

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

ЛІНІЙНА АЛГЕБРА: КУРС ЛЕКЦІЙ

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського
як навчальний посібник для здобувачів ступеня бакалавра
за освітньою програмою «Інтелектуальні сервіс-орієнтовані розподілені обчислення»
спеціальності 122 «Комп'ютерні науки»*

Київ

КПІ ім. Ігоря Сікорського

2022

Лінійна алгебра: Курс лекцій [Електронний ресурс]: курс лекц. для студ. спеціальності 122 «Комп'ютерні науки» / КПІ ім. Ігоря Сікорського; уклад. Ю. Є. Бохонов. – Електронні текстові дані (1 файл: 3,91 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2022. – 214 с.

*Гриф надано Методичною радою КПІ ім. Ігоря Сікорського (протокол № 5 від 26.05.2022 р.)
за поданням Вченої ради Інституту прикладного системного аналізу (протокол № 4 від 25.04.2022 р.)*

Електронне мережне навчальне видання

ЛІНІЙНА АЛГЕБРА: КУРС ЛЕКЦІЙ

Укладач:	<i>Бохонов Юрій Євгенович</i> , канд. ф. - мат. н, доцент
Відповідальний редактор	<i>Подколзін Г.Б.</i> , канд. ф. - мат. н., доцент
Рецензент:	<i>Голуб А.П.</i> , д-р ф.-мат. н., провідний науковий співробітник Інституту математики НАН України

Курс лекцій «Лінійна алгебра: Курс лекцій» призначено для вивчення студентами спеціальності 122 «Комп'ютерні науки» основ аналітичної геометрії та лінійної алгебри. Викладено теорію унітарних просторів, самоспряжених та унітарних операторів. Особливу увагу приділено приведенню матриць до нормальної форми Жордана та її застосуванню при визначенні матриць операторів. Наводяться численні приклади зведення до нормальної форми Жордана матриць третього і четвертого порядку. Розглядається також циклічна форма матриці. Вивчаються різні способи зведення квадратичних форм до канонічного вигляду. Алгебраїчний апарат застосовується до задач, що виникають у аналітичній геометрії. Викладено теорію поверхонь другого порядку, їхнього зведення до головних осей. Також викладено деякі питання теорії многочленів та дробово-раціональних функцій.

© КПІ ім. Ігоря Сікорського, 2022

ЗМІСТ

Вступ	5
1. ЛЕКЦІЯ II.1. АЛГЕБРА МНОГОЧЛЕНІВ, ДІЛЕННЯ З ОСТАЧЕЮ, АЛГОРИТМ ЕВКЛІДА	6
1.1.Схема Горнера ділення многочлена на біном	12
2. ЛЕКЦІЯ II.2. РЕЗУЛЬТАНТ	14
Дискримінант	24
3. ЛЕКЦІЯ II.3. ІНТЕРПОЛЯЦІЙНИЙ МНОГОЧЛЕН ЛАГРАНЖА; ДРОБОВО-РАЦІОНАЛЬНІ ФУНКЦІЇ	26
4. ЛЕКЦІЯ II.4. УНІТАРНІ ТА ЕВКЛІДОВІ ПРОСТОРИ	30
4.1.Нерівність Коші-Буняковського	32
4.2.Ортогональна сума підпросторів	44
4.3.ЛІНІЙНІ ФУНКЦІОНАЛИ	47
5. ЛЕКЦІЯ II.5. СПРЯЖЕНИЙ ОПЕРАТОР ТА ЙОГО МАТРИЦЯ. САМОСПРЯЖЕНІ, НОРМАЛЬНІ ТА УНІТАРНІ ОПЕРАТОРИ	52
6. Ізоморфізм між дуальними просторами	59
7. ЛЕКЦІЯ II.6. ВЛАСНІ ЧИСЛА І ВЕКТОРИ ЛІНІЙНОГО ОПЕРАТОРА	56
7.1.Анулюючі многочлени матриці лінійного оператора	60
8. ЛЕКЦІЯ II.7. КЛІТИНА ЖОРДАНА ТА ЇЇ ВЛАСТИВОСТІ	69
9. ЛЕКЦІЯ II.8. ЗВЕДЕННЯ МАТРИЦІ ДО НОРМАЛЬНОЇ ФОРМИ ЖОРДАНА	73
10.ЛЕКЦІЯ II.9. ЖОРДАНОВА НОРМАЛЬНА ФОРМА МАТРИЦЬ ТРЕТЬОГО ТА ЧЕТВЕРТОГО ПОРЯДКУ	79
11.ЛЕКЦІЯ II.10. ЖОРДАНОВА НОРМАЛЬНА ФОРМА МАТРИЦІ НОРМАЛЬНОГО ОПЕРАТОРА, ПОЛЯРНИЙ РОЗКЛАД	88
12.ЛЕКЦІЯ II.11. ОПЕРАТОРИ В ДІЙСНИХ ПРОСТОРАХ	105
13.ЛЕКЦІЯ II.12. ФУНКЦІЇ ВІД МАТРИЦЬ	113
14.ЛЕКЦІЯ II.13. ЦИКЛІЧНА НОРМАЛЬНА ФОРМА МАТРИЦІ	121
	135

15.ЛЕКЦІЯ II.14. БІЛІНІЙНІ ТА КВАДРАТИЧНІ ФОРМИ	138
15.1. Загальні відомості та закон інерції квадратичних форм	138
15.2. ЗВЕДЕННЯ КВАДРАТИЧНОЇ ФОРМИ ДО КАНОНІЧНОГО ВИГЛЯДУ МЕТОДОМ ЛАГРАНЖА	145
16.ЛЕКЦІЯ II.15. ЗВЕДЕННЯ КВАДРАТИЧНОЇ ФОРМИ ДО КАНОНІЧНОГО ВИГЛЯДУ МЕТОДОМ ЯКОБІ	146
16.1. Критерій Сільвестра	146
17.ЛЕКЦІЯ II.16. КАНОНІЧНІ РІВНЯННЯ ПОВЕРХОНЬ ДРУГОГО ПОРЯДКУ	154
18.ЛЕКЦІЯ II.17. ЛІНІЙЧАТІ ПОВЕРХНІ	170
19.ЛЕКЦІЯ II.18. ПРИВЕДЕННЯ ЗАГАЛЬНОГО РІВНЯННЯ ПОВЕРХНІ ДО КАНОНІЧНОГО ВИГЛЯДУ	177
19.1. Класифікація поверхонь другого порядку	183
20.ДОДАТОК 1. КУТИ ЕЙЛЕРА, ГРУПА $SO(3)$	193
21.ДОДАТОК II. ТЕНЗОРНІ ДОБУТКИ МАТРИЦЬ	197
22.ЛІТЕРАТУРА	208
23.ПРЕДМЕТНИЙ ПОКАЖЧИК	212

ВСТУП

Курс лекцій «Лінійна алгебра: Курс лекцій» поєднує в собі вивчення основних алгебраїчних структур, насамперед, основ теорії скінченновимірних лінійних унітарних просторів і лінійних операторів в них, теорію білінійних і квадратичних форм зі стандартними задачами аналітичної геометрії. При цьому потрібний алгебраїчний апарат вже вивчено для просторів довільної розмірності, що дає змогу їхнього ефективного застосування в геометрії.

При викладенні теорії лінійних унітарних просторів особлива увага приділяється властивостям скалярного добутку, зображенню лінійних функціоналів, методу Грама-Шмідта ортогоналізації систем векторів.

Викладено теорію самоспряжених, унітарних, ортогональних та нормальних операторів, їхнім власним числам.

Для загальних квадратних матриць викладено з усіма доведеннями теорію їхнього зведення до нормальної форми Жордана. Наводяться приклади такого зведення для матриць третього та четвертого порядків. Окремо викладено теорію нормальної форми у дійсному випадку. Також вивчається циклічна форма матриці. Викладена теорія застосовується до побудови функцій від матриць.

В курсі лекцій розглянуто різні методи зведення квадратичної форми до канонічного вигляду, зокрема, Лагранжа та Якобі, доводиться закон інерції квадратичних форм. Методика Якобі дозволяє довести критерій Сільвестра додатності симетричної матриці. Також доводиться критерій невід'ємності такої матриці.

Розвинений алгебраїчний апарат далі застосовується при дослідженні об'єктів аналітичної геометрії – поверхонь другого порядку. Розібрано канонічні рівняння таких поверхонь, окремо розглянуто лінійчаті поверхні.

Особливу увагу приділено зведенню поверхонь другого порядку до головних осей і вичерпній класифікації, в тому числі, у вироджених випадках.

Також в курсі приділено увагу теорії многочленів та дробово-раціональних функцій. Вивчається методика розв'язання системи двох цілих рівнянь з двома невідомими із застосуванням результанта.

В курсі вміщено додатки для самостійної роботи студентів – група обертань трьохвимірного простору за допомогою кутів Ейлера

ЛЕКЦІЯ II.1. АЛГЕБРА МНОГОЧЛЕНІВ, ДІЛЕННЯ З ОСТАЧЕЮ, АЛГОРИТМ ЕВКЛІДА

План

- 1). Ділення многочленів з остачею.
- 2). Формули Вієта.
- 3). Алгоритм Евкліда.
- 4). Схема Горнера.

В цій лекції будуть розглянуті питання існування коренів многочлена і висновки, що звідси випливають. Будемо розглядати многочлени над полем дійсних або комплексних чисел. Але поки що будемо давати означення у випадку коефіцієнтів многочлена з довільного поля K .

Нагадаємо, що лінійний простір многочленів є також кільцем, яке позначається $Z[x]$.

Означення 1.1. Поле називається алгебраїчно замкненим, якщо кожен многочлен з коефіцієнтами з нього має принаймні один корінь, який належить до цього поля.

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

Теорема 1.1. Довільний многочлен з комплексними коефіцієнтами має хоча б один комплексний корінь.

Цей факт називається основною теоремою алгебри. Його доведення виходить за межі нашого курсу.

Отже, згідно з основною теоремою алгебри, поле комплексних чисел алгебраїчно замкнене.

Читач повинен перевірити, що поля $\mathbb{R}, \mathbb{Q}$ алгебраїчно незамкнені.

Означення 1.2. Підмножина I кільця R називається ідеалом, якщо для довільного $a \in R$ і довільного $i \in I \Rightarrow ai \in I$. 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 $I = RI$. Дійсно, кожен елемент вигляду $\forall x \in R, a \in I \Rightarrow (xa^{-1})a = x \in I \Rightarrow I = R$. Отже, можна зробити важливий висновок: для того щоб ідеал не збігався з кільцем, необхідно і досить, щоб він не містив елементів, які мають обернені, зокрема одиниці кільця.

Приклади ідеалів:

1). $n\mathbb{Z} \subset \mathbb{Z}n$. Доведення очевидне.

2). Множина $P(x)Z[x]$, де $P(x)$ – деякий многочлен в кільці многочленів $Z[x]$ від однієї змінної. Доведення очевидне.

Означення 1.3. Ідеал називається головним, якщо він утворений одним елементом.

Так, ідеали в прикладах 1–2 головні.

Означення 1.4. Кільце, в якому всі ідеали головні, називається кільцем головних ідеалів.

Розглянемо більш детально кільце многочленів від однієї змінної $Z[x]$.

Нагадаємо, що многочлен $P(x)$ ділиться на многочлен $Q(x)$ з залишком, якщо

$P(x) = Q(x)S(x) + R(x).$	(1.1)
---------------------------	-------

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

Теорема 1.2 (Безу). Число a є коренем многочлена тоді і тільки тоді, коли він ділиться на $x - a$ без остачі, інакше він дає остачу при діленні, яка дорівнює значенню многочлена в точці a .

Доведення. Згідно з (1.1)

$$P(x) = (x - a)Q(x) + R.$$

Підставляючи сюди $x = a$, одержимо: $P(a) = R$. Це й доводить твердження.

З основної теореми алгебри і теореми Безу випливає, що многочлен з комплексними коефіцієнтами степені n має саме n коренів, причому, вони можуть повторюватись. При цьому

$P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_k x^{n-k} + \dots + a_n = a_0 (x - x_1)^{r_1} (x - x_2)^{r_2} \dots (x - x_m)^{r_m},$ $r_1 + r_2 + \dots + r_m = n.$	(1.2)
--	---------

Степінь r_k в розкладі (1.2) виразу $(x - x_k)^{r_k}$ називається кратністю кореня x_k ; якщо кратність дорівнює одиниці, корінь називається простим.

Нехай многочлен $P(x)$ степені n має n коренів $x_1, \dots, x_n$, причому, деякі з них можуть повторюватись. Тоді (1.2) перепишеться у вигляді

$$P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_k x^{n-k} + \dots + a_n = a_0 (x - x_1)(x - x_2) \dots (x - x_n).$$

Розкриваючи дужки, одержимо:

$$\begin{aligned} P(x) &= a_0 x^n + a_1 x^{n-1} + \dots + a_k x^{n-k} + \dots + a_n = \\ &= a_0 x^n - a_0 (x_1 + \dots + x_n) x^{n-1} + a_0 (x_1 x_2 + \dots + x_{n-1} x_n) x^{n-2} + \dots + \\ &+ (-1)^k (x_1 x_2 \dots x_k + \dots + x_{n-k+1} \dots x_{n-1} x_n) + \dots + (-1)^n a_0 x_1 x_2 \dots x_n. \end{aligned}$$

Оскільки ця рівність тотожна, коефіцієнти при однакових степенях многочленів в обох частинах рівні. Таким чином, приходимо до системи рівностей, що називаються формулами Вієта, і пов'язують корені многочлена з його коефіцієнтами:

$\left\{ \begin{array}{l} x_1 + \dots x_n = -\frac{a_1}{a_0}, \\ x_1 x_2 + \dots + x_{n-1} x_n = \frac{a_2}{a_0}, \\ \dots \\ x_1 x_2 \dots x_k + \dots + x_{n-k+1} \dots x_{n-1} x_n = (-1)^k \frac{a_k}{a_0}, \\ \dots \\ x_1 x_2 \dots x_n = (-1)^n \frac{a_n}{a_0}. \end{array} \right.$	(1.3)
--	-------

Найбільш простий вигляд мають формули Вієта для многочлена, старший коефіцієнт якого дорівнює одиниці:

$\left\{ \begin{array}{l} x_1 + \dots x_n = -a_1, \\ x_1 x_2 + \dots + x_{n-1} x_n = a_2, \\ \dots \\ x_1 x_2 \dots x_k + \dots + x_{n-k+1} \dots x_{n-1} x_n = (-1)^k a_k, \\ \dots \\ x_1 x_2 \dots x_n = (-1)^n a_n. \end{array} \right.$	(1.4)
--	-------

Наслідок. При $a_0 = 1, x_1 = \dots = x_n = -a$ одержимо відому формулу бінома Ньютона:

$$(x + a)^n = x^n + C_n^1 a x^{n-1} + C_n^2 a^2 x^{n-2} + \dots + C_n^k a^k x^{n-k} + \dots + a^n.$$

Нехай $P_1(x), P_2(x)$ мають спільний дільник $D(x)$. Найбільшим спільним дільником називається їхній спільний дільник найбільшої степені. Зрозуміло, що він визначається неоднозначно, з точністю до ненульового скалярного множника і позначається $(P_1(x), P_2(x)) = D(x)$.

Нехай степінь $P(x)$ більша або дорівнює степені $Q(x)$. Поділимо $P(x)$ на $Q(x)$ з залишком, який 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:

$$\begin{aligned} P(x) &= Q(x)S(x) + R_1(x); \\ Q(x) &= R_1(x)S_1(x) + R_2(x); \\ R_1(x) &= R_2(x)S_2(x) + R_3(x); \\ &\dots\dots\dots \\ R_{k-3}(x) &= R_{k-2}(x)S_{k-2}(x) + R_{k-1}(x); \\ R_{k-2}(x) &= R_{k-1}(x)S_{k-1}(x) + R_k(x); \\ R_{k-1}(x) &= R_k(x)S_k(x). \end{aligned} \tag{1.5}$$

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 вгору, 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

¹ **Евклід** (грець. Ευκλείδης; близько 325 - близько 270 до н. е.) — старогрецький математик і визнаний основоположник математики, автор перших теоретичних трактатів з математики, що дійшли до сучасності.

$D(x)$ і здійснимо 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

$P(x)U(x) + Q(x)V(x) = D(x).$	(1.6)
-------------------------------	---------

Доведення випливає з рівностей (1.5). Якщо покласти

$$R_k(x) = D(x), U_1(x) = 1, V_1(x) = -S_k(x), \text{ то з передостанньої рівності одержимо } D(x) = R_{k-2}(x)U_1(x) + R_{k-1}(x)V_1(x).$$

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), що й треба було довести.

Наслідок. Якщо многочлени $P_1(x), \dots, P_l(x)$ мають найбільший спільний дільник $D(x)$, то існують многочлени $U_1(x), \dots, U_l(x)$, через які $D(x)$ можна знайти за формулою

$$D(x) = \sum_{i=1}^l U_i(x)P_i(x). \tag{1.7}$$

У випадку, коли всі $P_1(x), \dots, P_l(x)$ взаємно прості,

$$\sum_{i=1}^l U_i(x)P_i(x) = 1. \tag{1.8}$$

Твердження доводиться за допомогою математичної індукції.

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

Приклад 1.1. $\mathbb{Z}$ – кільце головних ідеалів.

Для доведення треба розглянути довільний ідеал $I \subset \mathbb{Z}$ і довести, що він породжений одним елементом. Розглянемо найменше натуральне $a \in I$. Якщо воно не породжує I , то знайдеться таке $b \in I$, $b > a$, що $b \neq ap$. 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 .

Приклад 1.2. $\mathbb{Z}[x]$ – кільце головних ідеалів. Доведення спирається на можливість ділення з залишком у кільці многочленів. Радимо читачу провести його самостійно.

Схема Горнера² ділення многочлена на біном

Нехай треба поділити многочлен $P(x) = a_0x^n + a_1x^{n-1} + \dots + a_kx^{n-k} + \dots + a_{n-1}x + a_n$ на біном $x - a$. Позначимо неповну частку, многочлен степені $n-1$ через

$$Q(x) = b_0x^{n-1} + b_1x^{n-2} + \dots + b_kx^{n-k} + \dots + b_{n-2}x + b_{n-1}. \text{ Тоді } P(x) = (x - a)Q(x) + R:$$

$$a_0x^n + a_1x^{n-1} + \dots + a_kx^{n-k} + \dots + a_{n-1}x + a_n =$$

$$= (x - a)(b_0x^{n-1} + b_1x^{n-2} + \dots + b_kx^{n-k} + \dots + b_{n-2}x + b_{n-1}) + R =$$

$$b_0x^n + (b_1 - ab_0)x^{n-1} + (b_2 - ab_1)x^{n-2} + \dots + (b_k - ab_{k-1})x^{n-k} + \dots + (b_{n-1} - ab_{n-2})x - ab_{n-1} + R.$$

Прирівнюючи коефіцієнти при однакових степенях змінної x , одержимо:

² Горнер, Вільям Джордж (англ. William George Horner) (1786, Бристоль, - 22/09/1837, Гросвенор) – відомий математик і винахідник.

$\begin{cases} b_0 = a_0, \\ b_1 = ab_0 + a_1, \\ \dots \\ b_k = ab_{k-1} + a_k, \\ \dots \\ b_{n-1} = ab_{n-2} + a_{n-1}, \\ R = ab_{n-1} + a_n. \end{cases}$	(1.9)
--	-------

Ці обчислення зручно робити, користуючись таблицею:

a_0	a_1	a_2	$\dots$	a_{n-1}	a_n	
—	—	—	—	—	—	—
$a \mid b_0 =$	$b_1 =$	$b_2 =$	$\dots$	$b_{n-1} =$	$R =$	
$\mid = a_0 \mid$	$= ab_0 + a_1 \mid$	$= ab_1 + a_2 \mid$	$\dots$	$= ab_{n-2} + a_{n-1} \mid$	$= ab_{n-1} + a_n \mid$	

Приклад 1.2. Розділити многочлен $P(x) = x^4 + 3x^3 + 4x^2 - 5x - 47$ на біном $x + 3 = x - (-3)$.

Заповнюємо таблицю, користуючись таблицею:

	1	3	4	-5	-47	
—	—	—	—	—	—	—
$-3 \mid$	$1 \mid$	$0 \mid$	$4 \mid$	$-17 \mid$	$R = 4 \mid$	

Це означає, що

$$x^4 + 3x^3 + 4x^2 - 5x - 47 = (x + 3)(x^3 + 4x - 17) + 4.$$

І як висновок з теореми Безу $P(-3) = 4$.

Контрольні запитання

- 1). В чому полягає алгоритм Евкліда?
- 2). Чому кільце многочленів – це кільце головних ідеалів?
- 3). Як ділити многочлен на біном за схемою Горнера?

ЛЕКЦІЯ П.2. РЕЗУЛЬТАНТ

План

- 1). Постановка задачі про існування спільного кореня двох многочленів?
- 2). Результат і його властивості.
- 3). Формула Сільвестра для обчислення результанту.

Основною метою лекції є запропонувати алгоритм дослідження системи двох рівнянь з функціями, які є поліномами від двох змінних. Але поки що розглянемо питання існування спільних коренів у двох поліномів від однієї змінної.

Отже, нехай є поліноми з, взагалі кажучи, комплексними коефіцієнтами

$\begin{aligned}f(x) &= a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n, \\g(x) &= b_0x^m + b_1x^{m-1} + \dots + b_{m-1}x + b_m, \\a_0 \neq 0, b_0 \neq 0.\end{aligned}$	(2.1)
---	-------

З попередньої лекції відомо, що система двох таких рівнянь має розв'язки тоді і тільки тоді, коли ці поліноми мають спільний дільник. Його можна знайти, користуючись алгоритмом Евкліда. Зараз буде вказано інший шлях дослідження ситуації.

Нехай перший і другий поліноми мають, відповідно, n , і m , взагалі кажучи, комплексних коренів, множини яких ми позначимо $\{a\}$ і $\{b\}$:

$\begin{aligned}\{a\} &= \{\alpha_1, \dots, \alpha_n\}, f(x) = a_0 \prod_{i=1}^n (x - \alpha_i), \\ \{b\} &= \{\beta_1, \dots, \beta_m\}, g(x) = b_0 \prod_{j=1}^m (x - \beta_j).\end{aligned}$	(2.2)
---	-------

Означення 2.1. Вираз

$R(f, g) = a_0^m b_0^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j)$	(2.3)
--	-------

Називається результатом двох многочленів $f(x)$ і $g(x)$.

З (2.2) впливає властивість кососиметричності результанту:

$R(g, f) = b_0^n a_0^m \prod_{j=1}^m \prod_{i=1}^n (\beta_j - \alpha_i) = (-1)^{mn} a_0^m b_0^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j) = (-1)^{mn} R(f, g).$	(2.4)
--	-------

Завдяки (2.2), (2.3) і

$$g(\alpha_i) = b_0 \prod_{j=1}^m (\alpha_i - \beta_j), i = 1, \dots, n,$$

$$f(\beta_j) = a_0 \prod_{i=1}^n (\beta_j - \alpha_i), j = 1, \dots, m$$

можна записати:

$R(f, g) = a_0^m \prod_{i=1}^n g(\alpha_i),$	(2.5)
--	-------

$R(g, f) = b_0^n \prod_{j=1}^m f(\beta_j).$	(2.6)
---	-------

Твердження 2.1. Поліноми $f(x)$ і $g(x)$ мають спільний корінь тоді і тільки тоді, коли $R(f, g) = 0$.

Доведення очевидне.

Теорема 2.2. Результат двох поліномів може бути обчислений за допомогою наступного визначника (Сільвестра):

$ D = \begin{vmatrix} a_0 & a_1 & \dots & a_n & & & \\ & a_0 & a_1 & \dots & a_n & & \\ & & \cdot & \cdot & \dots & & \\ & & & a_0 & a_1 & \dots & a_n \\ b_0 & b_1 & \cdot & \dots & b_m & & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & b_0 & b_1 & \cdot & \dots & b_m \end{vmatrix}$	(2.7)
---	-------

Зауваження. На порожніх місцях визначника знаходяться нулі, він має $n + m$ рядків і стовпчиків, на головній діагоналі стоїть спочатку m значень a_0 , а потім n значень b_m . При цьому рядків з коефіцієнтами $a_0, a_1, \dots, a_n$ тут m , а з коефіцієнтами $b_0, b_1, \dots, b_m$ - n .

Для доведення розглянемо допоміжний визначник - визначник Вандермонда:

$ V = \begin{vmatrix} \beta_1^{n+m-1} & \beta_2^{n+m-1} & \dots & \beta_m^{n+m-1} & \alpha_1^{n+m-1} & \alpha_2^{n+m-1} & \dots & \alpha_n^{n+m-1} \\ \beta_1^{n+m-2} & \beta_2^{n+m-2} & \dots & \beta_m^{n+m-2} & \alpha_1^{n+m-2} & \alpha_2^{n+m-2} & \dots & \alpha_n^{n+m-2} \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot & \dots & \cdot \\ \beta_1^2 & \beta_2^2 & \dots & \beta_m^2 & \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \beta_1 & \beta_2 & \dots & \beta_m & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 1 & \dots & 1 & 1 & 1 & \dots & 1 \end{vmatrix}$	(2.8)
--	-------

Як відомо,

$ V = \prod_{1 \leq i < j \leq m} (\beta_i - \beta_j) \prod_{j=1}^m \prod_{i=1}^n (\beta_j - \alpha_i) \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j).$	(2.9)
---	-------

Звідси, з урахуванням (2.6), будемо мати:

$a_0^m b_0^n D V = D R(g, f) \prod_{1 \leq i < j \leq m} (\beta_i - \beta_j) \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j).$	(2.10)
---	--------

Тепер обчислимо той самий вираз, але з використанням властивості визначника добутку двох матриць: $|DV| = |D||V|$.

Докладно покажемо, як множаться окремі рядки і стовпчики цих матриць:

$$(a_0 \quad a_1 \quad \dots \quad a_n \quad 0 \quad \dots \quad 0) \begin{pmatrix} \beta_j^{n+m-1} \\ \beta_j^{n+m-2} \\ \dots \\ \beta_j^2 \\ \beta_j \\ 1 \end{pmatrix} = a_0 \beta_j^{n+m-1} + a_1 \beta_j^{n+m-2} + \dots + a_{n-1} \beta_j^{m+1} + a_n \beta_j^m + a_n \beta_j^{m-1} =$$

$$= \beta_j^{m-1} (a_0 \beta_j^n + a_1 \beta_j^{n-1} + \dots + a_{n-2} \beta_j^2 + a_{n-1} \beta_j + a_n) = \beta_j^{m-1} f(\beta_j),$$

Аналогічно

$$(0 \quad a_0 \quad \dots \quad a_n \quad 0 \quad \dots \quad 0) \begin{pmatrix} \beta_j^{n+m-1} \\ \beta_j^{n+m-2} \\ \dots \\ \beta_j^2 \\ \beta_j \\ 1 \end{pmatrix} = a_0 \beta_j^{n+m-2} + a_1 \beta_j^{n+m-3} + \dots + a_{n-1} \beta_j^{m-1} + a_n \beta_j^{m-2} =$$

$$= \beta_j^{m-2} f(\beta_j),$$

і, нарешті,

$$(0 \quad \dots \quad 0 \quad a_0 \quad a_1 \quad \dots \quad a_n) \begin{pmatrix} \beta_j^{n+m-1} \\ \beta_j^{n+m-2} \\ \dots \\ \beta_j^2 \\ \beta_j \\ 1 \end{pmatrix} = a_0 \beta_j^n + a_1 \beta_j^{n-1} + \dots + a_{n-1} \beta_j + a_n = f(\beta_j).$$

Далі

$$(a_0 \quad a_1 \quad \dots \quad a_n \quad 0 \quad \dots \quad 0) \begin{pmatrix} \alpha_j^{n+m-1} \\ \alpha_j^{n+m-2} \\ \dots \\ \alpha_j^2 \\ \alpha_j \\ 1 \end{pmatrix} = a_0 \alpha_j^{n+m-1} + a_1 \alpha_j^{n+m-2} + \dots + a_{n-2} \alpha_j^{m+1} + a_{n-1} \alpha_j^m + a_n \alpha_j^{m-1} =$$

$$= \alpha_j^{m-1} f(\alpha_j) = 0,$$

і аналогічно

$$(0 \quad \dots \quad 0 \quad a_0 \quad a_1 \quad \dots \quad a_n) \begin{pmatrix} \alpha_j^{n+m-1} \\ \alpha_j^{n+m-2} \\ \dots \\ \alpha_j^2 \\ \alpha_j \\ 1 \end{pmatrix} = a_0 \alpha_j^n + a_1 \alpha_j^{n-1} + \dots + a_{n-1} \alpha_j + a_n = f(\alpha_j) = 0.$$

Так само

$$(b_0 \quad b_1 \quad \dots \quad b_m \quad 0 \quad \dots \quad 0) \begin{pmatrix} \beta_j^{n+m-1} \\ \beta_j^{n+m-2} \\ \dots \\ \beta_j^2 \\ \beta_j \\ 1 \end{pmatrix} = b_0 \beta_j^{n+m-1} + b_1 \beta_j^{n+m-2} + \dots + b_{m-2} \beta_j^{n+1} + b_{m-1} \beta_j^n + b_m \beta_j^{n-1} =$$

$$= \beta_j^{n-1} (b_0 \beta_j^m + b_1 \beta_j^{m-1} + \dots + b_{m-2} \beta_j^2 + b_{m-1} \beta_j + b_m) = \beta_j^{n-1} g(\beta_j) = 0, \dots,$$

$$(0 \quad \dots \quad 0 \quad b_0 \quad b_1 \quad \dots \quad b_m) \begin{pmatrix} \beta_j^{n+m-1} \\ \beta_j^{n+m-2} \\ \dots \\ \beta_j^2 \\ \beta_j \\ 1 \end{pmatrix} = g(\beta_j) = 0,$$

$$(b_0 \quad b_1 \quad \dots \quad b_m \quad 0 \quad \dots \quad 0) \begin{pmatrix} \alpha_j^{n+m-1} \\ \alpha_j^{n+m-2} \\ \dots \\ \alpha_j^2 \\ \alpha_j \\ 1 \end{pmatrix} = b_0 \alpha_j^{n+m-1} + b_1 \alpha_j^{n+m-2} + \dots + b_{m-2} \alpha_j^{n+1} + b_{m-1} \alpha_j^n + b_m \alpha_j^{n-1} =$$

$$= \alpha_j^{n-1} g(\alpha_j), \dots,$$

$$(0 \quad \dots \quad 0 \quad b_0 \quad b_1 \quad \dots \quad b_m) \begin{pmatrix} \alpha_j^{n+m-1} \\ \alpha_j^{n+m-2} \\ \dots \\ \alpha_j^2 \\ \alpha_j \\ 1 \end{pmatrix} = b_0 \alpha_j^n + b_1 \alpha_j^{n-1} + \dots + b_{m-2} \alpha_j^2 + b_{m-1} \alpha_j + b_m = g(\alpha_j).$$

$$\text{Отже, } |DV| = |D||V| =$$

$$= \begin{vmatrix} \beta_1^{m-1} f(\beta_1) & \beta_2^{m-1} f(\beta_2) & \dots & \beta_m^{m-1} f(\beta_m) & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot & \dots & \cdot \\ \beta_1 f(\beta_1) & \beta_2 f(\beta_2) & \dots & \beta_m f(\beta_m) & 0 & 0 & \dots & 0 \\ f(\beta_1) & f(\beta_2) & \dots & f(\beta_m) & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \alpha_1^{n-1} g(\alpha_1) & \alpha_2^{n-1} g(\alpha_2) & \dots & \alpha_n^{n-1} g(\alpha_n) \\ 0 & 0 & \dots & 0 & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot & \alpha_1 g(\alpha_1) & \alpha_2 g(\alpha_2) & \dots & \alpha_n g(\alpha_n) \\ 0 & 0 & \dots & 0 & g(\alpha_1) & g(\alpha_2) & \dots & g(\alpha_n) \end{vmatrix} =$$

$$= \prod_{j=1}^m f(\beta_j) \prod_{i=1}^n g(\alpha_i) \begin{vmatrix} \beta_1^{m-1} & \beta_2^{m-1} & \dots & \beta_m^{m-1} & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot & \dots & \cdot \\ \beta_1 & \beta_2 & \dots & \beta_m & 0 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} \\ 0 & 0 & \dots & 0 & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot & \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \end{vmatrix} =$$

$$= \prod_{j=1}^m f(\beta_j) \prod_{i=1}^n g(\alpha_i) \prod_{1 \leq i < j \leq m} (\beta_i - \beta_j) \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j).$$

Враховуючи формули (2.5) і (2.6), одержимо:

$a_0^m b_0^n D V = R(f, g) R(g, f) \prod_{1 \leq i < j \leq m} (\beta_i - \beta_j) \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j).$	(2.11)
---	----------

Отже, праві частини (2.10) і (2.11) рівні між собою, в результаті чого одержимо:

$$R(f, g) R(g, f) \prod_{1 \leq i < j \leq m} (\beta_i - \beta_j) \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j) = |D| R(g, f) \prod_{1 \leq i < j \leq m} (\beta_i - \beta_j) \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)$$

Вважаючи, що всі корені многочленів різні і при умові $a_0 \neq 0, b_0 \neq 0$, приходимо до рівності

$R(f, g) = D = \begin{vmatrix} a_0 & a_1 & \dots & a_n & & \\ & a_0 & a_1 & \dots & a_n & \\ & & \cdot & \cdot & \dots & \\ & & & a_0 & a_1 & \dots & a_n \\ b_0 & b_1 & \cdot & \dots & b_m & & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ & & b_0 & b_1 & \cdot & \dots & b_m \end{vmatrix}.$	(2.12)
--	----------

Відмовимось тепер від умови $a_0 \neq 0, b_0 \neq 0$. З рівності нулю якогось з цих коефіцієнтів випливає, що результат, що обчислюється за формулою (2.3), дорівнює нулю, і з цього не можна робити висновок про існування у многочленів спільних коренів. З іншої сторони, визначник (2.12) і в цьому випадку має сенс, і за означенням будемо вважати його результатом.

Теорема 2.3. Результат двох многочленів (2.1) тоді і тільки тоді дорівнює нулю, коли вони мають спільний корінь або їхні старші коефіцієнти водночас дорівнюють нулю.

Доведення. Випадок існування спільного кореня вже розглядався, і в цій ситуації, як було пояснено, теорема вірна. Звичайно, при $a_0 = b_0 = 0$ результат дорівнює нулю – це випливає з властивості визначника (2.12). Залишилось розглянути випадок, коли тільки один із старших коефіцієнтів дорівнює нулю, а інший – ні. Нехай $a_0 \neq 0, b_0 = 0$. Якщо при цьому $b_0 = b_1 = \dots = b_m = 0$, тоді визначник (2.12) дорівнює нулю, оскільки один його рядок складається з нулів, і $g(x) \equiv 0$ тотожно. Отже, многочлени мають спільні корені. Нехай тепер $b_0 = b_1 = \dots = b_{k-1} = 0, b_k \neq 0, k \leq m$. Розглянемо многочлен

$\tilde{g}(x) = b_k x^{m-k} + b_{k+1} x^{m-k-1} + \dots + b_{m-1} x + b_m.$	(2.13)
---	----------

Якщо у визначнику (2.12) замінити елементи $b_0, b_1, \dots, b_{k-1}$ нулями, то отримаємо, очевидну рівність:

$R(f, g) = a_0^k R(f, \tilde{g}).$	(2.14)
------------------------------------	--------

Помітимо, що многочлени g і $\tilde{g}$ мають однакові корені, старші коефіцієнти многочленів f і $\tilde{g}$ не рівні нулю, тому для існування спільного кореня у f і $\tilde{g}$ необхідно і достатньо, щоб $R(f, \tilde{g}) = 0$, а ця умова еквівалентна умові $R(f, g) = 0$, що й треба було довести.

Розглянемо тепер два поліноми від двох змінних, причому зразу перепишемо їх як поліноми від однієї змінної x , коефіцієнти яких залежать від y :

$\begin{aligned} f(x) &= a_0(y)x^k + a_1(y)x^{k-1} + \dots + a_{k-1}(y)x + a_k(y), \\ g(x) &= b_0(y)x^l + b_1(y)x^{l-1} + \dots + b_{l-1}(y)x + b_l(y). \end{aligned}$	(2.15)
--	--------

Оскільки ці поліноми розглядаються як функції від x , то їхній результат будемо позначати $R_x(f, g)$. З розглянутої раніше теорії робимо висновок: при наявності спільного кореня $x = \alpha, y = \beta$ він дорівнює нулю або, якщо рівняння $R_x(f, g) \equiv F(y) = 0$ має розв'язок, то тоді ці поліноми мають спільний корінь або обидва їхніх старших коефіцієнти дорівнюють нулю.

Таким чином, знаходження спільних коренів двох поліноміальних рівнянь від двох змінних звелось до розв'язання поліноміального рівняння від однієї змінної. Цей підхід називається методом виключення і належить Ейлеру. Визначник (2.12) запропонований Сільвестром.

Приклад 2.1. Розв'язати систему рівнянь

$$\begin{cases} f(x, y) = x^2 y + 3xy + 2y + 3 = 0, \\ g(x, y) = 2xy - 2x + 2y + 3 = 0. \end{cases}$$

Запишемо поліноми у лівих частинах як функції від x з коефіцієнтами, що залежать від y :

$$\begin{cases} yx^2 + (3y)x + (2y + 3) = 0, \\ ((2y - 2)x + (2y + 3)) = 0. \end{cases}$$

$$R_x(f, g) = \begin{vmatrix} y & 3y & 2y + 3 \\ 2y - 2 & 2y + 3 & 0 \\ 0 & 2y - 2 & 2y + 3 \end{vmatrix} = 2y^2 + 11y + 12 = 2(y + 4)\left(y + \frac{3}{2}\right).$$

Старші коефіцієнти даних в умові поліномів не обертаються в нуль при значеннях $y = -4, y = -\frac{3}{2}$.

Знайдемо

$$\begin{aligned} f(x, -4) &= -4x^2 - 12x - 5 = -4\left(x + \frac{5}{2}\right)\left(x + \frac{1}{2}\right), \\ g(x, -4) &= -10x - 5 = -10\left(x + \frac{1}{2}\right). \end{aligned}$$

Многочлени мають спільний корінь $\alpha_1 = -\frac{1}{2}$.

Далі

$$\begin{aligned} f\left(x, -\frac{3}{2}\right) &= -\frac{3}{2}x^2 - \frac{9}{2}x = -\frac{3}{2}x(x + 3), \\ g\left(x, -\frac{3}{2}\right) &= -5x. \end{aligned}$$

Відповідь. Система має розв'язки: $(x, y) \in \left\{\left(-\frac{1}{2}, -4\right), \left(0, -\frac{3}{2}\right)\right\}$.

Дискримінант

Розглянемо питання про існування кратних коренів многочлена. Нехай $a_0 \neq 0$,

$$f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n$$

і нехай $\alpha_1, \alpha_2, \dots, \alpha_n$ - корені цього многочлена. Очевидно, серед коренів будуть рівні тоді і тільки тоді, коли дорівнює нулю добуток

$$\begin{aligned} (\alpha_2 - \alpha_1)(\alpha_3 - \alpha_1) \dots (\alpha_n - \alpha_1) \cdot (\alpha_3 - \alpha_2)(\alpha_4 - \alpha_2) \dots (\alpha_n - \alpha_2) \cdot \dots \cdot (\alpha_n - \alpha_{n-1}) = \\ = \prod_{n \geq i > j \geq 1} (\alpha_i - \alpha_j). \end{aligned}$$

Це еквівалентно умові

$D = a_0^{2n-2} \prod_{n \geq i > j \geq 1} (\alpha_i - \alpha_j)^2 = 0.$	(2.16)
---	--------

Останній вираз називається дискримінантом многочлена $f(x)$.

