_1, \ldots, \vec{a}_{s-1}$. Якщо $s < k$, то система $\mathcal{A}$ містить лінійно залежну підсистему за теоремою 1.3 є лінійно залежною. Якщо ж $s = k$, то система $\mathcal{A}$ є лінійно залежною за теоремою 1.2. ■

Теорема 1.5. *Система векторів*

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

є лінійно незалежною, причому кожний вектор $\vec{x} \in \mathbb{R}^n$ лінійно виражається через вектори $\vec{e}_1, \ldots, \vec{e}_n$.

Доведення. Оскільки рівність ($\star$), що складена для векторів $\vec{e}_1, \ldots, \vec{e}_n$, виконується лише при $c_1 = \ldots = c_n = 0$, то система лінійно незалежна. Нехай $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ — довільний вектор з простору $\mathbb{R}^n$. Тоді $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \ldots + x_n\vec{e}_n$, що і потрібно було довести. ∎

Наслідок (теорем 1.4 та 1.5)**.** *Кожна множина з m векторів простору $\mathbb{R}^n$ при $m > n$ є лінійно залежною.*

Теорема 1.6. *Нехай $\vec{a}_1, \ldots, \vec{a}_k$ — вектори-стовпці простору $\mathbb{R}^n$ і $A = \begin{bmatrix} \vec{a}_1 & \ldots & \vec{a}_k \end{bmatrix}$ — матриця, стовпцями якої є $\vec{a}_1, \ldots, \vec{a}_k$. Тоді система векторів $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ є лінійно залежною тоді і тільки тоді, коли однорідна система лінійних рівнянь $\left[A \mid \vec{0}\right]$ має ненульовий розв'язок.*

Доведення. Справді, однорідну систему лінійних рівнянь $\left[A \mid \vec{0}\right]$ з матрицею $A = \begin{bmatrix} \vec{a}_1 & \ldots & \vec{a}_k \end{bmatrix}$ можна записати у векторній формі $x_1\vec{a}_1 + \ldots + x_n\vec{a}_k = \vec{0}$. Якщо остання рівність виконується не при всіх x_i рівних нулю, то, з одного боку, система має ненульовий розв'язок, а з іншого — вектори $\vec{a}_1, \ldots, \vec{a}_k$ утвоюють лінійно залежну систему. ∎

2.3. Базис системи векторів.

Означення 1.14. Дві системи векторів $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ і $\left(\vec{b}_1, \ldots, \vec{b}_l\right)$ простору $\mathbb{R}^n$ називаються *еквівалентними*, якщо кожен вектор першої системи лінійно виражається через вектори другої і навпаки.

Теорема 1.7. *Дві еквівалентні лінійно незалежні системи містять однакову кількість векторів.*

Доведення. Нехай $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ і $\left(\vec{b}_1, \ldots, \vec{b}_l\right)$ — еквівалентні лінійно незалежні системи векторів. Якщо припустити, що $k > l$, то за основною теоремою про лінійну залежність система $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ мала б бути лінійно залежною, що суперечить умові. Аналогічно, якщо $l > k$, то лінійно залежною мала б бути система $\left(\vec{b}_1, \ldots, \vec{b}_l\right)$. Таким чином, $k = l$. ■

Лема 1.1. *Якщо система векторів $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ лінійно незалежна, а система векторів $\left(\vec{b}, \vec{a}_1, \ldots, \vec{a}_k\right)$ — лінійно залежна, то $\vec{b}$ лінійно виражається через $\vec{a}_1, \ldots, \vec{a}_k$.*

Доведення. Із лінійної залежності системи векторів $\left(\vec{b}, \vec{a}_1, \ldots, \vec{a}_k\right)$ слідує, що рівність

$$\alpha \cdot \vec{b} + \alpha_1 \vec{a}_1 + \ldots + \alpha_k \vec{a}_k = \vec{0}$$

можлива, коли не усі коефіцієнти одночасно дорівнюють нулю. Якщо припустити, що $\alpha = 0$, то цей відмінний від нуля коефіцієнт буде серед $\alpha_1, \ldots, \alpha_k$, а отже система $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ буде лінійно залежною, що суперечить умові. Отже, $\alpha \neq 0$ і

$$\vec{b} = -\frac{\alpha_1}{\alpha}\vec{a}_1 - \ldots - \frac{\alpha_k}{\alpha}\vec{a}_k.$$

■

Означення 1.15. Підсистема векторів $\left(\vec{a}_{i_1}, \ldots, \vec{a}_{i_l}\right)$ системи $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ називається *базисом цієї системи*, якщо виконуються умови:

1) підсистема $\left(\vec{a}_{i_1}, \ldots, \vec{a}_{i_l}\right)$ є лінійно незалежною;

2) для довільного вектора $\vec{a}_s$ системи $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ підсистема $\left(\vec{a}_s, \vec{a}_{i_1}, \ldots, \vec{a}_{i_l}\right)$ є лінійно залежною.

За лемою 1.1 вектор $\vec{a}_s$ лінійно виражається через вектори $\left(\vec{a}_{i_1}, \ldots, \vec{a}_{i_l}\right)$, тому можна дати таке означення базису системи векторів.

Означення 1.15′. Підсистема векторів $\left(\vec{a}_{i_1}, \ldots, \vec{a}_{i_l}\right)$ системи $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ називається її *базисом*, якщо ця підсистема є лінійно незалежною і через її вектори лінійно виражається будь-який вектор даної системи.

Приклад 1.16. Знайти базис системи векторів $\vec{a}_1 = (1, 0, 2)$, $\vec{a}_2 = (1, 0, 1)$, $\vec{a}_3 = (1, 0, 0)$.

Розв'язання. Рівність ($\star$), складена для цих векторів, приводить до системи

$$\begin{cases} \alpha_1+\alpha_2+\alpha_3 = 0, \\ \qquad\qquad\quad 0 = 0, \\ 2\alpha_1+\alpha_2 \qquad = 0, \end{cases}$$

яка має безліч розв'язків. Це означає, що вихідна система векторів є лінійно залежною, тому її базис містить два або менше векторів. Вектори, наприклад, $\vec{a}_1$ і $\vec{a}_3$ є лінійно незалежними (чому?), а тому вони утворюють базис даної системи.

Аналогічно в якості базису можна було б взяти підсистеми $(\vec{a}_1, \vec{a}_2)$, $(\vec{a}_2, \vec{a}_3)$. ■

Розглянемо деякі властивості базису системи векторів:

1. Будь-яка система векторів еквівалентна з базисом цієї системи.

2. Будь-які два базиси системи векторів еквівалентні між собою.

Справді, нехай $(\vec{a}_{i_1}, \ldots, \vec{a}_{i_l})$ і $(\vec{a}_{j_1}, \ldots, \vec{a}_{j_m})$ — два базиси однієї і тієї ж системи векторів. Оскільки перша підсистема — базис, то кожний вектор другої лінійно виражається через вектори першої. Навпаки, оскільки друга підсистема є базисом, то кожний вектор першої підсистеми лінійно виражається через вектори другої. Таким чином, ці два базиси є еквівалентними системами векторів.

3. Будь-які два базиси однієї і тієї ж системи векторів містять однакову кількість векторів.

Справді, за доведеною властивістю два будь-які два базиси однієї і тієї ж системи еквівалентні. А оскільки базис є лінійно незалежною системою, то за теоремою 1.7 вони містять однакову кількість векторів.

Остання властивість дає можливість ввести таке означення.

Означення 1.16. *Рангом* системи векторів $(\vec{a}_1, \ldots, \vec{a}_k)$ називається кількість векторів її базису.

Наприклад, ранг системи векторів $\vec{a}_1 = (1, 0, 2)$, $\vec{a}_2 = (1, 0, 1)$, $\vec{a}_3 = (1, 0, 0)$ (див. приклад 1.16) дорівнює 2, а ранг системи векторів $\vec{a}_1 = (1, 0, 0, 0)$, $\vec{a}_2 = (0, 1, 0, 0)$, $\vec{a}_3 = (0, 0, 1, 0)$ дорівнює 3.

Якщо в якості системи векторів розглянути *усі* вектори простору $\mathbb{R}^n$, то базис цієї системи називається *базисом простору* $\mathbb{R}^n$, а ранг цієї системи — *розмірністю* простору $\mathbb{R}^n$. З теореми 1.5 випливає, що розмірність простору $\mathbb{R}^n$ дорівнює n.

2.4. Елементарні перетворення системи векторів. Ранг матриці.

Означення 1.17. Нехай $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ — деяка система векторів простору $\mathbb{R}^n$. Наступні перетворення цієї системи називаються *елементарними*:

1) перестановка місцями будь-яких векторів системи;

2) множення будь-якого із векторів системи на число $c \neq 0$;

3) додавання до будь-якого із векторів системи іншого вектора, помноженого на c;

4) відкидання від векторів системи $\vec{0}$ (якщо він там є) або приписування до векторів системи $\vec{0}$.

Обґрунтування наступного твердження пропонуємо читачеві зробити самостійно.

Теорема 1.8. *При елементарних перетвореннях системи векторів її ранг не змінюється.*

Нехай маємо систему лінійних рівнянь з n невідомими з розширеною матрицею $\left[A \mid \vec{b}\right]$. Цю матрицю можна розглядати як систему векторів-рядків простору $\mathbb{R}^{n+1}$. При цьому виконання рядкового елементарного перетворення матриці відповідатиме елементарному перетворенню для системи цих векторів-рядків.

Теорема 1.9. *Ненульові вектори-рядки матриці у східчастій формі є лінійно незалежними.*

Доведення. Справді, нехай $\vec{v}_1, \ldots, \vec{v}_r$ — ненульові вектори-рядки матриці, що має східчасту форму з ведучими елементами $c_1, \ldots, c_r$. Складемо рівність $(\star)$ для цієї системи векторів:

$$\alpha_1\vec{v}_1 + \ldots + \alpha_r\vec{v}_r = \vec{0}.$$

Якщо записати цю рівність у координатній формі, то знайдемо:

$$\begin{cases} \alpha_1 c_1 & = 0, \\ \alpha_1 d_{21} + \alpha_2 c_2 & = 0, \\ \quad\ldots & \\ \alpha_1 d_{r1} + \alpha_2 d_{r2} + \ldots + \alpha_r c_r & = 0, \end{cases}$$

звідки $\alpha_1 = \ldots = \alpha_r = 0$, що і потрібно було довести. ■

Означення 1.18. *Рангом матриці* називається кількість ненульових рядків у її східчастій формі.

Ми позначатимемо ранг матриці A через $r(A)$ або $\operatorname{rank}(A)$.

Приклад 1.17. Знайти ранг матриці

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

Розв'язання. Приведемо задану матрицю до східчастої форми:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 3 & 1 \\ 2 & 1 & 1 & 2 & -1 \\ -2 & -1 & -1 & -2 & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -3 & -3 & -5 & -9 \\ 0 & -3 & -5 & -6 & -11 \\ 0 & 3 & 5 & 6 & 11 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -3 & -3 & -5 & -9 \\ 0 & 0 & -2 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -3 & -3 & -5 & -9 \\ 0 & 0 & -2 & -1 & -2 \end{bmatrix}.$$

Отже, $r(A) = 3$. ■

2.5. Стовпцевий ранг матриці.

Нехай маємо матрицю, що містить m рядків та n стовпчиків. Кожний з рядків матриці є n-вимірним вектором, а кожний з її стовпців — m-вимірним вектором. Тому з кожною матрицею зв'язуються дві системи векторів: система векторів-рядків і система векторів-стовпців. Кожна з цих двох систем має свій ранг. Як було показано вище, рангом матриці називається по суті ранг системи векторів-рядків (його ще називають *рядковим рангом матриці*). Подібно до цього *стовпцевим рангом матриці* назвемо ранг системи векторів-стовпців.

Виявляється, що рядковий і стовпцевий ранги матриць завжди співпадають. Для доведення цього факту використаємо таке допоміжне твердження.

Лема 1.2. *Якщо рядковий ранг матриці A дорівнює p, то існує система із p лінійно незалежних векторів-стовпців, через які лінійно виражається кожний вектор-стовпець цієї матриці.*

Доведення. Нехай ранг системи векторів-рядків $\vec{v}_1, \ldots, \vec{v}_m$ матриці

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

дорівнює p. Це означає, що базис цієї системи містить p векторів. Не порушуючи загальності вважатимемо, що базис утворюють вектори $\vec{v}_1, \ldots, \vec{v}_p$. Кожний з решти векторів рядків лінійно виражається через них:

$$\begin{aligned} \vec{v}_{p+1} &= c_{p+1,1}\vec{v}_1 + c_{p+1,2}\vec{v}_2 + \ldots + c_{p+1,p}\vec{v}_p, \\ \vec{v}_{p+2} &= c_{p+2,1}\vec{v}_1 + c_{p+2,2}\vec{v}_2 + \ldots + c_{p+2,p}\vec{v}_p, \\ &\ldots \\ \vec{v}_m &= c_{m1}\vec{v}_1 + c_{m2}\vec{v}_2 + \ldots + c_{mp}\vec{v}_p. \end{aligned}$$

Запишемо ці рівності в координатах:

$$\begin{aligned} a_{p+1,i} &= c_{p+1,1}a_{1i} + c_{p+1,2}a_{2i} + \ldots + c_{p+1,p}a_{pi}, \quad i = \overline{1,n}, \\ a_{p+2,i} &= c_{p+2,1}a_{1i} + c_{p+2,2}a_{2i} + \ldots + c_{p+2,p}a_{pi}, \quad i = \overline{1,n}, \\ &\ldots \\ a_{mi} &= c_{m1}a_{1i} + c_{m2}a_{2i} + \ldots + c_{mp}a_{pi}, \quad i = \overline{1,n}. \end{aligned}$$

Розглянемо вектори

$$\begin{aligned} \vec{u}_1 &= (\overbrace{1, 0, 0, \ldots, 0}^{p}, c_{p+1,1}, c_{p+2,1}, \ldots, c_{m1}), \\ \vec{u}_2 &= (0, 1, 0, \ldots, 0, c_{p+1,2}, c_{p+2,2}, \ldots, c_{m2}), \\ &\ldots \\ \vec{u}_p &= (0, 0, \ldots, 0, 1, c_{p+1,p}, c_{p+2,p}, \ldots, c_{mp}). \end{aligned}$$

Ця система векторів лінійно незалежна (подумайте чому). З іншого боку, j-й стовпець матриці A можна подати у вигляді їх лінійної комбінації:

$$\vec{a}_j = a_{1j}\vec{u}_1 + a_{2j}\vec{u}_2 + \ldots + a_{pj}\vec{u}_p,$$

що і потрібно було довести. ■

Теорема 1.10. *Рядковий і стовпцевий ранги матриці співпадають.*

Доведення. Нехай A — матриця розмірів $m \times n$, рядковий ранг якої дорівнює p, а стовпцевий — s. Нехай базис системи векторів-рядків утворюють вектори $\left(\vec{v}_1, \dots, \vec{v}_p\right)$, а векторів стовпців — вектори $\left(\vec{a}_1, \dots, \vec{a}_s\right)$.

За доведеною вище лемою кожний з векторів системи $\left(\vec{a}_1, \dots, \vec{a}_s\right)$ лінійно виражається через вектори $\left(\vec{u}_1, \dots, \vec{u}_p\right)$ (такі вектори $\left(\vec{u}_1, \dots, \vec{u}_p\right)$ існують і вони лінійно незалежні). Тоді за основною теоремою про лінійну залежність $s \leqslant p$ (в іншому випадку система векторів $\left(\vec{a}_1, \dots, \vec{a}_s\right)$ була б лінійно залежною, що неможливо, оскільки вона є базисом).

Розглянемо так звану *транспоновану* матрицю до A:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Її рядковий ранг дорівнює s, а стовпцевий — r. Повторивши попередні міркування стосовно матриці A^T одержимо, що $p \leqslant s$. Таким чином, $p = s$. ■

Таким чином, нами доведено такий факт: ранг довільної матриці A дорівнює рангу транспонованої матриці A^T.

2.6. Множення матриці на вектор і матрично-векторна форма запису системи лінійних рівнянь.

Означення 1.19. Якщо A є $m \times n$ матрицею зі стовпцями $\vec{a}_1, \dots, \vec{a}_n$ і $\vec{x} \in \mathbb{R}^n$, то *добутком матриці A на вектор $\vec{x}$* (позначається $A \cdot \vec{x}$ або $A\vec{x}$) називається лінійна комбінація стовпців матриці A з використанням відповідних координат вектора $\vec{x}$ як ваг:

$$A\vec{x} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n.$$

Наприклад,

$$\begin{bmatrix} 1 & 3 & 2 \\ -1 & -1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 11 \\ 9 \end{bmatrix}.$$

Нехай маємо систему лінійних рівнянь

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{1n}x_n & = & b_2, \\ \dots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m. \end{cases} \tag{1.31}$$

Введемо позначення:

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \vec{a}_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \vec{a}_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}, \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Тоді систему (1.31) можна подати у вигляді векторного рівняння

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n = \vec{b}. \tag{1.32}$$

Останню рівність називатимемо *векторною формою запису системи лінійних рівнянь* (1.31).

Якщо ж головну матрицю системи (1.31) позначити через A і ввести вектор невідомих $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, то систему (1.31) можна подати у так званому *матрично-векторному вигляді*

$$A\vec{x} = \vec{b}. \tag{1.33}$$

Однорідна система лінійних рівнянь у матрично-векторній формі, вочевидь, набуває вигляд $A\vec{x} = \vec{0}$.

Властивості введеної операції множення матриці на вектор висвітлено в наступній теоремі, доведення якої пропонується читачам виконати самостійно.

Теорема 1.11. *Якщо A — $m \times n$-матриця, $\vec{u}$ і $\vec{v} \in \mathbb{R}^n$, $c \in \mathbb{R}$, то:*
1) $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$;
2) $A(c\vec{u}) = c \cdot A\vec{u}$.

Крім цього, зазначимо, що результатом добутку матриці A на вектор $\vec{e}_j$ (вектор, j-та координата якого дорівнює 1, а решта — 0) є j-ий стовпець матриці A.

2.7. Критерій сумісності системи лінійних рівнянь (теорема Кронекера–Капеллі).

Теорема 1.12 (Теорема Кронекера[1]–Капеллі[2]). *Для того, щоб система лінійних рівнянь була сумісною, необхідно і достатньо, щоб ранг головної матриці цієї системи співпадав з рангом її розширеної матриці.*

Доведення. *Необхідність.* Нехай маємо сумісну систему лінійних рівнянь, векторна форма якої має вигляд

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \ldots + x_n\vec{a}_n = \vec{b}.$$

Із її сумісності випливає, що існують такі числа $\alpha_1, \alpha_2, \ldots, \alpha_n$, що

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n = \vec{b},$$

тобто $\vec{b}$ лінійно виражається через $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$. Отже, системи векторів $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right)$ та $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n, \vec{b}\right)$ є еквівалентними і тому мають однаковий ранг.

Достатність. Оскільки (стовпцеві) ранги головної та розширеної матриць співпадають, то ранг системи векторів $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\right)$ дорівнює рангу системи $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n, \vec{b}\right)$. Це означає, що $\vec{b}$ лінійно виражається через $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$. Тому існують такі $\alpha_1, \ldots, \alpha_n$, що

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n = \vec{b},$$

а це і означає, що $(\alpha_1, \ldots, \alpha_n)$ є розв'язком системи, тобто вона є сумісною. ■

§ 3. Поняття підпростору простору $\mathbb{R}^n$.

Означення 1.20. *Підпростором* арифметичного векторного простору $\mathbb{R}^n$ називається довільна множина S векторів з $\mathbb{R}^n$ така, що виконуються умови :

1) якщо $\vec{u}, \vec{v} \in S$, то $\vec{u} + \vec{v} \in S$ (говорять, що S замкнена відносно операції додавання);
2) якщо $\vec{u} \in S$, то для довільного $\alpha \in \mathbb{R}$ вектор $\alpha\vec{u}$ також належить S (говорять, що S замкнена відносно операції множення на число).

З другої умови означення підпростору простору $\mathbb{R}^n$ випливає, що підпростір завжди містить $\vec{0}$.

---

[1] Leopold Kronecker (1823 — 1891) — німецький математик.

[2] Alfredo Capelli (1855 – 1910) — італійський математик.

Приклад 1.18. 1. Розглянемо геометричну інтерпретацію простору $\mathbb{R}^3$. Нехай S — множина векторів, що розташовані на деякій прямій, яка проходить через початок координат. З геометричних тлумачень операцій додавання і множення на число векторів зрозуміло, що ця множина задовольняє означення підпростору простору $\mathbb{R}^3$.

2. Множина векторів, що розташовані на площині, яка проходить через початок координат, з аналогічних міркувань також утворює підпростір простору $\mathbb{R}^3$.

3. Множина, яка складається з одного нуль-вектора, очевидно, задовольняє означення 3.2. Сам векторний простір $\mathbb{R}^n$ також можна мислити як підпростір. Це так звані *тривіальні підпростори* арифметичного векторного простору.

Теорема 1.13. *Нехай $\vec{v}_1, \ldots, \vec{v}_k$ — деякі вектори простору $\mathbb{R}^n$. Тоді їх лінійна оболонка $L(\vec{v}_1, \ldots, \vec{v}_k)$ — підпростір простору $\mathbb{R}^n$.*

Доведення. Нехай $\vec{u}_1, \vec{u}_2 \in L(\vec{v}_1, \cdots \vec{v}_k)$. Це означає, що існують такі числа $\alpha_1, \ldots, \alpha_k$, $\beta_1, \ldots, \beta_k$, що

$$\vec{u}_1 = \alpha_1\vec{v}_1 + \ldots + \alpha_k\vec{v}_k,$$
$$\vec{u}_2 = \beta_1\vec{v}_1 + \ldots + \beta_k\vec{v}_k.$$

Але тоді їх сума $\vec{u}_1 + \vec{u}_2 = (\alpha_1 + \beta_1)\vec{v}_1 + \ldots + (\alpha_k + \beta_k)\vec{v}_k$, тобто $\vec{u}_1 + \vec{u}_2$ є лінійною комбінацією системи векторів $\vec{v}_1, \ldots, \vec{v}_k$. Отже, $L(\vec{v}_1, \cdots \vec{v}_k)$ замкнена відносно операції додавання векторів. Аналогічно доводиться, що вона замкнена і відносно операції множення на число. ■

Лінійна оболонка системи векторів є в певному смислі «найменшим» підпростором, що містить ці вектори. Тому лінійну оболонку векторів $\vec{v}_1, \cdots \vec{v}_k$ ми називатимемо також *підпростором, що породжений векторами* $\vec{v}_1, \ldots, \vec{v}_k$.

Приклад 1.19. Чи є підпростором множина векторів простору $\mathbb{R}^3$ виду (x, y, z) така, що $x = 5y$, $z = -2y$?

Розв'язання. Вказана множина векторів $\{(5y, y, -2y) \colon y \in \mathbb{R}\}$ може мислитись як лінійна оболонка вектора $\vec{a} = (5, 1, -2)$. За теоремою 1.13 ця множина є підпростором простору $\mathbb{R}^n$. ■

Приклад 1.20. Чи є підпростором множина векторів простору $\mathbb{R}^3$ виду (x, y, z) така, що $x = 5y$, $z = -2y + 1$?

Розв'язання. Множина векторів $\{(5y, y, -2y+1) \colon y \in \mathbb{R}\}$ не містить нуль-вектор простору $\mathbb{R}^n$, а тому не може бути підпростором простору $\mathbb{R}^n$. ■

3.1. Підпростори, пов'язані з матрицями.

Означення 1.21. Нехай A — $m \times n$ матриця. *Простором рядків* матриці A називається підпростір $\mathrm{row}(A)$ простору $\mathbb{R}^n$, що породжений векторами-рядками матриці A; *простором стовпців* матриці A називається підпростір $\mathrm{col}(A)$ простору $\mathbb{R}^m$, що породжений векторами-стовпцями матриці A.

Приклад 1.21. Чи належить вектор $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ простору стовпців матриці $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 3 & -2 \end{bmatrix}$?

Розв'язання. Перевіримо, чи існують α_1, α_2 такі, що

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}.$$

Одержали систему лінійних рівнянь з розширеною матрицею

$$\left[\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & -2 & 3 \end{array}\right],$$

яка рядково еквівалентна до матриці

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{array}\right].$$

Отже, система несумісна.

Відповідь. Не належить. ■

Розглянемо однорідну систему лінійних рівнянь $A\vec{x} = \vec{0}$, де A — $m \times n$-матриця.

Теорема 1.14. *Множина розв'язків однорідної системи лінійних рівнянь $A\vec{x} = \vec{0}$ утворює підпростір простору $\mathbb{R}^n$.*

Доведення. Нехай $\vec{u}$ та $\vec{v}$ — деякі два розв'язки системи $A\vec{x} = \vec{0}$. Це означає, що $A\vec{u} = A\vec{v} = \vec{0}$. Але тоді згідно з теоремою 1.11 про властивості добутку матриці на вектор $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} = \vec{0}$ і для довільного $c \in \mathbb{R}$ $A(c\vec{u}) = c \cdot A\vec{u} = \vec{0}$. Отже, множина розв'язків однорідної системи замкнена відносно операції додавання і множення на число, а тому є підпростором простору $\mathbb{R}^n$. ■

Означення 1.22. Нехай A — $m \times n$-матриця. *Нуль-простором* матриці A називається підпростір $\text{null}(A)$ простору $\mathbb{R}^n$, що складається з усіх розв'язків однорідної системи лінійних рівнянь $A\vec{x} = \vec{0}$.

Означення 1.23. *Базисом* підпростору S простору $\mathbb{R}^n$ називається кожна лінійно незалежна множина векторів підпростору, що породжує S.

Кількість векторів базису називається *розмірністю* підпростору S і позначається через $\dim S$.

Таким чином, рядковий ранг матриці A є розмірністю простору рядків цієї матриці ($\dim(\text{row}\, A)$), стовпцевий ранг — розмірністю простору стовпців ($\dim(\text{col}\, A)$). Розмірність нуль-простору матриці називається *дефект матриці* і позначається $\text{nullity}(A) = \dim(\text{null}\, A)$.

3.2. Зв'язок між розв'язками неоднорідної та однорідної систем лінійних рівнянь.

Нехай маємо систему лінійних рівнянь

$$A\vec{x} = \vec{b} \tag{1.34}$$

і відповідну до неї однорідну систему

$$A\vec{x} = \vec{0}. \tag{1.35}$$

Теорема 1.15. *Якщо $\vec{x}_{част}$ — деякий частинний (фіксований) розв'язок системи* (1.34), *Z — множина всіх розв'язків системи* (1.35), *то множина $\{\vec{z} + \vec{x}_{част} \colon \vec{z} \in Z\}$ — множина всіх розв'язків системи* (1.34).

Доведення. Нехай $\vec{z}$ — деякий розв'язок системи (1.35). Тоді $A(\vec{x}_{\text{част}} + \vec{z}) = A\vec{x}_{\text{част}} + A\vec{z} = \vec{b} + \vec{0} = \vec{b}$, тобто $\vec{x}_{\text{част}} + \vec{z}$ — деякий розв'язок системи (1.34). З іншого боку, різниця двох розв'язків системи (1.34) є розв'язком системи (1.35): $A(\vec{x}_1 - \vec{x}_2) = A\vec{x}_1 - A\vec{x}_2 = \vec{b} - \vec{b} = \vec{0}$. Тобто для кожного розв'язку $\vec{x}_1$ системи (1.34) можна вказати такий розв'язок системи (1.35), що $\vec{x}_1 = \vec{z} + \vec{x}_{\text{част}}$. Отже, множина $\{\vec{z} + \vec{x}_{\text{част}} : \vec{z} \in Z\}$ співпадає з множиною усіх розв'язків системи (1.34). ■

З останньої теореми випливає такий спосіб знаходження загального розв'язку системи лінійних рівнянь (1.34):

1) підібрати один з розв'язків цієї системи;
2) знайти усі розв'язки відповідної однорідної системи лінійних рівнянь;
3) до знайденого розв'язку системи (1.34) додати по черзі усі розв'язки системи (1.35). В результаті одержимо усі розв'язки системи (1.34).

Означення 1.24. Нехай U — довільний підпростір простору $\mathbb{R}^n$, $\vec{a} \in \mathbb{R}^n$. Множина

$$M = \vec{a} + U = \{x \in \mathbb{R}^n : \vec{x} = \vec{a} + \vec{u}, \vec{u} \in U\}$$

називається *лінійним многовидом* простору $\mathbb{R}^n$, що утворений паралельним перенесенням підпростору U на вектор $\vec{a}$.

Отже, множина всіх розв'язків неоднорідної системи лінійних рівнянь з n невідомими є лінійний многовид $\vec{a} + U$ простору $\mathbb{R}^n$, де $\vec{a}$ — деякий частинний розв'язок неоднорідної системи, U — підпростір розв'язків відповідної їй однорідної системи лінійних рівнянь.

3.3. Загальний розв'язок і фундаментальна система розв'язків однорідної системи лінійних рівнянь.

Ми показали, що множина розв'язків однорідної системи лінійних рівнянь утворює підпростір простору $\mathbb{R}^n$. Цей підпростір має свій базис. Ті розв'язки однорідної системи, які утворюють базис, називають її *фундаментальною системою розв'язків*.

Теорема 1.16. *Нехай A — $m \times n$ матриця; r — її ранг. Тоді фундаментальна система розв'язків системи $A\vec{x} = \vec{0}$ містить рівно $n - r$ векторів.*

Доведення. Якщо привести A до зведеної східчастої форми, то вона міститиме r ненульових рядків, а значить і r основних змінних, які виражаються через $n - r$ вільних змінних. Не порушуючи загальності вважатимемо, що основними є змінні $x_1, \ldots, x_r$, а решта —

вільні, тоді

$$\begin{aligned} x_1 &= \gamma_{11}x_{r+1} + \gamma_{12}x_{r+2} + \ldots + \gamma_{1,n-r}x_n, \\ x_2 &= \gamma_{21}x_{r+1} + \gamma_{22}x_{r+2} + \ldots + \gamma_{2,n-r}x_n, \\ &\vdots \\ x_r &= \gamma_{r1}x_{r+1} + \gamma_{r2}x_{r+2} + \ldots + \gamma_{r,n-r}x_n. \end{aligned}$$

Якщо вільним змінним надати значень $\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_n$, то дістанемо деякий розв'язок системи $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, де

$$\begin{aligned} \alpha_1 &= \gamma_{11}\alpha_{r+1} + \gamma_{12}\alpha_{r+2} + \ldots + \gamma_{1,n-r}\alpha_n, \\ \alpha_2 &= \gamma_{21}\alpha_{r+1} + \gamma_{22}\alpha_{r+2} + \ldots + \gamma_{2,n-r}\alpha_n, \\ &\vdots \\ \alpha_r &= \gamma_{r1}\alpha_{r+1} + \gamma_{r2}\alpha_{r+2} + \ldots + \gamma_{r,n-r}\alpha_n. \end{aligned}$$

Надамо вільним змінним послідовно таких значень:

$$\begin{aligned} &1, 0, \ldots, 0; \\ &0, 1, \ldots, 0; \\ &\vdots \\ &0, 0, \ldots, 1. \end{aligned}$$

Дістанемо $n - r$ розв'язків системи:

$$\begin{aligned} \vec{s}_1 &= (\gamma_{11}, \gamma_{21}, \ldots, \gamma_{r1}, 1, 0, \ldots, 0), \\ \vec{s}_2 &= (\gamma_{12}, \gamma_{22}, \ldots, \gamma_{r2}, 0, 1, \ldots, 0), \\ &\vdots \\ \vec{s}_{n-r} &= (\gamma_{1,n-r}, \gamma_{2,n-r}, \ldots, \gamma_{r,n-r}, 0, 0, \ldots, 1). \end{aligned}$$

Цей набір розв'язків є лінійно незалежною системою векторів. Крім цього, кожний інший розв'язок системи лінійно виражається через ці $n - r$ розв'язків. Справді, нехай $\vec{s} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ — довільний розв'язок системи. Маємо:

$$\alpha_{r+1}\vec{s}_1 + \alpha_{r+2}\vec{s}_2 + \ldots + \alpha_n\vec{s}_{n-r} = (\alpha_1, \alpha_2, \ldots, \alpha_n) = \vec{s},$$

тобто $\vec{s} \xrightarrow{\text{л.в.}} \big(\vec{s}_1, \ldots, \vec{s}_{n-r}\big)$. Тому $\vec{s}_1, \ldots, \vec{s}_{n-r}$ — базис підпростору розв'язків системи $A\vec{x} = \vec{0}$. ■

Наслідок. *Якщо A — $m \times n$ матриця, то сума ранга і ядра цієї матриці рівна n.*

Додаткові відомості і завдання для самостійної роботи

1. Нехай A — $m \times n$ матриця, елементи якої належать полю $\mathbb{Z}_p$. Довести, що кожна сумісна система лінійних рівнянь з головною матрицею A над полем $\mathbb{Z}_p$ має рівно $p^{n-\operatorname{rank}(A)}$ розв'язків.
2. Знайти a, b, c такі, що
$$\frac{x^2 - x + 3}{(x^2+2)(2x-1)} = \frac{ax+b}{x^2+2} + \frac{c}{2x-1}.$$
3. Довести, що будь-який лінійний многовид простору $\mathbb{R}^n$ збігається з множиною всіх розв'язків деякої сумісної системи лінійних рівнянь.
4. Довести, що якщо M — лінійний многовид простору $\mathbb{R}^n$, то для довільних $\vec{x}, \vec{y} \in M$ і $\alpha, \beta \in \mathbb{R}$, для яких $\alpha + \beta = 1$, вектор $\vec{z} = \alpha\vec{x} + \beta\vec{y} \in M$.
5. Нехай M — підмножина простору $\mathbb{R}^n$ з такою властивістю: якщо $\vec{x}, \vec{y} \in M$ і $\alpha \in \mathbb{R}$, то $\alpha\vec{x} + (1-\alpha)\vec{y}$ також належить M. Довести, що M — лінійний многовид.

Розділ 2

Матриці та визначники

Я не сказав, що буде легко. Я лише обіцяв відкрити правду.

х/ф «Матрица»

До цього моменту ми використовували матриці для спрощеного запису систем лінійних рівнянь. В цьому розділі матриці розглядатимуться як математичні об'єкти, а множина всіх матриць буде наділена структурою шляхом введення на ній операцій додавання, множення на число і множення матриць. Доцільність таких означень буде зрозумілою з розділу 4, де ми побачимо як матриці будуть виконувати роль зручного представлення одного типу відображень. Виявиться, що вони слугують не просто інструментом для зручного запису і подання інформації, а насправді представляють певний тип функцій, що відображають одні вектори у інші. Теорія матриць має численні застосування; з окремими з них читач може познайомитись в посібнику [12].

§ 1. Матриці. Дії над матрицями.

1.1. Поняття матриці. Типи матриць.

Означення 2.1. *Матрицею над полем P* називається прямокутна таблиця елементів цього поля:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad a_{ij} \in P, i = \overline{1,m}, j = \overline{1,n}.$$

Матриці зазвичай позначають великими латинськими літерами. Іноді для того, щоб підкреслити, що матриця A складається з m рядків та n стовпчиків, ми позначатимемо її так: $A_{m\times n}$ або $[a_{ij}]_{m\times n}$.

Матрицю з елементами a_{ij} позначають $[a_{ij}]$, а через $(A)_{ij}$ — елемент i-го рядка і j-го стовпця матриці A.

$$A_{m\times n} = [a_{ij}]_{m\times n} = \begin{array}{cc} \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} & \begin{array}{c} \\ \\ \leftarrow \boxed{i\text{-й рядок}} \\ \\ \\ \end{array} \\ \uparrow \\ \boxed{j\text{-й стовпець}} & \end{array}$$

Дві матриці називаються *рівними*, якщо їх розміри однакові і відповідні елементи однакові.

Якщо окремо не зауважено, то надалі ми розглядатимемо матриці над полем $\mathbb{R}$.

Стовпці матриці A є m-вимірними векторами, позначимо їх $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$. Тоді $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_n \end{bmatrix}$.

Якщо $m = n$ (кількість рядків співпадає з кількістю стовпців), то така матриця називається *квадратною матрицею порядку n*. Набір елементів $a_{11}, a_{22}, \ldots, a_{nn}$ утворює *головну діагональ*, а набір елементів $a_{1n}, a_{(2,n-1)}, \ldots, a_{n1}$ — *побічну діагональ* матриці .

$$\boxed{\text{Побічна діагональ}} \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \boxed{\text{Головна діагональ}}$$

Квадратна матриця, усі елементи якої нижче (вище) від головної діагоналі дорівнюють нулю, називають *верхньою* (*нижньою*) *трикутною матрицею*:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

Квадратна матриця, усі елементи якої, крім, можливо, елементів головної діагоналі, дорівнюють нулю, називають *діагональною матрицею*. Діагональну матриця з елементами $a_{11}, \dots, a_{nn}$ на головній діагоналі записують так: $\operatorname{diag}\{a_{11}, a_{22}, \dots, a_{nn}\}$.

$$\begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

1.2. Лінійні операції над матрицями.

Означення 2.2. *Сумою* двох $m \times n$ матриць $A = \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{bmatrix}$ та $B = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix}$ називається матриця $A + B$, стовпці якої є сумами відповідних стовпців матриць A та B: $A+B = \begin{bmatrix} \vec{a}_1 + \vec{b}_1 & \dots & \vec{a}_n + \vec{b}_n \end{bmatrix}$.

Означення 2.3. *Добутком матриці* $A = \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{bmatrix}$ *на число* $\lambda \in \mathbb{R}$ називається матриця $\lambda A = \begin{bmatrix} \lambda \vec{a}_1 & \dots & \lambda \vec{a}_n \end{bmatrix}$.

Приклад 2.1. Нехай

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 3 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}.$$

Тоді

$$A + B = \begin{bmatrix} 1+2 & 2+1 & 3+0 \\ 4+(-1) & 1+3 & 2+2 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 4 & 4 \end{bmatrix},$$

$$3A = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 3 & 6 \end{bmatrix}, \quad -B = \begin{bmatrix} -2 & -1 & 0 \\ 1 & -3 & -2 \end{bmatrix}.$$

Суми $A+C$ і $B+C$ невизначені, оскільки матриці A і C, B і C мають різні розміри.

Різницею двох матриць A і B однакових розмірів називається матриця $A - B = A + (-1) \cdot B$.

Матриця, усі елементи якої дорівнюють нулю (нульовому елементу поля P), називається *нульовою* і позначається O (або $O_{m \times n}$ щоб підкреслити її розміри).

Елементи $a_{11}, a_{22}, \dots, a_{nn}$ називаються *діагональними елементами* матриці.

Теорема 2.1 (Властивості лінійних операцій над матрицями). *Нехай A, B, C — матриці однакових розмірів, $\alpha, \beta \in P$. Тоді справедливі такі тотожності:*

1) $A + (B + C) = (A + B) + C$ *(асоціативність додавання матриць);*
2) $A + B = B + A$ *(комутативність додавання матриць);*
3) $A + O = A$ *(властивість нульової матриці);*
4) $A + (-A) = O$ *(властивість протилежної матриці);*
5) $\alpha(A + B) = \alpha A + \alpha B$ *(дистрибутивність множення матриці на число відносно додавання матриць);*
6) $(\alpha + \beta)A = \alpha A + \beta A$ *(дистрибутивність множення матриці на число відносно додавання чисел);*
7) $\alpha(\beta A) = (\alpha\beta)A$ *(асоціативність множення матриці на число);*
8) $1 \cdot A = A$.

Твердження останньої теореми випливає з того, що усі матриці мають однакові розміри і перевірка кожної з восьми умов зводиться до перевірки виконання операцій додавання і множення на елемент поля P на множині P, а оскільки $(P, +, \cdot)$ є полем, то усі умови виконуються згідно з означенням поля.

Позначимо через $M_{m\times n}(P)$ множину всіх матриць розмірів $m \times n$, елементи яких належать полю P.

1.3. Множення матриць.

Раніше нами було введено означення добутку матриці $A_{m\times n}$ на n-вимірний вектор (див. означення 1.19). Переформулюємо його у дещо іншому вигляді.

Означення 2.4. Нехай

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Тоді *добутком матриці A на вектор $\vec{x}$* називається вектор

$$A\vec{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix}.$$

Так означене множення матриці розмірності $m \times n$ на n-вимірний вектор-стовпець (матрицю розмірності $n \times 1$) фактично задає певне відображення простору $\mathbb{R}^n$ в простір $\mathbb{R}^m$.

Означення 2.5. *Добутком матриці $A_{m\times n}$ на матрицю $B_{n\times p}$* називається матриця

$$A \cdot B = A \cdot \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_p \end{bmatrix} = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \dots & A\vec{b}_p \end{bmatrix}.$$

Результатом множення є матриця розмірності $m \times p$.

Приклад 2.2. Обчислити $A \cdot B$ та $B \cdot A$, якщо

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

Розв'язання. Оскільки матриця A розмірності 2×2, а матриця B розмірності 2×3, то множення матриць $A \cdot B$ визначене. Позначимо через $\vec{b}_1, \vec{b}_2, \vec{b}_3$ стовпці матриці B. Тоді формуємо з векторів $A\vec{b}_1$, $A\vec{b}_2$, $A\vec{b}_3$ матрицю AB:

$$A\vec{b}_1 = \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad A\vec{b}_2 = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad A\vec{b}_3 = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

$$AB = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

Множення $B \cdot A$ невизначене, оскільки розміри матриць A та B не узгоджуються з означенням 2.5. ■

Зауваження. Добуток двох матриць не завжди існує, а саме: добуток матриці розмірності $m \times n$ на матрицю розмірності $k \times p$ існує тоді і тільки тоді, коли $n = k$. Разом з цим добуток двох квадратних матриць одного і того ж порядку завжди існує.

$$\begin{array}{ccccc} A & \cdot & B & = & C \\ m \times n & & n \times p & & m \times p \end{array}$$

однакові

розмірність AB

Для обчислення добутку матриць використовується також зручне правило множення «рядок на стовпець». Воно базується на понятті добутку рядка однієї матриці на стовпець іншої.

Нехай маємо дві матриці $A_{m\times n} = [a_{ij}]$ і $B_{n\times p} = [b_{ij}]$. Під *добутком i-го рядка матриці A на j-й стовпчик матриці B* розуміють число

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Тоді добутком матриці $A_{m\times n}$ на матрицю $B_{n\times p}$ є матриця C розмірності $m \times p$, кожний елемент c_{ij} якої дорівнює добутку i-го рядка матриці A на j-й стовпчик матриці B.

Наприклад,

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

$$= \begin{bmatrix} 2\cdot 4 + 3\cdot 1 & 2\cdot 3 + 3\cdot(-2) & 2\cdot 6 + 3\cdot 3 \\ 1\cdot 4 + (-5)\cdot 1 & 1\cdot 3 + (-5)\cdot(-2) & 1\cdot 6 + (-5)\cdot 3 \end{bmatrix}.$$

Лема 2.1. *Нехай дано матриці $A \in M_{m\times n}$, $B \in M_{n\times p}$ і вектор $\vec{x} \in \mathbb{R}^p$. Тоді $A \cdot (B\vec{x}) = (AB)\vec{x}$.*

Доведення. Нехай $B = \begin{bmatrix} \vec{b}_1 & \ldots & \vec{b}_p \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$. Тоді згідно з властивостями множення матриці на вектор (теорема 1.11) маємо:

$$A \cdot (B\vec{x}) = A \cdot (x_1\vec{b}_1 + \ldots + x_p\vec{b}_p) = A(x_1\vec{b}_1) + \ldots + A(x_p\vec{b}_p) =$$

$$= x_1(A\vec{b}_1) + \ldots + x_p(A\vec{b}_p) = \begin{bmatrix} A\vec{b}_1 & \ldots & A\vec{b}_p \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \ldots \\ x_p \end{bmatrix} = (AB)\vec{x},$$

що і потрібно було довести. ■

Теорема 2.2. *Нехай A, B, C — матриці таких розмірів, для яких визначені операції, що наведені нижче; $\alpha \in \mathbb{R}$. Тоді мають місце тотожності:*

1. *$A(BC) = (AB)C$ (асоціативність множення матриць);*
2. *$A(B+C) = AB+AC$ (ліва дистрибутивність множення матриць відносно додавання);*

3. $(A+B)C = AC + BC$ *(права дистрибутивність множення матриць відносно додавання);*

4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Доведення.

1. Нехай маємо три матриці $A_{m\times n}$, $B_{n\times p} = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_p \end{bmatrix}$, $C_{p\times q} = \begin{bmatrix} \vec{c}_1 & \dots & \vec{c}_q \end{bmatrix}$. Тоді згідно з означенням 2.5 $BC = \begin{bmatrix} B\vec{c}_1 & \dots & B\vec{c}_q \end{bmatrix}$. Враховуючи лему 2.1, знаходимо:
$$\begin{aligned} A(BC) &= \begin{bmatrix} A(B\vec{c}_1) & \dots & A(B\vec{c}_q) \end{bmatrix} = \\ &= \begin{bmatrix} (AB)\vec{c}_1 & \dots & (AB)\vec{c}_q \end{bmatrix} = (AB)C. \end{aligned}$$
2. Нехай маємо три матриці $A_{m\times n}$, $B_{n\times p} = \begin{bmatrix} \vec{b}_1 & \dots & \vec{b}_p \end{bmatrix}$, $C_{n\times p} = \begin{bmatrix} \vec{c}_1 & \dots & \vec{c}_p \end{bmatrix}$. Тоді використовуючи означення 2.5 множення матриць і властивості множення матриці на вектор (теорема 1.11), знаходимо:
$$\begin{aligned} A(B+C) &= A\begin{bmatrix} (\vec{b}_1 + \vec{c}_1) & \dots & (\vec{b}_p + \vec{c}_p) \end{bmatrix} = \\ &= \begin{bmatrix} A(\vec{b}_1 + \vec{c}_1) & \dots & A(\vec{b}_p + \vec{c}_p) \end{bmatrix} = \\ &= \begin{bmatrix} A\vec{b}_1 + A\vec{c}_1 & \dots & A\vec{b}_p + A\vec{c}_p \end{bmatrix} = \\ &= \begin{bmatrix} A\vec{b}_1 & \dots & A\vec{b}_p \end{bmatrix} + \begin{bmatrix} A\vec{c}_1 & \dots & A\vec{c}_p \end{bmatrix} = AB + AC. \end{aligned}$$
3. Нехай матриці $A, B \in M_{m\times n}$ і матриця $C \in M_{n\times p}$. Легко перевірити, що матриці $(A+B)C$ і $AC+BC$ мають однакові розміри $m \times p$. Доведемо тепер, що відповідні елементи цих матриць однакові. Елемент i-го рядка і j-го стовпчика матриці $(A+B)C$ дорівнює
$$\sum_{k=1}^{n} (a_{ik} + b_{ik})c_{kj},$$
а елемент i-го рядка і j-го стовпчика матриці $AC+BC$ дорівнює
$$\sum_{k=1}^{n} a_{ik}c_{kj} + \sum_{k=1}^{n} b_{ik}c_{kj},$$
тобто вони однакові, що і потрібно було довести.
4. Нехай дано матриці $A_{m\times n} = [a_{ij}]$ і $B_{n\times p} = [b_{ij}]$, $\alpha \in \mathbb{R}$. Очевидно, що матриці $\alpha(AB)$, $(\alpha A)B$, $A(\alpha B)$ однакових розмірів. Розглянемо

елементи i-го рядка і j-го стовпчика цих матриць:

$$(\alpha(AB))_{ij} = \alpha \cdot \sum_{k=1}^{n} a_{ik} b_{kj};$$

$$((\alpha A)B)_{ij} = \sum_{k=1}^{n} (\alpha a_{ik}) b_{kj};$$

$$(A(\alpha B))_{ij} = \sum_{k=1}^{n} a_{ik} (\alpha b_{kj}).$$

Очевидно, вони однакові, і, отже, усі вказані матриці рівні.

■

Зауваження. Про матриці варто запам'ятати наступне:

- взагалі кажучи, $AB \neq BA$;
- матричне множення не допускає скорочень: якщо $AB = AC$, то це не означає, що $B = C$;
- якщо добуток AB є нульовою матрицею, то не можна стверджувати, що $A = O$ або $B = O$.

Матрицю A можна помножити саму на себе лише тоді, коли ця матриця квадратна. Натуральний степінь k матриці A розуміють як

$$A^k = \underbrace{A \cdot \ldots \cdot A}_{k \text{ множників}}.$$

1.4. Поняття оберненої матриці.

Нехай $M_n(\mathbb{R})$ — множина всіх квадратних матриць n-го порядку, заданих над полем $\mathbb{R}$. Матриця n-го порядку

$$\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} = \begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \ldots & \vec{e}_n \end{bmatrix}$$

називається *одиничною*. Вона позначається через I (або I_n, щоб підкреслити розмірність)[1]. Назва цієї матриці зумовлена тим, що для

[1] В деяких літературних джерелах одинична матриця позначається E. Ми використовуємо більш вживане позначення I від англійського «identity matrix» (одинична матриця).

довільної квадратної матриці A виконується тотожність $AI = IA = A$. Справді,

$$AI = A\begin{bmatrix}\vec{e}_1 & \dots & \vec{e}_n\end{bmatrix} = \begin{bmatrix}A\vec{e}_1 & \dots & A\vec{e}_n\end{bmatrix} = \begin{bmatrix}\vec{a}_1 & \dots & \vec{a}_n\end{bmatrix} = A.$$

Означення 2.6. Матриця A' називається *оберненою* до матриці A, якщо має місце рівність

$$A \cdot A' = A' \cdot A = I.$$

Теорема 2.3. *Якщо до матриці існує обернена матриця, то лише одна.*

Доведення. Припустимо, що для деякої матриці A існує дві різні обернені матриці A' і A''. Тоді з одного боку

$$A'AA'' = (A'A)A'' = IA'' = A'',$$

а з іншого

$$A'AA'' = A'(AA'') = A'I = A',$$

звідки $A' = A''$, що суперечить припущенню. Теорему доведено. ■

Приклад 2.3. Доведемо, що матриця $A = \begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}$ не має оберненої.

Нехай $\begin{bmatrix}a & b\\ c & d\end{bmatrix}$ — матриця, обернена до A. Тоді

$$\begin{bmatrix}1 & 1\\ 2 & 2\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix};$$

$$\begin{bmatrix}a+c & b+d\\ 2(a+c) & 2(b+d)\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$\begin{cases} a+c & = 1,\\ 2(a+c) & = 0,\\ b+d & = 0,\\ 2(b+d) & = 1. \end{cases}$$

Остання система є несумісною. Таким чином матриця A оберненої не має.

Обернену матрицю до матриці A позначатимемо A^{-1} (не можна писати $\frac{1}{A}$!). Матриця, до якої існує обернена, називається *оборотною*.

Теорема 2.4 (Обертання матриці другого порядку)**.** *Якщо* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *і* $ad - bc \neq 0$, *то матриця* A *є оборотною, причому* $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. *Якщо ж* $ad = bc$, *то до матриці* A *не існує оберненої.*

Доведення. Розглянемо першу частину теореми. Безпосередньо обчислимо

$$AA^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad - bc} \begin{bmatrix} ad - bc & ab - ba \\ cd - dc & ad - bc \end{bmatrix} = \\ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Абсолютно аналогічно можна переконатись, що і $A^{-1}A = I$. Отже, A^{-1} справді матриця, обернена до матриці A.

Доведення іншої частини теореми тут наводити не будемо, оскільки це твердження згодом (теорема 2.6) буде доведено в більш загальному вигляді. ■

Зазначимо, що вираз $ad - bc$ називається *детермінантом* (або *визначником*) матриці A.

Теорема 2.5 (Властивості оборотних матриць)**.** *1. Якщо* A *— оборотна матриця, то матриця* A^{-1} *також оборотна, причому* $(A^{-1})^{-1} = A$.

2. Якщо A *— оборотна матриця,* $c \neq 0$, *то тоді матриця* cA *також оборотна, причому* $(cA)^{-1} = \frac{1}{c} \cdot A^{-1}$.

