Міністерство освіти і науки України
Харківський національний університет імені В. Н. Каразіна

Курінний Григорій Чарльзович
Шугайло Олена Олексіївна

ЛІНІЙНІ ОПЕРАТОРИ В СКІНЧЕННОВИМІРНОМУ ПРОСТОРІ ТА ЇХ МАТРИЦІ

Навчально-методичний посібник з алгебри
для студентів 1-го курсу механіко-математичного факультету

Харків – 2015

Зміст

1 Основні означення, приклади

1.1 Означення лінійного оператора

Визначення 1.1 *Лінійним перетворенням векторного простору L над полем F в векторний простір $\tilde{L}$ над полем F (лінійним оператором $f : L \to \tilde{L}$) називають закон f, згідно з яким кожному вектору $\vec{x} \in L$ ставиться у відповідність вектор $f(\vec{x}) \in \tilde{L}$, і який задовольняє дві умови*

- *умова однорідності:*

$$f(\lambda\vec{x}) = \lambda f(\vec{x}), \quad \forall\vec{x} \in L,\ \forall\lambda \in F;$$

- *умова адитивності:*

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}), \quad \forall\vec{x}, \vec{y} \in L.$$

Приклад. Нехай L — двовимірний арифметичний простір. Перевіримо, що відображення $f : L \to L$, що задане правилом

$$\{x,\ y\} \overset{f}{\mapsto} \{-2x + 3y,\ 5x - y\}$$

(тобто відображення f ставить у відповідність вектору з координатами $\{x,\ y\}$ вектор з координатами $\{-2x + 3y,\ 5x - y\}$) є адитивним і однорідним (отже є лінійним оператором), а відображення $g : L \to L$, що задане правилом

$$\{x,\ y\} \overset{g}{\mapsto} \{-2x + 3y + 1,\ 5x - y\}$$

не є ні адитивним ні однорідним і, відповідно, не є лінійним оператором.

Перевіряємо адитивність відображення f. Для цього вибираємо два вектори $\vec{a} = \{x_1,\ y_1\}$ та $\vec{b} = \{x_2,\ y_2\}$ і проводимо обчислення:

$$\vec{a} + \vec{b} = \{x_1 + x_2,\ y_1 + y_2\},$$

$$f(\vec{a}) = \{-2x_1 + 3y_1,\ 5x_1 - y_1\}, \quad f(\vec{b}) = \{-2x_2 + 3y_2,\ 5x_2 - y_2\},$$
$$f(\vec{a} + \vec{b}) = \{-2(x_1 + x_2) + 3(y_1 + y_2),\ 5(x_1 + x_2) - (y_1 + y_2)\},$$
$$f(\vec{a}) + f(\vec{b}) = \{-2x_1 + 3y_1,\ 5x_1 - y_1\} + \{-2x_2 + 3y_2,\ 5x_2 - y_2\} =$$
$$= \{-2x_1 + 3y_1 - 2x_2 + 3y_2,\ 5x_1 - y_1 + 5x_2 - y_2\}.$$

Очевидно, що $f(\vec{a} + \vec{b}) = f(\vec{a}) + f(\vec{b})$ і перевірка адитивності відображення f закінчена.

Перевіряємо неадитивність відображення g. Для цього вибираємо два вектори $\vec{a} = \{x_1,\ y_1\}$ та $\vec{b} = \{x_2,\ y_2\}$ і проводимо обчислення:

$$\vec{a} + \vec{b} = \{x_1 + x_2,\ y_1 + y_2\},$$

$$g(\vec{a}) = \{-2x_1 + 3y_1 + 1,\ 5x_1 - y_1\}, \quad g(\vec{b}) = \{-2x_2 + 3y_2 + 1,\ 5x_2 - y_2\},$$

$$g(\vec{a} + \vec{b}) = \{-2(x_1 + x_2) + 3(y_1 + y_2) + 1,\ 5(x_1 + x_2) - (y_1 + y_2)\},$$

$$g(\vec{a}) + g(\vec{b}) = \{-2x_1 + 3y_1 + 1,\ 5x_1 - y_1\} + \{-2x_2 + 3y_2 + 1,\ 5x_2 - y_2\} =$$

$$= \{-2x_1 + 3y_1 - 2x_2 + 3y_2 + 2,\ 5x_1 - y_1 + 5x_2 - y_2\}.$$

Оскільки

$$-2(x_1 + x_2) + 3(y_1 + y_2) + 1 \neq -2x_1 + 3y_1 - 2x_2 + 3y_2 + 2,$$

то $g(\vec{a} + \vec{b}) \neq g(\vec{a}) + g(\vec{b})$ і перевірка неадитивності відображення g закінчена.

Перевіряємо однорідність відображення f і неоднорідність відображення g. Для цього вибираємо вектор $\vec{a} = \{x_1,\ y_1\}$, число λ і проводимо обчислення:

$$\lambda\vec{a} = \{\lambda x_1,\ \lambda y_1\},\ f(\vec{a}) = \{-2x_1 + 3y_1,\ 5x_1 - y_1\},\ f(\lambda\vec{a}) = \{-2\lambda x_1 + 3\lambda y_1,\ 5\lambda x_1 - \lambda y_1\},$$

$$\lambda f(\vec{a}) = \{\lambda(-2x_1 + 3y_1),\ \lambda(5x_1 - y_1)\},\ g(\vec{a}) = \{-2x_1 + 3y_1 + 1,\ 5x_1 - y_1\},$$

$$g(\lambda\vec{a}) = \{-2\lambda x_1 + 3\lambda y_1 + 1,\ 5\lambda x_1 - \lambda y_1\},\ \lambda g(\vec{a}) = \{\lambda(-2x_1 + 3y_1 + 1),\ \lambda(5x_1 - y_1)\}.$$

Оскільки $f(\lambda\vec{a}) = \lambda f(\vec{a})$ при будь-яких $\vec{a}$ і λ, то f — однорідне відображеня. А оскільки $g(\lambda\vec{a}) \neq \lambda g(\vec{a})$ при $\lambda \neq 1$, то відображення g не однорідне.

1.2 Приклади лінійних операторів

В наведених нижче прикладах $L,\ \tilde{L}$ означають векторний (або лінійний[1]) простір над полем F; $\vec{x}, \vec{y}, \ldots$ означають вектори, елементи лінійного простору L або $\tilde{L}$; $\lambda, \mu, \ldots$ — числа, елементи поля F; $f, g \ldots$ — лінійні перетворення.

Приклади.

1. *Нульовий* оператор $0 : L \to \tilde{L} = \{\vec{0}\}$ ставить у відповідність кожному вектору $\vec{x} \in L$ нульовий вектор, тобто $0(\vec{x}) = \vec{0}$. Перевіркою того, що це справді є лінійним оператором є

$$0(\lambda\vec{x}) = \vec{0},\ 0(\vec{x}) = \vec{0} \ \Rightarrow\ 0(\lambda\vec{x}) = \lambda 0(\vec{x});$$

$$0(\vec{x} + \vec{y}) = \vec{0},\ 0(\vec{x}) = 0(\vec{y}) = \vec{0} \ \Rightarrow\ 0(\vec{x} + \vec{y}) = 0(\vec{x}) + 0(\vec{y}).$$

2. *Одиничний* оператор $\mathrm{id} : L \to L$ ставить у відповідність кожному вектору $\vec{x} \in L$ той же вектор, тобто $\mathrm{id}(\vec{x}) = \vec{x}$. Перевіркою того, що це справді є лінійним оператором є

$$\mathrm{id}(\lambda\vec{x}) = \lambda\vec{x},\ \mathrm{id}(\vec{x}) = \vec{x} \ \Rightarrow \mathrm{id}(\lambda\vec{x}) = \lambda\,\mathrm{id}(\vec{x});$$

$$\mathrm{id}(\vec{x} + \vec{y}) = \vec{x} + \vec{y},\ \mathrm{id}(\vec{x}) = \vec{x},\ \mathrm{id}(\vec{y}) = \vec{y} \ \Rightarrow \mathrm{id}(\vec{x} + \vec{y}) = \mathrm{id}(\vec{x}) + \mathrm{id}(\vec{y}).$$

[1] Лінійний простір і векторний простір є синонімічними поняттями. Вважаємо, що знайомство з лінійними просторами уже відбулося. Тому основні поняття лінійних просторів використовуються без нагадуваня означень і без пояснень

3. *Скалярний* оператор $\mu : L \to L$ ставить у відповідність кожному вектору $\vec{x}$ той же самий вектор помножений на μ, тобто $\mu(\vec{x}) = \mu\vec{x}$. Перевіркою того, що це справді є лінійним оператором є

$$\mu(\lambda\vec{x}) = \mu\lambda\vec{x},\ \mu(\vec{x}) = \mu\vec{x}\ \Rightarrow \mu(\lambda\vec{x}) = \lambda\mu(\vec{x});$$

$$\mu(\vec{x}+\vec{y}) = \mu\vec{x}+\mu\vec{y},\ \mu(\vec{x}) = \mu\vec{x},\ \mu(\vec{y}) = \mu\vec{y}\ \Rightarrow \mu(\vec{x}+\vec{y}) = \mu(\vec{x})+\mu(\vec{y}).$$

4. Оператор *проектування* на підпростір паралельно іншому підпростору. Припустимо, що лінійний простір L є прямою сумою двох підпросторів L_1 та L_2, тобто кожен вектор $\vec{z} \in L$ можна записати і тільки одним способом у вигляді

$$\vec{z} = \vec{x}+\vec{y},\ \vec{x} \in L_1,\ \vec{y} \in L_2.$$

Оператор $\mathrm{pr} : L \to L$ проектування на L_1 паралельно L_2 задають правилом

$$(\vec{z} = \vec{x}+\vec{y}, \quad \vec{x} \in L_1, \vec{y} \in L_2)\ \Rightarrow (\mathrm{pr}(\vec{z}) = \vec{x}).$$

Перевіримо, що оператор проектування pr дійсно є лінійним оператором, тобто він задовольняє умову однорідності та умову адитивності. Для цього виберемо два вектори $\vec{z}_1,\ \vec{z}_2 \in L$ і числа $\lambda,\ \mu \in F$. Розкладему вектори $\vec{z}_1,\ \vec{z}_2,\ \lambda\vec{z}_1,\ \mu\vec{z}_2$ в суму:

$$\vec{z}_1 = \vec{u}_1+\vec{v}_1,\ \vec{z}_2 = \vec{u}_2+\vec{v}_2 \quad (\vec{u}_1,\ \vec{u}_2 \in L_1;\ \vec{v}_1,\ \vec{v}_2 \in L_2).$$

Тоді

$$\lambda\vec{z}_1 = \lambda\vec{u}_1+\lambda\vec{v}_1,\ \mu\vec{z}_2 = \mu\vec{u}_2+\mu\vec{v}_2 \quad (\lambda\vec{u}_1,\ \mu\vec{u}_2 \in L_1;\ \lambda\vec{v}_1,\ \mu\vec{v}_2 \in L_2),$$

$$\mathrm{pr}(\vec{z}_1) = \vec{u}_1,\ \mathrm{pr}(\vec{z}_2) = \vec{u}_2,\ \mathrm{pr}(\lambda\vec{z}_1) = \lambda\vec{u}_1 = \lambda\,\mathrm{pr}(\vec{z}_1),\ \mathrm{pr}(\mu\vec{z}_2) = \mu\vec{u}_2 = \mu\,\mathrm{pr}(\vec{z}_2)$$

і

$$\mathrm{pr}(\lambda\vec{z}_1+\mu\vec{z}_2) = \lambda\,\mathrm{pr}(\vec{z}_1)+\mu\,\mathrm{pr}(\vec{z}_2).$$

5. Нехай $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ — базис тривимірного простору L. Тоді відображення

$$f : L \to L, \quad f(x_1\vec{e}_1+x_2\vec{e}_2+x_3\vec{e}_3) = x_3\vec{e}_3 \qquad (x_1, x_2, x_3 \in F)$$

є проектуванням цього тривимірного простору на одновимірний підпростір з базисом $\vec{e}_3$ паралельно двовимірному підпростору з базисом $\{\vec{e}_1, \vec{e}_2\}$.

6. Розглядається площина — двовимірний лінійний простір над полем дійсних чисел. Вона складається із векторів на дійсній площині, які виходять із заданої точки — початку відліку. Обертання цієї площини навколо початку відліку є лінійним оператором.

