Ю.І. Травкін

Лінійна **АЛГЕБРА** і аналітична **ГЕОМЕТРІЯ**

НАВЧАЛЬНИЙ ПОСІБНИК

ХАРКІВ «МАЙДАН» 2009

УДК 512 (075.8)
ББК 22.143я73
Т65

Затверджено до друку Вченою радою
Харківського національного аграрного університету ім. В.В. Докучаєва
(протокол № 3 від 25 березня 2009 року)

Рецензенти: ***Литвин О.М.****, д-р фіз.-мат. наук, проф. Української інженерно-педагогічної академії;*
Ольшанський В.П.*, д-р фіз.-мат. наук, проф. Харківського національного технічного ун-ту сільського господарства;*
Проценко В.С.*, д-р фіз.-мат. наук, проф. Національного аерокосмічного ун-ту ім. М.Є. Жуковського («ХАІ»)*

Травкін Ю. І.
Т65 Лінійна алгебра і аналітична геометрія: Навчальний посібник.— Х.: Майдан, 2009.— 416 с. Укр. мова.

ISBN 978-966-372-276-4.

Навчальний посібник вміщує об'єднаний курс лінійної алгебри і багатовимірної аналітичної геометрії, а також спеціальні питання: теорію систем лінійних різницевих і диференціальних рівнянь першого порядку з постійними коефіцієнтами, метод найменших квадратів, обчислювальні методи лінійної алгебри.

Навчальний посібник призначений для студентів інженерно-технічних , економічних і інших нематематичних спеціальностей. Він може бути корисним аспірантам і науковцям указаних спеціальностей.

УДК 512 (075.8)
ББК 22.143я73

ISBN 978-966-372-276-4

ЗМІСТ

Присвячується
світлій пам'яті
моєї дорогої дружини
Тамари Миколаївни

ПЕРЕДМОВА

Лінійна алгебра, як ніяка інша математична дисципліна, тісно пов'язана з обчислювальною практикою розв'язання прикладних наукових задач. Не буде великим перебільшенням стверджувати, що будь-яка науково-технічна або економічна проблема на тому чи іншому етапі зводиться до алгебраїчної задачі. Тому лінійна алгебра є важливою складовою частиною математичної підготовки кожного фахівця інженерно-технічного або економічного профілю.

Сучасний підхід до викладання лінійної алгебри полягає в об'єднанні її з багатовимірною аналітичною геометрією, що дозволяє більш глибоко висвітлити обидві дисципліни і підкреслити тісні зв'язки між ними. Існуючі підручники з об'єднаного курсу лінійної алгебри і багатовимірної аналітичної геометрії [2, 4, 7, 8] зорієнтовані на потреби навчального процесу на математичних факультетах університетів. Студентам нематематичних спеціальностей скористатися цими підручниками досить складно, а аналогічних підручників для таких студентів не існує.

Заповнити вказану прогалину покликана книга, що пропонується увазі читача. Книга є переробленим і доповненим виданням відповідних розділів підручника [15] для економічних спеціальностей. Вона може бути особливо корисною тим особам, які у своїй фаховій діяльності матимуть справу з багатовимірним статистичним аналізом, математичним програмуванням, обчислювальними методами.

Книгу можна умовно розділити на дві частини. Перша з них (глави 1 — 9) є елементарним вступом до об'єднаного курсу лінійної алгебри і багатовимірної аналітичної геометрії, який містить, в основному, традиційні питання, розглядувані в цих математичних дисциплінах. У другій частині (глави 10 —12) розглядаються спеціальні питання: системи лінійних різницевих і диференціальних рівнянь першого порядку з постійними коефіцієнтами та аналогічні рівняння вищих порядків, метод найменших квадратів, обчислювальні задачі лінійної алгебри. Для читання книги достатньо знати математику в обсязі шкільної програми і лише в окремих випадках для розуміння тексту потрібні додаткові відомості з математичного аналізу (правило Лопіталя, початкові відомості про ряди, теорема Вейєрштрасса для функцій кількох змінних). Необхідну інформацію із зазначених питань можна почерпнути з книги [10].

До всіх глав книги добавлені вправи, які вміщують додаткові теоретичні питання, а наприкінці книги наведені вказівки до багатьох з цих вправ. Читачу рекомендується після вивчення глави прочитати всі вправи до неї незалежно від того, які з них він буде виконувати. Читач повинен також виконати доведення тих тверджень, до яких у тексті відносяться фрази «неважко переконатися», «аналогічно можна довести» і т. ін.

Кожна глава книги складається з розділів. Наприклад, розділ 7.3 — це третій розділ глави 7. Для нумерації і посилання на формули в межах розділу використовується одне число, а при посиланні на формули з інших розділів указуються три числа — номери розділів і номер формули. Позначка « $\square$ » вказує на кінець доведення теореми або його відсутність, а « $\triangle$ » — на кінець розв’язку прикладу.

Автор буде вдячний усім читачам, які звернуть його увагу на недоліки книги.

Ю.Травкін

ЛІТЕРАТУРА

1. А л б е р т А. Регрессия, псевдоинверсия и рекуррентное оценивание. — М.: Наука, 1977.

2. А л е к с а н д р о в П.С. Курс аналитической геометрии и линейной алгебры. — М.: Наука, 1979.

3. Б е к л е м и ш е в Д.В. Дополнительные главы линейной алгебры. — М.: Наука, 1983.

4. В о е в о д и н В.В. Линейная алгебра. — М.: Наука, 1980.

5. Г а н т м а х е р Ф.Р. Теория матриц. — М.: Наука, 1988.

6. Г о л у б Дж., В а н Л о у н Ч. Матричные вычисления. — М.: Мир, 1999.

7. Е ф и м о в Н.В., Р о з е н д о р н Э.Р. Линейная алгебра и многомерная геометрия. — М.: Наука, 1970.

8. И л ь и н В.А., К и м Г.Д. Линейная алгебра и аналитическая геометрия. — М.: ТК Велби, Изд-во Проспект, 2007.

9. И л ь и н В.А., П о з н я к Э.Г. Линейная алгебра. — М.: ФИЗМАТЛИТ, 2005.

10. И л ь и н В.А., П о з н я к Э.Г. Основы математического анализа, ч.1. — М.: ФИЗМАТЛИТ, 2005.

11. К а х а н е р Д., М о у л е р К., Н э ш С. Численные методы и математическое обеспечение. — М.: Мир, 1998.

12. К р ы л о в В.И., Б о б к о в В.В., М о н а с т ы р н ы й П.И. Вычислительные методы, т.1.— М.: Наука, 1974.

13. Л а н к а с т е р П. Теория матриц. — М.: Наука, 1982.

14. С т р е н г Г. Линейная алгебра и её применения. — М.: Мир, 1980.

15. Т р а в к і н Ю.І., М а л я р е ц ь Л.М. Математика для економістів. — Х.: ВД «ІНЖЕК», 2005.

16. Т р а в к і н Ю.І., Т р а в к і н а Т.М. Збірник вправ з основ лінійної алгебри і її застосувань. — Х.: Основа, 1998.

17. Ф о р с а й т Дж., М а л ь к о л ь м М., М о у л е р К. Машинные методы математических вычислений. — М.: Мир, 1980.

18. Ф о р с а й т Дж., М о л е р К. Численное решение систем линейных алгебраических уравнений. — М.: Мир, 1969.

19. Х о р н Р., Д ж о н с о н Ч. Матричный анализ. — М.: Мир, 1989.

ГЛАВА 1

МАТРИЦІ І ВИЗНАЧНИКИ

Теорія матриць є важливою складовою частиною лінійної алгебри. У цій главі розглядаються матриці і вводяться операції над ними. Крім того, вивчаються визначники, які є однією з основних числових характеристик квадратних матриць, а також обернення матриць, яке застосовується при розв'язанні систем лінійних рівнянь. В останньому розділі глави викладається техніка оперування з блоковими матрицями, яка дозволяє в багатьох випадках суттєво спростити теоретичні міркування.

1.1. Матриці

1.1.1. Основні визначення. *Матрицею* $\boldsymbol{A}$ *розмірів* $m \times n$ ($m \times n$ - *матрицею*) називається сукупність чисел (*елементів* матриці), розміщених у вигляді прямокутної таблиці, яка має *m рядків* і *n стовпців*:

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

Перший індекс елемента a_{ij} є номером рядка, а другий — номером стовпця, на перетині яких розміщений елемент. Крім позначення a_{ij} для елемента матриці $\boldsymbol{A}$ використовують позначення $\{\boldsymbol{A}\}_{ij}$. Якщо $m=1$ або $n=1$, то відповідний індекс може бути відсутнім. У перших главах книги розглядаються матриці, елементами яких є дійсні числа. У подальшому будуть розглянуті комплексні матриці, елементами яких є комплексні числа (ці числа розглядаються в розділі 6.1). Про які саме матриці (дійсні чи комплексні) буде йти мова в тому чи іншому конкретному випадку, буде вказуватися або буде зрозумілим з контексту.

Матриці $\boldsymbol{A}$ і $\boldsymbol{B}$ називаються *рівними* $(\boldsymbol{A}=\boldsymbol{B})$, якщо вони мають однакові розміри і їх відповідні елементи рівні.

Матриця, всі елементи якої дорівнюють нулю, називається *нульовою* і позначається символом $\mathbf{0}$.

Матриця називається *квадратною* матрицею *порядку* n, якщо $n=m$. У загальному випадку матриця називається прямокутною.

Матриця, що складається з одного стовпця (рядка), називається *матрицею-стовпцем* (*матрицею-рядком*). Матрицю-стовпець (матрицю-рядок) називають також *вектором.*

Матриця $\boldsymbol{A}^T$ розмірів $n\times m$ називається *транспонованою* по відношенню до $m\times n$ - матриці $\boldsymbol{A}$, якщо $\boldsymbol{A}^T$ одержується з $\boldsymbol{A}$ перетворенням стовпців у рядки з тим же самим номером, тобто $\{\boldsymbol{A}^T\}_{ij}=\{\boldsymbol{A}\}_{ji}$ для всіх i, j. Матриця $\boldsymbol{A}$ називається *симетричною* (*кососиметричною*), якщо $\boldsymbol{A}=\boldsymbol{A}^T$ $(\boldsymbol{A}=-\boldsymbol{A}^T)$. Матриця, транспонована по відношенню до матриці-рядка, є матрицею-стовпцем, що складається з тих же елементів. Цією обставиною ми будемо користуватися для зручності запису. Замість матриці-стовпця

$$\boldsymbol{a}=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \dots\\ \alpha_n\end{pmatrix}$$

ми будемо писати:

$$\boldsymbol{a}=(\alpha_1\quad \alpha_2\quad \dots\quad \alpha_n)^T.$$

Сукупність елементів $a_{11}, a_{22},\dots$ складає *головну діагональ* матриці. Ці елементи називаються *діагональними*, а всі інші — *позадіагональними*. Матриця називається *діагональною*, якщо всі її позадіагональні елементи дорівнюють нулю. Діагональна матриця з елементами $a_{11}, a_{22},\dots, a_{mm}$ позначається через $\mathrm{diag}\,(a_{11}, a_{22},\dots, a_{mm})$. Діагональна квадратна матриця з рівними діагональними елементами називається *скалярною*. Якщо у такої матриці діагональні елементи дорівнюють одиниці, то вона називається *одиничною* і позначається через $\boldsymbol{E}$. При необхідності вказати, що одинична матриця має порядок r, використовують позначення $\boldsymbol{E}_r$.

Якщо всі елементи матриці, розміщені нижче (вище) головної діагоналі, дорівнюють нулю, то матриця називається *верхньою* (*нижньою*) *трикутною*. Діагональну матрицю можна вважати верхньою і нижньою трикутною матрицею одночасно.

1.1.2. Операції над матрицями. Основними операціями над матрицями є додавання матриць, множення матриці на число і множення матриць.

Сумою матриць $\boldsymbol{A}$ і $\boldsymbol{B}$ однакових розмірів називається матриця $\boldsymbol{A}+\boldsymbol{B}$, елементи якої дорівнюють сумам відповідних елементів матриць $\boldsymbol{A}$ і $\boldsymbol{B}$. Операція знаходження суми матриць називається додаванням матриць. Очевидно, що операція додавання матриць має ті ж властивості, що і операція додавання дійсних чисел :

1) $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$; 2) $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$.

Добутком матриці $\boldsymbol{A}$ *на число* α називається матриця $\alpha\boldsymbol{A}$, елементи якої є добутками відповідних елементів матриці $\boldsymbol{A}$ на число α. Операція знаходження добутку матриці на число називається множенням матриці на число. Неважко переконатися в тому, що ця операція має такі властивості:

1) $\alpha(\beta\boldsymbol{A})=(\alpha\beta)\boldsymbol{A}$; 2) $\alpha(\boldsymbol{A}+\boldsymbol{B})=\alpha\boldsymbol{A}+\alpha\boldsymbol{B}$;

3) $(\alpha+\beta)\boldsymbol{A}=\alpha\boldsymbol{A}+\beta\boldsymbol{A}$.

Різниця $\boldsymbol{A}-\boldsymbol{B}$ двох *матриць* однакових розмірів визначається як $\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}+(-1)\boldsymbol{B}$.

Добутком $m\times n$*-матриці* $\boldsymbol{A}$ *і* $n\times p$*-матриці* $\boldsymbol{B}$ називається $m\times p$-матриця $\boldsymbol{AB}$, елемент $\{\boldsymbol{AB}\}_{ij}$ якої дорівнює сумі добутків відповідних елементів i-го рядка матриці $\boldsymbol{A}$ і j-го стовпця матриці $\boldsymbol{B}$:

$$\{\boldsymbol{AB}\}_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},\quad i=\overline{1,\ m},\quad j=\overline{1,\ p}. \tag{1}$$

Операція знаходження добутку матриць називається множенням матриць. З визначення добутку матриць випливає, що матрицю $\boldsymbol{A}$ можна помножити не на будь-яку матрицю $\boldsymbol{B}$: необхідно, щоб кількість стовпців матриці $\boldsymbol{A}$ дорівнювала кількості рядків матриці $\boldsymbol{B}$. У подальшому, коли мова буде йти про множення матриць, ми будемо вважати, що множники мають відповідні розміри.

Як приклад застосування формули (1) розглянемо добуток двох верхніх (нижніх) трикутних матриць. Якщо $\boldsymbol{A}$ і $\boldsymbol{B}$ є верхніми трикутними матрицями, то $a_{ik} = 0$ при $i > k$, а також $b_{kj} = 0$ при $k > j$. Тому $\sum_{k=1}^{n} a_{ik} b_{kj} = 0$ при $i > j$. На підставі формули (1) $\{\boldsymbol{AB}\}_{ij} = 0$ при $i > j$, тобто матриця $\boldsymbol{AB}$ є верхньою трикутною. Якщо $i = j$, то

$$\{\boldsymbol{AB}\}_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki} = a_{ii} b_{ii}, \quad i = \overline{1,\ \min(m, n)},$$

тобто діагональні елементи добутку двох верхніх трикутних матриць дорівнюють добуткам відповідних діагональних елементів матриць $\boldsymbol{A}$ і $\boldsymbol{B}$. Аналогічні результати можна одержати для нижніх трикутних матриць.

З формули (1) випливають такі властивості добутку матриць:

1) $\alpha(\boldsymbol{AB}) = (\alpha\boldsymbol{A})\boldsymbol{B} = \boldsymbol{A}(\alpha\boldsymbol{B})$; 2) $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC} + \boldsymbol{BC}$;

3) $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{AB} + \boldsymbol{AC}$; 4) $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$.

Обмежимося доведенням останньої властивості (інші є цілком очевидними). Припустимо, що матриці $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ мають розміри $m \times n$, $n \times p$, $p \times q$ відповідно. Тоді

$$\{(\boldsymbol{AB})\boldsymbol{C}\}_{ij} = \sum_{l=1}^{p} \left(\sum_{k=1}^{n} a_{ik} b_{kl} \right) c_{lj}, \quad \{\boldsymbol{A}(\boldsymbol{BC})\}_{ij} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{p} b_{kl}\, c_{lj} \right),$$

де $i = \overline{1,\ m}$, $j = \overline{1,\ q}$. Праві частини цих рівностей відрізняються лише порядком підсумовування відносно k і l, тобто є рівними, тому $\{(\boldsymbol{AB})\boldsymbol{C}\}_{ij} = \{\boldsymbol{A}(\boldsymbol{BC})\}_{ij}$. На підставі визначення рівності матриць це означає, що $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$.

З визначення добутку матриць випливає, що, взагалі кажучи, множення матриць не має переставної властивості. Неважко зрозуміти, що коли $\boldsymbol{AB}$ і $\boldsymbol{BA}$ існують, то вони є квадратними матрицями. Але навіть коли $\boldsymbol{AB}$ і $\boldsymbol{BA}$ мають однаковий порядок, рівність $\boldsymbol{AB} = \boldsymbol{BA}$, взагалі кажучи, не виконується. Наприклад,

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

Якщо $\boldsymbol{AB} = \boldsymbol{BA}$, то матриці $\boldsymbol{A}$ і $\boldsymbol{B}$ називаються *переставними*. Серед переставних матриць важливе значення мають одинична і нульова матриці. Легко переконатися в тому, що $\boldsymbol{EA} = \boldsymbol{AE} = \boldsymbol{A},\ \boldsymbol{0A} = \boldsymbol{A0} = \boldsymbol{0}$. Ці формули характеризують особливу роль одиничної і нульової матриць, аналогічну тій ролі, яку відіграють числа 1 і 0 при множенні чисел.

Нехай $\boldsymbol{A}$ — квадратна матриця. Степінь матриці $\boldsymbol{A}^m$ визначається як добуток m множників, рівних $\boldsymbol{A}$. При цьому вважається, що $\boldsymbol{A}^0 = \boldsymbol{E}$. З властивості 4) добутку матриць випливає, що

$$\boldsymbol{A}^p\boldsymbol{A}^q = \boldsymbol{A}^{p+q},$$

де p і q — будь-які цілі невід'ємні числа. Для многочлена

$$f(\lambda) = a_0 + a_1\lambda + \ldots + a_{m-1}\lambda^{m-1} + a_m\lambda^m$$

многочлен від матриці $\boldsymbol{A}$ визначається так:

$$f(\boldsymbol{A}) = a_0\boldsymbol{E} + a_1\boldsymbol{A} + \ldots + a_{m-1}\boldsymbol{A}^{m-1} + a_m\boldsymbol{A}^m.$$

Нехай многочлен $f(\lambda)$ дорівнює добутку многочленів $g(\lambda)$ і $h(\lambda)$:

$$f(\lambda) = g(\lambda)\,h(\lambda)\,.$$

Многочлен $f(\lambda)$ одержується з $g(\lambda)$ і $h(\lambda)$ шляхом почленного множення і зведення подібних членів. При цьому використовується правило множення степенів, що визначається рівністю $\lambda^p\lambda^q = \lambda^{p+q}$. Вище ми бачили, що це правило залишається справедливим при заміні величини λ матрицею $\boldsymbol{A}$. Тому

$$f(\boldsymbol{A}) = g(\boldsymbol{A})\,h(\boldsymbol{A}).$$

Оскільки $g(\lambda)\,h(\lambda) = h(\lambda)\,g(\lambda)$, то

$$g(\boldsymbol{A})\,h(\boldsymbol{A}) = h(\boldsymbol{A})\,g(\boldsymbol{A}),$$

тобто два многочлена від однієї і тієї ж матриці є переставними між собою.

Повертаючись до формули (1), введемо спеціальні позначення для стовпців і рядків $m \times n$-матриці $\boldsymbol{A}$. Будемо позначати i-й рядок матриці $\boldsymbol{A}$ через $\boldsymbol{a}_{i*}$, а j-й стовпець — через $\boldsymbol{a}_{*j}$:

$$\boldsymbol{a}_{i*} = (a_{i1}\ \ a_{i2}\ \ \ldots\ \ a_{in}),\ \ \boldsymbol{a}_{*j} = (a_{1j}\ \ \ a_{2j}\ \ldots\ a_{mj})^T,\quad i = \overline{1,\ m},\quad j = \overline{1,\ n}.$$

Введемо поняття *лінійної комбінації* стовпців $\boldsymbol{a}_{*1}, \ldots, \boldsymbol{a}_{*n}$ як вираз $\sum_{j=1}^{n} \alpha_j \boldsymbol{a}_{*j}$, де $\alpha_1,\ \ldots, \alpha_n$ — числа. Аналогічно введемо поняття лінійної комбінації рядків. Тепер запишемо формулу (1) для матриці $\boldsymbol{AB}$ у двох виглядах:

$$\{\boldsymbol{AB}\}_{*j} = \sum_{k=1}^{n} b_{kj}\boldsymbol{a}_{*k},\quad j = \overline{1,\ p}, \tag{2}$$

$$\{\boldsymbol{AB}\}_{i*} = \sum_{k=1}^{n} a_{ik}\boldsymbol{b}_{k*},\quad i = \overline{1,\ m}. \tag{3}$$

Формула (2) показує, що будь-який j-й стовпець матриці $\boldsymbol{AB}$ є лінійною комбінацією стовпців першого множника, тобто матриці $\boldsymbol{A}$, причому коефіцієнти лінійної комбінації утворюють j-й стовпець другого множника. Аналогічно з рівності (3) випливає, що будь-який i-й рядок матриці $\boldsymbol{AB}$ є лінійною комбінацією рядків матриці $\boldsymbol{B}$, а коефіцієнти цієї лінійної комбінації є елементами i-го рядка матриці $\boldsymbol{A}$.

З формули (3) випливає, що при множенні зліва будь-якої матриці на квадратну діагональну матрицю $\operatorname{diag}(d_1,\ d_2,\ \ldots)$ рядки цієї матриці множаться на числа $d_1,\ d_2,\ \ldots$. Якщо ж деяка матриця множиться справа на квадратну діагональну матрицю, то на числа $d_1,\ d_2,\ \ldots$ множаться стовпці цієї матриці, в чому неважко переконатися за допомогою формули (2).

Сформулюємо правило транспонування добутку матриць. З урахуванням формули (1) маємо:

$$\{(\boldsymbol{AB})^T\}_{ij} = \{\boldsymbol{AB}\}_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki} = \sum_{k=1}^{n} \{\boldsymbol{B}^T\}_{ik} \{\boldsymbol{A}^T\}_{kj} = \{\boldsymbol{B}^T \boldsymbol{A}^T\}_{ij},$$

звідки випливає рівність

$$(\boldsymbol{AB})^T = \boldsymbol{B}^T \boldsymbol{A}^T. \tag{4}$$

Її можна узагальнити на випадок більшої кількості множників. Наприклад, для добутку трьох множників на підставі (4) маємо:

$$(\boldsymbol{ABC})^T = ((\boldsymbol{AB})\boldsymbol{C})^T = \boldsymbol{C}^T (\boldsymbol{AB})^T = \boldsymbol{C}^T \boldsymbol{B}^T \boldsymbol{A}^T. \tag{5}$$

Будь-яка матриця першого порядку є симетричною, тому при транспонуванні вона не змінюється. Наприклад, якщо $\boldsymbol{A}$ — матриця і $\boldsymbol{x}$, $\boldsymbol{y}$ — вектори, то $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{y}$ — матриця першого порядку(число). На підставі формули (5) маємо:

$$\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{y} = \boldsymbol{y}^T \boldsymbol{A}^T \boldsymbol{x}.$$

Цією рівністю ми будемо часто користуватися в подальшому.

1.1.3. Елементарні перетворення матриць. *Елементарними* називаються такі *перетворення* матриці: 1) множення деякого рядка (стовпця) на число, відмінне від нуля; 2) додавання до будь-якого рядка (стовпця) іншого рядка (стовпця), помноженого на довільне число; 3) переставлення двох будь-яких рядків (стовпців).

Дві матриці називаються *еквівалентними*, якщо кожна з них отримується з іншої за допомогою скінченної кількості елементарних перетворень. Якщо матриці $\boldsymbol{A}$ і $\boldsymbol{B}$ еквівалентні, то це позначається так: $\boldsymbol{A} \sim \boldsymbol{B}$.

Якщо будь-яке елементарне перетворення застосувати до одиничної матриці, то одержимо матрицю, яка називається *елементарною*. Ми будемо позначати через $\boldsymbol{L}_i(\lambda), \boldsymbol{L}_i(\lambda | j), \boldsymbol{L}_{ij}$ елементарні матриці, одержані з одиничної матриці відповідно множенням i-го рядка на число λ, додаванням до i-го рядка j-го рядка, помноженого на λ, переставленням i-го та j-го рядків. З формули (3) випливає, що кожне елементарне перетворення з рядками матриці рівносильне множенню її зліва на відповідну елементарну матрицю.

Елементарні перетворення зі стовпцями матриці $\boldsymbol{A}$, як випливає з формули (2), рівносильні множенню її справа на елементарні матриці, які за аналогією з матрицями $\boldsymbol{L}_i(\lambda)$, $\boldsymbol{L}_i(\lambda | j)$, $\boldsymbol{L}_{ij}$ ми позначимо через $\boldsymbol{M}_i(\lambda)$, $\boldsymbol{M}_i(\lambda | j)$, $\boldsymbol{M}_{ij}$ відповідно. Неважко показати, що

$$\boldsymbol{M}_i(\lambda) = \boldsymbol{L}_i(\lambda), \quad \boldsymbol{M}_i(\lambda \mid j) = \boldsymbol{L}_i^T(\lambda \mid j) = \boldsymbol{L}_j(\lambda \mid i), \quad \boldsymbol{M}_{ij} = \boldsymbol{L}_{ij}.$$

1.2. Визначники

1.2.1. Визначники другого і третього порядків. Будь-якій квадратній матриці $\boldsymbol{A}$ можна зіставити деяке число, що називається її визначником. Порядком визначника називається порядок матриці $\boldsymbol{A}$, якій він відповідає. Наприклад, визначник матриці другого порядку покладається рівним

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

а визначником матриці третього порядку є величина

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

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

Зазначимо ряд особливостей цих визначників.

1) Кожний доданок у правій частині є добутком такої кількості елементів, яка збігається з порядком матриці $\boldsymbol{A}$.

2) Кожний доданок алгебраїчної суми містить тільки один елемент з кожного рядка і тільки один елемент з кожного стовпця.

3) Кількість доданків у правих частинах дорівнює відповідно $2!$ і $3!$. У загальному випадку їх повинно бути $n!$.

4) Кожний доданок алгебраїчної суми містить елементи, розташовані таким чином, що їх перші індекси утворюють ряд натуральних чисел $1, 2, ..., n$. Другі індекси являють собою різні

переставлення з цих чисел. Оскільки таких переставлень існує $n!$, то ми одержимо $n!$ різних доданків, як зазначено в попередньому пункті.

5) Половина всіх доданків додатна, а половина від'ємна. Кажуть, що в переставленні сукупності натуральних чисел $1, 2, ..., n$ має місце інверсія, якщо серед двох чисел більше передує меншому. Кількістю інверсій у переставленні n натуральних чисел є кількість пар чисел, не обов'язково сусідніх, в яких більше число передує меншому. Переставлення називається парним (непарним), якщо кількість інверсій у ньому парна (непарна). Парним переставленням відповідає знак плюс, а непарним — мінус. Наприклад, другий доданок у формулі для визначника третього порядку складають елементи, другими індексами яких є числа $2, 1, 3$. Тут є тільки одна інверсія — число 2 стоїть раніше числа 1. Тому це переставлення є непарним і другий доданок має знак мінус. У наступному доданку другими індексами є числа $2, 3, 1$. Тут є дві інверсії, бо 2 стоїть раніше 1 і 3 теж розміщено раніше 1. Отже, цей доданок має знак плюс.

Ці п'ять особливостей обумовлюють загальний вигляд *визначника* матриці $\boldsymbol{A}$ порядку n:

$$\det \boldsymbol{A} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (\pm a_{1i} \ a_{2j} \ \dots \ a_{nm}) . \qquad (1)$$

Тут сума береться за всіма переставленнями других індексів множників, причому знак плюс мають доданки з парними переставленнями, а знак мінус — доданки з непарними переставленнями.

1.2.2. Властивості визначників. З'ясуємо властивості визначників n-го порядку. Ми покажемо, що: 1) визначник не змінюється при транспонуванні матриці; 2) переставлення будь-яких стовпців змінює знак визначника; 3) визначник з двома рівними стовпцями дорівнює нулю; 4) загальний множник усіх елементів деякого стовпця визначника можна винести за знак визначника.

Для доведення властивості 1) покладемо $\boldsymbol{B} = \boldsymbol{A}^T$ і на прикладі матриці третього порядку покажемо, як пов'язані визначники $\det \boldsymbol{B}$ і $\det \boldsymbol{A}$. Наприклад, доданок $b_{13}b_{21}b_{32}$ у розкладі визначника $\det \boldsymbol{B}$

відповідає доданку $a_{31}a_{12}a_{23}$ визначника $\det \boldsymbol{A}$. Знак цього доданка в $\det \boldsymbol{B}$ визначається кількістю інверсій у переставленні $3, 1, 2$ і є додатним. Знак доданка $a_{31}a_{12}a_{23}$ можна визначити, якщо розмістити множники в порядку збільшення перших індексів. Оскільки в переставленні $2, 3, 1$ є дві інверсії, то доданок $a_{31}a_{12}a_{23}$ буде мати той же знак у розкладі $\det \boldsymbol{A}$, що і в розкладі $\det \boldsymbol{B}$. Ці міркування можна навести для будь-якого доданка. Доведена властивість дозволяє обмежитись у подальшому формулюванням властивостей визначника стосовно лише стовпців.

Для доведення властивості 2) зауважимо, що, помінявши місцями два сусідні стовпці, ми перетворимо парне переставлення в непарне або навпаки. Припустимо, що ми поміняли j-й та k-й стовпці, причому $j < k$ і між числами j та k розміщені m інших чисел. Щоб досягти цього, треба перемістити k-й стовпець ліворуч, виконавши $m+1$ переставлень з сусідніми стовпцями, а j-й стовпець — праворуч після переставлення з сусідніми стовпцями m разів. Усього кількість інверсій зміниться $2m+1$ разів, тобто парність переставлень з n чисел зміниться і тому знак усіх доданків у розкладі (1) зміниться на протилежний.

З доведеної властивості випливає властивість 3), оскільки переставлення однакових стовпців змінює знак визначника, а, з іншого боку, визначник не змінюється, тобто $\det \boldsymbol{A} = -\det \boldsymbol{A}$, звідки $\det \boldsymbol{A} = 0$.

Властивість 4) випливає з того, що кожний доданок у розкладі (1) містить лише один елемент кожного стовпця.

1.2.3. Розклад визначника за елементами стовпця (рядка). *Мінором* елемента a_{ij} визначника $\det \boldsymbol{A}$ (або матриці $\boldsymbol{A}$) порядку n називається визначник порядку $n-1$, що отримується викреслюванням у матриці $\boldsymbol{A}$ стовпця і рядка, в яких розміщений цей елемент. Мінор елемента a_{ij} ми будемо позначати через M_{ij}. Величина $A_{ij} = (-1)^{i+j} M_{ij}$ називається *алгебраїчним доповненням* елемента a_{ij}.

Теорема 1.2.1. *Визначник дорівнює сумі добутків елементів деякого стовпця на їх алгебраїчні доповнення.*

Д о в е д е н н я. Якщо в розкладі (1) зібрати всі доданки, що містять елемент a_{11} як множник, то коефіцієнтом при a_{11} буде величина $\sum(\pm a_{2j}\, a_{3k}\, ...\, a_{nm})$, де підсумовування виконується за всіма $(n-1)!$ переставленнями чисел $2, 3, \dots, n$. Оскільки елемент a_{11} має одиничні індекси, то його добавлення множником до будь-якого доданка розкладу $\sum(\pm a_{2j}\, a_{3k}\, ...\, a_{nm})$ не впливає на знак цього доданка. Тому цей знак визначається переставленням $j, k, \dots, m$ натуральних чисел $2, 3, \dots, n$. Отже,

$$\sum(\pm a_{2j}\, a_{3k}\, ...\, a_{nm}) = \begin{vmatrix} a_{22} & ... & a_{2n} \\ ... & ... & ... \\ a_{n2} & ... & a_{nn} \end{vmatrix} = M_{11}.$$

З'ясуємо, чому дорівнює сума доданків у розкладі визначника $\det \boldsymbol{A}$, що містять як множник елемент a_{ij}. Розглянемо матрицю $\boldsymbol{B}$, яку одержимо з $\boldsymbol{A}$ таким способом: 1) міняємо місцями i-й рядок і попередній доти, поки після $i-1$ замін i-й рядок не стане першим; 2) міняємо місцями j-й стовпець і попередній доти, поки після $j-1$ замін j-й стовпець не опиниться на першому місці. Всього буде здійснено $i+j-2$ замін і визначник $\det \boldsymbol{B}$ на підставі властивості 2) визначників буде дорівнювати

$$\det \boldsymbol{B} = (-1)^{i+j-2} \det \boldsymbol{A} = (-1)^{i+j} \det \boldsymbol{A}\,. \tag{2}$$

Сума всіх доданків розкладу визначника $\det \boldsymbol{B}$, що містять елемент b_{11}, дорівнює $b_{11}N_{11}$, де N_{11} — мінор елемента b_{11}. Множення доданків розкладу визначника $\det \boldsymbol{B}$ на $(-1)^{i+j}$ перетворює їх у відповідності до формули (2) в доданки розкладу визначника $\det \boldsymbol{A}$, тому сума всіх доданків розкладу $\det \boldsymbol{A}$, що містять a_{ij}, дорівнює величині $(-1)^{i+j} a_{ij} M_{ij} = a_{ij} A_{ij}$.

Розглянемо вираз $a_{1j}A_{1j} + a_{2j}A_{2j} + ... + a_{nj}A_{nj}, \quad j = \overline{1,\ n}$. Він містить $(n-1)!n = n!$ доданків розкладу визначника $\det \boldsymbol{A}$. Всі доданки різні і мають ті ж знаки, що і в розкладі $\det \boldsymbol{A}$. Але в розкладі $\det \boldsymbol{A}$ є всього $n!$ різних доданків. Отже,

$$\det \boldsymbol{A} = \sum_{i=1}^{n} a_{ij} A_{ij}, \quad j = \overline{1,\ n}. \tag{3}$$

Ця формула дає розклад визначника $\det \boldsymbol{A}$ за елементами j-го стовпця. Аналогічно визначник може бути представлений як сума добутків елементів деякого рядка на їх алгебраїчні доповнення.

Висновок 1. *Сума добутків елементів деякого стовпця визначника на алгебраїчні доповнення іншого стовпця дорівнює нулю.*

Д о в е д е н н я. Оскільки $\sum_{i=1}^{n} a_{ik} A_{ij}$ є розкладом визначника з двома однаковими стовпцями, то

$$\sum_{i=1}^{n} a_{ik} A_{ij} = 0, \quad k \neq j. \tag{4}$$

Висновок 2. *Визначник не змінюється, якщо до будь-якого стовпця додати інший стовпець, помножений на деяке число.*

Д о в е д е н н я. Нехай $\det \boldsymbol{B}$ — визначник, що отримується з визначника $\det \boldsymbol{A}$ заміною j-го стовпця стовпцем елементів вигляду $a_{ij} + \lambda a_{ik}, \quad k \neq j$. На підставі теореми і формули (4) маємо:

$$\det \boldsymbol{B} = \sum_{i=1}^{n} (a_{ij} + \lambda a_{ik}) A_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij} + \lambda \sum_{i=1}^{n} a_{ik} A_{ij} = \det \boldsymbol{A}\,. \ \square$$

Якщо застосувати формулу (3) до трикутної матриці, то одержимо, що її визначник дорівнює добутку діагональних елементів.

Формули (3) і (4) можна об’єднати і записати так:

$$\sum_{i=1}^{n} a_{ik} A_{ij} = \delta_{kj} \det \boldsymbol{A}\,, \tag{5}$$

де через δ_{kj} позначений *символ Кронекера,* який дорівнює

$$\delta_{kj} = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases}$$

Аналогічно, розкладаючи визначник за елементами рядка, маємо:

$$\sum_{i=1}^{n} a_{ki} A_{ji} = \delta_{kj} \det \boldsymbol{A}. \tag{6}$$

1.3. Обернена матриця

1.3.1. Визначення оберненої матриці. В алгебрі чисел для $\alpha \neq 0$ мають місце рівності $\alpha\alpha^{-1} = \alpha^{-1}\alpha = 1$. Покажемо, що в матричній алгебрі за певної умови, аналогічної умові $\alpha \neq 0$, існує матриця $\boldsymbol{A}^{-1}$ така, що

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}. \tag{1}$$

Якщо матриця $\boldsymbol{A}^{-1}$ існує, то її називають *оберненою* матрицею по відношенню до $\boldsymbol{A}$. Очевидно, що $\boldsymbol{A}$ і $\boldsymbol{A}^{-1}$ можуть бути тільки квадратними матрицями однакового порядку. Квадратна матриця $\boldsymbol{A}$ називається *невиродженою* (*виродженою*), якщо $\det \boldsymbol{A} \neq 0$ ($\det \boldsymbol{A} = 0$).

Теорема 1.3.1. *Для невиродженої матриці n-го порядку*

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

обернена матриця дорівнює

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \tag{2}$$

де A_{ij} *— алгебраїчне доповнення елемента* a_{ij}, $i, j = \overline{1,\ n}$.

Д о в е д е н н я. З формул (1.1.1) і (1.2.6) випливає, що

$$\{\boldsymbol{A}\boldsymbol{A}^{-1}\}_{ij} = \sum_{k=1}^{n} a_{ik} \frac{A_{jk}}{\det \boldsymbol{A}} = \delta_{ij}, \quad i,\ j = \overline{1,\ n}.$$

Аналогічно з формул (1.1.1) і (1.2.5) знаходимо:

$$\{\boldsymbol{A}^{-1}\boldsymbol{A}\}_{ij} = \sum_{k=1}^{n} \frac{A_{ki}}{\det \boldsymbol{A}} a_{kj} = \delta_{ij}, \quad i,\ j = \overline{1,\ n}.$$

Отже, елементи матриць $\boldsymbol{A}\boldsymbol{A}^{-1}$ і $\boldsymbol{A}^{-1}\boldsymbol{A}$ дорівнюють відповідним елементам одиничної матриці, тому мають місце рівності (1). □

Знаходження оберненої матриці за допомогою формули (2) пов’язане з трудомісткими обчисленнями, тому ця формула важлива лише в теоретичному відношенні. У наступній главі буде запропонований більш прийнятний спосіб обернення матриці.

1.3.2. Властивості оберненої матриці. Поряд з властивістю (1) зазначимо інші властивості оберненої матриці.

Теорема 1.3.2. *Обернена матриця має такі властивості*:

1) *обернена матриця єдина*;

2) $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$;

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

Д о в е д е н н я. 1) Припустивши, що існує матриця $\boldsymbol{B}$ така, що $\boldsymbol{AB} = \boldsymbol{E}$, одержимо:

$$\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{E} = \boldsymbol{A}^{-1}\boldsymbol{AB} = \boldsymbol{B}.$$

Аналогічно, якщо існує матриця $\boldsymbol{C}$ така, що $\boldsymbol{CA} = \boldsymbol{E}$, маємо рівність $\boldsymbol{C} = \boldsymbol{A}^{-1}$.

2) Оскільки $\boldsymbol{E} = (\boldsymbol{A}^{-1}\boldsymbol{A})^T = \boldsymbol{A}^T(\boldsymbol{A}^{-1})^T$, то на підставі частини 1) теореми маємо рівність $(\boldsymbol{A}^{-1})^T = (\boldsymbol{A}^T)^{-1}$.

3) Оскільки $(\boldsymbol{AB})(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}) = \boldsymbol{E}$, то на підставі частини 1) теореми виконується рівність $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1} = (\boldsymbol{AB})^{-1}$. □

Для з’ясування подальших властивостей оберненої матриці доведемо дві допоміжні теореми.

Теорема 1.3.3. *Будь-яку невироджену матрицю можна представити у вигляді добутку елементарних матриць.*

Д о в е д е н н я. Припустимо, що елемент a_{11} матриці

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

не є нульовим. Помножимо перший рядок матриці $\boldsymbol{A}$ на число a_{11}^{-1}. Така операція є елементарним перетворенням матриці $\boldsymbol{A}$. Його можна виконати множенням зліва матриці $\boldsymbol{A}$ на елементарну матрицю $\boldsymbol{L}_1(a_{11}^{-1})$.

Після цього, додаючи до i-го рядка перший рядок, помножений на $-a_{i1}$, $i = \overline{2,\ n}$, перетворимо в нулі елементи першого стовпця, розміщені нижче його першого елемента. Ці перетворення можна виконати шляхом множення зліва матриці $\boldsymbol{L}_1(a_{11}^{-1})\boldsymbol{A}$ на елементарні матриці $\boldsymbol{L}_i(-a_{i1} \mid 1)$, $i = \overline{2,\ n}$. У результаті вказаних перетворень буде одержана матриця

$$\begin{pmatrix} 1 & a'_{12} & \dots & a'_{1n} \\ 0 & a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a'_{n2} & \dots & a'_{nn} \end{pmatrix} \tag{3}$$

Може так статися, що $a_{11} = 0$. Проте в першому стовпці є принаймні один ненульовий елемент, бо в протилежному випадку $\det \boldsymbol{A} = 0$, що неможливо через невиродженість матриці $\boldsymbol{A}$. Нехай a_{k1} є ненульовим елементом першого стовпця. Поміняємо місцями перший і k-й рядки. Це рівносильно множенню зліва матриці $\boldsymbol{A}$ на елементарну матрицю $\boldsymbol{L}_{1k}$. Після застосування до матриці $\boldsymbol{L}_{1k}\boldsymbol{A}$ таких же перетворень, що і для випадку $a_{11} \neq 0$, знову одержимо матрицю вигляду (3).

Ця матриця є невиродженою, бо ненульовий визначник матриці $\boldsymbol{A}$ не може стати нульовим при застосуванні до матриці $\boldsymbol{A}$ елементарних перетворень. З невиродженості матриці (3) випливає невиродженість матриці

$$\begin{pmatrix} a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots \\ a'_{n2} & \dots & a'_{n2} \end{pmatrix}.$$

розміщеної в правому нижньому куті матриці (3). Завдяки цій обставині ми можемо за допомогою елементарних перетворень з рядками матриці (3), розміщеними нижче першого рядка, звести цю матрицю до такого вигляду, коли на місці a'_{22} буде одиниця, а нижче цього елемента — нулі. Продовжуючи цей процес, ми будемо мати наприкінці верхню трикутну матрицю $\boldsymbol{R}$ з одиницями на головній діагоналі.

Подальші перетворення спрямовані на одержання нулів вище головної діагоналі. Для цього в останньому стовпці матриці $\boldsymbol{R}$ одержимо нулі вище діагонального елемента, додаючи до рядків матриці $\boldsymbol{R}$ останній її рядок, помножений на відповідні числа. Далі перетворимо в нулі елементи передостаннього стовпця, розміщені вище діагонального елемента, і т.д. Врешті решт ми одержимо одиничну матрицю. Нагадаємо, що зазначені перетворення рівносильні множенню зліва матриці $\boldsymbol{A}$ на послідовність елементарних матриць. Позначивши їх добуток через $\boldsymbol{L}$, одержимо, що

$$\boldsymbol{LA} = \boldsymbol{E}. \qquad (4)$$

Покажемо, що матриця $\boldsymbol{L}^{-1}$ існує і є добутком елементарних матриць.

Перш за все зауважимо, що з властивостей визначників випливають рівності

$$\det \boldsymbol{L}_i(\lambda) = \lambda, \quad \det \boldsymbol{L}_i(\lambda \mid j) = 1, \quad \det \boldsymbol{L}_{ij} = -1. \qquad (5)$$

Це означає, що всі елементарні матриці є невиродженими. Обернені до них матриці одержуються з одиничної матриці «оберненими» елементарними перетвореннями:

$$\boldsymbol{L}_i^{-1}(\lambda) = \boldsymbol{L}_i(\lambda^{-1}), \quad \boldsymbol{L}_i^{-1}(\lambda \mid j) = \boldsymbol{L}_i(-\lambda \mid j), \quad \boldsymbol{L}_{ij}^{-1} = \boldsymbol{L}_{ij}.$$

Отже, матриці, обернені до елементарних, є елементарними. Тому при оберненні матриці $\boldsymbol{L}$ за допомогою частини 3) теореми 1.3.2 ми одержимо матрицю $\boldsymbol{L}^{-1}$ як добуток елементарних матриць.

Тепер для завершення доведення помножимо зліва обидві частини рівності (4) на $\boldsymbol{L}^{-1}$. Тоді будемо мати:

$$\boldsymbol{A} = \boldsymbol{L}^{-1},$$

тобто матриця $\boldsymbol{A}$ є добутком елементарних матриць. □

Теорема 1.3.4. *Якщо принаймні одна з матриць* $\boldsymbol{A}$, $\boldsymbol{B}$ *є невиродженою, то*

$$\det(\boldsymbol{AB}) = \det \boldsymbol{A} \cdot \det \boldsymbol{B}. \qquad (6)$$

Д о в е д е н н я. З властивостей визначників і рівностей (5) маємо:

$$\det(\boldsymbol{L}_i(\lambda)\boldsymbol{B}) = \lambda\det\boldsymbol{B} = \det\boldsymbol{L}_i(\lambda)\cdot\det\boldsymbol{B},$$

$$\det(\boldsymbol{L}_i(\lambda \mid j)\boldsymbol{B}) = \det\boldsymbol{B} = \det\boldsymbol{L}_i(\lambda \mid j)\cdot\det\boldsymbol{B},$$

$$\det(\boldsymbol{L}_{ij}\boldsymbol{B}) = -\det\boldsymbol{B} = \det\boldsymbol{L}_{ij}\cdot\det\boldsymbol{B}.$$

Отже, теорема справедлива, якщо вважати матрицю $\boldsymbol{A}$ рівною одній з елементарних матриць. Наведені міркування легко поширюються на випадок, коли $\boldsymbol{A}$ — добуток елементарних матриць. З урахуванням теореми 1.3.3 це означає, що рівність (6) виконується при невиродженій матриці $\boldsymbol{A}$. Аналогічно можна довести теорему, якщо невиродженою є матриця $\boldsymbol{B}$. □

Тепер ми маємо можливість завершити розгляд властивостей оберненої матриці.

Теорема 1.3.5. *Для невиродженої матриці* $\boldsymbol{A}$ *виконуються рівності*:

1) $\det(\boldsymbol{A}^{-1}) = (\det\boldsymbol{A})^{-1}$;

2) $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.

Д о в е д е н н я. 1) На підставі теореми 1.3.4 з рівності $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{E}$ випливає, що $\det\boldsymbol{A}\cdot\det(\boldsymbol{A}^{-1}) = 1$, тобто має місце властивість 1).

2) З першої частини теореми випливає, що матриця $\boldsymbol{A}^{-1}$ є невиродженою. Помноживши зліва обидві частини рівності $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}$ на $(\boldsymbol{A}^{-1})^{-1}$, одержимо рівність $\boldsymbol{A} = (\boldsymbol{A}^{-1})^{-1}$. □

1.3.3. Формули Крамера. У подальшому ми неодноразово будемо мати справу з *системою лінійних алгебраїчних рівнянь* відносно *невідомих* $x_1, x_2, ..., x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n. \end{aligned} \tag{7}$$

Числа $a_{11}, a_{12}, ..., a_{nn}$ називаються *коефіцієнтами* системи, числа $b_1, b_2, ..., b_n$ — її *правими частинами*. Систему (7) записують також у вигляді:

$$\boldsymbol{Ax} = \boldsymbol{b}\,, \tag{8}$$

де

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix}.$$

Систему рівнянь (8) ми будемо часто розглядати як рівняння відносно невідомого вектора $\boldsymbol{x}$.

Якщо матриця $\boldsymbol{A}$, яку називають *матрицею системи* (7), є невиродженою, то з рівності (8) випливає, що $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$. Звідси з урахуванням формули (2) маємо:

$$x_i = \frac{1}{\det \boldsymbol{A}} \sum_{j=1}^{n} A_{ji} b_j, \quad i = \overline{1,\ n}\,.$$

Сума в останній формулі збігається з розкладом визначника $\det \boldsymbol{A}$ за елементами i-го стовпця, якщо елементи цього стовпця замінені правими частинами $b_1,\ b_2,\ \dots,\ b_n$. Отже, розв’язок системи (7) з невиродженою матрицею $\boldsymbol{A}$ визначається за *формулами Крамера*:

$$x_i = \frac{\Delta_i}{\det \boldsymbol{A}}, \quad i = \overline{1,\ n},$$

де Δ_i — визначник, що одержується з $\det \boldsymbol{A}$ заміною i-го стовпця матрицею-стовпцем $\boldsymbol{b}$.

Формулами Крамера рідко користуються при розв’язанні систем лінійних рівнянь через велику трудомісткість обчислень. Проте в теоретичному відношенні формули Крамера можуть бути корисними при з’ясуванні залежності розв’язку від елементів матриці $\boldsymbol{A}$ і правих частин.

1.4. Блокові матриці

1.4.1. Операції над блоковими матрицями. Часто доводиться мати справу з матрицями, поділеними вертикальними і горизонтальними прямими на прямокутні частини — «*блоки*», або «*підматриці*». Зокрема, будь-яку матрицю можна уявляти як блокову матрицю-стовпець, у якої підматрицями є рядки матриці, або як блокову матрицю-рядок, у якої підматрицями є її стовпці. Очевидно, що множення блокової матриці на число зводиться до множення на це число кожного блоку, а сумою двох матриць однакових розмірів і з однаковим розбиттям на блоки є блокова матриця, кожний блок якої дорівнює сумі відповідних блоків даних матриць.

Для можливості «блокового» множення матриць ми будемо вважати, що при розбитті на блоки всі горизонтальні розміри в першому множнику збігаються з відповідними вертикальними розмірами в другому.

Розглянемо, наприклад, $m \times n$ - матрицю $\boldsymbol{A}$ і $n \times p$ - матрицю $\boldsymbol{B}$ з таким розбиттям на блоки:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix}.$$

Будемо вважати, що блоки $\boldsymbol{A}_{11}$ і $\boldsymbol{B}_{11}$ мають розміри $q \times r$ і $r \times s$ відповідно. Тоді блоки $\boldsymbol{A}_{12}$ і $\boldsymbol{B}_{21}$ будуть мати розміри $q \times (n-r)$ і $(n-r) \times s$ відповідно. Покажемо, що

$$\boldsymbol{AB} = \begin{pmatrix} \boldsymbol{A}_{11}\boldsymbol{B}_{11} + \boldsymbol{A}_{12}\boldsymbol{B}_{21} & \boldsymbol{A}_{11}\boldsymbol{B}_{12} + \boldsymbol{A}_{12}\boldsymbol{B}_{22} \\ \boldsymbol{A}_{21}\boldsymbol{B}_{11} + \boldsymbol{A}_{22}\boldsymbol{B}_{21} & \boldsymbol{A}_{21}\boldsymbol{B}_{12} + \boldsymbol{A}_{22}\boldsymbol{B}_{22} \end{pmatrix}. \tag{1}$$

Щоб перевірити це співвідношення, скористаємося формулою (1.1.1) при $i = \overline{1,\ q}, \quad j = \overline{1,\ s}$:

$$\{\boldsymbol{AB}\}_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj} + \sum_{k=r+1}^{n} a_{ik}b_{kj}. \tag{2}$$

Перша сума правої частини рівності (2) є елементом i-го рядка і j-го стовпця добутку $\boldsymbol{A}_{11}\boldsymbol{B}_{11}$, друга — елементом i-го рядка і j-го стовпця добутку $\boldsymbol{A}_{12}\boldsymbol{B}_{21}$. Отже, вся права частина (2) є елементом i-го рядка і j-го стовця суми $\boldsymbol{A}_{11}\boldsymbol{B}_{11} + \boldsymbol{A}_{12}\boldsymbol{B}_{21}$. Аналогічно можна показати, що й інші підматриці добутку $\boldsymbol{AB}$ можна записати у відповідності до формули (1). Цю формулу можна поширити на іншу кількість підматриць у множників $\boldsymbol{A}$ і $\boldsymbol{B}$.

Отже, операції над блоковими матрицями виконуються за тими ж правилами, що і у випадку, коли замість блоків маємо числові елементи. Якщо, наприклад, вважати матрицю системи (1.3.8) блоковою матрицею-рядком, у якої підматрицями є її стовпці $\boldsymbol{a}_{*1}, \ldots, \boldsymbol{a}_{*n}$, то вказану систему можна записати у вигляді:

$$\sum_{i=1}^{n} x_i \boldsymbol{a}_{*i} = b.$$

Квадратна блокова матриця називається *квазідіагональною*, якщо її діагональні блоки є квадратними, а позадіагональні — нульовими. Діагональна матриця є окремим випадком квазідіагональної матриці. Очевидно, що при множенні казідіагональних матриць $\boldsymbol{A}$, $\boldsymbol{B}$ з однаковими порядками відповідних діагональних блоків добуток $\boldsymbol{AB}$ також має квазідіагональну форму, причому діагональні блоки матриці $\boldsymbol{AB}$ дорівнюють добуткам відповідних діагональних блоків множників $\boldsymbol{A}$ і $\boldsymbol{B}$.

1.4.2. Обчислення визначників і обернення матриць за допомогою розбиття на підматриці. У ряді випадків матрицю можна поділити на підматриці, які мають певні особливості (нульові, одиничні підматриці). Наявність таких особливостей може полегшити обчислення визначника і оберненої матриці, якщо скористатися операціями над блоковими матрицями.

Знайдемо вираз для визначника довільної квадратної матриці

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}$$

з квадратними підматрицями $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$.

При $\det \boldsymbol{A}_{11} \neq 0$ безпосередньою перевіркою неважко показати, що

$$\boldsymbol{B}_1 \boldsymbol{A} \boldsymbol{C}_1 = \boldsymbol{D}_1, \qquad (3)$$

де

$$\boldsymbol{B}_1 = \begin{pmatrix} \boldsymbol{E} & \boldsymbol{0} \\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{E} \end{pmatrix}, \quad \boldsymbol{C}_1 = \begin{pmatrix} \boldsymbol{E} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ \boldsymbol{0} & \boldsymbol{E} \end{pmatrix},$$

$$\boldsymbol{D}_1 = \begin{pmatrix} \boldsymbol{A}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \end{pmatrix}.$$

Обчислимо визначники матриць $\boldsymbol{B}_1, \boldsymbol{C}_1, \boldsymbol{D}_1$. Якщо обчислювати $\det \boldsymbol{B}_1$ розкладанням за елементами останнього стовпця, то одержимо добуток одиниці на визначник такої ж форми, але меншого на 1 порядку. Продовжуючи ці дії далі, одержимо, що $\det \boldsymbol{B}_1 = 1$. Аналогічно доводиться, що $\det \boldsymbol{C}_1 = 1$.

Представимо матрицю $\boldsymbol{D}_1$ у вигляді:

$$\boldsymbol{D}_1 = \begin{pmatrix} \boldsymbol{A}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{E} \end{pmatrix} \begin{pmatrix} \boldsymbol{E} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \end{pmatrix}.$$

У відповідності до теореми 1.3.4 $\det \boldsymbol{D}_1$ дорівнює добутку визначників матриць, які є множниками правої частини останньої формули. Ці визначники, обчислені за допомогою таких же міркувань, що і в разі визначників $\det \boldsymbol{B}_1$ і $\det \boldsymbol{C}_1$, дорівнюють $\det \boldsymbol{A}_{11}$ і $\det(\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})$.

Тепер на підставі теореми 1.3.4 з формули (3) знаходимо, що

$$\det \boldsymbol{A} = \det \boldsymbol{D}_1 = \det \boldsymbol{A}_{11} \cdot \det(\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}). \tag{4}$$

Для обчислення $\boldsymbol{A}^{-1}$ обернемо обидві частини рівності (3), скориставшись частиною 3) теореми 1.3.2. Припустивши, що $\boldsymbol{D}_1^{-1}$ існує, знаходимо, що $\boldsymbol{C}_1^{-1}\boldsymbol{A}^{-1}\boldsymbol{B}_1^{-1} = \boldsymbol{D}_1^{-1}$, звідки

$$\boldsymbol{A}^{-1} = \boldsymbol{C}_1\boldsymbol{D}_1^{-1}\boldsymbol{B}_1. \tag{5}$$

За допомогою формули (1.3.1) неважко переконатися в тому, що при $\det(\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}) \neq 0$ будемо мати:

$$\boldsymbol{D}_1^{-1} = \begin{pmatrix} \boldsymbol{A}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & (\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1} \end{pmatrix}.$$

Аналогічно знаходяться визначник $\det \boldsymbol{A}$ і обернена матриця $\boldsymbol{A}^{-1}$ у випадку, коли $\det \boldsymbol{A}_{22} \neq 0$ і $\det(\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}) \neq 0$. Опускаючи обчислення, наведемо остаточний результат:

$$
\begin{gathered}
\boldsymbol{B}_2\boldsymbol{A}\boldsymbol{C}_2 = \boldsymbol{D}_2, \qquad \boldsymbol{A}^{-1} = \boldsymbol{C}_2\boldsymbol{D}_2^{-1}\boldsymbol{B}_2, \\
\det \boldsymbol{A} = \det \boldsymbol{D}_2 = \det(\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}) \cdot \det \boldsymbol{A}_{22}.
\end{gathered} \tag{6}
$$

Тут

$$
\boldsymbol{B}_2 = \begin{pmatrix} \boldsymbol{E} & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} \\ \boldsymbol{0} & \boldsymbol{E} \end{pmatrix}, \quad \boldsymbol{C}_2 = \begin{pmatrix} \boldsymbol{E} & \boldsymbol{0} \\ -\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & \boldsymbol{E} \end{pmatrix},
$$

$$
\boldsymbol{D}_2 = \begin{pmatrix} \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{22} \end{pmatrix}, \quad \boldsymbol{D}_2^{-1} = \begin{pmatrix} (\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{22}^{-1} \end{pmatrix}.
$$

Вибір формул (4), (5) або, з іншого боку, формул (6) залежить від властивостей матриці $\boldsymbol{A}$ і її підматриць. У кожному окремому випадку використовують ті властивості, які можуть найбільшою мірою спростити обчислення.

Вправи до глави 1

1.1. Довести, що коли $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ для всіх векторів $\boldsymbol{x}$, то $\boldsymbol{A} = \boldsymbol{0}$.

1.2. Довести, що дві матриці є переставними, якщо одна з них є скалярною матрицею.

1.3. Показати, що кожну з елементарних матриць $\boldsymbol{L}_i(\lambda)$, $\boldsymbol{L}_i(\lambda \mid j)$, $\boldsymbol{L}_{ij}$ можна подати у вигляді суми $\boldsymbol{E} + \sigma\boldsymbol{u}\boldsymbol{v}^T$, де σ — число, $\boldsymbol{u}$ і $\boldsymbol{v}$ — вектори. Знайти σ, $\boldsymbol{u}$, $\boldsymbol{v}$ для кожної з указаних матриць.

1.4. Нехай матриця $\boldsymbol{A}$ є верхньою (нижньою) трикутною і має одиничні діагональні елементи. Довести, що матриця $\boldsymbol{A}^{-1}$ має такі ж властивості.

1.5. Показати, що обернення матриці $\boldsymbol{A}$ порядку n можна звести до розв’язання n систем лінійних рівнянь, кожна з яких вміщує n рівнянь відносно n невідомих, причому матрицею кожної системи є $\boldsymbol{A}$.

ГЛАВА 2

МЕТОД ГАУССА

Метод Гаусса є основним методом розв'язання систем лінійних алгебраїчних рівнянь. Існує багато різноманітних обчислювальних схем і модифікацій цього методу. Більш того, назву «метод Гаусса» застосовують до великої групи алгоритмів, пов'язаних не тільки з розв'язанням систем лінійних алгебраїчних рівнянь, але й оберненням матриць, обчисленням визначників, розкладанням матриці на множники і т.д. Основи застосування методу Гаусса до вказаних проблем викладаються в цій главі.

2.1. Метод виключення Гаусса

2.1.1. Розв'язання систем лінійних рівнянь. Звернемося до системи рівнянь вигляду (1.3.7). Розв'язком цієї системи називають упорядковану сукупність значень невідомих, що задовольняють кожному з рівнянь (1.3.7). Дві системи лінійних рівнянь відносно одних і тих же невідомих називаються *еквівалентними*, якщо кожний розв'язок однієї системи є розв'язком іншої системи або обидві вони не мають розв'язків. Очевидно, що ми будемо мати справу з еквівалентними системами, якщо до рівнянь системи застосувати такі перетворення: 1) множення деякого рівняння на число, відмінне від нуля; 2) додавання до деякого рівняння іншого рівняння, помноженого на довільне число; 3) переставлення двох рівнянь. Загальний спосіб розв'язання може полягати в перетворенні системи (1.3.7) до такої еквівалентної системи, для якої розв'язок знаходиться досить просто. Одним з таких способів є *метод виключення Гаусса*.

Розглянемо метод виключення на прикладі системи лінійних рівнянь третього порядку

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 ,$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 ,$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

з невиродженою матрицею системи. Будемо вважати, що $a_{11} \neq 0$, бо в протилежному випадку цього можна домогтися переставленням рівнянь системи. Першим кроком розв'язання є виключення невідомої x_1 з другого і третього рівнянь за допомогою першого рівняння. Для цього треба додати до другого і третього рівнянь перше, помножене на $l_{21} = -a_{21}a_{11}^{-1}$ і $l_{31} = -a_{31}a_{11}^{-1}$ відповідно. У результаті одержимо еквівалентну систему рівнянь

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a'_{22}x_2 + a'_{23}x_3 &= b'_2, \\ a'_{32}x_2 + a'_{33}x_3 &= b'_3. \end{aligned}$$

Коефіцієнт a_{11}, що стоїть при x_1 у першому рівнянні, називається *провідним елементом* першого кроку виключення. На другому кроці виключення ми не займаємо перше рівняння. До двох інших рівнянь можна застосувати процедуру виключення. Провідним елементом для другого кроку є коефіцієнт a'_{22}, який ми будемо вважати ненульовим (при $a'_{22} = 0$ треба переставити друге і третє рівняння). Додавши до третього рівняння друге, помножене на $l_{32} = -a'_{32}(a'_{22})^{-1}$, виключимо x_2 з третього рівняння, після чого одержимо трикутну систему

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a'_{22}x_2 + a'_{23}x_3 &= b'_2, \\ a''_{33}x_3 &= b''_3. \end{aligned}$$

Процес одержання цієї системи називається прямим ходом методу виключення. Порядок дій при подальшому розв'язанні очевидний. З останнього рівняння знаходимо x_3. Підставивши це значення в передостаннє рівняння, знаходимо x_2 і, нарешті, після підставлення знайдених величин x_2 і x_3 у перше рівняння визначаємо x_1. Процес розв'язання останньої системи називається оберненим ходом методу виключення.

Очевидно, що коли всі провідні елементи a_{11}, a'_{22}, a''_{33} відмінні від нуля, задача має єдиний розв'язок. Саме з цим випадком ми маємо справу, коли матриця системи невироджена. У цьому легко переконатися за допомогою міркувань, використаних при доведенні теореми 1.3.3.

З а у в а ж е н н я. Якщо мова йде при систему $\boldsymbol{Ax}=\mathbf{0}$ з невиродженою матрицею $\boldsymbol{A}$, то єдиним її розв'язком є нульовий розв'язок $\boldsymbol{x}=\mathbf{0}$, а при наявності нульових провідних елементів частина рівнянь перетворюється в тотожності і система буде мати меншу кількість рівнянь ніж невідомих. Про цей випадок мова буде йти в розділі 3.5. Тут ми лише зазначимо той очевидний факт, що крім нульового система буде мати ненульові розв'язки.

2.1.2. Метод Гаусса — Жордана. Матриця $\boldsymbol{A}_b=(\boldsymbol{A} \quad \boldsymbol{b})$, яка одержується з матриці системи приєднанням до неї стовпця правих частин, називається *розширеною* матрицею системи $\boldsymbol{Ax}=\boldsymbol{b}$. Указаним вище перетворенням системи до еквівалентних систем відповідають елементарні перетворення її розширеної матриці $\boldsymbol{A}_b$, які, у свою чергу, рівносильні множенню зліва матриці $\boldsymbol{A}_b$ на елементарні матриці.

Застосуємо наведені міркування до обґрунтування однієї з найбільш важливих схем методу виключення, якою є *метод Гаусса — Жордана*. Стосовно розв'язання системи лінійних рівнянь з невиродженою матрицею цей метод полягає в перетворенні розширеної матриці $\boldsymbol{A}_b$ до такого вигляду, коли на місці $\boldsymbol{A}$ виникає діагональна матриця з провідними елементами на головній діагоналі або навіть одинична матриця. В останньому випадку розв'язок системи буде знаходитися в останньому стовпці перетвореної матриці. Указані перетворення можна здійснити подібно до того, як це було зроблено при доведенні теореми 1.3.3. Розглянемо

П р и к л а д 1. Нехай треба розв'язати методом Гаусса — Жордана систему рівнянь $\boldsymbol{Ax} = \boldsymbol{b}$, де

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 \\ 3 & -1 & 4 \\ 2 & 2 & 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}.$$

Розширену матрицю системи $\boldsymbol{A}_b = (\boldsymbol{A} \quad \boldsymbol{b})$ за допомогою елементарних перетворень з рядками зведемо до такого вигляду, щоб на місці $\boldsymbol{A}$ виникла одинична матриця. Тоді на місці $\boldsymbol{b}$ буде розв'язок системи. Необхідні перетворення матриці $\boldsymbol{A}_b$ можна написати у вигляді:

$$\boldsymbol{A}_b = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 3 & -1 & 4 & 4 \\ 2 & 2 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & -7 & 1 & -5 \\ 0 & -2 & -1 & -4 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -1/7 & 5/7 \\ 0 & -2 & -1 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 9/7 & 11/7 \\ 0 & 1 & -1/7 & 5/7 \\ 0 & 0 & -9/7 & -18/7 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 9/7 & 11/7 \\ 0 & 1 & -1/7 & 5/7 \\ 0 & 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$$

Отже, розв'язок системи рівнянь дорівнює $x_1 = -1$, $x_2 = 1$, $x_3 = 2$. $\triangle$

Метод Гаусса — Жордана є ефективним засобом обчислення оберненої матриці. Для обернення матриці $\boldsymbol{A}$ складемо блокову матрицю $(\boldsymbol{A} \quad \boldsymbol{E})$ і застосуємо до неї елементарні перетворення з рядками таким чином, щоб на місці $\boldsymbol{A}$ виникла матриця $\boldsymbol{E}$. Покажемо, що при цьому на місці $\boldsymbol{E}$ в блоковій матриці буде $\boldsymbol{A}^{-1}$.

