МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

АЛГЕБРА
ТА ГЕОМЕТРІЯ:
ЛІНІЙНА АЛГЕБРА

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за освітньою програмою «Системний аналіз і управління»
спеціальності 124 «Системний аналіз»*

Київ

КПІ ім. Ігоря Сікорського

2021

Алгебра та геометрія: Лінійна алгебра. [Електронний ресурс]: навч. посіб. для студ. спеціальності 124 «Системний аналіз» / КПІ ім. Ігоря Сікорського; уклад. Ю. Є. Бохонов. – Електронні текстові дані (1 файл: 4,31 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. – 243 с.

Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 6 від 25.02.2021 р.) за поданням Вченої ради Інституту прикладного системного аналізу *(протокол № 2 від 22.02.2021 р.)*

Електронне мережне навчальне видання

АЛГЕБРА ТА ГЕОМЕТРІЯ:
ЛІНІЙНА АЛГЕБРА

Укладач: *Бохонов Юрій Євгенович*, канд. ф. - мат. н, доцент

Відповідальний редактор *Подколзін Г.Б.*, канд. ф. - мат. н., доцент

Рецензент: *Голуб А.П.*, д-р ф.-мат. н., провідний науковий співробітник Інституту математики НАН України

Посібник «Алгебра та геометрія: Лінійна алгебра» призначено для вивчення студентами спеціальності 124 «Системний аналіз» основ лінійної алгебри та її застосування до аналітичної геометрії. Особливу увагу приділено теорії лінійних просторів, операторів, матриць та визначників, причому, остання викладається з точки зору апарату алгебр Грасмана. Розглянуто питання зведення матриці до нормальних форм – Жордана в комплексному і дійсному випадках та циклічної. Детально розібразведення до нормальної форми Жордана усіх випадків для матриць третього, четвертого і п’ятого порядків. В посібнику приводяться всі конструкції з суміжних розділів – теорії груп, кілець, тензорів.

ЗМІСТ

ВСТУП

Пропонований посібник є результатом грунтовного розширення курсу лекцій, що читався автором на факультеті прикладної математики НТУУ «КПІ» студентам спеціальності 123 «Комп’ютерна інженерія». Разом з традиційними питаннями лінійної алгебри викладено основи теорії груп (глава 1), відомості якої використовуються в питаннях тензорного аналізу та теорії визначників і квадратичних форм.

Глава 1. Тут наведено деякі поняття та факти теорії груп та кілець. Особливу увагу приділено групам підстановок, які потім будуть використані в теорії визначників.

Глава 2 присвячена питанням теорії лінійних просторів та лінійних операторів. Наголос робиться на знаходженні матриці оператора в заданому базисі та зв’язку між його матрицями в різних базисах. В главі розглядається також спектр самоспряженого оператора. Як доповнення для допитливих викладено основи теорії кватерніонів та октав Келі.

В главі 3 викладено деякі поняття тензорного аналізу та алгебри Грасмана (розділ 3.2). При першому читанні цей розділ можна пропустити і переходити до 3.3, де пропонується інший підхід до викладення теорії алгебри Грасмана.

Глава 4. В ній будується теорія визначників з застосуванням нетрадиційного підходу, що базується на техніці зовнішнього множення. Тут також наведено основні факти теорії систем лінійних алгебраїчних рівнянь. Доповненнями в цій главі є знаходження об’єму багатовимірного паралелепіпеду та розгляд деяких властивостей функціональних визначників, зокрема, вронскіанів.

Глава 5. Викладено теорію зведення матриці оператора до нормальної форми Жордана як у комплексному, так і дійсному випадках. Теорія застосовується до визначення функцій від оператора. Наведено приклади, що вичерпують усі випадки у разі матриць третього та четвертого порядку. Викладено також питання зведення матриці оператора до циклічної форми (нормальної форми Фробеніуса).

Глава 6 містить основні поняття теорії білінійних та квадратичних форм. Доводиться закон інерції квадратичних форм, різні способи їхнього зведення до канонічного вигляду (методи Лагранжа та Якобі), а також критерій Сільвестра знаковизначеності матриць.

Викладена теорія застосовується для геометричної задачі зведення квадратичної форми до головних осей.

Пропонований посібник доповнює видання «Алгебра та геометрія. Конспект лекцій з курсу «Аналітична геометрія та лінійна алгебра»» [Електронний ресурс] : навчальний посібник для студентів спеціальності 124 «Системний аналіз» / КПІ ім. Ігоря Сікорського ; уклад. О. О. Калюжний, А. Ю. Мальцев, Г. Б. Подколзін, Ю. А. Чаповський ; КПІ ім. Ігоря Сікорського.

ГЛАВА 1. ДЕЯКІ ВІДОМОСТІ ПРО ГРУПИ, КІЛЬЦЯ, ПОЛЯ

1.1. ГРУПИ

Групою G називається множина елементів, на якій визначена одна операція (яка іноді називається законом композиції), що ставить у відповідність двом елементам $a,\ b \in G$ деякий елемент $c \in G$. Іноді, виходячи з міркувань зручності, будемо позначати цю операцію як добуток ($ab=c$), називаючи групу мудьтиплікативною, іноді – як суму ($a + b = c$), називаючи групу адитивною.

Операція повинна задовольняти такі аксіоми.

1. Для $a, b \in G$ $(ab)\,c = a(bc)$ – закон асоціативності.

2. Існує такий елемент е $\in G$, що для довільного $a \in G$ $ae = ea = a$. Цей елемент називається одиничним або тотожним, оскільки композиція з ним не змінює елемент a.

3. Для кожного $a \in G$ існує елемент b такий, що $ab=ba=e$. Елемент b називається оберненим до a і буде позначатися a^{-1}. Іноді кажуть, що в групі визначене ділення.

4. Якщо в групі виконується ще аксіома комутативності для довільних $a, b \in G$, тобто $ab = ba$, то група називається комутативною або абелевою на честь математика Нільса Хенріка Абеля.

Говорячи про абстрактні групи, як правило, групову операцію називають добутком. Якщо ж група комутативна, то часто зручніше абстрактну групу вважати адитивною і групову операцію називати сумою.

У групі може існувати тільки одна одиниця. Справді, якщо e_1 також одиниця, то, з одною боку, $ee_1 = e_1$, а з іншого – $ee_1 = e$, тобто $e_1= e$.

Для кожного елемента існує тільки один обернений. Дійсно, якщо $ab=ac=e$, то, розглянувши елемент bac і застосувавши аксіому 1, одержимо: $(ba)c = c = b(ac) = b$, отже, $b=c$.

Доведемо, що $\left(ab\right)^{-1} = b^{-1}a^{-1}$. Для доведення треба показати, що їх добуток дорівнює одиничному елементу. Справді, $abb^{-1}a^{-1} = aa^{-1} = e$. Оскільки обернений елемент єдиний, це і доводить твердження. Зрозуміло, що така ж формула справедлива для довільно кількості співмножників: $\left(a_1a_2...a_n\right)^{-1} = a_n^{-1}...a_2^{-1}a_1^{-1}$.

Через a^k будемо позначати добуток *a...a* (*k* разів). З попереднього випливає, що $\left(a^k\right)^{-1} = \left(a^{-1}\right)^k$, тому будемо вживати позначення $a^{-k} = \left(a^k\right)^{-1} = \left(a^{-1}\right)^k$.

Групи, як ми побачимо далі, можуть бути скінченними або нескінченними. Кількість елементів скінченної групи називається її порядком. Порядок нескінченної групи – це потужність множини її елементів.

Приклади. Далі будемо використовувати позначення: $\mathbb{Z}$ – множина цілих чисел, $\mathbb{Q}$ – множина раціональних чисел $\mathbb{R}$ – множина дійсних чисел, $\mathbb{C}$ – множина комплексних чисел.

1. Якщо закон композиції в групі – операція множення чисел, то мультиплікативними групами будуть такі множини: $\mathbb{R}\setminus 0$, $\mathbb{C}\setminus 0$. Очевидно, це абелеві групи. Одиничним елементом у них є число 1. Оберненим елементом до кожного елемента *a* (відповідно для кожної з множин) є обернене число $a^{-1} = \dfrac{1}{a}$.

2. $G = \mathbb{R}_+$ (множина дійсних додатних чисел) і $G = \mathbb{Q}_+$ (множина раціональних додатних чисел) утворюють абелеві мультиплікативні групи.

3. Якщо законом композиції є операція додавання, то групами будуть множини $\mathbb{Z}, \mathbb{R}, \mathbb{Q}, \mathbb{C}$. Очевидно, це абелеві групи. Одиничним елементом в них є число 0, оскільки композиція довільного числа з 0 (тобто додавання 0) не змінить його. Оберненим для *a*, очевидно, є число – *a*.

4. Вектори на площині або в трьохвимірному просторі теж утворюють групу (адитивну абелеву). Операція композиції – це додавання векторів за правилом “паралелограма”.

5. Множина з двох чисел { 1, –1 } – мультиплікативна абелева група.

6. $Z_2 = \{0,1\}$. Композиція – додавання, що визначається так: 1+1=0, одиниця групи – нуль.

Далі ми зустрінемось з іншими прикладами груп.

Означення підгрупи. Нехай G – група і $H \subseteq G$. Якщо H є групою і для довільних елементів a, b $\in H$, $ab \in H$, то H називається підгрупою групи G.

Очевидно, сама група G є підгрупою G. Її називають невласною підгрупою, $\{e\} \subset G$ – її підмножина, що складається з одного елемента, одиниці групи, – теж підгрупа групи G. Вона називається тривіальною підгрупою.

Наведемо цікавіші приклади:

1) $\mathbb{C} \supset \mathbb{R} \supset \mathbb{Q} \supset \mathbb{Z}$ (композиція – сума);

2) $\mathbb{C} \setminus 0 \supset \mathbb{R} \setminus 0 \supset \mathbb{Q} \setminus 0$ (композиція – добуток);

3) $n\mathbb{Z} \equiv \{nk, k \in \mathbb{Z}\} \subset \mathbb{Z}$ (композиція – сума) – підгрупа цілих чисел, що діляться на n;

4) числа {1, –1, i, –i} – корені четвертого степеня з одиниці утворюють комутативну абелеву групу, {1, –1} – її підгрупа. Легко бачити, що корені довільного натурального степеня n утворюють абелеву мультиплікативну групу. Якщо $n=2k$ то вона має підгрупу {1, –1};

5) множина $\mathbb{Q}(p)$ всіх p-ічних дробів, що мають вигляд $\frac{m}{pk}$, де p – просте число, k – натуральне, m, – ціле, утворює адитивну комутативну групу, що є підгрупою групи . Доведення очевидне.

Вправа 1.1.1. Нехай A, $B \subset G$ – підгрупи групи G. Довести, що добуток $AB \equiv \{ab$, $a \in A$, $b \in B)$ буде підгрупою в G тоді і тільки тоді, коли A і B комутують. Під комутуванням підгруп будемо розуміти, що для довільних $a_1 \in A$, $b_1 \in B$ знайдуться такі $a_2 \in A$, $b_2 \in B$, що $a_1 b_1 = b_2 a_2$.

Д о в е д е н н я. Необхідність. Нехай AB – підгрупа. Тоді добуток двох елементів з AB, які обов'язково мають вигляд $a_1b_1a_2b_2$, належить AB, тобто теж має такий самий вигляд – ab. Маємо: $a_1b_1a_2b_2 = ab$. Звідси $b_1a_2 = \left(a_1^{-1}a\right)\left(bb_2^{-1}\right) = a_3b_3$, тобто, A і B комутують.

Достатність. Нехай підгрупи A і B комутують. Тоді знайдуться такі $a_3, a_4, a_5 \in A$, $b_3, b_4, b_5 \in B$, що $(a_1b_1)(a_2b_2) = b_3a_3a_2b_2 = b_3a_4b_2 = a_5b_4b_2 = a_5b_5$ і добуток $(a_1b_1)(a_2b_2)$ належить AB. Завдяки комутуванню обернений для ab – елемент $(ab)^{-1}=b^{-1}a^{-1}=a_1b_1$ належить до AB. Одиниця, очевидно, належить до AB. Доведення закінчене.

Розглядаючи більш детально приклад 4 групи коренів четвертого степеня з одиниці, можна помітити, що довільний елемент цієї групи одержуємо як відповідну степінь i: $-1=i^2$, $-i=i^3$, $1=i^4$, тобто група утворена одним елементом i.

Означення. Група називається циклічною з твірним елементом a, якщо довільний її елемент x може бути зображений у вигляді деякого степеня a.

Z є нескінченною циклічною адитивною групою з твірним елементом 1.

У прикладі 3 $H=nZ$ – теж нескінченна циклічна адитивна група з твірним n.

Тут можна зробити цікаве зауваження. Виявляється, що кожна підгрупа H групи Z має вигляд $\{nk, k\in Z\}\equiv nZ$, де n – деяке додатне ціле число, тобто кожна підгрупа Z циклічна. Справді, візьмемо найменше додатне число в H і позначимо його через n. Тоді для довільного $x\in H$, як відомо, існують такі k, $r\in Z$, $0\le r<n$, що $x=nk+r$ (ділення з залишком). Оскільки $r=x-nk$, то $r\in H$, а з припущення, що n – найменше число в H, випливає $r=0$.

Якщо для елемента $a\in G$ існує число m, що $a^m=e$, то m називається показником елемента a. У такому разі, очевидно, $a^{mk}=e$. Найменший з показників елемента, натуральне число d, називається його періодом. Часто його називають порядком елемента.

Легко бачити, що всі степені $a^0=e$, a, $a^2,\ldots, a^{d-1}$ різні і утворюють циклічну підгрупу в G. Дійсно, якщо $a^k= a^m$, $0\le k\le m\le d-1$, то $a^{m-k}=e$ і, оскільки $0\le m-k\le d$, то в силу мінімальності d $m-k=0$, тобто, $m=k$. Якщо елемент a не має періоду, то всі його степені різні і група G необхідно нескінченна. Дійсно, нехай $a^k=a^m$ (для визначенності припустимо, що $k<m$). Тоді одержимо, що $a^{k-m}=e$, а це протирічить припущенню.

Степені довільного елемента, очевидно, утворюють підгрупу, яка називається циклічною підгрупою елемента.

Вправа 1.1.2. Довести, що, коли $ab=ba$, порядки елементів, числа m і k, взаємно прості, то порядок ab дорівнює mk.

Доведення. Помітимо, що $b \neq a^{-1}$ бо в разі $b=a$ $b^m=(a^{-1})^m=(a^m)^{-1}=e^{-1}=e$, що протирічить взаємній простоті m і k. Нехай n – таке найменше натуральне число, що $(ab)^n=e$. Тоді за умовою $(ab)^n=a^n b^n=e$ і $a^n=(b^n)^{-1}=(b^{-1})^n$. Оскільки $b \neq a^{-1}$, то звідси випливає, що $a^n=b^n=e$, тобто n – показник a і b, а тоді n повинно ділитися на m і k, що й треба було довести.

Нехай $H \subset G$ – підгрупа групи G *і* $a \in G$, $a \notin H$. Розглянемо множину, яку будемо позначати aH – це множина всіх елементів вигляду ah, де h – довільний елемент з H, тобто $aH \equiv \{ah,\ h \in H\}$. aH називається лівим класом суміжності (або лівим суміжним класом) елемента a по підгрупі H. Аналогічно будується правий клас суміжності (або правий суміжний клас) Ha.

Доведемо, що в aH всі елементи, які мають вигляд ah_1, ah_2,… – різні. Дійсно, коли $ah_i=ah_j$, то, множачи обидві частини рівності на a^{-1} одержуємо $h_i=h_j$. З доведеного випливає, що кількість елементів в aH (у випадку нескінченних G та H – потужність множини aH) дорівнює порядку підгрупи H.

Нехай $b \notin aH$. Доведемо, що $aH \cap bH=\varnothing$, Дійсно, коли $c \in aH \cap bH$, то знайдуться елементи $h, f \in H$, що $c=ah=bf$. Звідси випливає, що $b=ahf^{-1}$. Оскільки $hf^{-1} \in H$, то $b \in aH$, а це протирічить припущенню.

Очевидно, все сказане стосується і правих класів суміжності.

З наведених міркувань випливає, що група G може бути зображена як об'єднання лівих (чи правих) класів суміжності по підгрупі H: $G=\{H, a_1H,\ldots,a_kH,\ldots\}$ (або $G=\{h,Hb_1,\ldots,Hb_k,..\}$). Ці класи не перетинаються і мають однакову потужність, яка дорівнює потужності підгрупи H. Потужність множини класів суміжності називається індексом підгрупи. Нехай група G

скінченна, її порядок дорівнює n, порядок H дорівнює m, а індекс H дорівнює k. В такому разі з попередніх міркувань випливає, що $n=mk$. Таким чином, одержали важливий результат.

Теорема 1.1.1 (Лагранжа). Порядок групи ділить порядок підгрупи.

Іноді порядок групи позначають $(G:1)$, що символічно відповідає розкладу G по тривіальній підгрупі $\{e\}$. Аналогічно $(H:1)$ – порядок H, $(G:H)$ – індекс H. Доведена рівність в таких позначеннях має вигляд

$$(G:1)=(G:H)(H:1). \tag{1.1.1}$$

Наслідок. Порядок скінченної групи ділиться на період кожного свого елемента. Дійсно, період елемента збігається з порядком періодичної підгрупи, утвореної цим елементом, і за теоремою Лагранжа він є дільником порядку групи. Тобто кожен елемент в степені порядку групи дорівнює одиниці.

Наслідок. Група простого порядку p не має підгруп, крім одиничної та невласної – самої себе; більше того, вона циклічна, і її твірним є кожен елемент групи. Дійсно, період кожного елемента підгрупи повинен ділити порядок групи, а в разі простого порядку цей період може бути лише одиницею, що відповідає одиниці e групи, або дорівнювати p. Отже, всі степені довільного елемента різні, і вони утворюють всю групу.

З 'ясуємо, які підгрупи може мати циклічна група непростого порядку.

Теорема 1.1.2. Підгрупа циклічної групи циклічна.

Доведення. Якщо група нескінченна, цей факт доводиться так само, як і у випадку групи Z. Якщо група скінченна, то, як уже відомо, для того, щоб вона мала нетривіальні підгрупи, її порядок необхідно повинен бути непростим. Нехай H – підгрупа в G, $a\in G$ – твірний елемент G і m – найменший його натуральний степінь, що $a^m\in H$. Доведемо, що a^m – твірний в H, тобто $H=\{a^{mi}\}$. Якщо існує елемент $a^k\in H$ $(k>0)$ і $k\neq mi$, то, розділивши k з залишком на m– $k=mq+r$, $(r<m)$, одержимо, що елемент $a^{k-mq}=a^k(a^{mq})^{-1}=a^r\in H$, а це протирічить мінімальності m.

В подальшому нам буде потрібна властивість класу суміжності, яка полягає в тому, що його зображення не залежить від представника цього класу, тобто коли $a,\ a_1 \in aH$, то $aH=a_1H$, тому що для $h \in H$ знайдеться такий елемент $h_1 \in H$, що $ah=a_1h_1$ – треба взяти $h_1=aha_1^{-1}$.

Зауважимо, що, взагалі кажучи, ліві та праві класи суміжності не збігаються: $aH \neq Ha$.

В наступному параграфі розглянемо випадок, коли $aH=Ha$ для довільного $a \in G$.

1.2. НОРМАЛЬНІ ПІДГРУПИ ТА ГОМОМОРФІЗМИ

Означення. Якщо $aH=Ha$ для довільного $a \in G$ підгрупа H називається нормальною або інваріантною, або нормальним дільником.

Останню рівність слід розуміти таким чином: для довільного $a \in G$ існують елементи $h_1,\ h_2 \in H$, що $ah_1=h_2a$.

З означення H випливає, що $H=a^{-1}Ha$ (або $H=aHa^{-1}$). Часто таку рівність приймають за означення нормальної підгрупи. З неї випливає, що при $h \in H$ $g^{-1}hg \in H$ для довільного $g \in G$. Елемент $g^{-1}hg$ називається спряженим до h. Отже, підгрупа є нормальною тоді і тільки тоді, коли вона разом з кожним своїм елементом вміщає всі спряжені до нього.

Очевидно, тривіальна підгрупа $\{e\}$ нормальна. Вся група як своя невласна підгрупа теж нормальна. Дійсно, коли $a,g \in G$, то слід взяти $g_1=aga^{-1}$ і в такому разі $ag=g_1a$.

Означення. Група називається простою, якщо вона має лише дві нормальні підгрупи – $\{e\}$ і саму себе.

Розглянуті приклади нормальних дільників тривіальні і незмістовні. Наведемо більш цікавий приклад.

Означення. Підмножина C елементів групи G називається її центром, якщо її елементи комутують з усіма елементами з G.

Зауважимо, що в деяких групах $C=\{e]$. Легко бачити, що C – комутативна підгрупа в G, тому що її елементи комутують між собою; якщо $a,b \in C$, тобто для довільного $g\in G$ $ag=ga$, $bg=gb$, то $(ab)g=agb=g(ab)$. В такому разі елемент ab комутує з довільним елементом з G, а тому за означенням належить до C. Очевидно, $e\in C$, тому що за означенням одиничний елемент комутує з довільним елементом групи і діє на нього тотожно. Якщо $a\in C$, то $a^{-1}\in G$. Дійсно, для довільного $g\in G$ $ag=ga$, звідки $aga^{-1}=g$, і $ga^{-1}=a^{-1}g$, що й треба було довести.

З означення центра групи випливає, що для довільного $a\in G$ $aC=Ca$, оскільки це правильно для довільного $h\in C$.

Означення. Комутатором елементів a_1, a_2 називається елемент b, який визначається так: $b=a_1a_2a_1^{-1}a_2^{-1}$.

З цього означення випливає, що a_1 і a_2 комутують тоді і лише тоді, коли їх комутатор дорівнює одиниці групи.

Зауважимо, якщо G – комутативна група, то кожна її підгрупа нормальна.

Якщо $H\subset G$ – нормальна підгрупа, то можна ввести групову операцію на множині лівих класів суміжності G по H. Дійсно, $(aH)(bH)=aHbH=abHH=(ab)H$, тобто добуток суміжних класів – знову суміжний клас. Як ми знаємо з попереднього, зображення суміжного класу не залежить від його представника, тому замість групової операції над класами досить розглядати групову операцію для довільних представників цих класів. Роль одиничного елемента на множині суміжних класів грає сам нормальний дільник H, роль оберненого до aH - клас $a^{-1}H$, тому що $(aH)\left(a^{-1}H\right)=\left(aa^{-1}\right)HH=H$.

Теорема 1.2.1. Нехай H – нормальна підгрупа групи G. Тоді множина лівих суміжних класів групи G по H утворює групу. Ця група називається факторгрупою групи G по нормальній підгрупі H і позначається G/H.

Означення. Нехай G_1,G_2 – дві групи і існує відображення φ: $G_1\to G_2$, визначене на всій групі G_1, що для довільних $a,b\in G_1$ $\varphi(ab)=\varphi(a)\varphi(b)$. Таке відображення називається гомоморфізмом. Його образ, тобто $\varphi\left(G_1\right)\subset G_2$ позначається $\operatorname{Im}(\varphi)$.

Зазначимо, що він переводить групову операцію з G_1 в групову операцію в G_2. При гомоморфізмі одиниця переходить в одиницю, обернений елемент – в обернений. Дійсно, для довільного $x \in G_1$ $\varphi(xe_1) = \varphi(x) = \varphi(x)\varphi(e_1)$, звідки $\varphi(e_1) = e_2$. Далі, $e_2 = \varphi(e_1) = \varphi(xx^{-1}) = \varphi(x)\varphi(x^{-1})$, звідки $\varphi(x^{-1}) = (\varphi(x))^{-1}$. З цих міркувань бачимо, що образ гомоморфізму (позначається $Im(\varphi)$) утворює групу, яка є підгрупою в G_2 або може збігатися з нею.

Множина елементів, що переходить при гомоморфізмі в одиницю (вона позначається через $Ker(\varphi)$), утворює підгрупу в G. Дійсно, коли $a,b \in Ker(\varphi)$, то за означенням гомоморфізму $\varphi(ab) = \varphi(a)\ \varphi(b) = e$, тобто $ab \in Ker(\varphi)$. Якщо $a \in Ker(\varphi)$, то з попередніх міркувань $\varphi(a^{-1}) = (\varphi(a))^{-1} = e$, тобто, $a^{-1} \in Ker(\varphi)$.

Коли гомоморфізм φ має обернене відображення, що визначене на всій групі G_2, яке є гомоморфізмом, *то* φ називається ізоморфізмом, а відповідні групи ізоморфними. Цей обернений гомоморфізм позначається через φ^{-1} . Факт ізоморфізму G_1 і G_2 позначається так: $G_1 \simeq G_2$.

Гомоморфізм групи в себе називається ендоморфізмом, а ізоморфізм на себе – автоморфізмом.

Теорема 1.2.2. Ядро гомоморфізму – нормальна підгрупа групи G.

Доведення. Нехай $\varphi : G \to G'$ - гомоморфізм групи G в групу G'. Позначимо через e' одиницю групи G'. Нехай $h \in Ker(\varphi)$. Тоді $\varphi(h) = e'$. Доведемо, що $g^{-1}hg \in H$ для довільного $g \in G$. З означення гомоморфізму $\varphi(g^{-1}hg) = \varphi(g^{-1})\varphi(g) = (\varphi(g))^{-1}\varphi(g) = e'$, що і доводить твердження.

В умовах цієї теореми можна розглянути факторгрупу $G/Ker(\varphi)$.

Теорема 1.2.3. Нехай $G \to G'$ - гомоморфізм груп. Тоді фактор-група групи G по ядру гомоморфізму ізоморфна його образу:

$$G/Ker(\varphi)=Im(\varphi). \qquad (1.2.1)$$

Доведення. Розглянемо прообраз елемента $a \in Im(\varphi)$, тобто всі елементи $x \in G$, що $\varphi(x) = a$, і доведемо, що вони утворюють клас суміжності по $Ker(\varphi)$. Якщо *x,y* належать одному класу суміжності по $Ker(\varphi)$, то $x=yh$ при деякому $h \in Ker(\varphi)$. Тоді $\varphi(x) = \varphi(yh) = \varphi(y)\varphi(h) = \varphi(y)$. Навпаки, коли $\varphi(x) = \varphi(y)$, то $e' = \varphi(x)(\varphi(y))^{-1} = \varphi(x)\varphi(y^{-1}) = \varphi(xy^{-1})$, звідки $xy^{-1} \in Ker(\varphi)$. Тоді існує $h \in Ker(\varphi)$ такий, що $xy^{-1} = h$, звідки $x = hy$, тобто, x і y належать одному класу суміжності по $Ker(\varphi)$. В результаті одержали взаємно однозначну відповідність між класами суміжності по нормальній підгрупі $Ker(\varphi)$ і елементами з $Im(\varphi)$. Нехай $\varphi(x) = a, \varphi(y) = b$, тоді $\varphi(xH) = a, \varphi(yH) = b$ і за означенням гомоморфізму

$$\varphi\big((xH)(yH)\big) = \varphi(xH)\varphi(yH) = ab,$$

що й треба було довести. ■

Виникає питання, чи кожна нормальна підгрупа може бути ядром деякого гомоморфізму. Досить побудувати гомоморфізм групи *G* на факторгрупу по цій підгрупі – $G \to G/H$. Цей гомоморфізм кожному елементу з *G* ставить у відповідність його клас суміжності (лівий чи правий) по цій нормальній підгрупі, ядро якого збігається з *Я*. Побудований гомоморфізм називається природним або канонічним. Таким чином, теорема 1.2.3 стверджує, що довільний гомоморфізм групи *G* на деяку групу G_1 (образ – вся група G_1) по суті не відрізняється від канонічного гомоморфізму на свою факторгрупу по ядру цього гомоморфізму.

Приклади.

1. *G* – абелева мультиплікативна група, $\varphi(x) = x^n \ \forall x \in G, n \in \mathbb{Z}_+$ – гомоморфізм *G* в *G*. Дійсно, $\varphi(xy) = (xy)^n = x^n y^n = \varphi(x)\varphi(y)$.

Якщо $G=\mathbb{R}\setminus\{0\}, n=2k+1,$ то $Im(\varphi)=G$ і групи ізоморфні; при $n=2k$ $\mathrm{Im}(\varphi)=\mathbb{R}_+$.

2. G – одна з адитивних груп $\mathbb{Z},\mathbb{Q},\mathbb{R},\mathbb{C}$, a – довільне число (відповідно з кожної з груп). Радимо перевірити, чи $\varphi: x\to x+a$ – гомоморфізм.

3. $G_1=G_2=\mathbb{R}\setminus 0$; $\varphi(x)$ – знак числа x.

4. $G_1=\mathbb{R}_+$ – мультиплікативна група. $G_2=\mathbb{R}$ – адитивна група. Гомоморфізм визначається за допомогою функції $\ln$: $\ln(xy)=\ln x+\ln y$, $\ln 1=0$. Обернений гомоморфізм дається функцією *exp*: $\exp(0)=1, \exp(x+y)=\exp(x)\exp(y)$. Отже, групи ізоморфні.

5. $\mathbb{Q}(p)$ і $\mathbb{Q}(q)$ неізоморфні при $p\neq q$.

Доведення. Нехай є деякий гомоморфізм φ: $\mathbb{Q}(p)\to\mathbb{Q}(q)$, при якому $\varphi(1)=\dfrac{a}{q^k}$, $a=bp^n$. Тоді $\varphi(1)=\varphi\left(p^{n+1}\dfrac{1}{p^{n+1}}\right)=$ $=p^{n+1}\varphi\left(\dfrac{1}{p^{n+1}}\right)=\dfrac{a}{q^k}=\dfrac{bp^n}{q^k}$ звідки $\varphi\left(\dfrac{1}{p^{n+1}}\right)=\dfrac{bp^n}{q^k p^{n+1}}=\dfrac{b}{q^k p^n}$, тобто образ елемента $\dfrac{1}{p^{n+1}}\in\mathbb{Q}(p)$ не належить $\mathbb{Q}(q)$ при $p\neq q$.

Означення. Ланцюг гомоморфізмів груп

$$G_1\xrightarrow{g}G_2\xrightarrow{f}G_3 \tag{1.2.2}$$

називається точним в G_2, якщо $Im(f)=Ker(g)$.

Наведемо декілька прикладів для адитивних груп.

1. $0\xrightarrow{g}\mathrm{G}\xrightarrow{\mathrm{f}}F$. Точність в G означає, що $\mathrm{Im(g)=0=Ker}(f)$, тобто ядро гомоморфізму f нульове, і в такому разі f – взаємно однозначне відображення G на свій образ в F. Таке відображення називається вкладенням, або ін'єкцією.

2. $F \xrightarrow{f} \mathrm{G} \xrightarrow{g} 0$. Точність в G означає, що $Ker(g)=G=Im(f)$, тобто f – відображення на всю групу G. Таке відображення називається сюр'єкцією.

3. $0 \to \mathrm{G}_1 \xrightarrow{f} G_2 \xrightarrow{g} G_3 \to 0$. Точність цієї послідовності в кожному члені означає, що f – вкладення, тобто $Im(f) \cong \mathrm{G}_1$ і g – сюр'єкція, і $Ker(g)=Im(f) = G_1$. Оскільки за теоремою 1.2.3 $\mathrm{G}_3=\mathrm{G}_2/\mathrm{Ker(g)}=\mathrm{G}_2/\mathrm{Im}(f)=\mathrm{G}_2/\mathrm{G}_1$, то розглядувана послідовність – те ж саме, що і точна послідовність

$$0 \to G_1 \xrightarrow{f} G_2 \xrightarrow{g1} G_2/G_1 \to 0, \qquad (1.2.3)$$

де $g'=ig$ (i – ізоморфізм між G_3 і G_2/G_1).

Тепер дамо конструкцію прямого добутку груп.

Нехай G_1,G_2 – групи. Розглянемо їх прямий (декартовий) добуток як множин, тобто множину пар $\{x_1, x_2: x_1 \in G_1, x_2 \in G_2\}$. Визначимо для таких елементів "покомпонентне" множення:

$$(x_1, x_2)(y_1,y_2) = (x_1y_1, x_2y_2). \qquad (1.2.4)$$

Таке множення, очевидно, асоціативне, тому що асоціативне множення в кожній з компонент; елемент $(e_1,e_2) \equiv e$, де e_1 – одиниця в G_2, є одиниця в G_2 є одиниця відносно множення, визначеного за допомогою (1.2.4); оберненим для (x_1, x_2) буде, очевидно, елемент (x_1^{-1}, x_2^{-1}).

Отже, побудовано групу, що називається прямим добутком груп G_1 і G_2 і позначається G_1 x G_2.

Властивості зовнішнього добутку.

1. Множина $G_1 \times e_2 \equiv \{(x, e_2), x \in G_1\}$ – нормальна підгрупа в $G_1 \times G_2$, що ізоморфна G_1. Дійсно, добуток елементів такого вигляду є знову елемент такого ж вигляду: $(x_1, e_2)(y_1, e_2) = (x_1y_1, e_2)$. Ізоморфізм визначається відображенням $(x_1, e_2) \to x_1$, що, очевидно, взаємно однозначне і переводить групову операцію

з $G_1 \times G_2$ в групову операцію в G_1. Позначимо так побудовану підгрупу в $G_1 \times G_2$ через G_1. Для доведення того факту, що G_1 – нормальна підгрупа, знайдемо такий добуток:

$$(y_1, y_2)^{-1}(x_1, e_2)(y_1, y_2) = (y_1^{-1}, y_2^{-1})(x_1, e_2)(y_1, y_2) = (y_1^{-1}x_1y_1,\ y_2^{-1}e_2y_2) = (y_1^{-1}x_1y_1,\ e_2)$$

Останній вираз належить G_1, що й доводить твердження.

2. Аналогічні властивості має підгрупа $G_2 = \{(e_1, x_2),\ x_2 \in G_2\}$.

3. Елементи з підгруп G_1 *і* G_2 комутують між собою.

Дійсно, $(x_1, e_2)(e_1, x_2) = (x_1, x_2)$ і $(e_1, x_2)(x_1, e_2) = (x_1, x_2)$.

4. $G_1 \cap G_2 = e$. Доведення очевидне.

5. $G_1\ G_1G_2 = G_1 \times G_2$. Дійсно, довільний елемент з $G_1 \times G_2$ має вигляд (x_1, x_2). Він дорівнює $(x_1, e_2)(e_1, x_2)$.

Зауважимо, що таким самим чином можна визначити добуток довільної скінченної кількості груп.

Виникає питання: чи може деяка група бути подана у вигляді добутку своїх підгруп?

Теорема 1.2.4. Нехай H_1, H_2 – нормальні підгрупи групи G і $H_1 \cap H_2 = e$. Тоді елементи з H_1 комутують з елементами з H_2.

Доведення. Розглянемо комутатор $b = a_1a_2a_1^{-1}a_2^{-1}$ елементів $a_1 \in H_1,\ a_2 \in H_2$ і $a_2 \in H_2$. Запишемо його у вигляді $b = a_1(a_2a_1^{-1}a_2^{-1})$. Оскільки перший співмножник належить H_1 *і* другий теж належить H_1 завдяки нормальності H_1, то $b \in H_1$. Але b можна подати у вигляді $b = (a_1a_2a_1^{-1})a_2^{-1}$. Аналогічні міркування доводять, що $b \in H_2$. Отже, $b \in H_1 \cap H_2 = e$, а тому ці елементи комутують.

Теорема 1.2.5. Нехай *H₁* і *H₂* – нормальні підгрупи групи *G*, $H_1 \cap H_2 = e$ і $H_1H_2 = G$. Тоді *G* ізоморфна прямому добутку цих підгруп.

Доведення. Розглянемо прямий добуток $H_1 \times H_2$ і зіставимо кожен його елемент (x_1,x_2) з елементом x_1x_2 групи G. Це відображення – гомоморфізм. Дійсно, добуток пар $(x_1,x_2)(y_1,y_2) = (x_1y_1,x_2y_2)$ зіставляється з елементом $x_1y_1x_2y_2 \in G$. Але в силу теореми 1.2.4 $x_1y_1x_2y_2 = x_1x_2y_1y_2 = (x_1x_2)(y_1y_2)$, тобто добутку пар відповідає добуток їх образів при відображенні. Це відображення є сюр'єкція, тобто на всю групу, бо $H_1H_2 = G$. Доведемо, що воно взаємно однозначне. Для цього досить довести, що два різних елементи (x_1,x_2) і (y_1y_2) переходять в два різних елементи x_1x_2 *і* y_1y_2. При $x_1x_2 = y_1y_2$ $y_1^{-1}x_1 = y_2x_2^{-1}$. Ліва частина рівності належить H_1, права – H_2, тому завдяки умові $H_1 \cap H_2 = e$ $y_1^{-1}x_1 = y_2x_2^{-1} = e$, звідки $x_1 = y_1$, $x_2 = y_2$. Таким чином, відображення $(x_1,x_2) \rightarrow x_1x_2$ є ізоморфізм групи $H_1 \times H_2$ і G.

Зауваження. Якщо групи $G_1,G_2, \ldots ,G_n$ комутативні, то їх прямий добуток G будемо називати сумою і вживати таке позначення: $G = G_1 \oplus G_2 \oplus \ldots \oplus G_n$.

1.3. ВІЛЬНІ ГРУПИ

Розглянемо пряму суму адитивних груп цілих чисел: $\mathbb{Z}^n = (\mathbb{Z},\mathbb{Z},\ldots,\mathbb{Z})$. Це комутативна адитивна група з покомпонентною операцією додавання.

Означення. Абелевою вільною групою називається група, яка ізоморфна $\mathbb{Z}^n$.

В $\mathbb{Z}^n$ є твірні елементи e_1 - $(1,0,\ldots0)$, $e_2 = (0,1,0,\ldots0)$, $e_n = (0,0,\ldots0,1)$, тобто такі, що кожен елемент $x = (x_1,\ldots,x_n) \in \mathbb{Z}^n$ можна подати як суму деякого числа цих елементів, а саме $x = \Sigma\, x_ke_k$. Очевидно, між ними немає ніяких співвідношень вигляду $\Sigma\, x_ke_k - 0$ при ненульових x_k, тобто така сума дорівнює нулю тоді і лише тоді, коли $x_1 = \ldots = x_n = 0$. В такому разі кажуть, що система $\{e_1,\ldots,e_n\}$ лінійно незалежна, а така система твірних називається базисом вільної групи. Якщо G ізоморфна $\mathbb{Z}^n$, то це значить, що в G теж можна знайти деякі елементи f_1, f_2, ..., f_n, що кожен елемент $g \in G$ можна подати у вигляді аналогічної суми, яка називається лінійною комбінацією f_k. Отже, такі елементи твірні. Важливою їх властивістю є те, що між ними не може бути ніяких співвідношень вигляду

$\Sigma a_k f_k = 0$, оскільки їх немає між $e_1.....e_n$, тобто $\{f_1, f_2, ..., f_n\}$ – базис групи. Очевидно, базис визначається неоднозначно.

Отже, виходячи з довільної скінченної множини $S = \{s_1, ..., s_n]$, можна побудувати вільну групу $F(S)$, формально визначаючи групову операцію, що зручно вважати додаванням. Вона, як і раніше, полягає в покомпонентному формальному додаванні елементів s_k. В результаті кожен елемент $F(S)$ має вигляд $(a_1s_1, ..., a_ns_n)$, де $a_к$ – цілі числа.

Нехай M,N – множини, $\lambda : M \to N$ – відображення множин, $f_м : M \to F(M)$ (відображення M в вільну групу $F(M)$), $f_n : N \to F(N)$. Тоді маємо діаграму

$$\begin{array}{ccc} M & \overset{F_M}{\to} & F(M) \\ \lambda\downarrow & & \downarrow F(\lambda) \\ N & \overset{F_N}{\to} & F(N) \end{array} \tag{1.3.1}$$

Відображення F (λ) визначається так, щоб діаграма була комутативною, тобто щоб $F(\lambda)f_m = f_n : N \to F(N)$ або щоб обидва шляхи, які визначаються відповідними стрілками, приводили з M до F/N.

Якщо відображення λ сюр'єктивне, то F (λ) – теж сюр'єктивне. Дійсно, нехай $b \in F(N)$. Це означає, що існує елемент $n \in N$ такий, що $f_N(n) = b$. Але з сюр'єктивності λ випливає, що існує $m \in M$, що $\lambda(m) = n$. Звідси маємо, що коли подіяти відображенням F (λ) на елемент $f_M(m) = a \in F(M)$, то одержимо b в силу комутативності діаграми.

Теорема 1.3.1. Всяка комутативна група є факторгрупою деякої вільної групи.

Доведення. Розглянемо G як множину елементів без групової структури, що будемо позначати через S, $S = \{g_1, ..., g_n\}$. Побудуємо вільну групу $F(S)$ з базисом $g_1 = (g_1, 0, ..., 0), ..., g_n = (0, ..., 0, g_n)$.

Доведемо існування єдиного сюр'єктивного гомоморфізму $\xi : F(S) \to G$ такого, що ξ (g_k) = g_k. Покажемо, що відображення, задане лише на базисі, можна продовжити до гомоморфізму $F(S) \to G$. Покладемо $\xi(\Sigma x_k g_k) = \Sigma x_k \xi(g_k) = \Sigma x_k g_k$. Нехай тепер $a = \Sigma x_k g_k$, $b = \Sigma y_k g_k \in F(S)$. *Тоді*

$$\xi(a)+\xi(b)=\xi(\Sigma x_k g_k+\Sigma y_k g_k)=\xi(\Sigma(x_k+y_k)g_k)=\Sigma(x_k+y_k)\ \xi\ (g_k) = \xi(\Sigma(x_k+y_k)g_k) = \Sigma x_k g_k+\Sigma y_k g_k=\Sigma x_k \xi(g_k) + \Sigma y_k \xi(g_k)=\xi(\Sigma x_k g_k)+\xi(\Sigma y_k g_k)=\xi(a+b),$$

а це й означає, що ξ – гомоморфізм. Якщо є ще один гомоморфізм $\eta : F(S) \to G$ такий, що $\eta(g_k) = \xi(g_k)$, то з властивості гомоморфізму випливає

$$\eta(\Sigma x_k g_k)=\Sigma x_k \eta(g_k)=\Sigma x_k g_k=\Sigma x_k \zeta(g_k)=\zeta(\Sigma x_k g_k),$$

тобто $\zeta = \eta$. В такому разі за теоремою 1.2.3 $G \cong F(S)/\mathrm{Ker}(\zeta)$, що й треба було довести.

Приклад 1. Щоб краще зрозуміти доведену теорему, розглянемо, як можна побудувати адитивну групу $\mathbb{Z}_2 = \{0,1\}$. Розглянемо вільну групу $\mathbb{Z}$ і її підгрупу $2\mathbb{Z}$. Тоді, очевидно, $\mathbb{Z}_2 = \mathbb{Z}/(2\mathbb{Z})$. Взагалі кажучи, можна побудувати групу $\mathbb{Z}_n = \mathbb{Z}/(n\mathbb{Z})$, яка називається групою залишків за модулем n. Щоб знайти суму двох елементів $a,b \in \mathbb{Z}_n$, треба знайти залишок від ділення звичайної суми $a + b$ на n. Отже, якщо позначити суму в $\mathbb{Z}_n$ символом $\oplus$, то для $0 \le a, b \le n-1$ $a \oplus b = a + b$, якщо $a + b < n$, і $a \oplus b = a + b - n$, якщо $a + b \ge 0$. Для всіх інших a,b треба суму визначати як суму їх залишків за модулем n. Така ситуація вже обговорювалась при вивченні поняття факторгрупи: щоб визначити операцію над суміжними класами, досить визначити її для представників цих класів. Якщо два елементи a,b належать одному класу (нагадаємо: це значить, що їх залишки від ділення на n рівні або різниця ділиться на n без залишку), то часто, особливо в теорії чисел, цей факт записують так: $a \cong b (\mathrm{mod}\, n)$ і читають *"a порівняне з b по модулю n"*. Отже, можна вважати, що $\mathbb{Z}_n = \{0,1,\dots,n-1\}$. Згідно з теоремою Лагранжа сума n однакових доданків дорівнює нулю (тотожному елементу $\mathbb{Z}_n$). Якщо n – непросте натуральне число, $n = p_1 p_2 \dots p_n$, то кожен елемент p_{i1} ... p_{ik} при $k < n$ має порядок, менший за n. Наприклад, якщо позначити $n = p_1 q$, то елемент p_1 має порядок q, оскільки сума q доданків $p_1 + p_1 + ... + p_1 = p_1 q = n \equiv 0 (\text{mod } n)$ і за означенням $p_1 \oplus p_1 \oplus \ldots \oplus p_1 = 0$. Навпаки, коли a і n взаємно прості, то ні при якому $q < n$ не може виконуватися рівність $aq = 0$ в $\mathbb{Z}_n$, тобто всі елементи $a, 2a, ..., (n\text{-}1)a$ – різні, і такий елемент є твірний в $\mathbb{Z}_n$, група циклічна. Отже, всі елементи, які відповідають числам, меншим за n і взаємно простим з n, починаючи з 1, є твірним в $\mathbb{Z}_n$. Наприклад, 1 – твірний елемент. Кількість таких елементів обчислює функція Ейлера φ: $\varphi(n)$ за означенням дорівнює кількості натуральних чисел, менших за n, починаючи з одиниці. Наприклад, $\varphi(i) \equiv 1$, $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$, $\varphi(5) = 4$, $\varphi(6) = 2$, $\varphi(7) = 6$, $\varphi(8) = 4$, $\varphi(9) = 6$, 4, $\varphi\{10) = 4$, $\varphi(11) = 10$, $\varphi(12) = 4$ і т.ін.

Зауваження. Якщо n – просте число, то $\varphi(n)$ = n - 1. Доведення очевидне.

П р и к л а д 2. Група коренів степеня n (n натуральне) I_n. Корені степеня n утворюють групу, бо якщо ε_1, ε_2 – корені степеня n, тобто $(\varepsilon_1)^n = (\varepsilon_2)^n = 1$, то $(\varepsilon_1\varepsilon_2)^n = (\varepsilon_1)^n(\varepsilon_2)^n = 1$, $1^n = 1$, для кожного кореня ε стєпєня n
обернене число ε^{-1} – теж корінь степеня n: $(\varepsilon^{-1})^n = (\varepsilon^n)^{-1} = 1$. Як відомо, кількість різних коренів степен n дорівнює n. Їх можна знайти за формулою Ейлера

$$\varepsilon_k = \exp\left(i\frac{2\pi k}{n}\right), k = 0, 1, ..., n-1 \tag{1.3.2}$$

При множенні різних коренів їх показники додаються за правилом додавання в групі $\mathbb{Z}_n$, кожному $k \in \mathbb{Z}_n$ взаємно однозначно відповідає корінь ε_k, групова операція в I_n переходить в групову операцію в $\mathbb{Z}_n$. Твірними в I_n є ε_k, якщо k взаємно просте з n. Такі корені інколи називають первісними.

На завершення короткого екскурсу в загальну теорію груп з'ясуємо, які існують групи до п'ятого порядку включно. Будемо вважати, що групова операція – множення, але це не суттєво.

Очевидно, існує тільки по одній групі першого, другого, третього та п'ятого порядку, що випливає з теореми Лагранжа – порядок кожного елемента ділить порядок групи, а числа 2, 3 і 5 – прості, тому групи циклічні, вони ізоморфні відповідним групам коренів з одиниці, група першого порядку-складається з єдиного одиничного елемента. Існує також циклічна група четвертого порядку, що також ізоморфна групі коренів четвертого порядку з одиниці. Але оскільки $4 = 2 \times 2$, то можна сподіватися, що існують групи з елементами другого порядку. Отже, нехай є елементи *e, a,* де *a* такий, що $a^2 = e$, тобто $a^{-1} = a$. Очевидно, ці елементи не вичерпують групу. Отже, існує ще деякий елемент *b*, який не збігається ні з е, ні з *a.* За теоремою Лагранжа його порядок дорівнює 2 (він не може дорівнювати 4, бо тоді в групі було б більше чотирьох елементів – степеня елемента *b* і елемент а) і $b^{-1} = b$. За означенням групи в ній повинен існувати елемент $c = ab$, що повинен мати теж другий порядок: $c^2 = (a\ b)^2 = e$. Звідси випливає, що $abab = e$. Множачи зліва та справа цю рівність відповідно на *a* і *b* і враховуючи, що $a^2 = b^2 = e$, одержимо: $ba = ab = c$, тобто група виявляється комутативною. Легко бачити, що вона є прямим добутком двох підгруп другого порядку $G = \{e, a\} \times \{e, b\}$ з такою відповідністю (будемо вживати для елементів *G* і для елементів прямих співмножників, але це не повинно призвести до непорозуміння):
$e \Leftrightarrow (e,e)$, $a \Leftrightarrow (a,e)$, $b \Leftrightarrow (e,b)$, $c \Leftrightarrow (a,b)$. Побудована група називається четвертою групою Кляйна.

1.4. ДІЯ ГРУПИ НА МНОЖИНІ

Нехай *S* – множина і *G* – група. Розглянемо множину перетворень множини *S*. Серед них є такі, що переставляють місцями елементи цієї множини. Такі перетворення називають перестановками. Серед перестановок є тотожна, яка не змінює порядок елементів множини. Якщо виконано якусь перестановку *S*, то обернена перестановка приводить елементи *S* до попереднього стану. Отже, перестановки множини *S* утворюють групу.

Нехай кожному елементу $x \in G$ відповідає деяке перетворення f_x множини S. Це можна записати у вигляді $x \to f_x$: $S \to S$. Крім того, будемо вважати, що відображення $x \to f_x$ задовольняє умову

$$f_{xy} = f_x f_y \qquad (1.4.1)$$

для всіх $x, y \in G$.

При цьому f_e – тотожне перетворення, $f_{x^{-1}}$ – обернене для f_x. Отже, такі перетворення – перестановки елементів S.

Відображення (1.4.1) – гомоморфізм G в групу перестановок множини S, яке називається зображенням групи G групою перестановок множини S. При цьому будемо говорити, що визначена дія групи G на множині S.

Цікаво розглянути випадок, коли $S=G$, тобто дію групи на собі.

Приклади. 1. Нехай для $x, a \in G$ задано перетворення a_x: $G \to G$ формулою

$$\sigma_x(a) = xax^{-1} \qquad (1.4.2)$$

яке називається спряженням або трансформуванням. Перевіримо, що ця формула визначає дію G на собі. Дійсно, для $x \quad y \in G \sigma_{xy}(a) = (xy)a(xy)^{-1} = yx(yay^{-1})x^{-1} = \sigma_x\sigma_y(a)$, що відповідає (1.4.1). Крім того, виявляється, що σ_x – автоморфізм G. Дійсно, для x, $y, z \in G$ $\sigma_x(yz) = x(yz)x^{-1} = xyx^{-1}xzx^{-1} = (xyx^{-1})(xz^{-1}) = \sigma_x(y)\ \sigma_x(z)$, $(\sigma_x)^{-1} = \sigma^{-1}$, бо якщо $\sigma_x(a)=xax^{-1}$, т о $((\sigma_x)^{-1})(xax^{-1}) = a$ і $\left(xax^{-1}\right) = x^{-1}\left(xax^{-1}\right)\left(x^{-1}\right)^{-1} = x^{-1}xax^{-1}x = a$.

Таким чином, відображення $x \to \sigma_x$ є гомоморфізм G в її групу автоморфізмів. Ядро цього гомоморфізму – нормальна підгрупа в G, яка складається з таких елементів x $\in$ G, що σ_x діє тотожно на кожному елементі $y \in G$: $o_x(\mathrm{y}) = xyx^{-1} = y$, тобто з усіх x, які комутують з кожним елементом з G, а це значить – ядро гомоморфізму є центром групи G.

Можна розглянути також дію групи *G* за допомогою спряжень на множині своїх підмножин, що також визначається за формулою (1.4.2): для підмножини *A* з *G* $\sigma_x(A) = xAx^{-1} \equiv A_x$ – теж підмножина в *G*. Якщо *A* – підгрупа, то xAx^{-1} – теж підгрупа. Дійсно, для двох елементів xax^{-1}, xbx^{-1} з xAx^{-1} добуток $(xax^{-1})(xbx^{-1})$ = $xabx^{-1}$ теж має такий же самий вигляд, тобто належить до xAx^{-1}, $e = xex^{-1}$, $\left(xax^{-1}\right)^{-1} = xa^{-1}x^{-1}$. Отже, група діє за допомогою спряжень також на множині своїх підгруп.

2. Зсув. Для кожного x $\in$ *G* визначимо операцію зсуву або просто зсув L_x (лівий) на елемент x: $L_x(y) = xy$ $(y \in G)$. Правий зсув R_x визначається за формулою $R_x(y) = xy$.

Переконаємось, що лівий зсув – дія групи на собі (правий зсув читач повинен розглянути самостійно). Дійсно,

$L_xLy(a) = L_x(ya) = (xy)a = L_{xy}(a)$, $L_\varepsilon(a) = ea = a$,

тому Le – тотожне відображення, $(L_x)^{-1} = L_x^{-1}$, тому що $\left(L_{x^{-1}}L_x\right)a = L_{x^{-1}}(xa) = x^{-1}xa = a$, отже, добуток $L_{x^{-1}}L_x$ є тотожне відображення L_e.

Але L_x – не гомоморфізм групи *G*, це деяка перестановка цієї групи.

Як і у попередньому випадку, можна розглянути дію групи за допомогою зсувів на множині своїх підмножин. Зазначимо, що коли $N \subset G$ – підгрупа, то xH – не є підгрупою, а лівий клас суміжності по *H*. Тобто, група діє таким чином на множині лівих суміжних класів.

Як діє група на собі за допомогою зсувів, можна зрозуміти на прикладі адитивної групи $G = R : L_x(a) = x+a$ – це звичайний зсув точки на прямій на величину *a*; $L_{vt}(a) = vt + a$ – визначає положення в момент часу *t*, коли точка рухається з швидкістю *v*; $L_x(A) = x + A$ – зсув множини на вектор *x*. Аналогічно можна розглянути зсуви на площині. Радимо зробити це самостійно.

1.5. ГРУПИ ПІДСТАНОВОК

Перейдемо до вивчення дуже важливої групи – групи підстановок або симетричної групи, яка позначається S_n.

Нехай є деякі n натуральних чисел $\{i_1, i_2, \ldots, i_n\}$. Якщо якось переставити ці числа, одержимо перестановку $\{j_1, j_2, \ldots, j_n\}$. Відомо, що кількість всіх перестановок дорівнює $n! = 1 \cdot 2 \cdot \ldots \cdot n$. Перестановка, в якій $j_1 < j_2 < \ldots < j_n$, називається узгодженою або правильною.

Означення. Взаємно однозначне відображення деякої підмножини множини натуральних чисел на себе називається підстановкою. Підстановка визначається двома перестановками – початковою та її образом і записується у вигляді

$$\sigma = \begin{bmatrix} i_1, & i_2, & \cdots & i_n \\ j_1, & j_2, & \cdots & j_n \end{bmatrix} \tag{1.5.1}$$

З перестановками елементів деякої множини ми вже зустрічались, вивчаючи дію групи на цій множині за допомогою зсувів. Можна вважати, що підстановка вигляду (3.5.1) діє на номери елементів цієї множини. Якщо елементи групи було розміщено спочатку таким чином: $G = \left\{a_{i_1}, a_{i_2}, \ldots, a_{i_n}\right\}$, то потім у результаті перестановки вони будуть розміщені в порядку $G = \left\{a_{j_1}, a_{ij}, \ldots, a_{j_n}\right\}$. Тепер цілком зрозуміло, як такому перетворенню зіставити підстановку (1.5.1).

На практиці будемо зустрічатися, як правило, лише з випадком правильної верхньої перестановки, але в деяких конструкціях загальної теорії буде потрібен загальний випадок (1.5.1).

Зараз будемо вважати, що верхня перестановка - це $\{1,2, \ldots, n\}$.

Кількість чисел в обох перестановках називається степенем підстановки. Можна визначити добуток двох підстановок однакових степенів. Зазначимо, що дві перестановки, виконані по черзі, є знову перестановка. Треба тільки вказати формальне правило, за яким, знаючи кожну з підстановок, визначених цими перестановками, можна одержати результат суперпозиції двох перестановок.

Нехай є дві підстановки

$$A = \begin{bmatrix} 1, & 2, & \cdots & n \\ j_1, & j_2, & \cdots & j_n \end{bmatrix},\ B = \begin{bmatrix} 1, & 2, & \cdots & j_1, & \cdots & n \\ k_1, & k_2 & \cdots & k_{j_1} & \cdots & k_n \end{bmatrix}$$

Визначимо, куди потрапило перше число (1) першого рядка підстановки *A* після двох перестановок. Спочатку воно перейшло в число j_1 (після першої перестановки), потім знаходимо число j_1 в верхньому рядку *B*. Йому відповідає число k_{j_1} в нижньому рядку *B*. Таким чином, числу 1 в результаті двох перестановок відповідає число k_{j_1}, яке треба в підстановці $C \equiv AB$ підписати під 1. Далі те ж саме треба зробити з числами 2,3,..., *n* з верхнього рядка підстановки *A*. Приклад:

A= $\begin{bmatrix} 1,2,3,4 \\ 3,4,1,2 \end{bmatrix}$, B=$\begin{bmatrix} 1,2,3,4 \\ 2,3,4,1 \end{bmatrix}$.

Знайдемо $C = AB$:

1→3(A),3→4(B),1→4(C),

2→4(A),4→1(B),1→1(C),

3→1(A),1→2(B),1→2(C),

4→2(A),2→3(B),1→3(C).

Отже,

C = $\begin{bmatrix} 1,2,3,4 \\ 4,1,2,3 \end{bmatrix}$.

Тепер, не виписуючи діаграм, знайдемо, наприклад, $D=AB$:

$$\begin{bmatrix} 1,2,3,4 \\ 2,3,4,1 \end{bmatrix} \begin{bmatrix} 1,2,3,4 \\ 3,4,1,2 \end{bmatrix} = \begin{bmatrix} 1,2,3,4 \\ 4,1,2,3 \end{bmatrix}$$

Зазначимо, що $AB = BA$.

Легко помітити, що підстановка

$$E=\begin{bmatrix}1,2,...,n\\1,2,...,n\end{bmatrix}$$

виконує роль одиниці або тотожного елемента при такому множенні.

Нехай є підстановка

$$A=\begin{bmatrix}1,\ 2,...,\ n\\k_1,k_2,...,k_n\end{bmatrix}.$$

Знайдемо обернену для A підстановку. Легко бачити, що такою підстановкою $B \equiv A^{-1}$ буде

$$B=A^{-1}=\begin{bmatrix}k_1,k_2,...,k_n\\1,\ 2,...,\ n\end{bmatrix}.$$

Перепишемо тепер B так, щоб верхня перестановка в ній була правильною. Для цього треба переписати її таким чином, щоб у верхньому рядку в ній під k_1 знаходилось число 1, під k_2 – 2 і т.ін.

Таку підстановку B виписати нелегко. Обмежимося тим, що випишемо два числа в нижньому рядку, припустивши, що $k_1 < k_2$:

$$B=\begin{bmatrix}1, & \cdots & k_1, & \cdots & k_2, & \cdots & n\\ \cdots & \cdots & 1, & \cdots & 2, & \cdots & \cdots\end{bmatrix} \tag{1.5.2}$$

Легко помітити, що $AB = BA = E$.

Радимо читачеві перевірити самостійно асоціативність такого множення підстановок:

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

Тепер можна зробити висновок, що підстановки степеня n утворюють групу, яка називається симетричною групою степені n і позначається S_n. Неважко перевірити, що при $n > 2$ вона некомутативна.

Як уже зазначалось, елементам скінченної групи можна співставити деяку підстановку. Нехай $G = \{a_{i_1}, a_{i_2}, ..., a_{i_n}\}$. Якщо вплинути на всі елементи деяким елементом g за допомогою правого зсуву, позначаючи $a_{i_1} g = a_{j_1}$, одержимо: $G = \{a_{j_1}, a_{j_2}, ..., a_{j_n}\}$. В такому разі одержуємо взаємно однозначне співставлення: $g \Leftrightarrow \sigma_g$, де σ_g визначається формулою (1.5.1). Таке співставлення дійсно взаємно однозначне, тому що коли є один елемент g_1, якому теж відповідає підстановка σ_g, то це означає, що $a_{i_1} g = a_{i_1} g_1 = a$, звідки $g_1 = g$. Елементи групи G, діючи на собі за допомогою правих зсувів, діють так само, як і елементи групи S_n, що діють на перестановки, які визначаються номерами елементів групи, причому виконується співвідношення вигляду (1.4.1). Отже, одержали гомоморфізм G в S_n, що взаємнооднозначно відображає G на свій образ (що, як відомо, є підгрупа) в S_n. Сформулюємо результат.

Теорема 1.5.1 (Келі). Кожна скінченна група ізоморфна підгрупі групи підстановок.

Якщо переставити два довільні числа нижнього рядка E, одержимо підстановку, яка називається транспозицією:

$$\boldsymbol{T_{ij}^{(n)}} = \begin{bmatrix} 1, ..., i, ..., j, ..., n \\ 1, ..., j, ..., i, ..., n \end{bmatrix}.$$

(Нижні індекси показують, які числа в нижньому рядку переставлені; верхній – степінь підстановки.)

Таку транспозицію іноді будемо позначати (i,j).

Якщо переставити сусідні числа, одержимо елементарну транспозицію. Наприклад, елементарна транспозиція степені 4:

$$T_{12} = \begin{bmatrix} 1,2,3,4 \\ 2,1,3,4 \end{bmatrix}.$$

Легко перевірити, що довільна транспозиція T задовольняє тотожність $T^2 = E$. Нехай є підстановка

$$A = \begin{bmatrix} 1,2,3,4 \\ 3,2,4,1 \end{bmatrix}.$$

Щоб отримати підстановку

$$B = \begin{bmatrix} 1,2,3,4 \\ 1,2,4,3 \end{bmatrix},$$

що одержана з A перестановкою в нижньому рядку чисел 3 і 4, які стоять на першому та четвертому місцях, виявляється, треба обчислити добуток T_{14} A. Перевіримо це:

$$T_{14}A = \begin{bmatrix} 1,2,3,4 \\ 4,2,3,1 \end{bmatrix} \begin{bmatrix} 1,2,3,4 \\ 3,2,4,1 \end{bmatrix} = \begin{bmatrix} 1,2,3,4 \\ 1,2,4,3 \end{bmatrix} = B.$$

Таке правило неважко довести в загальному випадку. Треба тільки уважно прослідкувати за пересуванням чисел.

Очевидно, довільна підстановка може бути одержана з тотожної в результаті кількох транспозицій. Тому, враховуючи одержане правило, робимо висновок, що довільна підстановка може бути представлена як добуток кількох транспозицій. Зауважимо, що транспозиція може бути представлена як добуток елементарних транспозицій. Цікаво підрахувати їх число. Нехай $i < j$ і треба з E одержати транспозицію

$$T_{ij} = \begin{bmatrix} 1, & \cdots & i, & \cdots & j, & \cdots & n \\ 1 & \cdots & j, & \cdots & i, & \cdots & \cdots \end{bmatrix} \tag{1.5.3}$$

Для того щоб пересунути індекс j в E вправо на місце i, треба зробити $j - i$ перестановок сусідніх чисел: i переставити з числами $i+1, i+2, i+3, ..., j$. Потім пересувається індекс j, що вже знаходиться на $(j - 1)$-му місці, вліво на місце i, тобто переставити його з j -1 - i числами. Всього ми повинні були виконати $2(j-i)-1$ (непарне число) транспозицій сусідніх чисел. З попереднього відомо: це означає, що розглядувана транспозиція дорівнює добутку непарного числа елементарних транспозицій.

З цих спостережень легко зробити висновок, що деякі підстановки одержуються після непарного числа елементарних транспозицій (наприклад, довільні транспозиції), деякі – після їх парного числа. Це дає змогу поставити кожній підстановці $A \in S_n$ число $\varepsilon (A)$ з множини $Z_2 = \{0,1\}$ (адитивної групи з двох елементів, де 1 + 1 =0), яке називається парністю підстановки, причому будемо вимагати, щоб це відображення $\varepsilon : S_n \to Z_2$ було гомоморфізмом груп. Парність тотожної підстановки за означенням дорівнює нулю. Парність кожної елементарної транспозиції будемо вважати рівною одиниці. Тоді, наприклад, парність $\varepsilon (T)$ довільної транспозиції T_{ij}, буде теж рівною 1, бо T_{ij} є добуток непарного числа $2(j-i)-1$ елементарних транспозицій, а завдяки гомоморфізму парність непарного добутку елементарних транспозицій є непарною сумою одиниць, тобто одиницею.

Підстановка, яка є добутком непарного числа довільних транспозицій, також має парність і завдяки гомоморфізму – при знаходженні парності добутку парності співмножників додаються за правилом додавання в Z_2.

Підстановки, що можуть бути зображені як добуток парної кількості транспозицій, завдяки тому ж правилу мають парність 0. Так, $\varepsilon (E) = \varepsilon (T^2_{ij}) = \varepsilon (T_{ij}) \varepsilon (T_{IJ}) =$ 1 + 1=0.

Інколи зручніше парністю підстановки називати гомоморфізм групи підстановок в мультиплікативну групу {-1,1}, тобто парні підстановки, зокрема, тотожня, матимуть парність 1, а непарні -1. Як визначати парність, не має суттєвого значення, тому що група $\{-1,1\}$ ізоморфна $\mathbb{Z}_2$. З таким означенням парності легше працювати при вивченні алгебр Грасмана.

На прикладі парності можна зробити загальне зауваження щодо гомоморфізму: він може загубити деякі властивості початкової групи, на якій він визначений. Так, S_n при $n \geq 3$ – некомутативна група, а її образ при гомоморфізмі парність – комутативна група $\mathbb{Z}_2$.

Цікаво знайти підгрупи в S_n. Так, підгрупу утворюють два елементи $\{E,T_{ij}\}$ при довільних $1 \leq i,\ j \leq n$. Це мультиплікативна група, що, очевидно, ізоморфна адитивній групі $\mathbb{Z}_2 = \{0, 1\}$.

Розглянемо всі парні підстановки в S_n, множину яких позначимо через A_n. Оскільки $E \in A_n$ і для $x,y \in A_n$ $\varepsilon(x) = \varepsilon(y) = 0$ і $\varepsilon(xy) = \varepsilon(x) + \varepsilon(y) = 0 + 0 = 0$, то $xy \in A_n$.

Розглянемо підгрупу циклічних підстановок, які можна одержати з $E \subset S_n$ в такий спосіб: одиницю в нижньому рядку E переставляємо з кожним числом, доки вона не займе місця під n (всього робимо n - 1 елементарних транспозицій), В одержаній підстановці

$$\sigma_1 = \begin{bmatrix} 1,2,...,n-1,n \\ 2,3,...,n,1 \end{bmatrix},$$

те ж саме зробимо з числом 2, одержимо підстановку

$$\sigma_2 = \begin{bmatrix} 1,2,...,n-2,n-1,n \\ 3,4,...,n,1,2 \end{bmatrix}$$

і т.ін., доки не одержимо підстановку

$$\sigma_{n-1} = \begin{bmatrix} 1,2,3,...,n-1,n \\ n,1,2,...,n-2,n-1 \end{bmatrix}.$$

Зауважимо, що $\sigma_2 = \sigma_1^2$. Дійсно,

$$\sigma_1^2 = \begin{bmatrix} 1,2,...,n-1,n \\ 2,3,..., \ n, \ \ 1 \end{bmatrix} \begin{bmatrix} 1,2,...,n-2,n-1,n \\ 2,3,...,n-1, \ n, \ \ 1 \end{bmatrix} = \begin{bmatrix} 1,2,...,n-2,n-1,n \\ 3,4,..., n, \ \ 1, \ \ 2 \end{bmatrix}.$$

За індукцією доводимо, що $\sigma_k = \sigma_1^k$, $k \in \{1,2,...,n\text{-}1\}$, $\sigma_1^n = e \equiv \sigma_0$.

Якщо паперовий круг покласти на папір, його коло розділити на n рівних частин і біля кожної як на крузі, так і на папері написати числа від 1 до n за годинниковою стрілкою, числа на папері будуть відповідати верхньому рядку підстановки, на колі – нижньому. Таким чином, ми одержали тотожну підстановку. Повертаючи $(n-1)$ разів коло проти годинникової стрілки на кут $\frac{2\pi}{n}$, одержимо всі описані підстановки.

Повороти на такі ж кути за годинниковою стрілкою дають підстановки, відповідно обернені до вже одержаних, що легко перевірити самостійно. Цікаво, що таким поворотам можна протиставити мультиплікативну групу коренів n-го степеня із одиниці, які мають вигляд $exp(2\pi ki/n)$, $k \in \{0,1,...,n\text{-}1\}$, де $i = \sqrt{-1}$. У свою чергу, вона, як вже відомо, ізоморфна адитивній групі класів залишків по модулю n, тобто $Z_n = \{0,1,...,n - 1\}$.