Теорема 1.1 *Нехай L — $n-$вимірний лінійний простір, $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\} \in L$ — його базис і $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\} \in L$ — довільна система векторів із L. Тоді існує і до того ж єдиний лінійний оператор $f : L \to L$ такий, що*

$$f(\vec{e}_i) = \vec{a}_i, \quad i = 1, 2, \ldots, n. \tag{1}$$

Доведення. В припущенні, що лінійний оператор f, для якого виконується (1) існує, доводимо його єдиність. Справді, будь-який вектор $\vec{x} \in L$ може бути єдиним чином записаний у вигляді

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i, \quad x_1, x_2, \ldots, x_n \in F. \tag{2}$$

Однорідність і адитивність оператора f забезпечують нам, що образ $f(\vec{x})$ вектора $\vec{x}$ визначається єдиним чином:

$$f(\vec{x}) = f\left(\sum_{i=1}^{n} x_i \vec{e}_i\right) = \sum_{i=1}^{n} x_i f(\vec{e}_i) = \sum_{i=1}^{n} x_i \vec{a}_i.$$

Єдиність потрібного оператора перевірена. Переходимо до доведення існування цього оператора. Доведення конструктивне, тобто ми вказуємо на той оператор, довести існування якого нам потрібно, — для вектора $\vec{x}$ (2) розглянемо відображення

$$f(\vec{x}) = \sum_{i=1}^{n} x_i \vec{a}_i.$$

Перевіримо, що вказане нами відображення $f : L \to L$ є лінійним оператором для якого виконується умова (1).

Беремо два довільні вектори $\vec{x}, \vec{y} \in L : \vec{x} = \sum\limits_{i=1}^{n} x_i \vec{e}_i,\ \vec{y} = \sum\limits_{i=1}^{n} y_i \vec{e}_i$, і число $\lambda \in F$. Тоді

$$\vec{x} + \vec{y} = \sum_{i=1}^{n} (x_i + y_i) \vec{e}_i,\ \lambda \vec{x} = \sum_{i=1}^{n} \lambda x_i \vec{e}_i,$$

$$f(\vec{x} + \vec{y}) = \sum_{i=1}^{n} (x_i + y_i) \vec{a}_i = \sum_{i=1}^{n} x_i \vec{a}_i + \sum_{i=1}^{n} y_i \vec{a}_i = f(\vec{x}) + f(\vec{y}),$$

$$f(\lambda \vec{x}) = \sum_{i=1}^{n} (\lambda x_i) \vec{a}_i = \lambda \sum_{i=1}^{n} x_i \vec{a}_i = \lambda f(\vec{x}).$$

Однорідність та адитивність відображення f перевірені.

Оскільки

$$\vec{e}_1 = 1 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + \ldots + 0 \cdot \vec{e}_n,\ \vec{e}_2 = 0 \cdot \vec{e}_1 + 1 \cdot \vec{e}_2 + \ldots + 0 \cdot \vec{e}_n,\ \ldots$$

то $f(\vec{e}_i) = 1 \cdot \vec{a}_i = \vec{a}_i$ $(i = 1, 2, \ldots, n)$, умова (1) виконана і теорема доведена повністю.

$\blacksquare$

Приклади.

1. В чотиривимірному просторі з базисом $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$ візьмемо чотири вектори

$$\vec{a}_1 = 2\vec{e}_1 - 3\vec{e}_2 - 7\vec{e}_3 + \vec{e}_4, \quad \vec{a}_2 = -\vec{e}_2 + 3\vec{e}_4, \quad \vec{a}_3 = -\vec{e}_1 - \vec{e}_2 - \vec{e}_3 + 9\vec{e}_4, \quad \vec{a}_4 = -\vec{e}_2 + 3\vec{e}_4,$$

Лінійний оператор f, який переводить вектори $\vec{e}_i$ у вектори $\vec{a}_i$ $(i = 1, 2, 3, 4)$, ставить у відповідність кожному вектору $\vec{x} = (x_1, x_2, x_3, x_4)$ вектор з координатами

$$f(\vec{x}) = \{2x_1 - x_3,\ -3x_1 - x_2 - x_3 - x_4,\ -7x_1 - x_3,\ x_1 + 3x_2 + 9x_3 + 3x_4\}.$$

2. Розглядаємо лінійний простір L всіх дійсних многочленів від змінної x, які мають степінь n або менше. Похідна від многочлена із L є лінійним оператором, оскільки похідна від суми двох многочленів є сумою похідних доданків і скаляр (число) можна виносити за знак похідної.

3. Розглядаємо лінійний простір L всіх дійсних многочленів $P_n(x)$ від змінної x, які мають степінь n або менше. Оскільки визначений інтеграл від суми двох многочленів є сумою інтегралів доданків і скаляр (число) можна виносити за знак інтеграла, то відображення f

$$P_n(x) \overset{f}{\mapsto} Q_n(x) = \frac{1}{x}\int\limits_0^x P_n(t)\,dt$$

є лінійним оператором.

4. Розглядаємо лінійний простір L всіх функцій $f(x)$, які можна записати у вигляді $f(x) = a\sin x + b\cos x$. Знаходження похідної в цьому просторі є лінійним оператором.

5. Нехай L є лінійним простором стовпчиків заданої довжини n, елементами яких є числа, що належать певному полю. Тоді відображення f, що ставить у відповідність стовпчику добуток цього стовпчика зліва на задану матрицю $A_{n\times n}$ є лінійним оператором $f : L \to L$.

6. Нехай L дійсний двовимірний арифметичний лінійний простір. Знайти лінійний оператор $f : L \to L$, який переводить вектори $\vec{a}_1 = \{3, 2\}$, $\vec{a}_2 = \{1, 1\}$ відповідно у вектори $\vec{b}_1 = \{5, 5\}$, $\vec{b}_2 = \{4, 7\}$, тобто

$$\vec{b}_1 = f(\vec{a}_1), \quad \vec{b}_2 = f(\vec{a}_2).$$

Вираз "знайти лінійний оператор" означає, що потрібно знайти правило, згідно з яким кожному вектору $\vec{x} \in L$ ставиться у відповідність його образ $f(\vec{x})$.

Вектор $\vec{x} = \{x_1,\ x_2\}$ записуємо у вигляді $\vec{x} = t_1 \cdot \vec{a}_1 + t_2 \cdot \vec{a}_2$. В таких позначеннях

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}.$$

і

$$\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 - 2x_2 \\ -x_1 + 3x_2 \end{pmatrix},$$

або

$$t_1 = x_1 - 2x_2, \quad t_2 = -x_1 + 3x_2.$$

$$\text{Тепер } f(\vec{x}) = t_1 f(\vec{a}_1) + t_2 f(\vec{a}_2) = (x_1 - 2x_2)\vec{b}_1 + (-x_2 + 3x_2)\vec{b}_2 =$$

$$= \{(x_1 - 2x_2)5 + (-x_2 + 3x_2)4,\ (x_1 - 2x_2)5 + (-x_2 + 3x_2)7\}.$$

Отже відповіддю буде

$$f(\vec{x}) = \{x_1 + 2x_2,\ -2x_1 + 11x_2\}.$$

1.3 Ранг та дефект лінійного оператора

Зауважимо, що кожний лінійний оператор f переводить нульовий вектор в нульовий. Це випливає з однорідності:

$$f(0 \cdot \vec{0}) = 0 \cdot f(\vec{0}) = 0 \cdot \vec{0} = \vec{0}.$$

Визначення 1.2 *Для заданого оператора $f : L \to \tilde{L}$ множина*

$$\ker f = \{\vec{x} \in L |\ f(\vec{x}) = \vec{0}\}$$

називається ядром, а множина

$$\operatorname{im} f = \{\vec{y} \in \tilde{L} |\ \vec{y} = f(\vec{x}) \textit{ для деякого } \vec{x} \in L.\}$$

називається образом цього оператора. Образ називають також коядром, а ядро інколи називають кообразом оператора.

Теорема 1.2 *Нехай маємо лінійний оператор $f : L \to \tilde{L}$; L, $\tilde{L}$ — лінійні простори над полем F. Ядро і образ оператора f є підпросторами L та $\tilde{L}$ відповідно.*

Доведення. Оскільки $f(\vec{0}) = \vec{0}$, то і ядро і образ містять в собі нульовий вектор.

Виберемо будь-які два вектори $\vec{x}, \vec{y} \in L$ $\lambda, \mu \in F$. Якщо $\vec{x}, \vec{y} \in \ker f$, то $f(\vec{x}) = f(\vec{y}) = \vec{0}$,

$$f(\lambda\vec{x} + \mu\vec{y}) = \lambda f(\vec{x}) + \mu f(\vec{y}) = \vec{0},$$

і $\lambda\vec{x} + \mu\vec{y} \in \ker f$. Доведення того, що $\ker f$ є підпростором лінійного простору L завершене.

Виберемо тепер будь-які два вектори $\vec{x}, \vec{y} \in \tilde{L}$ $\lambda, \mu \in F$. Якщо $\vec{x}, \vec{y} \in \operatorname{im} f$, то $f(\vec{x'}) = \vec{x},\ f(\vec{y'}) = \vec{y}$ для деяких $\vec{x'}, \vec{y'} \in L$ і

$$f(\lambda\vec{x'} + \mu\vec{y'}) = \lambda\vec{x} + \mu\vec{y}.$$

А це означає, що $\lambda\vec{x}+\mu\vec{y} \in \operatorname{im} f$. Доведення того, що $\operatorname{im} f$ є підпростором лінійного простору $\tilde{L}$ завершене.

∎

Визначення 1.3 *Вимірність ядра оператора f називається дефектом і позначається* $\operatorname{def} f$*, а вимірність образу називається рангом лінійного перетворення f і позначається* $\operatorname{rang} f$:

$$\dim \ker f = \operatorname{def} f, \quad \dim \operatorname{im} f = \operatorname{rang} f.$$

Приклади.

1. Образом оператора обертання площини буде вся площина, оскільки кожен вектор є образом. Відповідно, ранг цього оператора дорівнює 2. В нульовий вектор при обертанні переходить лише нульовий вектор. Тому ядро в цього оператора збігається з нульовим підпростором (кажуть — нульове, або тривіальне ядро). Відповідно, дефект оператора обертання дорівнює нулю.

2. Дефект оператора проектування на вісь паралельно площині дорівнює 2, а ранг дорівнює 1.

3. Дефект нульового оператора $0 : L \to \tilde{L}$ дорівнює вимірності всього простору L, а ранг дорівнює нулю.

4. Ранг одиничного оператора $\operatorname{id} : L \to L$ дорівнює вимірності всього простору L, а дефект дорівнює нулю.

Теорема 1.3 *Нехай $f : L \to \tilde{L}$ — лінійний оператор, L — скінченновимірний простір над полем F. Тоді*

$$\operatorname{rang} f + \operatorname{def} f = \dim L. \tag{3}$$

Доведення. Нехай $\dim\, L = n,\ \operatorname{rang} f = p,\ \operatorname{def} f = q$.

Виберемо в L базис $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$. Оскільки кожен вектор $\vec{x} \in L$ можна записати у вигляді $\vec{x} = \sum\limits_{i=1}^{n} x_i\vec{e}_i$ $(x_1, x_2, \ldots, x_n \in F)$, то кожен вектор $f(\vec{x}) \in \operatorname{im} f$ можна записати у вигляді $f(\vec{x}) = \sum\limits_{i=1}^{n} x_i f(\vec{e}_i)$. Тому $\operatorname{im} f$ є лінійною оболонкою векторів $\{f(\vec{e}_1), f(\vec{e}_2), \ldots, f(\vec{e}_n)\} \in \tilde{L}$ і вектори $\vec{e}_1, \ldots, \vec{e}_n$ можна підібрати так, щоб вектори $\{f(\vec{e}_1), \ldots, f(\vec{e}_p)\}$ утворювали базис $\operatorname{im} f$.

Із означення базису випливає, що для будь-якого $i = p+1, p+2, \ldots, n$ вектори $f(\vec{e}_1), \ldots, f(\vec{e}_p)$, $f(\vec{e}_i)$ є лінійно залежними і для деяких $b_1, b_2, \ldots, b_p, b_i \in F$, серед яких є ненульове число, буде виконуватись рівність

$$b_1 f(\vec{e}_1) + b_2 f(\vec{e}_2) + \ldots + b_p f(\vec{e}_p) + b_i f(\vec{e}_i) = \vec{0}. \tag{4}$$

Число b_i в (4) є ненульовим — в протилежному випадку вектори $f(\vec{e}_1), \ldots, f(\vec{e}_p)$ були б лінійно залежними. Отже рівність (4) можна розділити на b_i і для деяких $a_{i1}, a_{i2}, \ldots, a_{ip} \in F$ можна записати

$$a_{i1}f(\vec{e}_1) + a_{i1}f(\vec{e}_1) + \ldots + a_{ip}f(\vec{e}_p) = f(\vec{e}_i), \quad i = p+1, p+2, \ldots, n,$$

або

$$f\left(\vec{e}_i - \sum_{j=1}^{p} a_{ij}\vec{e}_j\right) = \vec{0}, \quad i = p+1, p+2, \ldots, n, \tag{5}$$

Позначимо

$$\vec{u}_{i-p} = \vec{e}_i - \sum_{j=1}^{p} a_{ij}\vec{e}_j.$$

Рівності (5) показують, що $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-p}$ належать ядру. Із лінійної незалежності векторів $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ випливає лінійна незалежність векторів $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-p}$. Тому маємо доведену нерівність

$$\operatorname{def} f \geq n - p. \tag{6}$$

Позначимо через $L_1 \subseteq L$ підпростір з базисом $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_p\}$. Із лінійної незалежності векторів $f(\vec{e}_1), f(\vec{e}_2), \ldots, f(\vec{e}_p)$ випливає, що

$$L_1 \cap \ker f = \vec{0}.$$

Із відомої рівності

$$\dim(L_1 + \ker f) = \dim L_1 + \dim \ker f - \dim(L_1 \cap \ker f)$$

одержуємо рівність

$$\dim(L_1 + \ker f) = \dim L_1 + \dim \ker f,$$

нерівність

$$p + \operatorname{def} f \leq n$$

і, відповідно, нерівність

$$\operatorname{def} f \leq n - p. \tag{7}$$

Нерівності (6) та (7) доводять рівність (3).