3. Якщо квадратні матриці однакового порядку A *та* B *є оборотними, то матриця* AB *також оборотна, причому*

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

Доведення. 1. Щоб знайти обернену матрицю до матриці A^{-1}, потрібно знайти таку матрицю X, що $A^{-1}X = XA^{-1} = I$. Але якщо $X = A$, то ці рівності виконуються, а оскільки обернена матриця лише єдина, то це і є матриця A.

2. Маємо:

$$cA \cdot \frac{1}{c}A^{-1} = \frac{1}{c}A^{-1} \cdot cA = A \cdot A^{-1} = I.$$

3. Щоб знайти обернену матрицю до матриці AB потрібно знайти таку матрицю X, що $(AB)X = X(AB) = I$. Але якщо підставити $X = B^{-1}A^{-1}$, то враховуючи асоціативність множення матриць дістанемо:

$$ABB^{-1}A^{-1} = A(BB^{-1})A^{-1} = AIA^{-1} = (AI)A^{-1} = AA^{-1} = I;$$

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

■

Теорема 2.6. *Для того, щоб до матриці n-го порядку існувала обернена, необхідно і достатньо, щоб ранг цієї матриці дорівнював n.*

Доведення. *Необхідність.* Нехай маємо оборотну матрицю $A = \begin{bmatrix}\vec{a}_1 & \dots & \vec{a}_n\end{bmatrix}$. Це означає, що існує така матриця $X = \begin{bmatrix}\vec{x}_1 & \dots & \vec{x}_n\end{bmatrix}$, що $AX = XA = I$. Маємо:

$$AX = A\begin{bmatrix}\vec{x}_1 & \dots & \vec{x}_n\end{bmatrix} = \begin{bmatrix}A\vec{x}_1 & \dots & A\vec{x}_n\end{bmatrix},$$

звідки $A\vec{x}_i = \vec{e}_i$:

$$x_{1i}\vec{a}_1 + \ldots + x_{ni}\vec{a}_n = \vec{e}_i.$$

Таким чином, усі вектори $\vec{e}_i$, $i = \overline{1,n}$ лінійно виражаються через вектори-стовпці матриці A. Якщо припустити, що система векторів $\vec{a}_1, \ldots, \vec{a}_n$ є лінійно залежною, то тоді лінійно залежною мала б бути і система $\vec{e}_1, \ldots, \vec{e}_n$, що невірно. Отже, система векторів $\vec{a}_1, \ldots, \vec{a}_n$ є лінійно незалежною, а, значить, ранг матриці A дорівнює n.

Достатність. Нехай дано матрицю n-го порядку A, ранг якої дорівнює n. Шукатимемо матрицю $B = \begin{bmatrix}\vec{b}_1 & \dots & \vec{b}_n\end{bmatrix}$ таку, що $AB = I$. Маємо:

$$\begin{bmatrix}A\vec{b}_1 & \dots & A\vec{b}_n\end{bmatrix} = I,$$

звідки $A\vec{b}_i = \vec{e}_i$ для всіх $i = \overline{1,n}$. Це n систем, які записані в матрично-векторній формі. Оскільки ранг основної матриці цих систем дорівнює n, а ранг розширеної матриці більше n бути не може (і менше теж), то за теоремою Кронекера–Капеллі усі вони сумісні, а значить існує матриця B.

Розглянемо систему $B\vec{x} = \vec{0}$ як матричну рівність. Домножимо її зліва на A: $AB\vec{x} = A\vec{0}$. Маємо: $AB\vec{x} = (AB)\vec{x} = I\vec{x} = \vec{x}$, $A\vec{0} = \vec{0}$. Тобто $\vec{x} = 0$ і система $B\vec{x} = \vec{0}$ має лише нульовий розв'язок. Це означає, що ранг матриці B дорівнює n. Тоді за доведеним вище існує така матриця C, що $BC = I$.

Нарешті розглянемо добутки:

$$ABC = A(BC) = AI = A$$
$$ABC = (AB)C = IC = C,$$

звідки $C = A$. Отже, $AB = BA = I$, а це і означає, що матриця B є оберненою до матриці A. ■

Теорема 2.7 (Основна теорема оборотних матриць)**.** *Нехай $A \in M_n(\mathbb{R})$. Наступні твердження є еквівалентними:*

1) матриця A — оборотна;
2) ранг матриці A дорівнює n;
3) стовпці матриці A лінійно незалежні;
4) система $A\vec{x} = \vec{b}$ є визначеною для кожного $\vec{b} \in \mathbb{R}^n$;
5) однорідна система $A\vec{x} = \vec{0}$ має лише нульовий розв'язок;
6) зведена східчаста форма матриці A рівна одиничній матриці I_n.

Зауважимо, що множина $M_n(\mathbb{R})$ квадратних матриць n-го порядку разом з операціями додавання і множення матриць утворює *кільце*. Воно некомутативне (оскільки множення матриць некомутативне), але є кільцем з одиницею (роль одиничного елемента виконує одинична матриця I_n).

Множина оборотних матриць порядку n над полем $\mathbb{R}$ з операцією множення матриць утворює групу, яка називається *загальною лінійною групою* (або *повною лінійною групою*) і позначається $GL_n(\mathbb{R})$. Це випливає з того, що:

- добуток двох оборотних матриць є оборотною матрицею (третя частина теореми 2.5);
- множення матриць асоціативне;
- одиничним елементом групи є одинична матриця I_n;
- до кожної елемента A з цієї множини існує обернений A^{-1} (обернена матриця).

1.5. Елементарні матриці. Знаходження оберненої матриці.

Означення 2.7. Матриця n-го порядку називається *елементарною*, якщо вона одержана з відповідної одиничної матриці шляхом одного елементарного перетворення над її рядками.

Оскільки є три типи елементарних перетворень рядків матриці, то відповідно є і три типи елементарних матриць. Розглянемо їх на прикладі.

Приклад 2.4. Встановити, чи є елементарними матриці

$$E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

і обчислити добутки E_1A, E_2A, E_3A, де

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}.$$

Розв'язання. Матриці E_1, E_2, E_3 є елементарними, оскільки матриця E_1 одержана з одиничної перестановкою першого і третього рядків, E_2 — множенням другого рядка на 3, а E_3 — додаванням до третього рядка першого, помноженого на -2.

Далі знаходимо:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} g & h & i \\ d & e & f \\ a & b & c \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g-2a & h-2b & i-2c \end{bmatrix}.$$

■

Останній приклад наводить на думку, що множення матриці зліва на елементарну матрицю, яку одержану з одиничної шляхом деякого елементарного перетворення її рядків, здійснює з вихідною матрицею таке ж саме елементарне перетворення її рядків. Виявляється, що це не випадково і має місце загальне твердження, доведення якого ми опускаємо.

Теорема 2.8. *Нехай E — елементарна матриця n-го порядку, яку одержано з одиничної матриці I_n з допомогою певного рядкового елементарного перетворення. Якщо ж це саме перетворення застосувати до матриці A, то результатом буде матриця EA.*

За основною теоремою оборотних матриць 2.7 кожну оборотну матрицю A за допомогою елементарних перетворень можна звести до одиничної. З теореми 2.8 слідує, що тоді має місце рівність

$$E_k E_{k-1} \dots E_2 E_1 A = I$$

для деяких елементарних матриць $E_1, E_2, \dots, E_k$. Помноживши останню рівність справа на A^{-1}, дістанемо:

$$E_k E_{k-1} \dots E_2 E_1 I = A^{-1}.$$

Одержане співвідношення приводить до способу відшування оберненої матриці, який полягає в наступному. Випишемо поруч задану матрицю A та одиничну матрицю I і дістанемо розширену матрицю $[A \mid I]$. За допомогою рядкових елементарних перетворень зведемо цю матрицю до виду $[I \mid B]$. Тоді матриця B і є оберненою матрицею до матриці A:

$$[A \mid I] \sim \dots \sim \left[I \mid A^{-1}\right].$$

Приклад 2.5. Знайти матрицю, обернену до матриці

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

Розв'язання. Запишемо матрицю $[A \mid I]$ і з допомогою елементарних перетворень приведемо її до виду $\left[I \mid A^{-1}\right]$:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 6 & 2 & 1 & 0 & 1 & 0 \\ 5 & 1 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & -5 & -6 & 1 & 0 \\ 0 & 1 & -2 & -5 & 0 & 1 \end{array}\right] \sim$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -5 & 0 & 1 \\ 0 & 2 & -5 & -6 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -5 & 0 & 1 \\ 0 & 0 & -1 & 4 & 1 & -2 \end{array}\right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & 1 & -2 \\ 0 & 1 & 0 & -13 & -2 & 5 \\ 0 & 0 & -1 & 4 & 1 & -2 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & 1 & -2 \\ 0 & 1 & 0 & -13 & -2 & 5 \\ 0 & 0 & 1 & -4 & -1 & 2 \end{array}\right].$$

Таким чином,

$$A^{-1} = \begin{bmatrix} 5 & 1 & -2 \\ -13 & -2 & 5 \\ -4 & -1 & 2 \end{bmatrix}.$$

■

1.6. Матричний метод розв'язування систем лінійних рівнянь. Нехай маємо систему лінійних рівнянь $A\vec{x} = \vec{b}$, де A — квадратна матриця n-го порядку. Цю систему можна сприймати як матричне рівняння з невідомою матрицею $\vec{x}$ розмірів $n \times 1$. Припустимо, що A — оборотна. Тоді має місце рівність

$$A^{-1}(A\vec{x}) = A^{-1}\vec{b}.$$

Але звідси знаходимо: $A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} = I\vec{x} = \vec{x}$, тому

$$\vec{x} = A^{-1}\vec{b}.$$

Такий метод відшукання розв'язку системи лінійних рівнянь називається *матричним*.

Приклад 2.6. Розв'язати систему

$$\begin{cases} x_2+2x_3 = -4, \\ x_1 +3x_3 = -5, \\ 4x_1-3x_2+8x_3 = -12. \end{cases}$$

Розв'язання. Дана система може бути записана в матрично-векторному вигляді так:

$$A\vec{x} = \vec{b},$$

де

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -4 \\ -5 \\ -12 \end{bmatrix}.$$

Можна переконатись в тому, що A має обернену (зробіть перевірку самостійно), причому:

$$A^{-1} = \begin{bmatrix} -\frac{9}{2} & 7 & -\frac{3}{2} \\ -2 & 4 & -1 \\ \frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix}.$$

Тоді

$$\vec{x} = \begin{bmatrix} -\frac{9}{2} & 7 & -\frac{3}{2} \\ -2 & 4 & -1 \\ \frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} -4 \\ -5 \\ -12 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$$

Відповідь. $\{(1, 0, -2)\}$. ■

§ 2. Детермінанти (визначники) n-го порядку

Розглянемо квадратну матрицю $A = [a_{ij}]$ n-го порядку. Поставимо у відповідність даній матриці деяке число, яке певним чином буде характеризувати її властивості. Це число називатимемо детермінантом (визначником) матриці.

Існує кілька різних підходів до введення поняття визначника, зокрема комбінаторний (вираження детермінанта матриці через її елементи, див. [4, 16]), індуктивний (детермінант матриці n-го порядку означається через детермінант матриці $n-1$-го порядку, див. [10, 20, 22]), аксіоматичний (визначник як індикатор лінійної залежності системи векторів, див. [14]). Історично поняття детермінанту виникло набагато раніше, ніж була розроблена теорія матриць (вперше поняття детермінанта було введено Секі Такакадзу[1] в 1683 році і незалежно Годфрідом Лейбніцем[2] у 1693 році; теорія матриць формувалась у середині XIX століття). Ми зупинимось на індуктивному підході.

Означення 2.8. Детермінантом 1×1 матриці $A = [a_{11}]$ назвемо число a_{11}. *Детермінантом (визначником)* $n \times n$ матриці, $n > 1$ називається число

$$\sum_{k=1}^{n} a_{1k}(-1)^{1+k} M_{1k},$$

де M_{1k} — детермінант матриці порядку $n-1$, яка одержується з заданої шляхом викреслення першого рядка та k-го стовпця.

Позначатимемо визначник матриці A через $\det A$ або $|A|$. Іноді, щоб не говорити «детермінант матриці n-го порядку», ми вживатимемо термін «детермінант n-го порядку».

[1] Seki Takakazu, відомий також як Seki Kowa (1640–1708) — японський математик.

[2] Gottfried Wilhelm Leibniz (1646–1716) — німецький математик і філософ.

Перед тим, як перейти до формул для обчислення визначника другого та третього порядків сформулюємо декілька допоміжних означень.

Означення 2.9. *Мінором елемента* a_{ij} матриці $A = [a_{ij}]$ називається детермінант матриці, яка утворюється з матриці A викресленням i-го рядка та j-го стовпця. Позначається M_{ij}.

Означення 2.10. *Алгебраїчним доповненням елемента* a_{ij} матриці $A = [a_{ij}]$ називається число $(-1)^{i+j}M_{ij}$. Позначається A_{ij}.

Відповідно до вказаних означень можна в такий спосіб переформулювати означення 2.8.

Означення 2.11. *Детермінантом* матриці A n-го порядку ($n > 1$) називається сума добутків елементів першого рядка цієї матриці на їх алгебраїчні доповнення:

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n}.$$

2.1. Формули для обчислення визначників матриць другого та третього порядків.

Розглянемо визначник другого порядку:

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

$$= a_{11}a_{22} - a_{12}a_{21}.$$

Схематично обчислення визначника другого порядку можна зобразити в такий спосіб:

$$\begin{vmatrix} \bullet & \circ \\ \circ & \bullet \end{vmatrix} - \begin{vmatrix} \circ & \bullet \\ \bullet & \circ \end{vmatrix}$$

Обчислимо детермінант матриці третього порядку:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(-1)^{1+1}M_{11} + a_{12}(-1)^{1+2}M_{12} + a_{13}(-1)^{1+3}M_{13} =$$

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

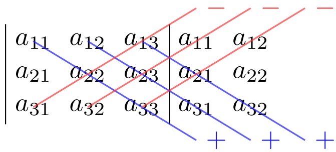
$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

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

Схематично запам'ятати формулу для обчислення детермінанту третього порядку допомагає наступна схема:

$$\begin{vmatrix} \bullet & \circ & \circ \\ \circ & \bullet & \circ \\ \circ & \circ & \bullet \end{vmatrix} + \begin{vmatrix} \circ & \bullet & \circ \\ \circ & \circ & \bullet \\ \bullet & \circ & \circ \end{vmatrix} + \begin{vmatrix} \circ & \circ & \bullet \\ \bullet & \circ & \circ \\ \circ & \bullet & \circ \end{vmatrix} - \begin{vmatrix} \circ & \circ & \bullet \\ \circ & \bullet & \circ \\ \bullet & \circ & \circ \end{vmatrix} - \begin{vmatrix} \circ & \bullet & \circ \\ \bullet & \circ & \circ \\ \circ & \circ & \bullet \end{vmatrix} - \begin{vmatrix} \bullet & \circ & \circ \\ \circ & \circ & \bullet \\ \circ & \bullet & \circ \end{vmatrix}.$$

Інше мнемонічне правило (правило Саррюса[1]) для запам'ятовування формули визначника третього порядку полягає в наступному. Випишемо визначник і поруч справа допишемо його перші два стовпці. Обчислимо добутки елементів на шести діагоналях утвореної таблиці чисел так, як показано на рисунку. Ті доданки, які лежать на головній діагоналі матриці або паралельно їй беруться зі знаком «+», а інші зі знаком «−».



2.2. Теорема Лапласа.

Означення 2.11 детермінанта матриці, тобто подання його через суму добутків елементів першого рядка на відповідні алгебраїчні доповнення, ще називають розкладом визначника за першим рядком. Виявляється, що сума добутків елементів довільного іншого рядка (або навіть стовпця) матриці на їх алгебраїчні доповнення також дорівнюють визначнику матриці. Цей фундаментальний факт теорії детермінантів носить назву «теорема Лапласа[2]», доведення якого ми наводимо в пункті 2.3.

Теорема 2.9 (Теорема Лапласа про розклад визначника)**.** *Детермінант матриці* $A = [a_{ij}]$ *дорівнює сумі добутків елементів*

---

[1]Pierre Frédéric Sarrus (1798–1861) — французький математик.

[2]Pierre Simon de Laplace (1749–1827) — французький математик, фізик і астроном.

довільного рядка (стовпця) матриці на їх алгебраїчні доповнення:

$$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}, \qquad i = \overline{1,n},$$
$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}, \qquad j = \overline{1,n}.$$

Приклад 2.7. Обчислити

$$\Delta = \begin{vmatrix} 3 & 7 & 1 & -1 & 4 \\ 0 & 2 & -2 & 5 & 3 \\ 0 & 0 & 8 & 1 & 0 \\ 0 & 0 & 3 & 3 & -1 \\ 0 & 0 & 0 & 2 & 0 \end{vmatrix}$$

Розв'язання. Розкладемо цей визначник за першим стовпцем:

$$\Delta = 3\cdot A_{11} + 0\cdot A_{21} + 0\cdot A_{31} + 0\cdot A_{41} + 0\cdot A_{51} = 3\cdot(-1)^{1+1}\cdot \begin{vmatrix} 2 & -2 & 5 & 3 \\ 0 & 8 & 1 & 0 \\ 0 & 3 & 3 & -1 \\ 0 & 0 & 2 & 0 \end{vmatrix}.$$

Отже, обчислення визначника п'ятого порядку звелось до обчислення визначника четвертого порядку. Розкладемо і його за першим стовпцем:

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

Одержаний визначник третього порядку обчислимо, розклавши його за третім рядком. Врешті-решт дістанемо:

$$\Delta = 2\cdot 6\cdot(-1)^{3+2}\cdot \begin{vmatrix} 8 & 0 \\ 3 & -1 \end{vmatrix} = 12\cdot(-1)\cdot(-8) = 96.$$

■

Теорема 2.10. *Детермінант трикутної матриці дорівнює добутку елементів, які розташовані на головній діагоналі цієї матриці.*

Доведення. Розглянемо доведення цієї теореми для випадку верхньої трикутної матриці. Доведення проведемо індукцією по розміру матриці. База індукції

$$\begin{vmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{vmatrix} = a_{11}a_{22}$$

виконується. Припустимо, що детермінант кожної верхньої трикутної матриці n-го порядку дорівнює добутку діагональних елементів. Розглянемо трикутну матрицю D $n+1$-го порядку. Розклавши детермінант цієї матриці за останнім рядком дістанемо, що він дорівнює добутку елемента останнього рядка і останного стовпчика на його алгебраїчне доповнення, яке за індуктивним припущенням рівне $a_{11}a_{22}\ldots a_{nn}$. Отже, $\det D = a_{n+1,n+1}\cdot(-1)^{2n+2}a_{11}\ldots a_{nn} = a_{11}a_{22}\ldots a_{n+1,n+1}$. ■

2.3. Доведення теореми Лапласа.

Для доведення теореми Лапласа використаємо два допоміжних твердження. Перше з них полягає в тому, що сума добутків елементів першого рядка на відповідні алгебраїчні доповнення дорівнює сумі добутків елементів першого стовпця на відповідні алгебраїчні доповнення.

Лема 2.2. *Нехай A — $n\times n$ матриця. Тоді*

$$a_{11}A_{11}+a_{12}A_{12}+\ldots+a_{1n}A_{1n} = \det A = a_{11}A_{11}+a_{21}A_{21}+\ldots+a_{n1}A_{n1}. \quad (2.4)$$

Доведення. Доведення леми проведемо індукцією по n. Для $n=1$ твердження очевидне. Припустимо, що твердження має місце для всіх матриць $n-1$-го порядку (припущення індукції). Нехай A — довільна матриця n-го порядку. Зазначимо, що ліва і права частини рівності (2.4) містять доданок $a_{11}A_{11}$, тому ним можна знехтувати.

В правій частині рівності (2.4) i-ий доданок $a_{i1}A_{i1}$ дорівнює $a_{i1}\cdot(-1)^{i+1}M_{i1}$, де

$$M_{i1} = \begin{vmatrix} a_{12} & a_{13} & \ldots & a_{1j} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i-1,2} & a_{i-1,3} & \ldots & a_{i-1,j} & \ldots & a_{i-1,n} \\ a_{i+1,2} & a_{i+1,3} & \ldots & a_{i+1,j} & \ldots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \ldots & a_{nj} & \ldots & a_{nn} \end{vmatrix}.$$

Розкладемо цей визначник за першим рядком. j-ий доданок цього розкладу має вигляд $a_{1j}\cdot(-1)^{1+j-1}M_{1i,1j}$, де позначення $M_{kl,rs}$ означає детермінант матриці, що одержується з матриці A викресленням рядків k та l і стовпців r та s. Тоді доданок правої частини рівності (2.4), що містить $a_{i1}a_{1j}$ набуває вигляду

$$a_{i1}(-1)^{i+1}a_{1j}(-1)^{1+j-1}M_{1i,1j} = (-1)^{i+j+1}a_{i1}a_{1j}M_{1i,ij}.$$

Тепер знайдемо доданок, що містить $a_{i1}a_{1j}$ в лівій частині рівності (2.4). Множник a_{1j} з'являється лише в j-му доданку $a_{1j}A_{1j} = a_{1j}(-1)^{1+j}M_{1j}$, де

$$M_{1j} = \begin{vmatrix} a_{21} & \dots & a_{2,j-1} & a_{2,j+1} & \dots & a_{2n} \\ a_{31} & \dots & a_{3,j-1} & a_{3,j+1} & \dots & a_{3n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{i,j-1} & a_{i,j+1} & \dots & a_{in} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}.$$

За припущенням індукції ми можемо здійснити розклад визначника M_{1j} за першим стовпцем. i-ий доданок цього розкладу дорівнює $a_{i1} \cdot (-1)^{(i-1)+1}M_{1i,1j}$, тому доданок, що містить множник $a_{i1}a_{1j}$ лівої частини рівності (2.4) дорівнює

$$a_{1j}(-1)^{1+j}a_{i1}(-1)^{(i-1)+1}M_{1i,1j} = (-1)^{i+j+1}a_{i1}a_{1j}M_{1i,1j},$$

звідки знаходимо, що ліва і права частини рівності (2.4) однакові. ■

Лема 2.3. *Нехай A — матриця n-го порядку, а B одержується шляхом перестановки двох рядків (стовпців) матриці A. Тоді*

$$\det B = -\det A.$$

Доведення. Знову доведення проведемо методом математичної індукції за розміром матриці. Твердження леми легко перевіряється для $n = 2$. Припустимо, що воно має місце для всіх матриць розмірів $(n-1) \times (n-1)$. Доведемо, що тоді і для матриць n-го порядку воно також має місце. Спочатку доведемо це для випадку, коли переставляються два сусідні рядки матриці A (нехай це r-й та $r+1$-й рядки).

За лемою 2.2 детермінант матриці B можна обчислити через розклад за першим стовпцем. i-ий доданок цього розкладу має вигляд $(-1)^{i+1}b_{i1}M_{i1}(B)$ (позначення $M_{ij}(B)$ означає мінор елемента i-го рядка і j-го стовпця матриці B). Якщо $i \neq r$ та $i \neq r+1$, то $b_{i1} = a_{i1}$ і $M_{i1}(B)$ є детермінантом $n-1$-го порядку, що співпадає з мінором $M_{i1}(A)$ за винятком двох сусідніх рядків, що переставлено. Але тоді за припущенням індукції $M_{i1}(B) = -M_{i1}(A)$ при $i \neq r$ та $i \neq r+1$.

Якщо $i = r$, то $b_{i1} = a_{r+1,1}$ і $M_{i1}(B) = M_{r+1,1}(A)$. Тоді r-ий доданок в розкладі $\det B$ має такий вигляд:

$$(-1)^{r+1}b_{r1}M_{r1}(B) = (-1)^{r+1}a_{r+1,1}M_{r+1,1}(A) = -(-1)^{(r+1)+1}a_{r+1,1}M_{r+1,1}(A).$$

Аналогічно, якщо $i = r+1$, то $b_{i1} = a_{r1}$, $M_{i1}(B) = M_{r1}(A)$ і $r+1$-ий доданок в розкладі $\det B$ дорівнює

$$(-1)^{(r+1)+1} b_{r+1,1} M_{r+1,1}(B) = (-1)^r a_{r1} M_{r1}(A) = -(-1)^{r+1} a_{r1} M_{r1}(A).$$

Іншими словами, r-ий та $r+1$-ий доданки в розкладі $\det B$ за першим стовпцем протилежні за знаком і рівні за абсолютною величиною відповідно з $r+1$-им та r-им доданками розкладу $\det A$ за першим стовпцем.

Підставимо усі ці результати в $\det B$ і знову застосуємо лему 2.2:

$$\det B = \sum_{i=1}^{n} (-1)^{i+1} b_{i1} M_{i1}(B) =$$
$$= \sum_{\substack{i=1 \\ i \neq r, r+1}}^{n} (-1)^{i+1} b_{i1} M_{i1}(B) + (-1)^{r+1} b_{r1} M_{r1}(B) +$$
$$+ (-1)^{(r+1)+1} b_{r+1,1} M_{r+1,1}(B) =$$
$$= \sum_{\substack{i=1 \\ i \neq r, r+1}}^{n} (-1)^{i+1} a_{i1}(-M_{i1}(A)) - (-1)^{(r+1)+1} a_{r+1,1} M_{r+1,1}(A) -$$
$$- (-1)^{r+1} a_{r1} M_{r1}(A) = -\sum_{i=1}^{n} (-1)^{i+1} a_{i1} M_{i1}(A) = -\det A.$$

Отже, ми довели твердження леми для випадку, коли здійснюється перестановка двох сусідніх рядків. Для доведення того, що воно має місце для перестановки двох довільних рядків лише відзначимо, що перестановку двох рядків з номерами, скажімо, r та s $(r < s)$ можна здійснити шляхом $2(s-r)-1$ перестановок сусідніх рядків. Оскільки це число перестановок непарне і кожна з них змінює знак детермінанта, зрештою детермінант матриці B буде протилежним за знаком із детермінантом матриці A. ■

Тепер можемо перейти до *доведення теореми Лапласа.*

Нехай матриця B одержується з A шляхом переміщення i-го рядка вгору на перше місце з допомогою $i-1$ перестановки сусідніх рядків матриці. За лемою 2.3 $\det B = (-1)^{i-1} \det A$. Але $b_{1j} = a_{ij}$ і $M_{1j}(B) = M_{ij}(A)$ для всіх $j = 1, \ldots, n$.

Тоді

$$\det A = (-1)^{i-1} \det B = (-1)^{i-1} \sum_{j=1}^{n} (-1)^{1+j} b_{1j} M_{1j}(B) =$$
$$= (-1)^{i-1} \sum_{j=1}^{n} (-1)^{1+j} a_{ij} M_{ij}(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}(A),$$

тобто ми одержали формулу розкладу детермінанта матриці A за i-м рядком.

Доведення розкладу детермінанта за стовпцем аналогічне з урахуванням леми 2.2.

2.4. Вираження детермінанта матриці через її елементи.

Індуктивне означення визначника матриці n-го порядку породжує запитання: а чи існує формула для обчислення визначника матриці через її елементи? Для відповіді на це запитання введемо деякі нові означення.

Означення 2.12. *Перестановкою n* елементів називається будь-яка упорядкована множина цих елементів.

Перестановки відрізняються між собою лише порядком елементів. Число усіх можливих перестановок з n елементів дорівнює $P_n = 1 \cdot 2 \cdot \ldots \cdot n = n!$. Якщо елементами перестановки є числа $1, 2, \ldots, n$, то ту з $n!$ перестановок цих елементів, в якій вони розміщені в порядку зростання, називатимемо *нормально впорядкованою*.

Означення 2.13. *Інверсією* називається таке розташування двох чисел в перестановці, коли більше число стоїть попереду (лівіше) меншого.

Наприклад, перестановка 2, 4, 3, 1 має чотири інверсії: 2, 4, 3 знаходяться лівіше від 1; 4 лівіше від числа 3. Перестановка, в якій кількість інверсій є парним числом називається *парною*, в іншому випадку — *непарною*.

Теорема 2.11. *Нехай $A = [a_{ij}]$ — матриця n-го порядку. Тоді*

$$\det A = \sum_{n!} (-1)^{N(\alpha_1,\ldots,\alpha_n)} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n}, \tag{2.5}$$

де $N(\alpha_1, \ldots, \alpha_n)$ — число інверсій у перестановці $\alpha_1, \ldots, \alpha_n$.

Доведення. Доведення проведемо методом математичної індукції.

Нехай $n = 2$. Тоді формула (2.5) набуває вигляду

$$\det A = (-1)^{N(1,2)} a_{11} a_{22} + (-1)^{N(2,1)} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21},$$

що узгоджується з раніше виведеною формулою.

Припустимо, що твердження теореми справедливе при $n = k - 1$. Це означає, що детермінант довільної матриці $A = [a_{ij}]$ порядку $k-1$ дорівнює сумі $\sum (-1)^{N(\alpha_1,\ldots,\alpha_{k-1})} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{k-1,\alpha_{k-1}}$.

Розглянемо тепер довільну матрицю $A = [a_{ij}]$ k-го порядку.

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1k}A_{1k} = \sum_{i=1}^{k} a_{1i}A_{1i} =$$

$$= \sum_{i=1}^{k} \left(a_{1i} \cdot (-1)^{1+i} \sum_{(k-1)!} (-1)^{N(\alpha_1^{(i)},\ldots,\alpha_{k-1}^{(i)})} a_{2\alpha_1^{(i)}} a_{3\alpha_2^{(i)}} \ldots a_{k,\alpha_{k-1}^{(i)}} \right),$$

де внутрішня сума береться по всім можливим перестановкам $(\alpha_1^{(i)}, \ldots, \alpha_{k-1}^{(i)})$ чисел $1, \ldots, i-1, i+1, \ldots, k$.

Внесемо множник $a_{1i} \cdot (-1)^{1+i}$ під знак внутрішньої суми. Тоді

$$\det A = \sum_{i=1}^{k} \left(\sum_{(k-1)!} (-1)^{(1+i)+N(\alpha_1^{(i)},\ldots,\alpha_{k-1}^{(i)})} a_{1i} a_{2\alpha_1^{(i)}} a_{3\alpha_2^{(i)}} \ldots a_{k,\alpha_{k-1}^{(i)}} \right).$$

Відзначимо, що якщо в перестановці $(\alpha_1^{(i)}, \ldots, \alpha_{k-1}^{(i)})$ було N інверсій, то в перестановці $(i, \alpha_1^{(i)}, \ldots, \alpha_{k-1}^{(i)})$ буде рівно $i - 1 + N$ інверсій. Але

$$(-1)^{(1+i)+N(\alpha_1^{(i)},\ldots,\alpha_{k-1}^{(i)})} = (-1)^{(i-1)+N(\alpha_1^{(i)},\ldots,\alpha_{k-1}^{(i)})} = (-1)^{N(i,\alpha_1^{(i)},\ldots,\alpha_{k-1}^{(i)})},$$

звідки

$$\det A = \sum_{i=1}^{k} \left(\sum_{(k-1)!} (-1)^{N(i,\alpha_1^{(i)},\ldots,\alpha_{k-1}^{(i)})} a_{1i} a_{2\alpha_1^{(i)}} a_{3\alpha_2^{(i)}} \ldots a_{k,\alpha_{k-1}^{(i)}} \right),$$

що і потрібно було показати. Отже, формула (2.5) справедлива. ■

Зауважимо, що формула (2.5) непридатна для обчислення визначників достатньо великих порядків. Якщо сучасний комп'ютер, що виконує 10^{12} операцій на секунду, буде так обраховувати значення детермінанта матриці 50×50 (для сучасних застосувань це відносно невисокий порядок детермінанта), то йому знадобиться виконати $50! \approx 3 \cdot 10^{64}$ операцій за близько $3 \cdot 10^{52}$ секунд, тобто 10^{45} років. Для порівняння: астрономи вважають, що «вік» Всесвіту становить близько 10^{10} років. Тобто формула (2.5) носить швидше теоретичний характер, аніж практичний. Проте є й інші методи обчислення детермінантів.

2.5. Властивості детермінантів.

Нехай $A = [a_{ij}]$ — квадратна матриця.

Властивість 1. *Якщо A містить нульовий рядок (стовпець), то* $\det A = 0$.

Доведення. Справді, нехай усі елементи i-го рядка матриці дорівнюють нулю. Тоді розклавши детермінант матриці за цим рядком знайдемо:

$$\det A = 0 \cdot A_{i1} + \dots + 0 \cdot A_{in} = 0.$$

■

Означення 2.14. Матриця $A^T = [a_{ji}]$ називається *транспонованою* до матриці $A = [a_{ij}]$.

Наприклад, якщо $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, то $A^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$. Більш детально властивості транспонування матриць будуть розглянуті у розділі 7 (стор. 170).

Властивість 2. *При транспонуванні детермінант матриці не змінюється:* $\det A = \det A^T$.

Доведення. Доведення проведемо методом математичної індукції. Для $n = 2$ твердження, вочевидь, має місце. Припустимо, що для кожної матриці A k-го порядку $\det A = \det A^T$. Розглянемо довільну матрицю $B = [a_{ij}]$ $k+1$-го порядку та транспоновану до неї B^T. Розклавши детермінант матриці B за першим рядком, а детермінант матриці B^T за першим стовпцем, дістанемо:

$$\det B = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1,k+1}A_{1,k+1},$$
$$\det B^T = a_{11}A_{11}^T + a_{12}A_{12}^T + \ldots + a_{1,k+1}A_{1,k+1}^T.$$

А оскільки для всіх j $A_{1j} = A_{1j}^T$ за індукційним припущенням, то $\det B = \det B^T$. ■

Наслідок. *Кожна властивість, доведена для рядків детермінанта, має місце і для стовпців.*

Властивість 3. *Якщо матриця B одержується з матриці A переставленням двох рядків (стовпців), то* $\det B = -\det A$.

Ця властивість нами вже доведена у лемі 2.3.

Властивість 4. *Якщо матриця A має два однакових рядки (стовпці), то* $\det A = 0$.

Доведення. Справді, якщо переставити ці два рядки такої матриці, то, з одного боку, матриця (а значить і детермінант) не зміниться, а з іншого — за властивістю 3 змінить знак на протилежний, що можливо лише в тому випадку, коли $\det A = 0$. ■

Властивість 5. *Якщо матриця B одержується з матриці A множенням деякого рядка (стовпця) матриці A на число k, то $\det B = k \cdot \det A$.*

Доведення. Нехай

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ ka_{i1} & \dots & ka_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}.$$

Тоді розкладаючи детермінант матриці B за i-им рядком, знаходимо:

$$\det B = ka_{i1}A_{i1} + \ldots + ka_{in}A_{in} = k \cdot \det A.$$

■

Властивість 6. *Якщо квадратні матриці A, B, C однакові за винятком елементів i-го рядка (стовпця), причому i-ий рядок (стовпець) матриці C дорівнює сумі i-их рядків (стовпців) матриць A та B, то*

$$\det C = \det A + \det B.$$

Доведення. Нехай

$$A = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & \dots & b_{1j} & \dots & a_{1n} \\ a_{21} & \dots & b_{2j} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & b_{nj} & \dots & a_{nn} \end{bmatrix},$$

$$C = \begin{bmatrix} a_{11} & \dots & a_{1j} + b_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j} + b_{2j} & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nj} + b_{nj} & \dots & a_{nn} \end{bmatrix}.$$

Тоді

$$\begin{aligned}\det C =&(a_{1j}+b_{1j})A_{1j}+\ldots+(a_{nj}+b_{nj})A_{nj}=\\ =&a_{1j}A_{1j}+\ldots+a_{nj}A_{nj}+b_{1j}A_{1j}+\ldots+b_{nj}A_{nj}=\det A+\det B.\end{aligned}$$

■

Властивість 7. *Якщо до деякого рядка (стовпця) матриці додати будь-який інший рядок (стовпець), помножений на довільне число, то детермінант матриці не зміниться.*

Доведення. Нехай

$$A=\begin{bmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{i1} & \dots & a_{in}\\ \vdots & \ddots & \vdots\\ a_{j1} & \dots & a_{jn}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix},\quad B=\begin{bmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{i1} & \dots & a_{in}\\ \vdots & \ddots & \vdots\\ a_{j1}+ka_{i1} & \dots & a_{jn}+ka_{in}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix}.$$

Слід довести, що $\det B=\det A$. Розкладаючи детермінант матриці B за j-м рядком, дістанемо після спрощень:

$$\det B=\det A+k\cdot\det\begin{bmatrix} a_{11} & \dots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{i1} & \dots & a_{in}\\ \vdots & \ddots & \vdots\\ a_{i1} & \dots & a_{in}\\ \vdots & \ddots & \vdots\\ a_{n1} & \dots & a_{nn}\end{bmatrix}=\det A$$

(тут ми використали властивість 4). ■

Розглянуті властивості дають ефективний інструмент для обчислення визначників. Розглянемо це на прикладах.

Приклад 2.8. Обчислити

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

Розв'язання. Маємо:

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

■

Приклад 2.9. Обчислити $\det A$, якщо

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

Розв'язання. Спочатку за властивістю 5 винесемо трійку з другого рядка визначника:

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

Далі переставимо місцями перший і другий рядки (властивість 3):

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

Тепер додамо до третього рядка перший, що помножений на -2, а до четвертого перший, що помножений на -5 (властивість 7):

$$\det A = -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -4 & 5 \\ 0 & 4 & 7 & 3 \\ 0 & -1 & 2 & -9 \end{vmatrix}.$$

Розкладемо детермінант за першим стовпцем:

$$\det A = -3 \begin{vmatrix} 2 & -4 & 5 \\ 4 & 7 & 3 \\ -1 & 2 & -9 \end{vmatrix}.$$

Додамо до першого рядка третій, що помножений на 2, а до другого третій, що помножений на 4:

$$\det A = -3\begin{vmatrix} 0 & 0 & -13 \\ 0 & 15 & -33 \\ -1 & 2 & -9 \end{vmatrix} = 3\begin{vmatrix} -1 & 2 & -9 \\ 0 & 15 & -33 \\ 0 & 0 & -13 \end{vmatrix} = 3\cdot(-1)\cdot 15\cdot(-13) = 585.$$

■

2.6. Детермінанти елементарних матриць. Критерій оборотності матриць.

Теорема 2.12. *Нехай E — елементарна матриця порядку n.*

1. *Якщо E одержується з одиничної матриці переставленням двох рядків, то* $\det E = -1$.
2. *Якщо E одержується з одиничної матриці множенням рядка на деяке число $k \neq 0$, то* $\det E = k$.
3. *Якщо E одержується з одиничної матриці додаванням до деякого рядка якогось іншого, помноженого на певне число, то* $\det E = 1$.

Доведення. Доведення теореми безпосередньо випливає з властивостей 3, 5, 7 визначників. ■

Наслідок. *Детермінант елементарної матриці відмінний від нуля.*

Лема 2.4. *Нехай A — деяка квадратна матриця, E — елементарна матриця того ж порядку. Тоді*

$$\det(EA) = \det E \cdot \det A.$$

Доведення. Матриця EA — це матриця, що одержана з матриці A з допомогою того елементарного перетворення, яке переводить одиничну матрицю I в E. Тоді, враховуючи властивості визначників:

1) якщо E утворена з одиничної перестановкою двох рядків, то $\det(EA) = -\det A = \det E \det A$;
2) якщо E утворена з одиничної множенням рядка матриці I на k, то $\det E = k$ і $\det(EA) = k \det A = \det E \det A$;
3) якщо E утворена з одиничної додаванням до деякого рядка іншого, помноженого на певне число, то $\det E = 1$ і $\det(EA) = \det E \det A$.

■

Теорема 2.13. *Квадратна матриця A оборотна тоді і тільки тоді, коли* $\det A \neq 0$.

Доведення. *Необхідність.* Нехай A оборотна матриця n-го порядку. Тоді за теоремами 2.6 та 2.7 її ранг дорівнює n, а зведеною східчастою формою матриці A є одинична матриця I. А це в свою чергу означає, що існують елементарні матриці $E_1, \dots E_m$ такі, що $E_m \dots E_1 A = I$, звідки, переходячи до детермінантів:

$$1 = \det I = \det(E_m \dots E_1 A) = \det E_m \dots \det E_1 \det A$$

і, отже, $\det A \neq 0$.

Достатність. Нехай маємо матрицю A, детермінант якої відмінний від нуля. Доведемо, що до неї існує обернена. Якщо A привести до зведеної східчастої форми, то можливі такі два випадки:

1. Зведена східчаста форма — це одинична матриця I. Тоді $\operatorname{rank} A = n$, а тому вона оборотна за теоремою 2.7.
2. Зведена східчаста форма B матриці A має нульові рядки. Тоді $\det B = 0$. Але має місце рівність $B = E_m \dots E_1 A$ для деяких елементарних матриць $E_1, \dots, E_m$. Тоді переходячи до визначників і використовуючи лему 2.4, знаходимо $\det E_m \dots \det E_1 \det A = 0$, звідки за наслідком з теореми 2.12 $\det A = 0$. Дістали суперечність, а тому цей випадок неможливий.

■

Означення 2.15. Матриця A, детермінант якої дорівнює нулю, називається *виродженою* (*особливою*), в іншому випадку — *невиродженою* (*неособливою*).

Отже, терміни «оборотна матриця» і «невироджена (неособлива) матриця» є синонімами.

2.7. Детермінанти і операції над матрицями.

Дослідимо питання про можливий зв'язок між детермінантами і основними матричними операціями, зокрема нашою метою буде встановлення формул для $\det(A^T)$, $\det(A + B)$, $\det(k \cdot A)$, $\det(AB)$, $\det(A^{-1})$ в термінах $\det A$ і $\det B$.

1. За властивістю 2 детермінантів має місце рівність для кожної квадратної матриці A:

$$\det(A^T) = \det A.$$

2. Інтуїтивно може скластись враження, що $\det(A+B) = \det A + \det B$. Підкреслимо, що взагалі кажучи це неправильно. Так, наприклад, якщо

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

то $\det A = \det B = 0$, але $\det(A+B) = \det I = 1$.

3. Має місце таке твердження.

Теорема 2.14. *Якщо A — матриця порядку n, $k \in \mathbb{R}$, то*

$$\det(kA) = k^n \det A.$$

Доведення. Твердження теореми випливає з властивості 5 визначників. ■

4. Наступне фундаментальне твердження теорії детермінантів вперше обґрунтував Коші[1].

Теорема 2.15. *Якщо A та B квадратні матриці однакових розмірів, то*

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

Доведення. Розглянемо окремо випадки, коли матриця A є оборотною, і коли необоротною.

а) Нехай A — оборотна. Тоді

$$\begin{aligned} E_m \dots E_1 \cdot A &= I; & &\Rightarrow \\ E_m \dots E_1 &= A^{-1}; & &\Rightarrow \\ (E_m \dots E_1)^{-1} &= A; & &\Rightarrow \\ A &= E_1^{-1} \dots E_m^{-1}; & &\Rightarrow \\ A &= E_1' \dots E_m', \end{aligned}$$

де $E_1', \dots, E_m'$ — елементарні, оскільки обернена матриця до елементарної, вочевидь, є елементарною.

Таким чином,

$$\det(AB) = \det(E_1' \dots E_m' B) = \det E_1' \cdot \ldots \cdot \det E_n' \cdot \det B =$$
$$= \det(E_1' \dots E_m') \cdot \det B = \det A \det B.$$

б) Нехай матриця A необоротна. Тоді AB також необоротна.

[1] Augustin Louis Cauchy (1789–1857) — видатний французький математик.

Справді, доведемо таке допоміжне твердження: якщо матриця AB оборотна, то матриці A та B також оборотні.

Якщо AB оборотна, то це означає, що існує така матриця C, що $(AB)C = C(AB) = I$. З асоціативності множення матриць знаходимо:

$$A(BC) = I, \text{а тому } A^{-1} = BC,$$
$$(CA)B = I, \text{а тому } B^{-1} = CA,$$

тобто матриці A та B оборотні.

Тоді маємо $\det A = 0$, $\det(AB) = 0$, а тому $\det(AB) = \det A \det B$.

■

5. Зв'язок між детермінантом матриці і оберненої до неї висвітлює таке твердження.

Теорема 2.16. *Якщо A — оборотна, то*

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

Доведення. Маємо: $A \cdot A^{-1} = I$. Тоді $\det(A \cdot A^{-1}) = \det I$, що рівносильно $\det A \cdot \det(A^{-1}) = 1$, звідки і дістаємо твердження теореми.

■

2.8. Правило Крамера для розв'язування систем лінійних рівнянь.

В цьому пункті ми розглянемо формули, що пов'язують детермінанти з розв'язками систем лінійних рівнянь.

Нехай маємо систему лінійних рівнянь $A\vec{x} = \vec{b}$, де $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{bmatrix}$. Через $A_i(\vec{b})$ позначатимемо матрицю, яка одержується з A заміною її i-ого стовпця на $\vec{b}$:

$$A_i(\vec{b}) = \begin{array}{c} \begin{bmatrix} \ \vec{a}_1 & \cdots & \vec{b} & \cdots & \vec{a}_n \ \end{bmatrix} \\ \uparrow \\ i\text{-й стовпець} \end{array}$$

Теорема 2.17 (Правило Крамера[1] розв'язування систем лінійних рівнянь)**.** *Якщо A — оборотна матриця n-го порядку, то система $A\vec{x} = \vec{b}$ має єдиний розв'язок*

$$x_i = \frac{\det A_i(\vec{b})}{\det A}, \quad i = 1, 2, \dots, n. \tag{2.6}$$

[1] Gabriel Cramer (1704–1752) — швейцарський математик.

Доведення. Маємо:

$$AI_i(\vec{x}) = A\begin{bmatrix}\vec{e}_1 & \dots & \vec{x} & \dots & \vec{e}_n\end{bmatrix} = \begin{bmatrix}A\vec{e}_1 & \dots & A\vec{x} & \dots & A\vec{e}_n\end{bmatrix} =$$
$$= \begin{bmatrix}\vec{a}_1 & \dots & \vec{b} & \dots & \vec{a}_n\end{bmatrix} = A_i(\vec{b}).$$

За теоремою 2.15:

$$\det A \cdot \det I_i(\vec{x}) = \det(AI_i(\vec{x})) = \det A_i(\vec{b}).$$

Разом з тим

$$\det I_i(\vec{x}) = \begin{vmatrix} 1 & 0 & \dots & x_1 & \dots & 0 & 0 \\ 0 & 1 & \dots & x_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & x_i & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & x_{n-1} & \dots & 1 & 0 \\ 0 & 0 & \dots & x_n & \dots & 0 & 1 \end{vmatrix} = x_i,$$

що слідує з розкладу цього визначника за i-им рядком. Таким чином, $\det A \cdot x_i = \det A_i(\vec{b})$, звідки діленням на $\det A$ ($\det A \neq 0$ оскільки A оборотна) завершується доведення теореми. ■

Приклад 2.10. Розв'язати систему:

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

Розв'язання. Обчислюємо визначник матриці системи:

$$\Delta = \begin{vmatrix} 3 & 2 \\ -5 & -1 \end{vmatrix} = -3 + 10 = 7.$$

$\Delta \neq 0$, тому система має єдиний розв'язок. Знаходимо далі

$$\Delta_1 = \begin{vmatrix} 2 & 2 \\ 1 & -1 \end{vmatrix} = -2 - 2 = -4,$$

$$\Delta_2 = \begin{vmatrix} 3 & 2 \\ -5 & 1 \end{vmatrix} = 3 - (-10) = 13.$$

Відповідь. $\left(-\frac{4}{7}, \frac{13}{7}\right)$. ■

Зауважимо, що розв'язувати системи лінійних рівнянь за допомогою формул (2.6) (вони називаються *формулами Крамера*) в загальному випадку неефективно. Ці формули зручні для розв'язування систем другого порядку, для систем вищих порядків зручніше використати метод Гауса. Незважаючи на це, формули Крамера мають велику теоретичну цінність.

2.9. Приєднана матриця. Формула для обчислення оберненої матриці.

Означення 2.16. Нехай $A_{n\times n} = [a_{ij}]$. Обчислимо алгебраїчні доповнення усіх елементів цієї матриці. Матриця

$$C^* = \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}$$

називається *приєднаною матрицею* .

Лема 2.5 (про фальшивий розклад визначника)**.** *Сума добутків елементів будь-якого рядка (стовпця) матриці на алгебраїчні доповнення іншого рядка (стовпця) матриці дорівнює нулю.*

Доведення. Нехай маємо матрицю $A_{n\times n} = [a_{ij}]$. Доведемо, що $a_{k1}A_{l1} + \ldots + a_{kn}A_{ln} = 0$ для довільних індексів $k \neq l$. Справді, розглянемо визначник матриці, яка одержується з матриці A заміною його l-го рядка k-им. Тоді з одного боку він дорівнює нулю, оскільки має два однакових рядки, а з іншого — розкладаючи його за l-им рядком знаходимо:

$$\Delta = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kn} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kn} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{k1}A_{l1} + \ldots + a_{kn}A_{ln}.$$

Отже, $a_{k1}A_{l1} + \ldots + a_{kn}A_{ln} = 0$, що і потрібно було довести. ■

Теорема 2.18. *Якщо A — оборотна, то*

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix} = \frac{1}{\det A}\cdot C^*.$$

Доведення. Безпосередньо обчислимо добуток $A\cdot\frac{1}{\det A}C^*$. Недіагональні елементи матриці $A\cdot C^*$ за лемою про фальшивий розклад визначника дорівнюють нулю, а діагональні — визначнику матриці, оскільки є сумами елементів певного рядка на алгебраїчні доповнення того ж рядка. Таким чином, $A\cdot\frac{1}{\det A}C^* = I$, а це і означає, що $A^{-1} = \frac{1}{\det A}\cdot C^*$. ∎

Отже, доведена теорема фактично стверджує, що обернена до матриці A є транспонованою матрицею алгебраїчних доповнень елементів A, які поділено на визначник матриці A.

Додаткові відомості і завдання для самостійної роботи

1. *Слідом матриці* n-го порядку $A = [a_{ij}]$ називається сума елементів, що розташовані на її головній діагоналі (позначається $\operatorname{tr} A$), тобто
$$\operatorname{tr} A = a_{11} + a_{22} + \ldots + a_{nn}.$$
Довести, що:
   1) $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B,\ A, B \in M_n(\mathbb{R})$;
   2) $\operatorname{tr}(\lambda A) = \lambda \operatorname{tr} A, k \in \mathbb{R}$;
   3) $\operatorname{tr}(AB) = \operatorname{tr}(BA),\ A, B \in M_n(\mathbb{R})$.
2. Квадратна матриця A називається *ідемпотентною*, якщо $A^2 = A$.
   1) Знайти усі ідемпотентні матриці другого порядку.
   2) Довести, що єдиною невиродженою ідемпотентною $n\times n$ матрицею є одинична матриця.
3. Квадратна матриця A називається *нільпотентною*, якщо $A^k = O$ для деякого $k > 1$. Чому може дорівнювати визначник нільпотентної матриці?
4. Нехай A та B — матриці другого порядку, сума елементів кожного рядка яких дорівнює 1. Довести, що тією ж властивістю володіє і матриця AB. Узагальніть цей результат на випадок матриць n-го порядку.
5. Довести, що якщо $n\times n$ матриці A та B мають ранг n, то матриця AB має також ранг n.
6. Довести, що $\operatorname{rank}(AB) \leqslant \operatorname{rank}(A)$ і $\operatorname{rank}(AB) \leqslant \operatorname{rank}(B)$.
7. Довести, що якщо матриця U оборотна, то $\operatorname{rank}(UA) = \operatorname{rank}(A)$. (*Підказка.* $A = U^{-1}(UA)$).
8. Довести, що якщо матриця U оборотна, то $\operatorname{rank}(AU) = \operatorname{rank}(A)$.

9. Довести, що для $m \times n$ матриць A, B виконується нерівність

$$\operatorname{rank}(A+B) \leqslant \operatorname{rank}(A) + \operatorname{rank}(B).$$

10. Нехай A — квадратна матриця, яка має такий блочний вигляд:

$$A = \left[\begin{array}{c|c} P & Q \\ \hline O & S \end{array}\right],$$

де P та S — квадратні матриці, O — нульова. Довести, що $\det A = \det P \cdot \det S$. (*Підказка.* Застосуйте метод математичної індукції по кількості рядків в P.)