Тепер дамо алгоритм для визначення парності підстановок. Для того щоб його краще зрозуміти, розглянемо приклад підстановки $\sigma \in \mathbb{Z}_8$.

$$о= \begin{bmatrix} 1,2,3,4,5,6,7,8 \\ 4,6,3,5,8,2,7,1 \end{bmatrix}.$$

З'ясуємо, які числа дійсно перемішуються, а які ні. Очевидно, на місці залишаються числа 3 і 7. Будемо говорити, що їм відповідають цикли довжиною 1. З'ясуємо, що число 1 переходить в 4, далі 4 – в 5, 5 – в 8, 8 – в 1. Отже, одержимо такий ланцюг: $1\rightarrow4\rightarrow5\rightarrow8\rightarrow1$. Якщо продовжувати цей процес, матимемо нескінченний ланцюг з повторюваним циклом довжиною 4

(1,4,5,8). Можемо починати з числа 4, тоді одержимо цикл (4,5,8,1), з числа 5 – цикл (5,8,1,4), з числа 8 – (8,1,4,5). Це нагадує приклад з підгрупою циклічних підстановок, але тут підстановки утворені не перестановками перших чотирьох упорядкованих чисел, а чотирма упорядкованими числами (1,4,5,8) і їх перестановками:

$$\sigma_0 = \begin{bmatrix} 1,4,5,8 \\ 1,4,5,8 \end{bmatrix},\ \sigma_1 = \begin{bmatrix} 1,4,5,8 \\ 4,5,8,1 \end{bmatrix},\ \sigma_2 = \begin{bmatrix} 1,4,5,8 \\ 5,8,1,4 \end{bmatrix},\ \sigma_3 = \begin{bmatrix} 1,4,5,8 \\ 8,1,4,5 \end{bmatrix}.$$

Як і в тому прикладі, одержали циклічну групу з твірним елементом σ_1 (перевірте!).

Продовжуючи розгляд підстановки σ, знаходимо ще один цикл з елементів, які пересуваються: (2,6). Ясно, що ці цикли не перетинаються і можна подати підстановку σ у вигляді композиції циклів C_1 = (1,4,5,8) і C_2 = (2,6) : $\sigma = C_1C_2$. Пояснимо цей запис. Циклу C = (1,4,5,8) відповідає підстановка

$$C_1 = \begin{bmatrix} 1,2,3,4,5,6,7,8 \\ 4,2,3,5,8,6,7,1 \end{bmatrix},$$

циклу C_2 –

$$C_2 = \begin{bmatrix} 1,2,3,4,5,6,7,8 \\ 1,6,3,4,5,2,7,8 \end{bmatrix}.$$

Запис $\sigma = C_1C_2$ треба з точки зору добутку підстановок уявляти як $\sigma = C_1C_2$, але для його скорочення будемо використовувати перший варіант, розуміючи його точний зміст. Як і в згадуваному прикладі циклічної підгрупи, визначаємо, що будь-який цикл довжиною k може бути одержаний як добуток k-1 транспозицій. В загальному випадку легко перевірити, що $(i_1, i_2, ..., i_k) = (i_1, i_2)(i_1, i_3)...(i_1, i_k)$.

Нехай в загальному випадку $\sigma = C_1 C_2$ C_1C_2 ... C_l, де C_i – цикл довжиною k_i. Тоді парність σ можна визначити як парність числа $k_1 - 1 + k_2 - 1 + ... + k_l - 1 = k_1 + k_2 + ... + k_l - l$. Якщо додати і відняти число m тих елементів, що не пересуваються, то одержана сума не зміниться, і при цьому k_1 + k_2 +...+ k_1 + m = n. Отже, досить визначити парність числа n – (1 + m), яке називається декрементом підстановки і визначається як сума кількості циклів і числа елементів, що не пересуваються. Остаточно

$$\varepsilon(\sigma) = \left[\frac{1 + (-1)^{n-(m+l)}}{2} \right] \tag{1.5.4}$$

де $[x]$ означає цілу частину числа x, тобто найбільше ціле число, яке менше x.

Якщо вважати, що парність непарної підстановки дорівнює – 1, а парність парної підстановки дорівнює 1, то одержану формулу можна переписати у вигляді

$$\varepsilon(\sigma) = (-1)^{n-(m+l)} \tag{1.5.5}$$

1.6. КІЛЬЦЯ І ПОЛЯ

Означення. Множина R, для двох довільних елементів якої визначені дві операції – додавання і множення, внаслідок яких знову одержуються елементи із R, називається кільцем, якщо ці операції задовольняють аксіомам 1–7:

1) $a + (b + c) = (a + b) + c$ (асоціативність додавання);

2) $a + b = b + a$ (комутативність додавання);

3) існує елемент $0 \in R$, що $a + 0 = 0 + a = a$ (існування нуля);

4) для $a \in R$ існує елемент $b \in R$, що $a + b = b + a = 0$; позначення: $b = - a$ (існування оберненого елемента відносно додавання);

5) $a(bc) = (ab)c$ (асоціативність добутку);

6) існує $e \in R$ (або інколи позначають $1 \in R$), що $ea = ae = a$ (існування одиниці);

7) $a(b + c) = ab + ac$ (дистрибутивність множення відносно додавання). Якщо виконується аксіома:

8) $ab = ba$ (комутативність добутку), кільце називається комутативним або абелевим.

Зауважимо, що, взагалі кажучи, елементи кільця не повинні мати обернені щодо операції множення. Якщо ж для деякого елемента a існує елемент r такий, що $ra=e$, то він називається правим оберненим до a. Аналогічно визначається лівий обернений: $al=e$. Доведемо, що такі лівий і правий обернений елементи співпадають. Це має місце завдяки асоціативності множення. Маємо: $e= ar$. Помножимо зліва обидві частини цієї рівності на l: $l= (la)r=r$. Тепер обернений до a елемент можна позначити через a^{-1}.

Якщо для деяких ненульових елементів a і b $ab=0$, то вони називаються дільниками нуля.

Якщо елемент a є дільником одиниці, то він не є дільником нуля. Дійсно, нехай $ab=0$. Помножимо обидві частин рівності на a^{-1} : $a^{-1}ab=b=0$.

Існують кільця без дільників нуля і одиниці, наприклад, кільце цілих чисел.

Якщо виконуються сім перших аксіом і аксіома:

9) для $a \in Z, a \neq 0$ існує елемент b такий, що $ab = ba = e, b \equiv a^{-1}$, то таке кільце називається тілом.

Якщо виконуються аксіоми 1–9, то таке комутативне тіло називається полем.

Відомо, що скінченне тіло є поле (теорема Зедерберна).

Підмножина $R_1 \subseteq R$ називається підкільцем кільця R, коли вона є кільцем і замкнута відносно всіх операцій кільця.

Нехай R_1 і R_2 – кільця. Відображення $\varphi: R_1 \to R_2$ називається гомоморфізмом, якщо воно переводить операції кільця R_1 в операції кільця R_2, тобто для $x,y \in R$ $\varphi(x + y) = \varphi(x) + \varphi(y)$, $\varphi(xy) = \varphi(x)\ \varphi(y)$. Радимо читачеві перевірити (як це робилося в випадку груп), що при гомоморфізмі нуль R_1 переходить в нуль R_2, а одиниця – в одиницю.

Гомоморфізм кільця в себе називається ендоморфізмом.

Взаємнооднозначний гомоморфізм називається ізоморфізмом; ізоморфізм кільця на себе називається автоморфізмом.

Приклади кілець:

1) $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, причому три останніх – поля;

2) *Z[x]* – кільце всіх многочленів від змінної *x;*

3) $\mathbf{Z_2} = \{0,1\}$.

Означення. Характеристикою поля *К* (char (*K*) називається таке найменше число *n,* що сума *n* одиниць дорівнює нулю.

Характеристика поля – просте число. Дійсно, якщо char(K) = kl, то $(kl) \times 1 = 0$, звідки або $k = 0$, або $l = 0$, що протирічить мінімальності char (K).

Очевидно, char $(Z_2) = 2$.

За означенням нескінченні поля мають характеристику 0.

Можна сказати, що поле складається з двох груп – адитивної та мультиплікативної.

Означення. Поле називається алгебраїчно замкнутим, якщо кожен многочлен з коефіцієнтами з нього має принаймні один корінь, який належить до цього поля.

З означення випливає, що многочлен степеня *n* має *n* коренів, які належать полю його коефіцієнтів, якщо воно алгебраїчно замкнуте.

З основної теореми алгебри випливає, що $\mathbb{C}$ – алгебраїчно замкнуте поле. Читач повинен перевірити, що поля $\mathbb{R}, \mathbb{Q}$ алгебраїчно незамкнуті.

Якщо розглянути корені всіх многочленів з раціональними коефіцієнтами, то одержимо так зване поле алгебраїчних чисел, яке є алгебраїчно замкнутим. Ми не маємо змоги в рамках даного посібника доводити ці твердження.

Продовжимо вивчення групи $\boldsymbol{Z_n}$. Для її елементів можна визначити операцію множення $\oplus$ – для цього треба, як і при визначенні додавання, взяти залишок від ділення на n звичайного їх добутку. Наприклад, в Z_6 $2 \otimes 2 = 4$, $2 \otimes 3 = 0$, $2 \otimes 4 = 2$, $2 \otimes 5 = 4$, $3 \otimes 3 = 3$, $3 \otimes 4 = 0$, $3 \otimes 5 = 3$, $4 \otimes 4 = 4$,$4 \otimes 5 = 2$,$5 \otimes 5=1$.

Легко бачити, що операції додавання і множення в Z_n задовольняють аксіоми кільця, і таке кільце теж будемо позначати Z_n.

Очевидно, якщо $n = pq$, то в Z_n $p \otimes q = 0$. Такі елементи називаються дільниками нуля; елементи, що мають обернений, називаються дільниками одиниці. Легко бачити, що числу від 1 до n - 1, яке взаємно просте з n, відповідає елемент з Z_n, який не є дільником нуля. Доведемо, що такий елемент є дільником одиниці. Нехай a взаємно просте з n, $1 < a < n$. Розглянемо елементи, що є добутками на a всіх елементів з Z_n $\{a, 2a, \ldots ,(n - 1)a\}$. Вони всі різні, бо якщо $ba = ca$, то $(b - c)a = 0$, а в такому разі при $b \neq c$ a був би дільником нуля. Отже, виписані елементи – це всі елементи Z_n без 0, тому серед них є елемент 1, тобто існує такий елемент a_1, що $aa_1 = 1$, що й треба було довести.

Зауважимо, що дільник нуля в Z_n не може бути дільником одиниці і навпаки. Якщо a – дільник нуля, то існує $b \neq 0$, що $ab=0$, і коли припустити, що a – також дільник одиниці, то, помноживши цю рівність на a^{-1}, одержимо $b = 0$. Обернене твердження доводиться аналогічно.

Розглянемо всі елементи з Z_n, які відповідають числам від 1 до n - 1, взаємно простим з n. Очевидно, вони утворюють підполе кільця Z_n. Порядок його мультиплікативної групи дорівнює кількості натуральних чисел, менших за n і взаємно простих з n: $\varphi(n)$ (функція Ейлера). З теореми Лагранжа випливає важливий результат для Z_n.

Теорема 1.6.1 (Ейлера):

$$a^{\varphi(n)} \equiv 1 \bmod(n). \qquad (1.6.1)$$

В частинному випадку простого n, що будемо позначати через p, $\varphi(p) = p - 1$ і тоді одержимо такий результат.

Теорема 1.6.2 (мала теорема Ферма). Якщо p – просте число, то

$$a^{p-1} \equiv 1 \bmod(p). \qquad (1.6.2)$$

Отже, при простому p Z_P є поле.

Вправа. Нехай p – просте число. Знайти в Z_p елементи, квадрати яких дорівнюють одиниці поля.

Отже, нехай $k^2 = 1$, $1 \le k \le p$-1. Очевидно, це рівняння задовольняють числа 1, а також -1, але $-1 \equiv p\text{-}1 \ (mod\ p)$. Покажемо, що інших розв'язків в Z_p немає. Рівність $k^2 = 1$ в Z_p означає, що існує таке $m \in Z$, що $k^2 = 1 + mp$. Звідси випливає, що $(k - 1)(k + 1) = mp$. Це справджується лише в випадках, коли або k - 1 ділиться на p, або k + 1 ділиться на p. Це означає, що або $k \equiv 1 \ (\mathrm{mod}\ p)$, або $k \equiv -1 \ (\mathrm{mod}\ p) =$ p-l(mod p).

Вправа. Нехай p – просте число. Довести, що $(p - 1)! + 1$ ділиться на p.

Доведення. Розглянемо добуток $(p\text{-}1)! = 1 \otimes 2 \otimes ... \otimes (p - 2) \otimes (p - 1)$ (як елемент Z_p). Тут виписані всі елементи Z_p. Кожен з них має обернений, тому в цей добуток кожен елемент, крім p-1, входить разом зі своїм оберненим (з попередньої вправи випливає, що обернений до p - 1 – сам елемент p - 1), і такі пари дають при множенні одиницю. Отже, $p!=p\text{-}1$ (в Z_p). Тому $p! + 1 = p\text{-}1 + 1 = p \equiv 0 \ (\mathrm{mod}\ p)$, що й треба було довести. Цей результат в теорії чисел називається теоремою Вільсона.

Означення. Підмножина I кільця R називається ідеалом, якщо для довільного $a \in R$ і довільного $i \in I$ $ai \in I$. Це також можна записати у вигляді $RI \subset I$. З визначення випливає, що ідеал кільця є його підкільцем.

Зауважимо, що якщо $a \in I$ і a має обернений a^{-1}, то $I = R$. Дійсно, кожен елемент вигляду $(xa^{-1})a = x$ при довільному $x \in R$ належить до Z, а також і до I за означенням ідеала. Це й означає, що $I = R$. Отже, можна зробити важливий висновок: для того щоб ідеал не збігався з кільцем, необхідно і досить, щоб він не мав елементів, які мають обернені, зокрема одиниці кільця.

Приклади ідеалів:

1) $n\,Z \subset Z$. Доведення очевидне;

2) множина $P(x)Z[x]$, де $P(x)$ – деякий многочлен в кільці многочленів $Z[x]$ від однієї змінної;

3) нехай n – непросте натуральне число, число a не взаємно просте з n і $1 < a < n$. Тоді множина $I= aZ_n$ – ідеал в Z_n. Для доведення треба лише помітити, що ab – дільник нуля при довільному $b \in Z_n$.

Означення. Ідеал називається головним, якщо він утворений одним елементом. Так, ідеали в прикладах 1–3 головні.

Означення. Кільце, в якому всі ідеали головні, називається кільцем головних ідеалів.

Вправа 1. Z – кільце головних ідеалів.

Для доведення треба розглянути довільний ідеал $I\subset Z$ і довести, що він породжений одним елементом. Розглянемо найменше натуральне $a \subset I$. Якщо воно не породжує I, то знайдеться таке $b \in I$, $b > a$, що $b \neq an$. Розділимо b з залишком на a: $b = ak + r$, де остача r, як відомо, менша за a. Оскільки $a, b \in I$, то $r \in I$, а це протирічить мінімальності a.

Вправа 2. $Z[x]$ – кільце головних ідеалів. Доведення спирається на можливість ділення з залишком у кільці многочленів. Радимо читачеві зробити його самостійно.

Розглянемо більш детально кільце многочленів від однієї змінної $Z[x]$. Нагадаємо, що многочлен $P(x)$ ділиться на многочлен $Q(x)$ з залишком, якщо

$$P(x) = Q(x)\, S(x) + R(x). \qquad (1.6.3)$$

При цьому $S(x)$ називається неповною часткою, а $R(x)$ – залишком від ділення. Відомо, що вони існують, і степінь $R(x)$ менший за степінь $Q(x)$. Коли $R(x) =0$, то кажуть, що $Q(x)$ є дільником $P(x)$.

Нехай $P(x)$ і$Q(x)$ мають спільний дільник. Найбільшим спільним дільником називається їх спільний дільник найбільшого степеня. Зрозуміло, що він визначається неоднозначно, а з точністю до ненульового множника і позначається *(P(x), Q(x))*.

Покажемо, як знайти найбільший спільний дільник. Методика його знаходження спирається на так званий алгоритм Евкліда, який полягає в виконанні таких дій.

Нехай степінь *P(x)* більший або дорівнює степеню *Q(x)*. Поділимо *P(x)* На *Q(x)* з залишком, що позначимо $R_1(x)$ (в окремому випадку він може дорівнювати нулю). Степінь *Q(x)* менший за степінь $R_1(x)$. Далі поділимо *Q(x)* на $R_1(x)$ і дістанемо деякий залишок $R_2(x)$, потім поділимо $R_1\ (x)$ на $R_2(x)$ і дістанемо залишок $R_3(x)$ і т. ін. Степені залишків весь час зменшуються, тому через скінченне число кроків процес зупиниться і на останньому кроці залишок дорівнюватиме нулю. Передостанній залишок в такому ланцюгу і буде найбільшим спільним залишком *P(x)* і *Q(x)*. Щоб це довести, запишемо весь ланцюг ділень:

$$\begin{gathered} P(x)=Q(x)S(x)+R_1(x);\\ Q(x)=R_1(x)S_1(x)+R_2(x);\\ R_1(x)=R_2(x)S_2(x)+R_3(x);\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ R_{k-2}(x)=R_{k-1}(x)S_{k-1}(x)+R_k(x);\\ R_{k-1}(x)=R_k(x)S_k(x). \end{gathered} \qquad (1.6.4)$$

Остання рівність показує, що $R_k(x)$ – дільник для $R_{k-1}(x)$, з попередньої рівності випливає, що $R_k(x)$ – дільник для $R_{k-2}(x)$. Далі, піднімаючись по ланцюгу рівностей вверх, впевнюємось таким же чином, що $R_k(x)$ – дільник для $R_{k-3}(x), R_2(x), R_1(x)$, а потім і для $P(x)$ і $Q(x)$. Для доведення, що це найбільший дільник, треба взяти деякий дільник *P(x)* і *Q(x)*, який позначимо *D(x)* і зробимо аналогічну процедуру зверху донизу. В такому разі одержимо, що *D(x)* ділить $R_k(x)$, що і доводить твердження.

Теорема 1.6.1. Якщо *D(x)* – найбільший спільний дільник *P(x)* і *Q(x)*, то існують многочлени *U(x)* і *V(x),* що справедлива рівність

$$P(x)\,U(x) + Q(x)\,V(x) = D(x). \qquad (1.6.5)$$

Доведення випливає з рівностей (1.6.4). Якщо покласти $R_k(x)=D(x), U_1(x)=1, V_1(x)=-S_k(x)$, то з передостанньої рівності одержимо $D(x)=R_{k-2}(x)U_1(x)+R_{k-1}(x)V_1(x)$.

Підставляючи сюди вираз $R_{k-1}(x)$ через $R_{k-2}(x)$ і $R_{k-3}(x)$ з попередньої рівності, запишемо:

$$D(x)=R_{k-3}(x)U_2(x)+R_{k-2}(x)V_2(x),$$

де

$$U_2(x)=V_1(x), V_2(x)=U_1(x)-V_1(x)S_{k-1}(x)$$

Продовжуючи підніматися вгору по ланцюгу рівностей, прийдемо врешті до рівності (1.6.5), що й треба було довести.

Наслідок. Якщо многочлени $P_1(x), \ldots, P_n(x)$ мають найбільший спільний дільник *D(x),* то існують многочлени $U_1(x), \ldots, U_n(x)$, через які *D(x)* можна знайти за формулою

$$D(x)=\sum_{i=1}^{n}U_i(x)P_i(x). \qquad (1.6.6)$$

У випадку, коли всі $P_1(x),\ldots,P_n(x)$ взаємно прості,

$$\sum_{i=1}^{n}U_i(x)P_i(x)=1. \qquad (1.6.7)$$

Твердження доводиться за допомогою математичної індукції.

Зауважимо, що коли многочлени $P_1(x), \ldots, P_n(x)$ в сукупності взаємно прості, це не обов'язково означає, що кожні два многочлени з них взаємно прості.

ГЛАВА 2. ЛІНІЙНІ ПРОСТОРИ І ОПЕРАТОРИ

2.1. ВЕКТОРНІ ЛІНІЙНІ ПРОСТОРИ

Означення 2.1.1. Нехай *E* — адитивна абелева група, для елементів якої визначена операція множення на елементи деякого поля *K,* причому виконуються аксіоми для $a,b \in K,\ x,y \in E$:

1) $a(bx) = (ab)x,\ 1x = x;$
2) $(a + b)x = ax + bx;$
3) $a(x + y) = ax + ay.$

Тоді *E* називається векторним лінійним простором або просто – лінійним простором.

Наведемо деякі наслідки з означення лінійного простору.

А) $0x = a0 = 0$ для всіх $x \in E$, $a \in K$. Дійсно, $0x + 0x = (0 + 0)x = 0x$, звідки, додаючи обернений до $0x$ елемент (що існує в абелевій групі) до обох частин рівності, одержимо $0x=0$. Аналогічно $a0 + a0 = a(0 + 0) = a0$ і $a0 = 0$.

Б) $(-1)\,x = -x$;. Дійсно, $x + (-1)\,x = 1\,x + (-1)\,x = (1+(-1))\,x = 0\,x$, тобто вектор $(-x)$ протилежний до x.

В) Якщо $ax = 0$, то або $a = 0$, або $x = 0$. Дійсно, якщо а $\neq$ 0, то $0 = a^{-1}(ax) = (a^{-1}a)x = 1x = x$, і $x=0$.

З першої аксіоми випливає, що $a\,(-x) = -ax$.

Множина $F \subset E$ називається підпростором простору *E,* якщо разом з довільними векторами $x, y \in F$ в неї входить кожна їхня лінійна комбінація $ax + by\ (a,b \in K)$.

Лінійні простори *E* і *F* називаються гомоморфними, якщо існує відображення $\varphi\ :\ E \to F$, яке задовольняє умову $\varphi(\alpha x + \beta y) = \alpha\varphi(x) + \beta\varphi(y)$ $\forall x, y \in E,\ \alpha, \beta \in K$.

Взаємно однозначний гомоморфізм при умові $\varphi(E) = F$ називається ізоморфізмом.

Система векторів $e = \{e_1, ..., e_n\}$ називається системою твірних лінійного простору, якщо кожен його елемент *x* можна подати у вигляді лінійної комбінації елементів e, тобто

$$x = \sum_{i=1}^{n} x_i e_i \tag{2.1.1}$$

Означення 2.1.2. Система векторів $e = \{e_1, ..., e_n\}$ називається лінійно незалежною, якщо рівність

$$\sum_{i=1}^{n} x_i e_i = 0 \tag{2.1.2}$$

можлива тоді і лише тоді, коли всі елементи x_i дорівнюють 0.
В противному випадку система називається лінійно залежною. Тобто, якщо існує хоча б одне $x_j \neq 0$ і при цьому $\sum_{i=1}^{n} x_i e_i = 0$, то система буде лінійно залежною. З цього, зокрема, випливає, що система, яка містить 0-вектор, обов'язково лінійно залежна. Дійсно, цей вектор досить помножити на довільне не рівне нулю число, а всі інші – на нулі, і тоді умова (2.1.2) буде справджуватись.

Очевидно, якщо взяти всі елементи *E,* то вони будуть утворювати систему твірних, але серед них, безумовно, є лінійно залежні.
Означення 2.1.3. Лінійно незалежна система твірних простору називається його базисом.
Доведемо, що кожен елемент простору має єдине зображення у вигляді (2.1.1), якщо елементи $e_1, e_2, ... e_n$ утворюють базис. Дійсно, коли є ще одне зображення $x = \sum_{i=1}^{n} y_i e_i$, тоді, віднімаючи ці рівності, одержимо, що

$$\sum_{i=1}^{n} (x_i - y_i) e_i = 0,$$

а це суперечить лінійній незалежності векторів базису.
Виявляється, що всі базиси лінійного простору мають однакові потужності (однакову кількість елементів).
Твердження 2.1.1. Нехай системи векторів $f = \{ f_1, ..., f_m \}$ і $g = \{ g_1, ..., g_n \}$ утворюють базис лінійного простору. Тоді $m=n$.
Доведення. Для визначеності будемо вважати, що $m<n$. Візьмемо вектор g_1 і перенесемо його в перший базис. Нова система буде лінійно залежною, і цей вектор зобразимо у вигляді лінійної комбінації векторів з *f*. Нехай, скажімо, перед першим вектором f_1 цієї лінійної комбінації коефіцієнт не дорівнює нулю, отже, f_1 можна виразити через g_1 і решту елементів $f_2, ..., f_m$. Розглянемо нову сукупність $\{ g_1, f_2, ..., f_m \}$. Очевидно, вона буде лінійно незалежною системою твірних, тобто, базисом простору. Дійсно, вектор g_1 лінійно виражається через вектори $f_1, ..., f_m$, а вони утворюють базис. Будемо діяти так само, аж поки в новій системі не опиняться вектори $\{ g_1, ..., g_m \}$, які будуть становити базис простору, а значить, вектори $\{ g_{n-m+1}, ..., g_n \}$, що залишились в системі *g*, будуть лінійно виражатись через вектори $\{ g_1, ..., g_m \}$, що неможливо, оскільки система *g* утворює базис.

Означення 2.1.4. Лінійний простір називається скінченновимірним, якщо він має скінченний базис. Кількість елементів базису називається вимірністю простору і позначається dim (*E*).
Зауважимо: можна довести, що базис завжди існує в просторі, котрий має скінченну систему твірних.

Означення 2.1.5. Нехай $g_1, g_2, ...g_m$ - деяка система векторів, серед яких існує максимальна лінійно незалежна підсистема r $(r \le m)$ векторів. Число r називається її рангом, а множина цих лінійно незалежних векторів – базою системи.

Нехай E_1 — підпростір в *E* і $e_1, e_2, ...e_m$ — базис в E_1. Тоді його можна доповнити до базису всього *E*. Дійсно, візьмемо будь-який базис в *E* $\left(f_1, f_2, ..., f_n\right)$ і розглянемо сукупність векторів $e_1, ..., e_m, f_1, ..., f_n$, що, безумовно, буде системою твірних *E*. Будемо викидати з цієї сукупності по черзі вектори, які є лінійними комбінаціями попередніх. При цьому базисні вектори з E_1 залишаться, бо вони лінійно незалежні між собою. Таким чином через скінченну кількість кроків буде побудовано базис простору *E*.
Нехай E_1 і E_2 — лінійні підпростори простору *E*. Зрозуміло, що множина $E_1 \cap E_2$ є підпростір простору *E*, а множина всіх лінійних комбінацій вигляду $\alpha_1 x_1 + \alpha_2 x_2, x_1 \in E_1, x_2 \in E_2, \alpha_1, \alpha_2 \in K$ також утворює підпростір простору *E*. Її будемо позначати $E_1 + E_2$. Позначимо $n_1 = \dim\left(E_1\right), n_2 = \dim\left(E_2\right)$, $n_0 = \dim\left(E_1 \cap E_2\right)$.

Теорема 2.1.1. Справедлива формула:

$$\dim\left(E_1 + E_2\right) = \dim\left(E_1\right) + \dim\left(E_2\right) - \dim\left(E_1 \cap E_2\right)$$

Доведення. Виберемо базис $\left\{h_1, ...h_m\right\} \subset E_1 \cap E_2$ і доповнимо його до базису E_1: $\left\{h_1, ...h_m, f_1, ..., f_k\right\}$ та E_1: $\left\{h_1, ...h_m, g_1, ..., g_k\right\}$. Очевидно, сукупність $\left\{h_1, ...h_m, f_1, ..., f_k, g_1, ..., g_l\right\}$ буде системою твірних в $E_1 + E_2$. Доведемо її лінійну незалежність. Нехай має місце рівність

$\gamma_1 h_1 + ... + \lambda_m h_m + \alpha_1 f_1, ..., \alpha_k f_k + \beta_1 g_1 + ... + \beta_l g_l = 0$. Звідси

$\gamma_1 h_1 + ... + \gamma_m h_m + \alpha_1 f_1, ..., \alpha_k f_k = -\beta_1 g_1 - ... - \beta_l g_l \equiv h$.

Права частина рівності, вектор, який позначено через h, належить E_2, а ліва - до E_1, значить, вирази в обох частинах належать до $E_1 \cap E_2$, отже, його можна розкласти за базисом $\{h_1,...h_{n_0}\}$ і тоді $-\beta_1 g_1 - ... - \beta_l g_l = \lambda_1 h_1 + ... + \lambda_m h_m$. Звідси маємо: $\lambda_1 h_1 + ... + \lambda_m h_m + \beta_1 g_1 + ... + \beta_l g_l = 0$, що протирічить лінійній незалежності векторів $\{h_1,...h_m, g_1,..., g_k\}$. Отже, $\beta_1 = ... = \beta_l = 0$, а тоді h=0, з чого випливає лінійна залежність векторів системи $\{h_1,...h_m, f_1,..., f_k\}$, а це неможливо, оскільки вона становить базис у E_1. Звідси випливає, що $\alpha_1 = ... = \alpha_k = \gamma_1 = ... = \gamma_m = 0$. Це й завершує доведення лінійної незалежності твірної системи $\{h_1,...h_m, f_1,..., f_k, g_1,..., g_l\}$ в $E_1 + E_2$.

Нехай E_1 і E_2 — лінійні підпростори простору E і $E = E_1 + E_2$. Кожен вектор $x \in E$ можна зобразити у вигляді суми векторів $x_1 \in E_1, x_2 \in E_2 : x = x_1 + x_2$. Взагалі кажучи, таке зображення може бути неєдиним. Це легко проілюструвати на прикладі трьохвимірного простору, в якому взято дві площини, що перетинаються під ненульовим кутом. Візьмемо на лінії їхнього перетину точку *O,* два вектори в одній площині, що виходять з *O* і перетинаються під ненульовим кутом, та аналогічно два вектори в другій площині, причому ні один з цих векторів не лежить на лінії перетину площин. Очевидно, побудовані вектори утворюють систему твірних простору, але будуть лінійно залежними і тому кожен вектор простору можна записати як їхню лінійну комбінацію неєдиним способом. Це означає, що кожен вектор простору можна зобразити як суму векторів цих площин неєдиним способом. Якщо ж взяти площину і пряму, що перетинає цю площину під ненульовим кутом, то довільний вектор простору можна зобразити як суму векторів площини і прямої єдиним способом. Для цього можна вибрати деякий базис в площині. Тоді разом з довільним вектором прямої він буде утворювати базис простору.

Означення 2.1.6. Нехай E_1, E_2 - підпростори в *E,* причому $E_1 + E_2 = E$ і для довільного вектору $a \in E$ зображення $a = x + y$, де $x \in E_1, y \in E_2$ єдине. Тоді будемо говорити, що простір *E* є пряма сума підпросторів E_1 *і* E_2 і позначати це так: $E = E_1 \oplus E_2$.

Аналогічно можна утворити пряму суму довільної скінченної множини лінійних просторів.

Виявляється, що досить перевіряти однозначність зображення тільки нульового вектору з *E* як суму нульових векторів з E_1 і E_2. Цього буде необхідно і досить для того, щоб сума підпросторів, які утворюють весь простір, була прямою. Дійсно, якщо сума пряма, то кожен вектор, в тому числі і нульовий, має єдине зображення. Навпаки, якщо нульовий вектор має єдине таке зображення, а деякий вектор $a \in E$ має два зображення: $a = x_1 + x_2 = y_1 + y_2$, тоді

$\left(x_1 - y_1\right) + \left(x_2 - y_2\right) = 0$, причому $x_1 - y_1 \in E_1,\ x_2 - y_2 \in E_2$ і з однозначності зображення нуля випливає, що $x_1 = y_1,\ x_2 = y_2$. Таким же чином аналогічне твердження доводиться для довільної скінченної прямої суми підпросторів. Радимо читачеві зробити це самостійно.

Нехай $e = \{e_1, ..., e_n\}$ — базис в E_1, $f = \{f_1, ..., f_n\}$— базис в E_2, тоді, очевидно, $e \cup f$ — базис в $E_1 \oplus E_2$.

Теорема 2.1.2. Нехай $E_1, ..., E_m \subset E$ — лінійні підпростори простору E. Тоді $E = E_1 \oplus ... \oplus E_m$, якщо і лише якщо

$$E = \sum_i E_i \text{ і } E_j \cap \left(\sum_{i \neq j} E_i \right) = \{0\}.$$

Доведення. Якщо сума підпросторів не пряма, то
нуль зображується у вигляді $\sum_i l_i$ неоднозначно, наприклад, при цьому $l_j \neq 0$.

Тоді

$$l_j = -\sum_{i \neq j} l_i \in E_j \cap \left(\sum_{i \neq j} E_i \right) \neq \{0\},$$

і умова теореми не виконується. Навпаки, якщо умова теореми не виконується, тоді зображення нуля неєдине.

Зауважимо, що у випадку двох підпросторів $E = E_1 \oplus E_2$ тоді і тільки тоді, коли $E_1 \cap E_2 = \{0\}$.

Приклад. Нехай $K = \mathbb{R}$ або $K = \mathbb{C}$. Розглянемо впорядковані набори з *n* чисел: $x = \left(x_1, ..., x_n\right),\ y = \left(y_1, ..., y_n\right)$. Будемо називати їх векторами-рядками. Числа $x_1, x_2, ..., x_n$ та $y_1, y_2, ..., y_n$ будемо називати компонентами або координатами цих векторів. Введемо операції додавання таких векторів та множення їх на число: $x + y = \left(x_1 + y_1, ..., x_n + y_n\right)$ $\alpha x = \left(\alpha x_1, ..., \alpha x_n\right),\ \alpha \in K$.

Очевидно, ці операції задовольняють аксіоми лінійного простору, тому така множина векторів утворює лінійний простір, що позначається K^n, зокрема, $\mathbb{R}^n$ при $K = \mathbb{R}$ і $\mathbb{C}^n$ — при $K = \mathbb{C}$.

Легко бачити, що вектори e_1 = (1,0,...,0), e_2 = (0,1,0,... ...,0),...,e_n = (0,...,0,1) утворюють базис простору.

Отже, можна вважати, що $K^n = K \times K \times ... \times K = K \oplus K \oplus ... ... \oplus K$, оскільки в скінченновимірному випадку, як уже відзначалося, прямі добутки та прямі суми абелевих груп збігаються.

Нехай E_n - *n*-вимірний лінійний простір. Зробимо важливий висновок: кожному вектору $\forall x \in E_n$ можна зіставити вектор-стовпчик з $\mathbb{R}^n$, що утворений з

коефіцієнтів розкладу його за деяким базисом (формула (2.1.1)) і навпаки: по кожному вектору-стовпчику можна побудувати вектор з E_n, користуючись формулою (2.1.1). Таким чином, кожен *n*-вимірний простір ізоморфний простору $\mathbb{R}^n$, і тому всі *n*-вимірні простори ізоморфні між собою.
Отже, нехай у n-вимірному лінійному просторі E_n вибрано базис $e=\{e_1,...,e_n\}$ і вектору $x\in E_n$ зіставляється вектор з простору $\mathbb{R}^n$, який будемо записувати в стовпчик і позначати

$$[x]_e=\begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{pmatrix} \qquad (2.1.3)$$

Для економії місця його інколи записують в рядок і позначають $[x]_e=(x_1,...,x_n)^T$, де T – операція транспонування, яка полягає в заміні рядка на відповідний стовпчик (або навпаки).

Простір називається нормованим, якщо на ньому задана числова функція, значення якої для $x\in E_n$ позначається $\|x\|$, що задовольняє умови:

1) $\|x\| \ge 0$, $\|x\|=0$ тоді і тільки тоді, коли x = 0;
2) $\|\alpha x\| = |\alpha|\ \|x\|$, $\alpha\in\mathbb{R}$;
3) $\|x+y\|\le\|x\|+\|y\|$.

Властивість 3) називається нерівністю трикутника.
Зрозуміло, що модуль вектору в E_2 і E_3 дає приклад норми.

УНІТАРНІ ТА ЕВКЛІДОВІ ПРОСТОРИ

Означення 2.1.7. Нехай **E** — дійсний лінійний простір ($K=\mathbb{R}$). Скалярним добутком векторів *x,y* $\in$ E називається числова функція, значення якої на цих векторах позначається *(x,y),* що задовольняє умови:
1) $(x,y)\ge 0$, $(x,x) = 0$ тоді і тільки тоді, коли x = 0;
2} $(x,y)=(y,x)$;
3) $(\alpha x+\beta y,z)=\alpha(x,z)+\beta(y,z)$.
З умов 2 і 3 випливає, що скалярний добуток лінійний як по першій, так і по другій змінній. В такому випадку кажуть, що скалярний добуток — це білінійна форма.
Якщо поле $K=\mathbb{C}$, то аксіома 2 має вигляд $(x,y)=\overline{(x,y)}$, звідки випливає, що в такому випадку скалярний добуток лінійний по першій і антилінійний по другій змінній:

$$(z, \alpha x + \beta y) = \overline{\alpha}(z, x) + \overline{\beta}(z, y).$$

Отже, скалярний добуток в комплексному просторі — це так звана півторалінійна форма.
Простір над $\mathbb{R}$ зі скалярним добутком називається евклідовим, над $\mathbb{C}$ — унітарним. В цих просторах можна задати норму за формулою $\|x\| = \sqrt{(x, x)}$.
Перші дві аксіоми норми, очевидно, виконуються. Пізніше буде доведено нерівність трикутника.
Два вектори *x,y* називаються ортогональними, якщо $(x,y) = 0$.
Нехай $a \in E$ — ортогональний кожному вектору $x \in E$. Тоді $a = 0$. Дійсно, з умови випливає, що а — також ортогональний вектору $x = a$, тобто $(a,a) = 0$.
Але з аксіоми 1 одержуємо, що а = 0.
Нерівність Коші-Буняковського: $\forall x, y \in E \; |(x, y)| \le \|x\| \|y\|$

$$|(x, y)| \le \|x\| \|y\|;$$

Рівність має місце тоді і тільки тоді, коли вектори лінійно залежні .
Доведення. Зауважимо, що при $(x, y) = 0$ нерівність очевидна. Доведемо її при умові $(x, y) \neq 0$. Нехай $\alpha, \beta \in \mathbb{C}$. За властивістю скалярного добутку $(\alpha x + \beta y, \alpha x + \beta y) \ge 0$. З іншого боку $(\alpha x + \beta y, \alpha x + \beta y)$ =
$= |\alpha|^2 (x + \beta y, x + \beta y) + |\beta|^2 (y, y) + \alpha\overline{\beta}(x, y) + \overline{\alpha}\beta\overline{(x, y)}$.
Будемо вважати α дійсним, а $\beta = (x, y)$. Після підстановки одержимо:
$\|x\|^2 \alpha^2 + 2|(x, y)|^2 \alpha + |(x, y)|^2 \|y\|^2 \ge 0$ - квадратний тричлен відносно α не менше нуля, отже, $\frac{D}{4} = |(x, y)|^4 - |(x, y)|^2 \|x\|^2 \|y\|^2 \le 0$, звідки і випливає очікувана нерівність. Випадок рівності означає, що дискримінант квадратного тричлена дорівнює нулю, і тоді він є повним квадратом, отже, існує таке α , що $(\alpha x + \beta y, \alpha x + \beta y) = 0$, а це й означає лінійну залежність цих векторів.

Доведемо, що вираз $\sqrt{(x, x)}$ визначає норму. Доведемо нерівність трикутника.
$\|x + y\|^2 = (x + y, x + y) = (x, x) + 2\operatorname{Re}(x, y) + (y, y) \le \|x\|^2 + 2|(x, y)| + \|y\|^2 \le$
$\le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$

Теорема 2.1.3. (Піфагора) Нехай $z = x + y, \; (x, y) = 0$. Тоді $\|z\|^2 = \|x\|^2 + \|y\|^2$
Доведення. $\|z\|^2 = (z, z) = (x + y, x + y) = (x, x) + (y, y) + (x, y) + (y, x) = \|x\|^2 + \|y\|^2$

Нехай вектори $e_1, e_2, \ldots, e_n$ попарно ортогональні, тобто

$$(e_i,e_j)=\|e_j\|^2\delta_{ij}\text{, де } \delta_{ij}=\begin{bmatrix}1\sim i=j\\0\sim i\neq j\end{bmatrix}\text{ (дельта Кронекера).}$$

Доведемо, що вони лінійно незалежні. Справді, нехай виконується рівність $\sum_{i=1}^{k}x_ie_i=0$. Помноживши скалярно обидві її частини на x_j (j= 1,2,...,k), дістанемо

$$\sum_{i=1}^{k}x_i\|e_j\|^2\delta_{ij}=x_j\|e_j\|^2=0,$$

тобто їх лінійна комбінація дорівнює нулю тоді і лише тоді, коли всі числа де, дорівнюють нулю, що і доводить твердження.

Зауважимо, що при діленні довільного ненульового вектору на його норму дістанемо вектор, норма якого дорівнює одиниці. Такий вектор називається ортом, а операція переходу до вектора одиничної норми — нормуванням. Можна пронормувати базис простору. Одержаний таким чином базис буде називатися нормованим. Якщо всі вектори базису ортогональні між собою, він називається ортогональним. Ортогональний і нормований водночас базис називається ортонормованим.

Нехай $e=\{e_1,...,e_m\}$ - ортонормований базис у m-вимірному підпросторі L n-вимірного простору E_n ($m<n$). Вектору $x\in E$ зіставимо вектор $x_e\in L$: $x_e=\sum_{k=1}^{m}(x,e_k)e_k$. Покажемо, що вектор $x-x_e$ ортогональний до L. Для цього досить довести, що $(x-x_e,e_k)0,\ k=1,...,m$. Маємо:

$$(x-x_e,e_k)=\left(x-\sum_{j=1}^{m}(x,e_j)e_j,e_k\right)=(x,e_k)-(x,e_k)=0.$$

Позначимо $h=x-x_e$. Оскільки h ортогональний до L, $x=x_e+h$, то за теоремою Піфагора $x^2=\|x\|_e^{\ 2}+\|h\|^2$. Вектор $x_e\in L$, що зіставляється вектору $x\in E$, називається ортогональною проекцією x на підпростір L. Очевидно, ця операція співставлення задає лінійний оператор: $P_Lx=x_e$.

Нехай $y=\sum_{k=1}^{m}\beta_ke_k$. Доведемо, що $\|g-y\|\geq\|g-g_e\|$.

$\|g-y\|^2=\|g_e+h-y\|^2=\|(g_e-y)+h\|^2=\|g_e-y\|^2+\|h\|^2\geq\|g_e-y\|^2$, що й доводить твердження. Ця властивість називається екстремальною властивістю проекції: $\|x - x_e\| = \inf\limits_{y \in L} \|x - y\|$. Вираз справа розумно назвати відстанню від x до підпростору L. Будемо позначати її через δ. Пізніше буде знайдено її точний вираз.

Доведемо, що, виходячи з довільного базису, можна побудувати ортонормований. Ця процедура називається метод ортогоналізації Грама - Шмідта.
Нехай $a = \{a_1, ..., a_n\}$ - базис. Покажемо, як за допомогою деяких лінійних перетворень одержати з нього ортонормований базис, що буде позначено $e = \{e_1, ..., e_n\}$.

Перший крок: $e_1 \equiv \dfrac{a_1}{\|a_1\|}$. *Другий крок:* візьмемо допоміжний вектор $b_2 = a_2 - \alpha_1 e_1$, де α_1 знаходиться з умови, що $(b_2, e_1) = 0$. Для цього скалярно помножимо обидві частини попередньої рівності на e_1. Звідси $\alpha_1 = (a_2, e_1)$. Тепер покладемо $e_2 \equiv b_2 / \|b_2\| = \dfrac{a_2 - \alpha_1 e_1}{\|a_2 - \alpha_1 e_1\|}$. Якщо вже зроблено m кроків, покажемо, як треба діяти на $(m+1)$-му кроці. Візьмемо допоміжний вектор $b_{m+1} = a_{m+1} - (a_{m=1}, e_1)e_1 - (a_{m=1}, e_2)e_2 - ... - (a_{m=1}, e_m)e_m$. Очевидно, він ортогональний всім векторам $a_1, ..., a_m$ (перевірте!). І остаточно $e_{m+1} \equiv b_{m+1} / \|b_{m+1}\|$.

Приклади:
1. Простір $\mathbb{R}^n$. Нехай x=($x_1, x_2, ..., x_n$), y=($y_1, y_2, ..., y_n$). Визначимо скалярний добуток таким чином:

$$(x,y) = \sum_i x_i y_i.$$

Читач повинен перевірити, чи виконуються аксіоми скалярного добутку.
2. Простір $\mathbb{C}^n$. Нехай

$$(x,y) = \sum_{i=1}^{n} x_i \overline{y}_i$$

($\overline{y_i}$, — число, комплексно спряжене до y_i).

Висновок. Якщо $\sum_{i=1}^{n} a_i x_i = 0$ при довільних $x_i \in \mathbb{R}^n$ або $\sum_{i=1}^{n} a_i x_i = 0$ при довільних $x_i \in \mathbb{C}^n$, то $a_i = 0, i = 1,...,n$.
Для доведення слід зазначити, що ці суми означають скалярні добутки і вектор $a = (a_1, a_2, ..., a_n)$ — ортогональний довільному вектору $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$. Тоді з попередніх міркувань випливає, що a = 0.
Нехай $E_1 \subset E$ - лінійний підпростір в *E*. Розглянемо множину E_2 векторів, які ортогональні до кожного вектора $x \in E_1$, тобто *(x,y)=0* для довільних $x \in E_1$ і $y \in E_2$. Якщо $y_1, y_2 \in E_2$, тобто $(y_1, x) = (y_2, x) = 0$, для кожного $x \in E_1$, то $\alpha x + \beta y \in E_2$. Дійсно, $(\alpha y_1 + \beta y_2, x) = \alpha(y_1, x) + \beta(y_2, x) = 0$. Це означає, що E_2 утворює підпростір в *E,* що називається ортогональним до E_1. В такому разі будемо писати $E_1 \perp E_2$. З побудови ортогонального підпростору випливає, що $E_1 \cap E_2 = \{0\}$, і в такому разі можна розглянути пряму суму $E_1 \oplus E_2$. Покажемо, що вона збігається з *E.* Для цього виберемо ортонормовані базиси $e \subset E_1$ і $f \subset E_2$. Доведемо, що є $e \cup f$ дає базис всього простору *E.* Дійсно, нехай це не так, і можна знайти лінійно незалежний з цими векторами вектор *x,* ортогональний до всіх векторів з $e \cup f$. Тоді, з одного боку, $x \in f$, тому що *x* ортогональний до e, а з іншого боку, $x \in e$, тому що x ортогональний до *f.* Але $E_1 \cap E_2 = \{0\}$ і тому *x* = 0, а це призводить до протиріччя. Отже, кожен вектор з *E* можна подати у вигляді лінійної комбінації векторів побудованого базису, а тому - у вигляді суми векторів з E_1 і E_2. В такому разі $E = E_1 \oplus E_2$ будемо називати ортогональною прямою сумою.

2.2. ЛІНІЙНІ ФУНКЦІОНАЛИ

Означення 2.2.1. Лінійним функціоналом на *E* називається відображення $f : E \to \mathbb{C}$**,** таке, що $\forall \alpha, \beta \in \mathbb{C}$ $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$. Іноді функціонали називають ковекторами.
Якщо простір *E* нормований, то можна визначити норму функціонала за формулою:

$$\| f \| = \sup_{\|x\|=1} | f(x) | \qquad (2.2.1)$$

З'ясуємо, який вигляд має лінійний функціонал у випадку унітарного простору.
Теорема 2.2.1. Будь-який лінійний функціонал в унітарному просторі має вигляд

$$f(x) = (x, y), \qquad (2.2.2)$$

де y – деякий елемент простору, що визначається за f однозначно. При цьому $\|f\| = \|y\|$.

Доведення. Розглянемо множину G всіх $g \in E$, для яких $f(g) = 0$. Оскільки функціонал лінійний, G є лінійний підпростір в E. Якщо $G = E$, функціонал всюди дорівнює нулю, і тоді треба взяти $y = 0$. Нехай тепер $G \neq E$ і G_1 - ортогональне доповнення в E до G. Тоді існує ненульовий елемент $z \in G_1$. $\forall x \in E$ розглянемо елементи вигляду $f(x)z - f(z)x$. Вони належать до G, оскільки $f\left(f(x)z - f(z)x \right) = f(x)f(z) - f(z)f(x) = 0$. Отже, $\left(f(x)z - f(z)x, z \right) = 0$, звідки $f(x)\left(z, z\right) = \left(x, \overline{f(z)}z\right)$. Якщо покласти $y = \dfrac{\overline{f(z)}}{(z,z)} z$, то отримаємо $f(x) = (x, y)$, що й треба було довести.

Доведемо тепер єдиність такого зображення. Нехай існує такий вектор y_1, що $(x, y_1) = (x, y)$. Тоді $(x, y_1 - y) = 0$ для кожного $x \in E$. Як відзначалось раніше, звідси випливає, що $y_1 - y = 0 \Rightarrow y_1 = y$. Знайдемо норму одержаного вектору y. За нерівністю Коші-Буняковського $|f(x)| = |(x, y)| \leq \|x\|\|y\|$, а тоді $\|f\| \leq \|y\|$. З іншого боку, якщо взяти $x = \dfrac{y}{\|y\|}$, одержимо: $\dfrac{f(y)}{\|y\|} = \|y\|$, звідки $\|f\| \geq \|y\|$. Отже, $\|f\| = \|y\|$, і теорему повністю доведено.

Множина лінійних функціоналів також утворює лінійний простір, що називається спряженим до E і позначається E^*. Треба визначити тільки лінійну комбінацію елементів $f, g \in E^*$:

$$(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) \quad \forall x \in E.$$

Нульовий функціонал діє на x таким чином: $0(x) = 0$. Якщо $f(x) = 0 \ \forall x \in E$, будемо вважати, що $f = 0$.

Для базисних векторів $e_1, ..., e_n \in E$ задамо лінійні функціонали $e^j \in E^*$, $j = 1, ..., n$ за формулами

$$e^j(e_i) = \delta_{ij} = \begin{bmatrix} 1, i = j \\ 0, i \neq j \end{bmatrix}.$$

δ_{ij} називається «дельтою Кронекера».

Доведемо, що ці лінійні функціонали утворюють базис в E.

Спочатку доведемо їхню лінійну незалежність, тобто $\sum \alpha_i e^j = 0$ виконується тоді і лише тоді, коли $\alpha_i = 0$, $i = 1,...,n$. Подіємо цим функціоналом на кожен вектор e_k, $k = 1,...,n$:

$$(\sum \alpha_i e^j)(e_k) = \sum \alpha_i \delta_{ik} = \alpha_k = 0.$$

Доведемо тепер, що кожен функціонал $f \in E^*$ можна зобразити як лінійну комбінацію $e^1,..., e^n$. Перевіримо, що

$$f = \sum_{i=1}^{n} f(e_i) e^i \tag{2.2.3}$$

Вплинемо лівою і правою частинами на довільний вектор $x = \sum_{i=1}^{n} x_j e_j$:

$$f\left(\sum x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i);$$

$$\left(\sum_i f(e_i) e^i\right)\left(\sum_j x_j e_j\right) = \sum_i f(e_i) e^i \left(\sum_j x_j e_j\right) = \sum_{i=1}^{n} f(e_i) x_i ,$$

що і треба було довести.

З цих міркувань можна зробити цікавий висновок: простори E і E^* лінійно ізоморфні між собою, бо ізоморфізм, визначений на базисних векторах, можна по лінійності продовжити на весь простір.

Слід зробити важливе зауваження, що цей ізоморфізм істотно залежить від вибору базису. Пояснимо це на прикладі одновимірного простору. Нехай $e \in E^1$ - базисний вектор, f - лінійний функціонал. Зробимо перетворення $e \to ae$ $(a \neq 0)$. Тоді в просторі лінійних функціоналів (одновимірному) треба зробити перетворення $f \to (1/a) f$, щоб не змінилось значення функціонала на базисному векторі: $(1/a)f(ae) = (a/a)f(e) = f(e)$. Очевидно, ізоморфізм не буде залежати від базису тоді і лише тоді, коли $a = 1/a$, тобто при $a=1$ або $a = -1$.

Ізоморфізм, який не залежить від базису, називається канонічним. Отже, ізоморфізм між E і E^* неканонічний.

Застосуємо тепер теорему про зображення лінійного функціоналу в унітарному просторі. Дію e^j на e_i можна визначити за допомогою скалярного добутку e_j на e_i :

$$e^j(e_i) = (e_i, e_j) = \delta_{ji}.$$

Тоді

$$e^j(x) = (x, e^j) = \left(\sum x_i e_i, e_j\right) = \sum x_i (e_i, e_j) = x_j .$$

Зробимо висновок: для визначення коефіцієнтів розкладу вектору за базисом треба мати спряжений простір або скалярний добуток, який, треба відзначити, може не існувати на деяких структурах, а от спряженість можна визначити в більшості випадків.

2.3. ЛІНІЙНІ ОПЕРАТОРИ

Означення 2.3.1. Нехай E_n і E_m — лінійні *n*- *і m*-вимірні простори над полем *K* ($\mathbb{R}$ або $\mathbb{C}$). Відображення *A:* $E_n \to E_m$ з областю визначення *D(A)* = E_n, що задовольняє умову $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$ $(x, y \in E_n, \alpha, \beta \in K)$, називається лінійним оператором з E_n в E_m.
Оператор *I*, який тотожно діє з E_n в , тобто, за правилом $Ix = x$, називається одиничним.
Нехай $A \in L(E_n, E_m)$ Множина векторів з E_n , яка переводиться оператором *A* в нульовий вектор, є лінійний підпростір (доведення очевидне). Він позначається Ker(*A*) і називається ядром оператора. Розмірність *Ker*(*A*) називається дефектом оператора.
Множина векторів $AE_n = \{y = Ax, x \in E_n\}$ називається образом оператора і позначається Im(*A*). Легко бачити, що це також лінійний підпростір в E_m. Його розмірність називається рангом оператора і позначається *rank (A).*
Лінійний оператор називається невиродженим, якщо він має нульове ядро.

У випадку $E^m = K$ одержуємо вже відомий нам лінійний функціонал.
Норми в E^n і E^m будемо позначати $\|x\|_n$ і $\|y\|_m$ або без індексів, якщо з контексту буде зрозуміло, про який простір йде мова.
Нормою оператора *A* називається число $\|A\|$, яке знаходиться таким чином:

$$\|A\| = \sup_{x \in E_n} \frac{\|Ax\|_m}{\|x\|_n}$$

Оскільки $\|z\|_n = \left\| \frac{x}{\|x\|_n} \right\|_n = \frac{\|x\|_n}{\|x\|_n} = 1$ це можна переписати у вигляді

$$\|A\| = \sup_{\|z\|=1} \|Az\|_m .$$

Можна визначити операцію сум операторів та добуток оператора на число і взагалі лінійну комбінацію операторів *A* і *B:* $E^n \to E^m$ за формулою

$$(\alpha A + \beta B)x = \alpha Ax + \beta Bx .$$

Нульовий оператор 0 визначається так: $0x = 0$ для $x \in E_n$. Треба тільки пам'ятати, що його образ — нульовий вектор в E^n. Очевидно, $\|0\| = 0, \|I\| = 1$.
Протилежний до A оператор є результат множення числа -1 на A: $-A$. Таким чином, оператори, що діють з E_n в E_m, утворюють лінійний нормований простір, який позначається $L(E_n, E_m)$ (нерівність $\|A+B\| \le \|A\| + \|B\|$ випливає з властивостей супремуму). Якщо $A \in L(E_k, E_m)$, $B \in L(E_n, E_k)$, можна визначній оператор $AB \in L(E_n, E_m)$: для $x \in E_n$

$$(AB)x = A(Bx).$$

Якщо $A \in L(E_m, E_k)$, $B \in L(E_n, E_m)$, $C \in L(E_l, E_n)$, то буде справедливий закон:

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

Взагалі кажучи, оператори між собою не комутують.
Розглядаючи оператори, що діють з E_n в E_n, побачимо, що вони утворюють кільце. Як вже відзначалось, в ньому є одиничний оператор – одиниця кільця. Взагалі кажучи, не у всякої квадратної матриці існує обернена, але, якщо вона існує, то, як було показано в 1.6, ліва і права обернені матриці для такої матриці співпадають.
Якщо деяка множина є векторним простором і водночас кільцем, вона називається алгеброю. Отже, $L(E_n, E_n)$ утворюють нормовану алгебру. З визначення супремуму випливає властивість норми: $\|AB\| \le \|A\|\|B\|$; $\|I\| = 1$.
Якщо A^{-1} — обернений до A оператор, то $1 = \|I\| = = \|AA^{-1}\| \le \|A\|\|A^{-1}\|$ і $\|A^{-1}\| \ge \|A\|^{-1}$.

МАТРИЦЯ ЛІНІЙНОГО ОПЕРАТОРА

Знайдемо вигляд оператора $A \in L(E_n, E_m)$ у конкретних базисах $e_1, ..., e_n \in E_n$, $f_1, ..., f_m \in E_m$. Подіємо оператором на всі базисні вектори в E_n і одержані вектори з E_m розкладемо за базисом в E_m.

$$\begin{array}{ccccccccc} Ae_1 & = & a_{11} f_1 & + & a_{21} f_2 & + & ... & + & a_{m1} f_m \\ Ae_2 & = & a_{12} f_1 & + & a_{22} f_2 & + & ... & + & a_{m2} f_m \\ . & . & . & . & . & . & . & . & . \\ Ae_n & = & a_{1n} f_1 & + & a_{2n} f_2 & + & ... & + & a_{mn} f_m \end{array}. \qquad (2.3.1)$$

Простежимо за перетвореннями координат вектору $x = \sum_{k=1}^{n} x_k e_k$ при дії на нього оператора *A*. Нехай $Ax = y = \sum_{j=1}^{m} y_j f_j \in E_m$ Тоді згідно з (2.3.1)

$$Ax = A\sum_{k=1}^{n} x_k e_k = \sum_{k=1}^{n} x_k (Ae_k) = \sum_{k=1}^{n} x_k \left(\sum_{j=1}^{m} a_{jk} f_j \right) = \sum_{j=1}^{m} \left(\sum_{k=1}^{n} a_{jk} x_k \right) f_j = \sum_{k=1}^{n} y_j f_j ,$$

тобто

$$y_j = \sum_{k=1}^{n} a_{jk} x_k \tag{2.3.2}$$

Останню рівність можна тлумачити як скалярний добуток векторів $(a_{j1}, a_{j2}, ..., a_{jn})$ і $(x_1, x_2, ..., x_n)$. Це дає змогу кожному оператору поставити у відповідність прямокутну таблицю з *m* рядками і *n* стовпчиками, що називається матрицею. Оскільки матриця оператора $A \in L(E_n, E_m)$ суттєво пов'язана з базисами просторів, в яких діє оператор, то її зручно позначати $[A]_{f,e}$ і (увага!) коефіцієнти в рядках у розкладі (2.3.1) записувати стовпчиками в матрицю:

$$[A]_{f,e} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \tag{2.3.3}$$

і так зручно робити завдяки формулі (2.3.2): в сумі змінюється другий індекс a_{jk} і індекс вектора-стовпчика *x*. Множину (насправді, лінійний простір) матриць, у яких *m* рядків, *n* стовпчиків будемо позначати $M_{m,n}$, у випадку *m*=*n* будемо писати просто M_n.

Елементи вигляду $a_{ii}, i = 1,...,\min\{m,n\}$ будемо називати діагональними.

Будемо говорити, що елементи $a_{ij}, i - j = k = const$ утворюють квазідіагональ.

Легко бачити, що кожна квазідіагональ паралельна діагоналі.

Означення 2.3.2. Квадратна матриця називається діагональною, якщо всі її недіагональні елементи дорівнюють нулю. Діагональна матриця з однаковими елементами на діагоналі називається скалярною. Скалярна матриця з одиницями на діагоналі називається одиничною. Матриця, всі елементи якої дорівнюють нулю, називається нульовою.

Легко бачити, що діагональні матриці комутують між собою. Матриці з ненульовими елементами на діагоналі утворюють комутативну мультиплікативну групу. Всі діагональні матриці утворюють комутативне кільце.

Означення 2.3.3. Транспонованою по відношенню до *A* називається матриця, кожен *i*-й рядок якої є *i*-м стовпчиком матриці *A*. Для транспонованої матриці будемо вживати позначення A^T. Очевидно, елементи цих двох матриць пов'язані співвідношенням:

$$a_{kj}^T = a_{jk} \tag{2.3.4}$$

З'ясуємо, як знайти елементи матриці, якщо оператор дії в унітарних просторах. Для цього скалярно помножимо обидві частини *j*-ї рівності в (2.3.1) на базисний вектор f_k:

$$\left(Ae_j, f_k\right) = a_{kj}, \tag{2.3.5}$$

$k = 1,...,m;\ j = 1,...,n$.

Для скорочення запису інколи матрицю позначають (a_{ik}), *i*= 1,...,*m*, *k* = 1,... ...,*n*. Зазначимо, що рядки в (2.3.3) записано як стовпчики в (2.3.1), а це зручно завдяки (2.3.2). Тепер можна визначити поняття добутку матриці *A* на вектор *x*, координати якого зручно записувати стовпчиком, згідно з (2.3.2). Його координатами будуть скалярні добутки векторів-рядків матриці на вектор-стовпчик *x:*

$$\left[Ax\right]_f = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ ... & ... & ... & ... \\ a_{j1} & a_{j2} & ... & a_{jn} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} a_{1k} x_k \\ .... \\ \sum_{k=1}^{n} a_{jk} x_k \\ \\ \sum_{k=1}^{n} a_{mk} x_k \end{pmatrix} = \begin{pmatrix} y_1 \\ . \\ y_j \\ . \\ y_m \end{pmatrix} = \left[y\right]_f \tag{2.3.6}$$

Отже, одержали формальне правило:

$$[y]_f = \left[Ax\right]_f = \left[A\right]_{fe} [x]_e. \tag{2.3.7}$$

Зауважимо, що оператору $A \in L(E_n, E_n)$ відповідає квадратна матриця. Очевидно, таке зображенні оператора A за допомогою матриці не єдине, бо вона істотно залежить від вибору базисів. В інших базисах йому відповідатиме інша матриця.

Приклад. Нехай E_4 - лінійний простір многочленів степені не вище третьої. В ньому можна вибрати базис $e = \{1, x, x^2, x^3\}$. Нехай D – оператор першої похідної. Знайдемо його матрицю. Запишемо рівності (2.3.1) :

$$\begin{array}{rcccccccc} D1 & = & 0 \cdot 1 & + & 0x & + & 0x^2 & + & 0x^3; \\ Dx & = & 1 \cdot 1 & + & 0x & + & 0x^2 & + & 0x^3; \\ Dx^2 & = & 0 \cdot 1 & + & 2x & + & 0x^2 & + & 0x^3; \\ Dx^3 & = & 0 \cdot 1 & + & 0x & + & 3x^2 & + & 0x^3. \end{array}$$

За формулою (2.3.3) маємо:

$$[D]_{e,e} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Нехай $P(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \sim [P(x)]_e = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}_e$ Тоді

$$(DP)(x) = a_1 + 2a_2 x + 3a_3 x^2 \sim [(DP)(x)]_e = \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{bmatrix}_e$$

Безпосередньою перевіркою впевнюємось, що $[(DP)(x)]_{e,e} = [D]_{e,e}[P(x)]_e$:

$$\begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{bmatrix}_e = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}_e ,$$

тобто, формулу (2.3.7) підтверджено. ■

Зробимо важливе зауваження. Позначимо через a^j (j=1,…,n) j-й вектор-стовпчик матриці *A*. Тоді формулу (2.3.2) можна записати у такий спосіб:

$$\left[Ax\right]_f = x_1a^1 + ... + x_na^n \tag{2.3.8}$$

Це означає, що довільний вектор з Im(A) можна зобразити у вигляді лінійної комбінації векторів-стовпчиків матриці *A*. Ці вектори, взагалі кажучи, можуть бути лінійно залежними. Розмірність бази векторів-стовпчиків, тобто, ранг цієї системи, називається стовпчиковим рангом матриці. Зрозуміло, що вона і є розмірністю Im(A), тобто $rank(A)$. Аналогічно можна ввести поняття її рядкового рангу. Пізніше буде доведено, що ці ранги співпадають.

Теорема 2.3.1. Нехай лінійний оператор A діє з простору *E* в *E*. Справедлива формула: $\{g_1,...,g_r\}$

$$\dim(E) = \dim(Ker(A)) + \dim(\mathrm{Im}(A)) \tag{2.3.9}$$

Доведення. Нехай $r = rank(A)$, тобто, в підпросторі $\mathrm{Im}(A)$ можна знайти базис з r векторів: $\{g_1,...,g_r\}$. В просторі існують вектори $\{f_1,..., f_r\}$ такі, що $\{g_1 = Af_1,...,g_r = Af_r\}$. Зауважимо, що їхній вибір, взагалі кажучи, не є однозначним. Якби система $\{f_1,..., f_r\}$ була лінійно залежною, тобто існувала б їхня нетривіальна лінійна комбінація, в якій не всі коефіцієнти дорівнюють нулю, що дорівнювала б нулю: $\alpha_1 f_1 + ...\alpha_r f_r = 0$, то, подіявши на обидві частини рівності оператором *A*, мали б, що $\alpha_1 g_1 + ...\alpha_r g_r = 0$, а це протирічить лінійній незалежності системи $\{g_1,...,g_r\}$. З цього випливає, що підпростір *F*, породжений системою $\{f_1,..., f_r\}$, не перетинається з $Ker(A)$. Покажемо, що його сума з $Ker(A)$ породжує весь простір *E*. Нехай $z \in E$. Розглянемо вектор $u = Az \in \mathrm{Im}(A)$. Тоді в підпросторі *F* знайдеться вектор *x*, такий що $Ax = u$ і він записується як лінійна комбінація векторів з системи $\{f_1,..., f_r\}$: $x = \lambda_1 f_1 + ... + \lambda_r f_r$. При цьому $u = Ax = \lambda_1 g_1 + ... + \lambda_r g_r$. Звідси $Az = Ax \Rightarrow A(z - x) = 0$, тобто, $z - x \in Ker(A)$. Але $z = x + (z - x)$. З доведеної раніше формули про розмірність суми двох просторів і випливає справедливість формули (2.3.9).

Зауваження. Нехай матриця діє з n-вимірного простору в себе, тобто, вона квадратна розмірності $n \times n$; r – її ранг. Число $d=n-r$ називається дефектом матриці. З формули (2.3.9) випливає, що дефект – це розмірність $Ker(A)$. З формули (2.3.9) випливає, що матриця з ненульовим ядром (з дефектом більшим за нуль) не може переводити простір на весь простір.

Нехай A — матриця розміром $m \times n$, B — матриця розміром $n \times k$. Тоді матриця $C=AB$ має розмір $m \times k$. З означення добутку двох операторів і попереднього співвідношення випливає, що елемент c_{ij} матриці C природно визначити за формулою

$$c_{ij} = \sum_{p=1}^{k} a_{ip} b_{pj}\,,\ i=1,\ldots,m;\ j=1,\ldots,n. \qquad (2.3.10)$$

Доведемо, що операція множення матриць асоціативна, тобто, для матриць $A(m \times n), B(n \times k), C(k \times l)$ справедлива рівність $(AB)C = A(BC)$.

Введемо позначення $A = \left(a_{i\mu}\right), i=1,...,m, \mu=1,...,n;\quad B=\left(b_{\mu\nu}\right), \mu=1,...,n, \nu=1,...,k;$

$C=\left(c_{\nu\lambda}\right), \nu=1,...,k, \lambda=1,...,l;$

$AB=U=\left(u_{i\nu}\right), i=1,...,m, \nu=1,...,k, (AB)C=UC=S=\left(s_{ij}\right), i=1,...,m,\ j=1,...,l$,

$BC=V=\left(v_{\mu j}\right), \mu=1,...,n,\ j=1,...,l,\ A(BC)=AV=T=\left(t_{ij}\right), i=1,...,m,\ j=1,...,l$.

Маємо:

$$u_{i\nu} = \sum_{\rho=1}^{n} a_{i\rho} b_{\rho\nu},\ i=1,...,m,\ \nu=1,...,k,$$

$$s_{ij} = \sum_{\nu=1}^{k} u_{i\nu} c_{\nu j} = \sum_{\nu=1}^{k}\left(\sum_{\rho=1}^{n} a_{i\rho} b_{\rho\nu}\right) c_{\nu j} = \sum_{\nu=1}^{k}\sum_{\rho=1}^{n} a_{i\rho} b_{\rho\nu} c_{\nu j},\ i=1,...,m,\ j=1,...,l.$$

Завдяки асоціативності додавання чисел в останньому виразі порядки сумування можна поміняти. Отже

$$s_{ij} = \sum_{\nu=1}^{k}\sum_{\rho=1}^{n} a_{i\rho} b_{\rho\nu} c_{\nu j} = \sum_{\rho=1}^{n}\sum_{\nu=1}^{k} a_{i\rho} b_{\rho\nu} c_{\nu j},\ i=1,...,m,\ j=1,...,l.$$

Аналогічно

$$v_{\rho j} = \sum_{\nu=1}^{k} b_{\rho\nu} c_{\nu j},\ \rho=1,...,n,\ j=1,...,l,$$

$$t_{ij}=\sum_{\rho=1}^{n}a_{i\rho}\sum_{\nu=1}^{k}b_{\rho\nu}c_{\nu j}=\sum_{\rho=1}^{n}\sum_{\tau=1}^{k}a_{i\rho}b_{\rho\nu}c_{\nu j} \quad .$$

Отже, $s_{ij}=t_{ij}\,, i=1,...,m,\, j=1,...,l$ і асоціативність множення матриць доведено.