Дійсно, застосування елементарних перетворень рівносильне множенню зліва матриці $(\boldsymbol{A} \quad \boldsymbol{E})$ на деяку невироджену матрицю $\boldsymbol{L}$. Отже,

$L(A \quad E)=(LA \quad L)$. Оскільки за припущенням $LA=E$, то $L=A^{-1}$ і ми маємо рівність $L(A \quad E)=(E \quad A^{-1})$, що і треба довести.

П р и к л а д 2. Виконаємо обернення матриці

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

методом Гаусса-Жордана. Блокову матрицю (*A* *E*) за допомогою елементарних перетворень з рядками зведемо до такого вигляду, коли на місці *A* буде матриця *E*. Тоді на місці *E* буде A^{-1}. Отже,

$$(A \quad E)=\begin{pmatrix} 0 & 3 & 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 0 & 1 & 0 \\ 3 & -2 & 0 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 1 & 1 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 1 & 1/2 & 0 & 1/2 & 0 \\ 0 & 3 & 1 & 1 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1/2 & 0 & 1/2 & 0 \\ 0 & 3 & 1 & 1 & 0 & 0 \\ 0 & -5 & -3/2 & 0 & -3/2 & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 1 & 1/2 & 0 & 1/2 & 0 \\ 0 & 1 & 1/3 & 1/3 & 0 & 0 \\ 0 & -5 & -3/2 & 0 & -3/2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1/6 & -1/3 & 1/2 & 0 \\ 0 & 1 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/6 & 5/3 & -3/2 & 1 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 1/6 & -1/3 & 1/2 & 0 \\ 0 & 1 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1 & 10 & -9 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -2 & 2 & -1 \\ 0 & 1 & 0 & -3 & 3 & -2 \\ 0 & 0 & 1 & 10 & -9 & 6 \end{pmatrix}.$$

Таким чином,

$$A^{-1}=\begin{pmatrix} -2 & 2 & -1 \\ -3 & 3 & -2 \\ 10 & -9 & 6 \end{pmatrix}. \ \triangle$$

2.1.3. Обчислення визначників. Метод Гаусса обчислення визначників заснований на таких міркуваннях. Нехай елемент a_{ij} матриці $\boldsymbol{A}$ порядку n є ненульовим. Якщо до k-го рядка, $k \neq i$, додати i-й рядок, помножений на довільне число α_k, то визначник $\det \boldsymbol{A}$ не зміниться. Будемо вважати, що $\alpha_k = -a_{kj} a_{ij}^{-1}$, і виконаємо зазначену процедуру для всіх $k \neq i$. Тоді всі елементи j-го стовпця, крім елемента a_{ij}, будуть дорівнювати нулю. Розкладаючи одержаний визначник за елементами j-го стовпця, ми зведемо обчислення визначника порядку n до обчислення одного визначника порядку $n-1$. До цього визначника застосуємо аналогічні міркування і т.д.

П р и к л а д 3. Нехай треба обчислити визначник матриці

$$\boldsymbol{A} = \begin{pmatrix} 3 & 4 & 2 \\ 2 & 0 & 3 \\ -2 & 3 & -2 \end{pmatrix}.$$

Перетворимо визначник до такого вигляду, коли в деякому стовпці всі елементи, крім одного, будуть нульовими. У матриці $\boldsymbol{A}$ вже є один нульовий елемент a_{22}, тому доцільно одержати ще один нуль у другому стовпці. Для цього додамо до третього рядка перший, помножений на –3/4. У результаті одержимо:

$$\det \boldsymbol{A} = \begin{vmatrix} 3 & 4 & 2 \\ 2 & 0 & 3 \\ -17/4 & 0 & -7/2 \end{vmatrix}.$$

Розкладемо визначник за елементами другого стовпця:

$$\det \boldsymbol{A} = -4 \begin{vmatrix} 2 & 3 \\ -17/4 & -7/2 \end{vmatrix}.$$

Обчислюючи визначник другого порядку, одержимо:

$$\det \boldsymbol{A} = -4\left(-7 + \frac{51}{4}\right) = -23. \ \triangle$$

2.2. *LR*- і *LDU*-розклади матриці

2.2.1. Визначення *LR*- і *LDU*-розкладів. Повернемося до системи лінійних рівнянь третього порядку, розглянутої на початку розділу 2.1, і прослідкуємо, як перетворюється при розв'язанні системи її матриця. Обмежимося випадком, коли при розв'язанні не виникає потреби переставляти рядки матриці, тобто єдиним елементарним перетворенням, що використовується, є додавання до рядків матриці інших рядків, помножених на деякі числа. Це означає, що матрицю $\boldsymbol{A}_b$ можна перетворити до трикутного вигляду множенням зліва на елементарні матриці $\boldsymbol{L}_i(\lambda \mid j)$. У нашому випадку матрицю $\boldsymbol{A}_b$ треба послідовно помножити на $\boldsymbol{L}_2(l_{21} \mid 1)$, $\boldsymbol{L}_3(l_{31} \mid 1)$, $\boldsymbol{L}_3(l_{32} \mid 2)$, де l_{21}, l_{31}, l_{32} — числа, визначені в розділі 2.1. Позначивши матрицю $\boldsymbol{L}_3(l_{32} \mid 2)\,\boldsymbol{L}_3(l_{31} \mid 1)\,\boldsymbol{L}_2(l_{21} \mid 1)$ через $\boldsymbol{L}^{-1}$, будемо мати:

$$\boldsymbol{L}^{-1}\boldsymbol{A}_b = \boldsymbol{L}^{-1}(\boldsymbol{A} \quad \boldsymbol{b}) = (\boldsymbol{L}^{-1}\boldsymbol{A} \quad \boldsymbol{L}^{-1}\boldsymbol{b}) = (\boldsymbol{R} \quad \boldsymbol{L}^{-1}\boldsymbol{b}).$$

Тут $\boldsymbol{R}$ — верхня трикутна матриця, діагональними елементами якої є провідні елементи методу виключення.

Отже, $\boldsymbol{L}^{-1}\boldsymbol{A} = \boldsymbol{R}$ і, оскільки матриця $\boldsymbol{L}^{-1}$ є невиродженою як добуток невироджених матриць,

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{R}\,. \tag{1}$$

Обчислюючи матрицю $\boldsymbol{L}$, знаходимо, що

$$\boldsymbol{L} = (\boldsymbol{L}_3(l_{32} \mid 2)\boldsymbol{L}_3(l_{31} \mid 1)\boldsymbol{L}_2(l_{21} \mid 1))^{-1} =$$

$$= \boldsymbol{L}_2^{-1}(l_{21} \mid 1)\boldsymbol{L}_3^{-1}(l_{31} \mid 1)\boldsymbol{L}_3^{-1}(l_{32} \mid 2) = \boldsymbol{L}_2(-l_{21} \mid 1)\boldsymbol{L}_3(-l_{31} \mid 1)\boldsymbol{L}_3(-l_{32} \mid 2).$$

Щоб одержати остаточний результат, представимо $\boldsymbol{L}$ у вигляді:

$$\boldsymbol{L} = \boldsymbol{L}_2(-l_{21} \mid 1)\boldsymbol{L}_3(-l_{31} \mid 1)\boldsymbol{L}_3(-l_{32} \mid 2)\,\boldsymbol{E}$$

і прослідкуємо за змінами матриці $\boldsymbol{E}$ при множенні її на елементарні матриці.

При множенні матриці $\boldsymbol{E}$ на $\boldsymbol{L}_3(-l_{32}|\,2)$ до її третього рядка додається другий, помножений на $-l_{32}$. Іншими словами, в матриці $\boldsymbol{E}$ нульовий елемент e_{32} замінюється величиною $-l_{32}$. Подальше множення на $\boldsymbol{L}_3(-l_{31}|1)$ і $\boldsymbol{L}_2(-l_{21}|1)$ додає до третього і другого рядків матриці $\boldsymbol{L}_3(-l_{32}|\,2)\,\boldsymbol{E}$ перший рядок матриці $\boldsymbol{E}$, помножений на $-l_{31}$ і $-l_{21}$ відповідно, тобто нульові елементи першого стовпця замінюються величинами $-l_{31}$ і $-l_{21}$ при незмінних інших величинах. Отже,

$$\boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 \\ -l_{21} & 1 & 0 \\ -l_{31} & -l_{32} & 1 \end{pmatrix}.$$

Аналогічний результат можна одержати для матриці будь-якого порядку.

Таким чином, якщо невироджену матрицю $\boldsymbol{A}$ можна звести до верхньої трикутної матриці $\boldsymbol{R}$ шляхом додавання до рядків матриці $\boldsymbol{A}$ інших її рядків, помножених на відповідні числа, то має місце рівність (1), де $\boldsymbol{L}$ — нижня трикутна матриця з одиницями на головній діагоналі і піддіагональними елементами, рівними взятим з протилежним знаком множникам l_{ij}, з якими j-й рядок матриці $\boldsymbol{A}$ додається до i-го рядка.

Рівність (1) називається ***LR**-розкладом* матриці $\boldsymbol{A}$. Якщо він відомий, то система $\boldsymbol{Ax} = \boldsymbol{b}$ зводиться до двох систем з трикутними матрицями:

$$\boldsymbol{Lc} = \boldsymbol{b}, \quad \boldsymbol{Rx} = \boldsymbol{c},$$

які легко розв'язати.

Обчислення оберненої матриці $\boldsymbol{A}^{-1}$ при відомому **LR**-розкладі суттєво спрощується, оскільки $\boldsymbol{A}^{-1} = \boldsymbol{R}^{-1}\boldsymbol{L}^{-1}$, а обернення трикутних матриць, наприклад, методом Гаусса — Жордана, не є важкою проблемою. Що стосується обчислення визначника, то $\det \boldsymbol{A} = \det \boldsymbol{R}$ і ми можемо знайти визначник як добуток діагональних елементів матриці $\boldsymbol{R}$, тобто $\det \boldsymbol{A} = d_1 d_2 \ldots d_n$.

В одному відношенні ***LR***-розклад є «несиметричним»: на головній діагоналі матриці $\boldsymbol{R}$ розміщені провідні елементи, в той час як у $\boldsymbol{L}$ на цих місцях стоять одиниці. Проте цю «несиметричність» можна усунути, якщо з матриці $\boldsymbol{R}$ «відділити» діагональну матрицю $\boldsymbol{D}$ з провідними елементами $d_1, d_2, \ldots, d_n$ на діагоналі:

$$\boldsymbol{R} = \operatorname{diag}(d_1,\ d_2,\ \ldots,\ d_n)\boldsymbol{U}, \quad \boldsymbol{U} = \begin{pmatrix} 1 & r_{12}d_1^{-1} & r_{13}d_1^{-1} & \ldots & r_{1n}d_1^{-1} \\ 0 & 1 & r_{23}d_2^{-1} & \ldots & r_{2n}d_2^{-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Тепер з ***LR***-розкладу (1) ми одержуємо ***LDU-розклад*** матриці $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{LDU}\,. \tag{2}$$

Покажемо, що ***LDU***-розклад невиродженої матриці є єдиним. Тим самим буде доведено, що і ***LR***-розклад є єдиним.

Теорема 2.2.1*. Якщо **LDU**-розклад невиродженої матриці існує, то він є єдиним.*

Д о в е д е н н я. Нехай $\boldsymbol{A} = \boldsymbol{L}_1\boldsymbol{D}_1\boldsymbol{U}_1 = \boldsymbol{L}_2\boldsymbol{D}_2\boldsymbol{U}_2$ — два розклади матриці $\boldsymbol{A}$. Тут $\boldsymbol{L}_1, \boldsymbol{L}_2$ — нижні трикутні матриці з одиницями на головних діагоналях, $\boldsymbol{U}_1, \boldsymbol{U}_2$ — верхні трикутні матриці з одиницями на головних діагоналях і $\boldsymbol{D}_1, \boldsymbol{D}_2$ — діагональні матриці з ненульовими діагональними елементами.

У подальших міркуваннях ми будемо враховувати, що $\boldsymbol{L}_1^{-1},\ \boldsymbol{L}_2^{-1}$ — нижні трикутні, а $\boldsymbol{U}_1^{-1}, \boldsymbol{U}_2^{-1}$ — верхні трикутні матриці з одиницями на головних діагоналях. У справедливості цих тверджень неважко переконатися за допомогою формули (1.3.2.). Крім того, легко перевірити за допомогою формули (1.3.1.), що будь-яка діагональна матриця з ненульовими діагональними елементами має обернену, причому

$$(\operatorname{diag}(d_1, d_2,\ \ldots, d_n))^{-1} = \operatorname{diag}(d_1^{-1},\ d_2^{-1}, \ldots, d_n^{-1})\,. \tag{3}$$

Повертаючись до рівності $\boldsymbol{L}_1\boldsymbol{D}_1\boldsymbol{U}_1 = \boldsymbol{L}_2\boldsymbol{D}_2\boldsymbol{U}_2$, помножимо її зліва на матрицю $\boldsymbol{D}_1^{-1}\boldsymbol{L}_1^{-1}$ і справа на $\boldsymbol{U}_2^{-1}$. Очевидно, що після цього ми будемо мати:

$$\boldsymbol{U}_1\boldsymbol{U}_2^{-1} = \boldsymbol{D}_1^{-1}\boldsymbol{L}_1^{-1}\boldsymbol{L}_2\boldsymbol{D}_2. \qquad (4)$$

Ліва частина рівності (4) є добутком двох верхніх трикутних матриць з одиницями на головній діагоналі, а права частина — добутком нижніх трикутних матриць $\boldsymbol{D}_1^{-1}\boldsymbol{L}_1^{-1}$ і $\boldsymbol{L}_2\boldsymbol{D}_2$. У розділі 1.1 було з'ясовано, що добуток двох верхніх (нижніх) трикутних матриць є верхньою (нижньою) трикутною матрицею і що діагональні елементи добутку дорівнюють добуткам відповідних діагональних елементів множників. Отже, ліва частина рівності (4) є верхньою трикутною матрицею з одиницями на головній діагоналі, а права частина — нижньою трикутною матрицею. Єдиною матрицею, що одночасно є верхньою трикутною і нижньою трикутною і має одиниці на головній діагоналі, є одинична матриця. Таким чином, з рівності (4) маємо:

$$\boldsymbol{U}_1\boldsymbol{U}_2^{-1} = \boldsymbol{E}, \qquad \boldsymbol{D}_1^{-1}\boldsymbol{L}_1^{-1}\boldsymbol{L}_2\boldsymbol{D}_2 = \boldsymbol{E}. \qquad (5)$$

З першої рівності (5) знаходимо, що $\boldsymbol{U}_1 = \boldsymbol{U}_2$, а другу можна перетворити до вигляду:

$$\boldsymbol{L}_1^{-1}\boldsymbol{L}_2 = \boldsymbol{D}_1\boldsymbol{D}_2^{-1}. \qquad (6)$$

Відносно рівності (6) можна застосувати міркування, подібні до використаних при розгляді рівності (4). Це дає можливість стверджувати, що $\boldsymbol{L}_1^{-1}\boldsymbol{L}_2 = \boldsymbol{E},\ \boldsymbol{D}_1\boldsymbol{D}_2^{-1} = \boldsymbol{E}$, тобто

$$\boldsymbol{L}_1 = \boldsymbol{L}_2, \qquad \boldsymbol{D}_1 = \boldsymbol{D}_2.$$

Висновок. ***LDU****-розклад симетричної матриці* ***A*** *визначається рівностями*

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^T = \boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U}. \qquad (7)$$

Д о в е д е н н я. Оскільки $\boldsymbol{A} = \boldsymbol{A}^T$, то $\boldsymbol{LDU} = \boldsymbol{U}^T\boldsymbol{DL}^T$. Тепер з єдиності $\boldsymbol{LDU}$-розкладу випливає, що $\boldsymbol{L} = \boldsymbol{U}^T$, тобто мають місце рівності (7). □

П р и к л а д 1. Знайдемо $\boldsymbol{LR}$- і $\boldsymbol{LDU}$-розклади матриць

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 2 \\ 4 & -1 & 6 \\ -4 & -2 & -2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 3 & 2 \\ -2 & 2 & 11 \end{pmatrix}.$$

Будемо перетворювати матриці до трикутного вигляду шляхом додавання до рядків матриць інших їх рядків, помножених на деякі числа. Для матриці $\boldsymbol{A}$ виконаємо ці перетворення в такій послідовності.

Додамо до другого рядка матриці $\boldsymbol{A}$ перший рядок, помножений на –2, і до третього — перший, помножений на 2. Тоді будемо мати:

$$\boldsymbol{A} \sim \begin{pmatrix} 2 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & -2 & 2 \end{pmatrix}$$

При цьому множники –2 і 2, взяті з протилежними знаками, є елементами l_{21} і l_{31} матриці $\boldsymbol{L}$, тобто $l_{21} = 2,\ l_{31} = -2.$ Далі додамо до третього рядка одержаної матриці другий рядок, помножний на –2. У результаті будемо мати:

$$\boldsymbol{A} \sim \begin{pmatrix} 2 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & -2 \end{pmatrix},$$

причому $l_{32} = 2$. Отже,

$$\boldsymbol{R} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & -2 \end{pmatrix}, \quad \boldsymbol{L} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix}.$$

Матрицю $\boldsymbol{R}$ можна представити у вигляді добутку діагональної матриці $\boldsymbol{D}$ і верхньої трикутної матриці $\boldsymbol{U}$ з одиницями на головній діагоналі:

$$\boldsymbol{R} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{DU}.$$

Тепер $\boldsymbol{LR}$- і $\boldsymbol{LDU}$-розклади матриці $\boldsymbol{A}$ мають вигляд:

$$\boldsymbol{A} = \boldsymbol{LR} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & -2 \end{pmatrix},$$

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

Матриця $\boldsymbol{B}$ є симетричною, тому для неї $\boldsymbol{LDU}$-розклад визначається рівністю $\boldsymbol{B} = \boldsymbol{U}^T\boldsymbol{DU}$. Він не містить в явному вигляді матрицю $\boldsymbol{L}$, що позбавляє нас від необхідності запам'ятовувати множники l_{ij} при перетворенні $\boldsymbol{B}$ до верхньої трикутної матриці. Отже,

$$\boldsymbol{B} \sim \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & 4 \\ 0 & 4 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & 4 \\ 0 & 0 & -1 \end{pmatrix} = \boldsymbol{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{DU}.$$

$\boldsymbol{LR}$- і $\boldsymbol{LDU}$-розклади є такими:

$$\boldsymbol{B} = \boldsymbol{U}^T\boldsymbol{R} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & 4 \\ 0 & 0 & -1 \end{pmatrix},$$

$$\boldsymbol{B} = \boldsymbol{U}^T\boldsymbol{DU} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}. \ \triangle$$

2.2.2. Метод виключення без переставлень рядків. З'ясуємо, за яких умов виключення Гаусса можливі без переставлень рядків. Очевидно, що перший провідний елемент матриці $\boldsymbol{A}$, рівний $d_1 = a_{11}$, повинен бути ненульовим. Другий провідний елемент, теж повинний бути ненульовим, дорівнює, як неважко перевірити, $d_2 = (a_{11}a_{22} - a_{12}a_{21})a_{11}^{-1}$, тобто залежить тільки від елементів a_{11}, a_{12}, a_{21}, a_{22}, а інша частина матриці $\boldsymbol{A}$ ніяк не впливає на визначення d_2. Нижче ми покажемо, що це є загальним правилом, тобто провідний елемент d_i, $i = \overline{1,\ n}$, повинен бути ненульовим і визначатися підматрицею $\boldsymbol{A}_i$, розміщеною в перших i рядках і перших i стовпцях матриці $\boldsymbol{A}$. Визначники таких підматриць називають *провідними* мінорами.

Теорема 2.2.2. *Якщо існує **LDU**-розклад матриці* $\boldsymbol{A}$ *порядку n, то для лівих верхніх квадратних підматриць* $\boldsymbol{A}_i, \boldsymbol{L}_i, \boldsymbol{D}_i, \boldsymbol{U}_i$ *відповідних матриць мають місце рівності*

$$\boldsymbol{A}_i = \boldsymbol{L}_i \boldsymbol{D}_i \boldsymbol{U}_i, \quad i = \overline{1,\ n}. \tag{8}$$

Д о в е д е н н я. Скористаємося блоковим представленням рівності $\boldsymbol{A} = \boldsymbol{LDU}$:

$$\begin{pmatrix} \boldsymbol{A}_i & * \\ * & * \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}_i & \boldsymbol{0} \\ * & * \end{pmatrix} \begin{pmatrix} \boldsymbol{D}_i & \boldsymbol{0} \\ \boldsymbol{0} & * \end{pmatrix} \begin{pmatrix} \boldsymbol{U}_i & * \\ \boldsymbol{0} & * \end{pmatrix}.$$

Тут позначка «$*$» вказує на підматриці, вигляд яких не має значення при доведенні теореми. Виконавши множення, одержимо рівність

$$\begin{pmatrix} \boldsymbol{A}_i & * \\ * & * \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}_i \boldsymbol{D}_i \boldsymbol{U}_i & * \\ * & * \end{pmatrix},$$

звідки випливає формула (8). ☐

Теорема 2.2.3. *Метод виключення без переставлень рядків може бути застосований до матриці* $\boldsymbol{A}$ *тоді і тільки тоді, коли всі її провідні мінори відмінні від нуля.*

Д о в е д е н н я. З рівності (8) випливає, що

$$\det \boldsymbol{A}_i = \det \boldsymbol{L}_i \cdot \det \boldsymbol{D}_i \cdot \det \boldsymbol{U}_i = \det \boldsymbol{D}_i = d_1 d_2 \ldots d_i.$$

Звідси маємо:

$$\frac{\det \boldsymbol{A}_i}{\det \boldsymbol{A}_{i-1}} = \frac{d_1 d_2 \ldots d_i}{d_1 d_2 \ldots d_{i-1}} = d_i\,. \tag{9}$$

Ця формула справедлива для $i = \overline{1, n}$, якщо вважати, що $\det \boldsymbol{A}_0 = 1$. Формула (9) свідчить про те, що провідні елементи відмінні від нуля тоді і тільки тоді, коли всі провідні мінори є ненульовими. Разом з тим ми бачимо, що провідний елемент d_i, $i = \overline{1, n}$, визначається виключно елементами провідного мінора *i*-го порядку. Отже, метод виключення може бути застосований без переставлень рядків. □

2.2.3. Переставлення рядків. Припустимо, що після здійснення декількох кроків виключення або на самому початку цього процесу виник нульовий провідний елемент. Якщо матриця системи $\boldsymbol{Ax} = \boldsymbol{b}$ є невиродженою, то це утруднення можна усунути шляхом переставлення рядків розширеної матриці. Результатом процесу виключення, як і раніше, буде деяка верхня трикутна матриця $\boldsymbol{R}$. Проте цей процес пов'язаний тепер не тільки з матрицями $\boldsymbol{L}_i(\lambda \mid j)$, але і з елементарними матрицями переставлення $\boldsymbol{L}_{ij}$. Оскільки ці матриці і обернені до них не є нижніми трикутними, то і матриця $\boldsymbol{L}$ не буде такою. Отже, ми не можемо отримати $\boldsymbol{LR}$-розклад матриці $\boldsymbol{A}$.

Проте існує дещо змінений розклад матриці $\boldsymbol{A}$. Припустимо, що у нас є перелік усіх переставлень рядків, які необхідно здійснити під час виключення. Можна довести, що, виконавши всі ці переставлення заздалегідь, до початку виключення, ми одержимо матрицю, для якої існує $\boldsymbol{LDU}$-розклад. Це означає, що ми замінюємо вихідну матрицю $\boldsymbol{A}$ матрицею $\boldsymbol{PA}$, де $\boldsymbol{P}$ — матриця переставлень, що має в кожному стовпці і в кожному рядку тільки один ненульовий елемент, рівний одиниці. Очевидно, що $\boldsymbol{P}$ є добутком елементарних матриць переставлення $\boldsymbol{L}_{ij}$ і об'єднує всі переставлення разом. Зрозуміло, що для даної матриці $\boldsymbol{A}$ може знайтися не одна, а багато матриць $\boldsymbol{P}$. Таким чином, замість $\boldsymbol{LDU}$-розкладу маємо:

$$\boldsymbol{PA} = \boldsymbol{LDU}\,. \tag{10}$$

Якщо матриця $\boldsymbol{A}$ невироджена, то провідні елементи матриці $\boldsymbol{PA}$ є ненульовими. При цьому існує єдиний розв'язок системи $\boldsymbol{Ax} = \boldsymbol{b}$, який можна знайти методом виключення з переставленнями рядків.

П р и к л а д 2. Вище зазначалося, що коли для невиродженої матриці $\boldsymbol{A}$ неможливо одержати $\boldsymbol{LDU}$-розклад, то цей розклад можна одержати для матриці $\boldsymbol{PA}$, де матрицю переставлень $\boldsymbol{P}$ можна знайти, виконавши всі необхідні переставлення заздалегідь. Проілюструємо цю можливість на прикладі матриці

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 1 & -1 & 2 & 3 \\ 1 & -1 & 2 & 4 \\ 1 & 0 & 3 & 4 \end{pmatrix}.$$

Перетворимо матрицю $\boldsymbol{A}$ до трикутного вигляду:

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 1 & -1 & 2 & 3 \\ 1 & -1 & 2 & 4 \\ 1 & 0 & 3 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 3 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 2 & 4 \\ 1 & 0 & 3 & 4 \end{pmatrix} \sim$$

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

Ці перетворення були виконані з послідовним застосуванням переставлень таких рядків: 1) першого і другого; 2) другого і четвертого; 3) третього і четвертого. Це означає, що матрицю $\boldsymbol{A}$ треба послідовно множити зліва на елементарні матриці $\boldsymbol{L}_{12}$, $\boldsymbol{L}_{24}$, $\boldsymbol{L}_{34}$. Тоді буде одержана матриця $\boldsymbol{PA} = \boldsymbol{L}_{34}\boldsymbol{L}_{24}\boldsymbol{L}_{12}\boldsymbol{A}$, для якої можна знайти $\boldsymbol{LDU}$-розклад. Отже,

$$\boldsymbol{P} = \boldsymbol{L}_{34}\boldsymbol{L}_{24}\boldsymbol{L}_{12} =$$

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

Вправи до глави 2

2.1. Як треба тлумачити випадок, коли при виконанні елементарних перетворень розширеної матриці системи одержується рядок, який складається з нулів? Що означає випадок, коли всі елементи в рядку, за винятком останнього, дорівнюють нулю?

2.2. За допомогою методу виключення довести властивість матриці $\boldsymbol{A}^{-1}$ з вправи 1.4.

2.3. Показати, що матриці $\boldsymbol{R}$ і $\boldsymbol{L}^{-1}$ можна знайти за такою схемою: 1) скласти блокову матрицю $(\boldsymbol{A} \quad \boldsymbol{E})$ і шляхом додавання до її рядків інших рядків, помножених на відповідні числа, перетворити її так, щоб одержати верхню трикутну матрицю; 2) вважати, що в перетвореній матриці на місці $\boldsymbol{A}$ стоїть $\boldsymbol{R}$, а на місці $\boldsymbol{E}$ — матриця $\boldsymbol{L}^{-1}$.

2.4. Описати і обґрунтувати спосіб обчислення добутку $\boldsymbol{A}^{-1}\boldsymbol{B}$ за допомогою елементарних перетворень рядків матриць $\boldsymbol{A}$ і $\boldsymbol{B}$. Описати і обґрунтувати спосіб обчислення добутку $\boldsymbol{A}\boldsymbol{B}^{-1}$ за допомогою елементарних перетворень стовпців матриць $\boldsymbol{A}$ і $\boldsymbol{B}$.

2.5. Довести, що коли $\boldsymbol{A}$ — матриця порядку n і для провідних мінорів виконуються умови $\det A_i \neq 0, \quad i = \overline{1,\ n-1}$, то існує єдиний $\boldsymbol{LDU}$-розклад матриці $\boldsymbol{A}$.

2.6. Довести, що коли $\det A_i = 0$ при $i \leq n-1$, то $\boldsymbol{LDU}$-розклад матриці $\boldsymbol{A}$ може не існувати, а якщо він існує, то не є єдиним.

2.7. Показати, що матриця $\boldsymbol{PAP}^T$ одержується з $\boldsymbol{A}$ одночасним переставленням рядків і стовпців з однаковими номерами.

ГЛАВА 3

ВЕКТОРНІ ПРОСТОРИ

У попередній главі було показано, як виконувати процедуру виключення Гаусса для спрощення і розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$ з квадратною невиродженою матрицею. Маючи на увазі системи з прямокутною матрицею, ми введемо до розгляду деякі геометричні поняття, які дадуть можливість досягти більш глибокого і різнобічного розуміння проблеми розв'язання систем лінійних рівнянь. З іншого боку, теорія систем лінійних рівнянь буде застосована до розв'язання задач багатовимірної аналітичної геометрії. У цій і наступних главах ми зможемо переконатися в існуванні настільки тісного зв'язку між лінійною алгеброю і аналітичною геометрією, що між ними важко провести чітку межу. Проте в цьому не існує потреби і тому в подальшому геометричні задачі розглядаються паралельно з викладенням необхідного алгебраїчного апарату.

3.1. Лінійна незалежність і базис

3.1.1. Аксіоми векторного простору. Множина R називається *лінійним* (*векторним*) *простором*, якщо для елементів цієї множини (*векторів*) виконані такі вимоги.

A. Кожній парі векторів $\boldsymbol{a}, \boldsymbol{b} \in R$ відповідає вектор $\boldsymbol{a} + \boldsymbol{b} \in R$, який називається *сумою* $\boldsymbol{a}$ і $\boldsymbol{b}$.

B. Кожній парі α і $\boldsymbol{a}$, де α — число і $\boldsymbol{a} \in R$, відповідає вектор $\alpha \boldsymbol{a} \in R$, який називається *добутком вектора* $\boldsymbol{a}$ *на число* α.

C. Операції додавання векторів і множення вектора на число задовольняють аксіомам ($\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ — вектори, α, β — числа):

1) $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$;

2) $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$;

3) існує «**0**» (*нульовий* вектор), тобто такий вектор, що $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$;

4) для кожного $\boldsymbol{a}$ існує такий вектор $-\boldsymbol{a}$ (що називається *протилежним* до $\boldsymbol{a}$), для якого $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$;

5) $1\cdot\boldsymbol{a}=\boldsymbol{a}$;

6) $\alpha(\beta\boldsymbol{a})=(\alpha\beta)\boldsymbol{a}$;

7) $(\alpha+\beta)\boldsymbol{a}=\alpha\boldsymbol{a}+\beta\boldsymbol{a}$;

8) $\alpha(\boldsymbol{a}+\boldsymbol{b})=\alpha\boldsymbol{a}+\alpha\boldsymbol{b}$.

Викладемо без доведення висновки з аксіом 1 – 8. Для довільного векторного простору мають місце такі твердження: 1) нульовий вектор є єдиним; 2) для кожного вектора існує єдиний протилежний вектор; 3) добуток числа 0 і будь-якого вектора $\boldsymbol{a}$ дорівнює нульовому вектору; 4) добуток числа –1 і будь-якого вектора $\boldsymbol{a}$ дорівнює протилежному вектору; 5) добуток будь-якого числа і нульового вектора дорівнює нульовому вектору.

У векторному просторі визначена операція віднімання: вектор $\boldsymbol{c}$ називається *різницею* векторів $\boldsymbol{a}$ і $\boldsymbol{b}$, якщо $\boldsymbol{c}+\boldsymbol{b}=\boldsymbol{a}$, або $\boldsymbol{c}=\boldsymbol{a}-\boldsymbol{b}$. Можна довести, що різниця будь-яких векторів $\boldsymbol{a}$ і $\boldsymbol{b}$ існує і є єдиною.

Завдяки аксіомам і висновкам з них операції над векторами виконуються подібно до перетворень в елементарній алгебрі, з тією різницею, що нема множення і ділення векторів і треба відрізняти число 0 і нульовий вектор **0**. Зокрема, можна переносити вектор з однієї частини векторної рівності до другої, замінюючи його протилежним вектором.

Важливим прикладом векторного простору є дійсний *арифметичний* простір R_n, в якому векторами є матриці-стовпці висоти n з дійсними елементами, а операції над векторами виконуються у відповідності до визначення операцій над матрицями. Легко перевірити для R_n справедливість усіх аксіом векторного простору. Очевидно, що нульовим вектором простору R_n є вектор $\mathbf{0} = (0\ \ 0\ \dots\ 0)^T$, а протилежним до вектора $(\alpha_1\ \ \alpha_2\ \ \dots\ \ \alpha_n)^T$ є вектор $(-\alpha_1\ \ -\alpha_2\ \ \dots\ \ -\alpha_n)^T$.

3.1.2. Лінійна залежність і незалежність векторів. Вектор $\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{a}_i$, де $\alpha_1, \dots, \alpha_n$ — деякі числа, називається *лінійною комбінацією* векторів $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$. Вектори $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ називаються *лінійно залежними*, якщо існують такі числа $\alpha_1, \dots, \alpha_n$, не рівні нулю одночасно, що

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{a}_i = \mathbf{0}. \tag{1}$$

Якщо ця рівність має місце лише тоді, коли всі числа $\alpha_1, \dots, \alpha_n$ дорівнюють нулю, то вектори $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ називаються *лінійно незалежними*.

Теорема 3.1.1. *Для того, щоб вектори $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ були лінійно залежними, необхідно і достатньо, щоб принаймні один з них був лінійною комбінацією інших.*

Доведення. Необхідність. Припустимо, що у формулі (1) $\alpha_n \neq 0$. Тоді

$$\boldsymbol{a}_n = -\frac{\alpha_1}{\alpha_n}\boldsymbol{a}_1 - \frac{\alpha_2}{\alpha_n}\boldsymbol{a}_2 - \dots - \frac{\alpha_{n-1}}{\alpha_n}\boldsymbol{a}_{n-1},$$

тобто $\boldsymbol{a}_n$ є лінійною комбінацією векторів $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{n-1}$.

Достатність. Якщо $\boldsymbol{a}_n = \beta_1 \boldsymbol{a}_1 + \dots + \beta_{n-1}\boldsymbol{a}_{n-1}$, то $\beta_1 \boldsymbol{a}_1 + \dots + \beta_{n-1}\boldsymbol{a}_{n-1} - \boldsymbol{a}_n = \mathbf{0}$. Оскільки серед чисел $\beta_1, \dots, \beta_{n-1}, -1$ принаймні одне відмінне від нуля, то вектори $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ є лінійно залежними. □

Якщо лінійно залежними є два вектори, то вони називаються *колінеарними*.

3.1.3. Вимірність і базис. Векторний простір називається *n-вимірним*, якщо в ньому існує n лінійно незалежних векторів і нема більшої кількості лінійно незалежних векторів. Вимірність простору R позначається через $\dim R$.

Сукупність лінійно незалежних векторів називається *базисом* векторного простору, якщо будь-який вектор простору є лінійною комбінацією цих векторів.

Розглянемо як приклад питання про базис арифметичного простору R_n. Покажемо, що стовпці одиничної матриці $\boldsymbol{E}_n$ утворюють базис цього простору. Для цього спочатку покажемо, що стовпці $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$ матриці $\boldsymbol{E}_n$ є лінійно незалежними. Складемо лінійну комбінацію цих векторів з коефіцієнтами $\alpha_1, ..., \alpha_n$. Очевидно, що

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{e}_{*i} = (\alpha_1 \ ... \ \alpha_n)^T . \tag{2}$$

Вектор $(\alpha_1 \ ... \ \alpha_n)^T$ є нульовим лише за умови $\alpha_1 = \ldots = \alpha_n = 0$, що означає лінійну незалежність векторів $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$.

Ми повинні також переконатися в тому, що довільний вектор $\boldsymbol{a} = (\alpha_1 \ ... \ \alpha_n)^T$ з R_n представляється у вигляді лінійної комбінації векторів $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$. Але саме це показує формула (2). Отже, стовпці одиничної матриці є базисом простору R_n. Цей базис називається *натуральним*.

З'ясуємо зв'язок між поняттями вимірності і базису векторного простору.

Теорема 3.1.2. *Будь-яка система n лінійно незалежних векторів n-вимірного векторного простору є базисом цього простору.*

Д о в е д е н н я. Нехай $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$ — лінійно незалежні вектори і $\boldsymbol{a}$ — будь-який вектор простору. З визначення вимірності векторного простору випливає, що вектори $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$, $\boldsymbol{a}$ є лінійно залежними, тобто існують числа $\alpha_0, \ \alpha_1, \ ..., \ \alpha_n$, не рівні нулю одночасно і такі, що

$$\alpha_{\mathbf{0}} \boldsymbol{a} + \alpha_{\mathbf{1}} \boldsymbol{e}_{\mathbf{1}} + \ldots + \alpha_n \boldsymbol{e}_n = \mathbf{0} . \tag{3}$$

При цьому $\alpha_0 \neq 0$, бо якщо $\alpha_0 = 0$, то $\sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = \boldsymbol{0}$, що суперечить лінійній незалежності векторів $\boldsymbol{e}_{*1}, ..., \boldsymbol{e}_{*n}$.

З рівності (3) знаходимо, що

$$\boldsymbol{a} = \sum_{i=1}^{n} \alpha'_i \boldsymbol{e}_i, \tag{4}$$

де $\alpha'_i = -\alpha_i \alpha_0^{-1}$. Оскільки $\boldsymbol{a}$ — довільний вектор простору, то рівність (4) доводить, що вектори $\boldsymbol{e}_{*1}, ..., \boldsymbol{e}_{*n}$ утворюють базис цього простору. □

Лінійна комбінація (4) називається *розкладом n-вимірного вектора* $\boldsymbol{a}$ *за базисом* $\boldsymbol{e}_{*1}, ..., \boldsymbol{e}_{*n}$, а числа $\alpha'_1, \ldots, \alpha'_n$ називаються *координатами* вектора $\boldsymbol{a}$ відносно цього базису.

Теорема 3.1.3. *Координати вектора відносно деякого базису визначаються єдиним способом.*

Доведення. Якщо $\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = \sum_{i=1}^{n} \alpha'_i \boldsymbol{e}_i$ — два різні розклади вектора $\boldsymbol{a}$ за базисом $\boldsymbol{e}_{*1}, ..., \boldsymbol{e}_{*n}$, то $\sum_{i=1}^{n} (\alpha_i - \alpha'_i) \boldsymbol{e}_i = \boldsymbol{0}$. Оскільки вектори $\boldsymbol{e}_{*1}, ..., \boldsymbol{e}_{*n}$ лінійно незалежні, то це можливо лише коли $\alpha_i - \alpha'_i = 0,\ i = \overline{1, n}$. Звідси випливають рівності $\alpha_i = \alpha'_i,\ i = \overline{1, n}$, що доводить єдиність розкладу. □

Аналогічні міркування застосовуються при доведенні наступної теореми.

Теорема 3.1.4. *Для колінеарності двох векторів необхідно і достатньо, щоб їх координати були пропорційними.*

Доведення. Необхідність. Нехай вектори

$$\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i, \quad \boldsymbol{b} = \sum_{i=1}^{n} \beta_i \boldsymbol{e}_i$$

є колінеарними. На підставі теореми 3.1.1 або $\boldsymbol{a} = \lambda \boldsymbol{b}$, або $\boldsymbol{b} = \lambda \boldsymbol{a}$, де λ — деяке ненульове число. Нехай $\boldsymbol{a} = \lambda \boldsymbol{b}$, тобто

$$\sum_{i=1}^{n}\alpha_i\boldsymbol{e}_i = \lambda\sum_{i=1}^{n}\beta_i\boldsymbol{e}_i.$$

Звідси випливає, що

$$\alpha_i = \lambda\beta_i, \quad i = \overline{1,\, n}, \tag{5}$$

тому

$$\frac{\alpha_1}{\beta_1} = \frac{\alpha_2}{\beta_2} = \ldots = \frac{\alpha_n}{\beta_n}.$$

Достатність. Припустимо, що координати пропорційні, тобто виконуються останні рівності. Позначивши загальне значення відношень через λ, будемо мати рівності (5). Тепер

$$\boldsymbol{a} = \sum_{i=1}^{n}\alpha_i\boldsymbol{e}_i = \sum_{i=1}^{n}\lambda\beta_i\boldsymbol{e}_i = \lambda\sum_{i=1}^{n}\beta_i\boldsymbol{e}_i = \lambda\boldsymbol{b}$$

і це означає колінеарність векторів $\boldsymbol{a}$ і $\boldsymbol{b}$. □

Головне значення базису полягає в тому, що операції над векторами перетворюються у відповідні операції над координатами цих векторів, оскільки, як неважко переконатися, при додаванні векторів їх координати додаються, а при множенні вектора на число його координати множаться на це число.

Отже, виходячи з найбільш загального визначення n-вимірного векторного простору, ми дійшли висновку, що цей простір улаштований у деякому сенсі так, як арифметичний простір R_n. Можна показати, що вимірність векторного простору є його єдиною характеристикою, тобто всі n-вимірні векторні простори є однаковими.

3.1.4. Перетворення координат при зміні базису. Припустимо, що $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$ і $\boldsymbol{e}_1'$, ..., $\boldsymbol{e}_n'$ є базисами n-вимірного векторного простору і визначені координати векторів «нового» базису $\boldsymbol{e}_1'$, ..., $\boldsymbol{e}_n'$ відносно «старого» базису $\boldsymbol{e}_{*1}$, ..., $\boldsymbol{e}_{*n}$:

$$\boldsymbol{e}'_j = \sum_{i=1}^{n} q_{ij}\boldsymbol{e}_i, \quad j = \overline{1, n}. \tag{6}$$

Нехай $\boldsymbol{a} = (\alpha_1 \;\dots\; \alpha_n)^T$ і $\boldsymbol{a}' = (\alpha'_1 \;\dots\; \alpha'_n)^T$ є матрицями-стовпцями координат довільного вектора у «старому» і «новому» базисах. Тоді

$$\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = \sum_{j=1}^{n} \alpha'_j \boldsymbol{e}'_j . \tag{7}$$

З рівностей (7) і (6) випливає, що

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = \sum_{j=1}^{n} \alpha'_j \sum_{i=1}^{n} q_{ij}\boldsymbol{e}_i = \sum_{i=1}^{n} \boldsymbol{e}_i \sum_{j=1}^{n} q_{ij}\alpha'_j ,$$

звідки на підставі теореми 3.1.3 маємо:

$$\alpha_i = \sum_{j=1}^{n} q_{ij}\alpha'_j, \quad i = \overline{1, n}. \tag{8}$$

Таким чином, ми знайшли залежність між координатами вектора $\boldsymbol{a}$ в різних базисах. Якщо позначити через $\boldsymbol{Q}$ матрицю n-го порядку з елементами q_{ij}, то формулу (8) можна записати у вигляді:

$$\boldsymbol{a} = \boldsymbol{Q}\boldsymbol{a}' . \tag{9}$$

Формула (6) показує, що стовпці матриці $\boldsymbol{Q}$ складають координати векторів «нового» базису в «старому» базисі. Матриця $\boldsymbol{Q}$ називається *матрицею перетворення координат* при зміні базису. Покажемо, що вона є невиродженою.

Дійсно, якщо вважати, що $\boldsymbol{a} = \boldsymbol{0}$, то з формули (9) одержимо систему рівнянь $\boldsymbol{Q}\boldsymbol{a}' = \boldsymbol{0}$, у якої нульовий розв'язок $\boldsymbol{a}' = \boldsymbol{0}$ повинен бути єдиним.У протилежному разі рівність $\sum_{j=1}^{n} \alpha'_j \boldsymbol{e}'_j = \boldsymbol{0}$, яка випливає з формули (7),буде свідчити про лінійну залежність векторів $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n$. Отже, матриця $\boldsymbol{Q}$ є невиродженою.

3.2. Підпростори векторних просторів

3.2.1. Підпростори. Непуста множина U векторів у векторному просторі R називається його *підпростором*, якщо вона разом з кожною парою векторів $\boldsymbol{a}, \boldsymbol{b}$ вміщує всі їх лінійні комбінації $\alpha \boldsymbol{a} + \beta \boldsymbol{b}$.

Кожний підпростір є векторним простором. Щоб переконатися в цьому, досить перевірити лише ті аксіоми векторного простору, які стосуються нульового і протилежного векторів, бо виконання інших аксіом є очевидним. Якщо взяти $\alpha = \beta = 0$, то $\alpha\boldsymbol{a} + \beta\boldsymbol{b} = \boldsymbol{0}$, тобто нульовий вектор належить до U. Нехай $\alpha = -1,\ \beta = 0$. Тоді $(-1)\boldsymbol{a} + 0\boldsymbol{b} = -\boldsymbol{a}$, тому разом з вектором $\boldsymbol{a}$ в U є протилежний до нього вектор. Отже, множина U є простором.

У кожному векторному просторі множина, що складається з одного нульового вектора, є підпростором. Він називається *нульовим* підпростором. Множина, яка складається з усіх векторів простору R, є підпростором, тобто весь простір R можна вважати підпростором. Нульовий підпростір і весь простір є двома крайніми випадками підпросторів. Ці два підпростори називаються тривіальними, а інші — нетривіальними.

Очевидно, що множина всяких лінійних комбінацій векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in R$ є підпростором. Про нього говорять, що він породжений векторами $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$, або що він є натягнутим на ці вектори. Векторний простір породжується векторами базису цього простору.

Якщо в просторі R вибрати базис $\boldsymbol{e}_{*1}, \ldots, \boldsymbol{e}_{*n}$, то базисні вектори підпростору U, взагалі кажучи, не можна вибрати з цих векторів, бо може статися, що ні один з них не належить до U. Проте справедливе в деякому сенсі протилежне твердження.

***Теорема 3.2.1**. Будь-який базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ підпростору можна доповнити до базису всього простору.*

Д о в е д е н н я. Якщо $\dim R = n$ і $m < n$, то знайдеться такий вектор $\boldsymbol{e}_{m+1} \in R$, що вектори $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}$ будуть лінійно незалежними, бо в протилежному випадку простір R виявився б m-вимірним. Якщо $m+1<n$, то міркування можна повторити і добавити до системи $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}$ ще один вектор, не порушуючи лінійної незалежності. Так можна продовжувати доти, поки кількість векторів у системі не досягне n. Тоді вона перетвориться в базис простору R. □

Нехай підпростори U, V простору R є такими, що $U \subseteq V$ і $\dim U = \dim V$. Покажемо, що підпростори U і V збігаються. Дійсно, якщо $U \subseteq V$, то вектори деякого базису в U належать до V, а з огляду на рівність вимірностей U і V ці вектори є базисом у V. Виходить, що підпростори U і V породжуються векторами одного і того ж базису, тому $U = V$.

3.2.2. Сума і переріз підпросторів. *Сумою* $U+V$ *підпросторів* U, V векторного простору R називається множина всіх векторів вигляду $\boldsymbol{a}=\boldsymbol{u}+\boldsymbol{v}$, де $\boldsymbol{u} \in U, \boldsymbol{v} \in V$.

Перерізом $U \bigcap V$ *підпросторів* U, V називається множина всіх векторів, які одночасно належать як до U, так і до V.

Сума підпросторів і їх переріз є непустими множинами, оскільки їм належить нульовий вектор простору R. Покажемо, що ці множини є підпросторами. Дійсно, будемо вважати, що $\boldsymbol{a}_1, \boldsymbol{a}_2 \in U+V$, тоді $\boldsymbol{a}_1=\boldsymbol{u}_1+\boldsymbol{v}_1$, $\boldsymbol{a}_2=\boldsymbol{u}_2+\boldsymbol{v}_2$, де $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$ і $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Розглянемо лінійну комбінацію $\alpha \boldsymbol{a}_1+\beta \boldsymbol{a}_2$. Очевидно, що $\alpha \boldsymbol{a}_1+\beta \boldsymbol{a}_2=(\alpha \boldsymbol{u}_1+\beta \boldsymbol{u}_2)+(\alpha \boldsymbol{v}_1+\beta \boldsymbol{v}_2)$. Оскільки $\alpha \boldsymbol{u}_1+\beta \boldsymbol{u}_2 \in U$ і $\alpha \boldsymbol{v}_1+\beta \boldsymbol{v}_2 \in V$, то $\alpha \boldsymbol{a}_1+\beta \boldsymbol{a}_2 \in U+V$. Тому $U+V$ є підпростором.

Нехай тепер $\boldsymbol{a}_1, \boldsymbol{a}_2 \in U \bigcap V$, тобто $\boldsymbol{a}_1, \boldsymbol{a}_2 \in U$ і $\boldsymbol{a}_1, \boldsymbol{a}_2 \in V$. Очевидно, що $\alpha \boldsymbol{a}_1+\beta \boldsymbol{a}_2 \in U$, і $\alpha \boldsymbol{a}_1+\beta \boldsymbol{a}_2 \in V$, тобто $\alpha \boldsymbol{a}_1+\beta \boldsymbol{a}_2 \in U \bigcap V$. Отже, $U \bigcap V$ є підпростором.

На найпростіших прикладах неважко переконатися в тому, що вимірність суми двох підпросторів залежить не тільки від вимірностей підпросторів, але і від вимірності їх спільної частини. Має місце

Теорема 3.2.2. *Для будь-яких підпросторів* U, V *векторного простору* R *справедлива формула Грассмана*

$$\dim(U+V)=\dim U+\dim V-\dim(U \bigcap V). \tag{1}$$

Д о в е д е н н я. Якщо $U \bigcap V$ — ненульовий підпростір, то візьмемо в ньому базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ і доповнимо його векторами $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ до базису в U і векторами $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$ до базису у V. Це можливо зробити на підставі теореми 3.2.1. Покажемо, що система векторів

$$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l \tag{2}$$

є базисом підпростору $U + V$.

Кожний вектор $\boldsymbol{a} \in U + V$ є лінійною комбінацією векторів (2). Це випливає з того, що $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in U, \boldsymbol{v} \in V$, причому $\boldsymbol{u}$ розкладається за базисом $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$, а $\boldsymbol{v}$ — за базисом $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$. Нам залишається довести лінійну незалежність системи векторів (2).

Візьмемо яку-небудь рівну нульовому вектору лінійну комбінацію векторів (2):

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{u}_i + \sum_{j=1}^{m} \beta_j \boldsymbol{e}_j + \sum_{p=1}^{l} \gamma_p \boldsymbol{v}_p = \mathbf{0}. \tag{3}$$

Позначимо через $\boldsymbol{b}$ вектор $\sum_{i=1}^{k} \alpha_i \boldsymbol{u}_i + \sum_{j=1}^{m} \beta_j \boldsymbol{e}_j$. Очевидно, що $\boldsymbol{b} \in U$. З іншого боку, рівність (3) доводить, що $\boldsymbol{b} = -\sum_{p=1}^{l} \gamma_p \boldsymbol{v}_p$, тобто $\boldsymbol{b} \in V$. Отже, $\boldsymbol{b} \in U \bigcap V$ і тому $\boldsymbol{b} = \sum_{q=1}^{m} \delta_q \boldsymbol{e}_q$. З двох одержаних представлень вектора $\boldsymbol{b}$ маємо:

$$\sum_{q=1}^{m} \delta_q \boldsymbol{e}_q + \sum_{p=1}^{l} \gamma_p \boldsymbol{v}_p = \mathbf{0}.$$

Звідси на підставі лінійної незалежності векторів $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$ робимо висновок про те, що $\delta_1 = \ldots = \delta_m = \gamma_1 = \ldots = \gamma_l = 0$. Після цього рівність (3) набуває вигляду:

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{u}_i + \sum_{j=1}^{m} \beta_j \boldsymbol{e}_j = \mathbf{0},$$

звідки $\alpha_1 = \ldots = \alpha_k = \beta_1 = \ldots = \beta_m = 0$ через лінійну незалежність векторів $\boldsymbol{u}_1, ..., \boldsymbol{u}_k, \boldsymbol{e}_1, ..., \boldsymbol{e}_m$.

Таким чином, ми показали, що рівність (3) можлива лише тоді, коли всі коефіцієнти $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_l$ дорівнюють нулю. Це доводить лінійну незалежність векторів (2).

Отже, система векторів (2) є базисом підпростору $U + V$, звідки випливає, що $\dim(U + V) = k + m + l$. З попередніх міркувань ми знаємо, що $\dim U = k + m, \dim V = m + l, \dim(U \bigcap V) = m$. Тому має місце формула (1).

Якщо $\dim(U \bigcap V) = 0$, то доведення виконується аналогічно, але в ньому не беруть участь вектори $\boldsymbol{e}_1, ..., \boldsymbol{e}_m$. ☐

П р и к л а д. Розглянемо деякі з можливих співвідношень між вимірностями підпросторів U, V, $U + V$, $U \bigcap V$ векторного простору R_4 (більш простий випадок простору R_3 читачеві рекомендується розглянути самостійно).

Два двовимірних підпростори U, V простору R_4 можуть перерізатися по нульовому вектору. Тоді їх сума збігається з усім простором, оскільки формула (1) набуває вигляду:

$$\dim(U + V) = 2 + 2 - 0 = 4 .$$

Якщо ці підпростори перерізаються по одновимірному підпростору, то

$$\dim(U + V) = 2 + 2 - 1 = 3 ,$$

тобто сума виявляється тривимірною. Підпростори U і V можуть збігатися, тобто їх переріз є двовимірним. Тоді сума теж є двовимірною, бо

$$\dim(U + V) = 2 + 2 - 2 = 2 .$$

Оскільки переріз двовимірних підпросторів не є більш ніж двовимірним, то інших варіантів не існує.

Два тривимірних підпростори або можуть перерізатися по двовимірному підпростору і тоді

$$\dim(U + V) = 3 + 3 - 2 = 4 ,$$

або збігаються:

$$\dim(U+V)=3+3-3=3.$$

Інших випадків перерізу двох тривимірних підпросторів не існує, оскільки сума підпросторів не є більш ніж чотиривимірною.

Якщо $\dim U = 2$ і $\dim V = 3$, то або U і V перерізаються по одновимірному підпростору:

$$\dim(U+V)=2+3-1=4,$$

або U вміщується у V:

$$\dim(U+V)=2+3-2=3.$$

Аналогічно можна розглянути випадки, коли один з підпросторів U, V є одновимірним. △

3.2.3. Пряма сума підпросторів. Якщо $U \cap V$ — нульовий підпростір, то сума підпросторів U, V називається *прямою сумою* і позначається через $U \oplus V$. При розгляді теореми 3.2.2 ми бачили, що вимірність прямої суми підпросторів дорівнює сумі вимірностей доданків і об'єднання будь-яких базисів, вибраних по одному в кожному доданку, утворює базис прямої суми.

Теорема 3.2.3. *Кожний вектор $\boldsymbol{a} \in U \oplus V$ можна розкласти єдиним способом у суму*

$$\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}, \tag{4}$$

де $\boldsymbol{u} \in U, \boldsymbol{v} \in V$.

Д о в е д е н н я. Припустимо, що є два таких представлення, тобто $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v} = \boldsymbol{u}' + \boldsymbol{v}'$. Звідси знаходимо, що $\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{v}' - \boldsymbol{v}$. Оскільки $\boldsymbol{u} - \boldsymbol{u}' \in U$, $\boldsymbol{v}' - \boldsymbol{v} \in V$ і оскільки єдиним спільним вектором підпросторів U, V є нульовий вектор, то $\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{0}$, $\boldsymbol{v}' - \boldsymbol{v} = \boldsymbol{0}$, тобто $\boldsymbol{u}' = \boldsymbol{u}$, $\boldsymbol{v}' = \boldsymbol{v}$. □

Нехай мова йде про представлення простору R у вигляді прямої суми його підпросторів U, V. Говорять, що простір R розщеплюється на два підпростори:

$$R = U \oplus V.$$

Теорема 3.2.3 показує, що умова $\dim(U \bigcap V) = 0$, яка накладається на доданки прямої суми, рівносильна умові єдиності представлення (4). У такому вигляді визначення розщеплення безпосередньо поширюється на будь-яку кількість підпросторів. Можна довести, шо в цьому випадку сума вимірностей підпросторів дорівнює вимірності простору R і його базис є об'єднанням базисів підпросторів. До речі, маючи певний базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ простору R, можна тлумачити його як деяке розщеплення R на n одновимірних підпросторів.

3.3. Афінний простір

3.3.1. Визначення афінного простору. У багатьох задачах у центрі уваги опиняються не дії над векторами, а геометричні факти, пов'язані з взаємним розміщенням фігур у просторі. Тому разом з векторним простором вводиться до розгляду так званий афінний простір.

Множина S_n називається *n-вимірним афінним простором*, а її елементи — *точками* цього простору, якщо кожній упорядкованій парі елементів $A, B \in S_n$ зіставляється єдиний вектор з R_n (який позначається через $\overrightarrow{AB}$) у відповідності до аксіом: 1) для кожної точки $A \in S_n$ і кожного вектора $\boldsymbol{a} \in R_n$ існує єдина точка $B \in S_n$ така, що $\overrightarrow{AB} = \boldsymbol{a}$ (перша з цих точок називається початком, а друга — кінцем вектора); 2) для будь-яких трьох точок A, B, C виконується рівність $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.

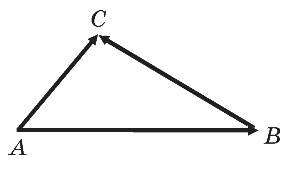

Рис. 1.

Перша аксіома дає можливість відкласти будь-який вектор з довільної точки, а друга аксіома визначає додавання векторів (рис. 1). За допомогою аксіом неважко показати, що кожній парі точок, які збігаються, відповідає нульовий вектор з R_n. Крім того, якщо $\overrightarrow{AB} = \boldsymbol{a}$, то $\overrightarrow{BA} = -\boldsymbol{a}$.

Векторний простір R_n можна розглядати як афінний простір S_n. Для цього достатньо вектори назвати точками і кожній парі векторів $\boldsymbol{a}, \boldsymbol{b}$, що розглядаються як точки простору S_n, зіставити вектор $\boldsymbol{b} - \boldsymbol{a} \in R_n$.

З іншого боку, афінний простір S_n можна розглядати як векторний. Для цього достатньо в просторі S_n вибрати яку-небудь точку O і довільній точці $M \in S_n$ зіставити вектор $\overrightarrow{OM}$, який називають *радіус-вектором* точки M. Множина радіус-векторів усіх точок простору S_n утворює простір R_n.

Наведені міркування надають можливість у подальшому рівноправно використовувати поняття вектора і точки.

3.3.2. Афінні координати. Виберемо в афінному просторі S_n деяку точку O, яку будемо називати початком координат, і в просторі R_n візьмемо який-небудь базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Початок координат разом з базисом утворюють *афінну систему координат* в афінному просторі.

Розклавши радіус-вектор довільної точки $M \in S_n$ за базисом $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, одержимо:

$$\overrightarrow{OM} = \alpha_1 \boldsymbol{e}_1 + \ldots + \alpha_n \boldsymbol{e}_n .$$

Коефіцієнти $\alpha_1, \ldots, \alpha_n$ цього розкладу називаються *афінними координатами* точки M відносно вибраної афінної системи координат з початком O і базисом $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Ці координати визначаються однозначно, оскільки єдиним є розклад вектора $\overrightarrow{OM}$ за базисом $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$.

Припустимо, що ми маємо іншу довільну точку N з афінними координатами $\beta_1, \ldots, \beta_n$. Знайдемо координати вектора $\overrightarrow{MN}$. Очевидно, що

$$\overrightarrow{MN} = \overrightarrow{MO} + \overrightarrow{ON} = \overrightarrow{ON} - \overrightarrow{OM} = (\beta_1 - \alpha_1)\boldsymbol{e}_1 + \ldots + (\beta_n - \alpha_n)\boldsymbol{e}_n .$$

Отже, щоб одержати координати вектора $\overrightarrow{MN}$, треба від координат його кінця відняти відповідні координати початку.

3.3.3. Площини. Якщо в афінному просторі S_n зафіксована деяка точка A і у векторному просторі R_n вибраний m-вимірний підпростір U_m, то множина всіх точок M, для яких $\overrightarrow{AM} \in U_m$, називається *m-вимірною площиною*, що проходить через точку A в напрямі підпростору U_m. Точки A і M називають відповідно *початковою* і *поточною точками* площини, а U_m — *спрямовуючим підпростором* цієї площини. При $m = 0$ площина складається з однієї точки (нульвимірна площина). Одновимірна площина називається *прямою лінією*. Площина вимірності $n - 1$ називається *гіперплощиною*, а при $m = n$ площина збігається з усім простором S_n.

Припустимо, що в просторі S_n вибрана яка-небудь афінна система координат з початком O і базисом $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Розглянемо площину, що проходить через точку A в напрямі підпростору U_m. Якщо U_m породжується лінійно незалежними векторами $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$, то радіус-вектор $\overrightarrow{OM}$ поточної точки площини дорівнює (рис. 2)

$$\overrightarrow{OM} = \overrightarrow{OA} + \overrightarrow{AM} = \overrightarrow{OA} + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i, \qquad (1)$$

де $\alpha_1, \ldots, \alpha_m$ — параметри, які набувають усяких числових значень. Позначимо радіус-вектор поточної точки площини через $\boldsymbol{x}$, а радіус-вектор початкової точки — через $\boldsymbol{f}_0$. Тоді з формули (1) одержимо *векторне рівняння* площини:

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i. \qquad (2)$$

Покажемо, що початкова точка визначається неєдиним способом.

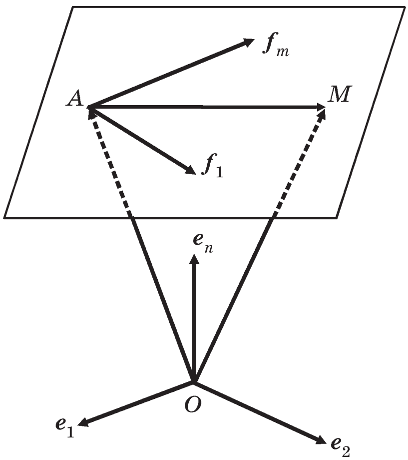

Рис. 2

Теорема 3.3.1. *Якщо вектори $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$ належать площині, то вектор $\boldsymbol{x}_1 - \boldsymbol{x}_2$ належить її спрямовуючому підпростору.*

Д о в е д е н н я. Вектори $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$ задовольняють рівнянню (2):

$$\boldsymbol{x}_1 = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i, \quad \boldsymbol{x}_2 = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i' \boldsymbol{f}_i.$$

Звідси випливає, що

$$\boldsymbol{x}_1 - \boldsymbol{x}_2 = \sum_{i=1}^{m} (\alpha_i - \alpha_i')\boldsymbol{f}_i .$$

Отже, $\boldsymbol{x}_1 - \boldsymbol{x}_2$ лінійно виражається через вектори $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$, тобто $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in U_m$.

Висновок. *Початковою точкою площини може бути будь-яка її точка.*

Д о в е д е н н я. Представимо рівняння (2) у вигляді:

$$\boldsymbol{x} = \boldsymbol{f}_0' + \boldsymbol{f}_0 - \boldsymbol{f}_0' + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i , \tag{3}$$

де $\boldsymbol{f}_0'$ — деяка точка розглядуваної площини. Оскільки $\boldsymbol{f}_0 - \boldsymbol{f}_0' \in U_m$, то $\boldsymbol{f}_0 - \boldsymbol{f}_0' = \sum_{i=1}^{m} \alpha_i'' \boldsymbol{f}_i$. Тому з рівняння (3) випливає, що

$$\boldsymbol{x} = \boldsymbol{f}_0' + \sum_{i=1}^{m} (\alpha_i + \alpha_i'') \boldsymbol{f}_i .$$

Порівнюючи це рівняння з (2), бачимо, що $\boldsymbol{f}_0'$ є початковою точкою площини. □

Будемо вважати, що вектори $\boldsymbol{x}, \boldsymbol{f}_0, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$ розкладені за базисом $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$, тобто

$$\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{e}_j , \quad \boldsymbol{f}_i = \sum_{j=1}^{n} f_{ij} \boldsymbol{e}_j ,$$

де $i = \overline{0,\ m}$. Тоді з векторного рівняння (2) одержуються параметричні рівняння площини в заданій системі координат:

$$x_j = f_{0j} + \sum_{i=1}^{m} \alpha_i f_{ij}, \quad j = \overline{1,\ n}. \tag{4}$$

Якщо вважати, що $f_{0j} = 0,\ j = \overline{1,\ n}$, то з формули (4) одержимо параметричні рівняння спрямовуючого підпростору площини.

В окремому випадку, коли $m = 1$, маємо векторне рівняння прямої лінії:

$$\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1$$

і її параметричні рівняння:

$$x_j = f_{0j} + \alpha_1 f_{1j}, \quad j = \overline{1,\ n}.$$

Якщо з параметричних рівнянь виключити параметр α_1, то одержимо рівняння прямої лінії, які називаються *канонічними*:

$$\frac{x_1 - f_{01}}{f_{11}} = \frac{x_2 - f_{02}}{f_{12}} = \ldots = \frac{x_n - f_{0n}}{f_{1n}}.$$

Якщо площина (2) є гіперплощиною, то її параметричні рівняння (4) мають вигляд:

$$x_j - f_{0j} - \sum_{i=1}^{n-1} \alpha_i f_{ij} = 0, \quad j = \overline{1,\ n}.$$

Це — система лінійних рівнянь відносно 1, $\alpha_1, \ \ldots, \ \alpha_{n-1}$ з нульовими правими частинами. Оскільки вказаний розв'язок не є нульовим, то матриця системи не може бути невиродженою, тобто визначник системи є нульовим:

$$\begin{vmatrix} x_1 - f_{01} & -f_{11} & -f_{21} & \ldots & -f_{n-1,1} \\ x_2 - f_{02} & -f_{12} & -f_{22} & \ldots & -f_{n-1,2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x_n - f_{0n} & -f_{1n} & -f_{2n} & \ldots & -f_{n-1,n} \end{vmatrix} = 0.$$

Це рівняння відносно $x_1, \ \ldots, \ x_n$ називається рівнянням гіперплощини в координатній формі.