■

1.4 Обмеження лінійного оператора на інваріантний підпростір

Надалі будемо розглядати тільки лінійні перетворення векторного простору L над полем F в себе, тобто лінійні оператори $f : L \to L$.

Визначення 1.4 *Підпростір називається інваріантним, коли образ кожного вектора із цього підпростору знову лежить в цьому підпросторі.*

Приклади.

1. Оскільки образ нульового вектора є нульовим вектором, то нульовий підпростір є інваріантним. Оскільки образи всіх векторів містяться у цілому просторі, то весь лінійний простір є інваріантним. Таким чином, тривіальні лінійні підпростори є інваріантними.

2. Образ і ядро лінійного оператора є інваріантними підпросторами.

Інваріантні підпрости дозволяють будувати обмеження лінійного оператора на підпростір — ці два оператори (весь і обмеження) відрізняються лише областю визначення. Один діє на всьому просторі, а другий — на інваріантному підпросторі.

Теорема 1.4 *Сума інваріантних підпросторів є інваріантним підпростором. Перетин двох інваріантних підпросторів є інваріантний підпростір.*

Доведення. Нехай L_1, L_2 — два інваріантні підпростори для лінійного оператора f. Якщо $\vec{z} \in L_1 + L_2$ і $\vec{z} = \vec{x} + \vec{y}$ для деяких $\vec{x} \in L_1,\ \vec{y} \in L_2$, то

$$f(\vec{x}) \in L_1,\ f(\vec{y}) \in L_2\ \Rightarrow f(\vec{z}) = f(\vec{x}) + f(\vec{y}) \in L_1 + L_2.$$

Отже сума інваріантних підпросторів є інваріантним підпростором.

Нехай $\vec{x} \in L_1 \cap L_2$. Тоді $\vec{x} \in L_1$ і $f(\vec{x}) \in L_1$; $\vec{x} \in L_2$ і $f(\vec{x}) \in L_2$, отже $f(\vec{x}) \in L_1 \cap L_2$.

Доведення теореми завершене.

■

Визначення 1.5 *Нехай L — лінійний простір, і M — інваріантний для лінійного оператора f підпростір. Тоді лінійний оператор $g : M \to M$, що визначається умовою $g(\vec{x}) = f(\vec{x})$ для будь-якого вектора $\vec{x} \in M$ називається обмеженням f на M і позначається*

$$g = f|_M\,.$$

2 Алгебра лінійних операторів.

2.1 Додавання та множення лінійних операторів

Визначення 2.1 *Якщо f, g — два лінійні перетворення векторного простору L, то відображення $h_1 : L \to L$, яке визначається формулою*

$$h_1(\vec{x}) = f(\vec{x}) + g(\vec{x}),$$

є лінійним. Це перетворення називається сумою лінійних перетворень f та g і позначається через

$$h_1 = f + g.$$

Також лінійним буде перетворення $h_2 : L \to L$, що визначається формулою

$$h_2(\vec{x}) = f(g(\vec{x})).$$

Лінійне перетворення h_2 називається добутком (або композицією) перетворень f та g і позначається

$$h_2 = fg.$$

Приклад. Нехай лінійний простір L розкладений у пряму суму двох підпросторів L_1 та L_2, лінійний оператор f є проектуванням на L_2 паралельно L_1, а g є проектуванням на L_1 паралельно L_2. Тоді $fg = gf = 0$, а $f + g = \mathrm{id}$.

Визначення 2.2 *Якщо для лінійного оператора $f^2 = f$, то такий оператор називають ідемпотентним. Якщо $f^2 = \mathrm{id}$, то такий оператор називають інволютивним. Якщо для деякого натурального n виконується рівність $f^n = 0$, то такий оператор називають нільпотентним.*

Приклади.

1. Всі оператори проектування є ідемпотентними. Оператор обертання площини на 180 градусів є інволютивним.

2. Обмеження ідемпотентного оператора на образ є одиничним оператором. Це випливає із тотожності $f(f(\vec{x})) = f(\vec{x})$.

Теорема 2.1 *Множення лінійних перетворень асоціативне, але не комутативне. Одиничне перетворення грає роль одиниці при множенні, тобто для будь-якого оператора f виконуються рівності*

$$\mathrm{id} \cdot f = f \cdot \mathrm{id} = f. \tag{8}$$

Нульовий оператор грає роль нуля при множенні, тобто для будь-якого оператора f виконуються рівності

$$0 \cdot f = f \cdot 0 = 0. \tag{9}$$

Доведення. Множення відображень асоціативне — це вважаємо відомим. А множення лінійних операторів є окремим випадком множення відображень. Тому множення лінійних операторів асоціативне.

Коли стверджується, що множення лінійних операторів не комутативне, то мається на увазі, що існують лінійні оператори f, g такі, що $fg \neq gf$. Прикладом, який доводить некомутативність, є оператор f проектуання площини на пряму паралельно іншій прямій і оператор обертання цієї площини на кут $\dfrac{\pi}{2}$.

Рівності (8), (9) є простим наслідком із означень операторів id, 0 та множення операторів.

■

Нагадаємо, що множина разом з асоціативною бінарною операцією називаєтсья *напівгрупою.* Таким чином, лінійні перетворення заданого простору разом з операцією множення утворюють напівгрупу. Вони утворюють напівгрупу також з операцією додавання.

Теорема 2.2 *Додавання лінійних операторів асоціативне, комутативне. Серед лінійних операторів існує нульовий і для кожного лінійного оператора існує протилежний.*

Доведення. Наявність заявлених властивостей у лінійних операторів на даному лінійному просторі випливає із наявності відповідних властивостей в лінійному просторі, на якому діють оператори.

■

Нагадаємо, що множина з двома бінарними асоціативними операціями множення та додавання називається *кільцем*, якщо разом з додаванням ця множина утворює комутативну групу, а між собою додавання та множення зв'язані дистрибутивними законами. Теорема 2.2 дозволяє стверджувати, що *лінійні оператори на заданому лінійному просторі утворюють кільце.*

2.2 Множення лінійного оператора на число

Визначення 2.3 *Для лінійного оператора f на лінійному просторі L і числа $\lambda \in F$ відображення*

$$h_3 : L \to L,\ \vec{x} \mapsto h_3(\vec{x}) = \lambda f(\vec{x})$$

називається добутком f на число λ і позначається

$$h_3 = \lambda f.$$

Теорема 2.3 *Всі лінійні оператори, що діють в лінійному просторі L над полем F разом з операціями додавання та множення на число утворюють векторний простір* $\mathrm{End}\ L$ *вимірності n^2, де n — вимірність L.*

Якщо $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ — базис в L, то базисом векторного простору лінійних операторів на L можна взяти лінійні оператори

$$f_{ij} : L \to L,$$

що визначені умовами

$$f_{ij}(\vec{e}_k) = \begin{cases} \vec{e}_j, & \text{якщо } i = k, \\ 0, & \text{якщо } i \neq k. \end{cases} \tag{10}$$

Доведення. Щоб довести, що лінійні оператори із End L утворюють лінійний простір, потрібно згадати аксіоми лінійного простору і перевірити їх виконання в End L.

Аксіоми лінійного простору

1. Разом з додаванням End L утворює комутативну групу:
   $\forall f, g, h \in \text{End } L: \ f + (g + h) = (f + g) + h$ (асоціативність додавання);
   $\exists\ 0 \in \text{End } L\ \forall f \in \text{End } L: \ f + 0 = 0 + f = f$ (існування 0);
   $\forall f \in \text{End } L\ \exists(-f) \in \text{End } L: \ f + (-f) = (-f) + f = 0$ (кожен елемент має протилежний);
   $\forall f, g, \in \text{End } L: \ f + g = g + f$ (комутативність додавання).

2. Аксіоми множення на число:
   $\forall \lambda, \mu \in F\ \forall f \in \text{End } L: \ (\lambda\mu)f = \lambda(\mu f)$ (асоціативність);
   $\forall f \in \text{End } L: \ 1 \cdot f = f$ (унітарність).

3. Дистрибутивні закони:
   $\forall \lambda, \mu \in F\ \forall f \in \text{End } L: \ (\lambda + \mu)f = \lambda f + \mu f;$
   $\forall \lambda \in F\ \forall f, g \in \text{End } L: \ \lambda(f + g) = \lambda f + \lambda g.$

Всі аксіоми перевіряються з використанням відповідних властивостей операцій над векторами. Доведемо виконання однієї аксіоми, а саме: асоціативністі додавання лінійних операторів.

Нехай маємо три лінійні оператори: $f, g, h \in \text{End } L$.. Доведемо, що

$$f + (g + h) = (f + g) + h. \tag{11}$$

Два лінійні оператори збігаються, коли образи одного і того ж елемента під дією цих операторів збігаються. Отже, для того, щоб довести (11) необхідно і достатньо довести рівність

$$(f + (g + h))(\vec{x}) = ((f + g) + h)(\vec{x}). \tag{12}$$

для кожного вектора $\vec{x} \in L$. З використанням асоціативності додавання векторів, та з використанням означення додавання лінійних операторів рівність (12) перевіряється наступним чином:

$$(f + (g + h))(\vec{x}) = f(\vec{x}) + (g + h)(\vec{x}) = f(\vec{x}) + (g(\vec{x}) + h(\vec{x})) =$$
$$= (f(\vec{x}) + g(\vec{x})) + h(\vec{x}) = (f + g)(\vec{x}) + h(\vec{x}) = ((f + g) + h)(\vec{x}).$$

Решта аксім перевіряються подібним чином. Отже, вважаємо доведеним, що лінійні оператори утворюють лінійний простір.

Переходимо до обгрутування того, що лінійні оператори f_{ij} $(i, j = 1, 2, \ldots, n)$ є базисом простору лінійних операторів. Спочатку відмітимо, що лінійні оператори

f_{ij} визначені коректно, це нам забезпечує теорема 1.1. Для доведення того, що оператори f_{ij} утворюють базис, необхідно і достатньо довести повноту і лінійну незалежність обраної системи лінійних операторів.

Повнота означає, що будь-який лінійний оператор f є лінійною комбінацією вибраних лінійних операторів. Перевіримо це. Нехай

$$f(\vec{e}_i) = \sum_{j=1}^{n} a_{ij}\vec{e}_j.$$

Тоді

$$f(\vec{e}_i) = \sum_{j=1}^{n} a_{ij} f_{ij}(\vec{e}_i),$$

і

$$f = \sum_{j=1}^{n} a_{ij} f_{ij}.$$

Повнота перевірена.

Переходимо до доведення лінійної незалежності лінійних операторів f_{ij}. Виберемо довільну лінійну комбінацію

$$\sum_{i,j=1}^{n} a_{ij} f_{ij} \tag{13}$$

цих векторів, яка дорівнює нулю (нульовому лінійному оператору), і доведемо, що всі числа $a_{k,j}$ $k, j = 1, 2, \dots, n$, дорівнюють нулю.

Оскільки лінійний оператор (13) нульовий, то для будь-якого базисного вектора $\vec{e}_k$ буде

$$\sum_{i,j=1}^{n} a_{ij} f_{ij}(\vec{e}_k) = \sum_{j=1}^{n} a_{kj}\vec{e}_j = 0 \tag{14}$$

Оскільки базисні вектори обов'язково незалежні, то остання рівність можлива в тому і тільки тому випадку, коли всі числа a_{kj} $(\forall k, j = \overline{1, n})$ дорівнюють нулю.

Теорема доведена повністю.