Нехай *X, Y* — лінійні простори і *A : X* →*Y*- лінійний оператор. Розглянемо спряжені простори *X* ,Y** . Нехай $f \in Y^*, x\in X^*$. Визначимо лінійний оператор, який будемо позначати через *A*, A** : *Y** → *X** за формулою

$$\left(A^* f\right)x=f\left(Ax\right) \tag{2.3.11}$$

Тепер розглянемо спряжений оператор у просторах зі скалярним добутком. Нехай E^n і E^m — простори зі скалярним добутком. Розглянемо для $\forall x\in E_n$, $y\in E_m$, $A\in L\left(E_n,E_m\right)$ вираз $\left(y,Ax\right)_m$. Вже відомо, що він визначає деякий лінійний функціонал *f* на E_n в разі $K=\mathbb{R}$ або антилінійний при $K=\mathbb{C}$: $f(x)=\left(y,Ax\right)_m$ (це випливає із властивостей лінійності оператора і відповідно лінійності або антилінійності по другій змінній скалярного добутку). Отже, за теоремою 2.2.1 (яка справедлива і для антилінійних функціоналів; він може бути зображений у вигляді скалярного добутку *x* на деякий елемент $y^*\in E_n$: $\left(y,Ax\right)_m=\left(y^*,x\right)_n$. Таким чином, одержали відповідність $E_m \ni y\to y^*\in E_n$, що буде лінійною завдяки властивостям скалярного добутку. Тобто одержали лінійний оператор *A*,* який буде спряженим до *A*. Отже, якщо $A\in L\left(E_n,E_m\right)$, то $A^*\in L\left(E_m,E_n\right)$ і

$$\left(y,Ax\right)_m=\left(A^*y,x\right)_n$$

З означення випливає, що *(AB)* = B*A* ,* а звідси $(A^{-1})^*=(A^*)^{-1}$.

Нехай (a_{ik})— матриця оператора *A* в деяких ортонормованих базисах. Знайдемо матрицю оператора *A** - (b_{ik}). Користуючись (2.3.5), маємо (при *K*= ***C***)

$$b_{jk}=\left(A^* f_k,e_j\right)=\left(f_k,Ae_j\right)=\overline{\left(Ae_j,f_k\right)}=\overline{a_{kj}}$$

Остаточно

$$b_{jk} = \overline{a_{kj}} \qquad (2.3.12)$$

j=l,...,m; k=1,...,n, тобто всі елементи матриці A треба замінити на спряжені і рядки записати стовпчиками. Таку матрицю зручно назвати спряжено-транспонованою. У випадку K = **R**, природно, одержимо транспоновану матрицю.

Оператор $A \in L(E_n, E_n)$, називається самоспряженим, якщо $A = A^*$. Йому, як ми бачили, повинна відповідати в разі K=**R** симетрична матриця, тобто така, що дорівнює своїй транспонованій. В разі K = **C** її елементи задовольняють рівність $a_{jk} = \bar{a}_{kj}$, тобто після транспонування і заміни всіх елементів на комплексно спряжені одержимо ту ж саму матрицю.

Нехай є два базиси в E_n: $e = \{e_1, ..., e_n\}$ і $e' = \{e'_1, ..., e'_n\}$, причому перехід від e до e' (від нового до старого) дається оператором T: $Te_k = e_k'$. З'ясуємо, як пов'язані координати довільного вектора $x \in E_n$ в цих базисах.

$$x = \sum_{j=1}^{n} x_j e_j = \sum_{k=1}^{n} x'_k e'_k .$$

Розкладемо $Te_k = e_k'$ по старому базису e:

$$Te_k = e_k' = \sum_{j=1}^{n} t_{jk} e_j .$$

Зауважимо, що найпростішим чином коефіцієнти t_{jk} знаходяться, якщо базис e ортонормований. Як вже відзначалося, у такому разі $t_{jk} = \left(e_k', e_j\right)$. При n=2 або n=3 цей скалярний добуток дає косинус кута між цими векторами.

В результаті

$$x = \sum_{k=1}^{n} x'_k \sum_{j=1}^{n} t_{jk} e_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} t_{jk} x'_k \right) e_j = \sum_{j=1}^{n} x_j e_j .$$

Отже, $x_j = \sum_{k=1}^{n} t_{jk} x'_k$, що відповідає дії матриці $[T]_{e,e'}$ оператора T зліва на вектор-стовпчик $[x]_{e'}$ координат $x \in E_n$ у новому базисі:

$$[x]_e = [T]_{e,e'} [x]_{e'} \qquad (2.3.13)$$

або в розгорнутому вигляді

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \dots & \dots & \dots & \dots \\ t_{m1} & t_{m2} & \dots & t_{mn} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} \qquad (2.3.14)$$

Зауваження. Будемо називати канонічним базисом простору систему векторів

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Нехай вектори $e_k^{'} = Te_k, k = 1,...,n$ мають координати $e_k^{'} = \begin{pmatrix} t_{1k} \\ t_{2k} \\ \vdots \\ t_{nk} \end{pmatrix}, k = 1,...,n$.

Тоді безпосередньою перевіркою впевнюємось, що матриця $[T]_{e,e'}$ оператора *T* переходу від старого базису до нового має вигляд

$$[T]_{e,e'} = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \dots & \dots & \dots & \dots \\ t_{m1} & t_{m2} & \dots & t_{mn} \end{pmatrix} \qquad (2.3.15)$$

Підкреслимо, що оператор *T* заміни базису переводить старий базис e у новий e', а його матриця $[T]_{e,e'}$ дії в протилежному напрямку, переводячи вектор-стовпчик $[x]_{e'}$ координат довільного вектора у новому базисі в вектор-стовпчик $[x]_e$ його координат у старому базисі. Виявляється, що це не випадково: якщо задані дві множини *X* і *Y*, простори *Φ(X)* і *Φ(Y)* числових функцій, визначених на цих множинах, і відображення $\varphi : X \to Y$, то простори функцій відображаються в оберненому порядку. Дійсно, нехай $f : Y \to R$ — функція з *Ф(Y).* Їй можна зіставити функцію $g(x) \equiv f(\varphi(x)) \equiv (f\varphi)(x)$ (композиція двох функцій), визначену на *X,* тобто ми побудували відображення *Ф(x)* в *Ф(y),* яке позначається φ *. Таким чином, маємо діаграму

$$\begin{array}{c} \Phi(X) \xleftarrow{\varphi^*} \Phi(Y) \\ X \xrightarrow{\varphi} Y \end{array} \tag{2.3.16}$$

Дійсно, ми вже зустрічалися з такою ситуацією, визначаючи спряжені оператори. Адже лінійні функціонали - окремий випадок функцій, і спряжений оператор діяв у протилежному напрямі. Такі ситуації - типові в теорії категорій, яка є мовою сучасної математики.

Теорема 2.3.2. Для того щоб лінійний оператор $A : E_n \to E_m$ мав обернений, необхідно і досить, щоб він переводив деякий базис $e = \{e_1, ..., e_n\}$ в базис $g = (g_1, ..., g_m)$.

Доведення. Нехай лінійний оператор *A* переводить базис *e* в базис *f*. Обернений оператор A^{-1} можна визначити спочатку на базисних векторах: $A^{-1} g_i = e_i$ i = l,...,m, а потім продовжити на весь простір по лінійності:

$$A^{-1} (\sum x_i g_i) \equiv \sum x_i (A^{-1} g_i) = \sum x_i e_i \ .$$

З останньої рівності випливає, що $AA^{-1} = A^{-1}A = I$. Дійсно,

$$AA^{-1}\left(\sum x_i g_i\right) = A\left(\sum x_i e_i\right) = \sum x_i A g_i = \sum x_i g_i \ .$$

Навпаки, нехай оператор *A* має обернений. Візьмемо деякий базис *e* і подіємо на базисні вектори оператором *A.* Тоді одержимо деяку систему векторів *f* , де $f_i = Ae_i$ *, i* = 1,...,*n*. Доведемо, що одержані вектори лінійно незалежні. Справді, з їхньої лінійної залежності $\sum \alpha_i f_i$ =0 випливає, що $A^{-1}\left(\sum \alpha_i f_i\right)$ = $\sum \alpha_i e_i$ = 0, а це протирічить лінійній незалежності векторів з *e*. Тепер зрозуміло, що система утворює базис, тому що кількість векторів у цій системі дорівнює розмірності простору.

Зауважимо, що оператор, який має обернений, визначений на всьому лінійному просторі, переводить цей простір на себе взаємно однозначно. Такий оператор називається ще невиродженим або неособливим.
Розглянемо питання, як зміниться вигляд матриці оператора $A \in L(E_n, E_m)$ при заміні базиса в цих просторах. Позначимо матрицю оператора в базисах $e = \{e_1, ..., e_n\} \subset E_n$, $g = (g_1, ..., g_m) \subset E_m$ через $[A]_{ge}$, а в нових базисах $e' = \{e_1', ..., e_n'\}$, $g' = (g_1', ..., g'_m)$ через $A_{g'e'}$. Позначимо оператори, які переводять базиси e в e', а g в g' відповідно через T $(Te_k = e'_k)$ і S $(Sg_k = g'_k)$.
Корисно скористатись наступною зрозумілою діаграмою:

$$\begin{tikzcd}
R^n \arrow[rrr, "{[A]_{g,e}}"] \arrow[dr, leftrightarrow] & & & R^m \\
& E_n \arrow[r, "A"] \arrow[d, "T"] & E_m \arrow[ur, leftrightarrow] \arrow[d, "T"'] & \\
& E_n \arrow[r, "A"] \arrow[dl, leftrightarrow] & E_m \arrow[dr, leftrightarrow] & \\
{[x]_{e'} \in R^n} \arrow[uuu, "{[T]_{e,e'}}"] \arrow[rrr, "{[A]_{g'e'}}"] & & & R^m \arrow[uuu, "{[S]_{e,e'}}"']
\end{tikzcd} \qquad (2.3.17)$$

Візьмемо вектор $[x]_{e'} \in R^n$ у лівому нижньому куті діаграми і переведемо його в простір R^m у правому верхньому куті. Це можна зробити двома способами: спочатку множимо його на матрицю $[T]_{e,e'}$, а потім на $[A]_{ge}$ або спочатку множимо на $[A]_{g'e'}$, а потім на $[S]_{e,e'}$. Результати будуть однакові, отже $[A]_{ge}[T]_{e,e'} = [S]_{e,e'}[A]_{g'e'}$. Звідси маємо:

$$[A]_{g'e'} = \left([S]_{e,e'}\right)^{-1}[A]_{ge}[T]_{e,e'} \qquad (2.3.18)$$

Аналогічно можна виписати формулу для знаходження матриці оператора в старих базисах через його матрицю в нових базисах.
Зауважимо, що при $n = m$, $g=e$, $g'=e'$ маємо $S = T$, і формула (2.3.10) набуває вигляду

$$[A]_{e'e'} = \left([T]_{e,e'}\right)^{-1}[A]_{ee}[T]_{e,e'} \qquad (2.3.19)$$

$$[A]_{ee}=([T]_{e,e'})[A]_{e'e'}[T]^{-1}{}_{e,e'} \qquad (2.3.20)$$

Матриці в E^n, пов'язані співвідношенням (2.3.19), називаються подібними. Зараз для з'ясування деяких властивостей будемо використовувати більш прості позначення для матриць, не вказуючи, в яких базисах вони записані.

Отже, нехай є дві пари матриць $A_1=T^{-1}AT$ і $B_1=T^{-1}BT$,то

$A_1B_1=T^{-1}ATT^{-1}BT=T^{-1}ABT$, тобто добутки подібних матриць теж подібні; одинична матриця *E* переходить в себе: $T^{-1}ET=T^{-1}T=E$. Тому відображення подібності є ендоморфізм кільця матриць. Якщо деяка матриця *A* подібна до нульової матриці 0, то $0=T^{-1}AT$ і $AT=0$ (після домноження зліва на *T)* і так само $A=0$, тобто ядро цього ендоморфізму нульове.

ПІДПРОСТОРИ, ІНВАРІАНТНІ ВІДНОСНО ЛІНІЙНОГО ОПЕРАТОРА

Означення 2.3.4. Підпростір $F\subset E$ називається інваріантним відносно лінійного оператора *A,* якщо $AF\subset F$, тобто, $\forall f\in F\Rightarrow Af\in F$.

Зауважимо, що в разі одновимірного підпростору *F* це означає, що для $\forall f\in F\Rightarrow Af=\lambda f$. При цьому вектор *f* називається власним вектором оператора *A*, а λ — його власним числом.

Якщо $f\in Ker(A)$, то $Af = 0$, тобто f — власний вектор оператора *A*, який відповідає власному числу 0.

Коли $G\subset E$ - такий підпростір, що $F\oplus G=E$ і *F* і *G* інваріантні відносно *A,* то будемо говорити, що оператор *A* приводиться підпростором *F* (а також *G).*

Зауважимо, що з інваріантності *F* відносно *A* ще не випливає інваріантність відносно *A* підпростору *G.*

З'ясуємо, якого вигляду набуває матриця оператора *A*.

У базисі $\{e_1,e_2,...,e_m,e_{m+1},...,e_n\}$, де m перших векторів утворюють базис *F,* а останні - базис *G,* якщо *F* інваріантний відносно *A* , а також у випадку, коли *F* приводить *A*.

Нехай *F* інваріантний відносно *A*. Це означає, що $Ae_i\in F$ при $i = 1,...,m$. В такому разі маємо:

$$Ae_i=a_{1i}e_1+a_{2i}e_2+...+a_{mi}e_m+0e_{m+1}+...+0e_n.$$

Оскільки Ae_i не належить лише, скажемо, *G* або *F,* то

$$Ae_i = a_{1i}e_1 + a_{2i}e_2 + ... + a_{mi}e_m + a_{m+1,i}e_{m+1} + ... + a_{ni}e_n,$$
$i = m + 1,...,n.$

При цьому матриця $[A]_{e,e}$ оператора має вигляд

$$\begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} & a_{1m+1} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2m} & a_{2m+1} & ... & a_{2n} \\ ... & ... & ... & ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mm} & a_{mm+1} & ... & a_{mn} \\ 0 & 0 & ... & 0 & a_{m+11m+1} & ... & a_{m+1n} \\ 0 & 0 & ... & 0 & a_{m+2m+1} & ... & a_{m+2n} \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & a_{nm+1} & ... & a_{nn} \end{pmatrix} \qquad (2.3.21)$$

Якщо підпростір *F* приводить *A,* то матриця $[A]_{e,e}$ набуває вигляду

$$[A]_{e,e} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1m} & 0 & ... & 0 \\ a_{21} & a_{22} & ... & a_{2m} & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mm} & 0 & ... & 0 \\ 0 & 0 & ... & 0 & a_{m+1,1m+1} & ... & a_{m+1,n} \\ 0 & 0 & ... & 0 & a_{m+2,m+1} & ... & a_{m+2,n} \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & a_{n,m+1} & ... & a_{nn} \end{pmatrix} \qquad (2.3.22)$$

Якщо $E = E_1 \oplus E_2$ і оператор *A* інваріантний відносно E_1 і E_2, то з (2.3.21) випливає існування операторів A_1 і A_2, що $Ax = A_1x_1 + A_2x_2$, де $x \in E, x_1 \in E_1, x_2 \in E_2$. У такому разі оператор *A* називається прямою сумою операторів A_1 і A_2.

З теореми 2.3.1 автоматично випливає висновок — оператор *A* не має оберненого тоді і лише тоді, коли він деякі лінійні незалежні елементи переводить в лінійно залежні, зокрема базис — в систему лінійно залежних векторів.

Відзначимо, що E може завжди бути розкладений в ортогональну пряму суму такого вигляду:

$$E = \mathrm{Ker}(A) \oplus \mathrm{Im}(A^*). \tag{2.3.23}$$

Дійсно, якщо де $x \in \mathrm{Ker}(A)$, то $Ax=0$ і тому для довільного $y \in E$ маємо: $0 = (Ax,y) = (x,A^*y)$, тобто $A^*y \perp x$, а як вже відомо, пряма сума деякого підпростору і його ортогонального доповнення дає весь простір.

З формули (2.3.14) одержуємо цікавий наслідок: якщо $A=A^*$ (самоспряжений оператор), то

$$E = \mathrm{Ker}(A) \oplus \mathrm{Im}(A). \tag{2.3.24}$$

В такому разі його матриця A, (в базисі e, про який йшла мова) має вигляд:

$$[A]_{e,e} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2m} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \tag{2.3.25}$$

У загальному випадку для довільного оператора A формула вигляду (2.3.24) невірна. Можна лише стверджувати, що

$$\dim (E) = \dim (\mathrm{Ker}(A) + \dim (\mathrm{Im}\,(A)). \tag{2.3.25}$$

Доведення. Нехай $\{e_1, e_2, ..., e_k\}$— базис в Ker(φ). Доповнимо його до базису $\{e_1, e_2, ..., e_k, e_{k+1}, ..., e_n\}$ всього простору. Таким чином, E — пряма сума Ker (A) і підпростору E_1, натягнутого на вектори $\{e_{k+1}, ..., e_n\}$. Подіявши на вектори

$e_{k+1},...,e_n$ оператором *A,* одержимо вектори $Ae_{k+1},...,Ae_n$, що є лінійно незалежними. Дійсно якщо $\sum_{i=k+1}^{n} a_i A a_i = 0$, то $A(\sum_{i=k+1}^{n} a_i e_i) = 0$, і це мало значити, що $\sum_{i=k+1}^{n} a_i e_i \in Ker(A)$, а це суперечить тому, що $E = Ker(A) \oplus E_1$.

Наведемо деякі цікаві приклади, які допоможуть краще зрозуміти викладену теорію.

Приклад 1. Перетворення координат на площині. Нехай є система координат *x*0*y*. Повернемо її на кут φ (φ >0, коли поворот здійснюється проти годинникової стрілки, φ <0, якщо за годинниковою стрілкою), і одержимо нову систему $x'0y'$. Знайдемо зв'язок між координатами вектора в старій і новій системах координат. Нехай $\{e_1, e_2\}$ і $\{e_1', e_2'\}$ - ортонормовані базиси, направляючі орти осей. З умови випливає, що кут між e_1' *і* e_1 дорівнює φ, між e_1' *і* e_2 дорівнює $\frac{\pi}{2} - \varphi$, між e_2' і e_1 дорівнює $\frac{\pi}{2} + \varphi$, між e_2' і e_2 дорівнює φ.

Застосуємо відому схему визначення матриці оператора повороту $U(\varphi)$

$$\begin{cases} e_1' = U(\varphi)e_1 = u_{11}e_1 + u_{21}e_2 \\ e_2' = U(\varphi)e_2 = u_{12}e_1 + u_{22}e_2 \end{cases}$$

Для визначення коефіцієнтів розкладу помножимо скалярно обидві частини кожного розкладу на вектори e_1 *і* e_2, враховуючи, що скалярний добуток одиничних векторів дорівнює косинусу кута між ними. В результаті одержимо:

$$\begin{cases} e_1' = U(\varphi)e_1 = (\cos\varphi)e_1 + (\sin\varphi)e_2 \\ e_2' = U(\varphi)e_2 = -(\sin\varphi)e_1 + (\cos\varphi)e_2 \end{cases}$$

$$[U(\varphi)]_{e,e'} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \tag{2.3.26}$$

Нехай r – довільний вектор площини, $r = x_1 e_1 + x_2 e_2 = x_1' e_1 + x_2' e_2$. Як було доведено,

$$[r]_e = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \left[U(\varphi)\right]_{e,e'} [r]_{e'} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix},$$

звідки одержимо:

$$\begin{cases} x_1 = x_1' \cos\varphi - x_2' \sin\varphi \\ x_2 = x_1' \sin\varphi + x_2' \cos\varphi \end{cases} \qquad (2.3.27)$$

Зауважимо, що виконання двох по черзі поворотів на кути φ, а потім ψ або спочатку на ψ, а потім на φ еквівалентне повороту на кут $\varphi + \psi$. Що ж до координат вектора при цих поворотах, то вони множаться на матрицю $U(\psi)U(\varphi)$ або на $U(\varphi + \psi)$ і результати повинні бути однакові, тобто

$$U(\varphi + \psi) = U(\psi)U(\varphi) \qquad (2.3.28)$$

Робимо висновок: такі матриці між собою комутують (читач може перевірити це безпосередньо). Знайшовши добуток, маємо

$$\begin{pmatrix} \cos\varphi\cos\psi - \sin\varphi\sin\psi & -\left(\sin\varphi\cos\psi + \cos\varphi\sin\psi\right) \\ \sin\varphi\cos\psi + \cos\varphi\sin\psi & \cos\varphi\cos\psi - \sin\varphi\cos\psi \end{pmatrix} =$$

$$= \begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix},$$

звідки одержуємо відомі з тригонометрії формули:

$$\cos(\varphi + \psi) = \cos\varphi\cos\psi - \sin\varphi\sin\psi,$$
$$\sin(\varphi + \psi) = \sin\varphi\cos\psi + \cos\varphi\sin\psi.$$

Зауважимо, що при $\varphi = 0$ повороту немає, і такому виродженому повороту відповідає одинична матриця. Якщо після повороту на кут φ виконаємо поворот на кут $-\varphi$, тобто, в протилежному напрямку, то в результаті виконається поворот на 0, а це для матриць набуває вигляду

$$U(-\varphi)U(\varphi) = I$$

Робимо висновок, що $\left(U(\varphi)\right)^{-1}=U(-\varphi)$ - обернена матриця, яку одразу можна знайти, користуючись властивостями тригонометричних функцій

$$(U(\varphi))^{-1}=U(-\varphi)=\begin{bmatrix}\cos(-\varphi) & -\sin(-\varphi)\\ \sin(-\varphi) & \cos(-\varphi)\end{bmatrix}=\begin{bmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{bmatrix} \qquad (2.3.29)$$

Отже, одержали транспоновану матрицю. Остаточно

$$\left(U(\varphi)\right)^{-1}=U(-\varphi)=\left(U(\varphi)\right)^{T} \qquad (2.3.30)$$

Тепер можна нові координати виразити через старі, користуючись співвідношенням (2.3.29):

$$\begin{pmatrix} x_1^{'}\\ x_2^{'}\end{pmatrix}=\left[U(\varphi)\right]_{e,e'}^{-1}\begin{pmatrix} x_1\\ x_2\end{pmatrix}=\begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix} x_1\\ x_2\end{pmatrix}=\begin{pmatrix} x_1\cos\varphi+x_2\sin\varphi\\ -x_1\sin\varphi+x_2\cos\varphi\end{pmatrix}.$$

Звідси

$$\begin{cases} x_1'=x_1\cos\varphi+x_2\sin\varphi\\ x_2'=-x_1\sin\varphi+x_2\cos\varphi\end{cases} \qquad (2.3.31)$$

Зміст формул (2.3.28), (2.3.29) можна висловити мовою теорії груп. Матриці такого типу утворюють мультиплікативну абелеву групу. Ми одержали ізоморфізм між групою поворотів кола на довільні кути і групою матриць. Зауважимо, що на комплексній площині поворот на кут φ відповідає множенню на число $e^{i\varphi}$, а такі числа утворюють підгрупу в $\mathbb{C}\setminus 0$. У цьому випадку одержимо також ізоморфізм цих груп і групи $\{e^{i\varphi},\varphi\in[0,2\pi)\}$. Згадавши, що довільне комплексне число $z=a+ib$ може бути подане у вигляді $a+ib=|z|e^{i\varphi}=|z|\cos\varphi+i|z|\sin\varphi$, ($a=|z|\cos\varphi$, $b=|z|\sin\varphi$), то тепер числу z можна зіставити матрицю

$$a+ib\sim|z|\begin{bmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{bmatrix}=\begin{bmatrix}a & -b\\ b & a\end{bmatrix} \qquad (2.3.32)$$

а комплексно спряженому ($\overline{z}=a-ib=|z|e^{-i\varphi}$) -

$$a - ib \sim \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \qquad (2.3.33)$$

Визначником матриці $A = (a_{ik})$ другого порядку називається число $a_{11}a_{22} - a_{12}a_{21}$ (пізніше ми дамо означення визначника довільної квадратної матриці). З цього прикладу ми бачимо, що визначник побудованої матриці є квадратом модуля комплексного числа, $|z| = \sqrt{a^2 + b^2}$, а визначники матриць повороту осей координат дорівнюють одиниці. Якщо взяти два вектори $x,y \in \mathbf{R}^2$, які утворюють кут α , то скалярний добуток між ними $(x, y) = \|x\|\|y\|\cos\alpha$ не зміниться після дії на кожен вектор оператором $U(\varphi)$, тому що при повороті евклідові норми вектора та кут між ними не змінюються: $(x,y)=(U(\varphi)x,U(\varphi)y)$.

Приклад 2. Матриці перестановок.

Так називаються матриці *P* з M_n, у яких у довільному рядку і у довільному стовпчику знаходиться рівно одна одиниця, а всі інші елементи – нулі. Покажемо на прикладі, як діє на вектор *a* матриця перестановок з M_4:

$$Pa = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} a_2 \\ a_3 \\ a_1 \\ a_4 \end{pmatrix}.$$

Як бачимо, одиниця в першому рядку матриці *P* стоїть на другому місці (номер стовпчика - 2), і при множенні друга a_2 координата вектора перемістилась на перше місце, в другому рядку матриці одиниця на третьому місці, і третя координата a_3 вектора перемістилась на друге місце. Читач може самостійно прослідкувати за переміщеннями останніх координат вектора. З іншого боку, перша координата a_1 пересунулась на третє місце, тому що в третьому рядку на першому місці знаходиться одиниця; друга координата a_2 пересунулась на переше місце, тому що в першому рядку на другому місці знаходиться одиниця і т.ін. Узагальнюючи проведені дослідження, неважко впевнитись, що виконуються правила, що будуть сформульовані нижче.

1). Нехай на місці (i, j) матриці *P* знаходиться одиниця. Тоді при множенні її на вектор на *i*-е місце переставиться *j*-та координата a_j вектора a .

2). i-та координата a_i вектора a пересунеться на місце j, якщо на місці (j,i) матриці P знаходиться одиниця.

Проста перевірка показує, що добуток матриць перестановок – також матриця перестановок (елементи в кожному стовпчику матриці, яка є правим множником, будуть переставлятись за вказаним правилом). Якщо розглядати різні степені цієї матриці, то через те, що вектор має скінченну кількість координат, значення $P^k a$ і $P^l a$ при певних різних k і l співпадуть, звідки $P^{|k-l|}a = a$, отже, P у деякому степені дорівнює одиничній матриці. Звідси випливає, що кожна матриця перестановок має обернену, тобто, вона є дільником одиниці, а тому, як було показано раніше, вона не є дільником нуля, значить, така матриця є невиродженою. Більш детальний аналіз показує, що $P^{-1} = P^T$. Для доведення треба скористатись правилом 2). Нехай в P одиниця стояла на місці (j, i) , то в P^T вона знаходиться на місці (i, j). Отже, при множенні вектора на матрицю P i-та координата a_i пересунеться на місце j, а при множенні одержаного вектора на P^T a_j пересунеться на i-те місце, тобто, кожна i-та координата залишиться на місці при множенні на $P^T P$. Отже, $P^T P = I$.

При множенні даної матриці на матриці перестановок вказані перетворення відбуваються з векторами-стовпчиками. Радимо читачу перевірити це самостійно.

Із всього сказаного випливає, що матриці перестановок утворюють групу, одиницею якої є одинична матриця, причому, некомутативну при $n > 2$, і вона ізоморфна групі перестановок з n елементів, що розглядалась в главі 1.

Розглянемо так звану основну циркулянтну матрицю перестановки:

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

Дослідимо степені цієї матриці:

$$C^2 = \begin{pmatrix} 0 & 0 & 1 & . & . & . & 0 \\ 0 & 0 & 0 & 1 & . & . & . \\ . & & . & . & . & & . \\ . & & & . & . & . & . \\ . & & & & . & . & 1 \\ 1 & 0 & . & . & . & 0 & 0 \\ 0 & 1 & . & . & . & 0 & 0 \end{pmatrix}.$$

Як бачимо, верхня «наддіагональ» пересувається вправо-вгору і з одиниці в лівому нижньому куті виникає нижня «піддіагональ». При піднесенні до кубу обидві квазідіагоналі зсуваються вправо-вгору. Звідси випливає. Що $C^n = I$. Будемо також вважати, що $C^0 = I$.

Приклад 3. Циркулянтні матриці.
$A \in M_n$

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & . & . & . & a_n \\ a_n & a_1 & a_2 & . & . & . & a_{n-1} \\ a_{n-1} & a_n & a_1 & . & . & . & a_{n-2} \\ . & . & . & . & & & . \\ . & . & & . & . & . & . \\ . & . & & & . & . & . \\ a_2 & a_3 & . & . & . & a_n & a_1 \end{pmatrix}.$$

Кожен наступний рядок – це результат циклічного зсуву попереднього, отже, кожен рядок – деяка циклічна перестановка першого рядка. Легко бачити, що кожна квазідіагональ цієї матриці складається з однакових елементів: головна діагональ – з елементів a_1, квазідіагональ на сусідній позиції справа від головної і елемент в лівому нижньому куті – з елементів a_2 - будемо їхні номери позначати (2,*n)*, на другій квазідіагоналі справа від головної і на другій квазідіагоналі зліва-унизу з елементів a_3 - номери (3,*n*-1) і т.ін. Отже, елементи матриці на вказаних квазідіагоналях можна одержати, якщо одиничну матрицю помножити на a_1, пару квазідіагоналей (2,*n)* основної циркулянтної матриці перестановок *C* помножити на a_2, елементи квазідіагоналей (3,*n*-1) C^2 помножити на a_3 і т.ін. В результаті одержимо формулу:

$$A=\sum_{k=0}^{n-1} a_{k+1} C^k .$$

Приклад 4. Матриці Тьоплиця.

$A \in M_{n+1}$.

Нехай є скінченна послідовність комплексних чисел $\{a_{-n}, a_{-n+1}, ..., a_{-1}, a_0, a_1, ..., a_{n-1}, a_n\}$. Розглянемо матрицю

$$A=\begin{pmatrix} a_0 & a_1 & a_2 & . & . & . & a_n \\ a_{-1} & a_0 & a_1 & . & . & . & a_{n-1} \\ a_{-2} & a_{-1} & a_0 & . & . & . & . \\ . & . & . & . & & & . \\ . & . & & . & . & . & . \\ a_{-n+1} & a_{-n+2} & & & . & . & a_1 \\ a_{-n} & a_{-n+1} & . & . & . & a_{-1} & a_0 \end{pmatrix}.$$

Розглянемо матриці Тьоплиця (в іншій главі вони будуть називатись клітинами Жордана):

$$R=\begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 0 & 0 & 1 & . & . & . & 0 \\ . & & . & . & & & . \\ . & & & . & . & & . \\ . & & & & . & . & 0 \\ 0 & 0 & . & . & . & 0 & 1 \\ 0 & 0 & . & . & . & 0 & 0 \end{pmatrix}, \quad L=\begin{pmatrix} 0 & 0 & & . & . & . & 0 \\ 1 & 0 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & . & . & . & 0 \\ . & & & . & . & . & . \\ . & & & & . & . & 0 \\ 0 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & . & . & . & 1 & 0 \end{pmatrix}.$$

Легко бачити, що справедливо наступне зображення матриці Тьоплиця:

$$A=\sum_{k=1}^{n} a_{-k} L^k + \sum_{k=0}^{n} a_k R^k .$$

Означення 2.3.5. Лінійний оператор $V \in L(E, E)$ називається ізометричним, якщо $(Vx, Vy) = (x, y)$.

З означення випливає, що $\|Vx\|^2=(Vx,Vx)=(x,x)=\|x\|^2$ норма ізометричного оператора дорівнює одиниці.

Означення 2.3.6. Ізометричний оператор називається унітарним, якщо його образ – весь простір.

Зауваження. В скінченновимірних просторах ізометричний оператор, визначений на всьому просторі, є обов'язково унітарним. Дійсно, він переводить базис у базис, тому що ніяку нетривіальну лінійну комбінацію він не може перевести в нуль через збереження норми векторів.

Лема 2.3.1. Якщо лінійний оператор, визначений на всьому просторі має властивість $\|Ux\|^2=\|x\|^2$, то він унітарний.

Доведення. $\forall\alpha\in\mathbb{C}$ маємо $(U(x+\alpha y),U(x+\alpha y))=(x+\alpha y,x+\alpha y)$.

Звідси

$(Ux,Ux)+(Ux,\alpha Uy)+(\alpha Uy,Ux)+|\alpha|^2(Uy,Uy)=(x,x)+(x,\alpha y)+(\alpha y,x)+|\alpha|^2(y,y)$,

і, враховуючи умову, одержимо:

$(Ux,\alpha Uy)+(\alpha Uy,Ux)=(x,\alpha y)+(\alpha y,x)$.

Підставимо по черзі $\alpha=1,\alpha=i$. Тоді будемо мати

$(Ux,Uy)+(Uy,Ux)=(x,y)+(y,x),$

$-i(Ux,Uy)+i(Uy,Ux)=-i(x,y)+i(y,x).$

В результаті

$(Ux,Uy)+(Uy,Ux)=(x,y)+(y,x),$

$-(Ux,Uy)+(Uy,Ux)=-(x,y)+(y,x).$

Після віднімання від першої рівності другу і скорочення, прийдемо до умови унітарності:

$$(Ux,Uy)=(x,y)$$

Нехай V – унітарний оператор. Тоді $(V^*Vx,y)=(x,y)\Rightarrow((V^*V-I)x,y)=0$.

Оскільки це справедливо для довільного вектора y, то $V^*V=I\Rightarrow V^*=V^{-1}$

Звідси випливає: щоб знайти матрицю оператора, оберненого до унітарного (базис ортонормований), треба матрицю даного оператора транспонувати і здійснити комплексне спряження для кожного елемента.

Оскільки $VE=E$, обернений оператор визначено на всьому просторі, і оскільки V^* - лівий обернений, то він є і правим оберненим: $VV^*=I$. Звідси випливає, що V^* - теж унітарний: $(V^*x,V^*y)=(VV^*x,y)=(x,y)$.

З правила множення матриць випливає, що множення i-го рядка $v_i=(v_{i1}v_{i2}...v_{in})$ матриці V^* на j-й стовпчик $(v_j)^*=(v^*_{1j}v^*_{2j}...v^*_{nj})^T$, одержимо 0, якщо $i\neq j$ і 1

при $i=j$. Тобто, $\sum_{k=1}^{n} v_{ik}v_{kj}^{*}=\sum_{k=1}^{n} v_{ik}\overline{v}_{jk}=\delta_{ij}$, але ця рівність говорить, що скалярний добуток різних рядків унітарної матриці дорівнює нулю, а квадрат норми кожного рядка дорівнює одиниці. Те ж саме стосується її стовпчиків.
Розглянемо визначник унітарної матриці. $\det\left(V^{*}\right)=\det\left(\overline{V}^{T}\right)=\det\left(\overline{V}\right)$. Оскільки $1=\det\left(VV^{*}\right)=\det\left(V\right)\det\left(V^{*}\right)=\det\left(V\right)\det\left(\overline{V}\right)=\det\left(V\right)\overline{\det\left(V\right)}=\left|\det\left(V\right)\right|^{2}$, отже, $\left|\det\left(V\right)\right|=1$.

Якщо повернутись до матриць $U(\varphi)$ з прикладу 1, побачимо, що вони здійснюють унітарне перетворення простору E_2. Матриці в дійсному просторі, які задовольняють умову $A^{-1}=A^{T}$, називаються ортогональними. Група ортогональних матриць в E_n над полем K позначається O(n,K). Матриці з визначником одиниця утворюють спеціальну лінійну групу $SL(n,K)$.
$O\left(n,K\right)\cap SL(n,K)=SO(n,K)$ - спеціальна ортогональна група. Отже, більш точно, матриці $U(\varphi)$ належать до $SO\left(2,\mathbb{R}\right)$.

Приклад 2. Заміна координатних осей. Здійснимо перетворення координат так, щоб осі Оx і Оy помінялися місцями, тобто дзеркальне відображення відносно бісектриси кутів першої-третьої чвертей. У такому разі вектор з координат $r=(x,y)$ перейдуть у вектор з координат $r'=(y,x)$. Такому перетворенню відповідає матриця

$$J=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \qquad (2.3.34)$$

(перевірте!).
Для того, щоб виконати дзеркальне відображення і поворот, треба перемножити матриці J і $U(\varphi)$: $U(\varphi)J$.

Приклад 3. Ортогональне проектування. Нехай вектор $r=(x,y,z)^{T}\in\mathbb{R}^{3}$. Спроектуємо його на горизонтальну площину. Оператор ортогонального проектування (ортопректор) P діє за таким законом: $P(x,y,z)^{\mathrm{T}}=(x,y,0)^{\mathrm{T}}$, Тоді його матриця матиме вигляд:

$$P=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}. \qquad (2.3.35)$$

Легко довести, що проектор (не обов'язково ортогональний) задовольняє властивість $P^2=P$, яка називається ідемпотентністю.

Радимо читачеві виписати матриці проекторів на площини xOy, xOz та координатні осі Ox, Oy, Oz.

Взагалі проектором називається лінійний оператор, визначений на лінійному просторі E, що задовольняє умову $P^2 = P$. Нехай I — одиничний оператор. Тоді оператор $Q = I - P$ також є проектором. Дійсно, $Q^2 = (I—P)^2 = = I\text{-}2P+P^2 = I\text{-}2P+P = I\text{-}P = Q$. Далі, підпростори $E_1 = PE$ і $E_2 = QE = (I - P)E$ мають нульовий перетин. Дійсно, якщо $x \in E_1 \cap E_2$, то існують вектори $y \in E_1$, $z \in E_2$, що $x = Py= (I\text{-}P)z$. Звідси $P(y+z) = z$. Подіявши на обидві частини цієї рівності проектором P, дістанемо: $Py+Pz=Pz$, звідки $x=Py=0$. З іншого боку, довільний вектор $x \in E$ можна подати у вигляді суми $x=Px+(I\text{-}P)\, x=Px+Qx$. Це означає, що $E = E_1 \oplus E_2$. Якщо $E_1 \perp E_2$, то P називається ортогональним проектором або ортопроектором.

Теорема 2.3.3. Для того щоб проектор P був ортогональним, необхідно і достатньо, щоб він був самоспряженим.

Доведення. Нехай P — ортогональний проектор. Тоді для довільних $x \in E$, $y \in E$ $Px \perp Qy$, тобто, $(Px,(I\text{-}P)y)= 0$, звідки $(Px, (I\text{-}P)y) = ((I\text{-}P^*)Px, y) = =((P\text{-}P^*P)x,y)= 0$. Оскільки y — довільний вектор, $(P\text{-}P^*P)x = 0$, звідки з довільності x випливає, що $P^*P = P$. З іншого боку, $(P^*P)^*=P^*P = P^*$ (якщо знайти спряжений оператор до обох частин рівності), звідки $P=P^*$. Навпаки, нехай $P =P^*$. Тоді для довільного $x \in E$ $(Px,(I\text{-}P)x) = ((I\text{-}P)Px,x) = =((P\text{-}PP)x,x)=((P\text{-}P)x,x)=0$, тобто $Px \perp Qx$.

ЕЛЕМЕНТАРНІ ПЕРЕТВОРЕННЯ

Приклад 4. Елементарні перетворення. Розглянемо матриці, які будемо називати елементарними:

1) «збурена одинична» - діагональна матриця

$$D_{ij}=\begin{pmatrix}
1 & & & & & & & & & & \\
 & . & & & & & & & & & \\
 & & . & & & & & 0 & & & \\
 & & & . & & & & & & & \\
 & & & & 1 & & & & & & \\
 & & & & & d_{ii} & & & & & \\
 & & & & & & 1. & & & & \\
 & & & & & & & . & & & \\
 & & & 0 & & & & & . & & \\
 & & & & & & & & & . & \\
 & & & & & & & & & & 1
\end{pmatrix} \tag{2.3.36}$$

2) матриця перестановки

$$P_{ij}=\begin{pmatrix}
1 & & & & & & & & & & & & & & & & \\
 & . & & & & & & & & & & & & & & & \\
 & & . & & & & & & & & & & & & 0 & & \\
 & & & . & & & & & & & & & & & & & \\
 & & & & 1 & & & & & & & & & & & & \\
 & & & & & 0 & . & & . & & . & 1 & & & & & \\
 & & & & & . & 1 & & & & & . & & & & & \\
 & & & & & & & . & & & & & & & & & \\
 & & & & & . & & & . & & & . & & & & & \\
 & & & & & & & & & . & & & & & & & \\
 & & & & & . & & & & & 1 & . & & & & & \\
 & & & & & 1 & . & & . & & . & 0 & & & & & \\
 & & & & & & & & & & & & 1 & & & & \\
 & & & & & & & & & & & & & . & & & \\
 & & & & & & & & & & & & & & . & & \\
 & & & & & & & & & & & & & & & . & \\
 & & & & & & & & & & & & & & & & 1
\end{pmatrix} \tag{2.3.37}$$

Ця матриця утворилась з одиничної шляхом перестановки рядків i-го і j-го $(i<j)$.

3) Нехай $j>i$. Визначимо матриці $M^{(l)}{}_{ji}$ та $M^{(r)}{}_{ji}$:

$$M^{(l)}{}_{ji} = \begin{pmatrix}
1 &&&&&&&&&&&&&&&& \\
& \cdot &&&&&&&&&&&&&&& \\
&& \cdot &&&&&&&&&&&& 0 && \\
&&& \cdot &&&&&&&&&&&&& \\
&&&& 1 &&&&&&&&&&&& \\
&&&&& 1 &&&&&&&&&&& \\
&&&&& \cdot & 1 &&&&&&&&&& \\
&&&&&&& \cdot &&&&&&&&& \\
&&&&& \cdot &&& \cdot &&&&&&&& \\
&&&&&&&&& \cdot &&&&&&& \\
&&&&& \cdot &&&&& 1 &&&&&& \\
&&&&& m_{ji} & \cdot && \cdot && \cdot & 1 &&&&& \\
&&&&&&&&&&&& 1 &&&& \\
&&&&&&&&&&&&& \cdot &&& \\
&&&&&&&&&&&&&& \cdot && \\
&&&&&&&&&&&&&&& \cdot & \\
&&&&&&&&&&&&&&&& 1
\end{pmatrix} \qquad (2.3.38)$$

Якщо $j < i$, то

$$
M^{(r)}{}_{ji} = \begin{pmatrix}
1 &&&&&&&&&&&&&&&& \\
& \cdot &&&&&&&&&&&&&&& \\
&& \cdot &&&&&&&&&&& 0 &&& \\
&&& \cdot &&&&&&&&&&&&& \\
&&&& 1 &&&&&&&&&&&& \\
&&&&& 1 & \cdot && \cdot && \cdot & m_{ji} &&&&& \\
&&&&&& 1 &&&&& \cdot &&&&& \\
&&&&&&& \cdot &&&&&&&&& \\
&&&&&&&& \cdot &&& \cdot &&&&& \\
&&&&&&&&& \cdot &&&&&&& \\
&&&&&&&&&& 1 & \cdot &&&&& \\
&&&&&&&&&&& 1 &&&&& \\
&&&&&&&&&&&& 1 &&&& \\
&&&&&&&&&&&&& \cdot &&& \\
&&&&&&&&&&&&&& \cdot && \\
&&&&&&&&&&&&&&& \cdot & \\
&&&&&&&&&&&&&&&& 1
\end{pmatrix}. \qquad (2.3.39)
$$

Ми бачимо, що при множенні на D_{ij} зліва довільної матриці A добуток $D_{ij}A$ відрізняється від A тим, що його i-й рядок — це результат множення на d_{ij} i-го рядка A. При множенні на D_{ij} справа досить i-й стовпчик A помножити на d_{ij}.

При множенні матриці A на P_{ij} зліва, очевидно, переставляються i-й та j-й рядки матриці A. При множенні справа на P_{ij} те ж саме відбуватиметься зі стовпчиками A.

При множенні матриці A на $M_{ji}^{(l)}$ зліва ($j>i$) до j-го рядка (нижчого) матриці A буде додаватись її i-й рядок (вищий), помножений на m_{ji}, і результат запишеться в j-му рядку. Аналогічно, якщо $j< i$, при множенні матриці A на M_{ji}^{r} зліва до вищого, j-го рядка, буде додаватись нижчий i-й рядок, помножений на m_{ji}.

Аналогічні правила діють при множенні матриці справа на $M_{ji}^{(l)}$ та M_{ji}^{r}. Треба тільки рядок замінити на стовпчик.

Очевидно, $D_{ii} = D^T{}_{ii}$ і $P_{ij} = P^T{}_{ij}$, тобто ці матриці симетричні, вони збігаються зі своїми спряженими. $M_{ji}^{(l)}$ та M_{ji}^{r} - не симетричні матриці, $\left(M_{ji}^{(l)}\right)^T = M_{ij}^{r}$.
Для елементарних матриць легко знайти обернені. Якщо $d_{ii} \neq 0$, то матриця D_{ii}^{-1} буде мати такий же самий вигляд, як і D_{ii}, тільки в ній замість елемента d_{ii} стоятиме елемент d_{ii}^{-1}.
$P_{ij}^{-1} = P_{ij} = P^T{}_{ij}$, тому що $P_{ij}^2 = E$. Це зрозуміло і без обчислень, тому що одноразове множення на P_{ij} переставить *i*-й та *j*-й рядки *A,* а друге множення поверне їх на попередні місця, тобто P_{ij}^2 діє як тотожне перетворення. M_{ij}^{-1}— матриця такого ж вигляду, як і P_{ij}, але замість елемента m_{ij} вона матиме елемент $-m_{ij}$. Це також можна зрозуміти без обчислень. Спочатку при множенні на M_{ij} зліва до *j*-го рядка додається *i*-й, помножений на m_{ij}, а далі при множенні на матрицю з елементом $-m_{ij}$ на тому ж місці до одержаного рядка додається *i*-й, помножений на $-m_{ij}$, і в результаті буде одержано *j*-й рядок матриці *A* в його первісному вигляді.
Розглянемо тепер процедуру множення матриці *A* на різні елементарні матриці, які вибираються в такий спосіб. Спочатку, якщо $a_{11} \neq 0$, помножимо *A* зліва на M_{21}, де $m_{21} = -\dfrac{a_{21}}{a_{11}}$. У результаті елемент $b_{21}^{(1)}$ одержаної матриці буде дорівнювати нулю. Зазначимо, що новий елемент на діагоналі в другому рядку і стовпчику буде обчислюватись за формулою
$a_{22}^{(1)} = a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} = \dfrac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}}$. Потім аналогічно підберемо M_{31} і т. ін. (якщо відповідні діагональні елементи не дорівнюватимуть нулю). Таким чином можна зробити так, що під діагоналлю одержаної матриці будуть знаходитися лише нулі. Якщо ж $a_{11} = 0$, то за допомогою множення на деяку матрицю P_{j1} можна переставити рядки так, що на місці (1,1) опиниться ненульовий елемент. (Якщо ж всі елементи першого стовпчика дорівнюють нулю, то процедура переноситься на наступний ненульовий стовпчик.) Тепер те ж саме робимо у випадку $b_{22}^{(1)} = 0$, але переставляємо другий рядок лише з рядками, які розташовані нижче другого (якщо в деякому з них у другому стовпчику є ненульовий елемент) і т. ін. У результаті прийдемо до матриці $B = LA$, елементи якої, що знаходяться нижче діагоналі, нульові (через $\tilde{L}$ позначено добуток усіх елементарних матриць, що відповідали вказаним перетворенням). Така матриця *B* називається верхньою трикутною (верхньотрикутною). Аналогічно при множенні *A* на елементарні матриці справа можна звести її до нижньотрикутного вигляду.

Зведення матриці до верхньотрикутного вигляду еквівалентне відомому методу Гауса, який використовується для розв'язку систем лінійних рівнянь.
Сформулюємо результат.

Теорема 2.3.4. Кожна матриця за допомогою елементарних перетворень може бути зведена до верхньотрикутного (нижньотрикутного) вигляду.

Якщо матриця зведена до верхньотрикутного вигляду, можна, помноживши справа на елементарні матриці, звести її до діагонального вигляду, тобто в одержаній матриці *C* нулі будуть під і над діагоналлю. Треба тільки пояснити, що робити, коли на діагоналі матриці B_k, що одержана з *B*, на *k*-му кроці таких перетворень, опиниться число нуль. Тоді треба переставити стовпчики *k*-й та деякий *l*-й *(l > k),* якщо $b_{kl} \neq 0$, а потім за допомогою $\tilde{b}_{kk} = b_{kl} \neq 0$, що опиниться на діагоналі, в результаті множення зліва на елементарні матриці зробити нулі під діагоналлю. Потім при множенні на відповідні елементарні матриці справа можна зробити нулі правіше діагоналі. Таким чином, прийдемо до діагональної матриці *C,* причому деякі елементи на діагоналі (у випадку виродженої матриці) можуть бути нульовими. Тепер, множачи на відповідні матриці D_{ii}, можна одержати на діагоналі на місці ненульових елементів одиниці. Отже, одержимо діагональну матрицю *D,* на діагоналі якої розташовані лише одиниці і нулі (у випадку виродженої початкової матриці *A).* Кількість одиниць є рангом матриці *A,* оскільки елементарні перетворення не змінюють рангу (це буде доведено в розділі 4.1) і тому не залежить від процедури зведення до діагонального вигляду. Всі одержані таким чином матриці можуть лише відрізнятися порядком розташування одиниць і нулів. Отже,

$$D = LAR, \tag{2.3.40}$$

де *L* і *R* — добутки елементарних матриць.
Сформулюємо результат.

Теорема 2.3.5. Кожна квадратна матриця може бути зведена до діагонального вигляду (2.3.40) за допомогою елементарних перетворень, причому на діагоналі будуть знаходитися лише одиниці і нулі (у випадку виродженості матриці), кількість яких не залежить від способу зведення. Таке зображення єдине з точністю до порядку розташування одиниць і нулів на діагоналі.

Зауваження. З цієї теореми випливає ще одне і досить просте доведення асоціативності множення квадратних матриць: для елементарних матриць асоціативність перевіряється легко, отже, зображуючи кожну матрицю як добуток елементарних, доводимо асоціативність множення для цих трьох матриць.

Інколи діагональну матрицю D з числами $d_{11},...,d_{nn}$ на діагоналі зручно позначати

$$D = diag\left(d_{11},...,d_{nn}\right). \qquad (2.3.41)$$

Зауважимо, що, взагалі кажучи, в зображенні матриці за формулою (2.3.40) L і R не обов'язково трикутні матриці – адже в процесі діагоналізації, можливо, переставлялись рядки. Але, якщо на кожному кроці ведучий діагональний елемент не дорівнював нулю, матрицю A множили зліва тільки на елементарні нижньотрикутні матриці $M_{ji}^{(l)}$ і не множили на P_{ij}. Добуток нижньотрикутних матриць – також нижньотрикутна матриця; позначимо її через $\tilde{L}$. Тоді $\tilde{L}A = R$, і R – верхньотрикутна матриця, звідки, помноживши обидві частини цієї рівності зліва на $L = \tilde{L}^{-1}$, будемо мати зображення матриці A у вигляді добутку нижньотрикутної і верхньотрикутної матриць:

$$A = LR. \qquad (2.3.42)$$

Більш точну умову можливості отримання зображення (2.3.42) читач знайде у главі 4, теорема 4.1.31.

Нехай ранг матриці $A(n \times n)$ дорівнює n. Тоді множенням зліва цієї матриці на елементарні зведемо її до верхньотрикутної з ненульовими елементами на діагоналі. Тоді, як було вказано раніше, виконуючи зворотній хід методу Гауса, тобто, роблячи нулі над діагоналлю, зведемо матрицю до діагональної:

$$D = \tilde{L}A. \qquad (2.3.43)$$

Зворотний хід також відповідає множенню зліва на елементарні матриці. Помножимо обидві частини останньої рівності на обернену до діагональної і одержимо одиничну матрицю:

$$D^{-1}\tilde{L}A = I. \qquad (2.3.43)$$

Отже, таким чином одержимо матрицю, обернену до A:

$$L = D^{-1}\tilde{L} = A^{-1}, LA = I. \qquad (2.3.44)$$

2.4. СПЕКТРАЛЬНИЙ АНАЛІЗ САМОСПРЯЖЕНИХ ОПЕРАТОРІВ

Нагадаємо, що самоспряжений оператор в унітарному просторі має дійсні власні числа, власні вектори, які відповідають різним власним числам, ортогональні між собою. В цьому розділі доведемо, так звану, екстремальну властивість власних чисел.

Спочатку доведемо ще одну формулу для знаходження норми самоспряженого оператора.

Теорема 2.4.1. Справедлива формула:

$$\|A\| = \max_{\|x\|=1} |(Ax, x)| = \max_{\|y\|=1} \|Ay\| \tag{2.4.1}$$

Доведення. Введемо позначення: $\alpha = \max_{\|x\|=1} |(Ax, x)|$. Оскільки $\forall x \in E, \|x\| = 1$ $|(Ax, x)| \le \|Ax\|\|x\| \le \|A\|\|x\|^2 = \|A\|$, то $\alpha \le \|A\|$. Доведемо, що $\|A\| \le \alpha$.

Нехай $a, b \in E$. Обчислимо:

$$\left(A(a+b), a+b\right) - \left(A(a-b), a-b\right) = \left(Aa, a\right) + \left(Ab, b\right) + 2\operatorname{Re}\left(Aa, b\right) -$$
$$-\left(Aa, a\right) - \left(Ab, b\right) + 2\operatorname{Re}\left(Aa, b\right) = 4\operatorname{Re}\left(Aa, b\right).$$

Нехай $z \in E,\ z \ne 0$. Позначимо: $\dfrac{\|Az\|}{\|z\|} = \lambda^2\ (\lambda > 0),\ u = \dfrac{1}{\lambda} Az = \sqrt{\dfrac{\|z\|}{\|Az\|}} Az$. Маємо:

$$\|Az\|^2 = \left(Az, Az\right) = \left(\lambda Az, \frac{1}{\lambda} Az\right) = \left(\lambda Az, u\right) = \frac{1}{4}\Big(A(\lambda z + u), (\lambda z + u) -$$

$$-A(\lambda z - u), (\lambda z - u)\Big) \le \frac{1}{4}\alpha\left(\|\lambda z + u\|^2 + \|\lambda z - u\|^2\right) = \frac{1}{2}\left(\lambda^2\|z\|^2 + \frac{1}{\lambda^2}\|Az\|^2\right) = \alpha\|Az\|\|z\|$$

При $\|z\| = 1$ одержимо: $\|Az\| \le \alpha$, звідки $\|A\| \le \alpha$, що й доводить формулу (2.4.1).

Для подальшого дослідження використаємо доведений факт.

Теорема 2.4.2. Самоспряжений оператор в унітарному просторі має власне число, модуль якого дорівнює нормі оператора:

$$|\lambda| = \|A\|. \tag{2.4.2}$$

В силу компактності одиничної сфери в скінченновимірному просторі існує орт x_0 такий, що $\left|\left(Ax_0, x_0\right)\right| = \|A\|$, тобто, $\lambda \equiv \left(Ax_0, x_0\right) = \|A\|$ або $\lambda \equiv \left(Ax_0, x_0\right) = -\|A\|$.
Зрозуміло, $\|Ax_0\| \le \|A\|$.

Оцінимо:

$\left(Ax_0 - \lambda x_0, Ax_0 - \lambda x_0\right) = \|Ax_0\|^2 + \lambda^2 - 2\lambda\left(Ag_0 g_0\right) = \|Ax_0\|^2 + \lambda^2 - 2\lambda^2 = \|Ax_0\|^2 - \lambda^2 \le$ $\le \|A\|^2 - \|A\|^2 = 0$. Отже, $\|Ax_0 - \lambda x_0\| = 0$ і $Ax_0 - \lambda x_0 = 0$, тобто, або $Ax_0 = \|A\| x_0$, або $Ax_0 = -\|A\| x_0$. Зрозуміло, що величина $\|A\|$ - це найбільше за модулем власне число самоспряженого оператора.

Виділимо власний підпростір, що відповідає цьому власному числу. Його ортогональне доповнення завдяки властивостям самоспряженого оператора інваріантне – теж його інваріантний підпростір, в якому, безумовно, він теж самоспряжений. Розглядаючи оператор на цьому підпросторі таким же чином знайдемо його власне число, яке за модулем буде менше попереднього. Діючи в такий спосіб, знайдемо всі власні числа і власні підпростори самоспряженого оператора. Ці власні підпростори, як відомо, попарно ортогональні.

2.5. ТІЛО КВАТЕРНІОНІВ

Розглянемо лінійний чотиривимірний простір над полем **R**, натягнутий на базисні вектори 1, *i, j, к,* де поки що *i, j, к* — деяк, і символи. Отже, елементами цього простору будуть всілякі лінійні комбінації вигляду
$a + bi + cj + dk$, де $a,b,c,d \in \mathbf{R}$.
Введемо структуру алгебри в цей простір, тобто визначимо операцію множення. Спочатку цю операцію визначимо для базисних векторів. Вважаємо, що *1* комутує з усіма базисними векторами і діє як звичайна одиниця,

$$i^2 = j^2 = k^2 = -1, ij = -ji = k, jk = -kj = i, ki = -ik = j\,,$$

тобто вони антикомутують, і кожен з них має обернений: $i^{-1} = -i$ і т. ін.
Вважаючи, що виконується властивість дистрибутивності, можна визначити множення для довільних лінійних комбінацій базисних векторів, тобто для довільних елементів простору. Читач повинен перевірити, виходячи з даної таблиці множення, асоціативність цієї операції, але для цього потрібно зробити 27 перевірок, тому що в рівності $(ab)c = a(bc)$ на кожному місці може знаходитись один з трьох елементів. Можна зробити інакше — знайти матричне зображення цієї алгебри. Зробимо такі зіставлення:

$$1 \sim I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, i \sim A(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$
$$j \sim A(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, k \sim A(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{2.5.1}$$

Тут *i* — уявна одиниця.
Зауважимо, що матриці $\delta_x = -iA(i)$, $\delta_y = -iA(j)$, $\delta_z = -iA(k)$ у теоретичній фізиці називаються матрицями Паулі.
Лема 2.5.1. Нехай $q_1, q_2 \in \{i, j, k\}$. Тоді $A(q_1 q_2) = A(q_1)A(q_2)$.
Перевіримо справедливість рівності для $q_1 = i$, $q_2 = j$.
Наприклад,

$$A(i)A(j) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} = A(k).$$

Читач може самостійно перевірити інші варіанти.
Тепер кватерніону $q = a + bi + cj + dk$ зівставимо матрицю

$$A(q) = \begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix} \qquad (2.5.2)$$

Спираючись на лему, легко перевірити рівність для кватерніонів загального вигляду (2.5.2).

Доведена рівність означає, що побудовано ізоморфізм тіла кватерніонів і тіла матриць спеціального вигляду. Оскільки для матриць множення асоціативне, асоціативною буде також операція множення і для ква терніонів. Той факт, що всі матриці такого виду, крім 0, мають обернену, випливає з того, що їх визначники дорівнюють $a^2+b^2+c^2+d^2$, але з проблемою оберненості ми познайомимось пізніше.

Існує зв'язок між кватерніонами і векторами в трьохвимірному просторі. Назвемо число a скалярною частиною, а $u \equiv bi + cj + dk$ — векторною частиною кватерніона q. Кватерніон з $a = 0$ називається вектором, тому що він справді є вектором трьохвимірного простору.

Обчислимо добуток в **H** векторів $u_1 = b_1 i + c_1 j + d_1 k$ і $u_2 = b_2 i + c_2 j + d_2 k$:

$$\begin{gathered} u_1 u_2 = b_1 b_2 i^2 + b_1 c_2 ij + b_1 d_2 ik + c_1 b_2 ji + c_1 c_2 j^2 + c_1 d_2 jk + d_1 b_2 ki + d_1 c_2 kj + d_1 d_2 k^2 = \\ -(b_1 b_2 + c_1 c_2 + d_1 d_2) + (c_1 d_2 - d_1 c_2) i + (d_1 b_2 - b_1 d_2) j + (b_1 c_2 - c_1 b_2) k = \\ = -(u_1, u_2) + [u_1, u_2], \end{gathered}$$

де (u_1, u_2) — скалярний добуток; $[u_1, u_2]$ — відомий з аналітичної геометрії векторний добуток. Очевидно,

$$u_2 u_1 = = -(u_2, u_1) + [u_2, u_1] \qquad (2.5.3)$$

(ми скористалися комутативністю скалярного і антикомутативністю векторного добутків, а можна було б знов виконати всі обчислення). Звідси

$$(u_1, u_2) = -1/2(u_1 u_2 + u_2 u_1),$$

$$[u_1, u_2] = 1/2(u_1 u_2 - u_2 u_1).$$

Зробимо таке обчислення, користуючись останньою формулою:

$$[[u_1, u_2], u_3] = 1/2([u_1, u_2] u_3 - u_3 [u_1, u_2]) =$$

$$= 1/4((u_1 u_2 - u_2 u_1) u_3 - u_3 (u_1 u_2 - u_2 u_1)) =$$

$$= 1/4(u_1 u_2 u_3 - u_2 u_1 u_3 - u_3 u_1 u_2 + u_3 u_2 u_1).$$

За допомогою одержаного правила без обчислень знайдемо:

$$[[u_2,u_3],u_1]=1/4(u_2u_3u_1-u_3u_2u_1-u_1u_2u_3+u_1u_3u_2),$$
$$[[u_3,u_1],u_2]=1/4(u_3u_1u_2-u_1u_3u_2-u_2u_3u_1+u_2u_1u_3).$$

Склавши почленно ці рівності, одержимо відому тотожність Якобі з теорії алгебр Лі:

$$[[u_1,u_2],u_3]+[[u_2,u_3],u_1]+[[u_3,u_1],u_2]=0. \qquad (2.5.4)$$

Означення 2.5.1. Лінійний простір, для довільних двох елементів *x,y* якого задана операція множення [*x,y*], що задовольняє аксіомам:
1) [*x,y*] = -[*y,x*];
2) [*x*,*y*],*z*]+[[*y*,*z*],*x*]+[[*z*,*x*],*y*] = 0 (тотожність Якобі) , називається алгеброю Лі.
Операція множення в алгебрі Лі для двох елементів називається ще скобкою Лі цих елементів.
Отже, вектори в трьохвимірному просторі з операцією векторного добутку утворюють алгебру Лі. Операція [.,.] не асоціативна!
З аксіоми 1 випливає, що [*x,x*] = 0, оскільки [*x,x*] = = -[*x,x*], отже 2[*x,x*] = 0 і [*x,x*] = 0. (Зауважимо, що це було б і не так над полем характеристики 2, але **R** має характеристику 0.)
Зазначимо, що операція множення в алгебрі кватерніонів, так би мовити, об'єднує операції скалярного та векторного добутків в $\mathbf{R}^3$.
Для даного кватерніона $q = a + bi + cj + dk$ назвем спряженим до нього кватерніон $q = \alpha\text{-}bi\text{-}\ cj\text{-}\ dk$. Взявши їхній добуток, знайдемо:

$$q\overline{q}=(a+u)(a-u)=a^2+au-au-u^2=a^2+(u,u)-[u,u]=a^2+uuu^2=$$
$$=a^2+b^2+c^2+d^2$$

Очевидно, $q\overline{q}$ = $\overline{q}q > 0$ при $q\neq$ 0. Число $\sqrt{q\overline{q}}$ називається модулем кватерніона і позначається $\|q\|$. Тепер можна знайти кватерніон, обернений для даного.
Знайдемо такий добуток:

$$\left((1/(q\overline{q}))\overline{q}\right)q=q\left((1/(q\overline{q}))\overline{q}\right)=\left(a^2+b^2+c^2+d^2\right)/\left(a^2+b^2+c^2+d^2\right)=1,$$

тобто $1/(q\overline{q})\overline{q}$ є правим і лівим оберненим для *q*.
Зауважимо, що алгебру кватерніонів можна було б будувати інакше. Подібно до того, як, виходячи з дійсних чисел, комплексні будувались як числа вигляду $a + bi$ ($i = \sqrt{-1}$), так і кватерніони можна було б будувати, як числа вигляду $(a + bi) + (c + bi)j$, а далі вводити $k = ij$. Така процедура називається подвоєнням.
Виявляється, що так само можна зробити і з кватерніонами — розглянути елементи вигляду q_1+q_2e, де q_1,q_2 — кватерніони, а *e*- деякий символ. Треба тільки визначити операції множення. В результаті можна одержати алгебру

елементів вигляду $a_0 + \sum_{l=1}^{7} a_l i_l$, яка називається алгеброю октав Келі. Ми не будемо займатися її вивченням. Зазначимо лише, що, як випливає з узагальненої теореми Фробеніуса, є тільки чотири алгебри з діленням: $\mathbb{R}, \mathbb{C}, H, O$. Цей факт має принципове значення в диференціальній геометрії при вивченні векторних полів на сферах.

ГЛАВА 3
ТЕНЗОРИ

3.1. ЗАГАЛЬНІ ВІДОМОСТІ ПРО ТЕНЗОРИ.

Означення 3.1.1. Нехай є прямий добуток скінченого сімейства скінченновимірних просторів: $L = L_1 \times ... \times L_p$ над полем $K = \mathbb{R}$ або $K = \mathbb{C}$.

Розглянемо простір, побудований на векторах з L, елементи якого задовольняють такі властивості:

1). $\left(l_1,...l_{j-1}, l_j^{'} + l_j^{''}, l_{j+1},...,l_p\right) = \left(l_1,...l_{j-1}, l_j^{'}, l_{j+1},...,l_p\right) + \left(l_1,...l_{j-1}, l_j^{''}, l_{j+1},...,l_p\right)$;

2). $\left(l_1,...l_{j-1}, al_j, l_{j+1},...,l_p\right) = a\left(l_1,...l_{j-1}, l_j, l_{j+1},...,l_p\right)$.

Множина таких векторів називається тензорним добутком просторів $L_{1,...,}L_p$ і позначається $L = L_1 \otimes ... \otimes L_p$.

Зауважимо: з 2) випливає, що число a можна писати у будь - якому місці, і результат не зміниться.

Існують глибокі концептуальні підходи для визначення і дослідження тензорних добутків просторів. З'ясуємо лише деякі елементарні властивості тензорного добутку:

1. Якщо один з просторів нульовий, то тензорний добуток нульовий. За властивістю 2) його вектори мали б вигляд:

$$\left(l_1,...l_{j-1}, 0, l_{j+1},...,l_p\right) = 0a\left(l_1,...,l_p\right) = 0$$

2. $L \otimes K \cong L$ (ізоморфізм).

Дійсно , коли $\alpha \in K$, то $l \otimes \alpha = \alpha\left(l \otimes 1\right)$, а $l \otimes 1$ можна ототожнити з l.

Для простоти зараз обмежимось двома просторами. Знайдемо базис в $L_1 \otimes L_2$. Нехай $e = \{e_1,...,e_m\}$, $f = \{f_1,..., f_n\}$ - базиси в L_1 і L_2.

Вектори $x \otimes y$, де $x \in L_1$, $y \in L_2$, породжують весь простір $L_1 \otimes L_2$. Але

$$x=\sum_{j=1}^{m} x_j e_j\,,\; y=\sum_{k=1}^{n} y_k f_k\,.$$

Тоді в силу властивостей 1), 2) маємо:

$$x\otimes y=\left(\sum_k x_k e_k\right)\otimes\left(\sum_j y_j f_j\right)=\sum_k x_k\left(e_k\otimes\left(\sum_j y_j f_j\right)\right)=\sum_{k,j} x_k y_j e_k\otimes f_j$$

тобто $\left\{e_j\otimes f_k\right\}$ – система твірних $L_1\otimes L_2$. Доведемо їхню лінійну незалежність. Нехай $\left\{e^1,...,e^m\right\}\subset L^1;\left\{f^1,...,f^n\right\}\subset L^2$ – дуальні базиси в просторах L_1 і L_2. Функціонал $\left(e^j,f^k\right)$діє на вектор $\left(x,y\right)$ за правилом $\left(e^j,f^k\right)(x,y)=e^j(x)\,f^k(y)$. Нехай є лінійна залежність

$$\sum_{r,s}\alpha_{rs}\left(e_r\otimes f_s\right)=0.$$

Подіявши на обидві частини функціоналом $\left(e^r,f^s\right)$, одержимо що $\alpha_{rs}=0$ при довільних r,s, що і закінчує доведення. Звідси випливає, що $\dim\left(L_1\otimes L_2\right)=\dim\left(L_1\right)\dim\left(L_2\right)$. Аналогічне твердження можна довести для тензорного добутку довільного скінченної кількості просторів $E_1,...,E_n$ (наприклад, за індукцією):

$$\dim\left(L_1\otimes...\otimes L_2\right)=\dim\left(L_1\right)...\dim\left(L_1\right) \tag{3.1.1}$$

Нехай $E_1,...,E_n,M$ – деякі лінійні простори. Відображення $f:E_1\times...\times E_p\to M$ називається полілінійним, якщо воно лінійне по кожному з аргументів при фіксованих інших. Множина таких відображень є лінійний простір, який позначається $L^p\left(E_1,...,E_p;M\right)$. В силу полілінійності тензорного добутку полілінійне відображення $E_1\times...\times E_p\to M$ можна тлумачити як лінійне відображення $E_1\otimes...\otimes E_p\to M$. Отже,

$$L^p\left(E_1,...,E_p;M\right)=L\left(E_1\otimes...\otimes E_p;M\right). \tag{3.1.2}$$

Так скалярний добуток можна вважати лінійною функцією f $f:L_1\otimes L_2\to K$:

$$f\left(x\otimes y\right)=\left(x,y\right).$$

У загальній конструкції тензорного добутку деякі простори можуть бути спряженими лінійними просторами до лінійних просторів. Якщо є p лінійних просторів E_i і q спряжених просторів E^j, то їхній тензорний добуток, природно, позначається $E_1 \otimes \ldots \otimes E_p \otimes E^1 \otimes \ldots \otimes E^q$, і його елементи називаються q разів коваріантними і p разів контраваріантними тензорами.