3.4. Ранг матриці

3.4.1. Основна теорема про ранг матриці. Будемо розглядати стовпці $m \times n$-матриці $\boldsymbol{A}$ як вектори арифметичного простору R_m, а рядки — як вектори простору R_n. Тоді ми зможемо говорити про лінійну залежність і незалежність стовпців матриці $\boldsymbol{A}$ або про лінійну залежність і незалежність її рядків.

Мінором i-го порядку $m \times n$ -матриці $\boldsymbol{A}$ називається визначник i -го порядку з елементами, що знаходяться на перетині будь-яких i стовпців і будь-яких i рядків матриці $\boldsymbol{A}$. Зрозуміло, що i не перевищує менше з чисел m, n. Мінорами першого порядку є числові значення елементів матриці.

Максимальний порядок відмінних від нуля мінорів матриці називається її *рангом*. Інакше кажучи, рангом матриці називається таке число r, що серед мінорів матриці існує мінор порядку r, відмінний від нуля, а всі мінори більш високого порядку дорівнюють нулю або не можуть бути складені. Ранг матриці $\boldsymbol{A}$ позначається через $r(\boldsymbol{A})$. Ранг нульової матриці вважається рівним нулю.

Будь-який відмінний від нуля мінор порядку $r(\boldsymbol{A})$ називається *базисним мінором* матриці, а стовпці і рядки, в яких він розміщений, називаються *базисними*. Зрозуміло, що у матриці $\boldsymbol{A}$ базисний мінор може бути не один. Будь-яка матриця, крім нульової, має принаймні один базисний мінор і, отже, принаймні одну систему базисних стовпців (рядків).

Доведемо таку основну теорему.

***Теорема 3.4.1* (*теорема про базисний мінор*).** *Базисні стовпці матриці є лінійно незалежними і будь-який стовпець є лінійною комбінацією базисних стовпців.*

Д о в е д е н н я. Припустимо, що базисні стовпці лінійно залежні. Тоді за теоремою 3.1.1 принаймні один з них є лінійною комбінацією інших базисних стовпців. Можна, не змінюючи базисного мінора, додати до цього стовпця взяту з протилежним знаком указану лінійну комбінацію. Тоді базисний мінор буде мати нульовий стовпець, тобто дорівнюватиме нулю, що неможливо. Отже, базисні стовпці є лінійно незалежними.

Доведемо тепер, що будь-який стовпець матриці $\boldsymbol{A}$ є лінійною комбінацією базисних стовпців. Оскільки властивість визначника дорівнювати чи не дорівнювати нулю зберігається при переставленні стовпців, а також рядків, ми можемо, не обмежуючи загальності, вважати, що базисний мінор розміщений у лівому верхньому куті матриці $\boldsymbol{A}$.

Покажемо, що написаний нижче визначник Δ_{ij} є нульовим:

$$\Delta_{ij} = \begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1j} \\ \dots & \dots & \dots & \dots \\ a_{r1} & \dots & a_{rr} & a_{rj} \\ a_{i1} & \dots & a_{ir} & a_{ij} \end{vmatrix} = 0, \quad i = \overline{1,\ m}, \quad j = \overline{1,\ n}. \qquad (1)$$

Якщо $i \le r$ або $j \le r$, то $\Delta_{ij} = 0$, оскільки у визначника є два однакових стовпці або два однакових рядки. Якщо $i, j > r$, то визначник Δ_{ij} дорівнює нулю як мінор $(r+1)$-го порядку матриці рангу r. Отже, рівність (1) доведена.

Зафіксуємо j і будемо вважати, що $i = \overline{1,\ m}$. Розкладемо визначник Δ_{ij} за елементами останнього рядка. Алгебраїчні доповнення цих елементів позначимо через $A_1, \dots, A_r, A_j$. Ці величини залишаються незмінними при $i = \overline{1,\ m}$. Шуканий розклад виглядає так:

$$a_{i1}A_1 + \dots + a_{ir}A_r + a_{ij}A_j = 0, \quad i = \overline{1,\ m}.$$

Оскільки A_j збігається з відмінним від нуля базисним мінором, то

$$a_{ij} = \alpha_1 a_{i1} + \dots + \alpha_r a_{ir}, \quad i = \overline{1,\ m}, \qquad (2)$$

де $\alpha_k = -A_k A_j^{-1}$, $k = \overline{1,\ r}$. Рівності (2) можна записати у векторному вигляді:

$$\boldsymbol{a}_{*j} = \alpha_1 \boldsymbol{a}_{*1} + \dots + \alpha_r \boldsymbol{a}_{*r}.$$

Отже, будь-який стовпець $\boldsymbol{a}_{*j}$ матриці є лінійною комбінацією базисних стовпців $\boldsymbol{a}_{*1}, \dots, \boldsymbol{a}_{*r}$.

Якщо застосувати одержаний результат до транспонованої матриці, то ми отримаємо доведення теореми для рядків матриці, бо при транспонуванні матриці її мінори не змінюються.

Висновок 1. *Вимірність простору, породженого деякою системою векторів, дорівнює рангу матриці, складеної з координат цих векторів відносно будь-якого базису.*

Висновок 2*. Максимальна кількість лінійно незалежних стовпців* (*рядків*) *матриці дорівнює рангу матриці.*

Д о в е д е н н я обох висновків є цілком очевидними. □

Як висновок з теореми 3.4.1 маємо наступну теорему.

Теорема 3.4.2. *Для того, щоб визначник матриці* $\boldsymbol{A}$ *порядку* n *дорівнював нулю, необхідно і достатньо, щоб його стовпці* (*рядки*) *були лінійно залежними.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Якщо $\det \boldsymbol{A} = 0$, то базисний мінор матриці $\boldsymbol{A}$ має порядок, менший ніж n. Тому є принаймні один стовпець, який не є базисним. За теоремою 3.4.1 він є лінійною комбінацією базисних стовпців. У цю лінійну комбінацію ми можемо включити всі інші небазисні стовпці з нульовими коефіцієнтами, після чого один зі стовпців буде дорівнювати лінійній комбінації інших стовпців матриці. За теоремою 3.1.1 стовпці матриці є лінійно залежними.

Д о с т а т н і с т ь. Якщо стовпці лінійно залежні, то один з них є лінійною комбінацією інших. Додаючи до цього стовпця вказану лінійну комбінацію, взяту з протилежним знаком, будемо мати визначник з нульовим стовпцем, тобто $\det \boldsymbol{A} = 0$.

Аналогічно теорема доводиться для рядків. □

3.4.2. Ступінчасті матриці. Ефективним заходом при визначенні рангу ненульової $m \times n$ -матриці і розв'язанні деяких задач є зведення її за допомогою елементарних перетворень з рядками до *ступінчастої* матриці, яка має таку будову: 1) ненульові рядки розміщені вище нульових; 2) кожний провідний елемент, який є першим ненульовим елементом у своєму рядку (рахуючи зліва направо), дорівнює одиниці; 3) кожний провідний елемент розміщений праворуч від провідного елемента попереднього рядка.

Перетворення матриці до ступінчастого вигляду є узагальненням процесу одержання верхньої трикутної матриці з одиницями на головній діагоналі, яка є окремим випадком ступінчастої матриці. Ця схема, описана в главі 2 стосовно невиродженої квадратної матриці, у застосуванні до довільної $m \times n$ -матриці виглядає так.

Шукаємо в матриці перший стовпець (рахуючи зліва направо), нехай це буде j-й, який має принаймні один ненульовий елемент. Якщо цей ненульовий елемент не знаходиться в першому рядку, то ми переставляємо перший рядок і той рядок, в якому є ненульовий елемент. Цей елемент може бути перетворений в одиницю шляхом множення рядка на належне число. Якщо в j-му стовпці існують інші ненульові елементи, то їх треба перетворити в нуль відповідно до методу виключення.

На другому кроці ми рухаємося від j-го стовпця до наступного, в якому нижче першого рядка є принаймні один ненульовий елемент, і діємо подібно до першого кроку. Продовжуючи цей процес, отримаємо ступінчасту матрицю.

Якщо позначити ступінчасту матрицю через $\boldsymbol{G}$, то

$$\boldsymbol{G} = \boldsymbol{LA}, \tag{3}$$

де $\boldsymbol{L}$ — добуток елементарних матриць, які відповідають елементарним перетворенням матриці $\boldsymbol{A}$. Елементарні перетворення не змінюють ранг матриці, оскільки вони не впливають на властивість визначників бути чи не бути рівними нулю. Тому $r(\boldsymbol{A}) = r(\boldsymbol{G})$.

Ранг матриці $\boldsymbol{G}$ дорівнює кількості r її ненульових рядків. Щоб це довести, достатньо, спираючись на висновок 2 з теореми 3.4.1, показати, що ненульові рядки матриці $\boldsymbol{G}$ є лінійно незалежними. Спробуємо визначити числа $\alpha_1, \ldots, \alpha_r$ таким чином, щоб лінійна комбінація ненульових рядків матриці $\boldsymbol{G}$ дорівнювала нульовому вектору, тобто

$$\sum_{i=1}^{r} \alpha_i \boldsymbol{g}_{i*} = \boldsymbol{0}. \tag{4}$$

Рядок $\boldsymbol{g}_{1*}$ має елемент, рівний одиниці, в тому стовпці, в якому інші рядки мають нулі. Тому рівність (4) можлива лише коли $\alpha_1 = 0$. Аналогічно доводиться, що $\alpha_2 = \ldots = \alpha_r = 0$. Отже, ненульові рядки матриці $\boldsymbol{G}$ лінійно незалежні і $r(\boldsymbol{G}) = r$.

П р и к л а д 1. Визначимо ранг матриці

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

Для цього за допомогою елементарних перетворень з рядками зведемо матрицю $\boldsymbol{A}$ до ступінчастого вигляду:

$$\boldsymbol{A} \sim \begin{pmatrix} 1 & 1/2 & -1/2 & 1 \\ 4 & 2 & 3 & 4 \\ 6 & 3 & 4 & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 7 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 7 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

У ступінчастої матриці два ненульових рядки, тому $r(\boldsymbol{A}) = 2$. △

3.4.3. Скелетний розклад матриці. Представлення $\boldsymbol{A} = \boldsymbol{BC}$ довільної $m \times n$-матриці $\boldsymbol{A}$ рангу $r > 0$ у вигляді добутку матриць $\boldsymbol{B}$ і $\boldsymbol{C}$ розмірів відповідно $m \times r$, $r \times n$ і рангу r називається *скелетним розкладом* матриці $\boldsymbol{A}$. Щоб одержати скелетний розклад матриці $\boldsymbol{A}$, треба спочатку звести її до ступінчастого вигляду, а потім подальшими перетвореннями з рядками одержати нулі вище провідних елементів (це рівносильно перетворенню матриці $\boldsymbol{A}$ за схемою Гаусса—Жордана). Одержана матриця після відкидання в ній нульових рядків вважається множником $\boldsymbol{C}$, а множник $\boldsymbol{B}$ складається з тих стовпців матриці $\boldsymbol{A}$, в яких розміщені провідні елементи.

Очевидно, що побудована таким способом матриця $\boldsymbol{C}$ має необхідні розміри і ранг, рівний r. Що стосується матриці $\boldsymbol{B}$, то слід зауважити, що лінійна незалежність між стовпцями матриці зберігається при елементарних перетвореннях з рядками і тому стовпці матриці $\boldsymbol{A}$, відповідні до провідних елементів методу виключення (їх кількість дорівнює r), є лінійно незалежними. Таким чином, для обгрунтування запропонованої процедури обчислення скелетного розкладу залишається перевірити виконання рівності $\boldsymbol{A} = \boldsymbol{BC}$.

Елементарні перетворення з рядками рівносильні множенню зліва матриці $\boldsymbol{A}$ на елементарні матриці, добуток яких ми позначимо через $\boldsymbol{L}$. В одержаній після цього матриці $\boldsymbol{LA}$ переставимо стовпці так, щоб одержати в лівому верхньому куті одиничну матрицю $\boldsymbol{E}_r$. Це рівносильно множенню справа матриці $\boldsymbol{LA}$ на матриці переставлень, добуток яких ми позначимо через $\boldsymbol{P}$. Отже,

$$\boldsymbol{LAP} = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{F} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \tag{5}$$

де $\boldsymbol{F}$ — деяка підматриця.

Другий множник скелетного розкладу одержується з матриці $(\boldsymbol{E}_r \quad \boldsymbol{F})$ після повернення до попереднього розміщення стовпців, що досягається множенням справа матриці $(\boldsymbol{E}_r \quad \boldsymbol{F})$ на матрицю $\boldsymbol{P}^{-1}$, тобто

$$\boldsymbol{C} = \left(\boldsymbol{E}_r \quad \boldsymbol{F}\right)\boldsymbol{P}^{-1}. \tag{6}$$

Переходячи до визначення першого множника скелетного розкладу, врахуємо, що матриця $\boldsymbol{B}$ складається з перших r стовпців матриці $\boldsymbol{AP}$, тобто в матриці $\boldsymbol{AP}$ треба відкинути зайві стовпці. Це досягається множенням її справа на матрицю $\begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix}$:

$$\boldsymbol{B} = \boldsymbol{AP}\begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix}.$$

Звідси з урахування рівності (5) маємо:

$$\boldsymbol{B} = \boldsymbol{L}^{-1}\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{F} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{L}^{-1}\begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix}. \tag{7}$$

Тепер можна обчислити добуток $\boldsymbol{BC}$ за допомогою формул (7), (6) і (5):

$$\boldsymbol{BC} = \boldsymbol{L}^{-1}\begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix}(\boldsymbol{E}_r \quad \boldsymbol{F})\boldsymbol{P}^{-1} = \boldsymbol{L}^{-1}\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{F} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\boldsymbol{P}^{-1} = \boldsymbol{A}.$$

Отже, множники $\boldsymbol{B}$ і $\boldsymbol{C}$, знайдені у відповідності до запропонованої схеми, є такими, що $\boldsymbol{A} = \boldsymbol{BC}$.

Довільну матрицю, взагалі кажучи, можна різними способами звести до ступінчастого вигляду. Тому скелетний розклад матриці не є єдиним. Зокрема, при $r = m$ нема потреби застосовувати викладену вище процедуру, а натомість, скориставшись визначенням скелетного розкладу, взяти його у вигляді:

$$\boldsymbol{A} = \boldsymbol{EA},$$

тобто вважати, що $\boldsymbol{B} = \boldsymbol{E}$, $\boldsymbol{C} = \boldsymbol{A}$. Аналогічно можна стверджувати, що при $r = n$ скелетний розклад буде таким:

$$\boldsymbol{A} = \boldsymbol{AE},$$

тобто $\boldsymbol{B} = \boldsymbol{A}$, $\boldsymbol{C} = \boldsymbol{E}$.

П р и к л а д 2. Знайдемо скелетний розклад матриці $\boldsymbol{A}$ з прикладу 1. Там матриця $\boldsymbol{A}$ була зведена до ступінчастого вигляду:

$$\boldsymbol{A} \sim \begin{pmatrix} 1 & 1/2 & -1/2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Оскільки ранг матриці є меншим ніж її розміри, то потрібно застосувати викладену вище процедуру знаходження множників скелетного розкладу. Для цього треба продовжити перетворення матриці $\boldsymbol{A}$, щоб одержати нуль вище провідого елемента a_{23}. Це виконується за допомогою другого рядка матриці:

$$\boldsymbol{A} \sim \begin{pmatrix} 1 & 1/2 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Другий множник скелетного розкладу $\boldsymbol{A} = \boldsymbol{BC}$ складається з ненульових рядків останньої матриці:

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

Провідні елементи знаходяться в першому і третьому стовпцях ступінчастої матриці. Тому перший множник скелетного розкладу складається з першого і третього стовпців матриці $\boldsymbol{A}$:

$$\boldsymbol{B} = \begin{pmatrix} 2 & -1 \\ 4 & 3 \\ 6 & 4 \end{pmatrix}.$$

Таким чином,

$$\boldsymbol{A} = \boldsymbol{BC} = \begin{pmatrix} 2 & -1 \\ 4 & 3 \\ 6 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \ \triangle$$

3.4.4. Зведення матриці до діагонального вигляду. У розділі 1.3 було показано, як невироджену квадратну матрицю за допомогою елементарних перетворень рядків можна звести до одиничної матриці. Це ж саме можна зробити за допомогою анологічних перетворень стовпців. Проте для виродженої матриці зведення до діагонального вигляду можливе лише при одночасних перетвореннях рядків і стовпців. Ми покажемо, як це можна здійснити у більш загальному випадку прямокутної матриці.

Теорема 3.4.3. *Для $m \times n$-матриці $\boldsymbol{A}$ рангу r існують такі невироджені матриці **L** і **M** порядків m і n відповідно, що*

$$\boldsymbol{LAM} = \Delta, \quad \Delta = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \tag{8}$$

Д о в е д е н н я. Нехай матриця $\boldsymbol{A}$ зведена до ступінчастого вигляду у відповідності до формули (3). За допомогою елементарних перетворень зі стовпцями обернемо в нуль ненульові елементи кожного рядка, розміщені праворуч від провідного елемента, а потім переставимо стовпці так, щоб перемістити провідні елементи матриці у верхню частину головної діагоналі.

Ці перетворення виконуються за допомогою множення справа матриці $\boldsymbol{G}$ на елементарні матриці. Якщо позначити добуток цих матриць через $\boldsymbol{M}$, то одержимо:

$$\boldsymbol{GM} = \begin{pmatrix} \boldsymbol{E}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{9}$$

З формул (3) і (9) випливає формула (8). У залежності від особливостей матриці $\boldsymbol{A}$ деякі або всі нульові підматриці в $\boldsymbol{\Delta}$ можуть не з'явитися. Зокрема, якщо матриця $\boldsymbol{A}$ порядку n є невиродженою, то $\boldsymbol{LAM} = \boldsymbol{E}_n$. □

Для обчислення матриць $\boldsymbol{L}$ і $\boldsymbol{M}$ можна скористатися такою схемою. За допомогою елементарних перетворень з рядками зведемо матрицю $(\boldsymbol{A} \quad \boldsymbol{E})$ до такого вигляду, коли на місці $\boldsymbol{A}$ буде ступінчаста матриця. Тоді на місці $\boldsymbol{E}$ буде матриця $\boldsymbol{L}$. Дійсно,

$$\boldsymbol{L}(\boldsymbol{A} \quad \boldsymbol{E}) = (\boldsymbol{LA} \quad \boldsymbol{L}) = (\boldsymbol{G} \quad \boldsymbol{L}).$$

Далі у відповідності до теореми 3.4.3 треба застосувати елементарні перетворення зі стовпцями до матриці $\begin{pmatrix} \boldsymbol{G} \\ \boldsymbol{E} \end{pmatrix}$ і отримати на місці $\boldsymbol{G}$ матрицю $\boldsymbol{\Delta}$. Тоді на місці $\boldsymbol{E}$ буде матиця $\boldsymbol{M}$. Насправді,

$$\begin{pmatrix} \boldsymbol{G} \\ \boldsymbol{E} \end{pmatrix} \boldsymbol{M} = \begin{pmatrix} \boldsymbol{GM} \\ \boldsymbol{M} \end{pmatrix} = \begin{pmatrix} \Delta \\ \boldsymbol{M} \end{pmatrix}.$$

Зрозуміло, що матриці $\boldsymbol{L}$ і $\boldsymbol{M}$ визначаються неоднозначно, бо неєдиною є ступінчаста матриця, до якої зводиться матриця $\boldsymbol{A}$.

3.4.5. Ранг добутку і суми матриць. Доведемо теорему про ранг добутку матриць.

Теорема 3.4.4. *Ранг добутку двох матриць не є більшим ніж ранг кожного множника, тобто*

$$r(\boldsymbol{AB}) \le r(\boldsymbol{A}), \quad r(\boldsymbol{AB}) \le r(\boldsymbol{B}). \tag{10}$$

Д о в е д е н н я. Введемо до розгляду матрицю $(\boldsymbol{AB} \quad \boldsymbol{A})$. Оскільки приєднання до матриці $\boldsymbol{AB}$ стовпців іншої матриці не може зменшити ранг матриці $\boldsymbol{AB}$, то $r(\boldsymbol{AB}) \le r(\boldsymbol{AB} \quad \boldsymbol{A})$. З іншого боку, на підставі формули (1.1.2) стовпці матриці $\boldsymbol{AB}$ є лінійними комбінаціями стовпців матриці $\boldsymbol{A}$. Тому $r(\boldsymbol{AB} \quad \boldsymbol{A}) = r(\boldsymbol{A})$. Отже, $r(\boldsymbol{AB}) \le r(\boldsymbol{A})$.

Аналогічно, розглядаючи матрицю $\begin{pmatrix} \boldsymbol{AB} \\ \boldsymbol{B} \end{pmatrix}$, можна довести другу нерівність (10).

Висновок*. Множення будь-якої матриці на невироджену матрицю не змінює її ранг.*

Д о в е д е н н я. Нехай матриця $\boldsymbol{B}$ невироджена. На підставі теореми маємо:

$$r(\boldsymbol{A}) = r((\boldsymbol{AB})\boldsymbol{B}^{-1}) \le r(\boldsymbol{AB}).$$

Звідси і з першої нерівності (10) випливає, що $r(\boldsymbol{AB}) = r(\boldsymbol{A})$.

Аналогічно можна розглянути випадок, коли невиродженою є матриця $\boldsymbol{A}$. □

У розділі 1.3 йшла мова про визначник добутку матриць. Проте це питання не було розглянуте повною мірою і тому ми знову повертаємося до нього.

Теорема 3.4.5. *Визначник добутку двох матриць порядку n дорівнює добутку їх визначників, тобто*

$$\det(\boldsymbol{AB}) = \det\boldsymbol{A} \cdot \det\boldsymbol{B}. \tag{11}$$

Д о в е д е н н я. З урахуванням теореми 1.3.4 можна обмежитися випадком вироджених матриць $\boldsymbol{A}$ і $\boldsymbol{B}$. При цьому ранги цих матриць є меншими ніж n. За теоремою 3.4.4 $r(\boldsymbol{AB}) < n$ і тому $\det(\boldsymbol{AB}) = 0$, тобто рівність (11) виконується. □

Далі розглянемо ще дві теореми, пов’язані з питанням про ранг добутку і суми матриць.

Теорема 3.4.6. *Для того, щоб $m \times n$-матриця $\boldsymbol{A}$ мала ранг* 1*, необхідно і достатньо, щоб вона була добутком двох ненульових матриць розмірів $m \times 1$ і $1 \times n$.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Нехай $r(\boldsymbol{A}) = 1.$ Тоді доведення випливає з можливості представити матрицю $\boldsymbol{A}$ у вигляді скелетного розкладу.

Д о с т а т н і с т ь випливає з теореми 3.4.4. ☐

Теорема 3.4.7. *Ранг суми двох матриць не перевищує суми рангів доданків, тобто*

$$r(\boldsymbol{A} + \boldsymbol{B}) \le r(\boldsymbol{A}) + r(\boldsymbol{B}). \tag{12}$$

Д о в е д е н н я випливає з теореми про базисний мінор при врахуванні тієї обставини, що будь-який стовпець матриці $\boldsymbol{A} + \boldsymbol{B}$ є лінійною комбінацією базисних стовпців матриць $\boldsymbol{A}$ і $\boldsymbol{B}$. ☐

3.5. Системи лінійних рівнянь

3.5.1. Основні визначення. У цьому розділі ми розглянемо *систему лінійних алгебраїчних рівнянь* відносно невідомих $x_1, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2, \\ \ldots\ldots\ldots\ \ldots\ldots\ldots\ \ldots\ \ldots\ &\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m. \end{aligned} \tag{1}$$

Окремий випадок цієї системи, коли $m = n$, розглядався в розділі 1.3 і главі 2. Як і раніше, матрицю $\boldsymbol{A}$, елементами якої є числа a_{ij}, ми будемо називати *матрицею системи*, а матрицю $\boldsymbol{A}_b = (\boldsymbol{A} \quad \boldsymbol{b})$, яка отримується з матриці системи приєднанням до неї стовпця правих частин — *розширеною* матрицею системи. При цьому для системи (1) ми будемо використовувати позначення $\boldsymbol{Ax} = \boldsymbol{b}$, де $\boldsymbol{x} = (x_1 \ \ldots \ x_n)^T$, $\boldsymbol{b} = (b_1 \ \ldots \ b_m)^T$.

Упорядкована сукупність значень невідомих $x_1, \ ..., \ x_n$, яка задовольняє кожному з рівнянь (1), називається *розв'язком* системи. Якщо система має принаймні один розв'язок, то вона називається *сумісною*. У протилежному випадку вона називається *несумісною*. Якщо система сумісна, то кожний її розв'язок називається *окремим*. Сукупність усіх окремих розв'язків називається *загальним* розв'язком.

3.5.2. Однорідні системи рівнянь. *Однорідною* називається така система лінійних рівнянь, в якій всі праві частини дорівнюють нулю, тобто $\boldsymbol{b} = \mathbf{0}$. Така система завжди сумісна, бо вона має розв'язок $\boldsymbol{x} = \mathbf{0}$. Цей розв'язок називається тривіальним (нульовим) розв'язком системи $\boldsymbol{Ax} = \mathbf{0}$. Наступна теорема з'ясовує структуру загального розв'язку однорідної системи.

Теорема 3.5.1. *Якщо* $\boldsymbol{A}$ *є* $m \times n$ *-матрицею рангу r, то множина всіх розв'язків системи* $\boldsymbol{Ax} = \mathbf{0}$ *утворює* $(n - r)$ *-вимірний підпростір простору* $\boldsymbol{R}_n$.

Д о в е д е н н я. Очевидно, що коли $\boldsymbol{x}'$ і $\boldsymbol{x}''$ є розв'язками системи $\boldsymbol{Ax} = \mathbf{0}$, то $\alpha\boldsymbol{x}' + \beta\boldsymbol{x}''$ теж є розв'язком цієї системи, тобто множина розв'язків складає підпростір простору $\boldsymbol{R}_n$. Базис цього підпростору називається *фундаментальною системою* розв'язків.

Нехай базисний мінор розміщений у лівому верхньому куті матриці $\boldsymbol{A}$ (цього можна завжди досягти зміною нумерації невідомих і переставленням рівнянь). Як випливає з теореми про базисний мінор, перші r рядки як матриці $\boldsymbol{A}$, так і матриці $\boldsymbol{A}_b$ є базисними рядками, через які інші рядки представляються як лінійні комбінації. Це означає, що кожне з рівнянь системи, починаючи з $(r + 1)$-го, є наслідком перших r рівнянь. Запишемо систему $\boldsymbol{Ax} = \mathbf{0}$ у блоковому вигляді:

$$\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} = \mathbf{0}, \tag{2}$$

де $\boldsymbol{A}_{11}$ — невироджена підматриця порядку r, $\boldsymbol{A}_{12}$ — підматриця розмірів $r \times (n - r)$, $\boldsymbol{y}$ — вектор вимірності r і т.д. Як сказано вище, останні $m - r$ рядків у формулі (2) є зайвими і ми можемо обмежитися рівнянням

$$\boldsymbol{A}_{11}\boldsymbol{y} + \boldsymbol{A}_{12}\boldsymbol{z} = \boldsymbol{0}. \tag{3}$$

У рівнянні (3) координати вектора $\boldsymbol{z}$ будемо називати *вільними* невідомими. Їм можна надати будь-які числові значення. Координати вектора $\boldsymbol{y}$ будемо називати *базисними* невідомими. Вони визначаються з (3) за допомогою формули $\boldsymbol{y} = -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{z}$.

Таким чином, усі розв'язки системи $\boldsymbol{Ax} = \boldsymbol{0}$ мають вигляд:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\boldsymbol{z} \\ \boldsymbol{z} \end{pmatrix} = \boldsymbol{Bz}, \quad \boldsymbol{B} = \begin{pmatrix} -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ \boldsymbol{E}_{n-r} \end{pmatrix}. \tag{4}$$

Матриця $\boldsymbol{B}$ має розміри $n \times (n - r)$ і ранг $n - r$, а $\boldsymbol{z}$ є довільним $(n - r)$-вимірним вектором. Тому всі вектори підпростору розв'язків на підставі формули (1.1.2) є лінійними комбінаціями $n - r$ лінійно незалежних стовпців матриці $\boldsymbol{B}$.

Отже, вимірність підпростору розв'язків системи дорівнює $n - r$ і за фундаментальну систему розв'язків можна взяти стовпці матриці $\boldsymbol{B}$.

***Висновок**. Система рівнянь* $\boldsymbol{Ax} = \boldsymbol{0}$ *з квадратною матрицею* $\boldsymbol{A}$ *має ненульовий розв'язок тоді і тільки тоді, коли* $\det \boldsymbol{A} = 0$.

Д о в е д е н н я випливає з розглядуваної теореми і теореми 3.4.2. □

П р и к л а д 1. Знайдемо загальний розв'язок і фундаментальну систему розв'язків однорідної системи рівнянь

$$\begin{aligned} x_1 + 2x_2 - x_3 \quad\quad\; + 3x_5 &= 0, \\ 3x_1 + 4x_2 + 3x_3 + 2x_4 + 5x_5 &= 0, \\ 5x_1 + 6x_2 + 7x_3 + 4x_4 + 7x_5 &= 0. \end{aligned}$$

Перетворимо розширену матрицю системи за схемою Гаусса — Жордана (при цьому нульовий стовпець можна не писати):

$$\begin{pmatrix} 1 & 2 & -1 & 0 & 3 \\ 3 & 4 & 3 & 2 & 5 \\ 5 & 6 & 7 & 4 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 & 0 & 3 \\ 0 & -2 & 6 & 2 & -4 \\ 0 & -4 & 12 & 4 & -8 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & -1 & 0 & 3 \\ 0 & 1 & -3 & -1 & 2 \\ 0 & -4 & 12 & 4 & -8 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 5 & 2 & -1 \\ 0 & 1 & -3 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Невідомі x_1 і x_2 є базисними, а x_3, x_4 і x_5 — вільними. У відповідності до останньої матриці напишемо систему рівнянь:

$$\begin{aligned} x_1 \quad\quad + 5x_3 + 2x_4 - x_5 &= 0, \\ x_2 - 3x_3 - x_4 + 2x_5 &= 0. \end{aligned}$$

Звідси знайдемо базисні невідомі:

$$\begin{aligned} x_1 &= -\,5x_3 - 2x_4 + x_5, \\ x_2 &= 3x_3 + x_4 - 2x_5. \end{aligned}$$

Добавивши тотожності $x_3 = x_3$, $x_4 = x_4$, $x_5 = x_5$, будемо мати загальний розв'язок системи:

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -5x_3 - 2x_4 + x_5 \\ 3x_3 + x_4 - 2x_5 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -5 & -2 & 1 \\ 3 & 1 & -2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix},$$

де x_3, x_4 і x_5 відіграють роль параметрів, які можуть набувати будь-яких числових значень. Фундаментальною системою розв'язків є вектори

$$(-5 \quad 3 \quad 1 \quad 0 \quad 0)^T, \; (-2 \quad 1 \quad 0 \quad 1 \quad 0)^T, \; (1 \quad -2 \quad 0 \quad 0 \quad 1)^T. \; \triangle$$

3.5.3. Неоднорідні системи рівнянь. Система рівнянь $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ називається *неоднорідною*, якщо $\boldsymbol{b} \neq \boldsymbol{0}$. При її аналізі перш за все треба відповісти на запитання про умову сумісності, про структуру множини розв'язків, якщо розв'язок неєдиний, та про умову єдиності розв'язку.

Наступна теорема дає відповідь на перше запитання.

Теорема 3.5.2 (теорема Кронекера — Капеллі). *Для того, щоб система* $\boldsymbol{Ax} = \boldsymbol{b}$ *була сумісною, необхідно і достатньо, щоб ранг матриці системи дорівнював рангу розширеної матриці системи, тобто*

$$r(\boldsymbol{A}) = r(\boldsymbol{A}_b).$$

Доведення. Необхідність. Якщо існує розв'язок системи, то $\boldsymbol{b}$ є лінійною комбінацією стовпців матриці $\boldsymbol{A}$. Тому $r(\boldsymbol{A}_b) = r(\boldsymbol{A})$.

Достатність. Якщо $r(\boldsymbol{A}_b) = r(\boldsymbol{A})$, то $\boldsymbol{b}$ є лінійною комбінацією стовпців матриці $\boldsymbol{A}$. Коефіцієнти цієї лінійної комбінації визначають розв'язок системи $\boldsymbol{Ax} = \boldsymbol{b}$. □

Сформульована в теоремі Кронекера — Капеллі умова сумісності системи $\boldsymbol{Ax} = \boldsymbol{b}$ не є досить зручною. На практиці нема потреби порівнювати ранги матриць $\boldsymbol{A}$ і $\boldsymbol{A}_b$. Не гаючи часу на таке порівняння, можна, застосувавши до системи метод виключення, одержати її розв'язок або довести її несумісність.

З'ясуємо зв'язок між розв'язками неоднорідної і відповідної однорідної систем. Нехай $\boldsymbol{f}_0$ — окремий розв'язок системи $\boldsymbol{Ax} = \boldsymbol{b}$. Якщо шукати її загальний розв'язок у вигляді суми $\boldsymbol{f}_0 + \boldsymbol{x}$, то після підставлення цієї суми в систему одержимо, що $\boldsymbol{Ax} = \boldsymbol{0}$. Загальний розв'язок цієї системи, як з'ясовано в теоремі 3.5.1, дорівнює лінійній комбінації її фундаментальних розв'язків. Отже, загальний розв'язок неоднорідної системи дорівнює сумі загального розв'язку відповідної однорідної системи і окремого розв'язку неоднорідної системи.

При $n - r = 0$ загальний розв'язок однорідної системи збігається з її тривіальним розв'язком. У цьому випадку загальний розв'язок системи $\boldsymbol{Ax} = \boldsymbol{b}$ буде дорівнювати $\boldsymbol{f}_0$ і може бути знайдений, наприклад, методом Гаусса. Отже, система $\boldsymbol{Ax} = \boldsymbol{b}$ має єдиний розв'язок $\boldsymbol{f}_0$, коли кількість невідомих дорівнює рангу матриці системи.

Питання про загальний розв'язок системи лінійних рівнянь і умову його існування буде наново досліджене на іншій основі в розділі 11.4.

П р и к л а д 2. Знайдемо загальний розв'язок неоднорідної систем рівнянь

$$\begin{aligned} x_1 + 2x_2 + 2x_3 - x_4 - x_5 &= 2, \\ x_1 + 3x_2 + 2x_3 - 2x_4 - x_5 &= -1, \\ x_1 - x_2 + 2x_3 - 3x_4 + 4x_5 &= -4. \end{aligned}$$

Перетворимо розширену матрицю системи за схемою Гаусса — Жордана:

$$\boldsymbol{A}_b \sim \begin{pmatrix} 1 & 2 & 2 & -1 & -1 & 2 \\ 1 & 3 & 2 & -2 & -1 & -1 \\ 1 & -1 & 2 & -3 & 4 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 2 & -1 & -1 & 2 \\ 0 & 1 & 0 & -1 & 0 & -3 \\ 0 & -3 & 0 & -2 & 5 & -6 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 2 & 1 & -1 & 8 \\ 0 & 1 & 0 & -1 & 0 & -3 \\ 0 & 0 & 0 & -5 & 5 & -15 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 & 1 & -1 & 8 \\ 0 & 1 & 0 & -1 & 0 & -3 \\ 0 & 0 & 0 & 1 & -1 & 3 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 0 & 2 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 3 \end{pmatrix}.$$

Базисними невідомими є x_1, x_2 і x_4, а вільними — x_3 і x_5. У відповідності до перетвореної матриці маємо систему рівнянь:

$$\begin{aligned} x_1 + 2x_3 &= 5, \\ x_2 - x_5 &= 0, \\ x_4 - x_5 &= 3. \end{aligned}$$

Звідси знаходимо базисні невідомі:

$$\begin{aligned} x_1 &= 5 - 2x_3, \\ x_2 &= x_5, \\ x_4 &= 3 + x_5. \end{aligned}$$

Добавивши сюди тотожності $x_3 = x_3,\ x_5 = x_5,$ одержимо загальний розв'язок системи:

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 - 2x_3 \\ x_5 \\ x_3 \\ 3 + x_5 \\ x_5 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{pmatrix} + \begin{pmatrix} -2 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_5 \end{pmatrix}.$$

Перший доданок правої частини є окремим розв'язком неоднорідної системи, а другий доданок — загальним розв'язком відповідної однорідної системи. △

3.5.4. Геометричний зміст множини розв'язків неоднорідної системи лінійних рівнянь. Теорія систем лінійних рівнянь дозволяє продовжити вивчення площин в афінному просторі. Далі в цій главі ми позначаємо площину вимірності m і її спрямовуючий підпростір через Π_m і U_m відповідно. Нехай в афінному просторі S_n введена афінна система координат. Тоді кожному розв'язку $\boldsymbol{x} = (x_1 \ \ldots \ x_n)^T$ системи (1) можна зіставити точку простору S_n з координатами $x_1, \ldots, x_n$. У двох наступних теоремах з'ясовується геометричний зміст множини розв'язків системи рівнянь $\boldsymbol{Ax} = \boldsymbol{b}$.

Теорема 3.5.3. *Усі розв'язки системи* $\boldsymbol{Ax} = \boldsymbol{b}$ *утворюють у просторі* S_n *площину вимірності* $n - r$, *де* $r = r(\boldsymbol{A})$.

Д о в е д е н н я. Вище було показано, що загальний розв'язок $\boldsymbol{x}$ системи $\boldsymbol{Ax} = \boldsymbol{b}$ дорівнює

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{n-r} \alpha_i \boldsymbol{f}_i ,$$

де $\boldsymbol{f}_0$ — окремий розв'язок цієї системи, $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_{n-r}$ — фундаментальна система розв'язків однорідної системи $\boldsymbol{Ax} = \boldsymbol{0}$, а числа $\alpha_1, \ldots, \alpha_{n-r}$ є довільними. На підставі формули (3.3.2) можна стверджувати, що загальний розв'язок системи $\boldsymbol{Ax} = \boldsymbol{b}$ визначає площину вимірності $n - r$ в просторі S_n. □

Теорема 3.5.4. *В афінному просторі* S_n *і в будь-яких афінних координатах усяка площина* Π_m *може бути задана системою* $\boldsymbol{Ax} = \boldsymbol{b}$ *з матрицею* $\boldsymbol{A}$ *рангу* $n - m$.

Д о в е д е н н я. З векторного рівняння площини

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i \tag{8}$$

випливає, що вектор $\boldsymbol{x} - \boldsymbol{f}_0$ є лінійною комбінацією лінійно незалежних векторів $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$, які породжують спрямовуючий підпростір площини. Для того, щоб вектор $\boldsymbol{x} - \boldsymbol{f}_0$ належав підпростору, породженому векторами $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$, необхідно і достатньо, щоб виконувалася рівність

$$r(\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_m) = r(\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_m \quad \boldsymbol{x} - \boldsymbol{f}_0).$$

Тут вектори $\boldsymbol{f}_0, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$ вважаються заданими своїми координатами в базисі, відповідному до афінної системи координат.

Перетворимо матриці $(\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_m)$ і $(\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_m \quad \boldsymbol{x} - \boldsymbol{f}_0)$ до ступінчастого вигляду. Ранг матриці $(\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_m)$ дорівнює m, тому в перетвореній матриці $(\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_m \quad \boldsymbol{x} - \boldsymbol{f}_0)$ останні $n - m$ елементи останнього стовпця повинні бути нульовими. Це означає, що ми будемо мати $n - m$ рівнянь відносно n координат вектора $\boldsymbol{x}$. При цьому кожне з цих рівнянь вміщуватиме лише одну з невідомих $x_{m+1}, \ldots, x_n$ з коефіцієнтом, рівним одиниці, тобто базисному мінору, відповідному до цих змінних, буде відповідати одинична матриця $\boldsymbol{E}_{n-m}$.

Отже, площина (8) визначається системою $n - m$ рівнянь відносно n невідомих з матрицею рангу $n - m$. ☐

З теореми випливає, що гіперплощина визначається одним лінійним рівнянням

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1.$$

Кожне з рівнянь системи, про яку йшла мова в теоремі 3.5.4, можна розглядати як рівняння деякої гіперплощини і кожну площину Π_m можна розглядати як перетин $n - m$ гіперплощин.

Якщо система лінійних рівнянь несумісна, то це означає, що не існує жодної точки, що належить усім гіперплощинам, які задаються рівняннями системи. Очевидно, що площину Π_m можна задати різними системами рівнянь, тобто її можна визначити як перетин різних наборів гіперплощин у кількості $n - m$. При цьому ранг матриці сумісної системи рівнянь цих гіперплощин повинен мати максимально можливе значення, рівне кількості рівнянь.

3.5.5. Базис суми і перерізу підпросторів. Скористаємося методом виключення для знаходження базису суми і перерізу підпросторів.Будемо розглядати випадок двох підпросторів, бо загальний випадок досліджується повторенням міркувань.

Нехай підпростори U і V векторного простору породжуються векторами $\boldsymbol{u}_1, ..., \boldsymbol{u}_k$ і $\boldsymbol{v}_1, ..., \boldsymbol{v}_l$ відповідно.На підставі визначення суми $U + V$ вона є множиною векторів, які представляються всілякими лінійними комбінаціями векторів $\boldsymbol{u}_1, ..., \boldsymbol{u}_k, \boldsymbol{v}_1, ..., \boldsymbol{v}_l$. Зрозуміло,що всі ці вектори виражаються через базисні вектори підпросторів U і V. Тому базис суми $U + V$ є об’єднанням базисів підпросторів U і V. Він є максимальною за чисельністю системою лінійно незалежних векторів,відібраних з векторів $\boldsymbol{u}_1, ..., \boldsymbol{u}_k, \boldsymbol{v}_1, ..., \boldsymbol{v}_l$, причому $\dim(U + V) = r(\boldsymbol{A})$, де $\boldsymbol{A}$ — матриця, рядками якої є координати вказаних векторів.

Якщо вести мову про переріз $U \cap V$, то найбільш просто вирішується питання про його базис у випадку, коли підпростори визначаються системами рівнянь $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{0}$, $\boldsymbol{A}_2\boldsymbol{x} = \boldsymbol{0}$. Вектор $\boldsymbol{x}$ належить перерізу $U \cap V$ тоді і тільки тоді, коли його координати задовольняють обом системам. Інакше кажучи, базисом перерізу $U \cap V$ є фундаментальна система розв’язків «об’єднаної» однорідної системи $\begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix}\boldsymbol{x} = \boldsymbol{0}$.

Особливості побудови базисів суми і перерізу підпросторів з’ясуємо, розглянувши два приклади.

П р и к л а д 3. Знайдемо вимірність і базис суми підпросторів U і V, натягнутих відповідно на вектори

$$\boldsymbol{u}_1 = (1 \;\; 1 \;\; -1 \;\; 0)^T, \quad \boldsymbol{u}_2 = (0 \;\; 1 \;\; 0 \;\; 2)^T;$$

$$\boldsymbol{v}_1 = (-1 \;\; 1 \;\; 1 \;\; 4)^T, \quad \boldsymbol{v}_2 = (3 \;\; 2 \;\; -3 \;\; -2)^T.$$

Складемо матрицю, рядками якої є координати векторів $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}_1, \boldsymbol{v}_2$. Напишемо також ці вектори в останньому стовпці. Одержану матрицю перетворимо до ступінчастого вигляду:

$$\left(\begin{array}{rrrr|l} 1 & 1 & -1 & 0 & \boldsymbol{u}_1 \\ 0 & 1 & 0 & 2 & \boldsymbol{u}_2 \\ -1 & 1 & 1 & 4 & \boldsymbol{v}_1 \\ 3 & 2 & -3 & -2 & \boldsymbol{v}_2 \end{array}\right) \sim \left(\begin{array}{rrrr|c} 1 & 1 & -1 & 0 & \boldsymbol{u}_1 \\ 0 & 1 & 0 & 2 & \boldsymbol{u}_2 \\ 0 & 2 & 0 & 4 & \boldsymbol{u}_1 + \boldsymbol{v}_1 \\ 0 & -1 & 0 & -2 & -3\boldsymbol{u}_1 + \boldsymbol{v}_2 \end{array}\right) \sim$$

$$\sim \left(\begin{array}{rrrr|c} 1 & 1 & -1 & 0 & \boldsymbol{u}_1 \\ 0 & 1 & 0 & 2 & \boldsymbol{u}_2 \\ 0 & 0 & 0 & 0 & \boldsymbol{u}_1 - 2\boldsymbol{u}_2 + \boldsymbol{v}_1 \\ 0 & 0 & 0 & 0 & -3\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{v}_2 \end{array}\right).$$

Вектори $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}_1, \boldsymbol{v}_2$ є лінійно залежними, оскільки ранг одержаної матриці (без останнього стовпця) дорівнює 2, тобто є меншим ніж 4. Щоб два останні рядки були нульовими, треба вважати, що $\boldsymbol{u}_1 - 2\boldsymbol{u}_2 + \boldsymbol{v}_1 = \mathbf{0}, \; -3\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{v}_2 = \mathbf{0}.$ Це — лінійні залежності між векторами $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v}_1, \boldsymbol{v}_2$. При цьому $\boldsymbol{u}_1$ і $\boldsymbol{u}_2$ є лінійно незалежними, тому вони утворюють базис підпростору $U + V$ і $\dim(U + V) = 2$. △

П р и к л а д 4. Знайдемо вимірність і базис перерізу підпросторів U і V з попереднього прикладу. Для цього спочатку відшукаємо системи однорідних рівнянь, які визначають підпростори U і V.

Для того, щоб вектор $\boldsymbol{x} = (x_1 \;\; x_2 \;\; x_3 \;\; x_4)^T$ належав підпростору U, натягнутому на вектори $\boldsymbol{u}_1$ і $\boldsymbol{u}_2$, необхідно і достатньо, щоб $r(\boldsymbol{u}_1 \;\; \boldsymbol{u}_2) = r(\boldsymbol{u}_1 \;\; \boldsymbol{u}_2 \;\; \boldsymbol{x})$. Складемо матрцю $(\boldsymbol{u}_1 \;\; \boldsymbol{u}_2 \;\; \boldsymbol{x})$ і перетворимо її до ступінчастого вигляду:

$$(\boldsymbol{u}_1 \;\; \boldsymbol{u}_2 \;\; \boldsymbol{x}) =$$

$$= \begin{pmatrix} 1 & 0 & x_1 \\ 1 & 1 & x_2 \\ -1 & 0 & x_3 \\ 0 & 2 & x_4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & x_1 \\ 0 & 1 & -x_1 + x_2 \\ 0 & 0 & x_1 + x_3 \\ 0 & 2 & x_4 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & x_1 \\ 0 & 1 & -x_1 + x_2 \\ 0 & 0 & x_1 + x_3 \\ 0 & 0 & 2x_1 - 2x_2 + x_4 \end{pmatrix}.$$

На підставі умови $r(\boldsymbol{u}_1 \;\; \boldsymbol{u}_2) = r(\boldsymbol{u}_1 \;\; \boldsymbol{u}_2 \;\; \boldsymbol{x})$ можна написати систему рівнянь, що визначає підпростір U:

$$\begin{aligned} x_1 \qquad\quad + x_3 \qquad &= 0, \\ 2x_1 - 2x_2 \qquad + x_4 &= 0. \end{aligned}$$

Аналогічно можна одержати систему рівнянь, яка описує підпростір V. Вона виявляється такою ж, як і попередня:

$$\begin{aligned} x_1 \qquad\quad + x_3 \qquad &= 0, \\ 2x_1 - 2x_2 \qquad + x_4 &= 0. \end{aligned}$$

Щоб знайти $U \bigcap V$, об’єднаємо одержані системи:

$$\begin{aligned} x_1 \qquad\quad + x_3 \qquad &= 0, \\ 2x_1 - 2x_2 \qquad + x_4 &= 0. \\ x_1 \qquad\quad + x_3 \qquad &= 0, \\ 2x_1 - 2x_2 \qquad + x_4 &= 0. \end{aligned}$$

Її фундаментальна система розв’язків, яка складається з векторів $(1 \;\; 0 \;\; -1 \;\; -2)^T$ і $(0 \;\; 1 \;\; 0 \;\; 2)^T$ є базисом перерізу $U \bigcap V$. Зрозуміло, що $\dim(U \bigcap V) = 2$. $\triangle$

3.6. Взаємне розміщення площин

3.6.1. Перетинні, паралельні, перехресні площини. При одночасному розгляді кількох площин виникають різні задачі і в першу чергу задача визначення їх взаємного розміщення. Властивості підпросторів і систем лінійних рівнянь дають можливість дослідити це питання.

Дві площини можуть бути перетинними, паралельними або перехресними.

Якщо площини Π_k і Π_l є *перетинними,* то їх перерізом є деяка площина Π_m. Наприклад, дві двовимірні площини перетинаються по прямій лінії, тобто $k = l = 2$, $m = 1$ (рис. 3). Може так статися, що Π_m складається з однієї точки. Це буде у випадку перетину двох прямих ліній або прямої лінії і площини (рис.4). У довільному просторі S_n в одній точці можуть перетинатися площини, сума вимірностей яких дорівнює n. Може бути так, що одна з площин цілком належить іншій. Зокрема, пряма лінія може належати площині, наприклад, при $k = m = 1$, $l = 2$ (рис.5).

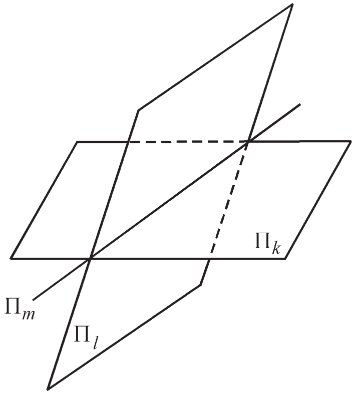

Рис. 3

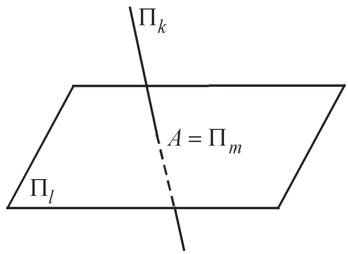

Рис. 4

Нижче ми побачимо, за яких умов пряма лінія належить площині, а зараз з'ясуємо, коли дві прямі лінії є перетинними. Нехай вони визначаються векторними рівняннями

$$\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1, \quad \boldsymbol{x} = \boldsymbol{g}_0 + \beta_1 \boldsymbol{g}_1, \qquad (1)$$

причому припускається, що вектори $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$ не є колінеарними.

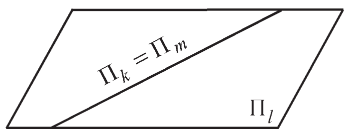

Рис. 5

Теорема 3.6.1. *Якщо вектор* $\boldsymbol{f}_0 - \boldsymbol{g}_0$ *лінійно виражається через вектори* $\boldsymbol{f}_1$ *і* $\boldsymbol{g}_1$, *то прямі лінії* (*1*) *перетинаються.*

Д о в е д е н н я. Умову теореми можна записати так:

$$\boldsymbol{f}_0 - \boldsymbol{g}_0 = \gamma_1 \boldsymbol{f}_1 + \gamma_2 \boldsymbol{g}_1, \tag{2}$$

де $\gamma_1,\ \gamma_2$ — деякі числа.

При будь-якому α_1 точка

$$\boldsymbol{x}_0 = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1$$

належить першій прямій лінії (1). Спробуємо визначити α_1 так, щоб точка $\boldsymbol{x}_0$ одночасно належала другій прямій лінії (1). Для цього підставимо в останню рівність вираз для $\boldsymbol{f}_0$, знайденний з (2). При цьому одержимо, що

$$\boldsymbol{x}_0 = \boldsymbol{g}_0 + (\alpha_1 + \gamma_1)\boldsymbol{f}_1 + \gamma_2 \boldsymbol{g}_1.$$

Будемо вважати, що $\alpha_1 = -\gamma_1$. Тоді

$$\boldsymbol{x}_0 = \boldsymbol{g}_0 + \gamma_2 \boldsymbol{g}_1,$$

звідки випливає, що точка $\boldsymbol{x}_0$ належить другій прямій лінії. Отже, прямі лінії (1) перетинаються. ☐

Для практичного розв'язання задачі треба скласти матрицю $\boldsymbol{A}$, рядками якої є координати векторів $\boldsymbol{a} = \boldsymbol{f}_0 - \boldsymbol{g}_0,\ \boldsymbol{f}_1,\ \boldsymbol{g}_1$, і обчислити її ранг. Він може дорівнювати 1, 2 або 3. При $r(\boldsymbol{A}) = 1$ вектор $\boldsymbol{a}$ виражається лінійно через $\boldsymbol{f}_1$ (або $\boldsymbol{g}_1$), тобто прямі лінії є перетинними. Неважко зрозуміти, що прямі лінії збігаються. Якщо $r(\boldsymbol{A}) = 2$ і вектори $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$ є колінеарними, то вектор $\boldsymbol{a}$ не виражається лінійно через $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$, тобто прямі лінії не перетинаються. При $r(\boldsymbol{A}) = 2$ і лінійно незалежних векторах $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$ можна стверджувати, що $\boldsymbol{a}$ є лінійною комбінацією векторів $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$. Це означає, що прямі лінії є перетинними. Нарешті, при $r(\boldsymbol{A}) = 3$ вектори $\boldsymbol{a},\ \boldsymbol{f}_1,\ \boldsymbol{g}_1$ є лінійно незалежними і тому вектор $\boldsymbol{a}$ неможливо представити у вигляді лінійної комбінації векторів $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$. Отже, прямі лінії не перетинаються при $r(\boldsymbol{A}) = 3$.

Нехай площини Π_k, Π_l визначаються точками A, B і підпросторами U_k, U_l. Будемо вважати, що $k \le l$. Площина Π_k називається *паралельною* до площини Π_l, якщо $U_k \subseteq U_l$. Говорять також, що площина Π_l паралельна до площини Π_k. Якщо площина Π_k паралельна до площини Π_l і $k = l$, то U_k збігається з U_l.

Неважко переконатися в тому, що при $n = 3$ окремі випадки, коли: 1) $k = l = 1$; 2) $k = l = 2$; 3) $k = 1$, $l = 2$, узгоджуються з поняттям паралельності прямих і площин, відомим з елементарної геометрії (рис. 6).

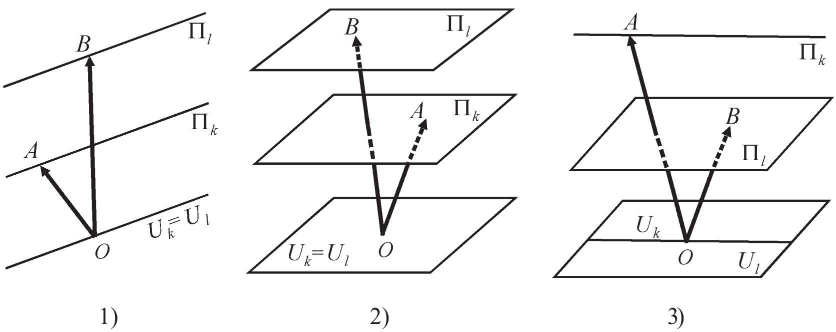

Рис. 6

Для паралельності двох площин однакової вимірності, які визначаються системами лінійних рівнянь, необхідно і достатньо, щоб відповідні однорідні системи були еквівалентними. Зокрема, дві гіперплощини

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ \dots \ + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \ \dots \ + a_{2n}x_n &= b_2 \end{aligned} \tag{3}$$

паралельні тоді і тільки тоді, коли коефіцієнти при $x_1, \dots, x_n$ є пропорційними:

$$\frac{a_{11}}{a_{21}} = \frac{a_{12}}{a_{22}} = \dots = \frac{a_{1n}}{a_{2n}}. \tag{4}$$

Вище зазначалося, що одна з двох перетинних площин може цілком належати іншій. Докладніше про це йдеться в наступній теоремі.

Теорема 3.6.2. *Якщо площина* Π_k *паралельна до площини* Π_l*, де* $l \geq k$*, і площини мають принаймні одну спільну точку, то площина* Π_k *належить площині* Π_l.

Д о в е д е н н я. Будемо вважати, що спільна точка площин є початковою точкою для кожної з них. Таке припущення є виправданим на підставі висновку з теореми 3.3.1. З урахуванням цього припущення запишемо векторні рівняння площин у вигляді:

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i, \quad \boldsymbol{x} = \boldsymbol{f}_0 + \sum_{j=1}^{l} \beta_j \boldsymbol{g}_j. \tag{5}$$

Покажемо, що будь-яка точка $\boldsymbol{x}_0 \in \Pi_k$ належить площини Π_l.

На підставі теореми 3.3.1 буде $\boldsymbol{x}_0 - \boldsymbol{f}_0 \in U_k$. Оскільки площини є паралельними, то $U_k \subseteq U_l$, тому $\boldsymbol{x}_0 - \boldsymbol{f}_0 \in U_l$. Інакше кажучи, вектор $\boldsymbol{x}_0 - \boldsymbol{f}_0$ лінійно виражається через вектори $\boldsymbol{g}_1, \ ..., \ \boldsymbol{g}_l$:

$$\boldsymbol{x}_0 - \boldsymbol{f}_0 = \sum_{j=1}^{l} \beta'_j \boldsymbol{g}_j.$$

Порівняння цієї рівності з другим рівнянням (5) свідчить про належність точки $\boldsymbol{x}_0$ площині Π_l. □

Як приклад розглянемо питання про умови, за яких пряма лінія $\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1$ належить площині $\boldsymbol{x} = \boldsymbol{g}_0 + \sum_{j=1}^{l} \beta_j \boldsymbol{g}_j$. Спрямовуючий підпростір U_1 прямої лінії породжується вектором $\boldsymbol{f}_1$, а спрямовуючий підпростір U_l площини — векторами $\boldsymbol{g}_1, \ ..., \ \boldsymbol{g}_l$. Якщо пряма лінія і площина паралельні, то $U_1 \subset U_l$, тобто вектор $\boldsymbol{f}_1$ повинен лінійно виражатися через $\boldsymbol{g}_1, \ ..., \ \boldsymbol{g}_l$. Крім того, при наявності спільної точки у прямої лінії і площини можна вважати, що такою точкою є $\boldsymbol{f}_0$. Тоді на підставі теореми 3.3.1 $\boldsymbol{f}_0 - \boldsymbol{g}_0 \in U_l$, що рівносильно вимозі, щоб вектор $\boldsymbol{f}_0 - \boldsymbol{g}_0$ лінійно виражався через $\boldsymbol{g}_1, \ ..., \ \boldsymbol{g}_l$.

Отже, потрібно перевірити, чи є вектори $\boldsymbol{f}_1$ і $\boldsymbol{f}_0 - \boldsymbol{g}_0$ лінійними комбінаціями векторів $\boldsymbol{g}_1, \ \dots, \ \boldsymbol{g}_l$. Це можна зробити, дотримуючись процедури, описаної в прикладі 3 з розділу 3.5.

Дві площини називаються *перехресними*, якщо вони не перетинаються і не є паралельними. У просторі S_3 дві прямі лінії, тобто одновимірні площини, можуть схрещуватися, проте пряма лінія і двовимірна площина в S_3 схрещуватися не можуть. Простори більшої вимірності є більш просторими і в них можна побудувати перехресні площини різних вимірностей, а не тільки одновимірні. Проте, як і в тривимірному просторі, гіперплощина не може схрещуватися з іншими площинами. Особливо просто в цьому пересвідчитися, розглядаючи дві гіперплощини.

Дійсно, якщо гіперплощини (3) не є паралельними, то принаймні одна з рівностей (4) не виконується. Припустимо, що $\dfrac{a_{11}}{a_{21}} \neq \dfrac{a_{12}}{a_{22}}$, тоді систему (3) можна розв'язати відносно x_1 і x_2, тобто гіперплощини перетинаються.

Розглянемо загальний випадок. Має місце

Теорема 3.6.3. *Якщо площина не має спільних точок з гіперплощиною, то площина паралельна до гіперплощини.*

Д о в е д е н н я. Нехай векторні рівняння площини Π_m, де $m < n-1$, і гіперплощини Π_{n-1} мають вигляд:

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i, \quad \boldsymbol{x} = \boldsymbol{g}_0 + \sum_{j=1}^{n-1} \beta_j \boldsymbol{g}_j.$$

Оскільки площина і гіперплощина не перетинаються, то вони можуть бути або паралельними, або перехресними. Покажемо, що вони є паралельними.

Умова теореми означає, що рівність

$$\boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i = \boldsymbol{g}_0 + \sum_{j=1}^{n-1} \beta_j \boldsymbol{g}_j \tag{6}$$

не може виконуватися ні при яких числових значеннях параметрів $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_{n-1}$. Якщо всі вектори $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$ лінійно виражаються через вектори $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}$, то площина Π_m паралельна до гіперплощини Π_{n-1}. Припустимо, що принаймні один з векторів $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$, нехай це буде $\boldsymbol{f}_m$, не є лінійною комбінацією векторів $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}$. Тоді вектори $\boldsymbol{f}_m$, $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}$ утворюють базис векторного простору. При цьому вектор $\boldsymbol{f}_0 - \boldsymbol{g}_0 + \sum_{i=1}^{m-1} \alpha_i \boldsymbol{f}_i$ можна розкласти за цим базисом:

$$\boldsymbol{f}_0 - \boldsymbol{g}_0 + \sum_{i=1}^{m-1} \alpha_i \boldsymbol{f}_i = \gamma_0 \boldsymbol{f}_m + \sum_{j=1}^{n-1} \gamma_j \boldsymbol{g}_j . \tag{7}$$

Але це є рівність (6) при $\gamma_0 = -\alpha_m$, $\gamma_j = \beta_j$, $j = \overline{1,\ n-1}$. Виконання рівності (7) суперечить умові невиконання рівності (6). Отже, неможливо припускати, що якийсь з векторів $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$ не є лінійною комбінацією векторів $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{n-1}$. Тому площина Π_m паралельна до гіперплощини Π_{n-1}. ☐

3.6.2. Системи лінійних нерівностей і опуклі многогранники. Нехай через точки *A* і *B*, радіус-вектори яких дорівнюють $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$, проходить пряма лінія. Її векторне рівняння має вигляд

$$\boldsymbol{x} = \boldsymbol{x}_1 + \alpha_1(\boldsymbol{x}_2 - \boldsymbol{x}_1) . \tag{8}$$

При $\alpha_1 = 0$ це рівняння визначає точку *A*, при $\alpha_1 = 1$ — точку *B*. Множина точок прямої лінії, що задовольняють нерівностям $0 \le \alpha_1 \le 1$, називається *відрізком AB*. Звичайно користуються формулою (8), записаною в симетричному вигляді:

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 , \tag{9}$$

де

$$\lambda_1 + \lambda_2 = 1, \quad \lambda_1,\ \lambda_2 \ge 0 . \tag{10}$$

Точка, відповідна до $\lambda_1 = \lambda_2 = 1/2$, називається серединою відрізка *AB*.

Формули (9), (10) визначають так звану опуклу комбінацію точок $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$. Аналогічно визначається опукла комбінація будь-якої кількості точок.

Множина точок дійсного афінного простору називається *опуклою*, якщо разом з кожними двома своїми точками A, B вона вміщує відрізок AB. Найпростішими прикладами опуклих множин є пряма лінія, відрізок, весь простір S_n. Множина, що складається з однієї точки, і пуста множина теж вважаються опуклими.

Нехай у просторі S_n задана гіперплощина $\sum\limits_{i=1}^{n} a_{1i}x_i = b$. Запишемо це рівняння у вигляді:

$$\boldsymbol{c}^T\boldsymbol{x} = b,$$

де $\boldsymbol{c} = (a_{11} \quad \ldots \quad a_{1n})^T$, $\boldsymbol{x} = (x_1 \quad \ldots \quad x_n)^T$. Гіперплощина розбиває простір на дві частини, точки яких характеризуються нерівностями $\boldsymbol{c}^T\boldsymbol{x} \le b$, $\boldsymbol{c}^T\boldsymbol{x} \ge b$. Кожна з цих частин називається *напівпростором*. Гіперплощина є спільною частиною напівпросторів.

Теорема 3.6.4. *Напівпростір є опуклою множиною.*

Д о в е д е н н я. Нехай точки A, B відрізку AB належать одному і тому ж напівпростору, тобто величини $\boldsymbol{c}^T\boldsymbol{x}_1 - b$ і $\boldsymbol{c}^T\boldsymbol{x}_2 - b$ мають однакові знаки. Якщо радіус-вектор $\boldsymbol{x}$ довільної точки C відрізку AB визначається формулою (8), то

$$\boldsymbol{c}^T\boldsymbol{x} - b = (1-\alpha_1)(\boldsymbol{c}^T\boldsymbol{x}_1 - b) + \alpha_1(\boldsymbol{c}^T\boldsymbol{x}_2 - b).$$

Звідси випливає, що при $0 \le \alpha_1 \le 1$ величина $\boldsymbol{c}^T\boldsymbol{x} - b$ має той же знак, що і величини $\boldsymbol{c}^T\boldsymbol{x}_1 - b$, $\boldsymbol{c}^T\boldsymbol{x}_2 - b$.

Отже, точка C належить тому ж напівпростору, що і точки A, B, тобто напівпростір є опуклою множиною. ☐

Переріз будь-якої сукупності опуклих множин є опуклою множиною, бо якщо точки A, B належать перерізу опуклих множин, то відрізок AB належить кожній з цих множин і, отже, їх перерізу.