Як відомо, многочлен має кратні корені тоді і тільки тоді, коли він має спільні корені зі своєю похідною, що, в свою чергу, еквівалентне умові рівності нулю їхнього результанта: $R(f, f') = 0$. Згідно з формулою (2.5)

$$R(f, f') = a_0^{n-1} \prod_{i=1}^n f'(\alpha_i),$$

Диференціюючи рівність $f(x) = a_0 \prod_{k=1}^n (x - \alpha_k)$, одержимо:

$$f'(x) = a_0 \sum_{k=1}^n \prod_{j \neq k}^n (x - \alpha_j).$$

Підставимо сюди $x = \alpha_i$ і помітимо, що тільки i -й доданок не буде рівний нулю:

$$f'(\alpha_i) = a_0 \prod_{j \neq i} (\alpha_i - \alpha_j),$$

звідки

$$R(f, f') = a_0^{n-1} a_0^n \prod_{i=1}^n \prod_{j \neq i} (\alpha_i - \alpha_j).$$

Помітимо, що при $i \neq j$ існує два множники $\alpha_i - \alpha_j$ і $\alpha_j - \alpha_i$. Їхній добуток дорівнює $-(\alpha_i - \alpha_j)^2$, і існує $\frac{n(n-1)}{2}$ пар індексів i, j , що задовольняють нерівності $n \geq i > j \geq 1$. Звідси

$R(f, f') = (-1)^{\frac{n(n-1)}{2}} \alpha_0^{2n-1} \prod_{n \geq i > j \geq 1}^n (\alpha_i - \alpha_j)^2 = (-1)^{\frac{n(n-1)}{2}} \alpha_0 D.$	(2.17)
--	----------

Приклад. Дискримінант квадратного тричлена $f(x) = ax^2 + bx + c$.

В цьому випадку $n = 2 \Rightarrow \frac{n(n-1)}{2} = 1 \Rightarrow D = -\frac{1}{a} R(f, f')$.

$$f'(x) = 2ax + b \Rightarrow R(f, f') = \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = a(-b^2 + 4ac). \text{ Звідси}$$

$D = b^2 - 4ac$, що співпадає з поняттям дискримінанта, відомого з курсу елементарної математики.

Контрольні запитання

1). Яка постановка задачі приводить до поняття «результант»?

- 2). Які властивості має результат?
- 3). Коли два многочлени мають спільний корінь? Дати відповідь, посилаючись на результат.
- 4). Як знаходити результат, зводячи його до визначника Сільвестра?

ЛЕКЦІЯ П.3. ІНТЕРПОЛЯЦІЙНИЙ МНОГОЧЛЕН ЛАГРАНЖА; ДРОБОВО-РАЦІОНАЛЬНІ ФУНКЦІЇ

План

- 1). Постановка задачі про знаходження інтерполяційного многочлену.
- 2). Інтерполяційний многочлен Лагранжа.
- 3). Дробово-раціональні функції і їхнє зображення у вигляді суми елементарних дробів.

Будемо розв'язувати задачу: знайти многочлен $P(x)$ степені n , який в точках $x_1, x_2, \dots, x_n, x_{n+1}$ приймає задані значення:

$P(x_1) = y_1, P(x_2) = y_2, \dots, P(x_n) = y_n, P(x_{n+1}) = y_{n+1}.$	(3.1)
--	-------

Розв'яжемо допоміжну задачу: знайти «індикаторний» многочлен $P_k(x)$, який у даній точці $x_k, k = 1, \dots, n+1$ приймає значення 1, а в інших точках – нульові значення. Легко бачити, що цей многочлен має вигляд

$P_k(x) = \frac{(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_{n+1})}{(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_{n+1})}.$	(3.2)
---	-------

Тепер шуканим є многочлен, що називається многочленом Лагранжа:

$P(x) = \sum_{k=1}^{n+1} y_k P_k(x) = \sum_{k=1}^{n+1} y_k \frac{(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_{n+1})}{(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_{n+1})}.$	(3.3)
--	-------

Зображення правильного раціонального дробу у вигляді суми елементарних дробів

Дробово-раціональна функція $\frac{P(x)}{Q(x)}$ – правильний дріб, якщо $\deg P(x) < \deg Q(x)$, а якщо навпаки – неправильний дріб. Далі будемо вважати, що старший коефіцієнт многочлена у знаменнику дорівнює 1.

Нехай знаменник розкладено на множники:

$$Q(x) = (x - a_1)^{r_1} \dots (x - a_k)^{r_k} (x^2 + p_1x + q_1)^{R_1} \dots (x^2 + p_lx + q_l)^{R_l}.$$

Очевидно, $r_1 + \dots + r_k + 2(R_1 + \dots + R_l) = \deg Q(x)$.

Елементарними називаються дроби таких типів:

$$\frac{A}{(x-a)^r}, \frac{\alpha x + \beta}{(x^2 + px + q)^R}, r, R \in \mathbb{N}, D = p^2 - 4q < 0.$$

Треба довести, що правильний раціональний дріб можна подати у вигляді суми елементарних дробів.

Доведемо допоміжні факти.

А). Нехай $Q(x) = (x - a)^k Q_1(x)$, де $Q_1(x)$ не ділиться на $(x - a)$. Тоді правильний дріб можна зобразити у вигляді

$$\frac{P(x)}{Q(x)} = \frac{A}{(x-a)^k} + \frac{P_1(x)}{(x-a)^{k-1} Q_1(x)}.$$

Звідси

$$P(x) - A Q_1(x) = (x - a) P_1(x).$$

Права частина ділиться на $(x - a)$, тому ліва частина повинна теж ділитися на цей двочлен. За теоремою Безу досить, щоб її значення при $x = a$ дорівнювало

нулю. Тоді невідоме A знайдеться як $A = \frac{P(a)}{Q_1(a)}$. Многочлен $P_1(x)$ тоді знайдеться

$$\text{як частка: } P_1(x) = \frac{P(x) - AQ_1(x)}{x - a}.$$

$$\text{В). } Q(x) = (x^2 + px + q)^m Q_1(x).$$

Покажемо, що

$$\frac{P(x)}{Q(x)} = \frac{Mx + N}{(x^2 + px + q)^m} + \frac{P_1(x)}{(x^2 + px + q)^{m-1} Q_1(x)}.$$

Звідси маємо:

$$P(x) - (Mx + N)Q_1(x) = (x^2 + px + q)P_1(x).$$

Знайдемо значення M, N , щоб ліва частина ділилась на $(x^2 + px + q)$. Нехай залишки від ділення $P(x), Q_1(x)$ на $(x^2 + px + q)$ дорівнювали, відповідно, $(\alpha x + \beta), (\gamma x + \delta)$. Тоді будемо вимагати, щоб на $(x^2 + px + q)$ ділився многочлен

$$\alpha x + \beta - (Mx + N)(\gamma x + \delta) = -\gamma Mx^2 + (\alpha - \delta M - \gamma N)x + (\beta - \delta N).$$

Виконавши ділення за відомим алгоритмом, будемо мати в остачі:

$$((p\gamma - \delta)M - \gamma N + \alpha)x + (q\gamma M - \delta N + \beta).$$

Невідомі M, N знайдемо, розв'язавши систему

$$\begin{cases} (p\gamma - \delta)M - \gamma N = -\alpha, \\ q\gamma M - \delta N = -\beta. \end{cases}$$

Знайдемо її визначник:

$$-(p\gamma - \delta)\delta + q\gamma^2 = \delta^2 - p\gamma\delta + q\gamma^2 = \gamma^2 \left(\left(-\frac{\delta}{\gamma} \right)^2 + p \left(-\frac{\delta}{\gamma} \right) + q \right).$$

Треба помітити, що вираз у дужках – це значення квадратного тричлена $(x^2 + px + q)$ при $x = -\frac{\delta}{\gamma}$, тому він строго більший за нуль. Отже, при $\gamma \neq 0$ цей визначник не дорівнює нулю.

Якщо тепер $\gamma = 0, \delta \neq 0$, визначник дорівнює $\delta^2 \neq 0$. Зауважимо, що випадок $\gamma = \delta = 0$ неможливий, оскільки $P(x)$ і $Q_1(x)$ за умовою не діляться на $x^2 + px + q$.

Знайшовши M, N , многочлен $P_1(x)$ знайдемо як частку:

$$P_1(x) = \frac{P(x) - (Mx + N)Q_1(x)}{x^2 + px + q}.$$

Застосуємо тепер до дробу описану процедуру А), будемо виділяти елементарні дробу вигляду

$$\frac{A_1}{x-a}, \dots, \frac{A_k}{(x-a)^k}.$$

Це стосується всіх дійсних коренів знаменника.

Аналогічно, згідно з В), виділимо елементарні дробу

$$\frac{M_1x + N_1}{x^2 + px + q}, \dots, \frac{M_mx + N_m}{(x^2 + px + q)^m}.$$

Це стосується всіх квадратних тричленів розкладу знаменника.

Приклад 3.1. Нехай корені многочлена $Q(x)$ в знаменнику прості:

$$Q(x) = (x - x_1)(x - x_2) \dots (x - x_n),$$

а $\deg P(x) < \deg Q(x)$. Не використовуючи метод невизначених коефіцієнтів, знайти чисельники елементарних дробів, сумою яких є даний дріб:

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \dots + \frac{A_n}{x - x_n}.$$

Звідси

$$P(x) = A_1(x - x_2) \dots (x - x_n) + \dots + A_k(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n) + \dots + A_n(x - x_1) \dots (x - x_{n-1}).$$

Звідси

$$P(x_k) = A_k(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n) \quad \text{і}$$

$$A_k = \frac{P(x_k)}{(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}.$$

Знайдемо похідну знаменника:

$$Q'(x) = (x - x_2) \dots (x - x_n) + \dots + (x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n) + (x - x_1) \dots (x - x_{n-1}).$$

Тоді

$$(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n) = Q'(x_k).$$

В результаті приходимо до формули:

$$A_k = \frac{P(x_k)}{Q'(x_k)}.$$

$$\text{Відповідь: } \frac{P(x)}{Q(x)} = \frac{P(x_1)}{Q'(x_1)(x - x_1)} + \frac{P(x_2)}{Q'(x_2)(x - x_2)} + \dots + \frac{P(x_n)}{Q'(x_n)(x - x_n)}.$$

Контрольні запитання

- 1). Яка постановка задачі приводить до інтерполяційного многочлена?
- 2). Що таке інтерполяційний многочлен Лагранжа?
- 3). Як дробово-раціональну функцію зобразити у вигляді суми елементарних дробів?

ЛЕКЦІЯ II.4. УНІТАРНІ ТА ЕВКЛІДОВІ ПРОСТОРИ

План

- 1). Скалярний добуток у лінійних просторах.
- 2). Нерівність Коші-Буняковського.
- 3). Унітарний та евклідовий простори.

- 4). Теорема Піфагора у унітарному просторі.
- 5). Метод ортогоналізації Грама-Шмідта.
- 6). Ортогональна сума підпросторів.
- 6). Лінійні функціонали та їхнє зображення з використанням скалярного добутку.

Означення 4.1. 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) $(x, x) \geq 0$, $(x, x) = 0$ 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 $(y, x) = \overline{(x, y)}$, звідки випливає, що в такому випадку скалярний добуток лінійний по першій і антилінійний по другій змінній:

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

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 a — також ортогональний вектору $x = a$, тобто, $(a, a) = 0$. Але з аксіоми 1 одержуємо, що $a = 0$.

Нерівність Коші³ - Буняковського⁴: $\forall x, y \in E \quad |(x, y)| \leq \|x\| \|y\|$

$$|(x, y)| \leq \|x\| \|y\|.$$

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 $(\alpha x + \beta y, \alpha x + \beta y) \geq 0$. З іншого боку $(\alpha x + \beta y, \alpha x + \beta y) = |\alpha|^2 (x, x) + |\beta|^2 (y, y) + \alpha \bar{\beta} (x, y) + \bar{\alpha} \beta \overline{(x, y)}$.

Будемо вважати α дійсним, а $\beta = (x, y)$. Після підстановки одержимо:

$\|x\|^2 \alpha^2 + 2|(x, y)|^2 \alpha + |(x, y)|^2 \|y\|^2 \geq 0$ - квадратний тричлен відносно α не

менше нуля, отже, $\frac{D}{4} = |(x, y)|^4 - |(x, y)|^2 \|x\|^2 \|y\|^2 \leq 0$, 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.

$$\|x + y\|^2 = (x + y, x + y) = (x, x) + 2 \operatorname{Re}(x, y) + (y, y) \leq \|x\|^2 + 2|(x, y)| + \|y\|^2 \leq$$

³ Коші, Огюстен Луї (фр. *Augustin Louis Cauchy*; [21/08/1789](#), [Париж](#) — [23/05/1857](#)) - великий математик, член Паризької академії наук (1816), Петербурзької академії наук (1831).

⁴ Буняковський Віктор Якович (3(15)/12/1804, Бар, Подільська губернія, – 30/11 (12/12)/1804, Санкт-Петербург, Російська імперія) – видатний математик, член Петербурзької Академії Наук та її віце-президент (1864—1889).

$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

Теорема 4.1. (Піфагора⁵) 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 $e_1, e_2, \dots, e_n$ попарно ортогональні, тобто

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

Доведемо, що вони лінійно незалежні. Справді, нехай виконується рівність

$\sum_{i=1}^n x_i e_i = 0$. Помноживши скалярно обидві її частини на e_j , $j = 1, \dots, n$, дістанемо

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

тобто їх лінійна комбінація дорівнює нулю тоді і лише тоді, коли всі коефіцієнти в ній 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.

⁵ **Піфагор** (давні.-гр. Πυθαγόρας, 570 до н. е., Сідон – 497 до н. е., Метапонт) — давньогрецький філософ, релігійний та політичний діяч, засновник піфагореїзму.

Нехай $e = \{e_1, \dots, e_m\}$ - 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 $(x - x_e, e_k) = 0, k = 1, \dots, m$. 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 зіставлення задає лінійний оператор: $P_L x = x_e$ і цей оператор має очевидну властивість: $P_L^2 = P_L$. Така властивість називається ідемпотентністю.

Нехай $y = \sum_{k=1}^m \beta_k e_k$. 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: $\|x - x_e\| = \inf_{y \in L} \|x - y\|$. Вираз справа розумно назвати відстанню від x до підпростору L . Будемо позначати її через δ . Пізніше буде знайдено її точний вираз.

З проведених міркувань випливає, що якщо $\{g_1, g_2, \dots, g_n\}$ - ортогональний базис лінійного простору,

$x = \sum_{k=1}^n x_k g_k$	(4.1)
----------------------------	-------

то для норми вектора x справедлива рівність, яка у функціональному аналізі називається рівністю Парсеваля:

$\ x\ ^2 = \sum_{k=1}^n x _k^2 \ g_k\ ^2.$	(4.2)
---	-------

$$\left(x, \frac{1}{\|g_j\|} g_j\right) = x_j \|g_j\| = p_{g_j}(x) - \text{величина проекції вектора } x \text{ на одновимірний}$$

підпростір, породжений вектором g_j .

Особливо просто рівність Парсеваля записується у випадку ортонормованого базису, який позначимо $\{e_1, e_2, \dots, e_n\}$:

$\ x\ ^2 = \sum_{k=1}^n x _k^2.$	(4.3)
-----------------------------------	-------

$(x, e_j) = x_j = p_{e_j}(x)$ - величина проекції вектора x на одновимірний підпростір, породжений ортом e_j .

Цей факт формулюється таким чином: квадрат норми довільного вектора дорівнює сумі квадратів величин його проекцій на вектори ортонормованого базису.

Нехай $g = \{g_1, \dots, g_n\} \subset E_n$ - деяка система векторів, взагалі кажучи, комплексного унітарного простору. Утворимо матрицю, що називається матрицею Грама цієї системи:

$$\Gamma(g_1, \dots, g_n) = \begin{pmatrix} (g_1, g_1) & (g_1, g_2) & \dots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \dots & (g_2, g_n) \\ \cdot & \cdot & \dots & \cdot \\ (g_n, g_1) & (g_n, g_2) & \dots & (g_n, g_n) \end{pmatrix}.$$

Цікаво помітити, що матриця Грама ортогонального базису діагональна, а ортонормованого – одинична.

Нехай система g лінійно залежна. Тоді існують не всі нульові числа $c_1, \dots, c_n$ такі, що

$$c_1 g_1 + \dots + c_n g_n = 0.$$

Помножимо зліва обидві частини рівності по черзі на всі вектори сукупності g :

$$\begin{cases} c_1(g_1, g_1) + c_2(g_1, g_2) + \dots + c_n(g_1, g_n) = 0, \\ c_1(g_2, g_1) + c_2(g_2, g_2) + \dots + c_n(g_2, g_n) = 0, \\ \dots \\ c_1(g_n, g_1) + c_2(g_n, g_2) + \dots + c_n(g_n, g_n) = 0. \end{cases}$$

Оскільки ця лінійна система відносно $c_1, \dots, c_n$ має нетривіальний розв'язок, її визначник, що є визначником Грама системи векторів $g = \{g_1, \dots, g_n\}$, дорівнює нулю.

Доведемо обернене твердження. Нехай визначник Грама системи $g = \{g_1, \dots, g_n\}$ дорівнює нулю. Тоді попередня лінійна система рівнянь має нетривіальний розв'язок. Її ліві частини можна записати у вигляді

$$\begin{cases} (g_1, \bar{c}_1 g_1 + \bar{c}_2 g_2 + \dots + \bar{c}_n g_n) = 0, \\ (g_2, \bar{c}_1 g_1 + \bar{c}_2 g_2 + \dots + \bar{c}_n g_n) = 0, \\ \dots \\ (g_n, \bar{c}_1 g_1 + \bar{c}_2 g_2 + \dots + \bar{c}_n g_n) = 0. \end{cases}$$

Помножимо рівняння системи відповідно на числа $\bar{c}_1, \bar{c}_2, \dots, \bar{c}_n$ і результати додамо. В результаті одержимо:

$$(\bar{c}_1 g_1 + \bar{c}_2 g_2 + \dots + \bar{c}_n g_n, \bar{c}_1 g_1 + \bar{c}_2 g_2 + \dots + \bar{c}_n g_n) = 0.$$

З властивості скалярного добутку звідси випливає: $\bar{c}_1 g_1 + \bar{c}_2 g_2 + \dots + \bar{c}_n g_n = 0$,

тобто дана система векторів лінійно залежна.

Висновок: Для того, щоб система векторів була лінійно незалежною, необхідно і досить, щоб її матриця Грама була невиродженою, тобто, визначник Грама не дорівнював нулю.

Нагадаємо: мінор визначника називається головним, якщо його головна діагональ лежить на діагоналі визначника.

Властивість мінорів визначника Грама. Якщо деякий головний мінор визначника Грама дорівнює нулю, то дорівнює нулю і сам визначник.

Доведення. Даний нульовий мінор можна сприймати як визначник деякої підсистеми векторів даної системи. Отже, ця підсистема лінійно залежна. Тоді лінійно залежна і вся система, звідки випливає рівність нулю її визначника Грама.

Приклад. Нехай є базис у $\mathbb{C}^n$:

$$g_1 = (g_{11}, g_{21}, \dots, g_{n1})^T, g_2 = (g_{12}, g_{22}, \dots, g_{n2})^T, \dots, g_n = (g_{1n}, g_{2n}, \dots, g_{nn})^T.$$

Скалярний добуток таких векторів обчислюється за формулою:

$$(g_i, g_j) = \left((g_{1i}, g_{2i}, \dots, g_{ni})^T, (g_{1j}, g_{2j}, \dots, g_{nj})^T \right) = \sum_{k=1}^n g_{ki} \bar{g}_{kj}.$$

Розглянемо матрицю, векторами-стовпчиками якої є вектори цього базису:

$$G = (g_1, g_2, \dots, g_n) = \begin{pmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \cdot & \cdot & \dots & \cdot \\ g_{n1} & g_{n2} & \dots & g_{nn} \end{pmatrix}.$$

Також розглянемо матрицю

$$\bar{G}^T = (\bar{g}_1, \bar{g}_2, \dots, \bar{g}_n)^T = \begin{pmatrix} g_{11} & \bar{g}_{21} & \dots & \bar{g}_{n1} \\ \bar{g}_{12} & g_{22} & \dots & \bar{g}_{n2} \\ \cdot & \cdot & \dots & \cdot \\ \bar{g}_{1n} & \bar{g}_{2n} & \dots & \bar{g}_{nn} \end{pmatrix}.$$

Зауважимо, що $g_{ii} = (g_i, g_i) > 0, i = 1, 2, \dots, n$. Обчислимо добуток цих матриць:

$$\begin{aligned} \bar{G}^T G &= \begin{pmatrix} \sum_{k=1}^n g_{1k} \bar{g}_{1k} & \sum_{k=1}^n g_{1k} \bar{g}_{2k} & \dots & \sum_{k=1}^n g_{1k} \bar{g}_{nk} \\ \sum_{k=1}^n g_{2k} \bar{g}_{1k} & \sum_{k=1}^n g_{2k} \bar{g}_{2k} & \dots & \sum_{k=1}^n g_{2k} \bar{g}_{nk} \\ \dots & \dots & \dots & \dots \\ \sum_{k=1}^n g_{nk} \bar{g}_{1k} & \sum_{k=1}^n g_{nk} \bar{g}_{2k} & \dots & \sum_{k=1}^n g_{nk} \bar{g}_{nk} \end{pmatrix} = \\ &= \begin{pmatrix} (g_1, g_1) & (g_1, g_2) & \dots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \dots & (g_2, g_n) \\ \dots & \dots & \dots & \dots \\ (g_n, g_1) & (g_n, g_2) & \dots & (g_n, g_n) \end{pmatrix} = \Gamma(g_1, g_2, \dots, g_n). \end{aligned}$$

Звідси

$$\begin{aligned} \det(\Gamma(g_1, g_2, \dots, g_n)) &= \\ &= \det(G) \det(\bar{G}^T) = \det(G) \det(\bar{G}) = \det(G) \overline{\det(G)} = |\det(G)|^2 > 0. \end{aligned}$$

Отже, визначник матриці Грама системи векторів, що утворюють базис простору $\mathbb{C}^n$, додатний. З цього випливає, що кожен її головний мінор також додатний. Дійсно, цей мінор – матриця Грама підсистеми базису, яка є базисом деякого підпростору.

Нехай тепер $g = \{g_1, g_2, \dots, g_n\} \subset \mathbb{R}^n$ - ортонормований базис,

$g_i = (g_{1i}, g_{2i}, \dots, g_{ni})^T, i = 1, \dots, n$. Запишемо ці вектори в матрицю так, щоб кожен вектор-стовпчик був стовпчиком матриці:

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

Розглянемо також транспоновану матрицю:

$$G^T = \begin{pmatrix} g_{11} & g_{21} & \cdots & g_{n1} \\ g_{12} & g_{22} & \cdots & g_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ g_{1n} & g_{2n} & \cdots & g_{nn} \end{pmatrix}.$$

Зазначимо: з ортонормованості базису випливає, що скалярні добутки різних векторів-рядків матриці дорівнюють нулю, а суми квадратів елементів у кожному рядку дорівнюють одиниці:

$$\sum_{i=1}^n g_{ij} g_{ik} = 0, j \neq k; \sum_{i=1}^n g_{ij}^2 = 1 (k = j).$$

Нехай

$$x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n, \|x\|^2 = \sum_{i=1}^n x_i^2 = 1.$$

Розкладемо цей вектор (орт) за базисом g :

$x = \sum_{i=1}^n (x, g_i) g_i = \sum_{i=1}^n (x_1 g_{1i} + x_2 g_{2i} + \dots + x_n g_{ni}) g_i.$	(4.4)
--	-------

Запишемо рівність Парсеваля для такого розкладу:

$\begin{aligned} & (x_1 g_{11} + x_2 g_{21} + \dots + x_n g_{n1})^2 + \dots + (x_1 g_{1n} + x_2 g_{2n} + \dots + x_n g_{nn})^2 = \\ & = x_1^2 + x_2^2 + \dots + x_n^2 = 1 \end{aligned}$	(4.5)
--	-------

Зауваження. Помітимо, що $p_{g_i}(x) = (x, g_i) = (g_i, x) = p_x(g_i)$. Тоді рівність Парсеваля можна інтерпретувати таким чином: сума квадратів проекцій векторів ортонормованого базису на довільний орт дорівнює одиниці:

$\sum_{i=1}^n (p_x(g_i))^2 = \sum_{i=1}^n (g_i, x)^2 = 1.$	(4.6)
--	-------

Розкладемо за базисом g кожен вектор канонічного базису:

$$e_j = (0, \dots, 0, 1, 0, \dots, 0), j = 1, \dots, n:$$

$$e_j = (0, \dots, 0, 1, 0, \dots, 0) = g_{j1}g_1 + g_{j2}g_2 + \dots + g_{jn}g_n.$$

Записуючи для цього розкладу рівність Парсеваля, одержимо:

$$\|e_j\|^2 = g_{j1}^2 + g_{j2}^2 + \dots + g_{jn}^2 = 1.$$

Перемножимо скалярно два ортогональні вектори e_j і e_k , $j \neq k$:

$$(e_j, e_k) = \sum_{i=1}^n g_{ji}g_{ki} = 0.$$

Ці дві рівності свідчать про те, що сума квадратів елементів матриці G^T у кожному стовпчику дорівнює одиниці, а скалярний добуток різних її векторів-рядків дорівнює нулю. Тобто, вектори-стовпчики і вектори-рядки матриці мають однакову властивість ортонормованості. Така матриця називається ортогональною.

З іншого боку,

$$G^T e_j = \begin{pmatrix} g_{11} & g_{21} & \dots & g_{n1} \\ g_{12} & g_{22} & \dots & g_{n2} \\ \dots & \dots & \dots & \dots \\ g_{1n} & g_{2n} & \dots & g_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} g_{j1} \\ g_{j2} \\ \dots \\ g_{jn} \end{pmatrix}, j = 1, \dots, n.$$

З попередніх досліджень випливає, що ця матриця не змінює скалярного добутку між векторами канонічного базису:

$$(G^T e_j, G^T e_k) = (e_j, e_k) = \delta_{jk} = \begin{cases} 1, j = k, \\ 0, j \neq k. \end{cases}$$

За лінійністю ця рівність розповсюджується на всі вектори простору:

$(G^T x, G^T y) = (x, y), \quad x, y \in \mathbb{R}^n.$	(4.7)
---	---------

Оператор, що переводить евклідовий (у випадку комплексного простору - унітарний) простір на весь простір і не змінює скалярного добутку, називається ортогональним (відповідно, унітарним) оператором.

Отже, матриця G^T - ортогональна. Легко бачити, що G - також ортогональна.

З властивості ортонормованості векторів-стовпчиків і векторів-рядків ортогональної матриці випливає її властивість, яку можна вважати другим означенням ортогональної матриці:

$$GG^T = G^T G = I \Rightarrow G^{-1} = G^T = G^*.$$

З такими операторами і їхніми матрицями ми ще зустрінемося в інших лекціях.

Доведемо, що, виходячи з довільного базису, можна побудувати ортонормований. Ця процедура називається **метод ортогоналізації Грама - Шмідта**.

Нехай $a = \{a_1, \dots, a_n\}$ - базис. Покажемо, як за допомогою деяких лінійних перетворень одержати з нього ортонормований базис, що буде позначено $e = \{e_1, \dots, e_n\}$.

Перший крок: $e_1 \equiv \frac{1}{\|a_1\|} a_1$. *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 $\alpha_1 = (a_2, e_1)$. Тоді

$b_2 = a_2 - (a_2, e_1) e_1 = a_2 - (p_{e_1} a_2) e_1$. Тепер покладемо

$$e_2 \equiv \frac{1}{\|b_2\|} b_2 = \frac{1}{\|a_2 - \alpha_1 e_1\|} (a_2 - \alpha_1 e_1).$$

Якщо вже зроблено m кроків, покажемо, як треба діяти на $(m+1)$ -му кроці.
Візьмемо допоміжний вектор

$$b_{m+1} = a_{m+1} - (a_{m+1}, e_1) e_1 - (a_{m+1}, e_2) e_2 - \dots - (a_{m+1}, e_m) e_m.$$

Очевидно, він ортогональний всім векторам $a_1, \dots, a_m$ (перевірте!). І остаточно

$$e_{m+1} \equiv \frac{1}{\|b_{m+1}\|} b_{m+1}.$$

Приклади:

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

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

Читаць повинен перевірити, чи виконуються аксіоми скалярного добутку.

2. Простір $\mathbb{C}^n$. Нехай

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

($\bar{y}_i$, — 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 $a_i = 0$, $i = 1, \dots, n$.

Для доведення слід зазначити, що ці суми означають скалярні добутки і вектор $a = (a_1, a_2, \dots, a_n)$ — ортогональний довільному вектору $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

Тоді з попередніх міркувань випливає, що $a = 0$.

Теорема 4.2. Нехай $a_1, \dots, a_n$ - довільний базис, $g_{ij} = (a_i, a_j)$. Тоді наступні вектори утворюють ортогональний базис:

$b_1 = a_1, b_k = \begin{vmatrix} g_{11} & g_{12} & \dots & g_{1k} \\ \cdot & \cdot & \dots & \cdot \\ g_{k-1,1} & g_{k-1,2} & \dots & g_{k-1,k} \\ a_1 & a_2 & \dots & a_k \end{vmatrix}, k = 2, \dots, n,$	(4.8)
--	-------

причому $\|b_k\| = \sqrt{G_{k-1} G_k}, G_0 = 1, G_k = \begin{vmatrix} g_{11} & \dots & g_{1k} \\ \cdot & \dots & \cdot \\ g_{k1} & \dots & g_{kk} \end{vmatrix}, k = 1, \dots, n.$

Доведення. Очевидно,

$$(a_i, b_k) = \begin{vmatrix} g_{11} & g_{12} & \dots & g_{1k} \\ \cdot & \cdot & \dots & \cdot \\ g_{k-1,1} & g_{k-1,2} & \dots & g_{k-1,k} \\ g_{i1} & g_{i2} & \dots & g_{ik} \end{vmatrix}.$$

При $i = 1, \dots, k-1$ маємо визначник, у якого два рядки рівні, тому $(a_i, b_k) = 0$.

Також

$$(a_k, b_k) = \begin{vmatrix} g_{11} & g_{12} & \dots & g_{1k} \\ \cdot & \cdot & \dots & \cdot \\ g_{k-1,1} & g_{k-1,2} & \dots & g_{k-1,k} \\ g_{k1} & g_{k2} & \dots & g_{kk} \end{vmatrix} = G_k.$$

Зрозуміло, що $G_k \neq 0$, оскільки це матриця Грама системи векторів, що утворюють базис.

Розкриваючи визначник (4.8) за останнім стовпчиком, одержимо:

$$b_k = G_{k-1} a_k + D_{k,k-1} a_{k-1} + \dots + D_{k,1} a_1,$$

де $D_{k,1}, \dots, D_{k,k-1}$ - алгебраїчні доповнення до елементів останнього рядка. Вектор $a \triangleq D_{k,k-1}a_{k-1} + \dots + D_{k,1}a_1 = b_k - G_{k-1}a_k$ належить підпростору, породженому векторами $a_1, \dots, a_{k-1}$, тому він ортогональний до b_k . Отже,

$$(a, b_k) = (b_k - G_{k-1}a_k, b_k) = 0 \Rightarrow (b_k, b_k) = G_{k-1}(a_k, b_k) = G_{k-1}G_k.$$

Це й доводить теорему.

Ортогональна сума підпросторів

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 $e \cup f$ 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.

Дамо інтерпретацію однорідної системи лінійних рівнянь з точки зору геометрії евклідового простору:

$ \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= 0, \\ &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= 0. \end{aligned} $	(4.9)
--	-------

Введемо позначення для наступних векторів:

$$a_1 = (a_{11}, a_{12}, \dots, a_{1n})^T, \dots, a_m = (a_{m1}, a_{m2}, \dots, a_{mn})^T, x = (x_1, x_2, \dots, x_n)^T.$$

Очевидно, кожне рівняння з (4.09) інтерпретується як умова ортогональності вектора x і кожного з векторів $a_k, k=1, \dots, m$. Як відомо, вектори $x \in \mathbb{R}^n$, координати яких задовольняють лінійній системі рівнянь, утворюють лінійний підпростір в $\mathbb{R}^n$. Будемо позначати його L . З (4.09) випливає, що він ортогональний підпростору, що є ортогональним доповненням до L , натягнутому на вектори $a_k, k=1, \dots, m$. Будемо позначати його через $L^\perp$. Один з його базисів – база системи цих векторів. Для зручності будемо вважати, що її утворюють перші r вектори $a_1, \dots, a_r$. При цьому $\mathbb{R}^n = L \oplus L^\perp$.

Запропонуємо кілька способів знаходження проекцій даного вектора $x \in \mathbb{R}^n$ на ці два ортогональні підпростори.

1. Знайдемо фундаментальну систему розв'язків лінійної системи (4.9). Для неї будемо вживати позначення: $a_{r+1}, \dots, a_n$. Розкладемо x за базисом (взагалі кажучи, неортогональним), що є об'єднанням систем $a_1, \dots, a_r$ і $a_{r+1}, \dots, a_n$:

$$x = x_1 a_1 + \dots + x_r a_r + x_{r+1} a_{r+1} + \dots + x_n a_n.$$

Для знаходження коефіцієнтів розкладу будемо множити попередню рівність по черзі на базисні вектори:

$ \begin{aligned} &(a_1, a_1)x_1 + \dots + (a_r, a_1)x_r + (a_{r+1}, a_1)x_{r+1} + \dots + (a_n, a_1)x_n = (x, a_1), \\ &\dots \\ &(a_1, a_r)x_1 + \dots + (a_r, a_r)x_r + (a_{r+1}, a_r)x_{r+1} + \dots + (a_n, a_r)x_n = (x, a_r), \\ &(a_1, a_{r+1})x_1 + \dots + (a_r, a_{r+1})x_r + (a_{r+1}, a_{r+1})x_{r+1} + \dots + (a_n, a_{r+1})x_n = (x, a_{r+1}), \\ &\dots \\ &(a_1, a_n)x_1 + \dots + (a_r, a_n)x_r + (a_{r+1}, a_n)x_{r+1} + \dots + (a_n, a_n)x_n = (x, a_n). \end{aligned} $	(4.10)
---	--------

Система має єдиний розв'язок, оскільки її визначник – це визначник Грама лінійно незалежної системи векторів, який, як відомо, не дорівнює нулю. Проекцією на підпростір L є лінійна комбінація $x_L \triangleq x_{r+1}a_{r+1} + \dots + x_na_n$, а проекцією на $L^\perp$ - $x_L^\perp \triangleq x_1a_1 + \dots + x_ra_r$.

Зауважимо, що система (4.10) розпадається у пряму суму систем, оскільки $(a_j, a_k) = 0 \forall j = 1, \dots, r; k = r+1, \dots, n$.

2. Система (4.10) буде мати діагональний вигляд, якщо системи базисних векторів окремо з L і окремо з $L^\perp$ ортогоналізувати за методом Грама-Шмідта.

3. Можна не розв'язувати систему (4.10), а шукати проекції на вказані підпростори безпосередньо. Будемо шукати $x_L^\perp$, проекцію вектора x на $L^\perp$ у наступному вигляді:

$$x_L^\perp = \alpha_1 a_1 + \dots + \alpha_r a_r.$$

Як відомо, $x - x_L^\perp = x_L \in L, x_L \perp L^\perp$. Запишемо умову ортогональності:

$$(x - x_L^\perp, a_j) = (x - \alpha_1 a_1 - \dots - \alpha_r a_r, a_j) = 0, j = 1, \dots, r.$$

Таким чином маємо лінійну систему рівнянь з визначником Грама у лівій частині:

$$\begin{aligned}(a_1, a_1)\alpha_1 + \dots + (a_r, a_1)\alpha_r &= (x, a_1), \\ \dots \\ (a_1, a_r)\alpha_1 + \dots + (a_r, a_r)\alpha_r &= (x, a_r).\end{aligned}$$

Знайшовши коефіцієнти $\alpha_1, \dots, \alpha_r$, будемо мати проекцію $x_L^\perp$ вектора на $L^\perp$, після чого знайдемо його проекцію на L : $x_L = x - x_L^\perp$.

ЛІНІЙНІ ФУНКЦІОНАЛИ

Означення 4.2. Лінійним функціоналом на E називається відображення $f: E \rightarrow \mathbb{C}$, таке, що $\forall \alpha, \beta \in \mathbb{C} \quad f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$. Іноді функціонали називають ковекторами.

Якщо простір E нормований, то можна визначити норму функціонала за формулою:

$\ f\ = \sup_{\ x\ =1} f(x) .$	(4.11)
----------------------------------	--------

З'ясуємо, який вигляд має лінійний функціонал у випадку унітарного простору.

Теорема 4.2. Будь-який лінійний функціонал в унітарному просторі має вигляд

$$f(x) = (x, y), \tag{4.12}$$

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 $f(f(x)z - f(z)x) = f(x)f(z) - f(z)f(x) = 0$.

Отже, $(f(x)z - f(z)x, z) = 0$, звідки $f(x)(z, z) = (x, \overline{f(z)}z)$. Якщо покласти $y = \frac{\overline{f(z)}}{(z, z)}z$, 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 $x = \frac{y}{\|y\|}, \|x\| = 1$, одержимо: $f(x) = \frac{f(y)}{\|y\|} = \left(\frac{y}{\|y\|}, y \right) = \|y\|$, звідки $\|f\| \geq \|y\|$. Отже, $\|f\| = \|y\|$, і теорему повністю доведено.

Зауваження. Подібний факт у функціональному аналізі називається теоремою Ріса.

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 $f(x) = 0 \quad \forall x \in E$, будемо вважати, що $f = 0$.

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

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

δ_{ij} називається «дельтою Кронекера».

Доведемо, що ці лінійні функціонали утворюють базис в E .

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

$$\left(\sum_{i=1}^n \alpha_i e^i \right) (e_k) = \sum_{i=1}^n \alpha_i e^i (e_k) = \sum_{i=1}^n \alpha_i \delta_{ik} = \alpha_k = 0.$$

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

$f = \sum_{i=1}^n f(e_i) e^i.$	(4.13)
--------------------------------	--------

Подіємо лівою і правою частинами рівності на довільний вектор $x = \sum_{i=1}^n x_i e_i$:

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

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

що й 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 за лінійністю продовжити на весь простір.

Зробимо підсумок. Побудовано ізоморфізм, який даному базису $e = \{e_1, e_2, \dots, e_n\}$ зіставляє спряжений (дуальний) базис $e^* = \{e_1^*, e_2^*, \dots, e_n^*\} : e_i^*(e_j) = \delta_{ij}$. З побудови зрозуміло, що такий ізоморфізм залежить від базису, тобто, не є канонічним (за

означенням канонічний ізоморфізм лінійних просторів не залежить від вибору базису). Якщо змінити базис простору, цей ізоморфізм, взагалі кажучи, зміниться. Пояснимо це на прикладі одновимірного простору.

Нехай $e \neq 0$ - базисний вектор в E_1 . Він єдиний через одновимірність E_1 . Нехай $e^* \in E^*, \dim(E^*)=1$ - вектор дуального базису, $e^*(e)=1$. Задамо відображення, що здійснює ізоморфізм: $f_1: E \rightarrow E^*, f_1(e)=e^*$. Зробимо зміну базису у $E_1: Te = te = g \in E_1, t \in \mathbb{R}, t \neq 0$. Цьому базису відповідає дуальний базис у $E^*: g^* = \frac{1}{t}e^*$, оскільки $g^*(g) = \frac{1}{t}e^*(te) = \frac{t}{t}e^*(e) = 1$.

Отже, $f_1(e)=e^*, f_2(te) \triangleq \frac{1}{t}e^* \Rightarrow f_2(e) = \frac{1}{t^2}e^*$. Звідси робимо висновок, що для співпадіння ізоморфізмів $f_1(e)=f_2(e)$ необхідно і достатньо, щоб виконувалась умова: $t^2=1 \Leftrightarrow t=\pm 1$. При інших значеннях t ізоморфізми різні, тобто, ізоморфізми між дуальними просторами суттєво залежать від базису.