Розглянемо приклад такого тензорного добутку: $E_1^{'} \otimes E_2$, де $E_1^{'}$ - спряжений до E_1 простір. Виявляється, що його елементами вигляду $x' \otimes y$ можна взаємно однозначно зіставити оператори з простору $L(E_1, E_2)$. Для цього треба показати, як задати таке відображення $\varphi : E_1^{'} \otimes E_2 \to L(E_1, E_2)$, щоб $\forall x' \in E_1^{'},\ y \in E_2\ \varphi(x' \otimes y)$ було б лінійним оператором з E_1 в E_2, причому таке відображення було б лінійним по обох аргументах - x' і y.

Задамо $\varphi(x' \otimes y)$ за допомогою дії на довільний елемент $x \in E_1$ в такий спосіб:

$$\varphi(x' \otimes y) = (x, x')\, y \in E_2$$

Очевидно, це відображення відповідає бажаним властивостям, які випливають з лінійності скалярного добутку та означення лінійного простору (дистрибутивність добутку числа на елемент простору). Взаємна однозначність випливає з того, що ніякий елемент з $E_1^{'} \otimes E_2$ крім нуля не переходить в нуль. Дійсно, коли $\forall x \neq 0\ (x, x')\, y = 0$, то або $y = 0$, або $x' = 0$ (функціонал, що переводить весь простір E_1 в нуль, є нульовий), а $0 \otimes y = x' \otimes 0 = 0$.

Цікаво тепер дати тлумачення матриці оператора $A \in L(E_n, E_m)$. Ми вже ретельно вивчили, як оператору в даному базисі поставити у відповідність матрицю. Нехай тепер $e = \{e_1, ..., e_n\}$ - базис в E_n, $f = \{f_1, ..., f_m\}$ - базис в E_m. В силу знайденого ізоморфізму можна матрицю оператора тлумачити як елемент $E_n^{'} \otimes E_m$, що має базисом $\{e^k \otimes f_j\},\ k = 1, ..., n;\ j = 1, ..., m$. У нього кожен вектор, а значить і матриця, має зображення

$$A = \sum_{k,j} a_k^j e^k \otimes f_j. \tag{3.1.3}$$

Індекси пишуть внизу , якщо вони відповідають векторам з базису спряженого до лінійного простору, вгорі – якщо вони відповідають векторам лінійного простору. Х цих зручних з точки зору тензорної алгебри позначень випливає, що позначення з двома індексами внизу не виражають істинної природи елементів матриці. але в багатьох випадках такий «правильний» запис не обов'язковий . З нової точки зору на простір білінійних відображень $L_2\left(E_1,E_2;M\right)$ ми знаємо, що він ізоморфний $L\left(E_1\otimes E_2;M\right)$. Тому будемо писати

$$L_2(E_1,E_{2;}M)\cong L(E_1\otimes E_{2;}M) \tag{3.1.4}$$

Розглянемо білінійне відображення *f*. Для довільних $x_1\in E_1,\, x_2\in E_2$ $f\left(x_1,x_2\right)=m\in M$. Якщо зафіксувати x_1, одержимо відображення $f_{x_1}:E_2\to M$, яке діє за правилом $f_{x_1}(x_2)=f(x_1,x_2)=m$. Таким чином, кожному $x_1\in E_1$ відповідає оператор $f_{x_1}(x_2)\in L(E_2,M)$, тобто ми одержали відображення E_1 в $L(E_2,M)$ – елемент з простору $L\left(E_1,L(E_2,M)\right)$. Легко перевірити, що це ізоморфізм. Отже,

$$L\left(E_1\otimes E_2;M\right)\cong L_2\left(E_1,E_2;M\right)\cong L\left(E_1,L(E_2,M)\right). \tag{3.1.5}$$

Розберемо цікавий наслідок цієї конструкції. Нехай E_1,E_2 – лінійні простори. Ми вже знаємо, що $L\left(E_2,\mathbb{R}\right)\cong E_2^*$ (простір лінійних функціоналів). Тому

$$L\left(E_1\otimes E_2;\mathbb{R}\right)\cong L_2\left(E_1,E_2;\mathbb{R}\right)\cong L\left(E_1,L(E_2,\mathbb{R})\right)\cong L\left(E_1\otimes E_2^*\right)$$

Нехай тепер $E_1=E_2\equiv E$. Тоді

$$L\left(E\otimes E;\mathbb{R}\right)\cong L_2\left(E,E;\mathbb{R}\right)\cong L\left(E\otimes E^*\right) \tag{3.1.6}$$

Пояснимо це на прикладі скалярного добутку (.,.), лінійної функції на $E\otimes E$ або білінійної на $E\times E$, що ставить у відповідність елементу $(x;y)\in E\otimes E$ за лінійним законом або $(x;y)\in E\times E$ за білінійним – число (x,y). Фіксуючи $y\in E$, можна трактувати $f_y=(.,y)$ як функціонал над Е, тобто елемента з *E**, який діє на *x* $\in$ *E* за правилом : $f_y(x)=(x,y)$. Таким чином, виходячи з скалярного добутку як елемента з $L_2\left(E,E;\mathbb{R}\right)$, одержали лінійне відображення з *E* в *E** $y\to f_y=(.,y)$. З таким міркуванням по суті ми вже зустрічалися, вивчаючи лінійні функціонали.

Означення 3.1.2. Тензорною алгеброю *T(E)* векторного простору Е називається простір, що є прямою сумою таких просторів:

$$T(E) = K \oplus E \oplus E^* \oplus E \otimes E \oplus E \otimes E^* \oplus E^* \otimes E^* ... \qquad (3.1.7)$$

Як бачимо, в кожному прямому доданку повинні бути всілякі комбінації тензорних добутків просторів E і E*. Її підалгеброю є алгебра контраваріантних тензорів, яка позначається *C(E)*:

$$C(E) = C^0(E) \oplus C^1(E) \oplus ... \oplus C^p(E) \oplus ... \qquad (3.1.8)$$

де $C^p(E) = E \otimes ... \otimes E$ *(p* разів), $C^0(E) = K$.

Нехай $x_1 \otimes ... \otimes x_p \in C^p(E)$. В такому просторі діє група підстановок S_p :

$$\sigma = \begin{Bmatrix} 1 & ... & p \\ i_1 & \cdots & i_p \end{Bmatrix},$$

$$\sigma(x_1 \otimes ... \otimes x_p) = x_{i_1} \otimes ... \otimes x_{i_p}$$

Дію σ можна по лінійності продовжити на весь $C^p(E)$. Кожному такому тензору $x \in C^p(E)$ зіставимо тензор

$$S(x) = \frac{1}{p!} \sum_{\sigma \in S_p} \sigma(x). \qquad (3.1.9)$$

Ця операція називається симетризацією. Її можна по лінійності продовжити на все $C^p(E)$. Одержимо підпростір $S^p(E) = S\left(C^p(E)\right)$ простору всіх тензорів, що називається простором симетричних тензорів степеню *p*. Симетрична алгебра простору E – це пряма сума таких просторів:

$$S(E) = K \oplus S^1 \oplus S^2 \oplus E \oplus \dots \qquad (3.1.10)$$

Очевидно, $S^1(E) = E$. Роз'яснимо більш детально , що таке операція симетризації. При $p = 2$: $S(x \otimes y) = \frac{1}{2}(x \otimes y + y \otimes x)$;

при $p = 3$: $S(x \otimes y \otimes z) =$

$$= \frac{1}{3!}\left(x \otimes y \otimes z + y \otimes z \otimes x + z \otimes y \otimes x + x \otimes y \otimes z + y \otimes x \otimes z + z \otimes x \otimes y\right)$$

Доведемо, що $S^2 = S$ (тобто операція симетризації ідемпотентна) і $\mathrm{Im}(S) = S^p(E)$. Результатом симетризації кожного тензора є симетричний тензор, тому що при симетризації кожного доданка з суми $S(x) = \frac{1}{p!}\sum_{\sigma\in S_p}\sigma(x)$ одержимо однакові набори доданків тільки, можливо, в різних порядках. Всього одержимо $p!$ груп по $p!$ різних доданків, в кожній. Звідси, зокрема, випливає: $\mathrm{Im}(S) = S^p(E)$. Навпаки, на симетричних тензорах симетризація є тотожною операцією, так що коли $T \in S^p(E)$, то $T = S(T)$. Це показує, що водночас $\mathrm{Im}(S) = S^p(E)$ і $S^2 = S$.

Ще раз підкреслимо, що симетричні тензори задовольняють умову

$$T = S(T) \tag{3.1.11}$$

3.2. ЗОВНІШНЯ АЛГЕБРА ГРАСМАНА

Назвемо тензор $T \in C^p(E)$ кососиметричним (або антисиметричним), якщо

$$\sigma(T) = \varepsilon(\sigma)T \tag{3.2.1}$$

де $\varepsilon(\sigma)$ – парність перестановки σ, тобто, число з множини $\{-1,1\}$. Очевидно, всі кососиметричні тензори утворюють лінійний підпростір в $C^p(E)$, який будемо позначати через $\Lambda^p(E)$. Тепер розглянемо на просторі $C(E)$ операцію альтернування, яка діє за правилом

$$A(x) = \frac{1}{p!}\sum_{\sigma\in S_p}\varepsilon(\sigma)\sigma(x) \tag{3.2.2}$$

Очевидно, що A: $C^p(E) \to C^p(E)$.

Таким чином, при $p = 2$: $A(x\otimes y) = \frac{1}{2}(x\otimes y - y\otimes x)$

$p = 3$: $A(x \otimes y \otimes z) =$

$$= \frac{1}{3!}\left(x \otimes y \otimes z + y \otimes z \otimes x + z \otimes y \otimes x - x \otimes y \otimes z - y \otimes x \otimes z - z \otimes x \otimes y \right)$$

Лема 3.2.1. $A^2 = A$ і $\operatorname{Im}(A) = \Lambda^p(E)$.

Передусім покажемо, що внаслідок альтернування одержимо кососиметричний тензор. Дійсно, оскільки ε – гомоморфізм (нагадаємо: $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$ і $\varepsilon^2(\sigma) = 1$), маємо для тензора T

$$\sigma(AT) = \sigma((1/q!)\sum_{\tau \in S_q} \varepsilon(\tau)\tau(T)) = \quad (1/q!)\sum_{\tau \in S_q} \varepsilon(\tau)\sigma\tau(T) =$$

$$= \varepsilon(\sigma)(1/q!)\sum_{\tau \in S_q} \varepsilon(\sigma\tau)\sigma\tau(T) = \varepsilon(\sigma)AT.$$

Далі, A є проектором, тому що

$$A^2 = (1/q!)^2 \sum_{\sigma,\tau \in Sq} \varepsilon(\sigma\tau)\sigma\tau = (1/q!)\sum_{\rho \in Sq} \varepsilon(\rho)\rho = A$$

Справді, кожен елемент $\rho \in S_q$ можна $q!$ способами зобразити у вигляді добутку $\sigma\tau$: виберемо σ довільним, а τ знайдемо з рівності $\tau = \sigma^{-1}\rho$. Повторюючи вищенаведене з доказу аналогічного факту про симетричні тензори міркування, одержуємо, що $\operatorname{Im}(A) = \Lambda^p(E)$.

Для елементів $a \in \Lambda^p(E)$, $b \in \Lambda^q(E)$, введемо операцію добутку, яка називається зовнішнім добутком і позначається символом $\wedge$: $a \wedge b$.

За означенням

$$a \wedge b = A\left(a \otimes b\right) \in \Lambda^{p+q}(E). \qquad (3.2.3)$$

Подивимось, як пов'язані між собою елементи $a \wedge b = A\left(a \otimes b\right)$ і $b \wedge a = A\left(b \otimes a\right)$.

Нехай спочатку $a,b \in \Lambda^1(E)=E$. Розкладемо їх за базисом: $a=\sum_i a_i e_i$,

$b=\sum_j b_j e_j$

По-перше,

$$e_i \wedge e_i = \frac{1}{2}(e_i \otimes e_i - e_i \otimes e_i) = 0$$

По-друге,

$$e_i \wedge e_j = \frac{1}{2}(e_i \otimes e_j - e_j \otimes e_i) = -\frac{1}{2}(e_j \otimes e_j - e_i \otimes e_j) = -e_j \wedge e_i$$

Отже, $e_i \wedge e_i = -e_j \wedge e_i$. Тепер

$$\sum_i a_i e_i \wedge \sum_j a_j e_j = \sum a_i^2 e_i \wedge e_i + \sum_{i \neq j} a_i a_j (e_i \wedge e_j + e_j \wedge e_i) = 0,$$

Тобто, $a \wedge a = 0$ (для векторів з $\Lambda^1(E)$).

Доведемо, що $a \wedge b = -b \wedge a$. Якщо $f_1, ..., f_k \in \Lambda^1(E)=E$ і $\sigma = \begin{pmatrix} 1 & ... & p \\ i_1 & ... & i_p \end{pmatrix}$, то

$$. \qquad f_{i_1} \wedge ... \wedge f_{i_k} = \varepsilon(\sigma) f_1 \wedge ... \wedge f_k \qquad (3.2.4)$$

Від $\sigma\tau$ можна перейти до тотожної підстановки $\sigma\tau_0$ після деякого числа транспозицій. При кожній з них знак зовнішнього добутку змінюється на протилежний. Таких змін знаків буде парна чи непарна кількість залежно від парності підстановки, що і доводить твердження. Тепер можна перейти до загальної ситуації. Щоб одержати $b \wedge a$ з $a \wedge b$, треба пересунути в кожному доданку в розкладі a і b за базисом перші q елементів за останні p елементів. У результаті зробимо pq елементарних транспозицій, і парність підстановки буде дорівнювати $(-1)^{pq}$**.** Тому $a \wedge b = (-1)^{pq} b \wedge a$**.**

Зауважимо: якщо dim$E$ = n, то $f_1 \wedge ... \wedge f_n \wedge f_{n+1} = 0$, бо при розкладі цих векторів за базисом в добутку обов'язково будуть рівні співмножники, а в такому разі він дорівнюватиме нулю.

Лема 3.2.2. Операція зовнішнього добутку асоціативна, тобто

$$(T_1 \wedge T)_2 \wedge T_3 = T_1 \wedge (T_2 \wedge T_3).$$

Доведення. За означенням $(T_1 \wedge T_2) \wedge T_3 = A\big(A(T_1 \otimes T_2) \otimes T_3\big)$, а $T_1 \wedge (T_2 \wedge T_3) = A\big(T_1 \otimes A(T_2 \otimes T_3)\big)$.

Доведемо, що для всіх тензорів $L_1 \in C^p(E), L_2 \in C^q(E)$ справджуються формули

$$A\big(A(L_1) \otimes L_2\big) = A\big(L_1 \otimes A(L_2)\big) = A(L_1 \otimes L_2).$$

Дійсно,

$$A(L_1) \otimes L_2 = (1/p!)\sum_{\sigma \in Sp} \varepsilon(\sigma)\sigma(L_1) \otimes L_2,$$

Звідки

$$A\big(A(L_1) \otimes L_2\big) = (1/p!)\sum_{\sigma \in Sp} \varepsilon(\sigma) A(\sigma(L_1) \otimes L_2).$$

Розглянемо вкладення $S_p \to S_{p+q}$, де $\sigma \to$õ, де $\tilde{\sigma}(i) = \sigma(i)$ при $1 \leq i \leq p$ і $\tilde{\sigma}(i) = i$ при і > р. Очевидно $\sigma(L_1) \otimes L_2 = \sigma(L_1 \otimes L_2)$ крім того, А і σ комутують, тому

$$A\big(A(L_1) \otimes L_2\big) = (1/p!)\sum_{\sigma \in Sp} \varepsilon^2(\sigma) A(L_1 \otimes L_2) = A(L_1 \otimes L_2)$$

Аналогічно доводиться друга рівність.

Тепер, підставляючи в першу доведену рівність $L_1 = T_1 \otimes T_2$, одержимо, що $(T_1 \Lambda T_2) \Lambda T_3 = A(A(T_1 \otimes T_2) \otimes T_3) = A(T_1 \otimes T_2 \otimes T_3)$, а потім, підставляючи в другу рівність $L_2 = T_2 \otimes T_3$, будемо мати:

$T_1 \Lambda (T_2 \Lambda T_3) = A(T_1 \otimes A(T_2 \otimes T_3)) = A(T_1 \otimes T_2 \otimes T_3)$, що і завершує доведення.

Означення 3.2.1. Зовнішньою алгеброю $\Lambda(E)$ простору E називається пряма сума

$$\Lambda(E) = \Lambda^0(E) \otimes \Lambda^1(E) \otimes \Lambda^2(E) \otimes ... \otimes \Lambda^n(E) \tag{3.2.5}$$

Елементи кожного з просторів - прямих доданків - називаються однорідними, а алгебра, що є прямою сумою однорідних просторів - градуйованою з операцією зовнішнього множення. Її елементами будуть всілякі лінійні комбінації вигляду:

$$a + \sum_i a_i e_i + \sum_{i<j} a_{ij} e_i \wedge e_j + ... + \sum_{i_1 < i_2 < ... < i_k} a_{i_1 i_2 ... i_k} e_{i_1} \wedge e_{i_2} ... \wedge e_{i_k} + ... + \\ + e_1 \wedge e_2 \wedge ... \wedge e_n \tag{3.2.6}$$

Доведемо, що елементи $e_{i_1} \wedge ... \wedge e_{i_k}$ одного степеню однорідності лінійно незалежні (те, що незалежними будуть елементи різного степеню, очевидно).

Нехай $\sum_{i_1 < i_2 < ... < i_k} a_{i_1 i_2 ... i_k} e_{i_1} \wedge ... \wedge e_{i_k}$. Візьмемо фіксований елемент $e_{i_1} \wedge ... \wedge e_{i_k}$.

Нехай $i_{k+1}, ..., i_n$ — індекси, доповнюючі до $i_1, ..., i_k$, так що разом вони утворюють деяку перестановку чисел 1,2,..., *n*. Очевидно, що кожен з інших елементів суми має хоча б один з доповнюючих індексів. Тому, коли обидві частини домножити на вектор $e_{i_{k+1}} \wedge ... \wedge e_{i_n}$, одержимо $a_{i_1 i_2 ... i_k} e_{i_1} \wedge ... \wedge e_{i_k} \wedge e_{i_{k+1}} \wedge ... \wedge e_{i_n} = 0$, звідки $a_{i_1 i_2 ... i_k} = 0$.

Теорема 3.2.1. Для того щоб вектори $x_1, ..., x_k \in E_n$ були лінійно залежними, необхідно і достатньо, щоб

$$x_1 \wedge \ldots \wedge x_k = 0 . \qquad (3.2.7)$$

Доведення. Якщо вектори лінійно залежні, то який-небудь з них вектор x_k можна вважати лінійною комбінацією інших: $x_k = \sum_{i=1}^{k-1} a_i x_i$. Тоді

$x_1 \wedge \ldots \wedge x_k = x_1 \wedge \ldots x_{k-1} \wedge \left(\sum_{i=1}^{k-1} a_i x_i \right) = \sum_{i=1}^{k-1} a_i (x_1 \wedge \ldots \wedge x_{k-1} \Lambda x_i) = 0$, тому що кожний доданок суми містить пару однакових співмножників.

Доведемо, що, навпаки, з $x_1 \wedge \ldots \wedge x_k = 0$ випливає лінійна залежність векторів. Для цього досить довести протилежне твердження: якщо вектори лінійно незалежні, то $x_1 \wedge \ldots \wedge x_k \neq 0$. В разі їхньої лінійної незалежності їх можна доповнити до базису E_n деякими векторами $x_{k+1}, \ldots, x_n$ Тоді $x_1 \Lambda \ldots x_k \Lambda x_{k+1} \Lambda \ldots \Lambda x_n \neq 0$, бо цей вектор є твірним для $\Lambda^n(E_n)$. Тим більше $x_1 \wedge \ldots \wedge x_k \neq 0$.

Наслідок. Якщо до деякого елемента лінійно незалежної системи додати деяку лінійну комбінацію інших елементів, то одержана таким чином система буде лінійно незалежною.

Доведення. Додамо, наприклад, до елемента x_k лінійну комбінацію $\sum_{i=1}^{k-1} a_i x_i$ і введемо позначення $\tilde{x} = x_k + \sum_{i=1}^{k-1} a_i x_i$. Знайдемо добуток елементів $x_1, \ldots, x_{k-1}, \tilde{x}_k$: $x_1 \wedge \ldots \wedge x_{k-1} \wedge \tilde{x}_k = x_1 \wedge \ldots \wedge x_{k-1} x_k + x_1 \wedge \ldots \wedge x_{k-1} \wedge (\sum_{i=1}^{k-1} a_i x_i)$.

Другий доданок дорівнює нулю, оскільки кожен елемент цієї суми містить два однакових елементи.

Означення 3.2.2. Елементи з $\Lambda^k(E^n)$ називаються *k*-векторами. Вектор вигляду $a_{i_1} \ldots a_{j_k} e_1 \wedge \ldots \wedge e_k$ називається розкладеним.

Лема 3.2.3. (Картана) Нехай вектори $x_1,...,x_k$ лінійно незалежні і існують вектори $y_1,...,y_k$ такі, що

$$x_1 \Lambda y_1 + ... + x_k \Lambda y_k = 0. \qquad (3.2.8)$$

Тоді

$$y_i = \sum_{j=1}^{k} a_{ij} x_j,$$

де $a_{ij} = a_{ji}$.

Доведення. Виберемо вектори $x_{k+1},...,x_n$, які доповнюють дані вектори до базису E. Тоді

$$y_i = \sum_{j=1}^{k} a_{ij} x_j + \sum_{j=k+1}^{n} b_{ij} x_j.$$

Підставляючи це в (3.2.8), одержимо

$$\sum_{i<j\le k} (a_{ij} - a_{ji}) x_i \Lambda x_j + \sum_{i<k,\, j>k} b_{ij} x_i \Lambda x_j = 0.$$

В силу лінійної незалежності $x_i \Lambda x_j$ в $\Lambda^2(E^n)$ $a_{ij} - a_{ji} = 0$ і $b_{ij} = 0$, звідки $a_{ij} = a_{ji}$.

3.3. ДРУГИЙ ПІДХІД ДО ВИЗНАЧЕННЯ ЗОВНІШНЬОЇ АЛГЕБРИ ГРАСМАНА

Нехай E_n — лінійний простір над полем K (характеристика якого не дорівнює 2 — частіше це будуть R або C) і $e = \{e_1,...,e_n\}$ — деякий базис в E_m. Будемо позначати $\Lambda^0(E_n) \equiv K$, $\Lambda^1(E_n) \equiv E_n$.

Нехай для векторів в e визначена операція множення, яка позначається $e_i \Lambda e_j$. Ця операція за означенням вважається антикомутативною, тобто $e_i \Lambda e_j = -e_j \Lambda e_i$, і дистрибутивною по відношенню до додавання:

$e_i\Lambda(e_j+e_k)=e_i\Lambda e_j+e_i\Lambda e_k$. На вектори такого вигляду можна натягти лінійний простір (тобто розглянути всілякі лінійні комбінації цих векторів). Позначимо його $\Lambda^2(E_n)$.

Зразу зауважимо, що $e_i\Lambda e_i=0$, бо $e_i\Lambda e_i=-e_i\Lambda e_i$, звідки $2e_i\Lambda e_i=0$, і за властивостями лінійного простору (оскільки $2\neq 0$) $e_i\Lambda e_i=0$.

Очевидно, його розмірність дорівнює $C_n^2=n(n-1)/2$.

Далі можна розглянути формальні добутки $e_i\Lambda e_j\Lambda e_k$, причому будемо вважати, що операція множення асоціативна:

$$e_i\wedge(e_j\wedge e_k)=(e_i\wedge e_j)\wedge e_k \tag{3.3.1}$$

З властивостей анти комутативності і асоціативності випливає, що

$$e_i\wedge e_j\wedge e_k=\varepsilon(\sigma)e_{\sigma(i)}\wedge e_{\sigma(j)}\wedge e_{\sigma(k)},$$

де $\sigma\in S_3$ — деяка підстановка. Справді, перестановку індексів можна зробити, застосувавши кілька транспозицій, які змінюють парність (i,j,k) в $\varepsilon(\sigma)$ разів.

На такі вектори, як на базисні, натягуємо лінійний простір - $\Lambda^3(E_n)$. Розглядаючи добутки довільної кількості співмножників, одержимо простір — $\Lambda^k(E_n)$. Зауважимо, що у виразі $e_{i_1}\wedge\ldots\wedge e_{i_k}$ немає рівних співмножників, бо їх добуток дорівнював би нулю. Очевидно, $\dim(\Lambda^k(E_n))=C_n^k$.

Таким чином дійдемо до $\Lambda^n(E_m)$ — одномірного простору, натягнутого на єдиний базисний вектор $e_n\wedge\ldots\wedge e_n$. Далі, взявши пряму суму $\Lambda^0(E_n)\oplus\Lambda^1(E_n)\oplus\ldots\oplus\Lambda^k(E_n)\oplus\ldots\oplus\Lambda^n(E_n)$, одержимо алгебру Грасмана $\Lambda(E_n)$ векторного простору E_n:

$$\dim(\Lambda(E^n)) = \sum_{k=0}^{n} \dim(\Lambda^k(E^n) = \sum_{k=0}^{n} C^k = 2^n$$

за властивістю біноміальних коефіцієнтів.

Елементи з $\Lambda^k(E_n)$ називаються однорідними порядку k або k-векторами Грасмана. Так, елементи з $\Lambda^1(E_n)$ будемо називати 1-векторами, з $\Lambda^2(E_n)$ — 2-векторами або бівекторами Грасмана і т.ін.

Такий підхід не потребує ділення на факторіали, тому може бути впроваджений, якщо $char(K) \neq 0$. Для кращого засвоєння викладеного матеріалу розглянемо декілька прикладів. Для простоти будемо в них позначати зовнішній добуток ab, а не $a\Lambda b$.

Приклад 3.3.1. Перемножити два вектори:

$$X_1 = a_{11}e_1 + a_{12}e_2\,,\ X_2 = a_{21}e_1 + a_{22}e_2\,.$$

Помножимо ці вирази за правилами множення многочленів (завдяки дистрибутивності), враховуючи властивості зовнішнього добутку:

$$X_1X_2 = (a_{11}e_1 + a_{12}e_2)(a_{21}e_1 + a_{22}e_2) = a_{11}a_{22}e_1e_2 + a_{12}a_{21}e_2e_1 =$$
$$= \left(a_{11}a_{22} - a_{12}a_{21}\right)e_1e_2$$

Приклад 3.3.2. Помножити три вектори:

$$X_1 = a_{11}e_1 + a_{12}e_2 + a_{13}e_3;\, X_2 = a_{21}e_1 + a_{22}e_2 + a_{23}e_3;\ X_3 = a_{31}e_1 + a_{32}e_2 + a_{33}e_3.$$

Маємо:

$$X_1X_2X_3 = (a_{11}e_1 + a_{12}e_2 + a_{13}e_3)(a_{21}e_1 + a_{22}e_2 + a_{23}e_3)(a_{31}e_1 +$$
$$+a_{32}e_2 + a_{33}e_3) = (\underline{a_{11}a_{21}e_1e_1} + a_{11}a_{22}e_1e_2 + a_{11}a_{23}e_1e_3 +$$
$$+a_{12}a_{21}e_2e_1 + \underline{a_{12}a_{22}e_2e_2} + a_{12}a_{23}e_2e_3 + a_{13}a_{21}e_3e_1 +$$
$$+a_{13}a_{22}e_3e_2 + \underline{a_{13}a_{23}e_2e_3})(a_{31}e_1 + a_{32}e_2 + a_{33}e_3) =$$
$$= \underline{a_{11}a_{22}a_{31}e_1e_2e_1} + \underline{a_{11}a_{22}a_{32}e_1e_2e_2} + a_{11}a_{22}a_{33}e_1e_2e_3 +$$
$$+\underline{a_{11}a_{23}a_{31}e_1e_3e_1} + a_{11}a_{23}a_{32}e_1e_3e_2 + a_{11}a_{23}a_{33}e_1e_3e_3 +$$
$$a_{12}a_{21}a_{31}e_2e_1e_1 + a_{12}a_{21}a_{32}e_2e_1e_2 + a_{12}a_{21}a_{33}e_2e_1e_3 +$$
$$+a_{12}a_{23}a_{31}e_2e_3e_1 + \underline{a_{12}a_{23}a_{32}e_2e_3e_2} + \underline{a_{12}a_{23}a_{33}e_2e_3e_3} +$$

$$+\underline{a_{13}a_{21}a_{31}e_3e_1e_1} + a_{13}a_{21}a_{32}e_3e_1e_2 + \underline{a_{13}a_{21}a_{33}e_3e_1e_3} +$$

$$+a_{13}a_{22}a_{31}e_3e_2e_1 + a_{13}a_{22}a_{32}e_3e_2e_2 + a_{13}a_{22}a_{33}e_3e_2e_3 .$$

Усі підкреслені члени містять по два однакових співмножники, тому вони перетворюються на нуль. У кожному з решти доданків упорядкуємо добутки базисних векторів, тобто зведемо до вигляду $e_1e_2e_3$. Враховуючи рівності

$$e_1e_3e_2 = -e_1e_2e_3 ,$$

$$e_2e_1e_3 = -e_1e_2e_3 ,$$

$$e_2e_3e_1 = -e_2e_1e_3 = e_1e_2e_3 ,$$

$$e_3e_1e_2 = -e_1e_3e_2 = e_1e_2e_3 ,$$

$$e_3e_2e_1 = -e_1e_2e_3 ,$$

після зведення подібних членів дістаємо

$$X_1X_2X_3 =$$

$$= (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{31}a_{22}a_{13})e_1e_2e_3 .$$

3.4. ЗОВНІШНІ ФОРМИ

Тепер нагадаємо проблему, яка обговорювалась при вивченні лінійного простору. Як знайти коефіцієнти розкладу по базису? Відповідь — треба мати спряжений простір. Таке ж саме питання постає і зараз. Для кожного з однорідних просторів $\Lambda^p(E_n)$ алгебри Грасмана треба визначити спряжений простір $(\Lambda^p(E_n))^*$, тобто лінійний простір функціоналів ω, які переводять елементи вигляду $\sum a_{i_1\ldots i_k} e_{i_1}\Lambda\ldots\Lambda e_{i_k}$ в елементи поля K (дійсні або комплексні числа):

$$\omega\left(\sum a_{i_1\ldots i_k} e_{i_1}\Lambda\ldots\Lambda e_{i_k}\right) \in K .$$

Очевидно, для цього доцільно взяти p-вектори з $\Lambda^p((E_n)^*)$, які називаються p-ковекторами, що можна одержати з $e^{i_1},...,e^{i_k} \in (E_n)^*$ за допомогою операції альтернування. Всілякі лінійні комбінації одержаних векторів і дадуть $\Lambda^p((E_n)^*)$

Не будемо втомлювати читача простими, але кропіткими перевірками і зразу наведемо результат: $(\Lambda^p(E_n))^* = \Lambda^p((E_n)^*)$. З іншого боку, можна, як і у випадку тензорних добутків, розглянути простір кососиметричних p-лінійних форм $L_p^a(E,...,E;K)$ таких, що для довільного його елемента ω справджується рівність

$$\omega(x_{i_1}, x_{i_2},..., x_{i_k}) = \varepsilon(\sigma)\omega(x_1,...,x_k),$$

де $\varepsilon(\sigma)$ — парність підстановки

$$\sigma = \begin{bmatrix} 1 & ... & k \\ i_1 & ... & i_k \end{bmatrix}$$

Кососиметричні p-лінійні форми можна трактувати як лінійні функціонали на алгебрі Грасмана $\Lambda^p(E_n)$. Виявляється, що справедливий аналогічний результат, як і у випадку полілінійних відображень:

$$(\Lambda^p(E_n))^* = \Lambda^p((E_n)^*) = L_p^a(E,...,E;K) \tag{3.4.1}$$

Щоб зрозуміти останню рівність, треба показати, як елемент $f^1\Lambda...\Lambda f^p \in \Lambda^p((E_n)^*)$ діє на вектори з $E \times E \times ... \times E$ і пояснити, чому він буде елементом з $L_p^a(E,...,E)$. Нехай, як і завжди,

$$\sigma = \begin{bmatrix} 1 & \cdots & p \\ i_1 & \cdots & i_p \end{bmatrix} \in S_p,$$

$$(f^1\Lambda...\Lambda f^p)(e_1,...,e_p) = A(f^1 \otimes ... \otimes f^p)(e_1,...,e_p) =$$

$$=(1/p!)\sum_{\sigma\in S_p}\varepsilon(\sigma)\,f^{\sigma(1)}(e_1)...f^{\sigma(p)}(e_p) \qquad (3.4.2)$$

При зміні порядку векторів $e_1,...,e_p$ треба перевірити, чи діє $f^1\Lambda...\Lambda f^p$ як кососиметрична форма, тобто, чи відрізняються на $\varepsilon(\sigma)$ знаки $(f^1\Lambda...\Lambda f^p)(e_1,...,e_p)$ і $(f_1\Lambda...\Lambda f_p)(e_{\sigma(1)},...,e_{\sigma(p)})$.

Обчислимо $(f^1\Lambda...\Lambda f^p)(e_{\sigma(1)},...,e_{\sigma(p)})=1/p!\sum_{\tau\in S_p}\varepsilon(\tau)\,f^{\tau(1)}(e_{\sigma(1)})...f^{\tau(p)}(e_{\sigma(p)})$ (підстановки τ перемішують індекси у векторів f^i згідно з означенням альтернування, а не у e_i). Але якщо підстановки τ пробігають всю групу S_p, то її будуть також пробігати і підстановки $\tau\sigma$. Тому останню суму можна подати у вигляді

$$(1/p!)\sum_{\tau\in S_p}\varepsilon(\tau\sigma)\,f^{\tau\sigma(1)}(e_{\sigma(1)})...f^{\tau\sigma(p)}(e_{\sigma(p)})$$

Тепер співмножники у доданках цієї суми можна переставляти як завгодно, бо це комутуючі дійсні числа, і на перше місце поставити $f^1(e_1)$, на друге - $f^2(e_2)$ та ін. Тоді, користуючись властивістю парності $\varepsilon(\tau\sigma)=\varepsilon(\tau)\varepsilon(\sigma)$, перепишемо останню суму у вигляді

$$\varepsilon(\sigma)(1/p!)\sum_{\tau\in S_p}\varepsilon(\tau)\,f^{\tau(1)}(e_1)...f^{\tau(p)}(e_p),$$

яка, судячи з попереднього, збігається з $\varepsilon(\sigma)(f^1\Lambda...\Lambda f^p)(e_1,...,e_p)$.

Таким чином, доведено, що $\Lambda^p((E_n)^*)\subset L_p^a(E,...,E,K)$ і це є вкладення, тобто тільки нульовий p-ковектор переходить в нульову форму. Ці простори насправді збігаються, бо якщо вибрати базис в $E_n\supset\{e_1,...,e_n\}$, то кожна кососиметрична p-лінійна форма ω на E_n однозначно обумовлюється своїми значеннями $\omega(e_{i_1},...,e_{i_p})$ при $1\le i_1<i_2<...<i_p\le n$, і їх можна вибрати довільно.

Тому розмірність простору таких форм дорівнює C_n^p. Зауважимо, що з формули (3.4.2) випливає.

$$(e^1 \Lambda ... \Lambda e^p)(e_{i_1} \Lambda ... \Lambda e_{i_p}) = \frac{1}{p!} \delta_{i,i_1} ... \delta_{p,i_p} \tag{3.4.1}$$

З цього видно, що при перестановці одномірних елементів f^i і f^j знак зовнішнього добутку зміниться: $f^i \Lambda f^j = -f^j \Lambda f^i$. Тепер, коли $\alpha \in \Lambda^p(E_n^*), \beta \in \Lambda^q(E_n^*)$, то $\alpha \wedge \beta = (-1)^{p+q} \beta \wedge \alpha$, оскільки довільну форму можна розкласти за базисом, а в кожній елементарній формі треба переставити кожен з p базисних векторів з q базисними векторами. В результаті буде зроблено p+q транспозицій, і парність підстановки індексів векторів зміниться на $(-1)^{p+q}$.

З формули (3.4.3) випливає, що в просторі p-форм є базис з елементів вигляду $e^{i_1} \wedge ... \wedge e^{i_p}$ і для $\omega \in \Lambda(E^{n*})$

$$\omega = \sum_{i_1 < ... < i_p} a_{i_1 ... i_p} e^{i_1} \wedge ... \wedge e^{i_n} \tag{3.4.1}$$

Наведемо приклади:

1) $p = 1, \omega \in \Lambda^1(E_n^*) = E_n^*$ — лінійний функціонал або ковектор:

$$\omega = \sum_{i=1}^{n} a_i e^i.$$

Якщо e^j — дуальний базис до базису $\{e_1, ..., e_n\}$, тобто $e^k(e_j) = \delta_{kj}$, то при $x = \sum x^k e_k$

$$\omega(x) = \left(\sum a_i e^j\right)\left(\sum x^k e_k\right) = \sum a_i x^i .$$

Така ситуація вже зустрічалась.

2) $p=2, n=2, \omega=e^1\Lambda e^2 \in \Lambda^2(E_2^{\,*})=L_2^a(E,E;K)$ — два ковектори, $\{e_1,e_2\}\subset E_2$ — базис, причому $e^i(e_j)=\delta_{ij}$.

Нехай $x_1=a_{11}e_1+a_{12}e_2, x_2=a_{21}e_1+a_{22}e_2$. Щоб використати (3.4.2), зауважимо, що S_2 має дві підстановки — тотожну σ_0 і транспозицію σ_1, $\varepsilon(\sigma_0)=1, \varepsilon(\sigma_1)=-1$. Отже, маємо:

$$(e^1\Lambda e^2)(x_1,x_2)=(1/2!)(\varepsilon(\sigma_0)e^1(x_1)e^2(x_2)+$$

$$\varepsilon(\sigma_1)e^1(x_2)e^2(x_1))=(1/2)(a_{11}a_{22}-a_{21}a_{12})\equiv(1/2)\det A$$

де A - матриця коефіцієнтів розкладу векторів x_1,x_2 по базису, а вираз в останніх дужках (дійсне число) називається визначником або детермінантом матриці A.

3.5. ВНУТРІШНІЙ ДОБУТОК ЗОВНІШНІХ ФОРМ

Зафіксуємо деякий елемент $x \in E_n$. Відображення

$$i(x): (\Lambda^p(E_n)^*) \to (\Lambda^{p-1}(E_n))^*,$$

яке визначається для довільної $\alpha \in (\Lambda^p(E_n))^*$ за формулою

$$(i(x)\alpha)(e_1,...,e_{p-1}) = \alpha(x, e_1,...,e_{p-1}) \in (\Lambda^p(E_n))^*, \qquad (3.5.1)$$

називається внутрішнім добутком форми α на елемент $x \in E_n$.

Приклад. Нехай $e \in E_n$, $e^* \in (E_n)^*$, $e^*(e) = 1$. Тоді $i(e)(e^*) \in K$,

$i(e)\left(e^*\right) = e^*(e) = 1$.

З означення легко одержати такі властивості внутрішнього добутку:

1) $i(x+y) = i(x) + i(y)$;

2) $i(ax) = ai(x), a \in K$;

3) $i(x)i(y) = -i(y)i(x)$;

4) $i(x)i(x) = 0$.

Доведення властивостей:

1) $(i(x+y)\alpha)(e_1,...,e_{p-1}) = \alpha(x+y, e_1,...,e_{p-1}) = \alpha(x, e_1,...,e_{p-1}) +$

$+\alpha(y, e_1,...,e_{p-1}) = (i(x)\alpha + i(y)\alpha)(e_1,...,e_{p-1})$;

2) $i(ax)(e_1,...,e_{p-1}) = \alpha(ax, e_1,...,e_{p-1}) = a\alpha(x, e_1,...,e_{p-1})$;

3) $(i(x)i(y)\alpha)(e_1,...,e_{p-2}) = (i(x)\alpha)(y, e_1,...,e_{p-2}) = \alpha(x, e_1,...,e_{p-2}) =$

$$= -\alpha(y, x, e_1,...,e_{p-2}) = -(i(y)i(x)\alpha)(e_1,...,e_{p-2})$$

(мінус з'явився завдяки властивості кососиметричності α);

4) з попередньої властивості випливає, що $i(x)i(x) = -i(x)i(x)$, звідки $i(x)i(x) = 0$;

5) для $\alpha \in (\Lambda^p(E_n))^*$, $\beta \in (\Lambda^q(E_n))^*$, $x \in E_n$ $i(x)(\alpha \Lambda \beta) =$

$$= (i(x)\alpha)\Lambda\beta + (-1)^p \alpha \Lambda (i(x)\beta).$$

Для зручності позначимо $x = e$.

При доведенні будемо користуватися формулою (3.4.2). Очевидно, всі підстановки розбиваються на об'єднання множин S' і S'' підстановок σ таких, що $\sigma(1) \le p$, і таких, що $\sigma(1) > p$.

За означенням

$$i(x)(\alpha \wedge \beta)(e_2,...,e_{p+q}) = (\alpha \wedge \beta)(e_1,...,e_{p+q}) =$$

$$= (1/\, p!q!) \sum_{s \in S_{p+q}} \varepsilon(s)\alpha(e_{s(1)},...,e_{s(p)})\beta(e_{s(p+1)},...,e_{p+q}).$$

Тепер цю суму можна розбити на дві і підсумувати за множинами $s \in S'$ і $s \in S''$. При підсумуванні за S' матимемо p варіантів — елемент $e_1 = x$ буде знаходитися на p різних місцях. Щоб одержати його на першому місці (для того щоб прийти до виразу $i(x)\alpha \wedge \beta$, треба зробити p транспозицій. Те ж саме можна сказати про сумування по S'' (з заміною p на q). Зауважимо, що для перестановки елемента x з першими p елементами треба зробити p транспозицій, парність яких дорівнює $(-1)^p$.

Додамо також, що $i(x)(\alpha\Lambda\beta)$ — елемент $\Lambda^{p+q-1}(s)$, і для визначення таких форм треба застосовувати підстановки з групи S_{p+q-1}.

Позначимо через t_s транспозицію з S_{p+q-1}, що міняє місцями 1 з $S(1)$ при $s(1) \le p$, і $t_s \in s_{p+q-1}$, що міняє $p + 1$ і $s(1)$ при $s(1) > p$.

Тоді

$$(1/\,(p!q!)) \sum_{s \in S} \varepsilon(S)\alpha(e_{s(1)},...,e_{s(p)})\beta(e_{s(p+1)},...,e_{s(p+p)}) =$$

$$= (1/\,(p!q!)) \sum_{s \in S} \varepsilon(t_s s)\alpha(x, e_{st_s(2)},...,e_{st_s(p)})\beta(e_{st_s(p+1)},...,e_{st_s(p+q)}) =$$

$$= (p\,/\,(p!q!)) \sum_{r \in S_{p+q-1}} \varepsilon(r)\alpha(x, e_{r(2)},...,e_{r(p)})\beta(e_{r(p+1)},...,e_{s(p+q)}) \equiv$$

$$\equiv (i(x)\alpha \wedge \beta)(e_2,...,e_{p+q})$$

Розглянемо аналогічну процедуру при $s \in S''$. Нехай підстановка σ переміщує елемент e_1 на місце: $p+1$: $\sigma(1)=p+1$, а на інших номерах вона діє таким чином:

$\sigma(i)=i-1$ при $r \le i \le p+1$,

$\sigma(i)=i$ при $r > p+1$

$\varepsilon(\sigma)=(-1)^p$.

Тоді

$$(1/(p!q!))\sum_{s\in S''}\varepsilon(S)\alpha(e_{s(1)},...,e_{s(p)})\beta(e_{s(p+1)},...,e_{s(p+q)})=$$

$$=(1/(p!q!))\sum_{s\in S''}\varepsilon(t_s{}'s)\alpha(e_{st'_s(2)},...,e_{st'_s(p)})\beta(x,e_{st'_s(p+1)},...,e_{st'_s(p+q)})=$$

$$=(1)^p/(p!q!)\sum_{s\in S''}\varepsilon(\sigma t'_s s)\alpha(e_{st_{s'}(2)},...,e_{st_{s'}\sigma(p+1)})\beta(x,e_{st'_s\sigma(p+2)},...,e_{s(p+q)})=$$

$$=(-1)^p(q/p!q!))\sum_{r\in S_{p+q-1}}\varepsilon(r)\alpha(e_{r(2)},...,e_{r(p)})\beta(x,e_{r(p+1)},...,e_{r(p+q)})=$$

$$=(-1)^p(\alpha\wedge(i(x)\beta))(e_2,...,e_{p+q}).$$

Наслідок. Нехай $x^1,...,x^p \in \Lambda^1(E_n)$, $e \in E_n$. Тоді

$$i(e)(x^1\wedge x^2\wedge...\wedge x^p)=$$

$$(i(e)x^1)\wedge x^2\wedge...\wedge x^p - x^1\wedge(i(e)x^2)\wedge e^3\wedge...\wedge x^p+...+ \tag{3.5.2}$$

$$+(-1)^p x^1\wedge x^2\wedge...\wedge(i(e)x^p).$$

ГЛАВА 4
ТЕОРІЯ ВИЗНАЧНИКІВ ТА ЇЇ ЗАСТОСУВАННЯ ДО ТЕОРІЇ ЛІНІЙНИХ СИСТЕМ АЛГЕБРАЇЧНИХ РІВНЯНЬ

4.1 ЗАГАЛЬНА ТЕОРІЯ ВИЗНАЧНИКІВ

Зіставимо матриці *A*:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \qquad 4.1.1)$$

вектори рядки

$$\begin{aligned} x_1 &= a_{11}e_1 + a_{12}e_2 + \dots + a_{1n}e_n, \\ x_2 &= a_{21}e_1 + a_{22}e_2 + \dots + a_{2n}e_n, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_n &= a_{n1}e_1 + a_{n2}e_2 + \dots + a_{nn}e_n, \end{aligned} \qquad (4.1.2)$$

Де $e = \{e_{1,\dots,}e_n\}$ базис в просторі E_n (або в підпросторі грасманової алгебри $\Lambda^p(E_n)$)

Обчислимо добуток $x_1\Lambda\dots\Lambda x_n$. Зауважимо, що ненульовими можуть бути тільки доданки вигляду $a_{1i_1}a_{2i_2}\dots a_{ni_n}e_{i_1}\Lambda e_{i_2}\Lambda\dots\Lambda e_{i_n}$ з попарно нерівними індексами i_k. Кожен з них можна, користуючись правилами множення в алгебрі Грасмана, привести до вигляду

$$\varepsilon(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)}e_1\Lambda e_2\Lambda\dots\Lambda e_n,$$

де

$$\sigma = \begin{bmatrix} 1 & \dots & n \\ i_1 & \dots & i_n \end{bmatrix} \in S_n$$

і $\varepsilon(\sigma)$ - парність цієї підстановки.

Підкреслимо: наша задача полягає в тому, щоб добуток базисних елементів був правильним, тобто співмножники були розташовані в порядку зростання індексів. Отже,

$$x_1 \wedge x_2 \wedge \dots \wedge x_n = \left(\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1_\sigma(1)} a_{2_\sigma(2)} \dots a_{n_\sigma(n)} \right) e_1 \wedge e_2 \wedge \dots \wedge e_n \qquad (4.1.3)$$

Означення. Число у дужках формули, коефіцієнт перед $e_1 \Lambda e_2 \Lambda \dots \Lambda e_n$, називається визначником або детермінантом матриці, і для нього вживається одне з позначень:

$$\Delta(A) = \det(A) = \begin{vmatrix} a_{11} \ a_{12} \dots \ a_{1\mathrm{n}} \\ \dots\dots\dots\dots\dots \\ a_{n1} \ a_{n2} \ \dots \ a_{nn} \end{vmatrix} = \left(\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1_\sigma(1)} a_{2_\sigma(2)} \dots a_{n_\sigma(n)} \right) \qquad (4.1.4)$$

З методу обчислення випливає, що в сумі (4.1.3) $n!$ доданків, бо кожен з них відповідає певній перестановці інших індексів – номерів стовпчиків, у яких розташовані елементи матриці, а перші індекси, які відповідають номерам рядків, знаходяться в порядку зростання, тобто, всього доданків стільки, скільки існує перестановок з n елементів - $n!$.

Зауважимо, що визначник можна визначити тільки для квадратної матриці.

З визначником квадратної матриці другого порядку ми вже зустрічались. Обчислимо його ще раз, користуючись правилами зовнішнього множення. З гідно з означенням для цього слід перемножити вектори $x_1 = a_{11}e_1 + a_{12}e_2$ і

$x_2 = a_{21}e_1 + a_{22}e_2$:

$x_1 \wedge x_2 = (a_{11}e_1 + a_{12}e_2) \wedge (a_{21}e_1 + a_{22}e_2) = =$

$a_{11}a_{21}e_1 \wedge e_2 + a_{11}a_{22}e_1 \wedge e_2 + a_{12}a_{21}e_2 \wedge e_1 + a_{12}a_{22}e_2 \wedge e_1 = (a_{11}a_{22} - a_{21}a_{12}) e_1 \wedge e_2$.

Визначник – це числовий коефіцієнт перед правильним добутком $e_1 \wedge e_2$, тобто, $\Delta = a_{11}a_{22} - a_{21}a_{12}$.

Як ми вже знаємо, якщо $e^1, e^2, ..., e^n$ - вектори дуального базису, то $p!\left(e^1 \wedge ... e^n\right) \wedge \left(x_1 \wedge ... \wedge x_n\right) = \det(A)\left(e^1 \wedge ... e^n\right) \wedge \left(e_1 \wedge ... \wedge e_n\right) = \det(A)$.

Аналізуючи вираз (4.1.3), можна прийти до стандартного означення визначника: це сума всіх добутків елементів матриці таких, що в кожному добутку присутній тільки один елемент з кожного рядка і один елемент кожного стовпчика матриці, причому, знак такого добутку – це знак парності підстановки, які утворюють перші і другі індекси елементів матриці.

Зазначимо, що в формулі (4.1.3) перші індекси, які відповідають номерам рядків, розташовані в правильну перестановку, тобто, в порядку зростання, тому їх доцільно записувати у верхньому рядку підстановки, а в нижньому – другі індекси, відповідні номерам стовпчиків матриці.

Можна помітити, що означення визначника в формі (4.1.3) досить штучне – матриці зіставляли вектори-рядки, а потім брали їхній зовнішній добуток. Виникає питання: що буде, якщо те саме зробити з векторами-стовпчиками. Виявляється, одержимо той самий результат. Доведемо це. Нехай

$$\begin{aligned}
&y_1 = a_{11}e_1 + a_{21}e_2 + ... + a_{n1}e_n,\\
&y_2 = a_{12}e_1 + a_{22}e_2 + ... + a_{n2}e_n,\\
&...\\
&y_n = a_{1n}e_1 + a_{2n}e_2 + ... + a_{nn}e_n,
\end{aligned}$$

Обчислимо значення зовнішнього добутку:

$$\begin{aligned}
y_1 \wedge ... \wedge y_n &= \left(a_{11}e_1 + a_{21}e_2 + ... + a_{n1}e_n\right) \wedge \left(a_{12}e_1 + a_{22}e_2 + ... + a_{n2}e_n\right) \wedge \\
&\wedge ... \wedge \left(a_{11}e_1 + a_{21}e_2 + ... + a_{n1}e_n\right) = \sum \left(a_{j_1 1}a_{j_2 2}...a_{j_n n}\right) e_{j_1} \wedge e_{j_2} \wedge ... \wedge e_{j_n} = \\
&= \sum_{\tau} \varepsilon(\tau)\left(a_{\tau(1)1}a_{\tau(2)2}...a_{\tau(n)n}\right) e_1 \wedge e_2 \wedge ... \wedge e_n
\end{aligned}$$

Треба тепер відзначити, що коли $\tau = \begin{bmatrix} j_1 & ... & j_n \\ 1 & ... & n \end{bmatrix} \in S_n$,

$j_1 = \tau(1), j_2 = \tau(2), ..., j_n = \tau(n)$, то, навпаки, $1 = \tau^{-1}(j_1), 2 = \tau^{-1}(j_2), ..., n = \tau^{-1}(j_n)$, і оскільки $\varepsilon(\tau^{-1}) = \varepsilon(\tau)$ (дійсно, $1 = \varepsilon(e) = \varepsilon(\tau^{-1}\tau) = \varepsilon(\tau^{-1})\varepsilon(\tau)$, звідки $\varepsilon(\tau^{-1}) = \left(\varepsilon(\tau)\right)^{-1} = \varepsilon(\tau)$), останній вираз можна записати у вигляді

$$\sum_{\tau^{-1} \in S_n} \varepsilon(\tau^{-1})\left(a_{1\tau^{-1}(1)}a_{2\tau^{-1}(2)}...a_{n\tau^{-1}(n)}\right) e_1 \wedge e_2 \wedge ... \wedge e_n,$$

тому що як τ, так і τ^{-1} однаково пробігають всю групу S_n. Отже, приходимо до виразу (4.1.3) після перепозначення $\sigma = \tau^{-1}$, тобто, обидва підходи дають один і той же вираз для визначника матриці. Будемо називати перший з них правилом рядків, а другий – стовпчиком.

Висновок. Маємо: $\det(A^T) = \det(A)$, де A^T - матриця транспонована до *A*. Справді, обчислення $\det(A^T)$ за правилом стовпчиків збігається з обчисленням $\det(A)$ за правилом рядків.

Покажемо тепер, як обчислити визначник за допомогою зовнішніх форм.

Нехай $\omega = e^1 \wedge e^2 \wedge ... \wedge e^n$, де $\left\{e^1, e^2, ..., e^n\right\}$ - дуальний до $\left\{e_1, e_2, ..., e_n\right\}$ базис, тобто, $e^i(e_j) = \delta_{ij}$. Нагадаємо, що з властивості лінійності зовнішньої форми $e^i\left(a_{j1}e_1 + a_{j2}e_2 + ... + a_{jn}e_n\right) = a_{ji}$. Тоді за означенням

$$n!\omega(x_1, x_2, ..., x_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) e^{\sigma(1)}(x_1)...e^{\sigma(n)}(x_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)}...a_{n\sigma(n)},$$

тобто, $\det(A)$.

Сформулюємо результат:

$$\det(A) = n!\omega(x_1, x_2, ..., x_n) = \begin{vmatrix} e^1(x_1) & ... & e^1(x_n) \\ ... & ... & ... \\ e^n(x_1) & ... & e^n(x_n) \end{vmatrix} \tag{4.1.5}$$

Ця формула дає змогу краще запам'ятати, як діє зовнішній добуток форм з $\left(\Lambda^1\left(E_n\right)\right)^*$, тобто, форма з $\left(\Lambda^n\left(E_n\right)\right)^*$, на вектори $x_1, ..., x_n$.

Властивості визначників.

Теорема 4.1.1. Визначник діагональної матриці дорівнює добутку діагональних елементів.

Доведення. Дійсно, нехай

$$A = \begin{vmatrix} a_{11} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & a_{nn} \end{vmatrix} \tag{4.1.6}$$

Зіставляючи цій матриці вектори-рядки $x_1 = a_{11}e_1, ..., x_n = a_{nn}e_n$, одержуємо: $x_1 \wedge ... \wedge x_n = a_{11}...a_{nn} e_1 \wedge ... \wedge e_n$. Це й доводить твердження.

Цей результат можна узагальнити. Нехай матриця *A* має вигляд

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \qquad (4.1.7)$$

Де - A_1 - матриця размірності $m \times m$, а A_2 - матриця размірності $n \times n$. Така матриця називається блочно-діагональною.

Теорема 4.1.2. Визначник блочно-діагональної матриці дорівнює добутку визначників матриць, що розташовані на її діагоналі.
Доведення. Зіставимо матриці *A* вектори-рядки

$$x_1 = a_{11}e_1 + a_{12}e_2 + ... + a_{1m}e_m,$$
$$...$$
$$x_m = a_{m1}e_1 + a_{m2}e_2 + ... + a_{mm}e_m,$$

$$x_{m+1} = a_{m+1,m+1}e_{m+1} + a_{m+1,m+2}e_{m+2} + ... + a_{m+1,m+n}e_{m+n},$$
$$...$$
$$x_{m+n} = a_{m+n,m+1}e_{m+1} + a_{m+n,m+2}e_{m+2} + ... + a_{m+n,m+n}e_{m+n},$$

Враховуючи те, що множини векторів $\{e_1, e_2, ..., e_m\}$ і $\{e_{m+1}, e_{m+2}, ..., e_{m+n}\}$ не перетинаються, після зовнішнього множення отримуємо потрібний результат: $\det(A) = \det(A_1)\det(A_2)$.

Тепер дамо інтерпретацію теореми 3.2.1 мовою теорії визначників.
Нехай $x_1, x_2, ..., x_n$ - вектори-рядки матриці *A*: $x_i = \sum_j a_{ij}e_j$. Теорема стверджує, що вони лінійно залежні тоді і тільки тоді, коли $x_1 \wedge ... \wedge x_n = 0$, але цей зовнішній добуток дорівнює $\det(A)e_1 \wedge ... \wedge e_n$, звідки випливає, що $\det(A) = 0$. Те саме стосується векторів-стовпчиків. Сформулюємо результат.
Теорема 4.1.3. Визначник матриці дорівнює нулю тоді і тільки тоді, коли вектори-рядки (або вектори-стовпчики) лінійно залежні.

Звідси з очевидністю випливає
Теорема 4.1.4. Вектори-рядки матриці або її вектори-стовпчики лінійно незалежні тоді і тільки тоді, коли її визначник не дорівнює нулю.

Означення. Матриця називається невиродженою, якщо вона переводить лінійно незалежні вектори в лінійно незалежні.
В такому разі вона переводить вектори базису знову у вектори базису. Тоді, як відомо, відповідний їй лінійний оператор має обернений, тому і його матриця має обернену (пізніше ми знайдемо її вигляд). Це означає, що система лінійних рівнянь $Ax = b$ має єдиний розв’язок: $x = A^{-1}b$. Зауважимо, що визначником

квадратної системи (кількість невідомих дорівнює кількості рівнянь) будемо називати визначник її матриці. Таким чином, одержано важливий результат.
Теорема 4.1.5. (Крамера). Квадратна система лінійних рівнянь має розв'язок і притому єдиний тоді і тільки тоді, коли її визначник не дорівнює нулю.

Наступні властивості визначників.
Теорема 4.1.6. Нехай є квадратні матриці *A* і A_1, і один вектор-рядок (вектор-стовпчик) матриці A_1 відрізняється від відповідного вектора-рядка (вектора-стовпчика) матриці *A* числовим множником α, а всі відповідні інші їхні елементи співпадають. Тоді $\det(A_1) = \alpha \det(A)$.
Доведення. Нехай вектор матриці *A*, про який йде мова має вигляд
$x_k = a_{k1}e_1 + a_{k2}e_2 + ... + a_{kn}e_n, 1 \le k \le n$, а відповідний вектор матриці A_1 -
$x'_k = \alpha a_{k1}e_1 + \alpha a_{k2}e_2 + ... + \alpha a_{kn}e_n = \alpha x_k$. Справедливість твердження випливає з властивості однорідності зовнішнього добутку векторів: $\det(A_1)e_1 \wedge ... \wedge e_n =$
$= x_1 \wedge ... \wedge x_{k-1} \wedge \alpha x_k \wedge x_{k+1} \wedge ... \wedge x_n = \alpha\left(x_1 \wedge ... \wedge x_{k-1} \wedge x_k \wedge x_{k+1} \wedge ... \wedge x_n\right) =$
$= \alpha \det(A)e_1 \wedge ... \wedge e_n$.

Означення. Нехай викреслено *k* деяких рядків і *k* стовпчиків матриці (не обов'язково квадратної) і утворено матрицю, елементами якої є елементи даної матриці, що знаходяться на перетині цих рядків і стовпчиків. Визначник так побудованої нової матриці називається мінором даної матриці. Число *k* називається порядком мінора.

Очевидно, в матриці можна знайти ненульовий мінор найбільшого порядку, і він, взагалі кажучи, не буде єдиним. Не зменшуючи загальності, можна вважати, що його елементи знаходяться у верхньому лівому куті матриці. Нехай його порядок дорівнює *r*. Застосовуючи теорему 4.1.4 до рядків мінора, побачимо, що вони лінійно незалежні. Те ж саме стосується його векторів-стовпчиків. Доповнюючи вектори-рядки мінора до відповідних векторів-рядків самої матриці, побачимо, що й вони будуть лінійно незалежними. Так само лінійно незалежними будуть відповідні стовпчики.
Нехай матриця *A* має *m* рядків і *n* стовпчиків і *r* – максимальна кількість лінійно незалежних векторів-рядків ($r \le m$), тобто, рядковий ранг матриці. Будемо вважати, що також $r \le n$. Розглянемо матрицю, що складається з цих *r* рядків:

$$\begin{pmatrix} a_{11} & ... & a_{1r} & ... & a_{1n} \\ . & . & . & . & . \\ a_{r1} & ... & a_{rr} & ... & a_{rn} \end{pmatrix}$$

Нехай $\{e_1, ..., e_n\}$ - деякий базис простору. Як і раніше, розглянемо зовнішній добуток

$$\left(a_{11}e_1+a_{12}e_2+\ldots+a_{1n}e_n\right)\wedge\ldots\wedge\left(a_{r1}e_1+a_{r2}e_2+\ldots+a_{rn}e_n\right)\neq 0$$

(нерівність нулю випливає з теореми 4.1.4). Цілком очевидно, що він буде лінійною комбінацією добутків вигляду $D_i e_{i_1}\wedge e_{i_2}\wedge\ldots\wedge e_{i_r}$, в яких номери базисних векторів утворюють зростаючу послідовність: $i_1<i_2<\ldots<i_r$, причому, беруться всі можливі такі комбінації, а коефіцієнти при них - це мінори матриці. Оскільки вектори $e_{i_1}\wedge e_{i_2}\wedge\ldots\wedge e_{i_r}$ утворюють базис у підпросторі алгебри Грасмана, то хоча б один з коефіцієнтів $D_i\neq 0$, тобто, знайдеться ненульовий мінор порядку *r*. А тоді і відповідні *r* вектори-стовпчики лінійно незалежні. Якщо *s* – стовпчиковий ранг матриці, то з доведеного випливає, що $r\leq s$. Відправляючись тепер від векторів-стовпчиків, так само доведемо, що $s\leq r$. Дійсно, беремо зовнішній добуток

$$\left(a_{11}e_1+a_{21}e_2+\ldots+a_{n1}e_n\right)\wedge\ldots\wedge\left(a_{1s}e_1+a_{2s}e_2+\ldots+a_{ns}e_n\right)\neq 0$$

Міркуючи, як і в попередньому випадку, пояснюємо, що знайдеться ненульовий мінор порядку *s* і відповідні *s* рядків будуть лінійно незалежними. Тоді $s\leq r$ і в результаті $s=r$.
Отже, доведено важливий факт:

Теорема 4.1.7. Рядковий і стовпчиковий ранги матриці співпадають і дорівнюють найбільшому з порядків її ненульових мінорів.

Означення. Рангом лінійної системи називається ранг матриці її лівої частини.

Означення. Нехай матриця квадратна, і в ній виділено деякий мінор. Визначник з елементів, що залишились після викреслювань, називається доповнюючим мінором. Нехай було викреслено рядки $i_1,i_2,\ldots,i_k, 1\leq k\leq n-1$ і стовпчики $j_1,j_2,\ldots,j_k$. Алгебраїчним доповненням до даного мінора називається доповнюючий мінор, помножений на $(-1)^{i_1+\ldots+i_k+j_1+\ldots+j_k}$.
З теореми 3.2.1 випливає, що ранг матриці не змінюється в результаті її елементарних перетворень, тобто, додавання до будь-якого рядка (стовпчика) лінійної комбінації інших рядків (відповідно стовпчиків).

Теорема 4.1.10. Якщо кожен елемент *j*-го рядка (стовпчика) визначника подано у вигляді суми двох доданків, то визначник дорівнює сумі двох визначників, у яких співпадають всі рядки (стовпчики) крім *j*-го рядка (стовпчика). При цьому у одного визначника *j*-го рядок (стовпчик) складається з перших доданків, а у іншого – з других доданків:

$$\begin{vmatrix} c_{11} & \cdot & \cdot & \cdot & c_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{j1}+b_{j1} & \cdot & \cdot & \cdot & a_{jn}+b_{jn} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n1} & \cdot & \cdot & \cdot & c_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} c_{11} & \cdot & \cdot & \cdot & c_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{j1} & \cdot & \cdot & \cdot & a_{jn} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n1} & \cdot & \cdot & \cdot & c_{nn} \end{vmatrix} + \begin{vmatrix} c_{11} & \cdot & \cdot & \cdot & c_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{j1} & \cdot & \cdot & \cdot & b_{jn} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n1} & \cdot & \cdot & \cdot & c_{nn} \end{vmatrix} \tag{4.1.8}$$

Зауваження. Формула у випадку сум у стовпчиках записується відповідним чином.
Доведення проведемо лише в одному варіанті; в другому все робиться аналогічно.
Вектор, який відповідає j-му рядку визначника, має вигляд $x_j = x'_j + x''_j$, тому доведення випливає з очевидної рівності

$$x_1 \wedge \ldots \wedge x_{j-1} \wedge \left(x'_j + x''_j\right) \wedge x_{j+1} \wedge \ldots \wedge x_n =$$
$$= x_1 \wedge \ldots \wedge x_{j-1} \wedge x'_j \wedge x_{j+1} \wedge \ldots \wedge x_n + x_1 \wedge \ldots \wedge x_{j-1} \wedge x''_j \wedge x_{j+1} \wedge \ldots \wedge x_n$$

Наслідок: Визначник не зміниться, якщо до деякого його рядка (стовпчика) додати лінійну комбінацію останніх рядків (стовпчиків). Дійсно, в силу теореми 4.1.10 він розпадеться на суму двох визначників, причому в одному з них рядок (стовпчик) буде лінійною комбінацією інших і буде дорівнювати нулю за теоремою 4.1.3.

Дамо інтерпретацію наслідку з властивості 5) внутрішнього добутку (див. 3.5) мовою визначників.

Кожен елемент a_{ij} матриці можна вважати її одновимірним мінором. Викреслюючи рядок і стовпчик, у яких знаходиться цей елемент, одержимо мінор, який позначимо M_{ji}.

Означення. Алгебраїчним доповненням елемента матриці A називається число $D_{ji} = (-1)^{i+j} M_{ji}$.

Теорема 4.1.11. Справедлива формула розкриття визначника по i-му рядку:

$$\det A = a_{i1}D_{1i} + a_{i2}D_{2i} + \ldots + a_{ik}D_{ki} + \ldots + a_{in}D_{ni} \tag{4.1.9}$$

Доведення. Для більшої наочності доведемо формул розкриття по першому рядку:

$$\det A = a_{11}D_{11} + a_{21}D_{12} + \ldots + a_{i1}D_{1i} + \ldots + a_{n1}D_{1n} \tag{4.1.10}$$

Розглянемо зовнішній добуток

Тим самим теорему доведено.

Аналогічна формула справедлива, якщо розкривати визначник по рядку.

Радимо самостійно довести формулу розкриття визначника по довільному рядку – з номером i, а також аналогічну формулу розкриття визначника по довільному стовпчику з номером j:

$$\det A = a_{1j}D_{j1} + a_{2j}D_{j2} + \ldots + a_{ij}D_{ji} + \ldots + a_{nj}D_{jn} \tag{4.1.11}$$

Зауважимо, що формулу розкриття визначника по j-му рядку можна довести, розкриваючи j-ту дужку в зовнішньому добутку $\sum_i a_{1i}e_i \wedge \ldots \wedge \sum_i a_{ni}e_i$. Радимо читачу зробити це самостійно.

Наслідок.

$$\sum_j a_{ij}D_{jk} = 0 \tag{4.1.12}$$

якщо D_{jk} – алгебраїчні доповнення до елементів j-того стовпчика ($k \neq i$).

Дійсно, цю суму можна інтерпретувати (з точністю до знака) як розкриття визначника по i-му рядку, але в такому визначнику i-й та k-й рядки однакові. У такому разі, як уже відомо, визначник дорівнює нулю. Звідси віпливає дуже важливий наслідок.

Зіставимо матриці A матрицю, елементи якої обчислюються в такий спосіб: в i-му стовпчику запишемо по черзі алгебраїчні доповнення до елементів i-го рядка. Тобто, кожен елемент матрицы B буде мати вигляд

$$b_{ij} = D_{ji},$$

Якщо тепер перемножити A і B у довільному порядку (AB або BA), то за попередніми властивостями одержимо скалярну матрицю, діагональ якої буде складатись з однакових чисел, а саме, $\det(A)$ (недіагональні елементи будуть дорівнювати нулю).

Означення. Матриця B, елементи якої знаходяться за формулами

$$b_{ij} = D_{ji}, \qquad (4.1.13)$$

називається приєднаною для A.

Отже, $AB = BA = \det(A)I$. Тоді з властивостей визначника випливає, що $\det(AB) = \det(A)$. Якщо $\det(A) = 0$, матриця A є дільником нуля і, як відомо, не може бути дільником одиниці.