Переріз скінченної кількості напівпросторів є опуклою множиною, яка у випадку непустої множини називається *опуклим многогранником*. З цього визначення випливає, що опуклий многогранник є множиною точок афінного простору, які задовольняють системі нерівностей $\boldsymbol{c}_i^T \boldsymbol{x} \le b_i, \quad i = \overline{1,\ m}$. Можна довести, що коли многогранник є *обмеженим*, тобто всі координати його точок задовольняють нерівності $|\, x_i \,| < M$ (M — деяке додатне число), то кількість m напівпросторів, перерізом яких є многогранник, обов'язково перевищує вимірність простору.

Вимірністю опуклого многогранника називають мінімальну з вимірностей площин, що його вміщують. Наприклад, вимірність відрізка дорівнює одиниці, вимірність трикутника — двом, вимірність трикутної піраміди — трьом.

Прикладом опуклого многогранника вимірності n є n*-вимірний паралелепіпед,* побудований на базисних векторах, з вершиною в початку координат. Такий паралелепіпед задається нерівностями

$$0 \le x_i \le 1, \qquad i = \overline{1,\ n}\,. \tag{11}$$

Якщо $n = 1$, то паралелепіпед являє собою відрізок, при $n = 2$ він є паралелограмом.

Частина паралелепіпеда (11), розміщена в одній з гіперплощин $x_i = 0$ або $x_i = 1$, є $(n-1)$-вимірним паралелепіпедом і називається $(n-1)$-вимірною *гранню* паралелепіпеда (11). Далі можна розглядати грані цих $(n-1)$-вимірних паралелепіпедів, грані їх граней і т.д. Усі вони називаються m-вимірними гранями паралелепіпеда (11), де $0 \le m \le n-1$. Одновимірні грані називаються *ребрами*, а їх кінці — *вершинами* паралелепіпеда. Вершини є нульвимірними гранями.

Аналогічні твердження щодо граней залишаються в силі для довільного опуклого многогранника. Проте треба мати на увазі, що мінімальна вимірність його грані на відміну від паралелепіпеда може бути додатною величиною. Наприклад, нерівності $x_1 \ge 0,\ \ x_1 \le 1$ визначають многогранник, який є частиною простору, обмеженою паралельними гіперплощинами $x_1 = 0,\ \ x_1 = 1$. Ці гіперплощини є єдиними гранями опуклого многогранника.

Множина $K \subset S_n$ називається *опуклим конусом*, якщо разом з будь-якими своїми точками $\boldsymbol{x}_1,\ \boldsymbol{x}_2$ вона вміщує всі точки $\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2$ при $\alpha_1,\ \alpha_2 \geq 0$. З цього визначення випливає, що конус K є опуклою множиною. Крім того, разом з вектором $\boldsymbol{x}_1$ він вміщує промінь $\alpha_1\boldsymbol{x}_1,\ \alpha_1 \geq 0$. Очевидно, що будь-який опуклий конус вміщує початок координат.

На двовимірній площині такі конуси — це кути, що не перевищують π. У тривимірному просторі прикладами опуклих конусів є двогранні кути, що не перевищують π, правильні піраміди, необмежено продовжені, і т.д.

Очевидно, що опуклим конусом є множина розв'язків системи нерівностей $\boldsymbol{Ax} \geq \boldsymbol{0}$. Покажемо, що множина Y точок $\boldsymbol{y} = \boldsymbol{Au}$, де $\boldsymbol{u} \geq \boldsymbol{0}$, є опуклим конусом. Дійсно, нехай $\boldsymbol{y}_1,\ \boldsymbol{y}_2 \in Y$, тобто $\boldsymbol{y}_1 = \boldsymbol{Au}_1,\ \boldsymbol{u}_1 \geq \boldsymbol{0}$ і $\boldsymbol{y}_2 = \boldsymbol{Au}_2,\ \boldsymbol{u}_2 \geq \boldsymbol{0}$. Тоді $\alpha_1\boldsymbol{y}_1 + \alpha_2\boldsymbol{y}_2 = \boldsymbol{Au}$, де $\boldsymbol{u} = \alpha_1\boldsymbol{u}_1 + \alpha_2\boldsymbol{u}_2 \geq \boldsymbol{0}$ при $\alpha_1,\ \alpha_2 \geq 0$. Отже, $\alpha_1\boldsymbol{y}_1 + \alpha_2\boldsymbol{y}_2 \in Y$.

Вправи до глави 3

3.1. Довести теореми про лінійну залежність (незалежність): 1) система векторів, що складається з одного вектора, лінійно залежна тоді і тільки тоді, коли цей вектор є нульовим; 2) якщо серед векторів $\boldsymbol{a}_1,\ \ldots,\ \boldsymbol{a}_n$ деякі утворюють лінійно залежну систему, що і вся система $\boldsymbol{a}_1,\ \ldots,\ \boldsymbol{a}_n$ є лінійно залежною; 3) якщо система векторів лінійно незалежна, то і будь-яка її частина є лінійно незалежною; 4) якщо система векторів $\boldsymbol{a}_1,\ \ldots,\ \boldsymbol{a}_n$ лінійно залежна, а система $\boldsymbol{a}_1,\ \ldots,\ \boldsymbol{a}_{n-1}$ лінійно незалежна, то вектор $\boldsymbol{a}_n$ дорівнює лінійній комбінації векторів $\boldsymbol{a}_1,\ \ldots,\ \boldsymbol{a}_{n-1}$.

3.2. Як показати, що: 1) вимірність перерізу двох підпросторів не перевищує мінімальної з вимірностей цих підпросторів; 2) вимірність суми двох підпросторів не є меншою ніж максимальна з вимірностей цих підпросторів?

3.3. Для дво- і тривимірного просторів навести приклади, коли: 1) вектор $\boldsymbol{a} \in U \oplus V$ розкладається єдиним способом у суму $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in U$, $\boldsymbol{v} \in V$; 2) вектор $\boldsymbol{a}$ розкладається в суму $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$ неєдиним способом, якщо сума підпросторів U, V не є прямою.

3.4. Як виглядають аналоги формул (2.2.2) і (2.2.10) для довільної $m \times n$-матриці $\boldsymbol{A}$?

3.5. Довести, що коли $\boldsymbol{A}$ — матриця порядку n, то $\det \boldsymbol{A} = 0$ тоді і тільки тоді, коли $r(\boldsymbol{A}) < n$.

3.6. Довести, що матриця $\boldsymbol{A}$ порядку n є невиродженою тоді і тільки тоді, коли $r(\boldsymbol{A}) = n$.

3.7. Показати, що будь-яку матрицю рангу r можна представити у вигляді суми r матриць рангу 1 і неможливо представити у вигляді суми меншої кількості таких матриць.

3.8. Одержати скелетний розклад матриці $\boldsymbol{A}$ за допомогою теореми 3.4.3.

3.9. Довести, що дві прямокутні матриці $\boldsymbol{A}, \boldsymbol{B}$ однакових розмірів еквівалентні тоді і тільки тоді, коли $r(\boldsymbol{A}) = r(\boldsymbol{B})$.

3.10. Довести, що матричне рівняння $\boldsymbol{AX} = \boldsymbol{B}$ розв'язуване тоді і тільки тоді, коли $r(\boldsymbol{A}) = r(\boldsymbol{A} \quad \boldsymbol{B})$.

3.11. Нехай дві площини мають спільний спрямовуючий підпростір і принаймні одну спільну точку. Показати, що ці площини збігаються.

3.12. Показати, що коли пряма лінія має дві спільні точки з площиною, то вона вміщується в цій площині.

3.13. За якої умови дві прямі лінії належать одній площині? Вважати, що прямі лінії задані своїми векторними рівняннями.

3.14. Довести, що будь-яка площина є опуклою множиною.

Г Л А В А 4

ЕВКЛІДІВ ПРОСТІР

У попередній главі був визначений векторний простір, в якому можна додавати вектори і множити їх на числа. Тепер ми введемо в цьому просторі метрику, тобто спосіб вимірювати відстані, кути і об'єми. Це буде зроблено шляхом використання поняття скалярного добутку векторів. Ми почнемо з розгляду дійсного векторного простору.

4.1. Векторний простір із скалярним добутком

4.1.1. Визначення евклідового простору. Дійсний векторний простір E називається *евклідовим*, якщо для векторів цього простору виконані такі вимоги.

A. Кожній парі векторів $\boldsymbol{a}, \boldsymbol{b} \in E$ зіставляється дійсне число $(\boldsymbol{a}, \boldsymbol{b})$, яке називається *скалярним добутком* цих векторів.

B. Скалярний добуток задовольняє аксіомам ($\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ — вектори, α — дійсне число):

1) $(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a})$;

2) $(\boldsymbol{a} + \boldsymbol{b}, \ \boldsymbol{c}) = (\boldsymbol{a}, \ \boldsymbol{c}) + (\boldsymbol{b}, \ \boldsymbol{c})$;

3) $(\alpha\boldsymbol{a}, \ \boldsymbol{b}) = \alpha(\boldsymbol{a}, \ \boldsymbol{b})$;

4) $(\boldsymbol{a}, \ \boldsymbol{a}) \geq 0$, $(\boldsymbol{a}, \ \boldsymbol{a}) = 0$ лише коли $\boldsymbol{a} = \boldsymbol{0}$.

Арифметичний простір R_n можна зробити евклідовим, якщо скалярний добуток векторів $\boldsymbol{a} = (\alpha_1 \ \ \alpha_2 \ \ \ldots \ \ \alpha_n)^T$, $\boldsymbol{b} = (\beta_1 \ \ \beta_2 \ \ \ldots \ \ \beta_n)^T$ визначити як величину

$$(\boldsymbol{a}, \ \boldsymbol{b}) = \boldsymbol{a}^T\boldsymbol{b} = \sum_{i=1}^{n} \alpha_i\beta_i \ . \tag{1}$$

Неважко переконатися в тому, що при цьому аксіоми скалярного добутку виконуються. Визначений таким способом евклідів простір позначається через E_n .

4.1.2 Довжина і ортогональність векторів. За допомогою скалярного добутку вводяться поняття, які в подальшому будуть відігравати важливу роль. Мова йде про довжину вектора і ортогональність векторів евклідового простору E .

Довжиною вектора $\boldsymbol{a}$ називається величина

$$|\,\boldsymbol{a}\,| = \sqrt{\boldsymbol{a}^T\boldsymbol{a}}\,. \tag{2}$$

На підставі аксіоми 4) скалярного добутку довжина ненульового вектора додатна, а довжина нульового вектора дорівнює нулю.

Вектори $\boldsymbol{a}$ і $\boldsymbol{b}$ називаються *ортогональними* $(\boldsymbol{a} \perp \boldsymbol{b})$, якщо їх скалярний добуток дорівнює нулю.Система векторів називається *ортогональною*, якщо або вона складається з одного вектора, або її вектори попарно ортогональні. Ортогональна система векторів називається *ортонормованою*, якщо вона складається з векторів довжини 1.

Теорема 4.1.1. *Ортогональна система ненульових векторів є лінійно незалежною.*

Д о в е д е н н я. Нехай вектори $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ попарно ортогональні і є ненульовими. Щоб довести їх лінійну незалежність, треба показати, що рівність $\sum\limits_{i=1}^{m} \alpha_i \boldsymbol{a}_i = \boldsymbol{0}$ можлива лише при виконанні умов $\alpha_1 = \ldots = \alpha_m = 0$. Помноживши справа обидві частини рівності по черзі на вектори $\boldsymbol{a}_1, \ \ldots, \ \boldsymbol{a}_m$, одержимо:

$$\sum_{i=1}^{m} \alpha_i \boldsymbol{a}_i^T \boldsymbol{a}_j = 0, \quad j = \overline{1,\ m}.$$

Оскільки $\boldsymbol{a}_j^T \boldsymbol{a}_j \neq 0$ і $\boldsymbol{a}_i^T \boldsymbol{a}_j = 0$ при $i \neq j$, то $\alpha_j = 0, \ \ j = \overline{1,\ m}$. □

4.1.3. Ортогоналізація Грама — Шмідта. Від довільної системи лінійно незалежних векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ можна перейти до ортогональної системи векторів шляхом так званої *ортогоналізації Грама — Шмідта.*

Теорема 4.1.2. *Нехай вектори* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ *лінійно незалежні. Тоді можна побудувати ортогональну систему ненульових векторів* $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$*, які є лінійними комбінаціями векторів* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$.

Доведення. За перший вектор $\boldsymbol{b}_1$ візьмемо $\boldsymbol{a}_1$ і припустимо у відповідності до методу математичної індукції, що ненульові вектори $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{i-1}$ вже побудовані. Шукаємо вектор $\boldsymbol{b}_i$ у вигляді:

$$\boldsymbol{b}_i = \boldsymbol{a}_i - \sum_{j=1}^{i-1} \beta_{ij} \boldsymbol{b}_j . \tag{3}$$

Виберемо коефіцієнти β_{ij} таким чином, щоб $\boldsymbol{b}_i^T \boldsymbol{b}_k = 0$ при $k = \overline{1,\ i-1}$. Тоді після множення справа обох частин рівності (3) на $\boldsymbol{b}_k$, $k = \overline{1,\ i-1}$, одержимо, що $0 = \boldsymbol{a}_i^T \boldsymbol{b}_k - \beta_{ik} \boldsymbol{b}_k^T \boldsymbol{b}_k$, звідки

$$\beta_{ij} = \frac{\boldsymbol{a}_i^T \boldsymbol{b}_j}{\boldsymbol{b}_j^T \boldsymbol{b}_j} .$$

Вектор $\boldsymbol{b}_i$ є лінійною комбінацією векторів $\boldsymbol{a}_i, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_{i-1}$, а кожний з векторів $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{i-1}$ за припущенням лінійно виражається через $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}$. Тому $\boldsymbol{b}_i$ є лінійною комбінацією векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i$. Очевидно, що $\boldsymbol{b}_i \neq \boldsymbol{0}$, бо якщо $\boldsymbol{b}_i = \boldsymbol{0}$, то це означало б лінійну залежність векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i$.

Отже, ми побудували вектор $\boldsymbol{b}_i$ за допомогою векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i$. Продовживши цей процес доти, поки не будуть вичерпані всі вектори $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$, ми одержимо m попарно ортогональних векторів $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$. □

Зауваження. Лінійна незалежність векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ використовувалася при доведенні теореми лише для того, щоб показати, що вектори $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ є ненульовими. Якщо ортогоналізацію застосовувати до лінійно залежних векторів, то на певному етапі ми одержимо нульовий вектор. Це станеться вперше на i-му кроці, якщо вектори $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}$ лінійно незалежні, а вектор $\boldsymbol{a}_i$ є їх лінійною комбінацією. Тому процес ортогоналізації може бути застосований для перевірки лінійної незалежності (залежності) векторів.

П р и к л а д. Одержимо ортогональну систему векторів $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$, застосувавши ортогоналізацію Грама — Шмідта до системи векторів

$$\boldsymbol{a}_1 = (1 \quad 0 \quad -1)^T, \quad \boldsymbol{a}_2 = (1 \quad 2 \quad 3)^T, \quad \boldsymbol{a}_3 = (3 \quad 4 \quad -1)^T.$$

Будемо вважати, що $\boldsymbol{b}_1 = \boldsymbol{a}_1$. Шукаємо вектор $\boldsymbol{b}_2$ у вигляді:

$$\boldsymbol{b}_2 = \boldsymbol{a}_2 - \beta_{21}\boldsymbol{b}_1.$$

Число β_{21} знайдемо з умови $\boldsymbol{b}_2^T\boldsymbol{b}_1 = 0.$ Оскільки

$$0 = \boldsymbol{b}_2^T\boldsymbol{b}_1 = \boldsymbol{a}_2^T\boldsymbol{b}_1 - \beta_{21}\boldsymbol{b}_1^T\boldsymbol{b}_1,$$

то

$$\beta_{21} = \frac{\boldsymbol{a}_2^T\boldsymbol{b}_1}{\boldsymbol{b}_1^T\boldsymbol{b}_1} = \frac{\boldsymbol{a}_2^T\boldsymbol{a}_1}{\boldsymbol{a}_1^T\boldsymbol{a}_1} = \frac{-2}{2} = -1,$$

$$\boldsymbol{b}_2 = \boldsymbol{a}_2 + \boldsymbol{b}_1 = \boldsymbol{a}_1 + \boldsymbol{a}_2 = (2 \quad 2 \quad 2)^T.$$

Далі шукаємо вектор $\boldsymbol{b}_3$ у вигляді:

$$\boldsymbol{b}_3 = \boldsymbol{a}_3 - \beta_{31}\boldsymbol{b}_1 - \beta_{32}\boldsymbol{b}_2.$$

Числа β_{31} і β_{32} визначимо з умов $\boldsymbol{b}_3^T\boldsymbol{b}_1 = 0,$ $\boldsymbol{b}_3^T\boldsymbol{b}_2 = 0.$ З урахуванням рівності $\boldsymbol{b}_2^T\boldsymbol{b}_1 = 0$ маємо:

$$0 = \boldsymbol{b}_3^T\boldsymbol{b}_1 = \boldsymbol{a}_3^T\boldsymbol{b}_1 - \beta_{31}\boldsymbol{b}_1^T\boldsymbol{b}_1, \quad 0 = \boldsymbol{b}_3^T\boldsymbol{b}_2 = \boldsymbol{a}_3^T\boldsymbol{b}_2 - \beta_{32}\boldsymbol{b}_2^T\boldsymbol{b}_2.$$

Звідси випливає, що

$$\beta_{31} = \frac{\boldsymbol{a}_3^T\boldsymbol{b}_1}{\boldsymbol{b}_1^T\boldsymbol{b}_1} = \frac{\boldsymbol{a}_3^T\boldsymbol{a}_1}{\boldsymbol{a}_1^T\boldsymbol{a}_1} = \frac{4}{2} = 2, \quad \beta_{32} = \frac{\boldsymbol{a}_3^T\boldsymbol{b}_2}{\boldsymbol{b}_2^T\boldsymbol{b}_2} = \frac{12}{12} = 1.$$

Тепер

$$\boldsymbol{b}_3 = \boldsymbol{a}_3 - 2\boldsymbol{b}_1 - \boldsymbol{b}_2 = -3\boldsymbol{a}_1 - \boldsymbol{a}_2 + \boldsymbol{a}_3 = (-1 \quad 2 \quad -1)^T.$$

Таким чином, шукана ортогональна система векторів є такою:

$$\boldsymbol{b}_1 = (1 \quad 0 \quad -1)^T, \quad \boldsymbol{b}_2 = (2 \quad 2 \quad 2)^T, \quad \boldsymbol{b}_3 = (-1 \quad 2 \quad -1)^T. \ \triangle$$

4.2. Ортонормовані базиси і ортогональні матриці

4.2.1. Ортонормований базис. Базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ евклідового простору E_n називається *ортогональним*, якщо вектори цього базису попарно ортогональні. Якщо, крім того, базисні вектори мають одиничну довжину, то базис називається *ортонормованим.* Отже, для ортонормованого базису маємо:

$$\boldsymbol{e}_i^T \boldsymbol{e}_j = \delta_{ij}, \quad i, j = \overline{1,\ n}. \tag{1}$$

З теореми 4.1.1 випливає, що в просторі E_n ортонормована система векторів вміщує не більше ніж n векторів. Найпростішим ортонормованим базисом простору E_n є натуральний базис.

Знайдемо вираз для скалярного добутку довільних векторів $\boldsymbol{a}$ і $\boldsymbol{b}$ через їх координати відносно ортонормованого базису. Нехай

$$\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i, \quad \boldsymbol{b} = \sum_{j=1}^{n} \beta_j \boldsymbol{e}_j.$$

На підставі аксіом скалярного добутку і формули (1) маємо:

$$\boldsymbol{a}^T \boldsymbol{b} = \left(\sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i \right)^T \left(\sum_{j=1}^{n} \beta_j \boldsymbol{e}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \boldsymbol{e}_i^T \boldsymbol{e}_j = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j \delta_{ij} = \sum_{i=1}^{n} \alpha_i \beta_i.$$

Отже, в ортонормованому базисі скалярний добуток двох векторів дорівнює сумі добутків відповідних координат цих векторів.

З'ясуємо, як треба тлумачити координати вектора $\boldsymbol{a}$ відносно ортонормованого базису $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. На підставі попередньої формули маємо:

$$\boldsymbol{a}^T \boldsymbol{e}_j = \alpha_j, \quad j = \overline{1,\ n}\ .$$

Це означає, що координати довільного вектора відносно ортонормованого базису дорівнюють скалярним добуткам цього вектора на відповідні базисні вектори.

З а у в а ж е н н я. У подальшому, якщо вектор простору E_n треба подати у вигляді розкладу за деяким базисом, ми будемо вважати цей базис ортонормованим.Це ж саме ми будемо припускати стосовно векторів унітарного простору, який вводиться в розділі 6.4.

4.2.2. Ортогональні матриці. Знайдемо матрицю перетворення координат при переході від ортонормованого базису $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ до іншого ортонормованого базису $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n$. З формул (3.1.6) і (1) випливає, що

$$(\boldsymbol{e}'_j)^T \boldsymbol{e}'_k = \left(\sum_{i=1}^n q_{ij}\boldsymbol{e}_i\right)^T \left(\sum_{l=1}^n q_{lk}\boldsymbol{e}_l\right) = \sum_{i=1}^n\sum_{l=1}^n q_{ij}q_{lk}\boldsymbol{e}_i^T\boldsymbol{e}_l = \sum_{i=1}^n\sum_{l=1}^n q_{ij}q_{lk}\delta_{il} = \sum_{i=1}^n q_{ij}q_{ik}.$$

Оскільки, з іншого боку, $(\boldsymbol{e}'_j)^T \boldsymbol{e}'_k = \delta_{jk}$, то

$$\sum_{i=1}^n q_{ij}q_{ik} = \delta_{jk}, \quad j, k = \overline{1, n}.$$

Останні рівності на підставі формули (1.1.1) можна подати у вигляді співвідношення для матриці перетворення координат:

$$\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{E}. \tag{2}$$

Матриця $\boldsymbol{Q}$ називається *ортогональною,* якщо $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{E}$. При цьому $\boldsymbol{Q}^T = \boldsymbol{Q}^{-1}$ і тому $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{E}$. Отже, матриця перетворення координат при переході від ортонормованого базису до іншого ортонормованого базису є ортогональною матрицею.

У наступних теоремах розглядаються властивості ортогональних матриць.

Теорема 4.2.1. *Матриця є ортогональною тоді і тільки тоді, коли її стовпці* (*рядки*) *утворюють ортонормовану систему векторів.*

Д о в е д е н н я. Ортонормованість стовпців є наслідком рівності (2), а ортонормованість рядків випливає з рівності $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{E}$.

Навпаки, з ортонормованості стовпців (рядків) випливають указані рівності. □

Теорема 4.2.2. *Ортогональна матриця має такі властивості*:

1) *добуток двох ортогональних матриць є ортогональною матрицею*;

2) *модуль визначника ортогональної матриці дорівнює одиниці.*

Д о в е д е н н я. 1) Нехай $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2$ — ортогональні матриці. Тоді для матриці $\boldsymbol{Q} = \boldsymbol{Q}_1\boldsymbol{Q}_2$ виконується рівність (2):

$$\boldsymbol{Q}^T\boldsymbol{Q} = (\boldsymbol{Q}_1\boldsymbol{Q}_2)^T\boldsymbol{Q}_1\boldsymbol{Q}_2 = \boldsymbol{Q}_2^T\boldsymbol{Q}_1^T\boldsymbol{Q}_1\boldsymbol{Q}_2 = \boldsymbol{Q}_2^T\boldsymbol{Q}_2 = \boldsymbol{E}.$$

2) На підставі теореми 3.4.5 з рівності (2) випливає, що $(\det\boldsymbol{Q}^T)\cdot\det\boldsymbol{Q} = 1$, тобто $|\det\boldsymbol{Q}| = 1$. ☐

4.3. Лінійні перетворення

4.3.1. Визначення лінійного перетворення. Правило, за яким кожному елементу x деякої непустої множини X зіставляється єдиний елемент y непустої множини Y, називається *відображенням* (*оператором, перетворенням*). Результат y застосування відображення $\tilde{A}$ до елемента x позначається символом

$$y = \tilde{A}x. \tag{1}$$

Говорять, що відображення $\tilde{A}$ діє з X у Y або відображає X в Y. Елемент y з (1) називається образом елемента x, а x — прообразом елемента y.

Нижче ми будемо розглядати так звані лінійні відображення, які ми будемо називати *лінійними перетвореннями*. У цьому випадку множини X і Y вважаються векторними просторами (підпросторами). Лінійні перетворення являють собою багатовимірне узагальнення лінійної функції одного числового аргумента $y = ax$. Їх різноманітність швидко зростає з підвищенням вимірності простору.

Властивість лінійності перетворення $\tilde{A}$ полягає в тому, що

$$\tilde{A}(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) = \alpha\tilde{A}\boldsymbol{x}_1 + \beta\tilde{A}\boldsymbol{x}_2$$

для будь-яких векторів $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ векторного простору і будь яких чисел α, β, які в загальному випадку є комплексними. Нижче ми будемо мати справу з лінійними перетвореннями, які діють з векторного простору (підпростору) R в R. Про таке лінійне перетворення говорять, що воно діє в просторі (підпросторі) R.

Найпростішими прикладами лінійних перетворень є нульове перетворення

$$\tilde{0}\boldsymbol{x} = \boldsymbol{0},$$

яке зіставляє кожному вектору $\boldsymbol{x} \in R$ нульовий вектор, і одиничне перетворення

$$\tilde{E}\boldsymbol{x} = \boldsymbol{x},$$

яке зіставляє будь-якому вектору $\boldsymbol{x} \in R$ той же вектор.

На множині всіх лінійних перетворень, діючих у R, введемо операції додавання лінійних перетворень, множення лінійного перетворення на число і множення лінійних перетворень.

Сумою лінійних перетворень $\tilde{A}$ і $\tilde{B}$ називається лінійне перетворення, яке зіставляє вектору $\boldsymbol{x}$ вектор $\tilde{A}\boldsymbol{x} + \tilde{B}\boldsymbol{x}$. Добутком лінійного перетворення $\tilde{A}$ на число α називається лінійне перетворення, яке зіставляє вектору $\boldsymbol{x}$ вектор $\alpha\tilde{A}\boldsymbol{x}$. Добутком лінійних перетворень $\tilde{A}$ і $\tilde{B}$ називається лінійне перетворення, яке полягає в послідовному застосуванні лінійного перетворення $\tilde{B}$, а потім — лінійного перетворення $\tilde{A}$.

4.3.2. Представлення лінійного перетворення матрицею. Виберемо в просторі R_n деякий базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ і позначимо через $x_1, \ldots, x_n$ і $y_1, \ldots, y_n$ координати векторів $\boldsymbol{x}$ і $\boldsymbol{y}$ у цьому базисі. Застосуємо лінійне перетворення до базисного вектора $\boldsymbol{e}_j$ і позначимо через $a_{1j}, \ldots, a_{nj}$ координати вектора $\tilde{A}\boldsymbol{e}_j$:

$$\tilde{A}\boldsymbol{e}_j = \sum_{i=1}^{n} a_{ij}\boldsymbol{e}_i, \quad j = \overline{1,\ n}. \tag{2}$$

З формул (1) і (2) маємо:

$$\boldsymbol{y} = \tilde{A}\boldsymbol{x} = \tilde{A}\sum_{j=1}^{n} x_j\boldsymbol{e}_j = \sum_{j=1}^{n} x_j\tilde{A}\boldsymbol{e}_j = \sum_{j=1}^{n} x_j\sum_{i=1}^{n} a_{ij}\boldsymbol{e}_i = \sum_{i=1}^{n}\boldsymbol{e}_i\sum_{j=1}^{n} a_{ij}x_j.$$

З ішого боку,

$$\boldsymbol{y} = \sum_{i=1}^{n} y_i\boldsymbol{e}_i.$$

З двох останніх формул знаходимо, що

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = \overline{1,\ n}. \tag{3}$$

Матриця $\boldsymbol{A}$ з елементами $a_{ij},\ i,\ j = \overline{1,\ n}$, називається *матрицею лінійного перетворення* в заданому базисі $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Отже, лінійне перетворення визначається своєю матрицею і співвідношення (3) можна записати в матричній формі:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \,. \tag{4}$$

Формула (2) показує, що j-й стовпець матриці $\boldsymbol{A}$ складається з координат вектора $\tilde{A}\boldsymbol{e}_j$, тобто є образом базисного вектора $\boldsymbol{e}_j$. Нагадаємо, що в розділі 3.1 для невиродженої матриці $\boldsymbol{A}$ ми тлумачили координати вектора $\boldsymbol{Ax}$ як нові координати вектора $\boldsymbol{x}$ при зміні базису.

Довільна матриця $\boldsymbol{A}$ може бути пов'язана з деяким лінійним перетворенням. Про це свідчить співвідношення (4), яке є лінійним відносно $\boldsymbol{x}$ і $\boldsymbol{y}$ при будь-якій матриці $\boldsymbol{A}$.

Таким чином, виявлена взаємно однозначна відповідність між лінійними перетвореннями і їх матрицями. Вона дозволяє в разі фіксованого базису простору ототожнювати лінійне перетворення і його матрицю подібно до того, як вектор ототожнюється з матрицею-стовпцем його координат. При такому ототожнюванні результат дії лінійного перетворення на вектор збігається з результатом множення матриці на матрицю-стовпець його координат, що підтверджується формулами (1) і (4). Можна довести, що операції над лінійними перетвореннями зводяться до відповідних операцій над їх матрицями.

У подальшому ми будемо позначати лінійне перетворення і його матрицю однією і тією ж літерою. Інколи лінійне перетворення ми будемо називати просто перетворенням.

4.3.3. Матриці лінійного перетворення в різних базисах. Розглянемо одночасно базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ і інший базис $\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n$. Аналогічно рівності (4) запишемо для другого базису:

$$\boldsymbol{y}' = \boldsymbol{A}'\boldsymbol{x}' , \tag{5}$$

де $\boldsymbol{x}'$ і $\boldsymbol{y}'$ — матриці-стовпці, складені з координат векторів $\boldsymbol{x}$ і $\boldsymbol{y}$ у базисі $\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n$, а $\boldsymbol{A}'$ — матриця лінійного перетворення в цьому базисі. На підставі формули (3.1.9) маємо:

$$\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{x}', \quad \boldsymbol{y} = \boldsymbol{Q}\boldsymbol{y}'. \tag{6}$$

З рівностей (6) і (4) знаходимо:

$$\boldsymbol{y}' = \boldsymbol{Q}^{-1}\boldsymbol{y} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}\boldsymbol{x}',$$

що разом з рівністю (5) дає:

$$\boldsymbol{A}' = \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}. \tag{7}$$

Матриці $\boldsymbol{A}$ і $\boldsymbol{A}'$ називаються *подібними* і перехід від $\boldsymbol{A}$ до $\boldsymbol{A}'$ називається *перетворенням подібності*, якщо $\boldsymbol{A}$ і $\boldsymbol{A}'$ пов'язані співвідношенням (7), де $\boldsymbol{Q}$ — невироджена матриця. Отже, дві матриці, відповідні до одного і того ж лінійного перетворення в різних базисах, подібні між собою, причому матриця $\boldsymbol{Q}$ є матрицею перетворення координат при переході від першого базису до другого. Можна стверджувати, що лінійному перетворенню відповідає множина подібних між собою матриць, які представляють лінійне перетворення в різних базисах.

Дві подібні матриці мають рівні визначники. Дійсно, на підставі теореми 3.4.5 з рівності (7) випливає, що

$$\det \boldsymbol{A}' = \det(\boldsymbol{Q}^{-1}) \cdot \det \boldsymbol{A} \cdot \det \boldsymbol{Q} = \det \boldsymbol{A}.$$

Подібні матриці мають також однакові ранги. Це є наслідком висновку з теореми 3.4.4 і формули (7).

Рівність визначників і рангів є необхідною, але недостатньою умовою подібності матриць $\boldsymbol{A}$ і $\boldsymbol{A}'$. Про необхідну і достатню умову подібності двох матриць мова буде йти в розділі 9.1.

4.3.4. Образ і ядро лінійного перетворення. *Образом* лінійного перетворення (матриці) $\boldsymbol{A}$ називається множина $\mathscr{R}(\boldsymbol{A})$ векторів $\boldsymbol{x}$ вигляду $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w}$, де $\boldsymbol{w}$ — деякий вектор. *Ядром* лінійного перетворення (матриці) $\boldsymbol{A}$ називається множина $\mathscr{N}(\boldsymbol{A})$ векторів $\boldsymbol{x}$, для яких $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$. Ядро $\mathscr{N}(\boldsymbol{A})$ завжди має принаймні один вектор — нульовий. Неважко пересвідчитися в тому, що $\mathscr{R}(\boldsymbol{A})$ і $\mathscr{N}(\boldsymbol{A})$ є підпросторами.

З визначення добутку матриць випливає, що $\mathscr{R}(\boldsymbol{A})$ збігається з підпростором, породженим стовпцями матриці $\boldsymbol{A}$. На підставі висновку 1 з теореми про базисний мінор можна стверджувати, що

$$\dim \mathscr{R}(\boldsymbol{A}) = r(\boldsymbol{A}). \tag{8}$$

Вимірність підпростору $\mathscr{R}(\boldsymbol{A})$ називають *рангом* лінійного перетворення $\boldsymbol{A}$. Рівність (8) показує, що ранг лінійного перетворення дорівнює рангу його матриці в деякому базисі простору R_n. Оскільки $r(\boldsymbol{A}') = r(\boldsymbol{A})$, то ранг лінійного перетворення не залежить від базису, в якому воно розглядається.

Вимірність підпростору $\mathscr{N}(\boldsymbol{A})$ можна визначити на підставі теореми 3.5.1:

$$\dim \mathscr{N}(\boldsymbol{A}) = n - r(\boldsymbol{A}). \tag{9}$$

4.3.5. Обернене лінійне перетворення. Лінійне перетворення, що діє у векторному просторі, називається *невиродженим*, якщо його ядро вміщує лише нульовий вектор. Лінійне перетворення, яке не є невиродженим, називається *виродженим*.

Невироджені лінійні перетворення мають певні особливості, на яких ми зупинимося. Для таких перетворень $\dim \mathscr{N}(\boldsymbol{A}) = 0$, тому з рівності (9) випливає, що ранг невиродженого лінійного перетворення збігається з вимірністю простору. При цьому, як показує формула (8), $\dim \mathscr{R}(\boldsymbol{A}) = n$. Це означає, що коли невироджене лінійне перетворення діє в просторі R_n, то кожний вектор з R_n є образом деякого вектора з R_n. Ця властивість невиродженого лінійного перетворення еквівалентна його визначенню.

Особливістю невиродженого лінійного перетворення є єдиність прообразу для будь-якого вектора простору. Припустимо, що для деякого вектора $\boldsymbol{y}$ існує два прообрази $\boldsymbol{u}$, $\boldsymbol{u}'$, тобто $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{u}$, $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{u}'$. Тоді $\boldsymbol{A}(\boldsymbol{u} - \boldsymbol{u}') = \boldsymbol{0}$. Оскільки ядро $\mathscr{N}(\boldsymbol{A})$ складається тільки з нульового вектора, то $\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{0}$, тобто $\boldsymbol{u} = \boldsymbol{u}'$. Ця властивість також еквівалентна визначенню невиродженого лінійного перетворення.

Оскільки відображення простору R_n в R_n, яке виконується невиродженим лінійним перетворенням $\boldsymbol{A}$, є взаємно однозначним, то для $\boldsymbol{A}$ існує *обернене* лінійне перетворення, яке зіставляє кожному вектору $\boldsymbol{y} \in R_n$ єдиний вектор $\boldsymbol{x} \in R_n$ такий, що задовольняє співвідношенню (4). Позначимо обернене лінійне перетворення через $\boldsymbol{A}^{-1}$. Якщо має місце рівність (4), то

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y}.$$

З визначення оберненого лінійного перетворення випливають співвідношення (1.3.1). Неважко довести лінійність і невиродженість оберненого лінійного перетворення.

Іноді з лінійним перетворенням, діючим у просторі R_n, можна пов'язати деяке невироджене лінійне перетворення навіть у тому випадку, коли $\boldsymbol{A}$ — вироджене. Розглянемо один приклад.

Припустимо, що $\dim \mathscr{N}(\boldsymbol{A}) \neq 0$ і $\dim(\mathscr{R}(\boldsymbol{A}) \bigcap \mathscr{N}(\boldsymbol{A})) = 0$. Звичайно, остання рівність має місце далеко не завжди, але при її виконанні з формули Грассмана і рівностей (8), (9) випливає, що

$$\dim(\mathscr{R}(\boldsymbol{A}) + \mathscr{N}(\boldsymbol{A})) = n,$$

тобто

$$R_n = \mathscr{R}(\boldsymbol{A}) \oplus \mathscr{N}(\boldsymbol{A}).$$

Нехай $\boldsymbol{y} \in \mathscr{R}(\boldsymbol{A})$, тобто $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$. На підставі теореми 3.2.3 довільний вектор $\boldsymbol{x} \in R_n$ можна представити у вигляді суми $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in \mathscr{R}(\boldsymbol{A})$, $\boldsymbol{v} \in \mathscr{N}(\boldsymbol{A})$. Тоді

$$\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}\boldsymbol{u},$$

тобто будь-який вектор $\boldsymbol{y} \in \mathscr{R}(\boldsymbol{A})$ має принаймні один прообраз $\boldsymbol{u} \in \mathscr{R}(\boldsymbol{A})$.

Насправді цей прообраз є єдиним. Припустимо, що для деякого вектора $\boldsymbol{y} \in \mathscr{R}(\boldsymbol{A})$ існують два прообрази $\boldsymbol{u}$, $\boldsymbol{u}' \in \mathscr{R}(\boldsymbol{A})$, тобто $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{u}$, $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{u}'$. Звідси випливає, що $\boldsymbol{A}(\boldsymbol{u} - \boldsymbol{u}') = \boldsymbol{0}$, тому $\boldsymbol{u} - \boldsymbol{u}' \in \mathscr{N}(\boldsymbol{A})$. З іншого боку, оскільки $\mathscr{R}(\boldsymbol{A})$ — підпростір, то $\boldsymbol{u} - \boldsymbol{u}' \in \mathscr{R}(\boldsymbol{A})$. Єдиним вектором, спільним для $\mathscr{R}(\boldsymbol{A})$ і $\mathscr{N}(\boldsymbol{A})$, є нульовий вектор. Тому $\boldsymbol{u} - \boldsymbol{u}' = \boldsymbol{0}$, тобто $\boldsymbol{u} = \boldsymbol{u}'$.

Отже, має місце взаємно однозначна відповідність між образами і прообразами при дії лінійного перетворення $\boldsymbol{A}$ в підпросторі

$\mathcal{R}(\boldsymbol{A})$. Це перетворення буде невиродженим, бо воно переводить у нульовий вектор лише нульовий вектор з $\mathcal{R}(\boldsymbol{A})$.

Таким чином, у розглянутому прикладі лінійне перетворення, яке діє в просторі R_n як вироджене, виявляється невиродженим, якщо вважати його діючим у підпросторі $\mathcal{R}(\boldsymbol{A})$. З подібною ситуацією ми будемо мати справу в розділі 12.3.

4.4. Ортогональні перетворення

4.4.1. Визначення ортогонального перетворення. Лінійне перетворення $\boldsymbol{y} = \boldsymbol{Qx}$, яке діє в просторі E_n, називається *ортогональним*, якщо матриця $\boldsymbol{Q}$ є ортогональною. Найважливіша особливість ортогонального перетворення полягає в тому, що для будь-яких двох векторів їх скалярний добуток дорівнює скалярному добутку їх образів при цьому перетворенні. Дійсно,

$$(\boldsymbol{Qu})^T \boldsymbol{Qv} = \boldsymbol{u}^T \boldsymbol{Q}^T \boldsymbol{Qv} = \boldsymbol{u}^T \boldsymbol{v} .$$

Звідси випливає, що при ортогональному перетворенні довжини векторів не змінюються і ортогональні вектори залишаються ортогональними.

Ортогональні перетворення відіграють важливу роль при розв'язанні багатьох задач лінійної алгебри. Далі ми розглянемо ортогональні перетворення обертання і відбиття.

4.4.2. Перетворення обертання. Нехай на площині $O\,x_1x_2$ кожний радіус-вектор $\boldsymbol{x}$ замінюється радіус-вектором $\boldsymbol{y}$, одержаним обертанням $\boldsymbol{x}$ на кут φ проти годинникової стрілки (рис. 7). Визначимо матрицю перетворення в натуральному базисі, який складається з ортогональних векторів $\boldsymbol{e}_1 = (1 \quad 0)^T$ і $\boldsymbol{e}_2 = (0 \quad 1)^T$. У результаті вказаного перетво-

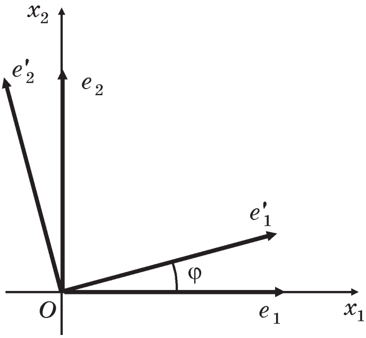

Рис. 7

рення базисний вектор $\boldsymbol{e}_1$ замінюється вектором $\boldsymbol{e}_1' = (\cos\varphi \quad \sin\varphi)^T$, а базисний вектор $\boldsymbol{e}_2$ — вектором $\boldsymbol{e}_2' = (-\sin\varphi \quad \cos\varphi)^T$. На підставі результатів, одержаних у розділі 4.3, матриця перетворення виглядає так:

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

Розглянуте перетворення є *перетворенням обертання* в просторі E_2. Його узагальнення на випадок простору E_n виконується за допомогою *матриці обертання*

$$\boldsymbol{T}_{ij}(\varphi) = \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & \iddots & \\ & & \cos\varphi & \dots & -\sin\varphi & \\ & & \vdots & & \vdots & \\ & & \sin\varphi & \dots & \cos\varphi & \\ & \iddots & & & \ddots & \\ 0 & & & & & 1 \end{pmatrix}.$$

Матриця $\boldsymbol{T}_{ij}(\varphi)$ відрізняється від одиничної матриці порядку n лише чотирма елементами, розміщеними на перетині рядків і стовпців з номерами i та j $(i < j)$, причому два з цих елементів, які є діагональними, дорівнюють $\cos\varphi$, а два інших дорівнюють $\sin\varphi$ і $-\sin\varphi$. Зокрема,

$$\boldsymbol{T}_{12}(\varphi) = \begin{pmatrix} \boldsymbol{T} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{E}_{n-2} \end{pmatrix}.$$

Очевидно, що матриці обертання є ортогональними.

Покажемо на прикладі, як вони використовуються. Нехай треба так перетворити матрицю $\boldsymbol{A}$ порядку n шляхом множення її зліва на $\boldsymbol{T}_{ij}(\varphi)$, щоб другий елемент першого стовпця став нульовим, а перший елемент того ж стовпця був додатним за умови, що елементи a_{11} і a_{21} не є нульовими одночасно. Очевидно, що ми повинні використати матрицю $\boldsymbol{T}_{12}(\varphi)$. Сформульовані вище вимоги до елементів перетвореної матриці $\boldsymbol{T}_{12}(\varphi)\boldsymbol{A}$ можна записати так:

$$a_{11}\cos\varphi - a_{21}\sin\varphi > 0, \quad a_{11}\sin\varphi + a_{21}\cos\varphi = 0.$$

Неважко переконатися в тому, що ця система обмежень відносно φ має розв'язок при вказаній вище умові для a_{11} і a_{21}.

Треба зауважити, що в добутку $\boldsymbol{T}_{ij}(\varphi)\boldsymbol{A}$ змінені в порівнянні з $\boldsymbol{A}$ тільки елементи *i*-го та *j*-го рядків. Якщо множити матрицю $\boldsymbol{A}$ справа на $\boldsymbol{T}_{ij}(\varphi)$, то змінюються лише елементи *i*-го та *j*-го стовпців матриці $\boldsymbol{A}$.

4.4.3. Перетворення відбиття. *Перетворенню відбиття* відповідає *матриця відбиття* (*матриця Хаусхолдера*)

$$\boldsymbol{H} = \boldsymbol{E} - 2\boldsymbol{u}\boldsymbol{u}^T,$$

де $\boldsymbol{u}$ — вектор одиничної довжини. Очевидно, що матриця $\boldsymbol{H}$ симетрична. Покажемо, що вона є ортогональною. Дійсно, оскільки $\boldsymbol{u}^T\boldsymbol{u} = 1$, то

$$\boldsymbol{H}^T\boldsymbol{H} = (\boldsymbol{E} - 2\boldsymbol{u}\boldsymbol{u}^T)^2 = \boldsymbol{E} - 4\boldsymbol{u}\boldsymbol{u}^T + 4\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{u}\boldsymbol{u}^T = \boldsymbol{E}.$$

З'ясуємо геометричний зміст перетворення відбиття у тривимірному випадку. Здійснимо відбиття відносно деякої площини, що проходить через початок координат. Перетворення повністю визначається вектором $\boldsymbol{u}$, який має одиничну довжину і є ортогональним до будь-якого вектора, що лежить у площині. Візьмемо довільний вектор $\boldsymbol{x}$ і розкладемо його на дві складові: паралельну до вектора $\boldsymbol{u}$, тобто $\boldsymbol{x}_u = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x})$, і перпендикулярну до нього. При відбитті вектора відносно площини його перпендикулярна складова залишається незмінною, а паралельна змінює знак (рис. 8), тому відбитий вектор $\boldsymbol{z}$ відрізняється від вихідного на подвоєну величину паралельної складової, тобто

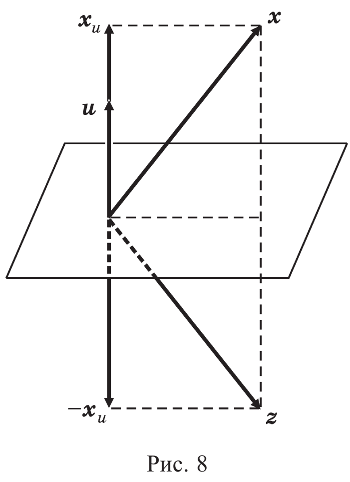

Рис. 8

$$\boldsymbol{z} = \boldsymbol{x} - 2\boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{x}) = (\boldsymbol{E} - 2\boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x} = \boldsymbol{H}\boldsymbol{x}.$$

Отже, ми маємо справу з перетворенням відбиття.

Геометричні міркування вказують на можливість вибрати вектор $\boldsymbol{u}$ так, щоб будь-який заданий вектор перетворився у вектор, паралельний до базисного. Саме про це свідчить

Теорема 4.4.1. *Припустимо, що*

$$\boldsymbol{y} = (1 \quad 0 \quad \ldots \quad 0)^T, \qquad \boldsymbol{u} = \frac{\boldsymbol{x} + \alpha\boldsymbol{y}}{|\,\boldsymbol{x} + \alpha\boldsymbol{y}\,|}, \qquad \alpha = \pm\,|\,\boldsymbol{x}\,|.$$

Тоді $\boldsymbol{Hx} = -\alpha\boldsymbol{y}$.

Д о в е д е н н я. Очевидно, що

$$\boldsymbol{Hx} = \boldsymbol{x} - 2\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{x} = \boldsymbol{x} - (\boldsymbol{x} + \alpha\boldsymbol{y})\beta,$$

де

$$\beta = \frac{2(\boldsymbol{x} + \alpha\boldsymbol{y})^T\boldsymbol{x}}{(\boldsymbol{x} + \alpha\boldsymbol{y})^T(\boldsymbol{x} + \alpha\boldsymbol{y})}.$$

Виконавши множення в чисельнику і знаменнику останньої формули з урахуванням рівностей $\boldsymbol{x}^T\boldsymbol{x} = \alpha^2$, $\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{x}$, одержимо, що $\beta = 1$. Тому $\boldsymbol{Hx} = -\alpha\boldsymbol{y}$. ☐

Для порівняння перетворень обертання і відбиття розглянемо задачу про зведення деякої квадратної матриці $\boldsymbol{A}$ до такого вигляду, коли перший її стовпець набуває вигляду першого стовпця верхньої трикутної матриці. Очевидно, що для цього треба послідовно помножити зліва матрицю $\boldsymbol{A}$ на матриці обертання $\boldsymbol{T}_{ij}(\varphi)$. При цьому кількість таких множень дорівнює кількості ненульових елементів першого стовпця, розміщених нижче елемента a_{11}.

З іншого боку, помножимо зліва матрицю $\boldsymbol{A}$ на матрицю Хаусхолдера $\boldsymbol{H}$, вважаючи, що в теоремі 4.4.1 $\boldsymbol{x} = \boldsymbol{a}_{*1}$, де $\boldsymbol{a}_{*1}$ — перший стовпець матриці $\boldsymbol{A}$. Тоді перший стовпець добутку $\boldsymbol{HA}$ буде дорівнювати $\boldsymbol{Ha}_{*1} = -\alpha\boldsymbol{y}$, де $\alpha = \pm\,|\,\boldsymbol{a}_{*1}\,|$. Отже, ми одержали перший стовпець верхньої трикутної матриці. Таким чином, перетворення, яке виконується кількома множеннями на матриці обертання, може бути одержане шляхом одного множення на матрицю Хаусхолдера.

Проте ця обставина зовсім не свідчить про безумовну перевагу перетворень відбиття в порівнянні з перетвореннями обертання. При належному виборі послідовності перетворень обертання і за деяких додаткових обставин загальна оцінка похибки обчислень може бути меншою ніж у перетворення відбиття, яке розв’язує ту ж саму задачу.

Розглянемо деякі із застосувань матриць відбиття до задач, які є «підготовчими кроками» при проведенні матричних обчислень.

4.4.4. *QR*-розклад. Перш за все покажемо, як отримати ***QR**-розклад* квадратної матриці у вигляді добутку ортогональної матриці $\boldsymbol{Q}$ і верхньої трикутної матриці $\boldsymbol{R}$. Існують різноманітні способи одержання $\boldsymbol{QR}$-розкладу. Проте найбільш досконалим є той, що використовується при доведенні наступної теореми.

Теорема 4.4.2. *Будь-яку матрицю **A** порядку n можна представити у вигляді*:

$$\boldsymbol{A} = \boldsymbol{QR}\,, \tag{1}$$

де $\boldsymbol{Q}$ — ортогональна і $\boldsymbol{R}$ — верхня трикутна матриці.

Д о в е д е н н я. Нехай $\boldsymbol{H}_1$ — матриця Хаусхолдера порядку n. Будемо вважати, що вона побудована за допомогою вектора $\boldsymbol{x}$, рівного першому стовпцю матриці $\boldsymbol{A}$. Як зазначено вище,

$$\boldsymbol{F}_1\boldsymbol{A} = \begin{pmatrix} -\alpha & \boldsymbol{J} \\ \boldsymbol{0} & \boldsymbol{L} \end{pmatrix}, \tag{2}$$

де $\boldsymbol{F}_1 = \boldsymbol{H}_1$, а $\boldsymbol{J}$ і $\boldsymbol{L}$ — деякі підматриці.

Помноживши зліва матрицю $\boldsymbol{F}_1\boldsymbol{A}$ на матрицю

$$\boldsymbol{F}_2 = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_2 \end{pmatrix},$$

де $\boldsymbol{H}_2$ — деяка матриця Хаусхолдера порядку $n-1$, будемо мати:

$$\boldsymbol{F}_2\boldsymbol{F}_1\boldsymbol{A} = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_2 \end{pmatrix}\begin{pmatrix} -\alpha & \boldsymbol{J} \\ \boldsymbol{0} & \boldsymbol{L} \end{pmatrix} = \begin{pmatrix} -\alpha & \boldsymbol{J} \\ \boldsymbol{0} & \boldsymbol{H}_2\boldsymbol{L} \end{pmatrix}.$$

Якщо для матриці $\boldsymbol{H}_2$ взяти за вектор $\boldsymbol{x}$ перший стовпець матриці $\boldsymbol{L}$, то в матриці $\boldsymbol{H}_2\boldsymbol{L}$ перший стовпець буде мати вигляд першого стовпця верхньої трикутної матриці порядку $n-1$. Тому в матриці $\boldsymbol{F}_2\boldsymbol{F}_1\boldsymbol{A}$ два перші стовпці є першими стовпцями верхньої трикутної матриці.

Продовжуючи ці перетворення, наприкінці будемо мати:

$$\boldsymbol{F}_{n-1}\ldots\boldsymbol{F}_2\boldsymbol{F}_1\boldsymbol{A}=\boldsymbol{R},\quad \boldsymbol{F}_i=\begin{pmatrix}\boldsymbol{E}_{i-1} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{H}_i\end{pmatrix},\quad i=\overline{2,\,n-1},\tag{3}$$

де $\boldsymbol{H}_i$ — матриця Хаусхолдера порядку $n+1-i$ з відповідним вибором вектора $\boldsymbol{x}$. Якщо на k-му кроці перетворення матриці $\boldsymbol{A}$ виявиться, що стовпець уже має необхідний вигляд, то вважаємо, що $\boldsymbol{F}_k=\boldsymbol{E}$.

Неважко переконатися в тому, що матриці $\boldsymbol{F}_1,\ldots,\boldsymbol{F}_{n-1}$ є ортогональними. На підставі частини 1) теореми 4.2.2 їх добуток є ортогональною матрицею і тому маємо рівність (1). □

Теорема може бути поширена на прямокутні матриці. Зокрема, якщо $m\times n$-матриця $\boldsymbol{A}$ при $m>n$ має ранг, рівний n, то, розмірковуючи як при доведенні рівності (1), одержимо:

$$\boldsymbol{A}=\boldsymbol{Q}\begin{pmatrix}\boldsymbol{R}\\ \mathbf{0}\end{pmatrix},\tag{4}$$

де порядки $\boldsymbol{Q}$ і $\boldsymbol{R}$ дорівнюють m і n відповідно, а нульова підматриця має розміри $(m-n)\times n$.

П р и к л а д 1. Знайдемо за допомогою перетворень відбиття $\boldsymbol{QR}$-розклад матриці

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

Для одержання нулів у першому стовпці матриці $\boldsymbol{A}$ нижче його першого елемента помножимо зліва матрицю $\boldsymbol{A}$ на матрицю $\boldsymbol{F}_1$, яка є матрицею Хаусхолдера $\boldsymbol{H}_1$, тобто

$$\boldsymbol{F}_1=\boldsymbol{H}_1=\boldsymbol{E}-\frac{2\boldsymbol{v}_1\boldsymbol{v}_1^T}{|\boldsymbol{v}_1|^2}.$$

Тут $\boldsymbol{v}_1=\boldsymbol{x}_1+\alpha_1\boldsymbol{y}_1$, $\boldsymbol{y}_1=(1\quad 0\quad 0)^T$, $\alpha_1=\pm|\boldsymbol{x}_1|$, а $\boldsymbol{x}_1$ є першим стовпцем матриці $\boldsymbol{A}$. Очевидно, що $\boldsymbol{x}_1=(0\quad -1\quad 1)^T$, $\alpha_1=\pm\sqrt{2}$,

$\boldsymbol{v}_1 = (\pm\sqrt{2} \quad -1 \quad 1)^T$. Звичайно знак величини α_1 вибирають так, щоб було більшим значення величини $|\boldsymbol{v}_1|$ (при цьому досягається більша точність наближених обчислень). Проте в нашому випадку величина $|\boldsymbol{v}_1|$ не залежить від знаку α_1, тому можна взяти будь-який знак. Нехай, наприклад, $\boldsymbol{v}_1 = (\sqrt{2} \quad -1 \quad 1)^T$. При цьому $|\boldsymbol{v}_1|^2 = 4$. Далі маємо:

$$\boldsymbol{v}_1\boldsymbol{v}_1^T = \begin{pmatrix} 2 & -\sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 1 & -1 \\ \sqrt{2} & -1 & 1 \end{pmatrix}, \quad \boldsymbol{F}_1 = \boldsymbol{H}_1 = \frac{1}{2}\begin{pmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \\ -\sqrt{2} & 1 & 1 \end{pmatrix},$$

$$\boldsymbol{F}_1\boldsymbol{A} = \frac{1}{2}\begin{pmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \\ -\sqrt{2} & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & -\sqrt{2} & \sqrt{2} \\ -1 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -\sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 1 \\ 0 & 2 & -1 \end{pmatrix}.$$

Щоб перетворити матрицю $\boldsymbol{F}_1\boldsymbol{A}$ у верхню трикутну матрицю $\boldsymbol{R}$, треба помножити її зліва на матрицю

$$\boldsymbol{F}_2 = \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_2 \end{pmatrix},$$

де $\boldsymbol{H}_2$ — матриця Хаусхолдера, побудована для вектора $\boldsymbol{x}_2$, координатами якого є елементи другого стовпця матриці $\boldsymbol{F}_1\boldsymbol{A}$ за винятком першого. Отже, $\boldsymbol{x}_2 = (0 \quad 2)^T$, $\alpha_2 = \pm 2$, $\boldsymbol{v}_2 = (\pm 2 \quad 2)^T$. Тут знову знак α_2 не випливає на величину $|\boldsymbol{v}_2|$, тому можна вважати, наприклад, що $\boldsymbol{v}_2 = (2 \quad 2)^T$. При цьому $|\boldsymbol{v}_2|^2 = 8$. Тепер маємо:

$$\boldsymbol{v}_2\boldsymbol{v}_2^T = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}, \quad \boldsymbol{H}_2 = \boldsymbol{E} - \frac{2\boldsymbol{v}_2\boldsymbol{v}_2^T}{|\boldsymbol{v}_2|^2} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

$$\boldsymbol{F}_2\boldsymbol{F}_1\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}\begin{pmatrix} -\sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 1 \\ 0 & 2 & -1 \end{pmatrix} = \begin{pmatrix} -\sqrt{2} & \sqrt{2} & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix} = \boldsymbol{R}.$$

Позначивши через $\boldsymbol{Q}^T$ матрицю $\boldsymbol{F}_2\boldsymbol{F}_1$, одержимо:

$$\boldsymbol{Q}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & 1 & 1 \\ -\sqrt{2} & 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & -1 & -1 \\ -\sqrt{2} & -1 & -1 \end{pmatrix}.$$

Отже, $\boldsymbol{Q}^T\boldsymbol{A} = \boldsymbol{R}$, звідки знаходимо $\boldsymbol{QR}$-розклад матриці $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{QR} = \frac{1}{2} \begin{pmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & -1 & -1 \\ -\sqrt{2} & -1 & -1 \end{pmatrix} \begin{pmatrix} -\sqrt{2} & \sqrt{2} & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{pmatrix}. \ \triangle$$

4.4.5. Дводіагональна форма матриці. *Дводіагональною* називається така верхня трикутна матриця, яка має нулі всюди, крім, можливо, двох діагоналей: головної і прилеглої до неї.

Теорема 4.4.3. *Будь-яку матрицю* $\boldsymbol{A}$ *порядку* n *можна подати у вигляді*:

$$\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{R}\boldsymbol{Q}_2\,, \tag{5}$$

де $\boldsymbol{Q}_1$ *і* $\boldsymbol{Q}_2$ *— ортогональні матриці, а* $\boldsymbol{R}$ *— дводіагональна матриця.*

Д о в е д е н н я. Помножимо зліва матрицю $\boldsymbol{A}$ на матрицю $\boldsymbol{F}_1$ таку ж саму, як і при одержанні $\boldsymbol{QR}$-розкладу матриці $\boldsymbol{A}$. Потім матрицю $\boldsymbol{F}_1\boldsymbol{A}$ помножимо справа на матрицю

$$\boldsymbol{G}_2 = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H}_2 \end{pmatrix},$$

де $\boldsymbol{H}_2$ — деяка матриця Хаусхолдера порядку $n-1$. На підставі формули (2) маємо:

$$\boldsymbol{F}_1\boldsymbol{A}\boldsymbol{G}_2 = \begin{pmatrix} -\alpha & \boldsymbol{J} \\ \mathbf{0} & \boldsymbol{L} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H}_2 \end{pmatrix} = \begin{pmatrix} -\alpha & \boldsymbol{J}\boldsymbol{H}_2 \\ \mathbf{0} & \boldsymbol{L}\boldsymbol{H}_2 \end{pmatrix}.$$

Підматриця $\boldsymbol{J}$ є $(n-1)$-вимірним вектором. Якщо його взяти за вектор $\boldsymbol{x}$ для матриці $\boldsymbol{H}_2$, то у відповідності до теореми 4.4.1 будемо мати:

$$JH_2 = (H_2 J^T)^T = -\alpha y^T,$$

де $\alpha = \pm |J|$.

Отже, матриця $F_1 A G_2$ має нульові елементи в першому стовпці і першому рядку за винятком, можливо, елементів $\{F_1 A G_2\}_{11}$ і $\{F_1 A G_2\}_{12}$.

Зауважимо, що даремно намагатися одержати нульовий елемент $\{F_1 A G_2\}_{12}$ шляхом множення $F_1 A$ на H_1. Хоч у цьому випадку цей елемент буде нульовим, але будуть втрачені нулі в першому стовпці.

Далі треба матрицю $F_1 A G_2$ помножити зліва на матрицю F_2, побудовану так само, як і для *QR*-розкладу, а потім помножити справа матрицю $F_2 F_1 A G_2$ на матрицю

$$G_3 = \begin{pmatrix} E_2 & 0 \\ 0 & H_3 \end{pmatrix},$$

склавши матрицю Хаусхолдера H_3 так, щоб обернулися в нулі елементи другого рядка, розміщені праворуч елемента, наступного за діагональним. Наприкінці цього процесу отримується розклад (5). □

Зрозуміло, що до дводіагонального вигляду можна звести також прямокутну матрицю.

4.4.6. Майже трикутні матриці. *Майже трикутною*, або *матрицею Хессенберга*, називається матриця, яка відрізняється від верхньої трикутної матриці тим, що діагональ, прилегла до головної і розміщена нижче від неї, не є нульовою.

Теорема 4.4.4. *Квадратну матрицю **A** порядку n за допомогою перетворення подібності з ортогональною матрицею можна звести до майже трикутної форми.*

Д о в е д е н н я. Нижче через $F_2, \ldots, F_{n-1}$ ми будемо позначати матриці, визначені другою формулою (3), і припустимо, що блокова матриця

$$\begin{pmatrix} I & J \\ K & L \end{pmatrix} \qquad (6)$$

є матрицею $\boldsymbol{A}$, причому підматриця $\boldsymbol{I}$ має порядок 1. Помноживши цю матрицю зліва на $\boldsymbol{F}_2$ і справа на $\boldsymbol{F}_2^T$, будемо мати:

$$\boldsymbol{F}_2\boldsymbol{A}\boldsymbol{F}_2^T=\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{H}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{J}\\ \boldsymbol{K} & \boldsymbol{L}\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{H}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{I} & \boldsymbol{J}\boldsymbol{H}_2\\ \boldsymbol{H}_2\boldsymbol{K} & \boldsymbol{H}_2\boldsymbol{L}\boldsymbol{H}_2\end{pmatrix}. \qquad (7)$$

Підматриця $\boldsymbol{K}$ є $(n-1)$-вимірним вектором, який ми візьмемо за вектор $\boldsymbol{x}$ для матриці Хаусхолдера $\boldsymbol{H}_2$. Тоді перший стовпець матриці $\boldsymbol{M}$, де $\boldsymbol{M}$ — права частина формули (7), буде відповідати формі Хессенберга.

Далі будемо вважати, що матриця $\boldsymbol{M}$ представлена у блоковому вигляді (6), де $\boldsymbol{I}$ — підматриця другого порядку. Помноживши $\boldsymbol{M}$ зліва на $\boldsymbol{F}_3$ і справа на $\boldsymbol{F}_3^T$, одержимо:

$$\begin{pmatrix}\boldsymbol{E}_2 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{H}_3\end{pmatrix}\begin{pmatrix}\boldsymbol{I} & \boldsymbol{J}\\ \boldsymbol{K} & \boldsymbol{L}\end{pmatrix}\begin{pmatrix}\boldsymbol{E}_2 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{H}_3\end{pmatrix}=\begin{pmatrix}\boldsymbol{I} & \boldsymbol{J}\boldsymbol{H}_3\\ \boldsymbol{H}_3\boldsymbol{K} & \boldsymbol{H}_3\boldsymbol{L}\boldsymbol{H}_3\end{pmatrix}. \qquad (8)$$

Підматриця $\boldsymbol{K}$ має розміри $(n-2)\times 2$ і її перший стовпець є нульовим. Тому перший стовпець підматриці $\boldsymbol{H}_3\boldsymbol{K}$ теж нульовий. Якщо взяти другий стовпець підматриці $\boldsymbol{K}$ за вектор $\boldsymbol{x}$ для матриці Хаусхолдера $\boldsymbol{H}_3$, то другий стовпець підматриці $\boldsymbol{H}_3\boldsymbol{K}$ буде мати вигляд першого стовпця верхньої трикутної матриці порядку $n-2$.

Отже, два перші стовпці правої частини формули (8) мають вигляд, відповідний до форми Хессенберга.

Продовжуючи цей процес, ми одержимо наприкінці майже трикутну матрицю $\boldsymbol{F}_{n-1}\ldots\boldsymbol{F}_3\boldsymbol{F}_2\boldsymbol{A}\boldsymbol{F}_2^T\boldsymbol{F}_3^T\ldots\boldsymbol{F}_{n-1}^T$, яку можна записати у вигляді $\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q}$, де $\boldsymbol{Q}$ — ортогональна матриця. □

Доведення теореми показує, в якій послідовності треба діяти, щоб перетворити матрицю $\boldsymbol{A}$ до майже трикутної форми. Для цього спочатку потрібно знайти матрицю $\boldsymbol{Q}$ з умови, що $\boldsymbol{Q}^T\boldsymbol{A}$ є матрицею Хессенберга (це робиться за аналогією з одержанням $\boldsymbol{QR}$-розкладу), а потім узяти матрицю $\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q}$.

Очевидно, що коли матриця $\boldsymbol{A}$ симетрична, то її форма Хессенберга є симетричною *тридіагональною* матрицею, у якої крім головної діагоналі ненульовими є дві діагоналі, прилеглі до головної.