■

Нагадаємо, що лінійний простір, який в той же час є кільцем, називаєть *лінійною алгеброю*. Отже многочлени утворюють лінійну алгебру, квадратні матриці заданого розміру утворюють лінійну алгебру. Ми довели, що також лінійні оператори утворюють і кільце і лінійний простір, отже маємо обгрунтованою теорему

Теорема 2.4 *Лінійні оператори, що діють в заданому лінійному просторі утворюють лінійну алгебру.*

2.3 Оператор, що обернений до даного

Визначення 2.4 *Коли для двох лінійних операторів f, g виконується рівність*

$$fg = gf = \mathrm{id},$$

то оператори f, g називають взаємно оберненими. (оператор g обернений до оператора f, а оператор f обернений до оператора g). В цьому випадку пишуть

$$f = g^{-1}, \qquad g = f^{-1}.$$

За означенням, якщо заданий оператор f має обернений оператор f^{-1}, то можна писати

$$\vec{y} = f(\vec{x}) \Leftrightarrow \vec{x} = f^{-1}(\vec{y}).$$

Теорема 2.5 *Лінійний оператор має обернений (тобто є оборотним) тоді і тільки тоді, коли він є бієктивним.*

Доведення. Будь-яке відображення має обернене тоді і тільки тоді, коли воно бієктивне — цей факт вважаємо відомим (наприклад, з курсу математичного аналізу). Отже, коли оператор має обернений, то він є бієктивним. Потрібно перевіряти лише, що коли він бієктивний, то в такому випадку він має обернений (тобто його обереренене відображення є лінійним).

Нехай $f : L \to L$ є бієктивним лінійним оператором, тобто бієктивним відображенням, і $g : L \to L$ є оберненим відображенням. Побрібно перевірити, що g є лінійним оператором, тобто виконання умов

$$\forall \vec{a}, \vec{b} \in L \; \forall \lambda \in F : \; g(\vec{a} + \vec{b}) = g(\vec{a}) + g(\vec{b}), \quad g(\lambda\vec{a}) = \lambda g(\vec{a}). \tag{15}$$

Виберемо $\vec{a}, \vec{b} \in L$ і $\lambda \in F$. Оскільки оператор f бієктивний, то для деяких векторів $\vec{a'}, \vec{b'} \in L$ будуть виконуватися рівності

$$f(\vec{a'}) = \vec{a}, \; f(\lambda\vec{a'}) = \lambda\vec{a}, \; f(\vec{b'}) = \vec{b}, \; f(\vec{a'} + \vec{b'}) = \vec{a} + \vec{b},$$

і, відповідно, рівності

$$\vec{a'} = g(\vec{a}), \quad \lambda\vec{a'} = g(\lambda\vec{a}), \quad \vec{b'} = g(\vec{b}), \quad \vec{a'} + \vec{b'} = g(\vec{a} + \vec{b}).$$

Грунтуючись на останніх рівностях робимо висновок про правильність (15).

■

Теорема 2.6 *Лінійний оператор $f : L \to L$ скінченновимірного пространу L має обернений тоді і тільки тоді, коли виконується одна із рівностей:*

$$\mathrm{rang}\, f = \dim L, \qquad \mathrm{def}\, f = 0.$$

Доведення.

Згідно з теоремою 1.3 умови $\operatorname{rang} f = \dim L$ та $\operatorname{def} f = 0$ рівносильні. А умова $\operatorname{rang} f = \dim L$ рівносильна умові бієктивності, яка рівносильна умові існування оберненого оператора згідно з теоремою 2.5.

∎

Приклад. Оберненим до оператора обертання площини є обертання на протилежний кут. Оберненим до оператора множення на число є оператор множення на обернене число.

Теорема 2.7 *Оборотні оператори (ті, що мають обернені) лінійного простору утворюють мультиплікативну групу.*

Доведення. Нагадаємо, що мультиплікативною групою називають непорожню множину разом з бінарною асоціативною операцією, яка названа множенням, причому ця множина містить одиницю і кожен елемент має обернений. Оскільки множення операторів асоціативне, одиничний оператор має обернений, і кожен оборотний оператор за означенням має обернений, то доводити потрібно лише те, що добуток оборотних операторів є оборотним оператором.

Нехай f, g — два оборотні оператори і f^{-1}, g^{-1} — обернені до цих операторів. Тоді

$$(f \cdot g) \cdot (g^{-1} \cdot f^{-1}) = f \cdot (g \cdot g^{-1}) \cdot f^{-1} = \mathrm{id}\,.$$

Отже оберненим до оператора $f \cdot g$ буде оператор $g^{-1} \cdot f^{-1}$.

Доведення теореми завершене.

∎

3 Матриця лінійного оператора

3.1 Означення матриці лінійного оператора

Визначення 3.1 *Матрицею лінійного оператора $f : L \to L$ в заданному базисі $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\} \in L$ називаєтся матриця, стовпчики якої утворені координатами образів базисних векторів при відображенні f. Тобто коли*

$$f(\vec{e}_1) = a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + \ldots + a_{n1}\vec{e}_n = \{a_{11}, a_{21}, \ldots, a_{n1}\},$$

$$f(\vec{e}_2) = a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + \ldots + a_{n2}\vec{e}_n = \{a_{12}, a_{22}, \ldots, a_{n2}\},$$

$$\ldots$$

$$f(\vec{e}_n) = a_{1n}\vec{e}_1 + a_{2n}\vec{e}_2 + \ldots + a_{nn}\vec{e}_n = \{a_{1n}, a_{2n}, \ldots, a_{nn}\},$$

то матрицею $A = A_f$ оператора f буде матриця

$$A = A_f = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Приклади.

1. Матрицею одиничного оператора є одинична матриця.

2. Матрицею нульового оператора є нульова матриця.

3. Матрицею обертання площини на заданий кут φ є матриця

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

4. Нехай L лінійний простір з базисом $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$ і $f : L \to L$ — лінійний оператор, для якого

$$f(\vec{e}_1) = f(\vec{e}_2) = 7\vec{e}_1 - \vec{e}_2 + 5\vec{e}_4, \quad f(\vec{e}_3) = 2\vec{e}_2 + \vec{e}_3 - 6\vec{e}_4, \quad f(\vec{e}_4) = \vec{e}_1 + \vec{e}_2 + \vec{e}_3.$$

Матрицею оператора f в базисі $\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4$ буде

$$\begin{pmatrix} 7 & 7 & 0 & 1 \\ -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 5 & 5 & -6 & 0 \end{pmatrix}$$

Теорема 3.1 *Коли в лінійному просторі L вибрано базис, матриця A є матрицею лінійного оператора $f : L \to L$ в цьому базисі, вектори $\vec{x}$, $\vec{y}$ мають координати $\vec{x} = \{x_1, x_2, \dots, x_n\}$, $\vec{y} = \{y_1, y_2, \dots, y_n\}$, то*

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Leftrightarrow \vec{y} = f(\vec{x}). \tag{16}$$

Доведення. Нехай $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ — базис L,

$$f(\vec{e}_i) = \sum_{j=1}^{n} a_{ji}\vec{e}_j,$$

і, відповідно,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Нехай також

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i, \quad \vec{y} = \sum_{i=1}^{n} y_i \vec{e}_i,$$

Тоді

$$\vec{y} = f(\vec{x}) = \sum_{i=1}^{n} x_i f(\vec{e}_i) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ji}\vec{e}_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} x_i \right) \vec{e}_j \Leftrightarrow y_j = \sum_{i=1}^{n} a_{ji} x_i (j = \overline{1, n}).$$

■

Теорема 3.2 *Коли в $n-$вимірному лінійному просторі є вибраний базис, то відображення, що ставить у відповідність кожному оператору його матрицю, є ізоморфізмом між алгеброю лінійних операторів і алгеброю квадратних матриць розміру $n \times n$, тобто*

- *це відображення є бієктивним;*
- *сумі двох операторів ставиться у відповідність сума матриць цих операторів;*
- *добутку двох операторів ставиться у відповідність добуток матриць цих операторів;*
- *нульовому оператору ставиться у відповідність нульова матриця;*
- *одиничному оператору ставиться у відповідність одинична матриця;*
- *протилежному оператору ставиться у відповідність протилежна матриця;*
- *оберненому оператору ставиться у відповідність обернена матриця.*

Доведення. Кроки доведення однотипні, тому зупинимося лише на одному пункті: добутку операторів ставиться у відповідність добуток відповідних матриць.

Нехай $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ базис лінійного простору L, нехай в цьому базисі лінійні оператори f, g, $h = f \cdot g$ мають матриці A, B, C

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix}, C = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix}$$

За означенням матриці лінійного оператора

$$f(\vec{e}_i) = \sum_{j=1}^{n} a_{ji}\vec{e}_j, \quad g(\vec{e}_i) = \sum_{j=1}^{n} b_{ji}\vec{e}_j, \quad h(\vec{e}_i) = \sum_{k=1}^{n} c_{ki}\vec{e}_k.$$

За означенням добутку двох операторів

$$h(\vec{e}_i) = f(g(\vec{e}_i)) = f\left(\sum_{j=1}^{n} b_{ji}\vec{e}_j\right) = \sum_{j=1}^{n} b_{ji} f(\vec{e}_j) = \sum_{j=1}^{n} b_{ji} \sum_{k=1}^{n} a_{kj}\vec{e}_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{kj}b_{ji}\right)\vec{e}_k$$

Звідси випливає, що $c_{ki} = \sum\limits_{j=1}^{n} a_{kj}b_{ji}$, тобто $C = A \cdot B$.

■

Теорема 3.3 *Матриця A лінійного оператора в лінійному просторі L вимірності n в деякому базисі має блочний вигляд*

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

де матриця A_1 має розмір $n_1 \times n_1$, а матриця A_2 має розмір $n_2 \times n_2$, $n_1 + n_2 = n$, тоді і тільки тоді, коли і лінійна оболонка перших n_1 базисних векторів і лінійна оболонка останніх n_2 векторів є інваріантними підпросторами.

Доведення. Теорема є прямим наслідком означення матриці лінійного оператора та означення інваріантного підпростору.

■

3.2 Зміна матриці лінійного оператора при переході до нового базису

Теорема 3.4 *Нехай в лінійному просторі L є два базиси: $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ — старий базис та $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ — новий базис, C — матриця переходу від старого базису до нового. Нехай також лінійний оператор f має матрицю A в старому базисі і матрицю B в новому базисі. Тоді*

$$B = C^{-1}AC.$$

Доведення. Нехай матриці A, B, C мають такі елементи

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix},\ B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix},\ C = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ c_{21} & c_{22} & \ldots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & c_{nn} \end{pmatrix}$$

За означенням матриці лінійного оператора

$$f(\vec{e}_j) = \sum_{k=1}^{n} a_{kj}\vec{e}_k, \quad f(\vec{u}_i) = \sum_{j=1}^{n} b_{ji}\vec{u}_j.$$

А за означенням матриці переходу від старого базису до нового

$$\vec{u}_i = \sum_{j=1}^{n} c_{ji}\vec{e}_j.$$

Тоді

$$f(\sum_{j=1}^{n} c_{ji}\vec{e}_j) = \sum_{j=1}^{n} b_{ji}\left(\sum_{k=1}^{n} c_{kj}\vec{e}_k\right), \quad \sum_{j=1}^{n} c_{ji} f(\vec{e}_j) = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} c_{kj}b_{ji}\right)\vec{e}_k,$$

$$\sum_{j=1}^{n} c_{ji}\sum_{k=1}^{n} a_{kj}\vec{e}_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} c_{kj}b_{ji}\right)\vec{e}_k, \quad \sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{kj}c_{ji}\right)\vec{e}_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} c_{kj}b_{ji}\right)\vec{e}_k,$$

$$\sum_{j=1}^{n} a_{kj}c_{ji} = \sum_{j=1}^{n} c_{kj}b_{ji}, \quad k, j = 1, 2, \ldots, n.$$

Останні рівності означають $AC = CB$, що й потрібно було довести.

■

4 Діагоналізація матриці лінійного оператора

4.1 Власні числа та власні вектори лінійного оператора

У розділі, як і раніше, через L позначаємо лінійний простір над полем F, а через f лінійний оператор в цьому просторі.

Визначення 4.1 *Число $\lambda \in F$ называаєтъся власним числом оператора f, якщо для деякого ненульового вектора $\vec{x} \in L$ виконуеъся рівність*

$$f(\vec{x}) = \lambda\vec{x}. \tag{17}$$

Визначення 4.2 *Вектор $\vec{x} \in L$ називаєтъся власним вектором лінійного оператора f, що відповідає власному числу λ, якщо виконуєтъся рівність (17).*

Підкреслимо, що власний вектор визначається після власного числа, і власне число може бути нульовим. Нульовий вектор є власним вектором для будь-якого власного числа. А при визначенні власного числа вимагається, щоб існував ненульовий власний вектор, який цьому числу відповідає. Порядок визначень тут не можна змінювати.

Приклад. Кожну квадратну матрицю розміру $n \times n$ можна розглядати як лінійний оператор у просторі стовпчиків довжини n: якщо A матриця, а $\vec{x}$ — стовчик елементів того поля, якому належать елементи матриці A, то

$$\vec{x} \mapsto A \cdot \vec{x}.$$

Природно, власні числа (відповідно, власні вектори) цього оператора називають власними числами (відповідно, власними векторами) матриці A.