Нехай тепер $\det(A) \neq 0$. З попередніх міркувань можна визначити обернену до A матрицю, якщо B – приєднана до неї:

$$A^{-1} = \left(\det(A)\right)^{-1} B. \qquad (4.1.14)$$

Застосуємо одержану формулу для розв‘язку квадратної системи лінійних рівнянь при умові $\det(A) \neq 0$.

Теорема 4.1.12. Нехай *A,B,* — матриці в R^n. Тоді

$$\det(AB) = \det(A)\det(B) \qquad (4.1.17)$$

Доведення. Нехай $C=AB$ і $e = \{e_1,...,e_n\}$ – деякий базис, в якому $g_k = \sum_i b_{ik} e_i$, $f_k = \sum_j a_{jk} g_j$.

Якщо $\{g_j\}$ не утворює базис, то A – вироджена матриця, а тому і AB – теж вироджена і $\det(A) = \det(AB) = 0$. Отже, в цьому випадку теорему доведено. Аналогічно теорема доводиться, коли B – вироджена матриця. Розглянемо тепер випадок невироджених A і B.

З наведених формул випливає, що $f_k = \sum_j c_{jk} e_j$. Тепер, з одного боку,

$$f_1 \Lambda ... \Lambda f_n = \det A, \;\; g_1 \Lambda ... \Lambda g_n = \det A \det B \;\; e_1 \Lambda ... \Lambda e_n,$$

з іншого –

$$f_1 \Lambda ... \Lambda f_n = \det C \;\; e_1 \Lambda ... \Lambda e_n$$

Звідки $\det C = \det AB = \det A \det B.$ Теорему доведено.

Височок. Нехай $AA^{-1} = A^{-1}A = E$. Тоді $\det E = 1 = \det A \det A^{-1}$, і

$$\det\left(A^{-1}\right) = (\det A)^{-1} \qquad (4.1.18)$$

З цієї формули можна одразу зробити висновок, що значення визначника не залежить від того, в якому базисі ми брали зображення оператора матрицею. Адже в іншому базисі матриця оператора має вигляд $\tilde{A} = TAT^{-1}$, де T – матриця заміни базису. Тоді

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

Отже,

$$\det TAT^{-1} = \det A. \tag{4.1.19}$$

Таким чином, можна говорити про визначник оператора як інваріант, тобто величину, незалежну від базису, в якому оператор зображується матрицею.

Вправа 4.1.1. Довести формулу (4.1.17), виходячи з зображення матриці в формі (2.3.40). Вказівка: спочатку довести формулу для довільних елементарних матриць.

Зауваження. Невироджені матриці, які діють на лінійному просторі E_n над полем *K*, утворюють мультиплікативну групу, що позначається *GL(n,K)*. Оскільки добуток таких матриць – теж невироджена матриця, кожна невироджена матриця має обернену, одинична матриця теж невироджена. Отже, завдяки формулі (4.1.17) можна зробити висновок, що функція $\det : GL(n,K) \to K$ є гомоморфізм з групи невироджених матриць на мультиплікативну групу поля *K* (в нашому випадку $\mathbb{R} \setminus \{0\}$ або $\mathbb{C} \setminus \{0\}$).

Повернемось до методу Гауса, за допомогою якого було доведено теорему 2.3.4 про можливість зведення матриці спочатку до верхньотрикутного вигляду, а потім до діагонального. Нагадаємо, що на кожному кроці треба було одержувати нулі під діагональним членом, і це вдавалось зробити, якщо він не дорівнював нулю. З'ясуємо зараз, при яких умовах можна звести невироджену квадратну матрицю до нижньотрикутного вигляду, не переставляючи рядки.
Розглянемо лінійну систему рівнянь

$$\begin{array}{ccccc} a_{11}x_1 + & a_{12}x_2 + & \dots & +a_{1n}x_n = & y_1 \\ a_{21}x_1 + & a_{22}x_2 + & \dots & +a_{2n}x_n = & y_1 \\ . & . & \dots & . & . \\ a_{n1}x_1 + & a_{n2}x_2 + & \dots & +a_{nn}x_n = & y_n \end{array} \tag{4.1.20}$$

Їй зручно зіставляти розширену матрицю

.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & y_n \\ a_{21} & a_{22} & \dots & a_{2n} & | & y_2 \\ . & . & \dots & . & | & \\ a_{1n} & a_{2n} & \dots & a_{nn} & | & y_n \end{pmatrix} \qquad (4.1.21)$$

Далі будемо використовувати позначення:
кутові мінори:

$$\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \dots, \Delta_n = \det(A), \qquad (4.1.22)$$

а також для мінорів, побудованих на перших *p* рядках і стовпчиках $k_1, k_2, ..., k_p$:

$$M\begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}. \qquad (4.1.23)$$

Зауважимо, що $M\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = \Delta_p$.

Далі будемо вважати, що всі кутові мінори не дорівнюють нулю:

$$\Delta_1 \neq 0,\ \Delta_2 \neq 0, ..., \Delta_n = \det(A) \neq 0 \qquad (4.1.24)$$

Додамо до другого рядка перший, помножений на $-\dfrac{a_{21}}{a_{11}}$. Потім так само до третього рядка додамо перший, помножений на $-\dfrac{a_{31}}{a_{11}}$ і т.д. Після цієї процедури під елементом a_{11} в першому стовпчику одержимо нулі. Новим діагональним елементом буде $a_{22}^{(1)} - \dfrac{a_{21}}{a_{11}} a_{12} = \dfrac{1}{a_{11}} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \dfrac{\Delta_2}{\Delta_1} \neq 0$.

Потім до 3-го рядка додамо 2-й, помножений на $-\dfrac{a_{31}}{a_{22}^{(1)}}$ і т.ін. Якщо діагональні елементи на кожному кроці не дорівнювали нулю, матрицю можна було звести до верхньотрикутного вигляду. Доведемо, що нерівності нулю всіх кутових мінорів достатньо для такого зведення.
Розглянемо *p*-й крок процедури. Матриця системи буде мати вигляд:

$$C_p = \begin{pmatrix} a_{11} & a_{11} & \dots & a_{1p} & a_{1,p+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2p}^{(1)} & a_{2,p+1}^{(1)} & \dots & a_{2n}^{(1)} \\ . & . & \dots & . & . & \dots & . \\ 0 & 0 & \dots & a_{pp}^{(p-1)} & a_{p,p+1}^{(p-1)} & \dots & a_{pn}^{(p-1)} \\ 0 & 0 & \dots & 0 & a_{p+1,p+1}^{(p)} & \dots & a_{p+1,n}^{(p)} \\ . & . & \dots & . & . & \dots & . \\ 0 & 0 & \dots & 0 & a_{n,p+1}^{(p)} & \dots & a_{n,n}^{(p)} \end{pmatrix}. \qquad (4.1.25)$$

Мінори цієї матриці будемо позначати

$$N\begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}. \qquad (4.1.26)$$

Матриця C_p була утворена з матриці A шляхом додавання до рядків A інших її рядків з меншими номерами, помножених на деякі числа, що, як відомо, не змінює мінорів. Тому

$$M\begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = N\begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}. \qquad (4.1.27)$$

Якщо $i > p$, то те саме можна сказати для мінорів $p+1$ - го порядку:

$$M\begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix} = N\begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix}.$$

Завдяки діагональному вигляду

$$M\begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = a_{11}a_{22}^{(1)}...a_{pp}^{(p-1)}, \qquad (4.1.28)$$

а також

$$M\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} = a_{11}a_{22}^{(1)}...a_{pp}^{(p-1)}a_{ik}^{(p)}. \qquad (4.1.29)$$

Поділивши другу рівність на першу, одержимо:

$$a_{ik}^{(p)}=\frac{M\begin{pmatrix}1 & 2 & \dots & p & i\\ 1 & 2 & \dots & p & k\end{pmatrix}}{M\begin{pmatrix}1 & 2 & \dots & p\\ k_1 & k_2 & \dots & k_p\end{pmatrix}},$$

$i,k=p+1,...,n$. Звідси при $i=k=p+1$

$$a_{p+1,p+1}^{(p)}=\frac{M\begin{pmatrix}1 & 2 & \dots & p & p+1\\ 1 & 2 & \dots & p & p+1\end{pmatrix}}{M\begin{pmatrix}1 & 2 & \dots & p\\ 1 & 2 & \dots & 2\end{pmatrix}}=\frac{\Delta_{p+1}}{\Delta_p} \qquad (4.1.30)$$

і, як наслідок, $\Delta_p=a_{11}a_{22}^{(1)}...a_{pp}^{(p-1)}$ (p=1,…,n). Отже, умов (4.1.24) необхідно і достатньо для зведення невиродженої матриці до верхньотрикутного вигляду. Як уже відзначалось при доведенні теореми 2.3.4, кожне з описаних елементарних перетворень матриці відповідає її множенню на елементарну нижньотрикутну невироджену матрицю. Добуток цих матриць, матриця $\tilde{L}$, – нижньотрикутна, обернена до якої – теж нижньотрикутна. Отже, таким чином прийдемо до верхньотрикутної матриці (при p=n) $R=C_n=\tilde{L}A$. Нехай $L=\left(\tilde{L}\right)^{-1}$. Тоді

$$A=LR \qquad (4.1.31)$$

Сформулюємо результат.

Теорема 4.1.13. Для того, щоб квадратну матрицю можна було звести до добутку нижньотрикутної і верхньотрикутної матриць, необхідно і достатньо, щоб всі її кутові мінори не дорівнювали нулю.

Вправа 4.1.1. Нехай $a_{11}\neq 0$. Зіставимо визначнику

$$\Delta_n=\begin{vmatrix}a_{11} & \dots & a_{1j} & \dots & a_{1n}\\ . & . & . & . & .\\ a_{i1} & \dots & a_{ij} & \dots & a_{in}\\ . & . & . & . & .\\ a_{n1} & \dots & a_{nj} & \dots & a_{nn}\end{vmatrix}$$

Визначник Δ_{n-1} порядку n-1 з елементами

$$a_{ij}^{(1)} = \begin{vmatrix} a_{11} & a_{1j} \\ a_{i1} & a_{ij} \end{vmatrix}, \qquad (4.1.33)$$

$i, j = 2,...,n$, тобто,

$$\Delta_{n-1} = \begin{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1j} \\ a_{21} & a_{2j} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{21} & a_{2n} \end{vmatrix} \\ . & . & . & . & . \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{i1} & a_{i2} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1j} \\ a_{i1} & a_{ij} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{i1} & a_{in} \end{vmatrix} \\ . & \cdots & . & \cdots & . \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{n1} & a_{n2} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1j} \\ a_{n1} & a_{nj} \end{vmatrix} & \cdots & \begin{vmatrix} a_{11} & a_{1n} \\ a_{n1} & a_{nn} \end{vmatrix} \end{vmatrix} \qquad (4.1.34)$$

Справедлива формула:

$$\Delta_n = \frac{1}{(a_{11})^{n-2}} \Delta_{n-1} \qquad (4.1.35)$$

Доведення. Будемо перетворювати визначник Δ_n за методом Гауса, віднімаючи від кожного i – го рядка ($i = 2,...,n$) перший, помножений на $\frac{a_{i1}}{a_{11}}$. В результаті кожен j-й елемент $\left(j = 2,...,n\right)$ i –го рядка буде мати вигляд:

$$a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}} = \frac{1}{a_{11}} \begin{vmatrix} a_{11} & a_{1j} \\ a_{i1} & a_{ij} \end{vmatrix}.$$

Таких елементів в i–му рядку буде n-1, отже, за визначник можна винести $\frac{1}{(a_{11})^{n-1}}$. В першому стовпчику будуть знаходитись всі нулі крім першого елемента a_{11}. Розкриваючи визначник за першим стовпчиком, отримаємо $\frac{a_{11}}{(a_{11})^{n-1}} \Delta_{n-1}^{(1)} = \frac{1}{(a_{11})^{n-2}} \Delta_{n-1}^{(1)}$, що й треба було довести.

4.2. ОБ'ЄМ *m*-ВИМІРНОГО ПАРАЛЕЛЕПІПЕДА.

Так називається тіло в просторі, побудоване на m лінійно незалежних векторах $g_1,...,g_m$. При m=2 паралелепіпед - це паралелограм, його об'м – це площа паралелограма, яка дорівнює добутку основи на висоту. Об'єм трьохвимірного паралелепіпеда дорівнює добутку площі основи на його висоту. Об'єм V_m m-вимірного паралелепіпеда дорівнює добутку об'єму m-1 – вимірної основи на висоту:

$$V_m = V_{m-1} h. \qquad (4.2.1)$$

Наступний визначник називається визначником Грама системи векторів $g_1,...,g_m$:

$$\Gamma(g_1,...g_m) = \begin{vmatrix} (g_1,g_1) & (g_2,g_1) & \cdots & (g_m,g_1) \\ (g_1,g_2) & (g_2,g_2) & & (g_m,g_2) \\ . & . & . & . \\ (g_1,g_m) & (g_2,g_m) & \cdots & (g_m,g_m) \end{vmatrix} \qquad (4.2.2)$$

Нехай $m<n$, $L_m \subset E_n$ - m-вимірний підпростір. Візьмемо $f \in E_n$. Нехай g – проекція f на підпростір L_m:

$$g = \lambda_1 g_1 + ... + \lambda_m g_m$$

Як відомо, $h = (f - g) \perp L_m$, $f = g + h = \lambda_1 g_1 + ... + \lambda_m g_m + h$, і таке зображення f єдине. Помножимо обидві частини цієї рівності на кожен вектор g_k:

$$(f, g_k) = \lambda_1 (g_1, g_k) + ... + \lambda_m (g_m, g_k),\ k=1,\dots,m.$$

Це лінійна система відносно $\lambda_1...,\lambda_m$. З попередніх міркувань вона має єдиний розв'язок. Визначник лінійної системи – це визначник Грама системи векторів $g_1,...,g_m$, тому він не дорівнює нулю.

Навпаки, якщо $g_1,...,g_m$ лінійно залежні, тоді $\Gamma(g_1,...g_m)=0$. Дійсно, в такому випадку існують не всі нульові значення $\lambda_1...,\lambda_m$ такі, що

$$\lambda_1 g_1 + ... + \lambda_m g_m = 0$$

Скалярно помноживши обидві частини рівності на кожен вектор g_k, одержимо:

$$\begin{aligned}
&\lambda_1(g_1,g_1)+\lambda_2(g_2,g_1)+...+\lambda_m(g_m,g_1)=0\\
&\lambda_1(g_1,g_2)+\lambda_2(g_2,g_2)+...+\lambda_m(g_m,g_2)=0\\
&...\\
&\lambda_1(g_1,g_m)+\lambda_2(g_2,g_m)+...+\lambda_m(g_m,g_m)=0
\end{aligned} \qquad (4.2.3)$$

Оскільки система має нетривіальний розв’язок, її визначник дорівнює нулю. Отже, доведено важливий факт:

Теорема . Система векторів лінійно незалежна тоді і тільки тоді, коли її визначник Грама не дорівнює нулю.

Знайдемо тепер величину $\delta=\inf\limits_{y\in L}\|f-y\|=\|h\|=\|f-g\|$ при умові, що $\lambda_1...,\lambda_m$ визначаються з системи (4.2.3). Значить

$\delta^2=(h,h)=(h,f)=(f,f)-\lambda_1(g_1,f)-\lambda_2(g_2,f)-...-\lambda_m(g_m,f)$. Значення $\lambda_1...,\lambda_m$ задовольняють лінійну систему рівнянь

$$\begin{aligned}
&(f,f)-\delta^2-\lambda_1(g_1,f)-\lambda_2(g_2,f)-...-\lambda_m(g_m,f)=0\\
&(f,g_1)-\lambda_1(g_1,g_1)-...-\lambda_m(g_m,g_1)=0\\
&...............\\
&(f,g_m)-\lambda_1(g_1,g_m)-...-\lambda_m(g_m,g_m)=0
\end{aligned} \qquad (4.2.4)$$

Ця система має нетривіальний розв’язок, тому її визначник дорівнює нулю:

$$\begin{vmatrix} (f,f)-\delta^2 & (g_1,f) & (g_2,f) & \cdots & (g_m,f) \\ (f,g_1) & (g_1,g_1) & (g_2,g_1) & \cdots & (g_m,g_1) \\ . & . & . & . & . \\ (f,g_m) & (g_1,g_m) & (g_2,g_m) & \cdots & (g_m,g_m) \end{vmatrix} = 0$$

Звідси

$$\begin{vmatrix} (f,f) & (g_1,f) & (g_2,f) & \cdots & (g_m,f) \\ (f,g_1) & (g_1,g_1) & (g_2,g_1) & \cdots & (g_m,g_1) \\ . & . & . & . & . \\ (f,g_m) & (g_1,g_m) & (g_2,g_m) & \cdots & (g_m,g_m) \end{vmatrix} = \delta^2 \begin{vmatrix} (g_1,g_1) & (g_2,g_1) & \cdots & (g_m,g_1) \\ (g_1,g_2) & (g_2,g_2) & & (g_m,g_2) \\ . & . & . & . \\ (g_1,g_m) & (g_2,g_m) & \cdots & (g_m,g_m) \end{vmatrix}$$

Тоді

$$\delta^2 = \frac{\Gamma(f,g_1,...g_m)}{\Gamma(g_1,...g_m)} \tag{4.2.5}$$

Обчислимо визначник Грама при *m*=2. Якщо кут між векторами g_1, g_2 дорівнює φ, $(g_1,g_2)=(g_2,g_1)=\|g_1\|\|g_2\|\cos\varphi$.

$$\Gamma(g_1,g_2)=\begin{vmatrix} (g_1,g_1) & (g_2,g_1) \\ (g_1,g_2) & (g_2,g_2) \end{vmatrix} = \begin{vmatrix} \|g_1\|^2 & \|g_1\|\|g_2\|\cos\varphi \\ \|g_1\|\|g_2\|\cos\varphi & \|g_2\|^2 \end{vmatrix} =$$

$$= \|g_1\|^2\|g_2\|^2\left(1-\cos^2\varphi\right)=\left(\|g_1\|\|g_2\|\sin\varphi\right)^2 = V^2$$

Теорема. $V^2 = \Gamma(g_1,...g_m)$

Доведення. При *m*=2 справедливість формули доведено. Припустимо, що для *m* формула справедлива. Для паралелепіпеда, побудованого на векторах $f, g_1,..., g_m$ згідно з (4.2.1) $\Gamma(f,g_1,...g_m)=\delta^2\Gamma(g_1,...g_m)=\left(\|h\|V_m\right)^2=\left(V_{m+1}\right)^2$.

4.3. ЗОВНІШНІЙ ДОБУТОК *k* ЛІНІЙНИХ КОМБІНАЦІЙ *m*-ВИМІРНИХ ВЕКТОРІВ

Нехай

$$\begin{aligned} & f_1 = a_{11}g_1 + a_{12}g_2 + \ldots + a_{1m}g_m \\ & f_2 = a_{21}g_1 + a_{22}g_2 + \ldots + a_{2m}g_m \\ & \ldots\ldots\ldots \\ & f_k = a_{k1}g_1 + a_{k2}g_2 + \ldots + a_{km}g_m \end{aligned} \qquad (4.3.1)$$

Позначимо: $M = \{1, 2, \ldots, m\}$, $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\} \subset M$, де $\gamma_1 < \gamma_2 < \ldots < \gamma_k$, $G_\gamma = g_{\gamma_1} \wedge g_{\gamma_2} \wedge \ldots \wedge g_{\gamma k}$. Введемо також позначення для мінора матриці A коефіцієнтів попередньої системи рівностей, що утворена стовпчиками з номерами $\gamma_1, \gamma_2, \ldots, \gamma_k$:

$$A_\Gamma = \begin{vmatrix} a_{1\gamma_1} & \cdots & a_{1\gamma k} \\ . & . & . \\ a_{k\gamma_1} & \cdots & a_{k\gamma_{1k}} \end{vmatrix} \qquad (4.3.2)$$

Лема.

$$f_1 \wedge f_2 \wedge \ldots \wedge f_k = \sum_\Gamma A_\Gamma G_\Gamma, \qquad (4.3.3)$$

де Γ пробігають всі *k*-елементні підмножини множини $M = \{1, 2, \ldots, m\}$.

Доведення.

$$f_1 \wedge f_2 \wedge \ldots \wedge f_k = \sum_{\alpha_1, \ldots, \alpha_k \in M} a_{1\alpha_1} a_{1\alpha_2} \ldots a_{1\alpha_k} g_{\alpha_1} \wedge g_{\alpha_2} \wedge \ldots \wedge g_{\alpha_k},$$

де $\alpha_1, \alpha_2, \ldots, \alpha_k$ попарно нерівні між собою. В такому випадку будемо писати:

$$f_1 \wedge f_2 \wedge \ldots \wedge f_k = \sum_{\alpha_i \neq \alpha_j} a_{1\alpha_1} a_{1\alpha_2} \ldots a_{1\alpha_k} g_{\alpha_1} \wedge g_{\alpha_2} \wedge \ldots \wedge g_{\alpha_k}$$

Цілком зрозуміло, що такі системи індексів утворюють всі можливі k-елементні підмножини M, причому, кожен такий набір зустрінеться $k!$ разів. Нехай $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\} \subset M$ - правильний набір, тобто, $\gamma_1 < \gamma_2 < \ldots < \gamma_k$. Введемо позначення для підстановки номерів $\alpha_1, \alpha_2, \ldots, \alpha_k$:

$\sigma_{\alpha_1,\alpha_2,\ldots,\alpha_k} = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_k \\ \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{pmatrix}$, $\varepsilon\left(\sigma_{\alpha_1,\alpha_2,\ldots,\alpha_k}\right) \in \{1,1\}$ - її парність. Тоді

$g_{\alpha_1} \wedge g_{\alpha_2} \wedge \ldots \wedge g_{\alpha_k} = \varepsilon\left(\sigma_{\alpha_1,\alpha_2,\ldots,\alpha_k}\right) g_{\gamma_1} \wedge g_{\gamma_2} \wedge \ldots \wedge g_{\gamma k}$. Отже,

$$f_1 \wedge f_2 \wedge \ldots \wedge f_k = \sum_{\alpha_i \neq \alpha_j} a_{1\alpha_1} a_{1\alpha_2} \ldots a_{1\alpha_k} g_{\alpha_1} \wedge g_{\alpha_2} \wedge \ldots \wedge g_{\alpha_k} =$$

$= \sum_{\Gamma} \left(\sum_{(\alpha_1,\alpha_2,\ldots,\alpha_k)} \varepsilon\left(\sigma_{\alpha_1,\alpha_2,\ldots,\alpha_k}\right) a_{1\alpha_2} \ldots a_{1\alpha_k} \right) g_{\gamma_1} \wedge g_{\gamma_2} \wedge \ldots \wedge g_{\gamma k}$, оскільки

$$\sum_{(\alpha_1,\alpha_2,\ldots,\alpha_k)} \varepsilon\left(\sigma_{\alpha_1,\alpha_2,\ldots,\alpha_k}\right) a_{1\alpha_2} \ldots a_{1\alpha_k} = \begin{vmatrix} a_{1\gamma_1} & \cdots & a_{1\gamma k} \\ . & . & . \\ a_{k\gamma_1} & \cdots & a_{k\gamma_{1k}} \end{vmatrix} = A_\Gamma$$

З формули () можна одержати кілька важливих наслідків.

Теорема Біне-Коші.

Нехай

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ . & . & . \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}, B = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ . & . & . \\ b_{k1} & \cdots & b_{km} \end{pmatrix}, C = \begin{pmatrix} c_{11} & \cdots & c_{1k} \\ . & . & . \\ c_{m1} & \cdots & c_{mk} \end{pmatrix}, \tag{4.3.4}$$

де $A=BC$, причому $k<m$. Нехай $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_k\} \subset M = \{1, 2, \ldots, m\}$, де $\gamma_1 < \gamma_2 < \ldots < \gamma_k$. Позначимо

$$B_\Gamma = \begin{vmatrix} b_{1\gamma_1} & \cdots & b_{1\gamma k} \\ . & . & . \\ b_{k\gamma_1} & \cdots & b_{k\gamma_{1k}} \end{vmatrix}, \ C^\Gamma = \begin{vmatrix} c_{\gamma_1 1} & \cdots & c_{\gamma_1 k} \\ . & . & . \\ c_{\gamma_k 1} & \cdots & c_{\gamma_k k} \end{vmatrix}. \tag{4.3.5}$$

Тоді

$$\det A = \sum_{\Gamma \in M} B_\Gamma C^\Gamma \tag{4.3.6}$$

Доведення. Нехай вектори $e_1, e_2, ..., e_k$ лінійно незалежні. Розглянемо вектори

$$\begin{array}{ll} f_1 = b_{11}g_1 + b_{12}g_2 + ... + b_{1m}g_m, & g_1 = c_{11}e_1 + c_{12}e_2 + ... + c_{1k}e_k, \\ f_2 = b_{21}g_1 + b_{22}g_2 + ... + b_{2m}g_m, & g_2 = c_{21}e_1 + c_{22}e_2 + ... + c_{2k}e_k, \\ .... & .... \\ f_k = b_{k1}g_1 + b_{k2}g_2 + ... + b_{km}g_m, & g_m = c_{m1}e_1 + c_{m2}e_2 + ... + c_{mk}e_k. \end{array} \tag{4.3.7}$$

Тоді

$$\begin{array}{l} f_1 = a_{11}e_1 + a_{12}e_2 + ... + a_{1k}e_k, \\ f_2 = a_{21}e_1 + a_{22}e_2 + ... + a_{2k}e_k, \\ .... \\ f_k = a_{k1}e_1 + a_{k2}e_2 + ... + a_{kk}e_k. \end{array} \tag{4.3.8}$$

Тоді $f_1 \wedge f_2 ... \wedge f_k = (\det A) e_1 \wedge e_2 \wedge ... \wedge e_k$, а також $f_1 \wedge f_2 ... \wedge f_k =$
$= \sum_\Gamma B_\Gamma g_{\gamma_1} \wedge g_{\gamma_2} \wedge ... \wedge g_{\gamma_k}$.

З (4.3.7) випливає, що

$$\begin{array}{l} g_{\gamma_1} = c_{\gamma_1 1}e_1 + c_{\gamma_1 2}e_2 + ... + c_{\gamma_1 k}e_k, \\ g_{\gamma_2} = c_{\gamma_2 1}e_1 + c_{\gamma_2 2}e_2 + ... + c_{\gamma_2 k}e_k, \\ .... \\ g_{\gamma_k} = c_{\gamma_k 1}e_1 + c_{\gamma_k 2}e_2 + ... + c_{\gamma_k k}e_k. \end{array}$$

звідки $g_{\gamma_1} \wedge g_{\gamma_2} \wedge ... \wedge g_{\gamma_k} = C^{\Gamma} e_1 \wedge e_2 \wedge ... \wedge e_k$ і в результаті $f_1 \wedge f_2 ... \wedge f_k =$
$= \sum_{\Gamma} B_{\Gamma} C^{\Gamma} e_1 \wedge e_2 \wedge ... \wedge e_k$. Порівнюючи коефіцієнти при $e_1 \wedge e_2 \wedge ... \wedge e_k$,

одержимо: $\det A = \sum_{\Gamma \in M} B_{\Gamma} C^{\Gamma}$, що і треба було довести.

Наслідок 1 (теорема Лапласа). Нехай квадратна матриця порядку *n* розділена на дві клітини, перша з яких має *k* рядків (а друга – *n*-*k* рядків). Нехай Γ - будь-яка *k*-елементна підмножина множини $N = \{1, 2, ..., n\}$, а $\Gamma' = N \setminus \Gamma$. Нехай $A_{1\Gamma}$ - мінор *k*-го порядку матриці A_1, номери стовпчиків якого утворюють Γ, а $A_{2\Gamma}$ - мінор *n*-*k*-го порядку матриці A_2, номери стовпчиків якого утворюють Γ'. Тоді

$$\det A = \sum_{\Gamma \subset N} (-1)^{\gamma_1 + ... + \gamma_k - \frac{k(k+1)}{2}} A_{1\Gamma} A_{2\Gamma'} \qquad (4.3.9)$$

Доведення. Нехай

$$\begin{aligned}
&f_1 = a_{11} e_1 + a_{12} e_2 + ... + a_{1n} e_n, \\
&...................................... \\
&f_k = a_{k1} e_1 + a_{k2} e_2 + ... + a_{kn} e_n, \\
&f_{k+1} = a_{k+1,1} e_1 + a_{k+1,2} e_2 + ... + a_{k+1,n} e_n, \\
&...................................... \\
&f_n = a_{n1} e_1 + a_{n2} e_2 + ... + a_{nk} e_n.
\end{aligned}$$

Тоді $(\det A) e_1 \wedge e_2 \wedge ... \wedge e_n = (f_1 \wedge f_2 ... \wedge f_k) \wedge (f_{k+1} \wedge f_{k+2} ... \wedge f_n)$. Як відомо, $f_1 \wedge f_2 ... \wedge f_k = \sum_{\Gamma} A_{1\Gamma} e_{\Gamma}$, $f_{k+1} \wedge f_{k+2} ... \wedge f_n = \sum_{\Gamma_2} A_{2\Gamma} e_{\Gamma_2}$, де Γ пробігає всі *k*-елементні підмножини, а Γ_2 - всі *n*-*k*- елементні підмножини множини *N*. Для обчислення повного добутку досить замість Γ_2 брати підмножини $\Gamma' = N \setminus \Gamma$, тому що в разі перетину Γ і Γ_2 добуток буде дорівнювати нулю. Отже,

$$(\det A) e_1 \wedge ... \wedge e_n = f_1 \wedge ... \wedge f_n = \sum_{\Gamma} A_{1\Gamma} A_{2\Gamma'} e_{\Gamma} \wedge e_{\Gamma'} = \sum_{\Gamma} \varepsilon(\Gamma, \Gamma') A_{1\Gamma} A_{2\Gamma'} e_{\Gamma} \wedge e_{\Gamma'}.$$

Обчислимо парність $\varepsilon(\Gamma,\Gamma')$ підстановки $\Gamma\cup\Gamma'$. Нехай $\Gamma=\{\gamma_1,\gamma_2,...,\gamma_k\}$, $\gamma_1<\gamma_2<...<\gamma_k$. Очевидно, всі елементи, що менші за γ_1 знаходяться в Γ', цьому γ_1 утворює γ_1-1 транспозицій з цими елементами з Γ'. Аналогічно, всі елементи, що менші за γ_2 окрім γ_1 знаходяться в Γ' і утворюють γ_2-2 транспозицій з елементами з Γ' і т.д. Таким чином

$$\varepsilon(\Gamma,\Gamma')=\gamma_1+...+\gamma_1-(1+2+...+k)=\gamma_1+...+\gamma_1-\frac{k(k+1)}{2} \qquad (4.3.10)$$

Отже, формулу доведено. Можна довести аналогічну формулу у випадку, коли береться не перші, а довільні *k* рядків.

Зауваження. Доведена раніше іншим способом формула розкриття визначника по рядку або стовпчику є частинним випадком формули Лапласа.

Наслідок 2 . Для того, щоб вектори $f_1=a_{11}e_1+a_{12}e_2+...+a_{1n}e_n$, …, $f_k=a_{k1}e_1+a_{k2}e_2+...+a_{kn}e_n$ були лінійно незалежними, необхідно і достатньо, щоб хоча б один мінор матриці коефіцієнтів (a_{ij}) не дорівнював нулю.

Доведення. Дійсно, для лінійної незалежності необхідно і достатньо, щоб $f_1\wedge f_2...\wedge f_k\neq 0$, але $f_1\wedge f_2\wedge...\wedge f_k=\sum_{\Gamma}A_\Gamma e_\Gamma$ і, оскільки всі e_Γ лінійно незалежні, необхідно і достатньо, щоб хоча б один з мінорів $A_\Gamma\neq 0$.

Наслідок 3. Нехай $A\in M_{m,n}$, $B\in M_{v,r}, C\in M_{r,n}$, $rank(A)=r_A$, $rank(B)=r_B$, $rank(C)=r_C$, $A=BC$. Тоді $r_A\leq r_B$, r_C.

Доведення очевидне.

Алгебра Грасмана в евклідовому просторі.

Нехай *E* – *n*-вимірний евклідовий простір і $\{e_1,e_2,...,e_n\}$ - його ортонормований базис. Задамо скалярний добуток базисних *k*-векторів:

$$\left(e_{i_1}\wedge\ldots\wedge e_{i_k}, e_{j_1}\wedge\ldots\wedge e_{j_k}\right)=\delta_{i_1,j_1}\ldots\delta_{j_k,j_k} \tag{4.3.11}$$

Будемо вважати, що однорідні елементи різних степенів однорідності ортогональні. Тепер скалярний добуток двох довільних елементів $x=\sum_{\Gamma}x_\Gamma e_\Gamma$ і $y=\sum_{\Gamma}y_\Gamma e_\Gamma$ обчислюється за формулою:

$$(x,y)=\left(\sum_{\Gamma}x_\Gamma e_\Gamma,\sum_{\Gamma}y_\Gamma e_\Gamma\right)=\sum_{\Gamma}x_\Gamma y_\Gamma$$

Теорема . Нехай $f_1,\ldots,f_k$ і $g_1,\ldots,g_k$ - дві системи векторів. Тоді

$$\left(f_1\wedge\ldots\wedge f_k, g_1\wedge\ldots\wedge g_k\right)=\begin{vmatrix}(f_1,g_1) & \cdots & (f_1,g_k)\\ . & \cdots & .\\ (f_k,g_1) & \cdots & (f_k,g_k)\end{vmatrix} \tag{4.3.12}$$

Доведення. Розглянемо

$$\begin{aligned}&f_1=b_{11}e_1+b_{12}e_2+\ldots+b_{1n}e_n,\\&f_2=b_{21}e_1+b_{22}e_2+\ldots+b_{2n}e_n,\\&\ldots\ldots\ldots\\&f_k=b_{k1}e_1+b_{k2}e_2+\ldots+b_{kn}e_n,\end{aligned}\quad\text{і}\quad\begin{aligned}&g_1=c_{11}e_1+c_{12}e_2+\ldots+c_{1n}e_n,\\&g_2=c_{21}e_1+c_{22}e_2+\ldots+c_{2n}e_n,\\&\ldots\ldots\ldots\\&g_k=c_{k1}e_1+c_{k2}e_2+\ldots+c_{kn}e_n.\end{aligned}$$

Нехай B_Γ,C_Γ - мінори, утворені k-елементною підмножиною $\Gamma\subset\{1,2,\ldots,n\}$ з елементів матриць $\left(a_{ij}\right),\left(b_{ij}\right)$. Як відомо,

$$F\equiv f_1\wedge f_2\wedge\ldots\wedge f_k=\sum_{\Gamma}B_\Gamma e_\Gamma, G\equiv g_1\wedge g_2\wedge\ldots\wedge g_k=\sum_{\Gamma}C_\Gamma e_\Gamma.$$

Звернемо увагу, що мінори C_Γ в другому з цих виразів відповідають в формулі Біне-Коші транспонованій матриці C^T. Отже, $\left(F,G\right)=\sum_{\Gamma}B_\Gamma C_\Gamma=\det\left(BC^\Gamma\right)$.

Запишемо добуток $BC^T = \begin{pmatrix} b_{11}c_{11}+...+b_{1n}c_{1n} & ... & b_{11}c_{k1}+...+b_{1n}c_{kn} \\ \cdots & \cdots & \cdots \\ b_{k1}c_{11}+...+b_{kn}c_{1n} & \cdots & b_{k1}c_{k1}+...+b_{kn}c_{kn} \end{pmatrix} =$

$$= \begin{pmatrix} (f_1, g_1) & \cdots & (f_1, g_k) \\ . & \cdots & . \\ (f_k, g_1) & \cdots & (f_k, g_k) \end{pmatrix}.$$

Звідси випливає формула для визначників.

Наслідок.

$$\| f_1 \wedge ... \wedge f_k \|^2 = (f_1 \wedge ... \wedge f_k, f_1 \wedge ... \wedge f_k) =$$

$$= \begin{vmatrix} (f_1, f_1) & \cdots & (f_1, f_k) \\ . & \cdots & . \\ (f_k, f_1) & \cdots & (f_k, f_k) \end{vmatrix} = \Gamma(f_1,..., f_k) \qquad (4.3.13)$$

Тобто, квадрат норми *k*-вектора $f_1 \wedge ... \wedge f_k$ дорівнює визначнику Грама, який, в свою чергу, визначає квадрат об'єма паралелепіпеда, утвореного векторами $f_1,..., f_k$.

Орієнтовний об'єм паралелепіпеда.

З попередньої формули випливає, що $\| f_1 \wedge ... \wedge f_k \| = \sqrt{\Gamma(f_1,..., f_k)} = V(f_1,..., f_k)$- об'єм паралелепіпеда, утвореного векторами $f_1,..., f_k$. Нехай

$$g_1 = c_{11} f_1 + c_{12} f_2 + ... + c_{1n}ef_k,$$
$$.........$$
$$g_k = c_{k1} f_1 + c_{k2} f_2 + ... + c_{kk} f_k.$$

Тоді, як відомо, $g_1 \wedge ... \wedge g_k = \det(C) f_1 \wedge ... \wedge f_k$. Отже,

$$V(g_1,..., g_k) = |\det(C)| V(f_1,..., f_k) \qquad (4.3.14)$$

Тепер можна дати визначення, що в разі $\det(C) > 0$ системи векторів $f_1 \wedge ... \wedge f_k$ і $g_1 \wedge ... \wedge g_k$ орієнтовані однаково, якщо $\det(C) < 0$, орієнтації протилежні.

Теорема . Нехай $f_1,..., f_k$ - лінійно незалежна системи векторів в n-вимірному евклідовому просторі ($k < n$) і $g_1,..., g_k$ - інша система. Для того, щоб мала місце рівність

$$f_1 \wedge ... \wedge f_k = g_1 \wedge ... \wedge g_k \tag{4.3.15}$$

Необхідно і достатньо, щоб ці системи породжували один і той же підпростір, були в ньому однаково орієнтовані і щоб $V\left(f_1,..., f_k \right) = V\left(g_1,..., g_k \right)$.

Доведення. Необхідність. Нехай рівність виконано. $\forall i : 1 \le i \le k$

$f_1 \wedge ... \wedge f_k \wedge g_i = g_1 \wedge ... \wedge g_k \wedge g_i = 0$, отже, оскільки $f_1,..., f_k$ - лінійно незалежні, то $g_i = \sum_{j=1}^{k} c_{ij} f_j, i = 1,...,k$. Тоді $g_1 \wedge ... \wedge g_k = \det(C) f_1 \wedge ... \wedge f_k$, тому $\det(C) = 1$, а це означає, що системи однаково орієнтовані і породжують один і той же підпростір. Оскільки $V\left(g_1,..., g_k\right) = \left|\det(C)\right| V\left(f_1,..., f_k \right)$, то об’єми рівні.

Достатність. Нехай виконано всі умови теореми. Доведемо рівність (). З умови випливає, що $g_i = \sum_{j=1}^{k} c_{ij} f_j, i = 1,...,k$, $\det(C) = 1$. Звідси випливає, що

$g_1 \wedge ... \wedge g_k = \det(C) f_1 \wedge ... \wedge f_k = f_1 \wedge ... \wedge f_k$.

Теорему доведено.

Проведені дослідження дають змогу інтерпретувати $f_1 \wedge ... \wedge f_k$ як орієнтовний об’єм паралелепіпеда, породженого векторами $f_1,..., f_k$.

Деякі важливі приклади і методи знаходження визначників.

При знаходження значень конкретних визначників, безумовно, використовуються всі доведені властивості, але інколи треба застосовувати деякі нестандартні способи. Покажемо це на прикладах.

Приклад 1. Визначник Вандермонда та його застосування.
Так називається визначник вигляду:

$$\Delta_n(x_1,...,x_n)=\begin{vmatrix}1 & x_1 & x_1^2 & ... & x_1^{n-1}\\ 1 & x_2 & x_2^2 & ... & x_2^{n-1}\\ . & . & . & ... & .\\ 1 & x_n & x_n^2 & ... & x_n^{n-1}\end{vmatrix} \qquad (4.3.16)$$

Покажемо, як знайти цей визначник, використовуючи рекурентні співвідношення.
Спочатку підрахуємо його при n=2. Маємо:

$$\Delta_2(x_1,x_2)=\begin{vmatrix}1 & x_1\\ 1 & x_2\end{vmatrix}=x_2-x_1 \qquad (4.3.17)$$

Обчислюючи тепер цей визначник при n=3:

$$\Delta_3(x_1,x_2,x_3)=\begin{vmatrix}1 & x_1 & x_1^2\\ 1 & x_2 & x_2^2\\ 1 & x_3 & x_3^2\end{vmatrix}.$$

віднімемо від третього стовпчика другий, помножимо на x_1, від другого – перший, помножений на x_1. Одержимо

$$\Delta_3(x_1,x_2,x_3)=\begin{vmatrix}1 & 0 & 0\\ 1 & x_2-x_1 & x_2{}^2-x_2x_1\\ 1 & x_3-x_1 & x_3{}^2-x_3x_1\end{vmatrix}=\begin{vmatrix}x_2-x_1 & x_2\left(x_2-x_1\right)\\ x_3-x_1 & x_3\left(x_3-x_1\right)\end{vmatrix}=$$

$$=\left(x_2-x_1\right)\left(x_3-x_1\right)\Delta_2(x_1,x_2)=\left(x_2-x_1\right)\left(x_3-x_1\right)\left(x_3-x_2\right).$$

Методом математичної індукції доведемо формулу:

$$\Delta_n(x_1,...,x_n)=\prod_{1\le j<i\le n}\left(x_i-x_j\right) \qquad (4.3.18)$$

Нехай формула справедлива для визначників порядку n-1. У визначнику порядку n віднімемо від кожного стовпчика попередній, помножений на x_1:

$$\Delta_n(x_1,x_2,...,x_n)=\begin{vmatrix} 1 & 0 & 0 & & 0 \\ 1 & x_2\text{-}x_1 & x_2^2-x_2x_1 & ... & x_2^{n-1}-x_2^{n-2}x_1 \\ 1 & x_3\text{-}x_1 & x_3^2-x_3x_1 & ... & x_3^{n-1}-x_3^{n-2}x_1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n\text{-}x_1 & x_n^2-x_nx_1 & ... & x_n^{n-1}-x_n^{n-2}x_1 \end{vmatrix}=$$

$$=(x_2-x_1)(x_3-x_1)...(x_n-x_1)\Delta_{n-1}(x_2,x_3,...,x_n)$$

Використання припущення індукції

$$\Delta_{n-1}(x_1,...,x_{n-1})=(x_2-x_1)(x_3-x_1)...(x_{n-1}-x_1)=\prod_{1\le j<i\le n-1}(x_i-x_j)$$

доводить формулу.

Дамо одне з застосувань визначника Вандермонда. Треба побудувати многочлен, який в точках $x_1,x_2,...,x_n$ набуває значення $y_1,y_2,...,y_n$. Усі значення аргументів вважаються попарно різними. Многочлен, що має n коефіцієнтів, має степінь n-1:

$$y=a_0+a_1x+...+a_{n-1}x^{n-1}\equiv f(x). \tag{4.3.19}$$

З умови випливає використання рівностей

$$\begin{aligned} &y_1=a_0+a_1x_1+...+a_{n-1}x_1^{n-1},\\ &y_2=a_0+a_1x_2+...+a_{n-1}x_2^{n-1},\\ &\dots\dots\dots\dots\dots\dots\\ &y_n=a_0+a_1x_n+...+a_{n-1}x_n^{n-1}. \end{aligned} \tag{4.3.20}$$

Одержали систему n лінійних рівнянь з n невідомими $a_0,a_1,...,a_{n-1}$. Її визначник – визначник Вандермонда:

$$\Delta_n(x_1,...,x_n)=\prod_{1\le j<j\le n}(x_i-x_j) \tag{4.3.21}$$

Для знаходження невідомих коефіцієнтів треба розв‘язати систему (4.3.21). Але ми не будемо робити цього безпосередньо, а використаємо метод розкладу правильного раціонального дробу на елементарні. Нехай $F(x)=(x-x_1)(x-x_2)...(x-x_n)$. Розглянемо дріб $\dfrac{f(x)}{F(x)}$(многочлен в знаменнику має прості корені), доведемо формулу розкладу її на елементарні дроби:

$$\frac{f(x)}{F(x)} = \sum_{k=1}^{n} \frac{f(x_k)}{F'(x_k)(x - x_k)}. \qquad (4.3.22)$$

З цієї формули одержимо інтерполяційну формулу Лагранжа:

$$f(x) = \sum_{k=1}^{n} \frac{y_k(x - x_1)...(x - x_{k-1})(x - x_{k+1})...(x - x_n)}{(x_k - x_1)...(x_k - x_{k-1})(x_k - x_{k+1})...(x_k - x_n)}. \qquad (4.3.23)$$

Для доведення (4.3.23) розкладемо $\frac{f(x)}{F(x)}$ в суму елементарних дробів з невизначеними коефіцієнтами:

$$\frac{f(x)}{(x - x_1)(x - x_2)...(x - x_n)} = \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + ... + \frac{A_n}{x - x_n}.$$

Помноживши на знаменник, одержимо:

$$f(x) = A_1(x - x_2)...(x - x_n) + A_2(x - x_1)(x - x_3)...(x - x_n) + ... + A_n(x - x_1)...(x - x_{n-1}).$$

Підставляючи по черзі $x = x_1$, $x = x_2$, $x = x_n$, одержимо

$$\begin{aligned} &f(x_1) = A_1(x_1 - x_2)(x_1 - x_3)...(x_1 - x_n), \\ &f(x_2) = A_2(x_2 - x_1)(x_2 - x_3)...(x_2 - x_n), \\ &.. \\ &f(x_n) = A_n(x_n - x_1)(x_n - x_2)...(x_n - x_{n-1}). \end{aligned}$$

Співмножники при цих коефіцієнтах у правих частинах відмінні від нуля і зображуються за допомогою похідної многочлена $F(x)$. Дійсно,

$$\begin{aligned} &F'(x) = \\ &= (x - x_2)...(x - x_n) + (x - x_1)(x - x_3)...(x - x_n) + ... + (x - x_1)...(x - x_{n-1}). \end{aligned} \qquad (4.3.24)$$

Підставляючи по черзі $x = x_1$, $x = x_2$, $x = x_n$, одержимо

$$\begin{aligned} &F'(x_1) = (x_1 - x_2)(x_1 - x_3)...(x_1 - x_n), \\ &F'(x_2) = (x_2 - x_1)(x_2 - x_3)...(x_2 - x_n), \\ &.. \\ &F'(x_n) = (x_n - x_1)(x_n - x_2)...(x_n - x_{n-1}). \end{aligned}$$

З двох систем рівностей маємо

$$A_1 = \frac{f(x_1)}{F'(x_1)},\ \ A_2 = \frac{f(x_2)}{F'(x_2)},\ \dots, A_n = \frac{f(x_n)}{F'(x_n)}, \tag{4.3.25}$$

що й доводить (4.3.23).

Приклад 2. Визначник, що зводиться до визначника Вандермонда.

$$D_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & x_2^n \\ . & . & . & \dots & \dots & . \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & x_n^n \end{vmatrix}.$$

Використаємо дещо штучний прийом. Нехай $x_1, x_2, \dots, x_n$ - корені многочлена степені *n*: $x^n + a_1 x^{n-1} + \dots + a_{n-1}x + a_n = 0$. Згідно з формулою Вієта $a_1 = -\left(x_1 + x_2 + \dots + x_n\right)$. Звідси $x^n = \left(x_1 + x_2 + \dots + x_n\right)x^{n-1} - \dots - a_{n-1}x - a_n$. Підставляючи в цю рівність $x = x_k, k = 1,\dots,n$, будемо мати: $x_k^n = \left(x_1 + x_2 + \dots + x_n\right)x_k^{n-1} - a_2 x_k^{n-2} - \dots - a_{n-1}x_k - a_n$. Після заміни елементів останнього стовпчика такими виразами, будемо мати:

$$D_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & \left(x_1 + \dots + x_n\right)x_1^{n-1} - a_2 x_1^{n-2} - \dots - a_{n-1}x_1 - a_n \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & \left(x_1 + \dots + x_n\right)x_2^{n-1} - a_2 x_2^{n-2} - \dots - a_{n-1}x_2 - a_n \\ . & . & . & \dots & \dots & . \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & \left(x_1 + \dots + x_n\right)x_n^{n-1} - a_2 x_n^{n-2} - \dots - a_{n-1}x_n - a_n \end{vmatrix} =$$

$$= \left(x_1 + \dots + x_n\right)\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & x_2^{n-1} \\ . & . & . & \dots & \dots & . \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & x_n^{n-1} \end{vmatrix} + \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & -a_2 x_1^{n-2} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & -a_2 x_2^{n-2} \\ . & . & . & \dots & \dots & . \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & -a_2 x_n^{n-2} \end{vmatrix} +$$

$$+ \dots + \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & -a_{n-1}x_1 - a_n \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & -a_{n-1}x_2 - a_n \\ . & . & . & \dots & \dots & . \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & -a_{n-1}x_n - a_n \end{vmatrix} + \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & -a_n \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & -a_n \\ . & . & . & \dots & \dots & . \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & -a_n \end{vmatrix} =$$

$$=\left(x_1+\ldots+x_n\right)\begin{vmatrix}1 & x_1 & x_1^2 & \ldots & x_1^{n-2} & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-2} & x_2^{n-1}\\ . & . & . & \ldots & \ldots & .\\ 1 & x_n & x_n^2 & \ldots & x_n^{n-2} & x_n^{n-1}\end{vmatrix}=\left(x_1+\ldots+x_n\right)\prod_{1\le j<i\le n-1}\left(x_i-x_j\right).$$

В попередньому виразі всі визначники крім першого дорівнюють нулю, оскільки вони мають пропорційні стовпчики. Визначник, що залишився, – це визначник Вандермонда, значення якого було знайдено в попередньому прикладі.

Приклад 3. Виділення лінійних множників у визначнику, який розлядається як многочлен від кількох змінних.

$$\Delta=\begin{vmatrix}0 & x & y & z\\ x & 0 & z & y\\ y & z & 0 & x\\ z & y & x & 0\end{vmatrix}.$$

Очевидно, якщо розкрити цей визначник, одержимо однорідний многочлен четвертої степені від трьох змінних, але для його знаходження не будемо використовувати прямі методи. Досить знайти всі лінійні дільники цього многочлена. Тоді він буде с точністю до константи буде дорівнювати їхньому добутку. Потім знайдемо цю константу, прирівнюючи коефіцієнти при старшій степені однієї з змінних або підставляючи замість його змінних певні числа, вибираючи їх таким чином, щоб визначник легко обчислювався.
Додамо до першого рядка другий, третій та четвертий рядки. Визначник від цього не зміниться:

$$\Delta=\begin{vmatrix}x+y+z & x & y & z\\ x+y+z & 0 & z & y\\ x+y+z & z & 0 & x\\ x+y+z & y & x & 0\end{vmatrix}=\left(x+y+z\right)\begin{vmatrix}1 & x & y & z\\ 1 & 0 & z & y\\ 1 & z & 0 & x\\ 1 & y & x & 0\end{vmatrix}.$$

Отже, один лінійний множник знайдено: $x+y+z$.
Тепер до першого стовпчика додамо другий і віднімемо третій та четвертий:

$$\Delta=\begin{vmatrix}0 & x & y & z\\ x & 0 & z & y\\ y & z & 0 & x\\ z & y & x & 0\end{vmatrix}=\begin{vmatrix}x-y-z & x & y & z\\ x-y-z & 0 & z & y\\ -x+y+z & z & 0 & x\\ -x+y+z & y & x & 0\end{vmatrix}=(y+z-x)\begin{vmatrix}-1 & x & y & z\\ -1 & 0 & z & y\\ 1 & z & 0 & x\\ 1 & y & x & 0\end{vmatrix}.$$

Далі за допомогою аналогічних дій знаходимо дільники $z+x-y$ та $x+y-z$. Радимо читачу зробити це самостійно. Добуток знайдених дільників – многочлен четвертої степені відносно трьох змінних. Отже,

$$\Delta=a\left(x+y+z\right)(y+z-x)\left(z+x-y\right)\left(x+y-z\right).$$

Старший його член, що містить четверту степінь змінної *x*, має вигляд $-ax^4$. Виходячи з означення визначника, формули (4.1.4), помічаємо, що підстановка індексів елементів нашого визначника, що дорівнюють *x*, має вигляд $\sigma=\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 1 & 4 & 3\end{pmatrix}$. З неї можна утворити тотожню підстановку, зробивши дві транспозиції сусідніх елементів: першого і другого та третього і четвертого, отже її парність дорівнює нулю, і коефіцієнт перед x^4 буде дорівнювати одиниці, з чого випливає, що $a=-1$. Отже,
$\Delta=-\left(x+y+z\right)(y+z-x)\left(z+x-y\right)\left(x+y-z\right)$.
Цікаво відзначити, що при виконанні умов $y+z>x>0, z+x>y>0, x+y>z>0$ величини x,y,z є сторонами деякого трикутника, і з формули Герона випливає, що квадрат його площі можна виразити через знайдений визначник:

$$S^2=\frac{x+y+z}{2}\frac{y+z-x}{2}\frac{z+x-y}{2}\frac{x+y-z}{2}=-\frac{\Delta}{16}.$$

Приклад 4. Розглянемо ще один приклад, де застосовується методика прикладу 3.

$$\Delta=\begin{vmatrix}a_0 & a_1 & a_2 & \dots & a_n\\ a_0 & x & a_2 & \dots & a_n\\ a_0 & a_1 & x & \dots & a_n\\ . & . & . & \dots & \\ a_0 & a_1 & a_2 & \dots & x\end{vmatrix}.$$

Якщо підставляти по черзі $x=a_1, x=a_2,\dots,x=a_n$, одержимо кожного разу визначник з двома однаковими стовпчиками, і він дорівнює нулю. Отже, наш визначник, який є многочленом від *x* степені *n*, може бути розкладений на

множники: $\Delta = k\left(x-a_1\right)\left(x-a_2\right)...\left(x-a_n\right)$. Порівнюючи коефіцієнти при x^n в розкладі визначника і після множення виразів у дужках в останньому многочлені, робимо висновок: $k = a_0$.
Відповідь: $\Delta = a_0\left(x-a_1\right)\left(x-a_2\right)...\left(x-a_n\right)$.

Приклад 5. Визначник кососиметричної матриці.
Так називається матриця, у якої $a_{ji} = -a_{ij}$. Звідси випливає, що $a_{ii} = -a_{ii} \Rightarrow 2a_{ii=0} \Rightarrow a_{ii} = 0$, тобто, її діагональні елементи дорівнюють нулю. Нехай матриця *A* має непарний порядок $2n-1$. Знайдемо її визначник:

$$\Delta(A)=\begin{vmatrix} 0 & \alpha & \beta & ... & \omega \\ -\alpha & 0 & \gamma & ... & . \\ -\beta & -\gamma & 0 & ... & . \\ . & . & . & ... & . \\ -\omega & . & . & ... & 0 \end{vmatrix}.$$

Цей визначник за відомою властивістю дорівнює визначнику транспонованої матриці:

$$\Delta\left(A\right)=\Delta\left(A^T\right)=\begin{vmatrix} 0 & -\alpha & -\beta & ... & -\omega \\ \alpha & 0 & -\gamma & ... & . \\ \beta & \gamma & 0 & ... & . \\ . & . & . & ... & . \\ \omega & . & . & ... & 0 \end{vmatrix}=(-1)^{2n-1}\Delta\left(A\right)=-\Delta\left(A\right),$$

Звідки випливає: $\Delta\left(A\right)=0$.

Приклад 6.
Зображення визначника у вигляді суми визначників за формулою (4.1.8).

$$\Delta_n = \begin{vmatrix} a_1+b_1 & a_1+b_2 & ... & a_1+b_n \\ a_2+b_1 & a_2+b_2 & ... & a_2+b_n \\ . & . & ... & . \\ a_n+b_1 & a_n+b_2 & ... & a_n+b_n \end{vmatrix}.$$

Розкладаючи визначник за цією формулою по першому рядку, потім по другому, будемо мати:

$$\Delta_n = \begin{vmatrix} a_1 & a_1 & \dots & a_1 \\ a_2+b_1 & a_2+b_2 & \dots & a_2+b_n \\ . & . & \dots & . \\ a_n+b_1 & a_n+b_2 & \dots & a_n+b_n \end{vmatrix} + \begin{vmatrix} b_1 & b_2 & \dots & b_n \\ a_2+b_1 & a_2+b_2 & \dots & a_2+b_n \\ . & . & \dots & . \\ a_n+b_1 & a_n+b_2 & \dots & a_n+b_n \end{vmatrix} =$$

$$= \begin{vmatrix} a_1 & a_1 & \dots & a_1 \\ a_2 & a_2 & \dots & a_2 \\ . & . & \dots & . \\ a_n+b_1 & a_n & \dots & a_n+b_n \end{vmatrix} + \begin{vmatrix} a_1 & a_1 & \dots & a_1 \\ b_1 & b_2 & \dots & b_n \\ . & . & \dots & . \\ a_n+b_1 & a_n+b_2 & \dots & a_n+b_n \end{vmatrix} +$$

$$+ \begin{vmatrix} b_1 & b_2 & \dots & b_n \\ a_2 & a_2 & \dots & a_2 \\ . & . & \dots & . \\ a_n+b_1 & a_n+b_2 & \dots & a_n+b_n \end{vmatrix} + \begin{vmatrix} b_1 & b_2 & \dots & b_n \\ b_1 & b_2 & \dots & b_n \\ . & . & \dots & . \\ a_n+b_1 & a_n+b_2 & \dots & a_n+b_n \end{vmatrix} = \dots$$

Застосовуючи ту ж формулу і далі, прийдемо до суми 2^n визначників, рядки яких будуть мати вигляд $\left(a_k \quad a_k \quad \dots \quad a_k\right), k=1,\dots,n$ або $\left(b_1 \quad b_2 \quad \dots \quad b_n\right)$. Очевидно, при $n \ge 3$ кожен з таких визначників буде мати не менше двох рядків першого або другого типів, причому, перші пропорційні, а другі рівні, отже, всі ці визначники дорівнюватимуть нулю. Слід окремо розглянути визначник 2-го порядку:

$$\Delta_2 = \begin{vmatrix} a_1+b_1 & a_1+b_2 \\ a_2+b_1 & a_2+b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_1 \\ a_2+b_1 & a_2+b_2 \end{vmatrix} + \begin{vmatrix} b_1 & b_2 \\ a_2+b_1 & a_2+b_2 \end{vmatrix} =$$

$$\Delta_2 = \begin{vmatrix} a_1+b_1 & a_1+b_2 \\ a_2+b_1 & a_2+b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_1 \\ a_2+b_1 & a_2+b_2 \end{vmatrix} + \begin{vmatrix} b_1 & b_2 \\ a_2+b_1 & a_2+b_2 \end{vmatrix} =$$

$$= \begin{vmatrix} a_1 & a_1 \\ a_2 & a_2 \end{vmatrix} + \begin{vmatrix} a_1 & a_1 \\ b_1 & b_2 \end{vmatrix} + \begin{vmatrix} b_1 & b_2 \\ a_2 & a_2 \end{vmatrix} + \begin{vmatrix} b_1 & b_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & 0 \\ b_1 & b_2-b_1 \end{vmatrix} + \begin{vmatrix} b_1 & b_2-b_1 \\ a_2 & 0 \end{vmatrix} =$$

$$= \left(b_2-b_1\right)\left(\begin{vmatrix} a_1 & 0 \\ b_1 & 1 \end{vmatrix} + \begin{vmatrix} b_1 & 1 \\ a_2 & 0 \end{vmatrix}\right) = \left(a_1-a_2\right)\left(b_2-b_1\right).$$

Приклад 7. Метод зміни всіх елементів визначника.

Розглянемо ситуацію, коли додавання до всіх елементів визначника певного числа значно полегшує подальші обчислення.
Нехай треба обчислити визначник

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

Обчислимо визначник

$$\tilde{\Delta}=\begin{vmatrix} a_{11}+x & a_{12}+x & \dots & a_{1n}+x \\ a_{21}+x & a_{22}+x & \dots & a_{2n}+x \\ . & . & \dots & . \\ a_{n1}+x & a_{n2}+x & \dots & a_{nn}+x \end{vmatrix}.$$

Як і в минулому прикладі будемо застосовувати формулу (4.1.8). В результаті прийдемо до 2^n визначників, елементи яких не будуть мати вигляд сум типу $a_{ij}+x$. Всі визначники, що містять не менше, як два рядки вигляду $\begin{pmatrix} x & x & \dots & x \end{pmatrix}$, дорівнюють нулю. Залишаться визначники, в яких є тільки один рядок такого вигляду, причому, в якомусь визначнику він буде перший, в іншому – другий і, нарешті, *n*-ий. Розкладемо кожен такий визначник по рядку такому рядку, виносячи *x* за дужки. При розкладі будемо мати суму алгебраїчних доповнень до елементів першого рядка, потім другого і, нарешті, *n*-го. Позначаючи ці алгебраїчні доповнення через A_{ij}, одержимо:

$$\tilde{\Delta}=\Delta+x\sum_{i,j=1}^{n}A_{ij}.$$

Цю формулу варто застосовувати, якщо зручно обчислюються алгебраїчні доповнення.

Проілюструємо можливість використання одержаної формули на прикладі.

$$\tilde{\Delta}=\begin{vmatrix} a_1 & x & x & \dots & x \\ x & a_2 & x & \dots & x \\ x & x & a_3 & \dots & x \\ . & . & . & \dots & . \\ x & x & x & \dots & a_n \end{vmatrix}.$$

Якщо до кожного його елемента додати $-x$, одержимо

$$\Delta=\begin{vmatrix} a_1-x & 0 & 0 & \dots & 0 \\ 0 & a_2-x & 0 & \dots & 0 \\ 0 & 0 & a_3-x & \dots & 0 \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & a_n-x \end{vmatrix}=(a_1-x)(a_2-x)\dots(a_n-x).$$

З одержаної формули $\tilde{\Delta}=(a_1-x)(a_2-x)\dots(a_n-x)+x\sum_{i,j=1}^{n}A_{ij}$. Підкреслимо: A_{ij} - алгебраїчні доповнення більш простого за формою визначника Δ. Алгебраїчні доповнення його кожного недіагонального елемента дорівнюють нулю (перевірте!). Алгебраїчні ж доповнення кожного його діагонального елемента дорівнюють добутку всіх інших елементів головної діагоналі. Отже,

$$\tilde{\Delta}=(a_1-x)(a_2-x)\dots(a_n-x)+x\sum_{i=1}^{n}(a_1-x)\dots(a_{i-1}-x)(a_{i+1}-x)\dots(a_n-x)=$$

$$=(a_1-x)(a_2-x)\dots(a_n-x)+x\sum_{i=1}^{n}(a_1-x)\dots(a_i-x)\dots(a_n-x)\frac{1}{a_i-x}=$$

$$=x(a_1-x)(a_2-x)\dots(a_n-x)\left(\frac{1}{x}+\frac{1}{a_1-x}+\frac{1}{a_2-x}+\dots+\frac{1}{a_n-x}\right).$$

Приклад 8. Трьохдіагональні матриці.

$A\in M_n$.

$$\begin{pmatrix} a_{11} & a_{12} & 0 & . & . & . & 0 \\ a_{21} & a_{22} & a_{23} & . & . & . & 0 \\ 0 & a_{32} & a_{33} & a_{34} & . & . & 0 \\ . & . & . & . & . & & . \\ . & . & & . & . & . & . \\ 0 & . & & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & . & . & . & 0 & a_{n,n-1} & a_{nn} \end{pmatrix}$$

Знайдемо рекурентну формулу для визначників трьохдіагональної матриці. Такий визначник зручно позначити $\det\left(A(1,...,k+1)\right)$.

$$\det\left(A(1,...,k+1)\right) = \begin{vmatrix} a_{11} & a_{12} & 0 & . & . & . & 0 \\ a_{21} & a_{22} & a_{23} & 0 & . & . & 0 \\ 0 & a_{32} & a_{33} & a_{34} & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & ... & 0 & a_{k-1,k-2} & a_{k-1,k-1} & a_{k-1,k} & 0 \\ 0 & ... & & 0 & a_{k,k-1} & a_{k,k} & a_{k,k+1} \\ 0 & . & . & . & 0 & a_{k+1,k} & a_{k+1,k+1} \end{vmatrix} =$$

$$= a_{k+1,k+1} \begin{vmatrix} a_{11} & a_{12} & 0 & . & . & . & 0 \\ a_{21} & a_{22} & a_{23} & 0 & . & . & 0 \\ 0 & a_{32} & a_{33} & a_{34} & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & ... & 0 & a_{k-2,k-3} & a_{k-2,k-2} & a_{k-2,k-1} & 0 \\ 0 & ... & & 0 & a_{k-1,k-2} & a_{k-1,k-1} & a_{k-1,k} \\ 0 & . & . & . & 0 & a_{k,k-1} & a_{k,k} \end{vmatrix} -$$

$$-a_{k,k+1} \begin{vmatrix} a_{11} & a_{12} & 0 & . & . & . & 0 \\ a_{21} & a_{22} & a_{23} & 0 & . & . & 0 \\ 0 & a_{32} & a_{33} & a_{34} & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & ... & 0 & a_{k-2,k-3} & a_{k-2,k-2} & a_{k-2,k-1} & 0 \\ 0 & ... & & 0 & a_{k-1,k-2} & a_{k-1,k-1} & a_{k-1,k} \\ 0 & . & . & . & 0 & 0 & a_{k+1,k} \end{vmatrix} =$$

$$= a_{k+1,k+1} \det\left(A(1,...,k)\right) - a_{k+1,k}\, a_{k,k+1} \det\left(A(1,...,k-1)\right).$$

Приклад 9. Розглянемо матрицю

$$A=\begin{vmatrix} a_{n-1} & 1 & 0 & & & ... & 0 \\ a_{n-2} & 0 & 1 & 0 & & ... & 0 \\ a_{n-3} & 0 & 0 & 1 & & & 0 \\ . & . & . & . & . & ... & . \\ a_1 & 0 & & & & 0 & 1 \\ a_0 & 0 & . & . & . & . & 0 \end{vmatrix}.$$

Розглянемо визначник $\chi(\lambda)=\det(\lambda I-A)$. Такі визначники будуть вивчатись у главі 5. $\det(\lambda I-A)$ називається характеристичним многочленом матриці *A*.

$$\chi(\lambda)=\begin{vmatrix} \lambda-a_{n-1} & -1 & 0 & & & ... & 0 \\ -a_{n-2} & \lambda & -1 & 0 & & ... & 0 \\ -a_{n-3} & 0 & \lambda & -1 & & & 0 \\ . & . & . & . & . & ... & . \\ -a_1 & 0 & & & & \lambda & -1 \\ -a_0 & 0 & . & . & . & . & \lambda \end{vmatrix}=$$

$$=\begin{vmatrix} \lambda & -1 & 0 & & & ... & 0 \\ 0 & \lambda & -1 & 0 & & ... & 0 \\ 0 & 0 & \lambda & -1 & & & 0 \\ . & . & . & . & . & ... & . \\ 0 & 0 & & & & \lambda & -1 \\ 0 & 0 & . & . & . & . & \lambda \end{vmatrix}-\begin{vmatrix} a_{n-1} & -1 & 0 & & & ... & 0 \\ a_{n-2} & \lambda & -1 & 0 & & ... & 0 \\ a_{n-3} & 0 & \lambda & -1 & & & 0 \\ . & . & . & . & . & ... & . \\ a_1 & 0 & & & & \lambda & -1 \\ a_0 & 0 & . & . & . & . & \lambda \end{vmatrix}=$$

$$=\lambda^n-a_{n-1}\begin{pmatrix} \lambda & -1 & & ... & 0 \\ 0 & \lambda & -1 & ... & 0 \\ . & . & . & ... & . \\ 0 & & & ... & -1 \\ 0 & . & . & 0 & \lambda \end{pmatrix}+a_{n-2}\begin{pmatrix} -1 & 0 & & ... & 0 \\ 0 & \lambda & -1 & ... & 0 \\ . & . & . & ... & . \\ 0 & & & ... & -1 \\ 0 & . & . & 0 & \lambda \end{pmatrix}-$$

$$-a_{n-3}\begin{vmatrix} -1 & 0 & & & ... & 0 \\ 0 & -1 & -1 & & ... & 0 \\ 0 & 0 & \lambda & -1 & ... & 0 \\ . & . & . & . & ... & \\ & & & & ... & -1 \\ 0 & & & & & \lambda \end{vmatrix}-\ldots-(-1)^{n+1}a_0\begin{vmatrix} -1 & 0 & ... & & & 0 \\ \lambda & -1 & ... & & & 0 \\ . & . & ... & . & . & . \\ 0 & & ... & -1 & 0 & 0 \\ 0 & & ... & \lambda & -1 & 0 \\ 0 & & & & \lambda & -1 \end{vmatrix}=$$

$$= \lambda^n - a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - ... - a_1\lambda - a_0.$$

В главі 5 ми побачимо, що визначник

$$\chi(\lambda) = \lambda^n - a_{n-1}\lambda^{n-1} - a_{n-2}\lambda^{n-2} - ... - a_1\lambda - a_0$$

є так званим характеристичним многочленом матриці

4.4. ФУНКЦІОНАЛЬНІ ВИЗНАЧНИКИ

Розглянемо визначник, елементи якого — диференційовні функції однієї незалежної змінної x:

$$\Delta_n(x) = \begin{vmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ . & . & \cdots & . \\ f_{n1}(x) & f_{21}(x) & \cdots & f_{nn}(x) \end{vmatrix} \tag{4.4.1}$$

Очевидно, що визначник також буде диференційовною функцією аргументу x. Якщо значення функції — елементи грасманової алгебри, то їх можна розкласти за базисом алгебри. Коефіцієнти цього розкладу будуть скалярними функціями. За означенням будемо вважати, що для диференціювання функції зі значеннями у грасмановій алгебрі треба продиференціювати коефіцієнти цього розкладу і результати додати.