Вправи до глави 4

4.1. Довести, що коли вектор $\boldsymbol{a}$ є ортогональним до кожного з векторів $\boldsymbol{b}_1, ..., \boldsymbol{b}_m$, то він є ортогональним до будь-якої лінійної комбінації цих векторів.

4.2. Нехай вектори $\boldsymbol{b}_1, ..., \boldsymbol{b}_n$ одержані за допомогою процедури ортогоналізації Грама — Шмідта системи векторів $\boldsymbol{a}_1, ..., \boldsymbol{a}_n$. Якщо $\boldsymbol{A}$ і $\boldsymbol{B}$ — квадратні матриці, стовпцями яких є координати векторів $\boldsymbol{a}_1, ..., \boldsymbol{a}_n$ і $\boldsymbol{b}_1, ..., \boldsymbol{b}_n$ відповідно, то як пов'язані між собою визначники $\det \boldsymbol{A}$ і $\det \boldsymbol{B}$?

4.3. Довести теорему: для того, щоб у деякому базисі простору E_n скалярний добуток двох довільних векторів дорівнював сумі добутків відповідних координат цих векторів, необхідно і достатньо, щоб базис був ортонормованим.

4.4. Довести, що в будь-якому евклідовому просторі існує: 1) ортогональний базис; 2) ортонормований базис.

4.5. Довести, що елементарна матриця $\boldsymbol{L}_{ij}$ є ортогональною.

4.6. За якої умови діагональна матриця є ортогональною?

4.7. Довести, що квазідіагональна матриця є ортогональною тоді і тільки тоді, коли ортогональними є її діагональні блоки.

4.8. Для матриці обертання довести, що

$$\boldsymbol{T}_{ij}(\varphi_1)\boldsymbol{T}_{ij}(\varphi_2) = \boldsymbol{T}_{ij}(\varphi_1 + \varphi_2), \quad \boldsymbol{T}_{ij}^{-1}(\varphi) = \boldsymbol{T}_{ij}(-\varphi).$$

4.9. Чи єдиною є матриця $\boldsymbol{Q}$ в рівності (4.3.7)?

4.10. Довести, що: 1) матриця подібна до самої себе; 2) коли $\boldsymbol{A}$ подібна до $\boldsymbol{B}$, то $\boldsymbol{B}$ подібна до $\boldsymbol{A}$; 3) коли $\boldsymbol{A}$ подібна до $\boldsymbol{B}$ і $\boldsymbol{B}$ подібна до $\boldsymbol{C}$, то $\boldsymbol{A}$ подібна до $\boldsymbol{C}$.

4.11. Довести, що коли матриці $\boldsymbol{A}$ і $\boldsymbol{B}$ подібні, то подібними є матриці: 1) $\boldsymbol{A}^m$ і $\boldsymbol{B}^m$ при будь-якому натуральному числі m; 2) $f(\boldsymbol{A})$ і $f(\boldsymbol{B})$, де $f(\lambda)$ — многочлен; 3) $\boldsymbol{A}^{-1}$ і $\boldsymbol{B}^{-1}$, якщо $\boldsymbol{A}$ і $\boldsymbol{B}$ є невиродженими.

4.12. Чи є еквівалентні квадратні матриці подібними? Чи є подібні матриці еквівалентними?

4.13. Як для заданої матриці $\boldsymbol{A}$ знайти ортонормовані базиси підпросторів $\mathcal{R}(\boldsymbol{A})$ і $\mathcal{N}(\boldsymbol{A})$?

4.14 Як треба змінити алгоритм $\boldsymbol{QR}$-розкладу, щоб він обчислював $\boldsymbol{QL}$-розклад, де $\boldsymbol{Q}$, $\boldsymbol{L}$ — ортогональна і нижня трикутна матриці відповідно?

4.15. Як можна застосувати процес ортогоналізації для розв'язання систем лінійних рівнянь?

4.16. Яким є узагальнення формули (4.4.4) для випадку, коли $r(\boldsymbol{A}) < n$?

4.17. Нехай $m \times n$-матриця $\boldsymbol{A}$ при $m < n$ має лінійно незалежні рядки. Показати, що аналогом формули (4.4.4) є представлення матриці $\boldsymbol{A}$ у вигляді рівності $\boldsymbol{A} = (\boldsymbol{L} \quad \boldsymbol{0})\boldsymbol{Q}$.

4.18. Чому при перетворенні матриці $\boldsymbol{A}$ до форми Хессенберга не починають множити на матриці Хаусхолдера $\boldsymbol{H}_1$ і $\boldsymbol{H}_1^T$?

Г Л А В А 5

ГЕОМЕТРІЯ ЕВКЛІДОВОГО ПРОСТОРУ

У цій главі розглядаються геометричні задачі, пов'язані з визначенням відстаней, кутів, об'ємів. Основою для розгляду цих питань є теорема про ортогональний розклад.

5.1. Теорема про ортогональний розклад

5.1.1. Розклад евклідового простору в пряму суму підпростору і його ортогонального доповнення. Дві множини U, V евклідового простору називаються ортогональними $(U \perp V)$, якщо кожний вектор з U ортогональний до кожного вектора з V. Множина всіх векторів, ортогональних до множини U, називається *ортогональним доповненням* множини U і позначається через $U^{\perp}$. Ортогональне доповнення $U^{\perp}$ є підпростором незалежно від того, чи є підпростором U. Дійсно, якщо $\boldsymbol{a}, \boldsymbol{b} \in U^{\perp}$, то $\boldsymbol{a}, \boldsymbol{b} \perp U$. Але тоді $\alpha\boldsymbol{a} + \beta\boldsymbol{b} \perp U$ для будь-яких чисел α і β, тобто $\alpha\boldsymbol{a} + \beta\boldsymbol{b} \in U^{\perp}$.

Теорема 5.1.1. *Якщо U — підпростір евклідового простору E_n, то*

$$\dim U + \dim U^{\perp} = n. \tag{1}$$

Д о в е д е н н я. Нехай $\dim U = m$. Виберемо в U базис і доповнимо його до базису всього простору E_n. До побудованого базису застосуємо ортогоналізацію Грама — Шмідта, а потім помножимо кожний з одержаних векторів на величину, обернену до його довжини. У результаті вказаних операцій ми одержимо ортонормований базис простору E_n, вектори якого ми позначимо через $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n$.

Очевидно, що вектори $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ складають базис підпростору U. Залишається переконатися в тому, що вектори $\boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n$ утворюють базис підпростору $U^{\perp}$. Дійсно, якщо

$$\boldsymbol{a} = \alpha_1 \boldsymbol{e}_1 + \ldots + \alpha_m \boldsymbol{e}_m + \alpha_{m+1} \boldsymbol{e}_{m+1} + \ldots + \alpha_n \boldsymbol{e}_n \in U^{\perp},$$

то після почергового множення справа на $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ одержимо:

$$\alpha_1 = \boldsymbol{a}^T \boldsymbol{e}_1 = 0, \quad \ldots, \quad \alpha_m = \boldsymbol{a}^T \boldsymbol{e}_m = 0,$$

тобто довільний вектор $\boldsymbol{a} \in U^{\perp}$ має вигляд лінійної комбінації лінійно незалежних векторів $\boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n$. Тому ці вектори є базисом $U^{\perp}$ і $\dim U^{\perp} = n - m$.

Отже, рівність (1) доведена. □

Розглянемо теорему, яка надалі буде відігравати важливу роль.

***Теорема 5.1.2* (*теорема про ортогональний розклад*).** *Якщо U є підпростором евклідового простору E_n, то*

$$E_n = U \oplus U^{\perp}, \tag{2}$$

$$U^{\perp\perp} = U. \tag{3}$$

де $U^{\perp\perp} = (U^{\perp})^{\perp}$.

Д о в е д е н н я. Припустимо, що вектор $\boldsymbol{a}$ одночасно належить до U і до $U^{\perp}$. Тоді $\boldsymbol{a}^T \boldsymbol{a} = 0$, звідки $\boldsymbol{a} = \boldsymbol{0}$, тобто єдиним спільним для U і $U^{\perp}$ вектором є нульовий вектор. Це означає, що $U \bigcap U^{\perp}$ є нульовим підпростором і тому сума підпросторів U і $U^{\perp}$ є прямою.

З формули Грассмана і теореми 5.1.1 випливає, що

$$\dim (U \oplus U^{\perp}) = \dim U + \dim U^{\perp} = n.$$

Отже, $U \oplus U^{\perp}$ має вимірність простору E_n і тому збігається з E_n, тобто справедлива рівність (2).

Для доведення формули (3) зауважимо, що у відповідності до визначення ортогонального доповнення будь-який вектор $\boldsymbol{a} \in U$ належить до $(U^{\perp})^{\perp}$. Але вимірність підпростору $(U^{\perp})^{\perp}$ на підставі формули (1) дорівнює

$$\dim (U^{\perp})^{\perp} = n - \dim U^{\perp} = n - (n - \dim U) = \dim U,$$

тому $(U^{\perp})^{\perp}$ збігається з U.

Висновок. *Якщо деякий вектор* $\boldsymbol{a} \in E_n$ *є ортогональним до всіх векторів простору* E_n, *то* $\boldsymbol{a} = \mathbf{0}$.

Д о в е д е н н я. Вектор буде ортогональним до всіх векторів евклідового простору, якщо він є ортогональним до векторів деякого базису простору. Припустимо, що $\boldsymbol{a} \neq \mathbf{0}$. Вектор $\boldsymbol{a}$ породжує одновимірний підпростір U, а вектори базису повинні належати підпростору $U^{\perp}$. Проте це неможливо, оскільки при цьому буде $\dim U^{\perp} = n - 1$. Отже, $\boldsymbol{a} = \mathbf{0}$. ☐

П р и к л а д. Припустимо, що підпростір U породжується векторами:

1) $\boldsymbol{u}_1 = (1 \quad 3 \quad 1 \quad -1)^T$, $\boldsymbol{u}_2 = (2 \quad 5 \quad 0 \quad 1)^T$;

2) $\boldsymbol{u}_1 = (a_{11} \quad a_{12} \quad a_{13} \quad \ldots \quad a_{1n})^T$.

Знайдемо ортогональне доповнення до підпростору U.

Припустимо, що в першому випадку довільний вектор $\boldsymbol{x} \in U^{\perp}$ має вигляд:

$$\boldsymbol{x} = (x_1 \quad x_2 \quad x_3 \quad x_4)^T.$$

Він повинен бути ортогональним до векторів $\boldsymbol{u}_1$ і $\boldsymbol{u}_2$, тобто

$$\boldsymbol{u}_1^T \boldsymbol{x} = 0, \quad \boldsymbol{u}_2^T \boldsymbol{x} = 0.$$

Звідси випливає, що координати вектора $\boldsymbol{x}$ задовольняють системі рівнянь

$$\begin{aligned} x_1 + 3x_2 + x_3 - x_4 &= 0, \\ 2x_1 + 5x_2 \qquad\ \ + x_4 &= 0. \end{aligned}$$

Загальний розв'язок цієї системи дорівнює

$$\boldsymbol{x} = x_3(5 \quad -2 \quad 1 \quad 0)^T + x_4(-8 \quad 3 \quad 0 \quad 1)^T.$$

Це — векторне рівняння підпростору, породженого векторами

$$\boldsymbol{u}_3 = (5 \quad -2 \quad 1 \quad 0)^T, \quad \boldsymbol{u}_4 = (-8 \quad 3 \quad 0 \quad 1)^T.$$

Отже, підпростір $U^{\perp}$ визначається своїм базисом $\boldsymbol{u}_3$, $\boldsymbol{u}_4$.

У другому випадку для координат вектора

$$\boldsymbol{x} = (x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_n)^T$$

маємо рівняння

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = 0 .$$

Принаймні одна з координат вектора $\boldsymbol{u}_1$ є ненульовою. Нехай, наприклад, $a_{11} \neq 0$. Тоді загальний розв'язок останнього рівняння має вигляд:

$$\boldsymbol{x} = x_2\boldsymbol{u}_2 + x_3\boldsymbol{u}_3 + \ldots + x_n\boldsymbol{u}_n ,$$

де

$$\boldsymbol{u}_2 = \begin{pmatrix} -\dfrac{a_{12}}{a_{11}} & 1 & 0 & \ldots & 0 \end{pmatrix}^T ,$$

$$\boldsymbol{u}_3 = \begin{pmatrix} -\dfrac{a_{13}}{a_{11}} & 0 & 1 & \ldots & 0 \end{pmatrix}^T ,$$

. . . .

$$\boldsymbol{u}_n = \begin{pmatrix} -\dfrac{a_{1n}}{a_{11}} & 0 & 0 & \ldots & 1 \end{pmatrix}^T ,$$

Аналогічно можна знайти загальний розв'язок, коли ненульовою є будь-яка інша координата вектора $\boldsymbol{u}_1$.

Таким чином, у просторі E_n ортогональним доповненням до одновимірного підпростору, породженого вектором $\boldsymbol{u}_1$, є спрямовуючий підпростір гіперплощини, перпендикулярної до $\boldsymbol{u}_1$. △

5.1.2. Властивості образу і ядра матриці. За допомогою теореми про ортогональний розклад можна дослідити властивості образу і ядра матриці, введених у розділі 4.3.

Теорема 5.1.3. *Для будь-якої матриці* $\boldsymbol{A}$ *має місце рівність*

$$\mathcal{N}(\boldsymbol{A}) = \mathcal{R}^{\perp}(\boldsymbol{A}^T) . \tag{4}$$

Д о в е д е н н я. Теорема буде доведена, якщо ми покажемо, що будь-який вектор $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A})$ належить до $\mathcal{R}^{\perp}(\boldsymbol{A}^T)$ і, навпаки, довільний вектор $\boldsymbol{x} \in \mathcal{R}^{\perp}(\boldsymbol{A}^T)$ належить до $\mathcal{N}(\boldsymbol{A})$.

Дійсно, якщо $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A})$, то $\boldsymbol{Ax} = \boldsymbol{0}$ і, отже, вектор $\boldsymbol{x}$ ортогональний до кожного рядка матриці $\boldsymbol{A}$. Тому $\boldsymbol{x} \perp \mathcal{R}(\boldsymbol{A}^T)$, тобто $\boldsymbol{x} \in \mathcal{R}^{\perp}(\boldsymbol{A}^T)$.

Нехай, навпаки, $\boldsymbol{x} \in \mathcal{R}^{\perp}(\boldsymbol{A}^T)$. Це означає, що $\boldsymbol{Ax} = \boldsymbol{0}$, і тому $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A})$. □

Теорема 5.1.3. показує, що утворений розв'язками системи $\boldsymbol{Ax} = \boldsymbol{0}$ підпростір $\mathcal{N}(\boldsymbol{A})$ є ортогональним доповненням до підпростору, натягнутого на рядки матриці $\boldsymbol{A}$. Вимірність останнього підпростору дорівнює $r(\boldsymbol{A})$. З формули (1) випливає, що вимірність $\mathcal{N}(\boldsymbol{A})$ дорівнює $n - r(\boldsymbol{A})$. У розділі 4.3 цей результат був одержаний на підставі теореми 3.5.1.

Теорема 5.1.4. *Будь-який вектор $\boldsymbol{a} \in E_n$ можна розкласти єдиним способом у суму $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in \mathcal{R}(\boldsymbol{A})$, $\boldsymbol{v} \in \mathcal{N}(\boldsymbol{A}^T)$ і матриця* ***A*** *має n рядків.*

Д о в е д е н н я. На підставі теореми про ортогональний розклад можна стверджувати, що $\boldsymbol{u} \in \mathcal{R}(\boldsymbol{A})$ і $\boldsymbol{v} \in \mathcal{R}^{\perp}(\boldsymbol{A})$. З рівності (4) випливає, що

$$\mathcal{R}^{\perp}(\boldsymbol{A}) = \mathcal{N}(\boldsymbol{A}^T). \tag{5}$$

Тому $\boldsymbol{v} \in \mathcal{N}(\boldsymbol{A}^T)$. □

Теорема 5.1.5. *Для будь-якої матриці* ***A*** *виконуються співвідношення*

$$\mathcal{N}(\boldsymbol{A}) = \mathcal{N}(\boldsymbol{A}^T\boldsymbol{A}), \quad \mathcal{N}(\boldsymbol{A}^T) = \mathcal{N}(\boldsymbol{A}\boldsymbol{A}^T), \tag{6}$$

$$\mathcal{R}(\boldsymbol{A}) = \mathcal{R}(\boldsymbol{A}\boldsymbol{A}^T), \quad \mathcal{R}(\boldsymbol{A}^T) = \mathcal{R}(\boldsymbol{A}^T\boldsymbol{A}). \tag{7}$$

Д о в е д е н н я. Щоб довести першу рівність (6), зауважимо, що коли $\boldsymbol{Ax} = \boldsymbol{0}$, то $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{0}$. З іншого боку, якщо $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{0}$, то $\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{Ax} = 0$, тобто $(\boldsymbol{Ax})^T\boldsymbol{Ax} = 0$. Звідси випливає, що $\boldsymbol{Ax} = \boldsymbol{0}$. Отже, $\boldsymbol{Ax} = \boldsymbol{0}$ тоді і тільки тоді, коли $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{0}$. Це означає, що $\mathcal{N}(\boldsymbol{A}) = \mathcal{N}(\boldsymbol{A}^T\boldsymbol{A})$.

Друга рівність (6) безпосередньо випливає з першої.

Рівності (7) є наслідком рівностей (6). Це легко доводиться за допомогою рівностей (5) і (3).

Висновок. *Мають місце співвідношення*

$$r(\boldsymbol{A}) = r(\boldsymbol{A}\boldsymbol{A}^T) = r(\boldsymbol{A}^T\boldsymbol{A}). \tag{8}$$

Д о в е д е н н я випливає з рівностей (7), якщо врахувати формулу (4.3.8). □

5.1.3. Теорема Фредгольма і альтернатива Фредгольма. При вивченні теореми Кронекера — Капеллі зазначалося, що вона не є досить зручною при дослідженні сумісності системи $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Існує інший підхід до цього питання, не пов'язаний з розглядом рангів матриць $\boldsymbol{A}$ і $\boldsymbol{A}_b$. Таку можливість надає

***Теорема 5.1.6* (*теорема Фредгольма*).** *Для того, щоб система* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *була сумісною, необхідно і достатньо, щоб вектор* $\boldsymbol{b}$ *був ортогональним до всіх розв'язків системи* $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$.

Д о в е д е н н я. Н е о б х і д н і с т ь. Якщо система $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ має розв'язок, то $\boldsymbol{b} \in \mathscr{R}(\boldsymbol{A})$, або, з урахуванням формули (5), $\boldsymbol{b} \in \mathscr{N}^{\perp}(\boldsymbol{A}^T)$. Це означає, що $\boldsymbol{b} \perp \mathscr{N}(\boldsymbol{A}^T)$, тобто $\boldsymbol{b}^T\boldsymbol{y} = 0$ для всіх векторів $\boldsymbol{y}$, що задовольняють рівнянню $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$.

Д о с т а т н і с т ь. Нехай $\boldsymbol{b}^T\boldsymbol{y} = 0$ для тих же векторів $\boldsymbol{y}$, тоді $\boldsymbol{b} \perp \mathscr{N}(\boldsymbol{A}^T)$, або, з урахуванням формули (5), $\boldsymbol{b} \in \mathscr{R}(\boldsymbol{A})$. Це означає, що можна знайти вектор $\boldsymbol{x}$ такий, що $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$. □

При дослідженні розв'язуваності лінійних систем буває корисною

***Теорема 5.1.7* (*альтернатива Фредгольма*).** *Або система* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *є сумісною при будь-якому векторі* $\boldsymbol{b}$, *або система* $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$ *має ненульові розв'язки.*

Д о в е д е н н я. Це — альтернативна теорема, оскільки одночасно знайти $\boldsymbol{x}$ і $\boldsymbol{y} \neq \boldsymbol{0}$ неможливо. Дійсно, помноживши зліва обидві частини рівності $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$ на $\boldsymbol{x}^T$, з урахуванням рівності $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ будемо мати:

$$0 = \boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{x})^T\boldsymbol{y} = \boldsymbol{b}^T\boldsymbol{y},$$

що суперечить висновку з теореми про ортогональний розклад.

Покажемо, що коли система $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ є несумісною, то система $\boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{0}$ має ненульовий розв'язок. Скориставшись теоремою 5.1.4, розкладемо вектор $\boldsymbol{b}$ у суму $\boldsymbol{b} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in \mathscr{R}(\boldsymbol{A})$, $\boldsymbol{v} \in \mathscr{N}(\boldsymbol{A}^T)$. Оскільки система $\boldsymbol{Ax} = \boldsymbol{b}$ не має розв'язків, то $\boldsymbol{b} \notin \mathscr{R}(\boldsymbol{A})$, тобто $\boldsymbol{v} \neq \boldsymbol{0}$. Це означає, що система $\boldsymbol{A}^T \boldsymbol{y} = \boldsymbol{0}$ має ненульовий розв'язок $\boldsymbol{y} = \boldsymbol{v}$. □

5.2. Відстані, кути, об'єми

5.2.1. Визначник Грама. Введемо до розгляду визначник, за допомогою якого будуть розв'язані деякі задачі цього розділу. Мова йде про визначник матриці Грама

$$\boldsymbol{\Gamma}_m = \begin{pmatrix} \boldsymbol{a}_1^T\boldsymbol{a}_1 & \boldsymbol{a}_1^T\boldsymbol{a}_2 & \ldots & \boldsymbol{a}_1^T\boldsymbol{a}_m \\ \boldsymbol{a}_2^T\boldsymbol{a}_1 & \boldsymbol{a}_2^T\boldsymbol{a}_2 & \ldots & \boldsymbol{a}_2^T\boldsymbol{a}_m \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{a}_m^T\boldsymbol{a}_1 & \boldsymbol{a}_m^T\boldsymbol{a}_2 & \ldots & \boldsymbol{a}_m^T\boldsymbol{a}_m \end{pmatrix}.$$

Він називається *визначником Грама* для векторів $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ і позначається через $G(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m)$.

Покажемо, як за допомогою визначника Грама можна дослідити лінійну залежність або незалежність векторів без розкладу за векторами базису.

Теорема 5.2.1. *Для того, щоб вектори $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ були лінійно незалежними, необхідно і достатньо, щоб визначник Грама цих векторів був додатним.*

Доведення. Необхідність. Будемо вважати, що $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in R_n$ і $m \le n$. Очевидно, що $\boldsymbol{\Gamma}_m = \boldsymbol{A}^T\boldsymbol{A}$, де $\boldsymbol{A}$ є $n \times m$-матрицею, стовпці якої складаються з координат векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$. На підставі $\boldsymbol{QR}$-розкладу (4.4.4) маємо:

$$\boldsymbol{\Gamma}_m = \boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix}^T \boldsymbol{Q}^T\boldsymbol{Q} \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix} = \boldsymbol{R}^T\boldsymbol{R}, \quad \det\boldsymbol{\Gamma}_m = (\det\boldsymbol{R})^2 \ge 0.$$

Якщо вектори $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ лінійно незалежні, то матриця $\boldsymbol{R}$ є невиродженою і тому $\det\boldsymbol{\Gamma}_m > 0$.

Д о с т а т н і с т ь. Припустимо, що $\det\boldsymbol{\Gamma}_m > 0$. Оскільки порядок матриці $\boldsymbol{A}^T\boldsymbol{A}$ дорівнює m, то $r(\boldsymbol{A}^T\boldsymbol{A}) = m$. З рівності $r(\boldsymbol{A}^T\boldsymbol{A}) = r(\boldsymbol{A})$, про яку йшла мова у висновку з теореми 5.1.5, випливає, що $r(\boldsymbol{A}) = m$. Це означає, що стовпці матриці $\boldsymbol{A}$ є лінійно незалежними.

Висновок 1. *Вектори* $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m \in R_n$ *є лінійно залежними тоді і тільки тоді, коли визначник Грама цих векторів дорівнює нулю.*

Д о в е д е н н я. Висновок доводиться від супротивного з урахуванням твердження теореми.

Висновок 2. *Має місце нерівність Коші — Шварца — Буняковського*

$$(\boldsymbol{a}_1^T\boldsymbol{a}_2)^2 \le (\boldsymbol{a}_1^T\boldsymbol{a}_1)(\boldsymbol{a}_2^T\boldsymbol{a}_2),$$

причому рівність виконується тоді і тільки тоді, коли вектори $\boldsymbol{a}_1$ *і* $\boldsymbol{a}_2$ *є лінійно залежними.*

Д о в е д е н н я випливає з теореми і висновку 1 при $m = 2$. □

5.2.2. Перпендикуляр і ортогональна проекція вектора на підпростір. У розділі 5.1 було показано, що будь-який вектор $\boldsymbol{a} \in E_n$ можна представити єдиним способом у вигляді суми $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in U$, $\boldsymbol{v} \in U^{\perp}$. Вектор $\boldsymbol{u}$ називається *ортогональною проекцією вектора* $\boldsymbol{a}$ *на підпростір* U, а $\boldsymbol{v}$ — *перпендикуляром*, опущеним з $\boldsymbol{a}$ на U. Якщо $\boldsymbol{a} \in U$, то $\boldsymbol{v} = \boldsymbol{0}$. Одночасно вектор $\boldsymbol{v}$ є ортогональною проекцією $\boldsymbol{a}$ на $U^{\perp}$, причому $\boldsymbol{u} = \boldsymbol{0}$, якщо $\boldsymbol{a} \in U^{\perp}$.

Нехай підпростір U має вимірність m і вектори $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ утворюють базис цього підпростору. Визначимо вектори $\boldsymbol{u}$ і $\boldsymbol{v}$. Представимо вектор $\boldsymbol{u}$ у вигляді розкладу

$$\boldsymbol{u} = \sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i, \qquad (1)$$

де $\alpha_1, \ldots, \alpha_m$ — деякі дійсні числа. Для їх визначення скористаємося ортогональністю вектора $\boldsymbol{v}$ до векторів $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$. Очевидно, що $\boldsymbol{v}^T\boldsymbol{u}_j = (\boldsymbol{a} - \boldsymbol{u})^T\boldsymbol{u}_j = 0, \quad j = \overline{1,\ m}$, звідки маємо систему рівнянь відносно $\alpha_1, \ldots, \alpha_m$:

$$\sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i^T \boldsymbol{u}_j - \boldsymbol{a}^T \boldsymbol{u}_j = 0, \qquad j = \overline{1,\ m}. \tag{2}$$

Визначник матриці цієї системи є визначником Грама векторів $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$. Оскільки він відмінний від нуля, то система (2) має єдиний розв'язок. Визначивши $\alpha_1, \ldots, \alpha_m$ з цієї системи, одержимо вектор $\boldsymbol{u}$ після підставлення величин $\alpha_1, \ldots, \alpha_m$ у формулу (1). Вектор $\boldsymbol{v}$ тепер визначається як різниця

$$\boldsymbol{v} = \boldsymbol{a} - \boldsymbol{u}. \tag{3}$$

Для знаходження векторів $\boldsymbol{u}$ і $\boldsymbol{v}$ можна застосувати інші міркування. Розглянемо разом з рівняннями (2) співвідношення

$$\sum_{i=1}^{m} \alpha_i \boldsymbol{u}_i - \boldsymbol{u} = \boldsymbol{0}, \tag{4}$$

яке випливає з (1). Замінимо в (3) вектори їх *k*-ми координатами в деякому базисі, де $k = \overline{1,\ n}$. Тоді система (2), (4) буде однорідною системою рівнянь відносно $\alpha_1, \ldots, \alpha_m, -1$. Запишемо умову існування ненульового розв'язку цієї системи. Для цього треба у відповідності до висновку з теореми 3.5.1 порівняти з нулем визначник матриці системи. Всього буде *n* таких умов і ми запишемо їх у вигляді однієї умови відносно векторів $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$, $\boldsymbol{u}$:

$$\begin{vmatrix} \boldsymbol{u}_1^T \boldsymbol{u}_1 & \boldsymbol{u}_2^T \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_m^T \boldsymbol{u}_1 & \boldsymbol{a}^T \boldsymbol{u}_1 \\ \boldsymbol{u}_1^T \boldsymbol{u}_2 & \boldsymbol{u}_2^T \boldsymbol{u}_2 & \ldots & \boldsymbol{u}_m^T \boldsymbol{u}_2 & \boldsymbol{a}^T \boldsymbol{u}_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{u}_1^T \boldsymbol{u}_m & \boldsymbol{u}_2^T \boldsymbol{u}_m & \ldots & \boldsymbol{u}_m^T \boldsymbol{u}_m & \boldsymbol{a}^T \boldsymbol{u}_m \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 & \ldots & \boldsymbol{u}_m & \boldsymbol{u} \end{vmatrix} = \boldsymbol{0}. \tag{5}$$

Рівність (5) рівносильна рівності

$$\boldsymbol{u} \det \boldsymbol{\Gamma}_m + \begin{vmatrix} & & & \boldsymbol{u}_1 \\ & \boldsymbol{\Gamma}_m & & \ldots \\ & & & \boldsymbol{u}_m \\ \boldsymbol{a}^T \boldsymbol{u}_1 & \ldots & \boldsymbol{a}^T \boldsymbol{u}_m & 0 \end{vmatrix} = \boldsymbol{0},$$

де $\boldsymbol{\Gamma}_m$ — матриця Грама векторів $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$. У цьому можна переконатися, якщо транспонувати визначник у лівій частині рівності (5), а потім розкласти його за елементами останнього стовпця. Оскільки $\det \boldsymbol{\Gamma}_m \neq 0$, то на підставі останньої рівності маємо:

$$\boldsymbol{u} = -\frac{1}{\det \boldsymbol{\Gamma}_m} \begin{vmatrix} & & & \boldsymbol{u}_1 \\ & \boldsymbol{\Gamma}_m & & \ldots \\ & & & \boldsymbol{u}_m \\ \boldsymbol{a}^T\boldsymbol{u}_1 & \ldots & \boldsymbol{a}^T\boldsymbol{u}_m & 0 \end{vmatrix}. \tag{6}$$

Ця рівність визначає ортогональну проекцію вектора $\boldsymbol{a}$ на підпростір U. Перпендикуляр, опущений з $\boldsymbol{a}$ на U, визначається за допомогою формули (3):

$$\boldsymbol{v} = \boldsymbol{a} - \boldsymbol{u} = \frac{1}{\det \boldsymbol{\Gamma}_m} \begin{vmatrix} & & & \boldsymbol{u}_1 \\ & \boldsymbol{\Gamma}_m & & \ldots \\ & & & \boldsymbol{u}_m \\ \boldsymbol{a}^T\boldsymbol{u}_1 & \ldots & \boldsymbol{a}^T\boldsymbol{u}_m & \boldsymbol{a} \end{vmatrix}.$$

При цьому

$$|\, \boldsymbol{v} \,|^2 = \boldsymbol{v}^T\boldsymbol{v} = \boldsymbol{v}^T(\boldsymbol{a} - \boldsymbol{u}) = \boldsymbol{v}^T\boldsymbol{a} = \frac{1}{\det \boldsymbol{\Gamma}_m} \begin{vmatrix} & & & \boldsymbol{u}_1^T\boldsymbol{a} \\ & \boldsymbol{\Gamma}_m & & \ldots \\ & & & \boldsymbol{u}_m^T\boldsymbol{a} \\ \boldsymbol{a}^T\boldsymbol{u}_1 & \ldots & \boldsymbol{a}^T\boldsymbol{u}_m & \boldsymbol{a}^T\boldsymbol{a} \end{vmatrix}$$

і довжина h перпендикуляра дорівнює

$$h = |\, \boldsymbol{v} \,| = \sqrt{\frac{G(\boldsymbol{u}_1,\, \boldsymbol{u}_2,\, \ldots,\, \boldsymbol{u}_m,\, \boldsymbol{a})}{G(\boldsymbol{u}_1,\, \boldsymbol{u}_2,\, \ldots,\, \boldsymbol{u}_m)}}. \tag{7}$$

З а у в а ж е н н я. Для розв'язуваності системи рівнянь (2) вектори $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ не обов'язково повинні утворювати базис підпростору U. Щоб переконатися в цьому, напишемо систему (2) у вигляді:

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{a}, \tag{8}$$

де $\boldsymbol{A} = (\boldsymbol{u}_1 \ \dots \ \boldsymbol{u}_m)$, $\boldsymbol{x} = (\alpha_1 \ \dots \ \alpha_m)^T$. Очевидно, що $\boldsymbol{A}^T\boldsymbol{a} \in \mathscr{R}(\boldsymbol{A}^T)$. З іншого боку, на підставі другої формули (5.1.7) $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} \in \mathscr{R}(\boldsymbol{A}^T)$. Тепер рівність (8) означає, що вектор $\boldsymbol{A}^T\boldsymbol{a} \in \mathscr{R}(\boldsymbol{A}^T)$ представляється у вигляді лінійної комбінації (базисних) векторів з $\mathscr{R}(\boldsymbol{A}^T)$, що завжди можливо. Отже, система рівнянь (2) є розв'язуваною. За коефіцієнти $\alpha_1, \ \dots, \ \alpha_m$ можна взяти будь-який її розв'язок.

П р и к л а д 1. Знайдемо ортогональну проекцію $\boldsymbol{u}$ вектора

$$\boldsymbol{a} = (3 \quad -1 \quad 4 \quad 0)^T$$

на підпростір, породжений векторами

$$\boldsymbol{u}_1 = (1 \quad 1 \quad 0 \quad 0)^T, \quad \boldsymbol{u}_2 = (0 \quad 0 \quad 1 \quad -1)^T, \quad \boldsymbol{u}_3 = (1 \quad 1 \quad -1 \quad 1)^T,$$

а також перпендикуляр $\boldsymbol{v}$, опущений з $\boldsymbol{a}$ на цей підпростір.

Шукаємо вектор $\boldsymbol{u}$ у вигляді суми $\boldsymbol{u} = \sum_{i=1}^{3} \alpha_i \boldsymbol{u}_i$. Оскільки

$$\boldsymbol{u}_1^T\boldsymbol{u}_1 = 2, \ \boldsymbol{u}_1^T\boldsymbol{u}_2 = 0, \ \boldsymbol{u}_1^T\boldsymbol{u}_3 = 2, \ \boldsymbol{u}_2^T\boldsymbol{u}_2 = 2, \ \boldsymbol{u}_2^T\boldsymbol{u}_3 = -2, \ \boldsymbol{u}_3^T\boldsymbol{u}_3 = 4,$$

$$\boldsymbol{a}^T\boldsymbol{u}_1 = 2, \ \boldsymbol{a}^T\boldsymbol{u}_2 = 4, \ \boldsymbol{a}^T\boldsymbol{u}_3 = -2,$$

то система рівнянь відносно $\alpha_1, \alpha_2, \alpha_3$ є такою:

$$\begin{aligned} 2\alpha_1 \qquad\quad + 2\alpha_3 &= \ \ 2, \\ 2\alpha_2 - 2\alpha_3 &= \ \ 4, \\ 2\alpha_1 - 2\alpha_2 + 4\alpha_3 &= -2. \end{aligned}$$

Її загальний розв'язок дорівнює $(1 \quad 2 \quad 0)^T + \alpha_3(-1 \quad 1 \quad 1)^T$. При $\alpha_3 = 0$ знаходимо:

$$\boldsymbol{u} = \alpha_1\boldsymbol{u}_1 + \alpha_2\boldsymbol{u}_2 =$$

$$= (1 \quad 1 \quad 0 \quad 0)^T + 2(0 \quad 0 \quad 1 \quad -1)^T = (1 \quad 1 \quad 2 \quad -2)^T.$$

Таким же буде вектор $\boldsymbol{u}$ при будь-якому іншому значенні α_3. Вектор $\boldsymbol{v}$ дорівнює

$$\boldsymbol{v} = \boldsymbol{a} - \boldsymbol{u} = (3 \quad -1 \quad 4 \quad 0)^T - (1 \quad 1 \quad 2 \quad -2)^T = (2 \quad -2 \quad 2 \quad 2)^T.$$

Щоб не мати справи з розв'язанням системи рівнянь, можна обчислити вектор $\boldsymbol{u}$ за допомогою формули (6). Проте в цьому разі треба подбати про знаходження базису підпростору U, бо вектори $\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \boldsymbol{u}_3$ є лінійно залежними. Щоб відшукати вказаний базис, треба серед векторів $\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \boldsymbol{u}_3$ знайти лінійно незалежні. Такими є, наприклад, вектори $\boldsymbol{u}_1$ і $\boldsymbol{u}_2$. Тепер матриця $\boldsymbol{\Gamma}_2$ у формулі (6) має вигляд

$$\boldsymbol{\Gamma}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

а її визначник дорівнює $G(\boldsymbol{u}_1,\ \boldsymbol{u}_2) = 4$. Обчислимо вектор $\boldsymbol{u}$ за формулою (6):

$$\boldsymbol{u} = -\frac{1}{4}\begin{vmatrix} 2 & 0 & \boldsymbol{u}_1 \\ 0 & 2 & \boldsymbol{u}_2 \\ 2 & 4 & 0 \end{vmatrix} = \boldsymbol{u}_1 + 2\boldsymbol{u}_2 = (1 \quad 1 \quad 2 \quad -2)^T.$$

Це є результат, одержаний вище. △

5.2.3. Відстані. Про визначення довжини і ортогональності векторів мова йшла в розділі 4.1. Зупинимося на узагальненнях цих понять. При цьому, розглядаючи афінний простір S_n, ми будемо вважати його простір векторів евклідовим. Ми будемо називати величину $|\boldsymbol{a} - \boldsymbol{b}|$ *відстанню між векторами* $\boldsymbol{a}$ і $\boldsymbol{b}$ або *відстанню між точками* простору S_n з радіус-векторами $\boldsymbol{a}$ і $\boldsymbol{b}$. Множина векторів $\boldsymbol{a}$, які задовольняють нерівності $|\boldsymbol{a} - \boldsymbol{b}| < \delta$, називається δ-околом точки $\boldsymbol{b}$.

Очевидно, що

$$|\boldsymbol{a} - \boldsymbol{b}|^2 = (\boldsymbol{a} - \boldsymbol{b})^T(\boldsymbol{a} - \boldsymbol{b}) = |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 - 2\boldsymbol{a}^T\boldsymbol{b}. \qquad (9)$$

Якщо $\boldsymbol{a} \perp \boldsymbol{b}$, то маємо *теорему Піфагора*:

$$|\boldsymbol{a} - \boldsymbol{b}|^2 = |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 .$$

Крім того, з формули (9) і нерівності Коші — Шварца — Буняковського випливає, що

$$|\boldsymbol{a} - \boldsymbol{b}| \le |\boldsymbol{a}| + |\boldsymbol{b}|, \qquad |\boldsymbol{a} - \boldsymbol{b}| \ge \|\boldsymbol{a}| - |\boldsymbol{b}\|.$$

Отже, для сторін трикутника виконуються відомі з елементарної геометрії нерівності.

Відстанню між вектором $\boldsymbol{a}$ *і підпростором* U називається найменша довжина вектора $\boldsymbol{a}-\boldsymbol{b}$ для всіх векторів $\boldsymbol{b}\in U$. Якщо підпростір U породжений векторами $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$, то задача рівносильна такому вибору коефіцієнтів $\alpha_1, \dots, \alpha_k$, при якому вектор

$$\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}-\sum_{i=1}^{k}\alpha_i\boldsymbol{u}_i$$

має найменшу довжину.

Знайдемо цю довжину, розклавши вектор $\boldsymbol{a}$ у суму $\boldsymbol{a}=\boldsymbol{u}+\boldsymbol{v}$, де $\boldsymbol{u}\in U$, $\boldsymbol{v}\in U^{\perp}$. Очевидно, що $\boldsymbol{u}-\boldsymbol{b}\in U$. З рівності $\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{u}-\boldsymbol{b}+\boldsymbol{v}$ на підставі теореми Піфагора маємо:

$$|\,\boldsymbol{a}-\boldsymbol{b}\,|^2=|\,\boldsymbol{u}-\boldsymbol{b}\,|^2+|\,\boldsymbol{v}\,|^2 .$$

Звідси випливає, що $|\,\boldsymbol{a}-\boldsymbol{b}\,|\geq|\,\boldsymbol{v}\,|$, причому рівність можлива тоді і тільки тоді, коли $\boldsymbol{b}=\boldsymbol{u}$.

Отже, серед усіх векторів $\boldsymbol{b}\in U$ ортогональна проекція вектора $\boldsymbol{a}$ на U найменше відхиляється від $\boldsymbol{a}$. Тому відстань між вектором $\boldsymbol{a}$ і підпростором U дорівнює $|\,\boldsymbol{a}-\boldsymbol{u}\,|=|\,\boldsymbol{v}\,|$, тобто довжині перпендикуляра, опущеного з $\boldsymbol{a}$ на U. Геометричні побудови для тривимірного простору зображені на рис. 9.

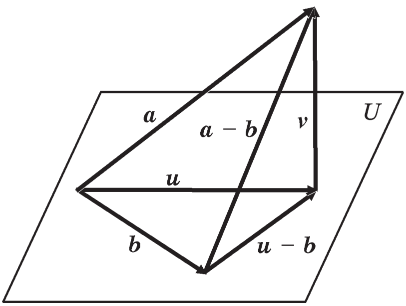

Рис. 9

Відстань між множинами Π і Σ векторів одного і того ж простору визначається як найменша довжина вектора $\boldsymbol{a}-\boldsymbol{b}$, де $\boldsymbol{a}\in\Pi$, $\boldsymbol{b}\in\Sigma$. Нехай, зокрема, множини Π і Σ є площинами, початкові точки яких мають радіус-вектори $\boldsymbol{f}_0$ і $\boldsymbol{g}_0$, а спрямовуючі підпростори U і V породжуються векторами $\boldsymbol{f}_1, \dots, \boldsymbol{f}_k$ і $\boldsymbol{g}_1, \dots, \boldsymbol{g}_l$ відповідно. Тоді довільні вектори $\boldsymbol{a}\in\Pi$, $\boldsymbol{b}\in\Sigma$ записуються у вигляді:

$$\boldsymbol{a} = \boldsymbol{f}_0 + \sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i, \quad \boldsymbol{b} = \boldsymbol{g}_0 + \sum_{j=1}^{l} \beta_j \boldsymbol{g}_j.$$

Відстанню між П і Σ є найменша довжина вектора

$$\boldsymbol{a} - \boldsymbol{b} = \boldsymbol{f}_0 - \boldsymbol{g}_0 + \sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i - \sum_{j=1}^{l} \beta_j \boldsymbol{g}_j,$$

яка може бути отримана належним вибором коефіцієнтів $\alpha_1, \dots, \alpha_k$ і $\beta_1, \dots, \beta_l$. Оскільки вектор $\sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i - \sum_{j=1}^{l} \beta_j \boldsymbol{g}_j$ є довільним вектором підпростору $U + V$, то це означає, що *відстань між площинами* П і Σ дорівнює відстані між вектором $\boldsymbol{f}_0 - \boldsymbol{g}_0$ і підпростором $U + V$.

В окремому випадку, коли один з підпросторів є нульовим, можна стверджувати, що ві*дстань між вектором* $\boldsymbol{a}$ *і площиною* (3.3.2) дорівнює відстані між вектором $\boldsymbol{a} - \boldsymbol{f}_0$ і спрямовуючим підпростором площини.

5.2.4. Нормальний вектор площини. *Нормальним вектором площини* П називається вектор $\boldsymbol{n}$ цієї площини, ортогональний до її спрямовуючого підпростору *U*. Знайдемо нормальний вектор площини

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i.$$

Оскільки $\boldsymbol{n} \in \Pi$, то

$$\boldsymbol{n} = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i. \tag{10}$$

Параметри $\alpha_1, \dots, \alpha_m$ знаходимо з умов $\boldsymbol{n}^T \boldsymbol{f}_j = 0, \ j = \overline{1, m}$, які рівносильні системі рівнянь

$$\sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i^T \boldsymbol{f}_j + \boldsymbol{f}_0^T \boldsymbol{f}_j = 0, \quad j = \overline{1, m}. \tag{11}$$

Ця система, як і система(2), має єдиний розв’язок. Обчисливши його, одержимо $\boldsymbol{n}$ за допомогою формули (10).

На підставі висновку з теореми 3.3.1 вектор $\boldsymbol{n}$ можна взяти за початкову точку площини. Тому її векторне рівняння може бути таким:

$$\boldsymbol{x} = \boldsymbol{n} + \boldsymbol{u}, \quad \boldsymbol{u} \in U.$$

Ця рівність є розкладом довільного вектора $\boldsymbol{x} \in \Pi$ на ортогональну проекцію $\boldsymbol{u}$ і перпендикуляр $\boldsymbol{n}$. Отже, нормальний вектор площини збігається з перпендикуляром, опущеним з будь-якого вектора площини на її спрямовуючий підпростір. Виходячи з властивості перпендикуляра, опущеного з вектора на підпростір, можна стверджувати, що нормальний вектор має найменшу довжину серед усіх векторів площини.

Розглянемо питання про ортогональну проекцію вектора $\boldsymbol{a}$ на площину Π. Якщо U — спрямовуючий підпростір площини, то, як зазначалося вище, $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in U$, $\boldsymbol{v} \in U^{\perp}$. Візьмемо вектор $\boldsymbol{y} \in \Pi$ такий, що $\boldsymbol{y} = \boldsymbol{n} + \boldsymbol{u}$, де $\boldsymbol{n}$ — нормальний вектор площини. Тепер маємо рівність $\boldsymbol{a} - \boldsymbol{y} = \boldsymbol{v} - \boldsymbol{n}$, звідки

$$\boldsymbol{a} = \boldsymbol{y} + (\boldsymbol{v} - \boldsymbol{n}).$$

Вектори $\boldsymbol{y}$, $\boldsymbol{v}$ і $\boldsymbol{n}$ визначаються однозначно, тому остання формула вказує на можливість представлення єдиним способом довільного вектора $\boldsymbol{a}$ простору у вигляді суми

$$\boldsymbol{a} = \boldsymbol{y} + \boldsymbol{z}, \tag{12}$$

де вектор $\boldsymbol{y}$ належить площині Π, а вектор $\boldsymbol{z} = \boldsymbol{v} - \boldsymbol{n}$ є ортогональним до спрямовуючого підпростору U. Вектор $\boldsymbol{y}$ у розкладі (12) називається *ортогональною проекцією вектора* $\boldsymbol{a}$ *на площину* Π, а $\boldsymbol{z}$ — *перпендикуляром*, опущеним з $\boldsymbol{a}$ на Π.

Розглянемо більш докладно питання про нормальний вектор площини і розклад (12) у випадку, коли площина є прямою лінією $\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1$. При цьому система рівнянь (11) зводиться до однієї умови $\alpha_1 \boldsymbol{f}_1^T \boldsymbol{f}_1 + \boldsymbol{f}_0^T \boldsymbol{f}_1 = 0$ і нормальний вектор, обчислений за допомогою формули (10), дорівнює

$$\boldsymbol{n} = \boldsymbol{f}_0 - \frac{\boldsymbol{f}_0^T \boldsymbol{f}_1}{|\boldsymbol{f}_1|^2} \boldsymbol{f}_1.$$

Ортогональна проекція $\boldsymbol{a}$ на спрямовуючий підпростір прямої лінії на підставі формули (6) є такою:

$$\boldsymbol{u} = -\frac{1}{\boldsymbol{f}_1^T \boldsymbol{f}_1}\begin{vmatrix} \boldsymbol{f}_1^T \boldsymbol{f}_1 & \boldsymbol{f}_1 \\ \boldsymbol{a}^T \boldsymbol{f}_1 & 0 \end{vmatrix} = \frac{\boldsymbol{a}^T \boldsymbol{f}_1}{|\boldsymbol{f}_1|^2}\boldsymbol{f}_1.$$

Звідси знаходимо перпендикуляр, опущений з $\boldsymbol{a}$ на спрямовуючий підпростір:

$$\boldsymbol{v} = \boldsymbol{a} - \boldsymbol{u} = \boldsymbol{a} - \frac{\boldsymbol{a}^T \boldsymbol{f}_1}{|\boldsymbol{f}_1|^2}\boldsymbol{f}_1.$$

Тепер перпендикуляр, опущений з $\boldsymbol{a}$ на пряму лінію, дорівнює

$$\boldsymbol{z} = \boldsymbol{v} - \boldsymbol{n} = \boldsymbol{a} - \boldsymbol{f}_0 - \frac{(\boldsymbol{a} - \boldsymbol{f}_0)^T \boldsymbol{f}_1}{|\boldsymbol{f}_1|^2}\boldsymbol{f}_1, \tag{13}$$

а ортогональна проекція $\boldsymbol{a}$ на пряму лінію є такою:

$$\boldsymbol{y} = \boldsymbol{a} - \boldsymbol{z} = \boldsymbol{f}_0 + \frac{(\boldsymbol{a} - \boldsymbol{f}_0)^T \boldsymbol{f}_1}{|\boldsymbol{f}_1|^2}\boldsymbol{f}_1. \tag{14}$$

П р и к л а д 2. Знайдемо ортогональну проекцію вектора

$$\boldsymbol{a} = (-1 \quad 2 \quad 3)^T$$

на пряму лінію

$$\frac{x_1 - 2}{1} = \frac{x_2 + 1}{1} = \frac{x_3}{-1},$$

а також перпендикуляр, опущений з $\boldsymbol{a}$ на неї.

Представимо канонічні рівняння прямої лінії у векторній формі:

$$\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1, \quad \boldsymbol{f}_0 = (2 \quad -1 \quad 0)^T, \quad \boldsymbol{f}_1 = (1 \quad 1 \quad -1)^T.$$

За допомогою формул (14) і (13) знаходимо шукані вектори:

$$\boldsymbol{y} = (1 \quad -2 \quad 1)^T, \quad \boldsymbol{z} = \boldsymbol{a} - \boldsymbol{y} = (-2 \quad 4 \quad 2)^T. \ \triangle$$

Знайдемо нормальний вектор і розклад (12) для гіперплощини П, заданої рівнянням $\boldsymbol{c}^T \boldsymbol{x} = b.$ Нехай $\boldsymbol{x}_0$ — початкова точка гіперплощини, тобто $\boldsymbol{c}^T \boldsymbol{x}_0 = b.$ Тоді рівняння гіперплощини можна подати у вигляді:

$$\boldsymbol{c}^T (\boldsymbol{x} - \boldsymbol{x}_0) = 0. \tag{15}$$

На підставі теореми 3.3.1 вектор $\boldsymbol{x} - \boldsymbol{x}_0$ належить спрямовуючому підпростору U. Остання рівність показує, що $\boldsymbol{c} \in U^{\perp}$. Але і $\boldsymbol{n} \in U^{\perp}$, тому $\boldsymbol{n} = \beta \boldsymbol{c}$, де $\beta \neq 0$ — деяке число. Оскільки $\boldsymbol{n} \in \Pi$, то $\boldsymbol{c}^T \boldsymbol{n} = b$, або $\beta |\boldsymbol{c}|^2 = b$. Отже,

$$\boldsymbol{n} = \frac{b}{|\boldsymbol{c}|^2}\boldsymbol{c}.$$

Розкладемо довільний вектор $\boldsymbol{a}$ у суму $\boldsymbol{a} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u} \in U$, $\boldsymbol{v} \in U^{\perp}$. Підпростір U визначається рівнянням $\boldsymbol{c}^T \boldsymbol{x} = 0$. Ортогональна проекція вектора $\boldsymbol{a}$ на U є вектором, спільним для гіперплощини $\boldsymbol{c}^T \boldsymbol{x} = 0$ і прямої лінії, що проходить через $\boldsymbol{a}$ паралельно до $\boldsymbol{c}$. Векторне рівняння цієї прямої лінії має вигляд:

$$\boldsymbol{x} = \boldsymbol{a} + \alpha_1 \boldsymbol{c}.$$

Підставивши цей вираз для $\boldsymbol{x}$ в рівняння $\boldsymbol{c}^T \boldsymbol{x} = 0$, одержимо відповідне значення параметра α_1:

$$\alpha_1 = -\frac{\boldsymbol{c}^T \boldsymbol{a}}{|\boldsymbol{c}|^2}.$$

Тепер маємо шукану ортогональну проекцію $\boldsymbol{a}$ на U:

$$\boldsymbol{u} = \boldsymbol{a} - \frac{\boldsymbol{c}^T \boldsymbol{a}}{|\boldsymbol{c}|^2}\boldsymbol{c}, \tag{16}$$

а також перпендикуляр, опущений з $\boldsymbol{a}$ на U:

$$\boldsymbol{v} = \boldsymbol{a} - \boldsymbol{u} = \frac{\boldsymbol{c}^T \boldsymbol{a}}{|\boldsymbol{c}|^2}\boldsymbol{c}. \tag{17}$$

Далі знаходимо перпендикуляр $\boldsymbol{z}$, опущений з $\boldsymbol{a}$ на Π:

$$\boldsymbol{z} = \boldsymbol{v} - \boldsymbol{n} = \frac{\boldsymbol{c}^T \boldsymbol{a} - b}{|\boldsymbol{c}|^2}\boldsymbol{c}. \tag{18}$$

При цьому ортогональна проекція вектора $\boldsymbol{a}$ на Π дорівнює

$$\boldsymbol{y} = \boldsymbol{a} - \boldsymbol{z} = \boldsymbol{a} - \frac{\boldsymbol{c}^T \boldsymbol{a} - b}{|\boldsymbol{c}|^2}\boldsymbol{c}.$$

5.2.5. Кути. *Кутом* $\varphi\{\boldsymbol{a},\ \boldsymbol{b}\}$ *між* ненульовими *векторами* $\boldsymbol{a}$, $\boldsymbol{b}$ евклідового простору називається кут, який визначається співвідношеннями:

$$\cos\varphi\,\{\boldsymbol{a},\ \boldsymbol{b}\} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{|\,\boldsymbol{a}\,||\,\boldsymbol{b}\,|},\quad 0 \le \varphi\{\boldsymbol{a},\ \boldsymbol{b}\} \le \pi. \tag{19}$$

Це визначення є коректним, бо з нерівності Коші — Шварца — Буняковського випливає, що $|\cos\varphi\{\boldsymbol{a},\ \boldsymbol{b}\}| \le 1$. Якщо серед векторів $\boldsymbol{a}$ і $\boldsymbol{b}$ є принаймні один нульовий, то кут між такими векторами вважається невизначеним.

Кутом $\varphi\,\{\boldsymbol{a},\ U\}$ *між* ненульовим *вектором* $\boldsymbol{a}$ *і* ненульовим *підпростором* U називається найменший з кутів між вектором $\boldsymbol{a}$ і векторами $\boldsymbol{b} \in U$. З урахуванням нерівності Коші — Шварца — Буняковського маємо:

$$\cos\varphi\,\{\boldsymbol{a},\ \boldsymbol{b}\} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{|\,\boldsymbol{a}\,||\,\boldsymbol{b}\,|} = \frac{(\boldsymbol{u}+\boldsymbol{v})^T\boldsymbol{b}}{|\,\boldsymbol{a}\,||\,\boldsymbol{b}\,|} = \frac{\boldsymbol{u}^T\boldsymbol{b}}{|\,\boldsymbol{a}\,||\,\boldsymbol{b}\,|} \le \frac{|\,\boldsymbol{u}\,|}{|\,\boldsymbol{a}\,|}.$$

Остання нерівність перетворюється в рівність тоді і тільки тоді, коли вектор $\boldsymbol{b}$ утворює нульовий кут з вектором $\boldsymbol{u}$. Отже, кут між вектором $\boldsymbol{a}$ і підпростором U збігається з кутом між вектором $\boldsymbol{a}$ і його ортогональною проекцією на U.

Кут $\varphi\,\{\boldsymbol{a},\ \Pi\}$ *між вектором* $\boldsymbol{a}$ *і площиною* Π вважається рівним куту $\varphi\,\{\boldsymbol{a},\ U\}$ між $\boldsymbol{a}$ і спрямовуючим підпростором U площини Π.

Кутом $\varphi\{U,\ V\}$ *між* ненульомими *підпросторами* U і V при $\dim(U \cap V) = 0$ називається найменший з кутів між довільним вектором одного підпростору і його ортогональною проекцією на інший. Нехай $\boldsymbol{a}$ — довільний вектор з U і $\boldsymbol{v}_0$ — його ортогональна проекція на V. Кут $\varphi\{U,\ V\}$ у відповідності до його визначення дорівнює найменшому з кутів $\varphi\{\boldsymbol{a}, \boldsymbol{v}_0\}$. Але найменший з кутів $\varphi\{\boldsymbol{a}, \boldsymbol{v}_0\}$ — це кут між $\boldsymbol{v}_0$ і підпростором U, тобто $\varphi\{\boldsymbol{v}_0,\ U\}$. Позначимо через $\boldsymbol{u}_0$ ортогональну проекцію вектора $\boldsymbol{v}_0$ на підпростір U. Тоді у відповідності до визначення кута між вектором і підпростором повинна виконуватися рівність $\varphi\{\boldsymbol{v}_0, U\} = \varphi\{\boldsymbol{v}_0, \boldsymbol{u}_0\}$.

Таким чином,

$$\varphi\{U,V\}=\varphi\{\boldsymbol{u}_0,\boldsymbol{v}_0\},\tag{20}$$

де $\boldsymbol{v}_0$ — ортогональна проекція довільного вектора $\boldsymbol{a}\in U$ на V і $\boldsymbol{u}_0$ — ортогональна проекція вектора $\boldsymbol{v}_0$ на U.

Нехай $\dim(U\bigcap V)\neq 0,$ причому переріз $U\bigcap V$ не збігається з U або V. *Кутом між підпросторами* U і V в цьому випадку називається кут між підпросторами $U'\subset U$ і $V'\subset V,$ які є ортогональними доповненнями до $U\bigcap V$, тобто $\varphi\{U,V\}=\varphi\{U',V'\}$. У випадку, коли $U\subseteq V$ (або $V\subseteq U$), вважається, що $\varphi\{U,V\}=0$.

Покажемо, що при такому визначенні кута $\varphi\{U,V\}$ його обчислення зводиться до розглянутого вище випадку, коли переріз підпросторів є нульовим підпростором. Щоб пересвідчитися в цьому, врахуємо співвідношення

$$U'=U\bigcap(U\bigcap V)^{\perp},\quad V'=V\bigcap(U\bigcap V)^{\perp},\tag{21}$$

які визначають підпростори U' і V'. Звідси впливає, що

$$U'\bigcap V'=U\bigcap(U\bigcap V)^{\perp}\bigcap V=(U\bigcap V)\bigcap(U\bigcap V)^{\perp}.$$

Оскільки єдиним спільним вектором для будь-якого підпростору і його ортогонального доповнення є нульовий вектор, то з останньої формули випливає рівність $\dim(U'\bigcap V')=0,$ яку треба довести.

Зазначимо, що при розгляді задачі про кут між підпросторами U і V можна обмежитися випадком, коли $\dim U>1$ і $\dim V>1$, бо, якщо принаймні один з підпросторів є одновимірним, ми будемо мати справу із задачею про кут між вектором і підпростором.

Кутом $\varphi\{\Pi,\ \Sigma\}$ *між площинами* Π і Σ називається кут між їх спрямовуючими підпросторами.

5.2.6. Об'єми. Якщо вважати, що $\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_m,\boldsymbol{a}$ — радіус-вектори точок афінного простору і на них побудований $(m+1)$-вимірний паралелепіпед, то величину $h=|\boldsymbol{v}|$ можна тлумачити як висоту цього паралелепіпеда, опущену з кінця ребра $\boldsymbol{a}$ на основу, утворену векторами $\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m$.

Позначимо $G(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ через G_m і $\sqrt{\dfrac{G_{m+1}}{G_m}}$ — через h_m.

З рівності $G(\boldsymbol{u}_1) = |\boldsymbol{u}_1|^2$ маємо:

$$\sqrt{G_1} = |\boldsymbol{u}_1| = V_1.$$

Тут через V_1 позначений «об'єм» одновимірного паралелепіпеда, утвореного вектором $\boldsymbol{u}_1$, тобто його довжина.

Далі за допомогою формули (7) знаходимо:

$$\sqrt{G_2} = \sqrt{G_1}h_1 = V_1 h_1 = V_2,$$

де V_2 — «об'єм» двовимірного паралелепіпеда, утвореного векторами $\boldsymbol{u}_1$ і $\boldsymbol{u}_2$, тобто площа паралелограма, побудованого на цих векторах. Аналогічно знаходимо, що

$$\sqrt{G_3} = \sqrt{G_2}h_2 = V_2 h_2 = V_3,$$

де V_3 — об'єм паралелепіпеда, побудованого на векторах $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

Продовжуючи далі, одержимо індуктивне визначення *об'єму m-вимірного паралелепіпеда*, побудованого на векторах $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$:

$$\sqrt{G_m} = \sqrt{G_{m-1}}h_{m-1} = V_{m-1}h_{m-1} = V_m.$$

Зокрема, якщо m збігається з вимірністю простору E_n, то

$$V_n = \sqrt{G_n} = \sqrt{\det \boldsymbol{\Gamma}_n} = \sqrt{\det(\boldsymbol{A}^T \boldsymbol{A})} = \sqrt{(\det \boldsymbol{A}^T) \cdot \det \boldsymbol{A}} = |\det \boldsymbol{A}|.$$

Тут $\boldsymbol{A}$ — матриця, стовпці якої складаються з координат n векторів простору E_n.

П р и к л а д 3. Знайдемо об'єм паралелепіпеда, побудованого на векторах:

1) $\boldsymbol{u}_1 = (1 \;\; 1 \;\; 1)^T, \quad \boldsymbol{u}_2 = (2 \;\; 3 \;\; 4)^T$;

2) $\boldsymbol{u}_1 = (1 \;\; 1 \;\; 0)^T, \quad \boldsymbol{u}_2 = (0 \;\; 0 \;\; 1)^T, \quad \boldsymbol{u}_3 = (1 \;\; 2 \;\; 2)^T$.

У першому випадку шуканий об'єм є площею паралелограма, побудованого на векторах $\boldsymbol{u}_1$ і $\boldsymbol{u}_2$. Він дорівнює

$$V_2 = \sqrt{G(\boldsymbol{u}_1,\ \boldsymbol{u}_2)}.$$

Обчисливши визначник Грама:

$$G(\boldsymbol{u}_1,\ \boldsymbol{u}_2) = \begin{vmatrix} \boldsymbol{u}_1^T\boldsymbol{u}_1 & \boldsymbol{u}_1^T\boldsymbol{u}_2 \\ \boldsymbol{u}_2^T\boldsymbol{u}_1 & \boldsymbol{u}_2^T\boldsymbol{u}_2 \end{vmatrix} = \begin{vmatrix} 3 & 9 \\ 9 & 29 \end{vmatrix} = 6,$$

одержимо, що $V_2 = \sqrt{6}.$ $\triangle$

У другому випадку об'єм паралелепіпеда дорівнює

$$V_3 = \sqrt{G(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)} = \sqrt{\begin{vmatrix} \boldsymbol{u}_1{}^T\boldsymbol{u}_1 & \boldsymbol{u}_1{}^T\boldsymbol{u}_2 & \boldsymbol{u}_1{}^T\boldsymbol{u}_3 \\ \boldsymbol{u}_2{}^T\boldsymbol{u}_1 & \boldsymbol{u}_2{}^T\boldsymbol{u}_2 & \boldsymbol{u}_2{}^T\boldsymbol{u}_3 \\ \boldsymbol{u}_3{}^T\boldsymbol{u}_1 & \boldsymbol{u}_3{}^T\boldsymbol{u}_2 & \boldsymbol{u}_3{}^T\boldsymbol{u}_3 \end{vmatrix}} = \sqrt{\begin{vmatrix} 2 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 9 \end{vmatrix}} = 1.\ \triangle$$

Наприкінці розглянемо нерівність Адамара.

Теорема 5.2.2. *Для довільної системи ненульових векторів* $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m$ *має місце нерівність Адамара*

$$G(\boldsymbol{u}_1,\ \boldsymbol{u}_2, \dots, \boldsymbol{u}_m) \le G(\boldsymbol{u}_1)G(\boldsymbol{u}_2)\dots G(\boldsymbol{u}_m), \tag{22}$$

причому рівність можлива тоді і тільки тоді, коли система векторів $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m$ *є ортогональною.*

Д о в е д е н н я. Нехай, як і раніше, $\boldsymbol{u}$ — ортогональна проекція вектора $\boldsymbol{a} \in E_n$ на підпростір U і $\boldsymbol{v}$ — перпендикуляр, опущений з $\boldsymbol{a}$ на U. З теореми Піфагора випливає очевидна рівність $|\boldsymbol{a}|^2 = |\boldsymbol{u}|^2 + |\boldsymbol{v}|^2$, тому

$$G(\boldsymbol{a}) = \boldsymbol{a}^T\boldsymbol{a} = |\boldsymbol{u}|^2 + |\boldsymbol{v}|^2 \ge |\boldsymbol{v}|^2 = h^2.$$

Звідси з урахуванням формули (7) маємо:

$$G(\boldsymbol{u}_1,\ \boldsymbol{u}_2, \dots, \boldsymbol{u}_m,\ \boldsymbol{a}) \le G(\boldsymbol{u}_1,\ \boldsymbol{u}_2, \dots, \boldsymbol{u}_m)G(\boldsymbol{a}),$$

причому рівність можлива тоді і тільки тоді, коли $\boldsymbol{u} = \boldsymbol{0}$, тобто вектор $\boldsymbol{a}$ є ортогональним до векторів $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$.

Після цього неважко зробити висновок про виконання нерівності (22), з якої випливає, що об'єм паралелепіпеда не перевищує добутку довжин його ребер і дорівнює цьому добутку лише тоді, коли паралелепіпед є прямокутним. $\square$

5.3. Розв'язання задач аналітичної геометрії

5.3.1. Обчислення відстаней. Скористаємося одержаними в попередньому розділі результатами для обчислення відстаней. Стосовно задач про відстані можливі два підходи при їх розв'язанні. Перший заснований на використанні формул (5.2.1) — (5.2.3) для знаходження вектора $\boldsymbol{v}$ з подальшим обчисленням шуканої відстані $h = |\boldsymbol{v}|$. Інший підхід передбачає використання готової формули (5.2.7) для h.За трудомісткістю обидва підходи мало відрізняються при невеликих m. Виходячи з цієї обставини, ми користуємося для простоти викладення формулою (5.2.7). Проте при досить великих m зростають труднощі, пов'язані з обчисленням визначників Грама. Тому в цьому випадку має сенс звернутися до першої схеми розв'язання задачі.