Отже, число λ є *власним числом матриці* A тоді і тільки тоді (за означенням), коли для деякого ненульового вектора (стовпчика елементів відповідного поля) виконується рівність

$$A \cdot \vec{x} = \lambda \vec{x}. \tag{18}$$

А вектор $\vec{x}$, що задовольняє рівнянню (18) є *власним вектором матриці* A, що відповідає власному числу λ матриці A.

Теорема 4.1 *Якщо в лінійному просторі вибрали базис і лінійному оператору f поставили у відповідність матрицю A, то власні вектори і власні числа матриці A і лінійного оператора f збігаються.*

Доведення. Нехай вектор $\vec{x}$ має координати $\vec{x} = \{x_1, x_2, \ldots, x_n\}$. Тоді (див. (16)) маємо

$$f(\vec{x}) = \lambda \cdot \vec{x} \;\Leftrightarrow\; A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

∎

Приклади.

1. Розглянемо лінійний простір дійсних функцій, що мають похідні всіх порядків у всіх точках інтервалу (0,1), і лінійний оператор в цьому просторі, що ставить у відповідність функції її похідну. Тоді власними числами будуть ті дійсні числа λ, для яких існує ненульова функція f така, що

$$f' = \lambda f. \tag{19}$$

Оскільки $\left(e^{\lambda x}\right)' = \lambda e^{\lambda x}$, то всі дійсні числа є власними для цього оператора. Оскільки розв'язками рівняння (19) для заданого λ є функції $C \cdot e^{\lambda x}$, то власними векторами введеного оператора знаходження похідної, що відповідають власному числу λ, будуть функції $C \cdot e^{\lambda x}$, де C — стале число.

2. Розглянемо оператор проектування на підпростір L_2 паралельно підпростору L_1. Якщо L_1, L_2 є ненульовими підпросторами, то власними числами є 0 та 1, причому, власними векторами, які відповідають власному числу 0 будуть вектори із L_1, а власними числами, які відповідають власном числу 1 будуть вектори із L_2.

Визначення 4.3 *Характеристичним многочленом матриці A називають многочлен*

$$\chi(\lambda) = \det(A - \lambda E).$$

Характеристичним многочленом лінійного оператора називають характеристичний многочлен його матриці в якомусь базисі.

Визначення характеристичного многочлена лінійного оператора коректне, тобто характеристичний многочлен матриці лінійного оператора не залежить від вибору базису. Дійсно, нехай є два базиси — новий і старий. В новому оператор має матрицю B, а в старому — матрицю A, нехай C — матриця переходу від старого базиса до нового. Тоді

$$\det(B - \lambda E) = \det(C^{-1}AC - \lambda C^{-1}EC) = \det(C^{-1}(A - \lambda E)C) =$$
$$= \det C^{-1} \cdot \det(A - \lambda E) \cdot \det C = \det(A - \lambda E).$$

Отже характеристичні многочлени збігаються і означення коректне.

Теорема 4.2 *Число λ є власним для лінійного оператора тоді і тільки тоді, коли воно є коренем характеристичного многочлена $\chi(\lambda) = \det(A - \lambda E)$ матриці A цього оператора в деякому базисі.*

Доведення. Твердження теореми випливає з того, що рівняння $A\vec{x} = \lambda \cdot \vec{x}$ має ненульовий розв'язок тоді і тільки тоді, коли має ненульовий розв'зок система лінійних рівнянь

$$A\vec{x} - \lambda\vec{x} = (A - \lambda E)\vec{x} = \vec{0},$$

а останнє рівняння має ненульовий розв'язок тоді і тільки тоді, коли визначник системи дорівнює нулю (або ранг системи менше кількості невідомих).

■

Теорема 4.3 *Всі власні вектори, що відповідають заданому власному числу λ, утворюють ненульовий інваріантний підпростір.*

Доведення. Всі власні вектори лінійного оператора $f : L \to L$, що відповідають власному числу λ, утворюють ядро оператора $f - \lambda \cdot \mathrm{id}$ і тому утворюють підпростір всього простору. Цей підпростір ненульовий тому, що за означенням власного числа, повинен бути бодай один ненульовий власний вектор, який цьому числу відповідає.

■

Теорема 4.4 *Сума власних підпросторів, що відповідають різним власним числа, є прямою сумою.*

Доведення. Нехай L — лінійний простір, $L_1, L_2, \ldots, L_k \subseteq L$, є власними підпросторами. L_i відповідає власному числу λ_i ($i = 1, 2, \ldots, k$) причому всі власні числа різні. Нагадаємо, що сума

$$L_1 + L_2 + \ldots + L_k$$

є прямою, якщо для кожного $i = 1, 2, \ldots, k$

$$\left(\sum_{j \neq i, j=1}^{k} L_j \right) \cap L_i = 0.$$

Оскільки нумерація власних підпросторів довільна, то можна доводити лише рівність

$$(L_1 + L_2 + \ldots + L_{k-1}) \cap L_k = 0.$$

Припустимо протилежне: цей перетин непустий. Тоді існують вектори $\vec{x}_i \in L_i$ для кожного $i = 1, 2, \ldots, k$, серед яких є бодай один ненульовий, такі, що виконується рівність:

$$\vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_k = \vec{0}. \tag{20}$$

Подієио на рівність (20) операторами $f, \ f^2, \ f^3, \ldots, \ f^{k-1}$. Одержимо систему рівностей

$$\begin{aligned} \vec{x}_1 + \vec{x}_2 + \ldots + \vec{x}_k &= \vec{0}, \\ \lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2 + \ldots + \lambda_k \vec{x}_k &= \vec{0}, \\ \lambda_1^2 \vec{x}_1 + \lambda_2^2 \vec{x}_2 + \ldots + \lambda_k^2 \vec{x}_k &= \vec{0}, \\ &\ldots \\ \lambda_1^{k-1} \vec{x}_1 + \lambda_2^{\vec{k}-1} \vec{x}_2 + \ldots + \lambda_k^{k-1} \vec{x}_k &= \vec{0}, \end{aligned}$$

яку можна записати в матричному вигляді $W \cdot X = O$, де O — нульова матриця розміру $k \times n$,

$$W = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_k \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \ldots & \lambda_k^{k-1} \end{pmatrix}_{k \times k}, \qquad X = \begin{pmatrix} \vec{x}_1 \\ \vec{x}_2 \\ \vdots \\ \vec{x}_k \end{pmatrix}_{k \times n}$$

(i-та строка матриці X — це координати вектора $\vec{x}_i$). Оскільки мариця W — це матриця Вандермонда, і всі власні числа λ_i різні, то ця матриця має обернену W^{-1} — вважаємо це відомим. Домноживши зліва рівність $W \cdot X = O$ на W^{-1} одержуємо $X = W^{-1} \cdot W \cdot X = W^{-1} \cdot O = O$. Тобто всі вектори $\vec{x}_i$ нульові. Протиріччя — теорема доведена.

■

Приклад. Знайдемо власні числа та власні підпростори лінійного оператора, що заданий матрицею

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}.$$

Спочатку виписуємо характеристичний многочлен:

$$\chi(\lambda) = \begin{vmatrix} 2 - \lambda & 3 \\ 4 & 6 - \lambda \end{vmatrix} = \lambda^2 - 8\lambda.$$

Потім шукаємо корені характеристичного многочлена:

$$\lambda_1 = 0, \qquad \lambda_2 = 8.$$

Отже власними числами лінійного оператора є числа 0, 8.

Далі для кожного власного числа шукаємо власний підпростір, що відповідає цьому числу. Знайти власний підпростір означає знайти базис цього підпростору.

Для власного числа $\lambda_1 = 0$ відпровідним власним підпростором є підпростір розв'язків рівняння

$$A\vec{x} = \vec{0}$$

Базис цього простору складається із одного вектора — ним можна взяти

$$\vec{e}_1 = \{-3, 2\}.$$

Для власного числа $\lambda_2 = 8$ відпровідним власним підпростором є підпростір розв'язків рівняння

$$(A - 8E)\vec{x} = \vec{0}, \quad \begin{pmatrix} -6 & 3 \\ 4 & -2 \end{pmatrix} \vec{x} = \vec{0}.$$

Базис цього простору складається із одного вектора — ним можна взяти

$$\vec{e}_2 = \{1, 2\}.$$

В базисі $\vec{e}_1,\ \vec{e}_2$ матриця B лінійного оператора має вигляд

$$B = C^{-1}AC,$$

де C — це матриця переходу від старого базису $\{1, 0\},\ \{0, 1\}$ до нового базису $\vec{e}_1,\ \vec{e}_2$, тобто

$$C = \begin{pmatrix} -3 & 1 \\ 2 & 2 \end{pmatrix}, \quad C^{-1} = -\frac{1}{8} \begin{pmatrix} 2 & -1 \\ -2 & -3 \end{pmatrix}$$

Отже

$$B = C^{-1}AC = -\frac{1}{8} \begin{pmatrix} 2 & -1 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 8 \end{pmatrix}.$$

4.2 Діагоналізація матриці лінійного оператора.

Теорема 4.5 *Матриця лінійного оператора має діагональный вид тоді і тільки тоді, коли базис лінійного простору складається із власних векторів.*

Доведення. Доведення однокрокове — виписуємо матрицю лінійного оператора у базисі із власних векторів і бачимо, що вона діагональна (доведення "туди"). Потім виписуємо дію оператора на базисні вектори у випадку, коли матриця діагональна і бачимо, що базис складається виключно із власних векторів (доведення "назад").

∎

Визначення 4.4 *Діагоналізувати матрицю лінійного оператора означає знайти базис, в якому матриця цього оператора має діагональний вигляд.*

Діагоналізувати матрицю A означає знайти оборотну матрицю C таку, що матриця $C^{-1}AC$ має діагональний вигляд.

Вправа. Діагоналізувати матрицю лінійного оператора, що в деякому базисі має матрицю

$$A = \begin{pmatrix} 10 & 18 \\ -6 & -11 \end{pmatrix}.$$

Знайти власні числа та власні підпростори.

Відповідь: в базисі $\vec{e}_1 = \{2, \ -1\}$, $\vec{e}_2 = \{3, \ -2\}$ лінійний оператор має матрицю

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

Власні числа: $\lambda_1 = 1, \ \lambda_2 = -2$. Власні підпростори: $L_1 = \{\vec{e}_1\}, \ L_2 = \{\vec{e}_2\}$.

Зауважимо, що не кожну матрицю можливо діагоналізувати.

Приклади.

1. Спробуємо діагоналізувати матрицю, яка має елементи в полі дійсних чисел:

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

Виписуємо характеристичний многочлен:

$$\chi(\lambda) = \begin{vmatrix} 1-\lambda & 3 \\ -3 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 9.$$

Цей многочлен не має дійсних коренів, матриця A не має дійсних власних чисел, і відповідно, власних підпросторів. Отже, не існує оборотної матриці C з дійсними елементами такої, що матриця $C^{-1}AC$ має діагональний вигляд, тобто в полі дійсних чисел дану матрицю A діагоналізувати не можливо.

Перевірте самостійно, що над полем комплексних чисел існує матриця

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

така, що

$$C^{-1}AC = \begin{pmatrix} 1+3i & 0 \\ 0 & 1-3i \end{pmatrix}.$$

2. Знайдемо власні числа та власні підпростори матриці

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

Виписуємо характеристичний многочлен:

$$\chi(\lambda) = \begin{vmatrix} 2-\lambda & 0 \\ 3 & 2-\lambda \end{vmatrix} = (2-\lambda)^2.$$

Шукаємо корені характеристичного многочлена:

$$\lambda_1 = \lambda_2 = 2.$$

Отже матриця має єдине власне число 2.

Відповідним власним підпростором є підпростір розв'язків рівняння

$$(A - 2E)\vec{x} = \vec{0}, \quad \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix} \vec{x} = \vec{0}.$$

Базис власного підпростору складається із одного вектора — ним можна взяти

$$\vec{e}_1 = \{0, 1\}.$$

Отже, весь двовимірний простір не можливо представити у вигляді суми власних підпросторів і матрицю A не можливо діагоналізувати ні в полі дійсних, ні в полі комплексних чисел.

4.3 Анулюючий та мінімальний многочлен

Визначення 4.5 *Многочлени, які мають своїм коренем заданний оператор (відповідно, матрицю), називають анулюючими цей оператор (відповідно, цю матрицю).*

Приклад. Оскільки $A^2 = 0$ для матриці $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, то многочлен x^2 є анулюючим для матриці A.

Теорема 4.6 *Для кожного оператора (відповідно, матриці) є ненульовий анулюючий многочлен. Серед анулюючих многочленів є многочлен найменшого степеня (один з точністю до сталого множника).*

Доведення. Відомо (див. теорему 2.3), що вимірність лінійного простору лінійних операторів, що діють в заданому скінченновимірному просторі, скінченна. Тому для заданого оператора f певна кількість степенів $f, f^2, f^3, \ldots, f^n$ буде лінійно залежною, тобто

$$a_1 f + a_2 f^2 + \ldots + a_n f^n = 0,$$

причому серед коефіцієнтів $a_1, a_2, \ldots, a_n$ є ненульовий. А це означає, що многочлен

$$a_1 x + a_2 x^2 + \ldots + a_n x^n$$

є анулюючим для оператора f.