Теорема 4.4.1. Похідна визначника, елементами якого є диференційовні функції, дорівнює сумі n визначників, кожен з яких можна одержати з даного шляхом диференціювання відповідного рядка:

$$\Delta'(x) =$$

$$= \begin{vmatrix} f'_{11}(x) & \cdots & f'_{1n}(x) \\ f_{21}(x) & \cdots & f_{2n}(x) \\ . & \cdots & . \\ f_{n1}(x) & \cdots & f_{nn}(x) \end{vmatrix} + \begin{vmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ f'_{21}(x) & \cdots & f'_{2n}(x) \\ . & \cdots & . \\ f_{n1}(x) & \cdots & f_{nn}(x) \end{vmatrix} + \begin{vmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & \cdots & f_{2n}(x) \\ . & \cdots & . \\ f'_{n1}(x) & \cdots & f'_{nn}(x) \end{vmatrix} \tag{4.4.2}$$

Перевіримо цю властивість при n=2:

$$\Delta_n'(x)=\begin{vmatrix} f_{11}(x) & f_{12}(x) \\ f_{21}(x) & f_{22}(x) \end{vmatrix}'=\left(f_{11}(x)f_{22}(x)-f_{12}(x)f_{21}(x)\right)'=$$

$$=\left(f_{11}'(x)f_{22}(x)-f_{12}'(x)f_{21}(x)\right)+\left(f_{11}(x)f_{22}'(x)-f_{12}(x)f_{21}'(x)\right)=$$

$$=\begin{vmatrix} f_{11}'(x) & f_{12}(x) \\ f_{21}(x) & f_{22}(x) \end{vmatrix}+\begin{vmatrix} f_{11}(x) & f_{12}(x) \\ f_{21}'(x) & f_{22}'(x) \end{vmatrix}.$$

Нехай формулу доведено для визначників (n-1)-го порядку. Доведемо її для визначників n-го порядку.

$$\Delta'(x)=\left(f_{11}(x)\begin{vmatrix} f_{22}(x) & \dots & f_{2n}(x) \\ . & \dots & . \\ f_{n2}(x) & \dots & f_{nn}(x) \end{vmatrix}\right)'+\ldots+(-1)^n\left(f_{1n}(x)\begin{vmatrix} f_{21}(x) & \dots & f_{2,n-1}(x) \\ . & \dots & . \\ f_{n2}(x) & \dots & f_{n,n-1}(x) \end{vmatrix}\right)'=$$

$$=\left(f_{11}'(x)\begin{vmatrix} f_{22}(x) & \dots & f_{2n}(x) \\ . & \dots & . \\ f_{n2}(x) & \dots & f_{nn}(x) \end{vmatrix}+\ldots+(-1)^n f_{1n}'(x)\begin{vmatrix} f_{21}(x) & \dots & f_{2,n-1}(x) \\ . & \dots & . \\ f_{n2}(x) & \dots & f_{n,n-1}(x) \end{vmatrix}\right)+$$

$$+\left(f_{11}(x)\begin{vmatrix} f_{22}(x) & \dots & f_{2n}(x) \\ . & \dots & . \\ f_{n2}(x) & \dots & f_{nn}(x) \end{vmatrix}'+\ldots+(-1)^n\begin{vmatrix} f_{21}(x) & \dots & f_{2,n-1}(x) \\ . & \dots & . \\ f_{n2}(x) & \dots & f_{n,n-1}(x) \end{vmatrix}'\right).$$

Вираз у першій дужці – це визначник, у якому в першому рядку знаходяться похідні даного визначника, а всі інші їхні рядки співпадають; для визначників в другій сумі справедлива формула диференціювання. Тобто, кожна похідна – це сума відповідних визначників. Перегруповуючи доданки всіх сум, помічаємо, що одержали розклади за першим рядком визначників, у яких рядки, починаючи з другого, - це рядки з похідних даного визначника. Отже, формулу доведено.

Визначник Вронського та його похідна. Визначник Вронського n диференційовних функцій $f_1(x), f_2(x),..., f_n(x)$, має вигляд

$$W(x)=\begin{vmatrix} f_1(x) & f_2(x) & \dots & f_n(x) \\ f_1'(x) & f_2'(x) & \dots & f_n'(x) \\ . & . & \dots & . \\ f_1^{n-2}(x) & f_2^{n-2}(x) & \dots & f_n^{n-2}(x) \\ f_1^{n-1}(x) & f_2^{n-1}(x) & \dots & f_n^{n-1}(x) \end{vmatrix}, \tag{4.4.3}$$

Зауваження. Інколи, щоб підкреслити, для якої системи функцій обчислюється визначник, вживають позначення $W\left(f_1(x), f_2(x),..., f_n(x)\right)$.

тобто в ньому елементами першого рядка є задані функції, а в кожному наступному — похідні відповідних елементів попереднього рядка.

Теорема 4.4.2. Похідна визначника Вронського – це визначник, перші n-1 рядки якого співпадають з відповідними рядками визначника Вронського, а останній – це продиференційований останній рядок визначника Вронського:

$$W'(x)=\begin{vmatrix} f_1(x) & f_2(x) & \dots & f_n(x) \\ f_1'(x) & f_2'(x) & \dots & f_n'(x) \\ . & . & \dots & . \\ f_1^{n-2}(x) & f_2^{n-2}(x) & \dots & f_n^{n-2}(x) \\ f_1^{n}(x) & f_2^{n}(x) & \dots & f_n^{n}(x) \end{vmatrix} \qquad (4.4.4)$$

Доведення випливає з формули (4.4.2), якщо помітити, що всі визначники, крім останнього, матимуть по парі рівних рядків і тому дорівнюватимуть нулю.

Теорема 4.4.3. Визначник Вронського від n функцій – однорідна степеню n функція, тобто, має місце співвідношення:

$$W(\varphi(x) f_1(x),\varphi(x) f_2(x),...,\varphi(x) f_n(x)) = \varphi^n(x)W(f_1(x), f_2(x),..., f_n(x)). \qquad (4.4.5)$$

$$W\left(\varphi(x) f_1(x),\varphi(x) f_2(x),...,\varphi(x) f_n(x)\right) =$$

$$=\begin{vmatrix} \varphi(x) f_1(x) & \varphi(x) f_2(x) & \dots & \varphi(x) f_n(x) \\ \left(\varphi(x) f_1(x)\right)' & \left(\varphi(x) f_2(x)\right)' & \dots & \left(\varphi(x) f_2(x)\right)' \\ . & . & \dots & . \\ \left(\varphi(x) f_1(x)\right)^{(n-1)} & \left(\varphi(x) f_2(x)\right)^{(n-1)} & \dots & \left(\varphi(x) f_n(x)\right)^{(n-1)} \end{vmatrix} =$$

$$= \varphi(x)\begin{vmatrix} f_1(x) & \dots & f_n(x) \\ \varphi'(x) f_1(x) + \varphi(x) f_1'(x) & \dots & \varphi'(x) f_n(x) + \varphi(x) f_n'(x) \\ \left(\varphi(x) f_1(x)\right)'' & \dots & \left(\varphi(x) f_n(x)\right)'' \\ . & \dots & . \\ \left(\varphi(x) f_1(x)\right)^{(n-1)} & \dots & \left(\varphi(x) f_n(x)\right)^{(n-1)} \end{vmatrix} =$$

$$= \varphi(x)\begin{vmatrix} f_1(x) & \dots & f_n(x) \\ \varphi'(x)\, f_1(x) & \dots & \varphi'(x)\, f_n(x) \\ \left(\varphi(x)\, f_1(x)\right)'' & \dots & \left(\varphi(x)\, f_n(x)\right)'' \\ . & \dots & . \\ \left(\varphi(x)\, f_1(x)\right)^{(n-1)} & \dots & \left(\varphi(x)\, f_n(x)\right)^{(n-1)} \end{vmatrix} +$$

$$+ \varphi(x)\begin{vmatrix} f_1(x) & \dots & f_n(x) \\ \varphi(x)\, f_1'(x) & \dots & \varphi(x)\, f_n'(x) \\ \left(\varphi(x)\, f_1(x)\right)'' & \dots & \left(\varphi(x)\, f_n(x)\right)'' \\ . & \dots & . \\ \left(\varphi(x)\, f_1(x)\right)^{(n-1)} & \dots & \left(\varphi(x)\, f_n(x)\right)^{(n-1)} \end{vmatrix}.$$

В першому доданку два перші рядки пропорційні, тому він дорівнює нулю. З другого винесемо спільний множник. В результаті

$$W\left(\varphi(x)\, f_1(x), \varphi(x)\, f_2(x), ..., \varphi(x)\, f_n(x)\right) = \varphi^2(x)\begin{vmatrix} f_1(x) & \dots & f_n(x) \\ f_1'(x) & \dots & f_n'(x) \\ \left(\varphi(x)\, f_1(x)\right)'' & \dots & \left(\varphi(x)\, f_n(x)\right)'' \\ . & \dots & . \\ \left(\varphi(x)\, f_1(x)\right)^{(n-1)} & \dots & \left(\varphi(x)\, f_n(x)\right)^{(n-1)} \end{vmatrix}$$

.

Нехай, діючи таким чином, прийшли до (k+1)-го рядка з похідними порядку k від його елементів. Використаємо формулу Лейбниця для j-го елемента в цьому рядку:

$$\left(\varphi(x)\, f_j(x)\right)^{(k)} = \left(\varphi(x)\right)^{(k)}\, f_j(x) + k\left(\varphi(x)\right)^{(k-1)}\, f_j'(x) + ... + k\varphi^{k-1}(x)\, f_j^{(k-1)}(x) + \varphi(x)\, f_j^{(k)}(x)$$

Визначник запишеться у вигляді суми визначників, причому тільки останній, з елементами $\varphi(x)\, f_j^{(k)}(x)$ у k-му рядку, не буде дорівнювати нулю, і з нього винесемо спільний множник $\varphi(x)$, а в кожному з попередніх знайдуться два пропорційні рядки, тому він буде дорівнювати нулю. Так дійдемо до останнього рядка, і в результаті за визначник буде винесено $\left(\varphi(x)\right)^n$. Теорему доведено.

Визначники Якобі (якобіани)

В рамках цього курсу немає можливості пояснювати, яким чином виникають наступні визначники. Ми обчислимо їх без інтерпретації, що виникає у математичному аналізі.

1). Якобіан у циліндричних координатах:

$$J=\begin{vmatrix}\cos\varphi & -\rho\sin\varphi & 0\\ \sin\varphi & \rho\cos\varphi & 0\\ 0 & 0 & 1\end{vmatrix}=\rho .$$

2). Якобіан у сферичних координатах:

$$J=\begin{vmatrix}\sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi\\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & \rho\sin\theta\cos\varphi\\ \cos\theta & -\rho\sin\theta & 0\end{vmatrix}=$$

$$=\ \rho^2\begin{vmatrix}\sin\theta\cos\varphi & \cos\theta\cos\varphi & -\sin\theta\sin\varphi\\ \sin\theta\sin\varphi & \cos\theta\sin\varphi & \sin\theta\cos\varphi\\ \cos\theta & -\sin\theta & 0\end{vmatrix}=$$

$$=\rho^2\left(\cos^2\theta\sin\theta\begin{vmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{vmatrix}+\sin^3\theta\begin{vmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{vmatrix}\right)=\rho^2\sin\theta .$$

4.5 ТЕОРІЯ ЛІНІЙНИХ СИСТЕМ АЛГЕБРАЇЧНИХ РІВНЯНЬ

Означення. Система рівнянь вигляду

$$\begin{aligned}&a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1;\\ &a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2;\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m\end{aligned}\qquad(4.5.1)$$

називається алгебраїчною лінійною неоднорідною. У випадку $b_i=0, i=1,\ldots,m$ вона називається однорідною. Якщо $m=n$, система називається квадратною.

Цілком зрозуміло, що якщо ввести позначення

$$A=\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, x=\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \in \mathbb{R}^n, b=\begin{pmatrix} b_1 \\ \dots \\ b_m \end{pmatrix} \in \mathbb{R}^m, \qquad (4.5.2)$$

цю систему можна записати у більш компактному матрично-векторному вигляді:

$$Ax = b. \qquad (4.5.3)$$

Однорідна система має вигляд $Ax = 0$.

Означення. Система називається сумісною, якщо вона має розв'язки.

Очевидно, однорідна система завжди має розв'язки: принаймні їй задовольняє нуль-вектор: $A0 = 0$. При деяких умовах вона може мати неєдиний розв'язок і тоді, як уже відомо, такі розв'язки утворюють підпростір $Ker(A)$ простору $\mathbb{R}^n$.

Означення. Розширеною матрицею системи алгебраїчних рівнянь називається матриця, що складається з векторів-стовпчиків матриці лівої частини і стовпчика з правої частини системи.

Рангом розширеної матриці називається найбільший з порядків серед всіх її ненульових мінорів. Це еквівалентно розмірності бази векторів-стовпчиків розширеної матриці.

Зрозуміло, що ранг розширеної матриці не менше рангу основної матриці системи.

Доведемо тепер основний результат теорії лінійних систем.

Теорема 4.2.1. (Кронекера-Капеллі). Для того, щоб лінійна система була сумісною, необхідно і достатньо, щоб ранги її основної і розширеної матриць співпадали.

Доведення. Нехай $a_1,\dots,a_n$ - вектори-стовпчики матриці системи, а b - вектор-стовпчик її правої частини. Запишемо систему у векторному вигляді:

$$x_1 a_1 + \dots + x_n a_n = b$$

Очевидно, для сумісності системи необхідно і досить, щоб вектор-стовпчик b був лінійною комбінацією векторів-стовпчиків $a_1,\dots,a_n$. Нехай система має розв'язки, тобто, існують $x_1,\dots,x_n$, при яких справедлива написана рівність. Звідси випливає, що ранг розширеної матриці не більший за ранг основної, але, як відомо, ранг розширеної матриці не менше рангу основної, отже, вони рівні. Навпаки, якщо вказані ранги однакові, то базис простору, натягнутого на вектори $a_1,\dots,a_n$ збігатиметься з базисом для простору, який натягнуто на $a_1,\dots,a_n,b$, а тому ці вектори лінійно залежні: $\alpha_1 a_1 + \dots + \alpha_n a_n + \beta b = 0$. При цьому $\beta \neq 0$, оскільки вектори $a_1,...,a_n$ лінійно незалежні. Отже, b можна подати як лінійну комбінацію векторів $a_1,...,a_n$.

Зауважимо, що взагалі система може мати або єдиний, або неєдиний розв'язок, або зовсім його не мати (якщо не виконано умови теореми Кронекера-Капеллі – це стосується неоднорідної системи). З'ясуємо, який вигляд має розв'язок системи у випадку його неєдиності. Частинним будемо називати будь-який розв'язок системи.
Розглянемо спочатку однорідні системи. Як відзначалось всяка така система сумісна, тому що має хоча б один, а саме, нульовий розв'язок. З'ясуємо, при яких умовах її розв'язок буде неєдиний. Існування $x \neq 0$ такого, що $Ax = 0$ означає, що вектори-стовпчики матриці системи лінійно залежні (необхідна і достатня умови), а це, в свою чергу, означає, що ранг матриці менше, ніж кількість стовпчиків або, іншими словами, менше кількості невідомих. Якщо ж система квадратна, при такій умові її визначник дорівнює нулю. Сформулюємо результат.

Теорема 4.2.2. Лінійна однорідна система має неєдиний розв'язок тоді і тільки тоді, коли ранг її матриці менше ніж розмірність вектора невідомих. У разі квадратної системи для цього необхідно і достатньо, щоб визначник матриці системи дорівнював нулю. ■

З доведеної теореми випливає, що квадратна однорідна система має єдиний розв'язок ($x = 0$) тоді і тільки тоді, коли $\det(A) \neq 0$.
Ще раз нагадаємо, що завдяки властивості лінійності матриці розв'язки однорідної системи утворюють підпростір $Ker(A)$ в $\mathbb{R}^n$. Виберемо в ньому деякий базис $g = \{g_1,...,g_d\}$. Тоді довільний розв'язок x_0 однорідної системи можна подати у вигляді лінійної комбінації цих базисних векторів:

$$x_0 = \sum_{i=1}^{d} \alpha_i g_i.$$

Означення. Базис підпростору $Ker(A)$ називається фундаментальною системою розв'язків однорідної системи.
Зауважимо, що існує безліч фундаментальних систем.

Теорема 4.2.3. Загальний розв'язок неоднорідної лінійної системи дорівнює сумі її частинного розв'язку і загального розв'язку відповідної однорідної системи (права частина дорівнює нулю).
Доведення. Нехай $\tilde{x}$ - деякий частинний розв'язок системи $Ax = b$, тобто, $A\tilde{x} = b$. Тоді $A(x - \tilde{x}) = b$, а це означає, що $x_0 = x - \tilde{x}$ - розв'язок однорідної системи. $x_0 \in Ker(A)$, в цьому ядрі є свій базис, по якому можна розкласти

довільний вектор, що йому належить, а це й відповідає загальному розв'язку однорідної системи. Отже, загальний розв'язок неоднорідої системи має вигляд: $x = \tilde{x} + x_0$.

Нехай – $\{g_1, ..., g_d\}$ - деяка фундаментальна система розв'язків однорідної системи. Тоді зображення загального розв'язку неоднорідної системи буде мати вигляд:

$$x = \tilde{x} + \sum_{k=1}^{d} \alpha_k g_k .$$

Ще раз пояснимо результат, доведений в попередньому розділі – Теорема 4.1.5 (Крамера). При $\det(A) \neq 0$ однорідна система має лише тривіальний розв'язок, і неоднорідна система має єдиний розв'язок. Його існування доведено в теоремі Крамера.

Застосуємо правило знаходження елементів оберненої матриці для розв'язку квадратної системи лінійних рівнянь при умові $\det(A) \neq 0$.

Нехай треба розв'язати неоднорідну систему лінійних рівнянь (4.5.3). Подіємо на обидві частини рівності матрицею A^{-1}. В результаті одержимо: $x = A^{-1}b$. Для знаходження j-ї компоненти невідомого вектора x або невідомої x_j треба, згідно з правилами множення матриці на вектор, j-й рядок матриці скалярно помножити на вектор-стовпчик b. У результаті одержимо

$$x_j = \left(\sum_{i=1}^{n} D_{ji} b_i \right) / \Delta. \qquad (4.5.4)$$

Вираз у чисельнику відповідає розкладу по i-му стовпчику визначника Δ_i, що має такий вигляд:

$$\Delta_j = \begin{vmatrix} a_{11} \; ... \; a_{1,j-1} & b_{1j} & a_{1,j+1} \; ... \; a_{1n} \\ a_{21} \; ... \; a_{2,j-1} & b_{2j} & a_{2,j+1} \; ... \; a_{2n} \\ .... & .... & .... \\ a_{n1} \; ... \; a_{n,j-1} & b_{nj} & a_{n,j+1} \; ... \; a_{nm} \end{vmatrix} \qquad (4.5.5)$$

Отже, маємо формулу для довільної компоненти вектора-розв'язку неоднорідної лінійної системи:

$$x_j = \frac{\Delta_j}{\Delta}, \; j = 1, ..., n. \qquad (4.5.6)$$

Вона називається формулою Крамера.

Метод Гауса розв'язання лінійних систем

Зіставимо системі (4.5.1) розширену матрицю:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ . & . & \dots & . & | & . \\ a_{m1} & a_{m2} & \dots & a_{mn} & | & b_m \end{pmatrix} \tag{4.5.7}$$

Виконуючи елементарні перетворення цієї матриці, які, нагадаємо, полягають в множенні на відповідні числа першого рядка, якщо діагональний елемент $a_{11} \neq 0$, і додавання до інших рядків, щоб під цим діагональним елементом в першому стовпчику зробити нулі. Якщо ж цей елемент дорівнює нулю, треба переставити цей рядок з деяким іншим рядком, в якому елемент першого рядка не дорівнює нулю (якщо такий існує), а потім діяти, як було вказано. Далі треба діяти так само з другим стовпчиком і т.ін. В результаті прийдемо до матриці трапецієвидної форми або до верхньотрикутної в залежності від розмірності матриці та її рангу. Якщо основна матриця системи квадратна і невироджена, її можна звести до діагональної форми з ненульовими елементами на діагоналі, і її розв'язки знаходяться зрозумілим чином:

$$\begin{pmatrix} d_{11} & 0 & \dots & 0 & | & \tilde{b}_1 \\ 0 & d_{22} & \dots & 0 & | & \tilde{b}_2 \\ . & . & \dots & . & | & . \\ 0 & 0 & \dots & d_{mn} & | & \tilde{b}_m \end{pmatrix}, \tag{4.5.7}$$

$$x_i = \frac{b_i}{d_{ii}}, i = 1, \dots, n. \tag{4.5.8}$$

Розглянемо сумісну систему (виконано умови теореми Кронекера-Капеллі) з матрицею $A(m \times n)$, з рангом r, причому $r \leq m < n$. Елементарними перетвореннями розширену матрицю приводимо до вигляду (ми зберегли ті ж самі позначення для елементів одержаної матриці для простоти запису):

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} & | & b_1 \\ 0 & a_{22} & \dots & a_{2r} & a_{2,r+1} & \dots & a_{2n} & | & b_2 \\ . & . & \dots & . & . & \dots & . & | & . \\ 0 & 0 & \dots & a_{rr} & a_{r,r+1} & \dots & a_{rn} & | & b_r \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & | & 0 \\ . & . & \dots & . & . & \dots & . & | & . \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & | & 0 \end{pmatrix} \tag{4.5.8}$$

Перенесемо доданки в кожному рівнянні, які містять невідомі $x_{r+1},...,x_n$. Останні тотожності $0=0$ можна не писати. В результаті одержимо:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} & | & b_1 - \sum\limits_{j=r+1}^{n} a_{1j}x_j \\ 0 & a_{22} & \dots & a_{2r} & | & b_2 - \sum\limits_{j=r+1}^{n} a_{2j}x_j \\ . & . & \dots & . & | & . \\ 0 & 0 & \dots & a_{rr} & | & b_r - \sum\limits_{j=r+1}^{n} a_{rj}x_j \end{pmatrix} \tag{4.5.9}$$

Елементарними перетворення приводимо основну матрицю систему до діагонального вигляду і розв’язуємо, як це було описано раніше. В цьому випадку перші *r* невідомих $x_1,...,x_r$ лінійно залежать від $x_{r+1},...,x_n$:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \\ x_{r+1} \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \beta_1 + \sum\limits_{i=r+1}^{n} \alpha_{1i} x_i \\ \beta_2 + \sum\limits_{i=r+1}^{n} \alpha_{2i} x_i \\ \dots \\ \beta_r + \sum\limits_{i=r+1}^{n} \alpha_{ri} x_i \\ x_{r+1} \\ \dots \\ x_n \end{pmatrix} \tag{4.5.10}$$

Це можна переписати як лінійну комбінацію частинного розв’язку неоднорідної системи і лінійної комбінації векторів фундаментальної системи розв’язків однорідної системи:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \\ x_{r+1} \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_r \\ 0 \\ \dots \\ 0 \end{pmatrix} + x_{r+1} \begin{pmatrix} \alpha_{1,r+1} \\ \alpha_{2,r+1} \\ \dots \\ \alpha_{r,r+1} \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \dots \\ \alpha_{rn} \\ 0 \\ 0 \\ \dots \\ 1 \end{pmatrix} \tag{4.5.11}$$

Випадок $m > n \ge r$ радимо читачу розглянути самостійно.

Може статись так, що матрицю буде приведено до так званого «східчастого» вигляду:

.

$$\begin{pmatrix}
0 & ... & 0 & a_{11} & a_{12} & ... & & & & & & & & & a_{1n} \\
0 & ... & 0 & 0 & 0 & ... & 0 & a_{2,j_2} & & & & & & & a_{2n} \\
0 & ... & 0 & 0 & 0 & ... & 0 & 0 & ... & 0 & a_{3,j_3} & ... & . & . & ... & a_{3n} \\
. & ... & . & . & . & ... & . & . & ... & . & . & ... & . & . & ... & . \\
0 & ... & & & & & & & & & & ... & 0 & a_{m,j_m} & ... & a_{mn} \\
0 & ... & & & & & & & & & & & & & ... & 0 \\
. & ... & & & & & & & & & & & & & & . \\
0 & ... & & & & & & & & & & & & & ... & 0
\end{pmatrix}.$$

Тобто, рядки матриці починаються з нулів, і кожного разу із зростанням номера рядка на одиницю кількість таких нулів буде монотонно збільшуватись не менше, ніж на одиницю. При розв'язанні такої системи треба діяти, як і у минулому випадку (при умові сумісності системи), залишаючи зліва ті стовпчики, які породжують мінор найбільшого порядку, що дорівнює рангу матриці, а інші доданки системи переносити в праву частину.

Приклад. $Ax = a,\ A(3\times 5), \dim(x) = 5, x = \begin{pmatrix} x_1, & x_2, & x_3, & x_4 & x_5 \end{pmatrix}^T, \dim(a) = 3$**.**

$$\begin{pmatrix} 1 & 1 & 3 & -2 & 3 & | & 1 \\ 2 & 2 & 4 & -1 & 3 & | & 2 \\ 3 & 3 & 5 & -2 & 3 & | & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 3 & -2 & 3 & | & 1 \\ 0 & 0 & -2 & 3 & -3 & | & 0 \\ 0 & 0 & -4 & 4 & -6 & | & -2 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 1 & 3 & -2 & 3 & | & 1 \\ 0 & 0 & 2 & -3 & 3 & | & 0 \\ 0 & 0 & -2 & 2 & -3 & | & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & 2 & -3 & 3 & | & 0 \\ 0 & 0 & 0 & -1 & 0 & | & -1 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & 2 & 0 & 3 & | & 3 \\ 0 & 0 & 0 & 1 & 0 & | & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & \frac{2}{3} & 0 & 1 & | & 1 \\ 0 & 0 & 0 & 1 & 0 & | & 1 \end{pmatrix} \to$$

Тепер перенесемо члени рівняння, що містять x_1, x_3 в праву частину.

$$\begin{pmatrix} 1 & 0 & 0 & | & -x_1-x_3 \\ 0 & 0 & 1 & | & 1-\frac{2}{3}x_3 \\ 0 & 1 & 0 & | & 1 \end{pmatrix} \Rightarrow \begin{cases} x_2=-x_1-x_3; \\ x_4=1; \\ x_5=1-\frac{2}{3}x_3. \end{cases}$$

Звідси $x=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}=\begin{pmatrix} x_1 \\ -x_1-x_3 \\ x_3 \\ 1 \\ 1-\frac{2}{3}x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}+x_1\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}+x_3\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ -\frac{2}{3} \end{pmatrix}.$

Читач повинен був впевнитись, що умову теореми Кронекера-Капеллі виконано: рангі основної і розширеної матриць співпадають і дорівнюють 3. Як бачимо, розв’язок системи є сумою частинного розв’язку неоднорідної системи і загального розв’язку однорідної системи – лінійної комбінації двох векторів, що становлять фундаментальну систему її розв’язків і яка, в свою чергу утворює базис $Ker(A);\dim\left(Ker(A)\right)=n-rank(A)=5-3=2$.

Матричні рівняння

Нехай є матриці таких розмірностей: $A(m\times m), X(m\times k), B(m\times k)$, $\det(A)\neq 0$. Запропонуємо методику розв’язання матричного рівняння

$$AX=B \tag{4.5.12}$$

Очевидно,

$$X=A^{-1}B \tag{4.5.13}$$

Але зараз покажемо, що таке розв’язання можна провести методом Гауса. Запропонуємо спочатку методику знаходження оберненої матриці методом Гауса.
В главі 2 було показано, що при виконанні умови $\det(A)\neq 0$ елементарними перетвореннями, множачи на елементарні матриці тільки зліва, дану матрицю можна звести до діагонального вигляду: $LA=I$, тобто, $L=A^{-1}$. Помножимо зліва на матрицю *L* тотожність $AA^{-1}=I$. Одержимо $LAA^{-1}=LI$ і в результаті - $A^{-1}=LI$. Це треба трактувати так: якщо з одиничною матрицею робити ті самі елементарні перетворення, що і з рядками матриці *A*, приводячи її до діагонального вигляду, одержимо з одиничної матриці обернену до *A* матрицю.

На практиці записують розширену матрицю: зліва A, а справа – одиничну і роблять з обома однакові елементарні перетворення, приводячи A до одиничної. Тоді справа одержимо A^{-1}.
Приклад.

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

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

$$\to\left(\begin{array}{ccc|ccc}1 & 2 & 2 & 1 & 0 & 0\\ 0 & -3 & -6 & -2 & 1 & 0\\ 0 & -3 & 3 & 0 & -1 & 1\end{array}\right)\to\left(\begin{array}{ccc|ccc}1 & 2 & 2 & 1 & 0 & 0\\ 0 & -3 & -6 & -2 & 1 & 0\\ 0 & 0 & 9 & 2 & -2 & 1\end{array}\right)\to$$

$$\to\left(\begin{array}{ccc|ccc}1 & 2 & 2 & 1 & 0 & 0\\ 0 & -3 & -6 & -2 & 1 & 0\\ 0 & 0 & 9 & 2 & -2 & 1\end{array}\right)\to\left(\begin{array}{ccc|ccc}1 & 2 & 2 & 1 & 0 & 0\\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0\\ 0 & 0 & 1 & \frac{2}{9} & -\frac{2}{9} & \frac{1}{9}\end{array}\right)\to$$

$$\to\left(\begin{array}{ccc|ccc}1 & 2 & 2 & 1 & 0 & 0\\ 0 & 1 & 2 & \frac{2}{3} & -\frac{1}{3} & 0\\ 0 & 0 & 1 & \frac{2}{9} & -\frac{2}{9} & \frac{1}{9}\end{array}\right)\to\left(\begin{array}{ccc|ccc}1 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & 0\\ 0 & 1 & 0 & \frac{2}{9} & \frac{1}{9} & -\frac{2}{9}\\ 0 & 0 & 1 & \frac{2}{9} & -\frac{2}{9} & \frac{1}{9}\end{array}\right)\to$$

$$\to\left(\begin{array}{ccc|ccc}1 & 0 & 0 & \frac{1}{9} & \frac{2}{9} & \frac{2}{9}\\ 0 & 1 & 0 & \frac{2}{9} & \frac{1}{9} & -\frac{2}{9}\\ 0 & 0 & 1 & \frac{2}{9} & -\frac{2}{9} & \frac{1}{9}\end{array}\right).$$

Отже,

$$A=\begin{pmatrix} \frac{1}{9} & \frac{2}{9} & \frac{2}{9} \\ \frac{2}{9} & \frac{1}{9} & -\frac{2}{9} \\ \frac{2}{9} & -\frac{2}{9} & \frac{1}{9} \end{pmatrix}.$$

Повернемось до рівняння (4.5.12). Його розв'язок дається формулою (4.5.13), яку можна трактувати як множення правої частини рівняння, матриці B на $L = A^{-1}$, тобто, з матрицею *B* треба зробити ті ж елементарні перетворення, що і з матрицею *A*, коли її приводять до одиничної. Очевидно, для цього можна застосувати для розширеної матриці ($A \mid B$) тільки що використану схему. Аналогічно можна діяти при розв'язанні рівняння $XA=B$: $X=BA^{-1}$ треба тільки одночасно здійснювати елементарні перетворення рядків – робити діагональними елементами нулі справа від них і привести матрицю *A* до одиничної. Тоді справа в розширеній матриці ($A \mid B$) одержимо розв'язок. Елементарні перетворення можна застосовувати також у випадку виродженої матриці, коефіцієнті при невідомій матриці.
Приклад.
$XA=B$:

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 3 & 6 \\ 4 & 8 \end{pmatrix}=\begin{pmatrix} 2 & 4 \\ 9 & 18 \end{pmatrix}.$$

$$\begin{pmatrix} 3 & 6 & | & 2 & 4 \\ 4 & 8 & | & 9 & 18 \end{pmatrix}\to\begin{pmatrix} 3 & 0 & | & 2 & 0 \\ 4 & 0 & | & 9 & 0 \end{pmatrix}.$$

Звідси

$$\begin{pmatrix} 3x_{11}+4x_{12} & 0 \\ 3x_{21}+4x_{22} & 0 \end{pmatrix}=\begin{pmatrix} 2 & 0 \\ 9 & 0 \end{pmatrix} \Rightarrow \begin{matrix} 3x_{11}+4x_{12}=2; \\ 3x_{21}+4x_{22}=0. \end{matrix} \Rightarrow \begin{matrix} x_{12}=\frac{1}{2}-\frac{3}{4}x_{11}; \\ x_{21}=\frac{4}{3}-\frac{4}{3}x_{22}. \end{matrix}$$

Відповідь: $X=\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}=\begin{pmatrix} x_{11} & \frac{1}{2}-\frac{3}{4}x_{11} \\ \frac{4}{3}-\frac{4}{3}x_{22} & x_{22} \end{pmatrix}.$

ГЛАВА 5
НОРМАЛЬНІ ФОРМИ ТА ФУНКЦІЇ МАТРИЦЬ

5.1. НОРМАЛЬНА ФОРМА ЖОРДАНА МАТРИЦІ

1. Нормальна форма Жордана. У цьому розділі розглянемо оператори та їхні матриці, що діють у комплексному n-вимірному просторі L (поле $K = \mathbb{C}$), при цьому матриці операторів, зрозуміло, будуть квадратними, розміру $n \times n$. Часто матрицю і сам оператор, якщо це не призведе до непорозуміння, будемо позначати однією і тією ж літерою.

Нехай $P(t)$ - многочлен з коефіцієнтом одиниця у старшому члені:

$$P(t) = t^m + a_1 t^{m-1} + \ldots + a_m. \tag{5.1.1}$$

Многочленом $P(A)$ від оператора A (або від матриці A) будемо називати вираз

$$P(A) = A^m + a_1 A^{m-1} + \ldots + a_m I, \tag{5.1.2}$$

де I – одиничний оператор (матриця).

Для кожної матриці A існує такий многочлен P, що $P(A) = 0$. (У такому разі будемо говорити, що P анулює матрицю або відповідно – оператор A, а P називати анулюючим многочленом матриці.) Дійсно, матриці в L утворюють лінійний простір розмірності n^2, тому степені A до порядку n^2 лінійно залежні. Отже, існує їхня нетривіальна лінійна комбінація, яка дорівнює нулю. Їй і відповідає анулюючий A многочлен. Очевидно також, що існують многочлени, анулюючі степені матриці.

Виявляється, що існують многочлени степені, нижчої за n, які анулюють деяку матрицю A.

Означення. Характеристичним многочленом матриці A називають многочлен

$$\chi(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & a_{23} & -a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & \lambda - a_{nn} \end{vmatrix} = \tag{5.1.3}$$

$$= \lambda^n + A_1 \lambda^{n-1} + \ldots + A_k \lambda^{n-k} + \ldots + A_n.$$

З'ясуємо, який вигляд має його коефіцієнт при λ^{n-1} (коефіцієнт при λ^n дорівнює, очевидно, одиниці) та вільний член. В розкладі визначника лише доданок $(\lambda - a_{11})...(\lambda - a_{nn})$ має (після множення) степені *n* та *n-1* змінної λ. Адже, якщо взяти довільний доданок розкладу визначника, в якому немає, скажімо, співмножника $\lambda - a_{kk}$, а є з того ж стовпчика елемент a_{ik}, то це також означає, що в цьому члені також немає $\lambda - a_{ii}$, тому що в кожному співмножнику не може бути двох елементів з одного стовпчика або з одного рядка. Отже, такий доданок має степінь, яка не перевищує n – 2. Тепер, якщо записати доданок $(\lambda - a_{11})...(\lambda - a_{nn})$ як многочлен, то можна побачити, що коефіцієнт при λ^{n-1} дорівнює сумі діагональних елементів матриці *A*, а це називається її слідом і позначається $sp(A)$ (*шпур* – нім.) або $tr(A)$ (*трейс* – англ.):

$$sp(A) = tr(A) = \sum_{k=1}^{n} a_{kk}. \tag{5.1.4}$$

Знайдемо інші коефіцієнти характеристичного многочлена. Для цього зручніше записати його у вигляді $\chi(\lambda) = \det(\lambda I - A) = (-1)^n \det(A - \lambda I)$. Нехай $a_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$,

e_i - *i*-й вектор-стовпчик одиничної матриці. Будемо використовувати таку форму запису для визначника:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & ... & a_{nn} - \lambda \end{vmatrix} = \left| a_1 - \lambda e_1,\ a_2 - \lambda e_2, ..., a_n - \lambda e_n \right|$$

(для зручності будемо розділяти вектори-стовпчики визначника комами, хоча це протирічить традиції).

Використовуючи лінійність визначника по кожному стовпчику , маємо:

$$\det(A - \lambda I) = \left| a_1 - \lambda e_1,\ a_2 - \lambda e_2,\ \cdots,\ a_n - \lambda e_n \right| =$$
$$= \left| a_1,\ a_2,\ \cdots,\ a_n \right| - \left(\left| e_1,\ a_2,\ \cdots,\ a_n \right| + \left| a_1,\ e_2, a_3, \cdots,\ a_n \right| + ... + \left| a_1,\ a_2,\ \cdots,\ a_{n-1}, e_n \right| \right) \lambda +$$
$$+ \left(\left| e_1,\ e_2,\ e_3,\ \cdots,\ a_n \right| + \left| a_1,\ e_2, a_3, \cdots,\ a_n \right| + ... + \left| a_1,\ a_2,\ \cdots,\ a_{n-2}, e_{n-1},\ e_n \right| \right) \lambda^2 +$$
$$+ (-1)^{n-k} \left(\left| e_1,\ e_2, ..., e_{n-k}, a_{n-k+1},\ \cdots,\ a_n \right| + ... + \left| a_1,\ a_2,\ \cdots, a_k,\ e_{n-k+1}, ..., e_{n-1},\ e_n \right| \right) \lambda^{n-k} +$$

$$+(-1)^n \left|e_1, e_2, \ldots, e_n\right| \lambda^n$$

Звідси видно, що коефіцієнт при λ^{n-1} - сума визначників вигляду $(-1)^{n-1}\left|e_1, e_2, \ldots, e_{k-1}, a_k, e_{k+1}, \cdots, e_n\right| = (-1)^{n-1} a_k$, тобто $(-1)^{n-1} Sp(A)$, а відповідний коефіцієнт в характеристичному многочлені тоді дорівнює $-Sp(A)$. Аналогічно вільний член характеристичного многочлена дорівнює $(-1)^n \left|a_1, a_2, \cdots, a_n\right| =$ $= (-1)^n \det(A)$, що вже було з'ясовано в інший спосіб. Зазначимо, що, наприклад, вираз $\left|e_1, e_2, \ldots, e_{n-k}, a_{n-k+1}, \cdots, a_n\right|$ - це мінор *k*-го порядку визначника, який одержано після викреслювання рядків і стовпчиків, що відповідають номерам $i = 1, \ldots, n-k$ одиничних векторів e_i. Діагональ цього мінора лежить на діагоналі визначника. Такий мінор називається головним. Отже, в дужках при λ^{n-k} міститься сума головних мінорів *k*-го порядку. Позначимо через $M_{i_1 \ldots i_k}^{i_1 \ldots i_k}, i_1 < \ldots < i_k$ головний мінор, на діагоналі якого знаходяться елементи $a_{i_1}, \cdots, a_{i_k}$, $i_1 < \ldots < i_k$. Тоді остаточно:

$$\chi(\lambda) = \det(\lambda I - A) =$$

$$= \lambda^n - Sp(A)\lambda^{n-1} + \ldots + (-1)^k \left(\sum_{i_1 \ldots i_k} M_{i_1 \ldots i_k}^{i_1 \ldots i_k} \right) \lambda^{n-k} + \ldots + (-1)^n \det(A) \tag{5.1.5}$$

Для того щоб знайти вільний член многочлена, очевидно, досить знайти його значення при $\lambda = 0$. Отже, вільний член $\chi(\lambda)$ дорівнює значенню $\det(\lambda I - A)$ при $\lambda = 0$, тобто $\det(-A) = (-1)^n \det(A)$.

Теорема 5.1.1. Характеристичний многочлен інваріантний відносно заміни базису в *L*.

Доведення випливає з формули (4.1.17). дійсно, при заміні базису за допомогою матриці переходу *T* матриця $\lambda E - A$ перейде в подібну матрицю $T^{-1}(\lambda E - A)T$, що випливає з формули (2.3.9), а

$$\det(T^{-1}(\lambda I - A)T) = \det(T^{-1})\det(\lambda I - A)\det(T) = (\det(T))^{-1}\det(T)\det(\lambda I - A) =$$
$$= \det(\lambda I - A).$$

Отже, можна говорити про характеристичний многочлен лінійного оператора, якому відповідають у різних базисах різні матриці.

Зауваження. Характеристичні многочлени подібних матриць співпадають.

Наслідок. Слід матриці, її визначник, а також суми головних мінорів однакового порядку не змінюються при замінах базису простору.

Вправа. Довести, що sp (*AB*) = sp (*BA*).

Доведення. Нехай *C* = *AB, D* = *BA.* Згідно з формулою (2.3.4)

$$c_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki}, \quad d_{ii} = \sum_{k=1}^{n} b_{ik} a_{ki},$$

звідки

$$sp(AB) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki}, \quad sp(BA) = \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ki} a_{ik} = sp(AB).$$

Означення. Число $\lambda \in K$ називається власним числом оператора *A*, якщо існує ненульовий вектор *x* такий, що $Ax = \lambda x$. *x* називається власним вектором, що відповідає власному числу λ.
Зауваження. Ця рівність еквівалентна рівності $\left(A - \lambda I\right)x = 0$.

Теорема 5.1.2. Число $\lambda \in K$ є власним для оператора *A* тоді і лише тоді, коли воно є коренем характеристичного многочлена оператора; в просторі над алгебраїчно замкненим полем лінійний оператор має хоча б один власний вектор.

Доведення. Оператор $\lambda I - A$ вироджений (тобто не має оберненого і в такому випадку має ненульове ядро $\mathrm{Ker}(\lambda I - A)$) тоді і лише тоді, коли його визначник дорівнює нулю: $\det(\lambda I - A)=0$. Отже існує ненульовий вектор *x*, що $(\lambda I - A)x = 0$, звідки $Ax = \lambda x$.
Зауваження. З лінійності оператора випливає, що разом з *x* оператор має власні вектори $\{tx\} \subset Ker(\lambda I - A),\ t \in K$.

Висновок. Якщо характеристичний многочлен над деяким полем *K* не має коренів, то оператор не має власних чисел. З основної теореми алгебри випливає, що кожен многочлен з комплексними коефіцієнтами має *n* комплексних коренів (деякі можуть бути кратними), тому кожен оператор у лінійному просторі над комплексним полем має власні числа, а значить, і власні вектори.

Нехай корені многочлена – комплексні числа λ_i кратностей $r_i,\ r_1 + ... + r_m = n$. Тоді

$$\det(\lambda E - A) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{r_i}. \tag{5.1.6}$$

Зауваження. Знаючи кратність деякого кореня, не можна нічого сказати про розмірність відповідного власного підпростору. Але якщо всі корені характеристичного многочлена прості, тобто мають кратність одиниця, то оператор має *n* власних векторів.

Теорема 5.1.3. Власні вектори, що відповідають різним власним числам, лінійно незалежні.

Доведення. Нехай $Ax=\lambda x$, $Ay=\mu y$, $\lambda\neq\mu$, $x,y,\neq 0$. Припустимо, що при деяких $\alpha,\beta\in K$, $\alpha x+\beta y=0$ і хоча б одне з цих чисел не дорівнює нулю, наприклад, $\alpha\neq 0$. Подіявши на цю рівність оператором *A*, одержимо: $\alpha\lambda x+\beta\mu y=0$. З цих двох рівностей після очевидних перетворень маємо $\alpha(\lambda-\mu)x=0$, що при зроблених припущеннях неможливо.

Висновок. Якщо всі корені характеристичного многочлена прості, то простір *L* розпадається на пряму суму *n* власних підпросторів, які відповідають власним числам оператора *A*. Оскільки кожен власний простір інваріантний відносно *A*, то в базисі з власних векторів оператор зображується діагональною матрицею з власними числами на діагоналі:

$$A=\begin{vmatrix}\lambda_1 & & & 0\\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n\end{vmatrix} \qquad (5.1.7)$$

Означення. Число λ називається регулярним числом оператора *A*, якщо оператор $\lambda I-A$ невироджений, тобто, має обернений.

Якщо оператор $\lambda I-A$ вироджений, то кажуть, що λ – точка спектра оператора *A*. Це позначається так: $\lambda\in spec(A)$.

Отже, спектр лінійного оператора, що діє у скінченновимірному лінійному просторі – це множина його власних чисел. Якщо корені характеристичного многочлена прості, то спектр називається простим.

Оператор з простим спектром має цікаву властивість.

Теорема 5.1.4. Нехай лінійний оператор *A* має простий спектр. Тоді існує такий вектор $e\in L$, що вектори $e,Ae,A^2e,...,A^{n-1}e$ утворюють базис *L*.

Доведення. Нехай $e_1,e_2,...,e_n$ - деякі ненульові власні вектори оператора *A*, які відповідають різним власним числам $\lambda_1,\lambda_2,...,\lambda_n$. Візьмемо вектор $e=e_1+e_2+...+e_n$ і припустимо, що $e,Ae,A^2e,...,A^{n-1}e$ лінійно залежні. У такому разі існують числа $a_1,a_2,...,a_n$ - не всі нульові, що $a_0e+a_1Ae+a_2A^2e+...+a_nA^ne=0,$ звідки, враховуючи зображення *e*, одержимо

$$\sum_{i=1}^{n}(a_0+a_1\lambda_i+a_2\lambda_i^2+...+a_{n-1}\lambda_i^{n-1})e_i=0.$$

З лінійної незалежності e_i випливає, що $a_0 + a_1\lambda_i + a_2\lambda_i^2 + ... + a_{n-1}\lambda_i^{n-1} = 0$ при кожному i = 1, …, n. Це означає: многочлен степені n – 1 має n різних коренів, що можливо тоді і лише тоді, коли $a_0 = a_1 = a_2 = ... = a_{n-1} = 0$, що і доводить лінійну незалежність системи векторів $e, Ae, A^2e, ..., A^{n-1}e$. Їхня кількість збігається з розмірністю простору, тому вони утворюють базис.

Означення. Вектор $e \in H \subseteq L$ називається твірним або циклічним вектором m-вимірного підпростору H, що відповідає оператору A, якщо система $e, Ae, A^2e, ..., A^{n-1}e$ – базис H.

Отже, Теорема 5.1.4 стверджує, що оператор A з простим спектром має твірні вектори всього простору. Твірний вектор e має вигляд $e = e_1 + e_2 + ... + e_n$, де $e_1, e_2, ..., e_n$ - ненульові власні вектори. Які відповідають різним власним числам.

Повернемось тепер до характеристичного многочлена (5.1.3).

Анулюючі многочлени матриці лінійного оператора

Теорема 5.1.5. (Гамільтона-Келі). Характеристичний многочлен оператора є його аннулюючим многочленом. Інакше кажучи, кожен оператор є коренем свого характеристичного многочлена.

Очевидно, матриця оператора при цьому є коренем свого характеристичного многочлена.

Доведення. Нехай B – матриця, приєднана до матриці $\lambda I - A$, тобто така, що кожен її елемент b_{ij} – це алгебраїчне доповнення до елемента a_{ji} матриці $\lambda I - A$. З розділу 4.1 відомо, що $B(\lambda I - A)$ = det $(\lambda I - A)E$. Очевидно, елементи B є многочлени від λ, не вищої за n-1, тому що при утворенні алгебраїчного доповнення b_{ij} до елемента елемента a_{ji} при $i \neq j$ треба викреслити j-й рядок та i-й стовпчик матриці $\lambda I - A$, а значить, і діагональні елементи $\lambda - a_{jj}$ та $\lambda - a_{ii}$, а це відповідатиме многочлену степені, не вищої за n-2. При знаходженні b_{ii} треба викреслити i-й рядок та j-й стовпчик, тому одержимо многочлен степені, не вищої за n-1.

Отже, B можна подати у вигляді

$$B = B_1\lambda^{n-1} + B_2\lambda^{n-2} + ... + B_{n-1}\lambda + B_n.$$

Використаємо згадувану рівність $B(\lambda I - A) = \chi(\lambda)I$. Нехай $\chi(\lambda) = \lambda^n + a_1\lambda^{n-1} + ... + a_{n-1}\lambda + a_n$. Тоді попередню рівність можна переписати у вигляді

$$(B_1\lambda^{n-1} + B_2\lambda^{n-2} + ... + B_{n-1}\lambda + B_n)(\lambda I - A) = (\lambda^n + a_1\lambda^{n-1} + ... + a_{n-1}\lambda + a_n)I\ .$$

Розкриваючи дужки та прирівнюючи коефіцієнти при рівних степенях λ, дістанемо

$$\begin{array}{l|l|l} \lambda^n & B_1 = I & \times A^n \\ \lambda^{n-1} & -B_1A + B_2 = a_1 I & \times A^{n-1} \\ \lambda^{n-2} & -B_2A + B_3 = a_2 I & \times A^{n-2} \\ \multicolumn{3}{c}{\dots\dots\dots\dots\dots\dots} \\ \lambda^2 & -B_{n-2}A + B_{n-1} = a_{n-2} I & \times A^2 \\ \lambda^1 & -B_{n-1}A + B_n = a_{n-1} I & \times A \\ \lambda^0 & -B_nA = a_n EI & \end{array} \tag{5.1.8}$$

Множачи рівності на степені A, як це показано у (5.1.8), і додаючи результати, дістанемо $0 = A^n + a_1A^{n-1} + ... + a_{n-1}A + a_n I$, тобто, $\chi(A) = 0$.

Тепер зрозуміло, що множина многочленів, яка анулює A, непорожня, до неї входить, принаймні, характеристичний многочлен A. Всі многочлени, анулюючі A, очевидно, утворюють ідеал кільця многочленів $K[t]$. Як вже відомо (вправа 2, розділ 1.6), кожен ідеал у кільці многочленів головний, тобто породжується одним многочленом, єдиним з точністю до постійного скалярного множника. Такий многочлен має найменшу степінь з усіх, які належать до ідеалу. Візьмемо многочлен, що породжує ідеал, має найменшу степінь і перший коефіцієнт одиницю; він анулює оператор A. Такий многочлен називається мінімальним многочленом оператора A. Будемо позначати його $M(t)$.

Приклад 5.1.1. $A = tE$. Тоді $\chi(t) = \det(\lambda E - A) = \det(\lambda I - tE) = (t-\lambda)^n$, а мінімальний многочлен - $M(t) = (t-\lambda)$, що випливає з вигляду A. При цьому всі вектори з простору L власні з власним числом $\lambda = t$. У кожному базисі оператор має одну й ту ж матрицю, бо $T(tI)T^{-1} = tTT^{-1} = tI$;

$$J_n(t) = \begin{pmatrix} t & 1 & 0 & \cdot & \cdot & 0 \\ 0 & t & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & t & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & 0 \\ \cdot & & & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & t \end{pmatrix} \tag{5.1.9}$$

Така матриця називається клітиною Жордана і позначається $J_n(t)$. З означення випливає, що $J_n(t) = J_n(0) + tI$. Характеристичний многочлен цієї матриці, очевидно, такий самий, як і в попередньому прикладі, $\chi(\lambda) = (t-\lambda)^n$. Отже, на головній діагоналі матриці стоять її власні числа, які дорівнюють t, а на діагоналі, що вище головної на одну позицію, - одиниці; всі інші елементи – нулі.

Знайдемо степені $J_n(0)$:

$$J^2{}_n(0) = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & 0 \\ \cdot & & & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & 0 \\ \cdot & & & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix} = \tag{5.1.10}$$

$$= \begin{pmatrix} 0 & 0 & 1 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & 1 \\ \cdot & & & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix}$$

Бачимо, що ланцюг з одиниць пересунувся на одну позицію вправо вгору. Застосовуючи індукцію, легко довести, що $J_n^k(0)$ одиниці стоять на k-й лінії паралельній головній діагоналі. Наприклад, всі елементи матриці $J_n^{n-1}(0)$ - нулі, за винятком останнього елемента в першому рядку, який дорівнює одиниці. З цих міркувань також випливає, що $J_n^n(0) = J_n^{n+1}(0) = \ldots = 0.$

Означення. Оператор B, який у деякій степені дорівнює нулю, називається нільпотентним.

З попереднього бачимо, що $J_n(0)$ - нільпотентний оператор. Очевидно, всі його степені теж нільпотентні.

Нехай $e_1 = (1,0,\ldots,0)$, $e_2 = (0,1,\ldots,0),\ldots,e_n = (0,0,\ldots,0,1)$ - базис E. Подивимось, як на ці вектори діє клітина Жордана. Маємо

$$J_n(t)e_1 = te_1,\ \ J_n(t)e_2 = e_1 + te_2,\ldots, J_n(t)e_n = e_{n-1} + te_n; \tag{5.1.11}$$

або

$$\left(J_n(t)-tI\right)e_1=0,\quad \left(J_n(t)-tI\right)e_2=e_1,\ldots,\left(J_n(t)-tI\right)e_n=e_{n-1} \qquad (5.1.12)$$

З цих формул можна зробити важливі наслідки: e_1 – власний вектор оператора $J_n(t)$ з власним числом $\lambda = t$, при $t = 0$, e_n – твірний вектор в L.

Базис, у якому оператор діє за правилом (5.1.10), називається жордановим. Треба запам'ятати, що при цьому номери базисних векторів зменшуються.

Оскільки $(tI-J_n(t))^k = J_n(0)^k \neq 0$ при $0 \le k \le n-1$, то мінімальний многочлен збігається з характеристичним, бо він ділить характеристичний многочлен $\chi(t)=(t-\lambda)^n$, а з попередньої нерівності випливає, що ануляторів $J_n(t)$ меншої степені немає.

В доповнення до викладеного покажемо, як діє нільпотентна клітина Жордана $J_n(0)(n\times n)$ на довільну матрицю $A(n\times m)$:

$$\begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot & \cdot & 0 \\ \cdot & & & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1m} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-1,1} & a_{n-1,2} & \cdot & \cdot & \cdot & a_{n-1,m} \\ a_{n,1} & a_{n,2} & \cdot & \cdot & \cdot & a_{n,m} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2m} \\ a_{31} & a_{32} & \cdot & \cdot & \cdot & a_{3m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & a_{n,2} & \cdot & \cdot & \cdot & a_{n,m} \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

Як бачимо, вона зсуває рядки A на одну позицію вгору. Аналогічно «нижня» клітина Жордана зсуває рядки A на одну позицію униз:

$$\begin{pmatrix} 0 & 0 & . & \cdot & \cdot & 0 \\ 1 & 0 & 0 & . & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot & \cdot \\ \cdot & . & \cdot & \cdot & \cdot & . \\ \cdot & . & . & 1 & 0 & 0 \\ 0 & \cdot & \cdot & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & . & \cdot & \cdot & a_{1m} \\ a_{21} & a_{22} & . & . & \cdot & a_{2m} \\ \cdot & \cdot & . & \cdot & \cdot & \cdot \\ \cdot & . & \cdot & \cdot & \cdot & . \\ a_{n-1,1} & a_{n-1,2} & . & . & \cdot & a_{n-1,m} \\ a_{n,1} & a_{n,2} & \cdot & \cdot & . & a_{n,m} \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & . & \cdot & \cdot & 0 \\ a_{11} & a_{12} & . & . & \cdot & a_{1m} \\ \cdot & \cdot & . & \cdot & \cdot & \cdot \\ \cdot & . & \cdot & \cdot & \cdot & . \\ a_{n-2,1} & a_{n-2,2} & . & . & \cdot & a_{n-2,m} \\ a_{n-1,1} & a_{n-1,2} & \cdot & \cdot & . & a_{n-1,m} \end{pmatrix}.$$

Приклад 5.1.2. Нехай $I_n(\lambda)$ - простір функцій вигляду $e^{\lambda x}P(x),$ де $\lambda \in C;$ $P(x)$ - довільний многочлен степені, яка не більше за n-1 (такі функції називаються квазімногочленами степеня n-1). Оскільки $(e^{\lambda x}P(x))' = e^{\lambda x}(\lambda P(x)) + P'(x)),$ то диференціювання – лінійний оператор в $I_n(\lambda)$. У $I_n(\lambda)$ можна вибирати базис - $e_i = (x^{i-1} / i!)e^{\lambda x}, \;\; i = 1,...,n$. Легко бачити, що

$$\begin{aligned} &\frac{d}{dx}(e_{i+1}) = (x^{i-1} / (i-1)!) + \lambda(x^i / i!)e^{\lambda x} = e_i + \lambda e_{i+1} \\ &(i = 1,...,n), \\ &\frac{d}{dx}(e_1) = \frac{d}{dx}\left(e^{\lambda x}\right) = \lambda e^{\lambda x} = \lambda e_1 \;\; (i = 0) \end{aligned} \tag{5.1.13}$$

при i=1,…,n. Коли i=1 $\dfrac{d}{dx}(e_1) = \lambda e^{\lambda x} = \lambda e_1$.

Таким чином, функції $\left\{(x^{i-1} / i!)e^{\lambda x}, \;\; i = 1,...,n\right\}$ утворюють базис, в якому оператор диференціювання *d/dx* має матрицю вигляду (5.1.9), жорданову клітину і базис є жордановим для оператора. Цей приклад має велике значення в теорії систем лінійних диференціальних рівнянь.

При розгляді клітини Жордана ми вже помітили, що

$Ker(tI - J_n(t)) = \{t_1 e_1, t_1 \in R\}$. Неважко переконатись, що $(tI - J_n(t))^2 e_2 = 0$ і, звичайно, $(tI - J_n(t))^2 e_1 = 0$. Отже, $Ker(tI - J_n(t))^2$ = $\{t_1 e_1, t_1 \in \mathbf{R}\} \oplus \{t_2 e_2, t_2 \in \mathbf{R}\}$ - пряма сума одновимірних підпросторів, натягнутих на базисні вектори e_1, e_2.

Взагалі,

$$Ker(tI - J_n(t))^k = \{t_1 e_1, t_1 \in \mathbb{R}\} \oplus ... \oplus \{t_k e_k, t_k \in \mathbb{R}\}, 1 \le k \le n.$$

Розглянемо ситуацію уважніше. Нехай $f_n \in Ker(tI - J_n(t))^n$ і

$f_n \notin Ker(tI - J_n(t))^m$, $m < n$. Це означає, що $(J_n(t) - tI)^n f_n = 0$. Звідси

$(J_n(t) - tI)^n f_n = (J_n(t) - tI)^{n-1}\big((J_n(t) - tI) f_n\big) = 0$, отже, вектор

$f_{n-1} = (J_n(t) - tI) f_n \in Ker(J_n(t) - tI)^{n-1}$. Взагалі

$f_{n-k} = (J_n(t) - tI) f_{n-k+1} = (J_n(t) - tI)^k f_n \in Ker(J_n(t) - tI)^{n-k}$ і нарешті

$f_1 = (J_n(t) - tI) f_2 = (J_n(t) - tI)^{n-1} f_n \in Ker(J_n(t) - tI)$.

Таким чином, маємо ланцюг вкладених підпросторів:

$$E = Ker(tI - J_n(t))^n \supset Ker(tI - J_n(t))^{n-1} \supset ... \supset Ker(tI - J_n(t)) \supset \{0\}. \qquad (5.1.14)$$

Так побудовані вектори $f_1, f_2, ..., f_n$ утворюють жордановий базис простору.

Означення. Вектори $f_1, f_2, ..., f_k \in E$ називаються лінійно незалежними відносно підростору $H \subset E$, якщо ніяка їхня лінійна комбінація з коефіцієнтами, що не всі дорівнюють нулю, не належать до *H*. Максимально така система називається базисом *E* відносно підпростору *H*.

Очевидно, лінійно незалежна система векторів відносно підпростору є лінійно незалежною. Доповнюючи відносний базис *E* (відносно *H*) базисом *H*, отримаємо базис *E*.

У попередньому прикладі 5.1.1 f_n - єдиний базисний вектор в *E* відносно $Ker(tI - J_n(t))^{n-1}$; f_n, f_{n-1} - базис відносно $Ker(tI - J_n(t))^{n-2}$ і т.ін.

Розглянемо тепер загальний оператор A з власним числами $\lambda_1,...,\lambda_m$, що є коренями характеристичного многочлена $\chi(\lambda)$ відповідно кратностей $r_1, r_2, ..., r_m$ $(r_1 + r_2 + ... + r_m = n)$. Тоді

$$\chi(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{r_i} \tag{5.1.15}$$

Означення. Кореневим (або приєднаним) вектором оператора A, який відповідає власному числу λ, називається такий ненульовий вектор f, що $(A - \lambda I)^k f = 0$. Найменше число h, що задовольняє цій рівності, називається висотою кореневого вектора.

Очевидно, власний вектор – це кореневий висоти 1.

У прикладі 5.1.1 $(J_n(t) - tI)^n e_n = 0$, тобто e_n - кореневий вектор висоти n, $(J_n(t) - tI)^{n-1} e_{n-1} = 0$, тому e_{n-1} - кореневий вектор висоти n-1,…; e_2 – кореневий вектор висоти 2; e_1 – кореневий вектор висоти 1, тобто, власний.

Очевидно, кореневі вектори, які відповідають одному власному числу λ (вони можуть мати різні висоти), утворюють підпростір в L, який будемо називати кореневим підпростором і позначити $L(\lambda)$.

Якщо f – кореневий вектор висоти h, то $f_1 = (A - \lambda I)^{h-1} f$ - власний вектор, відповідний власному числу λ:

$$(A - \lambda I) f_1 = (A - \lambda I)((A - \lambda I)^{h-1} f) = (A - IE)^h f = 0.$$

Звідси випливає, що кореневі вектори можуть відповідати тільки кореням характеристичного многочлена оператора, бо вони і тільки вони є власними його числами.

Введемо позначення:

$$\chi_j(\lambda) = \prod_{\substack{i=1 \\ i \neq j}}^{m} (\lambda - \lambda)^{ri} = \chi(\lambda)(\lambda - \lambda_j)^{-rj}. \tag{5.1.16}$$

Всі $\chi_j(\lambda)$ $(j = 1, ..., m)$ взаємно прості.

Покладемо

$$L_j \equiv \mathrm{Im}(\chi_j(A)). \tag{5.1.17}$$

У такому разі $L_j \subset L(\lambda_j) = Ker\left(\left(A - \lambda_j I\right)^{r_j}\right)$, тобто $((A - \lambda_j I)^{r_j})L_j = 0.$ Дійсно, якщо $x \in L_j$, тобто $x = \chi_j(A)y$ для деякого $y \in L,$ тоді

$$((A - \lambda_j I)^{r_j})x = ((A - \lambda_j I)^{r_j})\chi_j(A)y = \chi(A)y = 0$$

В силу теореми Гамільтона-Келі.

Лема 5.1.1. $L = L_1 \oplus L_2 \oplus ... \oplus L_{\mathrm{m}}$

Доведемо спочатку, що $L = L_1 + L_2 + ... + L_{\mathrm{m}}$. Многочлени $\chi_j(\lambda)$ $(j = 1,...,m)$ взаємно прості. Тому, як відомо (див. (1.6.7)), існують такі многочлени $u_i(t)$, що $\sum_{j=1}^{m} \chi_j(t) u_j(t) = 1$, тому, підставляючи замість *t* оператор *A*, одержимо $\sum_{j=1}^{m} \chi_j(A) u_j(A) = I$. Діючи цією рівністю на довільний вектор $l \in L$, маємо $l = \sum_{j=1}^{m} \chi_j(A)(u_j(A)l).$ Оскільки $\chi_j(A)(u_j(A)l) \in L_j,$ то $l \in \sum_{j=1} L_j$.

Тепер доведемо, що $L_i \cap (\sum_{j \neq i} L_j) = \{0\}$. Нехай *l* – вектор з цього перетину. Тоді $((A - \lambda_i EI^{r_j})l = 0,$ оскільки $l \in L_i \in L(\lambda_i)$. З іншого боку,

$$\chi_i(A)l = \prod_{j \neq i}^{m} ((A - \lambda_j I)^{r_j})l = 0,$$

оскільки $l \in \sum_{j \neq i} L_j$. Многочлени $(\lambda - \lambda_i)^{r_i}$ і $\chi_i(\lambda)$ взаємно прості, тому існують такі многочлени $u(\lambda),\ \upsilon(\lambda),$ що $u(\lambda)(\lambda - \lambda_i)^{ri} + \upsilon(\lambda)\chi_i(\lambda) = 1$. Підставляючи сюди *A* замість λ і діючи таким виразом на *l*, матимемо

$$u(A)(A - \lambda_i E)^{ri} l + \upsilon(A)\chi_i(A)l = u(A)0 + \upsilon(A)0 = l = 0.$$

Лема 5.1.2. $L(\lambda_j) = L_j$

Вже було доведено, що $L(\lambda_j) \subseteq L_j$. Доведемо обернене включення. Нехай $l \in L(\lambda_j)$. Зобразимо *l* як суму: $l = l' + l''$, де $l' \in L_j$; $l'' \in \bigoplus_{i \neq j} L_i$. Звідси $l'' = l - l'$ і $l'' \in L(\lambda_j)$, тоді як $l \in L(\lambda_j)$ і $l' \in L_j \subseteq L(\lambda_j)$. Тому $((A - \lambda_i E)^{ri})l'' = 0$ і, оскільки $l'' \in \bigoplus_{i \neq j} L_i$, то

$$\chi_j(A)l'' = \prod_{k\neq j}\left(A-\lambda_k I\right)^{r_k}\left(\sum_i \prod_{s\neq i}\left((A-\lambda_s I)^{r_s}\right)\right)x_s.$$

Многочлени $(\lambda-\lambda_i)^{r_i}$ і $\chi_i(\lambda)$ взаємно прості, тому існують такі многочлени $u(\lambda)$, $\upsilon(\lambda)$, що $u(\lambda)(\lambda-\lambda_i)^{ri}+\upsilon(\lambda)\chi_i(\lambda)=1$, звідки, підставляючи сюди A замість λ, одержимо $u(A)(A-\lambda_i E)^{ri}l''+\upsilon(A)\chi_i(A)l''=l''=0$, тому $l=l'\in L_j$.

Зауваження. З лем 1 і 2 одержимо формулу

$$L = L_1\oplus L_2\oplus\ldots\oplus L_m = L(\lambda_1)\oplus L(\lambda_2)\oplus\ldots\oplus L(\lambda_m) \qquad (5.1.18)$$

Цікаво з‘ясувати, як діє оператор $(A-\lambda_j I)^{r_j}$ у просторі $\tilde{L}_j \equiv \oplus_{i\neq j} L_i = \oplus_{i\neq j} L(\lambda_i)$.

Лема 5.1.3. Простір $\tilde{L}_j$ - інваріантний відносно оператора $(A-\lambda_j I)^{r_j}$; звуження його на $\tilde{L}_j$ має обернений (тобто $\lambda=\lambda_j$ не є власним вектором в $\tilde{L}_j$).

Доведення. Нехай $i\neq j$. Доведемо, що підпростір $L_i = L(\lambda_i)$ інваріантний відносно $(A-\lambda_j I)^{r_j}$.

За означенням

$$L_i = (A-\lambda_1 I)^{r_1}\ldots(A-\lambda_{i-1}I)^{r_{i-1}}(A-\lambda_{i+1}I)^{r_{r+1}}\ldots(A-\lambda_m I)^{r_m}L.$$

Серед цих співмножників є, безумовно, $(A-\lambda_j I)^{r_j}$. Тому, якщо подіяти на L_i оператором $(A-\lambda_j E)^{r_j}$, то такий співмножник з'явиться двічи, і тоді

$$\begin{gathered}(A-\lambda_j I)^{r_j}L_i = (A-\lambda_1 I)^{r_1}\ldots(A-\lambda_{i-1}I)^{r_{i-1}}(A-\lambda_{i+1}EI^{r_{i+1}}\ldots\\ \ldots(A-\lambda_m EI^{r_m}((A-\lambda_j I)^{r_j}L)\subset L_i\end{gathered}.$$

Це й доводить, що $\tilde{L}_j \equiv \oplus_{i\neq j} L_i = \oplus_{i\neq j} L(\lambda_i)$ інваріантний відносно $(A-\lambda_j I)^{r_j}$.

Тепер доведемо, що звуження цього оператора на L_j має обернений. Для цього досить довести, що ядро звуження $(A-\lambda_j I)^{r_j}$ на $\tilde{L}_j$ нульове. Дійсно, якщо вектор $\tilde{x}=\sum_{i\neq j}x_i\in\tilde{L}_j$ належить до ядра цього звуження, то $\tilde{x}=L(\lambda_j)\cap\tilde{L}_j=\{0\}$, оскільки сума (5.1.16) пряма. Тому в $\tilde{L}_j$ цей оператор діє як взаємно-однозначне відображення, а це й доводить, як відомо, оберненість оператора.

Запропонуємо ще одне, альтернативне доведення (5.1.16).

Лема 5.1.4. Простір L можна розкласти в пряму суму (взагалі кажучи, неортогональну) таких інваріантних відносно $(A-\lambda I)^r$ підпросторів:

$$L = L(\lambda) \oplus \mathrm{Im}((A-\lambda I)^r), \tag{5.1.19}$$

де λ - власне число оператора *A*, що має кратність *r*, як корінь характеристичного рівняння.