11. Довести таку формулу (*визначник Вандермонда* [1]):

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{i<j} (x_j - x_i).$$

12. Довести, що для будь-яких попарно різних чисел $a_1, \dots, a_{n+1}$ і довільних чисел $b_1, \dots, b_{n+1}$ існує єдиний многочлен $f(t)$ степеня не вище n такий, що

$$f(a_1) = b_1, \dots, f(a_{n+1}) = b_{n+1}.$$

Підказка. Система рівнянь для обчислення коефіцієнтів має основну матрицю з визначником Вандермонда.

[1] Alexandre-Theophile Vandermonde (1735—1796) — французький музикант і математик.

Розділ 3

Векторні простори

После того как мы усвоим несколько простых положений и выведем из них какое-либо иное, полезно обозреть их путем последовательного и непрерывного движения мысли, обдумать их взаимоотношения и отчетливо представить одновременно наибольшее их количество; благодаря этому наше знание сделается более достоверным и наш ум приобретет больший кругозор.

Р. Декарт

В розділах 1 і 2 ми бачили, що алгебри n-вимірних векторів і алгебри матриць мали багато спільного (порівняйте властивості операцій над векторами на сторінці 23 і результат теореми 2.1). В цьому розділі спільні властивості операцій над матрицями та векторами кладуться в основу узагальненого поняття «вектора» та множини таких «векторів» — лінійного (векторного) простору.

§ 1. Поняття векторного простору та його підпростору.

1.1. Означення векторного (лінійного) простору. Основні приклади векторних просторів.

Означення 3.1. Нехай V — це множина, на якій задана внутрішня бінарна алгебраїчна операція *додавання* і зовнішня бінарна алгебраїчна операція *множення* елементів множини V на елементи поля P. Якщо $\vec{a}, \vec{b} \in V$, то їх сума позначається через $\vec{a} + \vec{b}$, а добуток скаляра $\alpha \in P$ на $\vec{a}$ позначається через $\alpha \cdot \vec{a}$ або $\alpha\vec{a}$. Якщо наступні умови виконуються для всіх елементів $\vec{a}$, $\vec{b}$, $\vec{c}$ множини V і довільних елементів $\alpha_1, \alpha_2 \in P$, то множина V називається *векторним (лінійним) простором над полем* P, а її елементи називаються *векторами*:

1) $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$;
2) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$;
3) існує $\vec{0} \in V$ (він називається *нуль-вектором*), що $\vec{a} + \vec{0} = \vec{a}$;
4) для кожного $\vec{a} \in V$ існує такий елемент $(-\vec{a}) \in V$, що $\vec{a} + (-\vec{a}) = \vec{0}$;
5) $\alpha_1(\alpha_2\vec{a}) = (\alpha_1\alpha_2)\vec{a}$;
6) $\alpha_1(\vec{a} + \vec{b}) = \alpha_1\vec{a} + \alpha_1\vec{b}$;
7) $(\alpha_1 + \alpha_2)\vec{a} = \alpha_1\vec{a} + \alpha_2\vec{a}$;
8) $1\vec{a} = \vec{a}$.

Умови 1–8 з означення 3.1 називаються *аксіомами векторного простору*.

Зауваження. Найчастіше нами розглядатимуться векторні простори, задані над полем $\mathbb{R}$ дійсних чисел. Такі векторні простори називають ще *дійсними векторними просторами*. Іноді розглядаються векторні простори, задані над полем C комплексних чисел та полем $\mathbb{Z}_p$. Якщо окремо не зауважено, то вживаючи «векторний простір» ми матимемо на увазі «дійсний векторний простір».

Розглянемо деякі приклади векторних просторів.

Приклад 3.1. Для $n \geqslant 1$ n-вимірний арифметичний векторний простір $\mathbb{R}^n$ є дійсним векторним простором.

Приклад 3.2. Множина матриць розміру 2×3 з дійсними елементами разом з операціями додавання матриць і множення їх на дійсні числа є дійсним векторним простором (аксіоми векторного простору виконуються згідно з теоремою 2.1).

Взагалі, множина матриць фіксованої розмірності $m \times n$ з операціями додавання матриць і множення на скаляр є векторним простором. Цей векторний простір позначатимемо M_{mn}.

Приклад 3.3. Нехай $\mathcal{P}_2$ — множина многочленів від однієї змінної з дійсними коефіцієнтами, степінь яких не перевищує 2. Означимо в природний спосіб на цій множині операції додавання на множення на число, а саме: якщо

$$P(x) = a_0 + a_1x + a_2x^2, \quad Q(x) = b_0 + b_1x + b_2x^2,$$

то

$$P(x)+Q(x)=(a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2,$$
$$\alpha P(x)=\alpha a_0+(\alpha a_1)x+(\alpha a_2)x^2.$$

Очевидно, що ця множина є замкненою відносно введених операцій і усі аксіоми виконуються, в чому можна переконатись безпосередньо. Отже, $\mathcal{P}_2$ — векторний простір.

В загальному випадку $\mathcal{P}_n$ — множина многочленів від однієї змінної, степінь яких не перевищує n, є векторним простором.

Приклад 3.4. Нехай $\mathcal{F}_{[a,b]}$ — множина функцій дійсної змінної визначених на відрізку $[a,b]$. На цій множині введемо операції додавання функцій і множення функції на число, а саме: якщо $f(x)$ та $g(x)$ — дві функції, то їх сумою називається функція $(f+g)(x)=f(x)+g(x)$, а добутком функції на число α — $(\alpha f)(x)=\alpha\cdot f(x)$. Легко переконатись в тому, що множина $\mathcal{F}_{[a,b]}$ замкнена відносно введених операцій і умови 1–8 означення векторного простору виконуються. Отже $\mathcal{F}_{[a,b]}$ — векторний простір.

Приклад 3.5. Нехай $\mathcal{C}_{[a,b]}$ — множина неперервних на відрізку $[a,b]$ функцій. З курсу математичного аналізу відомо, що сума двох неперервних на $[a,b]$ функцій є неперервною функцією на $[a,b]$, і добуток неперервної на відрізку функції на число також є неперервною на цьому ж відрізку. При цьому виконуються аксіоми 1–8 векторного простору. Отже, $\mathcal{C}_{[a,b]}$ — векторний простір.

Приклад 3.6. Множина $\mathbb{Z}$ цілих чисел не є дійсним векторним простором. Для доведення цього слід вказати принаймні одну умову означення 3.1, яка не виконується. Справді, множина $\mathbb{Z}$ не є замкненою відносно операції множення на дійсне число, оскільки результатом такого множення може виявитись не ціле число.

Теорема 3.1 (Найпростіші властивості векторних просторів)**.** *Нехай V — векторний простір, $\vec{a}\in V$, $\alpha\in P$. Тоді:*
1) $0\vec{a}=\vec{0}$;
2) $\alpha\vec{0}=\vec{0}$;
3) $(-\alpha)\vec{a}=-\alpha\vec{a}$.

Доведення. 1) Проведемо ланцюжок рівносильних перетворень (число над знаком дорівнює позначає аксіому векторного простору, що використовується):

$$\vec{a} \overset{8}{=} (1+0)\vec{a} \overset{7}{=} 1\vec{a} + 0\vec{a} \overset{8}{=} \vec{a} + 0\vec{a}.$$

Додамо до обох частин вектор $-\vec{a}$. Тоді, використовуючи аксіоми 4 та 2, дістанемо:

$$\vec{0} = (0\vec{a} + \vec{a}) + (-\vec{a}) \overset{1}{=} 0\vec{a} + (\vec{a} + (-\vec{a})) \overset{4}{=} 0\vec{a} + \vec{0} \overset{3}{=} 0\vec{a}.$$

2) Зафіксуємо вектор $\vec{a}$. Тоді маємо:

$$\alpha\vec{a} \overset{3}{=} \alpha(\vec{a} + \vec{0}) \overset{6}{=} \alpha\vec{a} + \alpha\vec{0}.$$

Додамо до обох частин вектор $-\alpha\vec{a}$. Тоді:

$$\begin{aligned} \vec{0} &= (\alpha\vec{a} + \alpha\vec{0}) + (-\vec{\alpha}a) = \\ &= (\alpha\vec{0} + \alpha\vec{a}) + (-\vec{\alpha}a) = \\ &= \alpha\vec{0} + (\alpha\vec{a} + (-\vec{\alpha}a)) = \alpha\vec{0} + \vec{0} = \alpha\vec{0}. \end{aligned}$$

3) Проведемо ланцюжок рівносильних перетворень:

$$\vec{0} = (\alpha + (-\alpha))\vec{a} = \alpha\vec{a} + (-\alpha)\vec{a}.$$

Додамо до лівої та правої частини останньої рівності вектор $(-\alpha\vec{a})$:

$$\begin{aligned} -\alpha\vec{a} &= \alpha\vec{a} + (-\alpha)\vec{a} - \alpha\vec{a} = \\ &= (-\alpha)\vec{a} + (\alpha\vec{a} - \alpha\vec{a}) = (-\alpha)\vec{a} + \vec{0} = (-\alpha)\vec{a}, \end{aligned}$$

що і потрібно було довести. ■

Під *різницею* векторів $\vec{a}$ та $\vec{b}$ розумітимемо суму $\vec{a} + (-\vec{b})$ і позначатимемо $\vec{a} - \vec{b}$.

1.2. Поняття підпростору векторного простору.

Означення 3.2. Нехай V — векторний простір, заданий над полем P, U — підмножина множини V. Якщо множина U сама є векторним простором над полем P з тими ж самими операціями додавання і множення на скаляр, то U називається *підпростором* простору V.

Теорема 3.2 (Критерій підпростору)**.** *Для того, щоб непорожня підмножина U простору V заданого над полем P була його підпростором, необхідно і достатньо, щоб вона була замкнена відносно операції додавання векторів і відносно операції множення векторів на елементи поля P.*

Доведення. *Необхідність* очевидна.

Достатність. Нехай підмножина U векторного простору V замкнена відносно операцій додавання і множення на скаляр. Доведемо, що вона є підпростором цього простору. Для того, щоб це довести, слід показати, що сама множина U є простором, заданим над полем P, тобто довести, що усі 8 аксіом векторного простору мають місце.

Аксіоми 1, 2, 5–8 виконуються для всіх елементів простору V, а значить вони виконуються і для всіх елементів множини U.

Доведемо виконання третьої та четвертої аксіом векторного простору. Оскільки множина U непорожня, то вона містить принаймні один вектор $\vec{a}$. Але тоді $(-1)\vec{a} \in U$ також (оскільки U замкнена відносно операції множення на скаляр), або за властивістю 3 теореми 3.1 $-\vec{a} \in U$. Таким чином, до кожного елемента існує протилежний і четверта аксіома векторного простору виконується. Але тоді в силу замкненості множини U відносно операції додавання і $\vec{a} + (-\vec{a}) = \vec{0} \in U$, тобто третя аксіома також має місце.

Таким чином, U — векторний простір над полем P, а тому підпростір простору V. ■

Доведена теорема дає фактично спосіб перевірки того, чи є задана множина U векторним простором. Якщо вдається знайти векторний простір з тими ж самими операціями, для якого U є його підмножиною, то для перевірки того, що U є векторним простором достатньо перевірити лише її замкненість відносно операцій додавання та множення на скаляр.

Розглянемо це на прикладах.

Приклад 3.7. Нехай $\mathcal{P}$ — множина всіх многочленів від однієї змінної з дійсними коефіцієнтами. Легко переконатись в тому, що $\mathcal{P}$ — векторний простір. Тоді $\mathcal{P}_n$ — простір многочленів степінь, яких не перевищує n, є підпростором простору P.

Приклад 3.8. Нехай U — множина симетричних матриць n-го порядку (квадратна матриця $A = [a_{ij}]$ називається *симетричною*, якщо $a_{ij} = a_{ji}$ для всіх $i \neq j$, тобто якщо $A^T = A$). Легко переконатись в тому, що сума двох симетричних матриць є знову симетричною матрицею; добуток симетричної матриці на число є симетричною матрицею. Отже, U — підпростір векторного простору M_n квадратних матриць.

Приклад 3.9. Нехай $\mathcal{F}_{[a,b]}$ — множина усіх функцій $f\colon [a,b] \to \mathbb{R}$; $\mathcal{C}_{[a,b]}$ — множина усіх неперервних функцій $f\colon [a,b] \to \mathbb{R}$; $\mathcal{D}_{[a,b]}$ — множина усіх диференційовних функцій $f\colon [a,b] \to \mathbb{R}$. Має місце таке включення: $\mathcal{D}_{[a,b]} \subset \mathcal{C}_{[a,b]} \subset \mathcal{F}_{[a,b]}$. $\mathcal{F}_{[a,b]}$, як ми показали раніше, є векторним простором, а усі множини функцій $\mathcal{C}_{[a,b]}$ та $\mathcal{D}_{[a,b]}$ замкнені відносно операцій додавання і множення на скаляр. Таким чином, вони також є векторними просторами. Див. рис. 3.1.

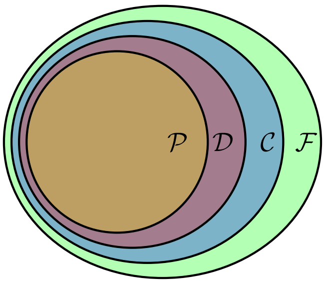


Рис. 3.1. Векторні простори $\mathcal{P}$, $\mathcal{D}$, $\mathcal{C}$, $\mathcal{F}$.

Приклад 3.10. Нехай S — множина функцій, що задовольняють рівняння $f'' + f = 0$. Зрозуміло, що $S \subset \mathcal{D}$ і S непорожня множина $(f(x) \equiv 0 \in S)$. Нехай f та g належать S і $\alpha \in \mathbb{R}$. Тоді

$$(f+g)'' + (f+g) = f'' + g'' + f + g = (f''+f) + (g''+g) = 0 + 0 = 0,$$
$$(\alpha f)'' + (\alpha f) = \alpha(f''+f) = \alpha \cdot 0 = 0,$$

тобто S замкнена відносно операцій додавання та множення на число, а тому є векторним простором.

Приклад 3.11. Якщо V векторний простір, то він може вважатись підпростором самого себе. $\{\vec{0}\}$ також є підпростором простору V — він називається *нуль-простором*. $\{\vec{0}\}$ та V називають *тривіальними* підпросторами простору V.

Часто для доведення того, що множина не є підпростором векторного простору використовується таке твердження, яке безпосередньо випливає з теореми 3.2.

Наслідок. *Якщо U — підпростір векторного простору V, то U містить $\vec{0}$ простору V.*

Приклад 3.12. Чи є множина W матриць виду

$$\begin{bmatrix} a & b \\ b+1 & a \end{bmatrix}, \quad a, b \in \mathbb{R}$$

підпростором векторного простору M_{22}?

Розв'язання. Оскільки нульова матриця (нульовий елемент простору M_{22}) вказаній множині не належить, то W не є підпростором простору M_{22}. ■

1.3. Лінійна оболонка системи векторів.

Подібно до того, як було введено поняття лінійної оболонки системи векторів простору $\mathbb{R}^n$, введемо поняття лінійної оболонки системи векторів довільного векторного простору V.

Означення 3.3. Нехай $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right)$ — система векторів простору V заданого над полем P. Вектор $\vec{b}$ простору V називається *лінійною комбінацією* векторів цієї системи, якщо існують такі $\alpha_1, \alpha_2, \ldots, \alpha_k \in P$, що

$$\vec{b} = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_k\vec{a}_k.$$

Те, що вектор $\vec{b}$ є лінійною комбінацією векторів $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ позначатимемо так: $\vec{b} \xrightarrow{\text{л.в.}} \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ (позначення «л.в.» є скороченням від «лінійно виражається»).

Означення 3.4. Нехай $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right)$ — система векторів простору V заданого над полем P. Сукупність усіх лінійних комбінацій векторів даної системи називається її *лінійною оболонкою* і позначається $L\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right)$ або $\mathrm{Span}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k)$.

Приклад 3.13. Векторний простір $\mathcal{P}_2$ є лінійною оболонкою векторів 1, x, x^2.

Приклад 3.14. Векторний простір M_{22} є лінійною оболонкою системи

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

оскільки довільна матриця з M_{22} подається у вигляді:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = aE_1 + bE_2 + cE_3 + dE_4.$$

Приклад 3.15. Чи належить функція $y = \sin 2x$ лінійній оболонці $L\left(\sin x, \cos x\right)$ векторного простору $\mathcal{F}$?

Розв'язання. Припустимо, що належить. Це означає, що існують такі a, b, що рівність $a \sin x + b \cos x = \sin 2x$ виконується для кожного x. Підставимо послідовно $x = 0$ та $x = \frac{\pi}{2}$. Дістанемо:

$$\begin{cases} a \cdot 0 + b \cdot 1 & = 0, \\ a \cdot 1 + b \cdot 0 & = 0, \end{cases}$$

звідки $a = b = 0$. Тоді $\sin 2x \equiv 0$, що, очевидно, є неправдою. Таким чином, наше припущення невірне і $\sin 2x$ не належить $L\left(\sin x, \cos x\right)$. ■

Теорема 3.3. *Лінійна оболонка системи векторів* $\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ *простору* V *є підпростором цього простору.*

Доведення. Згідно доведеного критерію підпростору потрібно показати, що множина $L\left(\vec{a}_1, \ldots \vec{a}_k\right)$ є замкненою відносно операцій додавання та множення на скаляр, тобто, що сумою двох лінійних комбінацій векторів системи $\vec{a}_1, \ldots \vec{a}_k$ є знову лінійна комбінація векторів цієї системи і добуток лінійної комбінації векторів даної системи на елемент $\alpha \in P$ є лінійною комбінацією векторів цієї системи.

Нехай

$$\vec{b}_1 = \alpha_{11}\vec{a}_1 + \ldots + \alpha_{1k}\vec{a}_k,$$
$$\vec{b}_2 = \alpha_{21}\vec{a}_1 + \ldots + \alpha_{2k}\vec{a}_k.$$

Тоді

$$\vec{b}_1 + \vec{b}_2 = \underbrace{\left(\alpha_{11} + \alpha_{21}\right)}_{\beta_1}\vec{a}_1 + \ldots + \underbrace{\left(\alpha_{1k} + \alpha_{2k}\right)}_{\beta_k}\vec{a}_k,$$
$$\alpha\vec{b}_1 = \underbrace{\left(\alpha\alpha_{11}\right)}_{\gamma_1}\vec{a}_1 + \ldots + \underbrace{\left(\alpha\alpha_{1k}\right)}_{\gamma_k}\vec{a}_k.$$

тобто одержали лінійні комбінації системи векторів $\vec{a}_1, \ldots \vec{a}_k$. Отже, $\vec{b}_1 + \vec{b}_2, \alpha\vec{b}_1 \in L\left(\vec{a}_1, \ldots \vec{a}_k\right)$. ■

Зауваження. $L\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ — найменший підпростір, що містить вектори $\vec{a}_1, \ldots \vec{a}_k$, тобто $L\left(\vec{a}_1, \ldots, \vec{a}_k\right)$ є підпростором кожного іншого простору, що містить ці вектори.

§ 2. Лінійна незалежність, базис і розмірність векторного простору

Введемо означення лінійно залежної та лінійно незалежної системи векторів довільного векторного простору подібно до того, як нами це було зроблено в просторі $\mathbb{R}^n$.

2.1. Лінійно залежні та лінійно незалежні системи векторів.

Означення 3.5. Система векторів $(\vec{a}_1, \ldots, \vec{a}_k)$ називається *лінійно залежною*, якщо існують такі числа $\alpha_1, \ldots, \alpha_k$, серед яких хоча б одне відмінне від нуля, що виконується рівність

$$\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k = \vec{0}. \tag{$\star$}$$

Якщо ж рівність $(\star)$ виконується лише при $\alpha_1 = \ldots = \alpha_k = 0$, то система векторів називається *лінійно незалежною*.

Теорема 3.4. *Для того, щоб система векторів $(\vec{a}_1, \ldots, \vec{a}_k)$ була лінійно залежною, необхідно і достатньо, щоб хоча б один вектор цієї системи лінійно виражався через інші.*

Доведення. Доведення проводиться цілком аналогічно як і для простору $\mathbb{R}^n$ (теорема 1.2). ■

Розглянемо деякі приклади.

Приклад 3.16. В просторі M_{22} задано матриці

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 3 \\ 0 & 1 \end{bmatrix}.$$

Система векторів цього простору $\{A, B, C\}$ є лінійно залежною, оскільки $C = A + B$.

Приклад 3.17. В просторі $\mathcal{F}$ система функцій $(\sin^2 x, \cos^2 x, \cos 2x)$ є лінійно залежною, оскільки $\cos 2x = \cos^2 x - \sin^2 x$.

Приклад 3.18. Довести, що в просторі $\mathcal{P}_n$ система многочленів $(1, x, x^2, \ldots, x^n)$ є лінійно незалежною.

Розв'язання. Складемо рівність $(\star)$ для вказаної системи векторів:

$$\alpha_0 \cdot 1 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_n x^n = 0.$$

Якщо покласти в цю рівність $x = 0$, то дістанемо $\alpha_0 = 0$. Продиференціювавши цю рівність і підставивши $x = 0$, знайдемо, що і $\alpha_1 = 0$. Диференціюючи рівність далі і підставляючи $x = 0$ послідовно дістанемо, що $\alpha_0 = \alpha_1 = \ldots = \alpha_n = 0$, тобто система многочленів є лінійно незалежною. ■

2.2. Базис векторного простору. Координати вектора в базисі.

Означення 3.6. Підмножина $\mathcal{B}$ векторного простору V називається *базисом* цього простору якщо виконуються умови:

1) $L(\mathcal{B}) = V$;
2) $\mathcal{B}$ — лінійно незалежна система.

Іншими словами, система векторів простору називається його базисом, якщо вона лінійно незалежна і через вектори цієї системи лінійно виражається кожний інший вектор цього простору.

Приклад 3.19. Вектори

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad , \ldots, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

утворюють базис простору $\mathbb{R}^n$. Його називають *стандартним базисом простору* $\mathbb{R}^n$.

Приклад 3.20. Система многочленів $\left(1, x, x^2, \ldots, x^n\right)$ утворює *стандартний базис простору* $\mathcal{P}_n$.

Приклад 3.21. Множина матриць

$$E_{11} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}, \ldots, E_{mn} = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

утворює так званий *стандартний базис простору* M_{mn}.

Приклад 3.22. Довести, що $\mathcal{B} = \left(1 + x, 1 + x^2, x + x^2\right)$ — базис простору $\mathcal{P}_2$.

Розв'язання. Покажемо, що кожний многочлен з простору $\mathcal{P}_2$ може бути поданий у вигляді лінійної комбінації многочленів з $\mathcal{B}$. Для цього доведемо, що існують такі коефіцієнти $\alpha_1, \alpha_2, \alpha_3$, що рівність

$$\alpha_1(1+x) + \alpha_2(1+x^2) + \alpha_3(x+x^2) = a_0 + a_1x + a_2x^2 \tag{3.1}$$

виконується для кожного x при фіксованих довільних a_0, a_1, a_2.

Розкривши дужки в рівності (3.1) і прирівнявши відповідні коефіцієнти одержаного многочлена з коефіцієнтами многочлена з правої частини, дістанемо систему:

$$\begin{cases} \alpha_1 + \alpha_2 & = a_0, \\ \alpha_1 + \alpha_3 & = a_1, \\ \alpha_2 + \alpha_3 & = a_2. \end{cases}$$

Визначник матриці цієї системи відмінний від нуля, а тому, по-перше, система має єдиний розв'язок при всіх значеннях a_0, a_1, a_2, і, по-друге, система многочленів $\mathcal{B}$ є лінійно незалежною. Таким чином, $\mathcal{B}$ є базисом простору $\mathcal{P}_2$. ■

Теорема 3.5. *Нехай V — векторний простір, $\mathcal{B} = \big(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\big)$ — деякий базис V. Кожний вектор $\vec{a} \in V$ однозначно виражається через вектори базису.*

Доведення. Те, що принаймні одне представлення вектора $\vec{a}$ у вигляді лінійної комбінації векторів базису існує, випливає з означення базису. Доведемо, що таке подання єдине.

Припустимо, що існує такий вектор $\vec{a}$, який має два різних подання через базисні вектори:

$$\begin{aligned} \vec{a} =& \alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n, \\ \vec{a} =& \beta_1\vec{e}_1 + \ldots + \beta_n\vec{e}_n. \end{aligned}$$

Віднявши ці дві рівності, дістанемо

$$\vec{0} = (\alpha_1 - \beta_1)\vec{e}_1 + \ldots + (\alpha_n - \beta_n)\vec{e}_n.$$

Одержали рівність ($\star$) для системи векторів базису. Вона за означенням виконується лише при $\alpha_1 - \beta_1 = \ldots = \alpha_n - \beta_n = 0$, тобто $\alpha_i = \beta_i$ для кожного i, що суперечить припущенню того, що представлення різні. ■

Означення 3.7. Нехай $\mathcal{B} = \big(\vec{e}_1, \ldots, \vec{e}_n\big)$ — базис векторного простору V і нехай $\vec{a} = \alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n$. Тоді $\alpha_1, \ldots, \alpha_n$ називаються *координатами вектора $\vec{a}$ в базисі $\mathcal{B}$*, а вектор-стовпець

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \equiv [\vec{a}]_{\mathcal{B}} -$$

вектором координат (координатним стовпцем) $\vec{a}$ в базисі $\mathcal{B}$.

Зауваження. Координатний стовпець — це вектор з простору $\mathbb{R}^n$. Тоді між впорядкованими n-ками дійсних чисел (елементами простору $\mathbb{R}^n$) і векторами простору V при вибраному базисі $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ існує взаємнооднозначна відповідність, а саме: кожному вектору $\vec{a} \in V$ ставиться у відповідність вектор його координат в базисі $\mathcal{B}$, а кожному вектору $\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \mathbb{R}^n$ — вектор $\alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n \in V$.

Приклад 3.23. Вектором координат многочлена $P(x) = 1 - 3x + 4x^2$ в базисі $\mathcal{B} = \left(1, x, x^2\right)$ векторного простору $\mathcal{P}_2$ є $\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$, а в базисі $\mathcal{B} = \left(x^2, x, 1\right)$ — $\begin{bmatrix} 4 \\ -3 \\ 1 \end{bmatrix}$.

Приклад 3.24. Координатами матриці $A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$ в стандартному базисі простору M_{22} є

$$\begin{bmatrix} 1 \\ 3 \\ 2 \\ -1 \end{bmatrix}.$$

Теорема 3.6. *Нехай $\mathcal{B}$ — базис векторного простору V заданого над деяким полем P, $\vec{a}, \vec{b} \in V$, $\lambda \in P$. Мають місце рівності:*

$$\left[\vec{a} + \vec{b}\right]_{\mathcal{B}} = [\vec{a}]_{\mathcal{B}} + \left[\vec{b}\right]_{\mathcal{B}};$$

$$[\lambda\vec{a}]_{\mathcal{B}} = \lambda \cdot [\vec{a}]_{\mathcal{B}}.$$

Доведення. Нехай вектор $\vec{a}$ в базисі $\mathcal{B} = \big(\vec{e}_1, \ldots, \vec{e}_n\big)$ задано координатним стовпцем $\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}$, а вектор $\vec{b}$ в тому ж базисі задано координатним стовпцем $\begin{bmatrix}\beta_1\\ \vdots\\ \beta_n\end{bmatrix}$. Це означає, що

$$\vec{a} = \alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n;$$
$$\vec{b} = \beta_1\vec{e}_1 + \ldots + \beta_n\vec{e}_n.$$

Тоді

$$\vec{a} + \vec{b} = (\alpha_1 + \beta_1)\vec{e}_1 + \ldots + (\alpha_n + \beta_n)\vec{e}_n,$$
$$\lambda\vec{a} = (\lambda\alpha_1)\vec{e}_1 + \ldots + (\lambda\alpha_n)\vec{e}_n.$$

Це означає, що векторам $\vec{a} + \vec{b}$, $\lambda\vec{a}$ відповідають координатні стовпці

$$\begin{bmatrix}\alpha_1 + \beta_1\\ \vdots\\ \alpha_n + \beta_n\end{bmatrix} \quad \text{та} \quad \begin{bmatrix}\lambda\alpha_1\\ \vdots\\ \lambda\alpha_n\end{bmatrix}$$

відповідно. Отже,

$$\left[\vec{a} + \vec{b}\right]_{\mathcal{B}} = \begin{bmatrix}\alpha_1 + \beta_1\\ \vdots\\ \alpha_n + \beta_n\end{bmatrix} = \begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix} + \begin{bmatrix}\beta_1\\ \vdots\\ \beta_n\end{bmatrix} = [\vec{a}]_{\mathcal{B}} + \left[\vec{b}\right]_{\mathcal{B}};$$

$$[\lambda\vec{a}]_{\mathcal{B}} = \begin{bmatrix}\lambda\alpha_1\\ \vdots\\ \lambda\alpha_n\end{bmatrix} = \lambda\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix} = \lambda \cdot [\vec{a}]_{\mathcal{B}}.$$

■

Наслідок. *Для довільних векторів $\vec{a}_1, \ldots, \vec{a}_k$ простору V і скалярів $\alpha_1, \ldots, \alpha_k$ при вибраному базисі $\mathcal{B}$ виконується рівність*

$$[\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k]_{\mathcal{B}} = \alpha_1\,[\vec{a}_1]_{\mathcal{B}} + \ldots + \alpha_k\,[\vec{a}_k]_{\mathcal{B}};$$

Теорема 3.7. *Нехай* $\mathcal{B} = \left(\vec{b}_1, \ldots, \vec{b}_n\right)$ *— базис векторного простору* V, $\vec{a}_1, \ldots, \vec{a}_k$ *— деяка система векторів цього простору. Ця система є лінійно незалежною тоді і тільки тоді, коли лінійно незалежною є система векторів* $[\vec{a}_1]_{\mathcal{B}}, \ldots, [\vec{a}_k]_{\mathcal{B}}$ *простору* $\mathbb{R}^n$.

Доведення. *Необхідність.* Нехай система векторів $\vec{a}_1, \ldots, \vec{a}_k$ є лінійно незалежною. Запишемо рівність ($\star$) для векторів простору $\mathbb{R}^n$

$$\alpha_1 [\vec{a}_1]_{\mathcal{B}} + \ldots + \alpha_k [\vec{a}_k]_{\mathcal{B}} = \vec{0}.$$

Враховуючи попередню теорему, цю рівність можна переписати у вигляді

$$[\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k]_{\mathcal{B}} = \vec{0},$$

тобто усі координати вектора $\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k$ в базисі $\mathcal{B}$ дорівнюють нулю, а тому

$$\alpha_1\vec{a}_1 + \ldots + \alpha_n\vec{a}_n = \vec{0}.$$

Але з лінійної незалежності системи векторів $\vec{a}_1, \ldots, \vec{a}_k$ випливає, що остання рівність виконується лише при $\alpha_1 = \ldots = \alpha_k = 0$. Отже, система векторів $[\vec{a}_1]_{\mathcal{B}}, \ldots, [\vec{a}_k]_{\mathcal{B}}$ також лінійно незалежна.

Достатність доводиться подібними міркуваннями. Залишаємо її читачам як вправу. ■

2.3. Розмірність векторного простору.

Теорема 3.8 (Основна теорема про лінійну залежність)**.** *Якщо кожний вектор системи з більшою кількістю векторів лінійно виражається через вектори системи з меншою кількістю векторів, то перша система лінійна залежна.*

Доведення. Доведення проводиться аналогічно як і для випадку простору $\mathbb{R}^n$ (теорема 1.4). ■

Теорема 3.9. *Нехай* $\mathcal{B} = \left(\vec{e}_1, \ldots, \vec{e}_n\right)$ *— базис векторного простору* V. *Тоді:*

1. *Кожна система, що містить більше* n *векторів, є лінійно залежною.*
2. *Лінійна оболонка кожної системи, що містить менше, ніж* n *векторів не співпадає з* V.

Доведення. Перша частина теореми слідує безпосередньо з основної теореми про лінійну залежність та означення базису.

Доведемо другу частину теореми. Нехай існує така система векторів $\vec{a}_1, \ldots \vec{a}_m$, $m < n$, що $L(\vec{a}_1, \ldots \vec{a}_m) = V$. Це означає, що кожний вектор простору V, а отже і $\vec{e}_1, \ldots, \vec{e}_n$, лінійно виражаються через ці вектори. Ми одержали, що кожний вектор системи з більшою кількість векторів $(\vec{e}_1, \ldots, \vec{e}_n)$ лінійно виражається через вектори системи з меншою кількістю векторів $(\vec{a}_1, \ldots, \vec{a}_m)$. За основною теоремою про лінійну залежність стверджуємо, що система $\vec{e}_1, \ldots, \vec{e}_n$ є лінійно залежною. Це суперечить тому, що $\vec{e}_1, \ldots, \vec{e}_n$ формують базис простору V. Таким чином, наше припущення невірне і $L(\vec{a}_1, \ldots \vec{a}_m) \neq V$. ■

Наслідок. *Кожний базис векторного простору містить однакову кількість векторів.*

Це обґрунтовує можливість такого означення.

Означення 3.8. Векторний простір V називається *скінченновимірним*, якщо він має базис, що складається зі скінченної кількості векторів. *Розмірністю* векторного простору називається кількість векторів його базису (позначається $\dim V$). Розмірність нульового векторного простору $\{\vec{0}\}$ вважають рівною нулю. Векторний простір, що не має скінченного базису називається *нескінченновимірним*.

Приклад 3.25. Розмірність векторного простору $\mathbb{R}^n$ дорівнює n, оскільки стандартний базис цього простору складається з n векторів (див. приклад 3.19).

Приклад 3.26. Розмірність векторного простору $\mathcal{P}_n$ дорівнює $n+1$, оскільки стандартний базис цього простору складається з $n+1$-го многочлена (див. приклад 3.20).

Приклад 3.27. Розмірність векторного простору M_{mn} дорівнює mn, оскільки стандартний базис цього простору складається з mn матриць (див. приклад 3.21).

Приклад 3.28. Векторні простори $\mathcal{P}$, $\mathcal{F}_{[a,b]}$, $\mathcal{C}_{[a,b]}$ є нескінченновимірними, оскільки усі вони містять нескінченну лінійно незалежну систему функцій $\big(1, x, x^2, x^3, \ldots\big)$.

Теорема 3.10. *Нехай V — n-вимірний векторний простір, $\vec{a}_1, \dots \vec{a}_k$, $k < n$ — лінійно незалежна система векторів цього простору. Тоді її можна доповнити до базису простору V.*

Доведення. Нехай $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ — базис простору V. Виберемо з цього базису такий вектор, який лінійно не виражається через вектори системи $\vec{a}_1, \dots \vec{a}_k$. Такий вектор обов'язково існує, оскільки в протилежному випадку кожний вектор базису $\vec{e}_1, \dots, \vec{e}_n$ лінійно виражався б через вектори системи $\vec{a}_1, \dots \vec{a}_k$, що за основною теоремою 3.8 про лінійну залежність свідчило б про лінійну залежність системи $\vec{e}_1, \dots, \vec{e}_n$, що неправильно. Нехай таким вектором є $\vec{e}_{i_1}$. Розглянемо систему $\vec{a}_1, \dots \vec{a}_k, \vec{e}_{i_1}$. Якщо $k + 1 = n$, то ця система буде базисом простору V. В іншому випадку, застосовуючи аналогічні міркування до нової системи векторів, знайдемо вектор $\vec{e}_{i_2}$, який лінійно не виражається через вектори цієї системи. Цей процес скінченний і за $n - k$ кроків ми побудуємо лінійно незалежну систему $\vec{a}_1, \dots \vec{a}_k, \vec{e}_{i_1}, \dots, \vec{e}_{n-k}$, яка буде базисом простору V. ■

§ 3. Перехід від одного базиса до іншого

В багатьох застосуваннях проблема, яка описана в одній системі координат, може бути ефективно розв'язана в іншій системи координат. В лінійній алгебрі система координат визначається базисом векторного простору, а вектори — координатами при вибраному базисі. Вибір «правильного» базису часто спрощує розв'язання задачі.

В якості прикладу розглянемо в просторі $\mathbb{R}^2$ два різних базиси (див. рис. 3.2): $\vec{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ та $\vec{c}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\vec{c}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Вектор $\vec{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ в базисі $\left(\vec{b}_1, \vec{b}_2\right)$ має координатний стовпець $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, а в базисі $\left(\vec{c}_1, \vec{c}_2\right)$ — $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Нас цікавить зв'язок між координатами одного і того ж самого вектора в різних базисах.

3.1. Матриця переходу від одного базису до іншого та її властивості.

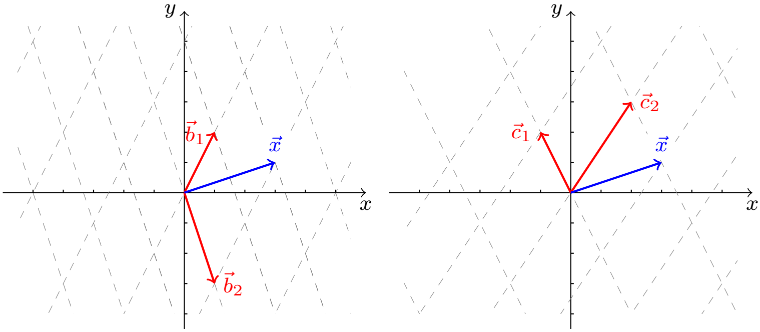


Рис. 3.2.

Нехай $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$, $\mathcal{C} = (\vec{c}_1, \ldots, \vec{c}_n)$ — два базиси векторного простору V. Виразимо вектори першого базису через вектори другого:

$$\begin{cases} \vec{b}_1 & = p_{11}\vec{c}_1 & + & p_{21}\vec{c}_2 & + & \ldots & + & p_{n1}\vec{c}_n, \\ \vec{b}_2 & = p_{12}\vec{c}_1 & + & p_{22}\vec{c}_2 & + & \ldots & + & p_{n2}\vec{c}_n, \\ \ldots \\ \vec{b}_n & = p_{1n}\vec{c}_1 & + & p_{2n}\vec{c}_2 & + & \ldots & + & p_{nn}\vec{c}_n, \end{cases} \tag{3.2}$$

Матриця

$$\begin{bmatrix} [\vec{b}_1]_{\mathcal{C}} & [\vec{b}_2]_{\mathcal{C}} & \ldots & [\vec{b}_n]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \ldots & p_{nn} \end{bmatrix}$$

називається *матрицею переходу від базису $\mathcal{B}$ до базису $\mathcal{C}$* і позначається $P_{\mathcal{C}\leftarrow\mathcal{B}}$.

Теорема 3.11 (Зв'язок між координатами вектора в різних базисах)**.** *Нехай $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$, $\mathcal{C} = (\vec{c}_1, \ldots, \vec{c}_n)$ — два базиси векторного простору V, $\vec{x} \in V$. Тоді*

$$[\vec{x}]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}\,[\vec{x}]_{\mathcal{B}}\,. \tag{3.3}$$

Доведення. Нехай $\vec{x} = \alpha_1\vec{b}_1 + \ldots + \alpha_n\vec{b}_n$.

Тоді, використовуючи наслідок з теореми 3.6, знаходимо:

$$\begin{aligned}[\vec{x}]_{\mathcal{C}} &= \left[\alpha_1 \vec{b}_1 + \ldots + \alpha_n \vec{b}_n\right]_{\mathcal{C}} = \\ &= \alpha_1 \left[\vec{b}_1\right]_{\mathcal{C}} + \ldots + \alpha_n \left[\vec{b}_n\right]_{\mathcal{C}} = \\ &= \left[\left[\vec{b}_1\right]_{\mathcal{C}} \quad \ldots \quad \left[\vec{b}_n\right]_{\mathcal{C}}\right] \cdot \begin{bmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{bmatrix} = P_{\mathcal{C}\leftarrow\mathcal{B}} [\vec{x}]_{\mathcal{B}} .\end{aligned}$$

■

Зауваження. Зауважимо, що матриця $P_{\mathcal{C}\leftarrow\mathcal{B}}$ є єдиною матрицею P такою, що рівність $[\vec{x}]_{\mathcal{C}} = P [\vec{x}]_{\mathcal{B}}$ виконується для кожного $\vec{x}$. Справді, нехай P така матриця n-го порядку, що для кожного $\vec{x}$ виконується рівність $[\vec{x}]_{\mathcal{C}} = P [\vec{x}]_{\mathcal{B}}$. Покладемо послідовно $\vec{x} = \vec{b}_i$, $i = \overline{1, n}$. Тоді $[\vec{x}]_{\mathcal{B}} = \vec{e}_i$, а результатом добутку $P\vec{e}_i$ є i-ий стовпець матриці P, який в свою чергу дорівнює $\left[\vec{b}_i\right]_{\mathcal{C}}$. Отже, стовпці матриць P та $P_{\mathcal{C}\leftarrow\mathcal{B}}$ співпадають, а тому вони рівні.

Зв'язок між координатами вектора $\vec{x}$ в різних базисах проілюстровано на рисунку 3.3.

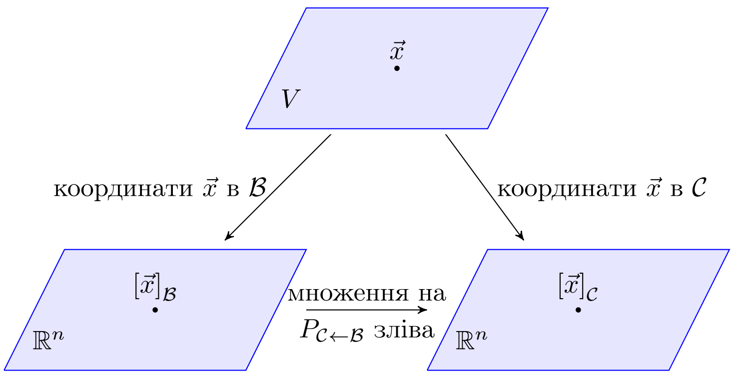


Рис. 3.3. Зв'язок між координатами вектора в різних базисах

Теорема 3.12. *Матриця переходу $P_{\mathcal{C}\leftarrow\mathcal{B}}$ оборотна, причому $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$.*

Доведення. За теоремою 3.7 знаходимо, що вектори-стовпці матриці переходу $\left[\left[\vec{b}_1\right]_{\mathcal{C}} \quad \left[\vec{b}_2\right]_{\mathcal{C}} \quad \ldots \quad \left[\vec{b}_2\right]_{\mathcal{C}}\right]$ є лінійно незалежними, а тому за основною теоремою оборотних матриць 2.7 вона є оборотною.

Тоді домноживши зліва обидві частини рівності 3.3 на $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}$ дістанемо

$$(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}\,[\vec{x}]_{\mathcal{C}} = [\vec{x}]_{\mathcal{B}}\,.$$

З іншого боку $[\vec{x}]_{\mathcal{B}} = P_{\mathcal{B}\leftarrow\mathcal{C}}\,[\vec{x}]_{\mathcal{C}}$. Оскільки дві останні рівності виконуються для довільного $\vec{x}$, то із зауваження до теореми 3.11 випливає, що

$$(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}.$$

■

Приклад 3.29. Знайти матрицю переходу від базису $\mathcal{B} = \left(1, x, x^2\right)$ до базису $\mathcal{C} = \left(1+x, x+x^2, 1+x^2\right)$ в векторному просторі $\mathcal{P}_2$. Знайти координати вектора $p(x) = 1 + 2x - x^2$ в базисі $\mathcal{C}$.

Розв'язання. Нам легше знайти матрицю переходу від базису $\mathcal{C}$ до базису $\mathcal{B}$. Справді,

$$[1+x]_{\mathcal{B}} = \begin{bmatrix}1\\1\\0\end{bmatrix},\ \left[x+x^2\right]_{\mathcal{B}} = \begin{bmatrix}0\\1\\1\end{bmatrix},\ \left[1+x^2\right]_{\mathcal{B}} = \begin{bmatrix}1\\0\\1\end{bmatrix},$$

а тому

$$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix}1 & 0 & 1\\1 & 1 & 0\\0 & 1 & 1\end{bmatrix}.$$

Тоді

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1} = \frac{1}{2}\begin{bmatrix}1 & 1 & -1\\-1 & 1 & 1\\1 & -1 & 1\end{bmatrix}.$$

$$[p(x)]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}\,[p(x)]_{\mathcal{B}} = \frac{1}{2}\begin{bmatrix}1 & 1 & -1\\-1 & 1 & 1\\1 & -1 & 1\end{bmatrix}\begin{bmatrix}1\\2\\-1\end{bmatrix} = \begin{bmatrix}2\\0\\-1\end{bmatrix}.$$

■

3.2. Обчислення матриці переходу від одного базиса до іншого методом Йордана–Гауса.

Приклад 3.29 показує, що дуже просто записати матрицю переходу до стандартного базису. Виникає питання: чи може це допомогти при знаходженні матриці переходу від одного базису до іншого?

Теорема 3.13. *Нехай*

$$\mathcal{B} = \left(\vec{b}_1, \ldots, \vec{b}_n\right),$$
$$\mathcal{C} = \left(\vec{c}_1, \ldots, \vec{c}_n\right),$$
$$\mathcal{E} = \left(\vec{e}_1, \ldots, \vec{e}_n\right) -$$

три базиси векторного простору V. Нехай нам відомо матриці переходу від базисів $\mathcal{B}$ та $\mathcal{C}$ до $\mathcal{E}$:

$$P_{\mathcal{E}\leftarrow\mathcal{B}} = \left[\left[\vec{b}_1\right]_{\mathcal{E}} \quad \ldots \quad \left[\vec{b}_n\right]_{\mathcal{E}}\right] \equiv B$$
$$P_{\mathcal{E}\leftarrow\mathcal{C}} = \left[[\vec{c}_1]_{\mathcal{E}} \quad \ldots \quad [\vec{c}_n]_{\mathcal{E}}\right] \equiv C.$$

Тоді матриця $[C|B]$ рядково евівалентна матриці $[I|P_{\mathcal{C}\leftarrow\mathcal{B}}]$.

Доведення. Нехай $P_{\mathcal{C}\leftarrow\mathcal{B}} = [p_{ij}]$ — матриця переходу від базису $\mathcal{B}$ до базису $\mathcal{C}$. Це означає, що $\vec{c}_k = p_{1k}\vec{b}_1 + \ldots + p_{nk}\vec{b}_n$. Перейдемо в цій рівності до базису $\mathcal{E}$:

$$\left[\vec{b}_k\right]_{\mathcal{E}} = [p_{1k}\vec{c}_1 + \ldots + p_{nk}\vec{c}_n]_{\mathcal{E}} = p_{1k}\,[\vec{c}_1]_{\mathcal{E}} + \ldots + p_{nk}\,[\vec{c}_n]_{\mathcal{E}}\,.$$

Це можна переписати у матричній формі

$$\left[\vec{b}_k\right]_{\mathcal{E}} = \left[[\vec{c}_1]_{\mathcal{E}} \quad \ldots \quad [\vec{c}_n]_{\mathcal{E}}\right] \cdot \begin{bmatrix} p_{1k} \\ \vdots \\ p_{nk} \end{bmatrix}. \tag{3.4}$$

Одержану систему лінійних рівнянь, що записана у матричній формі, можна розв'язати методом Йордана–Гауса, працюючи з розширеною матрицею цієї системи

$$\left[[\vec{c}_1]_{\mathcal{E}} \quad \ldots \quad [\vec{c}_n]_{\mathcal{E}} \quad | \quad \left[\vec{b}_k\right]_{\mathcal{E}}\right]$$

Таким чином, для знаходження усіх n стовпців матриці $P_{\mathcal{C}\leftarrow\mathcal{B}}$ нам потрібно розв'язати n систем лінійних рівнянь виду (3.4). Кожна з цих систем має одну і ту ж основну матрицю $\left[[\vec{c}_1]_{\mathcal{E}} \quad \ldots \quad [\vec{c}_n]_{\mathcal{E}}\right]$. Тому ми можемо здійснити одночасне розв'язування усіх систем, працюючи з розширеною матрицею виду

$$\left[[\vec{c}_1]_{\mathcal{E}} \quad \ldots \quad [\vec{c}_n]_{\mathcal{E}} \quad | \quad \left[\vec{b}_1\right]_{\mathcal{E}} \quad \ldots \left[\vec{b}_n\right]_{\mathcal{E}}\right].$$

Отже, якщо ліву половину цієї матриці звести за допомогою елементарних перетворень над рядками до одиничної, то права половина буде шуканою матрицею переходу від базису $\mathcal{B}$ до базису $\mathcal{C}$. ■

Зауваження. Якщо $\mathcal{E}$ — стандартний базис, то запропонований метод є дуже зручним, оскільки матриці $P_{\mathcal{E}\leftarrow\mathcal{B}}$ та $P_{\mathcal{E}\leftarrow\mathcal{C}}$ легко виписуються.

Теорема 3.14. *Нехай $\mathcal{B}$, $\mathcal{C}$, $\mathcal{E}$ — базиси скінченновимірного простору V. Тоді*

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{E}} P_{\mathcal{E}\leftarrow\mathcal{B}}. \tag{3.5}$$

Доведення. Справді, з означення матриці переходу знаходимо такі рівності:

$$\begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{bmatrix} = P^T_{\mathcal{C}\leftarrow\mathcal{E}} \begin{bmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{bmatrix}, \quad \begin{bmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{bmatrix} = P^T_{\mathcal{E}\leftarrow\mathcal{B}} \begin{bmatrix} \vec{e}_1 \\ \vdots \\ \vec{e}_n \end{bmatrix}, \quad \begin{bmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{bmatrix} = P^T_{\mathcal{C}\leftarrow\mathcal{B}} \begin{bmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{bmatrix}.$$

Тоді

$$\begin{bmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_n \end{bmatrix} = P^T_{\mathcal{E}\leftarrow\mathcal{B}} P^T_{\mathcal{C}\leftarrow\mathcal{E}} \begin{bmatrix} \vec{c}_1 \\ \vdots \\ \vec{c}_n \end{bmatrix}.$$

Тому

$$P^T_{\mathcal{E}\leftarrow\mathcal{B}} P^T_{\mathcal{C}\leftarrow\mathcal{E}} = P^T_{\mathcal{C}\leftarrow\mathcal{B}}.$$

Відомо, що для квадратних матриць A та B однакового порядку $(AB)^T = B^T A^T$ (властивості операції транспонування матриць нами будуть розглянуті в пункті 1.1, див. теорему 7.1). Тому має місце рівність (3.5). ■

Зауваження. Доведена теорема підкреслює доцільність позначення в матриці переходу написання справа наліво «від $\mathcal{B}$ до $\mathcal{C}$»: $\mathcal{C} \leftarrow \mathcal{B}$, оскільки в такому випадку працює «транзитивність» при записі $P_{\mathcal{C}\leftarrow\mathcal{E}} P_{\mathcal{E}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{B}}$.

§ 4. Перетин і сума підпросторів векторного простору

Означення 3.9. Нехай U і W — два підпростори векторного простору V заданого над полем P. *Перетином* цих двох підпросторів називається множина всіх тих векторів простору V, які належать як U, так і W:

$$U \cap W = \{\vec{x} \colon \vec{x} \in U, \vec{x} \in W\}.$$

Теорема 3.15. *Перетин двох підпросторів простору V є підпростір простору V.*

Доведення. Для доведення треба показати, що якщо $\vec{a}, \vec{b} \in U \cap W$, то $\vec{a} + \vec{b} \in U \cap W$ і для кожного $\alpha \in P$ $\alpha\vec{a} \in U \cap W$.

Нехай $\vec{a}, \vec{b} \in U \cap W$. Це означає, що $\vec{a}, \vec{b} \in U$ і $\vec{a}, \vec{b} \in W$. Оскільки U і W підпростори, то $\vec{a} + \vec{b} \in U$ і $\vec{a} + \vec{b} \in W$, звідки $\vec{a} + \vec{b} \in U \cap W$.

Далі, якщо $\vec{a} \in U \cap W$, то $\vec{a} \in U$ і $\vec{a} \in W$, тоді $\alpha\vec{a} \in U$, $\alpha\vec{a} \in W$, а тому і $\alpha\vec{a} \in U \cap W$. ■

Відмітимо, що об'єднання $U \cup W$ двох підпросторів простору V, взагалі кажучи, підпростором не буде, проте підпростором буде лінійна оболонка такого об'єднання.