Існування ненульових анулючих многочленів доведене.

Оскільки степені ненульових многочленів є нуль або натуральне число, то за принципом найменшого числа серед ненульових анулюючих многочленів є многочлен найменшого степеня. Доведемо, що такий многочлен єдиний з точністю до сталого множника.

Нехай $u(x), v(x)$ два ненульових анулюючих многочлени найменшого степеня для заданого оператора f. Розділимо $u(x)$ на $v(x)$ з остачею:

$$u(x) = v(x) \cdot q(x) + r(x), \qquad \deg r(x) < \deg v(x).$$

Оскільки

$$u(f) = v(f) \cdot q(f) + r(f),$$

і $u(f) = v(f) = 0$, то $r(f) = 0$, тобто $r(x)$ є анулюючим для f. Але $\deg r(x) < \deg v(x)$, тому $r(x) = 0$ і многочлен u ділиться на многочлен v.

■

Визначення 4.6 *Ненульовий анулюючий многочлен найменшого степеня називають мінімальним.*

Теорема 4.7 *Всі анулюючі многочлени ділятъся на мінімальний. Мінімальний многочлен єдиний з точністю до сталого множника.*

Доведення. Нехай $u(x)$ є анулюючим многочленом, а $m(x)$ є мінімальним многочленом для лінійного оператора f.

Розділимо $u(x)$ на $m(x)$ з остачею:

$$u(x) = m(x) \cdot q(x) + r(x), \qquad \deg r(x) < \deg m(x).$$

Оскільки

$$u(f) = m(f) \cdot q(f) + r(f),$$

і $u(f) = m(f) = 0$, то $r(f) = 0$, тобто $r(x)$ є анулюючим для f. Але $\deg r(x) < \deg m(x)$, тому $r(x) = 0$ і многочлен u ділиться на многочлен m.

■

Теорема 4.8 (Теорема Гамільтона-Келі) *Кожна матриця і кожен лінійний оператор є коренем свого характеристичного многочлена.*

Доведення. З огляду на ізоморфізм між алгеброю операторів і алгеброю матриць теорему можна доводити лише для матриць.

Матриця A є коренем многочлена $f(x) = a_0 + a_1x + \ldots + a_nx^n$ (за визначенням) коли

$$a_0E + a_1A + a_2A^2 + \ldots + a_nA^n = 0.$$

Нехай у нас є матриця A розміру $n \times n$ над деяким полем F. Нагадаємо, що визначення многочленів у нас алгебраїчне: многочлени — це вирази певного вигляду, і вони збігаються лише у випадку, коли всі коефіцієнти у них збігаються. Будуємо матрицю $A - \lambda E$ і називаємо її характеристичною. До характеристичної матриці будуємо приєднану матрицю B — її елементами є алгебраїчні доповнення певних елементів характеристичної матриці. Точніше, якщо

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

то b_{ij} — це мінор, що одержаний із характеристичної матриці викреслюванням $j-$го рядка і $i-$го стовпчика і помноженого на $(-1)^{i+j}$. Досить очевидно, що елементи матриці B є многочленами від λ степені, що не перевищує $n-1$. Тому можна написати

$$b_{ij} = b_{ij}^{(0)} + b_{ij}^{(1)}\lambda + b_{ij}^{(2)}\lambda^2 + \ldots + b_{ij}^{(n-1)}\lambda^{n-1}.$$

В таких позначеннях матриця B розкладається в суму

$$B = \sum_{k=0}^{n-1} \lambda^k \cdot \begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} & \dots & b_{1n}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} & \dots & b_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1}^{(k)} & b_{n2}^{(k)} & \dots & b_{nn}^{(k)} \end{pmatrix},$$

або

$$B = \sum_{k=0}^{n-1} \lambda^k \cdot B^{(k)}, \text{ де } B^{(k)} = \begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} & \dots & b_{1n}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} & \dots & b_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1}^{(k)} & b_{n2}^{(k)} & \dots & b_{nn}^{(k)} \end{pmatrix}.$$

Далі будемо користуватися відомою рівністю

$$B(A - \lambda E) = \det(A - \lambda E) \cdot E. \tag{21}$$

Характеристичний многочлен часто позначають χ, і ми так зробимо:

$$\chi(\lambda) = \det(A - \lambda \cdot E) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \lambda^n.$$

Вважаємо зрозумілим, що характеристичний многочлен має степінь n і старший коефіцієнт в ньому дорівнює 1.

Далі переписуємо рівняння (21) в нових позначеннях:

$$\sum_{k=0}^{n-1} \lambda^k \cdot B^{(k)}(A - \lambda E) = (\sum_{k=0}^{n} \lambda^k \alpha_k) \cdot E, \quad \alpha_n = 1.$$

Остання рівність виконується, коли збігаються коефіцієнти при степенях λ зліва і справа. Хоч тут і не многочлени над полем чи комутативним кільцем з одиницею, як ми їх визначали, все ж так можна стверджувати — ми можемо перейти на рівень окремих елементів і зовсім строго прослідкувати за правильністю такого твердження.

Отже прирівнюємо коефіцієнти зліва і справа.

$$\begin{array}{rcrcl}
-B^{(0)}A & & & = & \alpha_0 E, \\
-B^{(1)}A & + & B^{(0)} & = & \alpha_1 E, \\
-B^{(2)}A & + & B^{(1)} & = & \alpha_2 E, \\
\dots & & \dots & & \dots \\
-B^{(n-1)}A & + & B^{(n-2)} & = & \alpha_{n-1} E, \\
 & & B^{(n-1)} & = & E.
\end{array}$$

Ми виписали $n+1$ матричну рівність. Помножимо їх справа відповідно на $E, A, A^2, \dots, A^n$ і складемо. Зліва одержимо нульову матрицю, а справа характеристичний многочлен від матриці. Тобто ми одержали, що матриця є коренем свого характеристичного многочлена.

Теорема Гамільтона-Келі доведена.

■

5 Жорданова форма матриці лінійного оператора

Не кожну матрицю можна привести до діагонального вигляду (знайти базис із власних векторів). Але для кожної матриці A існує матриця переходу C така, що матриця $B = C^{-1}AC$ має блочний вигляд, в якому по головній діагоналі стоять клітини Жордана, а за межами діагоналі — нульові клітини. Кажуть, що так побудована матриця B має форму Жордана.

Визначення 5.1 *Клітини Жордана $J(\lambda)$ — це матриці, в яких по діагоналі стоїть одне число (власне), під діагоналлю йде рядок одиничок, а решта елементів – нулі, тобто*

$$J(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & \dots & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

5.1 Розкладення лінійного простору у пряму суму кореневих підпросторів.

В розділі будемо стало використовувати позначення:

- L — лінійний простір над полем $\mathbb{C}$ комплексних чисел;
- $\dim L = n \geq 1$; $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ — базис простору L;
- I — одиничний оператор, $I(\vec{x}) = \vec{x}$ для будь-якого вектора $\vec{x} \in L$;
- $f : L \to L$ — лінійний оператор;
- $\chi(\lambda) = (\lambda - \lambda_1)^{p_1}(\lambda - \lambda_2)^{p_2} \dots (\lambda - \lambda_k)^{p_k}$ — характеристичний многочлен оператора f;
- $\lambda_1, \lambda_2, \dots \lambda_k$ — різні власні числа оператора f;
- A — матриця оператора f базисі $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$.
- $\ker f$ — ядро оператора f, воно складається із тих векторів, які переводяться оператором f в $\vec{0}$.
- $\operatorname{im} f$ — образ лінійного простору під дією оператора f. Він складається із тих векторів $\vec{y} \in L$ для яких можна підібрати такий вектор $\vec{x} \in L$, що буде виконуватися рівність $\vec{y} = f(\vec{x})$.

Нагадаємо, що обмеження заданого оператора на інваріантний підпростір це новий оператор, який відрізняється від заданого лише областю визначення — він визначений лише на інваріантному підпросторі, але на векторах інваріантного підпростору діє так же, як і заданий.

Без нагадування будемо використовувати два факти:

1) добуток двох многочленів від одного оператора не залежить від порядку множників;

2) якщо $\vec{x}$ — власний вектор оператора f, що відповідає власному числу λ, і $\alpha(f)$ — многочлен від f, то $\vec{x}$ буде також власним вектором оператора $\alpha(f)$, який відповідає власному числу $\alpha(\lambda)$.

Визначення 5.2 *Кореневим підпростором, що відповідає власному значенню λ_i $(i = 1, 2, \ldots, k)$ називається підпростір $V_i \subseteq L$, елементами якого є ті вектори, які переводяться в $\vec{0}$ певним степенем оператора $f - \lambda_i I$.*

Іншими словами, кореневий підпростір V_i, що відповідає власному числу λ_i, це об'єднання ядер операторів $(f - \lambda_i I)^p$ при всіх $p = 1, 2, \ldots$

Визначення коректне, кореневий підпростір дійсно є підпростором. Для перевірки цього факту беремо два вектори $\vec{x}, \vec{y} \in V_i$, $a, b \in \mathbb{C}$. Потрібно переконатися, що $\vec{z} = a\vec{x} + b\vec{y} \in V_i$.

Дійсно, оскільки $\vec{x}, \vec{y} \in V_i$, то для деяких натуральних чисел $n_1 \geq n_2$,буде

$$(f - \lambda_i I)^{n_1}(\vec{x}) = (f - \lambda_i I)^{n_2}(\vec{y}) = \vec{0}.$$

Тоді

$$(f - \lambda_i I)^{n_1}(\vec{z}) = a(f - \lambda_i I)^{n_1}(\vec{x}) + b(f - \lambda_i I)^{n_2}(\vec{y}) = \vec{0}.$$

Розділ присвячений доведенню наступної теореми:

Теорема 5.1 *Лінійний простір є прямою сумою кореневих підпросторів, тобто*

$$L = V_1 \oplus V_2 \oplus \ldots \oplus V_k.$$

Спочатку одержимо декілька допоміжних результатів.

Лема 5.1 *Кожен кореневий підпростір є ненульовим інваріантним підпростором.*

Для доведення потрібно показати, що для вектора $\vec{x} \in V_i$ також $f(\vec{x}) \in V_i$. А це випливає з того, що коли $(f - \lambda_i I)^p(\vec{x}) = \vec{0}$, то

$$(f - \lambda_i I)^p(f(\vec{x})) = f((f - \lambda_i I)^p(\vec{x})) = f(\vec{0}) = \vec{0}.$$

Лема 5.2 *Для кожного кореневого підпростору V_i існує цілком певне число q_i таке, що*

$$V_i = \ker(f - \lambda_i I)^{q_i} = \{\vec{v} \in L|\ (f - \lambda_i I)^{q_i}(\vec{v}) = \vec{0}\}.$$

Доведення. Використовуємо інваріантність підпростору V_i. В цьому підпросторі оператор f має єдине власне число — λ_i. Обмеження f_i оператора f на підпростір V_i має характеристичний многочлен $(\lambda - \lambda_i I)^{q_i}$. А характеристичний многочлен є анулюючим. Тому оператор $(f_1 - \lambda_i I)^{q_i}$ є нульовим, а оператор $(f - \lambda_i I)^{q_i}$ переводить всі вектори підпростору V_i в нуль. Лема доведена.

■

Лема 5.3 *Нехай кореневий підпростір V_i є ядром оператора $(f - \lambda_i I)^{q_i}$. Тоді лінійний простір L є прямою сумою двох інваріантних підпросторів $V = V_i \oplus V_i'$, де $V_i' = \mathrm{im}(f - \lambda_i I)^{q_i}$.*

Доведення. Теорема 1.3 стверджує, що сума вимірностей ядра і образа оператора дорівнює вимірності усього простору.

В нашому випадку, якщо $V_i' = \mathrm{im}(f - \lambda_i I)^{q_i}$, то

$$\dim V_i + \dim V_i' = n.$$

Для доведення того, що $V = V_i \oplus V_i'$, лишилося перевірити рівність

$$V_i \cap V_i' = \vec{0}.$$

Припустимо, що $V_i \cap V_i' = U \neq \vec{0}$. Тоді U, як перетин двох інваріантних підпросторів, також є інваріантним підпростором. Обмеження g лінійного оператора f на лінійний підпростір U має характеристичний многочлен і, відповідно, власне число. Тому оператор f має власний вектор $\vec{x} \in U$, що відповідає певному власному числу μ. Якщо $\mu \neq \lambda_i$, то для будь-якого натурального m

$$(f - \lambda_i I)^m(\vec{x}) \neq 0$$

і $\vec{x} \notin V_i$. Якщо ж $\mu = \lambda_i$, то для деякого $\vec{y} \in V$ буде

$$\vec{x} = (f - \lambda_i I)^{q_i}(\vec{y}) \neq \vec{0}, \qquad (f - \lambda_i I)^{q_i}(\vec{x}) = (f - \lambda_i I)^{2q_i}(\vec{y}) = \vec{0}.$$

А це суперечить лемі 5.2.