Доведемо спочатку, що коли $x \in L(\lambda) \cap (A-\lambda I)^r L$, то *x* = 0. Справді, з одного боку, $(A-\lambda I)^h x = 0$ для деякого $h \le r$, з іншого – $x = (A-\lambda I)^r y$. У такому разі $(A-\lambda I)^{r+h} y = 0$, тому *y* – кореневий вектор, що відповідає власному числу λ і має висоту $h+r$. Але згідно з означенням $L(\lambda)$ всі кореневі вектори, які відповідають власному кореню λ, анулюються оператором $(A-\lambda I)^r$, тобто $x = (A-\lambda I)^r y = 0$. Тепер з формули (2.3.16) одержимо: $\dim(L) = \dim(L(\lambda)) + \dim(\mathrm{Im}((A-\lambda I)^r))$, тобто підпростори $L(\lambda)$ і $\mathrm{Im}\left((A-\lambda I)^r\right)$ породжують весь простір L.

Для доведення того, що $(A-\lambda I)^r$ має обернений у $\mathrm{Im}\left((A-\lambda I)^r\right)$, треба помітити, що ядро звуження оператора, діюче в цьому підпросторі, нульове. Дійсно, інакше б існував вектор $x \in \mathrm{Im}\left((A-\lambda I)^r\right)$, для якого $(A-\lambda I)^r x = 0$, але це свідчило б, що $x \in L(\lambda)$. Щойно було доведено, що в такому випадку *x* = 0. Цим лема повністю доведена.

Вправа. Довести останню рівність в формулі (5.1.16), тобто $L = L(\lambda_1) \oplus L(\lambda_2) \oplus \ldots \oplus L(\lambda_m)$ спираючись на результати леми 4.

Вказівка. Одержавши розклад (5.1.17) для $\lambda = \lambda_1$, треба в $\mathrm{Im}\left((A-\lambda_1 I)^{r_1}\right)$ відщепити підпростір $L(\lambda_2)$ і т.ін.

Підсумуємо отримані результати.

Якщо оператор *A* має власні числа $\lambda_1,\ldots,\lambda_m$, які є коренями характеристичного многочлена $\chi(\lambda) = \det(\lambda I - A)$, відповідно кратностей $r_1, r_2, \ldots, r_m$ $(r_1 + r_2 + \ldots + r_m = n)$, то простір L, в якому діє *A*, можна розкласти в пряму суму (5.1.16) підпросторів $L(\lambda_j)$, інваріантних відносно кожного з операторів $(A-\lambda_j I)^{r_j}$ $(j = 1,\ldots,m)$. Тому достатньо з'ясувати, як діє оператор *A* в кожному з цих підпросторів. В $L(\lambda_j) = L_j$ оператор $(A-\lambda_j I)^{r_j}$ нільпотентний, кожен вектор $x_j = L(\lambda_j)$ обов'язково задовольняє умову $(A-\lambda_j I)^{r_j} x_j = 0$. Але можливий випадок, коли існує число $k_j \le r_j$, що $(A-\lambda_j I)^{r_j} x_j = 0$ для кожного $x_j \in L(\lambda_j)$. При цьому $M_j(\lambda) = (\lambda - \lambda_j)^{k_j}$ є мінімальний многочлен звуження оператора *A* на просторі $L(\lambda_j)$.

Нехай λ - деяке власне число оператора *A*, що є коренем характеристичного многочлена кратності *r*. Якщо *y* – кореневий вектор висоти $h\,(h \le r)$, тобто $(A-\lambda I)^h y = 0$, то вектор $z=(A-\lambda I)y$ - кореневий вектор висоти *h*-1, вектор $z=(A-\lambda I)^2 y$ - кореневий вектор висоти *h*-2 і т.ін. Отже, можна побудувати ланцюг:

$$L(\lambda) = Ker(A-\lambda I)^h \supset Ker(A-\lambda I)^{h-1} \supset ... \supset Ker(A-\lambda I) \supset \{0\} \qquad (5.1.20)$$

Введемо позначення: $B = A - \lambda E$. Тоді цей ланцюг перепишеться у вигляді

$$L(\lambda) = Ker(B)^h \supset Ker(B)^{h-1} \supset ... \supset Ker(B) \supset \{0\}$$

Виберемо в максимальному з цих підпросторів, в $Ker(B)^h$, базис відносно $Ker(B)^h : e_1^h, ..., e_{q_r}^h$. Тоді вектори $Be_1^h, ..., Be_{q_r}^h \in Ker\left(B^{h-1}\right)$. Вони лінійно незалежні відносно $Ker\left(B^{h-2}\right)$. Дійсно, нехай $\alpha_1 e_1^h + ... + \alpha_q e_{q_r}^h \in Ker\left(B^{h-1}\right)$, а це протирічить їх лінійній незалежності відносно $Ker\left(B^{h-1}\right)$. Доповнимо одержані вектори до базису в $Ker\left(B^{h-1}\right)$ відносно $Ker\left(B^{h-2}\right)$:

$$Be_1^h, ..., Be_{q_r}^h, e_1^{h-1}, ..., e_{q_{r-1}}^{h-1}$$

Знову застосуємо до таких векторів оператор *B* і, доповнюючи їх образи до базису в $Ker\left(B^{h-2}\right)$ відносно $Ker\left(B^{h-3}\right)$, одержимо

$$B^2 e_1^h, ..., B^2 e_{q_r}^h, Be_1^{h-1}, ..., Be_{q_{r-1}}^{h-1}, e_1^{h-2}, ..., e_{q_{r-2}}^{h-2}$$

Продовжуючи цей процес, дойдемо нарешті до $Ker\left(B\right)$. Таким чином, маємо таблицю:

$$
\begin{aligned}
&e_1^h,\ldots,e_{q_r}^h,\\
&Be_1^h,\ldots,Be_{q_r}^h,e_1^{h-1},\ldots,e_{q_{r-1}}^{h-1},\\
&B^2e_1^h,\ldots,B^2e_{q_r}^h,Be_1^{h-1},\ldots,Be_{q_{r-1}}^{h-1},e_1^{h-2},\ldots,e_{q_{r-2}}^{h-2},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&B^{h-1}e_1^h,\ldots,B^{h-1}e_{q_r}^h,B^{h-2}e_1^{h-1},\ldots,B^{h-2}e_{q_{r-1}}^{h-1},\ldots,Be_1^2,\ldots,Be_{q_2}^2,e_1^{(1)},\ldots,e_{q_1}^1
\end{aligned}
\tag{5.1.21}
$$

Очевидно, вектори цієї таблиці утворюють базис у $L(\lambda)=Ker\left(B^h\right)$. Покажемо, що при цьому у базисі матриця оператора A має вигляд прямої суми клітин Жордана.

Розглянемо, наприклад, перший стовпчик цієї таблиці. Введемо позначення $f_1=B^{h-1}e_1^h,\ f_2=B^{h-2}e_1^h\ldots,\ f_h=e_1^h$. Ці вектори утворюють базис деякого підпростору. Очевидно, f_1 - власний вектор оператора A, відповідний власному числу λ. Легко бачити, що $Bf_i=f_{i-1}$ для $i=2,\ldots,k$ (порівняйте з прикладом 5.1.1). Очевидно, матриця A в такому базисі в підпросторі $L(\lambda)$ має вигляд (5.1.7) клітини Жордана розміру $h\times h$. (Матриця B має аналогічний вигляд з нулями на діагоналі.) Аналогічно знаходиться матриця оператора A в кожному з підпросторів, що натягнуто на вектори довільного стовпчика таблиці (5.1.19). очевидно, при просуванні вправо по таблиці матимемо підпростори, розмірності яких не збільшуються. Останнім векторам останнього рядка відповідають одновимірна клітини: (λ) – для матриці A і (0) – для матриці B.

Остаточно матриця A матиме вигляд прямої суми одержаних клітин Жордана з різними власними числами на діагоналі. Така матриця називається нормальною формою Жордана оператора A, а базис, у якому вона має такий вигляд – жордановим.

Оскільки ми прийшли до такого вигляду матриці оператора, певним чином вибравши базис в L, то, як відомо,

$$
A_J=TAT^{-1}
\tag{5.1.22}
$$

Зауважимо, що хоч ми довільним чином вибрали вектори відносних базисів (див. (5.1.19)), але їхня кількість – це є інваріанти, незалежні від вибору базису, а тільки від розмірностей ядер. Отже, всі зображення (5.1.20) будуть однакові з точністю до розташування в матриці A_j клітин Жордана.

Сформулюємо результат.

Теорема 5.1.6. Для оператора, що діє в комплексному скінченновимірному просторі, існує жордановий базис, тобто матриця оператора в деякому базисі може бути приведена заміною базиса до жорданової нормальної форми (5.1.20). Матриця A_J вигляду (5.1.20) визначена однозначно з точністю до розташування жорданових клітин.

Дамо рекомендації щодо застосування пропонованої методики приведення матриці до нормальної форми Жордана. Очевидно, для побудови жорданового базису достатньо знати лише розмірності ядер у (5.1.18). Оскільки $\dim\left(Ker(A-\lambda I)^i\right)=\dim L-\dim\left(\operatorname{Im}(A-\lambda I)^i\right)$, досить знайти $\dim\left(\operatorname{Im}(A-\lambda I)^i\right)=rank(A-\lambda I)^i$, а ранг матриці $\left(A-\lambda I\right)^i$ можна знайти елементарними перетвореннями, що не змінюють ранг матриці, тобто додаванням до деякого вектора-рядка (стовпчика) лінійної комбінації інших векторів-рядків (стовпчиків).

Жорданова нормальна форма матриць третього та четвертого порядка. Приклади

Розглянемо всі випадки зведення матриці до нормальної форми Жордана в залежності від кратностей її власних чисел. В цьому розділі будуть розглядатись матриці з дійсними власними числами. Випадок комплексних власних чисел буде розглянуто у розділі 5.4 для матриці 2-го порядку. Почнемо з матриць третього порядку.

Випадок 5.1.1. Всі три корені $\lambda_1,\lambda_2,\lambda_3$ характеристичного многочлена різні. Нагадаємо, що такі корені називаються простими. Очевидно, що $\dim Ker\left(A-\lambda_1 I\right)=\dim Ker\left(A-\lambda_2 I\right)=\dim Ker\left(A-\lambda_3 I\right)=1$, Простір розпадається в пряму суму власних підпросторів матриці:

$$\mathbb{R}^3=Ker\left(A-\lambda_1 I\right)\oplus Ker\left(A-\lambda_2 I\right)\oplus Ker\left(A-\lambda_3 I\right)$$

і жорданова форма J_A матиці має вигляд

$$J_A=\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}.$$

Випадок 5.1.2. $\chi(A,t)=\left(t-\lambda_1\right)^2\left(t-\lambda_2\right)$. В цьому випадку ядро оператора $A-\lambda_2 I$ може бути тільки одновимірне. Нехай $\dim Ker\left(A-\lambda_1 I\right)=2$. Це означає, що жордановий базис можна вибрати з власних векторів оператора *A*,

$$\mathbb{R}^3 = Ker\left(A-\lambda_1 I\right) \oplus Ker\left(A-\lambda_2 I\right)$$

В такому базисі її жорданова матриця має діагональний вигляд:

$$J_A = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Зауважимо, що в цьому випадку мінімальний многочлен матриці $\left(t-\lambda_1\right)\left(t-\lambda_2\right)$ не співпадає з характеристичним і є його дільником.
Для знаходження жорданового базису досить розв’язати лінійні системи

$$\left(A-\lambda_1 I\right)x = 0 \text{ та } \left(A-\lambda_2 I\right)x = 0.$$

Цікаво зауважити, що вектори $g_1, g_2 \in Ker(A-\lambda_1 I), g_3 \in Ker(A-\lambda_2 I)$ жорданового базису можна знайти, не розв’язуючи системи, а саме: нехай x і y – будь-які лінійно незалежні вектори з $\mathbb{R}^3$. Тоді $\left(A-\lambda_2 I\right)x, \left(A-\lambda_2 I\right)y$ (підбираємо x і y так, щоб ці одержані вектори були ненульовими) - лінійно незалежні і належать $Ker(A-\lambda_1 I)$, як це вже пояснювалось. Повторимо це ще раз: $\left(A-\lambda_1 I\right)\left(\left(A-\lambda_2 I\right)x\right) = \left(\left(A-\lambda_1 I\right)\left(A-\lambda_2 I\right)\right)x = 0$, аналогічно з другим вектором. Так само виберемо вектор z з $\mathbb{R}^3$ і $\left(A-\lambda_1 I\right)z$ (якщо він ненульовий) належить $Ker(A-\lambda_2 I)$.

Випадок 5.1.3. $\chi(A,t) = \left(t-\lambda_1\right)^2\left(t-\lambda_2\right)$. В цьому випадку знову ядро оператора $A-\lambda_2 I$ може бути тільки одновимірне. Нехай тепер $\dim Ker\left(A-\lambda_1 I\right) = 1$. Це означає, що оператор має одновимірний підпростір власних векторів і одновимірний відносно $Ker(A-\lambda_1 I)$ підпростір кореневих векторів $Ker(A-\lambda_1 I)^2$; його розмірніcь дорівнює 2. Маємо ланцюг ядер:

$$Ker\left(A-\lambda_1 I\right)^2 \supset Ker\left(A-\lambda_1 I\right) \supset \{0\},$$
$$\mathbb{R}^3 = Ker\left(A-\lambda_1 I\right)^2 \oplus Ker\left(A-\lambda_2 I\right)$$

Виберемо жордановий базис, взявши вектор $x \in \mathbb{R}^3$ так, що $(A-\lambda_2 I)x \neq 0$. Тоді $(A-\lambda_2 I)x \in Ker(A-\lambda_1 I)^2$. Слід вибирати *x* таким чином, щоб $(A-\lambda_1 I)(A-\lambda_2 I)x \neq 0$. Покладемо $g_2 = (A-\lambda_2 I)x$. Тоді $g_1 = (A-\lambda_1 I)g_2 = (A-\lambda_1 I)(A-\lambda_2 I)x \in Ker(A-\lambda_1 I)$. Власний вектор g_3 з $Ker(A-\lambda_2 I)$ знайдемо, розв'язавши систему $(A-\lambda_2 I)x = 0$ або, як і раніше, взявши вектор *z* з $\mathbb{R}^3$ так, щоб вектор $(A-\lambda_1 I)^2 z$ був ненульовий, тоді він належить $Ker(A-\lambda_2 I)$.

З виписаних рівностей маємо:

$$\begin{aligned} Ag_1 &= \lambda_1 g_2 \\ Ag_2 &= g_1 + \lambda_1 g_2 \\ Ag_3 &= \lambda_2 g_3 \end{aligned}$$

і

$$J_A = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

Випадок 5.1.4. $\chi(A,t) = (t-\lambda)^3$. Нехай $\dim Ker(A-\lambda I) = 1$, тобто, $rank(A-\lambda E) = 2$. Тоді маємо ланцюг (довгий) ядер:

$$\mathbb{R}^3 = Ker(A-\lambda I)^3 \supset Ker(A-\lambda I)^2 \supset Ker(A-\lambda I) \supset \{0\}$$

Беремо вектор g_3 з $\mathbb{R}^3$ так, щоб вектор $(A-\lambda I)^2 g_3$ був ненульовий, тоді $g_2 = (A-\lambda I)g_3$ належить $Ker(A-\lambda I)^2$ і $g_1 = (A-\lambda I)g_2$ належить $Ker(A-\lambda I)$. Звідси маємо:

$$\begin{aligned} Ag_1 &= \lambda g_1 \\ Ag_2 &= g_1 + \lambda g_2 \\ Ag_3 &= g_2 + \lambda g_3 \end{aligned}$$

і

$$J_A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

Випадок 5.1.5. $\chi(A,t) = (t-\lambda)^3$. Нехай $\dim Ker(A-\lambda I) = 2$, тобто, $rank(A-\lambda I) = 1$. Тоді маємо ланцюг (короткий) ядер:

$$\mathbb{R}^3 = Ker(A-\lambda I)^2 \supset Ker(A-\lambda I) \supset \{0\}$$

Беремо вектор g_3 з $\mathbb{R}^3$ так, щоб вектор $(A-\lambda I)g_3$ був ненульовий, тоді $g_2 = (A-\lambda I)g_3$ належить $Ker(A-\lambda I)$. Розв'язавши систему $(A-\lambda I)x = 0$, виберемо вектор g_1 лінійно незалежний з g_2. Маємо:

$$\begin{aligned} Ag_1 &= \lambda g_1 \\ Ag_2 &= \lambda g_2 \\ Ag_3 &= g_2 + \lambda g_3 \end{aligned}$$

В цьому базисі жорданова матриця має вигляд:

$$J_A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

Розглянемо тепер деякі випадки матриць 4-го порядку. Інші випадки читач, користуючись пропонованою технікою, зможе розглянути самостійно. Попередні приклади дають змогу не пояснювати процедуру зведення до нормальної форми Жордана занадто ретельно.

Випадок 5.1.5. Корені характеристичного рівняння прості. Матриця зводиться до діагонального вигляду у базисі, що складається з власних векторі матриці.

Випадок 5.1.5. $\chi(A,t) = (t-\lambda_1)^2(t-\lambda_2)(t-\lambda_3)$

а). Якщо $rank(A-\lambda_1 I) = 3$, а значить $\dim Ker(A-\lambda_1 I) = 1$, $\dim Ker(A-\lambda_1 I)^2 = 2$, жорданова нормальна форма матриці має вигляд:

$$J_A = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}.$$

b). Якщо $rank\left(A-\lambda_1 I\right)=2$, а значить $\dim Ker\left(A-\lambda_1 I\right)=2$, жорданова нормальна форма матриці має діагональний вигляд:

$$J_A = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}.$$

Випадок 5.1.6. $\chi(A,t)=\left(t-\lambda_1\right)^2\left(t-\lambda_2\right)^2$

a). $\dim Ker\left(A-\lambda_1 I\right)=\dim Ker\left(A-\lambda_1 I\right)=2$. Матриця зводиться до діагонального вигляду:

$$J_A = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

b). $\dim Ker\left(A-\lambda_1 I\right)=1,\ \dim Ker\left(A-\lambda_2 I\right)=2$. Тоді автоматично $\dim Ker\left(A-\lambda_1 I\right)^2=2$.

Матриця зводиться до нормальної форми Жордана:

$$J_A = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

c). $\dim Ker\left(A-\lambda_1 I\right)=\dim Ker\left(A-\lambda_2 I\right)=1$ Тоді автоматично $\dim Ker\left(A-\lambda_1 I\right)^2=\dim Ker\left(A-\lambda_2 I\right)^2=2$.

Матриця зводиться до нормальної форми Жордана:

$$J_A = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}.$$

Випадок 5.1.7. $\chi(A,t) = (t-\lambda)^4$

a). $\dim Ker(A-\lambda I) = 1$. Маємо довгий ланцюг ядер:

$\mathbb{R}^3 = Ker(A-\lambda I)^4 \supset Ker(A-\lambda I)^3 \supset Ker(A-\lambda I)^2 \supset Ker(A-\lambda I) \supset \{0\}$

і нормальну форму матриці:

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

b). $\dim Ker(A-\lambda I) = 2$. В такій ситуації ми не можемо нічого сказати про $\dim Ker(A-\lambda I)^2$, тому ядро $Ker(A-\lambda I)^2$ треба досліджувати окремо.

b1). Якщо $\dim Ker(A-\lambda I)^2 = 4$, то, знайшовши два лінійно незалежні вектори $f_1, f_3 \in Ker(A-\lambda I)$, побудуємо відносний (відносно $Ker(A-\lambda I)$) базис в $Ker(A-\lambda I)^2$ - $\{f_2, f_4\}$, щоб виконувались співвідношення $f_1 = (A-\lambda I) f_2$, $f_3 = (A-\lambda I) f_4$. В цьому базисі матриця зводиться до нормальної форми Жордана:

$$J_A = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix},$$

вона є прямою сумою двох двовимірних клітин Жордана.

b2). Нехай $\dim Ker(A-\lambda I) = 2$, $\dim Ker(A-\lambda I)^2 = 3$. Тоді маємо ланцюг ядер

$$\mathbb{R}^3 = Ker(A-\lambda I)^3 \supset Ker(A-\lambda I)^2 \supset Ker(A-\lambda I) \supset \{0\}.$$

Жордановий базис в такій ситуації будувати досить просто: беремо довільний вектор $g_4 \in \mathbb{R}^4 = Ker\left(A-\lambda I\right)^3$ і знаходимо:

$g_3 = \left(A-\lambda I\right) g_4 \in Ker\left(A-\lambda I\right)^2 \to g_2 = \left(A-\lambda I\right) g_3 \in Ker\left(A-\lambda I\right)$. Безумовно, треба вибирати g_4 так, щоб всі ці вектори були ненульовими. Вектор $g_1 \in Ker\left(A-\lambda I\right)$ знайдемо як лінійно незалежний з $g_2 \in Ker\left(A-\lambda I\right)$ розв'язок системи $\left(A-\lambda I\right)x = 0$. В цьому базисі матриця зводиться до нормальної форми Жордана:

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

Розглянемо кілька прикладів зведення матриці до нормальної форми Жордана. Кожного разу цікаво також виписати матрицю переходу від канонічного базису до жорданового, але ми не будемо цього робити, оскільки вигляд такої матриці зрозумілий, якщо знайдено жордановий базис, він дається формулою (2.3.8).

Приклад 5.1.3. Маємо

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

Після простих обчислень випишемо характеристичний многочлен цієї матриці: $\chi(\lambda) = (\lambda-1)^2(\lambda-2)$.

Оскільки $\dim(Ker(A-2I)) = 1$, треба знайти лише $\operatorname{rank}(A-I)$

$$A-E = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & \text{-1} & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & \text{-1} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Отже, $\operatorname{rank}(A-I) = \dim(\operatorname{Im}(A-I)) = 2$ і $\dim(Ker(A-I)) = 3-2 = 1$. Після цього можна не знаходити матрицю $(A\text{-}I)^2$, тому що $\dim(Ker(A-I)^2)$ може дорівнювати тільки 2, інших варіантів бути не може. У результаті можна побудувати такий базис:

$e = e_1 \in Ker(A-I)^2,\ \ e_1 \notin Ker(A-I),\ \ e_2 = (A-I)e_1 \in Ker(A-I),$ - власний вектор, тобто $Ae_2 = e_2$; e_3 – власний вектор A, що відповідає власному числу λ = 2, $Ae_3 = 2e_3$. У такому базисі

$$\begin{aligned} Ae_1 &= e_1 + e_2, \\ Ae_2 &= e_2, \\ Ae_3 &= 2e_3 \end{aligned} \tag{5.1.21}$$

Базис Жордана будується так: $f_1 = e_2,\ \ f_2 = e_1,\ \ f_3 = f_3$. У ньому оператор A діє таким чином:

$$\begin{aligned} Af_1 &= f_1, \\ Af_2 &= f_1 + f_2, \\ Af_3 &= 2f_3 \end{aligned} \tag{5.1.23}$$

Згадаємо, як виписати матрицю, виходячи з такої системи лінійних рівностей (формули (2.3.1)-(2.3.3)): треба коефіцієнти, що знаходяться у рядках, записати у стовпчики. При цьому матимемо таку ж саму матрицю, рядки якої збігаються з рядками попередньої системи. Це спостереження дає змогу одразу вписати матрицю в нормальній формі, не переходячи до жорданового базису.

Зауваження. Розв'язуючи інші приклади, будемо користуватись цим правилом. Це дасть змогу не виписувати систему вигляду (5.1.23), а обмежуватись системами вигляду (5.1.21)

Тепер можна виписати матрицю Жордана:

$$A_j = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Приклад 5.1.4.

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

Після простих обчислень знайдемо характеристичний многочлен цієї матриці: $\chi(\lambda) = (\lambda - 1)^3$.

Обчислимо $\operatorname{rank}(A-I)=\dim(\operatorname{Im}(A-I))$:

$$A-E=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix}.$$

звідки $\dim(\operatorname{Im}(A-I))=2$ і $\dim(Ker(A-I))=3-2=1$. Це означає, що розмірність $Ker(A-I)^2$ відносно $Ker(A-I)$ може дорівнювати тільки 1, тобто, $\dim\left(Ker(A-I)^2\right)=2$. В цьому можна впевнитись безпосереднім обчисленням:

$$(A-I)^2=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix}=\begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

звідки $\dim(\operatorname{Im}(A-I)^2)=1,\ \dim(Ker(A-I)^2)=3-1=2$. Тепер, зрозуміло, $\dim(Ker(A-I)^3)=3$, бо інших варіантів бути не може (в силу теореми Гамільтона–Келі і за допомогою безпосередніх обчислень: $\chi(A)=(A-A)^3=0$).

Маємо жордановий базис в L:

$e_3=(1,0,0)^T,\ \ e_2=(A-I)e_3,\ \ e_1=(A-I)e_2=(A-I)^2e_3;\ e_3$ - власний вектор.

Відзначимо, що вектор e_3 був «вгаданий»: проста перевірка показує, що він не міститься в ядрах $Ker(A-I)$ і $Ker(A-I)^2$. У цьому базисі

$$\begin{aligned} Ae_1 &= e_1, \\ Ae_2 &= e_1+e_2, \\ Ae_3 &= e_2+e_3 \end{aligned}$$

Звідси

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

Тобто, A_J - клітина Жордана.

Приклад 5.1.5. Маємо

$$A=\begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{pmatrix}.$$

Характеристичний многочлен $\chi(\lambda)=(\lambda+1)^4$. Знайдемо $rank(A+I)=\dim(\operatorname{Im}(A+I))$:

$$A=\begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -11 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 1 & 0 \\ -3 & 4 & -5 & 0 \\ 8 & -4 & 4 & 0 \\ 15 & -10 & 11 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & -1 & 1 & 0 \\ -3 & 4 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 15 & -10 & 11 & 0 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 2 & 0 & 1 & 0 \\ -3 & -1 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 15 & 1 & 11 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & 0 & 1 & 0 \\ 7 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -7 & 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

звідки $\dim(\operatorname{Im}(A+I))=2,\ \ \dim(Ker(A+I))=4-2=2.$

Знайдемо $\dim\left(\operatorname{Im}(A+I)^2\right)$:

$$(A+I)^2=\begin{pmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{pmatrix}\begin{pmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{pmatrix}=$$

$$=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & -1 & 1 \end{pmatrix},$$

Легко бачити, що $\dim(\operatorname{Im}(A+I)^2)=1$ і $\dim(Ker(A+I)^2)=4-1=3$. Очевидно, $\dim(Ker(A+I)^3)=4$.

Маємо $e_4=e\in Ker(A+I)^3$ - базис в $Ker(A+I)^3$ відносно $Ker(A+I)^2$; $e_3=(A+I)e_4$ базис в $Ker(A+I)^2$ відносно

$Ker(A+I)$; $e_2=(A+I)e_3$, e_2, e_4 - власні вектори, базис в $Ker(A+I)$. В цьому базисі

$$\begin{aligned}
Ae_1 &= -e_1,\\
Ae_2 &= -e_2,\\
Ae_3 &= e_2 - e_3\\
Ae_4 &= e_3 - e_4.
\end{aligned}$$

Звідси

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

Приклад 5.1.6. Маємо

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

Знаходимо характеристичний многочлен: $\chi(\lambda)=(\lambda-1)^4(\lambda+1)$.

Власному числу $\lambda=-1$ відповідає одновимірний власний простір і клітина (-1) у матриці Жордана A_j. Розглянемо більш детально $\lambda=1$. Знайдемо $rank(A-I)=\dim(\mathrm{Im}(A-I))$:

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

$$\to \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

звідки $\dim(\mathrm{Im}(A-I))=3,\ \dim(Ker(A-I))=5-3=2$. Знайдемо $\dim(Ker(A-I)^2)$:

$$(A-I)^2=\begin{pmatrix} 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 2 & 2 & -2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 2 & 2 & -2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix} =$$

$$=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -4 & 4 \\ 0 & 0 & 4 & -4 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

отже, $\dim(\mathrm{Im}(A-I)^2)=1,\ \ \dim(Ker(A-I))^2=5-1=4.$

Таким чином, маємо базис в $Ker(A-I)^2$ відносно $Ker(A-I)$: $e_3,\ e_4$ і базис в $Ker(A-I)$; $e_1=(A-I)e_3$; $e_2=(A-I)e_4$ і ще один власний вектор e_5, який треба знаходити безпосередньо, розв'язуючи систему $(A+I)x=0$. Запишемо дію оператора у такому базисі:

$$
\begin{aligned}
Ae_1 &= e_1,\\
Ae_2 &= e_1 + e_2,\\
Ae_3 &= e_3,\\
Ae_4 &= e_3 + e_4,\\
Ae_5 &= -e_5.
\end{aligned}
$$

Таким чином

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

5.2. ЖОРДАНОВА НОРМАЛЬНА ФОРМА МАТРИЦІ НОРМАЛЬНОГО ОПЕРАТОРА

Оператор, що діє в унітарному просторі називається нормальним, якщо він комутує зі своїм спряженим:

$$
AA^* = A^*A \tag{5.2.1}
$$

Лема 5.2.1. Нехай x – власний вектор нормального оператора A, що відповідає власному числу λ. Тоді x є також власним вектором оператора A^*, що відповідає власному числу $\bar{\lambda}$.

Доведення. Маємо:

$$
0 = \left((A-\lambda I)x, (A-\lambda I)x\right) = \left(x, (A^* - \bar{\lambda} I)(A-\lambda I)x\right) = \left(x, (A-\lambda I)(A^* - \bar{\lambda} I)x\right) =
$$

$$
= \left(\left(A^* - \bar{\lambda} I\right)x, \left(A^* - \bar{\lambda} I\right)x\right).
$$

Звідси $\left(A^* - \bar{\lambda} I\right)x=0$, що і треба було довести.

Лема 5.2.2. Власні вектори нормального оператора, що відповідають різним власним числам, ортогональні.

Доведення. Нехай $Ax = \lambda x,\ Ay = \mu y$. Тоді $A^* x = \bar{\lambda} x,\ A^* y = \bar{\mu} y$.

Маємо: $\lambda(x, y) = (Ax, y) = (x, A^* y) = \left(x, \bar{\mu} y\right) = \mu(x, y)$. Звідси $(\lambda - \mu)(x, y) = 0$ і $(x, y) = 0$.

Лема 5.2.3. Нехай A – нормальний оператор, λ - його власне число. Тоді $Ker(A-\lambda I)^2 = Ker(A-\lambda I)$.

Доведення. Нехай x – кореневий вектор висоти 2, що відповідає власному числу λ: $(A-\lambda I)^2 x = 0$, тобто, $(A-\lambda I)\big((A-\lambda I)x\big) = 0$. Тоді $(A-\lambda I)x$ також є власним вектором A^*, що відповідає власному числу $\overline{\lambda}$:
$(A^* - \overline{\lambda} I)\big((A-\lambda I)x\big) = 0$. Отже, $0 = \Big((A^* - \overline{\lambda} I)(A-\lambda I)x, x\Big)$ =
= $\big((A-\lambda I)x, (A-\lambda I)x\big)$, звідки $(A-\lambda I)x = 0$.

Теорема 5.2.1. В лінійному просторі можна вибрати ортонормований базис власних векторів нормального оператора, в якому його матриця буде мати діагональний вигляд.

Як свідчить лема, всі кореневі вектори нормального оператора мають висоту одиниця, тобто, всі вони власні. Власні вектори, що відповідають різним власним числам ортогональні, що доведено в лемі . В свою чергу, в кожному власному підпросторі можна вибрати ортонормований базис. Таким чином буде побудовано ортонормований базис власних векторів нормального оператора.

Зауважимо, що все це можна сказати про матрицю оператора A^* в тому самому базисі. На її діагоналі будуть знаходитись числа, комплексно спряжені до відповідних чисел матриці *A*.

Це можна сформулювати ще так: простір можна розкласти в пряму ортогональну суму власних підпросторів, що відповідають його власним числам:

$$E_n = Ker\big(A-\lambda_1 I\big) \oplus \ldots \oplus Ker\big(A-\lambda_m I\big) \qquad (5.2.2)$$

$$\dim Ker\big(A-\lambda_k I\big) = r_k,\ k = 1,\ldots,m \qquad (5.2.3)$$

Згідно з формулою (5.2.2), довільний вектор $x \in E_n$ можна подати (однозначно) у вигляді

$$x = x_1 + \ldots + x_m, x_k \in Ker(A-\lambda_k I).$$

Тоді

$$Ax = Ax_1 + \ldots + Ax_m = \lambda_1 x_1 + \ldots + \lambda_m x_m \qquad (5.2.4)$$

Нехай P_k - ортопроектор на власний підпростір $Ker\big(A-\lambda_k I\big)$. Згідно з формулою (5.2.3) оператор *A* можна подати у вигляді

$$A = \lambda_1 P_1 \oplus \ldots \oplus \lambda_m P_m \qquad (5.2.5)$$

Теорема 5.2.2. Нехай в ортогональному базисі оператор має діагональну матрицю. Тоді оператор нормальний.

Зауважимо, що ця теорема обернена до теореми 5.2.1.

Доведення. На діагоналі матриці спряженого оператора в тому ж базисі знаходяться комплексно спряжені числа. Діагональні матриці комутують, і це доводить теорему.

Теорема 5.2.3. Нехай *A* – матриця нормального оператора і J_A - її жорданова нормальна форма, тобто, діагональна матриця. Тоді існує унітарна матриця *U* оператора заміни базису, що

$$J_A = U^{-1}AU \tag{5.2.6}$$

Доведення. Справедливість твердження випливає з загальної форми (2.3.19) вигляду матриці оператора в новому базисі. Треба зазначити, що оператор заміни базису переводить канонічний (ортонормований) базис, в якому записано дану матрицю, в ортонормований базис з власних векторів, тобто, він зберігає норми базисних векторів і скалярні добутки між ними, а це властивість унітарного оператора. Потім він продовжується на весь простір за лінійністю.

Зробимо висновки: важливими прикладами нормальних операторів є самоспряжені і унітарні оператори. Отже, їхні матриці можна привести до діагональному вигляду. Уточнимо спочатку вигляд формули (5.2.5) для самоспряжених операторів.

Лема 5.2.4. Всі власні числа самоспряженого оператора дійсні.

Доведення. Нехай $Ax = \lambda x,\ x \neq 0$. Помножимо скалярно обидві частини цієї рівності на x: $\lambda(x,x) = (Ax,x) = (x,Ax) = (x,\lambda x) = \overline{\lambda}(x,x)$, звідки і випливає $\lambda = \overline{\lambda}$.

Тепер дослідимо власні числа унітарного оператора.

Лема 5.2.5. Всі власні числа унітарного оператора *V* за модулем дорівнюють одиниці.

Доведення. Нехай $Vx = \lambda x,\ x \neq 0$. Тоді $(x,x) = (Vx,Vx) = |\lambda|^2 (x,x)$, звідки $|\lambda| = 1$.

Означення 5.2.1. Лінійний оператор *A* називається додатним, якщо $\forall x \neq 0\ (Ax,x) > 0$.

Легко бачити, що всі власні числа додатного оператора додатні.

Очевидно, невироджений оператор *A*A* – додатний і самоспряжений. Справді,

$$\forall x \neq 0\ (A^*Ax,x) = (Ax,Ax) = (x,A^*Ax) > 0$$

Означення 5.2.2. Лінійний додатний самоспряжений оператор *B* називається квадратним коренем з додатного самоспряженого оператора *A*, якщо $A = B^2$.

Лема 5.2.5. Нехай *A* – самоспряжений додатний оператор. Тоді існує додатний самоспряжений оператор *B* такий, що $B^2 = A$.
Доведення. Існує ортонормований базис з власних векторів оператора *A*, в якому матриця оператора має діагональний вигляд з власними числами $\lambda_1,...,\lambda_n$.
Нехай оператор *B* має ті самі власні вектори з власними числами $\sqrt{\lambda_1},...,\sqrt{\lambda_n}$ на діагоналі. Він самоспряжений, оскільки має діагональну форму у ортогональному базисі і додатний, оскільки всі власні числа додатні. Очевидно, квадрат його матриці дорівнює матриці оператора *A*.

Теорема 5.2.3. Довільний невироджений оператор можна зобразити у вигляді добутку унітарного і додатного самоспряженого оператора.
Доведення. Нехай *A* – невироджений оператор. Оператор $A^* A$ додатний і самоспряжений. Нехай *B* – квадратний корінь з $A^* A$: $A^* A = B^2$. Помножимо обидві частини рівності зліва і справа на B^{-1}: $B^{-1}A^* AB^{-1} = I$. Тоді $B^{-1}A^* = \left(AB^{-1}\right)^{-1}$. Помітимо, що $\left(AB^{-1}\right)^* = B^{-1}A^*$ (оскільки *B* - самоспряжений), звідки $\left(AB^{-1}\right)^{-1} = \left(AB^{-1}\right)^*$, а це означає, що оператор $U = AB^{-1}$ - унітарний.
Отже,

$$A = UB \qquad (5.2.7)$$

що й треба було довести.
Зображення довільного лінійного оператора у вигляді добутку унітарного і додатного операторів називається полярним розкладом цього оператора.
Зауваження. Можна застосувати доведений факт до A^*: $A^* = UB$.
Застосувавши спряження до обох частин, одержимо: $A = \left(UB\right)^* = B^* U^* = BU^{-1}$.
Оператор U^{-1} також спряжений, як і U. Отже, одержали ще один полярний розклад:

$$A = BU^{-1} \qquad (5.2.8)$$

5.3. ЦИКЛІЧНА НОРМАЛЬНА ФОРМА МАТРИЦІ

Розглянемо деякий лінійний оператор A, що діє в лінійному *n*-вимірному просторі *L* над полем R. Отже, від поля не вимагається алгебраїчної замкнутості.

Означення 5.3.1. Вектор $l \in L$ називається циклічним, якщо вектори $l, Al, A^2 l,..., A^{n-1} l$ утворюють базис простору *L*.

Означення 5.3.2. Простір називається циклічним відносно оператора, якщо він має циклічний вектор.

Покладемо $e_i = A^{n-i}l,\ i=1,...,n$. Очевидно $Ae_i = AA^{n-i}l = A^{n-i+1}l,\ i=2,...,n$. Для того, щоб знайти $Ae_1 = A^n l$, зауважимо, що вектори $e_1 = A^{n-1}l$, $e_2 = A^{n-2}l$,..., $e_n = l$ і $A^n l = Ae_1$ — лінійно залежні. Отже, існують такі числа a_i, серед яких не всі дорівнюють нулю, що

$$A^n l = \sum_{i=0}^{n-1} a_i A^i l \tag{5.3.1}$$

Запишемо дію A на базисні вектори у вигляді таблиці:

$$\begin{aligned}
&Ae_1 = a_{n-1}e_1 + a_{n-1}e_2 + \ldots + a_1 e_{n-1} + a_0 e_n\ ,\\
&Ae_2 = e_1,\\
&Ae_3 = e_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&Ae_n = e_{n-1}.
\end{aligned} \tag{5.3.2}$$

Запишемо матрицю $[A]_{e,e}$ оператора A у побудованому циклічному базисі, що відповідає (5.2.2):

$$[A]_{e,e} = \begin{pmatrix} a_{n-1} & 1 & 0 & \ldots & 0 \\ a_{n-2} & 0 & 1 & \ldots & 0 \\ . & . & . & \ldots & . \\ a_1 & 0 & 0 & \ldots & 1 \\ a_0 & . & . & \ldots & 0 \end{pmatrix} \tag{5.3.3}$$

З формули (5.3.1) випливає, що $A^n l - \sum_{i=0}^{n-1} a_i A^i l = 0$. Доведемо, що многочлен $M(t)$ $= t^n - \sum_{i=0}^{n-1} a_i t^i = 0$ анулює оператор A. Для цього досить довести виконання попередньої рівності для базисних векторів простору L Маємо: $M(A)(A^i l) = A^i(M(A)l) = 0$, а вектори $A^i\, l$ утворюють базис. З іншого боку, якщо $N(t)$ — анулюючий оператор A многочлен степені, меншої за n, то, подіявши оператором $N(A)$ на вектор l, одержимо, що між базисними векторами $l, Al, A^2 l,..., A^{n-1} l$ є лінійна залежність. Отже, $M(t)$ — мінімальний многочлен оператора A, тому вигляд клітини (5.3.3) не залежить від вибору циклічного вектору *l*. З попередніх міркувань можна зробити ще один важливий висновок: якщо простір циклічний відносно *A*, то його розмірність збігається зі степенем мінімального многочлена оператора *A*, а тому мінімальний і характеристичний многочлени збігаються. Зауважимо, що в 4.3 (приклад 9) характеристичний многочлен матриці (5.3.3) було знайдено за допомогою розкладу визначника $\det\left(\left[A\right]_{e,e}-\lambda I\right)$ за першим стовпчиком.

Аналогічно процедурі приведення матриці до нормальної форми Жордана можна, використовуючи аналогічні методи, привести матрицю у спеціально побудованому базисі до прямої суми циклічних клітин. При цьому замість дільників $(t-\lambda_i)^{r_i}$ характеристичного многочлена слід мати справу з многочленами $p_i(t)^{r_i}$, де $p_i(t)$ - дільники характеристичного многочлена, які не мають коренів в R. Таке зображення єдине, якщо мінімальні многочлени всіх клітин не мають дійсних коренів.

5.4. ФУНКЦІЇ ВІД МАТРИЦЬ

Одна з основних проблем операторного (матричного числення) — це визначення функції від оператора (матриці). Розглянемо побудову аналітичних функцій від оператора (матриці). Нагадаємо, що аналітична функція на деякій відкритій множині $G\subset\mathbb{C}$ - це така функція, що її ряд Тейлора абсолютно збігається до неї в точках цієї множини. Відомо, що на кожній замкненій підмножині G збіжність рівномірна. Якщо *f(z)* аналітична, то

$$f(z)=\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}z^k$$

цей ряд абсолютно збігається в крузі радіуса збіжності *R* з центром у нулі.

Для оператора *A* покладемо

$$f(A)=\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}A^k \tag{5.4.1}$$

Зрозуміло, що функція від матриці визначається за тією ж формулою. Дослідимо цей ряд на збіжність. Для цього застосуємо критерій Коші. Оцінимо за нормою відрізок операторного ряду:

$$\left\|\sum_{k=n+1}^{n+p}\frac{f^{(k)}(0)}{k!}A^k\right\|\le\sum_{k=n+1}^{n+p}\frac{\left|f^{(k)}(0)\right|}{k!}\|A\|^k$$

Оскільки відповідний числовий ряд збігається, якщо $\|A\| < R$, то для нього виконано умови збіжності критерію Коші, отже, вони виконуються і для операторного ряду, і він збігається.

Радимо самостійно довести для операторних рядів основні теореми теорії числових та функціональних рядів. Ці доведення майже співпадають. Зокрема, важливими є ознака Вейєрштрасса рівномірної збіжності та теорема про почленне диференціювання операторного функціонального ряда.

Далі під оператором часто будемо розуміти оператор в $\mathbb{R}^n$, тобто, квадратну матрицю, і це буде зрозуміло з контексту. Проблему побудови функції від матриці, виявляється, можна звести до побудови такого ряду від матриці її нормальної форми. Дійсно, $A = TA_J T^{-1}$, $A^2 = TA_J T^{-1} TA_J T^{-1} = TA_J^2 T^{-1}$, звідки за індукцією $A^k = TA_J^k T^{-1}$. Тому

$$f(A) = T\left(\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A_J^k\right) T^{-1}.$$

Експонента матриці. За ознакою Вейєрштраса ряд

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2} A^2 + \dots \qquad (5.4.1)$$

збігається абсолютно і рівномірно на кожній множині $\{A\colon \|A\| \le a, a > 0\}$ алгебри матриць, тому що

$$\|\exp(A)\| = \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k \le e^a < \infty.$$

Зазначимо, що з (5.4.1) випливає

$$\exp(0) = I \qquad (5.4.2)$$

Доведемо, що для $t \in \mathrm{R}$

$$\frac{d}{dt} \exp(tA) = A\exp(A) \qquad (5.4.3)$$

Продиференціювавши ряд (5.4.4) по t формально, одержимо ряд з похідних:

$$\sum_{k=0}^{\infty} \frac{1}{k!} \frac{d}{dt}(t^k) A^k = \sum_{k=0}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k = A \sum_{t=0}^{\infty} \frac{t^t}{i!} A^i = A\exp(tA).$$

Цей ряд збігається рівномірно на кожній обмеженій області своїх аргументів, тобто при $\|A\| \le a,\ |t| \le T$. Тому за теорією про диференціювання ряду похідна суми існує і дорівнює сумі ряду з похідних.

Рівність(5.4.3) означає, що матрична функція

$$X(t) = exp(tA) \tag{5.4.4}$$

є розв'язком матричного диференціального рівняння

$$\frac{d}{dt} X(t) = AX(t) \tag{5.4.5}$$

Зауважимо, що для матриці

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{5.4.6}$$

$$A^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix} \tag{5.4.7}$$

Приклад 5.3.1. Функція від діагональної матриці (5.4.6):

$$f(A) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} A^k = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} \tag{5.4.8}$$

Зокрема, якщо *f(z) = exp (z)*, то

$$\exp(tA) = \begin{pmatrix} \exp(t\lambda_1) & 0 & \cdots & 0 \\ 0 & \exp(t\lambda_2) & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & \exp(t\lambda_n) \end{pmatrix} \tag{5.4.9}$$

Приклад 5.3.2. Функція від нільпотентної матриці.

Нехай m — найменше ціле число, що $A^m = 0$ (тоді, як раніше було сказано, A — нільпотентна матриця). У такому разі ряд (5.4.1) — просто скінченна сума. Найбільш цікавий приклад: $A = J(0)$, де $J(0)$ - клітина Жордана (5.1.7) з нулями на діагоналі. Тоді

$$f(J(0)) = \sum_{k=0}^{m-1} \frac{f^{(k)}(0)}{k!} \left(J(0)\right)^k. \tag{5.4.10}$$

Якщо згадати формулу (5.1.8) і подальші дослідження щодо вигляду степенів нільпотентної клітини Жордана, то можна одержати такий вираз для функції від клітини Жордана:

$$f(J(0)) = \begin{pmatrix} 1 & f(0) & \frac{1}{2!} f'(0) & \ldots & \frac{1}{(m-1)!} f^{m-1}(0) \\ 0 & 1 & f(0) & \ldots & \frac{1}{(m-2)!} f^{m-2}(0) \\ 0 & 0 & 1 & \ldots & \frac{1}{(m-3)!} f^{m-3}(0) \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & f(0) \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \tag{5.4.11}$$

Важливою функцією є експонента від оператора та його матриці, оскільки за її допомогою можна знаходити розв'язки систем лінійних диференціальних рівнянь:

$$\exp(zJ) = \begin{pmatrix} 1 & z & z^2/2! & \ldots & z^{m-1}/(m-1)! \\ 0 & 1 & z & \ldots & z^{m-2}/(m-2)! \\ 0 & 0 & 1 & \ldots & z^{m-3}/(m-3)! \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & z \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

Теорема 5.3.3. Нехай лінійні оператори A і B комутують: $AB=BA$. Тоді

$$exp(A+B) = exp(A)exp(B) = exp(B)exp(A). \qquad (5.4.12)$$

Доведення. Порівняємо ряди

$$exp(A)exp(B) = (I + A + \frac{1}{2}A^2 + ...)(I + A + \frac{1}{2}B^2 + ...) =$$
$$= I + (A + B) + \frac{1}{2}(A^2 + 2AB + B^2) + \ldots$$

і

$$exp\,(A + B) = I + (A + B) + \frac{1}{2}(A + B)^2 + ... =$$
$$= I + (A + B) + \frac{1}{2}(A^2 + AB + BA + B^2) + \ldots$$

Оскільки $AB = BA$, то $AB + BA = 2AB$, а також співпадають члени в дужках, що не виписані (перевірте!). Отже, ряди збігаються. З їхньої абсолютної збіжності (це ми вже перевірили раніше) випливає вірність (5.4.12).

Наслідок:

$$exp((t + s)A) = exp(tA)\,exp(sA), \qquad (5.4.13)$$

де $t, s \in C$. Дійсно, оператори tA і sA комутують, і доведення випливає з (5.4.12). У свою чергу, з (5.4.13) можна одержати важливий наслідок:

$$(exp(tA))^{-1} = exp(-tA). \qquad (5.4.14)$$

Дійсно, з (5.4.13) і (5.4.2) одержимо $exp\,((t - t)A) = exp\,(tA)\,exp\,(-tA) = E$, що й доводить (5.4.14).

Властивості (5.4.2) — (5.4.14) означають, що множина операторів $\{exp(tA),\, t \in C\}$ утворює мультиплікативну абелеву групу, яка широко застосовується в теорії диференціальних рівнянь та диференціальній геометрії, якщо $t \in \mathbb{R}$.

Теорему 5.3.1 можна застосувати для скалярної матриці λI і нільпотентної клітини Жордана $J(0)$, які комутують, оскільки скалярна матриця комутує з будь-якою матрицею. Їхня сума — клітина Жордана з власними числами λ на діагоналі: $\lambda E + J(0) = J(\lambda)$. Отже,

$$\exp(tJ(\lambda)) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \dots & \frac{t^{m-1}}{(m-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \dots & \frac{t^{m-2}}{(m-2)!}e^{\lambda t} \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \cdots & te^{\lambda t} \\ 0 & 0 & 0 & \dots & e^{\lambda t} \end{pmatrix} \qquad (5.4.15)$$

Цікаво помітити, що $\det\left(\exp\left(J(\lambda)\right)\right) = \exp\left(m\lambda\right)$. Звідси для експоненти від довільної матриці, жордановою нормальною формою якої є пряма сума жорданових клітин, маємо формулу:

$$\det\left(\exp\left(A\right)\right) = \det\left(T\exp\left(\sum_{k=1}^{n}\lambda_k\right)T^{-1}\right) = \exp\left(SpA\right)$$

Зауважимо, що верхній рядок (5.4.15) містить базисні вектори простору квазімногочленів степені, меншої за *m*, розглянутого раніше, формула (5.4.15) дає розв'язок диференціального рівняння (5.4.5) при $A = J(\lambda)$.
Можна розглянути задачу Коші для лінійного диференціального векторного рівняння відносно вектор-функції $x(t) \in \mathbb{R}^m$.

$$\frac{d}{dt}x(t) = Ax(t),\ x(0) = x_0 \qquad (5.4.16)$$

Розв'язок (5.4.16) при $A = J(\lambda)$ знаходиться за допомогою (5.4.15), розв'язку (5.4.5):

$$x(t) = exp\ (tJ(\lambda))x_0. \qquad (5.4.17)$$

Тепер неважко розв'язати задачу Коші (5.4.16) з довільною матрицею A. Нехай T — матриця переходу до жорданового базису матриці A і при цьому $A = TA_JT^{-1},\ A_J = T^{-1}AT$. Зробимо заміну $x = Ty$. Тоді $y = T^{-1}x$. Позначимо $y_0 = T^{-1}x_0$ і після підстановки y в (5.4.16) помножимо обидві його частини на T:

$$T^{-1}\frac{d}{dt}\left(Ty(t)\right) = T^{-1}ATy(t),$$

$$y(0) = T^{-1}x_0 = y_0,$$

звідки

$$\frac{d}{dt}y(t) = A_J y(t),$$
$$y(0) = y_0. \qquad (5.4.18)$$

Отже, одержали рівняння з матрицею A_J в правій частині, яка має вигляд прямої суми жорданових клітин. Тому (5.4.16) розпадається у пряму суму тільки що розглянутих рівнянь. Воно має розв'язок $y(t) = exp(tA_J)y_0$, де матриця $exp(tA_J)$ - це пряма сума матриць вигляду (5.4.15). Зробивши обернену заміну, знайдемо розв'язок (5.4.16):

$$x(t) = T\exp(tA_j)T^{-1}x_0 \qquad (5.4.19)$$

Зрозуміло, що елементи матриці $T\exp(tA_J)T^{-1}x_0$ - лінійні комбінації елементів матриці exp (tA_J). Більш детально: нехай *x(t)* - *n*-вимірний вектор, розв'язок задачі Коші; $\lambda_1,...,\lambda_k$ — власні числа *A* кратностей $r_1,...,r_k$. Тоді компонента $x_j(t)$ розв'язку має вигляд

$$x_j(t) = \sum_{l=1}^{k} e^{\lambda_l t} p_{j,l}(t), \qquad (5.4.20)$$

де $p_{j,l}(t)$, — многочлен степеня, меншого за r_l.

Якщо деякі власні числа комплексні і мають вигляд $\alpha_j \pm i\omega_j$ (у випадку дійсної матриці, якщо λ — корінь характеристичного рівняння, то $\overline{\lambda}$ - теж його корінь), тоді згідно з формулою Ейлера (5.4.20) можна подати в такому вигляді:

$$x_j(t) = \sum_l e^{\lambda_l t} p_{j,l}(t) + \left(\sum_k e^{a_k t} q_{j,k}(t)\cos(\omega_k t) + s_{j,k}(t)\sin(\omega_k t) \right), \qquad (5.4.21)$$

Де $p_{j,l}(t)$, $q_{j,k}(t)$, $s_{j,k}(t)$ - многочлени з дійсними коефіцієнтами степенів, менших за кратності відповідних власних чисел.

Аналогічний результат можна сформулювати для розв'язків одного лінійного диференціального рівняння n-го порядку:

$$x^{(n)} = a_1 x^{(n-1)} + \ldots + a_n x \qquad (5.4.22)$$

Це рівняння після змін $x_1 = x, x_2 = x',..., x_к = x^k,...,x_n = x^{(n-1)}$ можна звести до системи з матрицею

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & . & \cdots & 1 \\ a_n & a_2 & . & \cdots & a_1 \end{pmatrix} \qquad (5.4.23)$$

Ця матриця має такі ж властивості, як і циклічна матриця (5.2.3), її характеристичний многочлен збігається з мінімальним, що має вигляд

$$M(\lambda) = \lambda^n + a_1\lambda^{n-1} + ...(-1)^n a_n \qquad (5.4.24)$$

Отже, щоб записати характеристичне рівняння матриці (5.4.23), яке також називається характеристичним рівнянням для (5.4.22), треба замість кожної похідної порядку k формально підставити в (5.4.22) k-ту степінь від λ. Розв'язок рівняння матиме вигляд

$$x(t) = \sum_{j=1}^{k} e^{\lambda_j t} p_j(t), \qquad (5.4.25)$$

а у випадку тільки дійсних коефіцієнтів (5.4.22) і наявності комплексних власних чисел у матриці (5.4.24):

$$x(t) = \sum_l e^{\lambda_l t} p_l(t) + \left(\sum_k e^{a_k t} q_k(t)\cos(\omega_k t) + s_k(t)\sin(\omega_k t) \right), \qquad (5.4.26)$$

Розглянемо тепер степеневу функцію від оператора $J(\lambda) = \lambda I + J(0)$:

$\left(J(\lambda)\right)^a = \left(\lambda I + J(0)\right)^a$. Обчислимо похідні у нулі функції $f(z) = (\lambda + z)^a$:

$$f'(0) = a\lambda^{a-1}, f''(0) = a(a-1)\lambda^{a-2},..., f^{(k)}(0) = a(a-1)...(a-k+1)\lambda^{a-k}.$$

Тоді

(5.4.27)

$$
\left(\lambda I + J(0)\right)^a = \begin{pmatrix} \lambda^a & a\lambda^{a-1} & \frac{a(a-1)\lambda^{a-2}}{2!} & \dots & \frac{a(a-1)...(a-m+2)\lambda^{a-m+1}}{(m-1)!} \\ 0 & \lambda^a & a\lambda^{a-1} & \dots & \frac{a(a-1)...(a-m+3)\lambda^{a-m+2}}{(m-2)!} \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \cdots & a\lambda^{a-1} \\ 0 & 0 & 0 & \dots & \lambda^a \end{pmatrix}
$$

Приклад 5.3.4. Нехай задано числа $a_1,...a_k$ і $x_1,...,x_k$, а при $n > k$ члени послідовності $\{x_i\}$ знаходяться за формулами

$$
x_n = a_1 x_{n-1} + ... + a_k x_{n-k}. \tag{5.4.28}
$$

Така послідовність називається зворотною, а співвідношення (5.4.28) - рекурентним. Перевірка показує, що відрізок послідовності $z_{n-1} = \{x_{n-k},...,x_{n-1}\}$ довжини k переводиться у відрізок $z_n = \{x_{n-k+1},...,x_n\}$ тієї ж довжини за допомогою матриці, яка відповідає (5.4.28):

$$
Az_{n-1} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & . & \cdots & 1 \\ a_n & a_2 & . & \cdots & a_1 \end{pmatrix} \begin{pmatrix} x_{n-k} \\ . \\ . \\ . \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} x_{n-k+1} \\ . \\ . \\ . \\ x_n \end{pmatrix} = z_n.
$$

Оскільки за допомогою математичної індукції одержимо, що $z_n = A^n z_k$, де $z_k = \{x_1,...,x_k\}$, то, знайшовши вигляд A^n, можна знайти формулу довільного члена послідовності. Для цього треба спочатку привести A до нормальної форми Жордана, а тоді після вже відомих міркувань (порівняйте з випадком рівняння (5.4.22)), користуючись виглядом степеневої функції для клітини Жордана, одержати вигляд шуканої формули загального члена:

$$
x_n = \sum_{i=1}^{k} \lambda_i^n p_i(n), \tag{5.4.29}
$$

де, як завжди, λ_i — власні числа матриці A кратності r_i; $p_i(n)$ — многочлен степені, меншої за r_i. Коефіцієнти многочленів можна знайти за методом невизначених коефіцієнтів. Власні числа матриці А знаходяться як корені характеристичного многочлена (5.4.24), що збігається з мінімальним.

Приклад 5.3.5. Розглянемо послідовність, в якій $x_o = x_1 = 1$; $x_n = x_{n-2} + x_{n-1}$ для $n>1$. Вона називається послідовністю Фібоначчі.

Матриця А, очевидно, має вигляд

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

Її характеристичне рівняння: $\lambda^2 = \lambda + 1$ або $\lambda^2 - \lambda - 1 = 0$.

Корені рівняння: $\lambda_1 = \dfrac{1-\sqrt{5}}{2}, \lambda_2 = \dfrac{1+\sqrt{5}}{2}$. Отже, $x_n = \alpha\lambda_1{}^n + \beta\lambda_2{}^n$, де числа α і β знайдемо за методом невизначених коефіцієнтів. Для цього покладемо n=0, n=1 і використаємо значення двох перших членів послідовності:

$$\alpha + \beta = 1, \alpha\lambda_1 + \beta\lambda_2 = 1.$$

Розв'язуючи систему відносно α і β, знаходимо

$$\alpha = \frac{\sqrt{5}-1}{2\sqrt{5}}, \beta = \frac{\sqrt{5}+1}{2\sqrt{5}},$$

Звідки

$$x_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{\sqrt{5}-1}{2}\right)^n\right).$$

Резольвента оператора. Нехай $\lambda \in \mathbb{C}$ не є власним числом оператора A. Тоді існує оператор

$$R(A,\lambda) = (A - \lambda I)^{-1}, \tag{5.4.30}$$

який називається резольвентою A. Резольвента - це дробово-раціональна функція від оператора. За аналогією з формулою суми членів нескінченно спадної геометричної прогресії $\left(|q|<1\right)$

$$S_\infty = 1 + q + q^2 + \ldots + q^n + \ldots = \frac{1}{1-q}$$

будемо шукати резольвенту у вигляді

$$R(A,\lambda)=(A-\lambda I)^{-1}=-\lambda^{-1}(I-\lambda^{-1}A)=-\lambda^{-1}(I+\lambda^{-1}A+...+\lambda^{-n}A^n+...).$$
(5.4.31)

Ряд абсолютно збігається, якщо $|\lambda|^{-1}\|A\|<1$, тобто, при

$$|\lambda|\rhd\|A\|. \tag{5.4.32}$$

Доведемо, що ця формула насправді зображує обернений оператор до $A-\lambda I$. Дійсно, після перемноження і приведення подібних членів маємо

$$-\lambda^{-1}(A-\lambda I)(I+\lambda^{-1}A+...+\lambda^{-n}A^n+...)=$$
$$=-\lambda^{-1}(A-\lambda I+\lambda^{-1}A^2-A+\lambda^2A^3-\lambda^{-1}A^2+...)=\lambda^{-1}\lambda I=I$$

Зауваження. Радіус збіжності ρ (A) ряда (5.4.29) можна знайти більш точно, користуючись відомою формулою Адамара, що приводить до такого вигляду: $\rho(A)=\overline{\lim\limits_{n\to\infty}}\sqrt[n]{\|A^n\|}$ (верхня границя). Але в даному випадку можна довести більш точну формулу:

$$\rho(A)=\lim_{n\to\infty}\sqrt[n]{\|A^n\|} \tag{5.4.33}$$

Тобто, існує границя послідовності, яка, зрозуміло, збігається з верхньою границею. Зауважимо, що (5.4.31) справджується також у нескінченновимірному випадку і у функціональному аналізі називається формулою Гельфанда.

Доведення. Позначимо $r=\inf\limits_{n\to\infty}\left(\|A^n\|^{1/n}\right)$ Очевидніно, $r\le\overline{\lim\limits_{n\to\infty}}\left(\|A^n\|^{1/n}\right)$.

Доведемо , що $r=\overline{\lim\limits_{n\to\infty}}\left(\|A^n\|^{1/n}\right)$. Для кожного $\varepsilon>0$ виберемо таке натуральне m, що
$\|A^m\|^{1/m}\le r+\varepsilon$. Тоді $\|A^m\|\le(r+\varepsilon)^m$. Далі, для довільного натурального $n\ge m$ нехай q — залишок від ділення n на m. Тоді $n=pm+q$, $0\le q\le m-1$ (p - ціле). Використовуючи нерівність для норм операторів: $\|XY\|\le\|X\|\|Y\|$, одержимо

$$\|A^n\|^{1/n}=\|A^{mp+q}\|^{1/n}=\|A^{mp}A^q\|^{1/n}\le$$
$$\le\|A^m\|^{p/n}\|A\|^{q/n}\le(r+\varepsilon)^{mp/n}\|A\|^{q/n}.$$

Оскільки $\lim\limits_{n\to\infty}(pm/n)=\lim\limits_{n\to\infty}(n-q)/n=1,\ \lim\limits_{n\to\infty}(q/n)=0,$, то $\overline{\lim\limits_{n\to\infty}}\left(\left\|A^n\right\|^{1/n}\right)\le r+\varepsilon$.

Число ε було вибране довільно, тому $\overline{\lim\limits_{n\to\infty}}\left(\left\|A^n\right\|^{1/n}\right)\le r$, а це і доводить існування границі і тим самим - формулу (5.4.31), де ρ (A) = r.

Доведемо, що $\rho(A)\le\|A\|$. Справді,

$$\rho(A)=\lim_{n\to\infty}\left(\left\|A^n\right\|^{1/n}\right)\le\lim_{n\to\infty}\left(\left\|A\right\|^{n(1/n)}\right)=\|A\|.$$

Вправа. Для довільних регулярних λ, μ довести рівність

$$R(A,\lambda)-R(A,\mu)=(\lambda-\mu)R(A,\lambda)R(A,\mu). \tag{5.4.34}$$

Доведення. Мають місце очевидні рівності:

$$R(A,\lambda)=R(A,\lambda)(A-\mu I)R(A,\mu),$$
$$R(A,\mu)=R(A,\lambda)(A-\lambda I)R(A,\mu).$$

Після їх віднімання одержимо

$$R(A,\lambda)-R(A,\mu)=R(A,\lambda)R(A,\mu)(A-\mu I-A+\lambda I)=$$
$$=(\lambda-\mu)R(A,\lambda)R(A,\mu).$$

Доведене співвідношення називається тотожністю Гільберта.

Наслідок. Резольвенти при різних регулярних значеннях λ і μ комутують. Доведення очевидне.

Перетворення Келі. Це ще одна функція від оператора – дробово-лінійна.

Нехай A - самоспряжений оператор в унітарному просторі E, $z=x+iy\in\mathbb{C},\ y=\mathrm{Im}(z)\ne 0$, $f,g,h\in E$. Нехай

$$\left(A-\bar{z}I\right)h=f$$
$$\left(A-zI\right)h=g$$

Недійсне число z (так само $\bar{z}$) не може бути власним числом самоспряженого оператора, а значить існує обернений оператор $\left(A-\bar{z}I\right)^{-1}$, визначений на всьому просторі. Тоді $h=\left(A-\bar{z}I\right)^{-1}f$ і $g=\left(A-zI\right)\left(A-\bar{z}I\right)^{-1}f$. Розглянемо оператор

$$U(z)=\left(A-zI\right)\left(A-\bar{z}I\right)^{-1}.$$

Називається перетворенням Келі самоспряженого оператора A.
З попередньої рівності $g=U(z)f=(A-xI)h-iyh$.
Доведемо унітарність $U(z)$.

Маємо: $\|f\|^2=\left\|\left(A-\bar{z}I\right)h\right\|^2=\left(\left(A-xI\right)h+iyh,\left(A-xI\right)h+iyh\right)=$

$$=\left\|\left(A-xI\right)h\right\|^2+\left\|yh\right\|^2-iy\left(\left(A-xI\right)h,h\right)+yi\left(h,\left(A-xI\right)h\right)=$$

$$=\left\|\left(A-xI\right)h\right\|^2+\left\|yh\right\|^2-iy\left(\left(A-xI\right)h,h\right)+yi\left(\left(A-xI\right)h,h\right)=\left\|\left(A-xI\right)h\right\|^2+\left\|yh\right\|^2$$

Так само доводиться, що $\left\|g\right\|^2=\left\|\left(A-xI\right)h\right\|^2+\left\|yh\right\|^2$, звідки

$$\left\|g\right\|^2=\|f\|^2=\left\|U(z)f\right\|^2$$

Тепер унітарність випливає з леми 2.3.1.

5.5. ОПЕРАТОР В ДІЙСНИХ ПРОСТОРАХ

Розглянемо випадок матриці з дійсними елементами, характеристичний многочлен якої має вигляд $\chi(A,t)=(t-\lambda)^r(t-\bar{\lambda})^r$, $\lambda=\alpha+i\beta,\bar{\lambda}=\alpha-i\beta\ (\beta\neq 0)$, і нехай $\dim\left(Ker(A-\lambda E)^r\right)=\dim\left(Ker(A-\bar{\lambda}E)^r\right)=1$. В такому випадку маємо ланцюг ядер:

$$R^r=Ker(A-\lambda E)^r\supset Ker(A-\lambda E)^r\supset\ldots\supset Ker(A-\lambda E)^2\supset Ker(A-\lambda E)\supset\left\{0\right\}$$

і такий самий для $\bar{\lambda}$.
Жорданова нормальна форма даної матриці має вигляд:

$$\begin{pmatrix} \lambda & 1 & \cdots & 0 & 0 & \cdots & & & & & \\ 0 & \lambda & \cdots & 0 & 0 & \cdots & & & & & \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & & & 0 & & \\ 0 & 0 & \cdots & \lambda & 1 & \cdots & & & & & \\ 0 & 0 & \cdots & 0 & \lambda & \cdots & & & & & \\ . & . & . & . & . & . & . & . & . & . & . \\ & & & & & \cdots & \bar{\lambda} & 1 & \cdots & 0 & 0 \\ & & & & & \cdots & 0 & \bar{\lambda} & \cdots & 0 & 0 \\ & & 0 & & & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots \\ & & & & & \cdots & 0 & 0 & \cdots & \bar{\lambda} & 1 \\ & & & & & \cdots & 0 & 0 & \cdots & 0 & \bar{\lambda} \end{pmatrix} \qquad (5.5.1)$$

Покажемо, як зіставити даній матриці аналог нормальної форми – матрицю з дійсними елементами. Нехай $f_r \in Ker(A-\lambda E)^r$, Тобто, $Ker(A-\lambda E)^r f_r = 0$, $(A-\lambda E) f_r = f_{r-1} \in Ker(A-\lambda E)^{r-1}, \ldots, (A-\lambda E) f_2 = f_1 \in Ker(A-\lambda E)^1$, $(A-\lambda E) f_1 = 0$.

f_r має комплексні координати, тому його можна зобразити у вигляді $f_r = g_r + ih_r$, де вектори g_r, h_r мають дійсні координати. Враховуючи це, перепишемо рівності у вигляді

$\big(A-(\alpha+i\beta)E\big)\big(g_r + ih_r\big) = g_{r-1} + ih_{r-1},$

$\big(A-(\alpha+i\beta)E\big)\big(g_{r-1} + ih_{r-1}\big) = g_{r-2} + ih_{r-2},$

$\ldots$

$\big(A-(\alpha+i\beta)E\big)\big(g_2 + ih_2\big) = g_1 + ih_1,$

$\big(A-(\alpha+i\beta)E\big)\big(g_1 + ih_1\big) = 0$

Звідси одержимо:

$$\begin{gathered} Ag_1 = \alpha g_1 - \beta h_1 \\ Ah_1 = \beta g_1 + \alpha h_1 \\ Ag_2 = g_1 + \alpha g_2 - \beta h_2 \\ Ah_2 = h_1 + \beta g_2 + \alpha h_2 \\ \ldots \\ Ag_r = g_{r-1} + \alpha g_r - \beta h_r \\ Ah_r = h_{r-1} + \beta g_r + \alpha h_r \end{gathered} \qquad (5.5.2)$$

В результаті знаходимо матрицю в новому базисі $\{g_k, h_k\}, k = 1,\ldots,r$:

$$[A]_{(g,h),(g,h)} = \begin{pmatrix} \alpha & \beta & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\beta & \alpha & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \alpha & \beta & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -\beta & \alpha & 0 & 1 & \cdots & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & \cdots & \alpha & \beta & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -\beta & \alpha & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \alpha & \beta \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -\beta & \alpha \end{pmatrix} \qquad (5.5.3)$$

Ця матриця схожа на жорданову клітину. На її діагоналі знаходяться блоки $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$, а над діагоналлю – блоки з одиничних матриць $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Для зручності позначимо $\bar{J}_\lambda = [A]_{(g,h),(g,h)}$, K – блочно-діагональна матриця з блоками $K_1 = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$ на діагоналі, а інші її елементи дорівнюють нулю.