Означення 3.10. Лінійна оболонка $L(U \cup W)$ об'єднання двох підпросторів U та W простору V називається *сумою* цих підпросторів і позначається $U + W$.

Очевидно, що якщо вектор $\vec{a} \in U + W$, то його можна подати у вигляді $\vec{a} = \vec{u} + \vec{w}$, де $\vec{u} \in U$, $\vec{w} \in W$. Тому означення 3.10 можна сформулювати ще в такий спосіб.

Означення 3.10′. *Сумою* підпросторів U і W простору V називається множина всіх тих векторів простору V, які можна подати у вигляді $\vec{u} + \vec{w}$, де $\vec{u} \in U$, $\vec{w} \in W$.

Взагалі кажучи, різних представлень вектора $\vec{a} \in U + W$ у вигляді суми двох векторів з підпросторів U та W може бути кілька. Якщо ж для кожного вектора суми двох підпросторів існує єдине представлення у вигляді $\vec{a} = \vec{u} + \vec{w}$, то така сума називається *прямою* і позначається $U \oplus W$.

Теорема 3.16. *Сума двох підпросторів є прямою тоді і тільки тоді, коли перетин цих підпросторів містить лише $\vec{0}$.*

Доведення. *Необхідність.* Нехай сума двох підпросторів U та W простору V є прямою. Доведемо, що $U \cap W = \{\vec{0}\}$.

Припустимо, що $U \cap W \neq \{\vec{0}\}$. Це означає, що існує $\vec{b} \in U \cap W$, відмінний від $\vec{0}$. Очевидно, що $\vec{b} \in U$, $\vec{b} \in W$. Розглянемо вектор $\vec{b}$ як елемент простору $\vec{U} + \vec{W}$. Він справді належить цьому простору, оскільки може бути поданий у вигляді

$$\vec{b} = \underbrace{\vec{b}}_{\in U} + \underbrace{\vec{0}}_{\in W} \quad \text{або}$$

$$\vec{b} = \underbrace{\vec{0}}_{\in U} + \underbrace{\vec{b}}_{\in W},$$

тобто існує два різних представлення одного вектора, що суперечить тому, що сума підпросторів є прямою.

Достатність. Нехай U та W — підпростори простору V такі, що $U \cap W = \{\vec{0}\}$. Припустимо, що $U + W \neq U \oplus W$. Це означає, що деякий вектор $\vec{a} \in U + W$ має два різних подання

$$\vec{a} = \vec{u}_1 + \vec{w}_1,$$

$$\vec{a} = \vec{u}_2 + \vec{w}_2, \vec{u}_1, \vec{u}_2 \in U, \vec{w}_1, \vec{w}_2 \in W.$$

Віднявши ці дві рівності дістанемо:

$$\vec{0} = (\vec{u}_1 - \vec{u}_2) + (\vec{w}_1 - \vec{w}_2)$$

або

$$\underbrace{\vec{u}_1 - \vec{u}_2}_{\in U} = \underbrace{-(\vec{w}_1 - \vec{w}_2)}_{\in W}.$$

Таким чином, підпростори U і W мають спільний вектор, причому це не $\vec{0}$, що суперечить умові. ∎

Теорема 3.17. *Розмірність суми двох підпросторів U і W простору V дорівнює сумі розмірностей цих підпросторів без розмірності їх перетину:*

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Доведення. Введемо такі позначення:

$$\dim U = d_1, \quad \dim(U + W) = d$$

$$\dim W = d_2, \quad \dim(U \cap W) = d_0.$$

Нехай

$$\vec{a}_1, \ldots, \vec{a}_{d_0} \tag{3.6}$$

є базисом простору $U \cap W$. Цей базис є лінійно незалежною системою векторів простору U, тому його за теоремою 3.10 можна доповнити до базису цього простору. Нехай

$$\vec{a}_1, \ldots, \vec{a}_{d_0}, \vec{b}_{d_0+1}, \ldots, \vec{b}_{d_1} \tag{3.7}$$

є базисом підпростору U.

Крім цього, система (3.6) є лінійно незалежною системою векторів простору W, тому її також можна доповнити до базису простору W. Нехай цим базисом буде

$$\vec{a}_1, \ldots, \vec{a}_{d_0}, \vec{c}_{d_0+1}, \ldots, \vec{c}_{d_2}. \tag{3.8}$$

Розглянемо систему векторів

$$\vec{a}_1, \ldots, \vec{a}_{d_0}, \vec{b}_{d_0+1}, \ldots, \vec{b}_{d_1}, \vec{c}_{d_0+1}, \ldots, \vec{c}_{d_2}. \tag{3.9}$$

Ця система складається з $d_1 + d_2 - d_0$ векторів. Покажемо, що ця система і буде базисом простору $U + W$. Для цього потрібно показати, по перше, що кожний вектор підпростору $U + W$ лінійно виражається через вектори системи (3.9), і, по-друге, що система (3.9) є лінійно незалежною.

Нехай $\vec{b} \in U + W$. Це означає, що $\vec{b} = \vec{u}_1 + \vec{w}_1$, де $\vec{u}_1 \in U$, $\vec{w}_1 \in W$. Оскільки $\vec{u}_1 \xrightarrow{\text{л.в.}} (3.7)$, $\vec{w}_1 \xrightarrow{\text{л.в.}} (3.8)$, то $\vec{u}_1 + \vec{w}_1 \xrightarrow{\text{л.в.}} (3.9)$.

Тепер доведемо, що система (3.9) є лінійно незалежною. Для цього потрібно показати, що рівність ($\star$)

$$\begin{aligned} &\alpha_1\vec{a}_1 + \ldots + \alpha_{d_0}\vec{a}_{d_0} + \\ &+\beta_{d_0+1}\vec{b}_{d_0+1} + \ldots + \beta_{d_1}\vec{b}_{d_1} + \\ &+\gamma_{d_0+1}\vec{c}_{d_0+1} + \ldots + \gamma_{d_2}\vec{c}_{d_2} = \vec{0} \end{aligned} \tag{3.10}$$

виконується лише, коли усі коефіцієнти дорівнюють нулю.

Перепишемо останню рівність у такому вигляді:

$$\begin{aligned} &\underbrace{\alpha_1\vec{a}_1 + \ldots + \alpha_{d_0}\vec{a}_{d_0} + \beta_{d_0+1}\vec{b}_{d_0+1} + \ldots + \beta_{d_1}\vec{b}_{d_1}}_{\vec{a}\in U} = \\ &= \underbrace{-\gamma_{d_0+1}\vec{c}_{d_0+1} - \ldots - \gamma_{d_2}\vec{c}_{d_2}}_{\vec{a}\in W}. \end{aligned}$$

Таким чином, вектор $\vec{a}$, який позначено вище, повинен належати підпростору $U \cap W$, і, отже, $\vec{a} \xrightarrow{\text{л.в.}} \vec{a}_1, \ldots, \vec{a}_{d_0}$, тобто $\vec{a} = \lambda_1\vec{a}_1 + \ldots + \lambda_{d_0}\vec{a}_{d_0}$. Тоді рівність (3.10) переписується у вигляді

$$\lambda_1\vec{a}_1 + \ldots + \lambda_{d_0}\vec{a}_{d_0} + \gamma_{d_0+1}\vec{c}_{d_0+1} + \ldots + \gamma_{d_2}\vec{c}_{d_2} = \vec{0}.$$

Але система (3.8) є лінійно незалежною, тому остання рівність можлива лише за умови, коли усі коефіцієнти дорівнюють нулю, зокрема $\gamma_{d_0+1} = \ldots = \gamma_{d_2} = 0$.

Тепер рівність (3.10) набуває виду

$$\alpha_1\vec{a}_1 + \ldots + \alpha_{d_0}\vec{a}_{d_0} + \beta_{d_0+1}\vec{b}_{d_0+1} + \ldots + \beta_{d_1}\vec{b}_{d_1} = \vec{0}.$$

Але система (3.7) є лінійно незалежною, тому остання рівність виконується лише при $\alpha_1 = \ldots = \alpha_{d_0} = \beta_{d_0+1} = \ldots = \beta_{d_1} = 0$.

Таким чином, рівність (3.10) виконується лише коли усі коефіцієнти дорівнюють нулю, а тому система (3.9) є лінійно незалежною. Це завершує доведення того, що система (3.9) є базисом простору $U + W$, а тому розмірність цього простору дорівнює $\dim U + \dim W - \dim(U \cap W)$. ■

З доведеної теореми і теореми 3.16 випливає таке твердження.

Наслідок. *Розмірність прямої суми $U \oplus W$ підпросторів U і W простору V дорівнює сумі розмірностей цих підпросторів.*

§ 5. Векторні простори зі скалярним множенням

Поняття скалярного множення векторів, яке буде означено нижче, є основою для подальшого введення довжини і відстані у довільному векторному просторі. Ми обмежуємось розглядом лише випадку векторного простору, що задано над полем $\mathbb{R}$.

5.1. Скалярне множення у векторному просторі.

Нехай V — векторний простір, заданий над полем $\mathbb{R}$.

Означення 3.11. *Скалярним множенням* у дійсному векторному просторі V називається операція, яка кожній парі $\vec{a}, \vec{b} \in V$ співставляє таке дійсне число (позначається $(\vec{a}, \vec{b})$), що для довільних $\vec{a}, \vec{b}, \vec{c} \in V$, $\lambda \in \mathbb{R}$ виконуються умови (*аксіоми скалярного множення*):

1) $(\vec{a}, \vec{b}) = (\vec{b}, \vec{a})$;
2) $(\vec{a}, \vec{b} + \vec{c}) = (\vec{a}, \vec{b}) + (\vec{a}, \vec{c})$;
3) $(\lambda\vec{a}, \vec{b}) = \lambda(\vec{a}, \vec{b})$;
4) $(\vec{a}, \vec{a}) \geqslant 0$, причому $(\vec{a}, \vec{a}) = 0$ тоді і тільки тоді, коли $\vec{a} = \vec{0}$.

Зазвичай скалярний добуток векторів позначається $\vec{a} \cdot \vec{b}$ або $(\vec{a}, \vec{b})$. Для уникнення непорозумінь з операціями матричного множення та скалярного множення, ми надалі будемо позначати скалярний добуток двох векторів $\vec{a}$ та $\vec{b}$ через $(\vec{a}, \vec{b})$.

Дійсний векторний простір зі скалярним множенням називається *евклідовим*.

Зауважимо, що поняття скалярного множення вводиться і для векторного простору, заданого над довільним полем P. Означення скалярного добутку у векторному просторі, заданому над полем $\mathbb{C}$ комплексних чисел, дещо інше: ним називається відповідність, яка кожній парі векторів $\vec{a}, \vec{b}$ цього простору співвідносить таке комплексне число, що виконуються наступні умови:

1) $\big(\vec{a}, \vec{b}\big) = \overline{\big(\vec{b}, \vec{a}\big)}$ (тут риска означає комплексне спряження);
2) $\big(\vec{a}, \vec{b} + \vec{c}\big) = \big(\vec{a}, \vec{b}\big) + \big(\vec{a}, \vec{c}\big)$;
3) $\big(\lambda\vec{a}, \vec{b}\big) = \lambda\big(\vec{a}, \vec{b}\big)$;
4) $\big(\vec{a}, \vec{a}\big) \geqslant 0$, причому $\big(\vec{a}, \vec{a}\big) = 0$ тоді і тільки тоді, коли $\vec{a} = \vec{0}$.

Комплексний простір зі скалярним множенням називається *унітарним*.

Скалярне множення, визначене за правилом $\big(\vec{a}, \vec{b}\big) = 0$ для довільних $\vec{a}, \vec{b} \in V$, називається *нульовим* (*виродженим*). В подальшому ми виключатимемо цей випадок з розгляду.

Приклад 3.30. У арифметичному векторному просторі $\mathbb{R}^n$ скалярне множення векторів можна означити так: якщо $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, то $\big(\vec{x}, \vec{y}\big) = x_1y_1 + \ldots + x_ny_n$. Так означене скалярне множення називається *стандартним*. Зазначимо, що розглядаючи вектори простору $\mathbb{R}^n$ як матриці-стовпці, можемо стандартне скалярне множення записати у вигляді матричного множення:

$$\big(\vec{x}, \vec{y}\big) = \vec{x}^T \cdot \vec{y}.$$

Приклад 3.31. Розглянемо ще один варіант введення скалярного множення в просторі $\mathbb{R}^n$. Зафіксуємо додатні дійсні числа $\omega_1, \ldots, \omega_n$ і парі векторів $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ поставимо у відповідність число $\big(\vec{x}, \vec{y}\big) = \omega_1x_1y_1 + \ldots + \omega_nx_ny_n$. Така відповідність задовольняє означення 3.11 і називається *ваговим скалярним множенням в просторі* $\mathbb{R}^n$. Зауважимо, що вагове скалярне множення можна подати у такому вигляді:

$$\big(\vec{x}, \vec{y}\big) = \vec{x}^T \cdot W \cdot \vec{y}, \quad \text{де } W = \begin{bmatrix} \omega_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \omega_n \end{bmatrix}.$$

Приклад 3.32. В просторі $\mathcal{P}_2$ скалярне множення можна означити так: якщо $p(x) = a_0 + a_1x + a_2x^2$, $q(x) = b_0 + b_1x + b_2x^2$, то $\big(p(x), q(x)\big) = a_0b_0 + a_1b_1 + a_2b_2$. Це *стандартне скалярне множення* простору $\mathcal{P}_2$.

Приклад 3.33. Розглянемо простір неперервних функцій $\mathcal{C}_{[a,b]}$. Довільним двом функціям f та g з цього простору поставимо у відповідність число

$$\big(f, g\big) = \int_a^b f(x)g(x)\,dx. \tag{3.11}$$

З курсу математичного аналізу відомо, що добуток двох неперервних функцій на відрізку функцій є неперервна на цьому ж відрізку функція, а також те, що якщо функція неперервна на відрізку, то вона інтегровна на цьому відрізку. Тому означена відповідність $\mathcal{C}_{[a,b]} \times \mathcal{C}_{[a,b]} \to \mathbb{R}$ коректно визначена.

Покажемо, що вона задовольняє умови 1–4 скалярного множення. Очевидно, що $\int_a^b f(x)g(x)\,dx = \int_a^b g(x)f(x)\,dx$, тобто перша умова виконується. Друга і третя умови виконуються у відповідності до лінійних властивостей визначеного інтеграла, які відомі з курсу математичного аналізу. Нарешті якщо $\int_a^b f^2(x)\,dx = 0$, то $f(x) \equiv 0$, що також є відомим твердженням математичного аналізу.

Зауважимо, що таке скалярне множення поширюється на будь-який підпростір простору $\mathcal{C}_{[a,b]}$, наприклад в просторі $\mathcal{P}_n$ скалярне множення можна ввести за формулою (3.11).

5.2. Властивості скалярного множення.

Теорема 3.18. *Нехай $\vec{a}, \vec{b}, \vec{c}$ — вектори векторного простору V зі скалярним множенням, λ — скаляр. Тоді*

1°. $\big(\vec{a} + \vec{b}, \vec{c}\big) = \big(\vec{a}, \vec{c}\big) + \big(\vec{b}, \vec{c}\big)$.

2°. $\big(\vec{a}, \lambda\vec{b}\big) = \lambda\big(\vec{a}, \vec{b}\big)$.

3°. $\left(\vec{a}, \vec{0}\right) = 0$.

Доведення. Застосуємо ланцюжок перетворень (тут цифра над знаком $=$ позначає номер умови з означення скалярного множення, яка використовується).

1°. $\left(\vec{a} + \vec{b}, \vec{c}\right) \overset{1}{=} \left(\vec{c}, \vec{a} + \vec{b}\right) \overset{2}{=} \left(\vec{c}, \vec{a}\right) + \left(\vec{c}, \vec{b}\right) \overset{1}{=} \left(\vec{a}, \vec{c}\right) + \left(\vec{b}, \vec{c}\right)$.

2°. $\left(\vec{a}, \lambda\vec{b}\right) \overset{1}{=} \left(\lambda\vec{b}, \vec{a}\right) \overset{3}{=} \lambda\left(\vec{b}, \vec{a}\right) \overset{1}{=} \lambda\left(\vec{a}, \vec{b}\right)$.

3°. $\left(\vec{a}, \vec{0}\right) = \left(\vec{a}, \vec{0} + \vec{0}\right) \overset{2}{=} \left(\vec{a}, \vec{0}\right) + \left(\vec{a}, \vec{0}\right)$. Тоді $\left(\vec{a}, \vec{0}\right)$ є нульовим елементом поля $\mathbb{R}$, тобто дорівнює 0.

■

Теорема 3.19. *Будь-який дійсний скінченновимірний векторний простір завжди можна перетворити в евклідовий векторний простір.*

Доведення. Нехай V — n-вимірний дійсний векторний простір, $\vec{e}_1, \ldots, \vec{e}_n$ — деякий базис цього простору. Для векторів

$$\vec{a} = \alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n,$$

$$\vec{b} = \beta_1\vec{e}_1 + \ldots + \beta_n\vec{e}_n$$

покладемо

$$\left(\vec{a}, \vec{b}\right) = \alpha_1\beta_1 + \ldots + \alpha_n\beta_n.$$

Легко перевірити, що умови 1–4 означення скалярного множення виконуються. Цим самим векторний простір V перетворено в евклідовий. ■

5.3. Довжина, ортогональність і відстань у евклідовому векторному просторі.

Означення 3.12. *Нормою*, або *довжиною* вектора $\vec{a}$ евклідового векторного простору V називається дійсне число $\|\vec{a}\| = \sqrt{\left(\vec{a}, \vec{a}\right)}$. Вектор $\vec{a}$ називається *одиничним* (*нормованим*, або *ортом*), якщо $\|\vec{a}\| = 1$.

Вказане означення є коректним, враховуючи умову 4 означення скалярного множення.

Будь-який ненульовий вектор $\vec{a}$ евклідового векторного простору V можна перетворити в нормований вектор, множенням цього вектора на число $\frac{1}{\|\vec{a}\|}$. Справді,

$$\|\frac{1}{\|\vec{a}\|}\vec{a}\| = \sqrt{\big(\frac{1}{\|\vec{a}\|}\vec{a}, \frac{1}{\|\vec{a}\|}\vec{a}\big)} = \frac{1}{\|\vec{a}\|}\sqrt{\big(\vec{a},\vec{a}\big)} = \frac{1}{\|\vec{a}\|}\|\vec{a}\| = 1.$$

Означення 3.13. Вектори $\vec{a}, \vec{b}$ евклідового простору V називаються *ортогональними*, якщо $\big(\vec{a}, \vec{b}\big) = 0$.

Те, що вектори $\vec{a}$ та $\vec{b}$ ортогональні позначатимемо так: $\vec{a} \perp \vec{b}$.

Теорема 3.20 (Теорема Піфагора)**.** *Вектори $\vec{a}$ і $\vec{b}$ евклідового векторного простору V є ортогональними тоді і тільки тоді, коли*

$$\|\vec{a}+\vec{b}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2. \tag{3.12}$$

Доведення. За аксіомами і властивостями скалярного множення маємо:

$$\begin{aligned}\|\vec{a}+\vec{b}\|^2 &= \big(\vec{a}+\vec{b}, \vec{a}+\vec{b}\big) = \\ &= \big(\vec{a},\vec{a}\big) + \big(\vec{a},\vec{b}\big) + \big(\vec{b},\vec{a}\big) + \big(\vec{b},\vec{b}\big) = \\ &= \|\vec{a}\|^2 + 2\big(\vec{a},\vec{b}\big) + \|\vec{b}\|^2.\end{aligned}$$

Тепер очевидно, що рівність (3.12) виконується тоді і тільки тоді, коли $\big(\vec{a}, \vec{b}\big) = 0$. ■

Властивості норми:

1°. $\|\vec{a}\| > 0$ якщо $\vec{a} \neq \vec{0}$.

2°. $\|\lambda\vec{a}\| = |\lambda| \cdot \|\vec{a}\|$.

Справді, $\|\lambda\vec{a}\| = \sqrt{\big(\lambda\vec{a}, \lambda\vec{a}\big)} = \sqrt{\lambda^2\big(\vec{a},\vec{a}\big)} = |\lambda| \cdot \|\vec{a}\|$.

3°.

$$\Big|\big(\vec{a},\vec{b}\big)\Big| \leqslant \|\vec{a}\| \cdot \|\vec{b}\| \tag{3.13}$$

(нерівність Коші–Буняковського[1]–Шварца[2]).

Доведення. Потрібно довести, що

$$\big(\vec{a},\vec{b}\big) \leqslant \|\vec{a}\| \cdot \|\vec{b}\|,$$

$$-\big(\vec{a},\vec{b}\big) \leqslant \|\vec{a}\| \cdot \|\vec{b}\|.$$

---

[1]Віктор Якович Буняковський (1804–1889) — російський математик.

[2]Karl Hermann Amandus Schwarz (1843–1921) — німецький математик.

Якщо хоча б один з векторів $\vec{a}$ та $\vec{b}$ дорівнює $\vec{0}$, то нерівність, очевидно, виконується. Розглянемо випадок, коли обидва вектори ненульові. За властивістю 1° для довільних $\alpha, \beta \in \mathbb{R}$:

$$\big(\alpha\vec{a} - \beta\vec{b}, \alpha\vec{a} - \beta\vec{b}\big) \geqslant 0;$$
$$\alpha^2\big(\vec{a}, \vec{a}\big) - 2\alpha\beta\big(\vec{a}, \vec{b}\big) + \beta^2\big(\vec{b}, \vec{b}\big) \geqslant 0.$$

Покладемо $\alpha = \|\vec{b}\|$, $\beta = \|\vec{a}\|$. Тоді

$$\|\vec{b}\|^2\big(\vec{a}, \vec{a}\big) - 2\|\vec{b}\|\|\vec{a}\|\big(\vec{a}, \vec{b}\big) + \|\vec{a}\|^2\big(\vec{b}, \vec{b}\big) \geqslant 0,$$
$$\|\vec{b}\|^2\|\vec{a}\|^2 - 2\|\vec{b}\|\|\vec{a}\|\big(\vec{a}, \vec{b}\big) + \|\vec{a}\|^2\|\vec{b}\|^2 \geqslant 0; \quad |: \|\vec{a}\| \cdot \|\vec{b}\|,$$
$$\|\vec{b}\|\|\vec{a}\| - 2\big(\vec{a}, \vec{b}\big) + \|\vec{a}\|\|\vec{b}\| \geqslant 0,$$

звідки $\big(\vec{a}, \vec{b}\big) \leqslant \|\vec{a}\|\|\vec{b}\|$. Поклавши замість $\vec{a}$ в останню рівність вектор $-\vec{a}$ одержимо $-\big(\vec{a}, \vec{b}\big) \leqslant \|\vec{a}\|\|\vec{b}\|$, що і потрібно було довести. ∎

4°. $\|\vec{a} + \vec{b}\| \leqslant \|\vec{a}\| + \|\vec{b}\|$ (нерівність трикутника).

Доведення.

$$\|\vec{a} + \vec{b}\|^2 = \big(\vec{a} + \vec{b}, \vec{a} + \vec{b}\big) = \|\vec{a}\|^2 + 2\big(\vec{a}, \vec{b}\big) + \|\vec{b}\|^2 \leqslant$$
$$\leqslant \|\vec{a}\|^2 + 2\left|\big(\vec{a}, \vec{b}\big)\right| + \|\vec{b}\|^2 \overset{3^\circ}{\leqslant}$$
$$\leqslant \|\vec{a}\|^2 + 2\|\vec{a}\|\|\vec{b}\| + \|\vec{b}\|^2 = (\|\vec{a}\| + \|\vec{b}\|)^2,$$

звідки безпосередньо знаходимо нерівність, що доводиться. ∎

Надзвичайно важливим є означення поняття відстані між векторами у векторному просторі, адже це дає можливість в подальшому означити основну операцію математичного аналізу — граничний перехід. Відстанню назвемо кожне відображення, яке двом векторам $\vec{a}, \vec{b}$ простору V співвідносить таке дійсне число $d(\vec{a}, \vec{b})$, що виконуються наступні три умови для довільних векторів $\vec{a}, \vec{b}, \vec{c}$ простору:

1) $d(\vec{a}, \vec{b}) \geqslant 0$, причому $d(\vec{a}, \vec{b}) = 0$ лише тоді, коли $\vec{a} = \vec{b}$;
2) $d(\vec{a}, \vec{b}) = d(\vec{b}, \vec{a})$ (аксіома симетрії);
3) $d(\vec{a}, \vec{b}) \leqslant d(\vec{a}, \vec{c}) + d(\vec{b}, \vec{c})$ (аксіома трикутника).

Означення 3.14. *Відстанню* між векторами $\vec{a}$ і $\vec{b}$ називається число $d(\vec{a}, \vec{b}) = \|\vec{a} - \vec{b}\|$.

Вказане означення є коректним, оскільки воно задовольняє наведеним вище аксіомам відстані (подумайте, чому). Як бачимо, в різний спосіб означений скалярний добуток породжує різні тлумачення відстані у векторному просторі.

Означення 3.15. *Кутом* між ненульовими векторами $\vec{a}$ і $\vec{b}$ називається кут φ, що визначається співвідношенням

$$\cos\varphi = \frac{(\vec{a}, \vec{b})}{\|\vec{a}\| \cdot \|\vec{b}\|}, \quad 0 \leqslant \varphi \leqslant \pi.$$

Коректність означення кута між векторами випливає з нерівності Коші–Буняковського–Шварца.

5.4. Ортогональні системи векторів.

Означення 3.16. Система векторів $(\vec{a}_1, \ldots, \vec{a}_k)$ векторного простору V зі скалярним множенням називається *ортогональною*, якщо кожний вектор цієї системи ортогональний до будь-якого іншого вектора системи.

Теорема 3.21. *Ортогональна система ненульових векторів векторного простору зі скалярним множенням є лінійно незалежною.*

Доведення. Нехай $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ — ортогональна система ненульових векторів простору V. Потрібно показати, що складена для неї рівність $(\star)$

$$\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_k\vec{a}_k = \vec{0} \tag{3.14}$$

виконується лише при $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

Домножимо обидві частини останньої рівності скалярно на вектор $\vec{a}_1$ і скористаємось властивостями скалярного множення:

$$\big(\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_k\vec{a}_k, \vec{a}_1\big) = \big(\vec{0}, \vec{a}_1\big),$$
$$\alpha_1\big(\vec{a}_1, \vec{a}_1\big) + \alpha_2\underbrace{\big(\vec{a}_2, \vec{a}_1\big)}_{=0} + \ldots + \alpha_k\underbrace{\big(\vec{a}_k, \vec{a}_1\big)}_{=0} = 0,$$

звідки $\alpha_1\big(\vec{a}_1, \vec{a}_1\big) = 0$, що можливо лише при $\alpha_1 = 0$, оскільки $\vec{a}_1 \neq \vec{0}$.

Помноживши (3.14) послідовно на $\vec{a}_2, \ldots, \vec{a}_k$ одержимо, що $\alpha_2 = \ldots = \alpha_k = 0$, а тому система векторів $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k$ є лінійно незалежною. ■

Означення 3.17. Ортогональна система ненульових векторів $\vec{a}_1, \ldots, \vec{a}_n$ n-вимірного простору V називається *ортогональним базисом* цього простору.

Означення 3.18. Система векторів $\vec{a}_1, \ldots, \vec{a}_k$ евклідового векторного простору V називається *ортонормованою*, якщо вона ортогональна і кожний її вектор нормований.

Означення 3.19. Система ненульових векторів n-вимірного евклідового простору V називається *ортонормованим базисом* цього простору, якщо вона ортонормована і містить n векторів.

5.5. Ортогональне доповнення до підпростору векторного простору.

Нехай V — n-вимірний векторний простір зі скалярним множенням, U, W — підпростори цього простору.

Означення 3.20. Підпростір W називається *ортогональним* до підпростору U, якщо кожний вектор з W ортогональний до кожного вектора з U.

Приклад 3.34. Розглянемо геометричну інтерпретацію векторного простору $\mathbb{R}^3$ як множину тривимірних векторів з початком у початку координат. Площина xOy та вісь Oz є, очевидно, підпросторами цього простору. Крім цього, кожний вектор з осі Oz перпендикулярний (ортогональний) до кожного вектора з площини xOy. Таким чином, ці підпростори ортогональні один до одного.

Якщо вектор $\vec{a}$ ортогональний до кожного вектора підпростору U, то будемо говорити, що вектор $\vec{a}$ ортогональний до підпростору U.

Зрозуміло, що підпростір W буде ортогональним до підпростору U, якщо кожен вектор із W ортогональний до кожного вектора базису підпростору U. Справді, якщо кожний вектор із W ортогональний до базису $\vec{e}_1, \ldots, \vec{e}_r$ підпростору U, тобто $\vec{a} \perp \vec{e}_1, \ldots, \vec{a} \perp \vec{e}_r$, то якщо взяти вектор $\vec{b} = \alpha_1 \vec{e}_1 + \ldots + \alpha_r \vec{e}_r$, то

$$(\vec{a}, \vec{b}) = \alpha_1 \underbrace{(\vec{a}, \vec{e}_1)}_{=0} + \ldots + \alpha_r \underbrace{(\vec{a}, \vec{e}_r)}_{=0} = 0,$$

що говорить про те, що $\vec{a} \perp \vec{b}$.

Позначимо через $U^\perp$ множину всіх векторів простору V ортогональних до підпростору U.

Теорема 3.22. *Нехай U — підпростір простору V. Тоді $U^\perp$ також є підпростором простору V.*

Доведення. Згідно з критерієм підпростору достатньо довести, що $U^\perp$ є замкнутою множиною відносно операцій додавання векторів і множення їх на скаляри.

Нехай $\vec{a}, \vec{b} \in U^\perp$. Це означає, що для будь-якого $\vec{c} \in U$: $\big(\vec{a}, \vec{c}\big) = \big(\vec{b}, \vec{c}\big) = 0$. Тоді $\big(\vec{a}+\vec{b}, \vec{c}\big) = \big(\vec{a}, \vec{c}\big) + \big(\vec{b}, \vec{c}\big) = 0$, а тому $\vec{a}+\vec{b} \in U^\perp$. Також

$$\big(\alpha\vec{a}, \vec{c}\big) = \alpha\big(\vec{a}, \vec{c}\big) = \alpha \cdot 0 = 0,$$

звідки отримуємо $\alpha\vec{a} \in U^\perp$. Отже, $U^\perp$ — підпростір простору V. ■

Очевидно, що $U^\perp$ буде «найбільшим» підпростором, ортогональним до підпростору U в тому розумінні, що кожний інший ортогональний підпростір до підпростору U буде його підмножиною.

Означення 3.21. Нехай U — підпростір простору V. Підпростір $U^\perp$ називається *ортогональним доповненням* до підпростору U.

Серед властивостей ортогонального доповнення до підпростору відмітимо такі:

- $V = U \oplus U^\perp$;
- $(U^\perp)^\perp = U$;
- $U \cap U^\perp = \{\vec{0}\}$;
- $\dim U + \dim U^\perp = \dim V$.

Доведення першої властивості ми наведемо в п. 5.7 (теорема 3.24). Інші властивості є її простими наслідками.

5.6. Процес ортогоналізації лінійно незалежної системи векторів.

Дослідимо таке питання: чи існує, а якщо існує, то як побудувати, ортогональний базис довільного векторного простору? Розглянемо це на прикладі.

Приклад 3.35. Побудувати ортогональний базис підпростору $L(\vec{a}_1, \vec{a}_2)$ векторного простору $\mathbb{R}^3$, якщо

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Розв'язання. Оскільки координати векторів не пропорційні, то вони лінійно незалежні і утворюють базис підпростору $L(\vec{a}_1, \vec{a}_2)$. Але оскільки $\left(\vec{a}_1, \vec{a}_2\right) = -2 \neq 0$, то цей базис не ортогональний.

Основна ідея в побудові ортогонального базиса $\left(\vec{b}_1, \vec{b}_2\right)$ полягає в наступному. Нехай $\vec{b}_1 = \vec{a}_1$, а вектор $\vec{b}_2$ сконструюємо таким, щоб виконувались дві умови: 1) $\vec{b}_2 \perp \vec{b}_1$; 2) $\vec{b}_2 \in L(\vec{a}_1, \vec{a}_2)$.

Друга умова буде дотримана, якщо $\vec{b}_2$ подати у вигляді

$$\vec{b}_2 = \vec{a}_2 - \lambda \vec{b}_1, \quad \lambda \in \mathbb{R}.$$

Підберемо λ таким, щоб $\left(\vec{b}_2, \vec{b}_1\right) = 0$. З цією метою домножимо обидві частини останньої рівності скалярно на $\vec{b}_1$. Дістанемо:

$$0 = \left(\vec{a}_2, \vec{b}_1\right) - \lambda\left(\vec{b}_1, \vec{b}_1\right),$$

звідки $\lambda = \frac{\left(\vec{a}_2, \vec{b}_1\right)}{\left(\vec{b}_1, \vec{b}_1\right)}$. Обчислюємо:

$$\left(\vec{a}_2, \vec{b}_1\right) = -2;$$
$$\left(\vec{b}_1, \vec{b}_1\right) = 2.$$

Отже, $\lambda = -1$ і $\vec{b}_2 = \vec{a}_2 + \vec{b}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$.

Таким чином, ортогональним базисом підпростору $L(\vec{a}_1, \vec{a}_2)$ є, наприклад, $\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right)$. ■

Вказаний метод розв'язування задачі побудови ортогонального базису підпростору векторного простору лежить в основі більш загального результату, який носить назву *процесу ортогоналізації лінійно незалежної системи векторів* (або *процесу Грама*[1]*–Шмідта*[2]).

[1] Jörgen Pedersen Gram (1850–1916) — датський математик.

[2] Erhard Schmidt (1876–1959) — німецький математик.

Теорема 3.23. *Нехай* $\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right)$ *— лінійно незалежна система векторів простору* V*. Нехай також*

$$\begin{aligned}
\vec{b}_1 =& \vec{a}_1, \\
\vec{b}_2 =& \vec{a}_2 - \frac{\left(\vec{b}_1, \vec{a}_2\right)}{\left(\vec{b}_1, \vec{b}_1\right)}\vec{b}_1, \\
\vec{b}_3 =& \vec{a}_3 - \frac{\left(\vec{b}_1, \vec{a}_3\right)}{\left(\vec{b}_1, \vec{b}_1\right)}\vec{b}_1 - \frac{\left(\vec{b}_2, \vec{a}_3\right)}{\left(\vec{b}_2, \vec{b}_2\right)}\vec{b}_2, \\
&\vdots \\
\vec{b}_k =& \vec{a}_k - \frac{\left(\vec{b}_1, \vec{a}_k\right)}{\left(\vec{b}_1, \vec{b}_1\right)}\vec{b}_1 - \ldots - \frac{\left(\vec{b}_{k-1}, \vec{a}_k\right)}{\left(\vec{b}_{k-1}, \vec{b}_{k-1}\right)}\vec{b}_{k-1}.
\end{aligned} \tag{3.15}$$

Тоді $\left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_k\right)$ *є ортогональною системою векторів простору* V*, причому* $L\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right) = L\left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_k\right)$.

Доведення. Доведемо твердження теореми індукцією по кількості векторів. Якщо система складається з одного вектора, то, очевидно, твердження теореми має місце.

Припустимо, що для деякого довільного, але фіксованого $k < \dim V$ твердження теореми виконується, тобто для довільної лінійно незалежної системи векторів $\vec{a}_1, \ldots, \vec{a}_k$ система $\vec{b}_1, \ldots, \vec{b}_k$, яка побудована згідно з (3.15), є ортогональною, причому $L\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\right) = L\left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_k\right)$. Доповнимо систему $\vec{a}_1, \ldots, \vec{a}_k$ вектором $\vec{a}_{k+1}$ так, щоб не порушити її лінійну незалежність. Доведемо, що вектор

$$\vec{b}_{k+1} = \vec{a}_{k+1} - \frac{\left(\vec{b}_1, \vec{a}_{k+1}\right)}{\left(\vec{b}_1, \vec{b}_1\right)}\vec{b}_1 - \frac{\left(\vec{b}_2, \vec{a}_{k+1}\right)}{\left(\vec{b}_2, \vec{b}_2\right)}\vec{b}_2 - \ldots - \frac{\left(\vec{b}_k, \vec{a}_{k+1}\right)}{\left(\vec{b}_k, \vec{b}_k\right)}\vec{b}_k$$

ортогональний до кожного з векторів $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_k$. Справді, домноживши обидві частині останньої рівності на кожний з векторів $\vec{b}_1, \ldots, \vec{b}_k$, ми неодмінно одержимо, що $\left(\vec{b}_{k+1}, \vec{b}_i\right) = 0$ для всіх $i = \overline{1, k}$. Крім цього, $b_{k+1} \xrightarrow{\text{л.в.}} \vec{a}_1, \ldots, \vec{a}_{k+1}$, і, враховуючи лінійну незалежність системи векторів $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{k+1}$ і припущення індукції, стверджуємо, що $L\left(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_{k+1}\right) = L\left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_{k+1}\right)$. ■

Наслідок. *Раніше ми довели (див. теорему 3.10), що кожну лінійно незалежну систему можна доповнити до базису скінченновимірного векторного простору. Процес Грама–Шмідта показує, що цю систему можна «перетворити» в ортогональний базис цього простору. Таким чином, векторний простір зі скалярним множенням завжди має ортогональний базис. Ортонормований базис можна одержати шляхом нормування кожного вектора ортогонального базису.*

5.7. Ортогональна проекція вектора на підпростір. Спочатку доведемо таке твердження, яке було проанонсоване вище.

Теорема 3.24 (про ортогональний розклад простору)**.** *Для довільного підпростору U векторного простору V має місце розклад $V = U \oplus U^{\perp}$, тобто довільний вектор $\vec{x} \in V$ єдиним чином може бути поданий у вигляді $\vec{x} = \vec{u} + \vec{u}^{\perp}$, де $\vec{U} \in U$, $\vec{u}^{\perp} \in U^{\perp}$.*

Доведення. Нехай $\vec{e}_1, \ldots, \vec{e}_k$ — ортогональний базис підпростору U (факт його існування доведено в попередньому розділі 5.6). Ці вектори можна доповнити до базису простору V: $\vec{e}\,'_{k+1}, \ldots, \vec{e}\,'_n$, а застосувавши до цих векторів процес Грама–Шмідта дістанемо ортогональний базис $\vec{e}_1, \ldots, \vec{e}_n$. Разом з тим вектори $\vec{e}_{k+1}, \ldots, \vec{e}_n$ належать підпростору $U^{\perp}$ за означенням ортогонального доповнення. Отже, кожний вектор $\vec{x} \in V$ подається у вигляді

$$\vec{x} = \underbrace{\alpha_1\vec{e}_1 + \ldots + \alpha_k\vec{e}_k}_{\in U} + \underbrace{\alpha_{k+1}\vec{e}_{k+1} + \alpha_n\vec{e}_n}_{\in U^{\perp}},$$

тобто $V = U + U^{\perp}$.

Доведемо, що сума $U + U^{\perp}$ є прямою. Нехай $\vec{a} \in U \cap U^{\perp}$. Оскільки $\vec{a} \in U^{\perp}$, то $\big(\vec{a}, \vec{x}\big) = 0$ для всіх $\vec{x} \in U$. Можемо покласти $\vec{x} = \vec{a}$, оскільки $\vec{a} \in U$. Тоді $\big(\vec{a}, \vec{a}\big) = 0$, тобто $\vec{a} = \vec{0}$. Отже, $U \cap U^{\perp} = \{\vec{0}\}$ і за теоремою 3.16 сума цих підпросторів є прямою. ■

Доведена теорема дає можливість для такого означення.

Означення 3.22. Нехай U — підпростір векторного простору V. Для довільного вектора $\vec{x} \in V$ запишемо розклад $\vec{x} = \vec{v}_1 + \vec{v}_2$, де $\vec{v}_1 \in U$, $\vec{v}_2 \in U^{\perp}$. Тоді вектор $\vec{v}_1$ називається *ортогональною проекцією вектора $\vec{x}$ на підпростір U* і позначається $\mathrm{pr}_U\,\vec{x}$, а вектор

$\vec{v}_2$ — *ортогональною складовою вектора* $\vec{x}$ *відносно підпростору* U і позначається $\operatorname{ort}_U \vec{x}$.

Очевидно, що $\operatorname{ort}_U \vec{x} = x - \operatorname{pr}_U \vec{x} = \operatorname{pr}_{U^\perp} \vec{x}$.

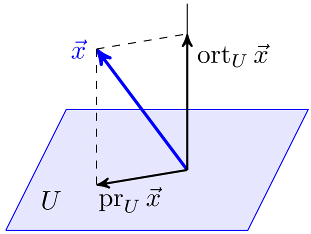


Рис. 3.4. Ортогональна проекція вектора на підпростір.

Теорема 3.25. *Нехай* $(\vec{e}_1, \ldots, \vec{e}_k)$ — *ортогональний базис підпростору* U *лінійного простору* V. *Тоді ортогональна проекція вектора* $\vec{x}$ *на* U *є такою лінійною комбінацією:*

$$\operatorname{pr}_U \vec{x} = \lambda_1 \vec{e}_1 + \ldots + \lambda_k \vec{e}_k,$$

де $\lambda_i = \dfrac{(\vec{e}_i, \vec{x})}{(\vec{e}_i, \vec{e}_i)}$ *для всіх* $i \in \{1, \ldots, k\}$.

Доведення. Справді, оскільки $\operatorname{pr}_U \vec{x} \in U$, то рівність $\operatorname{pr}_U \vec{x} = \lambda_1 \vec{e}_1 + \ldots + \lambda_k \vec{e}_k$ повинна виконуватись для деяких λ_i. Якщо обидві частини останньої рівності скалярно помножити на $\vec{e}_i$ і використати той факт, що $(\vec{e}_i, \vec{e}_j) = 0$ при $i \neq j$ (вектори $\vec{e}_1, \ldots, \vec{e}_k$ попарно ортогональні), то знайдемо значення λ_i. ■

Розглянемо ілюстрацію останньої теореми на прикладі.

Приклад 3.36. Знайти ортогональну проекцію вектора $\vec{x} = (3, -1, 2)$ на площину $x - y + 2z = 0$ простору $\mathbb{R}^3$.

Розв'язання. Розглянемо заданий в умові підпростір простору $\mathbb{R}^3$ $\{(x, x + 2z, z) \colon x, z \in \mathbb{R}\}$. Легко знайти базис цього підпростору, наприклад, це

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \vec{v}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.$$

Проте цей базис не ортогональний і до нього ми не можемо застосувати теорему 3.25. Перетворимо цей базис у ортогональний з допомогою процесу Грама–Шмідта: нехай $\vec{e}_1 = \vec{v}_1$,

$$\vec{e}_2 = \vec{v}_2 - \frac{(\vec{v}_2, \vec{v}_1)}{(\vec{v}_1, \vec{v}_1)}\vec{v}_1 = \begin{bmatrix}0\\2\\1\end{bmatrix} - \frac{2}{2}\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}-1\\1\\1\end{bmatrix}.$$

Обчислимо:

$$(\vec{e}_1, \vec{x}) = 2, \quad (\vec{e}_2, \vec{x}) = -2,$$
$$(\vec{e}_1, \vec{e}_1) = 2, \quad (\vec{e}_2, \vec{e}_2) = 3.$$

Тоді

$$\begin{aligned}\mathrm{pr}_U\,\vec{x} &= \frac{(\vec{e}_1, \vec{x})}{(\vec{e}_i, \vec{e}_i)}\vec{e}_1 + \frac{(\vec{e}_2, \vec{x})}{(\vec{e}_2, \vec{e}_2)}\vec{e}_2 = \\ &= \frac{2}{2}\begin{bmatrix}1\\1\\0\end{bmatrix} - \frac{2}{3}\begin{bmatrix}-1\\1\\1\end{bmatrix} = \begin{bmatrix}5/3\\1/3\\-2/3\end{bmatrix}.\end{aligned}$$

Між іншим, ортогональною складовою є вектор

$$\mathrm{ort}_U\,\vec{x} = \vec{x} - \mathrm{pr}_U\,\vec{x} = \begin{bmatrix}3\\-1\\2\end{bmatrix} - \begin{bmatrix}5/3\\1/3\\-2/3\end{bmatrix} = \begin{bmatrix}4/3\\-4/3\\8/3\end{bmatrix}.$$

Відповідь. $\{(\frac{5}{3}, \frac{1}{3}, -\frac{2}{3})\}$. ■

Додаткові відомості і завдання для самостійної роботи

1. Нехай U та V — векторні простори, задані над полем P. *Декартовим добутком* цих просторів називається
$$U \times V = \{(\vec{u}, \vec{v})\colon \vec{u} \in U, \vec{v} \in V\}.$$
Довести, що $U \times V$ є векторним простором.
2. Обчисліть $\dim(U \times V)$ в термінах $\dim U$ та $\dim V$.
3. Нехай f та g — неперервно диференційовні функції. Детермінант
$$W(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix}$$
називається *детермінантом Вронського* [1] (*вронскіаном*) функцій f та g. Довести, що якщо f та g є лінійно незалежними, то їх вронскіан не є тотожно рівним нулю (тобто існує таке x, що $W(x) \neq 0$).

[1] Józef Maria Hoene-Wroński (1776–1853) — польський математик і філософ.

4. В загальному випадку *вронскіаном функцій* $f_1, f_2, \ldots, f_n$, які є неперервними і мають неперервні похідні до $n-1$-го порядку включно, називається детермінант

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \ldots & f_n^{(n-1)}(x) \end{vmatrix}.$$

Довести, що якщо $f_1, \ldots, f_n$ лінійно незалежні, то $W(x)$ не є тотожно рівним нулю.

5. Довести, що коли розмірність суми двох підпросторів простору $\mathbb{R}^n$ на 1 більша від розмірності їх перетину, то сума співпадає з одним з них, а перетин — з іншим.
6. Довести, що n-вимірний векторний простір V при $n \geqslant 2$ можна подати у вигляді прямої суми n його одновимірних підпросторів.
7. Нехай U — довільний підпростір простору V, $\vec{a} \in V$. Множина

$$M = \vec{a} + U = \{x \in V \colon \vec{x} = \vec{a} + \vec{u}, \vec{u} \in U\}$$

називається *лінійним многовидом* простору V, що утворений паралельним перенесенням підпростору U на вектор $\vec{a}$, а розмірність $\dim U$ називають його *розмірністю.* Довести, що для різних векторів $\vec{a}, \vec{b}$ векторного простору V існує єдиний одновимірний лінійний многовид, який містить множину векторів $\{\alpha\vec{a} + \beta\vec{b} \operatorname{col} \alpha + \beta = 1\}$.

8. Нехай маємо два лінійні многовиди $S = \vec{a} + U$, $Q = \vec{b} + W$ векторного простору V. Довести, що $S = Q$ тоді і тільки тоді, коли $U = W$, $\vec{a} - \vec{b} \in U$.
9. Довести, що два лінійні многовиди $S = \vec{a} + U$ і $Q = \vec{b} + U$ векторного простору V або рівні, або не мають спільних елементів.
10. Довести, що в евклідовому векторному просторі всіх неперервних на відрізку $[-\pi, \pi]$ функцій зі скалярним множенням

$$\big(f(x), g(x)\big) = \int_{-\pi}^{\pi} f(x) g(x)\, dx$$

множина функцій $\{1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots, \sin nx, \cos nx\}$ є ортогональною.

11. Довести, що $\|\vec{u} + \vec{v}\|^2 + \|\vec{u} - \vec{v}\|^2 = 2\|\vec{u}\|^2 + 2\|\vec{v}\|^2$ для довільних $\vec{u}, \vec{v}$.
12. Довести, що $\big(\vec{u}, \vec{v}\big) = \frac{1}{4}\|\vec{u} + \vec{v}\|^2 - \frac{1}{4}\|\vec{u} - \vec{v}\|^2$ для довільних $\vec{u}, \vec{v}$.
13. *Визначником Грама* системи векторів $\big(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\big)$ евклідового векторного простору V називається детермінант

$$\begin{vmatrix} (\vec{a}_1, \vec{a}_1) & (\vec{a}_1, \vec{a}_2) & \ldots & (\vec{a}_1, \vec{a}_n) \\ (\vec{a}_2, \vec{a}_1) & (\vec{a}_2, \vec{a}_2) & \ldots & (\vec{a}_2, \vec{a}_n) \\ \vdots & \vdots & \ddots & \vdots \\ (\vec{a}_n, \vec{a}_1) & (\vec{a}_n, \vec{a}_2) & \ldots & (\vec{a}_n, \vec{a}_n) \end{vmatrix}.$$

Довести, що визначник Грама не змінюється в процесі ортогоналізації.

14. Довести, що визначник Грама будь-якої скінченної системи векторів евклідового векторного простору V:

    а) невід'ємний;

    б) дорівнює нулю тоді і тільки тоді, коли система лінійно залежна;

    в) не перевищує добутку квадратів довжин векторів системи;

    г) дорівнює добутку квадратів довжин векторів системи тоді і тільки тоді, коли вона ортогональна або один з них є нульовим вектором.

15. Довести, що в евклідовому векторному просторі V визначник Грама:

    а) двох векторів дорівнює квадрату площі паралелограма, побудованого на цих векторах;

    б) трьох векторів дорівнює квадрату об'єму паралелепіпеда, побудованого на цих векторах.

16. Довести, що ортогональне доповнення до підпростору евклідового векторного простору V має такі властивості:

    а) $(U+W)^\perp = U^\perp \cap W^\perp$;

    б) $(U\cap W)^\perp = U^\perp + W^\perp$.

17. (*Узагальнення теореми 3.25*) Нехай $(\vec{u}_1,\ldots,\vec{u}_k)$ — базис підпростору U лінійного простору V. Тоді проекція вектора $\vec{x}$ на U є такою лінійною комбінацією:

$$\mathrm{pr}_U\,\vec{x} = \lambda_1\vec{u}_1 + \ldots + \lambda_k\vec{u}_k,$$

    де коефіцієнти λ_i визначаються із системи

$$\begin{cases} (\vec{u}_1,\vec{u}_1)\lambda_1 & + & (\vec{u}_1,\vec{u}_2)\lambda_2 & + & \ldots & + & (\vec{u}_1,\vec{u}_k)\lambda_k & = & (\vec{u}_1,\vec{x}), \\ (\vec{u}_2,\vec{u}_1)\lambda_1 & + & (\vec{u}_2,\vec{u}_2)\lambda_2 & + & \ldots & + & (\vec{u}_2,\vec{u}_k)\lambda_k & = & (\vec{u}_2,\vec{x}), \\ \ldots \\ (\vec{u}_k,\vec{u}_1)\lambda_1 & + & (\vec{u}_k,\vec{u}_2)\lambda_2 & + & \ldots & + & (\vec{u}_k,\vec{u}_k)\lambda_k & = & (\vec{u}_k,\vec{x}). \end{cases}$$

18. Нехай $\mathcal{B} = \left(\vec{b}_1,\vec{b}_2,\ldots,\vec{b}_n\right)$ — ортонормований базис простору $\mathbb{R}^n$. Довести, що для довільних $\vec{x},\vec{y}\in\mathbb{R}^n$ має місце рівність (*рівність Парсеваля*[1]):

$$\left(\vec{x},\vec{y}\right) = \left(\vec{x},\vec{b}_1\right)\left(\vec{y},\vec{b}_1\right) + \left(\vec{x},\vec{b}_2\right)\left(\vec{y},\vec{b}_2\right) + \ldots + \left(\vec{x},\vec{b}_n\right)\left(\vec{y},\vec{b}_n\right).$$

19. Нехай $\left(\vec{b}_1,\vec{b}_2,\ldots,\vec{b}_k\right)$ — ортонормована система векторів простору $\mathbb{R}^n$. Довести, що для довільного $\vec{x}\in\mathbb{R}^n$ має місце нерівність (*нерівність Бесселя*[2]):

$$\|\vec{x}\|^2 \geqslant \left|\left(\vec{x},\vec{b}_1\right)\right|^2 + \left|\left(\vec{x},\vec{b}_2\right)\right|^2 + \ldots + \left|\left(\vec{x},\vec{b}_k\right)\right|^2.$$

    Довести також, що рівність виконується тоді і тільки тоді, коли $\vec{x}\in L\left(\vec{b}_1,\ldots,\vec{b}_k\right)$.