Пізніше будуть описані канонічні ізоморфізми між дуальними просторами розмірностей більших за одиницю.

Можна розглянути двічі спряжений простір до $E: (E^*)^*$, тобто, простір лінійних функціоналів над простором лінійних функціоналів над E . З попередніх досліджень випливає, що $n = \dim E = \dim E^* = \dim (E^*)^*$.

Очевидно, кожному вектору $x \in E$ можна зіставити елемент $(x^*)^* \in (E^*)^*$, що діє за правилом: $\forall f \in E^* \langle (x^*)^*, f \rangle = \langle f, x \rangle$. Лінійність відображення $x \rightarrow (x^*)^*$ очевидна. Взаємна однозначність випливає з рівності розмірностей цих просторів. Отже, побудовано ізоморфізм лінійного простору і його двічі спряженого, який не залежить від базису, що означає канонічність ізоморфізму.

Зробимо тепер важливе зауваження: завдяки теоремі Ріса про зображення лінійного функціоналу над унітарним простором, побудова якого не спиралась на базис, ізоморфізм між унітарним простором і його спряженим канонічний.

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

$e^j(x) = (x, e^j) = \left(\sum x_i e_i, e_j \right) = \sum x_i (e_i, e_j) = x_j.$	(4.14)
---	----------

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.

Знайдемо тепер формулу для обчислення скалярного добутку через коефіцієнти розкладу векторів за неортогональним базисом у дійсному просторі.

Нехай $g = \{g_1, g_2, \dots, g_n\}$ - неортогональний базис. Знайдемо вираз для скалярного добутку двох векторів $x, y \in E_n$, розкладених за цим базисом:

$$x = x_1 g_1 + x_2 g_2 + \dots + x_n g_n, \quad y = y_1 g_1 + y_2 g_2 + \dots + y_n g_n.$$

Зіставимо цим векторам вектори-стовпчики коефіцієнтів розкладу за цими базисами:

$$x \sim [x]_g = (x_1, x_2, \dots, x_n)^T, \quad y \sim [y]_g = (y_1, y_2, \dots, y_n)^T.$$

$$G_g = \begin{pmatrix} (g_1, g_1) & (g_1, g_2) & \dots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \dots & (g_2, g_n) \\ \cdot & \cdot & \dots & \cdot \\ (g_n, g_1) & (g_n, g_2) & \dots & (g_n, g_n) \end{pmatrix},$$

як вже відомо, є матрицею Грама системи $g = \{g_1, g_2, \dots, g_n\}$.

Маємо:

$$(x, y) = (x_1 g_1 + x_2 g_2 + \dots + x_n g_n, y_1 g_1 + y_2 g_2 + \dots + y_n g_n) = \sum_{j=1}^n \sum_{k=1}^n x_j y_k (g_j, g_k) =$$

$$= (x_1, x_2, \dots, x_n) \begin{pmatrix} (g_1, g_1) & (g_1, g_2) & \dots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \dots & (g_2, g_n) \\ \cdot & \cdot & \dots & \cdot \\ (g_n, g_1) & (g_n, g_2) & \dots & (g_n, g_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} =$$

$$= [x]_g^T G_g [y]_g$$

Остаточню:

$(x, y) = [x]_g^T G_g [y]_g = (x_1, x_2, \dots, x_n) \begin{pmatrix} (g_1, g_1) & (g_1, g_2) & \dots & (g_1, g_n) \\ (g_2, g_1) & (g_2, g_2) & \dots & (g_2, g_n) \\ \cdot & \cdot & \dots & \cdot \\ (g_n, g_1) & (g_n, g_2) & \dots & (g_n, g_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}.$	(4.15)
---	--------

Відзначимо, що $[x]_g^T = (x_1, x_2, \dots, x_n)$ - це ковектор.

Контрольні запитання

- 1). Які властивості має скалярний добуток у дійсних та комплексних просторах?
- 2). Що таке унітарний та евклідовий простори?
- 3). В чому полягає метод ортогоналізації Грама-Шмідта?
- 3). Чи є ортогональна сума підпросторів їхньою прямою сумою?
- 4). Що таке лінійний функціонал і в якій формі можна записати його у унітарному просторі?

**ЛЕКЦІЯ II.5. СПРЯЖЕНИЙ ОПЕРАТОР ТА ЙОГО МАТРИЦЯ.
САМОСПРЯЖЕНІ, НОРМАЛЬНІ ТА УНІТАРНІ ОПЕРАТОРИ**

План

- 1). Поняття спряженого оператора. Самоспряжені оператори.
- 2). Властивості спряженого оператора та його матриця.
- 3). Ізометричні та унітарні оператори. Спряжений оператор до унітарного.
- 3). Проектори.
- 4). Нормальні оператори.

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

$$(A^* f)(x) = f(Ax) \quad (5.1)$$

Зауважимо, що часто дію лінійного функціоналу записують у більш симетричному вигляді: $f(x) = \langle f, x \rangle$. Тоді співвідношення (5.1) можна записати у вигляді $\langle A^* f, x \rangle = \langle f, Ax \rangle$.

Тепер розглянемо спряжений оператор у просторах зі скалярним добутком.

Нехай E_n і E_m - простори зі скалярним добутком. Розглянемо для $\forall x \in E_n, y \in E_m, A \in L(E_n, E_m)$ вираз $(y, Ax)_m$. Вже відомо, що він визначає деякий лінійний функціонал f на E_n в разі $K = \mathbb{R}$ або антилінійний при $K = \mathbb{C}$: $f(x) = (y, Ax)_m$ (це випливає із властивостей лінійності оператора і відповідно лінійності або антилінійності по другій змінній скалярного добутку). Отже, за теоремою Ріса⁶ (яка справедлива і для антилінійних функціоналів) він може бути зображений у вигляді скалярного добутку x на деякий елемент $y^* \in E_n$: $(y, Ax)_m = (y^*, x)_n$. Таким чином, одержали відповідність $E_m \ni y \rightarrow y^* \in E_n$, що буде лінійною завдяки властивостям скалярного добутку. Тобто, одержали

⁶ Ріс, Фрідьєш (22/01/1880, Дьйор, Австро-Угорщина – 28/02/1956, Будапешт, Угорщина) – видатний математик, один з засновників функціонального аналізу.

лінійний оператор A^* , який буде спряженим до A . Отже, якщо $A \in L(E_n, E_m)$, то $A^* \in L(E_m, E_n)$ і

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

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

Нехай (a_{ik}) - матриця оператора A в ортонормованих базисах

$$e = \{e_1, \dots, e_n\}, f = \{f_1, \dots, f_m\}.$$

Нагадаємо, що її елементи знаходяться з рівностей (5.7), частина 1:

$$Ae_j = a_{1j} f_1 + \dots + a_{kj} f_k + \dots + a_{mj} f_m, k = 1, \dots, n.$$

Помноживши обидві частини рівності скалярно на $f_k, k = 1, \dots, m$, одержимо:

$a_{kj} = (Ae_j, f_k), k = 1, \dots, n, j = 1, \dots, m.$	(5.2)
---	-------

Знайдемо матрицю оператора $A^* - (a_{jk})$. Користуючись (5.7), частина 1, маємо (при $K = \mathbb{C}$)

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

Остаточно,

$$b_{jk} = \overline{a_{kj}} \tag{5.3}$$

$j = 1, \dots, m, k = 1, \dots, n$, тобто всі елементи матриці A треба замінити на спряжені і рядки записати стовпчиками. Таку матрицю зручно назвати спряжено-транспонованою. У випадку $K = \mathbb{R}$, природно, одержимо транспоновану матрицю.

Означення 5.1. Лінійний оператор $A \in L(E_n, E_n)$, називається самоспряженим або ермітовим⁷, якщо $A = A^*$. Тобто, для нього справедлива рівність $(Ax, y) = (x, Ay) \forall x, y \in E_n$.

Самоспряженому оператору, як ми бачили, повинна відповідати в разі $K = \mathbb{R}$ і в ортонормованому базисі симетрична матриця, тобто така, що дорівнює своїй транспонованій. В разі $K = \mathbb{C}$ її елементи задовольняють рівність $a_{jk} = \overline{a_{kj}}$, тобто, після транспонування і заміни всіх елементів на комплексно спряжені одержимо ту ж саму матрицю.

Твердження 5.1. Якщо деякий підпростір унітарного простору інваріантний відносно лінійного оператора, то його ортогональне доповнення інваріантне відносно оператора, спряженого до даного.

Доведення. Нехай $E_n = E' \oplus E''$, $E' \perp E''$, $AE' \subset E'$. Візьмемо довільні вектори: $x' \in E'$, $x'' \in E''$. Тоді $Ax' \in E' \Rightarrow (Ax', x'') = 0$. Тоді $0 = (Ax', x'') = (x', A^*x'')$.

Оскільки рівність справедлива для довільного вектора $x' \in E'$, то $A^*x'' \in E'$, що й доводить твердження.

Твердження 5.2. Неоднорідна система $Ax = b$, $x \in E_n$, $b \in E_m$ має розв'язки тоді і тільки тоді, коли вектор правої частини b ортогональний до $\ker A^*$.

Доведення. Необхідність. Нехай існує розв'язок системи $Ax = b$. Нехай $y \in \ker A^* \Rightarrow A^*y = 0$. Маємо: $(b, y) = (Ax, y) = (x, A^*y) = 0$, тобто, $b \perp \ker A^*$.

Достатність. Нехай $y \in \ker A^* \Rightarrow A^*y = 0$ і x - довільний вектор з E_n . Маємо:

$0 = (x, A^*y) = (Ax, y) = (b, y)$. Це означає, що вектор b такий, що $b = Ax$, ортогональний $\ker A^*$.

Означення 5.2. 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.

⁷ Шарль Ерміт (Charles Hermite); 24/12/1822, Д'йоз, Франція – 14/01/1901, Париж.

Означення 5.3. 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 5.1. 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 $v_i = (v_{i1} v_{i2} \dots v_{in})$ матриці V^* на j -й стовпчик $(v_j)^* = (v_{1j}^* v_{2j}^* \dots v_{nj}^*)^T$, одержимо 0, якщо $i \neq j$ і 1 при $i = j$. Тобто, $\sum_{k=1}^n v_{ik} v_{kj}^* = \sum_{k=1}^n v_{ik} \bar{v}_{jk} = \delta_{ij}$, але ця рівність свідчить, 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. $\det(V^*) = \det(\bar{V}^T) = \det(\bar{V})$. Оскільки $1 = \det(VV^*) = \det(V)\det(V^*) = \det(V)\det(\bar{V}) = \det(V)\overline{\det(V)} = |\det(V)|^2$, отже, $|\det(V)| = 1$.

Якщо повернутись до матриць повороту площини $U(\varphi)$, 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 $SL(n, K)$. $O(n, K) \cap SL(n, K) = SO(n, K)$ - спеціальна ортогональна група. Отже, більш точно, матриці повороту $U(\varphi)$ належать до $SO(2, \mathbb{R})$.

Приклад 5.1. Заміна координатних осей. Здійснимо перетворення координат так, щоб осі Ox і Oy 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

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (5.3)$$

(перевірте!). Ця матриця змінює орієнтацію:

$$Je_1 = J(1; 0)^T = (0; 1)^T, Je_2 = J(0; 1)^T = (1; 0)^T.$$

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

Приклад 5.2. 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: $P(x, y, z)^T = (x, y, 0)^T$. Тоді його матриця матиме вигляд:

$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$	(5.4)
--	-------

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, $Q^2 = (I - P)^2 = I - 2P + P^2 = I - 2P + P = I - P = Q$. 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 $x = Py = (I - P)z$. Звідси $P(y + z) = z$. Поділявши 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 $x = Px + (I - P)x = Px + Qx$. Це означає, що $E = E_1 \oplus E_2$. Якщо $E_1 \perp E_2$, то P називається ортогональним проектором або ортопроектором.

Теорема 5.1. Для того щоб проектор P був ортогональним, необхідно і достатньо, щоб він був самоспряженим.

Доведення. Нехай P — ортогональний проектор. Тоді для довільних $x \in E, y \in E$ $Px \perp Qx$, тобто, $(Px, (I - P)y) = 0$, звідки

$$(Px, (I - P)y) = ((I - P^*)Px, y) = ((P - P^*P)x, y) = 0.$$

Оскільки y — довільний вектор, $(P - P^*P)x = 0$, звідки з довільності x випливає, що $P^*P = P$. З іншого боку, $(P^*P)^* = P^*P = P^*$ (якщо застосувати спряження до обох частин рівності), звідки $P = P^*$. Навпаки, нехай $P = P^*$. Тоді для довільного $x \in E$ одержимо:

$$(Px, (I - P)x) = ((I - P)Px, x) = ((P - PP)x, x) = ((P - P)x, x) = 0,$$

Тобто $Px \perp Qx$.

Означення 5.4. Оператор, що діє в унітарному просторі називається нормальним, якщо він комує з своїм спряженням:

$$AA^* = A^*A. \quad (5.5)$$

Отже, самоспряжені (ермітові) і унітарні оператори нормальні.

Ізоморфізми між дуальними просторами

Повернемося до питання існування канонічних ізоморфізмів між E і E^* .

Нагадаємо, що вже відзначалось, що ізоморфізми між лінійними просторами (вони не передбачаються унітарними) E і E^* , взагалі кажучи, не є канонічними, тобто, залежать від вибору базису в E . Зупинимось на вивченні ситуації в дійсних просторах розмірностей більших за одиницю.

Нехай у просторі E вибрано базиси: $\{e_1, \dots, e_n\} \subset E, \{\varepsilon_1, \dots, \varepsilon_n\} \subset E$. Нехай T - матриця переходу від $\{e_1, \dots, e_n\}$ до $\{\varepsilon_1, \dots, \varepsilon_n\}$: $Te_k = \varepsilon_k, k = 1, \dots, n$. Нехай $\{e_1^*, \dots, e_n^*\} \subset E^*, \{\varepsilon_1^*, \dots, \varepsilon_n^*\} \subset E^*$ - дуальні базиси, $e_i^*(e_j) = \delta_{ij}, \varepsilon_i^*(\varepsilon_j) = \delta_{ij}$. Нехай S - матриця переходу від $\{e_1^*, \dots, e_n^*\}$ до $\{\varepsilon_1^*, \dots, \varepsilon_n^*\}$: $Se_k^* = \varepsilon_k^*, k = 1, \dots, n$. Знайдемо зв'язок між матрицями S і T .

Маємо: $\langle e_i^*, e_j \rangle = \delta_{ij} = \langle \varepsilon_i^*, \varepsilon_j \rangle = \langle Se_i^*, Te_j \rangle = \langle e_i^*, S^T Te_j \rangle$. Звідси

$$S^T T = I \Rightarrow T^T S = I \Rightarrow S = (T^T)^{-1}.$$

Тепер можна відповісти на питання про умови співпадіння ізоморфізмів $E \rightarrow E^*$, побудованих за допомогою базисів $\{e_1, \dots, e_n\}$ і $\{\varepsilon_1, \dots, \varepsilon_n\}$.

Нехай

$$f, g : E \rightarrow E^*, \langle f, e_i \rangle = e_i^*, \langle g, \varepsilon_i \rangle = \varepsilon_i^*, T\{e_1, \dots, e_n\} \rightarrow \{\varepsilon_1, \dots, \varepsilon_n\}, Te_k = \varepsilon_k, k = 1, \dots, n.$$

Тоді

$$\langle g, e_i \rangle = \langle g, T^{-1} \varepsilon_i \rangle = T^{-1} \langle g, \varepsilon_i \rangle = T^{-1} \varepsilon_i^* = T^{-1} (T^T)^{-1} e_i^*.$$

Порівнюючи з рівністю $\langle f, e_i \rangle = e_i^*$, приходимо до висновку, що $\langle f, e_i \rangle = \langle g, e_i \rangle$ тоді і тільки тоді, коли $T^{-1} (T^T)^{-1} e_i^* = e_i^* \Leftrightarrow T^{-1} (T^T)^{-1} = I \Leftrightarrow T^T = T^{-1}$, а це означає, що матриця T ортогональна.

Контрольні запитання

- 1). Як вводиться спряжений оператор?
- 2). Які властивості має спряжений оператор?
- 3). Які оператори називаються ізометричними? Унітарними?
- 4). Яку властивість має оператор спряжений до унітарного?
- 3). Що таке проектор? Ортогональний проектор?
- 4). Який оператор називається нормальним?

ЛЕКЦІЯ II.6. ВЛАСНІ ЧИСЛА І ВЕКТОРИ ЛІНІЙНОГО ОПЕРАТОРА

План

- 1). Поняття власного числа лінійного оператора і відповідного йому власного вектора. Спектр оператора.
 - 2). Характеристичний многочлен матриці та форма його запису.
 - 3). Побудова базису, в якому оператор у унітарному просторі має верхньо-трикутну матрицю. Теорема Шура.
 - 4). Оператори з простим спектром.
 - 5). Анулюючі многочлени лінійного оператора. Мінімальний многочлен.
- Теорема Гамільтона-Келі.

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:

$$P(t) = t^m + a_1 t^{m-1} + \dots + a_m. \quad (6.1)$$

Многочленом $P(A)$ від оператора A (або від матриці A) будемо називати вираз

$$P(A) = A^m + a_1 A^{m-1} + \dots + a_m I, \quad (6.2)$$

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 6.1. Характеристичним многочленом матриці A називають многочлен

$\chi(\lambda) = \det(\lambda I - A) =$	$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{21} & \dots & -a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix} =$
---	---

$= \lambda^n + A_1 \lambda^{n-1} + \dots + A_k \lambda^{n-k} + \dots + A_n.$	(6.3)
--	-------

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 $(\lambda - a_{11}) \dots (\lambda - a_{nn})$ 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 $(\lambda - a_{11}) \dots (\lambda - a_{nn})$ як многочлен, то можна побачити, що коефіцієнт при λ^{n-1} дорівнює сумі діагональних елементів матриці A , взятій із знаком «мінус», а така сума називається слідом матриці і позначається $sp(A)$ (*шпур* – нім.) або $tr(A)$ (*трейс* – англ.):

$$sp(A) = tr(A) = \sum_{k=1}^n a_{kk}. \quad (6.4)$$

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 $a_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$ -

i -й вектор-стовпчик одиничної матриці. Будемо використовувати таку форму запису для визначника:

$$(-1)^n \chi(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} =$$

$$= |a_1 - \lambda e_1 \quad a_2 - \lambda e_2 \quad \dots \quad a_n - \lambda e_n|.$$

Тут $e_i, i = 1, \dots, n$ - орти канонічного базису.

Використовуючи лінійність визначника по кожному стовпчику, маємо:

$$\begin{aligned} \det(A - \lambda I) &= |a_1 - \lambda e_1 \quad a_2 - \lambda e_2 \quad \dots \quad a_n - \lambda e_n| = \\ &= |a_1 \quad a_2 \quad \dots \quad a_n| - \\ &- (|e_1 \quad a_2 \quad \dots \quad a_n| + |a_1 \quad e_2 \quad a_3 \quad \dots \quad a_n| + \dots + |a_1 \quad \dots \quad a_{n-1} \quad e_n|) \lambda + \\ &+ (|e_1 \quad e_2 \quad a_3 \quad \dots \quad a_n| + \dots + |a_1 \quad \dots \quad a_{n-2} \quad e_{n-1} \quad e_n|) \lambda^2 + \dots + \\ &+ (-1)^{n-k} (|e_1 \quad e_2 \quad \dots \quad e_{n-k} \quad a_{n-k+1} \quad \dots \quad a_n| + \dots + |a_1 \quad a_2 \quad \dots \quad a_k \quad e_{k+1} \quad \dots \quad e_n|) \lambda^{n-k} + \\ &+ \dots + (-1)^n |e_1 \quad e_2 \quad \dots \quad e_n| \lambda^n. \end{aligned}$$

Звідси видно, що коефіцієнт при λ^{n-1} - сума визначників вигляду

$(-1)^{n-1} |e_1 \quad \dots \quad e_{k-1} \quad a_k \quad e_{k+1} \quad \dots \quad e_n| = (-1)^{n-1} a_k$, тобто, $(-1)^{n-1} \text{Sp}(A)$, а відповідний коефіцієнт в характеристичному многочлені тоді дорівнює $(-1)^n (-1)^{n-1} \text{Sp}(A) = -\text{Sp}(A)$. Аналогічно вільний член характеристичного многочлена дорівнює $(-1)^n |a_1 \quad a_2 \quad \dots \quad a_n| = (-1)^n \det(A)$, що вже було з'ясовано в інший спосіб.

Зазначимо, що, наприклад, вираз $|e_1 \quad e_2 \quad \dots \quad e_{n-k} \quad a_{n-k+1} \quad \dots \quad a_n|$ - це мінор k -го порядку визначника, який одержано після викреслювання рядків і стовпчиків, що відповідають номерам $i = 1, \dots, n-k$ 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 $M_{i_1 \dots i_k}^{i_1 \dots i_k}, i_1 < \dots < i_k$ головний мінор, на діагоналі якого знаходяться елементи $a_{i_1}, \dots, a_{i_k}, i_1 < \dots < i_k$. Тоді остаточно:

$$\chi(\lambda) = \det(\lambda I - A) = (-1)^n \det(A - \lambda I) = \quad (6.5)$$

$$= \lambda^n - Sp(A)\lambda^{n-1} + (-1)^k \left(\sum_{i_1 \dots i_k} M_{i_1 \dots i_k}^{i_1 \dots i_k} \right) \lambda^{n-k} + \dots + (-1)^n \det(A).$$

Вільний член $\chi(\lambda)$ дорівнює значенню $\det(\lambda I - A)$ при $\lambda = 0$, тобто, $\det(-A) = (-1)^n \det(A)$.

Теорема 6.1. Характеристичний многочлен інваріантний відносно заміни базису в L .

Доведення випливає з властивості добутку двох матриць. Дійсно, при заміні базису за допомогою матриці переходу T матриця $\lambda I - A$ перейде в подібну матрицю $T^{-1}(\lambda I - A)T$, що випливає з формули (5.22), частина 1, а

$$\begin{aligned} \det(T^{-1}(\lambda I - A)T) &= \det(T^{-1})\det(\lambda I - A)\det(T) = (\det(T))^{-1}\det(T)\det(\lambda I - A) = \\ &= \det(\lambda I - A). \end{aligned}$$

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 $Sp(A + B) = Sp(A) + Sp(B)$.

Доведення очевидне.

Вправа. Довести, що $Sp(AB) = Sp(BA)$.

Доведення. Нехай $C = AB, D = BA$. Згідно з формулою (4.5), частина 1,

$$c_{ii} = \sum_{k=1}^n a_{ik}b_{ki}, \quad d_{ii} = \sum_{k=1}^n b_{ik}a_{ki},$$

звідки

$$sp(AB) = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki}, \quad sp(BA) = \sum_{k=1}^n \sum_{i=1}^n b_{ki} a_{ik} = sp(AB).$$

Наслідок. Для двох матриць (комутуючих чи некомутуючих) їхній комутант – це матриця, слід якої дорівнює нулю:

$$[A, B] = AB - BA = C \Rightarrow sp(C) = 0.$$

Для доведення треба використати результати двох попередніх вправ.

Означення 6.2. 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 $(A - \lambda I)x = 0$.

Теорема 6.2. 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 $\text{Ker}(\lambda I - A)$) 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 $\{tx\} \subset \text{Ker}(\lambda I - A)$, $t \in K$.

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 , $r_1 + \dots + r_m = n$. Тоді

$$\det(\lambda E - A) = \prod_{i=1}^m (\lambda - \lambda_i)^{r_i}. \quad (6.6)$$

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 6.3. 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, $\alpha \neq 0$. Подіавши 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:

$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$	(6.7)
---	---------

Означення 6.3. 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: $\lambda \in \text{spec}(A)$.

Отже, спектр лінійного оператора, що діє у скінченновимірному лінійному просторі, – це множина його власних чисел. Якщо корені характеристичного многочлена прості, то спектр називається простим.

Твердження 6.1. Якщо спряжені оператори мають спільний власний вектор, то їхні власні числа, що відповідають цьому вектору, комплексно-спряжені.

Доведення. Нехай $Ax = \lambda x, A^*x = \mu x, x \neq 0$. Тоді

$$(Ax, x) = \lambda(x, x) = (x, A^*x) = (x, \mu x) = \bar{\mu}(x, x) \Rightarrow \lambda = \bar{\mu}.$$

Твердження 6.2. Нехай y - власний вектор оператора A^* і $E^{(n-1)}$ - ортогональне доповнення до одновимірного підпростору, породженому цим вектором. Тоді підпростір $E^{(n-1)}$ інваріантний відносно A .

Доведення. Спряжений до A^* оператор - це $A = (A^*)^*$. Тоді справедливість твердження випливає з твердження 5.1.

Тепер можна розглядати звуження оператора A на підпростір $E^{(n-1)}$. Завдяки твердженню 6.3. він діє в цьому підпросторі, і у нього там є спряжений, який має хоч один власний вектор. За доведеним оператор A має в $E^{(n-1)}$ інваріантний підпростір розмірності $n-2$, який позначимо $E^{(n-2)} : E^{(n-2)} \subset E^{(n-1)}$. Діючи за цією схемою і далі, одержимо ланцюг вкладених підпросторів:

$E^{(1)} \subset E^{(2)} \subset \dots \subset E^{(n-2)} \subset E^{(n-1)} \subset E_n.$	(6.7)
--	---------

Виберемо орт у одновимірному підпросторі $E^{(1)}$, який позначимо e_1 . У підпросторі $E^{(2)}$ виберемо ортогональний до нього орт e_2 - він існує, оскільки довільний базис в лінійному підпросторі можна ортогоналізувати за методом Грама-Шмідта. Продовжуючи цей процес, побудуємо ортонормований базис у E_n . Він має властивість: підпростори $E^{(k)}$, натягнуті на перші k векторів

$e_1, e_2, \dots, e_k$, інваріантні відносно оператора A . Випишемо матрицю оператора у побудованому базисі:

$$Ae_1 = a_{11}e_1,$$

$$Ae_2 = a_{12}e_1 + a_{22}e_2,$$

...

$$Ae_{n-1} = a_{1,n-1}e_1 + a_{2,n-1}e_2 + \dots + a_{n-1,n-1}e_{n-1},$$

$$Ae_n = a_{1,n}e_1 + a_{2,n}e_2 + \dots + a_{n-1,n}e_{n-1} + a_{n,n}e_n.$$

Користуючись правилом записування матриці лінійного оператора, якщо відомо розклад його дії на кожен базисний вектор, одержимо верхньо-трикутну матрицю.

$$[A]_{ee} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Таким чином, ми прийшли до важливого факту:

Теорема 6.4 (Шура). Для кожного лінійного оператора у унітарному лінійному просторі можна побудувати такий ортонормований простір, в якому його матриця є трикутною.

Розглянемо тепер більш детально оператор, що має простий спектр.

Теорема 6.5. Нехай лінійний оператор A має простий спектр. Тоді існує такий вектор $e \in L$, що вектори $e, Ae, A^2e, \dots, A^{n-1}e$ утворюють базис L .

Доведення. Нехай $e_1, e_2, \dots, e_n$ - деякі ненульові власні вектори оператора A , які відповідають різним власним числам $\lambda_1, \lambda_2, \dots, \lambda_n$. Візьмемо вектор $e = e_1 + e_2 + \dots + e_n$ і припустимо, що $e, Ae, A^2e, \dots, A^{n-1}e$ лінійно залежні. У такому разі існують числа $a_1, a_2, \dots, a_n$ - не всі нульові, що $a_0e + a_1Ae + a_2A^2e + \dots + a_{n-1}A^{n-1}e = 0$, звідки, враховуючи вигляд e , одержимо:

$$\sum_{i=1}^n (a_0 + a_1 \lambda_i + a_2 \lambda_i^2 + \dots + a_{n-1} \lambda_i^{n-1}) e_i = 0.$$

З лінійної незалежності e_i випливає, що $a_0 + a_1 \lambda_i + a_2 \lambda_i^2 + \dots + a_{n-1} \lambda_i^{n-1} = 0$ при кожному $i = 1, \dots, n$. Це означає: многочлен степені $n - 1$ має n різних коренів, що можливо тоді і лише тоді, коли $a_0 = a_1 = a_2 = \dots = a_{n-1} = 0$, що і доводить лінійну незалежність системи векторів $e, Ae, A^2e, \dots, A^{n-1}e$. Їхня кількість збігається з розмірністю простору, тому вони утворюють базис.

Означення 6.4. Вектор $e \in H \subseteq L$ називається твірним або циклічним вектором m -вимірного підпростору H , що відповідає оператору A , якщо система $e, Ae, A^2e, \dots, A^{n-1}e$ – базис H .

Отже, Теорема 6.5 стверджує, що оператор A з простим спектром має твірні вектори всього простору. Твірний вектор e має вигляд $e = e_1 + e_2 + \dots + e_n$, де $e_1, e_2, \dots, e_n$ – ненульові власні вектори, які відповідають різним власним числам.

Повернемося тепер до характеристичного многочлена (6.3).

Анулюючі многочлени лінійного оператора

Теорема 6.6. (Гамільтона⁸ - Келі⁹). 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.

⁸ **Сер Гамільтон, Уїльям Р'оуэн** (Англ.. *William Rowan Hamilton*; 4/08/1805 – 2/08/1865) – великий математик, механік-теоретик, фізик-теоретик.

⁹ **Келі, Артур** (англ. *Arthur Cayley*; 16/08/1821, Річмонд, - 20/01/1895, Кембрідж) - видатний математик.

$$\begin{array}{l} \lambda^n \left| B_1 = I \right| \times A^n \\ \lambda^{n-1} \left| -B_1 A + B_2 = a_1 I \right| \times A^{n-1} \\ \lambda^{n-2} \left| -B_2 A + B_3 = a_2 I \right| \times A^{n-2} \\ \dots\dots\dots \\ \lambda^2 \left| -B_{n-2} A + B_{n-1} = a_{n-2} I \right| \times A^2 \\ \lambda^1 \left| -B_{n-1} A + B_n = a_{n-1} I \right| \times A \\ \lambda^0 \left| -B_n A = a_n I \right| \end{array} \quad (6.8)$$

Множачи рівності на степені A , як це показано у (6.8), і додаючи результати, дістанемо $0 = A^n + a_1 A^{n-1} + \dots + a_{n-1} A + a_n I$, 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 (приклад 1.2, ЛЕКЦІЯ II.1), 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 $M(t)$.

Зауваження. Наявність у матриці порядку n анулюючого многочлену степені $m \leq n$ можна трактувати таким чином: m -та степінь матриці є лінійною комбінацією її менших степенів:

$A^m = a_1 A^{m-1} + \dots + a_{m-1} A + a_m I.$	(6.9)
--	-------

Приклад 6.1. $A = tI$. Цей оператор називається скалярним. Його характеристичний многочлен має наступний вигляд:

$\chi(t) = \det(\lambda I - tI) = \det((\lambda - t)I) = (\lambda - t)^n$, а мінімальний многочлен - $M(t) = (t - \lambda)$, що випливає з вигляду A . При цьому всі вектори з простору E_n власні з власним числом $\lambda = t$. У кожному базисі оператор має одну й ту ж матрицю, бо $T(tI)T^{-1} = tTT^{-1} = tI$.

Твердження 6.3. Комутуючі між собою лінійні оператори обов'язково мають спільний власний вектор.

Доведення. Нехай лінійні оператори A і B комутують і нехай $Ax = \lambda x, x \neq 0$. Тоді комутують і їхні степені, і

$AB^k x = B^k Ax = \lambda B^k x, k \in \mathbb{N} \cup \{0\}.$	(6.10)
---	----------

Як випливає з попереднього зауваження, степені матриці B залежні, і виконується співвідношення (6.9). Отже, відпростір, породжений сукупністю векторів $B_x \triangleq \{x, Bx, \dots, B^{m-1}x\}$, інваріантний відносно оператора B і тому він має тут власний вектор, який позначимо $u: Bu = \mu u, u \neq 0$. Зрозуміло, що u можна подати як лінійну комбінацію векторів системи B_x . З (6.10) випливає, що всі вектори $B^k x$ є власними векторами оператора A з власним числом λ . Оскільки u , власний вектор B , лінійно виражається через вектори $B^k x$, то він є власним для оператора A і буде відповідати власному числу λ . Тим самим і доведено існування спільного власного вектора комутуючих операторів A і B .

Теорема 6.7. Нехай P_n - многочлен, матриця $A(n \times n)$, λ - власне число матриці і x - відповідний власний вектор. Тоді $P_n(\lambda)$ - власне число матриці $P_n(A)$.

Доведення. З умови $Ax = \lambda x$. Будемо діяти на обидві частини рівності матрицею A : $A(Ax) = A^2x = \lambda Ax = \lambda^2 x \Rightarrow A^k x = \lambda^k x, \dots, A^n x = \lambda^n x$.

Звідси

$$P_n(A)x = a_0 A^n x + a_1 A^{n-1} x + \dots + a_{n-k} A^k x + \dots + a_{n-1} Ax + a_n x =$$

$$P_n(A)x = (a_0 \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-k} \lambda^k + \dots + a_{n-1} \lambda + a_n I)x.$$

Отже, $P_n(A)x = P_n(\lambda)x$, що й треба було довести.

Наслідок 1. Нехай $\lambda_1, \dots, \lambda_n$ - власні числа матриці $A, Sp(A) = \lambda_1 + \dots + \lambda_n$. Тоді $Sp(A^m) = \lambda_1^m + \dots + \lambda_n^m$.

Для доведення треба попередній результат застосувати до функції $P_m(t) = t^m$.

Наслідок 2. Нехай Q - анулюючий многочлен квадратної матриці: $Q(A) = 0$ і λ - власне число матриці. Тоді $Q(\lambda) = 0$.

Доведення. Нехай $x \neq 0$ - власний вектор, що відповідає власному числу λ . За доведеним $Q(\lambda)$ - власне число оператора $Q(A)$:

$$Q(A)x = Q(\lambda)x = 0 \Rightarrow Q(\lambda) = 0.$$

Контрольні запитання

- 1). Що таке власне число і власний вектор лінійного оператора?
- 2). Що таке спектр лінійного оператора?
- 3). Що таке простий спектр лінійного оператора?
- 4). Що таке циклічний вектор?
- 5). Що таке анулюючий многочлен лінійного оператора? Навести приклад.

ЛЕКЦІЯ II.7. КЛІТИНА ЖОРДАНА¹⁰ ТА ЇЇ ВЛАСТИВОСТІ

План

- 1). Клітина Жордана. Нільпотентна клітина Жордана.
- 2). Кореневі підпростори.
- 3). Підпростори, інваріантні відносно нільпотентної клітини Жордана.

$$J_n(t) = \begin{pmatrix} t & 1 & 0 & \cdot & \dots & 0 \\ 0 & t & 1 & 0 & \dots & 0 \\ 0 & 0 & t & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & t \end{pmatrix}. \quad (7.1)$$

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

¹⁰ Жордан, Марі Енм́он Камі́ль (фр. *Marie Ennemond Camille Jordan*, 5/01/1838, Ліон – 22/01/1922, Париж) – видатний математик.

діагоналі, що вище головної на одну позицію, - одиниці; всі інші елементи – нулі. Знайдемо степені

$$J_n(0) = \begin{pmatrix} 0 & 1 & 0 & . & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ . & . & . & . & \dots & . \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Маємо:

$$J_n^2(0) = \begin{pmatrix} 0 & 1 & 0 & . & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ . & . & . & . & \dots & . \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & . & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ . & . & . & . & \dots & . \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix} = \quad (7.2)$$

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

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:

$$J_n^{n-1}(0) = \begin{pmatrix} 0 & 0 & 0 & \cdot & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

З цих міркувань також випливає, що $J_n^n(0) = J_n^{n+1}(0) = \dots = 0$.

Означення 7.1. 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 $f \in \mathbb{R}^n : J^k(0)f \neq 0 \ \forall k=0,1,\dots,n-1$. З попереднього відомо, що $J^k(0)f = 0$. Введемо позначення:

$$f_n = f, f_{n-1} = J(0)f_n, \dots, f_{n-k} = J(0)f_{n-k+1} = J^k(0)f_n, \dots, f_1 = J(0)f_2 = J^{n-1}(0)f_n.$$

Звідси

$$J^k(0)f_l = f_{l-k} \ \forall l \geq k \ (f_0 \triangleq 0).$$

Зауважимо, що $f_1 = J^{n-1}(0)f$ - власний вектор, що відповідає власному числу 0.

Легко бачити, що довільний підпростір

$$L_k = \{\alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_k f_k \ \forall \alpha_1, \dots, \alpha_k \in \mathbb{R}\} = \text{Ker}(J^k(0)) \ \forall k=1, \dots, n$$

інваріантний відносно $J(0)$ (а значить і його степенів).

Теорема 7.1. У нільпотентної клітини Жордана немає інших інваріантних підпросторів крім ядер її степенів.

Доведення. Нехай $\mathcal{L}$ - інваріантний підпростір. Знайдемо для нього кореневий підпростір $\text{Ker}(J^k(0))$ такий, що $\text{Ker}(J^k(0)) \subset \mathcal{L}$, але $\text{Ker}(J^{k+1}(0)) \not\subset \mathcal{L}$.

Візьмемо довільний вектор з цього підпростору і розкладемо його по жордановому базису:

$$f = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_j f_j.$$

Якщо $j \leq k$, тоді $f \in \text{Ker}(J^k(0))$. Розглянемо випадок $j > k$.

Запишемо цей розклад більш докладно:

$$f = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_k f_k + \alpha_{k+1} f_{k+1} + \dots + \alpha_j f_j$$

Подіємо на обидві частини цієї рівності оператором B^{j-k-1} . Слід зазначити, що

$$B^{j-k-1} f_{j-1} = f_{j-1-j+k+1} = f_k \in \text{Ker}(J^k(0)) \subset \mathcal{L}. \text{ Тим більше}$$

$$B^{j-k-1} f_i \in \text{Ker}(J^k(0)) \subset \mathcal{L} \quad \forall i < j.$$

Знайдемо тепер

$$B^{j-k-1} f_j = f_{k+1} \in \text{Ker}(J^{k+1}(0)).$$

З одного боку, з вибору вектора f випливає, що $f_{k+1} \in \mathcal{L}$, але це означає, що всі лінійні комбінації векторів $f_1, \dots, f_k, f_{k+1}$ належать $\mathcal{L}$, а це протирічить умові $\text{Ker}(J^{k+1}(0)) \not\subset \mathcal{L}$. Теорему доведено.

Наслідок. З доведеної теореми випливає, що простір, в якому діє клітина Жордана тієї ж розмірності, не може бути зображений як пряма сума підпросторів, інваріантних відносно цієї клітини. Дійсно, як ми бачили, при дії клітини утворюється низка вкладених підпросторів, а в прямій сумі підпростори мають нульовий перетин.

Нехай $e = \{e_1, e_2, \dots, e_n\}$,

$$e_1 = (1 \ 0 \ \dots \ 0)^T, e_2 = (0 \ 1 \ 0 \ \dots \ 0)^T, \dots, e_n = (0 \ 0 \ \dots \ 0 \ 1)^T -$$

канонічний базис $\mathbb{R}^n$. Подивимось, як на ці вектори діє клітина Жордана.