Почнемо з обчислення відстані між точкою з радіус-вектором $\boldsymbol{a}$ і площиною

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{m} \alpha_i \boldsymbol{f}_i .$$

Ця відстань визначається як відстань між вектором $\boldsymbol{a} - \boldsymbol{f}_0$ і підпростором, натягнутим на вектори $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$. На підставі формули (5.2.7) вона дорівнює

$$h = \sqrt{\frac{G(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m, \boldsymbol{a} - \boldsymbol{f}_0)}{G(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m)}} . \tag{1}$$

При $m = 1$ маємо відстань між вектором і прямою лінією, а при $m = n - 1$ — між вектором і гіперплощиною. В останньому випадку користуватися цією формулою незручно, якщо гіперплощина представлена рівнянням $\boldsymbol{c}^T \boldsymbol{x} = b$. Покажемо, як розв'язується зазначена задача в цьому випадку.

Спочатку зазначимо, що відстань h між вектором $\boldsymbol{a}$ і спрямовуючим підпростором U дорівнює довжині перпендикуляра (5.2.17):

$$h = \frac{|\boldsymbol{c}^T \boldsymbol{a}|}{|\boldsymbol{c}|}. \tag{2}$$

У загальному випадку гіперплощини $\boldsymbol{c}^T \boldsymbol{x} = b$ відстань h між $\boldsymbol{a}$ і цією гіперплощинною дорівнює відстані між вектором $\boldsymbol{a} - \boldsymbol{x}_0$ і підпростором U. Тут $\boldsymbol{x}_0$ — початкова точка гіперплощини:

$$\boldsymbol{c}^T \boldsymbol{x}_0 = b. \qquad (3)$$

За допомогою формул (2) і (3) знаходимо:

$$h = \frac{\left|\boldsymbol{c}^T (\boldsymbol{a} - \boldsymbol{x}_0)\right|}{|\, \boldsymbol{c} \,|} = \frac{\left|\boldsymbol{c}^T \boldsymbol{a} - b\right|}{|\, \boldsymbol{c} \,|}.$$

Цей же результат можна одержати з формули (5.2.18), врахувавши, що $h = |\, \boldsymbol{z} \,|$.

Далі будемо розглядати відстані між площинами. Почнемо з обчислення відстані між прямими лініями

$$\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1, \qquad \boldsymbol{x} = \boldsymbol{g}_0 + \beta_1 \boldsymbol{g}_1, \qquad (4)$$

вважаючи, що вони не є паралельними. Шукана відстань h дорівнює відстані між вектором $\boldsymbol{f}_0 - \boldsymbol{g}_0$ і підпростором, що є сумою спрямовуючих підпросторів прямих ліній. Оскільки вектори $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$ є лінійно незалежними, то вони породжують указаний підпростір. На підставі формули (5.2.7) маємо:

$$h = \sqrt{\frac{G(\boldsymbol{f}_1,\ \boldsymbol{g}_1,\ \boldsymbol{f}_0 - \boldsymbol{g}_0)}{G(\boldsymbol{f}_1,\ \boldsymbol{g}_1)}}. \qquad (5)$$

Якщо прямі лінії перетинаються, то відстань між ними дорівнює нулю. З останньої формули випливає, що умовою перетинання двох прямих ліній є рівність

$$G(\boldsymbol{f}_1,\ \boldsymbol{g}_1,\ \boldsymbol{f}_0 - \boldsymbol{g}_0) = 0.$$

Це є умова лінійної залежності векторів $\boldsymbol{f}_1$, $\boldsymbol{g}_1$ і $\boldsymbol{f}_0 - \boldsymbol{g}_0$. У теоремі 3.6.1 ця умова була одержана іншим шляхом.

При визначенні відстані між площинами

$$\boldsymbol{x} = \boldsymbol{f}_0 + \sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i, \qquad \boldsymbol{x} = \boldsymbol{g}_0 + \sum_{j=1}^{l} \beta_j \boldsymbol{g}_j$$

треба знайти базис суми їх спрямовуючих підпросторів U, V. Це питання було висвітлене в розділі 3.5. Нехай $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ — базис суми $U+V$. Відстань між площинами на підставі формули (5.2.7) дорівнює

$$h = \sqrt{\frac{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{f}_0 - \boldsymbol{g}_0)}{G(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)}}.$$

Припустимо, що треба знайти відстань між паралельними площинами. У відповідності до визначення паралельності площин повинно бути $U \subseteq V$ (або $V \subseteq U$). Нехай $V \subseteq U$ тоді сума $U+V$ збігається з U і ми повинні знаходити відстань між вектором $\boldsymbol{f}_0 - \boldsymbol{g}_0$ і підпростором U.

У випадку паралельності прямих ліній (4) вектори $\boldsymbol{f}_1$ і $\boldsymbol{g}_1$ є колінеарними. Відстанню між прямими лініями є відстань між вектором $\boldsymbol{f}_0 - \boldsymbol{g}_0$ і підпростором, породженим вектором $\boldsymbol{f}_1$, тобто

$$h = \sqrt{\frac{G(\boldsymbol{f}_1, \boldsymbol{f}_0 - \boldsymbol{g}_0)}{G(\boldsymbol{f}_1)}}. \tag{6}$$

Розглянемо випадок, коли розшукується відстань між гіперплощинами $\boldsymbol{c}_1^T \boldsymbol{x} = b_1$, $\boldsymbol{c}_2^T \boldsymbol{x} = b_2$. Припустимо, що вектори $\boldsymbol{c}_1$ і $\boldsymbol{c}_2$ є колінеарними. Тоді гіперплощини будуть паралельними і відстань між ними можна знайти як відстань між деякою точкою однієї гіперплощини і іншою гіперплощиною. Якщо вектори $\boldsymbol{c}_1$ і $\boldsymbol{c}_2$ не є колінеарними, то гіперплощини перетинаються і тому відстань між ними є нульовою.

5.3.2. Обчислення кутів. Обчислимо кут φ між вектором $\boldsymbol{a}$ і гіперплощиною $\boldsymbol{c}^T \boldsymbol{x} = b$. Він дорівнює куту між вектором $\boldsymbol{a}$ і його ортогональною проекцією $\boldsymbol{u}$ на спрямовуючий підпростір гіперплощини. Ця проекція визначається формулою (5.2.16). Отже,

$$\cos\varphi = \frac{\boldsymbol{a}^T \boldsymbol{u}}{|\boldsymbol{a}|\,|\boldsymbol{u}|}.$$

Підставляючи сюди $\boldsymbol{u}$ з (5.2.16), одержимо після нескладних перетворень:

$$\cos\varphi = \frac{\sqrt{|\boldsymbol{a}|^2 |\boldsymbol{c}|^2 - (\boldsymbol{c}^T \boldsymbol{a})^2}}{|\boldsymbol{a}|\,|\boldsymbol{c}|},$$

або

$$\sin\varphi = \frac{|\boldsymbol{c}^T\boldsymbol{a}|}{|\boldsymbol{a}||\boldsymbol{c}|}.$$

Останню формулу можна отримати також, скориставшись тією обставиною, що сума кутів між вектором і його ортогональними проекціями на деякий підпростір і його ортогональне доповнення дорівнює $\pi/2$.

Знайдемо кут ψ між гіперплощинами $\boldsymbol{c}_1^T\boldsymbol{x} = b_1$, $\boldsymbol{c}_2^T\boldsymbol{x} = b_2$, де $\boldsymbol{c}_1$ і $\boldsymbol{c}_2$ — лінійно незалежні вектори. Кут між гіперплощинами визначається як кут між їх спрямовуючими підпросторами *U, V*. Оскільки гіперплощини перетинаються, то $\dim(U \cap V) \neq 0$. У цьому випадку, як зазначалося в розділі 5.2, кутом між *U* і *V* називається кут між підпросторами U' і V', які вміщуються в *U* та *V* і є ортогональними доповненнями до $U \cap V$.

Величина $\dim(U+V)$ може бути рівною $n-1$ в разі паралельності гіперплощин або бути більшою ніж $n-1$, якщо гіперплощини не є паралельними. Ми маємо справу з останнім випадком, тому $\dim(U+V) = n$. З формули Грассмана знаходимо, що $\dim(U \cap V) = n-2$. Тепер з формули (5.1.1) випливає, що $\dim(U \cap V)^{\perp} = 2$. Оскільки $\boldsymbol{c}_1 \perp U$, то одночасно $\boldsymbol{c}_1 \perp U \cap V$. Аналогічно можна стверджувати, що $\boldsymbol{c}_2 \perp U \cap V$. Таким чином, $\boldsymbol{c}_1,\ \boldsymbol{c}_2 \in (U \cap V)^{\perp}$.

Останнє співвідношення показує, що в підпросторі $(U \cap V)^{\perp}$ є вектори, які не належать до *U* і *V*. Тому вимірності підпросторів $U + (U \cap V)^{\perp}$, $V + (U \cap V)^{\perp}$ повинні бути більшими ніж вимірності *U* і *V* .Це означає, що

$$\dim(U + (U \cap V)^{\perp}) = n, \quad \dim(V + (U \cap V)^{\perp}) = n.$$

Тепер для підпросторів U' і V' на підставі співвідношень (5.2.21) і формули Грассмана маємо:

$$\dim U' = 1, \quad \dim V' = 1.$$

Як відомо, у випадку тривимірного простору підпростори U' і V' є прямими лініями — сторонами лінійного кута шуканого двограного кута. Останні рівності показують, що і в просторі довільної вимірності ці підпростори є прямими лініями. Геометричні побудуви, вказані вище, зображені на рис. 10 для випадку тривимірного простору.

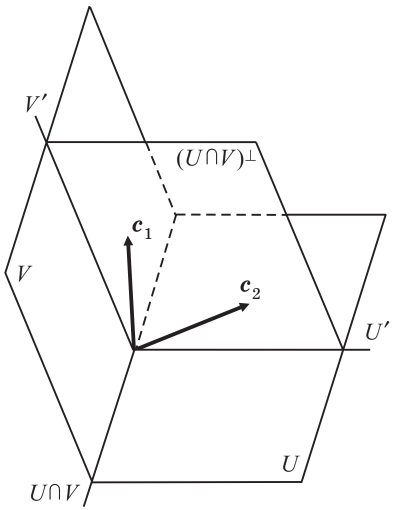

Рис. 10

Щоб визначити підпростори U' і V', скористаємося тією обставиною, що лінійно незалежні вектори $\boldsymbol{c}_1$ і $\boldsymbol{c}_2$, будучі належними до двовимірного підпростору $(U \bigcap V)^{\perp}$, утворюють базис цього підпростору. Це означає, що лінійна комбінація $\alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2$ є довільним вектором цього підпростору. Його ортогональні проекції на підпростори U і V є підпросторами U' і V'. Знайдемо ці проекції за допомогою формули (5.2.16), а потім знайдемо кут між ними.

Ортогональна проекція $\boldsymbol{u}_1'$ вектора $\alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2$ на підпростір U у відповідності до вказаної формули дорівнює

$$\boldsymbol{u}_1' = \alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2 - \frac{\boldsymbol{c}_1^T(\alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2)}{|\,\boldsymbol{c}_1\,|^2}\boldsymbol{c}_1 = -\frac{\alpha_2}{|\,\boldsymbol{c}_1\,|^2}((\boldsymbol{c}_1^T\boldsymbol{c}_2)\boldsymbol{c}_1 - |\,\boldsymbol{c}_1|^2\boldsymbol{c}_2).$$

Аналогічною є ортогональна проекція $\boldsymbol{u}_2'$ вектора $\alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2$ на підпростір V:

$$\boldsymbol{u}_2' = \alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2 - \frac{\boldsymbol{c}_2^T(\alpha_1\boldsymbol{c}_1 + \alpha_2\boldsymbol{c}_2)}{|\,\boldsymbol{c}_2|^2}\boldsymbol{c}_2 = \frac{\alpha_1}{|\,\boldsymbol{c}_2|^2}(|\,\boldsymbol{c}_2|^2\boldsymbol{c}_1 - (\boldsymbol{c}_1^T\boldsymbol{c}_2)\boldsymbol{c}_2).$$

Нескладні обчислення показують, що

$$\frac{(\boldsymbol{u}_1')^T\boldsymbol{u}_2'}{|\boldsymbol{u}_1'||\boldsymbol{u}_2'|} = \pm\frac{\boldsymbol{c}_1^T\boldsymbol{c}_2}{|\boldsymbol{c}_1||\boldsymbol{c}_2|}\,.$$

Це означає, що кут між гіперплощинами вимірюється кутом між векторами $\boldsymbol{c}_1$ і $\boldsymbol{c}_2$, перпендикулярними до гіперплощин. Ми одержали відомий з елементарної геометрії результат стосовно розглядуваної задачі для тривимірного простору.

5.3.3. Умови паралельності та перпендикулярності. З'ясуємо ці умови щодо гіперплощин і прямих ліній, а також стосовно гіперплощини і прямої лінії.

Гіперплощини $\boldsymbol{c}_1^T\boldsymbol{x} = b_1$, $\boldsymbol{c}_2^T\boldsymbol{x} = b_2$ є перпендикулярними за умови $\cos\psi = 0$, тобто при

$$\boldsymbol{c}_1^T\boldsymbol{c}_2 = 0.$$

Вище ми бачили, що гіперплощини паралельні за умови колінеарності векторів $\boldsymbol{c}_1$ і $\boldsymbol{c}_2$, тобто при

$$\boldsymbol{c}_1 = \alpha\boldsymbol{c}_2,$$

де α — деяке число, не рівне нулю.

Очевидно, що прямі лінії $\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1\boldsymbol{f}_1$, $\boldsymbol{x} = \boldsymbol{g}_0 + \beta_1\boldsymbol{g}_1$ паралельні при

$$\boldsymbol{f}_1 = \gamma\boldsymbol{g}_1$$

де γ — відмінне від нуля число, і перпендикулярні при

$$\boldsymbol{f}_1^T\boldsymbol{g}_1 = 0.$$

Пряма лінія $\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1\boldsymbol{f}_1$ перпендикулярна до гіперплощини $\boldsymbol{c}^T\boldsymbol{x} = b$, якщо вектори $\boldsymbol{f}_1$ і $\boldsymbol{c}$ є колінеарними, тобто

$$\boldsymbol{f}_1 = \delta\boldsymbol{c}, \tag{7}$$

де δ — число, відмінне від нуля. Пряма лінія є паралельною до гіперплощини за умови, що вектор $\boldsymbol{f}_1$ належить спрямовуючому підпростору гіперплощини. Тому він повинен бути ортогональним до $\boldsymbol{c}$. Отже, умова паралельності прямої лінії і гіперплощини є такою:

$$\boldsymbol{c}^T\boldsymbol{f}_1 = 0. \tag{8}$$

У цьому випадку відстань між прямою лінією і гіперплощиною знаходиться як відстань між точкою $\boldsymbol{f}_0$ і гіперплощиною.

5.3.4. Приклади. Розглянемо приклади застосування одержаних вище результатів.

П р и к л а д 1. Знайдемо відстань між вектором

$$\boldsymbol{a} = (-1 \quad 2 \quad 3)^T$$

і прямою лінією

$$\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1,$$

де

$$\boldsymbol{f}_0 = (2 \quad -1 \quad 0)^T, \quad \boldsymbol{f}_1 = (1 \quad 1 \quad -1)^T.$$

Тут $\boldsymbol{a} - \boldsymbol{f}_0 = (-3 \quad 3 \quad 3)^T$. У відповідності до формули (1) шукана відстань дорівнює

$$h = \sqrt{\frac{G(\boldsymbol{f}_1, \boldsymbol{a} - \boldsymbol{f}_0)}{G(\boldsymbol{f}_1)}} = \frac{\sqrt{G(\boldsymbol{f}_1, \boldsymbol{a} - \boldsymbol{f}_0)}}{|\boldsymbol{f}_1|}.$$

Далі маємо:

$$G(\boldsymbol{f}_1, \boldsymbol{a} - \boldsymbol{f}_0) = \begin{vmatrix} 3 & -3 \\ -3 & 27 \end{vmatrix} = 72, \quad |\boldsymbol{f}_1| = \sqrt{3}, \quad h = \frac{\sqrt{72}}{\sqrt{3}} = 2\sqrt{6}.$$

Задачу можна розв’язати інакше, якщо врахувати, що шукана відстань дорівнює довжині перпендикуляра, опущеного з $\boldsymbol{a}$ на пряму лінію. Цей перпендикуляр, обчислений у прикладі 2 з розділу 5.2, дорівнює $\boldsymbol{z} = (-2 \quad 4 \quad 2)^T$. Тому $h = |\boldsymbol{z}| = 2\sqrt{6}$, тобто маємо результат, одержаний вище. △

П р и к л а д 2. Знайдемо відстань між прямими лініями:

$$1)\ \frac{x_1 - 1}{1} = \frac{x_2 - 2}{1} = \frac{x_3 + 1}{1}, \quad \frac{x_1 - 2}{1} = \frac{x_2 - 1}{-2} = \frac{x_3 + 1}{0};$$

$$2)\ \frac{x_1 - 1}{1} = \frac{x_2 + 1}{-2} = \frac{x_3 + 2}{-2}, \quad \frac{x_1 - 2}{-1} = \frac{x_2 + 1}{2} = \frac{x_3 + 1}{2}.$$

У першому випадку вектори $\boldsymbol{f}_0$, $\boldsymbol{f}_1$, $\boldsymbol{g}_0$, $\boldsymbol{g}_1$ в рівняннях (4) є такими:

$$\boldsymbol{f}_0 = (1 \quad 2 \quad -1)^T, \quad \boldsymbol{f}_1 = (1 \quad 1 \quad 1)^T;$$

$$\boldsymbol{g}_0 = (2 \quad 1 \quad -1)^T, \quad \boldsymbol{g}_1 = (1 \quad -2 \quad 0)^T.$$

Очевидно, що прямі лінії не є паралельними, тому для обчислення відстані між ними треба взяти формулу (5). Урахувавши, що $\boldsymbol{f}_0 - \boldsymbol{g}_0 = (-1 \quad 1 \quad 0)^T$, одержимо:

$$G(\boldsymbol{f}_1, \boldsymbol{g}_1) = \begin{vmatrix} 3 & -1 \\ -1 & 5 \end{vmatrix} = 14, \quad G(\boldsymbol{f}_1, \boldsymbol{g}_1, \boldsymbol{f}_0 - \boldsymbol{g}_0) = \begin{vmatrix} 3 & -1 & 0 \\ -1 & 5 & -3 \\ 0 & -3 & 2 \end{vmatrix} = 1.$$

Тепер з формули (5) випливає, що відстань між прямими лініями дорівнює $h = \sqrt{1/14}$.

Для другої пари прямих ліній маємо:

$$\boldsymbol{f}_0 = (1 \quad -1 \quad -2)^T, \quad \boldsymbol{f}_1 = (1 \quad -2 \quad -2)^T;$$
$$\boldsymbol{g}_0 = (2 \quad -1 \quad -1)^T, \quad \boldsymbol{g}_1 = (-1 \quad 2 \quad 2)^T.$$

Тут прямі лінії є паралельними, тому скористаємося формулою (6). Оскільки

$$\boldsymbol{f}_0 - \boldsymbol{g}_0 = (-1 \quad 0 \quad -1)^T, \quad G(\boldsymbol{f}_1) = 9, \quad G(\boldsymbol{f}_1, \boldsymbol{f}_0 - \boldsymbol{g}_0) = \begin{vmatrix} 9 & 1 \\ 1 & 2 \end{vmatrix} = 17,$$

то шукана відстань дорівнює $h = \sqrt{17}/3$. $\triangle$

П р и к л а д 3. Знайдемо кут між підпросторами U і V, які породжуються відповідно векторами

$$\boldsymbol{u}_1 = (1 \quad 1 \quad 0 \quad 0)^T, \quad \boldsymbol{u}_2 = (0 \quad 0 \quad 1 \quad -1)^T;$$
$$\boldsymbol{v}_1 = (1 \quad 0 \quad 0 \quad 1)^T, \quad \boldsymbol{v}_2 = (0 \quad 1 \quad 1 \quad 0)^T.$$

Розмірковуючи, як при розгляді прикладу 1 з розділу 3.4, легко переконатися в тому, що $r(\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{v}_1 \quad \boldsymbol{v}_2) = 4$. Тому $\dim U = 2$, $\dim V = 2$, і ми повинні скористатися формулою (5.2.20).

Візьмемо вектор $\boldsymbol{u}_1$ і обчислимо його ортогональну проекцію $\boldsymbol{v}_0$ на підпростір V, скориставшись схемою розв’язання цієї задачі, викладеною при розгляді прикладу 1 з розділу 5.2. Вектор $\boldsymbol{v}_0$ дорівнює $\boldsymbol{v}_0 = (1/2 \quad 1/2 \quad 1/2 \quad 1/2)^T$. Аналогічно знаходимо ортогональну проекцію $\boldsymbol{u}_0$ вектора $\boldsymbol{v}_0$ на підпростір U. Вона дорівнює $\boldsymbol{u}_0 = (1/2 \quad 1/2 \quad 0 \quad 0)^T$.

Далі маємо:

$$\cos\varphi\{U, V\} = \cos\varphi\{\boldsymbol{u}_0, \boldsymbol{v}_0\} = \frac{\boldsymbol{u}_0^T \boldsymbol{v}_0}{|\boldsymbol{u}_o||\boldsymbol{v}_0|} = \frac{\sqrt{2}}{2},$$

тому $\varphi\{U, V\} = \pi/4$. $\triangle$

П р и к л а д 4. Знайдемо кут між підпросторами U і V, вважаючи, що вони визначаються відповідно системами однорідних рівнянь

$$1)\ \begin{aligned} x_1 \qquad + x_3 - x_4 &= 0, \\ x_1 + x_2 \qquad + x_4 &= 0; \end{aligned} \qquad 2)\ x_2 - x_3 - x_4 = 0.$$

Неважко побачити, що $\dim U = 2$, $\dim V = 3$. У розділі 3.5 було зазначено, що переріз $U \bigcap V$ визначається об'єднаною системою рівнянь

$$\begin{aligned} x_1 \qquad + x_3 - x_4 &= 0, \\ x_1 + x_2 \qquad + x_4 &= 0, \\ x_2 - x_3 - x_4 &= 0. \end{aligned}$$

Її фундаментальна система розв'язків складається з одного вектора $\boldsymbol{c} = (-1 \quad 1 \quad 1 \quad 0)^T$.

Отже, $\dim(U \bigcap V) = 1$ і ми повинні знаходити кут між підпросторами U' і V', які визначаються співвідношеннями (5.2.21).При розгляді прикладу з розділу 5.1 було з'ясовано, що ортогональним доповненням підпростору, натягнутого на вектор $\boldsymbol{c}$, є підпростір, що визначається рівнянням гіперплощини $\boldsymbol{c}^T \boldsymbol{x} = 0$. У нашому випадку це рівняння є таким:

$$-x_1 + x_2 + x_3 = 0.$$

Долучивши це рівняння до системи рівнянь, що визначає підпростір U, одержимо систему для знаходження підпростору U' у відповідності до першої формули (5.2.21):

$$\begin{aligned} x_1 \quad\quad + x_3 - x_4 &= 0, \\ x_1 + x_2 \quad\quad + x_4 &= 0, \\ -x_1 + x_2 + x_3 \quad\quad &= 0. \end{aligned}$$

Її фундаментальна система розв'язків складається з одного вектора $\boldsymbol{u}_1' = (0 \quad -1 \quad 1 \quad 1)^T$, який породжує підпростір U'.

Виходячи з другої формули (5.2.21), складемо систему рівнянь для знаходження підпростору V' шляхом об'єднання рівнянь, що визначають V і $(U \cap V)^{\perp}$:

$$\begin{aligned} x_2 - x_3 - x_4 &= 0, \\ -x_1 + x_2 + x_3 \quad\quad &= 0. \end{aligned}$$

Фундаментальною системою розв'язків цієї системи рівнянь є вектори $\boldsymbol{v}_1' = (2 \quad 1 \quad 1 \quad 0)^T$, $\boldsymbol{v}_2' = (1 \quad 1 \quad 0 \quad 1)^T$. Вони породжують підпростір V'.

Вектори $\boldsymbol{u}_1', \boldsymbol{v}_1', \boldsymbol{v}_2'$ є лінійно незалежними, тому можна стверджувати, що $\dim(U' \cap V') = 0$. Цей випадок розглядався в попередньому прикладі. Проте можна не вдаватися до обчислень, бо, як неважко перевірити, $U' \perp V'$. Тому $\varphi\{U, V\} = \pi/2$. △

П р и к л а д 5. Складемо рівняння гіперплощини, що проходить через точку $\boldsymbol{x}_0 = (2 \quad 1 \quad 1 \quad 0)^T$ перпендикулярно до прямої лінії

$$\frac{x_1 + 2}{1} = \frac{x_2}{-1} = \frac{x_3 - 1}{2} = \frac{x_4 + 3}{3}.$$

Скористаємося рівнянням гіперплощини у вигляді (5.2.15). Умова (7) показує, що вектор $\boldsymbol{c}$ можна взяти рівним

$$\boldsymbol{c} = \boldsymbol{f}_1 = (1 \quad -1 \quad 2 \quad 3)^T,$$

де $\boldsymbol{f}_1$ — вектор, що породжує спрямовуючий підпростір прямої лінії. Підставивши координати векторів $\boldsymbol{c}$ і $\boldsymbol{x}_0$ у формулу (5.2.15), одержимо рівняння гіперплощини

$$x_1 - x_2 + 2x_3 + 3x_4 = 3. \quad \triangle$$

П р и к л а д 6. Знайдемо рівняння гіперплощини, що проходить через точку $\boldsymbol{x}_0 = (1 \quad -1 \quad 1)^T$ паралельно до прямих ліній

$$\frac{x_1+1}{1} = \frac{x_2+1}{1} = \frac{x_3-3}{1}, \qquad \frac{x_1+3}{2} = \frac{x_2-1}{2} = \frac{x_3-2}{-1}.$$

Шукаємо рівняння гіперплощини у вигляді (5.2.15). Координати вектора $\boldsymbol{c} = (a_1 \quad a_2 \quad a_3)^T$ в цьому рівнянні знаходимо за допомогою умови (8), написаної для обох прямих ліній:

$$\begin{aligned} a_1 + \ \ a_2 + a_3 &= 0, \\ 2a_1 + 2a_2 - a_3 &= 0. \end{aligned}$$

Фундаментальна система розв'язків цієї системи рівнянь складається з вектора $(-1 \quad 1 \quad 0)^T$, який є шуканим вектором $\boldsymbol{c}$. Тепер на підставі (5.2.15) маємо рівняння гіперплощини

$$-x_1 + x_2 = -2. \ \triangle$$

Вправи до глави 5

5.1. Нехай матриця $\boldsymbol{A}$ має розміри $m \times n$, причому $m < n$ і $r(\boldsymbol{A}) = m$. Довести, що вектор $\boldsymbol{x}$ мінімальної довжини, що задовольняє умові $\boldsymbol{Ax} = \boldsymbol{b}$, дорівнює $\boldsymbol{x} = \sum_{j=1}^{m} c_j \boldsymbol{a}_{j*}^T$. Тут $\boldsymbol{a}_{1*}, \ldots, \boldsymbol{a}_{m*}$ — рядки матриці $\boldsymbol{A}$, а коефіцієнти $c_1, \ldots, c_m$ визначаються з системи рівнянь $\sum_{j=1}^{m} \boldsymbol{a}_{i*}\boldsymbol{a}_{j*}^T c_j = b_i$, $i = \overline{1, m}$, де величини b_i є координатами вектора $\boldsymbol{b}$.

5.2. У просторі S_n знайти пряму лінію, яка проходить через задану точку перпендикулярно до m ($m < n$) попарно непаралельних прямих ліній. При якому m задача має єдиний розв'язок? Окремо розглянути випадок тривимірного простору.

ГЛАВА 6

АЛГЕБРАЇЧНА ПРОБЛЕМА ВЛАСНИХ ЗНАЧЕНЬ

Алгебраїчна проблема власних значень є однією з найбільш складних і важливих у лінійній алгебрі. Про її складність свідчить хоч би той факт, що для будь-якої невиродженої матриці $\boldsymbol{A}$ порядку n розв'язок системи $\boldsymbol{Ax} = \boldsymbol{b}$ може бути поданий в аналітичному вигляді за допомогою формул Крамера, тоді як алгебраїчна проблема власних значень має аналітичний розв'язок для матриць не вище четвертого порядку.

Глава містить як початкові відомості про проблему і методи її розв'язання, так і деякі спеціальні питання. Зокрема, в останньому розділі глави одержане представлення довільної матриці, яке широко використовується при розв'язанні задач математичної статистики.

6.1. Власні значення і власні вектори

6.1.1. Визначення власних значень і власних векторів. Підпростір U векторного простору R називається *інваріантним* відносно лінійного перетворення $\boldsymbol{A}$, якщо для кожного вектора $\boldsymbol{x}$ з U вектор $\boldsymbol{Ax}$ теж належить до U. Тривіальними прикладами інваріантних підпросторів є нульовий підпростір і весь простір R. Неважко переконатися в тому, що образ і ядро матриці є інваріантними підпросторами простору R.

Нехай U — одновимірний інваріантний підпростір, породжений вектором $\boldsymbol{x}$. Це означає, що

$$\boldsymbol{Ax} = \lambda \boldsymbol{x}. \tag{1}$$

Вектор $\boldsymbol{x} \neq \mathbf{0}$, що задовольняє рівності (1), називається *власним вектором*, а відповідне число λ — *власним значенням* (*власним числом*) лінійного перетворення (матриці) $\boldsymbol{A}$. У подальшому ми будемо вести мову про власні значення, власні вектори і їх властивості, маючи на увазі передусім матрицю лінійного перетворення.

Рівності (1) можна надати вигляду:

$$(\lambda \boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \mathbf{0}. \tag{2}$$

Оскільки вектор $\boldsymbol{x}$ є ненульовим, то на підставі висновку з теореми 3.5.1 можна стверджувати, що матриця $\lambda \boldsymbol{E} - \boldsymbol{A}$ є виродженою, тобто

$$\det(\lambda \boldsymbol{E} - \boldsymbol{A}) = 0. \tag{3}$$

Це рівняння відоме як *характеристичне рівняння* матриці $\boldsymbol{A}$. Ураховуючи вираз для визначника матриці через її елементи, легко зрозуміти, що ліву частину (3) можна записати так:

$$\det(\lambda \boldsymbol{E} - \boldsymbol{A}) = a_0 + a_1\lambda + \ldots + a_{n-1}\lambda^{n-1} + \lambda^n. \tag{4}$$

Права частина рівності (4) називається *характеристичним многочленом* матриці $\boldsymbol{A}$, а його корені — *характеристичними коренями*.

Легко зрозуміти, що замість (2) і (3) можна взяти рівняння

$$(\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x} = \mathbf{0}, \quad \det(\boldsymbol{A} - \lambda \boldsymbol{E}) = 0. \tag{5}$$

Співвідношення (5) є такими ж уживанними, як і (2) та (3).

У деяких випадках треба за многочленом (4) побудувати матрицю $\boldsymbol{F}$, для якої цей многчлен є характеристичним. Такою матрицею може бути *матриця Фробеніуса*

$$\boldsymbol{F} = \begin{pmatrix} -a_{n-1} & -a_{n-2} & \ldots & -a_1 & -a_0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix}. \tag{6}$$

У цьому можна переконатися, розклавши визначник $\det(\lambda \boldsymbol{E} - \boldsymbol{F})$ за елементами першого рядка.

6.1.2. Комплексні числа. Поки ми мали справу з розв'язанням системи $\boldsymbol{Ax} = \boldsymbol{b}$, було зрозуміло, що розв'язок $\boldsymbol{x}$ є дійсним при дійсних $\boldsymbol{A}$ і $\boldsymbol{b}$. Інша річ, коли мова йде про алгебраїчну проблему власних значень. Матриця може бути дійсною, але власні значення і власні вектори можуть такими не бути. Для розв'язання в повному обсязі проблеми власних значень і власних векторів потрібно ввести до розгляду комплексні числа.

Уявною одиницею називається число i, що задовольняє рівнянню $i^2 = -1$. Усі числа bi, де b — дійсне, називають уявними. Сума дійсного і уявного чисел є *комплексним числом* $a + bi$, яке позначається однією літерою z. Число a називається дійсною частиною, а число b — уявною частиною комплексного числа $z = a + bi$. Для цих чисел використовують позначення:

$$a = \operatorname{Re}(a + bi) = \operatorname{Re} z, \quad b = \operatorname{Im}(a + bi) = \operatorname{Im} z.$$

Комплексне число $\bar{z} = a - bi$ називається *спряженим* по відношенню до комплексного числа $z = a + bi$. Очевидно, що $\overline{(\bar{z})} = z$ для будь-якого комплексного числа z, а рівність $\bar{z} = z$ виконується тоді і тільки тоді, коли z — дійсне число.

Два комплексних числа додаються за правилом:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

і множаться з використанням співвідношення $i^2 = -1$:

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Обчислення частки комплексних чисел зводиться до обчислення добутку шляхом множення чисельника і знаменника на число, спряжене зі знаменником. Зауважимо, що операція спряження переставна з арифметичними операціями над комплексними числами:

$$\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2} \ (z_2 \neq 0).$$

Ці рівності перевіряються безпосередньо.

Число $\sqrt{a^2+b^2}$ називається *модулем* комплексного числа $z=a+bi$ і позначається через $|z|$:

$$|z|=|a+bi|=\sqrt{a^2+b^2}.$$

Очевидно, що $|z|=|\bar{z}|$ і $z\bar{z}=|z|^2$. Зрозуміло, що $|z|>0$, причому $|z|=0$ тоді і тільки тоді, коли $z=0$.

Кут φ, який визначається зі співвідношень

$$\cos\varphi=\frac{a}{\sqrt{a^2+b^2}},\quad \sin\varphi=\frac{b}{\sqrt{a^2+b^2}}, \tag{7}$$

називається *аргументом* комплексного числа $z=a+bi$ (рис. 11). Аргумент числа z позначається через $\arg z$ і визначається з точністю до доданка $2m\pi$, де m — ціле число. З формул (7) випливає, що для $z\neq 0$ має місце *тригонометрична* форма комплексного числа:

$$z=|z|(\cos\varphi+i\sin\varphi).$$

Комплексне число $\cos\varphi+i\sin\varphi$ позначається символом $e^{i\varphi}$, тобто для будь-якого дійсного числа φ має місце *формула Ейлера*:

$$e^{i\varphi}=\cos\varphi+i\sin\varphi. \tag{8}$$

Функція $e^{i\varphi}$ має звичайні властивості показникової функції. Основними з них є такі:

$$e^{i\varphi_1}e^{i\varphi_2}=e^{i(\varphi_1+\varphi_2)},\quad \frac{e^{i\varphi_1}}{e^{i\varphi_2}}=e^{i(\varphi_1-\varphi_2)},$$

в чому неважко переконатися за допомогою формули (8). Ця ж формула дозволяє написати будь-яке комплексне число в *показниковій* формі:

$$z=|z|e^{i\varphi}.$$

Комплексне число можна зобразити точкою на комплексній площині (рис. 12). Горизонтальна вісь на комплексній площині називається дійсною віссю, а вертикальна — уявною.

Комплексне число z зображують також спрямованим відрізком (вектором) з початком у точці O і кінцем у точці z. Відповідність між комплексними числами і векторами комплексної площини з початком у точці O є взаємно однозначною. Тому вектор, що зображує комплексне число z, позначають тією ж літерою z. Модуль комплексного числа z дорівнює довжині вектора z, а аргумент — куту між дійсною віссю і вектором z.

За допомогою векторного тлумачення можна наочно ілюструвати додавання і віднімання комплексних чисел. З правила додавання комплексних чисел випливає, що число $z_1 + z_2$ зображується вектором, побудованим за звичайним правилом додавання векторів. Вектор $z_1 - z_2$ будується як сума векторів z_1 і $-z_2$ (рис. 12). На рис. 12 видно, що відстань між точками z_1 і z_2 дорівнює $|z_1 - z_2|$. Множина точок z, що задовольняє нерівності $|z_1 - z_0| \le R$, є кругом з колом включно радіуса R з центром у точці z_0.

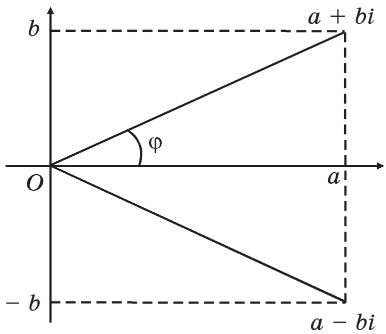

Рис. 11

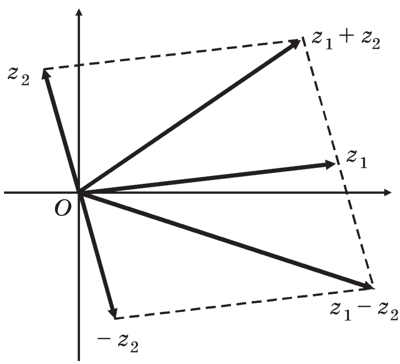

Рис. 12

Для будь-яких комплексних чисел z_1 і z_2 виконуються нерівності:

$$||z_1| - |z_2|| \le |z_1 + z_2| \le |z_1| + |z_2|. \tag{9}$$

Дійсно, довжини сторін трикутника з вершинами в точках O, z_1, $z_1 + z_2$ дорівнюють $|z_1|$, $|z_2|$ і $|z_1 + z_2|$ (рис. 12), тому нерівності (9) є відомими з елементарної геометрії нерівностями для довжин сторін трикутника.

Наприкінці покажемо, як виконується добування кореня n-го степеня з комплексного числа. Число $w = \sqrt[n]{z}$ називається коренем n-го степеня з числа z, якщо $w^n = z$. Нехай $w = |w| e^{i\varphi}$ і $z = |z| e^{i\theta}$, тоді

$$|w|^n e^{in\varphi} = |z| e^{i\theta}.$$

Два комплксних числа, записаних у показниковій формі, є рівними, коли їх модулі рівні, а аргументи відрізняються на $2k\pi$, де k — деяке ціле число. Отже,

$$|w|^n = |z|, \quad n\varphi = \theta + 2k\pi,$$

тому маємо множину коренів з числа z:

$$w_k = \sqrt[n]{|z|} e^{i(\theta + 2k\pi)/n}, \quad k = 0, \pm 1, \pm 2, \ldots . \tag{10}$$

Покажемо, що серед комплексних чисел (10) є лише n різних. Аргументи чисел $w_0, w_1, \ldots, w_{n-1}$ дорівнюють

$$\varphi_0 = \frac{\theta}{n}, \quad \varphi_1 = \frac{\theta + 2\pi}{n}, \quad \ldots, \quad \varphi_{n-1} = \frac{\theta + 2\pi(n-1)}{n}.$$

Усі вони різні і відрізняються один від одного менше ніж на 2π. Корінь w_n збігається з w_0, оскільки $|w_n| = |w_0|$ і $\varphi_n = \varphi_0 + 2\pi$. Аналогічно можна переконатися в тому, що $w_{n+1} = w_1$, $w_{-1} = w_{n-1}$ і т.д.

6.1.3. Характеристичний многочлен. У зв'язку з вивченням характеристичного многочлена нижче розглядаються деякі алгебраїчні властивості довільних многочленів. Сформулюємо теорему, на яку значною мірою спирається розв'язання алгебраїчної проблеми власних значень. Ми будемо користуватися визначенням коренів многочлена $f(\lambda)$ як значень аргумента λ, при яких многочлен обертається в нуль.

***Теорема 6.1.1* (*основна теорема алгебри*).** *Будь-який многочлен*

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

степеня $n \geq 1$ *з комплексними коефіцієнтами має принаймні один корінь, взагалі кажучи, комплексний.* □

Розглянемо найбільш важливий висновок з теореми. Нехай многочлен $f(\lambda)$ степеня $n \geq 1$ має принаймні один корінь λ_1. Тоді

$$f(\lambda) = (\lambda - \lambda_1)\, g(\lambda),$$

де $g(\lambda)$ — многочлен $(n-1)$-го степеня з комплексними коефіцієнтами. У свою чергу, $g(\lambda)$ має корінь λ_2 при $n \geq 2$:

$$g(\lambda) = (\lambda - \lambda_2)\, h(\lambda),$$

звідки знаходимо, що

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\, h(\lambda).$$

Продовжуючи ці міркування, наприкінці одержимо:

$$f(\lambda) = a(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n), \tag{11}$$

де a — деяке число. Неважко зрозуміти, що $a = a_n$. Число a_n називається старшим коефіцієнтом. Якщо $a_n = 1$, то многочлен $f(\lambda)$ називається *зведеним*.

Серед множників правої частини рівності (11) можуть бути однакові. Кількість однакових множників для деякого кореня називається його *кратністю*. Корінь кратності 1 називається *простим*. Можна дати інше еквівалентне визначення кратного кореня: число λ_0 є коренем многочлена $f(\lambda)$ кратності m, якщо

$$f(\lambda) = (\lambda - \lambda_0)^m\, g(\lambda), \tag{12}$$

де многочлен $g(\lambda)$ є таким, що $g(\lambda_0) \neq 0$. Про необхідні і достатні умови кратності кореня многочлена йдеться в наступній теоремі.

Теорема 6.1.2. *Для того, щоб число* λ_0 *було коренем многочлена* $f(\lambda)$ *кратності* m*, необхідно і достатньо, щоб виконувалися умови*

$$f^{(k)}(\lambda_0) = 0, \quad k = \overline{0,\ m-1}, \quad f^{(m)}(\lambda_0) \neq 0. \tag{13}$$

Д о в е д е н н я. Н е о б х і д н і с т ь. Продиференціюємо многочлен (12), тоді матимемо:

$$f'(\lambda) = m(\lambda - \lambda_0)^{m-1} g(\lambda) + (\lambda - \lambda_0)^m g'(\lambda) = (\lambda - \lambda_0)^{m-1} h(\lambda),$$

де

$$h(\lambda) = m\, g(\lambda) + (\lambda - \lambda_0) g'(\lambda), \quad h(\lambda_0) \neq 0.$$

Звідси випливає, що корінь многочлена $f(\lambda)$ кратності m є одночасно коренем многочлена $f'(\lambda)$ кратності $m-1$ і навпаки. Аналогічно можна стверджувати, що λ_0 є коренем кратності $m-2$ многочлена $f''(\lambda)$ і т.д. Наприкінці одержимо, що λ_0 не є коренем многочлена $f^{(m)}(\lambda)$, тобто $f^{(m)}(\lambda_0) \neq 0$.

Д о с т а т н і с т ь. Нехай виконуються умови (13). Оскільки $f^{(m-1)}(\lambda_0) = 0$, то число λ_0 є коренем многочлена $f^{(m-1)}(\lambda)$ кратності не нижче 1. Одночасно λ_0 є коренем многочлена $f^{(m-2)}(\lambda)$ кратності не нижче 2 і т.д.

Залишається показати, що кратність кореня λ_0 многочлена $f(\lambda)$ не перевищує m. Якби ця кратність була більшою ніж m, то кратність цього кореня для $f^{(m-1)}(\lambda)$ була б вищою ніж 1, звідки випливала б рівність $f^{(m)}(\lambda_0) = 0$, яка суперечить одній з умов (13). □

Таким чином, будь-який многочлен степеня $n \geq 1$ з комплексними коефіцієнтами має n коренів, якщо кожний корінь рахувати стільки разів, якою є його кратність. Таким є головний висновок з основної теореми алгебри.

Многочлен нульового степеня $f(\lambda) = a_0$, де $a_0 \neq 0$, не має коренів. Нульовий многочлен $f(\lambda) = a_0$, де $a_0 = 0$, є єдиним многочленом, який має скільки завгодно різних коренів. Якщо многочлени $f(\lambda)$ і $g(\lambda)$, степені яких не перевищують n, мають рівні значення більш ніж при n різних значеннях аргумента, то ці многочлени збігаються. Дійсно, многочлен $f(\lambda) - g(\lambda)$ має більше ніж n коренів, а його степінь не перевищує n. Тому $f(\lambda) - g(\lambda) = 0$, тобто $f(\lambda) = g(\lambda)$.

Нехай $f(\lambda)$ — многочлен з дійсними коефіцієнтами, тоді $\overline{f(\lambda)} = f(\overline{\lambda})$. Якщо $f(\lambda_0) = 0$, то $f(\overline{\lambda}_0) = \overline{f(\lambda_0)} = 0$. Це означає, що коли λ_0 є коренем многочлена з дійсними коефіцієнтами, то і комплексно спряжене число $\overline{\lambda}_0$ також є коренем цього многочлена. Звідси випливає, що многочлен $f(\lambda)$ ділиться на квадратний тричлен $(\lambda - \lambda_0)(\lambda - \overline{\lambda}_0)$, коефіцієнти якого є дійсними. Можна показати, що корені λ_0 і $\overline{\lambda}_0$ мають одну і ту ж кратність.

Очевидно, що многочлен з дійсними коефіцієнтами може мати тільки парну кількість коренів, що не є дійсними. Тому будь-який многочлен непарного степеня з дійсними коефіцієнтами має принаймні один дійсний корінь.

Повернемось до алгебраїчної проблеми власних значень. Нехай лінійне перетворення $\boldsymbol{A}$ діє в комплексному просторі. На підставі наведенних вище міркувань можна стверджувати, що характеристичний многочлен представляється у вигляді:

$$\det(\lambda \boldsymbol{E} - \boldsymbol{A}) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n), \tag{14}$$

де $\lambda_1, \ldots, \lambda_n$ — характеристичні корені. При цьому кожному характеристичному кореню, відмінному від інших коренів, відповідає принаймні один власний вектор.

Щодо власних значень ми будемо застосовувати ту ж термінологію, що і стосовно характеристичних коренів. Власні значення будемо називати *простими* (*кратними*), якщо вони є простими (кратними) характеристичними коренями. *Кратністю власного значення* будемо називати його кратність як характеристичного кореня. Сукупність n власних значень матриці $\boldsymbol{A}$ порядку n (кожне відмінне від інших власне значення береться стільки разів, якою є його кратність) називається її *спектром* і позначається через $\sigma(\boldsymbol{A})$.

Сформулюємо загальну схему розв'язання задачі про власні значення і власні вектори.

Спочатку з рівняння (3) знаходимо власні значення. Потім підставляємо кожне з них у систему рівнянь (2). Кожного разу ця система має нові числові коефіцієнти. Ранг матриці одержаної системи буде деяким числом $r < n$. У відповідності до теореми 3.5.1 ця система буде мати $n - r$ лінійно незалежних розв'язків. Визначивши їх, ми знайдемо тим самим $n - r$ лінійно незалежних власних векторів з одним і тим же власним значенням. Перебравши всі власні значення, ми знайдемо всі власні вектори матриці.

6.2. Властивості власних значень і власних векторів

6.2.1. Основні теореми. Почнемо з переліку цілком очевидних властивостей: 1) матриці $\boldsymbol{A}$ і $\boldsymbol{A}^T$ мають однакові власні значення; 2) для будь-якої трикутної (діагональної, скалярної, одиничної, нульової) матриці її власні значення збігаються з діагональними елементами; 3) для діагональної матриці будь-який вектор натурального базису є власним; 4) для скалярної (одиничної, нульової) матриці будь-який ненульовий вектор є власним.

Введемо до розгляду величину, рівну сумі діагональних елементів матриці $\boldsymbol{A}$. Вона називається *слідом* матриці $\boldsymbol{A}$ і позначається через $\operatorname{tr} \boldsymbol{A}$.

Теорема 6.2.1. *Добуток власних значень матриці дорівнює її визначнику, а сума власних значень — сліду матриці.*

Д о в е д е н н я. Будемо вважати, що $\lambda = 0$ в рівностях (6.1.4) і (6.1.14). Тоді одержимо:

$$(-1)^n a_0 = \det \boldsymbol{A} = \prod_{i=1}^{n} \lambda_i \,. \tag{1}$$

Порівнявши коефіцієнти при λ^{n-1} в правих частинах формул (6.1.4) і (6.1.14), будемо мати:

$$a_{n-1} = -\sum_{i=1}^{n} \lambda_i \,.$$

З іншого боку, знайдемо коефіцієнт при λ^{n-1} у визначнику $\det(\lambda \boldsymbol{E} - \boldsymbol{A})$, звернувши увагу на те, що всі доданки в $\det(\lambda \boldsymbol{E} - \boldsymbol{A})$, які містять λ^{n-1}, одержуються з добутків елементів головної діагоналі визначника. Виконавши необхідні обчислення, знайдемо, що шуканий коефіцієнт дорівнює $-\sum_{i=1}^{n} a_{ii}$.

Отже, ми показали, що

$$-a_{n-1} = \operatorname{tr} \boldsymbol{A} = \sum_{i=1}^{n} \lambda_i . \qquad (2)$$

Висновок. *Матриця* ***A*** *є невиродженою тоді і тільки тоді, коли вона не має нульового власного значення.*

Д о в е д е н н я випливає з формули (1). □

Далі з’ясуємо зв’язок між власними значеннями і власними векторами матриці $\boldsymbol{A}$ і матриць $\alpha\boldsymbol{A}$, $\boldsymbol{A}+\mu\boldsymbol{E}$, $\boldsymbol{A}^m$.

Формула (6.1.1) доводить, що при множенні матриці на ненульове число власні вектори не змінюються, а власні значення множаться на це число. Про матриці $\boldsymbol{A}+\mu\boldsymbol{E}$ і $\boldsymbol{A}^m$ йдеться у двох наступних теоремах.

Теорема 6.2.2. *Нехай* $\boldsymbol{x}$ *і* λ *— власний вектор і відповідне до нього власне значення матриці* ***A****. Тоді* $\boldsymbol{x}$ *є власним вектором матриці* $\boldsymbol{A}+\mu\boldsymbol{E}$*, відповідним до її власного значення* $\lambda+\mu$.

Д о в е д е н н я випливає зі співвідношення

$$[(\lambda+\mu)\boldsymbol{E}-(\boldsymbol{A}+\mu\boldsymbol{E})]\boldsymbol{x}=\boldsymbol{0},$$

яке є прямим наслідком рівності (6.1.2). □

Теорема 6.2.3. *Нехай* $\boldsymbol{x}$ *і* λ *— власний вектор і відповідне до нього власне значення матриці* ***A****. Тоді* $\boldsymbol{x}$ *є власним вектором матриці* $\boldsymbol{A}^m$*, відповідним до її власного значення* λ^m*, причому m є будь-яким цілим додатним або від’ємним числом, якщо матриця* ***A*** *невироджена, і цілим додатним числом, якщо* ***A*** *вироджена.*

Д о в е д е н н я. Помноживши зліва обидві частини рівності (6.1.1) на матрицю $\boldsymbol{A}$, будемо мати:

$$\boldsymbol{A}^2\boldsymbol{x}=\lambda\boldsymbol{A}\boldsymbol{x}=\lambda^2\boldsymbol{x}.$$

Продовжуючи множення на матрицю $\boldsymbol{A}$, одержимо:

$$\boldsymbol{A}^m\boldsymbol{x}=\lambda^m\boldsymbol{x}. \qquad (3)$$

Рівність (3) означає, що $\boldsymbol{x}$ є власним вектором матриці $\boldsymbol{A}^m$, а λ^m є її власним значенням. Покажемо, що рівність (3) залишається в силі і для цілих від'ємних чисел m, якщо матриця $\boldsymbol{A}^{-1}$ існує.

Зауважимо спочатку, що у відповідності до висновку з теореми 6.2.1 $\lambda \neq 0$ для невиродженої матриці $\boldsymbol{A}$. Тепер з рівності (6.1.1) знаходимо, що

$$\boldsymbol{A}^{-1}\boldsymbol{x} = \lambda^{-1}\boldsymbol{x}.$$

Помноживши зліва обидві частини цієї рівності на $\boldsymbol{A}^{-1}$, одержимо:

$$\boldsymbol{A}^{-2}\boldsymbol{x} = \lambda^{-1}\boldsymbol{A}^{-1}\boldsymbol{x} = \lambda^{-2}\boldsymbol{x}.$$

Цей результат можна узагальнити і показати, що має місце формула (3). □

Якщо $\boldsymbol{x}$ — власний вектор, відповідний до власного значення λ, то $\alpha\boldsymbol{x}$, де α — ненульове число, також є власним вектором. Очевидно, що коли деякому власному значенню відповідають кілька власних векторів, то їх лінійна комбінація теж є власним вектором. Властивість власних векторів, відповідних до різних власних значень, з'ясовується в наступній теоремі.

Теорема 6.2.4. *Власні вектори, відповідні до різних власних значень, є лінійно незалежними.*

Д о в е д е н н я. Нехай $\lambda_1, \ldots, \lambda_m$ — різні власні значення матриці $\boldsymbol{A}$ порядку n і $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ — відповідні власні вектори. Покажемо, що рівність

$$\sum_{i=1}^{m} \alpha_i \boldsymbol{x}_i = \mathbf{0} \tag{4}$$

можлива лише при $\alpha_1 = \ldots = \alpha_m = 0$.

Помноживши зліва обидві частини рівності (4) на $\lambda_1\boldsymbol{E} - \boldsymbol{A}$, знаходимо, що $\sum_{i=2}^{m} (\lambda_1 - \lambda_i)\alpha_i \boldsymbol{x}_i = \mathbf{0}$. Після подальшого множення на $\lambda_2\boldsymbol{E} - \boldsymbol{A}, \ldots, \lambda_{m-1}\boldsymbol{E} - \boldsymbol{A}$ одержимо:

$$(\lambda_1 - \lambda_m)(\lambda_2 - \lambda_m)\ldots(\lambda_{m-1} - \lambda_m)\alpha_m \boldsymbol{x}_m = \mathbf{0},$$

звідки випливає, що $\alpha_m = 0$.

Тепер рівність (4) набуває вигляду:

$$\sum_{i=1}^{m-1} \alpha_i \boldsymbol{x}_i = \mathbf{0}.$$

Далі можна довести, що $\alpha_{m-1} = 0$ і т. д. Отже, між векторами $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ нема лінійної залежності. □

З'ясуємо зв'язок між власними значеннями подібних матриць.

Теорема 6.2.5. *Подібні матриці мають однакові характеристичні многочлени.*

Д о в е д е н н я. На підставі формули (4.3.7) і теорем 3.4.5 та 1.3.5 маємо:

$$\det(\lambda \boldsymbol{E} - \boldsymbol{A}') = \det(\lambda \boldsymbol{Q}^{-1}\boldsymbol{E}\boldsymbol{Q} - \boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}) = \det(\boldsymbol{Q}^{-1}(\lambda \boldsymbol{E} - \boldsymbol{A})\boldsymbol{Q}) =$$

$$= \det(\boldsymbol{Q}^{-1}) \cdot \det(\lambda \boldsymbol{E} - \boldsymbol{A}) \cdot \det \boldsymbol{Q} = \det(\lambda \boldsymbol{E} - \boldsymbol{A}). \ \square$$

Як зазначалося в розділі 6.1, власні значення можуть бути знайдені як корені характеристичного многочлена. Проте, коли ми маємо справу лише з характеристичним многочленом, втрачається деяка інформація, бо існує багато різних матриць з одним і тим же характеристичним многочленом. Тому зрозумілим є той факт, що більш сильні результати одержуються при врахуванні індивідуальної будови матриці. Таку можливість надає теорема 6.2.5, на підставі якої власні значення матриці можна знаходити шляхом подібних перетворень її до трикутної форми, коли власні значення з'являються на головній діагоналі.

6.2.2. Матриці простої структури. Матриця порядку n називається *матрицею простої структури*, якщо вона має n лінійно незалежних власних векторів.

Оскільки кожному відмінному від інших власному значенню матриці відповідає принаймні один власний вектор, то матриця, всі власні значення якої є простими, на підставі теореми 6.2.4 має n лінійно незалежних власних векторів і тому є матрицею простої структури. У наступній теоремі йдеться про необхідну і достатню умову того, що матриця має просту структуру.

Теорема 6.2.6. *Матриця має просту структуру тоді і тільки тоді, коли вона подібна до діагональної матриці.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Припустимо, що матриця $\boldsymbol{A}$ має лінійно незалежні власні вектори $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$, відповідні до власних значень $\lambda_1, \dots, \lambda_n$ (не обов'язково різних), тобто

$$\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i, \quad i = \overline{1, n}. \tag{5}$$

Нехай $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ і $\boldsymbol{S} = (\boldsymbol{x}_1 \ \dots \ \boldsymbol{x}_n)$, тоді рівності (5) можна записати у вигляді матричної рівності $\boldsymbol{AS} = \boldsymbol{S\Lambda}$, звідки

$$\boldsymbol{A} = \boldsymbol{S\Lambda S}^{-1}. \tag{6}$$

Д о с т а т н і с т ь. Припустимо, що для діагональної матриці $\boldsymbol{\Lambda}$ і невиродженої матриці $\boldsymbol{S}$ виконується рівність (6). Тоді $\boldsymbol{AS} = \boldsymbol{S\Lambda}$ і $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$, $i = \overline{1, n}$, де через $\boldsymbol{x}_i$ і λ_i позначені стовпці матриці $\boldsymbol{S}$ і діагональні елементи $\boldsymbol{\Lambda}$.

Отже, стовпці матриці $\boldsymbol{S}$ можна вважати n лінійно незалежними власними векторами матриці $\boldsymbol{A}$ з власними значеннями λ_i, тобто матриця $\boldsymbol{A}$ має просту структуру. □

Представлення матриці $\boldsymbol{A}$ у вигляді (6) називають її *діагоналізацією* за допомогою перетворення подібності, а матрицю $\boldsymbol{S}$ називають діагоналізуючою.

П р и к л а д. За допомогою перетворення подібності діагоналізуємо матрицю

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

Матриця $\boldsymbol{A}$ має характеристичне рівняння

$$\begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = 0,$$

коренями якого є $\lambda_1 = i,\ \lambda_2 = -i$. Для визначення власного вектора $\boldsymbol{x}_1$, що відповідає власному значенню λ_1, маємо рівняння

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{E})\boldsymbol{x}_1 = \boldsymbol{0},$$

або

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \boldsymbol{x}_1 = \boldsymbol{0}.$$

Ранг матриці системи дорівнює одиниці, тому фундаментальна система розв'язків складається з одного вектора. Обчисливши його, одержимо, що $\boldsymbol{x}_1 = (-i \quad 1)^T$. Для власного значення λ_2 аналогічно знаходимо, що $\boldsymbol{x}_2 = (i \quad 1)^T$.

Складемо діагоналізуючу матрицю $\boldsymbol{S} = (\boldsymbol{x}_1 \quad \boldsymbol{x}_2)$ і знайдемо $\boldsymbol{S}^{-1}$:

$$\boldsymbol{S} = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}, \quad \boldsymbol{S}^{-1} = \frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}.$$

Тепер діагоналізація матриці $\boldsymbol{A}$ виглядає так:

$$\boldsymbol{A} = \boldsymbol{S\Lambda S}^{-1} = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\frac{1}{2}\begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}. \quad \triangle$$

Матриця $\boldsymbol{A}$ називається *збіжною*, якщо всі елементи матриці $\boldsymbol{A}^m$ прямують до нуля при $m \to \infty$. При цьому говорять, що матриця $\boldsymbol{A}^m$ прямує до нульової матриці, коли $m \to \infty$.

Теорема 6.2.7. *Нехай* $\boldsymbol{A}$ *— матриця простої структури. Для того, щоб* $\boldsymbol{A}^m \to \boldsymbol{0}$ *при* $m \to \infty$, *необхідно і достатньо, щоб усі власні значення матриці* $\boldsymbol{A}$ *були за модулем меншими ніж одиниця.*

Д о в е д е н н я. На підставі формули (6) маємо:

$$\boldsymbol{A}^m = (\boldsymbol{S\Lambda S}^{-1})(\boldsymbol{S\Lambda S}^{-1})\ldots(\boldsymbol{S\Lambda S}^{-1}) = \boldsymbol{S\Lambda}^m\boldsymbol{S}^{-1}. \tag{7}$$

Очевидно, що $\boldsymbol{A}^m \to \boldsymbol{0}$ тоді і тільки тоді, коли $\boldsymbol{\Lambda}^m \to \boldsymbol{0}$. Для збіжності $\boldsymbol{\Lambda}^m \to \boldsymbol{0}$ кожне власне значення за модулем повинно бути меншим ніж одиниця. □

6.2.3. Дефектні матриці. Матриця, у якої кількість лінійно незалежних власних векторів є меншою ніж її порядок, називається *дефектною*. Такою може бути тільки матриця з кратним власним значенням, оскільки йому може відповідати менша кількість лінійно незалежних власних векторів ніж кратність власного значення.

Стосовно дефектної матриці можна ставити питання про зведення її до деякої простої (жорданової) форми, яка є квазідіагональною матрицею з діагональними блоками у вигляді дводіагональних підматриць певної структури. Зазначимо, що побудова жорданової форми в обчислювальному сенсі є дуже нестійкою, бо незначна зміна матриці може повернути їй втрачені власні вектори і перетворити її в матрицю простої структури. Проте, всупереч цьому обмеженню теорія жорданової форми заслуговує на вивчення і надає багаті можливості для розуміння властивостей матриць.

У повному обсязі необхідні відомості про дефектні матриці і жорданову форму будуть одержані в главі 9, а тут ми обмежимося розглядом найпростішого випадку дефектної матриці $\boldsymbol{A}$ другого порядку.

Припустимо, що λ — двократне власне значення цієї матриці. Як зазначалося в розділі 6.1, система рівнянь (6.1.2) для знаходження власних векторів має $2 - r(\lambda \boldsymbol{E} - \boldsymbol{A})$ лінійно незалежних розв'язків. Їх буде два лише при $r(\lambda \boldsymbol{E} - \boldsymbol{A}) = 0$, тобто при $\boldsymbol{A} = \lambda \boldsymbol{E}$. Це означає, що $\boldsymbol{A}$ буде матрицею простої структури лише в разі, коли $\boldsymbol{A}$ є скалярною матрицею. У всіх інших випадках матриця буде мати один власний вектор. Отже, матриця другого порядку з кратним власним значенням є дефектною за умови, що вона не є скалярною.

Розглядаючи саме цей випадок, позначимо через $\boldsymbol{x}_1$ власний вектор, відповідний до кратного власного значення λ, і через $\boldsymbol{x}_2$ — вектор, що задовольняє рівнянню

$$\boldsymbol{A}\boldsymbol{x}_2 = \lambda \boldsymbol{x}_2 + \boldsymbol{x}_1. \tag{8}$$

Розмірковуючи, як при доведенні теореми 6.2.6, можна одержати аналог рівності (6) для дефектної матриці $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1}, \tag{9}$$

де

$$\boldsymbol{T} = (\boldsymbol{x}_1 \quad \boldsymbol{x}_2), \quad \boldsymbol{J} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}. \tag{10}$$

Матриця $\boldsymbol{J}$ є тим дводіагональним блоком, про який ішла мова вище. Вигляд блоків типу матриці $\boldsymbol{J}$ у більш складних випадках буде виявлений у главі 9.

Для обґрунтування справедливості рівності (9) треба показати, що розв'язок рівняння (8) існує і що він не є колінеарним до вектора $\boldsymbol{x}_1$. Остання умова забезпечує невиродженість матриці $\boldsymbol{T}$.

Помноживши зліва обидві частини рівності (8) на матрицю $\boldsymbol{A} - \lambda\boldsymbol{E}$, одержимо:

$$(\boldsymbol{A} - \lambda\boldsymbol{E})^2 \boldsymbol{x}_2 = \mathbf{0}.$$

Звідси випливає існування $\boldsymbol{x}_2$ як власного вектора матриці $(\boldsymbol{A} - \lambda\boldsymbol{E})^2$, відповідного до її нульового власного значення. Безпосередня перевірка показує, що рівняння (8) не задовольняється при $\boldsymbol{x}_2 = \alpha\boldsymbol{x}_1$, тобто вектор $\boldsymbol{x}_2$ не є колінеарним до вектора $\boldsymbol{x}_1$.

Отже, ми обґрунтували справедливість рівності (9). Координати вектора $\boldsymbol{x}_2$ можна знайти з (8) методом виключення.

6.3. Локалізація власних значень

6.3.1. Вступ. Алгебраїчна проблема власних значень має точний розв'язок лише в небагатьох виключних випадках. Як правило, для знаходження власних значень треба застосовувати наближені методи. Проте, якщо не намагатися знайти власні значення з високою точністю, а обмежитися шуканням їх оцінок, або, як кажуть, розв'язувати задачу локалізації власних значень, то тут існують прості аналітичні методи для матриць довільного порядку. У деяких задачах досить мати оцінки для власних значень, а не їх точні значення. Наприклад, у задачах, пов'язаних з розглядуваними в главі 10 системами різницевих і диференціальних рівнянь важливі висновки можна зробити без точного знання власних значень λ_i, а лише на підставі інформації про виконання чи невиконання нерівностей $|\lambda_i| < 1$ або $\operatorname{Re}\lambda_i < 0$.

Інтерес до проблеми локалізації власних значень пов'язаний також із застосуванням її при побудові чисельних методів. При знаходженні власних значень і власних векторів попередня локалізація власних значень звичайно суттєво полегшує їх обчислення. Навіть такі задачі, безпосередньо не пов'язані з власними значеннями, як системи лінійних рівнянь, потребують локалізації власних значень розглядуваних матриць. Стисле висвітлення деяких аспектів цієї проблеми можна знайти в главі 12.

Нижче розглядається один з найпростіших методів локалізації власних значень, заснований на використанні так званих кругів Гершгоріна. Він є придатним для будь-яких дійсних або комплексних матриць. Методи локалізації власних значень симетричних матриць вивчаються в розділах 7.4 і 7.5.

6.3.2. Матриці Адамара. Розглянемо основну з обговорюваного питання теорему.