[1] Marc-Antoine Parseval des Chênes (1755-1836) — французький математик.

[2] Friedrich Wilhelm Bessel (1784–1846) — німецький математик і астроном.

Розділ 4

Лінійні відображення векторних просторів

Бросая в воду камешки, смотри на круги, ими образуемые; иначе такое бросание будет пустою забавою.

Козьма Прутков. Мысль №156.

§ 1. Відображення між множинами. Приклади лінійних відображень векторних просторів

Нехай X, Y — дві множини. Розглянемо відображення $F\colon X \to Y$.

Означення 4.1. Відображення $F\colon X \to Y$ називається *сюр'єктивним* (*відображенням* X ***на*** Y), якщо кожний елемент $y \in Y$ є образом хоча б одного $x \in X$.

Означення 4.2. Відображення $F\colon X \to Y$ називається *ін'єктивним* , якщо різні елементи множини X відображаються в різні елементи множини Y.

Означення 4.3. Відображення $F\colon X \to Y$ називається *бієктивним* (*взаємно однозначним відображенням* X ***на*** Y), якщо у кожного елемента $y \in Y$ існує прообраз, причому єдиний.

Очевидно, що відображення є бієктивним тоді і тільки тоді, коли воно одночасно є і сюр'єктивним, і ін'єктивним. Іноді ін'єктивне відображення називають взаємно однозначним відображенням X ***в*** Y (тоді як бієктивне — аналогічно з прийменником «***на***» замість «***в***»). Для уникнення плутанини в термінах ми надалі під взаємно однозначним відображенням розумітимемо виключно бієктивне. На рисунку

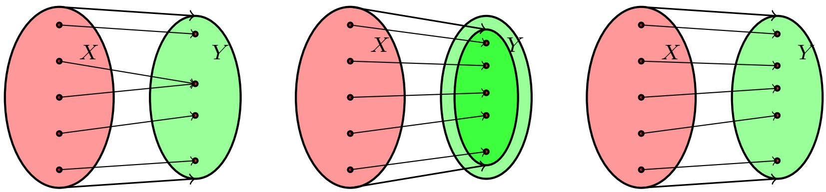


Рис. 4.1. Приклади сюр'єктивного (зліва), ін'єктивного (по центру) та бієктивного (справа) відображень

4.1 схематично показано приклади сюр'єктивного, ін'єктивного та бієктивного відображень.

1.1. Поняття ізоморфізму векторних просторів.

Означення 4.4. Нехай $(V, +, \cdot)$ і $(V', +, \cdot)$ — два векторні простори, які задані над одним і тим же полем P. Взаємно однозначне відображення $\varphi\colon V \to V'$ називається *ізоморфізмом* цих просторів (а самі простори називаються *ізоморфними*, що позначається $V \cong V'$), якщо для довільних $\vec{a}, \vec{b} \in V$, $\lambda \in P$ виконуються умови:

1) $\varphi(\vec{a} + \vec{b}) = \varphi(\vec{a}) + \varphi(\vec{b})$;
2) $\varphi(\lambda\vec{a}) = \lambda\varphi(\vec{a})$.

Серед властивостей ізоморфізмів векторних просторів відмітимо такі:

1°. При ізоморфізмі $\vec{0}$ простору V відображається в $\vec{0}$ простору V'.

Справді, для довільного $\vec{a} \in V$ $\vec{a} + \vec{0}_V = \vec{a}$ (тут ми позначили $\vec{0}$ простору V як $\vec{0}_V$). Тоді

$$\varphi(\vec{a} + \vec{0}_V) = \varphi(\vec{a}),$$

$$\varphi(\vec{a}) + \varphi(\vec{0}_V) = \varphi(\vec{a}),$$

звідки $\varphi(\vec{0}_V) = \vec{0}_{V'}$.

2°. При ізоморфізмі лінійно незалежна система векторів простору V переходить в лінійно незалежну систему векторів простору V'.

Доведення. Нехай $\vec{a}_1, \ldots, \vec{a}_k$ — лінійно незалежна система векторів простору V; $\varphi(\vec{a}_1), \ldots, \varphi(\vec{a}_k)$ — образи векторів цієї системи відносно ізоморфізму φ.

Припустимо, що система $\varphi(\vec{a}_1), \ldots, \varphi(\vec{a}_k)$ є лінійно залежною. Це означає, що рівність

$$\alpha_1\varphi(\vec{a}_1) + \ldots + \alpha_k\varphi(\vec{a}_k) = \vec{0}_{V'}$$

виконується не при всіх $\alpha_i = 0$. Але враховуючи означення ізоморфізму і властивість 1° останню рівність можна переписати у вигляді

$$\varphi(\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k) = \varphi(\vec{0}_V),$$

звідки знаходимо, що рівність $\alpha_1\vec{a}_1 + \ldots + \alpha_k\vec{a}_k = \vec{0}$ виконується не при всіх $\alpha_i = 0$, що суперечить лінійній незалежності системи векторів $\vec{a}_1, \ldots, \vec{a}_k$. ■

Наслідок. *При ізоморфізмі простору V на простір V' базис простору V переходить в базис простору V', а тому якщо простори V і V' ізоморфні, то вони мають однакову розмірність.*

Теорема 4.1. *Нехай $(V, +, \cdot)$ — n-вимірний векторний простір, заданий над полем $\mathbb{R}$. Тоді простір V ізоморфний простору $\mathbb{R}^n$.*

Доведення. Нехай $\vec{e}_1, \ldots, \vec{e}_n$ — деякий базис векторного простору V. В цьому базисі кожному вектору цього простору відповідає його координатний стовпець, який є вектором простору $\mathbb{R}^n$. Тобто ми маємо відповідність

$$\varphi(\vec{a}) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix},$$

якщо $\vec{a} = \alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n$ (див. зауваження на стор. 92). Це відображення задовольняє означення ізоморфізму. ■

Наслідок. *Будь-які два n-вимірні векторні простори, задані над полем $\mathbb{R}$, ізоморфні між собою, оскільки кожний з них ізоморфний з n-вимірним арифметичним векторним простором $\mathbb{R}^n$.*

Поняття ізоморфізму двох векторних просторів в певній мірі означає, що ці простори мають однакову структуру. В алгебрі структури вивчаються з точністю до ізоморфізму, тобто ізоморфні об'єкти фактично ототожнюються. Саме тому усі властивості, вивчені нами для арифметичного векторного простору $\mathbb{R}^n$, «переносяться» на випадок довільного n-вимірного векторного простору.

1.2. Лінійні відображення векторних просторів.

В будь-якій алгебраїчній теорії разом з поняттям ізоморфізму, яке означається в подібний спосіб для різноманітних алгебраїчних структур, вивчаються і більш загальні відображення, що в загальному випадку називаються *гомоморфізмами*, а у випадку векторних просторів — лінійними відображеннями.

Означення 4.5. Нехай V і W — векторні простори, задані над полем P. Відображення $T\colon V \to W$ називається *лінійним відображенням*, якщо для довільних $\vec{a}$, $\vec{b} \in V$, $\lambda \in P$ виконуються умови:
1) $T(\vec{a}+\vec{b}) = T(\vec{a}) + T(\vec{b})$;
2) $T(\lambda \cdot \vec{a}) = \lambda \cdot T(\vec{a})$.

Як бачимо, означення лінійного відображення просторів відрізняється від означення ізоморфізму просторів лише тим, що не вимагається бієктивності вказаного відображення.

Якщо відображення T задовольняє умови 1 та 2 з означення 4.5, то говорять, що воно зберігає операції додавання векторів і множення на їх на число.

Приклад 4.1. Нехай відображення $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ визначається правилом:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \end{bmatrix}$$

(таке відображення називається *проекцією*). Воно є лінійним, оскільки для $\vec{x}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ та $\vec{x}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$:

$$T(\vec{x}_1 + \vec{x}_2) = T\left(\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{bmatrix}\right) =$$

$$= \begin{bmatrix} x_1+x_2 \\ y_1+y_2 \end{bmatrix} = T(\vec{x}_1) + T(\vec{x}_2);$$

$$T(c\cdot\vec{x}_1) = T\left(\begin{bmatrix} cx_1 \\ cy_1 \\ cz_1 \end{bmatrix}\right) = \begin{bmatrix} cx_1 \\ cy_1 \end{bmatrix} = c\cdot T(\vec{x}_1).$$

Зауважимо, що вказане відображення T не є ізоморфізмом, оскільки воно не взаємно однозначне.

Приклад 4.2. Нехай A — фіксована матриця розмірності $m \times n$. Відображення $T\colon \mathbb{R}^n \to \mathbb{R}^m$ означимо таким чином: якщо $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, то $T(\vec{x}) = A \cdot \vec{x}$. Таке відображення назвемо *матричним*.

Вказане відображення є лінійним, оскільки для довільних $\vec{x}_1, \vec{x}_2$, $c \in \mathbb{R}$:

$$T(\vec{x}_1 + \vec{x}_2) = A \cdot (\vec{x}_1 + \vec{x}_2) = A \cdot \vec{x}_1 + A \cdot \vec{x}_2 = T(\vec{x}_1) + T(\vec{x}_2);$$
$$T(c \cdot \vec{x}_1) = A \cdot (c \cdot \vec{x}_1) = c \cdot (A \cdot \vec{x}_1) = c \cdot T(\vec{x}_1).$$

Приклад 4.3. Відображення $T\colon V \to W$ означимо так: якщо $\vec{x} \in V$, то $T(\vec{x}) = \vec{0}_W$. Таке відображення, очевидно, задовольняє означення лінійного відображення. Воно називається *нульовим*.

Приклад 4.4. Нехай $C^1_{[a,b]}$ — простір неперервно диференційовних на відрізку $[a, b]$ функцій, $C_{[a,b]}$ — простір неперервних на $[a, b]$ функцій. Відображення $D\colon C^1_{[a,b]} \to C_{[a,b]}$ означимо так: якщо $f \in C^1_{[a,b]}$, то $D(f) = f'$.

Як відомо з аналізу, для кожних $f, g \in C^1_{[a,b]}$, $c \in \mathbb{R}$ виконуються рівності $D(f + g) = (f + g)' = f' + g' = D(f) + D(g)$ та $D(c \cdot f) = c \cdot D(f)$. Таким чином, диференціювання є лінійним відображенням між відповідними векторними просторами.

Приклад 4.5. Відображення $S\colon C_{[a,b]} \to \mathbb{R}$ визначимо рівністю $S(f) = \int\limits_a^b f(x)\,dx$ для довільного елемента (функції) f простору $C_{[a,b]}$. Означення відображення коректне, оскільки, як відомо з аналізу, кожна неперервна на відрізку функція є інтегровною на ньому.

Оскільки

$$S(f + g) = \int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx = S(f) + S(g)$$

та

$$S(c \cdot f) = \int\limits_a^b (c \cdot f(x))\,dx = c\int_a^b f(x)\,dx = c \cdot S(f),$$

то вказане відображення S є лінійним.

Приклад 4.6. Довести, що відображення

1) $F_1 \colon M_{22} \to \mathbb{R}$ таке, що $F_1(A) = \det A$;

2) $F_2 \colon \mathbb{R}^1 \to \mathbb{R}^1$ таке, що $F_2(x) = 2^x$;

3) $F_3 \colon \mathbb{R}^1 \to \mathbb{R}^1$ таке, що $F_3(x) = x - 1$

не є лінійними відображеннями відповідних векторних просторів.

Розв'язання. Для розв'язання вказаної задачі слід вказати контрприклади, які б вказували на те, що якась із умов в означенні лінійного відображення порушується.

1) Відомо, що, взагалі кажучи, $\det(A+B) \neq \det A + \det B$. Справді, для доведення цього достатньо покласти $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ (тоді $\det A = \det B = 0$, $\det(A+B) = 1$, і $\det(A+B) \neq \det A + \det B$). Отже, перша умова з означення лінійного відображення векторних просторів не виконується і F_1 не є лінійним відображенням.

2) Нехай $x_1 = 1$, $x_2 = 2$. Тоді, з одного боку:

$$F_2(x_1 + x_2) = F_2(1+2) = F_2(3) = 2^3 = 8,$$

а з іншого:

$$F_2(x_1) + F_2(x_2) = 2^1 + 2^2 = 6,$$

тобто нами вказано такі $x_1, x_2 \in \mathbb{R}$, що $F_2(x_1 + x_2) \neq F_2(x_1) + F_2(x_2)$. Отже, F_2 не є лінійним відображенням.

3) Маємо, наприклад: $F_3(2) = 2 - 1 = 1$ і $2 \cdot F_3(1) = 2 \cdot 0 = 0$. Отже $F_3(2 \cdot 1) \neq 2 \cdot F_3(1)$ і не виконується друга умова лінійного відображення. ■

Зауваження. Останній приклад показує, що слід з обережністю вживати термін «лінійне відображення», коли мова йде про відображення $\mathbb{R} \to \mathbb{R}$. Функція $F_3(x) = x - 1$ називається лінійною. Однак вона не є лінійним відображенням векторного простору $\mathbb{R}$ в себе, оскільки не задовольняє відповідне означення.

1.3. Властивості лінійних відображень векторних просторів.

Нехай V і W — два векторні простори, що задані над деяким полем P.

Теорема 4.2. *Нехай $T \colon V \to W$ — лінійне відображення. Тоді:*

1) $T(\vec{0}_V) = \vec{0}_W$;

2) для довільного $\vec{x} \in V$ $T(-\vec{x}) = -T(\vec{x})$;

3) для довільних $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in V$ *і* $\alpha_1, \alpha_2, \ldots, \alpha_k \in P$*:*

$$T(\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \ldots + \alpha_k\vec{x}_k) = \alpha_1 T(\vec{x}_1) + \alpha_2 T(\vec{x}_2) + \ldots + \alpha_k T(\vec{x}_k).$$

Доведення. 1. Нехай $\vec{x}$ — довільний вектор простору V. Тоді $T(\vec{0}) = T(0 \cdot \vec{x}) = 0 \cdot T(\vec{x}) = \vec{0}$.

2. Маємо: $T(-\vec{x}) = T((-1) \cdot \vec{x}) = -1 \cdot T(\vec{x}) = -T(\vec{x})$.
3. Доведемо за індукцією. Те, що $T(\alpha_1\vec{x}_1) = \alpha_1 T(\vec{x}_1)$ випливає безпосередньо з того, що T є лінійним відображенням. Припустимо, що для довільних $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in V$ і $\alpha_1, \alpha_2, \ldots, \alpha_k \in P$ має місце

$$T(\alpha_1\vec{x}_1 + \ldots + \alpha_k\vec{x}_k) = \alpha_1 T(\vec{x}_1) + \ldots + \alpha_k T(\vec{x}_k).$$

Тоді для довільних $\vec{x}_{k+1} \in V$, $\alpha_{k+1} \in P$:

$$\begin{aligned}T(\alpha_1\vec{x}_1 + \ldots + \alpha_k\vec{x}_k + \alpha_k\vec{x}_{k+1}) &= T(\alpha_1\vec{x}_1 + \ldots + \alpha_k\vec{x}_k) + T(\alpha_k\vec{x}_{k+1}) = \\ &= \alpha_1 T(\vec{x}_1) + \ldots + \alpha_k T(\vec{x}_k) + \alpha_{k+1} T_{k+1},\end{aligned}$$

що і потрібно було довести.

■

Зауваження. З п. 1 теореми 4.2 безпосередньо випливає, що відображення F_2 і F_3 з прикладу 5 не є лінійними.

Найважливішою властивістю лінійного відображення $T\colon V \to W$ є та, що T повністю визначається своєю дією на базис простору V.

Теорема 4.3. *Нехай* $T_1, T_2\colon V \to W$ *— два лінійні відображення. Нехай також* $\big(\vec{e}_1, \vec{e}_2 \ldots, \vec{e}_n\big)$ *— базис простору* V*. Якщо для кожного* $i \in \{1, \ldots, n\}$ $T_1(\vec{e}_i) = T_2(\vec{e}_i)$*, то* $T_1 \equiv T_2$*.*

Доведення. Нехай $\vec{a} \in V$. Оскільки $\big(\vec{e}_1, \vec{e}_2 \ldots, \vec{e}_n\big)$ — базис, то існують $\alpha_1, \alpha_2, \ldots, \alpha_n$ такі, що $\vec{a} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n$.

Тоді

$$\begin{aligned}T_1(\vec{a}) &= T_1(\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n) = \\ &= \alpha_1 T_1(\vec{e}_1) + \alpha_2 T_1(\vec{e}_2) + \ldots + \alpha_n T_1(\vec{e}_n) = \\ &= \alpha_1 T_2(\vec{e}_1) + \alpha_1 T_2(\vec{e}_1) + \ldots + \alpha_n T_2(\vec{e}_n) = T_2(\vec{a}).\end{aligned}$$

Отже, для довільного $\vec{a} \in V$ $T_1(\vec{a}) = T_2(\vec{a})$, а це і означає, що $T_1 \equiv T_2$. ■

Теорема 4.4. *Нехай* V *і* W *— векторні простори і нехай* $\big(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\big)$ *— базис простору* V*. Нехай також* $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n$ *— довільний набір векторів простору* W *(не обов'язково різних). Тоді* *існує, причому єдине, лінійне відображення* $T\colon V \to W$ *таке, що* $T(\vec{e}_i) = \vec{w}_i$ *для кожного* $i = 1, 2, \ldots, n$.

Доведення. Якщо таке відображення існує, то воно за теоремою 4.3 єдине. Отже, залишається показати, що воно справді існує.

Нехай $\vec{x} \in V$. Оскільки $\big(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\big)$ — базис, то існує, причому єдиний, набір скалярів $\alpha_1, \alpha_2, \ldots, \alpha_n$ такий, що $\vec{x} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n$. Тоді нехай

$$T(\vec{x}) = T(\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n) = \alpha_1 T(\vec{e}_1) + \ldots + \alpha_n f(\vec{e}_n) = \\ = \alpha_1\vec{w}_1 + \ldots + \alpha_n\vec{w}_n.$$

Те, що побудоване відображення є лінійним, легко перевіряється. ■

В силу ізоморфізму скінченновимірних векторних просторів однакової розмірності особливу увагу звернемо на вивчення лінійних відображень $\mathbb{R}^n \to \mathbb{R}^m$.

З прикладу 4.2 випливає, що кожна матриця A розмірності $m \times n$ породжує лінійне відображення $T\colon \mathbb{R}^n \to \mathbb{R}^m$ за правилом $T(\vec{x}) = A \cdot \vec{x}$ для всіх векторів-стовпців $\vec{x}$ простору $\mathbb{R}^n$. Наступна теорема встановлює, що *усі* лінійні відображення $\mathbb{R}^n \to \mathbb{R}^m$ можна задати таким способом.

Теорема 4.5. *Нехай* $T\colon \mathbb{R}^n \to \mathbb{R}^m$. *Будемо записувати елементи простору* $\mathbb{R}^n$ *як стовпці.*

1. *Існує* $m \times n$ *матриця* A *така, що* $T(\vec{x}) = A \cdot \vec{x}$ *для всіх* $x \in \mathbb{R}^n$.

2. *Стовпці матриці* A *— відповідно* $T(\vec{e}_1), \ldots, T(\vec{e}_n)$, *де* $\big(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\big)$ *— стандартний базис простору* $\mathbb{R}^n$:

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \ldots & T(\vec{e}_n) \end{bmatrix}.$$

Доведення. Нехай $\big(\vec{e}_1, \vec{e}_2, \ldots \vec{e}_n\big)$ — стандартний базис простору $\mathbb{R}^n$ і нехай

$$T(\vec{e}_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(\vec{e}_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots, \quad T(\vec{e}_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Нехай $A = [a_{ij}] = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \dots & T(\vec{e}_n) \end{bmatrix}$. Нехай $\vec{x} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$, $\alpha_i \in \mathbb{R}$. Тоді

$$\begin{aligned}
T(\vec{x}) &= T(\alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \ldots + \alpha_n\vec{e}_n) = \\
&= \alpha_1 T(\vec{e}_1) + \alpha_2 T(\vec{e}_2) + \ldots + \alpha_n T(\vec{e}_n) = \\
&= \alpha_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \alpha_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + \alpha_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \\
&= \begin{bmatrix} a_{11}\alpha_1 + a_{12}\alpha_2 + \ldots + a_{1n}\alpha_n \\ a_{21}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{2n}\alpha_n \\ \vdots \\ a_{m1}\alpha_1 + a_{m2}\alpha_2 + \ldots + a_{mn}\alpha_n \end{bmatrix} = \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = A\vec{x}.
\end{aligned}$$

■

Матриця A з теореми 4.5 називається *стандартною матрицею лінійного відображення* T.

Приклад 4.7. Знайти стандартну матрицю лінійного відображення $T\colon \mathbb{R}^3 \to \mathbb{R}^2$, заданого за правилом:

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x - 2y + z \\ x - z \end{bmatrix}.$$

Розв'язання. З одного боку, очевидно, що

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

тобто матриця $A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{bmatrix}$ є стандартною матрицею заданого відображення.

З іншого боку, за п. 2 теореми 4.5 знаходимо:

$$T(\vec{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad T(\vec{e}_2) = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \qquad T(\vec{e}_3) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

■

1.4. Матриця лінійного відображення відносно різних базисів.

Нехай V — n-вимірний векторний простір, W — m-вимірний векторний простір. Нехай також $T\colon V \to W$ — лінійне відображення.

Нехай $\mathcal{B} = \left(\vec{b}_1, \vec{b}_2, \ldots \vec{b}_n\right)$ — базис простору V; $\mathcal{C} = \left(\vec{c}_1, \vec{c}_2, \ldots \vec{c}_n\right)$ — базис простору W.

Тоді кожний вектор $\vec{x} \in V$ можна представити у вигляді лінійної комбінації векторів базису $\mathcal{B}$:

$$\vec{x} = \alpha_1 \vec{b}_1 + \cdots + \alpha_n \vec{b}_n.$$

Координати вектора $\vec{x}$ в базисі $\mathcal{B}$ ми будемо позначати через $[\vec{x}]_{\mathcal{B}}$:

$$[x]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}_{\mathcal{B}}.$$

Маємо:

$$T(\vec{x}) = T(\alpha_1 \vec{b}_1 + \cdots + \alpha_n \vec{b}_n) = \alpha_1 T(\vec{b}_1) + \cdots + \alpha_n T(\vec{b}_n),$$

оскільки T — лінійне. Перепишемо останню рівність у координатах відповідних векторів в базисі $\mathcal{C}$. Одержимо:

$$[T(\vec{x})]_{\mathcal{C}} = \alpha_1 \left[T(\vec{b}_1)\right]_{\mathcal{C}} + \cdots + \alpha_n \left[T(\vec{b}_n)\right]_{\mathcal{C}}.$$

Цю рівність можна переписати у матричному вигляді

$$[T(\vec{x})]_{\mathcal{C}} = M \cdot [\vec{x}]_{\mathcal{B}}, \tag{4.1}$$

де

$$M = \Big[\big[T(\vec{b}_1)\big]_{\mathcal{C}} \quad \big[T(\vec{b}_2)\big]_{\mathcal{C}} \quad \dots \quad \big[T(\vec{b}_n)\big]_{\mathcal{C}}\Big] \tag{4.2}$$

Матриця M називається *матрицею лінійного відображення T відносно базисів $\mathcal{B}$ та $\mathcal{C}$*. Те, що матриця M є матрицею лінійного відображення T відносно базисів $\mathcal{B}$ та $\mathcal{C}$ ми будемо позначати так:

$$M = [T]_{\mathcal{C}\leftarrow\mathcal{B}}\,.$$

Стандартну ж матрицю лінійного відображення T ми також позначатимемо так: $[T]$.

Отже, нами побудовано перехід при вивченні лінійних відображень n-вимірного простору V в m-вимірний простір W до лінійних відображень $\mathbb{R}^n \to \mathbb{R}^m$ (рис. 4.2). Якщо потрібно визначити, яким є образ вектора $\vec{x}$ при лінійному відображенні T в заданому базисі $\mathcal{C}$ простору W потрібно знайти координати вектора $\vec{x}$ в базисі $\mathcal{B}$ (елемент простору $\mathbb{R}^n$) і помножити зліва на матрицю лінійного відображення T відносно базисів $\mathcal{B}$ та $\mathcal{C}$. Схематично це зображено на рис. 4.2.

Теорема 4.6. *Вибір базисів в n-вимірному просторі V та m-вимірному просторі W встановлює взаємно однозначну відповідність між лінійними відображеннями $V \to W$ та матрицями розмірності $m \times n$.*

Доведення. Справді, кожному лінійному відображенню $T\colon V \to W$ при вибраних базисах можна поставити у відповідність його матрицю відносно цих базисів. З іншого боку, кожна матриця $A = \big[\vec{a}_1 \quad \dots \quad \vec{a}_n\big]$ розмірів $m \times n$ породжує наступне лінійне відображення: якщо $\vec{b}_i$ — i-й вектор базису $\mathcal{B}$ простору V, то $[T(\vec{b}_i)]_{\mathcal{C}} = \vec{a}_i$. Існування та єдиність такого відображення доведено в теоремі 4.4. ■

Приклад 4.8. Нехай $\mathcal{B} = \big(\vec{b}_1, \vec{b}_2\big)$ — базис векторного простору V, $\mathcal{C} = \big(\vec{c}_1, \vec{c}_2, \vec{c}_3\big)$ — базис векторного простору W; $T\colon V \to W$ — лінійне відображення таке, що $T(\vec{b}_1) = 3\vec{c}_1 - 2\vec{c}_2 - 4\vec{c}_3$, $T(\vec{b}_2) = \vec{c}_1 + 5\vec{c}_2 - \vec{c}_3$. Знайти матрицю M лінійного відображення T відносно базисів $\mathcal{B}$ та $\mathcal{C}$.

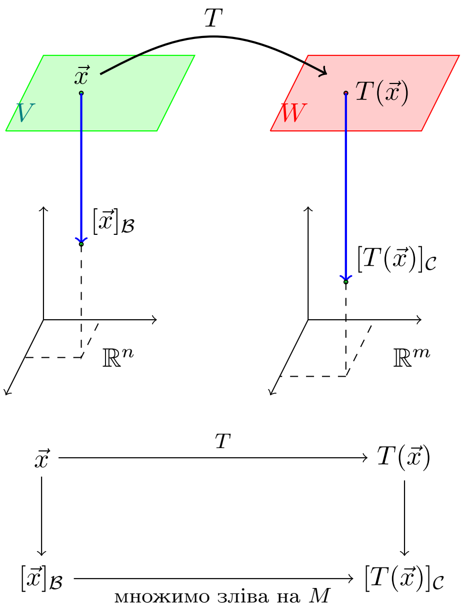


Рис. 4.2. Перехід від лінійного відображення $T\colon V \to W$ до відображення $\mathbb{R}^n \to \mathbb{R}^m$ при вибраних базисах $\mathcal{B}$ та $\mathcal{C}$ цих просторів.

Розв'язання. Із умови задачі безпосередньо випливає, що

$$\Big[T(\vec{b}_1)\Big]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ -4 \end{bmatrix}, \qquad \Big[T(\vec{b}_2)\Big]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 5 \\ -1 \end{bmatrix},$$

і, отже,

$$M = \begin{bmatrix} 3 & -1 \\ -2 & 5 \\ -4 & -1 \end{bmatrix}.$$

■

§ 2. Операції над лінійними відображеннями векторних просторів

В цьому розділі ми введемо в стандартний спосіб операції над лінійними відображеннями векторних просторів: додавання, множення на число та композицію. Оскільки між лінійними відображеннями і матрицями встановлюється взаємно однозначна відповідність, то операціям над лінійними відображеннями відповідатимуть певні операції над матрицями. При цьому виявиться, зокрема, що матриця композиції двох відображень є добутком матриць цих відображень. Це не випадково: історично означення добутку матриць A та B вводилось як матриця композиції двох лінійних відображень з матрицями A та B (в певному порядку).

2.1. Сума відображень та добуток відображення на скаляр.

Нехай V і W — два векторні простори, задані над полем P, і нехай $T_1 \colon V \to W$, $T_2 \colon V \to W$ — лінійні відображення.

Означення 4.6. *Сумою* лінійних відображень T_1 і T_2 називається відображення, що позначається $T_1 + T_2$, таке, що $(T_1 + T_2)(\vec{x}) = T_1(\vec{x}) + T_2(\vec{x})$ для кожного $\vec{x} \in V$.

Легко переконатись, що відображення $T_1 + T_2$ є лінійним. Справді, нехай $\vec{x}_1, \vec{x}_2$ — довільні вектори простору V, $\lambda \in P$ тоді:

$$\begin{aligned}
(T_1 + T_2)(\vec{x}_1 + \vec{x}_2) &\overset{df}{=} T_1(\vec{x}_1 + \vec{x}_2) + T_2(\vec{x}_1 + \vec{x}_2) = \\
&= T_1(\vec{x}_1) + T_1(\vec{x}_2) + T_2(\vec{x}_1) + T_2(\vec{x}_2) = \\
&= (T_1 + T_2)(\vec{x}_1) + (T_1 + T_2)(\vec{x}_2); \\
(T_1 + T_2)(\lambda\vec{x}_1) &\overset{df}{=} T_1(\lambda\vec{x}_1) + T_2(\lambda\vec{x}_1) = \\
&= \lambda T_1(\vec{x}_1) + \lambda T_2(\vec{x}_1) = \lambda(T_1 + T_2)(\vec{x}_1).
\end{aligned}$$

Якщо M_1 — матриця лінійного відображення T_1 відносно базисів $\mathcal{B}$ і $\mathcal{C}$, M_2 — матриця лінійного відображення T_2 відносно цих самих базисів, то легко переконатись в тому, що матриця відображення $T_1 + T_2$ відносно вказаних базисів має вигляд $M_1 + M_2$. Справді,

маємо:
$$(\forall \vec{x} \in V):\ [T_1(\vec{x})]_{\mathcal{C}} = M_1 \cdot [\vec{x}]_{\mathcal{B}}\,;$$
$$[T_2(\vec{x})]_{\mathcal{C}} = M_2 \cdot [\vec{x}]_{\mathcal{B}}\,.$$
Тоді
$$\begin{aligned}[(T_1+T_2)(\vec{x})]_{\mathcal{C}} =& [T_1(\vec{x})+T_2(\vec{x})]_{\mathcal{C}} = [T_1(\vec{x})]_{\mathcal{C}} + [T_2(\vec{x})]_{\mathcal{C}} = \\ =& M_1 \cdot [\vec{x}]_{\mathcal{B}} + M_2 \cdot [\vec{x}]_{\mathcal{B}} = (M_1+M_2)\cdot [\vec{x}]_{\mathcal{B}}\,,\end{aligned}$$
що випливає з властивостей множення матриць. Зокрема, якщо A_1 — стандартна матриця лінійного відображення T_1, A_2 — стандартна матриця лінійного відображення T_2, то A_1+A_2 — стандартна матриця лінійного відображення T_1+T_2.

Означення 4.7. *Добутком* лінійного відображення T на число (скаляр) $\lambda \in P$ називається відображення, що позначається λT, таке, що $(\lambda T)(\vec{x}) = \lambda T(\vec{x})$ для довільного $\vec{x} \in V$.

Для того, щоб переконатись, що відображення λT є лінійним перевіримо виконання двох умов означення лінійного відображення, а саме, для довільних $\vec{x}_1, \vec{x}_2 \in V$, $c \in P$:
$$\begin{aligned}(\lambda T)(\vec{x}_1+\vec{x}_2) \overset{df}{=}& \lambda \cdot T(\vec{x}_1+\vec{x}_2) = \lambda \cdot T(\vec{x}_1) + \lambda \cdot T(\vec{x}_2) = \\ =& (\lambda T)(\vec{x}_1) + (\lambda T)(\vec{x}_2);\\ (\lambda T)(c\cdot\vec{x}_1) \overset{df}{=}& \lambda \cdot T(c\cdot\vec{x}_1) = \lambda\cdot c\cdot T(\vec{x}_1) = c\lambda\cdot T(\vec{x}_1) = \\ =& c\cdot(\lambda T)(\vec{x}_1).\end{aligned}$$

Якщо лінійне відображення T задано матрицею M відносно базисів $\mathcal{B}$ та $\mathcal{C}$, то матрицею відображення λT відносно цих самих базисів є $\lambda \cdot M$. Справді,
$$\forall \vec{x} \in V:\ [T(\vec{x})]_{\mathcal{C}} = M \cdot [\vec{x}]_{\mathcal{B}}\,.$$
Тоді
$$[(\lambda T)(\vec{x})]_{\mathcal{C}} = \lambda \cdot [T(\vec{x})]_{\mathcal{C}} = (\lambda M)\,[\vec{x}]_{\mathcal{B}}\,.$$

Зокрема, якщо A — стандартна матриця лінійного відображення T, то λA — стандартна матриця відображення λT.

2.2. Композиція лінійних відображень.

Означення 4.8. Нехай U, V, W — векторні простори задані над деяким полем, $T\colon U \to V$, $S\colon V \to W$ — лінійні відображення. *Композицією (добутком)* відображень T та S називається таке відображення $S \circ T\colon U \to W$, що для довільного $\vec{x} \in U$: $(S \circ T)(\vec{x}) = S(T(\vec{x}))$.

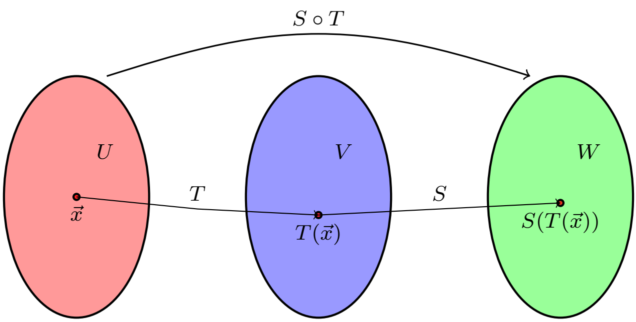


Рис. 4.3. Композиція відображень

Приклад 4.9. Нехай $T\colon \mathbb{R}^2 \to \mathcal{P}_1$, $S\colon \mathcal{P}_1 \to \mathcal{P}_2$ такі, що:

$$T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = a + (a+b)x; \qquad S(p(x)) = x \cdot p(x).$$

Знайти $(S \circ T)\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right)$, $(S \circ T)\left(\begin{bmatrix} a \\ a \end{bmatrix}\right)$.

Розв'язання. Маємо:

$$(S \circ T)\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) = S\left(T\left(\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right)\right) = S(3+x) = x(3+x) = x^2 + 3x.$$

$$(S \circ T)\left(\begin{bmatrix} a \\ a \end{bmatrix}\right) = S\left(T\left(\begin{bmatrix} a \\ a \end{bmatrix}\right)\right) = S(a+2ax) = x(a+2ax) = 2ax^2 + ax.$$

■

Доведемо, що $S \circ T$ — лінійне відображення. Справді,

$$\begin{aligned}(S \circ T)(\vec{x}_1 + \vec{x}_2) &= S(T(\vec{x}_1 + \vec{x}_2)) = S(T(\vec{x}_1) + T(\vec{x}_2)) = \\ &= S(T(\vec{x}_1)) + S(T(\vec{x}_2)) = (S \circ T)(\vec{x}_1) + (S \circ T)(\vec{x}_2); \\ (S \circ T)(\lambda \vec{x}_1) &= S(T(\lambda \vec{x}_1)) = S(\lambda T(\lambda \vec{x}_1)) = \\ &= \lambda S(T(\lambda \vec{x}_1)) = \lambda (S \circ T)(\vec{x}_1).\end{aligned}$$

Теорема 4.7. *Нехай U, V та W — скінченновимірні векторні простори з базисами $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ відповідно. Нехай $T\colon U \to V$, $S\colon V \to W$ — лінійні відображення. Тоді:*

$$[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}} [T]_{\mathcal{C}\leftarrow\mathcal{B}} .$$

Зауваження. Теорема стверджує, що фактично матриця композиції двох відображень дорівнює добутку матриць цих відображень.

Доведення. Доведемо, що відповідні стовпці матриць $[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}}$ та $[S]_{\mathcal{D}\leftarrow\mathcal{C}} [T]_{\mathcal{C}\leftarrow\mathcal{B}}$ рівні. З цією метою розглянемо образ i-го базисного вектора $\vec{b}_i$ базиса $\mathcal{B}$ при відображенні $S \circ T$ (i-ий стовпець матриці $[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}}$):

$$\left[(S \circ T)(\vec{b}_i)\right]_{\mathcal{D}} = \left[S(T(\vec{b}_i))\right]_{\mathcal{D}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}} \cdot \left[T(\vec{b}_i)\right]_{\mathcal{C}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}} \cdot [T]_{\mathcal{C}\leftarrow\mathcal{B}} \left[\vec{b}_i\right]_{\mathcal{B}}$$

(тут ми двічі застосували формулу (4.1) — зв'язок між координатами вектора та його образу в різних базисах). Але $\left[\vec{b}_i\right]_{\mathcal{B}} = \vec{e}_i = \begin{bmatrix}0 & \dots & 0 & 1 & 0 & \dots & 0\end{bmatrix}^T$, тому

$$[S]_{\mathcal{D}\leftarrow\mathcal{C}} \cdot [T]_{\mathcal{C}\leftarrow\mathcal{B}} \left[\vec{b}_i\right]_{\mathcal{B}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}} \cdot [T]_{\mathcal{C}\leftarrow\mathcal{B}} \vec{e}_i,$$

що, в свою чергу, є i-им стовпцем матриці $[S]_{\mathcal{D}\leftarrow\mathcal{C}} \cdot [T]_{\mathcal{C}\leftarrow\mathcal{B}}$, що і потрібно було довести.

■

Наслідок. *Якщо A — стандартна матриця лінійного відображення $T\colon U \to V$, B — стандартна матриця лінійного відображення $S\colon V \to W$, то BA є стандартною матрицею відображення $S \circ T$.*

Приклад 4.10. Знайти стандартну матрицю відображення $F\colon \mathbb{R}^2 \to \mathbb{R}^2$, яке спочатку повертає точку площини на 90° проти годинникової стрілки, а потім одержану точку симетрично відображає відносно осі Ox.

Розв'язання. *Спосіб 1.* Розглянемо геометричну інтерпретацію простору $\mathbb{R}^2$ як множину векторів площини з початком в точці O.

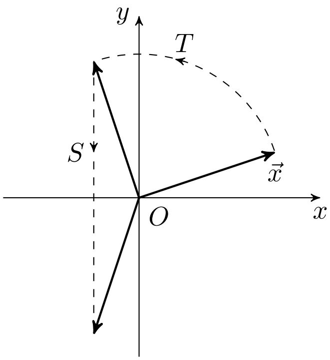


Рис. 4.4.

Знайдемо образи векторів стандартного базису $\vec{e}_1$ та $\vec{e}_2$. Вектор $\vec{e}_1 = \begin{bmatrix}1\\0\end{bmatrix}$ при повороті спочатку переходить у вектор $\begin{bmatrix}0\\1\end{bmatrix}$, а потім при симетрії відносно осі Ox — у вектор $\begin{bmatrix}0\\-1\end{bmatrix}$. Таким чином, $F(\vec{e}_1) = \begin{bmatrix}0\\-1\end{bmatrix}$. Аналогічно $F(\vec{e}_2) = \begin{bmatrix}-1\\0\end{bmatrix}$. Тоді шукана матриця $A = \begin{bmatrix}F(\vec{e}_1) & F(\vec{e}_2)\end{bmatrix} = \begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}$.

Спосіб 2. Помітимо, що задане відображення F є композицією відображень T повороту на кут 90° і S симетрії відносно осі Ox (рис. 4.4). Стандартні матриці перетворень T і S мають вигляд $A_1 = \begin{bmatrix}T(\vec{e}_1) & T(\vec{e}_2)\end{bmatrix} = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$ та $A_2 = \begin{bmatrix}S(\vec{e}_1) & S(\vec{e}_2)\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ відповідно. Тоді за наслідком теореми 4.7 знаходимо стандартну матрицю відображення $F\colon S \circ T$:

$$A = A_2 \cdot A_1 = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix} \cdot \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}0 & -1\\-1 & 0\end{bmatrix}.$$

■

Теорема 4.8. *Нехай V, W — векторні простори, задані над полем P. Позначимо через $\mathcal{L}(V,W)$ множину усіх лінійних відображень $V \to W$. Тоді $\mathcal{L}(V,W)$ відносно операцій суми відображень і їх добутку на число є векторним простором.*

Доведення. Очевидно, що множина $\mathcal{L}(V,W)$ замкнена відносно операцій суми відображень і добутку на число. Тому для доведення теореми потрібно показати, що для множини $\mathcal{L}(V,W)$ виконуються 8 умов з означення векторного простору, в чому легко переконатись безпосередньою перевіркою (нульовим елементом цього простору є нульове відображення). ■

§ 3. Ядро і образ лінійного відображення векторних просторів

Нехай V і W — векторні простори, задані над деяким полем.

Означення 4.9. *Ядром* лінійного відображення $T\colon V \to W$ (позначається $\operatorname{Ker} T$) називається множина усіх векторів простору V, образами яких є $\vec{0}$ простору W:

$$\operatorname{Ker} T = \left\{ \vec{x} \in V : T(\vec{x}) = \vec{0}_W \right\}.$$

Означення 4.10. *Образом (областю значень)* лінійного відображення $T\colon V \to W$ (позначається $\operatorname{Im} T$) називається множина усіх векторів простору W, що є образами векторів простору V при відображенні T:

$$\operatorname{Im} T = \{T(\vec{x}) \colon \vec{x} \in V\}.$$

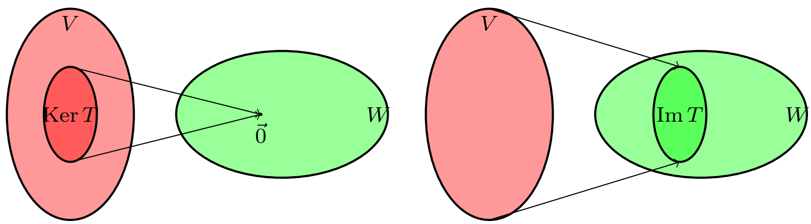


Рис. 4.5. Ядро і образ лінійного відображення $T\colon V \to W$.

Теорема 4.9. *Ядро лінійного відображення $T\colon V \to W$ — підпростір простору V, а образ — підпростір простору W.*

Доведення. Оскільки $T(\vec{0}_V) = \vec{0}_W$ (див. п. 1 теореми 4.2), то множини $\operatorname{Ker} T$ та $\operatorname{Im} T$ непорожні і містять $\vec{0}_V$ та $\vec{0}_W$ відповідно. Для доведення теореми достатньо показати, що множини $\operatorname{Ker} T$ та $\operatorname{Im} T$ замкнені відносно операцій додавання векторів і множення на число.

Нехай $\vec{x}_1, \vec{x}_2 \in \operatorname{Ker} T$. Тоді $T(\vec{x}_1) = \vec{0}$, $T(\vec{x}_2) = \vec{0}$. В силу лінійності T:

$$T(\vec{x}_1 + \vec{x}_2) = T(\vec{x}_1) + T(\vec{x}_2) = \vec{0} + \vec{0} = \vec{0};$$

$$T(\lambda\vec{x}_1) = \lambda T(\vec{x}_1) = \lambda \cdot \vec{0} = \vec{0},$$

і тому $\operatorname{Ker} T$ є підпростором простору V.

Нехай тепер $T(\vec{x}_1)$ і $T(\vec{x}_2)$ належать $\operatorname{Im} T$. Тоді $T(\vec{x}_1) + T(\vec{x}_2) = T(\vec{x}_1 + \vec{x}_2)$ і, отже, $T(\vec{x}_1) + T(\vec{x}_2)$ є образом вектора $\vec{x}_1 + \vec{x}_2$, а тому належить $\operatorname{Im} T$. Нехай c — скаляр. Тоді $c \cdot T(\vec{x}_1) = T(c\vec{x}_1) \in \operatorname{Im} T$, оскільки є образом елемента $c\vec{x}_1$. Теорему доведено. ■

Означення 4.11. Для даного лінійного відображення $T\colon V \to W$ розмірність підпростору $\operatorname{Ker} T$ називається *дефектом* відображення T, а розмірність підпростору $\operatorname{Im} T$ — *рангом* відображення T.

З огляду на означення 4.1 та 4.2 можемо стверджувати, що лінійне відображення $T\colon V \to W$ є сюр'єктивним, якщо $\operatorname{Im} T = W$, та ін'єктивним, якщо з того, що $T(\vec{x}_1) = T(\vec{x}_2)$ випливає, що $\vec{x}_1 = \vec{x}_2$. Крім цього, має місце таке твердження.

Теорема 4.10. *Лінійне відображення $T\colon W \to W$ ін'єктивне тоді і тільки тоді, коли $\operatorname{Ker} T = \{\vec{0}\}$.*

Доведення. Справді, нехай T є відображенням «в». Тоді оскільки $T(\vec{0}) = \vec{0}$ і за означенням ін'єктивного відображення різні вектори відображаються в різні, то $\operatorname{Ker} T = \{\vec{0}\}$. Навпаки, нехай $\operatorname{Ker} T = \{\vec{0}\}$ і припустимо, що існують такі $\vec{x}_1 \neq \vec{x}_2$, що $T(\vec{x}_1) = T(\vec{x}_2)$. Але тоді

$$\vec{0} = T(\vec{x}_1) - T(\vec{x}_2) = T(\vec{x}_1 - \vec{x}_2),$$

звідки $\vec{x}_1 - \vec{x}_2 = \vec{0}$, що суперечить припущенню. ■

Лема 4.1. *Нехай A — $m \times n$ матриця. Множина $\{A\vec{x}\colon \vec{x} \in \mathbb{R}^n\}$ є підпростором простору $\mathbb{R}^m$ і розмірність цього підпростору дорівнює рангу матриці A.*

Доведення. Розглядатимемо ранг матриці A як її стовпцевий ранг. Нагадаємо, що стовпцевим рангом матриці називається максимальна кількість її лінійно незалежних векторів-стовпців. Нехай $A = \begin{bmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{bmatrix}$. Тоді

$$\{A\vec{x}\colon \vec{x} \in \mathbb{R}^n\} = \left\{x_1\vec{a}_1 + \ldots + x_n\vec{a}_n\colon x_i \in \mathbb{R}, i = \overline{1, n}\right\}$$

і ця множина є лінійною оболонкою системи векторів $\left(\vec{a}_1, \ldots, \vec{a}_n\right)$. Добре відомо, що лінійна оболонка системи векторів є підпростором (теорема 1.13) і розмірність цього підпростору визначається максимальною кількістю лінійно незалежних векторів серед векторів $\vec{a}_1, \ldots, \vec{a}_n$, що і потрібно було довести. ■

Теорема 4.11. *Якщо A — матриця лінійного відображення $T\colon \mathbb{R}^n \to \mathbb{R}^m$, то ранг відображення T дорівнює рангу матриці A.*

Доведення. Справді, для кожного відображення $T\colon \mathbb{R}^n \to \mathbb{R}^m$ існує така матриця A, що $T(\vec{x}) = A\vec{x}$. Таким чином, маємо справу з відображенням, образом $\operatorname{Im} T$ якого є множина $\{A\vec{x}\colon \vec{x} \in \mathbb{R}^n\}$, яка за попереду доведеною лемою 4.1 є підпростором, розмірність якого співпадає з рангом матриці A. ■

Наслідок. *Якщо A — матриця лінійного відображення $T\colon V \to W$, то ранг відображення T дорівнює рангу матриці A.*

Приклад 4.11. Нехай дано лінійне відображення $T\colon \mathbb{R}^3 \to \mathbb{R}^3$:

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y+z \\ x-z \end{bmatrix}.$$

Знайти ядро та образ цього лінійного відображення, а також обчислити його дефект і ранг.

Розв'язання. Маємо:

$$\operatorname{Ker} T = \{(x,y,z)\colon x+y = y+z = x-z = 0\};$$
$$\operatorname{Im} T = \{(x+y, y+z, x-z)\colon x,y,z \in \mathbb{R}\}.$$

Розв'яжемо систему

$$\begin{cases} x+y = 0, \\ y+z = 0, \\ x-z = 0. \end{cases}$$

Її розв'язком є $(t,-t,t), t \in \mathbb{R}$. Отже, $\operatorname{Ker} T = \{(t,-t,t)\colon t \in \mathbb{R}\}$.

Тепер розглянемо систему

$$\begin{cases} x+y = a, \\ y+z = b, \\ x-z = c. \end{cases}$$

$$
\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 0 & 1 & 1 & b \\ 1 & 0 & -1 & c \end{array}\right]
\begin{matrix} \cdot(-1) \\ \\ + \end{matrix}
\sim
\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 0 & 1 & 1 & b \\ 0 & -1 & -1 & c-a \end{array}\right]
\begin{matrix} \\ \\ + \end{matrix}
\sim
$$

$$
\sim \left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 0 & 1 & 1 & b \\ 0 & 0 & 0 & b+c-a \end{array}\right].
$$

Таким чином, $b + c - a = 0$ і $c = a - b$. Отже, $\operatorname{Im} T = \{a, b, a - b \colon a, b \in \mathbb{R}\}$.

Бачимо, що ядро лінійного відображення є одновимірним підпростором простору $\mathbb{R}^3$, а образ — двовимірним. Отже, дефект дорівнює 1, а ранг — 2. ■

Теорема 4.12. *Нехай $T \colon V \to W$ — лінійне відображення скінченновимірного простору V у простір W. Тоді сума рангу і дефекту лінійного відображення T дорівнює розмірності простору V:*

$$\dim \operatorname{Ker} T + \dim \operatorname{Im} T = \dim V.$$

Доведення. За теоремою 4.9 ядро лінійного відображення є скінченновимірним підпростором простору V. Нехай $\big(\vec{e}_1, \ldots, \vec{e}_d\big)$ — базис $\operatorname{Ker} T$ (тобто дефект дорівнює d). Оскільки $\big(\vec{e}_1, \ldots, \vec{e}_d\big)$ є лінійно незалежною системою векторів, то вона може бути доповнена до базису простору V (теорема 3.10). Нехай таким базисом буде $\mathcal{B} = \big(\vec{e}_1, \ldots, \vec{e}_d, \vec{e}_{d+1}, \ldots, \vec{e}_n\big)$. Доведемо, що система векторів $\mathcal{C} = \big(T(\vec{e}_{d+1}), \ldots, T(\vec{e}_n)\big)$ є базисом $\operatorname{Im} T$. Для цього слід показати, що кожний вектор з $\operatorname{Im} T$ є лінійною комбінацією векторів системи $\mathcal{C}$ та довести лінійну незалежність системи $\mathcal{C}$.

Нехай $T(\vec{x}) \in \operatorname{Im} T$. Тоді $\vec{x} \in V$ і існують такі скаляри $\alpha_1, \ldots, \alpha_n$, що

$$\vec{x} = \alpha_1 \vec{e}_1 + \ldots + \alpha_d \vec{e}_d + \alpha_{d+1} \vec{e}_{d+1} + \ldots + \alpha_n \vec{e}_n.$$

Оскільки $\vec{e}_1, \ldots, \vec{e}_d \in \operatorname{Ker} T$, то $T(\vec{e}_1) = \ldots = T(\vec{e}_d) = 0$, і, отже,

$$
\begin{aligned}
T(\vec{x}) &= T(\alpha_1 \vec{e}_1 + \ldots + \alpha_d \vec{e}_d + \alpha_{d+1} \vec{e}_{d+1} + \ldots + \alpha_n \vec{e}_n) = \\
&= \alpha_1 T(\vec{e}_1) + \ldots + \alpha_d T(\vec{e}_d) + \alpha_{d+1} T(\vec{e}_{d+1}) + \ldots + \alpha_n T(\vec{e}_n) = \\
&= \alpha_{d+1} T(\vec{e}_{d+1}) + \ldots + \alpha_n T(\vec{e}_n).
\end{aligned}
$$

Тепер доведемо, що система векторів $\mathcal{C}$ є лінійно незалежною. Запишемо рівність

$$\alpha_{d+1}T(\vec{e}_{d+1}) + \ldots + \alpha_n T(\vec{e}_n) = \vec{0}.$$

З неї знаходимо, що $T(\alpha_{d+1}\vec{e}_{d+1} + \ldots + \alpha_n\vec{e}_n) = \vec{0}$, тобто $\alpha_{d+1}\vec{e}_{d+1} + \ldots + \alpha_n\vec{e}_n \in \operatorname{Ker} T$. Це означає, що цей вектор можна виразити через вектори базису $\operatorname{Ker} T$:

$$\alpha_{d+1}\vec{e}_{d+1} + \ldots + \alpha_n\vec{e}_n = \alpha_1\vec{e}_1 + \ldots + \alpha_d\vec{e}_d,$$

звідки

$$\alpha_1\vec{e}_1 + \ldots + \alpha_d\vec{e}_d - \alpha_{d+1}\vec{e}_{d+1} - \ldots - \alpha_n\vec{e}_n = \vec{0},$$

а остання рівність в силу лінійної незалежності системи векторів $\mathcal{B}$ виконується лише за умови, коли усі коефіцієнти дорівнюють нулю, зокрема і $\alpha_{d+1} = \ldots = \alpha_n = 0$. Отже, система $\mathcal{C}$ є лінійно незалежною.