Маємо

$$J_n(t)e_1 = te_1, \quad J_n(t)e_2 = e_1 + te_2, \dots, J_n(t)e_n = e_{n-1} + te_n; \quad (7.2)$$

або

$$(J_n(t) - tI)e_1 = 0, \quad (J_n(t) - tI)e_2 = e_1, \dots, (J_n(t) - tI)e_n = e_{n-1} \quad (7.3)$$

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 (7.3), 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 $0 \leq k \leq n-1$, 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 $A(n \times m)$:

$$\begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1m} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-1,1} & a_{n-1,2} & \cdot & \cdot & \cdot & a_{n-1,m} \\ a_{n,1} & a_{n,2} & \cdot & \cdot & \cdot & a_{n,m} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2m} \\ a_{31} & a_{32} & \cdot & \cdot & \cdot & a_{3m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n,1} & a_{n,2} & \cdot & \cdot & \cdot & a_{n,m} \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{pmatrix}.$$

Як бачимо, вона зсуває рядки A на одну позицію вгору. Аналогічно «нижня» клітина Жордана зсуває рядки A на одну позицію униз:

$$\begin{pmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 0 & 0 \\ 0 & \cdot & \cdot & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1m} \\ a_{21} & a_{22} & \cdot & \cdot & \cdot & a_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-1,1} & a_{n-1,2} & \cdot & \cdot & \cdot & a_{n-1,m} \\ a_{n,1} & a_{n,2} & \cdot & \cdot & \cdot & a_{n,m} \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & \cdot & \cdot & \cdot & 0 \\ a_{11} & a_{12} & \cdot & \cdot & \cdot & a_{1m} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n-2,1} & a_{n-2,2} & \cdot & \cdot & \cdot & a_{n-2,m} \\ a_{n-1,1} & a_{n-1,2} & \cdot & \cdot & \cdot & a_{n-1,m} \end{pmatrix}.$$

Приклад 7.1. Нехай $I_n(\lambda)$ - простір функцій вигляду

$e^{\lambda x} P(x)$, де $\lambda \in \mathbb{C}$; $P(x)$ - довільний многочлен степені, яка не більша 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 вибрати базис - $e_j = \frac{x^{j-1}}{(j-1)!} e^{\lambda x}$, $j = 1, \dots, n$. Легко бачити,

що

$\frac{d}{dx}(e_{j+1}) = \frac{x^{j-1}}{(j-1)!} e^{\lambda x} + \lambda \frac{x^j}{j!} e^{\lambda x} = e_j + \lambda e_{j+1}, j = 1, \dots, n;$ $j = 1: \frac{d}{dx}(e_1) = \lambda e^{\lambda x} = \lambda e_1.$	(7.4)
---	-------

Таким чином, функції $e_j = \frac{x^{j-1}}{(j-1)!} e^{\lambda x}$, $j = 1, \dots, n$ утворюють базис, в якому

оператор диференціювання d/dx має матрицю вигляду (7.1), жорданову клітину,

і базис є жордановим для оператора. Цей приклад має велике значення в теорії систем лінійних диференціальних рівнянь.

Контрольні запитання

- 1). Яка матриця називається клітиною Жордана?
- 2). Яка клітина Жордана має властивість нільпотентності?
- 3). Які інваріантні підпростори має Клітина Жордана?

ЛЕКЦІЯ II.8. ЗВЕДЕННЯ МАТРИЦІ ДО НОРМАЛЬНОЇ ФОРМИ ЖОРДАНА

План

- 1). Ланцюг ядер, що відповідає одній клітині Жордана.
- 2). Кореневі підпростори.
- 3). Зображення лінійного простору у вигляді прямої суми корневих підпросторів.
- 4). Вибір жорданового базису і зведення матриці до нормальної форми Жордана.

При розгляді клітини Жордана ми вже помітили, що

$$\text{Ker}(tI - J_n(t)) = \{t_1 e_1, t_1 \in \mathbb{R}\}.$$

Неважко переконатись, що $(tI - J_n(t))^2 e_2 = 0$ і, звичайно, $(tI - J_n(t))^2 e_1 = 0$.

Отже, $\text{Ker}(tI - J_n(t))^2 = \{t_1 e_1, t_1 \in \mathbb{R}\} \oplus \{t_2 e_2, t_2 \in \mathbb{R}\}$ - пряма сума одновимірних підпросторів, натягнутих на базисні вектори e_1, e_2 .

Взагалі,

$\text{Ker}(tI - J_n(t))^k = \{t_1 e_1, t_1 \in \mathbb{R}\} \oplus \dots \oplus \{t_k e_k, t_k \in \mathbb{R}\}, 1 \leq k \leq n.$	(8.1)
---	-------

Розглянемо ситуацію уважніше. Нехай $f_n \in \text{Ker}(tI - J_n(t))^n$ і $f_n \notin \text{Ker}(tI - J_n(t))^m, m < n$. Це означає, що $(J_n(t) - tI)^n f_n = 0$. Звідси $(J_n(t) - tI)^n f_n = (J_n(t) - tI)^{n-1}((J_n(t) - tI) f_n) = 0$, отже, вектор

$$f_{n-1} = (J_n(t) - tI) f_n \in \text{Ker}(J_n(t) - tI)^{n-1}.$$

Взагалі

$$f_{n-k} = (J_n(t) - tI) f_{n-k+1} = (J_n(t) - tI)^k f_n \in \text{Ker}(J_n(t) - tI)^{n-k}$$

і, нарешті,

$$f_1 = (J_n(t) - tI) f_2 = (J_n(t) - tI)^{n-1} f_n \in \text{Ker}(J_n(t) - tI).$$

Таким чином, маємо ланцюг вкладених підпросторів:

$$E = \text{Ker}(tI - J_n(t))^n \supset \text{Ker}(tI - J_n(t))^{n-1} \supset \dots \supset \text{Ker}(tI - J_n(t)) \supset \{0\}. \quad (8.2)$$

Так побудовані вектори $f_1, f_2, \dots, f_n \in E$ утворюють жордановий базис простору.

Означення 8.1. Вектори $f_1, f_2, \dots, f_k \in E$ називаються лінійно незалежними відносно підпростору $H \subset E$, якщо ніяка їхня лінійна комбінація з коефіцієнтами, що не всі дорівнюють нулю, не належать до H . Максимальна 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 7.1 f_n - єдиний базисний вектор в E відносно $\text{Ker}(tI - J_n(t))^{n-1}$; f_n, f_{n-1} - базис відносно $\text{Ker}(tI - J_n(t))^{n-2}$ і т.ін.

Розглянемо тепер загальний оператор A з власним числами $\lambda_1, \lambda_2, \dots, \lambda_m$, що є коренями характеристичного многочлена $\chi(\lambda)$ відповідно кратностей $r_1, r_2, \dots, r_m, r_1 + r_2 + \dots + r_m = n$. Тоді

$$\chi(\lambda) = \prod_{i=1}^m (\lambda - \lambda_i)^{r_i}. \quad (8.3)$$

Означення 8.2. 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 7.1 $(J_n(t) - tI)^n e_n = 0$, тобто e_n – кореневий вектор висоти n , $(J_n(t) - tI)^{n-1} e_n \neq 0$, тому e_n – кореневий вектор висоти n ; 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 E , який будемо називати кореневим підпростором і позначати $L(\lambda)$.

Якщо f – кореневий вектор висоти h , то $f_1 = (A - \lambda I)^{h-1} f$ – власний вектор, відповідний власному числу λ :

$$(A - \lambda I) f_1 = (A - \lambda I) \left((A - \lambda I)^{h-1} f \right) = (A - \lambda I)^h f = 0.$$

Звідси випливає, що кореневі вектори можуть відповідати тільки кореням характеристичного многочлена оператора, бо вони і тільки вони є власними його числами.

Введемо позначення:

$\chi_j(\lambda) = \prod_{\substack{i=1 \\ i \neq j}}^n (\lambda - \lambda_i)^{r_i} = \chi(\lambda) (\lambda - \lambda_j)^{-r_j}.$	(8.4)
--	---------

Всі $\chi_j(\lambda) (j=1, \dots, m)$ взаємно прості.

Покладемо

$$L_j \triangleq \text{Im}(\chi_j(A)). \quad (8.5)$$

У такому разі $L_j \subset L(\lambda_j) = \text{Ker}((A - \lambda_j I)^{r_j})$, 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

$$((A - \lambda_j I)^{r_j})x = ((A - \lambda_j I)^{r_j})\chi_j(A)y = \chi_j(A)y = 0$$

в силу теореми Гамільтона-Келі.

Лема 8.1. $E = L_1 \oplus L_2 \oplus \dots \oplus L_m$.

Доведемо спочатку, що $E = L_1 + L_2 + \dots + L_m$. Многочлени $\chi_j(\lambda)$ ($j=1, \dots, m$) взаємно прості. Тому, як відомо (див. (1.8)), існують такі многочлени $u_j(t)$, 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 $l \in E$, 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 $l \in \sum_{j=1}^m L_j$.

Тепер доведемо, що $L_i \cap (\sum_{j \neq i} L_j) = \{0\}$. Нехай l – вектор з цього перетину.

Тоді $(A - \lambda_i I)^{r_i} l = 0$, оскільки $l \in L_i \subset L(\lambda_i)$. 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 $u(\lambda)$, $v(\lambda)$, що $u(\lambda)(\lambda - \lambda_i)^{r_i} + v(\lambda)\chi_i(\lambda) = 1$. Підставляючи сюди A замість λ і діючи таким виразом на l , матимемо

$$u(A)(A - \lambda_i E)^{r_i} l + v(A) \chi_i(A) l = u(A) 0 + v(A) 0 = l = 0.$$

Лема 8.2. $L(\lambda_j) = L_j$.

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 $(A - \lambda_j I)^{r_j} l'' = 0$ і, оскільки $l'' \in \bigoplus_{i \neq j} L_i$, то

$$\chi_j(A) l'' = \prod_{k \neq j}^n (A - \lambda_k I)^{r_k} \left(\sum_{i \neq j} \prod_{v \neq i} (A - \lambda_v I)^{r_v} \right) l'' = 0.$$

Многочлени $(\lambda - \lambda_i)^{r_i}$ і $\chi_i(\lambda)$ взаємно прості, тому існують такі многочлени $u(\lambda)$, $v(\lambda)$, що $u(\lambda)(\lambda - \lambda_i)^{r_i} + v(\lambda)\chi_i(\lambda) = 1$, звідки, підставляючи сюди A замість λ , одержимо $u(A)(A - \lambda_i E)^{r_i} l'' + v(A)\chi_i(A) l'' = l'' = 0$, тому $l = l' \in L_j$.

Зауваження. З лем 8.1 і 8.2 одержимо формулу

$$L = L_1 \oplus L_2 \oplus \dots \oplus L_m = L(\lambda_1) \oplus L(\lambda_2) \oplus \dots \oplus L(\lambda_m) \quad (8.6)$$

Цікаво з'ясувати, як діє оператор $(A - \lambda_j I)^{r_j}$ у просторі $\tilde{L}_j \equiv \bigoplus_{i \neq j} L_i = \bigoplus_{i \neq j} L(\lambda_i)$.

Лема 8.3. 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

$$L_i = (A - \lambda_1 I)^{r_1} \dots (A - \lambda_{i-1} I)^{r_{i-1}} (A - \lambda_{i+1} I)^{r_{i+1}} \dots (A - \lambda_m I)^{r_m} L.$$

Серед цих співмножників є, безумовно, $(A - \lambda_j I)^{r_j}$. Тому, якщо подіяти на L_i оператором $(A - \lambda_j I)^{r_j}$, то такий співмножник з'явиться двічі, і тоді

$$(A - \lambda_j I)^{r_j} L_i = (A - \lambda_1 I)^{r_1} \dots (A - \lambda_{i-1} I)^{r_{i-1}} (A - \lambda_{i+1} I)^{r_{i+1}} \dots (A - \lambda_m I)^{r_m} \left((A - \lambda_j I)^{r_j} L \right) \subset L_i.$$

Це й доводить, що $\tilde{L}_j \equiv \bigoplus_{i \neq j} L_i = \bigoplus_{i \neq j} L(\lambda_i)$ 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 (7.10) 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 (8.6).

Лема 8.4. Простір L можна розкласти в пряму суму (взагалі кажучи, неортогональну) таких інваріантних відносно $(A - \lambda I)^r$ підпросторів:

$$L = L(\lambda) \oplus \text{Im}((A - \lambda I)^r), \quad (8.7)$$

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 $h \leq r$, 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 (5.15) одержимо:

$$\dim(L) = \dim(L(\lambda)) + \dim(\operatorname{Im}((A - \lambda I)^r)),$$

тобто підпростори $L(\lambda)$ і $\operatorname{Im}((A - \lambda I)^r)$ породжують весь простір L .

Для доведення того, що $(A - \lambda I)^r$ має обернений у $\operatorname{Im}((A - \lambda I)^r)$, треба помітити, що ядро звуження оператора, діюче в цьому підпросторі, нульове. Дійсно, інакше б існував вектор $x \in \operatorname{Im}((A - \lambda I)^r)$, для якого $(A - \lambda I)^r x = 0$, це свідчило б, що $x \in L(\lambda)$. Щойно було доведено, що в такому випадку $x = 0$. Цим лему повністю доведено.

Вправа. Довести останню рівність в формулі (8.6), тобто $L = L(\lambda_1) \oplus L(\lambda_2) \oplus \dots \oplus L(\lambda_m)$ спираючись на результати леми 4.

Вказівка. Одержавши розклад (8.6) для $\lambda = \lambda_1$, треба в $\operatorname{Im}((A - \lambda_1 I)^{r_1})$ відщепити підпростір $L(\lambda_2)$ і т.ін.

Підсумуємо отримані результати.

Якщо оператор A має власні числа $\lambda_1, \dots, \lambda_m$, які є коренями характеристичного многочлена $\chi(\lambda) = \det(\lambda I - A)$, відповідно кратностей $r_1, r_2, \dots, r_m$ ($r_1 + r_2 + \dots + r_m = n$), то простір L , в якому діє A , можна розкласти в пряму суму (8.6) підпросторів $L(\lambda_j)$, інваріантних відносно кожного з операторів $(A - \lambda_j I)^{r_j}$ ($j = 1, \dots, m$). 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 $x_j \in L(\lambda_j)$ обов'язково задовольняє умову $(A - \lambda_j I)^{r_j} x_j = 0$. Але можливий випадок, коли існує число $k_j \leq r_j$, 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 ($h \leq r$), 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:

$$L(\lambda) = \text{Ker}(A - \lambda I)^h \supset \text{Ker}(A - \lambda I)^{h-1} \supset \dots \supset \text{Ker}(A - \lambda I) \supset \{0\} \quad (8.8)$$

Введемо позначення: $B = A - \lambda E$. Тоді цей ланцюг переписеться у вигляді

$$L(\lambda) = \text{Ker}(B)^h \supset \text{Ker}(B)^{h-1} \supset \dots \supset \text{Ker}(B) \supset \{0\}$$

Виберемо в максимальному з цих підпросторів, в $\text{Ker}(B^h)$, базис відносно $\text{Ker}(B^{h-1})$: $e_1^h, \dots, e_{q_h}^h$. Тоді вектори $Be_1^h, \dots, Be_{q_h}^h \in \text{Ker}(B^{h-1})$. Вони лінійно незалежні відносно $\text{Ker}(B^{h-2})$. Дійсно, нехай

$$\alpha_1 Be_1^h + \dots + \alpha_{q_h} Be_{q_h}^h = B(\alpha_1 e_1^h + \dots + \alpha_{q_h} e_{q_h}^h) \in \text{Ker}(B^{h-2}),$$

а це протирічить лінійній незалежності $e_1^h, \dots, e_{q_h}^h$ відносно $\text{Ker}(B^{h-1})$.

Доповнимо одержані вектори до базису в $\text{Ker}(B^{h-1})$ відносно $\text{Ker}(B^{h-2})$:

$$Be_1^h, \dots, Be_{q_r}^h, e_1^{h-1}, \dots, e_{q_{r-1}}^{h-1}$$

Знову застосуємо до таких векторів оператор B і, доповнюючи їх образи до базису в $\text{Ker}(B^{h-2})$ відносно $\text{Ker}(B^{h-3})$, одержимо

$$B^2 e_1^h, \dots, B^2 e_{q_r}^h, Be_1^{h-1}, \dots, Be_{q_{r-1}}^{h-1}, e_1^{h-2}, \dots, e_{q_{r-2}}^{h-2}$$

Продовжуючи цей процес, дійдемо нарешті до $\text{Ker}(B)$. Таким чином, маємо таблицю:

Зауважимо, що хоч ми довільним чином вибрали вектори відносних базисів (див. (8.9)), але їхня кількість – це є інваріанти, незалежні від вибору базису, а тільки від розмірностей ядер. Отже, всі зображення (8.10) будуть однакові з точністю до розташування в матриці A_j клітин Жордана.

Сформулюємо результат.

Теорема 8.1. Для оператора, що діє в комплексному скінченновимірному просторі, існує жордановий базис, тобто матриця оператора в деякому базисі може бути приведена заміною базису до жорданової нормальної форми – прямої суми клітин Жордана і діагональних матриць. Матриця A_j вигляду (8.10) 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 (8.6). Оскільки

$$\dim(Ker(A - \lambda I)^i) = \dim L - \dim(Im(A - \lambda I)^i),$$

досить знайти

$$\dim(Im(A - \lambda I)^i) = rank(A - \lambda I)^i,$$

а ранг матриці $(A - \lambda I)^i$ можна знайти елементарними перетвореннями, що не змінюють рангу матриці, тобто додаванням до деякого вектора-рядка (стовпчика) лінійної комбінації інших векторів-рядків (стовпчиків).

Контрольні запитання

- 1). Що таке кореневий підпростір?
- 2). Як побудувати ланцюг корневих підпросторів?
- 3). Який перетин у корневих підпросторів, що відповідають різним кореням характеристичного многочлена?
- 4). Як побудувати жордановий базис і звести матрицю до нормального вигляду?

ЛЕКЦІЯ II.9. ЖОРДАНОВА НОРМАЛЬНА ФОРМА МАТРИЦЬ ТРЕТЬОГО ТА ЧЕТВЕРТОГО ПОРЯДКУ

План

- 1). Матриці третього порядку, що мають різні або кратні корені характеристичного рівняння. Аналіз усіх можливих ситуацій.
- 2). Матриці четвертого порядку, що мають різні або кратні корені характеристичного рівняння. Аналіз усіх можливих ситуацій.
- 3). Матриці п'ятого порядку, що мають різні або кратні корені характеристичного рівняння. Аналіз усіх можливих ситуацій.

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 9.1. Всі три корені $\lambda_1, \lambda_2, \lambda_3$ характеристичного многочлена різні. Нагадаємо, що такі корені називаються простими. Очевидно, що $\dim \text{Ker}(A - \lambda_1 I) = \dim \text{Ker}(A - \lambda_2 I) = \dim \text{Ker}(A - \lambda_3 I) = 1$, Простір розпадається в пряму суму власних підпросторів матриці:

$$\mathbb{R}^3 = \text{Ker}(A - \lambda_1 I) \oplus \text{Ker}(A - \lambda_2 I) \oplus \text{Ker}(A - \lambda_3 I)$$

і жорданова форма J_A матриці має вигляд

$$J_A = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Випадок 9.2. $\chi(A, t) = (t - \lambda_1)^2 (t - \lambda_2)$. В цьому випадку ядро оператора $A - \lambda_2 I$ може бути тільки одновимірне. Нехай $\dim \text{Ker}(A - \lambda_1 I) = 2$. Це означає, що жордановий базис можна вибрати з власних векторів оператора A ,

$$\mathbb{R}^3 = \text{Ker}(A - \lambda_1 I) \oplus \text{Ker}(A - \lambda_2 I)$$

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 $(t - \lambda_1)(t - \lambda_2)$ не співпадає з характеристичним і є його дільником.

Для знаходження жорданового базису досить розв'язати лінійні системи

$$(A - \lambda_1 I)x = 0 \text{ та } (A - \lambda_2 I)x = 0.$$

Цікаво зауважити, що вектори $g_1, g_2 \in \text{Ker}(A - \lambda_1 I), g_3 \in \text{Ker}(A - \lambda_2 I)$ жорданового базису можна знайти, не розв'язуючи системи, а саме: нехай x і y – будь-які лінійно незалежні вектори з $\mathbb{R}^3$. Тоді $(A - \lambda_2 I)x, (A - \lambda_2 I)y$ (підбираємо x і y так, щоб ці одержані вектори були ненульовими) – лінійно незалежні і належать $\text{Ker}(A - \lambda_1 I)$, як це вже пояснювалось. Повторимо це ще раз: $(A - \lambda_1 I)((A - \lambda_2 I)x) = ((A - \lambda_1 I)(A - \lambda_2 I))x = 0$, аналогічно з другим вектором. Так само виберемо вектор z з $\mathbb{R}^3$ і $(A - \lambda_1 I)z$ (якщо він ненульовий) належить $\text{Ker}(A - \lambda_2 I)$.

Випадок 9.3. $\chi(A, t) = (t - \lambda_1)^2 (t - \lambda_2)$. В цьому випадку знову ядро оператора $A - \lambda_2 I$ може бути тільки одновимірне. Нехай тепер $\dim \text{Ker}(A - \lambda_1 I) = 1$. Це означає, що оператор має одновимірний підпростір власних векторів і одновимірний відносно $\text{Ker}(A - \lambda_1 I)$ підпростір корневих векторів $\text{Ker}(A - \lambda_1 I)^2$; його розмірність дорівнює 2. Маємо ланцюг ядер:

$$\text{Ker}(A - \lambda_1 I)^2 \supset \text{Ker}(A - \lambda_1 I) \supset \{0\},$$

$$\mathbb{R}^3 = \text{Ker}(A - \lambda_1 I)^2 \oplus \text{Ker}(A - \lambda_2 I)$$

Виберемо жордановий базис, взявши вектор $x \in \mathbb{R}^3$ так, що $(A - \lambda_2 I)x \neq 0$. Тоді $(A - \lambda_2 I)x \in \text{Ker}(A - \lambda_1 I)^2$. Слід вибрати x таким чином, щоб $(A - \lambda_1 I)(A - \lambda_2 I)x \neq 0$. Покладемо $g_2 = (A - \lambda_2 I)x$. Тоді $g_1 = (A - \lambda_1 I)g_2 = (A - \lambda_1 I)(A - \lambda_2 I)x \in \text{Ker}(A - \lambda_1 I)$. Власний вектор g_3 з $\text{Ker}(A - \lambda_2 I)$ 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 $\text{Ker}(A - \lambda_2 I)$.

З виписаних рівностей маємо:

$$Ag_1 = \lambda_1 g_2$$

$$Ag_2 = g_1 + \lambda_1 g_2$$

$$Ag_3 = \lambda_2 g_3$$

i

$$J_A = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

Випадок 9.4. $\chi(A, t) = (t - \lambda)^3$. Нехай $\dim \text{Ker}(A - \lambda I) = 1$, тобто,

$\text{rank}(A - \lambda I) = 2$. Тоді маємо ланцюг (довгий) ядер:

$$\mathbb{R}^3 = \text{Ker}(A - \lambda I)^3 \supset \text{Ker}(A - \lambda I)^2 \supset \text{Ker}(A - \lambda I) \supset \{0\}$$

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 $\text{Ker}(A - \lambda I)^2$ і $g_1 = (A - \lambda I)g_2$ належить $\text{Ker}(A - \lambda I)$. 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 9.5. $\chi(A, t) = (t - \lambda)^3$. Нехай $\dim \text{Ker}(A - \lambda I) = 2$, тобто, $\text{rank}(A - \lambda I) = 1$. Тоді маємо ланцюг (короткий) ядер:

$$\mathbb{R}^3 = \text{Ker}(A - \lambda I)^2 \supset \text{Ker}(A - \lambda I) \supset \{0\}$$

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 $\text{Ker}(A - \lambda I)$. 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 9.6. Корені характеристичного рівняння прості. Матриця зводиться до діагонального вигляду у базисі, що складається з власних векторів матриці.

Випадок 9.7. $\chi(A, t) = (t - \lambda_1)^2 (t - \lambda_2) (t - \lambda_3)$

а). Якщо $\text{rank}(A - \lambda_1 I) = 3$, а значить $\dim \text{Ker}(A - \lambda_1 I) = 1$,

$\dim \text{Ker}(A - \lambda_1 I)^2 = 2$, жорданова нормальна форма матриці має вигляд:

$$J_A = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 \\ 0 & 0 & 0 & \lambda_3 \end{pmatrix}.$$

б). Якщо $\text{rank}(A - \lambda_1 I) = 2$, а значить $\dim \text{Ker}(A - \lambda_1 I) = 2$, 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 9.8. $\chi(A, t) = (t - \lambda_1)^2 (t - \lambda_2)^2$

a). $\dim \text{Ker}(A - \lambda_1 I) = \dim \text{Ker}(A - \lambda_1 I) = 2$. 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). $\dim \text{Ker}(A - \lambda_1 I) = 1$, $\dim \text{Ker}(A - \lambda_2 I) = 2$. Тоді автоматично $\dim \text{Ker}(A - \lambda_1 I)^2 = 2$.

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). $\dim \text{Ker}(A - \lambda_1 I) = \dim \text{Ker}(A - \lambda_2 I) = 1$ Тоді автоматично $\dim \text{Ker}(A - \lambda_1 I)^2 = \dim \text{Ker}(A - \lambda_2 I)^2 = 2$.

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 9.9. $\chi(A, t) = (t - \lambda)^4$

a). $\dim \text{Ker}(A - \lambda I) = 1$. Маємо довгий ланцюг ядер:

$$\mathbb{R}^3 = \text{Ker}(A - \lambda I)^4 \supset \text{Ker}(A - \lambda I)^3 \supset \text{Ker}(A - \lambda I)^2 \supset \text{Ker}(A - \lambda I) \supset \{0\}$$

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). $\dim \text{Ker}(A - \lambda I) = 2$. В такій ситуації ми не можемо нічого сказати про $\dim \text{Ker}(A - \lambda I)^2$, тому ядро $\text{Ker}(A - \lambda I)^2$ треба досліджувати окремо.

b1). Якщо $\dim \text{Ker}(A - \lambda I)^2 = 4$, то, знайшовши два лінійно незалежні вектори $f_1, f_3 \in \text{Ker}(A - \lambda I)$, побудуємо відносний (відносно $\text{Ker}(A - \lambda I)$) базис в $\text{Ker}(A - \lambda I) - \{f_2, f_4\}$, щоб виконувались співвідношення :

$$f_1 = (A - \lambda I) f_2, f_3 = (A - \lambda I) f_4 \Rightarrow Af_2 = f_1 + \lambda f_2, Af_4 = f_3 + \lambda f_4.$$

Завдяки також рівностям

$$Af_1 = \lambda f_1, Af_3 = \lambda f_3$$

записуємо нормальну форму Жордана даної матриці:

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

Вона є прямою сумою двох двовимірних клітин Жордана.

b2). Нехай $\dim \text{Ker}(A - \lambda I) = 2$, $\dim \text{Ker}(A - \lambda I)^2 = 3$. Тоді маємо ланцюг ядер

$$\mathbb{R}^3 = \text{Ker}(A - \lambda I)^3 \supset \text{Ker}(A - \lambda I)^2 \supset \text{Ker}(A - \lambda I) \supset \{0\}.$$

Жордановий базис в такій ситуації будувати досить просто: беремо довільний вектор $g_4 \in \mathbb{R}^4 = \text{Ker}(A - \lambda I)^3$ і знаходимо:

$$g_3 = (A - \lambda I)g_4 \in \text{Ker}(A - \lambda I)^2 \rightarrow g_2 = (A - \lambda I)g_3 \in \text{Ker}(A - \lambda I).$$

Безумовно, треба вибирати g_4 так, щоб всі ці вектори були ненульовими. Вектор $g_1 \in \text{Ker}(A - \lambda I)$ знайдемо як лінійно незалежний з $g_2 \in \text{Ker}(A - \lambda I)$ розв'язок системи $(A - \lambda I)x = 0$. 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 9.1. 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 $\dim(\text{Ker}(A - 2I)) = 1$, треба знайти лише $\text{rank}(A - I)$

$$A - E = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Отже, $\text{rank}(A - I) = \dim(\text{Im}(A - I)) = 2$ і $\dim(\text{Ker}(A - I)) = 3 - 2 = 1$. Після цього можна не знаходити матрицю $(A - I)^2$, тому що $\dim(\text{Ker}(A - I)^2)$ може дорівнювати тільки 2, інших варіантів бути не може. У результаті можна побудувати такий базис:

$e = e_1 \in \text{Ker}(A - I)^2$, $e_1 \notin \text{Ker}(A - I)$, $e_2 = (A - I)e_1 \in \text{Ker}(A - I)$,
 e_2 - власний вектор, тобто, $Ae_2 = e_2$; e_3 - власний вектор A , що відповідає власному числу $\lambda = 2$, $Ae_3 = 2e_3$. У такому базисі

$$\begin{aligned} Ae_1 &= e_1 + e_2, \\ Ae_2 &= e_2, \\ Ae_3 &= 2e_3 \end{aligned} \tag{9.1}$$

Базис Жордана будується так: $f_1 = e_2$, $f_2 = e_1$, $f_3 = e_3$. У ньому оператор A діє таким чином:

$$\begin{aligned} Af_1 &= f_1, \\ Af_2 &= f_1 + f_2, \\ Af_3 &= 2f_3 \end{aligned} \tag{9.2}$$

Згадаємо, як виписати матрицю, виходячи з такої системи лінійних рівностей (формули (5.1), (5.9)): 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 виписати 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 (9.1), а обмежуватись системами вигляду (9.2)

Тепер можна виписати матрицю Жордана:

$$A_j = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Приклад 9.2.

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

Після простих обчислень знайдемо характеристичний многочлен цієї матриці: $\chi(\lambda) = (\lambda - 1)^3$.

Обчислимо $\text{rank}(A - I) = \dim(\text{Im}(A - I))$:

$$A - E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & 0 \end{pmatrix}.$$

звідки $\dim(\text{Im}(A - I)) = 2$ і $\dim(\text{Ker}(A - I)) = 3 - 2 = 1$. Це означає, що розмірність $\text{Ker}(A - I)^2$ відносно $\text{Ker}(A - I)$ може дорівнювати тільки 1, тобто, $\dim \text{Ker}((A - I)^2) = 2$. В цьому можна впевнитись безпосереднім обчисленням:

$$(A - I)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

звідки $\dim(\text{Im}(A - I)^2) = 1$, $\dim(\text{Ker}(A - I)^2) = 3 - 1 = 2$. Тепер, зрозуміло, $\dim(\text{Ker}(A - I)^3) = 3$, 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 :

$$e_3 = (1, 0, 0)^T, \quad e_2 = (A - I)e_3, \quad e_1 = (A - I)e_2 = (A - I)^2 e_3;$$

e_3 - власний вектор. Відзначимо, що вектор e_3 був «вгаданий»: проста перевірка показує, що він не міститься в ядрах $\text{Ker}(A - I)$ і $\text{Ker}(A - I)^2$. 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 9.3. 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

$\text{rank}(A + I) = \dim(\text{Im}(A + I))$:

$$\begin{aligned} A + I &= \begin{pmatrix} 1 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -11 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 1 & 0 \\ -3 & 4 & -5 & 0 \\ 8 & -4 & 4 & 0 \\ 15 & -10 & 11 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 1 & 0 \\ -3 & 4 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 15 & -10 & 11 & 0 \end{pmatrix} \rightarrow \\ &\rightarrow \begin{pmatrix} 2 & 0 & 1 & 0 \\ -3 & -1 & -5 & 0 \\ 0 & 0 & 0 & 0 \\ 15 & 1 & 11 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & 1 & 0 \\ 7 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -7 & 1 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Звідси $\dim(\text{Im}(A+I))=2$, $\dim(\text{Ker}(A+I))=4-2=2$.

Знайдемо $\dim(\text{Im}(A+I)^2)$:

$$(A+I)^2 = \begin{pmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{pmatrix} =$$
$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & -1 & 1 \end{pmatrix},$$

Легко бачити, що $\dim(\text{Im}(A+I)^2)=1$ і $\dim(\text{Ker}(A+I)^2)=4-1=3$. Очевидно, $\dim(\text{Ker}(A+I)^3)=4$.

Маємо $e_4 = e \in \text{Ker}(A+I)^3$ - базис в $\text{Ker}(A+I)^3$ відносно

$\text{Ker}(A+I)^2$; $e_3 = (A+I)e_4$ базис в $\text{Ker}(A+I)^2$ відносно

$\text{Ker}(A+I)$; $e_2 = (A+I)e_3$, e_2, e_4 - власні вектори, базис в $\text{Ker}(A+I)$. В цьому базисі

$$Ae_1 = -e_1,$$

$$Ae_2 = -e_2,$$

$$Ae_3 = e_2 - e_3$$

$$Ae_4 = e_3 - e_4.$$

Звідси

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

Жорданова нормальна форма матриці має вигляд прямої суми одноклітинної (діагональної) матриці і трьохвимірної клітини Жордана.

Приклад 9.4. Нехай $G = J^2(0)$ - квадрат нільпотентної клітини Жордана:

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

Матриця має чотирьохкратний корінь $\lambda = 0$ характеристичного рівняння.

Очевидно, $\dim \text{Ker} G = 4 - \text{rank Im} G = 4 - 2 = 2$.

Легко бачити, що $G^2 = 0$, отже, відносна розмірність $\text{Ker} G^2$ відносно $\text{Ker} G$ дорівнює 2. Також очевидно, що

$$e_1 = (1; 0; 0; 0)^T, e_2 = (0; 1; 0; 0)^T \in \text{Ker} G,$$

$$Ge_3 = G(0; 0; 1; 0)^T = e_1 = (1; 0; 0; 0)^T, Ge_4 = G(0; 0; 0; 1)^T = e_2 = (0; 1; 0; 0)^T \in \text{Ker} G.$$

Тепер можна побудувати жордановий базис:

$$g_1 = e_1, g_2 = e_3, g_3 = e_2, g_4 = e_4.$$

Тоді маємо:

$$Gg_1 = 0g_1, Gg_2 = g_1, Gg_3 = 0g_3, Gg_4 = g_3.$$

Матриця Жордана буде мати вигляд:

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

Тобто, вона є прямою сумою клітин Жордана розмірностей 2.

Приклад 9.5. Нехай $G = J^2(0)$ - квадрат нільпотентної клітини Жордана в $\mathbb{R}^5$:

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

Звідси

$$G^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, G^3 = 0.$$

Матриця має п'ятикратний корінь $\lambda = 0$ характеристичного рівняння.

При цьому

$$\dim \text{Ker} G = 5 - \text{rank } \text{Im} G = 5 - 3 = 2,$$

$$\dim \text{Ker} G^2 = 5 - \text{rank } \text{Im} G^2 = 5 - 1 = 4.$$

Відносна розмірність $\text{Ker} G^2$ відносно $\text{Ker} G$ дорівнює 2. Подібно попередньому прикладу жордановий базис знаходиться усно:

$$g_1 = e_1, g_2 = e_3, g_3 = e_5, g_4 = e_2, g_5 = e_4.$$

В цьому базисі матриця діє таким чином:

$$Gg_1 = 0,$$

$$Gg_2 = g_1,$$

$$Gg_3 = g_2,$$

$$Gg_4 = 0,$$

$$Gg_5 = g_4.$$

Матриця Жордана буде мати вигляд:

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

Тобто, вона є прямою сумою клітин Жордана розмірностей 3 і 2.

Зауваження. Аналогічний результат має місце для квадрату нільпотентної клітини Жордана в $\mathbb{R}^n$. Результат залежить від парності розмірності простору.

Приклад 9.6. 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 $\text{rank}(A - I) = \dim(\text{Im}(A - I))$:

$$\begin{aligned} A - E &= \begin{pmatrix} 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 2 & 2 & -2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 3 & 0 \\ 0 & 0 & -2 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \\ &\rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

звідки $\dim(\text{Im}(A - I)) = 3$, $\dim(\text{Ker}(A - I)) = 5 - 3 = 2$. Знайдемо $\dim(\text{Ker}(A - I)^2)$

:

$$(A-I)^2 = \begin{pmatrix} 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 2 & 2 & -2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & -2 & 3 & -3 \\ 0 & 0 & 2 & 2 & -2 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -4 & 4 \\ 0 & 0 & 4 & -4 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

отже, $\dim(\text{Im}(A-I)^2) = 1$, $\dim(\text{Ker}(A-I))^2 = 5 - 1 = 4$.

Таким чином, маємо базис в $\text{Ker}(A-I)^2$ відносно $\text{Ker}(A-I)$: e_3, e_4 і базис в $\text{Ker}(A-I)$: $e_1 = (A-I)e_3$; $e_2 = (A-I)e_4$ 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

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

Контрольні запитання

- 1). Як визначити розмірність підпростору власних векторів?
- 2). Як побудувати ланцюг кореневих підпросторів, для матриці, що має єдиний трьохкратний корінь характеристичного рівняння?
- 3). Які можуть бути варіанти ланцюгів кореневих підпросторів для матриць третього порядку, що мають єдиний трьохкратний корінь характеристичного рівняння?
- 4). Які можуть бути варіанти ланцюгів кореневих підпросторів для матриць третього порядку, що мають єдиний трьохкратний корінь характеристичного рівняння?

ЛЕКЦІЯ II.10. ЖОРДАНОВА НОРМАЛЬНА ФОРМА МАТРИЦІ НОРМАЛЬНОГО ОПЕРАТОРА ПОЛЯРНИЙ РОЗКЛАД

План

- 1). Властивості власних чисел та власних векторів нормального оператора.
- 2). Властивості власних чисел та власних векторів унітарного оператора.
- 3). Додатні оператори.
- 4). Полярний розклад лінійного оператора.

Оператор, що діє в унітарному просторі називається нормальним, якщо він комутує зі своїм спряженим:

$$AA^* = A^*A \quad (10.1)$$

Лема 10.1. 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:

$$0 = ((A - \lambda I)x, (A - \lambda I)x) = (x, (A^* - \bar{\lambda} I)(A - \lambda I)x) = (x, (A - \lambda I)(A^* - \bar{\lambda} I)x) =$$

$$= \left((A^* - \bar{\lambda}I)x, (A^* - \bar{\lambda}I)x \right).$$

Звідси $(A^* - \bar{\lambda}I)x = 0$, що і треба було довести.

Лема 10.2. 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: $\lambda(x, y) = (Ax, y) = (x, A^*y) = (x, \bar{\mu}y) = \bar{\mu}(x, y)$. Звідси $(\lambda - \bar{\mu})(x, y) = 0$ і $(x, y) = 0$.

Лема 10.3. Нехай A – нормальний оператор, λ – його власне число. Тоді $\text{Ker}(A - \lambda I)^2 = \text{Ker}(A - \lambda I)$.

Доведення. Нехай x – кореневий вектор висоти 2, що відповідає власному числу λ : $(A - \lambda I)^2 x = 0$, тобто, $(A - \lambda I)((A - \lambda I)x) = 0$. Тоді $(A - \lambda I)x$ також є власним вектором A^* , що відповідає власному числу $\bar{\lambda}$: $(A^* - \bar{\lambda}I)(A - \lambda I)x = 0$.

Отже, $0 = ((A^* - \bar{\lambda}I)(A - \lambda I)x, x) = ((A - \lambda I)x, (A - \lambda I)x)$, звідки $(A - \lambda I)x = 0$.

Теорема 10.1. 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 10.2. 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:

$$E_n = \text{Ker}(A - \lambda_1 I) \oplus \dots \oplus \text{Ker}(A - \lambda_m I) \quad (10.2)$$

$$\dim \text{Ker}(A - \lambda_k I) = r_k, k = 1, \dots, m \quad (10.3)$$

Згідно з формулою (10.2), довільний вектор $x \in E_n$ можна подати (однозначно) у вигляді

$$x = x_1 + \dots + x_m, x_k \in \text{Ker}(A - \lambda_k I).$$

Тоді

$$Ax = Ax_1 + \dots + Ax_m = \lambda_1 x_1 + \dots + \lambda_m x_m \quad (10.4)$$

Нехай P_k - ортопроектор на власний підпростір $\text{Ker}(A - \lambda_k I)$. Згідно з формулою (10.4) оператор A можна подати у вигляді

$$A = \lambda_1 P_1 \oplus \dots \oplus \lambda_m P_m \quad (10.5)$$

Теорема 10.2. Нехай в ортогональному базисі оператор має діагональну матрицю. Тоді оператор нормальний.