■

Лема 5.4 *Кореневий підпростір V_i має вимірність p_i. Характеристичним многочленом обмеження оператора f на інваріантний підпростір $V_i' = \mathrm{im}(f - \lambda_i I)^{q_i}$ є многочлен*

$$\frac{\chi(\lambda)}{(\lambda - \lambda_i)^{p_i}}.$$

Доведення. Правильність леми випливає з того, що

1) обмеження f_1 оператора f на V_i має єдине власне число: λ_i;

2) обмеження f_2 оператора f на V_i' не має власних векторів, що відповідають власному числу λ_i, бо в протилежному випадку підпростори V_i та V_i' мали б ненульовий перетин;

3) характеристичний многочлен оператора f є добутком характеристичних многочленв операторів f_1 та f_2.

■

Тепер заготовано все необхідне для швидкого доведення теореми 5.1.

Доведення. Доведення проводиться індукцією по кількості різних власних чисел.

База індукції: одне власне число. В такому разі весь простір є єдиним кореневим підпростором і теорема правильна.

Індуктивне припущення: нехай ми можемо розкласти лінійний простір у пряму суму кореневих підпросторів у випадку, коли власних чисел менше ніж k.

Індуктивній перехід: нехай оператор f має k різних власних чисел. Тоді розкладаємо лінійний простір у пряму суму $V = V_1 \oplus V_1'$ (в позначеннях леми 5.4). А далі користуємося лемою 5.4 і розкладаємо уже V_1' у суму кореневих, що дозволене індуктивним припущенням.

Теорема 5.1 доведена.

■

5.2 Жорданова форма матриці з одним власним числом

Користуємося позначеннями попереднього розділу про кореневі підпростори.

Протягом цього розлілу оператор f буде мати лише одне власне число. Починаємо ми з випадку, коли цим власним числом є 0 — основна частина матеріалу стосується цього найважливішого окремого випадку. В цьому випадку характеристичний многочлен оператора має вигляд λ^n і f^n є нульовим оператором.

Визначення 5.3 *Лінійний оператор, який в певному степені є нульовим, називають нільпотентним. Найменше натуральне число, в якому нільпотентний оператор дорівнює нулю називають висотою нільпотентності оператора.*

Визначення 5.4 *Матриця*

$$J(\lambda) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

називається клітиною Жордана, яка відповідає власному числу λ, коли

$$a_{ii} = \lambda, i = 1, 2, \dots, n; \quad a_{i+1,i} = 1, i = 1, 2, \dots n-1; \quad a_{ij} = 0 \text{ якщо } j \notin \{i, i-1\}.$$

Прикладами клітин Жордана, що відповідають власному числу λ є матриці

$$(\lambda), \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}, \dots$$

Теорема 5.2 *Для нільпотентного оператора f існує базис, в якому матриця цього оператора блочно-діагональна і по діагоналі стоять клітини Жордана $J(0)$, які відповідають власному числу 0.*

Далі вважаємо, що оператор f нільпотентний і має висоту p. Нульова матриця є блочно-діагональною, де по діагоналі стоять клітини Жордана $J(0)$. Отже для нульового оператора теорема очевидна. Далі вважаємо, що $f \neq 0$, $p > 1$, $f^p = 0$, $f^{p-1} \neq 0$.

Позначимо через $V_i \subset V$ ядро оператора f^i, $i = 1, 2, , \ldots, p$. Таким чином, $V_p = V$ і V_1 — власний підпростір, що відповідає власному числу 0. Якщо для деякого натурального i і для деякого вектора $\vec{x}$ виконується рівність $f^i(\vec{x}) = \vec{0}$, то також буде $f^{i+1}(\vec{x}) = \vec{0}$. Тому побудовані підпростори $V_1, V_2, \ldots, V_p$ вкладені попередній в наступний:

$$\vec{0} = V_0 \subseteq V_1 \subseteq V_2 \subseteq \ldots \subseteq V_p = L.$$

Через U_i $(i = 2, 3, \ldots, p)$ позначимо доповнення підпростору V_{i-1} до підпростору V_i. Таким чином,

$$V_1 = U_1,\ V_2 = V_1 \oplus U_2,\ V_3 = V_2 \oplus U_3, \ldots,\ L = V_p = V_{p-1} \oplus U_p.$$

Лема 5.5 *Припустимо, що для деякого* $1 < q \leq p$ *вектори* $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k \in U_q$ *лінійно незалежні в* U_q. *Тоді для будь-якого натурального* r, $q > r \geq 1$ *вектори* $f^r(\vec{u}_1) = \vec{v}_1, f^r(\vec{u}_2) = \vec{v}_2, \ldots, f^r(\vec{u}_k) = \vec{v}_k \in V_{q-r}$ *лінійно незалежні і* $\mathrm{Lin}(\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k) \cap V_{q-r-1} = \vec{0}$.

Доведення. В цій лемі міститься три твердження:

1. Вектори $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in V_{q-r}$.
2. Вектори $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_k$ лінійно незалежні.
3. Лінійна оболонка $\mathrm{Lin}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k)$ векторів $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ має нульовий перетин із підпростором V_{q-r-1}

По черзі перевіряємо наведені твердження.

1. Оскільки $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k \in U_q$ і $U_q \subseteq V_q$, то $f^q(\vec{u}_i) = \vec{0}$ для $i = 1, 2, \ldots, k$. Звідси випливає, що $f^{q-r}(\vec{v}_i) = f^{q-r}(f^r(\vec{u}_i)) = f^q(\vec{u}_i) = \vec{0}$ для $i = 1, 2, \ldots, k$ і $\vec{v}_i \in V_{q-r}$. Перша частина перевірена.

2. Візьмемо лінійну комбінацію $\vec{v}$ векторів $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$:

$$\vec{v} = \sum_{i=1}^{k} \alpha_i \vec{v}_i = \sum_{i=1}^{k} \alpha_i f^r(\vec{u}_i) = f^r\left(\sum_{i=1}^{k} \alpha_i \vec{u}_i\right).$$

Позначимо вектор $\sum\limits_{i=1}^{k} \alpha_i \vec{u}_i$ через $\vec{u}$. Маємо

$$\vec{u} \in U_q,\ \vec{u} \in V_q,\ \vec{u} \notin V_{q-1},\ f^r(\vec{u}) = \vec{0} \ \Rightarrow\ f^{q-1}(\vec{u}) = \vec{0} \ \Rightarrow\ \vec{u} = \vec{0}.$$

Оскільки вектори $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ лінійно незалежні, то ми можемо написати

$$\vec{v} = \vec{0} \ \Rightarrow \ \vec{u} = \vec{0} \ \Rightarrow \ \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0,$$

що і означає лінійну незалежність векторів $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$. Доведення другої частини завершено.

3. Використовуючи позначення із доведення другої частини, маємо лінійну комбінацію $\vec{v} \in V_{q-r}$. Припустимо, що також $\vec{v} \in V_{q-r-1}$. Тоді

$$\vec{0} = f^{q-r-1}(\vec{v}) = f^{q-r-1}\left(\sum_{i=1}^{k} \alpha_i f^r(\vec{u}_i)\right) = f^{q-1}\left(\sum_{i=1}^{k} \alpha_i \vec{u}_i\right).$$

Звідси випливає, що $\vec{u} = \vec{0}$ і $\vec{v} = \vec{0}$. Третья частина леми доведена.

■

Приклад. З'ясуємо, який вигляд може мати матриця оператора f у 7-вимірному просторі, якщо ранг його матриці дорівнює 4 і $f^2 = 0$.

В нашому випадку

$$\dim V_2 = 7, \ \dim V_1 = \dim U_1 = 4, \ \dim U_2 = \dim V_2 - \dim V_1 = 3.$$

Вибираємо базис $\vec{u}_1, \ \vec{u}_2, \ \vec{u}_3 \in U_2$. Лінійно незалежні вектори $f(\vec{u}_1), \ f(\vec{u}_2), \ f(\vec{u}_3)$ доповнюємо до базису U_1 вектором $\vec{u}_4$. Об'єнання базисів U_1 і U_2 дає базис всього простору. В базисі

$$\vec{e}_1 = \vec{u}_1, \ \vec{e}_2 = f(\vec{u}_1), \ e_3 = u_2, \ \vec{e}_4 = f(\vec{u}_2), \ \vec{e}_5 = \vec{u}_3, \ \vec{e}_6 = f(\vec{u}_3), \ \vec{e}_7 = \vec{u}_4$$

матриця A оператора f має вигляд

$$A = \left(\begin{array}{cc|cc|cc|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

тобто матриця A блочно-діагональна, в ній по діагоналі стоять клітини жордана, що відповідають власному числу 0.

Доведення. Теорему 5.2 доводимо конструктивно: будуємо базис в якому матриця лінійного нільпотентного оператора f має жорданову форму.

Базис будується кроками.

На першому кроці будується базис $\vec{e}_{11}, \ \vec{e}_{12}, \ldots$ лінійного простору U_p. Оскільки ранг лінійного оператора (ранг його матриці) менше ніж n, і вимірність простору

V_{p-1} дорівнює рангу f, то

$$\dim U_p = \dim V_p - \dim V_{p-1} \geq 1.$$

На другому кроці вектори $f(\vec{e}_{11}),\ f(\vec{e}_{12}), \ldots$ доповнюємо векторам $\vec{e}_{21},\ \vec{e}_{22}, \ldots$ до базису U_{p-1}.

На $i-$ому кроці $(i = 2, 3, \ldots, p-1)$ беремо базис лінійного простору U_{p-i+2} діємо на його елементи (базисні вектори) оператором f і одержані лінійно незалежні вектори доповнюємо до базису лінійного підпростору U_{p-i+1}.

На останньому $p-$му кроці беремо базис підпростору U_2 діємо най його елементи оператором f. Одержані образи лінійно незалежні і ми доповнюємо їх векторами $\vec{e}_{p1},\ \vec{e}_{p2}, \ldots$ до базису $U_1 = V_1$.

Можливість здійснення вказаних кроків забезпечуєтсья лемою 5.5.

Об'єднання побудованих базисів є потрібним нам базисом Жордана, потрібно лише розташувати одержані вектори в належному порядку. Те, що об'єднання базисів підпросторів U_i для всіх $i = 1, 2, \ldots p$ дасть базис усього простору випливає з рівності $U_1 \oplus \ldots \oplus U_p = L$.

Одержані вектори розташовуються таким чином. Спочатку вибирається перший базисний вектор $\vec{e}_{11}$ із U_p, а далі ставляться його образи: $f(\vec{e}_{11}),\ f^2(\vec{e}_{11}), \ldots, f^{p-1}(\vec{e}_{11})$. Такі послідовності називаємо ланцюгами Жордана. Отже впорядкування базису починається із створення першого ланцюга Жордана

$$\vec{e}_{11},\ f(\vec{e}_{11}),\ f^2(\vec{e}_{11}), \ldots, f^{p-1}(\vec{e}_{11}).$$

Далі ставиться ланцюг Жордана, що починається із другого базисного вектора $\vec{e}_{12}$ підпростору U_p і так далі:

$$\vec{e}_{12}, f(e_{12}), f^2(e_{12}), \ldots, f^{p-1}(e_{12}),$$

$$\vec{e}_{13}, f(e_{13}), f^2(e_{13}), \ldots, f^{p-1}(e_{13}),$$

$$\ldots$$

Потім виписуються ланцюги Жордана, що починаються із базисних векторів підпростору U_{p-1}:

$$e_{21}, f(e_{21}), f^2(e_{21}), \ldots, f^{p-2}(e_{21}).$$

$$e_{22}, f(e_{22}), f^2(e_{22}), \ldots, f^{p-2}(e_{22}).$$

$$\ldots$$

Коли закінчуються ланцюги з початками в підпрострі U_i, переходимо до ланцюгів, що починаються в U_{i-1}. Таким чином, ми множину всіх базисних векторів розбили на ланцюги і ці ланцюги записали по черзі. Оце і є кінцевий результат — базис Жордана.

Ретельніше розглянемо одержаний базис Жордана.

Ланцюг Жордана є базисом інваріантного підпростору — це досить ясно. Також ясно, що в цьому інваріантному підпросторі у вибраному базисі матриця обмеження оператора буде клітиною Жордана. Весь простір розбиваєтсья в пряму суму таких інваріантних підпросторів, а матриця лінійного оператора стає блочно-діагональною, і по діагоналі стоять клітини Жордана.