Таким чином, нами доведено, що система векторів $\mathcal{C}$ утворює базис підпростору $\operatorname{Im} T$, а тому розмірність цього підпростору рівна $n - d$, що і завершує доведення теореми. ■

Додаткові відомості і завдання для самостійної роботи

1. Нехай $\vec{a}_1, \ldots, \vec{a}_k$ — вектори з лінійного простору V, $T\colon V \to W$ — лінійне відображення.

   а) довести, що якщо $\big(T(\vec{a}_1), \ldots, T(\vec{a}_k)\big)$ — лінійно незалежна система, то $\big(\vec{a}_1, \ldots, \vec{a}_k\big)$ також лінійно незалежна;

   б) довести (шляхом наведення контрприкладу), що твердження, обернене до а), взагалі кажучи є неправильним. Які обмеження слід накласти на відображення T, щоб воно стало правильним?
2. Нехай R, S, T — лінійні відображення такі, що наступні операції мають сенс. Довести, що:

   а) $R \circ (S + T) = R \circ S + R \circ T$;

   б) $\lambda(R \circ S) = (\lambda R) \circ S = R \circ (\lambda S)$ для довільного скаляра λ.
3. Нехай $S\colon V \to W$ і $T\colon U \to V$ — взаємно однозначні лінійні відображення. Довести, що відображення $S \circ T$ також взаємно однозначне.
4. Довести, що векторні простори $M_{mn}(\mathbb{R})$ та $\mathbb{R}^{mn}$ є ізоморфними.
5. Довести, що:

   а) ранг матриці лінійного сюр'єктивного відображення дорівнює кількості її рядків;

   а) ранг матриці лінійного ін'єктивного відображення дорівнює кількості її стовпців.

Розділ 5

Лінійні оператори векторного простору

> — Боюся, пані, я не зумію висловитися ясніше, — щонайчемніше відповіла Аліса. — Я й сама не все розумію: стільки перетворень за один день будь-кого зіб'ють з пантелику.
>
> *Л. Керрол «Аліса в країні чудес»*

§ 1. Поняття лінійного оператора (перетворення) векторного простору

Нехай маємо n-вимірний векторний простір V. Під *перетворенням* цього простору будемо розуміти відображення, яке кожному вектору $\vec{x}$ цього простору ставить у відповідність деякий вектор $\vec{x}'$ цього ж простору.

Означення 5.1. Перетворення T векторного простору V називається *лінійним перетворенням* (або *лінійним оператором*), якщо суму двох довільних векторів цього простору воно переводить в суму образів цих векторів, а добуток будь-якого вектора $\vec{x}$ на довільне число α — в добуток образа вектора $\vec{x}$ на це ж число α.

Таким чином, лінійний оператор векторного простору є лінійним відображенням векторного простору в себе. Тоді усі результати розділу 4 можна переформулювати в більш вузьких термінах лінійних операторів векторного простору.

Зокрема, один з основних результатів розділу 4 — теорема 4.5 — на мові лінійних операторів формулюється так:

Теорема 5.1. *Нехай T — лінійний оператор простору $\mathbb{R}^n$.*

1. *Існує $n \times n$ матриця A така, що $T(\vec{x}) = A \cdot \vec{x}$ для всіх $x \in \mathbb{R}^n$.*

2. *Стовпці матриці A — відповідно $T(\vec{e}_1), \dots, T(\vec{e}_n)$, де $\left(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\right)$ — стандартний базис простору $\mathbb{R}^n$:*
$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \dots & T(\vec{e}_n) \end{bmatrix}.$$

Матриця A з теореми 5.1 називається *стандартною матрицею лінійного оператора T.*

Розглянемо деякі приклади.

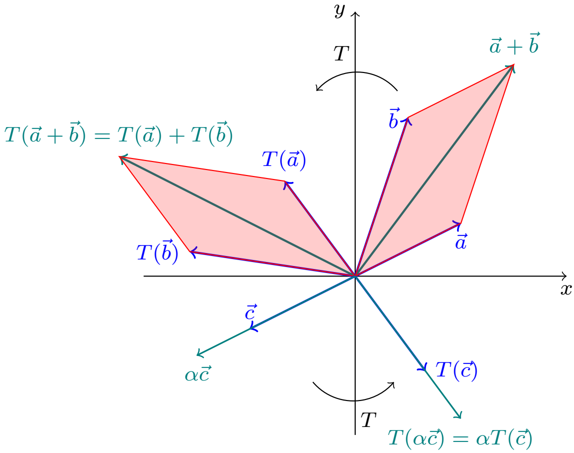
Приклад 5.1. Відображення $I \colon V \to V$ таке, що $I(\vec{x}) = \vec{x}$ для кожного $\vec{x} \in V$, є лінійним оператором простору V. Такий оператор називається *тотожним*. Стандартною матрицею цього лінійного оператора є одинична матриця n-го порядку.

Приклад 5.2. Нехай $R_O^{\varphi} \colon W_2 \to W_2$[1] — перетворення простору W_2, яке повертає кожну точку цього на кут φ проти годинникової стрілки відносно початку координат. З суто геометричних міркувань зрозуміло, що це перетворення є лінійним (див. рис. 5.1(a)).

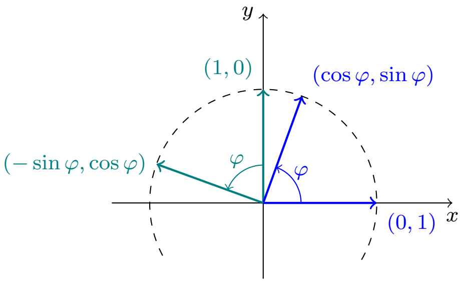
Знайдемо матрицю цього лінійного оператора. Для цього знайдемо образи базисних векторів:
$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}, \qquad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}.$$
За теоремою 5.1 знаходимо, що матриця оператора повороту простору W_2 має вигляд
$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$

Означення 5.2. *Ядром* лінійного оператора T простору V (позначається $\operatorname{Ker} T$) називається множина усіх векторів простору V, які відображаються у $\vec{0}$:
$$\operatorname{Ker} T = \left\{\vec{x} \in V : T(\vec{x}) = \vec{0}\right\}.$$

[1] W_2 — векторний простір усіх векторів площини з початком у початку координат відносно звичайних операцій додавання векторів і множення їх на числа. Цей простір ізоморфний простору $\mathbb{R}^2$.



а) Лінійність оператору повороту



б) Дія оператору повороту на стандартний базис простору

Рис. 5.1. Оператор повороту $T = R_O^{\varphi}$.

Означення 5.3. *Образом* лінійного оператора T простору V (позначається $\operatorname{Im} T$) називається множина усіх векторів простору, для яких є прообраз при відображенні T:

$$\operatorname{Im} T = \{T(\vec{x})\colon \vec{x} \in V\}\,.$$

Очевидно, що $\operatorname{Ker} T$ та $\operatorname{Im} T$ є підпросторами простору V. Розмірність ядра лінійного оператора називається його *дефектом*, а розмірність образу — *рангом*. Сума рангу і дефекту дорівнює розмірності простору V (див. теорему 4.12).

§ 2. Матриці лінійного оператора в різних базисах

2.1. Зв'язок між матрицями лінійного оператора в різних базисах.

Нехай V — n-вимірний векторний простір, заданий над полем P; $\mathcal{B} = \big(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\big)$ і $\mathcal{C} = \big(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n\big)$ — два базиси цього простору. Нехай також T — лінійний оператор простору V.

Як ми показали раніше (див. пункт 1.4), існує така матриця M, яка пов'язує координати довільного вектора $\vec{x} \in V$ в базисі $\mathcal{B}$ з координатами його образу $T(\vec{x})$ в базисі $\mathcal{C}$, а саме:

$$[T(\vec{x})]_{\mathcal{C}} = M \cdot [\vec{x}]_{\mathcal{B}}\,,$$

причому $M = \left[\Big[T(\vec{b}_1)\Big]_{\mathcal{C}} \quad \Big[T(\vec{b}_2)\Big]_{\mathcal{C}} \quad \ldots \quad \Big[T(\vec{b}_n)\Big]_{\mathcal{C}}\right]$. Таку матрицю M ми ще позначали так: $M = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$.

Часто розглядають випадок, коли $\mathcal{C} = \mathcal{B}$. Тоді матриця $M = [T]_{\mathcal{B}\leftarrow\mathcal{B}}$ позначається просто $[T]_{\mathcal{B}}$ і називається *матрицею лінійного оператора T в базисі $\mathcal{B}$*. У випадку ж, коли B є стандартним базисом, її називаються *стандартною матрицею оператора T* і позначають $[T]$.

Як і для лінійних відображень векторних просторів, має місце таке твердження (див. теорему 4.6).

Теорема 5.2. *Вибір базису в n-вимірному просторі V встановлює взаємно однозначну відповідність між лінійними операторами цього простору та квадратними матрицями n-го порядку.*

Нехай маємо лінійний оператор $T\colon V \to V$, два базиси $\mathcal{B} = \big(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\big)$ і $\mathcal{C} = \big(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n\big)$ простору V і нехай $[T]_{\mathcal{B}}$ і $[T]_{\mathcal{C}}$ — матриці цього лінійного оператора в базисах $\mathcal{B}$ і $\mathcal{C}$ відповідно.

Виникає питання: чи пов'язані якимось чином матриці $[T]_{\mathcal{B}}$ і $[T]_{\mathcal{C}}$?

Відповідь на поставлене питання дає таке твердження.

Теорема 5.3. *Нехай V — скінченновимірний векторний простір, $\mathcal{B} = \big(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\big)$, $\mathcal{C} = \big(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n\big)$ — два базиси цього простору.*

Нехай T — лінійний оператор простору V. Тоді матриці цього лінійного оператора в базисах $\mathcal{B}$ і $\mathcal{C}$ пов'язані співвідношенням:

$$[T]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T]_{\mathcal{B}} \cdot (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1},$$

де $P_{\mathcal{C}\leftarrow\mathcal{B}}$ — матриця переходу від базису $\mathcal{B}$ до базису $\mathcal{C}$.

Доведення. Для доведення теореми використаємо чотири співвідношення, які виконуються для довільного вектора $\vec{x}$ простору V:

$$[\vec{x}]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [\vec{x}]_{\mathcal{B}}, \tag{5.1}$$

$$[T(\vec{x})]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T(\vec{x})]_{\mathcal{B}}, \tag{5.2}$$

$$[T(\vec{x})]_{\mathcal{B}} = [T]_{\mathcal{B}} \cdot [\vec{x}]_{\mathcal{B}}, \tag{5.3}$$

$$[T(\vec{x})]_{\mathcal{C}} = [T]_{\mathcal{C}} \cdot [\vec{x}]_{\mathcal{C}}. \tag{5.4}$$

Тоді підставимо (5.2) в (5.4):

$$P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T(\vec{x})]_{\mathcal{B}} = [T]_{\mathcal{C}} \cdot [\vec{x}]_{\mathcal{C}}; \tag{5.5}$$

(5.3) в (5.5):

$$P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T]_{\mathcal{B}} \cdot [\vec{x}]_{\mathcal{B}} = [T]_{\mathcal{C}} \cdot [\vec{x}]_{\mathcal{C}}; \tag{5.6}$$

(5.1) в (5.6):

$$(P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T]_{\mathcal{B}}) \cdot [\vec{x}]_{\mathcal{B}} = ([T]_{\mathcal{C}} \cdot P_{\mathcal{C}\leftarrow\mathcal{B}})\, [\vec{x}]_{\mathcal{B}},$$

звідки

$$P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T]_{\mathcal{B}} = [T]_{\mathcal{C}} \cdot P_{\mathcal{C}\leftarrow\mathcal{B}},$$

тобто

$$[T]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot [T]_{\mathcal{B}} \cdot (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}.$$

■

2.2. Подібні матриці та їх властивості.

Нехай A та B — дві квадратні матриці n-го порядку.

Означення 5.4. Говорять, що матриця A *подібна* до матриці B (позначається $A \sim B$), якщо існує невироджена матриця Q така, що $A = QBQ^{-1}$.

Теорема 5.4. *Детермінанти подібних матриць однакові.*

Доведення. Нехай A і B — подібні матриці. Це означає, що існує невироджена матриця Q така, що $A = QBQ^{-1}$. Тоді

$$\det A = \det(QBQ^{-1}) = \det Q \cdot \det B \cdot \det Q^{-1} = \det Q \cdot \det Q^{-1} \cdot \det B =$$
$$= \det(Q \cdot Q^{-1}) \cdot \det B = \det B.$$

■

З теореми 5.3 випливає, що матриці лінійного оператора T в різних базисах подібні між собою.

В просторі $M_n(\mathbb{R})$ можна розглядати відношення «бути подібним» (відношення подібності матриць).

Теорема 5.5. *Відношення подібності матриць в просторі $M_n(\mathbb{R})$ є відношенням еквівалентності.*

Доведення. Слід довести, що відношення подібності матриць є рефлексивним, симетричним і транзитивним.

1. Рефлексивність ($A \sim A$) виконується, оскільки $A = I \cdot A \cdot I^{-1}$, де I — одинична матриця.
2. Нехай $A \sim B$. Це означає, що існує така матриця Q, що $A = QBQ^{-1}$. Тоді, помноживши останню рівність справа на Q, а зліва на Q^{-1}, дістанемо: $B = Q^{-1}AQ = Q^{-1}A(Q^{-1})^{-1}$, тобто $B \sim A$ і симетричність виконується.
3. Нехай $A \sim B$, $B \sim C$. Це означає, що існують невироджені матриці Q та S такі, що $A = QBQ^{-1}$, $B = SCS^{-1}$. Маємо:

$$A = QBQ^{-1} = Q(SCS^{-1})Q^{-1} = (QS) \cdot C \cdot (S^{-1}Q^{-1}) = (QS) \cdot C \cdot (QS)^{-1},$$

звідки $A \sim C$.

■

Отже, відношення подібності матриць розбиває простір квадратних матриць на класи еквівалентності. Важливою задачею є відшукання в різних класах еквівалентності в певному розумінні найпростіших представників. Такими представниками могли би бути, наприклад, діагональні матриці, оскільки ними зручно оперувати. На мові лінійних операторів це б означало відшукання такого базису простору V, при якому матриця даного лінійного оператора мала б діагональний вигляд. Але, як ми покажемо пізніше, не в кожному з класів еквівалентності можна обрати представником діагональну матрицю.

2.3. Алгебра лінійних операторів.

Означення 5.5. *Алгеброю над полем P* називається множина L з визначеними на ній внутрішніми бінарними операціями додавання ($+$) і множення ($\bullet$) та зовнішньою бінарною операцією множення на елементи поля P ($\cdot$) такими, що виконуються умови:

1) $(L, +, \cdot)$ — векторний простір;
2) $(L, +, \bullet)$ — кільце;
3) операції $\bullet$ та $\cdot$ пов'язані співвідношенням: $(\lambda \cdot \vec{a}) \bullet \vec{b} = \vec{a} \bullet (\lambda \cdot \vec{b}) = \lambda \cdot (\vec{a} \bullet \vec{b})$ для довільних $\lambda \in P$, $\vec{a}, \vec{b} \in L$.

Приклад 5.3. 1. Множина C комплексних чисел з визначеними операціями додавання, множення комплексних чисел і множення на дійсні числа є алгеброю над полем $\mathbb{R}$ дійсних чисел.

2. Множина $M_n(\mathbb{R})$ ($M_n(C)$) квадратних матриць n-го порядку з дійсними (комплексними) коефіцієнтами і визначеними на ній операціями додавання, множення матриць і множення матриці на дійсні (комплексні) числа є алгеброю над відповідним полем.
3. $(\mathcal{L}(V), +, \circ, \cdot)$ — множина лінійних операторів n-вимірного векторного простору V з операцією додавання, множення відображень (композиція) та множення на число є алгеброю на полем $\mathbb{R}$.

Відображення алгебр називається *ізоморфізмом*, якщо воно одночасно є ізоморфізмом відповідних векторних просторів та кілець.

Теорема 5.6. *Алгебри $(M_n(\mathbb{R}), +, \bullet, \cdot)$*[1] *та $(\mathcal{L}(\mathbb{R}^n), +, \circ, \cdot)$ ізоморфні.*

Остання теорема стверджує, що усі результати, які в подальшому нами будуть одержані для алгебри квадратних матриць n-го порядку «переносяться» на алгебру лінійних операторів n-вимірного векторного простору.

§ 3. Невироджені лінійні оператори

3.1. Вироджені та невироджені лінійні оператори.

Означення 5.6. Лінійний оператор T n-вимірного простору V називається *невиродженим*, якщо його ранг дорівнює n. Якщо ж його ранг менший n, то оператор називається *виродженим*.

---

[1]Знаком $\bullet$ позначена операція множення матриць, а знаком $\cdot$ — операція множення матриці на число

З огляду на теорему 4.12 одержуємо:

$$T\text{ — невироджений} \quad \Leftrightarrow \quad \dim \operatorname{Im} T = n \quad \Leftrightarrow \quad \dim \operatorname{Ker} T = 0$$

Приклад 5.4. Нехай $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ — симетрія відносно осі Ox. Стандартна матриця цього відображення $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$; її ранг дорівнює 2, що співпадає з розмірністю простору $\mathbb{R}^2$. Отже, T — невироджений оператор.

Приклад 5.5. Нехай $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ — проекція на вісь Ox. Стандартна матриця цього відображення $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$; ранг дорівнює 1 і менше розмірності простору $\mathbb{R}^2$. Таким чином, T — вироджений оператор.

Теорема 5.7. *Лінійний оператор T простору V є невиродженим тоді і тільки тоді, коли матриця цього оператора в довільному базисі є невиродженою.*

Доведення. Справді, ранг лінійного оператора дорівнює рангу матриці цього оператора в довільному базисі (див. наслідок до теореми 4.11), а матриця n-го порядку є невиродженою тоді і тільки тоді, коли її ранг дорівнює n. ■

Теорема 5.8. *Невироджений лінійний оператор простору є взаємнооднозначним відображенням цього простору.*

Доведення. Справді, оскільки ранг оператора дорівнює розмірності простору, то $\operatorname{Ker} T = \left\{\vec{0}\right\}$, звідки T є відображенням «в» (теорема 4.10). Разом з тим, оскільки $\operatorname{Im} T = V$, то T також є і відображенням «на». ■

Таким чином, невироджений лінійний оператор є ізоморфізмом векторного простору на себе.

3.2. Оборотні лінійні оператори.

Нагадаємо, що лінійний оператор $I\colon V \to V$ називається *тотожним*, якщо для кожного $\vec{x} \in V$ $I(\vec{x}) = \vec{x}$.

Означення 5.7. Лінійний оператор T' називається *оберненим* до лінійного оператора T, якщо $T \circ T' = T' \circ T = I$. Оператор T

називається *оборотним*, якщо до нього існує обернений (позначається T^{-1}).

Наприклад, в просторі $\mathbb{R}^2$ лінійний оператор повороту проти годинникової стрілки на кут 60° є оборотним, причому обернений до нього — оператор повороту на 60° за годинниковою стрілкою.

Теорема 5.9. *Лінійний оператор T є оборотним тоді і тільки тоді, коли він невироджений.*

Доведення. *Необхідність.* Нехай A — матриця оборотного лінійного оператора T; B — матриця оператора T^{-1}. Тоді за теоремою 4.7 про матрицю композиції відображень знаходимо, що $AB = BA = I$, тобто $B = A^{-1}$ і, отже, $\det A \neq 0$ і T невироджений за теоремою 5.7.

Достатність. Нехай T — невироджений лінійний оператор. Тоді матриця цього оператора також невироджена (теорема 5.7), а тому до неї існує обернена матриця A^{-1}. В силу того, що між лінійними відображеннями та матрицями відповідної розмірності існує взаємно однозначна відповідність, стверджуємо, що існує оператор T^{-1}, такий що $T \circ T^{-1} = T^{-1} \circ T = I$. ■

Таким чином, якщо A — матриця невиродженого оператора T в певному базисі, то A^{-1} — матриця оператора T^{-1} в цьому ж базисі.

Додаткові відомості і завдання для самостійної роботи

1. Довести, що якщо хоча б одна з двох матриць A, B невироджена, то матриці AB і BA подібні.
2. Довести, що існують вироджені матриці A, B такі, що матриці AB і BA не подібні.
3. Нехай U і W — підпростори n-вимірного векторного простору V, причому $\dim U + \dim W = n$. Довести, що існує лінійний оператор T простору V такий, що $U = \operatorname{Im} T$, $W = \operatorname{Ker} T$.
4. Нехай $T\colon V \to V$ — лінійний оператор. Довести, що множина всіх лінійних операторів S простору V таких, що $S \circ T = O$ (O — нульовий оператор: $O(\vec{x}) = \vec{0}$ для всіх $\vec{x} \in V$) є підпростором векторного простору $\mathcal{L}(V)$.

Розділ 6

Власні значення і власні вектори

> Теперь, создав свой собственный мир, он мог устроиться в нем поудобнее.
>
> *Антуан де Сент-Экзюпери. «Ночной полёт»*

Фактично власним вектором називається такий ненульовий вектор, який при лінійному відображенні переходить у пропорціональний («колінеарний») вектор. Це поняття має численні застосування до розв'язування різноманітних задач, деякі з яких ми розглянемо наприкінці цього розділу (для детальнішого ознайомлення можна використати посібник [12]).

§ 1. Власні значення і власні вектори лінійного оператора

Розглянемо спочатку приклад.

Приклад 6.1. Нехай відображення $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ задано своєю стандартною матрицею: $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$. Знайти образи векторів $\vec{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Розв'язання. Знаходимо:

$$T(\vec{u}) = A \cdot \vec{u}_1 = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix};$$

$$T(\vec{v}) = A \cdot \vec{u}_1 = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2\vec{v}$$

(див. також рис. 6.1). ■

Помічаємо, що при відображенні T з наведеного прикладу вектор $\vec{v}$ відображається у колінеарний йому вектор. Більше того, кожний

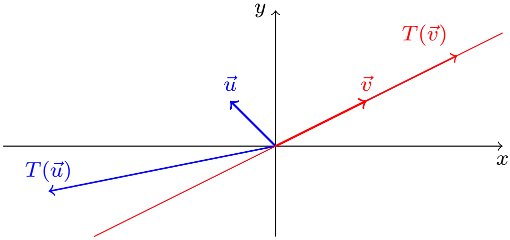


Рис. 6.1.

вектор, що колінеарний вектору $\vec{v}$ відображається у колінеарний вектор, причому довжина вектора збільшиться також в два рази. Справді, якщо $T(\vec{v}) = \lambda\vec{v}$, то $T(\lambda\vec{v}) = \lambda T(\vec{v}) = \lambda \cdot 2\vec{v} = 2 \cdot (\lambda\vec{v})$.

Отже, ми одержали таку множину $U = \{\lambda\vec{v} \colon \lambda \in \mathbb{R}\}$ (а вона є підпростором простору $\mathbb{R}^2$), що $T(U) \subset U$.

Вказане спостереження приводить до такого означення.

Означення 6.1. Підпростір U простору V називається *інваріантним відносно оператора* T, якщо для довільного $\vec{a} \in U$ $T(\vec{a}) \in U$, тобто $T(U) \subset U$.

Приклад 6.2. 1. $U = \{\vec{0}\}$ і $U = V$ — це тривіальні інваріантні підпростори кожного векторного простору V.

2. Нехай T — оператор повороту на кут α навколо осі Oz в просторі $\mathbb{R}^3$. Легко переконатись в тому, що таке перетворення простору $\mathbb{R}^3$ є лінійним. Прикладом інваріантного одновимірного підпростору простору $\mathbb{R}^3$ відносно оператора T є вісь Oz, а прикладом інваріантного двовимірного підпростору — площина xOy.

Особливо цікавим є випадок, коли інваріантні підпростори простору V є одновимірними. При цьому якщо U — інваріантний підпростір простору V відносно оператора T і $\vec{a} \in U$, то $T(\vec{a}) = \lambda\vec{a}$, тобто існує таке $\lambda \in P$, що $T(\vec{a}) = \lambda\vec{a}$.

Означення 6.2. Ненульовий вектор $\vec{u}$ простору V називається *власним вектором* лінійного оператора T цього простору, якщо існує таке $\lambda \in P$, що $T(\vec{u}) = \lambda\vec{u}$. При цьому число λ називається *власним значенням* лінійного оператора T, що відповідає власному вектору $\vec{u}$.

Приклад 6.3. Чи є вектор $\vec{x} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$ власним вектором лінійного оператора T, заданого своєю стандартною матрицею

$$A = \begin{bmatrix} 6 & -5 & -3 \\ 3 & -2 & -2 \\ 2 & -2 & 0 \end{bmatrix}?$$

Знайдемо образ вектора $\vec{x}$ при заданому відображенні:

$$T(\vec{x}) = A\vec{x} = \begin{bmatrix} 6 & -5 & -3 \\ 3 & -2 & -2 \\ 2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} = 2\vec{x},$$

тобто $\vec{x}$ є власним вектором лінійного оператора.

З означення 6.2 випливає, що $\vec{x}$ є власним вектором, коли рівняння $A \cdot \vec{x} = \lambda\vec{x}$ має ненульовий розв'язок. Але останнє матричне рівняння рівносильне такому:

$$(A - \lambda I)\vec{x} = \vec{0},$$

що є матричною формою запису однорідної системи лінійних рівнянь. Відомо, що однорідна система лінійних рівнянь має нетривіальний розв'язок тоді і тільки тоді, коли матриця цієї системи є виродженою.

Приклад 6.4. Нехай $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$. Довести, що 7 є власним значенням цієї матриці і знайти власні вектори, які йому відповідають.

Справді, матриця $A - 7I = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$ є виродженою. Власні вектори знаходимо з системи:

$$(A - 7I)\vec{x} = \vec{0}.$$

Маємо:

$$\begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

$$\begin{cases} -6x_1 + 6x_2 = 0, \\ 5x_1 - 5x_2 = 0. \end{cases}$$

Власними векторами є усі вектори виду $\begin{bmatrix} a \\ a \end{bmatrix}$, $a \in \mathbb{R}$.

Як відомо (див. теорему 1.14), множина розв'язків однорідної системи лінійних рівнянь з n невідомими (загальний розв'язок системи) утворює підпростір простору $\mathbb{R}^n$. Якщо λ — власне значення матриці A, то загальний розв'язок системи $(A-\lambda I)\vec{x}=\vec{0}$ називається *власним підпростором*, що відповідає власному значенню λ. Для визначення простору достатньо знати його базисні вектори. Тому надалі, розв'язуючи задачу відшукання власних векторів лінійного оператора (матриці), що відповідають власному значенню λ, ми будемо обмежуватись знаходженням лінійно незалежних власних векторів, що відповідають цьому значенню (знаходити фундаментальну систему розв'язків системи $(A-\lambda I)\vec{x}=\vec{0}$).

Приклад 6.5. Довести, що $\lambda = 6$ є власним значенням матриці
$$\begin{bmatrix} 7 & 1 & -2 \\ -3 & 3 & 6 \\ 2 & 2 & 2 \end{bmatrix}$$
і знайти базис власного підпростору, що відповідає цьому значенню.

Розв'язання. Оскільки матриця
$$A-\lambda I = \begin{bmatrix} 7 & 1 & -2 \\ -3 & 3 & 6 \\ 2 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 \\ -3 & -3 & 6 \\ 2 & 2 & -4 \end{bmatrix}$$
є виродженою (її ранг дорівнює 1), то 6 справді є власним значенням матриці A.

Власний підпростір U, що відповідає власному значенню 6 — це загальний розв'язок системи
$$\begin{bmatrix} 1 & 1 & -2 \\ -3 & -3 & 6 \\ 2 & 2 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$
$$\begin{cases} x_1 + x_2 - 2x_3 = 0, \\ -3x_1 - 3x_2 + 6x_3 = 0, \\ 2x_1 + 2x_2 - 4x_3 = 0, \end{cases}$$
Одержуємо $x_1 = -x_2 + 2x_3$, звідки знаходимо загальний розв'язок системи
$$U = \left\{ \begin{bmatrix} -a+2b \\ a \\ b \end{bmatrix} : a, b \in \mathbb{R} \right\} = L\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right).$$

Отже, в якості базису власного підпростору можна взяти власні вектори

$$\vec{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \text{ та } \vec{u}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

■

§ 2. Знаходження власних значень та власних векторів

2.1. Характеристичне рівняння лінійного оператора.

Припустимо, що лінійний оператор T n-вимірного простору V задано своєю матрицею A в деякому базисі $\mathcal{B} = \left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right)$. Нехай λ — власне значення цього лінійного оператора. Тоді існує ненульовий вектор $\vec{x}$ такий, що $T(\vec{x}) = \lambda\vec{x}$, тобто $A\,[\vec{x}]_{\mathcal{B}} = \lambda\,[\vec{x}]_{\mathcal{B}}$. Остання рівність рівносильна однорідній системі рівнянь, яка у матричній формі має вигляд

$$(A - \lambda I)\vec{x} = \vec{0}.$$

Нам відомо, що ця система має ненульовий розв'язок, отже визначник матриці $A - \lambda I$ дорівнює нулю.

Таким чином, усі власні значення лінійного оператора T — це корені λ рівняння $\det(A - \lambda I) = 0$. $\det(A - \lambda I)$ є многочленом n-го степеня відносно λ. Він називається *характеристичним многочленом лінійного оператора T*, що задано в деякому базисі $\mathcal{B}$ матрицею A.

Рівняння ж $\det(A - \lambda I) = 0$ назвемо *характеристичним рівнянням лінійного оператора T*, який в базисі $\mathcal{B}$ задано матрицею A.

Теорема 6.1. *Характеристичний многочлен лінійного оператора T не залежить від вибору базису.*

Доведення. Нехай лінійний оператор T n-вимірного простору V в базисі $\mathcal{B} = \left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right)$ подається матрицею A, а в базисі $\mathcal{C} = \left(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n\right)$ — матрицею A'.

Згідно з теоремою 5.3 матриці A та A' пов'язані співвідношенням

$$A' = P_{\mathcal{C}\leftarrow\mathcal{B}} \cdot A \cdot \left(P_{\mathcal{C}\leftarrow\mathcal{B}}\right)^{-1},$$

де $P_{\mathcal{C}\leftarrow\mathcal{B}}$ — матриця переходу від базису $\mathcal{B}$ до $\mathcal{C}$ (для спрощення позначимо її через P). Тоді, застосовуючи властивості операцій над матрицями та теорему про детермінант добутку матриць, знаходимо:

$$\begin{aligned}\det(A' - \lambda I) &= \det(P \cdot A \cdot P^{-1} - \lambda I) = \det(PAP^{-1} - \lambda \cdot PIP^{-1}) = \\ &= \det(P \cdot (A - \lambda I) \cdot P^{-1}) = \det P \cdot \det(A - \lambda I) \cdot \det P^{-1} = \\ &= \det(P \cdot P^{-1}) \cdot \det(A - \lambda I) = \det(A - \lambda I),\end{aligned}$$

що і потрібно було довести. ■

Зауваження. У зв'язку з тим, що за доведеною теоремою характеристичний многочлен лінійного оператора не залежить від вибору базису, замість слів «характеристичний многочлен лінійного оператора T в базисі $\mathcal{B}$» будемо говорити просто *«характеристичний многочлен лінійного оператора T»*.

Ізоморфізм алгебр лінійних операторів простору $\mathbb{R}^n$ і матриць $M_n(\mathbb{R})$ дає підстави до означення характеристичного многочлена матриці.

Означення 6.3. *Характеристичним многочленом* матриці A називається многочлен відносно змінної λ $\det(A - \lambda I)$.

Таким чином, для знаходження власних значень лінійного оператора (матриці) необхідно знаходити корені характеристичного многочлена. Відповідь на питання про можливу кількість таких коренів дають такі класичні теореми.

Теорема 6.2. *Многочлен n-го степеня з комплексними коефіцієнтами не може мати більше ніж n коренів.*

Теорема 6.3 (Основна теорема алгебри многочленів)**.** *Кожний многочлен степеня $n \geqslant 1$ з комплексними коефіцієнтами має принаймні один комплексний корінь.*

Остання теорема є фундаментальною в математиці. Їй відведена особлива назва: вона називається *основною теоремою алгебри многочленів* (або просто *основна теорема алгебри*). Її доведення можна знайти в підручниках з теорії функцій комплексної змінної[1].

---

[1] Див., наприклад, *Маркушевич А. И.* Теория аналитических функций. Том 1. — М., ФМЛ, 1967. — С. 312.

Отже, з основної теореми алгебри випливає, що якщо векторний простір розглядається над полем комплексних чисел, то кожний лінійний оператор цього простору має принаймні один власний вектор. Якщо ж векторний простір розглядається над полем дійних чисел, то лінійний оператор може взагалі не мати власних векторів.

2.2. Алгоритм відшукання власних значень і власних векторів матриці. Нехай A — матриця n-го порядку.

1. Складаємо характеристичний многочлен матриці A: $\det(A - \lambda I)$.
2. Розв'язуємо характеристичне рівняння $\det(A-\lambda I) = 0$ і знаходимо власні значення $\lambda_1, \ldots, \lambda_m$.
3. Для кожного власного значення λ_i, $1 \leqslant i \leqslant m$ знаходимо базис простору розв'язків однорідної системи $(A - \lambda_i I)\vec{x} = \vec{0}$ (фундаментальну систему розв'язків цієї системи) — сукупність лінійно незалежних власних векторів, що відповідають власному значенню λ_i.

   Власним підпростором, що відповідає λ_i, є загальний розв'язок системи $(A - \lambda_i I)\vec{x} = \vec{0}$.

Зауважимо, що знайти власні значення і власні вектори лінійного оператора — означає знайти власні значення і власні вектори матриці цього лінійного оператора в довільному базисі.

Приклад 6.6. Знайти власні значення і власні вектори матриці

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

Розв'язання. Знайдемо характеристичний многочлен матриці A:

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

$$= -\lambda(-4\lambda + \lambda^2 + 5) + 2 = -\lambda^3 + 4\lambda^2 - 5\lambda + 2.$$

Складемо і розв'яжемо характеристичне рівняння:

$$-\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0;$$

$$\lambda^3 - 4\lambda^2 + 5\lambda - 2 = 0.$$

Розклавши многочлен в лівій частині на множники, знаходимо:

$$(\lambda-1)^2(\lambda-2)=0,$$

тобто характеристичне рівняння має двократний корінь 1 ($\lambda_1=\lambda_2=1$) і однократний корінь 2 ($\lambda_3=2$). Отже, 1 і 2 — власні значення матриці A.

Знайдемо власні вектори, що відповідають власному значенню 1. Для цього слід розв'язати систему $(A-1\cdot I)\vec{x}=\vec{0}$. Маємо:

$$\begin{bmatrix}-1 & 1 & 0\\ 0 & -1 & 1\\ 2 & -5 & 3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix};$$

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

Зауважимо, що ми напевно знаємо, що ця система має безліч розв'язків (матриця цієї системи є виродженою). Це означає, що ранг матриці системи менший 3. Оскільки перше і друге рівняння незалежні, то третє рівняння є наслідком перших двох. Тоді $x_1=x_2=x_3$ і загальним розв'язком системи (власним підпростором, що відповідає власному значенню 1) є множина $\{(a,a,a)|a\in\mathbb{R}\}$. В якості базисного вектору можна взяти вектор $\vec{u}_1=\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}$.

Аналогічно знайдемо власні вектори, що відповідають власному значенню 2. Для цього розв'яжемо систему $(A-2\cdot I)\vec{x}=\vec{0}$:

$$\begin{bmatrix}-2 & 1 & 0\\ 0 & -2 & 1\\ 2 & -5 & 2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix};$$

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

звідки знаходимо загальний розв'язок $\{(a,2a,4a)|a\in\mathbb{R}\}$. Отже, усі власні вектори, що відповідають власному значенню 2, колінеарні вектору $\vec{u}_2=\begin{bmatrix}1\\ 2\\ 4\end{bmatrix}$. ■

2.3. Деякі властивості власних значень та власних векторів.

Теорема 6.4. *Власними значеннями трикутної матриці є ті її елементи, що розташовані на головній діагоналі.*

Доведення. Доведення проведемо для випадку верхньої трикутної матриці (доведення для нижньої трикутної матриці цілком аналогічне). Нехай A — верхня трикутна матриця, тобто матриця виду:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \dots & a_{2,n-1} & a_{2n} \\ 0 & 0 & \dots & a_{3,n-1} & a_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & a_{nn} \end{bmatrix}$$

(усі елементи, що розташовані під головною діагоналлю дорівнюють нулю). Тоді матриця $A - \lambda I$ є також верхньою трикутною. Відомо, що детермінант трикутної матриці дорівнює добутку елементів, розташованих на головній діагоналі. Тому

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda)\dots(a_{nn} - \lambda).$$

Очевидно, що коренями характеристичного рівняння є $a_{11}, a_{22}, \dots, a_{nn}$. ■

Теорема 6.5. *Квадратна матриця A є оборотною (неособливою) тоді і тільки тоді, коли 0 не є власним значенням цієї матриці.*

Доведення. *Необхідність.* Нехай матриця A оборотна. Припустимо, що 0 є її власним значенням. Це означає, що детермінант матриці $A - 0 \cdot I = A$ дорівнює нулю, що суперечить тому, що матриця оборотна. Отже, 0 не може бути власним значенням оборотної матриці.

Достатність. Нехай 0 не є власним значенням матриці A. Припустимо, що до матриці A не існує оберненої. Тоді вона вироджена і її детермінант дорівнює нулю. Тоді $0 = \det A = \det(A - 0 \cdot I)$, тобто 0 є власним значенням цієї матриці. Дістали суперечність, отже наше припущення було невірним і матриця A є оборотною. ■

Теорема 6.6. *Нехай A — квадратна матриця, λ — її власне значення, що відповідає власному вектору $\vec{u}$.*

1. *Для кожного $k \in \mathbb{N}$ $\vec{u}$ є власним вектором матриці A^k, а λ^k є власним значенням, що відповідає цьому власному вектору.*
2. *Якщо A — оборотна, то $\frac{1}{\lambda}$ є власним значенням матриці A^{-1}, що відповідає власному вектору $\vec{u}$.*

Доведення. 1. Доведення проведемо методом математичної індукції. База індукції очевидна. Припустимо, що для довільного, але фіксованого l, твердження теореми має місце, тобто λ^l є власним значенням матриці A^l, а $\vec{u}$ — власним вектором, що відповідає цьому власному значенню, тобто $A^l\vec{u} = \lambda^l\vec{u}$. Тоді

$$A^{l+1}\vec{u} = A^l \cdot A\vec{u} = A^l \cdot \lambda\vec{u} = \lambda \cdot A^l\vec{u} = \lambda^{l+1}\vec{u},$$

тобто твердження теореми має місце і для $k = l + 1$.

2. Нехай матриця A оборотна і λ — її власне значення, що відповідає власному вектору $\vec{u}$. З теореми 6.5 випливає, що $\lambda \neq 0$. Маємо: $A\vec{u} = \lambda\vec{u}$. Домножимо обидві частини останньої рівності на λ^{-1} і зліва на A^{-1}. Дістанемо:

$$\frac{1}{\lambda}A^{-1}A\vec{u} = A^{-1}\vec{u},$$

звідки $\frac{1}{\lambda}\vec{u} = A^{-1}\vec{u}$, що і означає, що $\frac{1}{\lambda}$ є власним значенням матриці A^{-1}, що відповідає власному вектору $\vec{u}$.

■

Теорема 6.7. *Система $\vec{u}_1, \ldots, \vec{u}_k$ власних векторів лінійного оператора T n-вимірного простору V $(k \leqslant n)$, яким відповідають попарно різні власні значення, є лінійно незалежною.*

Доведення. Доведемо методом математичної індукції за числом векторів системи.

При $k = 1$ теорема, вочевидь, має місце. Припустимо, що теорема має місце при довільному, але фіксованому $k = s < n$, тобто що система векторів $\vec{u}_1, \ldots, \vec{u}_s$ є лінійно незалежною. Приєднаємо до цієї системи ще один вектор $\vec{u}_{s+1}$. Потрібно показати, що утворена нова система векторів $\vec{u}_1, \ldots, \vec{u}_s, \vec{u}_{s+1}$ є лінійно незалежною.

Припустимо супротивне, а саме: нехай вектори $\vec{u}_1, \ldots, \vec{u}_s, \vec{u}_{s+1}$ утворюють лінійно залежну систему. Тоді оскільки $\vec{u}_1, \ldots, \vec{u}_s$ — лінійно незалежна система, то вектор $\vec{u}_{s+1}$ лінійно виражається через вектори $\vec{u}_1, \ldots, \vec{u}_s$:

$$\vec{u}_{s+1} = \alpha_1\vec{u}_1 + \cdots + \alpha_s\vec{u}_s. \tag{6.1}$$

Подіємо на обидві частини останньої рівності оператором T:

$$T(\vec{u}_{s+1}) = T(\alpha_1\vec{u}_1 + \dots + \alpha_s\vec{u}_s),$$

звідки з урахуванням лінійності T і того, що $\vec{u}_i$, $1 \leqslant i \leqslant s+1$ є власними векторами:

$$\lambda_{s+1}\vec{u}_{s+1} = \alpha_1\lambda_1\vec{u}_1 + \dots + \alpha_s\lambda_s\vec{u}_s. \tag{6.2}$$

З іншого боку, помножимо (6.1) на λ_{s+1} (λ_{s+1} не може дорівнювати нулю, оскільки в протилежному випадку рівність (6.2) свідчила б про лінійну залежність системи векторів $\vec{u}_1, \dots, \vec{u}_s$) і дістанемо:

$$\lambda_{s+1}\vec{u}_{s+1} = \alpha_1\lambda_{s+1}\vec{u}_1 + \dots + \alpha_s\lambda_{s+1}\vec{u}_s. \tag{6.3}$$

Віднімемо від рівності (6.2) рівність (6.3):

$$\vec{0} = \alpha_1(\lambda_1 - \lambda_{s+1})\vec{u}_1 + \dots + \alpha_s(\lambda_s - \lambda_{s+1})\vec{u}_s.$$

Остання рівність для лінійно незалежної системи векторів $\vec{u}_1, \dots, \vec{u}_s$ можлива лише, коли усі коефіцієнти дорівнюють при векторах нулю. Оскільки власні значення є попарно різними, то $\alpha_1 = \ldots = \alpha_s = 0$. Тоді з (6.1) одержуємо $\vec{u}_{s+1} = \vec{0}$, що суперечить тому, що власний вектор обов'язково ненульовий. Таким чином, наше припущення невірне і система векторів $\vec{u}_1, \dots, \vec{u}_s, \vec{u}_{s+1}$ є лінійно незалежною. ■

Приклад 6.7. Обчислити $\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}^{10} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

Розв'язання. Одним із способів розв'язання цієї задачі є спосіб безпосереднього обчислення степенів матриці $\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ до десятого степеня включно з подальшим множенням на матрицю $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$. Проте ми розглянемо інший спосіб розв'язання цієї задачі, при якому показник 10 може бути замінено довільним іншим натуральним числом.

Введемо такі позначення: $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$. Отже, нам необхідно обчислити $A^{10}\vec{x}$. Знайдемо власні значення і власні вектори матриці A.

Характеристичне рівняння $\begin{vmatrix} -\lambda & 1 \\ 2 & 1-\lambda \end{vmatrix} = 0$ має корені $\lambda_1 = -1$, $\lambda_2 = 2$. Власні вектори, що відповідають цим власним значенням відповідно $\vec{u}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ та $\vec{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Вони лінійно незалежні, і, більше того, утворюють базис простору $\mathbb{R}^2$. Знайдемо координати вектора $\vec{x}$ в цьому базисі. Нехай $\vec{x} = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2$. Тоді

$$\begin{bmatrix} 5 \\ 1 \end{bmatrix} = \alpha_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix};$$

$$\begin{cases} \alpha_1 + \alpha_2 = 5, \\ -\alpha_1 + 2\alpha_2 = 1. \end{cases}$$

Знаходимо: $\alpha_1 = 3$, $\alpha_2 = 2$. Таким чином, $\vec{x} = 3\vec{u}_1 + 2\vec{u}_2$. Тоді

$$A^{10} \cdot \vec{x} = A^{10} \cdot (3\vec{u}_1 + 2\vec{u}_2) = 3(A^{10} \cdot \vec{u}_1) + 2(A^{10} \cdot \vec{u}_2) =$$

$$= 3(\lambda_1^{10})\vec{u}_1 + 2(\lambda_2^{10})\vec{u}_2 = 3(-1)^{10} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2(2^{10}) \begin{bmatrix} 1 \\ 2 \end{bmatrix} =$$

$$= \begin{bmatrix} 3 + 2^{11} \\ -3 + 2^{12} \end{bmatrix} = \begin{bmatrix} 2051 \\ 4093 \end{bmatrix}.$$

■

§ 3. Діагоналізація матриць

Як було сказано вище, актуальною є проблема відшукання в класі подібних матриць в певному розумінні найпростішого представника, оскільки матриці лінійного оператора в різних базисах подібні. Таким представником можна було б вважати діагональні матриці, оскільки над ними легко виконувати операції. В цьому параграфі ми розглянемо умови, за яких задана матриця є подібною до діагональної, а також встановимо вигляд такої діагональної матриці.

3.1. Умови, за яких матриця зводиться до діагонального виду.

Означення 6.4. Матриця A n-го порядку називається *діагоналізовною* (*зводиться до діагонального виду*), якщо вона подібна до діагональної матриці, тобто існують невироджена матриця P і діагональна D такі, що $A = PDP^{-1}$ (або, що те ж саме, $AP = PD$).

Приклад 6.8. Чи зводиться до діагональної матриця $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$?

Розв'язання. Доведемо, що відповідь позитивна, і шуканими є матриці $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ і $P = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix}$, де $\vec{u}_1$, $\vec{u}_2$ — власні вектори, що відповідають власним значенням λ_1 і λ_2 відповідно.

Знайдемо власні значення матриці A. Для цього розв'яжемо характеристичне рівняння:

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

Маємо: $\lambda^2 - 3\lambda - 28 = 0$, звідки $\lambda_1 = -4$, $\lambda_2 = 7$.

Власний вектор, що відповідає власному значенню $\lambda_1 = -4$, є розв'язком системи

$$\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4x_1 \\ -4x_2 \end{bmatrix}.$$

Таким вектором є вектор $\vec{u}_1 = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$. Аналогічно знаходимо, що власним вектором, що відповідає власному значенню $\lambda_2 = 7$, є вектор $\vec{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Формуємо матриці $P = \begin{bmatrix} 6 & 1 \\ -5 & 1 \end{bmatrix}$, $D = \begin{bmatrix} -4 & 0 \\ 0 & 7 \end{bmatrix}$.

Нарешті, обчислюємо:

$$AP = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 & 1 \\ -5 & 1 \end{bmatrix} = \begin{bmatrix} -24 & 7 \\ 20 & 7 \end{bmatrix};$$

$$PD = \begin{bmatrix} 6 & 1 \\ -5 & 1 \end{bmatrix} \begin{bmatrix} -4 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -24 & 7 \\ 20 & 7 \end{bmatrix}.$$

Таким чином матриця A зводиться до діагональної. ■

Наведений приклад дозволяє висунути гіпотезу про алгоритм зведення матриці до діагонального виду. Наступна теорема встановлює, коли саме це можливо здійснити і обґрунтовує вигляд матриць P та D.

Теорема 6.8. *Матриця n-го порядку зводиться до діагональної тоді і тільки тоді, коли вона має n лінійно незалежних власних векторів.*

Доведення. *Необхідність.* Нехай матриця A зводиться до діагональної. Це означає, що існують невироджена матриця P та діагональна D такі, що $AP = PD$. Нехай $P = \begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n\end{bmatrix}$,

$$D = \begin{bmatrix}\alpha_1 & 0 & \dots & 0\\ 0 & \alpha_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \alpha_n\end{bmatrix}.$$

Тоді

$$AP = A\cdot\begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n\end{bmatrix} = \begin{bmatrix}A\vec{v}_1 & A\vec{v}_2 & \dots & A\vec{v}_n\end{bmatrix};$$

$$PD = \begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n\end{bmatrix} D = \begin{bmatrix}\alpha_1\vec{v}_1 & \alpha_2\vec{v}_2 & \dots & \alpha_n\vec{v}_n\end{bmatrix}.$$

Тоді $A\vec{v}_i = \alpha_i\vec{v}_i$ для кожного $1 \leqslant 1 \leqslant n$. Це означає, що $\vec{v}_1, \dots, \vec{v}_n$ — власні вектори матриці A і вони є лінійно незалежними, оскільки формують стовпці невиродженої матриці P.

Достатність. Нехай, $\vec{u}_1, \dots, \vec{u}_n$ — лінійно незалежні власні вектори матриці A, $\lambda_1, \dots, \lambda_n$ — відповідні їм власні значення (зауважимо, що серед власних значень можуть бути і однакові числа). Складемо матрицю $P = \begin{bmatrix}\vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_n\end{bmatrix}$, $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$. Тоді

$$AP = \begin{bmatrix}A\vec{u}_1 & A\vec{u}_2 & \dots & A\vec{u}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\vec{u}_1 & \lambda_2\vec{u}_2 & \dots & \lambda_n\vec{u}_n\end{bmatrix} = PD,$$

що і означає, що матриця A зводиться до діагонального виду. ■

Алгоритм діагоналізації квадратної матриці A:

1. Знаходимо усі власні значення матриці A. Якщо коренів менше n (з урахуванням кратності), то матриця не зводиться до діагонального виду; якщо коренів рівно n, то переходимо до наступного кроку.

2. Знаходимо усі лінійно незалежні власні вектори $\vec{u}_1, \dots, \vec{u}_k$, що відповідають знайденим власним значенням. Якщо $k = n$ (n — розмірність матриці), то матриця A зводиться до діагонального виду і переходимо до кроку 3, якщо ж $k < n$ — то не зводиться.

3. Складаємо матриці $P = \begin{bmatrix}\vec{u}_1 & \dots & \vec{u}_n\end{bmatrix}$ та $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$, причому власному вектору $\vec{u}_k$ повинно відповідати його власне значення λ_k, $k = \overline{1, n}$. Тоді
$$A = P\cdot D\cdot P^{-1}.$$

Приклад 6.9. Звести до діагонального виду матрицю

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

Розв'язання. Зазначимо, що ця матриця розглядалась в прикладі 6.6, де нами було встановлено, що вона має лише два лінійно незалежні власні вектори $\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ та $\vec{u}_2 = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$. Отже, за теоремою 6.8 стверджуємо, що дана матриця не зводиться до діагонального виду. ■

Зауваження. Як бачимо з прикладу, кількість лінійно незалежних власних векторів лінійного оператора може бути меншою n і в цьому випадку він не може бути зведений до діагональної форми. Виникає питання: як знайти базис простору, в якому матриця такого перетворення має якомога простіший вигляд? Такий простіший вигляд називають *жордановою нормальною формою*, яка може бути знайдена для кожної матриці, а у випадку, коли вона має n лінійно незалежних власних векторів — співпадає з діагональною. Детальніше з цим можна ознайомитись, наприклад, в посібниках [2, 16].