***Теорема 6.3.1* (*теорема Адамара*).** *Якщо для матриці* $\boldsymbol{A}$ *виконуються нерівності*

$$| a_{ii} | > \sum_{\substack{j=1 \\ j\neq i}}^{n} | a_{ij} |, \quad i = \overline{1,\ n}, \tag{1}$$

то матриця $\boldsymbol{A}$ *є невиродженою.*

Д о в е д е н н я. Припустимо, що матриця $\boldsymbol{A}$ вироджена. Тоді існує ненульовий вектор $\boldsymbol{x} = (x_1 \quad \dots \quad x_n)^T$ такий, що

$$\sum_{j=1}^{n} a_{ij} x_j = 0, \quad i = \overline{1,\ n}. \tag{2}$$

Виберемо індекс i такий, якому відповідає максимальний $| x_i |$. З формули (2) для вибраного i одержимо:

$$a_{ii} x_i = -\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} x_j ,$$

звідки

$$| a_{ii} || x_i | \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} | a_{ij} || x_j | \leq | x_i | \sum_{\substack{j=1 \\ j\neq i}}^{n} | a_{ij} |.$$

Скоротивши на $| x_i |$, одержимо нерівність, яка суперечить (1). Це означає, що матриця $\boldsymbol{A}$ є невиродженою. □

Оскільки $\det \boldsymbol{A} = \det(\boldsymbol{A}^T)$, то, замінивши $\boldsymbol{A}$ на $\boldsymbol{A}^T$, одержимо умову невиродженості матриці $\boldsymbol{A}$ для стовпців:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}|, \quad i = \overline{1,\ n}. \tag{3}$$

Матриця $\boldsymbol{A}$ називається *матрицею Адамара*, якщо її елементи задовольняють умовам (1) або (3). Кожна з цих умов означає, що модуль діагонального елемента перевищує суму модулів всіх інших елементів того ж рядка або стовпця.

6.3.3. Круги Гершгоріна. Власні значення діагональної матриці $\mathrm{diag}\,(a_{11},\ \ldots,\ a_{nn})$ локалізувати легко: це точки $a_{11},\ \ldots,\ a_{nn}$ на комплексній площині. Природно припустити, що при «малих» позадіагональних елементах власні значення будуть знаходитися в деяких малих околах точок $a_{11},\ \ldots,\ a_{nn}$. Таке припущення є обґрунтованим, як показує

Теорема 6.3.2 (теорема Гершгоріна). *Будь-яке власне значення матриці* $\boldsymbol{A}$ *лежить принаймні в одному з кругів*

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad i = \overline{1,\ n}, \tag{4}$$

або

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}|, \quad i = \overline{1,\ n}, \tag{5}$$

на комплексній λ*-площині.*

Д о в е д е н н я. Нехай нерівності (4) не виконуються для деякого власного значення λ. Тоді

$$|\lambda - a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad i = \overline{1,\ n}.$$

На підставі теореми Адамара це означає невиродженість матриці $\lambda\boldsymbol{E} - \boldsymbol{A}$, що суперечить припущенню, що λ — власне значення матриці $\boldsymbol{A}$.

Аналогічно, посилаючись на умови (3), можна довести нерівності (5). □

Круги на комплексній λ-площині, які визначаються нерівностями (4) і (5), називаються кругами Гершгоріна. Якщо позадіагональні елементи деякого рядка або стовпця є нульовими, то відповідний круг Гершгоріна є точкою. Наприклад, для одиничної матриці сукупність кругів Гершгоріна складається з однієї точки.

Теорема 6.3.2 не вказує, яке i відповідає власному значенню λ. У теоремі лише стверджується, що для кожного власного значення λ знайдеться круг з центром a_{ii} відповідного радіуса, який вміщує λ. Отже, λ належить об'єднанню таких кругів. Нижче ми одержимо додаткову інформацію про локалізацію власних значень, яка уточнює теорему 6.3.2. Для цього введемо поняття кратності кругів Гершгоріна і зв'язності області комплексної площини.

Якщо один і той же круг Гершгоріна породжується k рядками або k стовпцями матриці, то говорять, що круг має кратність k.

Множина точок комплексної площини називається зв'язною областю, якщо вона разом з двома своїми точками вміщує ламану лінію, яка з'єднує їх.

Має місце

Теорема 6.3.3. *З урахуванням кратності кругів Гершгоріна і власних значень має місце таке твердження: m кругів, які утворюють зв'язну область комплексної площини і не перетинаються з іншими $n - m$ кругами, вміщують рівно m власних значень.*

Д о в е д е н н я. Разом з матрицею $\boldsymbol{A}$ розглянемо матрицю $\boldsymbol{A}(t)$ для $t \in [0,\ 1]$ з елементами $a_{ii}(t) = a_{ii}$ і $a_{ij}(t) = ta_{ij}$ при $i \neq j$. Очевидно, що $\boldsymbol{A}(1) = \boldsymbol{A}$.

Нехай m кругів утворюють зв'язну область γ. Їй відповідає область $\gamma(t)$ для матриці $\boldsymbol{A}(t)$. Вона є об'єднанням кругів

$$|\,\lambda - a_{ii}\,| \le t s_j, \quad i = \overline{1,\ n},$$

де s_j — праві частини рівностей (4) або (5). Очевидно, що $\gamma(t) \subseteq \gamma$ і $\gamma(1) = \gamma$. Зазначимо, що область $\gamma(t)$ на відміну від γ не обов'язково є зв'язною для всіх $t \in [0,\ 1]$.

Відомо, що власні значення неперервно залежать від елементів матриці. Тому для матриці $\boldsymbol{A}(t)$ вони є неперервними функціями від t. Неперервно змінюючись, вони не виходять за межі області γ. Тому кількість власних значень матриці, належних до γ, залишається незмінною.

Але власними значеннями матриці $\boldsymbol{A}(0)$ є діагональні елементи a_{ii}, що є центрами кругів, належних до γ. Тому кількість власних значень матриці $\boldsymbol{A}(0)$ і, отже, матриці $\boldsymbol{A}(t)$, належних до γ, дорівнює кількості кругів, об'єднанням яких є γ.

Висновок 1. *Якщо деякий круг Гершгоріна не перетинається з іншими і має кратність* 1, *то він вміщує одне власне значення.*

Доведення є очевидним.

Висновок 2. *Якщо деякий круг Гершгоріна дійсної матриці не перетинається з іншими і має кратність* 1, *то він вміщує дійсне власне значення.*

Доведення. Якби власне значення було комплексним, то в крузі, центр якого лежить на дійсній осі, знаходилося б і комплексно спряжене власне значення, що неможливо. □

За допомогою кругів Гершгоріна можна отримати різноманітні оцінки для власних чисел. Одними з найпростіших є оцінки, які ми одержимо за допомогою нерівностей (6.1.9). З (4) і (6.1.9) випливає, що

$$||\lambda| - |a_{ii}|| \le \sum_{\substack{j=1 \\ j\ne i}}^{n} |a_{ij}|, \quad i = \overline{1,\ n},$$

або

$$|\lambda| \le |a_{ii}| + \sum_{\substack{j=1 \\ j\ne i}}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{ij}|, \quad i = \overline{1,\ n}. \tag{6}$$

Аналогічно на підставі нерівностей (5) і (6.1.9) маємо:

$$|\lambda| \le \sum_{j=1}^{n} |a_{ji}|, \quad i = \overline{1,\ n}. \tag{7}$$

Нехай α — максимальна з сум модулів елементів у рядках і β — максимальна з сум модулів елементів у стовпцях. Тоді нерівності (6) і (7) можна об'єднати:

$$|\lambda| \le \min(\alpha, \beta)\,. \tag{8}$$

Узагальнення нерівності (8) буде зроблено в теоремі 12.1.6.

У наступній теоремі йдеться про локалізацію власних значень матриці Адамара.

Теорема 6.3.4. *Якщо діагональні елементи матриці Адамара **A** є додатними* (*від'ємними*), *то всі її власні значення мають додатну* (*від'ємну*) *дійсну частину.*

Доведення. Центри всіх кругів Гершгоріна лежать на дійсній осі і ні один з них не вміщує точку $\lambda = 0$, оскільки матриця $\boldsymbol{A}$ є невиродженою. Тому всі круги (а з ними і власні значення) належать напівплощині $\operatorname{Re}\lambda > 0$ $(\operatorname{Re}\lambda < 0)$, якщо всі діагональні елементи a_{ii} є додатними (від'ємними). □

На підставі теореми 6.2.5 від задачі локалізації власних значень матриці $\boldsymbol{A}$ можна перейти до тієї ж задачі для матриці $\boldsymbol{S}^{-1}\boldsymbol{AS}$. Вдалий вибір матриці $\boldsymbol{S}$ може дати дуже точну оцінку. Особливо зручним є вибір матриці $\boldsymbol{S}$ у вигляді $\boldsymbol{S} = \operatorname{diag}(s_1,\ s_2,\ \dots,\ s_n)$, де числа $s_1,\ s_2,\ \dots,\ s_n$ є додатними. Неважко побачити, що матриця $\boldsymbol{S}^{-1}\boldsymbol{AS}$ має елементи $s_i^{-1}a_{ij}s_j$, а нерівності (4) і (5) для цієї матриці мають вигляд:

$$\begin{aligned} |\lambda - a_{ii}| &\le s_i^{-1}\sum_{\substack{j=1\\ j\ne i}}^{n} s_j\,|a_{ij}|, \quad i = \overline{1,\ n}, \\ |\lambda - a_{ii}| &\le s_i \sum_{\substack{j=1\\ j\ne i}}^{n} s_j^{-1}\,|a_{ji}|, \quad i = \overline{1,\ n}. \end{aligned} \tag{9}$$

Формули (9) показують, що діагональні елементи матриці $\boldsymbol{S}^{-1}\boldsymbol{AS}$ є такими ж, як і в матриці $\boldsymbol{A}$. Тому центри кругів Гершгоріна залишаються на місці. Одночасно радіуси кругів змінюються і задача полягає в такому виборі $\boldsymbol{S}$, що б вони зменшилися.

Приклад. Будемо розв'язувати задачу локалізації власних значень матриці

$$A = \begin{pmatrix} 1 & 3/8 \\ 3/8 & 2 \end{pmatrix}.$$

Оскільки матриця A симетрична, то її власні значення дійсні і кругами Гершгоріна є відрізки дійсної осі $|\lambda - 1| \le 3/8$, $|\lambda - 2| \le 3/8$, або $5/8 \le \lambda \le 11/8$, $13/8 \le \lambda \le 19/8$. Ці відрізки ізольовані, тому кожний з них вміщує по одному власному значенню (вони дорівнюють $\lambda_1 = 7/8$, $\lambda_2 = 17/8$).

Матриця SAS^{-1}, де $S = \operatorname{diag}(1,\ 2)$, дорівнює

$$SAS^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 3/8 \\ 3/8 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3/4 \\ 3/16 & 2 \end{pmatrix}.$$

На підставі формули (4) маємо нерівності $|\lambda - 1| \le 3/4$, $|\lambda - 2| \le$ $\le 3/16$, або $1/4 \le \lambda \le 7/4$, $29/16 \le \lambda \le 35/16$. Ці відрізки дійсної осі ізольовані, тому вони визначають оцінки для λ_1 і λ_2. Порівняння з попередніми оцінками свідчить про те, що оцінка для λ_2 поліпшилася за рахунок погіршення оцінки для λ_1. Якщо скористатися формулою (5), то одержимо нерівності $|\lambda - 1| \le 3/16$, $|\lambda - 2| \le 3/4$, або $13/16 \le \lambda \le 19/16$, $5/4 \le \lambda \le 11/4$. Тут, навпаки, оцінка для λ_1 поліпшилась, а оцінка для λ_2 стала гіршою.

Будемо вважати, що $S = \operatorname{diag}(1,\ 3)$, тоді

$$S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} 1 & 3/8 \\ 3/8 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 9/8 \\ 1/8 & 2 \end{pmatrix}$$

і далі у відповідності до формули (4) маємо нерівності $|\lambda - 1| \le$ $\le 9/8$, $|\lambda - 2| \le 1/8$, або $-1/8 \le \lambda \le 17/8$, $15/8 \le \lambda \le 17/8$. Тут відрізки дійсної осі не є ізольованими і тому можна стверджувати, що обидва власні значення вміщуються в об'єднанні цих відрізків, тобто $\lambda_1,\ \lambda_2 \in [-1/8,\ 17/8]$. У випадку формули (5) маємо нерівності $|\lambda - 1| \le 1/8$, $|\lambda - 2| \le 9/8$, або $7/8 \le \lambda \le 9/8$, $7/8 \le \lambda \le 25/8$. Тут відрізки дійсної осі також не є ізольованими і $\lambda_1,\ \lambda_2 \in [7/8,\ 25/8]$.

Таким чином, розглянувши другий варіант вибору матриці S, ми одержали значно гірший результат у порівнянні з першим варіантом, що свідчить про обмежені можливості поліпшення оцінок розглядуваним методом. $\triangle$

6.4. Спряжена матриця

6.4.1. Унітарний простір. Комплексний n-вимірний простір C_n називається *унітарним*, якщо для векторів цього простору виконані такі вимоги.

А. Кожній парі векторів $\boldsymbol{a},\ \boldsymbol{b} \in C_n$ зіставляється комплексне число $(\boldsymbol{a},\ \boldsymbol{b})$, яке називається *скалярним добутком* цих векторів.

В. Скалярний добуток задовольняє аксіомам ($\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ — вектори, α — комплексне число):

1) $(\boldsymbol{a},\ \boldsymbol{b}) = \overline{(\boldsymbol{b},\ \boldsymbol{a})}$;

2) $(\boldsymbol{a} + \boldsymbol{b},\ \boldsymbol{c}) = (\boldsymbol{a},\ \boldsymbol{c}) + (\boldsymbol{b},\ \boldsymbol{c})$;

3) $(\alpha \boldsymbol{a},\ \boldsymbol{b}) = \overline{\alpha}(\boldsymbol{a},\ \boldsymbol{b})$;

4) $(\boldsymbol{a},\ \boldsymbol{a}) \geq 0$, $(\boldsymbol{a},\ \boldsymbol{a}) = 0$ лише коли $\boldsymbol{a} = \mathbf{0}$.

Для векторів $\boldsymbol{a} = (\alpha_1 \ \ \dots \ \ \alpha_n)^T$, $\boldsymbol{b} = (\beta_1 \ \ \dots \ \ \beta_n)^T$ аксіоми 1) — 4) виконуються при такому визначенні скалярного добутку:

$$(\boldsymbol{a},\ \boldsymbol{b}) = \overline{\boldsymbol{a}}^T \boldsymbol{b} = \sum_{i=1}^{n} \overline{\alpha}_i \beta_i\ .$$

Вектор $\overline{\boldsymbol{a}}^T$ називається *спряженим* по відношенню до вектора $\boldsymbol{a}$ і позначається через $\boldsymbol{a}^*$. Очевидно, що

$$(\boldsymbol{a},\ \boldsymbol{b}) = \boldsymbol{a}^* \boldsymbol{b}, \quad (\boldsymbol{a},\ \alpha \boldsymbol{b}) = \boldsymbol{a}^* \alpha \boldsymbol{b} = \alpha(\boldsymbol{a},\ \boldsymbol{b}).$$

Подібно до спряженого вектора вводиться *спряжена* матриця $\boldsymbol{A}^* = \overline{\boldsymbol{A}}^T$. Легко перевіряються властивості спряженої матриці:

1) $(\boldsymbol{A} + \boldsymbol{B})^* = \boldsymbol{A}^* + \boldsymbol{B}^*$; 2) $(\alpha \boldsymbol{A})^* = \overline{\alpha} \boldsymbol{A}^*$; 3) $(\boldsymbol{AB})^* = \boldsymbol{B}^* \boldsymbol{A}^*$;

4) $(\boldsymbol{A}^*)^* = \boldsymbol{A}$; 5) $(\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^*)^{-1}$.

Тут у перших чотирьох формулах $\boldsymbol{A}$ і $\boldsymbol{B}$ — довільні матриці і α — комплексне число, а в останній формулі матриця $\boldsymbol{A}$ є невиродженою.

В унітарному просторі кут між векторами не визначається. Усі інші визначення і результати, наведені в главах 4 і 5 для евклідового простору, можна легко поширити на унітарний простір, якщо користуватися вказаним вище узагальненням скалярного добутку і поняттям спряженої матриці.

6.4.2. Біортонормовані системи векторів. Системи векторів $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ і $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ називаються *біортонормованими*, якщо

$$\boldsymbol{x}_i^* \boldsymbol{y}_j = \delta_{ij}, \quad i, j = \overline{1, n}. \tag{1}$$

Кожна з цих систем є лінійно незалежною. Щоб довести це, треба показати, що рівність $\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i = \mathbf{0}$ можлива лише при нульових значеннях коефіцієнтів α_i. Помноживши її справа по черзі на вектори $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$, одержимо з урахуванням (1), що $\alpha_i = 0, \quad i = \overline{1, n}$.

Кожна з біортонормованих систем є базисом у C_n. Якщо виражати довільний вектор через один з базисів, то коефіцієнти розкладу легко знайти за допомогою іншого базису. Наприклад, якщо вектор $\boldsymbol{a} \in C_n$ представлений у вигляді:

$$\boldsymbol{a} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i,$$

то коефіцієнти α_i знаходяться після множення справа цієї рівності по черзі на $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$. У результаті такої операції одержимо, що $\alpha_j = \boldsymbol{a}^* \boldsymbol{y}_j, \quad j = \overline{1, n}$.

6.4.3. Властивості власних значень і власних векторів матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$. Будемо досліджувати спільні властивості матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$. Перш за все зауважимо, що характеристичні многочлени матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$ мають комплексно спряжені коефіцієнти, тому власні значення матриці $\boldsymbol{A}^*$ будуть комплексно спряженими з власними значеннями $\boldsymbol{A}$. Якщо матриця $\boldsymbol{A}$ є дійсною, то власні значення матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$ збігаються. Про зв’язок між власними векторами цих матриць ідеться в наступній теоремі.

Теорема 6.4.1. *Якщо* $\boldsymbol{x}_i$ *— власний вектор матриці* $\boldsymbol{A}$*, відповідний до власного значення* λ_i*, і* $\boldsymbol{y}_j$ *— власний вектор матриці* $\boldsymbol{A}^*$*, відповідний до її власного значення* $\mu_j \neq \overline{\lambda}_i$*, то*

$$\boldsymbol{x}_i^* \boldsymbol{y}_j = 0. \tag{2}$$

Д о в е д е н н я. Обчислимо величину $\boldsymbol{x}_i^* \boldsymbol{A}^* \boldsymbol{y}_j$ двома способами. Очевидно, що

$$\boldsymbol{x}_i^* \boldsymbol{A}^* \boldsymbol{y}_j = \boldsymbol{x}_i^* \mu_j \boldsymbol{y}_j = \mu_j \boldsymbol{x}_i^* \boldsymbol{y}_j .$$

З іншого боку,

$$\boldsymbol{x}_i^* \boldsymbol{A}^* \boldsymbol{y}_j = (\boldsymbol{A}\,\boldsymbol{x}_i)^* \boldsymbol{y}_j = (\lambda_i \boldsymbol{x}_i)^* \boldsymbol{y}_j = \overline{\lambda}_i \boldsymbol{x}_i^* \boldsymbol{y}_j .$$

Тому

$$(\mu_j - \overline{\lambda}_i)\, \boldsymbol{x}_i^* \boldsymbol{y}_j = 0 .$$

Оскільки $\mu_j - \overline{\lambda}_i \neq 0,$ то маємо рівність (2). □

Нижче в теоремі 6.4.3 буде з’ясована особливість систем власних векторів матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$ в одному важливому випадку. Для цього нам знадобиться

Теорема 6.4.2. *Якщо деякий підпростір* U *простору* C_n *є інваріантним відносно лінійного перетворення* $\boldsymbol{A}$*, то* $U^{\perp}$ *є інваріантним відносно лінійного перетворення* $\boldsymbol{A}^*$*.*

Д о в е д е н н я. Нехай $\boldsymbol{x} \in U$, $\boldsymbol{y} \in U^{\perp}$. Тоді $\boldsymbol{A}\,\boldsymbol{x} \in U$ і $(\boldsymbol{A}\,\boldsymbol{x})^* \boldsymbol{y} = 0$. Звідси випливає, що $\boldsymbol{x}^* \boldsymbol{A}^* \boldsymbol{y} = 0,$ тобто $\boldsymbol{A}^* \boldsymbol{y} \in U^{\perp}$. □

Теорема 6.4.3. *Якщо* $\boldsymbol{A}$ *— матриця простої структури, то* $\boldsymbol{A}^*$ *також має просту структуру, причому можна так вибрати системи власних векторів матриць* $\boldsymbol{A}$ *і* $\boldsymbol{A}^*$*, щоб вони були біортонормованими.*

Д о в е д е н н я. Нехай $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ — лінійно незалежні власні вектори матриці $\boldsymbol{A}$, U — підпростір, породжений векторами

$$\boldsymbol{x}_1, \ \ldots, \ \boldsymbol{x}_{i-1}, \ \boldsymbol{x}_{i+1}, \ \ldots, \ \boldsymbol{x}_n , \tag{3}$$

і $\boldsymbol{y}_i \neq \boldsymbol{0}$ — вектор, який породжує одновимірний підпростір $U^{\perp}$. Оскільки на підставі теореми 6.4.2 підпростір $U^{\perp}$ є інваріантним відносно лінійного перетворення $\boldsymbol{A}^*$, то $\boldsymbol{A}^*\boldsymbol{y}_i = \mu\,\boldsymbol{y}_i$. Вектор $\boldsymbol{y}_i$, будучи ортогональним до векторів (3), не може бути одночасно ортогональним до вектора $\boldsymbol{x}_i$, оскільки на підставі висновку з теореми про ортогональний розклад вектор $\boldsymbol{y}_i$ був би нульовим. Отже, $\boldsymbol{x}_i^*\boldsymbol{y}_i \neq 0$.

Вектор $\boldsymbol{y}_i$ є власним вектором матриці $\boldsymbol{A}^*$. Аналогічно доводиться існування інших її власних векторів, їх загальна кількість дорівнює n.

Після множення векторів $\boldsymbol{x}_i$, $\boldsymbol{y}_i$ на належні числові множники можна одержати рівності $\boldsymbol{x}_i^*\boldsymbol{y}_j = \delta_{ij}$, $i,\ j = \overline{1,\ n}$. □

Виведемо тепер загальне представлення матриці простої структури, яке називають білінійним розкладом.

Теорема 6.4.4. *Нехай $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ і $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n$ — системи власних векторів матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$ відповідно, а $\lambda_1, \ldots, \lambda_n$ — власні значення матриці $\boldsymbol{A}$. Має місце білінійний розклад матриці $\boldsymbol{A}$*:

$$\boldsymbol{A} = \sum_{i=1}^{n}\lambda_i \boldsymbol{x}_i\,\boldsymbol{y}_i^*.$$

Д о в е д е н н я. З формули (6.2.6) знаходимо:

$$\boldsymbol{A}^* = (\boldsymbol{S}^{-1})^*\overline{\boldsymbol{\Lambda}}\,\boldsymbol{S}^*.$$

Звідси випливає, що матриця $(\boldsymbol{S}^{-1})^*$ складається з власних векторів матриці $\boldsymbol{A}^*$, відповідних до її власних значень $\overline{\lambda}_1, \ldots, \overline{\lambda}_k$. Позначимо її через $\boldsymbol{T}$, тоді

$$\boldsymbol{T} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n), \quad \boldsymbol{S}^{-1} = \boldsymbol{T}^*.$$

Тепер формулу (6.2.6) можна подати у вигляді:

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{T}^*.$$

Звідси знаходимо, що

$$\boldsymbol{A} = (\boldsymbol{x}_1 \ \ldots \ \boldsymbol{x}_n)\,\mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)\begin{pmatrix} \boldsymbol{y}_1^* \\ \ldots \\ \boldsymbol{y}_n^* \end{pmatrix} =$$

$$= (\boldsymbol{x}_1 \ \ldots \ \boldsymbol{x}_n)\begin{pmatrix} \lambda_1\,\boldsymbol{y}_1^* \\ \ldots \\ \lambda_n\,\boldsymbol{y}_n^* \end{pmatrix} = \sum_{i=1}^{n}\lambda_i\boldsymbol{x}_i\,\boldsymbol{y}_i^*. \ \square \qquad (4)$$

6.5. Спеціальні типи матриць

При діагоналізації матриць важливим для теоретичних досліджень і застосувань є той випадок, коли діагоналізуюча матриця є ортогональною. Це можливо в разі існування у матриці n-го порядку n ортонормованих власних векторів. Найпростішим і найважливішим випадком таких матриць є симетричні матриці. Проте ми не обмежимося вивченням цих матриць, а розглянемо в повному обсязі проблему існування вказаної вище ортонормованої системи. Це приведе нас до так званих нормальних матриць і їх важливих окремих випадків.

6.5.1. Унітарні, ермітові і косоермітові матриці. Розглянемо узагальнення деяких типів матриць на комплексний випадок. Узагальненням ортогональної матриці є *унітарна* матриця $\boldsymbol{U}$, яка задовольняє умові $\boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{E}$. Звідси випливає, що $\boldsymbol{U}^* = \boldsymbol{U}^{-1}$ і $\boldsymbol{U}\boldsymbol{U}^* = \boldsymbol{E}$. Властивості унітарної матриці подібні до властивостей ортогональної матриці, про які йшла мова в розділі 4.2.

Матриця $\boldsymbol{H}$ називається *ермітовою*, якщо вона задовольняє рівнянню $\boldsymbol{H} = \boldsymbol{H}^*$. У дійсному випадку ермітова матриця є симетричною матрицею. Оскільки $\det \boldsymbol{H} = \det \boldsymbol{H}^* = \det \overline{\boldsymbol{H}}$, то визначник ермітової матриці є дійсним числом.

Матриця $\boldsymbol{K}$ називається *косоермітовою*, якщо вона задовольняє співвідношенню $\boldsymbol{K} = -\boldsymbol{K}^*$. Косоермітова матриця є комплексним аналогом кососиметричної матриці. Очевидно, що діагональні елементи косоермітової матриці є уявними числами (нуль допускається). Якщо матриця $\boldsymbol{K}$ косоермітова, то матриця $\boldsymbol{H} = \boldsymbol{K}i$ буде ермітовою, бо $\boldsymbol{H}^* = \boldsymbol{K}^*(-i) = \boldsymbol{K}i = \boldsymbol{H}$. Тому властивості косоермітової матриці можна формулювати, виходячи з властивостей ермітової матриці.

6.5.2. Нормальні матриці. Матриця $\boldsymbol{A}$ називається *нормальною*, якщо вона переставна зі своєю спряженою, тобто

$$\boldsymbol{A}\boldsymbol{A}^* = \boldsymbol{A}^*\boldsymbol{A}. \tag{1}$$

Легко перевірити, що унітарні, ермітові, косоермітові і діагональні матриці є нормальними.

Теорема 6.5.1. *Якщо трикутна матриця є нормальною, то вона діагональна.*

Д о в е д е н н я наведемо для верхньої трикутної матриці $\boldsymbol{R}$. Випадок нижньої трикутної матриці можна розглянути аналогічно.

Скористаємося умовою рівності нулю діагональних елементів матриці $\boldsymbol{RR^* - R^*R}$. Оскільки

$$\{\boldsymbol{RR^* - R^*R}\}_{11} = |r_{11}|^2 + |r_{12}|^2 + \ldots + |r_{1n}|^2 - |r_{11}|^2 = 0,$$

то $r_{12} = \ldots = r_{1n} = 0$, тобто позадіагональні елементи першого рядка матриці $\boldsymbol{R}$ є нульовими.

З урахуванням рівності $r_{12} = 0$ далі маємо:

$$\{\boldsymbol{RR^* - R^*R}\}_{22} = |r_{22}|^2 + |r_{23}|^2 + \ldots + |r_{2n}|^2 - |r_{22}|^2 = 0,$$

тобто $r_{23} = \ldots = r_{2n} = 0$. Отже, нульовими є позадіагональні елементи другого рядка матриці $\boldsymbol{R}$.

Продовжуючи ці міркування далі, можна переконатися в тому, що матриця $\boldsymbol{R}$ є діагональною. ☐

6.5.3. Теорема Шура і розклад Шура. Для подальшого дослідження властивостей нормальних матриць нам знадобиться теорема про зведення матриці до трикутного вигляду за допомогою перетворення подібності з унітарною матрицею. При її доведенні використовується допоміжна теорема. Перш ніж перейти до її розгляду, сформулюємо визначення *квазітрикутної* матриці. Це є така квадратна блокова матриця, у якої діагональні блоки квадратні, а блоки, розміщені нижче або вище діагональних, є нульовими.

Теорема 6.5.2. *Нехай для комплексних матриць **A**, **B**, **X** виконується рівність*

$$\boldsymbol{AX} = \boldsymbol{XB}, \tag{2}$$

*де **X** є $n \times p$-матрицею рангу $p < n$, а **A** і **B** — квадратні матриці. Тоді існує унітарна матриця $\boldsymbol{U}_0$ така, що $\boldsymbol{U}_0^*\boldsymbol{A}\boldsymbol{U}_0$ є верхньою квазітрикутною матрицею*:

$$\boldsymbol{U}_0^*\boldsymbol{A}\boldsymbol{U}_0 = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{0} & \boldsymbol{C}_{22} \end{pmatrix}, \tag{3}$$

де $\boldsymbol{C}_{11}$ *— підматриця порядку p і* $\sigma(\boldsymbol{C}_{11}) = \sigma(\boldsymbol{B})$.

Д о в е д е н н я. Скористаємося комплексним аналогом формули (4.4.4) для матриці $\boldsymbol{X}$, який має вигляд:

$$\boldsymbol{X} = \boldsymbol{U}_0 \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix}. \tag{4}$$

Тут $\boldsymbol{U}_0$ — унітарна, а $\boldsymbol{R}$ — верхня трикутна матриці. Підставивши $\boldsymbol{X}$ з (4) у формулу (2), будемо мати:

$$\boldsymbol{U}_0^*\boldsymbol{A}\boldsymbol{U}_0 \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{B}. \tag{5}$$

Представимо матрицю $\boldsymbol{U}_0^*\boldsymbol{A}\boldsymbol{U}_0$ у блоковому вигляді:

$$\boldsymbol{U}_0^*\boldsymbol{A}\boldsymbol{U}_0 = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{pmatrix},$$

де $\boldsymbol{C}_{11}$ — підматриця порядку p. Тоді формулу (5) можна записати так:

$$\begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{B}.$$

Звідси знаходимо, що

$$\boldsymbol{C}_{11}\boldsymbol{R} = \boldsymbol{R}\boldsymbol{B}, \quad \boldsymbol{C}_{21}\boldsymbol{R} = \boldsymbol{0}. \tag{6}$$

Матриця $\boldsymbol{R}$ має порядок p і є невиродженою. Тому першу формулу (6) можна записати у вигляді рівності $\boldsymbol{C}_{11} = \boldsymbol{R}\boldsymbol{B}\boldsymbol{R}^{-1}$. З теореми 6.2.5 випливає, що $\sigma(\boldsymbol{C}_{11}) = \sigma(\boldsymbol{B})$. З другої рівності (6) через невиродженість матриці $\boldsymbol{R}$ випливає рівність $\boldsymbol{C}_{21} = \boldsymbol{0}$. □

Теорема 6.5.3 (теорема Шура). *Для довільної матриці* $\boldsymbol{A}$ *порядку n існує унітарна матриця* $\boldsymbol{U}$ *така, що матриця*

$$\boldsymbol{R} = \boldsymbol{U}^*\boldsymbol{A}\boldsymbol{U} \tag{7}$$

є верхньою трикутною.

Д о в е д е н н я. Якщо в теоремі 6.5.2 вважати, що $\boldsymbol{X}$ є власним вектором матриці $\boldsymbol{A}$, відповідним до власного значення λ_1 і $\boldsymbol{C}_{11} = \boldsymbol{B} = (\lambda_1)$, то

$$\boldsymbol{U}_0^* \boldsymbol{A} \boldsymbol{U}_0 = \begin{pmatrix} \lambda_1 & \boldsymbol{C}_{12} \\ \boldsymbol{0} & \boldsymbol{C}_{22} \end{pmatrix}. \tag{8}$$

Ця формула дозволяє довести теорему методом математичної індукції.

Очевидно, що теорема справедлива для $n = 1$. Припустимо, що вона має місце для матриць порядку $n - 1$. Це означає, що для матриці $\boldsymbol{C}_{22}$ можна знайти унітарну матрицю $\boldsymbol{U}_1$ таку, що матриця $\boldsymbol{U}_1^* \boldsymbol{C}_{22} \boldsymbol{U}_1$ буде верхньою трикутною. Неважко перевірити, що матриця $\boldsymbol{U} = \boldsymbol{U}_0 \operatorname{diag}(1, \boldsymbol{U}_1)$ є унітарною. Обчислимо $\boldsymbol{U}^* \boldsymbol{A} \boldsymbol{U}$, скориставшись формулою (8):

$$\boldsymbol{U}^* \boldsymbol{A} \boldsymbol{U} = \operatorname{diag}(1,\ \boldsymbol{U}_1^*) \boldsymbol{U}_0^* \boldsymbol{A} \boldsymbol{U}_0 \operatorname{diag}(1,\ \boldsymbol{U}_1) =$$

$$= \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_1^* \end{pmatrix} \begin{pmatrix} \lambda_1 & \boldsymbol{C}_{12} \\ \boldsymbol{0} & \boldsymbol{C}_{22} \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & \boldsymbol{C}_{12} \boldsymbol{U}_1 \\ \boldsymbol{0} & \boldsymbol{U}_1^* \boldsymbol{C}_{22} \boldsymbol{U}_1 \end{pmatrix}. \tag{9}$$

Оскільки матриця $\boldsymbol{U}_1^* \boldsymbol{C}_{22} \boldsymbol{U}_1$ за припущенням є верхньою трикутною, то верхньою трикутною є і матриця з правої частини рівностей (9). Отже, у відповідності до методу математичної індукції теорема вважається доведеною. □

З формули (7) випливає, що

$$\boldsymbol{A} = \boldsymbol{U} \boldsymbol{R} \boldsymbol{U}^*, \tag{10}$$

тобто матриця $\boldsymbol{A}$ представляється у вигляді добутку трьох множників, які, взагалі кажучи, є комплексними. Існує дійсний аналог цього представлення.

***Теорема 6.5.4* (*розклад Шура*).** *Для дійсної матриці* $\boldsymbol{A}$ *порядку* n *існує ортогональна матриця* $\boldsymbol{Q}$ *така, що* $\boldsymbol{Q}^T \boldsymbol{A} \boldsymbol{Q}$ *є верхньою квазітрикутною матрицею*:

$$Q^T AQ = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ 0 & R_{22} & \dots & R_{2m} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & R_{mm} \end{pmatrix}, \qquad (11)$$

де кожний діагональний блок R_{jj} *є або підматрицею першого порядку, або підматрицею другого порядку з комплексно спряженими власними значеннями.*

Д о в е д е н н я. Комплексні власні значення дійсної матриці A повинні входити до спектру $\sigma(A)$ матриці A комплексно спряженими парами. Нехай k — кількість комплексно спряжених пар у $\sigma(A)$. Доведемо теорему методом математичної індукції.

При $k = 0$ всі власні значення є дійсними. Доведення теорем 6.5.2 і 6.5.3 можна виконати, вважаючи унітарну матрицю U ортогональною. При цьому рівність (11) випливає з рівності (7).

Розглянемо випадок, коли $k \geq 1$. Нехай $\lambda = \mu + \nu i$ — власне значення матриці A і $y + z\,i$ — відповідний до нього власний вектор, де μ, ν — дійсні числа, а y, z — дійсні вектори. З рівності $A(y + z\,i) = (\mu + \nu\,i)\,(y + z\,i)$ знаходимо, що

$$A(y \quad z) = (y \quad z)\begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix}. \qquad (12)$$

Застосуємо до рівності (12) теорему 6.5.2. Зважаючи на те, що в цій рівності всі матриці є дійсними, унітарна матриця U_0 у формулі (3) теж буде дійсною, тобто ортогональною. Позначаючи її через Q_0, одержимо:

$$Q_0^T AQ_0 = \begin{pmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{pmatrix}, \qquad (13)$$

де C_{11} — підматриця другого порядку з власними значеннями $\lambda, \overline{\lambda}$.

У відповідності до припущення індукції існує ортогональна матриця Q_1 така, що $Q_1^T C_{22} Q_1$ є матрицею вказаної вище структури. Скориставшись формулою (13) і ортогональною матрицею $Q = Q_0 \operatorname{diag}(E_2,\ Q_1)$, обчислимо $Q^T AQ$:

$$\boldsymbol{Q}^T\boldsymbol{AQ} = \operatorname{diag}(\boldsymbol{E}_2,\ \boldsymbol{Q}_1^T)\,\boldsymbol{Q}_0^T\boldsymbol{AQ}_0\operatorname{diag}(\boldsymbol{E}_2,\ \boldsymbol{Q}_1) =$$

$$= \begin{pmatrix} \boldsymbol{E}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_1^T \end{pmatrix} \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{0} & \boldsymbol{C}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{E}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12}\boldsymbol{Q}_1 \\ \boldsymbol{0} & \boldsymbol{Q}_1^T\boldsymbol{C}_{22}\boldsymbol{Q}_1 \end{pmatrix}.$$

Права частина цих рівностей має структуру, яка відповідає умові теореми. Отже, теорема доведена методом математичної індукції. □

6.5.4. Властивості спеціальних типів матриць. Тепер ми доведемо дві теореми, які вміщують основний результат стосовно нормальних матриць і їх окремих випадків.

Теорема 6.5.5. *Для того, щоб матриця* $\boldsymbol{A}$ *порядку* n *була нормальною, необхідно і достатньо, щоб вона мала* n *ортонормованих власних векторів.*

Д о в е д е н н я. Н е о б х і д н і с т ь. У відповідності до теореми Шура нормальну матрицю $\boldsymbol{A}$ зведемо до верхньої трикутної матриці $\boldsymbol{R}$ за допомогою перетворення подібності з унітарною матрицею $\boldsymbol{U}$. З рівності (7) випливає, що

$$\boldsymbol{RR}^* - \boldsymbol{R}^*\boldsymbol{R} = \boldsymbol{U}^*(\boldsymbol{AA}^* - \boldsymbol{A}^*\boldsymbol{A})\boldsymbol{U}\,. \tag{14}$$

Оскільки матриця $\boldsymbol{A}$ нормальна, тобто виконується рівність (1), то з формули (14) випливає, що матриця $\boldsymbol{R}$ теж є нормальною. Унаслідок теореми 6.5.1 вона є діагональною.

Тепер на підставі рівності (10) і теореми 6.2.6 стовпці діагоналізуючої матриці $\boldsymbol{U}$ можна вважати ортонормованими власними векторами матриці $\boldsymbol{A}$.

Д о с т а т н і с т ь. Припустимо, що стовпці унітарної матриці $\boldsymbol{U}$ є власними векторами матриці $\boldsymbol{A}$. Тоді за допомогою перетворення подібності матриця $\boldsymbol{A}$ може бути зведена до діагональної матриці $\boldsymbol{R}$, тобто $\boldsymbol{U}^*\boldsymbol{AU} = \boldsymbol{R}$. Діагональна матриця нормальна, тому нормальність матриці $\boldsymbol{A}$ випливає з рівності (14). □

З теорем 6.5.5 і 6.2.6 випливає, що будь-яку нормальну матрицю можна діагоналізувати за допомогою унітарної матриці.

У наступній теоремі мова йде про особливості власних значень окремих типів матриць.

Теорема 6.5.6. *Власні значення унітарної, ермітової і косоермітової матриць є відповідно числами, рівними за модулем одиниці, дійсними або уявними.*

Д о в е д е н н я. Нехай матриця $\boldsymbol{A}$ є унітарною, ермітовою або косоермітовою, тобто задовольняє одній з рівностей

$$\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{E}, \quad \boldsymbol{A} = \boldsymbol{A}^*, \quad \boldsymbol{A} = -\boldsymbol{A}^*. \tag{15}$$

Скористаємося представленям (10) нормальної матриці $\boldsymbol{A}$, маючи на увазі, що $\boldsymbol{R}$ — діагональна матриця. Підставляючи $\boldsymbol{A}$ з рівності (10) у формули (15), будемо мати відповідно:

$$\boldsymbol{R}^*\boldsymbol{R} = \boldsymbol{E}, \quad \boldsymbol{R} = \boldsymbol{R}^*, \quad \boldsymbol{R} = -\boldsymbol{R}^*, \tag{16}$$

тобто якщо матриця $\boldsymbol{A}$ є унітарною, ермітовою або косоермітовою, то такою ж є і матриця $\boldsymbol{R}$. Але з рівностей (16) випливає, що діагональна матриця буде такою тоді і тільки тоді, коли її діагональні елементи є відповідно числами, рівними за модулем одиниці, дійсними або уявними. Оскільки діагональні елементи $\boldsymbol{R}$ є власними значеннями матриці $\boldsymbol{A}$, то теорема доведена. □

6.5.5. Симетричні матриці. Як зазначалося вище, симетричні матриці є найбільш важливим випадком нормальних матриць. Розглянемо окремо для таких матриць питання про їх діагоналізацію. Будемо вважати у відповідності до теореми 6.5.5, що всі власні вектори симетричної матриці утворюють ортонормовану систему.

Теорема 6.5.7. *Для того, щоб симетрична матриця* $\boldsymbol{A}$ *була представлена у вигляді*:

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T, \tag{17}$$

де $\boldsymbol{Q}$ *— ортогональна і* $\boldsymbol{\Lambda}$ *— діагональна матриці, необхідно і достатньо, щоб стовпці матриці* $\boldsymbol{Q}$ *були власними векторами, відповідними до власних значень — діагональних елементів матриці* $\boldsymbol{\Lambda}$.

Д о в е д е н н я аналогічне доведенню теореми 6.2.6, в якій треба вважати діагоналізуючу матрицю $\boldsymbol{S}$ рівною $\boldsymbol{Q}$. □

Необхідно дати пояснення стосовно власних векторів, відповідних до кратного власного значення. Їх чисельність дорівнює кратності власного значення і вони ортогональні до інших власних векторів, але, будучи визначеними з відповідної системи лінійних рівнянь, ці вектори, взагалі кажучи, не є попарно ортогональними. Проте за допомогою ортогоналізації Грама — Шмідта їх можна перетворити в ортонормовану систему, оскільки будь-яка ненульова лінійна комбінація власних векторів, відповідних до деякого власного значення, також є власним вектором, відповідним до цього власного значення. Розглянемо відповідний

П р и к л а д. За допомогою перетворення подібності з ортогональною матрицею діагоналізуємо симетричну матрицю

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

Обчислення показують, що власними значеннями і відповідними власними векторами матриці $\boldsymbol{A}$ є

$$\lambda_1 = -1, \quad \boldsymbol{x}_1 = (1 \quad -1 \quad 1)^T,$$

$$\lambda_2 = \lambda_3 = 2, \quad \boldsymbol{x}_2 = (1 \quad 1 \quad 0)^T, \quad \boldsymbol{x}_3 = (-1 \quad 0 \quad 1)^T.$$

Вектор $\boldsymbol{x}_1$ є ортогональним до векторів $\boldsymbol{x}_2$ і $\boldsymbol{x}_3$, проте вектори $\boldsymbol{x}_2$ і $\boldsymbol{x}_3$ не є ортогональними. Виходячи з векторів $\boldsymbol{x}_2$ і $\boldsymbol{x}_3$, побудуємо методом Грама — Шмідта ортогональну систему векторів

$$\boldsymbol{b}_1 = \boldsymbol{x}_2 = (1 \quad 1 \quad 0)^T, \quad \boldsymbol{b}_2 = \frac{1}{2}\boldsymbol{x}_2 + \boldsymbol{x}_3 = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}^T.$$

За допомогою ортогональної системи векторів $\boldsymbol{x}_1, \boldsymbol{b}_1, \boldsymbol{b}_2$ утворимо ортогональну матрицю $\boldsymbol{Q}$:

$$\boldsymbol{Q} = \begin{pmatrix} \dfrac{\boldsymbol{x}_1}{|\boldsymbol{x}_1|} & \dfrac{\boldsymbol{b}_1}{|\boldsymbol{b}_1|} & \dfrac{\boldsymbol{b}_2}{|\boldsymbol{b}_2|} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & \sqrt{2/3} \end{pmatrix}.$$

Тепер шукане представлення матриці $\boldsymbol{A}$ виглядає так:

$$A = Q\Lambda Q^T =$$

$$= \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & \sqrt{2/3} \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & 1/\sqrt{6} & \sqrt{2/3} \end{pmatrix}. \triangle$$

З теореми 6.4.4 випливає, що для симетричної матриці $\boldsymbol{A}$ з власними значеннями $\lambda_1, \ldots, \lambda_n$ і ортонормованими власними векторами $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ виконується співвідношення

$$\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i \boldsymbol{x}_i^T . \tag{18}$$

Воно може бути також одержане безпосередньо з формули (17) аналогічно рівності (6.4.4).

Формула (18) визначає *спектральний розклад* симетричної матриці. Аналогічний вигляд має спектральний розклад будь-якої нормальної матриці.

6.5.6. Ортогональні і антисиметричні матриці. З теорем 6.5.5 і 6.2.6 випливає, що симетричну, антисиметричну і ортогональну матриці можна діагоналізувати за допомогою унітарної матриці. Одночасно, розгядаючи теорему 6.5.7, ми побачили, що симетричну матрицю можна діагоналізувати за допомогою ортогональної матриці і представити її у вигляді добутку трьох дійсних матриць. Для ортогональної і антисиметричної матриць таке представлення в загальному випадку не існує, бо серед їх власних значень можуть бути комплексні (уявні) числа. Проте дійсний аналог представлення (17) для цих матриць існує подібно до того, як розклад Шура (11) є дійсним аналогом розкладу (7). Наступні теореми присвячені цьому питанню.

Теорема 6.5.8. *Для ортогональної матриці $\boldsymbol{A}$ існує ортогональна матриця $\boldsymbol{Q}$ така, що $\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q}$ є квазідіагональною матрицею*:

$$\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q} = \mathrm{diag}(\boldsymbol{R}_{11}, \boldsymbol{R}_{22}, \ldots, \boldsymbol{R}_{mm}), \tag{19}$$

де кожний діагональний блок $\boldsymbol{R}_{jj}$ є або підматрицею першого порядку, рівною одному з власних значень ± 1 матриці $\boldsymbol{A}$, або підматрицею обертання другого порядку $\boldsymbol{T}_{12}(\varphi_j)$, відповідною до комплексно спряжених власних значень $e^{\pm i\varphi_j}$ матриці $\boldsymbol{A}$.

Д о в е д е н н я. У відповідності до теореми 6.5.6 власні значення ортогональної матриці $\boldsymbol{A}$ дорівнюють 1, –1 або входять до її спектру $\sigma(\boldsymbol{A})$ комплексно спряженими парами $e^{i\varphi_j}, e^{-i\varphi_j}$. Якщо всі власні значення матриці $\boldsymbol{A}$ є дійсними, то такими ж є і її ортонормовані власні вектори. Тоді матрицю $\boldsymbol{A}$ можна діагоналізувати за допомогою ортогональної матриці $\boldsymbol{Q}$ і це доводить формулу (19).

Розглянемо випадок, коли матриця $\boldsymbol{A}$ має комплексні власні значення. З формули (13), обидві частини якої є дійсними матрицями, одержується рівність

$$(\boldsymbol{Q}_0^T\boldsymbol{A}\boldsymbol{Q}_0)^T = \begin{pmatrix} \boldsymbol{C}_{11}^T & \boldsymbol{0}^T \\ \boldsymbol{C}_{12}^T & \boldsymbol{C}_{22}^T \end{pmatrix}. \tag{20}$$

Перемножаючи рівності (13) і (20), будемо мати:

$$\boldsymbol{Q}_0^T\boldsymbol{A}\boldsymbol{Q}_0(\boldsymbol{Q}_0^T\boldsymbol{A}\boldsymbol{Q}_0)^T = \begin{pmatrix} \boldsymbol{C}_{11}\boldsymbol{C}_{11}^T + \boldsymbol{C}_{12}\boldsymbol{C}_{12}^T & \boldsymbol{C}_{12}\boldsymbol{C}_{22}^T \\ \boldsymbol{C}_{22}\boldsymbol{C}_{12}^T & \boldsymbol{C}_{22}\boldsymbol{C}_{22}^T \end{pmatrix}.$$

З огляду на ортогональність матриці $\boldsymbol{Q}_0^T\boldsymbol{A}\boldsymbol{Q}_0$ ліва частина цієї рівності є одиничною матрицею, тому

$$\boldsymbol{C}_{11}\boldsymbol{C}_{11}^T + \boldsymbol{C}_{12}\boldsymbol{C}_{12}^T = \boldsymbol{E}, \quad \boldsymbol{C}_{12}\boldsymbol{C}_{22}^T = \boldsymbol{0}, \quad \boldsymbol{C}_{22}\boldsymbol{C}_{22}^T = \boldsymbol{E}. \tag{21}$$

Третя рівність (21) показує, що $\boldsymbol{C}_{22}$ — ортогональна матриця, а з другої рівності випливає, що $\boldsymbol{C}_{12} = \boldsymbol{0}$. Тепер перша рівність (21) означає ортогональність матриці $\boldsymbol{C}_{11}$. Отже,

$$\boldsymbol{Q}_0^T\boldsymbol{A}\boldsymbol{Q}_0 = \operatorname{diag}(\boldsymbol{C}_{11}, \boldsymbol{C}_{22}), \tag{22}$$

причому спектр матриці $\boldsymbol{C}_{11}$ на підсаві теореми 6.5.2 складається з чисел $e^{i\varphi}$ і $e^{-i\varphi}$.

Застосуємо до обох частин рівності (22) перетворення подібності з матрицею $\boldsymbol{Q}_1 = \mathrm{diag}(\boldsymbol{Q}_{11}, \boldsymbol{E})$, де $\boldsymbol{Q}_{11}$ — ортогональна матриця така, що $\boldsymbol{Q}_{11}^T \boldsymbol{C}_{11} \boldsymbol{Q}_{11} = \boldsymbol{T}_{12}(\varphi)$. Остання рівність є можливою, бо спектри матриць $\boldsymbol{C}_{11}$ і $\boldsymbol{T}_{12}(\varphi)$ є однаковими. У результаті вказаного перетворення подібності одержимо:

$$\boldsymbol{Q}_1^T \boldsymbol{Q}_0^T \boldsymbol{A} \boldsymbol{Q}_0 \boldsymbol{Q}_1 = \mathrm{diag}(\boldsymbol{T}_{12}(\varphi), \boldsymbol{C}_{22}). \tag{23}$$

Очевидно, що матриця $\boldsymbol{Q}_1$ є отогональною. Тому рівність (23) свідчить про можливість звести матрицю $\boldsymbol{A}$ за допомогою перетворення подібності з отогональною матрицею $\boldsymbol{Q}_0 \boldsymbol{Q}_1$ до квазідіагональної матриці з двома діагональними блоками, один з яких є матрицею обертання.

Далі доведення завершується у відповідності до методу математичної індукції подібно до теореми 6.5.4. Отже, для будь-якої ортогональної матриці $\boldsymbol{A}$ в загальному випадку існує представлення

$$\boldsymbol{A} = \boldsymbol{Q}\mathrm{diag}(\boldsymbol{T}_{12}(\varphi_1), ..., \boldsymbol{T}_{12}(\varphi_m), 1, ..., 1, -1, ..., -1)\boldsymbol{Q}^T. \ \square \tag{24}$$

Теорема 6.5.9. *Для антисиметричної матриці* $\boldsymbol{A}$ *існує ортогональна матриця* $\boldsymbol{Q}$ *така, що виконується рівність* (*19*)*, де кожний діагональний блок* $\boldsymbol{R}_{jj}$ *є або нульовою підматрицею першого порядку, або*

$$\boldsymbol{R}_{jj} = \boldsymbol{R}_{jj}(\nu_j) = \begin{pmatrix} 0 & \nu_j \\ -\nu_j & 0 \end{pmatrix}, \tag{25}$$

причому деякі з величин ν_j *можуть бути нулями.*

Д о в е д е н н я. Транспонуємо обидві частини розкладу Шура. З урахуванням рівності $\boldsymbol{A}^T = -\boldsymbol{A}$ одержимо:

$$-\boldsymbol{Q}^T \boldsymbol{A} \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{R}_{11}^T & \mathbf{0}^T & ... & \mathbf{0}^T \\ \boldsymbol{R}_{12}^T & \boldsymbol{R}_{22}^T & ... & \mathbf{0}^T \\ ... & ... & ... & ... \\ \boldsymbol{R}_{1m}^T & \boldsymbol{R}_{2m}^T & ... & \boldsymbol{R}_{mm}^T \end{pmatrix}.$$

Порівнюючи цю рівність з (11),легко побачити, що позадіагональні підматриці є нульомими, а діагональні — антисиметричними. При цьому підматриці першого порядку повинні бути нульовими, а підматриці другого порядку — матрицями (25).

Напишемо вирази для антисиметричних матриць парного і непарного порядків.У першому випадку маємо:

$$\boldsymbol{A} = \boldsymbol{Q}\,\mathrm{diag}(\boldsymbol{R}_{11}(\nu_1), \boldsymbol{R}_{22}(\nu_2), ..., \boldsymbol{R}_{mm}(\nu_m))\boldsymbol{Q}^T, \tag{26}$$

а в другому —

$$\boldsymbol{A} = \boldsymbol{Q}\,\mathrm{diag}\big(\boldsymbol{R}_{11}(\nu_1), \boldsymbol{R}_{22}(\nu_2), ..., \boldsymbol{R}_{mm}(\nu_m), \boldsymbol{0}\big)\boldsymbol{Q}^T. \tag{27}$$

В останній рівності враховано, що антисиметрична матриця непарного порядку обов’язково має принаймні одне нульове власне значення.Як випливає з доведення теореми 6.5.4, деякі з величин ν_j у формулах (26) і (27) можуть бути нулями.

Висновок. *Ранг кососиметричної матриці є парним числом.*

Д о в е д е н н я випливає з формул (26) і (27) з урахуванням висновку з теореми 3.4.4.□

Наприкінці вкажемо на одну особливість власних векторів ортогональної і антисиметричної матриць, відповідних до комплексних (уявних) власних значень.

Представимо рівність (24) у вигляді:

$$\boldsymbol{AQ} = \boldsymbol{Q}\mathrm{diag}(\boldsymbol{T}_{12}(\varphi_1), ...).$$

Нехай $\boldsymbol{y}$ і $\boldsymbol{z}$ — два перші стовпці матриці $\boldsymbol{Q}$. З останньої рівності випливає рівність(12), де $\mu = \cos\varphi_1, v = -\sin\varphi_1$. Це означає, що дійсна та уявна частини власного вектора матриці $\boldsymbol{A}$, відповідного до власного значення $e^{-i\varphi_1}$, є ортогональними. Ці міркування легко поширити на власні вектори, відповідні до інших комплексних власних значень.

Аналогічного висновку можна дійти стосовно власного вектора антисиметричної матриці $\boldsymbol{A}$, відповідного до уявного власного значення $\nu i \neq 0$, якщо скористатися формулами (26) і (27).

6.6. Сингулярний розклад матриці

6.6.1. Сингулярний розклад. Основою для подальшого викладення є

Теорема 6.6.1. *Для будь-якої дійсної матриці* $\boldsymbol{A}$ *порядку n існують ортогональні матриці* $\boldsymbol{Q}_1$ *і* $\boldsymbol{Q}_2$ *такі, що*

$$\boldsymbol{A} = \boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{Q}_2, \qquad \boldsymbol{\Sigma} = \mathrm{diag}\,(\sigma_1, \sigma_2, \ldots, \sigma_n), \tag{1}$$

причому

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > \sigma_{r+1} = \ldots = \sigma_n = 0, \tag{2}$$

де $r = r(\boldsymbol{A})$. *Зокрема, якщо матриця* $\boldsymbol{A}$ *невироджена, то*

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0. \tag{3}$$

Д о в е д е н н я. Припустимо, що λ_i — будь-яке власне значення матриці $\boldsymbol{A}\boldsymbol{A}^T$ і $\boldsymbol{x}_i$ — відповідний до нього власний вектор, тобто $\boldsymbol{A}\boldsymbol{A}^T \boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$. Звідси маємо:

$$\boldsymbol{x}_i^T \boldsymbol{A}\boldsymbol{A}^T \boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i^T \boldsymbol{x}_i = \lambda_i |\, \boldsymbol{x}_i|^2. \tag{4}$$

З іншого боку,

$$\boldsymbol{x}_i^T \boldsymbol{A}\boldsymbol{A}^T \boldsymbol{x}_i = (\boldsymbol{A}^T \boldsymbol{x}_i)^T \boldsymbol{A}^T \boldsymbol{x}_i = |\, \boldsymbol{A}^T \boldsymbol{x}_i|^2 \geq 0. \tag{5}$$

З (4) і (5) випливає, що всі власні значення матриці $\boldsymbol{A}\boldsymbol{A}^T$ невід'ємні, тому їх можна позначити через $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, одночасно врахувавши умови (2) або (3).

З теореми 6.5.7 випливає існування ортогональної матриці $\boldsymbol{Q}_1$ такої, що

$$\boldsymbol{Q}_1^T \boldsymbol{A}\boldsymbol{A}^T \boldsymbol{Q}_1 = \boldsymbol{\Sigma}^2, \tag{6}$$

причому на підставі висновку з теореми 3.4.4 і формули (5.1.8) кількість ненульових діагональних елементів матриці $\boldsymbol{\Sigma}^2$ дорівнює *r*.

Визначимо матрицю $\boldsymbol{G}$ як

$$\boldsymbol{G} = \boldsymbol{Q}_1^T \boldsymbol{A}. \tag{7}$$

Оскільки $\boldsymbol{G}\boldsymbol{G}^T = \boldsymbol{Q}_1^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{Q}_1 = \boldsymbol{\Sigma}^2$, то рядки $\boldsymbol{g}_{1*},\ldots,\boldsymbol{g}_{r*}$ матриці $\boldsymbol{G}$ є ортогональними, причому $|\boldsymbol{g}_{i*}| = \sigma_i,\ i = \overline{1,\ r}$, а інші рядки є нульовими.

Побудуємо ортонормовану систему векторів $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n$ таким чином, щоб $\boldsymbol{f}_i = \sigma_i^{-1}\boldsymbol{g}_{i*}$ для $i = \overline{1,\ r}$, а інші вектори $\boldsymbol{f}_i$ виберемо з умови, щоб уся сукупність $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n$ складалася з ортонормованих векторів. Таку побудову можна виконати, користуючись схемою Грама — Шмідта.

Якщо взяти за рядки матриці $\boldsymbol{Q}_2$ вектори $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n$, то така матриця буде ортогональною. Оскільки $\boldsymbol{g}_{i*} = \sigma_i\boldsymbol{f}_i,\ i = \overline{1,\ n}$, то

$$\boldsymbol{G} = \boldsymbol{\Sigma}\boldsymbol{Q}_2. \tag{8}$$

З рівностей (7) і (8) випливає перша рівність (1). □

Розклад (1) називається *сингулярним розкладом* матриці $\boldsymbol{A}$, а числа $\sigma_1,\ldots,\sigma_n$ — її *сингулярними числами*. Теорема про сингулярний розклад може бути поширена на прямокутну матрицю.

Теорема 6.6.2. *Для будь-якої дійсної $m\times n$-матриці $\boldsymbol{A}$ рангу r існує сингулярний розклад*

$$\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{\Sigma}\boldsymbol{Q}_2, \tag{9}$$

де $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2$ — ортогональні матриці, $\boldsymbol{\Sigma}$ — діагональна $m\times n$-матриця з незростаючими невід'ємними елементами σ_i (сингулярними числами матриці $\boldsymbol{A}$) на головній діагоналі, причому $\sigma_i = 0$ при $i > r$.

Д о в е д е н н я. Нехай $m < n$, тоді в рівності (6) порядок матриць $\boldsymbol{Q}_1$, $\boldsymbol{A}\boldsymbol{A}^T$ і $\boldsymbol{\Sigma}$ дорівнює *m*. Матрицю $\boldsymbol{\Sigma}$ далі будемо позначати через $\boldsymbol{\Sigma}_0$. За допомогою блокових матриць порядку *n*

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{0} \end{pmatrix},\ \boldsymbol{S} = \begin{pmatrix} \boldsymbol{Q}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},\ \Pi^2 = \begin{pmatrix} \boldsymbol{\Sigma}_0^2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \tag{10}$$

запишемо (6) у вигляді рівності $\boldsymbol{S}^T\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{S} = \Pi^2$. Якщо визначити матрицю $\boldsymbol{G}$ як

$$\boldsymbol{G} = \boldsymbol{S}^T \boldsymbol{B}, \tag{11}$$

то $\boldsymbol{G}\boldsymbol{G}^T = \boldsymbol{S}^T \boldsymbol{B}\boldsymbol{B}^T \boldsymbol{S} = \Pi^2$. Подібно до попередньої теореми маємо:

$$\boldsymbol{G} = \Pi \boldsymbol{Q}_2, \tag{12}$$

де $\boldsymbol{Q}_2$ — ортогональна матриця порядку n.

З рівностей (11) і (12) випливає, що $\boldsymbol{S}^T \boldsymbol{B} \boldsymbol{Q}_2^T = \boldsymbol{\Pi}$. Підставивши в цю формулу вирази для матриць $\boldsymbol{S}$, $\boldsymbol{B}$ і Π з (10), одержимо:

$$\begin{pmatrix} \boldsymbol{Q}_1^T \boldsymbol{A} \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}_2^T = \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

звідки $\boldsymbol{Q}_1^T \boldsymbol{A} \boldsymbol{Q}_2^T = (\boldsymbol{\Sigma}_0 \quad \boldsymbol{0})$, або

$$\boldsymbol{A} = \boldsymbol{Q}_1 (\boldsymbol{\Sigma}_0 \quad \boldsymbol{0}) \boldsymbol{Q}_2. \tag{13}$$

Матриця $(\boldsymbol{\Sigma}_0 \quad \boldsymbol{0})$ є тією матрицею $\boldsymbol{\Sigma}$, про яку йде мова у формулюванні теореми. Тому формула (13) визначає сингулярний розклад матриці $\boldsymbol{A}$.

Випадок $m > n$ зводиться до випадку $m < n$, якщо застосувати формулу (13) до матриці $\boldsymbol{A}^T$, а потім виконати транспонування обох частин рівності. ☐

6.6.2. Сингулярні базиси. Розглянемо висновки, які випливають з існування сингулярного розкладу матриці.

Теорема 6.6.3. *Якщо А — довільна матриця і* $\mathbf{A} = \boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{Q}_2$ *— її сингулярний розклад, то*:

1) ненульові власні значення матриць $\boldsymbol{A}\boldsymbol{A}^T$ *і* $\boldsymbol{A}^T \boldsymbol{A}$ *однакові*;

2) стовпці матриці $\boldsymbol{Q}_1$ *є ортонормованими власними векторами матриці* $\boldsymbol{A}\boldsymbol{A}^T$;

3) стовці матриці $\boldsymbol{Q}_2^T$ *є ортонормованими власними векторами матриці* $\boldsymbol{A}^T \boldsymbol{A}$.

Д о в е д е н н я. За допомогою формули (9) знаходимо:

$$\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{Q}_2 (\boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{Q}_2)^T = \boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{\Sigma}^T \boldsymbol{Q}_1^T, \tag{14}$$

$$\boldsymbol{A}^T \boldsymbol{A} = (\boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{Q}_2)^T \boldsymbol{Q}_1 \boldsymbol{\Sigma} \boldsymbol{Q}_2 = \boldsymbol{Q}_2^T \boldsymbol{\Sigma}^T \boldsymbol{\Sigma} \boldsymbol{Q}_2. \tag{15}$$

Матриці $\boldsymbol{\Sigma\Sigma}^T$ і $\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}$ є діагональними і їх ненульовими діагональними елементами є числа $\sigma_1^2, \ldots, \sigma_r^2$. На підставі теореми 6.5.7 ці числа є власними значеннями матриць $\boldsymbol{AA}^T$ і $\boldsymbol{A}^T\boldsymbol{A}$. Отже, ненульові власні значення матриць $\boldsymbol{AA}^T$ і $\boldsymbol{A}^T\boldsymbol{A}$ збігаються.

На підставі тієї ж теореми з формули (14) випливає, що стовпці матриці $\boldsymbol{Q}_1$ є ортонормованими власними векторами матриці $\boldsymbol{AA}^T$, а формула (15) доводить, що стовпці матриці $\boldsymbol{Q}_2^T$ є ортонормованими власними векторами матриці $\boldsymbol{A}^T\boldsymbol{A}$. □

Ортонормовані системи векторів, які утворюються стовпцями матриць $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2^T$, називаються *сингулярними базисами* матриці **A**. Зв'язок між векторами сингулярних базисів зв'ясовується в наступній теоремі.