Нехай $\rho = \sqrt{\alpha^2 + \beta^2}$. Тоді існує таке $\varphi \in [0, 2\pi)$, що $\cos\varphi = \dfrac{\alpha}{\rho} = \dfrac{\alpha}{\sqrt{\alpha^2 + \beta^2}}$, $\sin\varphi = \dfrac{\beta}{\rho} = \dfrac{\beta}{\sqrt{\alpha^2 + \beta^2}}$. Тоді K – блочно-діагональна матриця з блоками $\rho\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$ на діагоналі. Легко бачити, що K^m - блочно-діагональна матриця, на діагоналі якої знаходяться блоки $\rho^m\begin{pmatrix} \cos m\varphi & \sin m\varphi \\ -\sin m\varphi & \cos m\varphi \end{pmatrix}$.

Позначимо також

$$\bar{J}_0 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (5.5.4)$$

Легко бачити, що матриця $\bar{J}_0$ має такі ж властивості, як і нільпотентна клітина Жордана: при піднесенні її до квадрату блоки $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ зсуваються вправо-вгору і т.ін, її степінь r-1 має вигляд

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad (5.5.5)$$

тобто, блок $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ в ній займає місце в верхньому правому куті, а її степінь r дорівнює нульовій матриці.

Розглянемо матрицю $C = \begin{pmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{pmatrix}$ в канонічному базисі

$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Перейдемо до нового базису

$g_1 = Te_1 = \frac{1}{2}\left(e_1 + ie_2\right)$, $g_2 = Te_2 = \frac{1}{2i}\left(e_1 - ie_2\right)$. Для простоти позначимо матрицю цього оператора також через *T*. Маємо:

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{i}{2} & -\frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix},\ T^{-1} = \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}$$

і

$$T^{-1}CT = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}\begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = K_1. \quad (5.5.6)$$

Тоді $e^{K_1} = T^{-1}e^{C}T = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}\begin{pmatrix} e^{a+ib} & 0 \\ 0 & e^{a-ib} \end{pmatrix}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} =$

$$= \frac{e^a}{2}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}\begin{pmatrix} \cos b + i\sin b & 0 \\ 0 & \cos b - i\sin b \end{pmatrix}\begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = e^a\begin{pmatrix} \cos b & \sin b \\ -\sin b & \cos b \end{pmatrix}.$$

Отже, одержано експоненту матриці K_1:

$$\exp(K_1) = \exp\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = e^a\begin{pmatrix} \cos b & \sin b \\ -\sin b & \cos b \end{pmatrix} \qquad (5.5.7)$$

В теорії диференціальних рівнянь потрібна така експонента:

$$\exp(tK_1) = \exp t\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = e^{at}\begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix} \qquad (5.5.8)$$

Очевидно, $\exp(K)$ - блочна матриця, блоками якої є матриці $\exp(K_1)$. Знаходження $\exp\left(t\bar{J}_0\right)$ проводиться за такою самою схемою, як і при знаходженні експоненти нільпотентної клітини Жордана:

$$
\exp\left(t\bar{J}_0\right) = \begin{pmatrix}
1 & 0 & t/1! & 0 & t^2/2! & 0 & \cdots & t^{r-1}/(r-1)! & 0 \\
0 & 1 & 0 & t/1! & 0 & t^2/2! & \cdots & 0 & t^{r-1}/(r-1)! \\
0 & 0 & 1 & 0 & t/1! & 0 & \cdots & t^{r-2}/(r-2)! & 0 \\
0 & 0 & 0 & 1 & 0 & t/1! & \cdots & 0 & t^{r-2}/(r-2)! \\
. & . & . & . & . & . & . & . & . \\
0 & 0 & 0 & 0 & \cdots & 1 & 0 & t/1! & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & t/1! \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1
\end{pmatrix}. \tag{5.5.9}
$$

Остаточно

$$
\exp\left(t\bar{J}_\lambda\right) = e^{at}\begin{pmatrix}
\cos bt & \sin bt & . & . & . \\
-\sin bt & \cos bt & . & . & . \\
. & . & . & . & . \\
. & . & . & \cos bt & \sin bt \\
. & . & . & -\sin bt & \cos bt
\end{pmatrix} \times
$$

$$
\times \begin{pmatrix}
1 & 0 & t/1! & 0 & t^2/2! & 0 & \cdots & t^{r-1}/(r-1)! & 0 \\
0 & 1 & 0 & t/1! & 0 & t^2/2! & \cdots & 0 & t^{r-1}/(r-1)! \\
0 & 0 & 1 & 0 & t/1! & 0 & \cdots & t^{r-2}/(r-2)! & 0 \\
0 & 0 & 0 & 1 & 0 & t/1! & \cdots & 0 & t^{r-2}/(r-2)! \\
. & . & . & . & . & . & . & . & . \\
0 & 0 & 0 & 0 & \cdots & 1 & 0 & t/1! & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & t/1! \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1
\end{pmatrix} \tag{5.5.10}
$$

Проілюструємо у випадку матриці другого порядку з парою комплексно-спряжених власних чисел можливість зведення її до блочного вигляду. Нехай $g = g_1$ - власний вектор, що відповідає власному числу λ, тоді

$$Ag = \lambda g$$

Звідси автоматично випливає: $A\bar{g} = \bar{\lambda}\bar{g}$, і це означає, що $\bar{g} = g_2$ - другий власний вектор, який відповідає іншому власному числу $\bar{\lambda}$. Виділимо в векторі g дійсну і уявну частини: $g = f_1 + if_2$. Нехай $\lambda = \alpha + i\beta$. Перепишемо попередню рівність у вигляді $A\left(f_1 + if \right)_2 = (\alpha + i\beta)\left(f_1 + if \right)$. Це еквівалентне системі

$$Af_1 = \alpha f_1 - \beta f_2$$
$$Af_2 = \beta f_1 + \alpha f_2$$

В новому базисі $\{ f_1, f_2 \}$ оператор має матрицю

$$[A]_{f,f} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Проілюструємо все це на конкретному прикладі. Нехай

$$A = \begin{pmatrix} 0 & 5 \\ -5 & 8 \end{pmatrix}.$$

Її характеристичний многочлен має вигляд

$$\chi(A,\lambda) = \lambda^2 - 8\lambda + 25 = (\lambda - (4+3i))(\lambda - (4-3i)),$$

Тобто, вона має два комплексно-спряжені власні числа $\lambda = 4+3i,\ \bar{\lambda} = 4-3i$.

Розв'язуючи системи $\left(A - \left(4+3i\right)E\right)x = 0,\ \left(A - \left(4-3i\right)E\right)x = 0$, знаходимо два власні вектори, що відповідають цим власним числам:

$$g_1 = \begin{pmatrix} 5 \\ 4+3i \end{pmatrix},\ g_2 = \begin{pmatrix} 5 \\ 4-3i \end{pmatrix}.$$

В цьому базисі її нормальна форма Жордана є діагональна матриця

$$J_A = \begin{pmatrix} 4+3i & 0 \\ 0 & 4-3i \end{pmatrix}$$

Виділивши дійсні і уявні частини в базисних векторах, одержимо два базисні вектори в дійсному просторі:

$$g_1 = \begin{pmatrix} 5 \\ 4+3i \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} + i\begin{pmatrix} 0 \\ 3 \end{pmatrix},\ g_2 = \begin{pmatrix} 5 \\ 4-3i \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} - i\begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

Позначимо: $f_1 = \begin{pmatrix} 5 \\ 4 \end{pmatrix},\ f_2 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. Як тільки що було показано, в цьому базисі матриця оператора має вигляд

$$\begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix}.$$

Проведемо більш детальні дослідження. Знайдемо матриці переходу між базисами і т.н.
Нехай оператор *T* переводить базис g_1, g_2 в f_1, f_2, $Tg_1 = f_1, Tg_2 = f_2$. Обернений оператор T^{-1} діє навпаки. Попередні рівності означають, що

$$g_1 = T^{-1} f_1 = f_1 + if_2,$$
$$g_2 = T^{-1} f_2 = f_1 - if_2.$$

Отже, маємо матриці операторі T^{-1} і *T*:

$$[T]^{-1}{}_{f,g} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, [T]_{g,f} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Використовуючи формулу (2.3.11), знайдемо матрицю оператора у базисі f_1, f_2 :

$$[A]_{f,f} = \left([T]_{g,f}\right)^{-1} [A]_{g,g} [T]_{g,f}$$

$$[A]_{f,f} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} 4+3i & 0 \\ 0 & 4-3i \end{pmatrix}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix},$$

тобто, знову одержали матрицю у вигляді (5.5.6).

5.6. СИНГУЛЯРНИЙ РОЗКЛАД МАТРИЦІ

Розглянемо тепер прямокутні, не обов'язково квадратні матриці і узагальнимо деякі результати, одержані в розділі 5.2.

Теорема 5.6.1 (скелетний розклад матриці). Нехай $A \in M_{m,n}$, $rank(A) = r$. Існують матриці $B \in M_{m,r}$, $C \in M_{r,n}$ такі, що $A = BC$.

Доведення. Нехай *r* стовпчиками матриці *B* будуть незалежні стовпчики матриці *A*. Довільний *j*-ий стовпчик матриці *A* буде лінійною комбінацією цих стовпчиків з коефіцієнтами $c_{1,j}, c_{2,j}, ..., c_{r,j}$. Ці числа і будуть утворювати *j*-ий стовпчик матриці *C*, $j = 1,...,n$.

Як відомо (наслідок з формули Біне-Коші), $rank(A) \le rank(B), rank(C)$, але в даному випадку ранги матриць *B* і *C* не можуть перевищувати $r = rank(A)$, тому що це один з розмірів *B* і *C*. З цього випливає, що матриці B^*B, CC^* невироджені. Справді, нехай *x* – розв'язок системи $B^*Bx = 0$. Помножимо зліва обидві частини цієї рівності на $x^T \equiv x^*$: $x^*B^*Bx = 0$, звідки $(Bx)^*Bx = 0$, тобто, $\|Bx\|^2 = 0$ і $Bx = 0$. Оскільки ліва частина цієї рівності – це лінійна комбінація векторів-стовпчиків матриці B, а вони лінійно незалежні, то $x = 0$, що й доводить невиродженість матриці B^*B, звідки, в свою чергу, $\det(B^*B) \ne 0$. Аналогічно $\det(CC^*) \ne 0$.

Лема 5.6.1. Нехай $A \in M_{m,n}$, причому, $m \le n$, $rank(A) = k \le m$. Існують матриці: унітарна $X \in M_{m,m}$, діагональна $D \in M_{m,m}$ з невід'ємними діагональними упорядкованими елементами $\lambda_1 \ge \lambda_2 \ge ... \ge \lambda_k > \lambda_{k+1} = ... = \lambda_m = 0$ і $Y \in M_{m,n}$ з ортонормованими рядками, і має місце зображення:

$$A = XDY. \tag{5.6.1}$$

При цьому матриця D визначена однозначно і $\{\lambda_1^2, \lambda_2^2, ..., \lambda_m^2\}$ - це власні числа матриці AA^*; вектори-стовпчики матриці X - власні вектори матриці AA^*. Якщо всі власні числа AA^*, то матриця X визначена однозначно с точністю до правого множника $\tilde{D} = diag\{\exp(i\theta_1),...,\exp(i\theta_n)\}, \theta_1,...\theta_n \in \mathbb{R}$. При фіксованій матриці X матриця Y однозначно визначена, якщо $rank(A) = m$. Якщо A - дійсна матриця, тоді X, Y можуть бути вибрані дійсними.

Доведення. Нехай $A = XDY$ - розклад описаного типу. В такому разі $AA^* = XDYY^*DX^* = XD^2X^*$. Позначимо через x_j $(j = 1,...,m)$ вектори-стовпчики матриці AA^* і $D^2 = diag\{\lambda_1^2, \lambda_2^2, ..., \lambda_m^2\}$. Тоді

$AA^*x_j = \lambda_j^2 x_j, (j = 1,...,m)$, причому вектори x_j утворюють ортонормовану систему. Оскільки за домовленістю власні числа упорядковані за незростанням, то матриця D однозначно визначена матрицею AA^*. Якщо усі власні числа λ_j^2 різні, тоді ортонормовані власні вектори матриці AA^* визначені однозначно с точністю до множників одиничного модуля вигляду $\exp(i\theta)$. Якщо є інше зображення (5.6.1) з унітарною матрицею $X_1 \in M_{m,m}$, то $X_1 = X\tilde{D}$, де $\tilde{D}$ - діагональна матриця з діагональними елементами, норми яких дорівнюють одиниці, тобто, $\tilde{D} = diag\{\exp(i\theta_1),...,\exp(i\theta_n)\}, \theta_1,...\theta_n \in \mathbb{R}$.

Якщо матриця AA^* має кратні власні числа, то базис із власних векторів, що відповідає конкретному власному числу, вибирається неоднозначно, але якщо його вибрано і він ортонормований, то тим самим унітарну матрицю X фіксовано. В такій ситуації матриця $Y \in M_{m,n}$ в разі невиродженої матриці D (якщо $k = rank(A) = m$) знаходиться єдиним чином: $Y = D^{-1}X^{-1}A$. Перевіримо, що її вектори-рядки ортонормовані.

$$YY^* = D^{-1}X^{-1}AA^*XD^{-1} = D^{-1}X^{-1}XD^2X^{-1}XD^{-1} = D^{-1}D^2D^{-1} = I,$$

що й доводить ортонормованість рядків. Зазначимо, що у випадку не виродженої матриці A (а значить, і D) $Y = D^{-1}X^{-1}A = D^{-1}(A^*X)^*$. З останнього виразу випливає, що довільний j-й рядок y_j матриці Y можна вибирати як відповідний стовпчик y_j^*:

$$y_j = \lambda_j^{-1}(A^*x_j), j = 1,...,m. \tag{5.6.2}$$

Враховуючи, що вектори x_j утворюють ортонормований базис, маємо:

$$\left[\lambda_j^{-1}(A^*x_j)\right]^*\left[\lambda_k^{-1}(A^*x_k)\right] = (\lambda_j\lambda_k)^{-1}x_j^*AA^*x_k = (\lambda_j\lambda_k)^{-1}(\lambda_k)^2x_j^*x_k =$$

$$= \frac{\lambda_k}{\lambda_j}x_j^*x_k = \delta_{jk}.$$

Таким чином і вектори y_j утворюють ортонормований базис.

Розглянемо випадок виродженої матриці. Тепер $k = rank(A) < m$. Тепер вектори y_j можна знаходити за формулами (5.6.2) тільки при $j = 1,...,k < m$; їх можна доповнити до базису простору $\mathbb{C}^n$ (неоднозначно) деякими векторами $y_{k+1},...,y_m$. Тоді матриця $Y^* = [y_1, y_2,...y_k, y_{k+1},..., y_m] \in M_{n.m}$ матиме *m* ортонормованих стовпчиків.
Зазначимо, що $X^*A = DY$. Дійсно, з формул (5.6.2) випливає, що перші *k* рядків в цих матрицях рівні. Наступні *m-k* рядків в правій частині дорівнюють нулю, тому що відповідні діагональні елементи матриці D нульові. В лівій частині ті ж рядки нульові, оскільки з $AA^*x_j = 0$ випливає, що $x_j^* AA^* x_j = (A^* x_j)^* (A^* x_j) = 0$, тобто $A^* x_j = 0$.
Якщо матриця $A \in M_{m,n}$ дійсна, то тим більше AA^* дійсна і має дійсні (навіть невід'ємні) власні числа. Тоді перші *k* власних векторів, що утворюють матрицю X, можуть бути взяті дійсними. Інші *m-k* ортонормовані вектори теж можна вибрати дійсними. Отже, всі матриці в (5.6.1) можуть бути вибрані дійсними.
Теорема 5.6.2. Нехай $A \in M_{m,n}$, $m \le n$. Тоді A можна зобразити у вигляді

$$A = PU, \tag{5.6.3}$$

де матриця $P \in M_{m,m}$ невід'ємна, $rank(P) = rank(A)$ і матриця $U \in M_{m,m}$ унітарна. Матриця P визначається однозначно, $P = (AA^*)^{\frac{1}{2}}$; матриця P визначається однозначно, якщо $rank(A) = m$. Для дійсної матриці A і P, і U можна вибрати дійсними.
Доведення. Застосовуючи лему 5.6.1, подамо матрицю A у вигляді $A = XDY = XDX^*XY$. Покладемо $P = XDX^*$, $U = XY$. Тоді P невід'ємна і $UU^* = XYY^*X^* = XX^* = I$, отже, U унітарна. З леми (5.6.1) випливає, що $P = (AA^*)^{\frac{1}{2}}$, але взагалі, коли $A = PU$, то $AA^* = PUU^*P = P^2$, і P повинна бути коренем (єдиним) невід'ємним з AA^*. Якщо $rank(A) = m$, то P невироджена і матриця $U = P^{-1}A$ визначається однозначно. Але, як це було показано в лемі 5.6.1, при $rank(A) < m$ рядки матриці Y, що відповідають нульовим власним числам матриці P, припускають неєдиний вибір, тому і матрицю $U = XY$ визначено неоднозначно.
Наслідок. Матриця $A \in M_{n,n}$ може бути зображена у вигляді $A = PU$, де матриця $P \in M_{n,n}$ невід'ємна, визначається однозначно: $P = (AA^*)^{\frac{1}{2}}$; матриця $U \in M_{n,n}$ унітарна, вона визначається однозначно, якщо матриця A невироджена. Для дійсної матриці A і P, і U можна вибрати дійсними.

Зауваження. Цей факт для квадратних матриць вже було одержано в минулому розділі (формула (5.2.8)).

Теорема 5.6.3. Матриця $A \in M_{m,n}$ рангу k може бути зображена у вигляді

$$A = VSW^*, \tag{5.6.4}$$

де $V \in M_{m,m}, W \in M_{n,n}$ - унітарні матриці. Матриця $S = (s_{ij}), i = 1,...,m; \ j = 1,...,n$ має такі елементи: $s_{ij} = \delta_{ij}$, $s_{11} \geq s_{22} \geq ... \geq s_{kk} > s_{k+1,k+1} = ... = s_{qq}$, де $q = \min\{m,n\}$. Числа s_{ii} - це квадратні корені з власних чисел матриці AA^*, тобто, визначені однозначно. Стовпчики матриці V - власні вектори матриці AA^*, стовпчики матриці W - власні вектори матриці A^*A, обидві системи векторів впорядковані в залежності від власних значень s_{ii}^2. Якщо $m \leq n$ і всі власні числа матриці AA^* різні, то матриця V визначена з точністю з точністю до правого діагонального множника $\tilde{D} = diag\{\exp(i\theta_1),...,\exp(i\theta_n)\}, \theta_1,...\theta_n \in \mathbb{R}$. Якщо $m < n$, то матриця W визначається неоднозначно, якщо $m = n$ і матриця V фіксована, то W знаходиться однозначно. Якщо $n \leq m$ твердження щодо єдності матриць V і W можна одержати, застосовуючи наведені факти до матриці A^*. В разі дійсної матриці A всі матриці V,S,W можуть бути вибрані дійсними.

Доведення. Можна розглянути випадок $m \leq n$ (в противному разі всі міркування треба проводити для матриці A^*). Застосовуючи лему (5.6.1), подамо матрицю A у вигляді $A = XDY$, де $X \in M_{m,m}$, $D \in M_{m,m}$, $Y \in M_{m,n}$. Покладемо $V = X$. Візьмемо $S = (D \quad | \quad 0) \in M_{m,n}$ (тобто, дві матриці записано у рядок). Визначимо W як матрицю вигляду $S = (Y^* \quad | \quad S^*) \in M_{n,n}$, стовпчики якої повинні утворювати ортонормований базис простору $\mathbb{C}^n$. Стовпчики Y^* вже ортонормовані, тому при $m < n$ можна підібрати (неєдиним способом) стовпчики матриці $S^* \in M_{n,n-m}$ так, щоб матриця W була унітарною.

Очевидно, що $VSW^* = XDY = A$. Твердження відносно єдності і дійсності випливають з відповідних тверджень леми.

Елементи $s_{ii}\left(i = 1,...,q = \min\{m,n\}\right)$ матриці S називаються сингулярними числами матриці A. Стовпчики матриць V і W називаються лівим і правим сингулярними векторами матриці A. Розклад (5.6.4) називається сингулярним розкладом матриці A.

ГЛАВА 6
БІЛІНІЙНІ ТА КВАДРАТИЧНІ ФОРМИ

6.1. ЗАГАЛЬНІ ВІДОМОСТІ ТА ЗАКОН ІНЕРЦІЇ КВАДРАТИЧНИХ ФОРМ.

Означення 6.1.1. Відображення $\Omega : E_n \times E_n \to \mathbb{C}$, якщо

$$\begin{aligned}
&\forall x_1, x_2, y \in E_n,\ \alpha_1, \alpha_2 \in \mathbb{C} \\
&\Omega(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 \Omega(x_1, y) + \alpha_2 \Omega(x_2, y), \\
&\forall x, y_1, y_2 \in E_n,\ \beta_1, \beta_2 \in \mathbb{C} \\
&\Omega(x, \beta_1 y_1 + \beta_2 y_2) = \overline{\beta_1} \Omega(x, y_1) + \overline{\beta_2} \Omega(x, y_2)
\end{aligned} \tag{6.1.1}$$

$$\begin{aligned}
&\forall x_1, x_2, y \in E_n,\ \alpha_1, \alpha_2 \in \mathbb{C} \\
&\Omega(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 \Omega(x_1, y) + \alpha_2 \Omega(x_2, y), \\
&\forall x, y_1, y_2 \in E_n,\ \beta_1, \beta_2 \in \mathbb{C} \\
&\Omega(x, \beta_1 y_1 + \beta_2 y_2) = \overline{\beta_1} \Omega(x, y_1) + \overline{\beta_2} \Omega(x, y_2)
\end{aligned}$$

Якщо E_n - унітарний простір, можна записати білінійну форму у зручному вигляді.

Теорема 6.1.1. Білінійній формі в унітарному просторі відповідає лінійний оператор $A : E_n \to E_n$ такий, що $\Omega(x, y) = (Ax, y)$.

Доведення. Зафіксуємо елемент $x \in E_n$. Тоді вираз $\overline{\Omega(x, y)}$ буде лінійним функціоналом по другому аргументу. Тоді за теоремою 6.1.1 знайдеться такий елемент $z \in E_n$, що $\overline{\Omega(x, y)} = (y, z)$ або $\Omega(x, y) = \overline{(y, z)} = (z, y)$. Тобто, кожному $x \in E_n$ відповідає свій $z \in E_n$. Отже, задано оператор: $Ax = z$. Доведемо його лінійність. За означенням 6.1.1 білінійної форми
$\Omega(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 \Omega(x_1, y) + \alpha_2 \Omega(x_2, y)$. Тоді
$(A(\alpha_1 x_1 + \alpha_2 x_2), y) = \alpha_1 (Ax_1, y) + \alpha_2 (Ax_2, y)$, звідки
$(A(\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 Ax_1 + \alpha_2 Ax_2), y) = 0$ при довільному $y \in E_n$. Отже,
$A(\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 Ax_1 + \alpha_2 Ax_2) = 0$ і $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2$.
Далі будемо працювати в дійсному просторі. Тоді $\Omega(x, \beta y) = \beta \Omega(x, y)$.
Нехай $e = \{e_1, ..., e_n\}$ - базис в E_n (не обов'язково унітарному). Покладемо $\Omega(e_j, e_k) = a_{kj}$. Ми зустрічались з подібною ситуацією в унітарному просторі при знаходженні елементів матриці – формула (2.3.5), яка в данному випадку має вигляд $(Ae_j, e_k) = a_{kj}$. Остаточно:

$$\Omega\left(e_j, e_k\right) = \left(Ae_j, e_k\right) = a_{kj} \tag{6.1.2}$$

Нехай $x = \sum_{j=1}^{n} x_j e_j, y = \sum_{k=1}^{n} y_k e_k$. Тоді

$$\Omega\left(\sum_{j=1}^{n} x_j e_j, \sum_{k=1}^{n} y_k e_k\right) = \sum_{j=1}^{n} x_j \sum_{k=1}^{n} y_k \Omega\left(e_j, e_k\right) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{kj} x_j\right) y_k$$

Останній вираз можна інтерпретувати як добуток:

$$y^T Ax = \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ . & . & . \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{6.1.3}$$

Означення 6.1.2.. Рангом білінійної форми називається ранг її матриці.
Означення 6.1.3. Білінійна Ω форма називається симетричною, якщо $\Omega(y, x) = \Omega(x, y)$. Далі будемо розглядати саме такі форми.

Легко бачити, що в разі унітарного простору оператор *A*, що відповідає симетричній формі, буде самоспряженим.
Замість другого аргументу *y* в форму підставимо *x*. Будемо мати вираз $\Omega(x, x)$.

Означення 6.1.4. $\Omega(x, x)$ називається квадратичною формою.

Перепишемо в разі квадратичної форми формулу (6.1.3) і здійснемо множення, враховуючи, що $a_{kj} = a_{jk}$:

$$\begin{gathered} x^T Ax = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{1n} \\ . & . & . \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j,k=1}^{n} a_{kj} x_j x_k = \\ = a_{11}x_1^2 + \ldots + a_{nn}x_n^2 + 2x_1x_2 + \ldots + 2x_{n-1}x_n \end{gathered} \tag{6.1.4}$$

Як і в теорії самоспряжених операторів, в теорії квадратичних форм постає задача зведення її до канонічного вигляду:

$$\Omega(x, x) = a_{11}x_1^2 + \ldots + a_{nn}x_n^2. \tag{6.1.5}$$

Для цього треба, як і раніше, вибрати відповідний базис. Це можна зробити неєдиним способом. В попередньому виразі, взагалі кажучи, можуть бути ненульовими не всі коефіцієнти, якщо її ранг $r < n$. Тобто, форма зведеться до виразу

$$\Omega(x, x) = a_{11}x_1^2 + \ldots + a_{rr}x_r^2. \tag{6.1.6}$$

Якщо деякий коефіцієнт $a_{jj} > 0$, введемо заміну $\xi_j = \sqrt{a_{jj}} x_j$. Тоді $a_{jj}\left(x_j\right)^2 = \xi_j^2$; якщо $a_{kk} < 0$ - заміну $\eta_k = \sqrt{|a_{kk}|} x_k$. Тоді $a_{kk}\left(x_k\right)^2 = -|a_{kk}|\left(x_k\right)^2 = -\left(\sqrt{|a_{kk}|} x_k\right)^2 = -\eta_k^2$.
Остаточно, квадратична форма в новому базисі буде мати форму

$$\Omega(x,x) = \xi_1^2 + \ldots + \xi_l^2 - \eta_{l+1}^2 - \ldots - \eta_r^2, \tag{6.1.7}$$

яка називається нормальною.
Означення 6.1.5. Число додатних і від'ємних членів в формулі (6.1.7) називається відповідно додатним і від'ємним індексами форми. Різниця між додатним індексом і від'ємним називається сігнатурою форми.; вона позначається σ.
Теорема 6.1.2. Додатний і від'ємний індекси не залежать від вибору базису, в якому форма має нормальний вигляд, тобто, вони є інваріанти форми.
Доведення. Нехай квадратична форма може бути зведена до нормальних форм

$$\Omega(x,x) = y_1^2 + \ldots + y_l^2 - y_{l+1}^2 - \ldots - y_r^2 \tag{6.1.8}$$

і

$$\Omega(x,x) = z_1^2 + \ldots + z_m^2 - z_{m+1}^2 - \ldots - z_r^2 \tag{6.1.9}$$

Припустимо, що $l > m$. Звичайно, змінні z_i $\left(i = 1,\ldots,n\right)$ можуть бути лінійно виражені через y_j:

$$z_i = \sum_{j=1}^{n} t_{ij} y_j\ , \tag{6.1.10}$$

причому, матриця коєфіцієнтів невироджена.
Підставимо вирази (6.1.10) в (6.1.8). Маємо тотожність:

$$\begin{aligned} &y_1^2 + \ldots + y_l^2 - y_{l+1}^2 - \ldots - y_r^2 = \\ &= z_1^2 + \ldots + z_m^2 - z_{m+1}^2 - \ldots - z_r^2 \end{aligned} \tag{6.1.11}$$

В сумах (6.1.10) при $i = 1,\ldots,m$ візьмемо перші l доданків і прирівняємо їх до нуля. Будемо мати однорідну систему m рівнянь з l невідомими, $l > m$:

$$\sum_{j=1}^{l} t_{ij} y_j = 0,\ i = 1,\ldots,m. \tag{6.1.12}$$

Вона має нетривіальний розв'язок $y_1,\ldots,y_l$. Покладемо додатково $y_{l+1} = \ldots = y_r = y_{r+1} = \ldots = y_n = 0$. В результаті одержимо:

$$y_1^2+\ldots+y_l^2=-z_{m+1}^2-\ldots-z_r^2,$$
але це неможливо, оскільки ліва частина строго додатна, а права невід'ємна. Отже, припущення $l>m$ невірне. Аналогічно спростовується протилежне припущення. Отже, $l=m$, а це й треба було довести.

6.2. ЗВЕДЕННЯ КВАДРАТИЧНОЇ ФОРМИ ДО КАНОНІЧНОГО ВИГЛЯДУ

Нехай є квадратична форма (6.1.4) з симетричною матицею в базисі $e=\{e_1,\ldots,e_n\}$. Розглянемо її мінори, що називаються кутовими:

$$\Delta_1=a_{11},\Delta_2=\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix},\Delta_3=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix},\ldots,\Delta_n=\det(A) \tag{6.2.1}$$

Для зручності введемо величину $\Delta_0=1$.

Будемо шукати новий базис $e'=\{e_1',\ldots,e_n'\}$, вектори якого виражаються через вектори базису e за формулами:

$$\begin{cases}e_1'=t_{11},\\ e_2'=t_{12}e_1+t_{22}e_2,\\ \ldots\\ e_k'=t_{1k}e_1+t_{2k}e_2+\ldots+t_{kk}e_k,\\ \ldots\\ e_n'=t_{1n}e_1+t_{2n}e_2+\ldots+t_{nn}e_n.\end{cases} \tag{6.2.2}$$

Для зведення квадратичної форми до канонічного вигляду треба накласти умови:

$$\Omega\left(e_j',e_k'\right)=a_{kj}'=0, \tag{6.2.3}$$

$j=1,\ldots,k-1$. Тоді $a_{jk}'=0$ завдяки симетричності форми. Розглянемо більш детально умову (6.2.3).

$$\Omega\left(e_j',e_k'\right)=\Omega\left(t_{1j}e_1+t_{2j}e_2+\ldots+t_{kk}e_k,e_k'\right)=t_{1k}\Omega\left(e_1,e_k'\right)+t_{2k}\Omega(e_2,e_k')+\ldots+t_{kk}\Omega\left(e_k,e_k'\right).$$

Зрозуміло, що для виконання умови (6.2.3), достатньо виконання

$$\Omega\left(e_j,e_k'\right)=0, \tag{6.2.4}$$

$j=1,...,k-1;\ k=1,...,n.$ Оскільки треба звести квадратичну форму до нормального вигляду, треба накласти умову $\Omega\left(e_k^{'},e_k^{'}\right)=a_{kj}^{'}=1$. Але, зважаючи на (6.2.4), досить вимагати

$$\Omega\left(e_k,e_k^{'}\right)=1, \tag{6.2.5}$$

Доведемо, що для цієї процедури зведення можна вибирати коефіцієнти в (6.2.2) так, що для діагональних коефіцієнтів буде виконуватись властивість

$$t_{kk}=\frac{\Delta_{k-1}}{\Delta_k}, \tag{6.2.6}$$

$k=1,...,n$.

При $k=1$ вона виконується, оскільки $1=\Omega\left(e_1,e_1^{'}\right)=t_{11}\Omega\left(e_1,e_1\right)=t_{11}a_{11}$, звідки

$$t_{11}=\frac{1}{a_{11}}=\frac{\Delta_0}{\Delta_1} \tag{6.2.7}$$

Нехай визначено всі коефіцієнти, що входять в перші $k-1$ рядків в (6.2.2) і доведено формулу $t_{k-1,k-1}=\frac{\Delta_{k-2}}{\Delta_{k-1}}$. Для знаходження коефіцієнтів у *k*-му рядку запишемо умови (6.2.4) і (6.2.5):

$$\Omega\left(e_1,e_k^{'}\right)=\Omega\left(e_2,e_k^{'}\right)=...=\Omega\left(e_{k-1},e_k^{'}\right)=0,\ \Omega\left(e_k,e_k^{'}\right)=1. \tag{6.2.8}$$

Розглянемо ситуацію більш детально. При цьому будемо використовувати властивість $a_{jk}=a_{kj}$.

$\Omega\left(e_1,e_k^{'}\right)=\Omega\left(e_1,t_{11}e_1+t_{21}e_2+...+t_{k1}e_k\right)=t_{11}\Omega\left(e_1,e_1\right)+t_{21}\Omega\left(e_1,e_2\right)+...+t_{k1}\Omega\left(e_1,e_k\right)=$

$=t_{11}\Omega\left(e_1,e_1\right)+t_{21}\Omega\left(e_1,e_2\right)+...+t_{k1}\Omega\left(e_1,e_k\right)=$

$t_{11}\Omega\left(e_1,e_1\right)+t_{21}\Omega\left(e_1,e_2\right)+...+t_{k1}\Omega\left(e_1,e_k\right)=a_{11}t_{11}+a_{12}t_{21}+...+a_{1k}t_{k1},$

…

$t_{1,k-1}\Omega\left(e_{k-1},e_1\right)+t_{2,k-1}\Omega\left(e_{k-1},e_2\right)+...+t_{k,k-1}\Omega\left(e_{k-1},e_k\right)=a_{k-1,1}t_{1,k-1}+a_{k-1,2}t_{2,k-1}+...+a_{k-1,k}t_{k,k-1}$

,

$t_{1,k}\Omega\left(e_k,e_1\right)+t_{2,k}\Omega\left(e_k,e_2\right)+...+t_{k,k}\Omega\left(e_k,e_k\right)=a_{k1}t_{1k}+a_{k2}t_{2k}+...+a_{kk}t_{kk}.$

Отже, маємо:

$$\begin{cases} a_{11}t_{11} + a_{12}t_{21} + ... + a_{1k}t_{k1} = 0, \\ \dots \\ a_{k-1,1}t_{1,k-1} + a_{k-1,2}t_{2,k-1} + ... + a_{k-1,k}t_{k,k-1} = 0, \\ a_{k1}t_{1k} + a_{k2}t_{2k} + ... + a_{kk}t_{kk} = 1. \end{cases} \tag{6.2.9}$$

Коефіцієнти $t_{k1},...,t_{k,k-1},t_{kk}$ знаходяться з цієї системи. Використаємо формулу Крамера для знаходження t_{kk}:

$$t_{kk} = \frac{1}{\Delta_k}\begin{vmatrix} a_{11} & \dots & a_{1k} \\ . & \dots & . \\ a_{k-1,1} & \dots & a_{k-1,k-1} \end{vmatrix} = \frac{\Delta_{k-1}}{\Delta_k}, \tag{6.2.10}$$

що й треба було довести.
Підрахуємо визначник Δ трикутної системи (6.2.2):
$\Delta = t_{11}t_{22...}t_{kk} = \frac{\Delta_0}{\Delta_1}\frac{\Delta_1}{\Delta_2}...\frac{\Delta_{k-2}}{\Delta_{k-1}}\frac{\Delta_{k-1}}{\Delta_k} = \frac{1}{\Delta_k}$. Отже, перехід від старого базису до нового здійснюється за допомогою невиродженої матриці.

Означення 6.2.1. Квадратична форма називається додатно визначеною, якщо

$$\Omega(x,x) > 0 \ \forall x \neq 0. \tag{6.2.11}$$

Означення 6.2.2. Квадратична форма називається від'ємно визначеною, якщо

$$\Omega(x,x) < 0 \ \forall x \neq 0 \tag{6.2.12}$$

Якщо простір унітарний, то квадратична форма записується у вигляді скалярного добутку: $\Omega(x,x) = (Ax,x)$.

Означення 6.2.3. Оператор *A* називається додатним, якщо відповідна йому квадратична форма додатно визначена, і від'ємним – якщо відповідна йому квадратична форма від'ємно визначена.

Легко довести дві необхідні умови додатності матриці.

1. Нехай матриця додатна, тоді всі її діагональні елементи додатні.

Доведення. Справді, з додатної визначеності відповідної квадратичної форми випливає: $a_{ii} = (Ae_i, e_i) > 0 \ \forall e_i = (0,...,0,1,...,0)$ (одиниця знаходиться на *i*–му місці).

2. Нехай матриця додатна, тоді її визначник додатний.

Доведення. Симетричну матрицю A можна привести до діагональної матриці A' за допомогою унітарної матриці U (5.2.6): $A = UA'U^{-1}$. Визначник діагональної матриці дорівнює добутку діагональних елементів, які за попередньою властивістю додатні. Отже

$$\det(A) = \det(U)\det(A')\det(U^{-1}) = \det(U)\det(A')(\det(U))^{-1} = \det(A') > 0 .$$

Теорема 6.2.1. (Критерій Сільвестра). Для того, щоб квадратична форма була додатно визначеною, необхідно і достатньо, щоб усі її кутові мінори були додатні.

Доведення. Необхідність. Звузимо квадратичну форму на лінійний підпростір $E_k \subset E_n$, базис в якому $\{e_1,...,e_k\} \subset \{e_1,...,e_n\}$. Координати вектора з E_k будуть мати вигляд $[x]_e = (x_1,...,x_k,0,...,0)^T$, а визначник матриці звуження форми

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ . & \dots & . \\ a_{k,1} & \dots & a_{k,k} \end{vmatrix} > 0 \tag{6.2.13}$$

що випливає з тільки що доведеної властивості. Це справедливо для $k = 1,...,n$, отже, необхідність доведено.

Достатність. Методом Якобі квадратична форма зводиться до вигляду

$$\Omega(x,x) = \frac{\Delta_0}{\Delta_1}\left(x_1'\right)^2 + \frac{\Delta_1}{\Delta_2}\left(x_2'\right)^2 + ... + \frac{\Delta_{n-1}}{\Delta_n}\left(x_n'\right)^2 . \tag{6.2.14}$$

Якщо $x \neq 0$, то хоча б одна координата $x_k' \neq 0$, а тоді $\Omega(x,x) > 0$.

Наслідок 1. Для від'ємної визначенності квадратичної форми необхідно і достатньо, щоб знаки мінорів чергувались таким чином:

$$\Delta_0 < 0,\ \Delta_1 > 0, \Delta_2 < 0,... \tag{6.2.15}$$

Для доведення досить розглянути додатну квадратичну форму $-\Omega(x,x)$ з додатною матрицією $-A$.

Наслідок 2. Розглянемо двохвимірний випадок.

$$\Omega(x,x) = a_{11}\left(x_1\right)^2 + 2a_{12}x_1x_2 + a_{22}\left(x_2\right)^2 . \tag{6.2.16}$$

Для додатної визначенності цієї квадратичної форми необхідно і достатньо, щоб

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - \left(a_{12}\right)^2 > 0 \tag{6.2.17}$$

Зауважимо, що з цього автоматично випливає, що $a_{22} > 0$.

ЗВЕДЕННЯ КВАДРАТИЧНОЇ ФОРМИ ДО КАНОНІЧНОГО ВИГЛЯДУ МЕТОДОМ ЛАГРАНЖА

Нехай в деякому базисі задано квадратичну форму, яку запишемо таким чином:

$$\Omega(x,x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + ... + 2a_{1n}x_1x_n + F\left(x_2,...,x_n\right), \tag{6.2.18}$$

причому, форма $F\left(x_2,...,x_n\right)$ не містить x_1.
Введемо нові змінні за формулами

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n, \\ y_i &= x_i \ \forall i = 2,...,n. \end{aligned} \tag{6.2.19}$$

Обчислимо

$$\frac{1}{a_{11}}y_1^2 = \frac{1}{a_{11}}\left(a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n\right)^2 = a_{11}x_1^2 + 2a_{12}x_1x_2 + ... + 2a_{1n}x_1x_n + G\left(x_2,...,x_n\right),$$

де $G\left(x_2,...,x_n\right)$ - квадратична форма, що не містить x_1.
Отже,

$$\Omega(x,x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + ... + 2a_{1n}x_1x_n + F\left(x_2,...,x_n\right) = \frac{1}{a_{11}}y_1^2 - G\left(x_2,...,x_n\right) + F\left(x_2,...,x_n\right)$$

і з квадратичною формою $\Phi\left(x_2,...,x_n\right) = F\left(x_2,...,x_n\right) - G\left(x_2,...,x_n\right)$, якщо в ній коефіцієнт при x_2^2 не дорівнює нулю, можна зробити таку ж саму процедуру.
Якщо при всіх квадратах змінних коефіцієнти не дорівнюватимуть нулю, то квадратична форма приведеться до діагонального вигляду.
Розглянемо випадок, коли всі коефіцієнти при квадратах змінних дорівнюють нулю. Тоді квадратична форма буде мати вигляд

$$\Omega(x,x) = 2a_{12}x_1x_2 + ... + 2a_{1n}x_{n-1}x_n. \tag{6.2.20}$$

Введемо заміни:

$$
\begin{aligned}
y_1 &= x_1 + x_2 \\
y_2 &= x_1 - x_2 \\
y_k &= x_k \ \forall k = 3, \ldots, n
\end{aligned}
\tag{6.2.21}
$$

Зауважимо, що старі змінні можна однозначно виразити через нові, тому перетворення невироджене. В результаті квадратична форма буде мати вигляд

$$\Omega(x, x) = 2a_{12} y_1^2 - 2a_{12} y_2^2 + \ldots \tag{6.2.22}$$

Зауважимо, що члени, що не виписані, не містять y_1^2, тому доданок $2a_{12} y_1^2$ не може зникнути при зведенні подібних. Далі ми приходимо до випадку, розглянутого раніше. Таким чином, квадратична форма зведеться до канонічного вигляду.

6.3. ЗВЕДЕННЯ РІВНЯННЯ КРИВОЇ ДРУГОГО ПОРЯДКУ ДО ГОЛОВНИХ ОСЕЙ

Розглянемо рівняння другого порядку від двох змінних

$$a_{11}\xi_1^2 + 2a_{12}\xi_1\xi_2 + a_{22}\xi_2^2 + 2a_{13}\xi_1 + 2a_{23}\xi_2 + a_{33} = 0\,. \tag{6.3.1}$$

Ввівши позначення

$$
\begin{gathered}
A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \\
\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},\ a = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix},\ b = a_{33}
\end{gathered}
\tag{6.3.2}
$$

перепишемо рівняння у вигляді

$$\left(A\xi, \xi\right) + 2\left(a, \xi\right) + b = 0 \tag{6.3.3}$$

A – симетрична матриця, в якій $a_{12} \neq 0$ (інакше квадратична форма $\Omega(x,x) \equiv (Ax,x)$ мала б діагональний вигляд) .
Зробимо перетворення зсуву так, щоб позбавитись від лінійних членів. Нехай $\xi = x + r$. Рівняння перепишеться у вигляді

$$\left(Ax,x\right) + 2\left(Ar + a, x\right) + \left(\left(Ar,r\right) + 2\left(a,r\right) + b\right) = 0. \tag{6.3.4}$$

Якщо $\det(A) \neq 0$, можна знайти вектор зсуву з умови

$$Ar = -a\,. \tag{6.3.5}$$

(Випадок $\det(A) = 0$ розглянемо окремо.)
Якщо ввести позначення $\left(Ar,r\right) + 2\left(a,r\right) + b = -C$, рівняння перепишеться у вигляді

$$\left(Ax,x\right) = C \tag{6.3.6}$$

Треба шляхом повороту координатних осей звести його до канонічного вигляду. Задача відповідає зведенню до діагональної форми матриці A.
Зведемо квадратичну форму у новому базисі $x' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$ до канонічного вигляду з матрицею

$$A' = \begin{pmatrix} a_{11}' & 0 \\ 0 & a_{22}' \end{pmatrix}. \tag{6.3.7}$$

Тоді рівняння перепишеться у вигляді

$$(A'x', x') = a_{11}'\left(x_{11}'\right)^2 + a_{22}'\left(x_{22}'\right)^2 = C \tag{6.3.8}$$

Слід відзначити, що при $C = 0$ і різних знаках власних чисел λ_1, λ_2 рівняння визначає пару прямих, що перетинаються; якщо знаки λ_1, λ_2 однакові, то дійсних розв'язків рівняння не має, і кажуть про пару уявних прямих.
З загальної теорії відомо, що $a_{11}' = \lambda_1,\ a_{22}' = \lambda_2$, де λ_1, λ_2 - власні числа матриці; вони знаходяться як розв'язки характеристичного рівняння

$$\lambda^2 - \left(a_{11} + a_{22}\right)\lambda + a_{11}a_{22} - a_{12}^2 = 0. \tag{6.3.9}$$

Якщо зразу виписати канонічну форму у вигляді (6/3/8), немає гарантії, що здійснено саме поворот координатних осей. Можливий випадок, коли дане перетворення – це поворот і дзеркальне відображення, що не відповідатиме геометричній постановці задачі.
Нехай здійснено поворот осей на кут φ. Матиця, яка нові координати переводить у старі, згідно з (2.3.26), має вигляд:

$$U(\varphi)=\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \tag{6.3.10}$$

$$U(\varphi)x' = x. \tag{6.3.11}$$

Запишемо в нашій ситуації матрицю в старих координатах через матрицю в нових і навпаки (див. (2.3.19), (2.3.20)):

$$A = U(\varphi)A'U^{-1}(\varphi) = U(\varphi)A'U^*(\varphi), \tag{6.3.12}$$

$$A' = U^{-1}(\varphi)AU(\varphi) = U^*(\varphi)AU(\varphi). \tag{6.3.13}$$

(Раніше для ортогональної матриці було доведено, що $U^*(\varphi)=U^{-1}(\varphi)$.)
Враховуючи, що $x'=U^{-1}(\varphi)x=U^*(\varphi)x$, маємо:

$$(Ax,x)=\left(U(\varphi)A'U^*(\varphi)x,x\right)=\left(A'U^*(\varphi)x,U^*(\varphi)x\right)=(A'x',x')=$$
$$=\left(U^*(\varphi)AU(\varphi)x',x'\right).$$

Виберемо кут φ таким чином, щоб добуток матриць $U^*(\varphi)AU(\varphi)$ мав би вигляд (6.3.7). Знайдемо

$$U^*(\varphi)AU(\varphi)=\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}=$$
$$=\begin{pmatrix} a_{11}\cos^2\varphi+2a_{12}\cos\varphi\sin\varphi+a_{22}\sin^2\varphi & a_{12}\cos^2\varphi+\left(a_{22}-a_{11}\right)\cos\varphi\sin\varphi-a_{12}\sin^2\varphi \\ a_{12}\cos^2\varphi+\left(a_{22}-a_{11}\right)\cos\varphi\sin\varphi-a_{12}\sin^2\varphi & a_{22}\cos^2\varphi-2a_{12}\cos\varphi\sin\varphi+a_{11}\sin^2\varphi \end{pmatrix}$$

Тепер зручно записати

$$A'=U^*(\varphi)AU(\varphi)\ =$$
$$=\cos^2\varphi\begin{pmatrix} a_{11}+2a_{12}tg\varphi+a_{22}tg^2\varphi & a_{12}+\left(a_{22}-a_{11}\right)tg\varphi-a_{12}tg^2\varphi \\ a_{12}+\left(a_{22}-a_{11}\right)tg\varphi-a_{12}tg^2\varphi & a_{22}-2a_{12}tg\varphi+a_{11}tg^2\varphi \end{pmatrix}=$$
$$=\frac{1}{1+tg^2\varphi}\begin{pmatrix} a_{11}+2a_{12}tg\varphi+a_{22}tg^2\varphi & a_{12}+\left(a_{22}-a_{11}\right)tg\varphi-a_{12}tg^2\varphi \\ a_{12}+\left(a_{22}-a_{11}\right)tg\varphi-a_{12}tg^2\varphi & a_{22}-2a_{12}tg\varphi+a_{11}tg^2\varphi \end{pmatrix}$$

Одержана матриця повинна мати діагональний вигляд. Для цього розв’яжемо рівняння:

Одержана матриця повинна мати діагональну форму. Для цього досить, щоб $tg\varphi$ задовольняв рівнянню

$$a_{12}tg^2\varphi-\left(a_{22}-a_{11}\right)tg\varphi-a_{12}=0. \tag{6.3.14}$$

Ми покажемо, що його зовсім не обов'язково розв'язувати.

Знайдемо, наприклад, $a'_{11}=\lambda_1=\dfrac{a_{11}+2a_{12}tg\varphi+a_{22}tg^2\varphi}{tg^2\varphi+1}$.

при умові $a_{12}tg^2\varphi=\left(a_{22}-a_{11}\right)tg\varphi+a_{12}$.

$$\lambda_1=\frac{a_{11}+2a_{12}tg\varphi+a_{22}tg^2\varphi}{1+tg^2\varphi}=a_{22}+a_{12}\frac{2a_{12}tg\varphi+a_{11}-a_{22}}{\left(a_{22}-a_{11}\right)tg\varphi+2a_{12}}$$

Звідси, пропускаючи елементарні перетворення, маємо:

$$tg\varphi=a_{12}\frac{\left(a_{11}+a_{22}-2\lambda\right)}{\left(a_{22}-a_{11}\right)\lambda+a_{11}a_{22}-a_{22}^2-2a_{12}^2}=\frac{a_{12}}{\lambda-a_{22}}\frac{\left(a_{11}+a_{22}-2\lambda\right)\left(\lambda-a_{22}\right)}{\left(a_{22}-a_{11}\right)\lambda+a_{11}a_{22}-a_{22}^2-2a_{12}^2}=$$

$$=\frac{a_{12}}{\lambda-a_{22}}\frac{-2\left(\lambda^2-\left(a_{11}+a_{22}\right)\lambda+a_{11}a_{22}-a_{12}^2\right)+\left(a_{22}-a_{11}\right)\lambda+a_{11}a_{22}-a_{22}^2-2a_{12}^2}{\left(a_{22}-a_{11}\right)\lambda+a_{11}a_{22}-a_{22}^2-2a_{12}^2}=$$

$$=\frac{a_{12}}{\lambda-a_{22}}.$$

Отже,

$$tg\varphi=\frac{a_{12}}{\lambda-a_{22}}, \tag{6.3.15}$$

або аналогічно

$$tg\varphi=\frac{\lambda-a_{11}}{a_{12}}. \tag{6.3.16}$$

Користуючись формулами (2.3.31), можна записати рівняння нових координатних осей Ox_1', Ox_2', в координатах x_1, x_2.

Вісь Ox_1' має рівняння $x_2'=0$, вісь Ox_2' - $x_1'=0$.

$$Ox_1': x_2=\left(tg\varphi\right)x_1 \tag{6.3.17}$$

$$Ox_2^{'} : x_2 = -\frac{1}{tg\varphi} x_1$$

Оскільки $\xi = x + r$, $x = \xi - r$, маємо остаточно :

$$O\xi_1^{'} : \xi_2 = (tg\varphi)(\xi_1 - r_1) + r_2$$

$$O\xi_2^{'} : \xi_2 = -\frac{1}{tg\varphi}(\xi_1 - r_1) + r_2 \qquad (6.3.18)$$

Треба визначити, який кут повороту вибрати.

1). $a'_{11} = \lambda_1, a'_{22} = \lambda_2$ мають однакові знаки. Дійсній кривій відповідає випадок

$$\frac{\lambda_1}{C} > 0, \frac{\lambda_2}{C} > 0.$$

Зрозуміло, що це еліпс. Щоб він мав канонічний вигляд, треба, щоб $|\lambda_1| = |a'_{11}| < |a'_{22}| = |\lambda_2|$ - тоді його горизонтальна піввісь буде більша за вертикальну. Саме з цих міркувань вибирається $tg\varphi$, наприклад, в формулі (6.3.15).

2). a'_{11}, a'_{22} мають різні знаки. Тоді крива є гіперболою. Для того, щоб вона мала канонічне рівняння, треба, щоб

$$\frac{\lambda_1}{C} > 0, \frac{\lambda_2}{C} < 0$$

3). Параболічний випадок . Він відповідає $tg\varphi = -\frac{a_{12}}{a_{22}}$.

Приклад 1. $5x^2 + 8xy + 5y^2 - 18x - 18y + 9 = 0$.

$$A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}, \ \lambda_1, \lambda_2 \in \{1;9\}, \ a = \begin{pmatrix} -9 \\ -9 \end{pmatrix}.$$

Власні числа додатні, тому ця крива є еліпс. Отже вибираємо $\lambda_1 < \lambda_2 \Rightarrow \lambda_1 = 1, \lambda_2 = 9$.

Для здійснення зсуву за формулою (6.3.5), знайдемо обернену матрицю:

$$A^{-1} = \frac{1}{9}\begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}$$

Тепер $r = -A^{-1}a = -\frac{1}{9}\begin{pmatrix} 5 & -4 \\ -4 & 5 \end{pmatrix}\begin{pmatrix} -9 \\ -9 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$-C = (Ar, r) + 2(a, r) + b = \left(\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) - 18\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) + 9 = -9.$$

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x'+1 \\ y'+1 \end{pmatrix}$ і в системі $\begin{pmatrix} x' \\ y' \end{pmatrix}$ рівняння запишеться у вигляді:

$$\left(\begin{pmatrix}5 & 4\\ 4 & 5\end{pmatrix}\begin{pmatrix}x'\\ y'\end{pmatrix},\begin{pmatrix}x'\\ y'\end{pmatrix}\right)=9.$$

За формулою (6.3.15) $tg\varphi=\frac{4}{1-5}=-1$. Зручно вважати, що $\cos\varphi>0$, тому $\varphi=-\frac{\pi}{4}$, $\cos\varphi=\frac{1}{\sqrt{2}}$, $\sin\varphi=-\frac{1}{\sqrt{2}}$

Отже, записуємо канонічне рівняння в координатах

$$\begin{cases}x''=\frac{1}{\sqrt{2}}x'-\frac{1}{\sqrt{2}}y';\\ y''=\frac{1}{\sqrt{2}}x'+\frac{1}{\sqrt{2}}y'\end{cases}$$

$$\frac{\left(x''\right)^2}{3^2}+\frac{\left(y''\right)^2}{1^2}=1.$$

Рівняння нових координатних осей в системі координат (x,y):

$$Ox'':\ y=-\left(x-1\right)+1\Rightarrow y=-x+2,$$

$$Oy'':\ y=x-1+1\Rightarrow y=x.$$

Приклад 2. $6xy+8y^2-12x-26y+11=0$.

$A=\begin{pmatrix}0 & 3\\ 3 & 8\end{pmatrix}$, $a=\begin{pmatrix}-6\\ -13\end{pmatrix}$, $b=11$.

Знайдемо вектор зсуву системи координат:

$$\begin{pmatrix}0 & 3\\ 3 & 8\end{pmatrix}\begin{pmatrix}r_1\\ r_2\end{pmatrix}=\begin{pmatrix}6\\ 13\end{pmatrix}.$$

Звідси $r_1=-1,\ r_2=2$; $\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x'-1\\ y'+2\end{pmatrix}$, $\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}x+1\\ y-2\end{pmatrix}$.

$\left(Ar,r\right)+2(a,r)+b=-C$,

$$\left(\begin{pmatrix}0 & 3\\ 3 & 8\end{pmatrix}\begin{pmatrix}-1\\ 2\end{pmatrix},\begin{pmatrix}-1\\ 2\end{pmatrix}\right)+2\begin{pmatrix}-6\\ 2\end{pmatrix}\begin{pmatrix}-1\\ -13\end{pmatrix}+11=-9=-C.$$

Звідси $C=9$. В нових координатах рівняння запишеться у вигляді

$$\left(\begin{pmatrix}0 & 3\\ 3 & 8\end{pmatrix}\begin{pmatrix}x'\\ y'\end{pmatrix},\begin{pmatrix}x'\\ y'\end{pmatrix}\right)=9.$$

Характеристичне рівняння: $\lambda^2+8\lambda-9=0$, власні числа матриці $\lambda_1,\lambda_2\in\{-1,9\}$. Отже, крива – гіпербола. Для того, щоб вона мала канонічний вигляд, треба взяти $\lambda_1=9,\lambda_2=-1$. Отже, прийдемо до канонічного рівнянн:

$$9\left(x''\right)^2-\left(y''\right)^2=9\Rightarrow\frac{\left(x''\right)^2}{1}-\frac{\left(y''\right)^2}{3^2}=1;$$

$$tg\varphi = \frac{a_{12}}{\lambda_1 - a_{22}} = \frac{3}{9-8} = 3.$$

Рівняння осі $Ox'': y - 2 = 3(x+1) \Rightarrow y = 3x + 5$,

осі $Oy'': y - 2 = -\frac{1}{3}(x+1) \Rightarrow y = -\frac{1}{3}x + \frac{5}{3}$.

Параболічний випадок.

Маємо рівняння кривої другого порядку $(Ax, x) + 2(a, x) + b = 0$, де A - симетрична матриця, $\det(A) = 0$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Одне з власних чисел матриці буде дорівнювати нулю, друге позначимо λ.

Як і раніше, перейдемо до нової системи координат $x' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$ за допомогою ортогональної матриці

$$U(\varphi) = \begin{pmatrix} \cos\varphi & = \sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$$

Рівняння перепишеться у вигляді

$$(U^*(\varphi) A U(\varphi) x', x') + 2(a, U(\varphi) x') + b = 0.$$

Виберемо $U(\varphi)$ так, що $A' = U^*(\varphi) A U(\varphi) = \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix}$.

Як відомо, $tg\varphi = \frac{a_{12}}{\lambda - a_{22}} = -\frac{a_{12}}{a_{22}}$ - тут врахували, що одне власне число матриці дорівнює нулю.

Позначимо $a' = U^*(\varphi) a$.Рівняння перепишеться у вигляді

$$\lambda (x_2')^2 + 2a_1' x_1' + 2a_2' x_2' + b = 0$$

а). Нехай $a_1' \neq 0$.

Після виділення повного квадрату і простих перетворень одержимо:

$$\lambda \left(x_2' + \frac{a_2'}{\lambda} \right)^2 + 2\frac{a_1'}{\lambda} \left(x_1' + \frac{1}{2a_1'} \left(b - \frac{(a_2')^2}{\lambda} \right) \right) = 0.$$

Отже, ввівши заміни

$$x_2'' = x_2' + \frac{a_2'}{\lambda},\ x_1'' = x_1' + \frac{1}{2a_1'} \left(b - \frac{(a_2')^2}{\lambda} \right)$$

і позначення $p=\frac{a_1^{'}}{\lambda}$, прийдемо до канонічного рівняння параболи:

$$\left(x_2^{''}\right)^2=2px_1^{''}.$$

b). $a_1^{'}=0$.

Після виділення повного квадрату і простих перетворень одержимо:

$$\left(x_2^{'}+\frac{a_2^{'}}{\lambda}\right)^2+\frac{1}{\lambda^2}\left(b-\frac{\left(a_2^{'}\right)^2}{\lambda}\right)=0$$

і

$$x_2^{'}=\frac{1}{\lambda}\left(-a_2^{'}\pm\sqrt{\frac{\left(a_2^{'}\right)^2}{\lambda}-b}\right).$$

Це рівняння пари прямих, якщо $\frac{\left(a_2^{'}\right)^2}{\lambda}-b\neq 0$. Воно вироджується в одну пряму, якщо $\frac{\left(a_2^{'}\right)^2}{\lambda}-b=0$.

Приклад 3. Звести рівняння другого порядку до головних осей:

$$x^2+2xy+y^2-8x+4=0.$$

$A=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$, власні числа цієї матриці: $\lambda_1=0,\ \lambda_2=2$, $tg\varphi=-\frac{a_{12}}{a_{22}}=-1.$

Візьмемо $\cos\varphi=\frac{1}{\sqrt{2}}$, $\sin\varphi=-\frac{1}{\sqrt{2}}$. Тоді

$$U(\varphi)=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix},\quad U^*(\varphi)=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$$

$$\begin{pmatrix}x\\ y\end{pmatrix}=U(\varphi)\begin{pmatrix}x_1^{'}\\ x_2^{'}\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}x_1^{'}\\ x_2^{'}\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}x_1^{'}+x_2^{'}\\ x_1^{'}-x_2^{'}\end{pmatrix}$$

$$\begin{pmatrix}x_1^{'}\\ x_2^{'}\end{pmatrix}=U^*(\varphi)\begin{pmatrix}x\\ y\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}x-y\\ x+y\end{pmatrix}$$

$$\begin{pmatrix}a_1^{'}\\ a_2^{'}\end{pmatrix}=U^*(\varphi)\begin{pmatrix}a_1\\ a_2\end{pmatrix}=-\frac{4}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}=\begin{pmatrix}-2\sqrt{2}\\ 2\sqrt{2}\end{pmatrix}.$$

$$x_2'' = x_2' + \frac{a_2'}{\lambda} = x_2' + \sqrt{2}\,, \quad x_1'' = x_1' + \frac{1}{2a_1'}\left(b - \frac{\left(a_2'\right)^2}{\lambda} \right) = x_1' - \frac{1}{4\sqrt{2}}\left(4 - \frac{8}{2}\right) = x_1'\,.$$

В результаті прийдемо до канонічного рівняння:

$$\left(x_2''\right)^2 = 2\sqrt{2}\,x_1',$$

де

$$x_2'' = x_2' + \frac{a_2'}{\lambda} = \frac{x + y + 2}{\sqrt{2}}$$

$$x_1'' = \frac{x - y}{\sqrt{2}}$$

ЛІТЕРАТУРА

1. Дубовик В. П. Вища математика: навч. посіб. для студ. вищ. навч. зак. / В.П Дубовик., І. І. Юрик. - 4-те вид. - К. : Ігнатекс-Україна., 2013. - 648 с.

2. Мартиненко М. А. Теорія функцій комплексної змінної. Операційне числення: навч. посіб. Для студ. Вищ. Навч. зак./М. А. Мартиненко, І. І. Юрик. – К.: Видавничий Дім «Слово», 2007. – 296 с.

3. Вища математика. Збірник задач : навч. посіб. / В. П. Дубовик, І. І. Юрик, І. П. Вовкодав, В. І. Дев'ятко, Р. К. Клименко, В. В. Крочук, М. А. Мартиненко ; за ред. В. П. Дубовика, І. І. Юрика. – К. : А.С.К., 2011. – 480 с.

4. ДОДАТКОВА

1. Александров П.С. Курс аналитической геометрии и линейной алгебры: учебник / П. С. Александров. - М.: «Наука», 1979, 511 с.
2. Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры: учебник / Д. В. Беклемишев. - М.: «Наука», 1984, 175 с.
3. Беллман Р. Введение в теорию матриц: учебник / Р. Беллман. - М.: «Наука», 1969, 367 с.
4. Блох Э. Л. Основы линейной алгебры и некоторые ее приложения: учебное пособие / Э. Л. Блох, Л. И. Лошинский, В. Я. Турин. - М.: Высшая школа, 1971. 256 с.
5. Бубнов В.А. Линейная алгебра: компьютерный практикум / В.А. Бубнов, Г. С. Толстова, О. Е. Клемешева. - М.: ЛБЗ, 2012. 168 с.
6. Ван дер Варден Б. Л. Алгебра: учебник / Б. Л. Ван дер Варден. - М.: «Наука», 1976. 648 с.
7. Винберг Э.Б. Курс алгебры: учебник / Э. Б. Винберг. - М.: Изд-во «Факториал Пресс», 2001. 544 с.
8. Воеводин В.В. Линейная алгебра: учебник / В. В. Воеводин. - М.: «Наука», 1974. 336 с.
9. Гантмахер Ф.Р. Теория матриц: монография / Ф. Р. Гантмахер. - М.: «Наука», 1967. 575 с.
10. Гельфанд И.М. Лекции по линейной алгебре: учебник / И. М. Гельфанд. - М.: «Наука», 1971. 271 с.
11. Глазман И.М., Любич Ю.И. Конечномерный линейный анализ: учебник / И. М. Глазман, Ю. И. Любич. -: «Наука», 1969. 475 с.
12. Головина Л.И. Линейная алгебра и некоторые ее приложения: учебник / Л. И. Головина. - М.: Наука, 1985. 392 с.

13.Гомонов, С.А. Математика. Линейная алгебра: Учебно-справочное пособие / С. А. Гомонов. - М.: Форум, НИЦ ИНФРА-М, 2013. 144 с.
14.Горлач, Б.А. Линейная алгебра: Учебное пособие / Б.А. Горлач. - СПб.: Лань, 2012. 480 с.
15.Ефимов Н.В. Линейная алгебра и многомерная геометрия: учебник / Н. В. Ефимов, Э. Р. Розендорн. - М.: «Наука», 1970. 528 с.
16.Икрамов Х.Д. Задачи по линейной алгебре: сборник задач / Х. Д. Икрамов. - М.: «Наука», 1975. 319 с.
17.Ильин В.А., Позняк Э.Г. Линейная алгебра: учебник / В.А. Ильин, З. Г. Позняк. - М.: «Наука», 1974. 296 с.
18.Каган В.Ф. Основания теории определителей: учебник / В. Ф. Каган. - Гос. изд-во Украины, 1922. 78 с.
19.Кострикин А. И. Линейная алгебра и геометрия: учебник / А. И. Кострикин, Ю. И. Манин. - М.: «Наука», 1986. 303 с.
20.Кочетков, Е.С. Линейная алгебра: учебное пособие /Е. С. Кочетков, А. В. Осокин. - М.: Форум, 2012. 416 с.
21.Курбатова Г.И. Курс лекций по алгебре: учебник / Г. И. Курбатова, В.Б.Филиппов. - СПб: Изд-во «Лань», 2015. 655 с.
22.Курош А. Г. Курс высшей алгебры: учебник / А. Г. Курош. - М.: «Наука», 1968. 431 с.
23.Курош А.Г. Лекции по общей алгебре: учебник / А. Г. Курош. - М.: «Наука», 1973. 400 с.
24.Ланкастер П. Теория матриц: монография/ М. Ланкастер.- «Наука», 1978. 280 с.
25. Ленг С. Алгебра: учебник / С. Ленг. - М.: «МИР», 1968. 564 с.
26.Лефор Г. Алгебра и анализ. Задачи: сборник задач / Г. Лефор. - М.: «Наука», 1973. 462 с.
27.Мальцев А.И. Основы линейной алгебры: учебник / А. И. Мальцев. - М.: «Наука», 1970. 400 с.
28.Попов А.М. Лекции по линейной алгебре: учебник / А. М. Попов. - М.: Изд-во РУДН, 2007. 183 с.
29.Постников М.М. Лекциии по геометрии. Семестр II. Линейная алгебра: учебник / М. М. Постников. - Москва, «Наука», 1986. 400 с.
30.Прасолов В.В. Задачи и теоремы линейной алгебры: учебное пособие / В. В. Прасолов. - М.: МЦНМО, 2015. 576 с.
31.Проскуряков И.В. Сборник задач по линейной алгебре. / И. В. Проскуряков. - М.: «Наука», 1978. 384 с.
32.Рудык, Б.М. Линейная алгебра: Учебное пособие / Б. М. Рудык, - М.: НИЦ ИНФРА, 2013. 318 с.

33. Самойленко А.М., Перестюк М.О., Парасюк І.О. Диференціальні рівняння: підручник / А. М. Самойленко, М. О. Перестюк, І. О. Парасюк. - К.: «Либідь», 2003. 599 с.
34. Смирнов В. И. Курс высшей математики: учебник в 8 т. / В. И. Смирнов. – М.: Физмат «Наука», 1974, Т.3, ч.1. 323 с.
35. Стренг Г. Линейная алгебра и ее приложения: учебное пособие / Г. Стренг. - М.: Мир, 1980. 454 с.
36. Фаддеев Д. К. Лекции по алгебре: учебник / Д. К. Фаддеев. - М.: «Наука», 1984. 416 с.
37. Фаддеев Д.К. Сборник задач по высшей алгебре / Д. К. Фаддеев, И. С. Соминский. - М.: «Наука», 1977. 288 с.
38. Хорн Р. Матричный анализ: монография / Р. Хорн, Ч. Джонсон. - М.: «Мир», 1989. 655 с.
39. Шафаревич И.Р., Ремизов А.О. Линейная алгебра и геометрия: учебник / И. Р. Шафаревич, А. О. Ремизов. - М.: «Физматлит», 2009. 511 с.
40. Шевцов, Г.С. Линейная алгебра: теория и прикладные аспекты: Учебное пособие / Г. С. Шувцов. - М.: Магистр, НИЦ ИНФРА-М, 2013. 528 с.
41. Шилов Г.Е. Математический анализ. Конечномерные линейные пространства : учебник / Г. Е. Шилов. - М.: «Наука», 1969. 431 с.