Зауважимо, що ця теорема обернена до теореми 10.1.

Доведення. На діагоналі матриці спряженого оператора в тому ж базисі знаходяться числа, комплексно спряжені до власних чисел даного оператора. Діагональні матриці комутують, і це доводить теорему.

Теорема 10.3. Нехай A – матриця нормального оператора і J_A - її жорданова нормальна форма, тобто, діагональна матриця. Тоді існує унітарна матриця U оператора заміни базису, що

$J_A = U^{-1}AU.$	(10.6)
-------------------	--------

Доведення. Справедливість твердження випливає з загальної форми (5.23) 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.

Підведемо підсумки. Нормальний оператор і його спряжений (теж нормальний) мають спільну ортонормовану систему власних векторів $e = \{e_1, e_2, \dots, e_n\}$, що становить базис простору. Дія обох операторів цілком визначається системою рівностей:

$Ae_k = \lambda_k e_k, A^*e_k = \bar{\lambda}_k e_k, k = 1, \dots, n.$	(10.7)
--	--------

Зупинимось на випадку, в якому нормальний оператор має простий спектр і розглянемо два інтерполяційні многочлени Лагранжа $P(\lambda)$ і $Q(\lambda)$, які задаються умовами:

$$P(\lambda_k) = \bar{\lambda}_k, Q(\bar{\lambda}_k) = \lambda_k, k = 1, 2, \dots, n.$$

Тоді, завдяки (10.7) $P(A) = A^*, Q(A^*) = A$.

Отже, у випадку нормального оператора A кожен з двох операторів A і A^* є многочленом від іншого. При цьому обидва ці многочлени визначаються власними числами оператора A .

Зауваження. Випадок кратних коренів оператора A потребує більш розвиненої техніки з застосуванням диференціального числення, і в даному курсі ми його не розглядаємо.

Зробимо висновки: важливими прикладами нормальних операторів є самоспряжені і унітарні оператори. Отже, їхні матриці можна привести до діагональному вигляду. Уточнимо спочатку вигляд формули (10.5) для самоспряжених операторів.

Лема 10.4. Всі власні числа самоспряженого оператора дійсні.

Доведення. Нехай $Ax = \lambda x$, $x \neq 0$. Помножимо скалярно обидві частини цієї рівності на x : $\lambda(x, x) = (Ax, x) = (x, Ax) = (x, \lambda x) = \bar{\lambda}(x, x)$, звідки і випливає $\lambda = \bar{\lambda}$.

Тепер дослідимо власні числа унітарного оператора.

Лема 10.5. 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 10.1. 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 10.2. Лінійний додатний самоспряжений оператор B називається квадратним коренем з додатного самоспряженого оператора A , якщо $A = B^2$.

Лема 10.6. 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 $\lambda_1, \dots, \lambda_n$.

Нехай оператор B має ті самі власні вектори з власними числами $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ 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 10.4. Довільний лінійний оператор $A \in L(E_n, E_n)$, що діє в унітарному (евклідовому) просторі можна зобразити у вигляді добутку унітарного (ортогонального) і додатного самоспряженого оператора.

Доведення. Нехай спочатку A – невироджений оператор. Оператор AA^* додатний і самоспряжений. Нехай B – квадратний корінь з $AA^* : AA^* = B^2$. Помножимо обидві частини рівності зліва і справа на $B^{-1} : B^{-1}AA^*B^{-1} = I$. Тоді $A^*B^{-1} = (B^{-1}A)^{-1}$. Помітимо, що $(B^{-1}A)^* = A^*B^{-1}$ (оскільки B - самоспряжений), звідки $A^*B^{-1} = (B^{-1}A)^{-1} = (B^{-1}A)^*$, а це означає, що оператор $U \triangleq B^{-1}A$ - унітарний. Отже,

$$A = BU, \quad (10.8)$$

що й треба було довести.

Зображення довільного лінійного оператора у вигляді добутку додатного і унітарного (в дійсному просторі - ортогональний) операторів називається полярним розкладом цього оператора.

Зауваження. Можна застосувати доведений факт до $A^* : A^* = B_1V$, де B_1 - додатний, а V - ортогональний оператори. Взявши спряження від обох частин рівності, одержимо: $(A^*)^* = A = (B_1V)^* = V^{-1}B_1$ (при цьому скористались самоспряженістю B_1 і унітарністю V). Оператор $\tilde{U} = V^{-1}$ також унітарний, як і V . Отже, одержали ще один полярний розклад:

$$A = V^{-1}B = \tilde{U}B. \quad (10.9)$$

Але для практичних застосувань дамо ще й пряме доведення цього факту.

Оператор A^*A додатний, самоспряжений і невироджений. Нехай B_1 - квадратний корінь з нього: $A^*A = B_1^2$. Помножимо зліва і справа обидві частини цієї рівності на $B_1^{-1} : B_1^{-1}A^*AB_1^{-1} = I$. Звідси випливає: $B_1^{-1}A^* = (AB_1^{-1})^{-1}$. Знайдемо спряжений до оператора в дужках у правій частині цієї рівності: $(AB_1^{-1})^* = B_1^{-1}A^*$ (оскільки B - самоспряжений), звідки $B_1^{-1}A^* = (AB_1^{-1})^* = (AB_1^{-1})^{-1}$, а це означає, що $\tilde{U} = AB_1^{-1}$ - унітарний. Звідси $A = \tilde{U}B$.

Зауваження. З доведення випливає, що в кожному з випадків унітарний оператор при умові невиродженості оператора A знаходиться однозначно.

Розглянемо тепер випадок, в якому оператор A може бути вироджений. Покажемо, що повна ортогональна система власних векторів оператора A^*A переводиться в знову ортогональну систему.

Нехай

$$A^*Ax_k = \lambda_k^2 x_k, \lambda_k \geq 0, (x_k, x_j) = \delta_{kj}, k, j = 1, \dots, n.$$

Тоді

$$(Ax_k, Ax_j) = (A^*Ax_k, x_j) = \lambda_k^2 (x_k, x_j) = 0, k \neq j.$$

Якщо $j = k$, маємо

$$\|Ax_k\|^2 = (Ax_k, Ax_k) = \lambda_k^2 \Rightarrow \|Ax_k\| = \lambda_k (\lambda_k \geq 0).$$

Це означає, що існує ортонормована система z_k така, що

$Ax_k = \lambda_k z_k, (z_j, z_k) = \delta_{jk}.$	(10.10)
---	---------

Визначимо лінійні оператори рівностями

$Ux_k = z_k, Bz_k = \lambda_k z_k.$	(10.11)
-------------------------------------	---------

З (10.10) і (10.11) випливає, що $A = BU$.

Зауважимо, що B - невід'ємний симетричний оператор, оскільки завдяки (10.11) він має невід'ємні числа і діагональну форму з цих чисел у ортонормованому базисі його власних векторів. Оператор U унітарний, оскільки він переводить ортонормований базис $\{x_1, x_2, \dots, x_n\}$ у ортонормований базис $\{z_1, z_2, \dots, z_n\}$.

Отже, одержано полярний розклад оператора типу (10.8). Аналогічним чином доводиться полярний розклад (10.9). При цьому квадратний корінь B з

оператора A^*A по ньому визначається однозначно, а унітарний оператор U визначається однозначно лише у випадку невідродженого оператора A .

Наслідок. З рівності $A = BU$ одержуємо:

$$\begin{aligned} AA^* &= BU(BU)^* = BUU^*B^* = B^2, \\ A^*A &= (BU)^*BU = U^*B^2U = U^{-1}B^2U. \end{aligned}$$

Якщо $AA^* = A^*A$, тобто, оператор нормальний, то звідси випливає:

$$U^{-1}B^2U = B^2 \Rightarrow B^2U = UB^2.$$

Звідси неважко довести (використовуючи інтерполяційний многочлен Лагранжа, що не розглядається в даному курсі), що $BU = UB$. Навпаки, якщо рівність $BU = UB$ має місце, то $A^*A = U^{-1}B^2U = U^{-1}BUB = U^{-1}UB^2 = B^2$. Але $B^2 = AA^*$, звідки $AA^* = A^*A$, що доводить нормальність оператора A .

Таким чином доведено **твердження**: оператор A нормальний тоді і тільки тоді, коли множники U і B в (10.8) або V і B в (10.9) комутують.

Зауваження. Зображення лінійного оператора (10.8) і (10.9) нагадують експоненційне зображення комплексного числа.

Доведемо зараз твердження, яке має аналогом зображення комплексного числа у алгебраїчній формі.

Теорема 10.5. Довільний лінійний оператор можна зобразити у наступному вигляді:

$A = \frac{1}{2}(A_1 + iA_2),$	(10.12)
--------------------------------	---------

де A_1, A_2 - ермітові оператори (ермітові компоненти оператора A), які визначаються оператором A однозначно. Оператор A є нормальним тоді і тільки тоді, коли його ермітові компоненти комутують.

Доведення. Формули $A_1 = \frac{1}{2}(A + A^*)$, $A_2 = \frac{1}{2i}(A - A^*)$ визначають два ермітових оператори. Справді,

$$(A_1)^* = \frac{1}{2}(A + A^*)^* = \frac{1}{2}(A^* + A) = A_1,$$

$$(A_2)^* = \left(\frac{1}{2i}(A - A^*) \right)^* = -\frac{1}{2i}(A^* - A) = \frac{1}{2i}(A - A^*) = A_2.$$

Звідси з очевидністю випливає: $A = \frac{1}{2}(A_1 + iA_2)$.

Нехай тепер має місце зображення (10.12). Тоді $A^* = \frac{1}{2}(A_1 - iA_2)$. З цих двох зображень маємо: $A_1 = \frac{1}{2}(A + A^*)$, $A_2 = \frac{1}{2i}(A - A^*)$ і ермітовість одержаних двох операторів вже було доведено.

Нехай тепер A - нормальний: $AA^* = A^*A$. Тоді

$$AA^* = \frac{1}{4i}(A_1 + iA_2)(A_1 - iA_2) = \frac{1}{4i}(A_1^2 + A_2^2 + i(A_2A_1 - A_1A_2)),$$

$$A^*A = \frac{1}{4i}(A_1 - iA_2)(A_1 + iA_2) = \frac{1}{4i}(A_1^2 + A_2^2 + i(A_1A_2 - A_2A_1)).$$

Очевидно, оператор A буде нормальним тоді і тільки тоді, коли буде мати місце рівність:

$$A_2A_1 - A_1A_2 = A_1A_2 - A_2A_1 \Leftrightarrow A_2A_1 = A_1A_2, \text{ що й треба було довести.}$$

Контрольні запитання

- 1). Які властивості мають власні числа нормального оператора?
- 2). Які властивості мають власні числа унітарного оператора?

ЛЕКЦІЯ II.11. ОПЕРАТОРИ В ДІЙСНИХ ПРОСТОРАХ

План

1). Випадок двох комплексно спряжених коренів характеристичного рівняння.

Матриця Жордана в цьому випадку.

2). Вибір дійсного базису у випадку комплексно спряжених коренів характеристичного рівняння.

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 розмірність відносно ядра меншого степеня

$$\dim(Ker(A - \lambda E)^r) = \dim(Ker(A - \bar{\lambda} E)^r) = 1.$$

У такому випадку маємо ланцюг ядер:

$$R^r = Ker(A - \lambda E)^r \supset Ker(A - \lambda E)^{r-1} \supset \dots \supset Ker(A - \lambda E)^2 \supset Ker(A - \lambda E) \supset \{0\}$$

і такий самий для $\bar{\lambda}$.

Жорданова нормальна форма даної матриці має вигляд:

$$\begin{pmatrix} \lambda & 1 & \dots & 0 & 0 & \dots & & & & & \\ 0 & \lambda & \dots & 0 & 0 & \dots & & & & & \\ \dots & \dots & \ddots & \dots & \dots & \dots & & & & & \\ 0 & 0 & \dots & \lambda & 1 & \dots & & & & & \\ 0 & 0 & \dots & 0 & \lambda & \dots & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & & & \dots & \bar{\lambda} & 1 & \dots & 0 & 0 \\ & & & & & & \dots & 0 & \bar{\lambda} & \dots & 0 & 0 \\ & & & & & 0 & \dots & \dots & \dots & \ddots & \dots & \dots \\ & & & & & & \dots & 0 & 0 & \dots & \bar{\lambda} & 1 \\ & & & & & & \dots & 0 & 0 & \dots & 0 & \bar{\lambda} \end{pmatrix} \quad (11.9)$$

Покажемо, як зіставити даній матриці аналог нормальної форми – матрицю з

дійсними елементами. Нехай $f_r \in Ker(A - \lambda E)^r$, Тобто, $Ker(A - \lambda E)^r f_r = 0$,

$$(A - \lambda E) f_r = f_{r-1} \in \text{Ker}(A - \lambda E)^{r-1}, \dots, (A - \lambda E) f_2 = f_1 \in \text{Ker}(A - \lambda E)^1,$$

$$(A - \lambda E) f_1 = 0.$$

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

$$(A - (\alpha + i\beta)E)(g_r + ih_r) = g_{r-1} + ih_{r-1},$$

$$(A - (\alpha + i\beta)E)(g_{r-1} + ih_{r-1}) = g_{r-2} + ih_{r-2},$$

...

$$(A - (\alpha + i\beta)E)(g_2 + ih_2) = g_1 + ih_1,$$

$$(A - (\alpha + i\beta)E)(g_1 + ih_1) = 0$$

Звідси одержимо:

$$Ag_1 = \alpha g_1 - \beta h_1$$

$$Ah_1 = \beta g_1 + \alpha h_1$$

$$Ag_2 = g_1 + \alpha g_2 - \beta h_2$$

$$Ah_2 = h_1 + \beta g_2 + \alpha h_2$$

...

(11.10)

$$Ag_r = g_{r-1} + \alpha g_r - \beta h_r$$

$$Ah_r = h_{r-1} + \beta g_r + \alpha h_r$$

В результаті знаходимо матрицю в новому базисі $\{g_k, h_k\}, k = 1, \dots, r$:

$$[A]_{(g,h),(g,h)} = \begin{pmatrix} \alpha & \beta & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -\beta & \alpha & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \alpha & \beta & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & -\beta & \alpha & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & \alpha & \beta & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & -\beta & \alpha & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \alpha & \beta \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & -\beta & \alpha \end{pmatrix} \quad (11.11)$$

Ця матриця схожа на жорданову клітину. На її діагоналі знаходяться блоки

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \text{ а над діагоналлю – блоки з одиничних матриць } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

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 $\rho = \sqrt{\alpha^2 + \beta^2}$. Існує таке $\varphi \in [0, 2\pi)$, що $\cos \varphi = \frac{\alpha}{\rho} = \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}}$,

$\sin \varphi = \frac{\beta}{\rho} = \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}$. 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

$$\bar{J}_0 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \end{pmatrix} \quad (11.12)$$

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

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (11.14)$$

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

$g_1 = Te_1 = \frac{1}{2}(e_1 + ie_2), g_2 = Te_2 = \frac{1}{2i}(e_1 - ie_2)$. Для простоти позначимо матрицю

цього оператора також через T . Маємо:

$$T = \begin{pmatrix} \frac{1}{2} & \frac{1}{2i} \\ \frac{i}{2} & -\frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix}$$

і

$$T^{-1}CT = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} \alpha + i\beta & 0 \\ 0 & \alpha - i\beta \end{pmatrix} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = K_1. \quad (11.15)$$

$$\text{Тоді } e^{K_1} = T^{-1}e^CT = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} e^{a+ib} & 0 \\ 0 & e^{a-ib} \end{pmatrix} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} =$$

$$= \frac{e^a}{2} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} \begin{pmatrix} \cos b + i \sin b & 0 \\ 0 & \cos b - i \sin b \end{pmatrix} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = e^a \begin{pmatrix} \cos b & \sin b \\ -\sin b & \cos b \end{pmatrix}.$$

Отже, одержано експоненту матриці K_1 :

$$\exp(K_1) = \exp \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = e^{\alpha} \begin{pmatrix} \cos b & \sin b \\ -\sin b & \cos b \end{pmatrix} \quad (11.16)$$

В теорії диференціальних рівнянь потрібна така експонента:

$$\exp(tK_1) = \exp t \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = e^{\alpha t} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix} \quad (11.17)$$

Очевидно, $\exp(K)$ - блочна матриця, блоками якої є матриці $\exp(K_1)$.

Знаходження $\exp(t\bar{J}_0)$ проводиться за такою самою схемою, як і при знаходженні експоненти нільпотентної клітини Жордана:

$$\exp(t\bar{J}_0) = \begin{pmatrix} 1 & 0 & t/1! & 0 & t^2/2! & 0 & \dots & t^{r-1}/(r-1)! & 0 \\ 0 & 1 & 0 & t/1! & 0 & t^2/2! & \dots & 0 & t^{r-1}/(r-1)! \\ 0 & 0 & 1 & 0 & t/1! & 0 & \dots & t^{r-2}/(r-2)! & 0 \\ 0 & 0 & 0 & 1 & 0 & t/1! & \dots & 0 & t^{r-2}/(r-2)! \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & t/1! & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & t/1! \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (11.18)$$

Остаточно

$$\exp(t\bar{J}_\lambda) = e^{at} \begin{pmatrix} \cos bt & \sin bt & . & . & . \\ -\sin bt & \cos bt & . & . & . \\ . & . & . & . & . \\ . & . & . & \cos bt & \sin bt \\ . & . & . & -\sin bt & \cos bt \end{pmatrix} \times$$

$$\times \begin{pmatrix} 1 & 0 & t/1! & 0 & t^2/2! & 0 & \dots & t^{r-1}/(r-1)! & 0 \\ 0 & 1 & 0 & t/1! & 0 & t^2/2! & \dots & 0 & t^{r-1}/(r-1)! \\ 0 & 0 & 1 & 0 & t/1! & 0 & \dots & t^{r-2}/(r-2)! & 0 \\ 0 & 0 & 0 & 1 & 0 & t/1! & \dots & 0 & t^{r-2}/(r-2)! \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & t/1! & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & t/1! \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{pmatrix} \quad (11.19)$$

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 $A(f_1 + if_2)_2 = (\alpha + i\beta)(f_1 + if_2)$. 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 $(A - (4 + 3i)E)x = 0$, $(A - (4 - 3i)E)x = 0$, 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 :

$$[T]_{f,g}^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, [T]_{g,f} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Використовуючи формулу (5.23), знайдемо матрицю оператора у базисі f_1, f_2 :

$$[A]_{f,f} = ([T]_{g,f})^{-1} [A]_{g,g} [T]_{g,f}$$

$$[A]_{f,f} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} 4+3i & 0 \\ 0 & 4-3i \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix},$$

тобто, знову одержали матрицю у вигляді (11.15).

Контрольні запитання

- 1). Який вигляд має клітина Жордана у випадку двох комплексно спряжених коренів характеристичного рівняння?
- 2). Як побудувати дійсний базис у випадку комплексно спряжених коренів характеристичного рівняння?
- 3). Який вигляд має матриця лінійного оператора у дійсному жордановому базисі?

ЛЕКЦІЯ II.12. ФУНКЦІЇ ВІД МАТРИЦЬ

План

- 1). Аналітичні функції від лінійних операторів, зокрема, експоненційна.
- 2). Аналітична функція від клітини Жордана.
- 3). Експонента від клітини Жордана.
- 4). Степенева функція від клітини Жордана.
- 5). Застосування нормальної форми для визначення 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

$$f(A) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A^k \quad (12.1)$$

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:

$$\left\| \sum_{k=n+1}^{n+p} \frac{f^{(k)}(0)}{k!} A^k \right\| \leq \sum_{k=n+1}^{n+p} \frac{|f^{(k)}(0)|}{k!} \|A\|^k$$

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 ряду.

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, $A = T A_J T^{-1}$, $A^2 = T A_J T^{-1} T A_J T^{-1} = T A_J^2 T^{-1}$, звідки за індукцією $A^k = T A_J^k T^{-1}$. Тому

$$f(A) = T \left(\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A_J^k \right) T^{-1}.$$

Експонента матриці. За ознакою Вейєрштрасса ряд

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2!} A^2 + \dots + \frac{1}{n!} A^n + \dots \quad (12.2)$$

збігається абсолютно і рівномірно на кожній множині $\{A: \|A\| \leq a, a > 0\}$ алгебри матриць, тому що

$$\|\exp(A)\| = \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k \leq e^a < \infty.$$

Зазначимо, що з (12.1) випливає

$$\exp(0) = I. \quad (12.3)$$

Доведемо, що для $t \in \mathbb{R}$

$$\frac{d}{dt} \exp(tA) = A \exp(tA). \quad (12.4)$$

Продиференціювавши ряд (12.2) по t формально, одержимо ряд з похідних:

$$\sum_{k=0}^{\infty} \frac{1}{k!} \frac{d}{dt} (t^k) A^k = \sum_{k=0}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k = A \sum_{i=0}^{\infty} \frac{t^i}{i!} A^i = A \exp(tA).$$

Цей ряд збігається рівномірно на кожній обмеженій області своїх аргументів, тобто при $\|A\| \leq a, |t| \leq T$. Тому за теорією про диференціювання ряду похідна суми існує і дорівнює сумі ряду з похідних.

Рівність (12.4) означає, що матрична функція

$$X(t) = \exp(tA) \quad (12.5)$$

є розв'язком матричного диференціального рівняння

$$\frac{d}{dt} X(t) = AX(t). \quad (12.6)$$

Зауважимо, що для матриці

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (12.7)$$

$$A^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \lambda_n^k \end{pmatrix} \quad (12.8)$$

Приклад 12.1. Функція від діагональної матриці (12.7):

$$f(A) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} A^k = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} \quad (12.9)$$

Зокрема, якщо $f(z) = \exp(z)$, то

$$\exp(tA) = \begin{pmatrix} \exp(t\lambda_1) & 0 & \cdots & 0 \\ 0 & \exp(t\lambda_2) & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \exp(t\lambda_n) \end{pmatrix} \quad (12.10)$$

Приклад 12.2. 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 (12.1) — просто скінченна сума.

Найбільш цікавий приклад: $A = J(0)$, де $J(0)$ - клітина Жордана з нулями на діагоналі. Тоді

$$f(J(0)) = \sum_{k=0}^{m-1} \frac{f^{(k)}(0)}{k!} (J(0))^k. \quad (12.11)$$

Якщо згадати формулу (7.1) і подальші дослідження щодо вигляду степенів нільпотентної клітини Жордана, то можна одержати такий вираз для функції від клітини Жордана:

$$f(J(0)) = \begin{pmatrix} 1 & f(0) & \frac{1}{2!} f'(0) & \dots & \frac{1}{(m-1)!} f^{m-1}(0) \\ 0 & 1 & f(0) & \dots & \frac{1}{(m-2)!} f^{m-2}(0) \\ 0 & 0 & 1 & \dots & \frac{1}{(m-3)!} f^{m-3}(0) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & f(0) \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (12.12)$$

Розглянемо тепер функцію

$$f(\lambda) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!} (\lambda - \alpha)^k. \quad (12.13)$$

Нашою метою є обчислення її значення від матричного аргументу розмірності $n \times n$:

$$J(\alpha) = \begin{pmatrix} \alpha & 1 & 0 & \dots & 0 & 0 \\ 0 & \alpha & 1 & \dots & 0 & 0 \\ 0 & 0 & \alpha & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha & 1 \\ 0 & 0 & 0 & \dots & 0 & \alpha \end{pmatrix} = \alpha I + J(0) \quad (12.14)$$

Підставляючи замість λ матрицю $J(\alpha)$, одержимо в правій частині степені нільпотентної матриці $J(0)$, тому ряд буде містити скінченну кількість доданків:

$$f(J(\lambda)) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!} J^k(0) = \quad (12.15)$$

$$= f(\alpha) + \frac{f'(\alpha)}{1!} J(0) + \frac{f''(\alpha)}{2!} J^2(0) + \dots + \frac{f^{(n-1)}(\alpha)}{(n-1)!} J^{n-1}(0).$$

Запишемо результат у вигляді матриці:

$$f(J(\alpha)) = \begin{pmatrix} 1 & f(\alpha) & \frac{1}{2!} f'(\alpha) & \dots & \frac{1}{(m-1)!} f^{m-1}(\alpha) \\ 0 & 1 & f(\alpha) & \dots & \frac{1}{(m-2)!} f^{m-2}(\alpha) \\ 0 & 0 & 1 & \dots & \frac{1}{(m-3)!} f^{m-3}(\alpha) \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & f(\alpha) \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (12.16)$$

Важливою функцією є експонента від оператора та його матриці, оскільки за її допомогою можна розв'язувати системи лінійних диференціальних рівнянь.

$$\exp(zJ(0)) = \begin{pmatrix} 1 & z & z^2 / 2! & \dots & z^{m-1} / (m-1)! \\ 0 & 1 & z & \dots & z^{m-2} / (m-2)! \\ 0 & 0 & 1 & \dots & z^{m-3} / (m-3)! \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & z \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (12.17)$$

Теорема 12.1. Нехай лінійні оператори A і B комутують: $AB=BA$. Тоді

$$\exp(A+B) = \exp(A)\exp(B) = \exp(B)\exp(A). \quad (12.18)$$

Доведення. Порівняємо ряди

$$\begin{aligned}\exp(A)\exp(B) &= \left(I + A + \frac{1}{2!}A^2 + \dots \right) \left(I + B + \frac{1}{2!}B^2 + \dots \right) = \\ &= I + (A + B) + \frac{1}{2!}(A^2 + 2AB + B^2) + \dots\end{aligned}$$

і

$$\exp(A + B) = I + (A + B) + \frac{1}{2!}(A^2 + AB + BA + B^2) + \dots$$

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 (12.18).

Наслідок:

$\exp(t + s)A = \exp(tA)\exp(sA),$	(12.19)
------------------------------------	---------

де $t, s \in \mathbb{C}$. Дійсно, оператори tA і sA комутують, і доведення випливає з (12.18). У свою чергу, з (12.18) можна одержати важливий наслідок:

$$(\exp(tA))^{-1} = \exp(-tA). \quad (12.20)$$

Дійсно, з (12.19) і (12.2) одержимо $\exp(t - t)A = \exp(tA)\exp(-tA) = I$, що й доводить (12.20).

Властивості (12.19) - (12.20) означають, що множина операторів $\{\exp(tA), t \in \mathbb{C}\}$ утворює мультиплікативну абелеву групу, яка широко застосовується в теорії диференціальних рівнянь та диференціальній геометрії, якщо $t \in \mathbb{R}$.

Теорему 12.1 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,

$$\exp(tJ(\lambda)) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \dots & \frac{t^{m-1}}{(m-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \dots & \frac{t^{m-2}}{(m-2)!}e^{\lambda t} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & te^{\lambda t} \\ 0 & 0 & 0 & \dots & e^{\lambda t} \end{pmatrix} \quad (12.21)$$

Цікаво помітити, що $\det(\exp(J(\lambda))) = \exp(m\lambda)$. Звідси для експоненти від довільної матриці, жордановою нормальною формою якої є пряма сума жорданових клітин, маємо формулу:

$\det(\exp(A)) = \det\left(T \exp\left(\sum_{k=1}^n \lambda_k\right) T^{-1}\right) = \exp(\text{Sp}A).$	(12.22)
---	---------

Зауважимо, що верхній рядок (12.21) містить базисні вектори простору квазімногочленів степені, меншої за m , розглянутого раніше, формула (12.21) дає розв'язок диференціального рівняння (12.6) при $A = J(\lambda)$.

Степенева функція

Розглянемо тепер степеневу функцію від оператора $J(\lambda) = \lambda I + J(0)$:

$(J(\lambda))^a = (\lambda I + J(0))^a$. Обчислимо похідні у нулі функції $f(z) = (\lambda + z)^a$:

$$f'(0) = a\lambda^{a-1}, f''(0) = a(a-1)\lambda^{a-2}, \dots, f^{(k)}(0) = a(a-1)\dots(a-k+1)\lambda^{a-k}.$$

Тоді

$$(\lambda I + J(0))^a = \begin{pmatrix} \lambda^a & a\lambda^{a-1} & \frac{a(a-1)\lambda^{a-2}}{2!} & \dots & \frac{a(a-1)\dots(a-m+2)\lambda^{a-m+1}}{(m-1)!} \\ 0 & \lambda^a & a\lambda^{a-1} & \dots & \frac{a(a-1)\dots(a-m+3)\lambda^{a-m+2}}{(m-2)!} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & a\lambda^{a-1} \\ 0 & 0 & 0 & \dots & \lambda^a \end{pmatrix} \quad (12.23)$$

Приклад 12.3. Нехай задано числа $a_1, \dots, a_k$ і $x_1, \dots, x_k$, а при $n > k$ члени послідовності $\{x_i\}$ знаходяться за формулами

$$x_n = a_1 x_{n-1} + \dots + a_k x_{n-k}. \quad (12.24)$$

Така послідовність називається зворотною, а співвідношення (12.24) - рекурентним. Перевірка показує, що відрізок послідовності $z_{n-1} = \{x_{n-k}, \dots, x_{n-1}\}$ довжини k переводиться у відрізок $z_n = \{x_{n-k+1}, \dots, x_n\}$ тієї ж довжини за допомогою матриці, яка відповідає (12.24):

$$Az_{n-1} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \cdot & \dots & 1 \\ a_n & a_2 & \cdot & \dots & a_1 \end{pmatrix} \begin{pmatrix} x_{n-k} \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} x_{n-k+1} \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = z_n.$$

Оскільки за допомогою математичної індукції одержимо, що $z_n = A^n z_k$, де $z_k = \{x_1, \dots, x_k\}$, 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, користуючись виглядом степеневі функції для клітини Жордана, одержати вигляд шуканої формули загального члена:

$$x_n = \sum_{i=1}^k \lambda_i^n p_i(n), \quad (12.25)$$

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 A знаходяться як корені характеристичного многочлена, що збігається з мінімальним.

Приклад 12.4. Розглянемо послідовність, в якій $x_0 = x_1 = 1$; $x_n = x_{n-2} + x_{n-1}$ для $n > 1$. Вона називається послідовністю Фібоначчі.

Матриця A , 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: $\lambda_1 = \frac{1-\sqrt{5}}{2}$, $\lambda_2 = \frac{1+\sqrt{5}}{2}$. 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:

$$\alpha + \beta = 1, \quad \alpha \lambda_1 + \beta \lambda_2 = 1.$$

Розв'язуючи систему відносно α і β , знаходимо

$$\alpha = \frac{\sqrt{5}-1}{2\sqrt{5}}, \quad \beta = \frac{\sqrt{5}+1}{2\sqrt{5}},$$

Звідки

$$x_n = \frac{1}{\sqrt{5}} \left(\left(\frac{\sqrt{5}+1}{2} \right)^n - \left(\frac{\sqrt{5}-1}{2} \right)^n \right).$$

Резольвента оператора. Нехай $\lambda \in \mathbb{C}$ не є власним числом оператора A . Тоді існує оператор

$$R(A, \lambda) = (A - \lambda I)^{-1}, \quad (12.26)$$

який називається резольвентою A . Резольвента - це дробово-раціональна функція від оператора. За аналогією з формулою суми членів нескінченно спадної геометричної прогресії ($|q| < 1$)

$$S_{\infty} = 1 + q + q^2 + \dots + q^n + \dots = \frac{1}{1 - q}$$

будемо шукати резольвенту у вигляді

$$R(A, \lambda) = (A - \lambda I)^{-1} = -\lambda^{-1}(I - \lambda^{-1}A) = -\lambda^{-1}(I + \lambda^{-1}A + \dots + \lambda^{-n}A^n + \dots). \quad (12.27)$$

Ряд абсолютно збігається, якщо $|\lambda|^{-1} \|A\| < 1$, тобто, при

$$|\lambda| > \|A\|. \quad (12.28)$$

Доведемо, що ця формула насправді зображує обернений оператор до $A - \lambda I$. Дійсно, після перемноження і приведення подібних членів маємо

$$\begin{aligned} & -\lambda^{-1}(A - \lambda I)(I + \lambda^{-1}A + \dots + \lambda^{-n}A^n + \dots) = \\ & = -\lambda^{-1}(A - \lambda I + \lambda^{-1}A^2 - A + \lambda^{-1}A^2 - \lambda^{-1}A^2 + \dots) = \lambda^{-1}\lambda I = I \end{aligned}$$

Зауваження. Радіус збіжності $\rho(A)$ ряду (12.27) можна знайти більш точно, користуючись відомою формулою Адамара, що приводить до такого вигляду:

$\rho(A) = \overline{\lim}_{n \rightarrow \infty} \sqrt[n]{\|A^n\|}$ (верхня границя). Але в даному випадку можна довести більш точну формулу:

$$\rho(A) = \lim_{n \rightarrow \infty} \sqrt[n]{\|A^n\|}. \quad (12.29)$$

Тобто, існує границя послідовності, яка, зрозуміло, збігається з верхньою границею. Зауважимо, що (12.29) має місце також у нескінченновимірному випадку і у функціональному аналізі називається формулою Гельфанда.

Доведення. Позначимо $r = \inf_{n \in \mathbb{N}} \|A^n\|^{\frac{1}{n}}$. Очевидно, $r \leq \overline{\lim}_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}}$. Доведемо, що $r = \overline{\lim}_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}}$. Для кожного $\varepsilon > 0$ виберемо таке натуральне m , що $\|A^m\|^{\frac{1}{m}} < r + \varepsilon$. Тоді $\|A^m\| < (r + \varepsilon)^m$. Далі, для довільного натурального $n \geq m$ нехай q - залишок від ділення n на m . Тоді $n = pm + q$, $0 \leq q \leq m-1$ (p - ціле). Використовуючи нерівність для норм операторів: $\|XY\| \leq \|X\| \|Y\|$, одержимо:

$$\|A^n\|^{\frac{1}{n}} = \|A^{mp+q}\|^{\frac{1}{n}} = \|A^{mp} A^q\|^{\frac{1}{n}} \leq \|A^{mp}\|^{\frac{1}{n}} \|A^q\|^{\frac{1}{n}} \leq \|A^m\|^{\frac{p}{n}} \|A\|^{\frac{q}{n}} < (r + \varepsilon)^{\frac{mp}{n}} \|A\|^{\frac{q}{n}}.$$

Оскільки $\lim_{n \rightarrow \infty} \frac{pm}{n} = \lim_{n \rightarrow \infty} \frac{n-q}{n} = 1$ то $\overline{\lim}_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}} \leq r + \varepsilon$. Число $\varepsilon > 0$ вибирається довільним чином, тому $\overline{\lim}_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}} \leq r$, а це і доводить існування границі, і тим самим, - формулу (5.4.31), де $\rho(A) = r$.

Доведемо, що $\rho(A) \leq \|A\|$. Справді,

$$\rho(A) = \lim_{n \rightarrow \infty} \sqrt[n]{\|A^n\|} \leq \lim_{n \rightarrow \infty} \sqrt[n]{\|A\|^n} = \|A\|.$$

Вправа. Для довільних регулярних λ, μ довести рівність

$$R(A, \lambda) - R(A, \mu) = (\lambda - \mu)R(A, \lambda)R(A, \mu). \quad (12.30)$$

Доведення. Мають місце очевидні рівності:

$$\begin{aligned} R(A, \lambda) &= R(A, \lambda)(A - \mu I)R(A, \mu), \\ R(A, \mu) &= R(A, \lambda)(A - \lambda I)R(A, \mu). \end{aligned}$$

Після їх віднімання одержимо

$$\begin{aligned} R(A, \lambda) - R(A, \mu) &= R(A, \lambda)R(A, \mu)(A - \mu I - A + \lambda I) = \\ &= (\lambda - \mu)R(A, \lambda)R(A, \mu). \end{aligned}$$

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 , $z = x + iy \in \mathbb{C}$, $y = \text{Im}(z) \neq 0$, $f, g, h \in E$. Нехай

$$\begin{aligned} (A - \bar{z}I)h &= f \\ (A - zI)h &= g \end{aligned}$$

Недійсне число z (так само $\bar{z}$) не може бути власним числом самоспряженого оператора, а значить існує обернений оператор $(A - \bar{z}I)^{-1}$, визначений на всьому просторі. Тоді $h = (A - \bar{z}I)^{-1}f$ і $g = (A - zI)(A - \bar{z}I)^{-1}f$. Розглянемо оператор

$$U(z) = (A - zI)(A - \bar{z}I)^{-1}$$

називається перетворенням Келі самоспряженого оператора A .

З попередньої рівності $g = U(z)f = (A - xI)h - iyh$.

Доведемо унітарність $U(z)$.

$$\begin{aligned}\text{Маємо: } \|f\|^2 &= \|(A - \bar{z}I)h\|^2 = ((A - xI)h + iyh, (A - xI)h + iyh) = \\ &= \|(A - xI)h\|^2 + \|yh\|^2 - iy((A - xI)h, h) + yi(h, (A - xI)h) = \\ &= \|(A - xI)h\|^2 + \|yh\|^2 - iy((A - xI)h, h) + yi((A - xI)h, h) = \|(A - xI)h\|^2 + \|yh\|^2\end{aligned}$$

Так само доводиться, що $\|g\|^2 = \|(A - xI)h\|^2 + \|yh\|^2$, звідки

$$\|g\|^2 = \|f\|^2 = \|U(z)f\|^2$$

Тепер унітарність випливає з леми 5.1.

Контрольні запитання

- 1). Як можна визначити аналітичні функції від лінійних операторів?
- 2). Який вигляд має аналітична функція від клітини Жордана?
- 3). Як побудувати аналітичну функцію від матриці, застосовуючи її нормальну форму?

ЛЕКЦІЯ II.13. ЦИКЛІЧНА НОРМАЛЬНА ФОРМА МАТРИЦІ

План

- 1). Циклічний вектор оператора та породжений їм базис.
- 2). Матриця оператора у циклічному базисі.
- 3). Анулюючий многочлен матриці.

Розглянемо деякий лінійний оператор A , що діє в лінійному n -вимірному просторі L над полем $\mathbb{R}$. Отже, від поля не вимагається алгебраїчної замкненості.

Означення 13.1. Вектор $l \in L$ називається циклічним, якщо вектори $l, Al, \dots, A^{n-1}l$ утворюють базис простору L .

Означення 13.2. Простір називається циклічним відносно оператора, якщо він має циклічний вектор.

Покладемо $e_i = A^{n-i}l, i = 1, \dots, n, e \triangleq \{e_1, \dots, e_n\}$. Очевидно,

$$Ae_i = AA^{n-i}l = A^{n-i+1}l, i = 2, \dots, n.$$

Для того, щоб знайти $Ae_1 = A^n l$, зауважимо, що вектори $e_1 = A^{n-1}l, e_2 = A^{n-2}l, \dots, e_n = l$ і $A^n l = Ae_1$ лінійно залежні. Отже, існують такі числа a_i , серед яких не всі дорівнюють нулю, що

$$A^n l = \sum_{i=0}^{n-1} a_i A^i l. \quad (13.1)$$

Подамо у вигляді таблиці результати дії A на базисні вектори:

$Ae_1 = a_{n-1}e_1 + a_{n-2}e_2 + \dots + a_1e_{n-1} + a_0e_n,$ $Ae_2 = e_1,$ $Ae_3 = e_2,$ $\dots\dots\dots$ $Ae_n = e_{n-1}.$	(13.2)
---	--------

Запишемо матрицю $[A]_{e,e}$ оператора A у побудованому циклічному базисі (нормальна форма Фробеніуса¹¹):

¹¹ Фробеніус Фердінанд Георг. 26/10/1849 – 3/08/1917. Видатний німецький математик.