3.2. Застосування діагоналізації матриць до знаходження степенів матриць.

Подання матриці у вигляді добутку інших матриць часто допомагає при розв'язанні різноманітних задач. Покажемо, як діагоналізація матриць допомагає у піднесенні матриць до степеня.

Теорема 6.9. *Нехай A — квадратна матриця n-го порядку, яка зводиться до діагонального виду: $A = PDP^{-1}$. Тоді для довільного натурального n:*

$$A^n = P \cdot D^n \cdot P^{-1}.$$

Доведення. Доведення проведемо методом математичної індукції. Для $n = 1$ твердження очевидне. Припустимо, що твердження має місце для довільного, але фіксованого k, тобто $A^k = P \cdot D^k \cdot P^{-1}$.

Тоді

$$A^{k+1} = A^k \cdot A = (P \cdot D^k \cdot P^{-1}) \cdot (P \cdot D \cdot P^{-1}) = \\ = PD^k(P^{-1}P)DP^{-1} = P(D^kD)P^{-1} = PD^{k+1}P^{-1},$$

що і потрібно було довести. ■

Приклад 6.10. Знайти A^{100}, якщо $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$.

Розв'язання. Матриця A фігурувала у прикладі 6.7. В тому прикладі було обчислено власні значення та власні вектори цієї матриці:

$$\lambda_1 = -1, \quad \vec{u}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix};$$

$$\lambda_2 = 2, \quad \vec{u}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Оскільки матриця A другого порядку, а лінійно незалежних власних векторів також 2, то за теоремою 6.8 її можна звести до діагонального виду, причому

$$A = PDP^{-1}, \qquad \text{де} \quad D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}.$$

Тоді

$$A^{100} = PD^{100}P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} (-1)^{100} & 0 \\ 0 & 2^{100} \end{bmatrix} \cdot \frac{1}{3} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \\ = \frac{1}{3} \begin{bmatrix} 1 & 2^{100} \\ -1 & 2^{101} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2^{100} + 2 & 2^{100} - 1 \\ 2^{101} - 2 & 2^{101} + 1 \end{bmatrix}.$$

■

3.3. Застосування діагоналізації матриць до розв'язування різницевих рівнянь.

Інше застосування теорії матриць загалом і діагоналізації матриць зокрема, а саме розв'язування різницевих рівнянь, проілюструємо на прикладі.

Приклад 6.11. Послідовність чисел (x_n) задана рекурентно: $x_{n+2} = 6x_n + x_{n+1}$, $n \geqslant 0$, $x_0 = x_1 = 1$. Знайти формулу для вираження x_n через n.

Розв'язання. Нехай $\vec{v}_n = \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix}$, $n \geqslant 0$. Тоді

$$\vec{v}_{n+1} = \begin{bmatrix} x_{n+1} \\ x_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 6 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix}.$$

Позначимо $A = \begin{bmatrix} 0 & 1 \\ 6 & 1 \end{bmatrix}$, тоді $\vec{v}_{n+1} = A \cdot \vec{v}_n$. Індуктивно знаходимо, що $\vec{v}_n = A^n \cdot \vec{v}_0$. А оскільки $\vec{v}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, то розв'язання задачі зводиться до знаходження A^n. З цією метою використаємо метод, що був запропонований в п. 3.2.

Характеристичне рівняння матриці A

$$\begin{vmatrix} -\lambda & 1 \\ 6 & 1-\lambda \end{vmatrix} = 0$$

має корені $\lambda_1 = -2$, $\lambda_2 = 3$. Власні вектори, що відповідають цим власним значенням — відповідно $\vec{u}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ та $\vec{u}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Таким чином,

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

$$\begin{aligned} A^n = PD^nP^{-1} &= \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} (-2)^n & 0 \\ 0 & 3^n \end{bmatrix} \cdot \frac{1}{5} \begin{bmatrix} 3 & -1 \\ 2 & 1 \end{bmatrix} = \\ &= \frac{1}{5} \begin{bmatrix} 3(-2)^n + 2\cdot 3^n & -(-2)^n + 3^n \\ 3(-2)^{n+1} + 2\cdot 3^{n+1} & -(-2)^{n+1} + 3^{n+1} \end{bmatrix}. \end{aligned}$$

Повернемось до $\vec{v}_n$:

$$\begin{aligned} \vec{v}_n &= \frac{1}{5} \begin{bmatrix} 3(-2)^n + 2\cdot 3^n & -(-2)^n + 3^n \\ 3(-2)^{n+1} + 2\cdot 3^{n+1} & -(-2)^{n+1} + 3^{n+1} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \\ &= \frac{1}{5} \begin{bmatrix} (-1)^n \cdot 2^{n+1} + 3^{n+1} \\ (-1)^{n+1} \cdot 2^{n+2} + 3^{n+2} \end{bmatrix}. \end{aligned}$$

Таким чином, $x_n = \frac{1}{5}((-1)^n \cdot 2^{n+1} + 3^{n+1})$. ■

Додаткові відомості і завдання для самостійної роботи

1. Якщо $p(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ і A — квадратна матриця n-го порядку, то ми можемо означити матрицю $p(A)$ рівністю:

$$p(A) = A^n + a_{n-1}A^{n-1} + \ldots + a_1A + a_0I.$$

Довести, що якщо $p(x)$ — характеристичний многочлен матриці A, то $p(A) \equiv O$ (іншими словами, кожна матриця задовольняє своє характеристичне рівняння). Цей факт відомий як *теорема Гамільтона*[1]*–Келі*[2]. Її доведення можна знайти в монографії Ф. Р. Гантмахера «Теория матриц» (М.: Наука, 1967. — 576 с.).

2. Довести, що нільпотентна матриця не має відмінних від нуля власних значень.
3. Довести, що якщо матриця A зводиться до діагонального виду, то A^T також зводиться до діагонального виду.
4. Довести, що якщо оборотна матриця A зводиться до діагонального виду, то A^{-1} також зводиться до діагонального виду.
5. Нехай A і B — $n \times n$ матриці, кожна з яких має n різних власних значень. Довести, що A і B мають однакові власні вектори тоді і тільки тоді, коли $AB = BA$.
6. Довести, що сума власних підпросторів лінійного оператора є прямою сумою.
7. Нехай A — квадратна матриця, яка має такий блочний вигляд:
$$A = \left[\begin{array}{c|c} P & Q \\ \hline O & S \end{array}\right],$$
де P та S — квадратні матриці, O — нульова. Довести, що характеристичний многочлен матриці A дорівнює добутку характеристичних многочленів матриць P та S.
8. Довести, що матриця $A \in M_n(\mathbb{R})$, яка має n дійсних характеристичних коренів, рахуючи кожен з його кратністю, подібна до деякої трикутної матриці (*теорема Шура*[3]).
9. Довести, що кожне лінійне перетворення дійсного ненульового векторного простору має принаймні один інваріантний підпростір розмірності 1 або 2 ([11, с. 221]).

---

[1] William Rowan Hamilton (1806–1865) — видатний ірландський математик.

[2] Arthur Cayley (1821–1895) — англійський математик.

[3] Issai Schur (1875–1941) — німецький та ізраїльський математик.

Розділ 7

Лінійні оператори в евклідовому просторі

...В ту пору часто, закрыв учебник,
я от амбиций моих ущербных
провозглашал решенным вопрос любой.
И заключал, что двойного смысла
иметь не могут слова и числа,
и пребывал отчаянно горд собой.

Но проходила неделя, две ли,
слова смещались куда хотели,
как А и Б, сидевшие на трубе.
И числа вновь обретали сложность.
И сознавал я свою ничтожность,
и изнывал от ненависти к себе...

М. Щербаков, «После детства»

В цьому розділі ми вивчаємо особливості лінійних операторів у дійсному просторі зі скалярним множенням. Ми введемо поняття оператора, спряженого до даного, ортогонального оператора, а також розглянемо властивості матриць, що їм відповідають, а наприкінці розділу встановимо цікаві властивості симетричних матриць.

§ 1. Поняття оператора, спряженого до даного

1.1. Транспоновані матриці та деякі їх властивості.

Нагадаємо деякі означення з розділу 2.

Означення 7.1. *Транспонованою* до матриці A розміру $m \times n$ називається матриця A^T розміру $n \times m$, стовпці якої є рядками матриці A.

Приклад 7.1. Якщо

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix},$$

то

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}.$$

Означення 7.2. Матриця A називається *симетричною*, якщо $A^T = A$.

Матриця C з попереднього прикладу є симетричною. Матриця $A = [a_{ij}]$ є симетричною тоді і тільки тоді, коли $a_{ij} = a_{ji}$ для усіх i, j.

Властивості операції транспонування матриці з іншими операціями над матрицями висвітлює наступне твердження.

Теорема 7.1. *Нехай A, B — матриці (таких розмірів, щоб операції, які описані нижче, виконувались), $\lambda \in \mathbb{R}$. Тоді:*
1) $(A^T)^T = A$;
2) $(A + B)^T = A^T + B^T$;
3) $(\lambda A)^T = \lambda \cdot A^T$;
4) $(AB)^T = B^T A^T$;
5) якщо A оборотна, то $(A^{-1})^T = (A^T)^{-1}$;
6) $A \cdot A^T$ та $A^T \cdot A$ — симетричні матриці.

Доведення. Властивості 1–3 є достатньо очевидними і доводяться безпосередньо перевіркою.

Доведемо властивість 4. Спочатку переконаємось, що матриці $(AB)^T$ та $B^T A^T$ мають однакову розмірність. Справді, нехай $A \in M_{m\times n}$, $B \in M_{n\times k}$, тоді $AB \in M_{m\times k}$, $(AB)^T \in M_{k\times m}$, а також $A^T \in M_{n\times m}$, $B^T \in M_{k\times n}$ і $B^T A^T \in M_{k\times m}$. Тепер доведемо, що відповідні елементи матриць $(AB)^T$ та $B^T A^T$ однакові. Позначимо через $[A]_{ij}$ — елемент i-го рядка і j-го стовпчика матриці A. Тоді:

$$\begin{aligned}
\left[(AB)^T\right]_{ij} =& [AB]_{ji} = \\
=& (j\text{-ий рядок матриці } A) \cdot (i\text{-ий стовпець матриці } B) = \\
=& (j\text{-ий стовпець матриці } A^T) \cdot (i\text{-ий рядок матриці } B^T) = \\
=& (i\text{-ий рядок матриці } B^T) \cdot (j\text{-ий стовпець матриці } A^T) = \\
=& [B^T A^T]_{ij}
\end{aligned}$$

(тут $\cdot$ позначає скалярний добуток векторів, координати яких є рядки чи стовпці відповідних матриці).

5. Нехай A — оборотна матриця. Це означає, що існує матриця A^{-1} така, що $A^{-1} \cdot A = I$. Тоді $(A^{-1} \cdot A)^T = I^T$. За попереду доведеною властивістю 4 дістаємо $A^T (A^{-1})^T = I$. Остання рівність означає, що матриця $(A^{-1})^T$ є оберненою до матриці A^T, що і потрібно було довести.

6. Справді, на основі доведених вище властивостей 4 та 1, знаходимо:

$$(A \cdot A^T)^T = (A^T)^T \cdot A^T = A \cdot A^T,$$

тобто матриця AA^T є симетричною. Аналогічно знаходимо, що і AA^T також симетрична. ■

1.2. Оператор, спряжений до даного.

Перейдемо до розгляду лінійних операторів, які задані у евклідовому векторному просторі, тобто дійсному векторному просторі, в якому визначене скалярне множення векторів.

Нехай E_n — евклідовий векторний простір, $\mathcal{E} = \big(\vec{e}_1, \ldots, \vec{e}_n\big)$ — ортонормований базис цього простору. Нехай також T — лінійний оператор простору E_n, $A = [a_{ij}]$ — матриця цього лінійного оператора в базисі $\mathcal{E}$.

Означення 7.3. Оператор T^* називається *спряженим* до лінійного оператора T, якщо для довільних $\vec{x}, \vec{y} \in E_n$:

$$\big(T(\vec{x}), \vec{y}\big) = \big(\vec{x}, T^*(\vec{y})\big) \tag{7.1}$$

Спочатку доведемо таке допоміжне твердження.

Лема 7.1. *Якщо T і S — лінійні оператори простору E_n і $\big(\vec{x}, T(\vec{y})\big) = \big(\vec{x}, S(\vec{y})\big)$ для довільних $\vec{x}, \vec{y} \in E_n$, то $T \equiv S$.*

Доведення. Маємо, що

$$\big(\vec{x}, T(\vec{y})\big) - \big(\vec{x}, S(\vec{y})\big) = \big(\vec{x}, (T-S)\vec{y}\big) = 0$$

для всіх $\vec{x}, \vec{y}$, а значить і для $\vec{x} = (T-S)\vec{y}$. Тоді

$$\big((T-S)\vec{y}, (T-S)\vec{y}\big) = 0,$$

звідки $(T-S)\vec{y} = \vec{0}$ для довільного $\vec{y}$, що і означає, що $T \equiv S$. ■

Наслідок. *Кожний лінійний оператор T має не більше одного спряженого.*

Нарешті доведемо, що кожний лінійний оператор справді має спряжений. Нехай T^* — це такий оператор простору E_n, матрицею якого в базисі $\mathcal{E}$ є A^T. Покажемо, що саме він і є спряженим до оператора T.

Знаходимо:

$$\big(T(\vec{e}_i), \vec{e}_j\big) = \big(A\vec{e}_i, \vec{e}_j\big) = \big(a_{1i}\vec{e}_1 + a_{2i}\vec{e}_2 + \cdots + a_{ni}\vec{e}_n, \vec{e}_j\big) = a_{ji},$$
$$\big(\vec{e}_i, T^*(\vec{e}_j)\big) = \big(\vec{e}_i, A^T\vec{e}_j\big) = \big(\vec{e}_i, a_{j1}\vec{e}_1 + a_{j2}\vec{e}_2 + \cdots + a_{jn}\vec{e}_n\big) = a_{ji}$$

для всіх $i, j = \overline{1, n}$ (тут використано аксіоми скалярного множення, а також те, що $\mathcal{E}$ — ортонормований базис і $\big(\vec{e}_i, \vec{e}_j\big) = 0$ при $i \neq j$). Нарешті, якщо

$$\vec{x} = \alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n,$$
$$\vec{y} = \beta_1\vec{e}_1 + \ldots + \beta_n\vec{e}_n,$$

то

$$\begin{aligned}\big(T(\vec{x}), \vec{y}\big) &= \big(T(\alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n), \beta_1\vec{e}_1 + \ldots + \beta_n\vec{e}_n\big) = \\ &= \sum_{i,j=1}^{n} \alpha_i\beta_j\big(T(\vec{e}_i), \vec{e}_j\big) = \sum_{i,j=1}^{n} \alpha_i\beta_j a_{ji}, \\ \big(\vec{x}, T^*(\vec{y})\big) &= \big(\alpha_1\vec{e}_1 + \ldots + \alpha_n\vec{e}_n, T^*(\beta_1 + \ldots + \beta_n\vec{e}_n)\big) = \\ &= \sum_{i,j=1}^{n} \alpha_i\beta_j\big(\vec{e}_i, T^*(\vec{e}_j)\big) = \sum_{i,j=1}^{n} \alpha_i\beta_j a_{ji}.\end{aligned}$$

Таким чином, T^* — спряжений до оператора T. Нашими міркуваннями ми по суті довели таке твердження.

Теорема 7.2. *Для кожного лінійного оператора T евклідового векторного простору E_n існує єдиний спряжений до нього оператор T^*, матриця якого в будь-якому ортонормованому базисі є транспонованою до матриці перетворення T.*

Приклад 7.2. Нехай $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ — оператор повороту на кут φ. Як відомо (див. приклад 5.2), стандартна матриця цього перетворення має вигляд:

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$

Тоді оскільки стандартна матриця є матрицею оператора у ортонормованому базисі і

$$A^T = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} = \begin{bmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{bmatrix},$$

то оператор, спряжений до T — це поворот на кут $-\varphi$.

Властивості спряжених лінійних операторів аналогічні властивостям транспонованих матриць.

Теорема 7.3. *Нехай T та S — лінійні оператори евклідового векторного простору E_n, $\lambda \in \mathbb{R}$. Тоді:*
1) $(T^*)^* = T$;
2) $(T + S)^* = T^* + S^*$;
3) $(\lambda T)^* = \lambda T^*$;
4) $(S \circ T)^* = T^* \circ S^*$;
5) Якщо оператор T невироджений, то спряжений до нього оператор також невироджений, причому $(T^{-1})^* = (T^*)^{-1}$.

1.3. Самоспряжений лінійний оператор.

Означення 7.4. Оператор T називається *самоспряженим*, якщо $T^* = T$.

Іншими словами оператор T є самоспряженим, якщо для довільних векторів $\vec{x}, \vec{y}$ простору E_n $\big(T(\vec{x}), \vec{y}\big) = \big(\vec{x}, T(\vec{y})\big)$.

Серед властивостей самоспряжених операторів відмітимо такі:

1) Тотожній оператор є самоспряженим.
2) Сума самоспряжених операторів є самоспряжений оператор, оскільки якщо $T = T^*$, $S = S^*$, то $(T + S)^* = T^* + S^* = T + S$.
3) Якщо оператор T — самоспряжений, то для довільного $\lambda \in \mathbb{R}$ оператор λT є самоспряженим: якщо $T^* = T$, то $(\lambda T)^* = \lambda T^* = \lambda T$.
4) Оператор, обернений до невиродженого самоспряженого оператора, є самоспряженим: якщо $T^* = T$, то $(T^{-1})^* = (T^*)^{-1} = T^{-1}$.
5) Для того, щоб композиція самоспряжених операторів T і S була самоспряженим оператором необхідно і достатньо, щоб перетворення T і S були перестановочними між собою, тобто, щоб $S \circ T = T \circ S$. Справді, якщо $T = T^*$, $S = S^*$, то $(S \circ T)^* = T^* \circ S^* = T \circ S$, і це

дорівнюватиме $S \circ T$ тоді і тільки тоді, коли T і S перестановочні між собою.

6) Композиція лінійного оператора T і спряженого до нього оператора T^* є самоспряженим лінійним оператором (це випливає з п. 6 теореми 7.1).

Очевидно, що якщо оператор T є самоспряженим, то в ортонормованому базисі його матриця є симетричною.

Виявляється, що усі корені характеристичного многочлена самоспряженого оператора є дійсними. Доведення цього факту ми наведемо нижче (це прямий наслідок теореми 7.15). Крім цього, в подальшому параграфі 3 ми доведемо деякі інші властивості симетричних матриць, з яких слідуватиме, що *матриця самоспряженого перетворення в деякому ортонормованому базисі зводитьcя до діагонального вигляду*. Цей базис, як відомо, буде складатись із власних векторів самоспряженого оператора. Тоді якщо $\vec{v}_1, \ldots, \vec{v}_n$ — цей ортонормований базис для оператора T, і

$$\vec{x} = \alpha_1 \vec{v}_1 + \ldots + \alpha_n \vec{v}_n,$$

то

$$T(\vec{x}) = \alpha_1 \lambda_1 \vec{v}_1 + \ldots + \alpha_n \lambda_n \vec{v}_n.$$

Цей факт дає геометричний зміст самоспряженого оператора: в ортонормованому базисі, що складається з його власних векторів, дія лінійного оператора зводиться до n розтягів вздовж базисних векторів з коефіцієнтами $\lambda_1, \ldots, \lambda_n$, де $\lambda_1, \ldots, \lambda_n$ — відповідні власні значення. На рисунку 7.1 зображено дію в евклідовому просторі $\mathbb{R}^2$ на фігуру F самоспряженого оператора з власними значеннями $\lambda_1 = 2$, $\lambda_2 = \frac{1}{2}$.

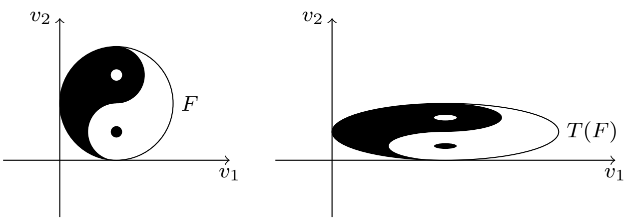


Рис. 7.1.

§ 2. Ортогональні лінійні оператори

2.1. Ортогональні матриці та деякі їх властивості.

Розглянемо матриці, стовпці яких утворюють ортонормовану систему векторів.

Приклад 7.3. Стовпці матриці $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \\ 1 & -5 \end{bmatrix}$ — ортогональні вектори. Для побудови матриці з ортонормованими стовпцями можна пронормувати стовпці матриці A:

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \|\vec{u}_1\| = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6}; \quad \vec{u}_1' = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix};$$

$$\vec{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -5 \end{bmatrix}, \quad \|\vec{u}_2\| = \sqrt{(-1)^2 + 3^2 + (-5)^2} = \sqrt{35}; \quad \vec{u}_2' = \frac{1}{\sqrt{35}} \begin{bmatrix} -1 \\ 3 \\ -5 \end{bmatrix}.$$

Матриця $A_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{35}} \\ \frac{2}{\sqrt{6}} & \frac{3}{\sqrt{35}} \\ \frac{1}{\sqrt{6}} & -\frac{5}{\sqrt{35}} \end{bmatrix}$ має ортонормовані стовпці.

Легко переконатися в тому, що $A_1^T \cdot A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Виявляється, що це не випадково.

Теорема 7.4. *Стовпці матриці $Q \in M_{m\times n}$ формують ортонормовану систему векторів тоді і тільки тоді, коли $Q^T Q = I_n$.*

Доведення. *Необхідність.* Нехай $Q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_n \end{bmatrix}$, причому для всіх $i, j \in \{1, 2, \dots, n\}$

$$(\vec{q}_i, \vec{q}_j) = \vec{q}_i^{\,T} \cdot \vec{q}_j = \begin{cases} 0, & \text{якщо } i \neq j \\ 1, & \text{якщо } i = j. \end{cases}$$

Тоді $Q^T = \begin{bmatrix} \vec{q}_1^{\,T} \\ \vec{q}_2^{\,T} \\ \vdots \\ \vec{q}_n^{\,T}\cdot \end{bmatrix}$ Маємо:

$$Q^T \cdot Q = \begin{bmatrix} \vec{q}_1^{\,T} \\ \vdots \\ \vec{q}_n^{\,T}\cdot \end{bmatrix} \cdot \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_n \end{bmatrix} =$$

$$= \begin{bmatrix} \vec{q}_1^{\,T}\vec{q}_1 & \vec{q}_1^{\,T}\vec{q}_2 & \dots & \vec{q}_1^{\,T}\vec{q}_n \\ \vec{q}_2^{\,T}\vec{q}_1 & \vec{q}_2^{\,T}\vec{q}_2 & \dots & \vec{q}_2^{\,T}\vec{q}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{q}_n^{\,T}\vec{q}_1 & \vec{q}_n^{\,T}\vec{q}_2 & \dots & \vec{q}_n^{\,T}\vec{q}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Достатність доводиться аналогічними міркуваннями в зворотному порядку. ■

Означення 7.5. Квадратна матриця Q називається *ортогональною*, якщо її стовпці утворюють ортонормовану систему векторів.

З теореми 7.4 безпосередньо одержуємо таке твердження.

Теорема 7.5. *Для того, щоб матриця Q була ортогональною необхідно і достатньо, щоб $Q^{-1} = Q^T$ (тобто $Q^TQ = I$).*

Приклад 7.4. Матриці

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{та} \quad B = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

є ортогональними, тому

$$A^{-1} = A^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{та} \quad B^{-1} = B^T = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}.$$

Наступна теорема встановлює характеристичні властивості ортогональних матриць.

Теорема 7.6. *Нехай Q — матриця n-го порядку. Наступні твердження є еквівалентними:*

1) Q — ортогональна матриця.

2) $\|Q\vec{x}\| = \|\vec{x}\|$ *для довільного* $\vec{x} \in \mathbb{R}^n$*.*
3) $\big(Q\vec{x}, Q\vec{y}\big) = \big(\vec{x}, \vec{y}\big)$ *для довільних* $\vec{x}, \vec{y} \in \mathbb{R}^n$*.*

Доведення. Доведення будемо проводити таким чином: покажемо, що з 1) слідує 3), з 3) — 2), з 2) — 1). Тим самим буде доведена рівносильність вказаних тверджень.

- 1) $\Rightarrow$ 3).

  Нехай матриця Q є ортогональною. Тоді

$$\big(Q\vec{x}, Q\vec{y}\big) = (Q\vec{x})^T \cdot Q\vec{y} = \vec{x}^T \cdot Q^T Q \cdot \vec{y} = \vec{x}^T \cdot I \cdot \vec{y} = \vec{x}^T \cdot \vec{y} = \big(\vec{x}, \vec{y}\big).$$

- 3) $\Rightarrow$ 2). Нехай для довільних $\vec{x}, \vec{y} \in \mathbb{R}^n$ виконується рівність $\big(Q\vec{x}, Q\vec{y}\big) = \big(\vec{x}, \vec{y}\big)$. Покладемо $\vec{y} = \vec{x}$. Тоді $\big(Q\vec{x}, Q\vec{x}\big) = \big(\vec{x}, \vec{x}\big)$, звідки $\|Q\vec{x}\| = \|\vec{x}\|$.
- 2) $\Rightarrow$ 1). Спочатку покажемо, що

$$\big(\vec{x}, \vec{y}\big) = \frac{1}{4}\left(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2\right).$$

  Справді,

$$\frac{1}{4}\big(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2\big) = \frac{1}{4}\left((\vec{x}+\vec{y})^2 - (\vec{x}-\vec{y})^2\right) =$$
$$= \frac{1}{4}\Big(\big(\vec{x},\vec{x}\big) + 2\big(\vec{x},\vec{y}\big) + \big(\vec{y},\vec{y}\big) - \big(\vec{x},\vec{x}\big) + 2\big(\vec{x},\vec{y}\big) - \big(\vec{y},\vec{y}\big)\Big) =$$
$$= \big(\vec{x},\vec{y}\big).$$

  Тоді для довільних $\vec{x}, \vec{y} \in \mathbb{R}^n$ з того, що $\|Q \cdot \vec{x}\| = \|\vec{x}\|$:

$$\big(\vec{x},\vec{y}\big) = \frac{1}{4}\left(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2\right) = \frac{1}{4}\left(\|Q(\vec{x}+\vec{y})\|^2 - \|Q(\vec{x}-\vec{y})\|^2\right) =$$
$$= \frac{1}{4}\left((Q\vec{x}+Q\vec{y}))^2 - (Q\vec{x}-Q\vec{y}))^2\right) = \frac{1}{4}\Big(\big(Q\vec{x},Q\vec{x}\big) + 2\big(Q\vec{x},Q\vec{y}\big) +$$
$$+ \big(Q\vec{y},Q\vec{y}\big) - \big(Q\vec{x},Q\vec{x}\big) + 2\big(Q\vec{x},Q\vec{y}\big) - \big(Q\vec{y},Q\vec{y}\big)\Big) = \big(Q\vec{x},Q\vec{y}\big)$$

  (чим ми фактично довели імплікацію 2) $\Rightarrow$ 3)).

  Нехай $\vec{q}_i$ — i-ий стовпець матриці Q. Тоді $\vec{q}_i = Q \cdot \vec{e}_i$, де $\vec{e}_i$ — i-ий стовпець одиничної матриці. Далі, для $i, j \in \{1, \ldots, n\}$:

$$\big(\vec{q}_i, \vec{q}_j\big) = \big(Q\vec{e}_i, Q\vec{e}_j\big) = \big(\vec{e}_i, \vec{e}_j\big) = \begin{cases} 0, & \text{якщо } i \neq j, \\ 1, & \text{якщо } i = j, \end{cases}$$

  що і означає, що матриця Q ортогональна.

■

Виявляється, що у ортогональної матриці вектори-рядки також утворюють ортонормовану систему векторів.

Теорема 7.7. *Якщо матриця Q ортогональна, то її рядки утворюють ортонормовану систему векторів.*

Доведення. Справді, за теоремою 7.5 $Q^{-1} = Q^T$. Тоді:

$$(Q^T)^{-1} = (Q^{-1})^{-1} = Q = (Q^T)^T,$$

а значить за тією ж теоремою 7.5 матриця Q^T також ортогональна і її стовпці (а вони є рядками матриці Q) утворюють ортонормовану систему векторів. ■

Деякі властивості ортогональних матриць відображені у наступному твердженні.

Теорема 7.8. *Нехай Q — ортогональна матриця. Тоді:*
1. Q^{-1} також ортогональна.
2. $\det Q = \pm 1$.
3. Якщо λ — власне значення матриці Q, то $|\lambda| = 1$.
4. Якщо Q_1 і Q_2 — ортогональні матриці, то матриця Q_1Q_2 також ортогональна.

Доведення. Перше твердження достатньо очевидне, оскільки з попередньої теореми випливає, що матриця Q^T є ортогональною, а $Q^{-1} = Q^T$ за теоремою 7.5.

Далі, знаходимо: $Q^T \cdot Q = I$. Перейшовши до визначників, одержуємо:

$$\det(Q^TQ) = \det Q^T \cdot \det Q = (\det Q)^2 = 1,$$

звідки знаходимо, що $\det Q = \pm 1$.

Нехай $\vec{u}$ — власний вектор матриці Q, що відповідає власному значенню λ. Тоді $Q\vec{u} = \lambda\vec{u}$. Знаходимо: $\|Q\vec{u}\| = \|\lambda\vec{u}\|$. З п. 2 теореми 7.6 і властивостей норми вектора, одержуємо

$$\|Q\vec{u}\| = \|\vec{u}\| = |\lambda| \cdot \|\vec{u}\|,$$

звідки $|\lambda| = 1$.

Нарешті нехай Q_1 і Q_2 — дві ортогональні матриці однакового порядку. Тоді $Q_1^TQ = I$, $Q_2^TQ = I$. На основі властивостей транспонованих матриць (теорема 7.1, п. 4) знаходимо:

$$(Q_1Q_2)^T \cdot Q_1Q_2 = Q_2^TQ_1^T \cdot Q_1Q_2 = Q_2^T(Q_1^T \cdot Q_1)Q_2 = Q_2^TIQ_2 = Q_2^TQ_2 = I.$$

Тоді за критерієм ортогональності матриці (теорема 7.5) стверджуємо, що Q_1Q_2 — ортогональна матриця. ■

2.2. Ортогональні матриці другого порядку.

Наступне твердження описує, який вигляд мають ортогональні матриці другого порядку.

Теорема 7.9. *Ортогональні матриці другого порядку мають вигляд*

$$\begin{bmatrix} a & b \\ b & -a \end{bmatrix} \quad \text{або} \quad \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

де $\begin{bmatrix} a \\ b \end{bmatrix}$ *— одиничний вектор.*

Доведення. Справді, нехай маємо ортогональну матрицю $\begin{bmatrix} a & x \\ b & y \end{bmatrix}$. Тоді виконуються такі умови:

$$\begin{cases} a^2 + b^2 = 1, \\ x^2 + y^2 = 1, \\ ax + by = 0. \end{cases}$$

Розглянемо два випадки. Якщо $a = 0$, то з першого рівняння $b = \pm 1$, а тоді з третього $y = 0$ і з другого $x = \pm 1$. Тобто в цьому випадку ортогональні матриці мають вигляд

$$\begin{bmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{bmatrix},$$

що цілком узгоджується з твердженням теореми.

Нехай $a \neq 0$. Тоді $x = -\frac{by}{a}$. Підставивши одержане значення в друге рівняння системи, знаходимо $\left(-\frac{by}{a}\right)^2 + y^2 = 1$, або, після спрощення з використанням того, що $a^2 + b^2 = 1$, дістанемо: $y^2 = a^2$. Отже, або $y = a$ (тоді $x = -b$), або $y = -a$ (тоді $x = b$). ■

Наслідок. *Кожна ортогональна матриця другого порядку має вигляд*

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \quad \text{або} \quad B = \begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}.$$

Матриця A задає перетворення повороту простору $\mathbb{R}^2$ на кут φ, а матриця B — композицію симетрії відносно першого координатного *вектора та повороту на кут* φ*. При цьому зауважимо, що* $\det A = 1$, $\det B = -1$.

2.3. Ортогональні лінійні оператори та їх властивості. Нехай E_n — евклідовий векторний простір, T — лінійний оператор цього простору.

Означення 7.6. Лінійний оператор T називається *ортогональним*, якщо він зберігає скалярний добуток векторів, тобто для довільних $\vec{x}, \vec{y} \in E_n$ $\big(T(\vec{x}), T(\vec{y})\big) = \big(\vec{x}, \vec{y}\big)$.

Теорема 7.10. *Ортогональний лінійний оператор зберігає довжини векторів та кут між ними.*

Доведення. Справді,

$$\|T(\vec{x})\|^2 = \big(T(\vec{x}), T(\vec{x})\big) = \big(\vec{x}, \vec{x}\big) = \|\vec{x}\|^2,$$

звідки $\|T(\vec{x})\| = \|\vec{x}\|$. Тоді також:

$$\cos(T(\vec{x})\widehat{,}T(\vec{y})) = \frac{\big(T(\vec{x}), T(\vec{y})\big)}{\|T(\vec{x})\|\|T(\vec{y})\|} = \frac{\big(\vec{x}, \vec{y}\big)}{\|\vec{x}\|\|\vec{y}\|} = \cos(\vec{x}\widehat{,}\vec{y}).$$

■

Зберігання довжин векторів є характеристичною властивістю ортогонального оператора, тому його ще називають *ізометрією* (від грец. isos $\iota\sigma o\zeta$ — рівний, $\mu\epsilon\tau\rho\epsilon\omega$ — виміряти).

Теорема 7.11. *Матриця ортогонального лінійного оператора в ортонормованому базисі є ортогональною.*

Доведення. Нехай $\mathcal{E} = \big(\vec{e}_1, \dots, \vec{e}_n\big)$ — ортонормований базис простору E_n, T — ортогональний лінійний оператор цього простору. Нехай A — матриця оператора T в базисі $\mathcal{E}$:

$$A = \begin{bmatrix} T(\vec{e}_1) & \dots & T(\vec{e}_n) \end{bmatrix}.$$

Розглянемо добуток:

$$A^T \cdot A = \begin{bmatrix} (T(\vec{e}_1))^T \\ \vdots \\ (T(\vec{e}_n))^T \end{bmatrix} \begin{bmatrix} T(\vec{e}_1) & \dots & T(\vec{e}_n) \end{bmatrix} =$$

$$= \begin{bmatrix} (T(\vec{e}_1))^T \cdot T(\vec{e}_1) & (T(\vec{e}_1))^T \cdot T(\vec{e}_2) & \dots & (T(\vec{e}_1))^T \cdot T(\vec{e}_n) \\ (T(\vec{e}_2))^T \cdot T(\vec{e}_1) & (T(\vec{e}_2))^T \cdot T(\vec{e}_2) & \dots & (T(\vec{e}_2))^T \cdot T(\vec{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ (T(\vec{e}_n))^T \cdot T(\vec{e}_1) & (T(\vec{e}_n))^T \cdot T(\vec{e}_2) & \dots & (T(\vec{e}_n))^T \cdot T(\vec{e}_n) \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = I,$$

звідки за теоремою 7.5 знаходимо, що матриця A є ортогональною. ■

Властивості ортогональних лінійних операторів аналогічні властивостям ортогональних матриць:

Теорема 7.12. *Нехай T — ортогональний лінійний оператор евклідового векторного простору. Тоді він невироджений, причому T^{-1} — також ортогональний.*

Теорема 7.13. *Якщо λ — власне значення ортогонального лінійного оператора T, то $|\lambda| = 1$.*

Теорема 7.14. *Композиція двох ортогональних лінійних операторів є ортогональний оператор.*

§ 3. Діагоналізація симетричних матриць

В параграфі 3 розділу 6 ми досліджували умови, за яких матрицю A можна подати у вигляді добутку PDP^{-1}, де D — діагональна матриця, і встановили критерій того, коли це можливо (див. теорему 6.8), а також вигляд матриць P та D (див. алгоритм діагоналізації матриці).

Розглянемо на прикладі випадок, коли матриця A є симетричною.

Приклад 7.5. Якщо можливо, діагоналізувати матрицю
$$A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}.$$

Розв'язання. Знайдемо власні значення матриці A. Характеристичне рівняння $\det(A - \lambda I) = 0$ після спрощень набуває вигляду
$$(\lambda - 3)(\lambda - 6)(\lambda - 8) = 0,$$
а тому власними значеннями є $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 8$. Оскільки матриця A третього порядку має три різних власних значення, то вона зводиться до діагонального виду (див. теореми 6.7 та 6.8).

В стандартний спосіб знаходимо власні вектори, що відповідають власним значенням. Власним значенням $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 8$ відповідають відповідно власні вектори
$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{u}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \quad \vec{u}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Отже, в якості матриці P можна взяти матрицю
$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 2 & 0 \end{bmatrix}.$$

Залишається складність в обчисленні матриці P^{-1}. Але помітимо, що власні вектори попарно ортогональні. Тому пронормувавши вектори $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ ми можемо з одержаних векторів скласти ортогональну матрицю P', обернена до якої легко будується (оскільки є транспонованою до P'):
$$\vec{u}'_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{u}'_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \quad \vec{u}'_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix};$$
$$P' = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \end{bmatrix}.$$

Враховуючи, що $(P')^{-1} = P'^T$, дістаємо таке подання симетричної матриці A: $A = P'D(P')^{-1} = P'DP'^T$, де $D = \operatorname{diag}\{3, 6, 8\}$. ■

Одержане подання матриці з останнього прикладу приводить до такого нового поняття.

Означення 7.7. Квадратна матриця A називається *ортогонально діагоналізовною*, якщо існують ортогональна матриця Q і діагональна матриця D такі, що

$$A = QDQ^T.$$

Зазначимо також, що остання рівність може бути записана у вигляді $D = Q^T AQ$.

Виявляється, що лише симетричні матриці є ортогонально діагоналізовними. Але перед доведенням цього твердження спочатку відмітимо деякі цікаві властивості симетричних матриць.

Теорема 7.15. *Власні значення дійсної симетричної матриці є дійсними числами.*

Доведення. Ідея доведення ґрунтується на наступному. Комплексне число $z = a + bi$ є дійсним тоді і тільки тоді, коли воно співпадає із комплексно спряженим числом $\overline{z} = a - bi$ (бо тоді $2bi = 0$, тобто $b = 0$). Далі використаємо таке позначення: якщо $A = [a_{ij}]$, через $\overline{A}$ позначимо матрицю з комплексно спряженими елементами ($\overline{A} = [\overline{a}_{ij}]$). Зрозуміло, що якщо матриця A дійсна, то $A = \overline{A}$. Можна переконатись в тому, що $\overline{AB} = \overline{A} \cdot \overline{B}$.

Нехай λ — власне значення дійсної симетричної матриці A, що відповідає власному вектору $\vec{v}$. Тоді $A\vec{v} = \lambda\vec{v}$, $\overline{A\vec{v}} = \overline{\lambda\vec{v}}$. Але тоді, оскільки матриця A дійсна, то

$$A\overline{\vec{v}} = \overline{A}\overline{\vec{v}} = \overline{A\vec{v}} = \overline{\lambda\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}.$$

Транспонуємо останню рівність:

$$\overline{\vec{v}}^T A = \overline{\vec{v}}^T A^T = (A\overline{\vec{v}})^T = (\overline{\lambda}\,\overline{\vec{v}})^T = \overline{\lambda}\overline{\vec{v}}^T.$$

Тому

$$\lambda(\overline{\vec{v}}^T \vec{v}) = \overline{\vec{v}}^T (\lambda\vec{v}) = \overline{\vec{v}}^T (A\vec{v}) = (\overline{\vec{v}}^T A)\vec{v} = (\overline{\lambda}\overline{\vec{v}}^T)\vec{v} = \overline{\lambda}(\overline{\vec{v}}^T \vec{v}),$$

або $(\lambda - \overline{\lambda})(\overline{\vec{v}}^T \vec{v}) = 0$.

Тепер якщо $\vec{v} = \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix}$, то $\overline{\vec{v}} = \begin{bmatrix} a_1 - b_1 i \\ \vdots \\ a_n - b_n i \end{bmatrix}$, а тому

$$\overline{\vec{v}}^T \vec{v} = (a_1^2 + b_1^2) + \ldots + (a_n^2 + b_n^2) \neq 0,$$

оскільки $\vec{v} \neq \vec{0}$. Отже, $\lambda = \overline{\lambda}$ і λ — дійсне число. ■

Теорема 7.16. *Якщо матриця A симетрична, то будь-які два власні вектори, яким відповідають різні власні значення, є ортогональними.*

Доведення. Нехай A — симетрична матриця, λ_1, λ_2 — її різні власні значення, $\vec{u}_1, \vec{u}_2$ — власні вектори, що цим власним значенням відповідають:

$$A\vec{u}_1 = \lambda_1\vec{u}_1,$$
$$A\vec{u}_2 = \lambda_2\vec{u}_2.$$

Нам потрібно показати, що $\big(\vec{u}_1, \vec{u}_2\big) = 0$. Розглянемо таке скалярне множення:

$$\begin{aligned}\big(\lambda_1\vec{u}_1, \vec{u}_2\big) &= \big(A\vec{u}_1, \vec{u}_2\big) = (A\vec{u}_1)^T \cdot \vec{u}_2 = \vec{u}_1^T A^T \cdot \vec{u}_2 = \\ &= \vec{u}_1^T \cdot (A^T\vec{u}_2) = \vec{u}_1^T \cdot (A\vec{u}_2) = \vec{u}_1^T \cdot (\lambda_2\vec{u}_2) = \\ &= \big(\vec{u}_1, \lambda_2\vec{u}_2\big).\end{aligned}$$

Отже, $\big(\lambda_1\vec{u}_1, \vec{u}_2\big) = \big(\vec{u}_1, \lambda_2\vec{u}_2\big)$, звідки $(\lambda_1 - \lambda_2)\big(\vec{u}_1, \vec{u}_2\big) = 0$. А оскільки $\lambda_1 \neq \lambda_2$, то $\big(\vec{u}_1, \vec{u}_2\big) = 0$, а це і означає, що вектори $\vec{u}_1, \vec{u}_2$ є ортогональними. ■

Теорема 7.17. *Якщо матриця є ортогонально діагоналізовною, то вона симетрична.*

Доведення. Нехай A — ортогонально діагоналізовна матриця. Це означає, що існують ортогональна матриця Q і симетрична D такі, що $A = QDQ^T$. Використовуючи властивості транспонованих матриць (теорема 7.1), знаходимо, що

$$A^T = (QDQ^T)^T = (Q^T)^T \cdot (QD)^T = (Q^T)^T \cdot D^T Q^T = QDQ^T = A.$$

Отже, A — симетрична. ■

Виявляється, що обернене твердження також має місце.

Теорема 7.18. *Кожна симетрична матриця $A \in M_n(\mathbb{R})$ ортогонально діагоналізовна.*

Доведення. Доведення проведемо методом математичної індукції по n. Для $n = 1$ нічого доводити, оскільки 1×1 матриця є симетричною. Припустимо, що кожна $k \times k$ дійсна симетрична матриця ортогонально діагоналізовна. Нехай $n = k + 1$ і A — дійсна симетрична матриця n-го порядку.

Нехай λ_1 — одне з власних значень матриці A. Воно дійсне за теоремою 7.15, і йому відповідає дійсний власний вектор $\vec{v}_1$ (подумайте, чому). Можна вважати вектор $\vec{v}_1$ одиничним, оскільки його завжди можна пронормувати. Використовуючи процес Грама–Шмідта можна побудувати ортонормований базис $\left(\vec{v}_1, \ldots, \vec{v}_n\right)$ простору $\mathbb{R}^n$. Складемо матрицю $Q_1 = \begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n\end{bmatrix}$. Q_1 — ортогональна матриця, а тому

$$Q_1^T A Q_1 = \begin{bmatrix}\vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_n^T\end{bmatrix} A \begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n\end{bmatrix} = \begin{bmatrix}\vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_n^T\end{bmatrix} \begin{bmatrix}A\vec{v}_1 & A\vec{v}_2 & \ldots & A\vec{v}_n\end{bmatrix} =$$

$$= \begin{bmatrix}\vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_n^T\end{bmatrix} \begin{bmatrix}\lambda_1\vec{v}_1 & A\vec{v}_2 & \ldots & A\vec{v}_n\end{bmatrix} = \left[\begin{array}{c|c}\lambda_1 & \star \\ \hline O & A_1\end{array}\right] = B,$$

оскільки $\vec{v}_1^T(\lambda_1\vec{v}_1) = \lambda_1(\vec{v}_1^T\vec{v}_1) = \lambda_1$, а $\vec{v}_k^T(\lambda_1\vec{v}_1) = \lambda_1(\vec{v}_k^T\vec{v}_1) = 0$ для $k > 1$.

Але

$$B^T = (Q_1^T A Q_1)^T = Q_1^T A^T (Q_1^T)^T = Q_1^T A Q_1 = B,$$

тому B — симетрична і має такий блочний вигляд

$$B = \left[\begin{array}{c|c}\lambda_1 & O \\ \hline O & A_1\end{array}\right],$$

з симетричною матрицею A_1. Також відмітимо, що матриця A_1 з дійсними елементами. Таким чином, A_1 є $k \times k$ дійсною симетричною матрицею, тому за припущенням індукції її можна ортогонально діагоналізувати. Це означає, що існує ортогональна матриця P_2 така, що матриця $P_2^T A_1 P_2$ є діагональною, скажімо, D_1. Розглянемо матрицю

$$Q_2 = \left[\begin{array}{c|c}1 & O \\ \hline O & P_2\end{array}\right].$$

Це ортогональна матриця $k+1$-го порядку, а тому такою також є і матриця $Q = Q_1Q_2$. Тоді

$$Q^T AQ = (Q_1Q_2)^T A(Q_1Q_2) = (Q_2^T Q_1^T)A(Q_1Q_2) = Q_2^T(Q_1^T AQ_1)Q_2 =$$
$$= Q_2^T BQ_2 = \left[\begin{array}{c|c} 1 & O \\ \hline O & P_2^T \end{array}\right]\left[\begin{array}{c|c} \lambda_1 & O \\ \hline O & A_1 \end{array}\right]\left[\begin{array}{c|c} 1 & O \\ \hline O & P_2 \end{array}\right] =$$
$$= \left[\begin{array}{c|c} 1 & O \\ \hline O & P_2^T A_1 P_2 \end{array}\right] = \left[\begin{array}{c|c} 1 & O \\ \hline O & D_1 \end{array}\right],$$

яка є діагональною матрицею. Це завершує крок індукції і можна стверджувати, що для всіх $n \geqslant 1$ дійсна симетрична матриця порядку n є ортогонально діагоналізовною. ■

Останні теореми об'єднані в лінійній алгебрі під назвою «спектальна теорема для симетричних матриць». Взагалі *спектром матриці* називається множина її власних значень.

Теорема 7.19 (Спектральна теорема для симетричних матриць)**.** *Нехай A — дійсна симетрична матриця n-го порядку. Тоді:*

1. A має n дійсних власних значень, враховуючи кратність.

2. A ортогонально діагоналізовна.

3. Якщо λ — власне значення кратності k, то цьому власному значенню відповідає k попарно ортогональних власних векторів (власний підпростір має ортонормований базис, що складається з власних векторів).

4. Власні підпростори, що відповідають різним власним значенням, попарно ортогональні.

Приклад 7.6. Ортогонально діагоналізувати симетричну матрицю

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}.$$

Розв'язання. Характеристичне рівняння матриці A має вигляд

$$-(\lambda - 7)^2 \cdot (\lambda + 2) = 0.$$

Звичайне обчислення власних векторів матриці приводить до таких результатів:

1) власному значенню $\lambda_1 = -2$ відповідає власний вектор $\vec{u}_1 = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$;

2) власному значенню $\lambda_2 = 7$ відповідають, наприклад, такі лінійно незалежні вектори $\vec{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{u}_3 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$.

Ми бачимо, що $\vec{u}_1 \perp \vec{u}_2$ і $\vec{u}_1 \perp \vec{u}_3$ (це узгоджується з теоремою 7.16). Але $\vec{u}_2$ не ортогональний до $\vec{u}_3$. Це пояснюється тим, що обрані нами вектори $\vec{u}_2$ та $\vec{u}_3$ формують базис власного підпростору, що відповідає власному значенню 7, але цей базис не ортогональний. Для знаходження матриці Q слід знайти ортонормований базис цього простору. Це можна досягти застосувавши процес ортогоналізації до векторів $\vec{u}_2$ та $\vec{u}_3$:

$$\vec{u}_3' = \vec{u}_3 - \frac{(\vec{u}_2, \vec{u}_3)}{(\vec{u}_2, \vec{u}_2)}\vec{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} - \frac{-1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 2 \\ \frac{1}{2} \end{bmatrix},$$

тобто вектор $\vec{u}_3$ можна замінити вектором $2\vec{u}_3' = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$. Після цього власні вектори стали попарно ортогональними. Пронормувавши їх, складемо матрицю Q:

$$\vec{q}_1 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}; \quad \vec{q}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}; \quad \vec{q}_3 = \frac{1}{3\sqrt{2}}\begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix};$$

$$Q = \begin{bmatrix} -\frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} \\ -\frac{1}{3} & 0 & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} \end{bmatrix}.$$

Матриця D має вигляд

$$D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

Отже, $A = QDQ^T$. ■

§ 4. Подання невиродженої матриці у вигляді добутку ортогональної та симетричної матриць

Подання матриць у вигляді добутку матриць специфічного типу (так звані *факторизації матриць*) знаходять численні застосування. Нами вже розглянуто питання про можливість подання матриці у формах PDP^{-1} та QDQ^T, де D — діагональна, Q — ортогональна матриці. Виявляється, що кожна невироджена матриця може бути подана у вигляді добутку ортогональної та симетричної матриць.

Теорема 7.20. *Кожну невироджену матрицю можна подати у вигляді добутку ортогональної та симетричної матриць.*

Доведення. Нехай $A \in M_n(\mathbb{R})$ — невироджена матриця. Доведення теореми конструктивне: ми покажемо, як побудувати ортогональну матрицю Q і симетричну матрицю B такі, що $A = QB$.

1. Побудуємо матрицю A^TA. На основі п. 6 теореми 7.1 стверджуємо, що ця матриця є симетричною, а тому за теоремою 7.18 її можна подати у вигляді $A^TA = P \cdot C \cdot P^T$, де C — діагональна матриця, на головній діагоналі якої знаходяться власні значення матриці A^TA. Виявляється, що вони завжди додатні. Доведемо це.

   Нехай λ — власне значення матриці A^TA, $\vec{u}$ — нормований власний вектор, що відповідає цьому власному значенню: $(A^TA)\vec{u} = \lambda\vec{u}$. Розглянемо скалярний добуток $\big(\vec{u}, (A^TA)\vec{u}\big)$ під двома кутами зору:

$$\begin{aligned}
\big(\vec{u}, (A^TA)\vec{u}\big) &= \big(\vec{u}, \lambda\vec{u}\big) = \lambda\big(\vec{u}, \vec{u}\big) = \lambda;\\
\big(\vec{u}, (A^TA)\vec{u}\big) &= \big(\vec{u}, A^T(A\vec{u})\big) = \vec{u}^T \cdot (A^T(A\vec{u})) =\\
&= (\vec{u}^TA^T) \cdot (A\vec{u})) = (A\vec{u})^T \cdot (A\vec{u})) = \big(A\vec{u}, A\vec{u}\big) \geqslant 0.
\end{aligned}$$

   Отже, $\lambda \geqslant 0$. А оскільки матриця A невироджена, то і матриця A^TA є також невиродженою (бо $\det(A^TA) = \det A^T \cdot \det A = (\det A)^2 \neq 0$), а тому 0 не може бути її власним значенням (див. теорему 6.5). Таким чином, $\lambda > 0$.

   Якщо $\lambda_1, \ldots, \lambda_n$ — власні значення матриці A^TA, то ми можемо побудувати значення $\sigma_k = \sqrt{\lambda_k}$, $k = \overline{1, n}$. Їх прийнято називати *сингулярними* значеннями матриці A.