■

Приклад. Маємо нільпотентний оператор f в десятивимірному просторі L. Його матриця в жордановій формі має три клітини — одна розміру 4 на 4 і дві клітини розміру 3 на 3. Які вимірності підпросторів

$$V_4 = \ker f^4,\ V_3 = \ker f^3,\ V_2 = \ker f^2,\ V_1 = \ker f?$$

Аналіз ситуації. Базис, в якому матриця оператора має жорданову форму складається із трьох ланцюгів Жордана (за кількістю клітин) — один ланцюг довжини 4 і два ланцюги довжиною 3. Отже f^4 є нульовим оператором і

$$\dim V_4 = \dim \ker f^4 = \dim L = 10.$$

Кількість жорданових ланцюгів довжини 4 — це вимірність простору U_4, який є доповненням підпростору V_3 до V_4. Тому

$$1 = \dim U_4 = \dim V_4 - \dim V_3 = 10 - \dim V_3 \ \Rightarrow\ \dim V_3 = 9.$$

Кількість ланцюгів довжини 3 — це вимірність U_3 без вимірності U_4. Тому

$$2 = \dim U_3 - \dim U_4 = (\dim V_3 - \dim V_2) - \dim U_4 = 9 - \dim V_2 - 1 \Rightarrow \dim V_2 = 6.$$

Кількість ланцюгів довжини 2 — це вимірність U_2 без вимірності U_3. $\dim U_3 = (\dim V_3 - \dim V_2) = 9 - 6 = 3$. Тому

$$0 = \dim U_2 - \dim U_3 = (\dim V_2 - \dim V_1) - 3 = 6 - \dim V_1 - 3 \Rightarrow \dim V_1 = 3.$$

Зауважимо, що $V_1 = \ker f$ є власним підпростором, і $\dim \ker f$ збігається із кількістю клітин Жордана в жордановій формі нільпотентного оператора.

5.3 Теорема Жордана

Теорема 5.3 (Теорема Жордана) *Для кожного лінійного оператора f існує базис, в якому матриця цього оператора блочно-діагональна і по діагоналі стоять клітини Жордана.*

Доведення. Спочатку лінійний простір L розкладаємо в пряму суму кореневих підпросторів лінійного оператора f. На кореневому підпросторі, що відповідає власному числу λ_0, розглядаємо нільпотентний оператор $g = f - \lambda_0 I$. Для оператора g на кореневому підпросторі будуємо жорданів базис. В цьому базисі оператор g має матрицю, що знаходиться в жордановій формі. Клітини в цій формі відповідають власному числу 0. Тоді оператор $f = g + \lambda_0 I$ на кореневому підпросторі, що відповідає власному числу λ_0 буде мати в цьому базисі матрицю, що знаходиться в формі Жордана і клітини відповідають власному числу λ_0.

Кореневі підпростори інваріантні, матриця оператора блочно-діагональна, і по діагоналі стоять матриці, що знаходяться в формі Жордана. Отже і вся матриця знаходиться в формі Жордана.

Кінець доведення.

$\blacksquare$

Приклад. Привести до жорданової форми матрицю

$$A = \begin{pmatrix} -262 & -590 & 951 & -821 & -619 \\ 327 & 740 & -1196 & 1041 & 749 \\ 54 & 122 & -197 & 171 & 125 \\ -73 & -166 & 269 & -236 & -162 \\ -21 & -48 & 78 & -69 & -45 \end{pmatrix}$$

з одним власним числом 0.

Відповідь: Матриця A має жорданову форму

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

Знаходження жорданової форми. Вважаємо відомим, що $A^2 = 0$ (власне, це не важко перевірити). Шукаємо базис ядра і вимірність образу оператора f. Ядро — це підпростір розв'язків однорідної системи рівнянь $A\vec{x} = \vec{0}$ із основною матрицею A. Розв'язуємо методом Гауса: виписуємо матрицю A і приводимо її до спрощеного вигляду.

$$A = \begin{pmatrix} -262 & -590 & 951 & -821 & -619 \\ 327 & 740 & -1196 & 1041 & 749 \\ 54 & 122 & -197 & 171 & 125 \\ -73 & -166 & 269 & -236 & -162 \\ -21 & -48 & 78 & -69 & -45 \end{pmatrix} \sim \begin{pmatrix} -1 & -4 & 8 & -11 & 9 \\ 0 & 2 & -5 & 9 & -13 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} -1 & 0 & -2 & 7 & -17 \\ 0 & 2 & -5 & 9 & -13 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B$$

За матрицею B розділяємо основні змінні: x_1, x_2 і вільні змінні: x_3, x_4, x_5, виписуємо систему рівнянь, що відповідає матричному рівнянню $B\vec{x} = \vec{0}$

$$\begin{cases} x_1 + 2x_3 - 7x_4 + 17x_5 &= 0, \\ 2x_2 - 5x_3 + 9x_4 - 13x_5 &= 0. \end{cases}$$

Переносимо вільні змінні направо:

$$\begin{cases} x_1 &= -2x_3 + 7x_4 - 17x_5, \\ 2x_2 &= 5x_3 - 9x_4 + 13x_5. \end{cases}$$

Надаємо вільним відомих належних значень, обчислюємо відповідні значення основних змінних, результати записуємо в табличку

	x_1	x_2	x_3	x_4	x_5
$\vec{e}_1$	-4	5	2	0	0
$\vec{e}_2$	14	-9	0	2	0
$\vec{e}_3$	-34	13	0	0	2

Ранг матриці B і відповідно, ранг матриці A дорівнюють двом. Отже вимірність образу оператора f дорівнює 2. Ядро оператора f тривимірне (кажуть ще, що дефект оператора f дорівнює 3). Базис ядра оператора f складають вектори

$$\vec{e}_1 = \{-4, 5, 2, 0, 0\}, \quad \vec{e}_2 = \{14, -9, 0, 2, 0\}, \quad \vec{e}_3 = \{-34, 13, 0, 0, 2\}.$$

В позначеннях, що використані при доведенні теореми, маємо V_2 — увесь 5-вимірний простір, V_1 — ядро оператора f — тривимірний простір. Потрібно знайти базис доповнення U_2:

$$V_2 = U_2 \oplus V_1.$$

Ми знайємо, що $\dim U_2 = \dim V_2 - \dim V_1 = 2$.

Перший спосіб полягає в тому, щоб подумки ввести скалярний добуток і знайти ортогональне доповнення, тобто знайти базис підпростору розв'язків однорідної системи рівнянь

$$\begin{cases} \langle \vec{e}_1,\ \vec{x} \rangle = 0, \\ \langle \vec{e}_2,\ \vec{x} \rangle = 0, \\ \langle \vec{e}_3,\ \vec{x} \rangle = 0. \end{cases}$$

Другий спосіб полягає в грубому підбиранні потрібних векторів. В якості кандидатів на випробування можна брати стандартні базисні вектори

$$\vec{u}_1 = \{1, 0, 0, 0, 0\},\ \vec{u}_2 = \{0, 1, 0, 0, 0\},\ \vec{u}_3 = \{0, 0, 1, 0, 0\},$$

$$\vec{u}_4 = \{0, 0, 0, 1, 0\},\ \vec{u}_5 = \{0, 0, 0, 0, 1\}.$$

Випробовуємо перший вектор $\vec{u}_1$. Він підходить, якщо ранг системи векторів $\{\vec{u}_1,\ \vec{e}_1,\ \vec{e}_2,\ \vec{e}_3\}$ дорівнює 4. Виписуємо матрицю із координат цих векторів

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -4 & 5 & 2 & 0 & 0 \\ 14 & -9 & 0 & 2 & 0 \\ -34 & 13 & 0 & 0 & 2 \end{pmatrix}$$

Мінор, що стоїть у першому, третьому, четвертому та 5 стовпчиках дорівнює 8 — не нульовий. Отже ранг матриці дорівнює 4 і вектор $\vec{u}_1$ лежить в доповненні.

Випробовуємо другий вектор $\vec{u}_2$. Цей вектор є другим вектором базиса доповнення, якщо ранг системи векторів $\{\vec{u}_1,\ \vec{u}_2,\ \vec{e}_1,\ \vec{e}_2,\ \vec{e}_3\}$ дорівнює 5. Виписуємо матрицю із координат цих векторів

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -4 & 5 & 2 & 0 & 0 \\ 14 & -9 & 0 & 2 & 0 \\ -34 & 13 & 0 & 0 & 2 \end{pmatrix}$$

Визначник цієї матриці дорівнює 8 — не нульовий. Отже ранг матриці дорівнює 5 і вектор $\vec{u}_2$ є другим базисним вектором доповнення.

Базисні вектори U_2 є початками ланцюгів Жордана довжини 2. Таким чином ми одержуємо два ланцюги Жордана довжини 2:

$$\vec{u}_1 = \{1, 0, 0, 0, 0\},\ f(\vec{u}_1) = \{-262, 327, 54, -73, -21\},$$

$$\vec{u}_2 = \{0, 1, 0, 0, 0\},\ f(\vec{u}_2) = \{-590, 740, 122, -166, -48\}.$$

Переходимо до пошуку 5-го, останнього вектора (позначимо його через $\vec{w}$) базиса Жордана, який утворює ланцюг довжини 1. Цей вектор повинен лежати в підпросторі V_1, базисними векторами якого є $\vec{e}_1,\ \vec{e}_2,\ \vec{e}_3$, і не лежати в підпросторі $f(U_2)$, базисним векторами якого є $f(\vec{u}_1),\ f(\vec{u}_2)$. Знову ж, для пошуку вектора $\vec{w}$ маємо дві можливості.

Перший спосіб полягає в записі вектора $\vec{w}$ у вигляді $\vec{w} = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3$, складанні умови ортогональності вектора $\vec{w}$ векторам $f(\vec{u}_1),\ f(\vec{u}_2)$, і розв'язуванні одержаної системи рівнянь:

$$\langle \vec{w},\ f(\vec{u}_1) \rangle = 0, \quad \langle \vec{w},\ f(\vec{u}_2) \rangle = 0.$$

Другий спосіб полягає в брутальному підбиранні потрібного базисного вектора. Кандидатами для випробування є базисні вектори $\vec{e}_1,\ \vec{e}_2,\ \vec{e}_3$. Підемо цим шляхом. Випробовуємо вектор $w = e_1$. Він підходить нам, якщо ранг системи векторів $f(\vec{u}_1),\ f(\vec{u}_2),\ \vec{w}$ дорівнює трьом. Виписуємо матрицю із координат наших векторів

$$\begin{pmatrix} -262 & 327 & 54 & -73 & -21 \\ -590 & 740 & 122 & -166 & -48 \\ -4 & 5 & 2 & 0 & 0 \end{pmatrix}$$

Мінор, що стоїть в останніх трьох стовпчиках дорівнює $2{\cdot}(73{\cdot}48-166{\cdot}21) = 36 \neq 0$. Тому ранг побудованої матриці дорівнює 3-м і вектор $\vec{e}_1$ можна взяти 5-м базисним вектором у базисі Жордана.

Записуємо одержані базисні вектори в належному порядку

$$\begin{array}{rcccl} \vec{v}_1 & = & \vec{u}_1 & = & \{1,\ 0,\ 0,\ 0,\ 0\}, \\ \vec{v}_2 & = & f(\vec{u}_1) & = & \{-262,\ 327,\ 54,\ -73,\ -21\}, \\ \vec{v}_3 & = & \vec{u}_2 & = & \{0,\ 1,\ 0,\ 0,\ 0\}, \\ \vec{v}_4 & = & f(\vec{u}_2) & = & \{-590,\ 740,\ 122,\ -166,\ -48\}, \\ \vec{v}_5 & = & \vec{e}_1 & = & \{-4,\ 5,\ 2,\ 0,\ 0\}. \end{array}$$

Оце і є базис Жордана

Матрицею переходу до нового базису буде матриця

$$C = \begin{pmatrix} 1 & -262 & 0 & -590 & -4 \\ 0 & 327 & 1 & 740 & 5 \\ 0 & 54 & 0 & 122 & 2 \\ 0 & -73 & 0 & -166 & 0 \\ 0 & -21 & 0 & -48 & 0 \end{pmatrix}$$

Оберненою до матриці C буде матриця

$$C^{-1} = \begin{pmatrix} 1 & 0 & 2 & -7 & 17 \\ 0 & 0 & 0 & -\frac{8}{3} & \frac{83}{9} \\ 0 & 1 & -\frac{5}{2} & \frac{9}{2} & -\frac{13}{2} \\ 0 & 0 & 0 & \frac{7}{6} & -\frac{73}{18} \\ 0 & 0 & \frac{1}{2} & \frac{5}{6} & -\frac{29}{18} \end{pmatrix}$$

Можна зробити перевірку того, що в процесі підрахунків не допущено якоїсь помилки. Для цього потрібно перемножити три матриці $C^{-1} \cdot A \cdot C$. Оскільки

$$C^{-1} \cdot A \cdot C = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

то в обчисленнях помилку не зробили.

Покажчик