$[A]_{e,e} = \begin{pmatrix} a_{n-1} & 1 & 0 & \dots & 0 \\ a_{n-2} & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ a_1 & 0 & 0 & \dots & 1 \\ a_0 & 0 & 0 & \dots & 0 \end{pmatrix}$	(13.3)
---	--------

З формули (11.1) випливає, що $A^n l - \sum_{i=0}^{n-1} a_i A^i l = 0$. Доведемо, що многочлен

$$M(t) = t^n - \sum_{i=0}^{n-1} a_i t^i$$

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. Отже, $M(A)l = 0$. З іншого боку, якщо $N(t)$ — анулюючий оператор A многочлен степені, меншої за n , то, подівавши оператором $N(A)$ на вектор l , одержимо, що між базисними векторами $l, Al, A^2 l, \dots, A^{n-1} l$ є лінійна залежність. Отже, $M(t)$ — мінімальний многочлен оператора A , тому вигляд клітини (13.3) не залежить від вибору циклічного вектора 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 Додатку II, Частина 1, приклад 14 характеристичний многочлен матриці (13.3) було знайдено за допомогою розкладу визначника $\det([A]_{e,e} - \lambda I)$ 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 $(p_i(t))^{f_i}$, де $p_i(t)$ - дільники характеристичного многочлена, які не мають коренів в $\mathbb{R}$. Циклічною формою матриці буде пряма сума матриць вигляду (11.3). Таке зображення єдине, якщо мінімальні многочлени всіх клітин не мають дійсних коренів.

Контрольні запитання

- 1). Який вектор називається циклічним?
- 2). Чи можна привести матрицю до циклічної нормальної форми (Фробеніуса) над алгебраїчно незамкненим полем?
- 3) Який вигляд має циклічна нормальна форма матриці (форма Фробеніуса)?

ЛЕКЦІЯ II.14. БІЛІНІЙНІ ТА КВАДРАТИЧНІ ФОРМИ ЗАГАЛЬНІ ВІДОМОСТІ ТА ЗАКОН ІНЕРЦІЇ КВАДРАТИЧНИХ ФОРМ

План

- 1). Поняття білінійної і квадратичної форми.
- 2). Зображення білінійної форми у унітарному просторі.
- 3). Сігнатура матриці.
- 4). Зведення квадратичної форми до канонічного вигляду методом Лагранжа.
- 5). Зведення квадратичної форми до канонічного вигляду методом Якобі.
- 6). Критерій Сільвестра.

Означення 14.1. Відображення $\Omega: E_n \times E_n \rightarrow \mathbb{C}$ називається білінійною формою, якщо $\forall x, y, x_1, x_2, y_1, y_2 \in E_n, \alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ воно має властивості:

$\begin{aligned}\Omega(\alpha_1 x_1 + \alpha_2 x_2, y) &= \alpha_1 \Omega(x_1, y) + \alpha_2 \Omega(x_2, y), \\ \Omega(x, \beta_1 y_1 + \beta_2 y_2) &= \bar{\beta}_1 \Omega(x, y_1) + \bar{\beta}_2 \Omega(x, y_2).\end{aligned}$	(14.1)
--	--------

Якщо E_n - унітарний простір, можна записати білінійну форму у зручному вигляді.

Теорема 14.1. Білінійній формі в унітарному просторі відповідає лінійний оператор $A: E_n \rightarrow E_n$ такий, що

$\Omega(x, y) = (Ax, y).$	(14.2)
---------------------------	--------

Доведення. Зафіксуємо елемент $x \in E_n$. Вираз $\overline{\Omega(x, y)}$ буде лінійним функціоналом по другому аргументу. Тоді за теоремою 14.1 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 14.1 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

$$\begin{aligned} (A(\alpha_1 x_1 + \alpha_2 x_2), y) - (\alpha_1 (Ax_1, y) + \alpha_2 (Ax_2, y)) = \\ = (A(\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 Ax_1 + \alpha_2 Ax_2), y) = 0 \end{aligned}$$

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 $e = \{e_1, \dots, e_n\}$ - базис в E_n (не обов'язково унітарному). Покладемо $\Omega(e_j, e_k) = a_{kj}$, $j, k = 1, \dots, n$. Ми зустрічались з подібною ситуацією в унітарному просторі при знаходженні елементів матриці – формула (5.2), яка в даному випадку має вигляд $(Ae_j, e_k) = a_{kj}$, $j, k = 1, \dots, n$. Остаточно:

$$\Omega(e_j, e_k) = (Ae_j, e_k) = a_{kj}, \quad j, k = 1, \dots, n. \quad (14.3)$$

Нехай $x = \sum_{j=1}^n x_j e_j$, $y = \sum_{k=1}^n y_k e_k$. Тоді

$$\Omega\left(\sum_{j=1}^n x_j e_j, \sum_{k=1}^n y_k e_k\right) = \sum_{j=1}^n x_j \sum_{k=1}^n y_k \Omega(e_j, e_k) = \sum_{k=1}^n \left(\sum_{j=1}^n a_{kj} x_j\right) y_k.$$

Останній вираз можна інтерпретувати як добуток:

$$y^T A x = (y_1 \quad \dots \quad y_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \dots & \cdot \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (Ax, y) = \Omega(x, y). \quad (14.4)$$

При цьому матриця $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \dots & \cdot \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$ називається матрицею білінійної форми (14.4).

Означення 14.2. Рангом білінійної форми називається ранг її матриці.

Означення 14.3. 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 14.4. $\Omega(x, x)$ називається квадратичною формою.

Перепишемо в разі квадратичної форми формулу (14.4) і здійснимо множення, враховуючи, що $a_{kj} = a_{jk}$:

$\Omega(x, x) = x^T A x = (x_1 \quad \dots \quad x_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (Ax, x) =$ $= \left(\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \sum_{k,j=1}^n a_{kj} x_j x_k =$ $= a_{11}x_1^2 + a_{22}x_2^2 + \dots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \dots + 2a_{n-1,n}x_{n-1}x_n.$	(14.5)
---	--------

Як і в теорії самоспряжених операторів, в теорії квадратичних форм постає задача зведення її до канонічного вигляду:

$\Omega(x, x) = a_{11}x_1^2 + \dots + a_{nn}x_n^2.$	(14.6)
---	--------

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

$$\Omega(x, x) = a_{11}x_1^2 + \dots + a_{rr}x_r^2. \quad (14.7)$$

Пояснимо це більш докладно. Нехай $[U]_{eg}$ - ортогональна (унітарна у комплексному випадку) матриця оператора U переходу від канонічного базису $e = \{e_1, e_2, \dots, e_n\}$ до базису власних векторів симетричного (ермітового) оператора $A: g = \{g_1, g_2, \dots, g_n\}, Ue_k = g_k, k = 1, \dots, n, Ag_k = \lambda_k g_k, k = 1, \dots, n \Rightarrow [A]_{ee} [g_k]_e = \lambda_k [g_k]_e$. Тоді, як відомо, матриця переходу переводить вектор-стовпчик з $\mathbb{R}^n$, що відповідає розкладу довільного вектора $[y]_g$ у новому базисі з власних векторів у вектор-стовпчик $[x]_e$ розкладу того самого вектора у старому, канонічному базисі: $[x]_e = [U]_{eg} [y]_g$. Використаємо це при перетворенні квадратичної форми:

$$\begin{aligned} \Omega(x, x) &= ([A]_{ee} [x]_e, [x]_e) = ([A]_{ee} [U]_{eg} [y]_g, [U]_{eg} [y]_g) = \\ &= \left(([U]_{eg})^* [A]_{ee} [U]_{eg} [y]_g, [y]_g \right) = \left(([U]_{eg})^{-1} [A]_{ee} [U]_{eg} [y]_g, [y]_g \right) = ([A]_{gg} [y]_g, [y]_g). \end{aligned}$$

Тут був врахований той факт, що спряжений до унітарного оператора – це його обернений, а також формула переходу до матриці оператора у іншому базисі. Матриця $[A]_{gg}$ діагональна з власними числами на головній діагоналі. Таким чином одержано зведення матриці до канонічного вигляду, що в розгорнутому вигляді відповідає (14.7). Випишемо окремо одержані формули, оскільки вони дуже важливі в теорії і на практиці:

$[x]_e = [U]_{eg} [y]_g, [y]_g = ([U]_{eg})^{-1} [x]_e,$ $[A]_{gg} = ([U]_{eg})^{-1} [A]_{ee} [U]_{eg}.$	(14.8)
--	--------

Якщо в (14.7) деякий коефіцієнт $a_{jj} > 0$, введемо заміну $\xi_j = \sqrt{a_{jj}} x_j$. Тоді

$a_{jj} (x_j)^2 = \xi_j^2$; якщо $a_{kk} < 0$ - заміну $\eta_k = \sqrt{|a_{kk}|} x_k$. Тоді

$$a_{kk} (x_k)^2 = -|a_{kk}| (x_k)^2 = -\left(\sqrt{|a_{kk}|} x_k\right)^2 = -\eta_k^2.$$

Остаточно, квадратична форма в новому базисі буде мати форму

$$\Omega(x, x) = \xi_1^2 + \dots + \xi_l^2 - \eta_{l+1}^2 - \dots - \eta_r^2. \quad (14.9)$$

яка називається нормальною, $r \leq n$ - ранг квадратичної форми.

Означення 14.5. Число додатних і від'ємних членів в формулі (12.9) називається відповідно додатним і від'ємним індексами форми. Різниця між додатним індексом і від'ємним називається сигнатурою форми; вона позначається σ .

Теорема 14.2 (закон інерції квадратичної форми). Додатний і від'ємний індекси квадратичної форми не залежать від вибору базису, в якому вона має нормальний вигляд, тобто, вони є інваріанти форми.

Доведення. Нехай квадратична форма може бути зведена до нормальних форм

$$\Omega(x, x) = y_1^2 + \dots + y_l^2 - y_{l+1}^2 - \dots - y_r^2. \quad (14.10)$$

і

$$\Omega(x, x) = z_1^2 + \dots + z_m^2 - z_{m+1}^2 - \dots - z_r^2. \quad (14.11)$$

Припустимо, що $l > m$. Звичайно, змінні $z_i, i=1, \dots, n$ можуть бути лінійно виражені через y_j :

$$z_i = \sum_{j=1}^n t_{ij} y_j = \sum_{j=1}^l t_{ij} y_j + \sum_{j=l+1}^n t_{ij} y_j, \quad (14.12)$$

З невідродженою матрицею коефіцієнтів.

Підставимо вирази (14.12) в (14.11). Маємо тотожність:

$\begin{aligned} y_1^2 + \dots + y_l^2 - y_{l+1}^2 - \dots - y_r^2 = \\ = z_1^2 + \dots + z_m^2 - z_{m+1}^2 - \dots - z_r^2, \end{aligned}$	(14.13)
---	-----------

Тобто,

$y_1^2 + \dots + y_l^2 - y_{l+1}^2 - \dots - y_r^2 =$ $= \left(\sum_{j=1}^n t_{1j} y_j \right)^2 + \dots + \left(\sum_{j=1}^n t_{mj} y_j \right)^2 - z_{m+1}^2 - \dots - z_r^2.$	(14.14)
--	---------

В сумах (14.14) при $j=1, \dots, m$ візьмемо перші l доданків і прирівняємо їх до нуля. Будемо мати однорідну систему m рівнянь з l невідомими, $l > m$:

$$\sum_{j=1}^l t_{ij} y_j = 0, i = 1, \dots, m. \quad (14.15)$$

Оскільки $l > m$, вона має нетривіальний розв'язок $y_1, \dots, y_l \Rightarrow y_1^2 + \dots + y_l^2 > 0$. При цьому (14.14) запишеться у вигляді

$y_1^2 + \dots + y_l^2 - y_{l+1}^2 - \dots - y_r^2 =$ $= \left(\sum_{j=l+1}^n t_{1j} y_j \right)^2 + \dots + \left(\sum_{j=l+1}^n t_{mj} y_j \right)^2 - z_{m+1}^2 - \dots - z_r^2.$	(14.16)
--	---------

Покладемо додатково $y_{l+1} = \dots = y_r = y_{r+1} = \dots = y_n = 0$. В результаті з (14.16) одержимо:

$$y_1^2 + \dots + y_l^2 = -z_{m+1}^2 - \dots - z_r^2,$$

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.

ЗВЕДЕННЯ КВАДРАТИЧНОЇ ФОРМИ ДО КАНОНІЧНОГО ВИГЛЯДУ МЕТОДОМ ЛАГРАНЖА

Нехай в деякому базисі задано квадратичну форму, яку запишемо таким чином:

$$\Omega(x, x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n + F(x_2, \dots, x_n), \quad (14.17)$$

причому, форма $F(x_2, \dots, x_n)$ не містить x_1 .

Введемо нові змінні за формулами

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n, \\ y_i &= x_i \quad \forall i = 2, \dots, n. \end{aligned} \quad (14.18)$$

Обчислимо

$$\frac{1}{a_{11}}y_1^2 = \frac{1}{a_{11}}(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n)^2 = a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n + G(x_2, \dots, x_n),$$

Де $G(x_2, \dots, x_n)$ - квадратична форма, що не містить x_1 .

Отже, $\Omega(x, x) =$

$$= a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n + F(x_2, \dots, x_n) = \frac{1}{a_{11}}y_1^2 - G(x_2, \dots, x_n) + F(x_2, \dots, x_n),$$

і з квадратичною формою $\Phi(x_2, \dots, x_n) = F(x_2, \dots, x_n) - G(x_2, \dots, x_n)$, 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

$$\Omega(x, x) = 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n. \quad (14.19)$$

Введемо заміни:

$$\begin{aligned}x_1 &= y_1 - y_2, \\x_2 &= y_1 + y_2, \\x_k &= y_k \quad \forall k = 3, \dots, n.\end{aligned}\tag{14.20}$$

Зауважимо, що старі змінні можна однозначно виразити через нові, тому перетворення не вироджене. В результаті квадратична форма буде мати вигляд

$$\Omega(x(y), x(y)) = 2a_{12}y_1^2 - 2a_{12}y_2^2 + \dots\tag{14.21}$$

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.

Контрольні запитання

- 1). Що таке білінійна і квадратична форма?
- 2). Як зобразити білінійну форму у вигляді скалярного добутку?
- 3). Як зіставити білінійній формі матрицю?
- 4). Як привести квадратичну форму до канонічного вигляду методом Лагранжа?
- 2). Чи можна привести квадратичну форму до канонічного вигляду методом Лагранжа у випадку поля характеристики 2?

ЛЕКЦІЯ II.15. ЗВЕДЕННЯ КВАДРАТИЧНОЇ ФОРМИ ДО КАНОНІЧНОГО ВИГЛЯДУ МЕТОДОМ ЯКОБІ¹² КРИТЕРІЙ СІЛЬВЕСТРА¹³

¹² **Якобі, Карл Густав Якоб** (нім. *Carl Gustav Jacob Jacobi*; 10/12/1804, Потсдам -18/02/1851, Берлін) - видатний математик.

¹³ **Сильвестр, Джеймс Джозеф** (англ. *James Joseph Sylvester*; 3/09/1814, Лондон, - 15/03/1897, Оксфорд) – відомий математик.

План

- 1). Побудова базису, в якому матриця лінійного оператора має трикутний вигляд.
- 2). Властивість діагональних елементів побудованого базису.
- 3). Знаковизначена квадратична форма.
- 4). Критерій Сільвестра додатності і від'ємності квадратичної форми.
- 5). Критерій Сільвестра нестрогої знаковизначеності квадратичної форми.

Нехай ϵ квадратична форма (14.5) з симетричною матрицею в базисі $e = \{e_1, \dots, e_n\}$. Розглянемо її мінори, що називаються кутовими:

$\Delta_1 = a_{11}, \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \dots, \Delta_n = \det(A).$	(15.1)
---	--------

Для зручності введемо величину $\Delta_0 = 1$.

Будемо шукати новий базис $e' = \{e'_1, e'_2, \dots, e'_n\}$, вектори якого виражаються через вектори базису e за формулами:

$\begin{cases} e'_1 = t_{11}e_1, \\ e'_2 = t_{12}e_1 + t_{22}e_2, \\ \dots \\ e'_k = t_{1k}e_1 + t_{2k}e_2 + \dots + t_{kk}e_k, \\ \dots \\ e'_n = t_{1n}e_1 + t_{2n}e_2 + \dots + t_{nn}e_n. \end{cases}$	(15.2)
--	--------

Для зведення квадратичної форми до канонічного вигляду треба накласти умови:

$$\Omega(e'_j, e'_k) = a'_{jk} = 0, \quad j = 1, \dots, k-1, \quad k = 2, \dots, n. \quad (15.3)$$

Тоді $a'_{kj} = 0$ завдяки симетричності форми. Розглянемо більш детально умову (15.3):

$$\Omega(e'_j, e'_k) = \Omega(t_{1j}e_1 + t_{2j}e_2 + \dots + t_{jj}e_j, e'_k) = t_{1j}\Omega(e_1, e'_k) + t_{2j}\Omega(e_2, e'_k) + \dots + t_{jj}\Omega(e_j, e'_k).$$

Зрозуміло, що для виконання умови (15.3), достатньо виконання

$$\Omega(e_j, e'_k) = 0, \quad (15.4)$$

$j = 1, \dots, k-1, \quad k = 2, \dots, n$. Оскільки треба звести квадратичну форму до канонічного вигляду, накладемо умову $\Omega(e'_k, e'_k) = a'_{kk} = 1$. Зважаючи на (15.4), будемо вимагати виконання

$$\Omega(e_k, e'_k) = 1. \quad (15.5)$$

Тепер для визначення коефіцієнтів системи (15.2) використаємо формули (15.4) і (15.5):

$$\Omega(e_j, t_{1k}e_1 + t_{2k}e_2 + \dots + t_{kk}e_k) = a_{j1}t_{1k} + a_{j2}t_{2k} + \dots + a_{jk}t_{kk} = 0 \quad \forall j \leq k-1,$$

$$\Omega(e_k, t_{1k}e_1 + t_{2k}e_2 + \dots + t_{kk}e_k) = a_{k1}t_{1k} + a_{k2}t_{2k} + \dots + a_{kk}t_{kk} = 1.$$

Отже, для знаходження коефіцієнтів розкладу векторів нового базису по старому (15.2) одержимо систему лінійних рівнянь:

$$\begin{cases} a_{11}t_{1k} + a_{12}t_{2k} + \dots + a_{1k}t_{kk} = 0, \\ a_{21}t_{1k} + a_{22}t_{2k} + \dots + a_{2k}t_{kk} = 0, \\ \dots \\ a_{k-1,1}t_{1k} + a_{k-1,2}t_{2k} + \dots + a_{k-1,k}t_{kk} = 0, \\ a_{k1}t_{1k} + a_{k2}t_{2k} + \dots + a_{kk}t_{kk} = 1. \end{cases} \quad (15.6)$$

Доведемо, що для цієї процедури зведення можна вибирати коефіцієнти в (15.2) так, що для діагональних коефіцієнтів буде виконуватись властивість

$$t_{kk} = \frac{\Delta_{k-1}}{\Delta_k}, \quad (15.7)$$

$k=1, \dots, n$.

При $k=1$ вона виконується, оскільки $1 = \Omega(e_1, e'_1) = t_{11}\Omega(e_1, e_1) = t_{11}a_{11}$, звідки

$$t_{11} = \frac{1}{a_{11}} = \frac{\Delta_0}{\Delta_1}.$$

Нехай визначено всі коефіцієнти, що входять в перші $k-1$ рядків в (15.2) і доведено формулу $t_{k-1,k-1} = \frac{\Delta_{k-2}}{\Delta_{k-1}}$. За правилом Крамера з (15.6) знайдемо t_{kk} :

$$t_{kk} = \frac{1}{\Delta_k} \begin{vmatrix} a_{11} & \dots & a_{1,k-1} & 0 \\ \cdot & \dots & \cdot & \cdot \\ a_{k-1,1} & \dots & a_{k-1,k-1} & 0 \\ a_{k,1} & \dots & a_{k,k-1} & 1 \end{vmatrix} = \frac{\Delta_{k-1}}{\Delta_k},$$

що й треба було довести.

Матриця перетворення базису (15.2) має трикутну форму, тому її визначник дорівнює добутку діагональних елементів:

$$\Delta \triangleq t_{11} t_{22} \dots t_{kk} = \frac{\Delta_0}{\Delta_1} \frac{\Delta_1}{\Delta_2} \dots \frac{\Delta_{n-1}}{\Delta_n} = \frac{1}{\Delta_n} \neq 0.$$

Отже, перехід від старого базису до нового здійснюється за допомогою невинродженої матриці.

Нагадаємо, що коефіцієнти при квадратах нових змінних дорівнюють $t_{kk}, k=1, \dots, n$. Згідно з (15.3) всі інші коефіцієнти при добутках неоднакових змінних дорівнюють нулю. Знайдемо $a'_{kk}, k=1, \dots, n$, використовуючи (15.4), (15.5) і (15.7):

$$a'_{kk} = \Omega(e'_k, e'_k) = \Omega(t_{1k}e_1 + t_{2k}e_2 + \dots + t_{kk}e_k, e'_k) = t_{kk}\Omega(e_k, e'_k) = t_{kk} = \frac{\Delta_{k-1}}{\Delta_k}.$$

Отже, в новому базисі квадратична форма буде мати канонічну форму:

$\Omega(x(y), x(y)) = \frac{\Delta_0}{\Delta_1} y_1^2 + \frac{\Delta_1}{\Delta_2} y_2^2 + \dots + \frac{\Delta_{n-1}}{\Delta_n} y_n^2.$	(15.8)
---	--------

Означення 15.1. Квадратична форма називається додатно визначеною, якщо

$$\Omega(x, x) > 0 \forall x \neq 0. \quad (15.9)$$

Означення 15.2. Квадратична форма називається від'ємно визначеною, якщо

$$\Omega(x, x) < 0 \forall x \neq 0. \quad (15.10)$$

Якщо простір унітарний, то квадратична форма записується у вигляді скалярного добутку: $\Omega(x, x) = (Ax, x)$.

Означення 15.3. 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: $a_{ii} = \Omega(e_i, e_i) = (Ae_i, e_i) \forall e_i = (0; \dots; 0; 1; 0; \dots; 0)$ (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 (10.6): $A = UA'U^{-1}$. 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 15.1. (Критерій Сільвестра). Для того, щоб квадратична форма була додатно визначеною, необхідно і достатньо, щоб усі її кутові мінори були додатні: $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$.

Доведення. Необхідність. Звужимо квадратичну форму на лінійний підпростір $E_k \subset E_n$, базис у якому $\{e_1, \dots, e_k\} \subset \{e_1, \dots, e_n\}$. Координати вектора з E_k будуть мати вигляд $[x]_e = (x_1; \dots; x_k; 0 \dots 0)^T$, а визначник матриці звуження форми

$$\Delta_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \dots & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} > 0, \quad (15.11)$$

що випливає з тільки що доведеної властивості. Це справедливо для $k = 1, \dots, n$, отже, необхідність доведено.

Достатність. Методом Якобі квадратична форма зводиться до вигляду

$$\Omega(x, x) = \frac{\Delta_0}{\Delta_1} (x'_1)^2 + \frac{\Delta_1}{\Delta_2} (x'_2)^2 + \dots + \frac{\Delta_{n-1}}{\Delta_n} (x'_n)^2. \quad (15.12)$$

Якщо $x \neq 0$, то хоча б одна координата $x'_k \neq 0$, а тоді $\Omega(x, x) > 0$.

Наслідок 1. Для від'ємної визначеності квадратичної форми необхідно і достатньо, щоб знаки мінорів чергувались таким чином:

$$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots \quad (15.13)$$

Для доведення досить розглянути додатну квадратичну форму $-\Omega(x, x)$ з додатною матрицею $-A$.

Наслідок 2. Розглянемо двовимірний випадок.

$$\Omega(x, x) = a_{11} (x_1)^2 + 2a_{12} x_1 x_2 + a_{22} (x_2)^2. \quad (15.14)$$

Для додатної визначеності цієї квадратичної форми необхідно і достатньо, щоб

$\Delta_1 = a_{11} > 0, \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = a_{11}a_{22} - (a_{12})^2 > 0.$	(15.15)
---	-----------

Зауважимо, що з цього автоматично випливає, що $a_{22} > 0$.

Наслідок 3. Усі головні мінори матриці додатно визначеної квадратичної форми додатні.

Очевидно, можна так перенумерувати змінні, що відповідний головний мінор стане кутовим. Це відповідає переставлянню рядків і стовпчиків матриці з однаковими номерами, що не змінює знаків мінорів.

Зауваження. Не слід вважати, що з невід'ємності кутових мінорів матриці квадратичної форми випливає її невід'ємна визначеність. Приклад: нехай у

квадратичної форми $\Omega(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$ коефіцієнти приймають такі значення: $a_{11} = a_{12} = 0, a_{22} < 0$. У неї $\Delta_1 = \Delta_2 = 0$, але форма не є невід'ємною. Виявляється, що для дослідження на нестрогу невід'ємність (недодатність) треба долучати всі головні мінори, а не тільки головні.

Теорема 15.2. Для того, щоб квадратична форма $(Ax, x) = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_k x_j$ була невід'ємно визначеною, необхідно і достатньо, щоб усі головні мінори її матриці були невід'ємними.

Доведення. Розглянемо допоміжну квадратичну форму

$$(A_\varepsilon x, x) = (Ax, x) + \varepsilon \sum_{j=1}^n x_j^2 \quad (\varepsilon > 0). \quad (15.16)$$

Очевидно, вона строго додатна, і з наслідку 3 теореми 15.1 усі її головні мінори додатні. Очевидно, $\lim_{\varepsilon \rightarrow 0} A_\varepsilon(x, x) = A(x, x)$. Тоді з строгої нерівності кожного головного мінору при переході до границі одержуємо нестрогу нерівність. Таким чином, **необхідність** доведена

Достатність. Позначимо через $M \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ i_1 & i_2 & \dots & i_m \end{pmatrix}$ головний мінор матриці даної квадратичної форми і через $M_\varepsilon \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ i_1 & i_2 & \dots & i_m \end{pmatrix}$ - головний мінор матриці збуреної квадратичної форми (15.16). З умови впливає строга додатність збуреного мінору, оскільки

$$M_\varepsilon \begin{pmatrix} i_1 & i_2 & \dots & i_m \\ i_1 & i_2 & \dots & i_m \end{pmatrix} = \varepsilon^m + \dots > \varepsilon^m > 0 \quad \forall 1 \leq i_1 < i_2 < \dots < i_m \leq n.$$

Тоді за критерієм Сільвестра $(A_\varepsilon x, x) > 0$, і після переходу до границі при $\varepsilon \rightarrow 0$, одержимо $A(x, x) = \lim_{\varepsilon \rightarrow 0} A_\varepsilon(x, x) \geq 0$.

Наслідок. Для того, щоб квадратична форма була недодатно визначеною, необхідно і достатньо, щоб форма $((-A)x, x)$ була невід’ємно визначеною.

Контрольні запитання

- 1). Як побудувати базис, у якому матриця лінійного оператора буде мати верхньо трикутний вигляд?
- 2). Що таке додатна, невід’ємна, від’ємна, недодатна матриці?
- 3). В чому полягає критерій Сільвестра?

ЛЕКЦІЯ II.16. КАНОНІЧНІ РІВНЯННЯ ПОВЕРХОНЬ ДРУГОГО ПОРЯДКУ

План

- 1). Поверхні обертання.
- 2). Канонічні рівняння поверхонь обертання кривих другого порядку – еліпсоїда, гіперболоїда і параболоїда.
- 3). Загальні канонічні рівняння поверхонь другого порядку.

Поверхні обертання

Нехай на площині zOx трьохвимірному простору задано криву

$$\begin{cases} F(x, z) = 0, \\ y = 0. \end{cases}$$

Величина $r = |x|$ - це відстань від точки $M_0(x, 0, z)$ до осі Oz . Якщо тепер точка починає обертатись навколо осі Oz , відстань кожної точки на колі до цієї осі залишається сталою, але тепер вона обчислюється за формулою $r = \sqrt{x^2 + y^2}$. Після підстановки цього виразу в рівняння кривої одержимо рівняння поверхні обертання:

$F\left(\pm\sqrt{x^2+y^2}, z\right)=0.$	(16.1)
---	--------

Знак «+» перед коренем береться, якщо $x \geq 0$, знак «-» береться для $x < 0$.

Приклад 16.1.
$$\begin{cases} z = x + 1, \\ y = 0. \end{cases} \Rightarrow z - 1 = x \Rightarrow z - 1 = \pm\sqrt{x^2 + y^2} \Rightarrow (z - 1)^2 = x^2 + y^2.$$

Зрозуміло, що при обертанні прямої, що лежить у координатній площині, навколо однієї з осей цієї площини утворюється конус. Отже, одержали рівняння конуса з вершиною в точці $(0,0,1)$.

Цей приклад показує, що є випадки, коли різниці в остаточному рівнянні між знаками перед коренем можна позбутись.

Читач сам напише рівняння поверхонь обертання навколо інших осей.

Канонічне рівняння еліпсоїда

Розглянемо в координатній площині zOx канонічне рівняння еліпса:

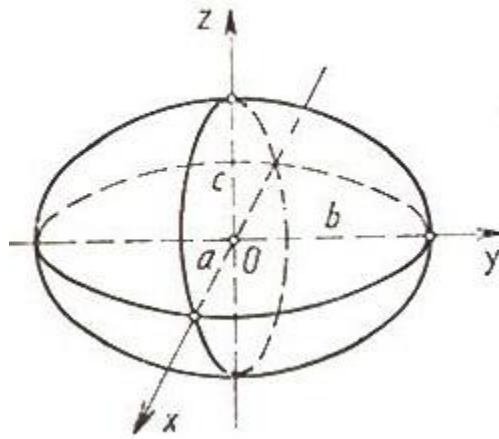
$\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1$. При його обертанні навколо осі Oz одержимо поверхню, яка називається еліпсоїдом обертання і має рівняння:

$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{c^2} = 1.$	(16.2)
--	--------

Якщо тепер здійснити перетворення гомотетії із змінною «у», тобто, розтягіння або стискання в певну кількість разів, то отримаємо канонічне рівняння еліпсоїда:

$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$	(16.3)
--	--------

Див. Мал. 16.1:



Мал. 16.1.

Канонічне рівняння гіперболоїдів

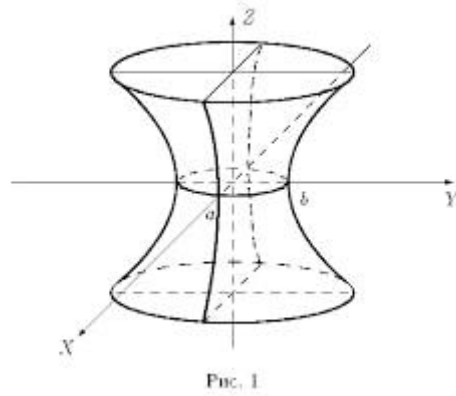
Обертанням гіперболи $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$ одержимо рівняння, так званого, однопорожнинного гіперболоїда:

$\frac{x^2}{a^2} + \frac{y^2}{a^2} - \frac{z^2}{c^2} = 1.$	(16.4)
--	--------

Звідси, як і в попередньому випадку, одержимо канонічне рівняння однопорожнинного гіперболоїда:

$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$	(16.5)
--	--------

Див. Мал. 16.2:



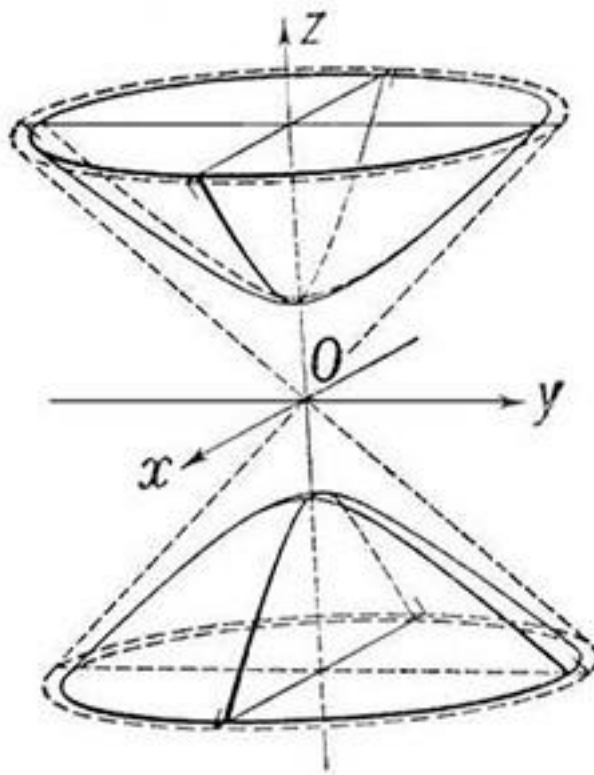
Мал. 16.2:

Можна обернути спряжену гіперболу $\frac{x^2}{a^2} - \frac{z^2}{c^2} = -1$.

В результаті отримаємо канонічне рівняння двопорожнинного гіперболоїда:

$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1.$	(16.6)
---	--------

Очевидно, рівняння має зміст при $|z| \geq c$. Це і пояснює назву. Див. Мал. 16.3:



Мал. 16.3.

Параболоїди

Еліптичний параболоїд

Запишемо таким чином канонічне рівняння параболі в площині xOy : $\frac{x^2}{p} = 2z$

Параболоїд обертання навколо осі Oz має рівняння:

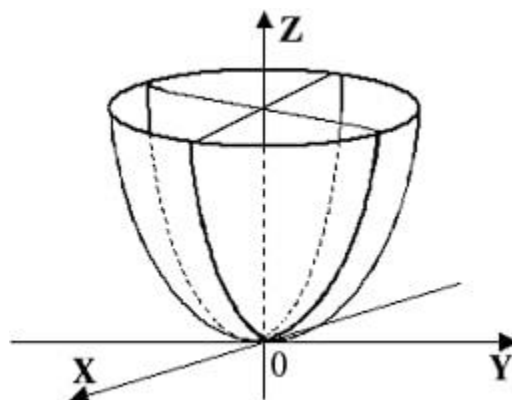
$\frac{x^2}{p} + \frac{y^2}{p} = 2z.$	(16.7)
---------------------------------------	--------

Після гомотетії по відношенню до змінної «у», одержимо:

$\frac{x^2}{p} + \frac{y^2}{q} = 2z.$	(16.8)
---------------------------------------	--------

Зрозуміло, що лінією перетину цієї поверхні з горизонтальною площиною $z = z_0 > 0$ є еліпс, що й обумовлює назву поверхні (Див. Мал. 16.4):

$$\frac{x^2}{2z_0 p} + \frac{y^2}{2z_0 q} = 1.$$



Мал. 16.4:

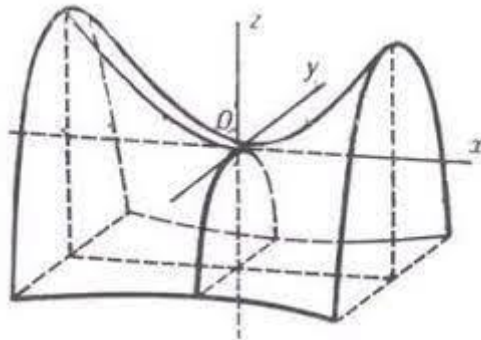
Гіперболічний параболоїд

Ця поверхня не утворюється в результаті обертання.

Нехай, наприклад, у площині zOx задано нерухому (напрямну) параболу:

$\begin{cases} x^2 = 2pz, \\ y = 0. \end{cases}$	(16.9)
--	--------

По цій параболі рухаються вершини парабол, а площини, в яких розташовані рухомі параболі, паралельні між собою і паралельні площині yOz . Вісь нерухомої параболі співпадає з віссю Oz , а осі рухомих парабол теж паралельні цій осі, але мають різні напрями з нею. Див. Мал. 16.5:



Мал. 16.5

Рівняння рухомих парабол:

$\begin{cases} y^2 = 2q(h - z), \\ x = d. \end{cases}$	(16.10)
--	---------

Тут d відстань площини рухомої параболі від координатної площини yOz .

Вершина рухомої параболі має координати $(\pm d, 0, h)$. Її координати задовольняють рівняння нерухомої параболі, тому маємо:

$$d^2 = 2ph \Rightarrow h = \frac{d^2}{2p}.$$

Підставивши одержане значення h в рівняння рухомої параболі, будемо мати:

$\begin{cases} y^2 = 2q\left(\frac{d^2}{2p} - z\right), \\ x = d. \end{cases}$	(16.11)
--	---------

Після виключення d отримаємо:

$$y^2 = 2q\left(\frac{x^2}{2p} - z\right),$$

звідки остаточно

$\frac{x^2}{p} - \frac{y^2}{q} = 2z.$	(16.12)
---------------------------------------	---------

Це рівняння називається канонічним рівнянням гіперболічного параболоїда. При перетині цієї поверхні горизонтальною площиною одержуємо гіперболу.

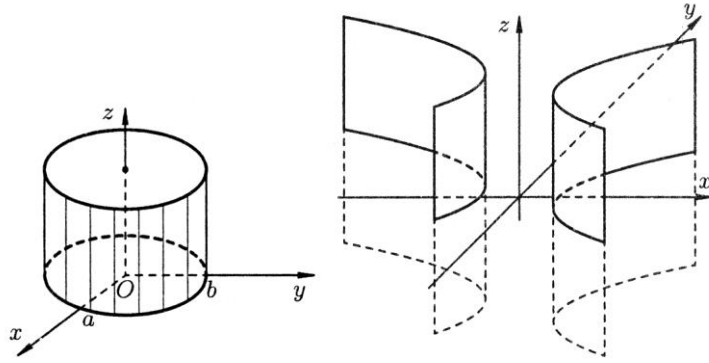
Задача. Одержати канонічне рівняння еліптичного параболоїда, поставивши задачу так само, як і у випадку гіперболічного параболоїда, але при цьому вісі рухомих і нерухомої параболі повинні мати однакові напрямки.

Циліндр

Означення 16.1. Циліндричною називається поверхня, що утворена рухом паралельних прямих, що перетинають задану криву.

Кожна така пряма називається твірною, а крива — напрямною циліндра.

Приклади циліндрів нв Мал. 16.6.



Мал. 16.6.

Нехай напрямну криву у просторі задано як перетин двох поверхонь

$\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0. \end{cases}$	(16.13)
--	---------

Нехай задано вектор $\vec{a} \sim (l, m, n)^T$ - напрямний для сім'ї прямих, що перетинають дану криву. Якщо позначити через (x_0, y_0, z_0) координати точок кривої, в яких її перетинає певна така пряма, то рівняння кожної твірної циліндра запишеться таким чином:

$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}.$	(16.14)
--	---------

Підкреслимо: $M(x, y, z)$ - координати довільної точки прямої.

Виключаючи (x_0, y_0, z_0) з системи (16.13), одержимо рівняння циліндра.

Нехай на координатній площині xOy задано криву (направну циліндра), що задається рівнянням $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$. Будемо вважати, що квадратична функція не розпадається у добуток двох лінійних. Тоді це рівняння визначає циліндр у трьохвимірному просторі. Дійсно, будь яка точка $M(x, y, z)$, де $z \in \mathbb{R}$, а x, y задовольняють рівняння, лежить на прямій, ортогональній

площині xOy , тобто, всі такі прямі паралельні. Умова, що квадратична функція не є добутком двох многочленів першої степені, забезпечує умову, що поверхня не є об'єднанням двох площин.

Конус

Означення 16.2. Конічною поверхнею або конусом називається поверхня, що утворена рухом прямої, яка проходить через фіксовану точку простору, що називається вершиною, і перетинає задану криву, що називається напрямною.

Вказана пряма називається твірною конуса.