2. Із сингулярних значень матриці A складемо діагональну матрицю $D = \mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$. Очевидно, що $D^2 = C$.
3. Утворимо матрицю $B = PDP^T$. Вона симетрична, оскільки, як бачимо, ортогонально діагоналізовна (див. теорему 7.17). Крім цього,

$$B^2 = PDP^T \cdot PDP^T = PD(P^{-1}P)DP^T = PD^2P^T = PCP^T = A^TA,$$

   і, отже, B неособлива.
4. Сконструюємо матрицю $Q = A \cdot B^{-1}$. Доведемо, що вона ортогональна. Для цього достатньо показати, що $Q^TQ = I$ (див. теорему 7.5):

$$Q^TQ = (AB^{-1})^T \cdot AB^{-1} = (B^{-1})^TA^T \cdot AB^{-1} = (B^{-1})^T(A^T \cdot A)B^{-1} =$$
$$= (B^T)^{-1}(B \cdot B)B^{-1} = (B)^{-1} \cdot B = I$$

   (тут ми використали різні властивості транспонування матриць — див. теорему 7.1).
5. Таким чином, шукані матриці Q (ортогональна) та B (симетрична) побудовано, оскільки $A = Q \cdot B$.

■

Доведення теореми фактично дає алгоритм подання невиродженої матриці A n-го порядку у вигляді добутку ортогональної та симетричної:

1) Складаємо матрицю A^TA і ортогонально діагоналізуємо її: $A^TA = PCP^T$.
2) Знаходимо сингулярні значення σ_k, $k = \overline{1,n}$ матриці A. Для цього знаходимо власні значення σ_k^2 матриці A^TA. Складаємо матрицю $D = \mathrm{diag}\{\sigma_1, \ldots, \sigma_n\}$.
3) Формуємо матрицю $B = PDP^T$.
4) Знаходимо B^{-1}.
5) Знаходимо матрицю $Q = AB^{-1}$.

Тоді $A = QB$ — шукане подання.

В силу ізоморфізму між алгебрами матриць та лінійних операторів (теорема 5.6) можна стверджувати, що кожний невироджений лінійний оператор евклідового простору можна подати у вигляді композиції самоспряженого і ортогонального операторів.

Додаткові відомості і завдання для самостійної роботи

1. Довести, що якщо симетрична матриця є оборотною, то обернена до неї також симетрична.
2. Довести, що для симетричних матриць A, B матриця AB є симетричною тоді і тільки тоді, коли $AB = BA$. (Довести, що композиція двох самоспряжених операторів є самоспряженим оператором тоді і тільки тоді, коли вони переставовочні).
3. Квадратна матриця W називається *кососиметричною*, якщо $W^T = -W$. Нехай A — довільна матриця.
   а) Довести, що $A - A^T$ — кососиметрична матриця; б) знайти симетричну матрицю S і кососиметричну W таку, що $A = S + W$; в) довести, що матриці S та W такі, що $A = S + W$, однозначно визначаються матрицею A.
4. Доведіть, що образ лінійного оператора T евклідового простору співпадає з ортогональним доповненням ядра оператора, спряженого до нього:
$$\operatorname{Im} T = (\operatorname{Ker} T^*)^{\perp}.$$
5. Довести, що у спряжених один одному операторах евклідового простору співпадають: 1) ранги; 2) характеристичні многочлени; 3) власні значення.

Розділ 8

Білінійні та квадратичні форми

И я выхожу из пространства
В запущенный сад величин
И мнимое рву постоянство
И самосознанье причин

О. Мандельштам. Восьмистишия

§ 1. Лінійні, білінійні та квадратичні форми

1.1. Лінійні функції векторного аргументу.

Означення 8.1. Нехай V — деякий векторний простір. Кожне відображення $f\colon V \to \mathbb{R}$ називається *числовою функцією векторного аргументу*.

Приклад 8.1. Числовими функціями векторного аргументу є, наприклад, такі відображення:
1) $f\colon \mathbb{R}^2 \to \mathbb{R}$:

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 - 3x_2.$$

2) $f\colon \mathbb{R}^4 \to \mathbb{R}$:

$$f(\vec{x}) = 1.$$

Означення 8.2. Числова функція векторного аргументу f, визначена на скінченновимірному векторному просторі V, називається *лінійною*, якщо виконується умова

$$f(\lambda_1\vec{x}_1 + \lambda_2\vec{x}_2) = \lambda_1 f(\vec{x}_1) + \lambda_2 f(\vec{x}_2)$$

для довільних $\vec{x}_1, \vec{x}_2 \in V$, $\lambda_1, \lambda_2 \in \mathbb{R}$.

Нехай $\big(\vec{e}_1, \ldots, \vec{e}_n\big)$ — базис простору V, f — лінійна числова функція векторного аргументу, визначена на цьому просторі. Тоді кожний вектор $\vec{x}$ лінійно виражається через базисні вектори:

$$\vec{x} = x_1\vec{e}_1 + \ldots + x_n\vec{e}_n, \quad x_k \in \mathbb{R}, k = \overline{1, n}$$

В силу лінійності f знаходимо:

$$f(\vec{x}) = x_1 f(\vec{e}_1) + \ldots + x_n f(\vec{e}_n).$$

Нехай відомо образи базисних векторів: $f(\vec{e}_i) = a_i$. Тоді

$$f(\vec{x}) = a_1x_1 + \ldots + a_nx_n.$$

Таким чином, при фіксованому базисі лінійна функція f подається *лінійною формою* (слово «форма» означає «однорідний многочлен» — многочлен, що є сумою одночленів одного і того ж степеня).

Якщо $\mathcal{E}$ — фіксований базис, то лінійну функцію можна подати формулою:

$$f(\vec{x}) = A \cdot \vec{x},$$

де $A = \begin{bmatrix} f(\vec{e}_1) & \ldots & f(\vec{e}_n) \end{bmatrix} = \begin{bmatrix} a_1 & \ldots & a_n. \end{bmatrix}$

Зауважимо, що часто лінійну функцію векторного аргументу так і називають: «лінійна форма».

1.2. Білінійні функції, білінійні форми.

Нехай V — скінченновимірний векторний простір.

Означення 8.3. Відображення $T\colon V \times V \to \mathbb{R}$ називається «білінійною функцією», якщо виконуються умови:

(1) $T(\vec{x}_1 + \vec{x}_2, \vec{y}) = T(\vec{x}_1, \vec{y}) + T(\vec{x}_2, \vec{y})$ $(\forall \vec{x}_1, \vec{x}_2, \vec{y} \in V)$;
(2) $T(\lambda\vec{x}, \vec{y}) = \lambda T(\vec{x}, \vec{y})$ $(\forall \vec{x}, \vec{y} \in V,\ \forall \lambda \in \mathbb{R})$;
(3) $T(\vec{x}, \vec{y}_1 + \vec{y}_2) = T(\vec{x}, \vec{y}_1) + T(\vec{x}, \vec{y}_2)$ $(\forall \vec{x}, \vec{y}_1, \vec{y}_2 \in V)$;
(4) $T(\vec{x}, \lambda\vec{y}) = \lambda T(\vec{x}, \vec{y})$ $(\forall \vec{x}, \vec{y} \in V,\ \forall \lambda \in \mathbb{R})$.

Приклад 8.2. Скалярний добуток векторів як відображення $V \times V \to \mathbb{R}$ є білінійною функцією.

Нехай $(\vec{e}_1, \ldots, \vec{e}_n)$ — базис простору V і нехай

$$\vec{x} = x_1\vec{e}_1 + \ldots + x_n\vec{e}_n;$$

$$\vec{y} = y_1\vec{e}_1 + \ldots + y_n\vec{e}_n.$$

Тоді білінійна функція $T(\vec{x},\vec{y})$ набуде вигляду

$$T(\vec{x},\vec{y}) = T(x_1\vec{e}_1 + \ldots + x_n\vec{e}_n, y_1\vec{e}_1 + \ldots + y_n\vec{e}_n) =$$
$$= x_1y_1T(\vec{e}_1,\vec{e}_1) + x_1y_2T(\vec{e}_1,\vec{e}_2) + \ldots + x_ny_nT(\vec{e}_n,\vec{e}_n) =$$
$$= \sum_{i,j=1}^{n} T(\vec{e}_i,\vec{e}_j)x_iy_j = \sum_{i,j=1}^{n} a_{ij}x_iy_j,$$

де $a_{ij} = T(\vec{e}_i,\vec{e}_j)$ не залежать від $\vec{x},\vec{y}$ (тут при перетвореннях ми використали аксіоми з означення білінійної форми).

Бачимо, що білінійна функція задається *білінійною формою* (тобто однорідним многочленом виду $\sum_{i,j=1}^{n} a_{ij}x_iy_j$). Часто замість слів «білінійна функція» ми так і говоритимемо: «білінійна форма».

Матриця

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}, \qquad a_{ij} = T(\vec{e}_i,\vec{e}_j)$$

називається *матрицею білінійної форми.*

Теорема 8.1. *Значення білінійної форми $T(\vec{x},\vec{y})$ з матрицею A в деякому базисі обчислюється за формулою:*

$$T(\vec{x},\vec{y}) = \vec{x}^T A\vec{y}.$$

Доведення. Безпосереднім підрахунком знаходимо:

$$\begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}x_1 + \ldots + a_{n1}x_n + \ldots + a_{1n}x_1 + \ldots + a_{nn}x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} =$$

$$= \sum_{i,j=1}^{n} a_{ij}x_iy_j,$$

що і потрібно було показати. ■

Зауваження. Кожну білінійну форму можна подати у вигляді $\vec{x}^T A\vec{y}$, де A — деяка матриця з дійсними коефіцієнтами. Це означає, що в якості означення білінійної функції (білінійної форми) можна взяти таке, яке еквівалентне означенню 8.3: *білінійною функцією* називається відображення $T\colon V \times V \to \mathbb{R}$ виду $T(\vec{x},\vec{y}) = \vec{x}^T \cdot A \cdot \vec{y}$, де A — деяка фіксована матриця.

Нехай $T(\vec{x},\vec{y})$ — білінійна форма, що в деякому базисі $\mathcal{B}$ задається матрицею A. Дослідимо, як зміниться матриця білінійної форми при зміні базису.

Отже, нехай $\mathcal{C}$ — деякий інший базис простору V, $P_{\mathcal{C}\leftarrow\mathcal{B}} \equiv P$ — матриця переходу від базису $\mathcal{B}$ до базису $\mathcal{C}$. Тоді координатами вектора $\vec{x}$ в новому базисі будуть $P\vec{x}$, а вектора $\vec{y}$ — $P\vec{y}$. Тоді білінійна форма T набуде виду:

$$T(\vec{x},\vec{y}) = (P\vec{x})^T \cdot A \cdot (P\vec{y}) = \vec{x}^T (P^T AP)\vec{y}.$$

Таким чином, матриця білінійної форми при зміні базису набуває виду $P^T AP$, де P — матриця переходу до нового базису.

Відомо, що якщо A — деяка матриця рангу r, а B — невироджена, то ранг матриці AB також дорівнює r (доведіть це). Оскільки матриця переходу до нового базису невироджена, то ранг матриці білінійної форми не залежить від вибору базису. Тому коректним є таке означення.

Означення 8.4. *Рангом* білінійної форми називається ранг її матриці в кожному базисі.

Означення 8.5. Білінійна форма $T(\vec{x},\vec{y})$ називається *симетричною*, якщо $T(\vec{x},\vec{y}) = T(\vec{y},\vec{x})$.

Для симетричної білінійної форми з матрицею A виконується співвідношення

$$\vec{x}^T A\vec{y} = \vec{y}^T A\vec{x}$$

для довільних векторів $\vec{x},\vec{y}$. Оскільки в останній рівності зліва і справа написані числа, то, розглядаючи їх як 1×1 матриці, дістаємо:

$$\vec{x}^T A\vec{y} = \vec{y}^T A\vec{x} = (\vec{y}^T A\vec{x})^T = \vec{x}^T A^T \cdot (\vec{y}^T)^T,$$

звідки $A = A^T$. Таким чином, матриця симетричної білінійної форми є симетричною.

Приклад 8.3. Нехай $T\colon V_2 \times V_2 \to \mathbb{R}$:

$$T(\vec{x}, \vec{y}) = 2x_1y_2 + x_2y_1 - x_2y_2.$$

Знайдемо матрицю цієї білінійної форми. Оскільки $T(\vec{x}, \vec{y}) = 0 \cdot x_1y_1 + 2 \cdot x_1y_2 + 1 \cdot x_2y_1 - 1 \cdot x_2y_2$, то робимо висновок, що $a_{11} = 0$, $a_{12} = 2$, $a_{21} = 1$, $a_{22} = -1$. Отже,

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

1.3. Квадратичні форми.

Нехай $T(\vec{x}, \vec{y})$ — симетрична білінійна форма. Якщо покласти $\vec{y} = \vec{x}$, то дістанемо так звану *квадратичну форму* $T(\vec{x}, \vec{x})$.

Означення 8.6. *Квадратичною формою* від n змінних $x_1, \dots, x_n$ називається функція $T\colon \mathbb{R}^n \to \mathbb{R}$, значення якої обчислюється за формулою $T(\vec{x}) = \vec{x}^T A\vec{x}$, де A — деяка симетрична матриця n-го порядку (вона називається *матрицею квадратичної форми*).

Квадратична форма називається *невиродженою*, якщо її матриця є невиродженою. *Рангом* квадратичної форми називається ранг її матриці.

Приклад 8.4. Знайти квадратичну форму, матриця якої: а) одинична; б) $A = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$.

Розв'язання.

а)

$$T(\vec{x}) = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + x_2^2 + \ldots + x_n^2.$$

б)

$$T(\vec{x}) = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - 3x_2 & -3x_1 + 5x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$
$$= 2x_1^2 - 6x_1x_2 + 5x_2^2.$$

■

Приклад 8.5. Знайти матрицю квадратичної форми

$$T(x_1, x_2, x_3) = 2x_1^2 - x_2^2 + 5x_3^2 + 6x_1x_2 - 3x_1x_3.$$

Розв'язання. Побудуємо таку матрицю:

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

Вона будується в такий спосіб: коефіцієнти при x_i^2 «йдуть» на головну діагональ матриці як елементи a_{ii}. Коефіцієнти при x_ix_j розподіляються порівну між a_{ij} та a_{ji}. Те, що вказана матриця є шуканою, легко перевірити безпосередньо, переконавшись, що

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \cdot A \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = T(x_1, x_2, x_3).$$

■

§ 2. Зведення квадратичної форми до канонічного вигляду

Говорять, що квадратична форма $T(x_1, \ldots, x_n)$ має *канонічний вигляд*, якщо вона не містить доданки виду ax_ix_j $(i \neq j)$. Наприклад, $T(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 5x_3^2$ має канонічний вигляд, а $T(x_1, x_2, x_3) = x_1^2 - 3x_2^2 + 5x_3^2 - x_1x_2$ — ні. Очевидно, що матриця квадратичної форми, що має канонічний вигляд, є діагональною.

Нехай $T(x_1, \ldots, x_n)$ — квадратична форма. Виникає питання: чи можна певним чином перейти до нових змінних (ввести заміну) так, щоб квадратична форма набула канонічного вигляду?

Виявляється, що відповідь на це питання завжди позитивна, причому досягти це можна різними способами.

2.1. Зведення квадратичної форми до канонічного вигляду за допомогою ортогональної діагоналізації її матриці.

Нехай $T(\vec{x}) = \vec{x}^T A \vec{x}$ — квадратична форма.

Оскільки A — симетрична матриця, то її завжди можна ортогонально діагоналізувати (див. теорему 7.18), тобто існують ортогональна матриця Q і діагональна D такі, що $A = QDQ^T$, причому діагональними елементами матриці D є власні значення матриці A (враховуючи кратність), а вектори-стовпці матриці Q — це відповідні їм власні вектори. Зазначимо, що $D = Q^T AQ$.

Якщо тепер ввести таку підстановку: $\vec{x} = Q\vec{y}$, або, що рівносильно, $\vec{y} = Q^T\vec{x}$, то дістанемо:

$$\begin{aligned}\vec{x}^T A\vec{x} =&(Q\vec{y})^T A(Q\vec{y}) = \vec{y}^T(Q^T AQ)\vec{y} = \\ &= \vec{y}^T D\vec{y},\end{aligned}$$

тобто вказана заміна змінної зводить вихідну квадратичну форму до такого канонічного вигляду:

$$\vec{y}^T D\vec{y} = \lambda_1\vec{y}_1^2 + \ldots + \lambda_n\vec{y}_n^2.$$

Вказаний процес називається *діагоналізацією квадратичної форми*. Нашими міркуваннями ми довели таке твердження, яке називається «теорема про головні осі»[1].

Теорема 8.2 (Теорема про головні осі)**.** *Кожну квадратичну форму можна звести до канонічного вигляду. Якщо A — матриця квадратичної форми $\vec{x}^T A\vec{x}$, ортогональна матриця Q і діагональна D такі, що $A = QDQ^T$, то заміна змінної $\vec{x} = Q\vec{y}$ перетворює квадратичну форму $\vec{x}^T A\vec{x}$ в квадратичну форму $\vec{y}^T D\vec{y}$ канонічного вигляду:*

$$\vec{x}^T A\vec{x} = \vec{y}^T D\vec{y} = \lambda_1\vec{y}_1^2 + \ldots + \lambda_n\vec{y}_n^2,$$

де $\lambda_1, \ldots, \lambda_n$ — власні значення матриці A.

Приклад 8.6. Звести квадратичну форму

$$T(\vec{x}) = 5x_1^2 + 4x_1x_2 + 2x_2^2$$

до канонічного вигляду.

Розв'язання. Складемо матрицю квадратичної форми:

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}.$$

У звичайний спосіб знаходимо, що власними значеннями цієї матриці є $\lambda_1 = 6$, $\lambda_2 = 1$, а власними векторами, що відповідають цим власним значенням є такі одиничні вектори:

$$\vec{q}_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}, \quad \vec{q}_2 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}.$$

[1]Доцільність такої назви буде зрозумілою з параграфу 5 (див. стор. 211)

Таким чином, виконавши таку заміну змінної: $\vec{x} = Q\vec{y}$, тобто

$$\begin{cases} x_1 = \dfrac{2}{\sqrt{5}}y_1 + \dfrac{1}{\sqrt{5}}y_2, \\ x_2 = \dfrac{1}{\sqrt{5}}y_1 - \dfrac{2}{\sqrt{5}}y_2, \end{cases}$$

квадратична форма $T(\vec{x})$ набуває канонічний вигляд $6y_1^2 + y_2^2$. ■

2.2. Зведення квадратичної форми до канонічного вигляду методом Лагранжа.

Існує інший спосіб зведення квадратичної форми до канонічного вигляду. Розглянемо його на попередньому прикладі.

Приклад 8.7. Звести квадратичну форму

$$T(\vec{x}) = 5x_1^2 + 4x_1x_2 + 2x_2^2$$

до канонічного вигляду.

Розв'язання. Здійснимо такі перетворення, виділивши повний квадрат:

$$T(\vec{x}) = 5x_1^2 + 4x_1x_2 + 2x_2^2 = 5\left(x_1^2 + \frac{4}{5}x_1x_2 + \frac{4}{25}x_2^2\right) - \frac{4}{5}x_2^2 + 2x_2^2 =$$

$$= 5\left(x_1 + \frac{2}{5}x_2\right)^2 + \frac{6}{5}x_2^2.$$

Отже, якщо покласти

$$\begin{cases} y_1 = x_1 + \dfrac{2}{5}x_2, \\ y_2 = x_2, \end{cases} \quad \text{або, що теж саме,} \quad \begin{cases} x_1 = y_1 - \dfrac{2}{5}y_2, \\ x_2 = y_2, \end{cases}$$

то квадратична форма буде зведена до канонічного вигляду $5y_1^2 + \frac{6}{5}y_2^2$.
■

Вказаний спосіб виділення повного квадрату (він називається методом Лагранжа) має той недолік, що матриця перетворення, взагалі кажучи, не є ортогональною. Це не завжди влаштовує при розв'язуванні задач.

Для розуміння загальної схеми виділення повних квадратів, розглянемо ще два приклади:

Приклад 8.8. Звести квадратичну форму

$$T(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + x_2^2 + x_1x_3$$

до канонічного вигляду.

Розв'язання.

$$
\begin{aligned}
T(x_1,x_2,x_3) &= x_1^2 + 2x_1x_2 + x_2^2 + x_1x_3 = \\
&= \left(x_1 + x_2 + \frac{1}{2}x_3\right)^2 - x_2x_3 - \frac{1}{4}x_3^2 = \\
&= (x_1 + x_2 + \frac{1}{2}x_3)^2 - \frac{1}{4}\left(x_3^2 + 4x_2x_3 + 4x_2^2\right) + x_2^2 = \\
&= (x_1 + x_2 + \frac{1}{2}x_3)^2 - \frac{1}{4}(x_3 + 2x_2)^2 + x_2^2 = y_1^2 + y_2^2 - \frac{1}{4}y_3^2,
\end{aligned}
$$

де $y_1 = x_1 + x_2 + \frac{1}{2}x_3$, $y_2 = x_2$, $y_3 = 2x_2 + x_3$. ■

Отже, основна ідея методу Лагранжа полягає в наступному. Припустимо, що квадратична форма містить хоча б один доданок виду $a \cdot x_i^2$. Тоді здіснюємо групування всіх доданків, що містять x_i і виділяємо повний квадрат таким чином, щоб він «поглинув» усі доданки виду ax_ix_j та ax_i^2. Тоді з доданками, які залишилась (а серед них вже немає доданків зі змінною x_i) повторюємо вказані дії.

Тепер припустимо, що квадратична форма не містить доданків виду $a \cdot x_i^2$. Не порушуючи загальності, припустимо, що вона містить доданок виду ax_1x_2. Тоді, ввівши проміжну заміну виду $x_1 = x_1'$, $x_2 = x_1' + x_2'$, $x_k = x_k'$, $k > 2$, одержимо квадратичну форму, в якій є доданок виду ax_1^2. Після цього виконуємо дії, описані вище.

Приклад 8.9. Звести квадратичну форму

$$T(x,y,z) = xy + yz + zx$$

до канонічного вигляду.

Розв'язання. Бачимо, що квадратична форма не містить доданків, необхідних для виділення повного квадрату. Тому введемо таку заміну:

$$
\begin{cases} x = u, \\ y = u + v, \\ z = w; \end{cases} \qquad \text{тобто} \qquad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}.
$$

Тоді квадратична форма набуде такий вигляд:

$$
\begin{aligned}
T(u,v,w) &= u(u+v) + (u+v)w + uw = u^2 + uv + 2uw + vw = \\
&= \left(u + \frac{1}{2}v + w\right)^2 - \frac{1}{4}v^2 - w^2.
\end{aligned}
$$

Тепер, ввівши таку заміну

$$\begin{cases} x_1 = u + \dfrac{1}{2}v + w, \\ y_1 = v, \\ z_1 = w, \end{cases}$$

дістанемо квадратичну форму в канонічному вигляді. Таким чином, заміна

$$\begin{cases} x_1 = \dfrac{1}{2}u + \dfrac{1}{2}(u+v) + w = \dfrac{1}{2}x + \dfrac{1}{2}y + z, \\ y_1 = u + v - u = y - x, \\ z_1 = z, \end{cases}$$

зводить вихідну квадратичну форму до канонічного вигляду $x_1^2 - \frac{1}{4}y_1^2 - z_1^2$. ■

Зауважимо, що квадратичну форму з останнього прикладу можна звести до канонічного вигляду і за допомогою діагоналізації її матриці

$$\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}.$$

Оскільки власні значення цієї матриці $\lambda_1 = 1$, $\lambda_2 = \lambda_3 = -\frac{1}{2}$, то після виконання відповідної заміни дістанемо таку квадратичну форму $y_1^2 - \frac{1}{2}y_2^2 - \frac{1}{2}y_3^2$. Помітимо, що як перший, так і другий спосіб зведення квадратичної форми до канонічного вигляду зводить її до виду, в якому лише один доданок з додатним коефіцієнтом, а два — з від'-ємним. Виявляється це не випадково, оскільки має місце загальне правило, яке називається *законом інерції квадратичних форм*.

§ 3. Закон інерції квадратичних форм

Припустимо, що квадратична форма $T(\vec{x})$, $\vec{x} = (x_1, x_2, \ldots, x_n)$ зведена до канонічного вигляду $T(\vec{x}) = \lambda_1 y_1^2 + \ldots + \lambda_r y_r^2$, де λ_i — відмінні від нуля коефіцієнти, причому перші q з них — додатні, а решта — від'ємні:

$$\lambda_1 > 0, \ldots, \lambda_q > 0, \lambda_{q+1} < 0, \ldots, \lambda_r < 0,$$

r — ранг квадратичної форми.

Розглянемо таке перетворення:

$$z_1 = \frac{1}{\sqrt{|\lambda_1|}} y_1, \ \ldots, \ z_r = \frac{1}{\sqrt{|\lambda_r|}} y_r.$$

В результаті такої заміни квадратична форма $T(\vec{x})$ набуде так званий *нормальний вигляд*:

$$T(\vec{x}) = z_1^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_r^2.$$

Теорема 8.3 (Закон інерції квадратичних форм)**.** *Кількість доданків з додатними (від'ємними) коефіцієнтами в нормальному вигляді квадратичної форми не залежить від способу зведення квадратичної форми до цього вигляду.*

Доведення. Припустимо, що підстановками

$$\vec{y} = P_1\vec{x}, \quad \vec{z} = P_2\vec{x}, \tag{8.1}$$

$\det P_1 \neq 0$, $\det P_2 \neq 0$, квадратична форма $T(\vec{x})$ зведена до нормального вигляду двома способами:

$$T(\vec{x}) = y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2,$$
$$T(\vec{x}) = z_1^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_r^2.$$

Слід довести, що $p = q$. Доведення проведемо методом від супротивного. Не порушуючи загальності, припустимо, що $p < q$. Покажемо, що існує такий ненульовий вектор $\vec{x}$, що $y_1 = \ldots = y_p = 0$, $z_{q+1} = \ldots = z_r = 0$. Справді, враховуючи (8.1), система

$$\begin{cases} y_1 = 0, \\ \vdots \\ y_p = 0, \\ z_{q+1} = 0, \\ \vdots \\ z_r = 0, \end{cases}$$

є однорідною системою відносно невідомих $x_1, \ldots, x_n$, причому кількість рівнянь в ній менша n (оскільки $p < q$), а значить вона має ненульовий розв'язок $\vec{x}_0$.

З іншого боку, тоді $f(\vec{x}_0) = -y_{p+1}^2 - \ldots - \vec{y}_r^2$ і $f(\vec{x}_0) = z_1^2 + \ldots + z_q^2$, що можливо лише при $y_{p+1} = \ldots = y_r = z_1 = \ldots = z_q = 0$. Тобто існує такий базис, в якому координати вектора $\vec{x}_0$ усі дорівнюють нулю. Це можливо лише для $\vec{x}_0 = \vec{0}$. Одержали протиріччя. Випадок $p > q$ розглядається аналогічно. Таким чином, $p = q$, що і потрібно було довести. ■

§ 4. Класифікація квадратичних форм

Означення 8.7. Квадратична форма $T(\vec{x})$ називається *додатно визначеною* (*від'ємно визначеною*), якщо для кожного ненульового вектора $\vec{x}$: $T(\vec{x}) > 0$ ($T(\vec{x}) < 0$). Якщо квадратична форма набуває як додатних, так і від'ємних значень, то вона називається *невизначеною*.

Виникає питання: чи можна визначити до якого з трьох наведених класів належить задана квадратична форма? Відповідь на це питання дають дві наступні теореми.

Теорема 8.4. *Квадратична форма $T(\vec{x}) = \vec{x}^T A\vec{x}$ є:*

1) *додатно визначеною тоді і тільки тоді, коли усі власні значення матриці A додатні;*
2) *від'ємно визначеною тоді і тільки тоді, коли усі власні значення матриці A від'ємні;*
3) *невизначеною тоді і тільки тоді, коли матриця A має як додатні, так і від'ємні власні значення.*

Доведення. За теоремою про головні осі (теорема 8.2) для квадратичної форми $T(\vec{x})$ існує заміна змінної $\vec{x} = Q \cdot \vec{y}$ така, яка зводить вихідну квадратичну форму до канонічного вигляду: $\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$. Оскільки матриця Q — невироджена (Q ортогональна, а визначник ортогональної матриці, як відомо, дорівнює 1 або -1), то рівність $\vec{x} = Q\vec{y}$ встановлює взаємно однозначну відповідність між ненульовими векторами $\vec{x}$ та $\vec{y}$. Тому для будь-якого $\vec{x}$ значення $T(\vec{x})$ співпадає із значенням виразу $\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$, знак якого цілком визначається числами $\lambda_1, \ldots, \lambda_n$.

Якщо деякі власні значення матриці A додатні, а деякі — від'ємні, то завжди можна підібрати такі вектори $\vec{y}_1$, $\vec{y}_2$, а разом з ними і відповідні їм $\vec{x}_1$, $\vec{x}_2$ такі, що $T(\vec{x}_1) > 0$, $T(\vec{x}_2) < 0$. ■

Означення 8.8. Симетрична матриця A називається *додатно визначеною*, *від'ємно визначеною*, *невизначеною*, якщо квадратична форма $\vec{x}^T A\vec{x}$ володіє відповідною властивістю.

Приклад 8.10. Дослідити на знаковизначеність квадратичну форму

$$T(\vec{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3.$$

Розв'язання. Квадратична форма інтуїтивно здається додатно визначеною. Перевіримо чи це справді так. Її матриця

$$\begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

має такі власні значення (перевірте!) $\lambda_1 = 5$, $\lambda_2 = 2$, $\lambda_3 = -1$. Таким чином, насправді квадратична форма є невизначеною. Справді, наприклад, $T(1,0,0) = 3 > 0$, $T(0,1,-1) = -1 < 0$. ■

Виявляється, що перевірку знаковизначеності квадратичної форми можна здійснити не вдаючись до знаходження власних значень матриці квадратичної форми.

Нехай квадратична форма $T(\vec{x})$ в деякому базисі задається матрицею:

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

Введемо такі позначення головних («кутових») мінорів матриці A:

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \Delta_n = \det A.$$

Теорема 8.5 (Критерій Сильвестра додатної визначеності квадратичної форми). *Для того, щоб квадратична форма $T(\vec{x}) = \vec{x}^T A\vec{x}$ була додатно визначеною, необхідно і достатньо, щоб усі головні мінори матриці A були додатними:*

$$\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0.$$

Доведення. *Необхідність.* Нехай $T(\vec{x})$ — додатно визначена квадратична форма, A — її матриця. Тоді за теоремою про головні осі існує невироджена матриця P така, що $A = PDP^T$, $D = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$, а квадратична форма заміною $\vec{x} = P\vec{y}$ зводиться до канонічного вигляду $\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2$. Зазначимо, що за теоремою 8.4 $\lambda_1 > 0, \dots, \lambda_n > 0$.

Маємо:

$$P^T AP = D;$$
$$\det(P^T AP) = \det D;$$
$$\det A \cdot (\det P)^2 = \lambda_1 \cdot \ldots \cdot \lambda_n > 0,$$

звідки $\det A = \Delta_n > 0$.

Отже ми показали, що детермінант матриці додатно визначеної квадратичної форми є додатним числом.

Розглянемо тепер вектор $\vec{x} = (x_1, \ldots, x_k, 0, \ldots, 0)$, $k \in \{1, \ldots, n-1\}$. Тоді квадратична форма $T(\vec{x})$ набуває такого вигляду, що може мислитись як квадратична форма відносно змінних $x_1, \ldots, x_k$, і, крім того, вона є додатно визначеною (за умовою). Тоді за доведеним вище детермінант її матриці (як квадратичної форми відносно змінних $x_1, \ldots, x_k$) є додатним числом: $\Delta_k > 0$, $k \in \{1, \ldots, n-1\}$, що і потрібно було довести.

Достатність. Нехай усі головні мінори матриці A квадратичної форми $T(\vec{x})$ є додатними. Доведемо, що $T(\vec{x})$ додатно визначена.

Доведення проведемо за індукцією по кількості змінних, що входять в форму. Якщо $T(\vec{x})$ залежить від однієї змінної, то $T(\vec{x}) = T(x_1) = a_{11}x_1^2$ і з того, що $\Delta_1 = a_{11} > 0$ дістаємо, що $T(x_1) > 0$, $\vec{x}_1 \neq 0$.

Тепер припустимо, що якщо для деякого довільного, але фіксованого $m > 2$ виконуються співвідношення $\Delta_1 > 0, \ldots, \Delta_{m-1} > 0$, то кожна квадратична форма $T(x_1, \ldots, x_{m-1})$ є додатно визначеною.

Розглянемо квадратичну форму $T(x_1, \ldots, x_m)$, для якої $\Delta_k > 0$, $k = \overline{1, m}$. З припущення індукції випливає, що та «підчастина» квадратичної форми, що не містить доданків з множником x_m, є додатно визначеною, тобто з допомогою певного перетворення зводиться до вигляду $(x_1')^2 + \ldots + (x_{m-1}')^2$.

Отже, зробивши відповідну заміну і підставивши $x_m = x_m'$, одержуємо:

$$T(\vec{x}) = (x_1')^2 + \ldots + (x_{m-1}')^2 + 2b_{1m}x_1'x_m' + \ldots + 2b_{m-1,m}x_{m-1}'x_m' + a_{mm}(x_m')^2,$$

де b_{im} — деякі нові коефіцієнти. Виділивши повні квадрати, знаходимо:

$$T(\vec{x}) = (x_1' + b_{1m}x_m')^2 + \ldots + (x_{m-1}' + b_{m-1,m}x_m')^2 + b \cdot (x_m')^2,$$

де $b = a_{mm} - b_{1m}^2 - \ldots - b_{m-1,m}^2$.

Тепер, виконавши заміну

$$\begin{cases} y_1 = x_1' + b_{1m}x_m', \\ \ldots \\ y_{m-1} = x_{m-1}' + b_{m-1,m}x_m', \\ y_m = x'm, \end{cases}$$

зведемо квадратичну форму до вигляду $T(\vec{x}) = y_1^2 + \ldots + y_{m-1}^2 + by_m^2$. Визначник матриці цієї форми дорівнює b і його знак співпадає зі знаком матриці квадратичної форми Δ_m (якщо $B = P^{-1}AP$, то $\det A$ та $\det B$ одного знаку). Тому $b > 0$, а значить $T(\vec{x})$ додатно визначена. Теорему доведено. ■

Наслідок (Критерій Сильвестра від'ємної визначеності квадратичної форми)**.** *Для того, щоб квадратична форма була від'ємно визначеною необхідно і достатньо, щоб знаки головних мінорів її матриці чергувались починаючи зі знаку «−», тобто*

$$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots.$$

Доведення. Квадратична форма $T(\vec{x}) = \vec{x}^T A\vec{x}$ є від'ємно визначеною тоді і тільки тоді, коли квадратична форма $T(\vec{x}) = -\vec{x}^T A\vec{x} = \vec{x}^T(-A)\vec{x}$ є додатно визначеною. Тоді за попередньою теоремою

$$-a_{11} > 0, \begin{vmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{vmatrix} > 0, \begin{vmatrix} -a_{11} & -a_{12} & -a_{13} \\ -a_{21} & -a_{22} & -a_{23} \\ -a_{31} & -a_{32} & -a_{33} \end{vmatrix} > 0, \ldots,$$

тобто

$$\Delta_1 = a_{11} < 0,\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0,\ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0, \ldots$$

що і потрібно було довести. ■

Приклад 8.11. За критерієм Сильвестра дослідимо на знаковизначеність квадратичну форму з прикладу 8.10: $T(\vec{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$. Її матриця

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

має такі головні мінори

$$\Delta_1 = 2 > 0,\ \Delta_2 = \begin{vmatrix} 3 & 2 \\ 2 & 2 \end{vmatrix} = 2 > 0,\ \Delta_3 = \det A = -10 < 0.$$

Таким чином, квадратична форма є невизначеною.

§ 5. Зведення рівняння лінії другого порядку до канонічного вигляду

Як відомо, *кривою другого порядку* називається геометричне місце точок площини, декартові координати яких задовольняють рівняння

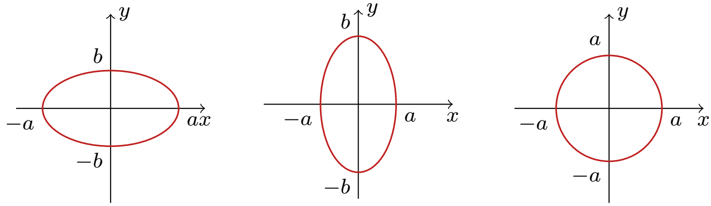
$$a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + 2a_1x + 2a_2y + a_0 = 0, \qquad (8.2)$$

де хоча б одне з чисел a_{11}, a_{22}, a_{12} відмінне від нуля.

З курсу аналітичної геометрії відомо, що найважливішими кривими другого порядку є еліпс, гіпербола та парабола. Це так звані невироджені криві другого порядку.

Також графіком рівняння (8.2) можуть бути так звані вироджені криві другого порядку: точка (вироджений еліпс); пара прямих, що перетинаються (вироджена гіпербола); пара паралельних прямих або одна пряма (вироджена парабола).

Шляхом введення нової системи координат кожну криву другого порядку можна звести до так званого канонічного вигляду (тоді говорять, що графік кривої другого порядку у стандартному положенні). Нагадаємо канонічні рівняння еліпса, гіперболи, параболи та їх графіки (див. рис. 8.1, 8.2, 8.3).



а) $a > b$. б) $a < b$. в) $a = b$.

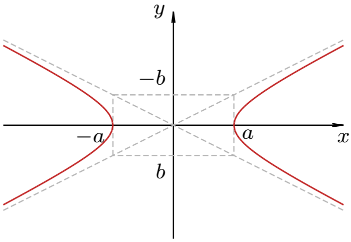
Рис. 8.1. Еліпс $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Далі розглянемо деякі приклади.

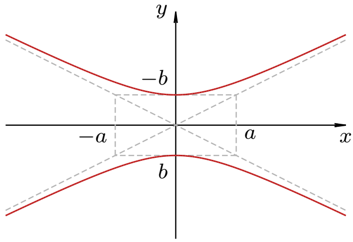
Приклад 8.12. Побудувати графіки кривих другого порядку:

(1) $4x^2 + 9y^2 = 36$;

(2) $4x^2 - 9y^2 + 1 = 0$;

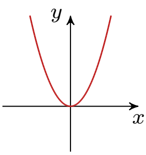


а) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$

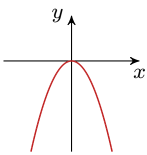


б) $\frac{x^2}{a^2} - \frac{y^2}{b^2} = -1.$

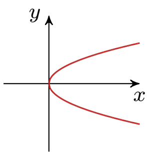
Рис. 8.2. Гіпербола.



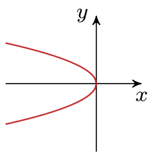
а) $y = ax^2$, $a > 0$.



б) $y = ax^2$, $a < 0$.



в) $x = ay^2$, $a > 0$.



г) $x = ay^2$, $a < 0$.

Рис. 8.3. Парабола.

(3) $4x^2 - 9y = 0$.

Розв'язання. Перше рівняння перепишемо у вигляді

$$\frac{x^2}{3^2} + \frac{y^2}{2^2} = 1,$$

звідки зрозуміло, що її графіком є еліпс з великою піввіссю 3 і малою піввіссю 2 (див. рис. 8.1(а) для $a = 3$, $b = 2$).

Друге рівняння набуває вигляд

$$\frac{x^2}{\left(\frac{1}{2}\right)^2} - \frac{y^2}{\left(\frac{1}{3}\right)^2} = -1,$$

а тому її графіком є гіпербола (див. рис. 8.2(б) для $a = \frac{1}{2}$, $b = \frac{1}{3}$).

Третє рівняння є графіком параболи $y = \frac{4}{9}x^2$ (див. рис. 8.3(a)). ■

Тепер розглянемо приклад, коли слід побудувати графік кривої другого порядку, в якої коефіцієнт при xy дорівнює нулю.

Приклад 8.13. Побудувати графік рівняння:

$$x^2 + 2y^2 - 6x + 8y + 9 = 0.$$

Розв'язання. Виділимо в лівій частині рівняння повні квадрати:

$$(x-3)^2 + 2(y+2)^2 = 8.$$

Введемо заміну

$$\begin{cases} x' = x - 3, \\ y' = y + 2. \end{cases}$$

Дістанемо рівняння $\frac{x'^2}{8} + \frac{y'^2}{4} = 1$. Одержали канонічне рівняння еліпса з півосями $2\sqrt{2}$ і 2 в системі координат $x'O'y'$. Початок цієї системи координат розташований в точці $(3, -2)$ (див. рис. 8.4). ■

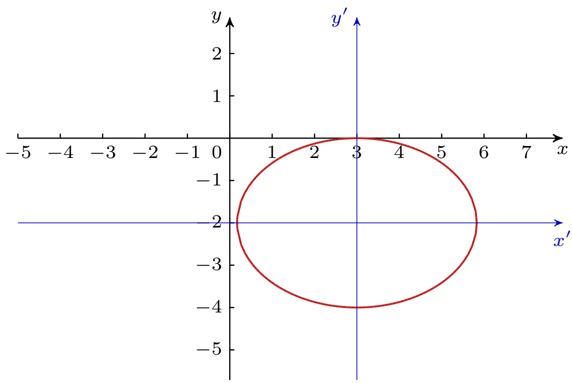


Рис. 8.4. Графік кривої $x^2 + 2y^2 - 6x + 8y + 9 = 0$.

Якщо в рівнянні (8.2) $a_{12} \neq 0$, то графіком цього рівняння є крива другого порядку, яка повернута на певний кут відносно стандартного положення.

Приклад 8.14. Побудувати графік кривої, що задається рівнянням:

$$5x^2 + 4xy + 2y^2 = 6.$$

Розв'язання. Розглянемо квадратичну форму $T(x,y) = 5x^2 + 4xy + 2y^2$. Матрицею цієї квадратичної форми є

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}.$$

Стандартним обчисленням знаходимо власні значення $\lambda_1 = 6$, $\lambda_2 = 1$ та відповідні їм нормовані власні вектори $\vec{q}_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$ та $\vec{q}_2 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}$.

З власних векторів складемо матрицю

$$Q = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}.$$

Помітимо, що детермінант матриці Q дорівнює -1. Це не зручно, оскільки ортогональна матриця з детермінантом рівним -1 задає перетворення композиції повороту та симетрії (див. наслідок на сторінці 180). Тому перепозначимо $\lambda_1 = 1$, $\lambda_2 = 6$ і

$$Q = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

Тепер $\det Q = 1$ і матриця Q визначає поворот на кут $\varphi = \arccos \frac{1}{\sqrt{5}}$. Таким чином, заміна змінної $\begin{bmatrix} x \\ y \end{bmatrix} = Q \begin{bmatrix} x' \\ y' \end{bmatrix}$ зводить квадратичну форму до канонічного вигляду $(x')^2 + 6(y')^2$. Зрештою рівняння лінії набуває вигляд

$$\frac{(x')^2}{6} + \frac{(y')^2}{1} = 1.$$

Тепер побудуємо графік. Для цього слід вияснити, якими є нові базисні вектори $\vec{e}_1'$ та $\vec{e}_2'$ нової системи координат (вони визначать напрям нових осей координат). Оскільки $\begin{bmatrix} x \\ y \end{bmatrix} = Q \begin{bmatrix} x' \\ y' \end{bmatrix}$, то

$$Q\vec{e}_1' = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix};$$

$$Q\vec{e}_2' = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix},$$

тобто новими базисними векторами є стовпці матриці Q. Графік рівняння на рисунку 8.5. ■

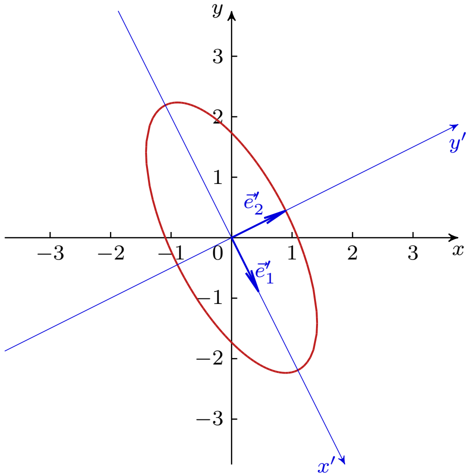


Рис. 8.5. Графік лінії $5x^2 + 4xy + 2y^2 = 6$.

Пояснимо тепер, чому теорема про головні осі так називається. Якщо симетрична матриця A є матрицею квадратичної форми, яка входить до рівняння другого порядку, то власні вектори цієї матриці встановлюють напрямки головних осей лінії другого порядку, що задається цим рівнянням (головною віссю лінії другого порядку називається її вісь симетрії: еліпс та гіпербола мають дві взаємно перпендикулярні осі симетрії, парабола — одну).

Приклад 8.15. Побудувати графік рівняння
$$5x^2 + 4xy + 2y^2 - \frac{28}{\sqrt{5}}x - \frac{4}{\sqrt{5}}y + 4 = 0.$$

Розв'язання. Запишемо дане рівняння в матричній формі:
$$\vec{x}^T A\vec{x} + B\vec{x} + 4 = 0,$$
де
$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} -\frac{28}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

В попередньому прикладі ми встановили, що заміна змінної $\vec{x} = Q\vec{x}'$, де

$$Q = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

зводить квадратичну форму $\vec{x}^T A\vec{x}$ до канонічного вигляду $(x')^2 + 6(y')^2$. Крім цього,

$$B\vec{x} = BQ\vec{x} = \begin{bmatrix} -\frac{28}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \vec{x} =$$
$$= -4x' - 12y'.$$

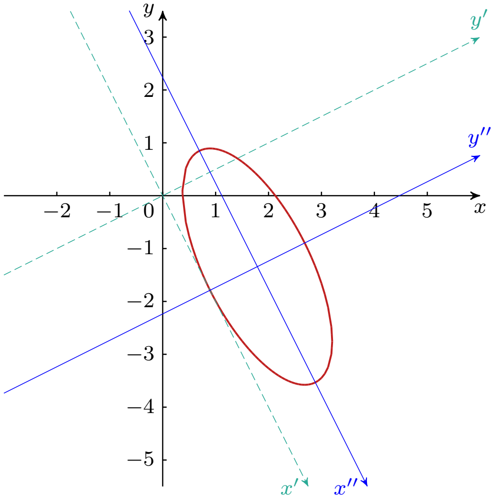


Рис. 8.6. Графік рівняння $5x^2 + 4xy + 2y^2 - \frac{28}{\sqrt{5}}x - \frac{4}{\sqrt{5}}y + 4 = 0$.

Отже, вказана заміна приводить до рівняння $(x')^2 + 6(y')^2 - 4x' - 12y' + 4 = 0$, яке можна переписати у вигляді $(x' - 2)^2 + 6(y' - 1)^2 = 6$. Ввівши ще одну заміну

$$\begin{cases} x'' = x' - 2, \\ y'' = y' - 1, \end{cases}$$

дістанемо, що в системі координат $x''y''$ графіком лінії є еліпс $\frac{(x'')^2}{6} + (y'')^2 = 1$.

Перейдемо до побудови графіка. Система координат $x'y'$ є такою ж як і в попередньому прикладі. Система координат $x''y''$ будується шляхом паралельного перенесення на вектор $(2, 1)$ системи координат $x'y'$.

■

Додаткові відомості і завдання для самостійної роботи

1. Нехай B — оборотна матриця. Довести, що матриця $A = B^T B$ є додатно визначеною.
2. Нехай A — додатно визначена симетрична матриця. Довести, що існує оборотна матриця B така, що $A = B^T B$.
3. Нехай A і B — додатно визначені симетричні матриці n-го порядку, $\lambda > 0$. Довести, що матриці λA, A^2, $A + B$, A^{-1} також додатно визначені.
4. Нехай A — додатно визначена симетрична матриця. Довести, що існує додатно визначена симетрична матриця B така, що $A = B^2$ (така матриця B називається *коренем квадратним* з A).
5. Доведіть, що дві квадратичні форми, одна з яких додатно визначена, можна одночасно звести до канонічного вигляду шляхом невиродженого лінійного перетворення змінних.

Предметний покажчик

Рекомендована література

1. *Воеводин В. В.* Линейная алгебра / В. В. Воеводин. — М. : Наука, 1974. — 336 с.
2. *Гельфанд И. М.* Лекции по линейной алгебре, 4-е изд., доп. / И. М. Гельфанд — М. : Наука, 1971. — 272 с.
3. *Головина Л. И.* Линейная алгебра и некоторые её приложения / Л. И. Головина.— М. : Наука, 1975. — 408 с.
4. *Завало С. Т.* Алгебра і теорія чисел. Частина I / С. Т. Завало, В. М. Костарчук, Б. І. Хацет. — К. : Вища школа, 1974. — 464 с.
5. *Збірник* задач з алгебри. Частина 1 : Навчальний посібник для студентів фізико-математичного факультету / за ред. І. О. Рокіцького. — Вінниця, 2002. — 176 с.
6. *Збірник* задач з алгебри. Частина 2 : Навчальний посібник для студентів фізико-математичного факультету / за ред. І. О. Рокіцького. — Вінниця, 2003. — 200 с.
7. *Ильин В. А.* Линейная алгебра : Учеб. для вузов, 4-е изд. / В. А. Ильин, Э. Г. Позняк — М. : Наука, Физматлит, 1999. — 296 с.
8. *Калужнін Л. А.* Лінійні простори / Л. А. Калужнін, В. А. Вишенський, Ц. О. Шуб — К. : Вища школа, 1971. — 344 с.
9. *Курош А. Г.* Курс высшей алгебры, изд. 9-е. / А. Г. Курош — М. : Наука, Гл. ред. физ.-мат. лит., 1968. — 431 с.
10. 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.
11. *Мальцев А. И.* Основы линейной алгебры / А. И. Мальцев. — М. : Наука, 1970. — 400 с.
12. *Рокіцький І. О.* Застосування лінійної алгебри / І. О. Рокіцький, О. Б. Панасенко. — Вінниця : ФОП Легкун, 2012. — 240 с.

13. *Стренг Г.* Линейная алгебра и её применения / Г. Стренг. — М. : Мир, 1980. — 454 с.
14. *Тыртышников Е. Е.* Матричный анализ и линейная алгебра : учебное пособие / Е. Е. Тыртышников. — М. : ФИЗМАТЛИТ, 2007. — 477 с.
15. *Чарін В. С.* Лінійна алгебра / В. С. Чарін. — К. : Техніка, 2004. — 416 с.
16. *Шевцов Г. С.* Линейная алгебра: теория и прикладные аспекты : Учеб. пособие / Г. С. Шевцов. — М. : Финансы и статистика, 2003. — 576 с.
17. *Элементы* линейной алгебры / под общ. ред. Р. Ф. Апатенок. — Мн. : Вышейш. школа, 1977. — 256 с.
18. *Hefferon J.* Linear Algebra / J. Hefferon. — Saint Michael's College Colchester, Vermont, 2008. — 445 p.
19. *Jurlewicz T.* Algebra liniowa 1. Definicje, twierdzenia, wzory / T. Jurlewicz, Z. Skoczylas. — Wroclaw: Oficyna Wydawnicza GiS, 2003. — 163 str.
20. *Lay D. C.* Linear Algebra and its Applications, 3-rd ed. / D. C. Lay. — Boston, 2005. — 560 p.
21. *Nicholson W. K.* Linear Algebra With Applications, 3rd Edition / W. K. Nicholson, 1995. — 540 p.
22. *Poole D.* Linear Algebra: A Modern Introduction, 2nd edition / D. Poole. — Brooks/Cole, 2006. — 712 p.

Навчальне видання

ПАНАСЕНКО Олексій Борисович

Лекції з лінійної алгебри

Видання друге, доповнене

Навчальний посібник

Ваші побажання з приводу змісту книги
можна надіслати автору на електронну адресу
panalbor@gmail.com

Підписано до друку 8.12.2015
Формат 64x90 1/16. Папір офсетний

Гарнітура Computer Modern. Друк цифровий.

Наклад 50 прим.