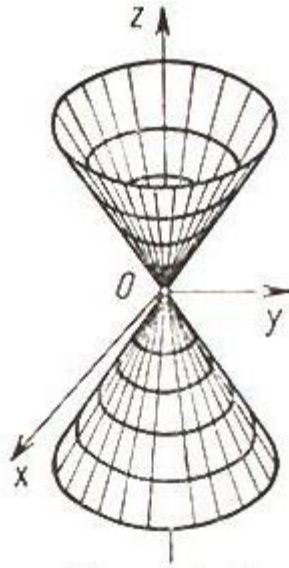
Нехай задано вершину $S(a, b, c)$ і напрямну як перетин двох площин

$\begin{cases} F_1(x, y, z) = 0, \\ F_2(x, y, z) = 0. \end{cases}$	(16.15)
--	---------

Якщо позначити через (x_0, y_0, z_0) координати точок кривої, в яких її перетинає певна така твірна пряма, то рівняння кожної твірної конуса запишеться таким чином:

$\frac{x-a}{x_0-a} = \frac{y-b}{y_0-b} = \frac{z-c}{z_0-c}.$	(16.16)
--	---------

Тут $M(x, y, z)$ - координати довільної точки прямої. Виключаючи (x_0, y_0, z_0) з системи (16.15), одержимо рівняння конуса. Див. Мал. 16.7:



Мал. 16.7

Приклад 16.2. Знайти рівняння конуса з вершиною у початку координат, напрямна якого задається як перетин сфери і площини:

$\begin{cases} x^2 + y^2 + z^2 = 1, \\ x + y + z = 1. \end{cases}$	(16.17)
--	---------

Канонічне рівняння твірної конуса має вигляд:

$\frac{x}{x_0} = \frac{y}{y_0} = \frac{z}{z_0}.$	(16.18)
--	---------

Розв'яжемо систему

$$\begin{cases} \frac{x}{x_0} = \frac{y}{y_0} = \frac{z}{z_0}, \\ x_0 + y_0 + z_0 = 1. \end{cases}$$

Позначивши рівні відношення у першому рівнянні через t одержимо:

$$x_0 = \frac{x}{t}, y_0 = \frac{y}{t}, z_0 = \frac{z}{t}, \frac{1}{t}(x + y + z) = 1.$$

Звідси

$$t = x + y + z, x_0 = \frac{x}{x + y + z}, y_0 = \frac{y}{x + y + z}, z_0 = \frac{z}{x + y + z}.$$

Підставляючи знайдені значення в перше рівняння (16.17), прийдемо до рівняння конуса:

$$\frac{x^2}{(x + y + z)^2} + \frac{y^2}{(x + y + z)^2} + \frac{z^2}{(x + y + z)^2} = 1.$$

Остаточно:

$xy + yz + zx = 0.$	(16.19)
---------------------	---------

Варто відзначити, що усі координатні осі лежать на цьому конусі.

Конус дотичних до поверхні другого порядку

Розглянемо поверхні другого порядку. Спочатку центральні:

$\frac{x^2}{A} + \frac{y^2}{B} + \frac{z^2}{C} = 1$	(16.20)
---	---------

Нехай точка $M'(x', y', z')$ не належить поверхні і

$x = x' + lt, y = y' + mt, z = z' + nt$ -	(16.21)
---	---------

параметричні рівняння дотичних до поверхні.

При деякому значенні параметру t точка $M(x, y, z)$ належить поверхні.

Підставивши її координати у рівняння поверхні, одержимо:

$\left(\frac{l^2}{A} + \frac{m^2}{B} + \frac{n^2}{C}\right)t^2 + 2\left(\frac{lx'}{A} + \frac{my'}{B} + \frac{nz'}{C}\right)t + \left(\frac{(x')^2}{A} + \frac{(y')^2}{B} + \frac{(z')^2}{C} - 1\right) = 0.$	(16.22)
---	---------

Умова дотику прямої (16.21) до поверхні (16.20) полягає в тому, що корінь t рівняння (16.22) двократний. Це еквівалентне рівності нулю дискримінанту квадратного рівняння:

$\left(\frac{lx'}{A} + \frac{my'}{B} + \frac{nz'}{C}\right)^2 - \left(\frac{l^2}{A} + \frac{m^2}{B} + \frac{n^2}{C}\right)\left(\frac{(x')^2}{A} + \frac{(y')^2}{B} + \frac{(z')^2}{C} - 1\right) = 0.$	(16.23)
---	---------

Підставимо в цю умову $l = \frac{x - x'}{t}, m = \frac{y - y'}{t}, n = \frac{z - z'}{t}$:

$\left(\frac{x'(x - x')}{A} + \frac{y'(y - y')}{B} + \frac{z'(z - z')}{C}\right)^2 -$ $-\left(\frac{(x - x')^2}{A} + \frac{(y - y')^2}{B} + \frac{(z - z')^2}{C}\right)\left(\frac{(x')^2}{A} + \frac{(y')^2}{B} + \frac{(z')^2}{C} - 1\right) = 0.$	(16.24)
--	---------

Зауважимо, що у випадку, коли точка $M'(x', y', z')$ належить поверхні, (16.24) – це рівняння дотичної площини до поверхні в цій точці.

Помітимо також, що у випадку, коли $M'(x', y', z')$ - внутрішня точка еліпсоїда, що аналітично записується нерівністю $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 < 0$ рівняння (16.24), очевидно, не визначає жодної дійсної прямої, тобто, з внутрішньої точки неможливо провести дотичну до еліпсоїда.

Рівняння (16.24) однорідне другого порядку відносно виразів

$x - x', m = y - y', z - z'$, тому воно є рівнянням конуса.

Розглянемо тепер нецентральну поверхню другого порядку (еліптичний або гіперболічний параболоїд):

$\frac{x^2}{P} + \frac{y^2}{Q} = 2z.$	(16.25)
---------------------------------------	---------

Діючи аналогічним чином, прийдемо до квадратного рівняння відносно t :

$\left(\frac{l^2}{P} + \frac{m^2}{Q}\right)t^2 + 2\left(\frac{lx'}{P} + \frac{my'}{Q} - n\right)t + \left(\frac{(x')^2}{P} + \frac{(y')^2}{Q} - 2z'\right) = 0.$	(16.26)
--	---------

Умова дотику еквівалентна кратності кореня рівняння, що еквівалентне рівності нулю дискримінанта:

$\left(\frac{lx'}{P} + \frac{my'}{Q} - n\right)^2 - \left(\frac{l^2}{P} + \frac{m^2}{Q}\right)\left(\frac{(x')^2}{P} + \frac{(y')^2}{Q} - 2z'\right) = 0.$	(16.27)
--	---------

Застосовуючи формули (16.21), виключимо звідси l, m, n і в результаті одержимо:

$\left(\frac{x'(x-x')}{P} + \frac{y'(y-y')}{Q} + (z-z') \right)^2 -$ $- \left(\frac{(x-x')^2}{P} + \frac{(y-y')^2}{Q} \right) \left(\frac{(x')^2}{P} + \frac{(y')^2}{Q} - 2z' \right) = 0.$	(16.28)
--	---------

Отже, знову одержали рівняння конуса другого порядку.

Означення 16.3. Асимптотичним для даної поверхні називається конус, утворений дотичними до неї, з точками дотику на нескінченності.

Теорема 16.1. Дотичний до поверхні конус з вершиною в її центрі (вважається, що центр – це початок координат) є асимптотичним конусом цієї поверхні.

Доведення. Введемо в рівнянні, з якого можна знайти умови дотику прямої до центральної поверхні, новий параметр $\tau = \frac{1}{t} \Rightarrow t = \frac{1}{\tau}$. Після множення обох частин на τ^2 одержимо:

$\left(\frac{l^2}{A} + \frac{m^2}{B} + \frac{n^2}{C} \right) + 2 \left(\frac{lx'}{A} + \frac{my'}{B} + \frac{nz'}{C} \right) \tau + \left(\frac{(x')^2}{A} + \frac{(y')^2}{B} + \frac{(z')^2}{C} - 1 \right) \tau^2 = 0.$	(16.29)
--	---------

Умова дотику на нескінченності прямої до поверхні еквівалентна існуванню двократного кореня $\tau = 0$ у рівняння (16.29). Це приводить до умов

--	--

$\begin{cases} \frac{l^2}{A} + \frac{m^2}{B} + \frac{n^2}{C} = 0, \\ \frac{lx'}{A} + \frac{my'}{B} + \frac{nz'}{C} = 0. \end{cases}.$	(16.30)
---	---------

Підставляючи в друге рівняння координати центру, $x' = y' = z' = 0$, приходимо до єдиного рівняння:

$$\frac{l^2}{A} + \frac{m^2}{B} + \frac{n^2}{C} = 0.$$

Для довільної точки на дотичній, що проходить через початок координат і яка в даному випадку є асимптотою, справджуються співвідношення:

$$l = \frac{x}{s}, m = \frac{y}{s}, n = \frac{z}{s}, s \in \mathbb{R}.$$

Після підстановки цих значень в рівняння і множенні обох частин на s^2 , одержимо:

$$\frac{x^2}{A} + \frac{y^2}{B} + \frac{z^2}{C} = 0.$$

Отже, одержали рівняння асимптотичного конуса для центральної кривої другого порядку.

Задача. Дослідити наявність асимптотичного конуса для нецентральных поверхонь другого порядку.

Розглянемо еліптичний і гіперболічний параболоїди, які запишемо одним рівнянням:

$$\frac{x^2}{P} + \frac{y^2}{Q} = 2z.$$

Як і раніше, підставивши в це рівняння $x = x' + lt, y = y' + mt, z = z' + nt$, одержимо:

$$\left(\frac{l^2}{P} + \frac{m^2}{Q}\right)t^2 + 2\left(\frac{lx'}{P} + \frac{my'}{Q} - n\right)t + \left(\frac{(x')^2}{P} + \frac{(y')^2}{Q} - 2z'\right) = 0$$

і після заміни $\tau = \frac{1}{t} \Rightarrow t = \frac{1}{\tau}$ -

$$\left(\frac{l^2}{P} + \frac{m^2}{Q}\right) + 2\left(\frac{lx'}{P} + \frac{my'}{Q} - n\right)\tau + \left(\frac{(x')^2}{P} + \frac{(y')^2}{Q} - 2z'\right)\tau^2 = 0.$$

Як і в попередньому випадку, умова дотику на нескінченності приводить до двох умов:

$$\begin{cases} \frac{l^2}{P} + \frac{m^2}{Q} = 0, \\ \frac{lx'}{P} + \frac{my'}{Q} = n. \end{cases}$$

Як вже було показано, перше рівняння дає рівняння конуса

$\frac{(x - x')^2}{P} + \frac{(y - y')^2}{Q} = 0.$	(16.31)
--	---------

Означення 16.4. (16.31) називається рівнянням конуса асимптотичних напрямів.

Це буде уявний конус для еліптичного параболоїда

$$\frac{(x - x')^2}{p^2} + \frac{(y - y')^2}{q^2} = 0.$$

У випадку гіперболічного параболоїда рівняння розпадається на два лінійних

$$\frac{(x-x')^2}{p^2} - \frac{(y-y')^2}{q^2} = 0 \Leftrightarrow \frac{x-x'}{p} - \frac{y-y'}{q} = 0, \frac{x-x'}{p} + \frac{y-y'}{q} = 0,$$

які є рівняннями двох площин, що проходять через пряму, паралельну осі

$$Oz: \begin{cases} x = x', \\ y = y'. \end{cases}$$

Отже, ці нецентральні поверхні на відміну від центральних не мають асимптотичних конусів.

Контрольні запитання

- 1). Який принцип створення рівняння поверхні обертання?
- 2). Який вигляд мають канонічні рівняння еліпсоїда, гіперболоїда і параболоїда?

ЛЕКЦІЯ II.17. ЛІНІЙЧАТІ ПОВЕРХНІ

План

- 1). Означення лінійчатої поверхні.
- 2). Однопорожнинний гіперболоїд як приклад лінійчатої поверхні.
- 3). Лінійчата поверхня гіперболічний параболоїд.

Означення 17.1. Лінійчатою називається поверхня, яку утворено рухом прямої. Всі такі прямі називаються твірними.

Теорема 17.1. Однопорожнинний гіперболоїд - лінійчата поверхня, що має дві серії твірних.

Доведення. Запишемо рівняння одно порошинного гіперболоїда у вигляді

$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2}.$	(17.1)
--	--------

Розклавши обидві частини на множники після зрозумілих перетворень одержимо:

$$\left(\frac{x}{a} - \frac{z}{c}\right)\left(\frac{x}{a} + \frac{z}{c}\right) = \left(1 - \frac{y}{b}\right)\left(1 + \frac{y}{b}\right) \Rightarrow$$

$\frac{\frac{x}{a} - \frac{z}{c}}{1 - \frac{y}{b}} = \frac{1 + \frac{y}{b}}{\frac{x}{a} + \frac{z}{c}} \triangleq u$	(17.2)
--	--------

Звідси одержимо пряму як лінію перетину двох площин:

$\begin{cases} \frac{x}{a} - \frac{z}{c} = u \left(1 - \frac{y}{b}\right), \\ u \left(\frac{x}{a} + \frac{z}{c}\right) = 1 + \frac{y}{b}. \end{cases}$	(17.3)
--	--------

Очевидно, пряма лежить на поверхні: якщо перемножити ці рівняння, прийдемо до рівняння одно пороожнинного гіперболоїда.

Помінявши в пропорції (17.2) середні члени, одержимо:

$\frac{\frac{x}{a} - \frac{z}{c}}{1 + \frac{y}{b}} = \frac{1 - \frac{y}{b}}{\frac{x}{a} + \frac{z}{c}} \triangleq v.$	(17.4)
---	--------

Звідси аналогічно

$\begin{cases} \frac{x}{a} - \frac{z}{c} = v \left(1 + \frac{y}{b} \right), \\ v \left(\frac{x}{a} + \frac{z}{c} \right) = 1 - \frac{y}{b}. \end{cases}$	(17.5)
--	--------

Очевидно, пряма лежить на поверхні. Доведення таке саме, як і для попередньої серії прямих

Отже, одержали дві системи твірних (17.3) і (17.5). З'ясуємо їхнє взаємне розміщення у просторі.

Теорема 17.2. Будь-які твірні кожної серії однопорожнинного гіперболоїда мимобіжні. Будь-які його твірні з різних серій перетинаються.

Доведення. Розглянемо дві твірні, наприклад, першої серії, що відповідають різним значенням параметра u_1 і u_2 :

$$\begin{cases} \frac{x}{a} - \frac{z}{c} = u_1 \left(1 - \frac{y}{b} \right), \\ u_1 \left(\frac{x}{a} + \frac{z}{c} \right) = 1 + \frac{y}{b}, \end{cases} \quad \begin{cases} \frac{x}{a} - \frac{z}{c} = u_2 \left(1 - \frac{y}{b} \right), \\ u_2 \left(\frac{x}{a} + \frac{z}{c} \right) = 1 + \frac{y}{b}. \end{cases}$$

Вважаючи, що у цих прямих є спільна точка, одержимо:

$$\begin{cases} u_1 \left(1 - \frac{y}{b} \right) = u_2 \left(1 - \frac{y}{b} \right), \\ \frac{1}{u_1} \left(1 + \frac{y}{b} \right) = \frac{1}{u_2} \left(1 + \frac{y}{b} \right) \end{cases} \Rightarrow \begin{cases} u_1 = u_2, \\ \frac{u_1 - u_2}{u_1 u_2} = 0. \end{cases}$$

Отже, при різних значеннях параметрів u_1 і u_2 прямі не перетинаються.

Зрозуміло, що такий самий висновок одержимо для прямих другої серії.

Від рівняння прямої як лінії перетину двох площин можна перейти до канонічного рівняння. Як відомо, напрямний вектор можна знайти як векторний добуток векторів нормалей цих площин. Зробивши це, побачимо, що при різних

значеннях параметрів напрямні вектори будуть не колінеарними, тобто, прямі не будуть паралельними, а значить, мимобіжними. Читачу цю перевірку радимо зробити самостійно.

Розглянемо тепер дві прямі з різних серій:

$$\begin{cases} \frac{x}{a} - \frac{z}{c} = u \left(1 - \frac{y}{b} \right), & \frac{x}{a} - \frac{z}{c} = v \left(1 + \frac{y}{b} \right), \\ u \left(\frac{x}{a} + \frac{z}{c} \right) = 1 + \frac{y}{b}, & v \left(\frac{x}{a} + \frac{z}{c} \right) = 1 - \frac{y}{b}. \end{cases}$$

Якщо у цих систем є спільний розв'язок тоді
$$\begin{cases} u \left(1 - \frac{y}{b} \right) = v \left(1 + \frac{y}{b} \right), \\ \frac{1}{u} \left(1 + \frac{y}{b} \right) = \frac{1}{v} \left(1 - \frac{y}{b} \right), \end{cases}$$

Але ці рівняння еквівалентні, звідки робимо висновок, що при будь-яких значеннях пари параметрів знайдеться спільна точка у двох прямих, оскільки з першого рівняння можна знайти y , а потім з рівняння будь-якої прямої знайти x .

Теорема 17.3. Гіперболічний параболоїд - лінійчата поверхня, що має дві серії твірних.

Як відомо, канонічне рівняння гіперболічного параболоїда має вигляд:

$$\frac{x^2}{p} - \frac{y^2}{q} = 2z.$$

Розкладемо ліву його частину на множники:

$$\left(\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} \right) \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2z.$$

Звідси одержимо пропорцію:

$\frac{\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}}}{z} = \frac{2}{\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}}} \triangleq u.$	(17.6)
---	--------

Запишемо рівняння одержаної прямої як лінію перетину двох площин:

$\begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = uz, \\ u \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2. \end{cases}$	(17.7)
--	--------

Очевидно, пряма лежить на поверхні: якщо перемножити ці рівняння, прийдемо до рівняння гіперболічного параболоїда.

Якщо поміняти місцями середні члени пропорції (17.6), прийдемо до другої серії твірних гіперболічного параболоїда. Той факт, що ці прямі лежать на його поверхні, доводиться так само, як і для першої серії. Остаточно, рівняння другої серії мають вигляд:

$\begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = 2v, \\ v \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = z. \end{cases}$	(17.8)
--	--------

Теорема 17.4. Будь-які твірні кожної серії гіперболічного параболоїда мимобіжні. Будь-які його твірні з різних серій перетинаються.

Доведення. Розглянемо дві твірні, наприклад, першої серії, що відповідають різним значенням параметра u_1 і u_2 :

$$\begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = u_1 z, \\ u_1 \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2, \end{cases} \quad \begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = u_2 z, \\ u_2 \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2. \end{cases}$$

Дослідимо систему цих чотирьох рівнянь на сумісність. Якщо система має розв'язок, тобто, прямі однієї серії перетинаються, то звідси одержимо:

$$\begin{cases} u_1 z = u_2 z, \\ \frac{2}{u_1} = \frac{2}{u_2} \end{cases} \Leftrightarrow u_1 = u_2.$$

Це й означає, що різні твірні не перетинаються. Як і у випадку однопорожнинного гіперболоїда радимо читачу перевірити, що розглянуті твірні не паралельні, отже, мимобіжні.

Візьмемо тепер дві твірні з різних серій:

$$\begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = uz, \\ u \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2, \end{cases} \quad \begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = 2v, \\ v \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = z. \end{cases}$$

З їхньої сумісності випливає:

$$\begin{aligned} uz &= 2v, \\ \frac{2}{u} &= \frac{z}{v}, \end{aligned}$$

але ці рівняння еквівалентні, звідки, як і у випадку однопорожнинного гіперболоїда, робимо висновок про існування точок перетину між довільними твірними різних серій.

Теорема 17.5. Будь-які твірні гіперболічного параболоїда певної серії паралельні одній площині.

Доведення. Розглянемо спочатку твірні першої серії як лінії перетину площин

$$\begin{cases} \frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} = uz, \\ u \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2 \end{cases}$$

і доведемо, що вони паралельні площині

$\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} = 0.$	(17.9)
--	--------

Оскільки напрямний вектор твірної можна знайти як векторний добуток векторів нормалей обох цих площин, то цей напрямний вектор ортогональний векторам нормалей. Тоді, в разі його паралельності іншій площині, необхідно і достатньо, щоб він був ортогональний його вектору нормалі. Отже, в такому разі вектори нормалей трьох площин будуть компланарними, що еквівалентно рівності нулю їхнього мішаного добутку. Знайдемо цей добуток:

$$\begin{vmatrix} \frac{1}{\sqrt{p}} & -\frac{1}{\sqrt{q}} & -u \\ \frac{u}{\sqrt{p}} & \frac{u}{\sqrt{q}} & 0 \\ \frac{1}{\sqrt{p}} & \frac{1}{\sqrt{q}} & 0 \end{vmatrix} = 0.$$

Рівність нулю визначника впливає з того, що його два рядки пропорційні.

Читач повинен самостійно пересвідчитись в тому, що твірні другої серії паралельні площині

$\sqrt{qx} - \sqrt{py} = 0.$	(17.10)
------------------------------	---------

Контрольні запитання

- 1). Що таке лінійчата поверхня?
- 2). Як знайти серії прямих у однопорожнинного гіперболоїда?
- 3). Як знайти серії прямих у гіперболічного параболоїда?

ЛЕКЦІЯ II.18. ПРИВЕДЕННЯ ЗАГАЛЬНОГО РІВНЯННЯ ПОВЕРХНІ ДО КАНОНІЧНОГО ВИГЛЯДУ

План

- 1). Перетворення зсуву рівняння другого порядку з трьома змінними.
- 2). Ортогональне перетворення рівняння поверхні другого порядку.
- 3). Інваріанти поверхонь другого порядку.
- 4). Класифікація поверхонь другого порядку.

Методика дослідження за допомогою інваріантів рівнянь другого степеня від трьох змінних ідейно нічим не відрізняється від двовимірного випадку, тому деякі технічні деталі не будуть пояснюватись занадто прискіпливо, але даватимуться посилання, коли це буде потрібно.

В розгорнутому вигляді загальне рівняння 2-го порядку від трьох змінних має вигляд:

$2\Omega(\xi, \xi) \triangleq a_{11}\xi_1^2 + a_{22}\xi_2^2 + a_{33}\xi_3^2 + 2a_{12}\xi_1\xi_2 + 2a_{13}\xi_1\xi_3 + 2a_{23}\xi_2\xi_3 +$ $+ 2a_{14}\xi_1 + 2a_{24}\xi_2 + 2a_{34}\xi_3 + a_{44} = 0.$	(18.1)
---	--------

Ввівши позначення

$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix},$	(18.2)
---	--------

$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}, a = \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix}, b = a_{44},$	(18.3)
--	--------

перепишемо рівняння у вигляді

$(A\xi, \xi) + 2(a, \xi) + b = 0,$	(18.4)
------------------------------------	--------

де A - симетрична матриця, в якій хоча б один позадіагональний елемент не дорівнює нулю (інакше квадратична форма $\Omega(\xi, \xi) = (A\xi, \xi)$ мала б діагональний вигляд) .

Здійснимо перетворення зсуву так, щоб позбавитись лінійних членів. Нехай

$x = \xi + r \Rightarrow \xi = x - r.$	(18.5)
--	--------

Тут x - вектор старих координат, x' - вектор нових координат, $r \sim (x_{01}, x_{02}, x_{03})^T$ - вектор зсуву нової системи до старої. Відзначимо також, що r - радіус-вектор нового початку координат системи $Ox_1x_2x_3$ в системі $O\xi_1\xi_2\xi_3$.

Після заміни (18.5) рівняння (18.4) перепишеться у вигляді

$(Ax, x) + 2(Ar + a, x) + ((Ar, r) + 2(a, r) + b) = 0.$	(18.6)
---	--------

Будемо тепер вимагати виконання співвідношення

$Ar + a = 0.$	(18.7)
---------------	--------

Запишемо також цю умову у координатній формі:

$a'_{13} = \frac{1}{2} \frac{\partial F(x_{01}, x_{02}, x_{03})}{\partial \xi_1} = a_{11}x_{01} + a_{12}x_{02} + a_{13}x_{03} + a_{14} = 0,$ $a'_{23} = \frac{1}{2} \frac{\partial F(x_{01}, x_{02}, x_{03})}{\partial \xi_2} = a_{12}x_{01} + a_{22}x_{02} + a_{23}x_{03} + a_{24} = 0,$ $a'_{33} = \frac{1}{2} \frac{\partial F(x_{01}, x_{02}, x_{03})}{\partial \xi_3} = a_{12}x_{01} + a_{22}x_{02} + a_{23}x_{03} + a_{24} = 0.$	(18.8)
--	--------

Тоді

$$a'_{33} = 2F(x_{01}, x_{02}, x_{03}) = a_{11}x_{01}^2 + a_{22}x_{02}^2 + a_{33}x_{03}^2 + 2a_{12}x_{01}x_{02} + 2a_{13}x_{01}x_{03} + 2a_{23}x_{02}x_{03} + 2a_{14}x_{01} + 2a_{24}x_{02} + 2a_{34}x_{03} + a_{44}.$$

Якщо $\det(A) \neq 0$, можна знайти вектор зсуву з умови

$Ar = -a \Rightarrow r = -A^{-1}a.$	(18.9)
-------------------------------------	--------

(Випадок $\det(A) = 0$ розглянемо окремо.)

Якщо ввести позначення $(Ar, r) + 2(a, r) + b = -C$, рівняння перепишеться у вигляді

$(Ax, x) = C.$	(18.10)
----------------	---------

При перетворенні симетрії у просторі

$x = -\tilde{x}$ рівняння (18.10) не змінить свого вигляду:

$$(Ax, x) = (-A\tilde{x}, -\tilde{x}) = (A\tilde{x}, \tilde{x}).$$

Отже, поверхня симетрична відносно початку координат і тому називається центральною. Вона характеризується умовою $\det(A) \neq 0$. Приклади центральних поверхонь другого порядку – це еліпсоїд, гіперболоїд. Сюди можна також включити конус, що має вершину в початку координат. Він також має рівняння другого степеня: $x_3^2 = \alpha x_1^2 + \beta x_2^2$.

Введемо однорідні координати за формулами

$$\xi_1 = \frac{x'_1}{x'_4}, \xi_2 = \frac{x'_2}{x'_4}, \xi_3 = \frac{x'_3}{x'_4}.$$

Після підстановки цих виразів у рівняння (18.1) і множення на $(x')^2$ одержимо:

$2\Omega(x',') \triangleq a_{11}x_1'^2 + a_{22}x_2'^2 + a_{33}x_3'^2 + 2a_{12}x_1'x_2' + 2a_{13}x_1'x_3' + 2a_{23}x_2'x_3' +$ $+ 2a_{14}x_1'x_4' + 2a_{24}x_2'x_4' + 2a_{34}x_3'x_4' + a_{44}x_4'^2 = 0.$	(18.11)
---	---------

Після введення симетричної матриці і чотирьохвимірному вектора

$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{pmatrix}, x' = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix}$	(18.12)
---	---------

Рівняння (18.11) запишеться у вигляді:

$(\tilde{A}x', x') = 0.$	(18.13)
--------------------------	---------

Зсуву (18.5) у декартових координатах при переході від них до однорідних координат відповідає матриця

$U_0 = \begin{pmatrix} 1 & 0 & 0 & x_{01} \\ 0 & 1 & 0 & x_{02} \\ 0 & 0 & 1 & x_{03} \\ 0 & 0 & 0 & 1 \end{pmatrix}, U_0 x' = \begin{pmatrix} 1 & 0 & 0 & x_{01} \\ 0 & 1 & 0 & x_{02} \\ 0 & 0 & 1 & x_{03} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} = \begin{pmatrix} x'_1 + x_{01}x'_4 \\ x'_2 + x_{02}x'_4 \\ x'_3 + x_{03}x'_4 \\ x'_4 \end{pmatrix}.$	(18.14)
---	---------

Зауважимо, що $\det(U_0) = \det(U_0^*) = 1$.

Ортогональне перетворення рівняння поверхні другого порядку

Нехай ортогональному перетворенню простору відповідає матриця:

$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}.$$

Нагадаємо, що матриця ортогонального перетворення трьохвимірного простору має такі властивості: спряжений до ортогонального оператора теж є ортогональним і до того ж оберненим до даного оператора; сума квадратів елементів у будь-якому рядку чи стовпчику матриці дорівнює одиниці; скалярний добуток різних рядків чи стовпчиків дорівнює нулю; модуль визначника дорівнює одиниці, а якщо матриця не змінює орієнтацію системи векторів, на які вона діє, то її визначник дорівнює одиниці.

Введемо заміну $\xi = Ux$.

Тоді рівняння (18.4) запишеться таким чином:

$$(A\xi, \xi) + 2(a, \xi) + b = 0 \Rightarrow (AUx, Ux) + 2(a, Ux) + b = 0 \Rightarrow$$

$\begin{aligned} (U^*AUx, x) + 2(U^*a, x) + b &= 0 \Leftrightarrow \\ \Leftrightarrow (U^{-1}AUx, x) + 2(U^{-1}a, x) + b &= 0. \end{aligned}$	(18.15)
---	---------

Як відомо з теорії самоспряжених операторів, його матрицю можна шляхом вибору базису, що здійснюється за допомогою ортогональної матриці, привести до діагонального вигляду з власними числами на діагоналі. При перетворенні квадратичної форми як раз і відбувається перехід до нового базису завдяки властивості ортогональної матриці переходу: $U^* = U^{-1}$. Це й показано у (18.15). Базис, в якому нова матриця має діагональну форму, складається з попарно

ортогональних власних векторів. У випадку ненульових власних чисел (еліптичний або гіперболічний типи) можна зсунути початок координат в центр поверхні, використовуючи формулу (18.9). Випадок, коли серед власних чисел є нульові, буде розглянуто окремо.

Для класифікації поверхонь введемо їхні інваріанти, як це робилось у двовимірному випадку. Нагадаємо, що характеристичний многочлен матриці інваріантний відносно вибору базису, тому всі його коефіцієнти є інваріантами:

$I_1 = sp(A) = a_{11} + a_{22} + a_{33} = \lambda_1 + \lambda_2 + \lambda_3,$ $I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix},$ $I_3 = \det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \lambda_1 \lambda_2 \lambda_3.$	(18.16)
--	---------

Тут $\lambda_1, \lambda_2, \lambda_3$ - власні числа матриці.

Легко бачити, що як і у двовимірному випадку, ортогональне перетворення і зсув у однорідних координатах здійснюється матрицею

$\tilde{U} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & x_{01} \\ u_{21} & u_{22} & u_{23} & x_{02} \\ u_{31} & u_{32} & u_{33} & x_{03} \\ 0 & 0 & 0 & 1 \end{pmatrix}$	(18.17)
--	---------

Очевидно, якщо ортогональне перетворення – це поворот, а він не змінює орієнтацію системи координат, то $\det(\tilde{U}) = 1$. Якщо ввести заміну

$x' = \tilde{U}x'',$	(18.18)
----------------------	---------

квадратична форма у однорідних координатах перетвориться чином:

$$(\tilde{A}x', x') = (\tilde{A}\tilde{U}x'', \tilde{U}x'') = (\tilde{U}^T \tilde{A}\tilde{U}x'', x'') = (\tilde{A}''x'', x'').$$

Звідси $\det(\tilde{A}'') = \det(\tilde{U}^T \tilde{A}\tilde{U}) = \det(\tilde{U}^T) \det(\tilde{A}) \det(\tilde{U}) = \det(\tilde{A}) \triangleq I_4$.

Отже, одержали ще один інваріант - $I_4 \triangleq \det(\tilde{A})$ (див. (18.12)).

Випишемо, нарешті, більш докладно формули перетворення однорідних координат під дією матриці (18.17):

$x' = \tilde{U}x'' \Rightarrow \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & x_{01} \\ u_{21} & u_{22} & u_{23} & x_{02} \\ u_{31} & u_{32} & u_{33} & x_{03} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x''_1 \\ x''_2 \\ x''_3 \\ x''_4 \end{pmatrix} = \begin{pmatrix} u_{11}x''_1 + u_{12}x''_2 + u_{13}x''_3 + x_{01}x''_4 \\ u_{21}x''_1 + u_{22}x''_2 + u_{23}x''_3 + x_{02}x''_4 \\ u_{31}x''_1 + u_{32}x''_2 + u_{33}x''_3 + x_{03}x''_4 \\ x''_4 \end{pmatrix}.$	(18.19)
--	---------

Знайшовши всі інваріанти поверхні, запишемо характеристичне рівняння матриці:

$\lambda^3 - I_1\lambda^2 + I_2\lambda - I_3 = 0.$	(18.20)
--	---------

Класифікація поверхонь другого порядку

I. Нехай всі власні числа матриці квадратичної форми, тобто, розв'язки характеристичного рівняння (18.20) не дорівнюють нулю. Це означає, що

$$I_3 = \delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} \neq 0.$$

Це, в свою чергу, означає, що поверхня центральна. Повернемо осі таким чином, щоб вони сумістились з власними векторами матриці квадратичної форми. В таких нових координатах рівняння буде мати вигляд:

$$\lambda_1 (x'_1)^2 + \lambda_2 (x'_2)^2 + \lambda_3 (x'_3)^2 + 2a'_{14}x'_1 + 2a'_{24}x'_2 + 2a'_{34}x'_3 + a_{44} = 0$$

Паралельним перенесенням пересунемо початок координат нової системи у центр поверхні, в результаті чого зникнуть лінійні члени. Рівняння буде мати вигляд:

$$\lambda_1 (x''_1)^2 + \lambda_2 (x''_2)^2 + \lambda_3 (x''_3)^2 + a''_{44} = 0.$$

Вільний його член визначимо, застосовуючи інваріанти:

$$I_4 = \Delta = \begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & a''_{44} \end{vmatrix} = \lambda_1 \lambda_2 \lambda_3 a''_{44} = I_3 a''_{44}.$$

Звідси

$a''_{44} = \frac{I_4}{I_3}.$	(18.21)
-------------------------------	---------

Таким чином приходимо до канонічного рівняння:

$\lambda_1 (x''_1)^2 + \lambda_2 (x''_2)^2 + \lambda_3 (x''_3)^2 + \frac{I_4}{I_3} = 0.$	(18.22)
--	---------

Тут існують ще випадки: поверхня вироджена чи ні.

I.1. $I_4 \neq 0$. Поверхня не вироджена. Рівняння запишемо таким чином:

$\frac{(x_1'')^2}{-\frac{I_4}{\lambda_1 I_3}} + \frac{(x_2'')^2}{-\frac{I_4}{\lambda_2 I_3}} + \frac{(x_3'')^2}{-\frac{I_4}{\lambda_3 I_3}} = 1.$	(18.23)
---	---------

В залежності від розв'язків характеристичного рівняння це може бути рівнянням еліпсоїда – дійсного чи уявного або гіперboloїда – однопорожнинного чи двопорожнинного.

I.2. $I_4 = 0$. Це рівняння конуса – дійсного чи уявного.

II. Два розв'язки характеристичного рівняння ненульові, наприклад, $\lambda_1 \lambda_2 \neq 0$, а третій $\lambda_3 = 0$. В цій ситуації $I_3 = 0$. Поверхня нецентральна або має нескінченну кількість центрів. Виберемо новими осями координат осі з напрямними векторами, які є власними для матриці A , вважаючи, що вони утворюють праву трійку. Перетворене шляхом повороту рівняння матиме вигляд:

$\lambda_1 (x_1')^2 + \lambda_2 (x_2')^2 + 2a'_{14}x_1' + 2a'_{24}x_2' + 2a'_{34}x_3' + a'_{44} = 0.$	(18.24)
---	---------

II.1. $I_4 \neq 0$. Поверхня не вироджена.

$$I_4 = \Delta = \begin{vmatrix} \lambda_1 & 0 & 0 & a'_{14} \\ 0 & \lambda_2 & 0 & a'_{24} \\ 0 & 0 & 0 & a'_{34} \\ a'_{14} & a'_{24} & a'_{34} & a'_{44} \end{vmatrix} = -a'_{34} \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ a'_{14} & a'_{24} & a'_{34} \end{vmatrix} = -(a'_{34})^2 \lambda_1 \lambda_2.$$

Звідси $a'_{34} \neq 0$.

Складемо систему рівнянь (18.8)

$$\frac{\partial F}{\partial x'_1} = \frac{\partial F}{\partial x'_2} = \frac{\partial F}{\partial x'_3} = 0$$

для знаходження центрів поверхні і випишемо матрицю розширеної такої системи:

$$\left(\begin{array}{ccc|c} \lambda_1 & 0 & 0 & a'_{14} \\ 0 & \lambda_2 & 0 & a'_{24} \\ 0 & 0 & 0 & a'_{34} \end{array} \right)$$

Ранги основної матриці дорівнює 2, а розширеної – 3. Отже, розв’язків вона не має, і поверхня нецентральна.

Паралельним переміщенням перенесемо початок координат нової системи на параболоїд. За формулами (18.8) знайдемо, як зміняться коефіцієнти рівняння при членах першої степені і вільні члени:

$$\begin{aligned} a''_{14} &= \lambda_1 x'_{01} + a'_{14}, \\ a''_{24} &= \lambda_2 x'_{02} + a'_{24}, \\ a''_{34} &= a'_{34} \neq 0, \\ a''_{44} &= \lambda_1 (x'_{01})^2 + \lambda_2 (x'_{02})^2 + 2a'_{14}x'_{01} + 2a'_{24}x'_{02} + 2a'_{34}x'_{03} + a'_{44}. \end{aligned}$$

Виберемо вектор зсуву так, щоб початок координат опинився на поверхні, і щоб його координати в старій системі задовольняли системі:

$\begin{aligned} \lambda_1 x'_{01} + a'_{14} &= 0, \\ \lambda_2 x'_{02} + a'_{24} &= 0, \\ \lambda_1 (x'_{01})^2 + \lambda_2 (x'_{02})^2 + 2a'_{14}x'_{01} + 2a'_{24}x'_{02} + 2a'_{34}x'_{03} + a'_{44} &= 0. \end{aligned}$	(18.25)
---	---------

В новому рівнянні вже не буде коефіцієнтів при перших степенях x''_1, x''_2 і вільного члена. Коефіцієнт a''_{34} знайдемо з застосуванням інваріанта:

$$I_4 = \Delta = \begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & 0 & a''_{34} \\ 0 & 0 & a''_{34} & 0 \end{vmatrix} = -\lambda_1 \lambda_2 (a''_{34})^2.$$

Звідси

$a''_{34} = \pm \sqrt{\frac{I_4}{-\lambda_1 \lambda_2}}.$	(18.26)
---	---------

Тоді рівняння (18.24) запишеться у вигляді

$$\lambda_1 (x''_1)^2 + \lambda_2 (x''_2)^2 \pm 2 \sqrt{\frac{I_4}{-\lambda_1 \lambda_2}} x''_3 = 0.$$

Від вибору знаку перед коренем залежить орієнтація осі Ox''_3 .

Запишемо канонічне рівняння поверхні:

$\frac{(x''_1)^2}{\mp \frac{1}{\lambda_1} \sqrt{\frac{I_4}{-\lambda_1 \lambda_2}}} + \frac{(x''_2)^2}{\mp \frac{1}{\lambda_2} \sqrt{\frac{I_4}{-\lambda_1 \lambda_2}}} = 2x''_3.$	(18.27)
---	---------

Це еліптичний або гіперболічний параболоїд, причому, завжди дійсний, оскільки підкореневий вираз завжди додатний, тому що він, як ми бачили, дорівнює квадрату коефіцієнта a''_{34} .

II.2. $I_4 = 0$. Поверхня вироджена.

З (18.27) випливає, що $a''_{34} = 0$. Розширена матриця для знаходження центрів має вигляд:

$$\left(\begin{array}{ccc|c} \lambda_1 & 0 & 0 & a'_{14} \\ 0 & \lambda_2 & 0 & a'_{24} \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Система сумісна і містить два лінійно незалежні рівняння, тобто, поверхня має лінію центрів. Паралельним переміщенням системи координат перенесемо початок в один з центрів, а вільний член знайдемо за формулою

$a''_{44} = \lambda_1 (x'_{01})^2 + \lambda_2 (x'_{02})^2 + 2a'_{14}x'_{01} + 2a'_{24}x'_{02} + a'_{44}.$	(18.28)
---	---------

Тут можуть бути різні випадки.

II.2.a. Лінія центрів не лежить на поверхні. Тоді $a''_{44} \neq 0$.

В результаті прийдемо до рівняння

$$\lambda_1 (x''_1)^2 + \lambda_2 (x''_2)^2 + a''_{44} = 0.$$

Звідси одержимо канонічне рівняння:

$\frac{(x''_1)^2}{-\frac{a''_{44}}{\lambda_1}} + \frac{(x''_2)^2}{-\frac{a''_{44}}{\lambda_2}} = 1.$	(18.29)
--	---------

Це рівняння циліндра еліптичного (дійсного або уявного) або гіперболічного.

II.2.b. Лінія центрів лежить на поверхні. Тоді $a''_{44} = 0$.

Канонічне рівняння має вигляд:

$$\lambda_1 (x''_1)^2 + \lambda_2 (x''_2)^2 = 0.$$

Якщо корені характеристичного рівняння мають однакові знаки, поверхня вироджується у точку, якщо вони різні, наприклад, $\lambda_1 > 0$, $\lambda_2 < 0$, многочлен у лівій частині можна розкласти на множники, в результаті чого приходимо до рівнянь двох площин, які перетинаються:

$\begin{aligned}\sqrt{\lambda_1} x_1'' + \sqrt{-\lambda_2} x_2'' &= 0, \\ \sqrt{\lambda_1} x_1'' - \sqrt{-\lambda_2} x_2'' &= 0.\end{aligned}$	(18.30)
--	---------

III. Тільки один з коренів характеристичного рівняння не дорівнює нулю.

Нехай, наприклад, $\lambda_1 = \lambda_3 = 0, \lambda_2 \neq 0$.

Поверхня має напрям, що визначається власним вектором матриці коефіцієнтів, і цей вектор відповідає ненульовому власному числу. Два інших власних вектори вибираємо ортогональними, і щоб вони утворювали праву трійку.

Тоді рівняння запишеться у вигляді:

$\lambda_2 (x_2')^2 + 2a'_{14}x_1' + 2a'_{24}x_2' + 2a'_{34}x_3' + a'_{44} = 0.$	(18.31)
--	---------

Пропускаючи kropitki викладки, сформулюємо результати.

III.1. Хоча б один коефіцієнт з двох наступних не дорівнює нулю:

$a'_{14} \neq 0 \vee a'_{34} \neq 0$. Нехай, наприклад, $a'_{14} \neq 0$. Тоді прийдемо до параболічного циліндру:

$(x_2'')^2 = -2 \frac{a'_{14}}{\lambda_2} x_1''.$	(18.32)
---	---------

III.2. $a'_{14} = a'_{34} = 0$.

Прийдемо до двох паралельних площин:

$(x_2'')^2 = -\frac{a''_{44}}{\lambda_2}.$	(18.33)
--	---------

Вони можуть співпадати, якщо $a''_{44} = 0$.

Таблиця типів поверхонь другого порядку

В основу покладено дослідження характеристичного рівняння

$$\det(\lambda I - A) = \lambda^3 - I_1 \lambda^2 + I_2 \lambda - I_3 = 0,$$

де

$$I_1 = \text{sp}(A) = a_{11} + a_{22} + a_{33} = \lambda_1 + \lambda_2 + \lambda_3,$$

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{23} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{13} & a_{33} \end{vmatrix},$$

$$I_3 = \delta = \det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = \lambda_1 \lambda_2 \lambda_3.$$

а також

$$I_4 = \Delta = \det(\tilde{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{vmatrix}.$$

I. $\lambda_1, \lambda_2, \lambda_3 \neq 0$.

I.1. $\Delta \neq 0, a''_{44} = \frac{\Delta}{\delta} \neq 0$ - поверхня не вироджена. Її канонічне рівняння має вигляд:

$$\frac{(x_1'')^2}{\left(-\frac{\Delta}{\delta \lambda_1}\right)} + \frac{(x_2'')^2}{\left(-\frac{\Delta}{\delta \lambda_2}\right)} + \frac{(x_3'')^2}{\left(-\frac{\Delta}{\delta \lambda_3}\right)} = 1.$$

Можливі випадки:

А). $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0, \frac{\Delta}{\delta} < 0$ - еліпсоїд.

Б). $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0, \frac{\Delta}{\delta} > 0$ - уявний еліпсоїд.

В). $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0, \frac{\Delta}{\delta} < 0$ - однопорожнинний гіперболоїд.

Г). $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0, \frac{\Delta}{\delta} > 0$ - двопорожнинний гіперболоїд.

І.2. $\Delta = 0, a''_{44} = \frac{\Delta}{\delta} = 0$ - поверхня вироджена. Канонічне рівняння має вигляд:

$$\lambda_1 (x''_1)^2 + \lambda_2 (x''_2)^2 + \lambda_3 (x''_3)^2 = 0.$$

А). $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 < 0$ - дійсний конус.

Б). $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$ - уявний конус.

ІІ. $\lambda_1 \neq 0, \lambda_2 \neq 0, \lambda_3 = 0$.

Можливі випадки:

ІІ.1. Якщо $\Delta \neq 0, a_{34} = \pm \sqrt{\frac{\Delta}{\lambda_1 \lambda_2}}$ - поверхня не вироджена. Канонічне рівняння має форму

$$\frac{(x''_1)^2}{\frac{1}{\lambda_1} \sqrt{-\frac{\Delta}{\lambda_1 \lambda_2}}} + \frac{(x''_2)^2}{\frac{1}{\lambda_2} \sqrt{-\frac{\Delta}{\lambda_1 \lambda_2}}} = 2x''_3.$$

Можливі випадки:

А). $\lambda_1 > 0, \lambda_2 > 0$ - еліптичний параболоїд.

Б). $\lambda_1 > 0, \lambda_2 < 0$ - гіперболічний параболоїд.

ІІ.2. Якщо $\Delta = 0, a_{34} = \pm \sqrt{\frac{\Delta}{\lambda_1 \lambda_2}} = 0$ - поверхня вироджена.

1. $a''_{44} = 2F(x'_{01}, x'_{02}, x'_{03}) \neq 0$. Канонічне рівняння має форму

$$\frac{(x_1'')^2}{\left(-\frac{2F(x_{01}', x_{02}', x_{03}')}{\lambda_1}\right)} + \frac{(x_2'')^2}{\left(-\frac{2F(x_{01}', x_{02}', x_{03}')}{\lambda_2}\right)} = 1.$$

А). $\lambda_1 > 0, \lambda_2 > 0, a_{44}'' < 0$ - еліптичний циліндр.

Б). $\lambda_1 > 0, \lambda_2 > 0, a_{44}'' > 0$ - уявний еліптичний циліндр.

В). $\lambda_1 > 0, \lambda_2 < 0, a_{44}'' < 0$ - гіперболічний циліндр.

$$2. a_{44}'' = 2F(x_{01}', x_{02}', x_{03}') = 0$$

Поверхня розпадається на пару площин.

А). $\lambda_1 > 0, \lambda_2 < 0$ - дійсні площини.

Б). $\lambda_1 > 0, \lambda_2 > 0$ - уявні площини.

ІІІ. $\lambda_1 = 0, \lambda_2 \neq 0, \lambda_3 = 0$.

1). Якщо $a_{14}' \neq 0$, канонічне рівняння має форму

$$(x_2'')^2 = -2 \frac{a_{14}''}{\lambda_2} x_1''.$$

Поверхня – параболічний циліндр. При $a_{14}' = 0, a_{34}' \neq 0$ одержимо поверхню такого ж вигляду.

2). $a_{14}' = a_{34}' = 0$. Канонічне рівняння має вигляд

$$(x_2'')^2 = -2 \frac{F(x_{01}', x_{02}', x_{03}')}{\lambda_2} \Leftrightarrow x_2'' = \pm \sqrt{-\frac{2F(x_{01}', x_{02}', x_{03}')}{\lambda_2}}.$$

Поверхня – це пара прямих дійсних або уявних – паралельних або таких, що суміщаються.

Других варіантів не існує. Класифікація повна.

Контрольні запитання

- 1). Яких членів позбавляються у рівнянні другого порядку за допомогою зсувів?
- 2). Яких членів можна позбавитись у рівнянні другого порядку за допомогою повороту координатних осей?
- 3). Що таке інваріанти поверхні другого порядку?
- 4). Як проводити класифікацію поверхонь другого порядку в залежності від значень їхніх інваріантів?

ДОДАТОК 1. КУТИ ЕЙЛЕРА¹⁴, ГРУПА $SO(3)$

Нехай у трьохвимірному просторі вибрано дві праві системи координат, що мають спільний початок: стара $Oxyz$ і нова - $Ox'y'z'$. Напрямні вектори старої системи $e = \{\vec{i}, \vec{j}, \vec{k}\}$, нової - $e' = \{\vec{i}', \vec{j}', \vec{k}'\}$.

Матриця оператора повороту від старої системи до нової:

$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}$	(I.1)
--	-------

Між векторами-стовпчиками з координат векторів в цих системах, як відомо, існує залежність:

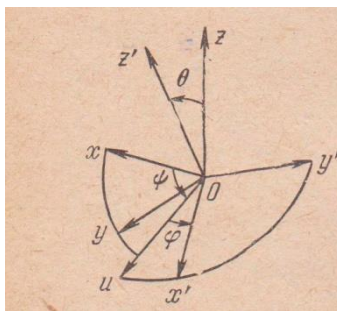
$r = Ur', r = \begin{pmatrix} x & y & z \end{pmatrix}^T, r' = \begin{pmatrix} x' & y' & z' \end{pmatrix}^T :$ $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$	(I.2)
--	-------

¹⁴ **Ейлер, Леона́рд**, правильніше **Ойлер** (нім. *Leonhard Euler*; 15/04/1707, Базель, Швейцарія - 7(18)/09/1783, Санкт-Петербург, Російська імперія – геніальний математик та фізик.

Задача полягає в знаходженні поворотів, за допомогою яких старі напрямні орти координатних осей $e = \{\vec{i}, \vec{j}, \vec{k}\}$ можна перевести в нові $e' = \{\vec{i}', \vec{j}', \vec{k}'\}$.

Нехай $\vec{u}$ - орт прямої перетину координатних площин Oxy і $Ox'y'$. Напрямок його вибирається таким чином, щоб трійка $\{\vec{k}, \vec{k}', \vec{u}\}$ була правою. Пряма з напрямним ортом $\vec{u}$ в астрономії називається лінією вузлів. Позначимо через ψ кут (в астрономії він називається прецесією) у площині Oxy , що його утворює орт $\vec{u}$ з ортом $\vec{i}$, і який відраховується в сторону найкоротшого повороту від $\vec{i}$ до $\vec{j}$. Позначимо через θ , $\theta \in [0, \pi]$ кут між ортами $\vec{k}$ і $\vec{k}'$ (в астрономії називається нутацією). І, нарешті, нехай φ - кут між $\vec{u}$ і $\vec{i}'$ в площині $Ox'y'$, який відраховується у сторону найкоротшого повороту від $\vec{i}'$ до $\vec{j}'$. φ називається кутом чистого обертання. Ці три кути називаються кутами Ейлера.

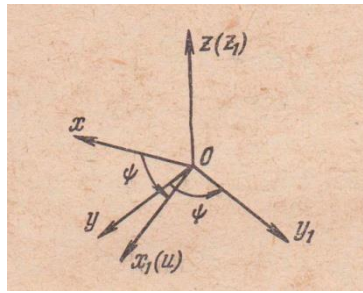
Легко бачити, що ці кути однозначно визначають напрямки нових осей Ox', Oy', Oz' за відомими осями Ox, Oy, Oz .



Мал. I.1.

Маючи стару систему і знаючи кути Ейлера здійснимо поворот старої системи до нової в три етапи.

1). Поворот системи $Oxyz$ навколо осі Oz на кут ψ . Позначимо одержану систему через $Ox_1y_1z_1$. При цьому вісь Ox_1 має напрямним ортом $\vec{u}$, осі Oz_1 і Oz співпадають.

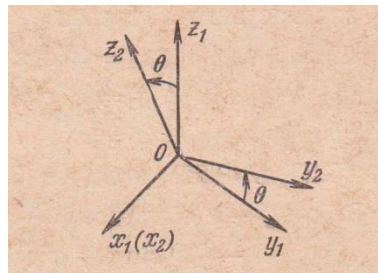


Мал. I.2.

Запишемо формули переходу, позначивши матрицю переходу через $U(\psi, 0, 0)$:

$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} x_1 \cos \psi - y_1 \sin \psi \\ x_1 \sin \psi + y_1 \cos \psi \\ z_1 \end{pmatrix}.$	(I.3)
---	-------

2). Повернемо систему $Ox_1y_1z_1$ на кут θ навколо осі Ox_1 , тобто, орта $\vec{u}$. Одержимо проміжну систему $Ox_2y_2z_2$. При цьому нова вісь Oz_2 утворить з віссю Oz_1 кут θ . Зауважимо, що вісь Oz_2 і буде віссю Oz' .



Мал. I.3.

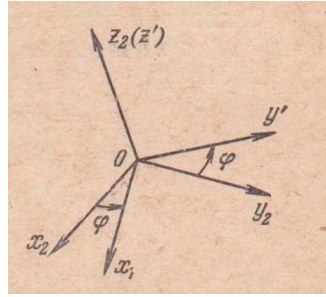
Формули переходу:

$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \cos \theta - z_2 \sin \theta \\ y_2 \sin \theta + z_2 \cos \theta \end{pmatrix}.$	(I.4)
---	-------

Позначимо матрицю переходу через $U(0, \theta, 0)$

3). Повернемо систему $Ox_2y_2z_2$ на кут φ навколо осі $Oz_2 = Oz'$.

В результаті прийдемо до системи $Ox'y'z'$.



Мал. І.3.

Формули переходу:

$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x' \cos \varphi - y' \sin \varphi \\ x' \sin \varphi + y' \cos \varphi \\ z' \end{pmatrix}.$	(І.5)
---	-------

Позначимо матрицю переходу від нової системи $Ox'y'z'$ до старої $Oxyz$ через $U(\varphi, \theta, \psi)$. Ця матриця, очевидно, є добутком трьох розглянутих матриць:

$U(\varphi, \theta, \psi) = U(\varphi, 0, 0)U(0, \theta, 0)U(\psi, 0, 0).$	(І.6)
--	-------

Остаточно:

$U(\varphi, \theta, \psi) =$ $= \begin{pmatrix} \cos \varphi \cos \psi - \sin \varphi \sin \psi \cos \theta & -\cos \varphi \sin \psi - \sin \varphi \cos \psi \cos \theta & \sin \varphi \sin \theta \\ \sin \varphi \cos \psi + \cos \varphi \sin \psi \cos \theta & -\sin \varphi \sin \psi + \cos \varphi \cos \psi \cos \theta & -\cos \varphi \sin \theta \\ \sin \psi \sin \theta & \cos \psi \sin \theta & \cos \theta \end{pmatrix}$	(І.7)
---	-------

Отже, шукана матриця переходу U - це $U(\varphi, \theta, \psi)$.

Ці матриці утворюють групу обертань трьохвимірного простору, і позначається ця група, як вже вказувалось, через $SO(3)$.

ДОДАТОК II. ТЕНЗОРНІ ДОБУТКИ

Тензорний добуток лінійних просторів

Нехай $E^{(1)}, E^{(2)}$ - лінійні простори. Розглянемо формальні добутки векторів з цих просторів $f' \otimes f'', f' \in E^{(1)}, f'' \in E^{(2)}$. Будемо вважати, що $\forall f', g' \in E^{(1)}, f'', g'' \in E^{(2)}, \lambda \in \mathbb{R}(\mathbb{C})$ виконуються властивості:

$$1). (f' + g') \otimes g' = f' \otimes g' + f'' \otimes g',$$

$$1'). f' \otimes (g' + g'') = f' \otimes g' + f' \otimes g'',$$

$$2). (\lambda f') \otimes g' = f' \otimes (\lambda g').$$

Завдяки 2) можна записати $(\lambda f') \otimes g' = f' \otimes (\lambda g') \triangleq \lambda (f' \otimes g')$.

Означення II.1. Лінійна оболонка, утворена елементами вигляду $f' \otimes f''$, називається тензорним добутком лінійних просторів $E^{(1)}, E^{(2)}$ і позначається $E^{(1)} \otimes E^{(2)}$.

Нехай $f = \{f_1^{(1)}, \dots, f_m^{(1)}\} \subset E^{(1)}, g = \{g_1^{(2)}, \dots, g_n^{(2)}\} \subset E^{(2)}$ - базиси. Неважко впевнитись в тому, що базисом у $E^{(1)} \otimes E^{(2)}$ буде сукупність векторів $\{f_i^{(1)} \otimes g_j^{(2)}, i=1, \dots, m; j=1, \dots, n\} \subset E^{(1)} \otimes E^{(2)}$.

Якщо простори $E^{(1)}, E^{(2)}$ унітарні (ортогональні), на їхньому тензорному добутку можна ввести скалярний добуток за формулою

$(f' \otimes f''), (g' \otimes g'') = (f' \otimes g')(f'' \otimes g'').$	(II.1)
--	--------

Нехай $e^{(1)} = \{e_1^{(1)}, \dots, e_m^{(1)}\} \subset E^{(1)}, e^{(2)} = \{e_1^{(2)}, \dots, e_n^{(2)}\} \subset E^{(2)}$ - ортонормовані базиси,

$$x = \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} e_i^{(1)} \otimes e_j^{(2)}, y = \sum_{i=1}^m \sum_{j=1}^n \beta_{ij} e_i^{(1)} \otimes e_j^{(2)} \in E^{(1)} \otimes E^{(2)}. \text{ Тоді}$$

$(x, y) = \left(\sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} e_i^{(1)} \otimes e_j^{(2)}, \sum_{i=1}^m \sum_{j=1}^n \beta_{ij} e_i^{(1)} \otimes e_j^{(2)} \right) = \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} \bar{\beta}_{ij}.$	(II.2)
---	--------

Звідси випливає:

$\ x\ ^2 = \left \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} e_i^{(1)} \otimes e_j^{(2)} \right ^2 = \sum_{i=1}^m \sum_{j=1}^n \alpha_{ij} ^2.$	(II.3)
---	--------

Нехай векторам $x \in E^{(1)}$, $y \in E^{(2)}$ у вибраних базисах f, g відповідають вектори-стовпчики

$$[x]_f = (x_1; \dots; x_m)^T \in \mathbb{R}^m, [y]_g = (y_1; \dots; y_n)^T \in \mathbb{R}^n.$$

Тоді вектору $x \otimes y \in E^{(1)} \otimes E^{(2)}$ зіставляється вектор-стовпчик

$$\begin{aligned} [x \otimes y]_{f \otimes g} &= [x]_f \otimes [y]_g = (x_1[y]_g; \dots; x_m[y]_g)^T = \\ &= (x_1 y_1; \dots; x_1 y_n; \dots; x_m y_1; \dots; x_m y_n)^T \in \mathbb{R}^{mn}. \end{aligned}$$

Будемо розглядати квадратні матриці $A \in L(\mathbb{R}^m, \mathbb{R}^m) \triangleq M_{mm}, B \in L(\mathbb{R}^n, \mathbb{R}^n) \triangleq M_{nn}$.

Означення II.2. Тензорним (кронекерівським) добутком $A \otimes B$ двох матриць $A \in L(\mathbb{R}^n, \mathbb{R}^n) \triangleq M_{nn}$ і $B \in L(\mathbb{R}^m, \mathbb{R}^m) \triangleq M_{mm}$ називається блочна матриця

$C = A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \cdot & \cdot & \dots & \cdot \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{pmatrix}.$	(II.4)
---	--------

Зрозуміло, що це квадратна матриця розмірності $mn \times mn$. В ній m «матричних» рядків і стільки ж «матричних» стовпчиків, в кожному такому рядку і стовпчику n звичайних рядків, відповідно, стовпчиків. Нехай елемент матриці $A \otimes B$ знаходиться в r -му матричному рядку в s -му рядку в матриці B і в i -му матричному стовпчику і в j -му стовпчику в матриці B . Тоді як елемент матриці $A \otimes B$ він має номери $(k, l) = ((r-1)n + s, (i-1)n + j)$ і дорівнює він добутку $c_{kl} = a_{ri} b_{sj}$.

Нехай $I_m \in M_{m,m}$ - одинична матриця. Тоді

$I_m \otimes B = \begin{pmatrix} B & 0 & \dots & 0 \\ 0 & B & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & B \end{pmatrix}.$	(II.2)
--	--------

Звідси випливає:

$\det(I_m \otimes B) = (\det B)^m.$	(II.6)
-------------------------------------	--------

Легко, виходячи з означення, перевіряються властивості тензорного добутку:

- 1). $\forall \lambda \in \mathbb{R} (\lambda \in \mathbb{C}) (\lambda A) \otimes B = A \otimes (\lambda B) \triangleq \lambda (A \otimes B).$
- 2). $(A + B) \otimes C = A \otimes C + B \otimes C.$
- 3). $A \otimes (B + C) = A \otimes B + A \otimes C.$
- 4). $(A \otimes B) \otimes C = (A \otimes B) \otimes C.$
- 5). $(A \otimes B)^T = A^T \otimes B^T.$

Зауваження. Взагалі кажучи, навіть у випадку $A, B \in M_{n,n}$ $A \otimes B \neq B \otimes A$.

Твердження II.1. Нехай $I_m \in M_{m,m}, I_n \in M_{n,n}$ - одиничні матриці. Тоді

$$(A \otimes I_n)(I_m \otimes B) = (I_m \otimes B)(A \otimes I_n) = A \otimes B.$$

(II.7)

Зауваження. В цих виразах фігурують звичайні добутки матриць $A \otimes I_n$ і $I_m \otimes B$.

Доведення.

$$A \otimes I_n = \begin{pmatrix} a_{11}I_n & a_{12}I_n & \dots & a_{1m}I_n \\ a_{21}I_n & a_{22}I_n & \dots & a_{2m}I_n \\ \cdot & \cdot & \dots & \cdot \\ a_{m1}I_n & a_{m2}I_n & \dots & a_{mm}I_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & 0 & a_{1m} & \dots & 0 \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ 0 & \dots & a_{11} & 0 & \dots & a_{1m} \\ \dots & \dots & \cdot & \dots & \dots & \cdot \\ a_{m1} & \dots & 0 & a_{mm} & \dots & 0 \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ 0 & \dots & a_{m1} & 0 & \dots & a_{mm} \end{pmatrix},$$

З (II.5) маємо:

$$I_m \otimes B = \begin{pmatrix} B & 0 & \dots & 0 \\ 0 & B & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & B \end{pmatrix} = \begin{pmatrix} b_{11} & \dots & b_{1m} & 0 & \dots & 0 \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ b_{m1} & \dots & b_{mm} & 0 & \dots & 0 \\ \dots & \dots & \cdot & \dots & \dots & \cdot \\ 0 & \dots & 0 & b_{11} & \dots & b_{1m} \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ 0 & \dots & 0 & b_{m1} & \dots & b_{mm} \end{pmatrix}.$$

Звідси $(A \otimes I_n)(I_m \otimes B) =$

$$= \begin{pmatrix} a_{11} & \dots & 0 & a_{1m} & \dots & 0 \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ 0 & \dots & a_{11} & 0 & \dots & a_{1m} \\ \dots & \dots & \cdot & \dots & \dots & \cdot \\ a_{m1} & \dots & 0 & a_{mm} & \dots & 0 \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ 0 & \dots & a_{m1} & 0 & \dots & a_{mm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1m} & 0 & \dots & 0 \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ b_{m1} & \dots & b_{mm} & 0 & \dots & 0 \\ \dots & \dots & \cdot & \dots & \dots & \cdot \\ 0 & \dots & 0 & b_{11} & \dots & b_{1m} \\ \cdot & \ddots & \cdot & \cdot & \ddots & \cdot \\ 0 & \dots & 0 & b_{m1} & \dots & b_{mm} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}b_{11} & \dots & a_{11}b_{1m} & & a_{1m}b_{11} & \dots & a_{1m}b_{1m} \\ & \ddots & & & & \ddots & \\ a_{11}b_{m1} & \dots & a_{11}b_{mm} & & a_{1m}b_{m1} & \dots & a_{1m}b_{mm} \\ & & & \ddots & & & \\ \dots & & & & \dots & & \\ a_{m1}b_{11} & \dots & a_{m1}b_{1m} & & a_{mm}b_{11} & \dots & a_{mm}b_{1m} \\ & \ddots & & & & \ddots & \\ a_{m1}b_{m1} & \dots & a_{m1}b_{mm} & & a_{mm}b_{m1} & \dots & a_{mm}b_{mm} \end{pmatrix} = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \vdots & \vdots & \dots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{pmatrix} =$$

$$= A \otimes B.$$

Також $(I_m \otimes B)(A \otimes I_n) =$

$$= \begin{pmatrix} b_{11} & \dots & b_{1m} & & 0 & \dots & 0 \\ & \ddots & & & & \ddots & \\ b_{m1} & \dots & b_{mm} & & 0 & \dots & 0 \\ & & & \ddots & & & \\ \dots & & & & \dots & & \\ 0 & \dots & 0 & & b_{11} & \dots & b_{1m} \\ & \ddots & & & & \ddots & \\ 0 & \dots & 0 & & b_{m1} & \dots & b_{mm} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & 0 & & a_{1m} & \dots & 0 \\ & \ddots & & & & \ddots & \\ 0 & \dots & a_{11} & & 0 & \dots & a_{1m} \\ & & & \ddots & & & \\ \dots & & & & \dots & & \\ a_{m1} & \dots & 0 & & a_{mm} & \dots & 0 \\ & \ddots & & & & \ddots & \\ 0 & \dots & a_{m1} & & 0 & \dots & a_{mm} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}b_{11} & \dots & a_{11}b_{1m} & & a_{1m}b_{11} & \dots & a_{1m}b_{1m} \\ & \ddots & & & & \ddots & \\ a_{11}b_{m1} & \dots & a_{11}b_{mm} & & a_{1m}b_{m1} & \dots & a_{1m}b_{mm} \\ & & & \ddots & & & \\ \dots & & & & \dots & & \\ a_{m1}b_{11} & \dots & a_{m1}b_{1m} & & a_{mm}b_{11} & \dots & a_{mm}b_{1m} \\ & \ddots & & & & \ddots & \\ a_{m1}b_{m1} & \dots & a_{m1}b_{mm} & & a_{mm}b_{m1} & \dots & a_{mm}b_{mm} \end{pmatrix} = A \otimes B.$$

Теорема II.1. Нехай $A, C \in M_{m,m}, B, D \in M_{n,n}$. Тоді

$(A \otimes B)(C \otimes D) = AC \otimes BD.$	(II.8)
---	--------

Доведення. $(A \otimes B)(C \otimes D) = (A \otimes I_n)(I_m \otimes B)(C \otimes I_n)(I_m \otimes D) =$
 $= (A \otimes I_n)(C \otimes I_n)(I_m \otimes B)(I_m \otimes D) = (AC \otimes I_n)(I_m \otimes BD) = AC \otimes BD.$

Наслідок 1. Нехай $A \in M_{m,m}, B \in M_{n,n}$. Тоді

$\det(A \otimes B) = \det^n(A) \det^m(B).$	(II.9)
--	--------

Доведення. Покажемо спочатку, що $\det(A \otimes I_n) = \det^n(A)$.

Очевидно, шляхом елементарних перетворень матрицю $A \otimes I_n =$

$$= \begin{pmatrix} a_{11} & \dots & 0 & & a_{1m} & \dots & 0 \\ & \ddots & & & & \ddots & \\ 0 & \dots & a_{11} & & 0 & \dots & a_{1m} \\ & & & \ddots & & & \\ a_{m1} & \dots & 0 & & a_{mm} & \dots & 0 \\ & \ddots & & & & \ddots & \\ 0 & \dots & a_{m1} & & 0 & \dots & a_{mm} \end{pmatrix}$$

можна привести до блочно-діагональної матриці

$$I_n \otimes B = \begin{pmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & A \end{pmatrix}.$$

Як відомо, це означає, що існують матриця перестановки P така, що $|\det P| = 1$ і $P(A \otimes I_n)P^T = I_n \otimes A$, оскільки з рядками і стовпчиками матриці $A \otimes I_n$ треба здійснити аналогічні елементарні перетворення для зведення її до блочно-

діагонального вигляду. При цьому $\det P \det P^T = \det^2 P = 1$. Тоді з формули (II.6) і властивості визначника випливає:

$$\det(P(A \otimes I_n)Q) = \det(P) \det(A \otimes I_n) \det(P^T) = \det(A \otimes I_n) = \det(I_n \otimes A) = \det^n(A).$$

З формули (II.7) випливає:

$$\det(A \otimes B) = \det((A \otimes I_n)(I_m \otimes B)) = \det(A \otimes I_n) \det(I_m \otimes B) = \det^n(A) \det^m(B).$$

Наслідок 2. Нехай $\exists A^{-1} \in M_{m,m}, B^{-1} \in M_{n,n}$. Тоді

$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$	(II.10)
---	---------

Доведення. З формули (II.8) випливає:

$$(A^{-1} \otimes B^{-1})(A \otimes B) = A^{-1}A \otimes B^{-1}B = I_m \otimes I_n = I_{mn},$$

що й доводить твердження.

Власні значення

Нагадаємо: множина власних чисел $\lambda_1, \dots, \lambda_m$ матриці A становить її спектр, що позначається $\sigma(A)$.

Нехай $P(x, y) = \sum_{j,k=1}^m c_{jk} x^j y^k, c_{jk} \in \mathbb{C}$ і $A \in M_{m,m}, B \in M_{n,n}$. Розглянемо матриці з

$M_{mn, mn}$, що мають вигляд

$$P(A, B) = \sum_{j,k=1}^m c_{jk} A^j B^k.$$

Приклад. $P(x, y) =$

$$= 3x^2y^2 - 5xy + 2x - 7y + 2 = 3x^2y^2 - 5x^1y + 2x^1y^0 - 7x^0y + 2x^0y^0.$$

Тоді

$$P(A, B) = 3A^2 \otimes B^2 - 5A \otimes B + 2A \otimes I - 7I \otimes B + 2I \otimes I.$$

Теорема II.2. Нехай $A \in M_{m,m}, B \in M_{n,n}$ і $\{\lambda_1, \dots, \lambda_m\} = \sigma(A)$, $\{\mu_1, \dots, \mu_n\} = \sigma(B)$.

Тоді $\sigma(P(A, B)) = \{P(\lambda_r, \mu_s), r=1, \dots, m, s=1, \dots, n\}$.

Доведення. Приведемо дані матриці до нормальної форми Жордана.

Одержимо, відповідно, матриці нормальної жорданової форми

$J_1 \in M_{m,m}, J_2 \in M_{n,n}$. Тоді існують матриці $T_1 \in M_{m,m}, T_2 \in M_{n,n}$ такі, що

$$T_1 A T_1^{-1} = J_1, T_2 B T_2^{-1} = J_2.$$

З означення тензорного добутку випливає, що $J_1^j J_2^k$ - верхньо-трикутні матриці, на головних діагоналях яких знаходяться числа $\lambda_r^j \mu_s^k, r=1, \dots, m, k=1, \dots, n$. Але діагональні елементи верхньо-трикутної матриці - це її власні числа. Отже, власні числа матриці $P(J_1, J_2)$ - це $P(\lambda_r, \mu_s), r=1, \dots, m, k=1, \dots, n$. Доведемо, що спектри матриць $P(J_1, J_2)$ і $P(A, B)$ співпадають. Застосовуючи доведені властивості тензорного добутку, одержимо:

$$\begin{aligned} J_1^j \otimes J_2^k &= T_1 A^j T_1^{-1} \otimes T_2 B^k T_2^{-1} = (T_1 \otimes T_2) (A^j \otimes B^k) (T_1^{-1} \otimes T_2^{-1}) = \\ &= (T_1 \otimes T_2) (A^j \otimes B^k) (T_1 \otimes T_2)^{-1}. \end{aligned}$$

Тобто, матриці $J_1^j \otimes J_2^k$ і $A^j \otimes B^k$ подібні, звідки і випливає співпадіння їхніх характеристичних многочленів, а значить і спектрів. Теорему доведено.

Наслідок 1. $\sigma(A \otimes I_n \pm I_m \otimes B) = \{\lambda_r \pm \mu_s, r=1, \dots, m; s=1, \dots, n\}$.

Для доведення досить застосувати теорему II.2 до многочленів $P(x, y) = x \pm y$.

Означення 16.1. Матриця $A \otimes I_n + I_m \otimes B$ називається кронекеровською сумою матриць A і B .

Наслідок 2. $\sigma(A \otimes B) = \{\lambda_r \mu_s, r = 1, \dots, m; s = 1, \dots, n\}$.

Для доведення досить застосувати теорему II.2 до многочлена $P(x, y) = xy$.

Матричні рівняння

Спочатку розв'яжемо рівняння відносно матриці $X \in M_{n,n}$, де $A, B, C \in M_{n,n}$:

$AXB = C$	(II.11)
-----------	---------

Введемо позначення для r -го рядка матриці A : $A_{r*} = (a_{r1} \ a_{r2} \ \dots \ a_{rn})$, а також для s -го стовпчика: $A_{*s} = (a_{1s} \ a_{2s} \ \dots \ a_{ns})^T$. Матриці X можна зіставити вектор $x = (X_{1*} \ X_{2*} \ \dots \ X_{n*})^T$ і аналогічно матриці C - вектор-стовпчик $c = (C_{1*} \ C_{2*} \ \dots \ C_{n*})^T$. Ліва частина матричного рівняння визначає лінійний оператор, матрицю $G \in M_{n^2, n^2}$. Виділимо елемент матриці AXB , що знаходиться в r -му рядку і s -му стовпчику:

$$A_{r*} x B_{*s} = \sum_{i,j=1}^n a_{ri} x_{ij} b_{js} = c_{rs}$$

Звідси зрозуміло, що коефіцієнт при x_{ij} дорівнює $a_{ri} b_{js}$. Робимо висновок: матриця G визначається елементами $a_{ri} b_{js}$ у рядку з номером $(r-1)n + s$ і у стовпчику з номером $(i-1)n + j, 1 \leq r, s, i, j \leq n$. Порівнюючи цей елемент з елементом тензорного добутку двох матриць, пересвідчуємось, що він є елементом матриці $A \otimes B^T$. Отже, $G = A \otimes B^T$.

Одержану конструкцію можна узагальнити на випадок більш загального рівняння

$A_1XB_1 + A_2XB_2 + \dots + A_kXB_k = C.$	(II.12)
--	---------

Воно еквівалентне рівнянню

$Gx = (A_1 \otimes B_1^T + A_2 \otimes B_2^T + \dots + A_k \otimes B_k^T)x = C.$	(II.13)
--	---------

Для аналізу існування розв'язків (II.13) можна застосовувати схему дослідження лінійних систем, наприклад, теорему Кронекера-Капеллі, тощо.

Задача 18.1. Знайти спектр оператора $L_A X \triangleq AX - XA$, $A, X \in M_{n,n}$, якщо відомий спектр матриці A : $\sigma(A) = \{\mu_1, \dots, \mu_n\}$.

Розв'язання. Треба знайти такі значення λ , при яких існують нетривіальні розв'язки рівняння $L_A X = AX - XA = \lambda X$. Згідно з (II.12)-(II.13) запишемо систему $(A \otimes I - I \otimes A^T)x = \lambda x$. Згідно з наслідком 1 з теореми II.2

$$\lambda \in \sigma(L_A) = \{\mu_r - \mu_s, r, s = 1, \dots, n\}.$$

Матричне рівняння $AX + XB = C$

Використовуючи перехід від рівняння типу (II.12) до рівняння типу (II.13), одержимо еквівалентне рівняння $Gx = (A \otimes I + I \otimes B^T)x = C$. Для того, щоб це рівняння мало єдиний розв'язок, необхідно і достатньо, щоб матриця G не мала нульових власних чисел. Згідно з наслідком 1 з теореми II.2, якщо $\{\lambda_1, \dots, \lambda_n\}$ - власні числа матриці A , а $\{\mu_1, \dots, \mu_n\}$ - власні числа матриці B , то $\{\lambda_r + \mu_s, r, s = 1, \dots, n\}$ - власні числа матриці $A \otimes I + I \otimes B$. При цьому слід

зауважити, що власні числа матриць B і B^T однакові. Звідси випливає наступний факт.

Теорема П.3. Матричне рівняння $AX + XB = C$ має єдиний розв'язок тоді і тільки тоді, коли матриці A і $-B$ не мають спільних власних чисел.

Зауваження. Рівняння $AX + XB = C$ називають рівнянням Ляпунова¹⁵.

¹⁵ Ляпунов Олександр Михайлович, видатний математик, 6/06/1857, Ярославль – 3/11/1918, Одеса.

ЛІТЕРАТУРА

1. 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. – 296 с.
2. 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.
3. Завало С.Т. Курс алгебри: підручник / С.Т. Завало. – К., Вища школа, 1985.
4. Завало С.Т. Алгебра і теорія чисел: навч. посіб. / С.Т. Завало, С.С. Левищенко та ін. – К., Вища школа, 1986. – 264 с.
5. Дубовик В. П. Вища математика: навч. посіб. для студ. вищ. навч. закл. / В.П. Дубовик., І. І. Юрик. - 4-те вид. - К. : Ігнатекс-Україна., 2013. - 648 с.
6. 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.
7. Клетеник Д. В. Сборник задач по аналитической геометрии: учеб. пособ. для вузов /Д.В. Клетеник. – СПб.: Специальная литература, 1998. – 200 с.

8. 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, Э. 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. Лекции 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.
- 42.Кряквин В.Д. Линейная алгебра. Пособие к решению задач и большая коллекция вариантов заданий /В.Д. Кряквин. – М.: Вузовская книга, 2004. – 519 с.

ПРЕДМЕТНИЙ ПОКАЖЧИК

Алгебра многочленів 9

Алгоритм Евкліда 10

Базис

- відносний 70
- канонічний 132
- ортонормований 31

Вектор 31

- нормований 31
- твірний (циклічний) 59

Вектори

- ортогональні 31
- лінійно незалежні відносно підпростору 70

Визначник 16

- Вандермонда 16
- Сільвестра 16

Група

- ортогональна 45
- спеціальна ортогональна 48

Добуток

- векторний 163
- скалярний 29
- тензорний просторів 187
- тензорний матриць 189

Ідеал кільця 6

Ізоморфізм

- лінійних просторів 42
- лінійних просторів канонічний 42

Кільце 6

- головних ідеалів 7

Критерій Сільвестра

знаковизначенності квадратичної форми 137

Конус дотичних до поверхні 155

Матриця

- Грама 33
- ортогональна 35

Метод зведення квадратичної форми до канонічного вигляду

- Лагранжа 135
- Якобі 136

Міnor 54

Многочлен

- інтерполяційний Лагранжа 24

Норма вектора 30

Нормальна форма

- Жордана 79
- Жордана нормального оператора 95
- Циклічна (Фробеніуса) 126

Образ унітарного оператора 47

Оператор

- ідемпотентний 32
- ізометричний 46
- ортогональний 36
- проектор 49
- самоспряжений 46
- спряжений 45
- унітарний 36

Ортогональні вектори 128

Поверхні другого порядку

- гіперболоїд 146
- еліпсоїд 145
- загальне рівняння 168
- конус 152
- лінійчата 161
- еліптичний 148
- обертання 144
- параболоїд гіперболічний 149
- циліндрична 151

Поле

- алгебраїчно замкнене 6
- дійсних чисел 6
- комплексних чисел 6

Проекція вектора на підпростір 32

Ранг білінійної форми 131

Раціональний дріб 25

Результант 14

Рівняння

- матричне 197
- Ляпунова 198
- характеристичне матриці

Рівняння поверхні другого порядку

- канонічне 145
- обертання 145

Розмірність

- відносна 91
- власного підпростору оператора 56

Слід матриці 52

Схема Горнера 6

Теорема

- Гамільтона-Келі 60
- Піфагора 31
- Шура 58

Форма

- білінійна 128
- канонічна оператора
- квадратична 131
- симетрична 131

Ядро лінійного оператора 55