Теорема 6.6.4. *Нехай матриця **A** має розміри $m \times n$ і ранг r. Припустимо, що $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ і $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ — ортонормовані власні вектори матриць $\boldsymbol{AA}^T$ і $\boldsymbol{A}^T\boldsymbol{A}$ відповідно, а $\sigma_1, \sigma_2, \ldots$ — сингулярні числа матриці **A**. Тоді*

$$\boldsymbol{A f}_i = \begin{cases} \sigma_i \boldsymbol{e}_i, & i \le r, \\ \boldsymbol{0}, & i > r, \end{cases} \qquad \boldsymbol{A}^T\boldsymbol{e}_i = \begin{cases} \sigma_i \boldsymbol{f}_i, & i \le r, \\ \boldsymbol{0}, & i > r. \end{cases} \tag{16}$$

Д о в е д е н н я. Запишемо сингулярний розклад (9) у вигляді:

$$\boldsymbol{A Q}_2^T = \boldsymbol{Q}_1 \boldsymbol{\Sigma}.$$

Ця рівність еквівалентна першій формулі (16), в чому неважко переконатися за допомогою теореми 6.6.3. Друга формула (16) випливає з рівності

$$\boldsymbol{A}^T\boldsymbol{Q}_1 = \boldsymbol{Q}_2^T\boldsymbol{\Sigma}^T,$$

одержуваній з (9), і тієї ж теореми 6.6.3. □

П р и к л а д. Обчислимо сингулярний розклад матриці

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

Для цього спочатку знайдемо матрицю $\boldsymbol{AA}^T$ та її власні значення і ортонормовані власні вектори:

$$\boldsymbol{AA}^T = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ -1 & -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 3 & 0 \\ 3 & 3 & -3 \\ 0 & -3 & 6 \end{pmatrix},$$

$$\lambda_1 = 9, \quad \lambda_2 = 6, \quad \lambda_3 = 0,$$

$$\boldsymbol{e}_1 = \frac{1}{\sqrt{3}}(-1 \quad -1 \quad 1)^T, \quad \boldsymbol{e}_2 = \frac{1}{\sqrt{2}}(1 \quad 0 \quad 1)^T, \quad \boldsymbol{e}_3 = \frac{1}{\sqrt{6}}(-1 \quad 2 \quad 1)^T.$$

Сингулярні числа матриці $\boldsymbol{A}$ дорівнюють

$$\sigma_1 = 3, \quad \sigma_2 = \sqrt{6}, \quad \sigma_3 = 0.$$

У відповідності до теореми 6.6.3 вектори $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ \boldsymbol{e}_3$ є стовпцями ортогональної матриці $\boldsymbol{Q}_1$. Стовпцями ортогональної матриці $\boldsymbol{Q}_2^T$ на підставі тієї ж теореми 6.6.3 є ортонормовані вектори $\boldsymbol{f}_i$, які в разі ненульових сингулярних чисел можна знайти за допомогою другої формули (16):

$$\boldsymbol{f}_1 = \frac{1}{\sigma_1}\boldsymbol{A}^T\boldsymbol{e}_1 = \frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \end{pmatrix},$$

$$\boldsymbol{f}_2 = \frac{1}{\sigma_2}\boldsymbol{A}^T\boldsymbol{e}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix},$$

Вектори $\boldsymbol{f}_i$, відповідні до нульового сингулярного числа, знайдемо, безпосередньо розглянувши матрицю $\boldsymbol{A}^T\boldsymbol{A}$. Ця матриця і її власні вектори, відповідні до власного значення $\lambda_3 = 0$, дорівнюють

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ -1 & -2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 & 0 & 3 \\ 3 & 5 & -2 & 1 \\ 0 & -2 & 2 & 2 \\ 3 & 1 & 2 & 5 \end{pmatrix},$$

$$\boldsymbol{x}_1 = (-1 \quad 1 \quad 1 \quad 0)^T, \quad \boldsymbol{x}_2 = (-2 \quad 1 \quad 0 \quad 1)^T.$$

Вектори $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$ не є ортогональними. Методом Грама — Шмідта побудуємо ортогональну систему векторів

$$\boldsymbol{b}_1 = \boldsymbol{x}_1, \quad \boldsymbol{b}_2 = -\boldsymbol{x}_1 + \boldsymbol{x}_2 = (-1 \quad 0 \quad -1 \quad 1)^T.$$

Тепер третій і четвертий стовпці матриці $\boldsymbol{Q}_2^T$ можна вважати рівними

$$\boldsymbol{f}_3 = \frac{\boldsymbol{b}_1}{|\boldsymbol{b}_1|} = \frac{1}{\sqrt{3}}(-1 \quad 1 \quad 1 \quad 0)^T, \quad \boldsymbol{f}_4 = \frac{\boldsymbol{b}_2}{|\boldsymbol{b}_2|} = \frac{1}{\sqrt{3}}(-1 \quad 0 \quad -1 \quad 1)^T.$$

Сингулярний розклад матриці $\boldsymbol{A}$ має вигляд:

$$\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{\Sigma}\boldsymbol{Q}_2 =$$

$$= \begin{pmatrix} -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ -1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & \sqrt{6} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{1}{\sqrt{3}} \begin{pmatrix} -1 & -1 & 0 & -1 \\ 0 & -1 & 1 & 1 \\ -1 & 1 & 1 & 0 \\ -1 & 0 & -1 & 1 \end{pmatrix}.$$

Зазначимо, що після знаходження векторів $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{f}_1$, $\boldsymbol{f}_2$ вектори $\boldsymbol{e}_3$, $\boldsymbol{f}_3$ і $\boldsymbol{f}_4$ можна визначити неєдиним способом. Дійсно, як неважко переконатися, замість $\boldsymbol{e}_3$ можна взяти $-\boldsymbol{e}_3$. Можна також замінити знак векторів $\boldsymbol{f}_3$ і (або) $\boldsymbol{f}_4$, а також поміняти їх місцями в матриці $\boldsymbol{Q}_2^T$. $\triangle$

Вправи до глави 6

6.1. Якими є умови того, що нерівності (6.1.9) виконуються як рівності?

6.2. Написати явні вирази для характеристичних многочленів матриць порядків 1, 2 і 3. Як виражаються коефіцієнти характеристичних многочленів через суми всіх головних мінорів матриці?

6.3. Показати, що характеристичний многочлен квазідіагональної (квазітрикутної) матриці дорівнює добутку характеристичних многочленів діагональних блоків.

6.4. Переконатися в справедливості тотожності

$$\begin{pmatrix} \lambda\boldsymbol{E}_m - \boldsymbol{AB} & \boldsymbol{A} \\ \boldsymbol{0} & \lambda\boldsymbol{E}_n \end{pmatrix}\begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{E}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{E}_n \end{pmatrix}\begin{pmatrix} \lambda\boldsymbol{E}_m & \boldsymbol{A} \\ \boldsymbol{0} & \lambda\boldsymbol{E}_n - \boldsymbol{BA} \end{pmatrix},$$

де $\boldsymbol{A}$ і $\boldsymbol{B}$ — матриці розмірів $m \times n$ і $n \times m$ відповідно. За допомогою тотожності довести, що: 1) характеристичні многочлени матриць $\boldsymbol{AB}$ і $\boldsymbol{BA}$ відрізняються лише множником λ^{m-n}; 2) $\sigma(\boldsymbol{AB}) = \sigma(\boldsymbol{BA})$ для квадратних матриць $\boldsymbol{A}$ і $\boldsymbol{B}$.

6.5. Знайти власні значення і власні вектори матриці $\boldsymbol{A} = \boldsymbol{ab}^T$, де $\boldsymbol{a}$ і $\boldsymbol{b}$ є n-вимірними векторами. Показати для матриці рангу 1, що будь-який її ненульовий стовпець є її власним вектором.

6.6. Як пов'язані власні вектори подібних матриць?

6.7. Довести, що коли принаймні одна з матриць $\boldsymbol{A}$, $\boldsymbol{B}$ є невиродженою, то матриці $\boldsymbol{AB}$ і $\boldsymbol{BA}$ подібні. Навести приклад двох вироджених матриць $\boldsymbol{A}$ і $\boldsymbol{B}$, для яких матриці $\boldsymbol{AB}$ і $\boldsymbol{BA}$ не є подібними.

6.8. Яка матриця діагоналізує матрицю $\boldsymbol{A}^T$, якщо матрицю $\boldsymbol{A}$ діагоналізує матриця $\boldsymbol{S}$?

6.9. Чи може мати просту структуру матриця порядку n при наявності характеристичного кореня кратності n?

6.10. Чому ранг матриці простої структури дорівнює кількості її ненульових власних значень?

6.11. Довести, що матриці простої структури подібні тоді і тільки тоді, коли вони мають однакові характеристичні многочлени.

6.12. Довести, що будь-який многочлен від матриці простої структури сам має просту структуру.

6.13. Має місце *теорема Келі — Гамільтона*: якщо $f(\lambda)$ — характеристичний многочлен матриці $\boldsymbol{A}$, то $f(\boldsymbol{A}) = \boldsymbol{0}$. Довести теорему для матриці простої структури, скориставшись результатом з вправи 6.12.

6.14. Як скористатися теоремою Гершгоріна для локалізації коренів многочлена?

6.15. Матриця ***B*** називається *домінантною*, якщо існують додатні числа d_i такі, що

$$d_i \mid b_{ii} \mid > \sum_{\substack{j=1 \\ j \neq i}}^{n} d_j \mid b_{ij} \mid, \quad i = \overline{1,\, n}.$$

Довести, що домінантна матриця є невиродженою. Написати аналогічні нерівності для стовпців матриці ***B***.

6.16. Довести, що матриця ***B*** є домінантною, якщо існують додатні числа d_i такі, що

$$d_i \mid b_{ii} \mid \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} d_j \mid b_{ij} \mid, \quad i = \overline{1,\, n},$$

де принаймні для одного *i* виконується строга нерівність. Написати аналогічні нерівності для стовпців матриці ***B***.

6.17. Як, не користуючись результатами з розділу 6.5, показати, що: 1) власні значення симетричної матриці є дійсними і відповідні до них власні вектори можуть бути дійсними; 2) власні вектори симетричної матриці, відповідні до різних власних значень, є ортогональними?

6.18. Як, не користуючись результатами з розділу 6.5, довести теорему 6.5.7 у випадку простих характеристичних коренів?

6.19. Довести для ортогональної матриці, що коли λ — власне значення, то λ^{-1} теж є власним значенням.

6.20. Довести для кососиметричної матриці, що коли λ — власне значення, то $-\lambda$ теж є власним значенням.

6.21. Для матриці Хаусхолдера ***H*** довести, що: 1) власному значенню –1 відповідає власний вектор ***u***; 2) усі вектори, ортогональні до ***u***, є власними і їм відповідає власне значення 1. Чи є інші власні значення у матриці Хаусхолдера?

6.22. Як поширити процедуру ортогоналізації Грама — Шмідта на вектори унітарного простору?

6.23. Довести, що визначником унітарної матриці є комплексне число, за модулем рівне одиниці.

6.24. Довести, що довільну квадратну матрицю $\boldsymbol{A}$ можна представити у вигляді суми $\boldsymbol{A} = \boldsymbol{H}_1 + \boldsymbol{H}_2 i$, де $\boldsymbol{H}_1$ і $\boldsymbol{H}_2$ — ермітові матриці. Довести, що з нормальності матриці $\boldsymbol{A}$ випливає переставність матриць $\boldsymbol{H}_1$ та $\boldsymbol{H}_2$ і навпаки.

6.25. Довести, що коли матриця $\boldsymbol{A}$ є нормальною, то такою ж є і матриця $\alpha\boldsymbol{A} + \beta\boldsymbol{E}$, де α, β — деякі числа.

6.26. Довести, що будь-який власний вектор $\boldsymbol{x}$ нормальної матриці $\boldsymbol{A}$, відповідний до власного значення λ, є одночасно власним вектором матриці $\boldsymbol{A}^*$, відповідним до власного значення $\overline{\lambda}$.

6.27. Довести, що сингулярні числа нормальної матриці дорівнюють модулям її власних значень.

6.28. Довести, що коли матриці $\boldsymbol{A}$ і $\boldsymbol{A}^*$ мають спільний власний вектор, то власні значення цих матриць, відповідні до спільного власного вектора, є комплексно спряженими.

ГЛАВА 7

КВАДРАТИЧНІ ФОРМИ І ДОДАТНО ВИЗНАЧЕНІ МАТРИЦІ

У цій главі розглядається питання про знаки власних значень симетричних матриць і пов'язані з ними властивості так званих квадратичних форм, а також вивчаються деякі методи локалізації власних значень таких матриць.

7.1. Додатно визначені матриці

7.1.1. Основні визначення. *Квадратичною формою* від змінних $x_1, \ldots, x_n$ називається многочлен відносно цих змінних, який вміщує тільки їх другі степені. Найбільш поширеною формою запису квадратичної форми f є подання її у вигляді $f = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$, де $\boldsymbol{x}$ — вектор евклідового простору з координатами $x_1, \ldots, x_n$ і $\boldsymbol{A}$ — симетрична матриця, яка називається матрицею квадратичної форми. Ми будемо розглядати насамперед дійсні квадратичні форми. Саме цей випадок є найбільш важливим у застосуваннях.

Обмеження, внаслідок якого матриця $\boldsymbol{A}$ повинна бути симетричною, не є суттєвим, бо якщо матриця $\boldsymbol{B}$ несиметрична і квадратична форма серед інших вміщує доданки $b_{ij}x_i x_j + b_{ji}x_j x_i$, то, поклавши $a_{ij} = a_{ji} = (b_{ij} + b_{ji})/2$, ми будемо мати в квадратичній формі замість зазначених доданків один доданок $2a_{ij}x_i x_j$, а матриця квадратичної форми стане симетричною.

Квадратична форма $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ називається: 1) *додатно* (*невід'ємно, від'ємно, недодатно*) *визначеною*, якщо для будь-якого ненульового вектора $\boldsymbol{x}$ виконується нерівність $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} > 0$ ($\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \geq 0$, $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} < 0$, $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \leq 0$);

2) *знакозмінною*, якщо існують вектори $\boldsymbol{x}$ і $\boldsymbol{y}$ такі, що $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$ і $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} < 0$. *Додатно* (*невід'ємно, від'ємно, недодатно*) *визначеною* називається симетрична матриця, яка відповідає додатно (невід'ємно, від'ємно, недодатно) визначеній квадратичній формі. Очевидно, що від'ємно (недодатно) визначені квадратичні форми отримуються з додатно (невід'ємно) визначених зміною знаку. Тому досить обмежитися розглядом додатно і невід'ємно визначених форм.

7.1.2. Основні ознаки додатної (невід'ємної) визначеності. Дві наступні теореми вміщують найбільш уживані ознаки додатної (невід'ємної) визначеності матриці.

Теорема 7.1.1. *Для того, щоб симетрична матриця **A** порядку n була додатно визначеною, необхідно і достатньо, щоб виконувалась одна з умов*:

1) $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$ *для всіх ненульових векторів* $\boldsymbol{x}$;
2) *усі власні значення матриці **A** додатні*;
3) *усі провідні мінори матриці **A** додатні* (*критерій Сільвестра*);
4) *усі провідні елементи* (*без переставлень рядків*) *додатні*;
5) *існує невироджена матриця **W** така, що* $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}$.

Д о в е д е н н я будується на з'ясуванні еквівалентності різних умов, зазначених вище.

Умова 1) є визначенням додатно визначеної матриці. Покажемо її еквівалентність умові 2). Нехай λ_i — власне значення матриці $\boldsymbol{A}$ і $\boldsymbol{x}_i$ — відповідний до нього власний вектор одиничної довжини, тобто $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i\boldsymbol{x}_i$ і $|\boldsymbol{x}_i| = 1$. Звідси знаходимо, що

$$\boldsymbol{x}_i^T\boldsymbol{A}\boldsymbol{x}_i = \boldsymbol{x}_i^T\lambda_i\boldsymbol{x}_i = \lambda_i .$$

Оскільки $\boldsymbol{x}_i^T\boldsymbol{A}\boldsymbol{x}_i > 0$, то $\lambda_i > 0$, тобто має місце умова 2).

Вважаючи, що $\lambda_i > 0$, $i = \overline{1, n}$, покажемо, що $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$ для будь-якого ненульового вектора $\boldsymbol{x}$. Оскільки на підставі теореми 6.5.7 симетрична матриця має повний набір ортонормованих власних векторів, то будь-який ненульовий вектор $\boldsymbol{x}$ можна подати у вигляді лінійної комбінації власних векторів, тобто $\boldsymbol{x} = \sum_{i=1}^{n}\alpha_i\boldsymbol{x}_i$, причому всі коефіцієнти $\alpha_1, \ldots, \alpha_n$ не можуть бути нульовими. Тепер маємо:

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \left(\sum_{i=1}^{n}\alpha_i\boldsymbol{x}_i\right)^T \boldsymbol{A}\sum_{j=1}^{n}\alpha_j\boldsymbol{x}_j = \left(\sum_{i=1}^{n}\alpha_i\boldsymbol{x}_i^T\right)\left(\sum_{j=1}^{n}\alpha_j\lambda_j\boldsymbol{x}_j\right) = \sum_{i=1}^{n}\alpha_i^2\lambda_i > 0.$$

Отже, з умови 2) випливає умова 1).

Покажемо, що з умови 1) випливає умова 3). З формули (6.2.1) і умови 2) випливає, що $\det \boldsymbol{A} > 0$. Щоб виявити цей результат для всіх провідних мінорів, розглянемо всі вектори, в яких останні $n-i$ координати дорівнюють нулю, де $1 \le i \le n-1$. Отже,

$$\boldsymbol{x} = \begin{pmatrix}\boldsymbol{x}_i\\ \boldsymbol{0}\end{pmatrix}, \quad \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \begin{pmatrix}\boldsymbol{x}_i\\ \boldsymbol{0}\end{pmatrix}^T \begin{pmatrix}\boldsymbol{A}_i & *\\ * & *\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_i\\ \boldsymbol{0}\end{pmatrix} = \boldsymbol{x}_i\boldsymbol{A}_i\boldsymbol{x}_i,$$

де $\boldsymbol{A}_i$ — підматриця, що відповідає провідному мінору i-го порядку, а позначка «$*$» вказує на довільний вигляд відповідних підматриць матриці $\boldsymbol{A}$. Таким чином, якщо $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$ для всіх ненульових векторів $\boldsymbol{x}$, то $\boldsymbol{x}_i^T\boldsymbol{A}_i\boldsymbol{x}_i > 0$ для всіх ненульових векторів $\boldsymbol{x}_i$, де $1 \le i \le n-1$. Це означає, що умова 1) справедлива для $\boldsymbol{A}_i$, тобто стосовно підматриць $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_{n-1}$ можна висловити ті ж міркування, що і по відношенню до матриці $\boldsymbol{A}$. Зокрема, їх власні значення і визначники, тобто провідні мінори матриці $\boldsymbol{A}$, повинні бути додатними.

Умова 4) випливає з формули (2.2.9) і умови 3).

З умови 5) випливає, що

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{W}^T\boldsymbol{W}\boldsymbol{x} = (\boldsymbol{W}\boldsymbol{x})^T\boldsymbol{W}\boldsymbol{x} = |\,\boldsymbol{W}\boldsymbol{x}\,|^2.$$

Оскільки матриця $\boldsymbol{W}$ невироджена, то рівність $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{0}$ можлива лише при $\boldsymbol{x} = \boldsymbol{0}$, тобто з умови 5) випливає умова 1).

Нарешті доведемо, що з умови 4) можна отримати умову 5). Для цього скористаємося $\boldsymbol{LDU}$-розкладом матриці $\boldsymbol{A}$, який визначається формулою (2.2.7):

$$\boldsymbol{A} = \boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U}, \qquad \boldsymbol{D} = \mathrm{diag}(d_1, \ldots, d_n).$$

Тут $\boldsymbol{U}$ — верхня трикутна матриця з одиничними діагональними елементами, а провідні елементи $d_1, \ldots, d_n$ є додатними. Завдяки останній обставині можна вважати, що

$$\boldsymbol{D} = \boldsymbol{D}^{1/2}\boldsymbol{D}^{1/2}, \qquad \boldsymbol{D}^{1/2} = \mathrm{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n}).$$

Тепер

$$\boldsymbol{A} = \boldsymbol{U}^T \boldsymbol{D}^{1/2} \boldsymbol{D}^{1/2} \boldsymbol{U} = (\boldsymbol{D}^{1/2} \boldsymbol{U})^T \boldsymbol{D}^{1/2} \boldsymbol{U}$$

і, отже, за матрицю $\boldsymbol{W}$ можна взяти матрицю $\boldsymbol{D}^{1/2}\boldsymbol{U}$. Це означає, що умова 5) є наслідком умови 4). □

З а у в а ж е н н я 1. Міркування, використані при обґрунтуванні критерія Сільвестра, легко поширити з провідних на *головні* мінори, розміщені в стовпцях і рядках з однаковими номерами. Тоді можна почати з будь-якого діагонального елемента a_{ii}, взявши його за головний мінор $\det \boldsymbol{A}_1$, і потім добавляти кожного разу необхідні елементи з нового стовпця і нового рядка. Це означає, що діагональні елементи повинні бути додатними. Проте, як свідчать численні приклади, ця умова не є достатньою (див. також теорему 7.1.3).

З а у в а ж е н н я 2. При доведенні теореми 7.1.1 ми одержали *розклад Холецького* для додатно визначеної матриці $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{W}^T \boldsymbol{W}, \quad \boldsymbol{W} = \operatorname{diag}(\sqrt{d_1}, \ldots, \sqrt{d_n})\boldsymbol{U}. \tag{1}$$

Матрицю $\boldsymbol{W}$ з розкладу $\boldsymbol{A} = \boldsymbol{W}^T \boldsymbol{W}$ можна побудувати ще двома способами, якщо скористатися теоремою 6.5.7:

$$\boldsymbol{W} = \operatorname{diag}(\sqrt{\lambda_1}, \ \ldots, \ \sqrt{\lambda_n})\boldsymbol{Q}^T, \quad \boldsymbol{W} = \boldsymbol{Q}\operatorname{diag}(\sqrt{\lambda_1}, \ \ldots, \ \sqrt{\lambda_n})\boldsymbol{Q}^T. \tag{2}$$

Остання матриця задовольняє умові $\boldsymbol{W}^2 = \boldsymbol{A}$ і називається *квадратним коренем* із симетричної додатно визначеної матриці $\boldsymbol{A}$. Вона теж є симетричною додатно визначеною матрицею і позначається через $\boldsymbol{A}^{1/2}$. Можна довести, що матриця $\boldsymbol{A}^{1/2}$ є єдиною для даної матриці $\boldsymbol{A}$.

Теорема 7.1.2. *Для того, щоб симетрична матриця* $\boldsymbol{A}$ *порядку* n *була невід'ємно визначеною, необхідно і достатньо, щоб виконувалась одна з умов*:

1) $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \geq 0$ *для всіх ненульових векторів* $\boldsymbol{x}$;

2) *усі власні значення матриці* $\boldsymbol{A}$ *невід'ємні*;

3) *усі головні мінори* (*а не тільки провідні*) *невід'ємні*;

4) *усі провідні елементи* (*можливо, з переставленнями рядків і стовпців з однаковими номерами*) *невід'ємні*;

5) *існує вироджена матриця* $\boldsymbol{W}$ *така, що* $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}$. □

При застосуванні до конкретної матриці теорем 7.1.1 і 7.1.2 найбільш зручною є умова 4), а найбільш важкою — умова 2). Проте умова 2) є найбільш зручною для теоретичних цілей.

Доведення теореми 7.1.2 не наводилось, оскільки воно мало відрізняється від доведення теореми 7.1.1. Натомість зробимо декілька зауважень.

З а у в а ж е н н я 3. Умова 3) формулюється для головних мінорів, а не тільки для провідних, бо в протилежному випадку ми не змогли б відрізняти матриці, в яких усі провідні мінори є нульовими. Прикладом такої матриці є матриця

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 0 & \alpha \end{pmatrix}.$$

Дійсно, власні значення дорівнюють 0 і α, тобто при $\alpha > 0$ матриця є невід'ємно визначеною. Це ж випливає з умови 3), оскільки головні мінори першого порядку дорівнюють 0 та α і визначник $\det \boldsymbol{A}$ дорівнює нулю. Проте невід'ємність тільки провідних мінорів ($\det \boldsymbol{A}_1 = \det \boldsymbol{A} = 0$) не забезпечує невід'ємну визначеність матриці, бо вона не має місця при $\alpha < 0$.

З а у в а ж е н н я 4. Нехай не існує $\boldsymbol{LDU}$ -розклад матриці $\boldsymbol{A}$, але існує невироджена матриця $\boldsymbol{P}$ така, що можна одержати $\boldsymbol{LDU}$ -розклад матриці $\boldsymbol{PAP}^T$, тобто

$$\boldsymbol{PAP}^T = \boldsymbol{U}^T\boldsymbol{DU}. \qquad (3)$$

Матрицею $\boldsymbol{P}$ може бути одна з елементарних матриць $\boldsymbol{L}_{ij}$, $\boldsymbol{L}_i(\lambda \mid j)$ (відповідні приклади можна знайти в наступному розділі). Оскільки при множенні на ці матриці величини головних мінорів матриці $\boldsymbol{A}$ не змінюються, то з умови 3) випливає умова 4).

З а у в а ж е н н я 5. З формули (3) знаходимо аналог формул (1):

$$\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}, \qquad \boldsymbol{W} = \boldsymbol{D}^{1/2}\boldsymbol{U}(\boldsymbol{P}^T)^{-1}.$$

Що стосується формул (2), то вони залишаються незмінними. Це означає, зокрема, що квадратний корінь з симетричної матриці існує не тільки для додатно визначеної, але і для невід'ємно визначеної матриці. До речі, матриця $\boldsymbol{W}$ в рівності $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}$ не обов'язково повинна бути квадратною, оскільки нерівність $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} \geq 0$ виконується для будь-якої матриці $\boldsymbol{W}$.

З а у в а ж е н н я 6. Умови невід'ємної визначеності, пов'язані зі знаками власних значень і визначників, можна одержати з теореми 7.1.1 за допомогою таких міркувань. Додавши до $\boldsymbol{A}$ матрицю $\varepsilon\boldsymbol{E}$, де $\varepsilon > 0$, одержимо додатно визначену матрицю $\boldsymbol{A} + \varepsilon\boldsymbol{E}$. Нехай $\varepsilon \to 0$, тоді власні значення і визначники, які неперервно залежать від ε, повинні бути додатними, поки ε не є нулем, і повинні стати невід'ємними при $\varepsilon = 0$.

7.1.3. Спеціальні ознаки додатної визначеності. Розглянемо деякі особливі властивості додатно визначених матриць. У наступній теоремі мова буде йти про умову додатної визначеності матриці Адамара.

Теорема 7.1.3. *Симетрична матриця Адамара з додатними діагональними елементами є додатно визначеною.*

Д о в е д е н н я. Власні значення симетричної матриці є дійсними. На підставі теореми 6.3.4 вони є додатними, тобто матриця є додатно визначеною у відповідності до частини 2) теореми 7.1.1. □

Наступна теорема дозволяє узагальнити частину 5) теореми 7.1.1.

Теорема 7.1.4. *Нехай матриця $\boldsymbol{B}$ порядку n є додатно визначеною, а матриця $\boldsymbol{W}$ розмірів $n \times r$,* де $r \leq n$*, має ранг r. Тоді матриця $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{B}\boldsymbol{W}$ є додатно визначеною.*

Д о в е д е н н я. Дійсно,

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T(\boldsymbol{W}^T\boldsymbol{B}\boldsymbol{W})\boldsymbol{x} = (\boldsymbol{W}\boldsymbol{x})^T\boldsymbol{B}(\boldsymbol{W}\boldsymbol{x}) > 0,$$

бо при $\boldsymbol{x} \neq \boldsymbol{0}$ вектор $\boldsymbol{W}\boldsymbol{x}$ є ненульовим на підставі теореми 3.5.1. Отже, симетрична матриця $\boldsymbol{A}$ є додатно визначеною за ознакою 1) з теореми

7.1.1, причому це має місце тільки тоді, коли $r(\boldsymbol{W}) = r$, бо в протилежному випадку існував би вектор $\boldsymbol{x} \neq \boldsymbol{0}$ такий, що $\boldsymbol{Wx} = \boldsymbol{0}$.

Висновок 1. *Якщо матриця **W** розмірів* $n \times r$, де $r \leq n$*, має ранг r, то матриця* $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}$ *є додатно визначеною.*

Доведення отримується при $\boldsymbol{B} = \boldsymbol{E}$.

Висновок 2. *Якщо матриця **B** додатно визначена, то матриця* $\boldsymbol{B}^{-1}$ *теж є додатно визначеною.*

Доведення отримується при $\boldsymbol{W} = \boldsymbol{B}^{-1}$. □

Остання теорема дозволяє дослідити питання про додатну визначеність блокових матриць.

Теорема 7.1.5. *Нехай **A** — квадратна матриця*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^T & \boldsymbol{A}_{22} \end{pmatrix},$$

де $\boldsymbol{A}_{11}$ *і* $\boldsymbol{A}_{22}$ *— симетричні підматриці. Для додатної визначеності матриці **A** необхідно і достатньо, щоб додатно визначеними були матриці* $\boldsymbol{A}_{11}$ *і* $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$.

Доведення. Необхідність. Для симетричної матриці $\boldsymbol{A}$ формулу (1.4.3) можна написати у вигляді рівності

$$\boldsymbol{C}_1^T\boldsymbol{A}\boldsymbol{C}_1 = \boldsymbol{D}_1. \qquad (5)$$

Матриця $\boldsymbol{D}_1$ є додатно визначеною на підставі теореми 7.1.4, а її діагональні блоки $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$ є додатно визначеними на підставі теореми 7.1.1.

Достатність. При додатно визначених діагональних блоках матриці $\boldsymbol{D}_1$ вона є додатно визначеною. З рівності (5) випливає, що $\boldsymbol{A} = (\boldsymbol{C}_1^{-1})^T\boldsymbol{D}_1\boldsymbol{C}_1^{-1}$, тому на підставі теореми 7.1.4 матриця $\boldsymbol{A}$ є додатно визначеною. □

Аналогічно за допомогою першої формули (1.4.6) можна розглянути випадок, коли додатно визначеними є матриці $\boldsymbol{A}_{22}$ і $\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{12}^T$. Про невід'ємну визначеність блокових матриць мова буде йти в теоремі 11.1.6.

7.2. Квадратичні форми

7.2.1. Перетворення квадратичних форм. З'ясуємо, як змінюється матриця квадратичної форми $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ при лінійному перетворенні $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{y}$, де $\boldsymbol{C}$ — задана матриця і $\boldsymbol{y}$ — вектор з координатами $y_1, \ldots, y_n$. Після підставлення виразу для $\boldsymbol{x}$ у квадратичну форму вона стає рівною $\boldsymbol{y}^T\boldsymbol{B}\boldsymbol{y}$, де

$$\boldsymbol{B} = \boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}. \tag{1}$$

Ця формула виражає матрицю перетвореної квадратичної форми через матрицю початкової форми і матрицю перетворення $\boldsymbol{C}$.

Далі ми будемо користуватися виключно невиродженими перетвореннями, коли $\det\boldsymbol{C} \neq 0$. При цьому на підставі висновку з теореми 3.4.4 буде $r(\boldsymbol{A}) = r(\boldsymbol{B})$, тобто при таких перетвореннях ранг матриці квадратичної форми залишається незмінним. Він називається *рангом квадратичної форми*.

Дві симетричні матриці $\boldsymbol{A}$ і $\boldsymbol{B}$, пов'язані співвідношенням (1), в якому $\det\boldsymbol{C} \neq 0$, називаються *конгруентними*, а перехід від $\boldsymbol{A}$ до $\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}$ називається *перетворенням конгруентності*. Таким чином, з кожною квадратичною формою пов'язаний цілий клас попарно конгруентних симетричних матриць.

Покажемо, як скористатися лінійними перетвореннями для зведення квадратичної форми до алгебраїчної суми квадратів, яка називається *канонічним виглядом* квадратичної форми.

Теорема 7.2.1. *Будь-яку квадратичну форму ортогональним перетворенням змінних можна звести до канонічного вигляду.*

Д о в е д е н н я. Нехай $\boldsymbol{Q}$ — ортогональна матриця з теореми 6.5.7, тобто $\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^T$. Тоді

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{Q}\Lambda\boldsymbol{Q}^T\boldsymbol{x} = \boldsymbol{y}^T\Lambda\boldsymbol{y} = \sum_{i=1}^{n}\lambda_i y_i^2, \tag{2}$$

де $\boldsymbol{y} = (y_1 \quad \ldots \quad y_n)^T = \boldsymbol{Q}^T\boldsymbol{x}$. Отже, квадратична форма за допомогою ортогонального перетворення $\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}$ зведена до канонічного вигляду (2). □

Квадратичну форму можна звести до канонічного вигляду неєдиним способом. Можна, наприклад, скористатися $\boldsymbol{LDU}$ -розкладом матриці $\boldsymbol{A}$. Якщо $\boldsymbol{A} = \boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U}$, то

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U}\boldsymbol{x} = \boldsymbol{z}^T\boldsymbol{D}\boldsymbol{z} = \sum_{i=1}^{n} d_i z_i^2, \tag{3}$$

де $\boldsymbol{z} = (z_1 \quad \dots \quad z_n)^T = \boldsymbol{U}\boldsymbol{x}$. При цьому $\boldsymbol{x} = \boldsymbol{U}^{-1}\boldsymbol{z}$, тобто матрицею лінійного перетворення є $\boldsymbol{C} = \boldsymbol{U}^{-1}$.

Розглянемо особливі випадки застосування $\boldsymbol{LDU}$ -розкладу, пов'язані з його неєдиністю або неможливістю одержати його без переставлення рядків і стовпців.

П р и к л а д 1. Перетворимо до канонічного вигляду квадратичну форму

$$f = 2x_1^2 + 2x_2^2 + 2x_3^2 - 4x_1x_2 + 4x_1x_3 - 4x_2x_3 .$$

Матриця цієї квадратичної форми дорівнює

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

Покажемо, що її $\boldsymbol{LDU}$ -розклад не є єдиним.

Перетворимо матрицю $\boldsymbol{A}$ у верхню трикутну матрицю за допомогою елементарних перетворень, пов'язаних з додаванням до рядків матриці інших її рядків, помножених на деякі числа. У результаті знайдемо матрицю $\boldsymbol{R}$ з $\boldsymbol{LR}$-розкладу:

$$\boldsymbol{A} \sim \begin{pmatrix} 2 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{R}.$$

Представимо $\boldsymbol{R}$ у вигляді добутку діагональної матриці і верхньої трикутної матриці з одиничними діагональними елементами:

$$\boldsymbol{R} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & u \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{DU}.$$

Тут u — будь-яке число. $\boldsymbol{LDU}$ -розклад має вигляд:

$$\boldsymbol{A} = \boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & u & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & u \\ 0 & 0 & 1 \end{pmatrix}.$$

Тепер знаходимо, що

$$f = \boldsymbol{x}^T\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U}\,\boldsymbol{x} = \boldsymbol{y}^T\boldsymbol{A}\,\boldsymbol{y} = 2\boldsymbol{y}_1^2.$$

Тут $\boldsymbol{y} = (y_1 \quad y_2 \quad y_3)^T = \boldsymbol{U}\,\boldsymbol{x}$, тому $\boldsymbol{x} = \boldsymbol{U}^{-1}\boldsymbol{y}$ і матрицею лінійного перетворення змінних є

$$\boldsymbol{C} = \boldsymbol{U}^{-1} = \begin{pmatrix} 1 & 1 & -1-u \\ 0 & 1 & -u \\ 0 & 0 & 1 \end{pmatrix}.$$

Матриця лінійного перетворення не є єдиною, проте це не позначилося на канонічному вигляді. △

П р и к л а д 2. Зведемо до канонічного вигляду квадратичну форму

$$f = x_2^2 + 2x_1x_2 + 2x_2x_3.$$

Матриця квадратичної форми дорівнює

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

Для неї неможливо одержати $\boldsymbol{LDU}$-розклад без переставлень рядків і стовпців. Тому перейдемо до матриці $\boldsymbol{B}$, переставивши в матриці $\boldsymbol{A}$ перший і другий рядки, а також перший і другий стовпці:

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

Матриця $\boldsymbol{B}$ пов’язана з матрицею $\boldsymbol{A}$ співвідношенням

$$\boldsymbol{B} = \boldsymbol{L}_{12}\boldsymbol{A}\,\boldsymbol{M}_{12} = \boldsymbol{L}_{12}\boldsymbol{A}\,\boldsymbol{L}_{12},$$

де $\boldsymbol{L}_{12}$ і $\boldsymbol{M}_{12}$ — елементарні матриці. Звідси знаходимо:

$$\boldsymbol{A} = \boldsymbol{L}_{12}^{-1}\boldsymbol{B}\,\boldsymbol{L}_{12}^{-1} = \boldsymbol{L}_{12}\boldsymbol{B}\,\boldsymbol{L}_{12}.$$

Далі маємо:

$$\boldsymbol{B} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{DU}.$$

LDU-розклад є таким:

$$\boldsymbol{B} = \boldsymbol{U}^T\boldsymbol{DU} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Тепер знаходимо, що

$$f = \boldsymbol{x}^T\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{L}_{12}\boldsymbol{B}\,\boldsymbol{L}_{12}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{L}_{12}\boldsymbol{U}^T\boldsymbol{DUL}_{12}\boldsymbol{x} = \boldsymbol{y}^T\boldsymbol{D}\,\boldsymbol{y} = y_1^2 - y_2^2,$$

де $\boldsymbol{y} = (y_1 \quad y_2 \quad y_3)^T = \boldsymbol{UL}_{12}\boldsymbol{x}$. Звідси випливає, що

$$\boldsymbol{x} = (\boldsymbol{UL}_{12})^{-1}\boldsymbol{y} = \boldsymbol{L}_{12}\boldsymbol{U}^{-1}\boldsymbol{y}.$$

Обчисливши $\boldsymbol{U}^{-1}$, знаходимо матрицю лінійного перетворення змінних $\boldsymbol{C}_1$:

$$\boldsymbol{C}_1 = \boldsymbol{L}_{12}\boldsymbol{U}^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Задачу зведення даної квадратичної форми до канонічного вигляду можна розв'язати інакше. Перетворимо матрицю $\boldsymbol{A}$, додавши другий рядок до першого рядка і другий стовбець до першого стовпця. У результаті одержимо матрицю

$$\boldsymbol{F} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

пов'язану з матрицею $\boldsymbol{A}$ співвідношенням

$$\boldsymbol{F} = \boldsymbol{L}_1(1\,|\,2)\boldsymbol{A}\,\boldsymbol{M}_1(1\,|\,2) = \boldsymbol{L}_1(1\,|\,2)\boldsymbol{A}\,\boldsymbol{L}_1^T(1\,|\,2),$$

де $\boldsymbol{L}_1(1\,|\,2)$ і $\boldsymbol{M}_1(1\,|\,2)$ — елементарні матриці. Звідси випливає, що

$$\boldsymbol{A} = \boldsymbol{L}_1^{-1}(1\,|\,2)\boldsymbol{F}(\boldsymbol{L}_1^T(1\,|\,2))^{-1} = \boldsymbol{L}_1(-1\,|\,2)\boldsymbol{F}\boldsymbol{L}_1^T(-1\,|\,2).$$

Далі знаходимо, що

$$\boldsymbol{F} \sim \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1/3 & 1/3 \\ 0 & 1/3 & -1/3 \end{pmatrix} \sim \begin{pmatrix} 3 & 2 & 1 \\ 0 & -1/3 & 1/3 \\ 0 & 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2/3 & 1/3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{D}\boldsymbol{U}.$$

$\boldsymbol{LDU}$-розклад виглядає так:

$$\boldsymbol{F} = \boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U} = \begin{pmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1/3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2/3 & 1/3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Тепер

$$f = \boldsymbol{x}^T\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{L}_1(-1\,|\,2)\boldsymbol{F}\,\boldsymbol{L}_1^T(-1\,|\,2)\,\boldsymbol{x} =$$

$$= \boldsymbol{x}^T\boldsymbol{L}_1(-1\,|\,2)\boldsymbol{U}^T\boldsymbol{D}\boldsymbol{U}\boldsymbol{L}_1^T(-1\,|\,2)\,\boldsymbol{x} = \boldsymbol{z}^T\boldsymbol{D}\boldsymbol{z} = 3z_1^2 - \frac{1}{3}z_2^2,$$

де $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{L}_1^T(-1\,|\,2)\,\boldsymbol{x}$. Звідси випливає, що

$$\boldsymbol{x} = (\boldsymbol{U}\boldsymbol{L}_1^T(-1\,|\,2))^{-1}\boldsymbol{z} = \boldsymbol{L}_1^T(1\,|\,2)\boldsymbol{U}^{-1}\boldsymbol{z}.$$

Після обчислення матриці $\boldsymbol{U}^{-1}$ знаходимо матрицю лінійного перетворення $\boldsymbol{C}_2$:

$$\boldsymbol{C}_2 = \boldsymbol{L}_1^T(1\,|\,2)\boldsymbol{U}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2/3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2/3 & -1 \\ 1 & 1/3 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad \triangle$$

7.2.2. Закон інерції квадратичних форм. Коефіцієнти при змінних у канонічному вигляді квадратичної форми називають *канонічними коефіцієнтами*. Взагалі кажучи, не всі вони є ненульовими. Очевидно, що кількість ненульових канонічних коефіцієнтів, яка дорівнює рангу матриці квадратичної форми, не залежить від вибору лінійного перетворення, за допомогою якого форма зводиться до канонічного вигляду. Більш того, кількість додатних (від'ємних) канонічних коефіцієнтів, як буде показано нижче, залишається незмінною при різних способах зведення квадратичної форми до канонічного вигляду. Ця властивість називається законом інерції квадратичних форм. Перш ніж перейти до обґрунтування закону інерції, перетворимо квадратичну форму до так званого нормального вигляду.

Нехай квадратична форма $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ зведена до канонічного вигляду:

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_r y_r^2 ,$$

причому відмінні від нуля канонічні коефіцієнти $\lambda_1, \ldots, \lambda_r$ занумеровані так, що $\lambda_1, \ldots, \lambda_p$ є додатними, а $\lambda_{p+1}, \ldots, \lambda_r$ — від'ємними. За допомогою лінійного перетворення змінних

$$y_1 = \frac{s_1}{\sqrt{\lambda_1}}, \ldots, y_p = \frac{s_p}{\sqrt{\lambda_p}}, y_{p+1} = \frac{s_{p+1}}{\sqrt{-\lambda_{p+1}}}, \ldots, y_r = \frac{s_r}{\sqrt{-\lambda_r}},$$

$$y_{r+1} = s_{r+1}, \ldots, y_n = s_n$$

квадратичну форму можна звести до вигляду:

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = s_1^2 + \ldots + s_p^2 - s_{p+1}^2 - \ldots - s_r^2 ,$$

який називається *нормальним виглядом* квадратичної форми.

***Теорема 7.2.2* (*закон інерції квадратичних форм*).** *Кількість доданків з додатними* (*від'ємними*) *коефіцієнтами в нормальному вигляді квадратичної форми не залежить від способу зведення форми до цього вигляду*.

Д о в е д е н н я. Нехай квадратична форма $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ двома способами зведена до нормального вигляду:

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = s_1^2 + \ldots + s_p^2 - s_{p+1}^2 - \ldots - s_r^2 = t_1^2 + \ldots + t_q^2 - t_{q+1}^2 - \ldots - t_r^2 . \quad (4)$$

Припустимо, що $p < q$, і розглянемо рівності

$$s_1 = \ldots = s_p = t_{q+1} = \ldots = t_n = 0. \quad (5)$$

Змінні $s_1, \ldots, s_p, t_{q+1}, \ldots, t_n$ одержані невиродженими перетвореннями змінних $x_1, \ldots, x_n$, тому співвідношення (5) можна розглядати як систему $n - q + p$ лінійних однорідних рівнянь з n невідомими $x_1, \ldots, x_n$. Кількість рівнянь у цій системі менша за кількість невідомих, тому система має ненульовий розв'язок, тобто існує вектор $\boldsymbol{x} \neq \boldsymbol{0}$, для якого виконуються рівності (5).

Підрахуємо значення квадратичної форми $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ для цього вектора $\boldsymbol{x}$. За допомогою формул (4) і (5) знаходимо:

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = -s_{p+1}^2 - \ldots - s_r^2 = t_1^2 + \ldots + t_q^2.$$

Остання рівність можлива лише при $s_{p+1} = \ldots = s_r = t_1 = \ldots = t_q = 0$.

Таким чином, система n лінійних однорідних рівнянь $t_i = 0$, $i = \overline{1, n}$, з n невідомими $x_1, \ldots, x_n$ має ненульовий розв'язок, тобто визначник цієї системи повинен дорівнювати нулю. Але це означає, що лінійне перетворення, за допомогою якого здійснюється перехід від змінних $x_1, \ldots, x_n$ до змінних $t_1, \ldots, t_n$, не є невиродженим. Отже, припущення $p < q$ веде до суперечності. З аналогічних міркувань неможлива нерівність $p > q$. Тому $p = q$. $\square$

Таким чином, кількість додатних і кількість від'ємних членів канонічного (нормального) вигляду квадратичної форми залишаються незмінними при будь-якому перетворенні квадратичної форми за допомогою лінійного перетворення. Вони називаються додатним і від'ємним *індексом інерції* відповідно. Різниця між додатним і від'єдним ідексами інерції називається *сигнатурою* квадратичної форми. Нехай r і σ — ранг і сигнатура квадратичної форми, μ і ν — додатний і від'ємний індекси інерції. Тоді $r = \mu + \nu$ і $\sigma = \mu - \nu$.

Для знаходження сигнатури квадратичної форми скористаємося формолою (3), в яку підставимо вирази для провідних елементів з формули (2.2.9). Позначаючи $\det \boldsymbol{A}_i$ через D_i, одержимо вираз для квадратичної форми рангу *n*:

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n} \frac{D_i}{D_{i-1}} z_i^2, \quad D_0 = 1.$$

Ця формула дозволяє знайти індекси інерції квадратичної форми. Якщо D_{i-1} і D_i мають однакові знаки, то коефіцієнт при z_i^2 є додатним, якщо ж їх знаки різні, то цей коефіцієнт є від'ємним. Виходить, що μ дорівнює кількості знакопостійностей, а ν дорівнює кількості знакозмін у ряді чисел $1, D_1, D_2, ..., D_n$. У цьому полягає *правило Якобі* для підрахунку індексів інерції. Аналогічно формулюється правило Якобі для кваратичної форми будь-якого рангу $r < n$ за умови, що ряд $1, D_1, D_2, ..., D_r$ складається з ненульових чисел.

7.2.3. Еліпсоїди в *n*-вимірному просторі. Теорема 7.1.1 дозволяє одержати геометричне тлумачення поняття додатної визначеності квадратичної форми. Розглянемо рівняння $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = 1$ і покажемо, що в *n*-вимірному афінному просторі воно визначає поверхню, яка є $(n-1)$-вимірним *еліпсоїдом* з центром у початку координат.

Почнемо з випадку двовимірного простору, розглянувши як приклад рівняння $7x_1^2 + 7x_2^2 - 6x_1x_2 = 1$. Перетворимо до канонічного вигляду квадратичну форму з лівої частини рівності. Матриця квадратичної форми дорівнює

$$\boldsymbol{A} = \begin{pmatrix} 7 & -3 \\ -3 & 7 \end{pmatrix},$$

а її діагоналізація за допомогою перетворення подібності з ортогональною матрицею виглядає так:

$$\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^T, \quad \boldsymbol{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 4 & 0 \\ 0 & 10 \end{pmatrix}.$$

Ортогональне перетворення $\boldsymbol{x} = \boldsymbol{Q}\,\boldsymbol{y}$ зберігає відстані і кути між векторами і зводить рівняння до вигляду:

$$4y_1^2 + 10y_2^2 = 1, \tag{6}$$

де $y_1 = \dfrac{x_1}{\sqrt{2}} + \dfrac{x_2}{\sqrt{2}}$, $y_2 = -\dfrac{x_1}{\sqrt{2}} + \dfrac{x_2}{\sqrt{2}}$.

Рівняння (6) визначає *еліпс*, тобто еліпсоїд у двовимірному просторі (рис. 13). Кінцем його довгої осі є точка $y_1 = 1/2$, $y_2 = 0$, а кінцем короткої — точка $y_1 = 0$, $y_2 = 1/\sqrt{10}$. Отже, половина довжини більшої осі дорівнює $1/\sqrt{\lambda_1}$, а половина довжини меншої осі — $1/\sqrt{\lambda_2}$. На площині Ox_1x_2 ми бачимо, що напрями осей еліпса збігаються з напрямами відповідних власних векторів $\boldsymbol{x}_1 = (1/\sqrt{2} \quad 1/\sqrt{2})^T$, $\boldsymbol{x}_2 = (-1/\sqrt{2} \quad 1/\sqrt{2})^T$ (рис. 14). Таким чином, геометрія додатної визначеності пов'язана з власними значеннями і власними векторами матриці $\boldsymbol{A}$.

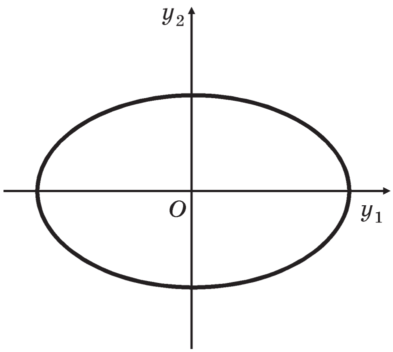

Рис. 13

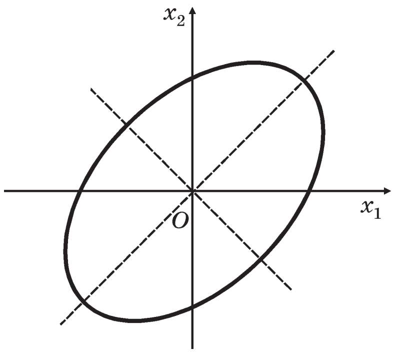

Рис. 14

У n-вимірному випадку теж здійснюємо діагоналізацію матриці квадратичної форми і за допомогою ортогонального перетворення $\boldsymbol{x} = \boldsymbol{Q}\,\boldsymbol{y}$ одержуємо:

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 = 1. \tag{7}$$

Розмірковуючи, як у двовимірному випадку, можна стверджувати, що рівняння $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = 1$ визначає еліпсоїд, кінець i-ї осі якого знаходиться в точці з координатами $y_1, \ldots, y_n$ де y_i визначається з рівності $\lambda_i y_i^2 = 1$, а інші координати дорівнюють нулю. Ці точки лежать на прямих лініях, які задаються власними векторами матриці $\boldsymbol{A}$, причому половина довжини відповідної осі дорівнює $1/\sqrt{\lambda_i}$.

При $\lambda_1 = \ldots = \lambda_n = R^{-2}$ формула (7) визначає *сферу* радіуса R в n-вимірному просторі.

7.3. Спільне перетворення двох квадратичних форм

При спільному перетворенні двох квадратичних форм до канонічного вигляду треба відрізняти два випадки в залежності від того, яким є лінійне перетворення змінних: ортогональним чи неортогональним. Для розгляду першого випадку нам знадобляться додаткові відомості стосовно матриць простої структури, а при дослідженні другого випадку виникає потреба в узагальненні визначення скалярного добутку векторів.

7.3.1. Властивості матриць простої структури. У двох наступних теоремах розглядаються властивості матриць простої структури, які стосуються проблеми спільного перетворення двох квадратичних форм, а також мають самостійне значення.

Теорема 7.3.1. *Для того, щоб квазідіагональна матриця*

$$\boldsymbol{A} = \operatorname{diag}(\boldsymbol{A}_{11},\ \boldsymbol{A}_{22},\ \ldots,\ \boldsymbol{A}_{rr})$$

була матрицею простої структури, необхідно і достатньо, щоб матрицею простої структури був кожний блок $\boldsymbol{A}_{jj}$. *При цьому власні вектори матриці* $\boldsymbol{A}$ *одержуються з власних векторів блоків* $\boldsymbol{A}_{jj}$ *добавленням нульових координат.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Будмо розглядати випадок двох блоків у матриці $\boldsymbol{A}$, оскільки загальний випадок одержується повторенням міркувань.

Нехай блоки $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$ матриці

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}_{22} \end{pmatrix}$$

мають порядки p і q $(p + q = n)$ відповідно. Оскільки $\boldsymbol{A}$ є матрицею простої структури, то на підставі формули (6.2.5) будемо мати:

$$\begin{pmatrix} \boldsymbol{A}_{11} & 0 \\ 0 & \boldsymbol{A}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \end{pmatrix} = \lambda_i \begin{pmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \end{pmatrix}, \quad i = \overline{1,\ n}. \tag{1}$$

Тут $\boldsymbol{x}_{i1}$ і $\boldsymbol{x}_{i2}$ — блоки власного вектора $\boldsymbol{x}_i$ і λ_i, $i = \overline{1,\ n}$ — власні значення $\boldsymbol{A}$ (не обов'язково різні). З (1) випливає, що

$$\boldsymbol{A}_{11}\boldsymbol{x}_{i1} = \lambda_i \boldsymbol{x}_{i1}, \quad \boldsymbol{A}_{22}\boldsymbol{x}_{i2} = \lambda_i \boldsymbol{x}_{i2}, \quad i = \overline{1,\ n}. \tag{2}$$

Якщо підматриці $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$ мають p і q власних векторів відповідно, то на підставі формул (2) власні вектори матриці $\boldsymbol{A}$ можна взяти у вигляді:

$$\begin{pmatrix} \boldsymbol{x}_{11} \\ \boldsymbol{0} \end{pmatrix}, \ \ldots, \ \begin{pmatrix} \boldsymbol{x}_{p1} \\ \boldsymbol{0} \end{pmatrix}, \ \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{x}_{12} \end{pmatrix}, \ \ldots, \ \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{x}_{q2} \end{pmatrix}. \tag{3}$$

Вони є лінійно незалежними, оскільки такими є вектори $\boldsymbol{x}_{11}, \ \ldots, \ \boldsymbol{x}_{p1}$ і $\boldsymbol{x}_{12}, \ \ldots, \ \boldsymbol{x}_{q2}$. Покажемо, що підматриці $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$ завжди мають указану кількість власних векторів.

Припустимо, що є k лінійно незалежних власних векторів підматриці $\boldsymbol{A}_{11}$ і l лінійно незалежних власних векторів підматриці $\boldsymbol{A}_{22}$, причому $k \le p$ і $l \le q$, але $k + l < n$. Тоді власних векторів (3) буде менше ніж n. Тому крім них існує принаймні один власний вектор $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix}$, відповідний до деякого власного значення λ і такий, що

$$\boldsymbol{A}_{11}\boldsymbol{x}_1 = \lambda \boldsymbol{x}_1, \quad \boldsymbol{A}_{22}\boldsymbol{x}_2 = \lambda \boldsymbol{x}_2, \tag{4}$$

причому вектори $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$ не є нульовими.

Рівності (4) показують, що $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$ є власними векторами підматриць $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$. Проте за припущенням ми вже знайшли всі лінійно незалежні власні вектори цих підматриць. Тому $\boldsymbol{x}_1$ є лінійною комбінацією векторів $\boldsymbol{x}_{11}, \ \dots, \ \boldsymbol{x}_{k1}$, а $\boldsymbol{x}_2$ — лінійною комбінацією векторів $\boldsymbol{x}_{12}, \ \dots, \ \boldsymbol{x}_{l2}$. Отже, $\boldsymbol{x}$ лінійно виражається через вектори

$$\begin{pmatrix} \boldsymbol{x}_{11} \\ \mathbf{0} \end{pmatrix}, \ \dots, \ \begin{pmatrix} \boldsymbol{x}_{k1} \\ \mathbf{0} \end{pmatrix}, \ \begin{pmatrix} \mathbf{0} \\ \boldsymbol{x}_{12} \end{pmatrix}, \ \dots, \ \begin{pmatrix} \mathbf{0} \\ \boldsymbol{x}_{l2} \end{pmatrix},$$

тобто $\boldsymbol{x}$ не є новим лінійно незалежним вектором. Одержана суперечність доводить, що $k = p, \ \ l = q$.

Таким чином, підматриці $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$ мають p і q лінійно незалежних власних векторів відповідно. Тому вони є матрицями простої структури. З (3) випливає, що власні вектори матриці $\boldsymbol{A}$ відповідають умові теореми щодо їх нульових координат.

Д о с т а т н і с т ь. Якщо підматриці $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22}$ мають просту структуру, то у відповідності до формули (6.2.6)

$$\boldsymbol{A}_{11} = \boldsymbol{S}_1 \boldsymbol{\Lambda}_1 \boldsymbol{S}_1^{-1}, \quad \boldsymbol{A}_{22} = \boldsymbol{S}_2 \boldsymbol{\Lambda}_2 \boldsymbol{S}_2^{-1}.$$

Тоді

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{S}_1 \boldsymbol{\Lambda}_1 \boldsymbol{S}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{S}_2 \boldsymbol{\Lambda}_2 \boldsymbol{S}_2^{-1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{S}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{S}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{S}_2^{-1} \end{pmatrix},$$

тобто матрицю $\boldsymbol{A}$ можна діагоналізувати. Діагоналізуюча матриця є квазідіагональною, тому умова щодо вигляду власних векторів матриці $\boldsymbol{A}$ виконується. □

Теорема 7.3.2. *Для того, щоб дві матриці простої структури можна було діагоналізувати за допомогою одного і того ж перетворення подібності, необхідно і достатньо, щоб ці матриці були переставними.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Нехай матриці $\boldsymbol{A}$ і $\boldsymbol{B}$ можуть бути діагоналізовані за допомогою одного і того ж перетворення подібності, тобто

$$\boldsymbol{A} = \boldsymbol{S} \boldsymbol{\Lambda} \boldsymbol{S}^{-1}, \quad \boldsymbol{B} = \boldsymbol{S} \boldsymbol{M} \boldsymbol{S}^{-1}.$$

Тут Λ, $\boldsymbol{M}$ — діагональні матриці, які є, очевидно, переставними. Переставність матриць $\boldsymbol{A}$ і $\boldsymbol{B}$ доводиться так:

$$\boldsymbol{AB} = \boldsymbol{S}\Lambda\boldsymbol{S}^{-1}\boldsymbol{S}\boldsymbol{M}\boldsymbol{S}^{-1} = \boldsymbol{S}\Lambda\boldsymbol{M}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{M}\Lambda\boldsymbol{S}^{-1} =$$

$$= \boldsymbol{S}\boldsymbol{M}\boldsymbol{S}^{-1}\boldsymbol{S}\Lambda\boldsymbol{S}^{-1} = \boldsymbol{B}\boldsymbol{A}.$$

Достатність. Застосуємо до матриць $\boldsymbol{A}$ і $\boldsymbol{B}$ перетворення подібності таке, що зводить одну з матриць, наприклад $\boldsymbol{A}$, до діагонального вигляду:

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \Lambda, \quad \boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S} = \boldsymbol{C}. \tag{5}$$

Тут матриця $\boldsymbol{C}$ не є, взагалі кажучи, діагональною. Зважаючи на те, що $\boldsymbol{AB} = \boldsymbol{B}\boldsymbol{A}$, далі знаходимо:

$$\Lambda\boldsymbol{C} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}\boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{B}\boldsymbol{S} = \boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S}\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{C}\Lambda.$$

Отже, матриці Λ і $\boldsymbol{C}$ є переставними. Скориставшись правилом множення довільної матриці на діагональну матрицю, одержимо:

$$\{\Lambda\boldsymbol{C}\}_{ij} = \lambda_{ii}c_{ij}, \quad \{\boldsymbol{C}\Lambda\}_{ij} = c_{ij}\lambda_{jj}, \quad i,\, j = \overline{1,\ n}.$$

Звідси знаходимо, що

$$c_{ij}(\lambda_{ii} - \lambda_{jj}) = 0, \quad i,\, j = \overline{1,\ n}. \tag{6}$$

Якщо всі діагональні елементи матриці Λ є різними, то, як показує формула (6), позадіагональні елементи матриці $\boldsymbol{C}$ повинні бути нульовими, тобто $\boldsymbol{C}$ — діагональна матриця. Отже, ми довели достатність умови переставності матриць $\boldsymbol{A}$ і $\boldsymbol{B}$, якщо принаймні одна з цих матриць має різні власні значення.

Нехай матриця $\boldsymbol{A}$ має власне значення λ кратності m. Не порушуючи загальності, можна вважати, що m перших діагональних елементів матриці Λ дорівнюють λ. З формули (6) випливає, що, взагалі кажучи, відмінними від нуля повинні бути елементи c_{ij} для $i,\, j = \overline{1,\ m}$. Це означає, що матриця $\boldsymbol{C}$ є квазідіагональною, причому порядок кожного блоку дорівнює кратності відповідного власного значення матриці $\boldsymbol{A}$.

Для завершення доведення треба скористатися теоремою 7.3.1. З цією метою зазначимо, що матриця $\boldsymbol{C}$ має просту структуру. Дійсно, нехай $\boldsymbol{S}_B$ — діагоналізуюча матриця для $\boldsymbol{B}$, тоді

$$\boldsymbol{B} = \boldsymbol{S}_B \boldsymbol{M} \boldsymbol{S}_B^{-1}.$$

Підставивши цей вираз у другу рівність (5), одержимо, що

$$\boldsymbol{S}^{-1} \boldsymbol{S}_B \boldsymbol{M} \boldsymbol{S}_B^{-1} \boldsymbol{S} = \boldsymbol{C},$$

або

$$\boldsymbol{S}^{-1} \boldsymbol{S}_B \boldsymbol{M} (\boldsymbol{S}^{-1} \boldsymbol{S}_B)^{-1} = \boldsymbol{C}.$$

Остання рівність доводить, що $\boldsymbol{C}$ є матрицею простої структури. На підставі теореми 7.3.1 просту структуру мають також діагональні блоки матриці $\boldsymbol{C}$.

Застосуємо до другої рівності (5) перетворення подібності, яке діагоналізує матрицю $\boldsymbol{C}$. Як і в теоремі 7.3.1, розглянемо випадок двох блоків у перетворюваної матриці. Інакше кажучи, помножимо справа обидві частини другої рівності (5) на матрицю

$$\boldsymbol{S}_0 = \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_{22} \end{pmatrix},$$

а зліва — на $\boldsymbol{S}_0^{-1}$. Тут $\boldsymbol{S}_{11}$ і $\boldsymbol{S}_{22}$ — діагоналізуючі матриці для блоків матриці $\boldsymbol{C}$. Після вказаного множення будемо мати:

$$(\boldsymbol{S}\,\boldsymbol{S}_0)^{-1} \boldsymbol{B}\,\boldsymbol{S}\,\boldsymbol{S}_0 = \boldsymbol{M}. \tag{7}$$

Аналогічне перетворення першої рівності (5) дає:

$$\boldsymbol{S}_0^{-1} \boldsymbol{S}^{-1} \boldsymbol{A}\,\boldsymbol{S}\,\boldsymbol{S}_0 = \boldsymbol{S}_0^{-1} \Lambda\,\boldsymbol{S}_0 =$$

$$= \begin{pmatrix} \boldsymbol{S}_{11}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_{22}^{-1} \end{pmatrix} \begin{pmatrix} \Lambda_1 & \boldsymbol{0} \\ \boldsymbol{0} & \Lambda_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{S}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{S}_{11}^{-1} \Lambda_1 \boldsymbol{S}_{11} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{S}_{22}^{-1} \Lambda_2 \boldsymbol{S}_{22} \end{pmatrix}. \tag{8}$$

Тут Λ_1 і Λ_2 — діагональні блоки матриці Λ. Оскільки $\Lambda_1 = \lambda \boldsymbol{E}$, то $\boldsymbol{S}_{11}^{-1} \Lambda_1 \boldsymbol{S}_{11} = \Lambda_1$. Виходячи з аналогічних міркувань, можна стверджувати, що $\boldsymbol{S}_{22}^{-1} \Lambda_2 \boldsymbol{S}_{22} = \Lambda_2$.

Отже, на підставі формули (8) маємо рівність

$$(\boldsymbol{S}\,\boldsymbol{S}_0)^{-1}\boldsymbol{A}\,\boldsymbol{S}\,\boldsymbol{S}_0 = \Lambda. \tag{9}$$

Формули (9) і (7) показують, що матриці $\boldsymbol{A}$ і $\boldsymbol{B}$ діагоналізовані за допомогою одного і того ж перетворення подібності. □

7.3.2. Додатна визначеність і скалярний добуток. Покажемо, як, користуючись поняттям додатної визначеності, можна узагальнити формулу (4.1.1) для скалярного добутку векторів.

Якщо $\boldsymbol{H}$ — додатно визначена матриця, то число

$$(\boldsymbol{a}, \boldsymbol{b})_{\boldsymbol{H}} = \boldsymbol{a}^T \boldsymbol{H}\boldsymbol{b} \tag{10}$$

можна розглядати як скалярний добуток векторів $\boldsymbol{a}, \boldsymbol{b} \in E_n$. У цьому неважко пересвідчитися, перевіряючи виконання аксіом скалярного добутку. Зокрема, аксіома 4) виконується завдяки додатній визначеності квадратичної форми $\boldsymbol{a}^T \boldsymbol{H}\boldsymbol{a}$. Зрозуміло, що скалярному добутку (4.1.1) відповідає матриця $\boldsymbol{H} = \boldsymbol{E}$ у формулі (10).

На випадок скалярного добутку (10) легко узагальнюються поняття ортогональності і ортонормованості векторів. Вектори $\boldsymbol{a}$, $\boldsymbol{b}$ ми називаємо ***H**-ортогональними*, якщо $\boldsymbol{a}^T \boldsymbol{H}\boldsymbol{b} = 0$. Систему векторів називають ***H**-ортогональною*, якщо вона складається з одного вектора або її вектори є попарно $\boldsymbol{H}$-ортогональними. $\boldsymbol{H}$-ортогональна система ненульвих векторів є лінійно незалежною. У цьому легко переконатися, розмірковуючи подібно до того, як це було зроблено при доведенні теореми 4.1.1.

За допомогою скалярного добутку (10) можна узагальнити формулу (4.1.2) для довжини вектора. Тепер вона дорівнюватиме $|\,\boldsymbol{a}\,|_{\boldsymbol{H}} = \sqrt{\boldsymbol{a}^T \boldsymbol{H}\boldsymbol{a}}$. Очевидно, що $|\,\boldsymbol{a}\,|_{\boldsymbol{H}} > 0$ при $\boldsymbol{a} \neq \boldsymbol{0}$ і $|\,\boldsymbol{0}\,|_{\boldsymbol{H}} = 0$. Систему векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ називають ***H**-ортонормованою*, якщо $\boldsymbol{a}_i^T \boldsymbol{H}\boldsymbol{a}_j = \delta_{ij}$ при $i, j = \overline{1, m}$.

Узагальнюючи процедуру ортогоналізації Грама — Шмідта на випадок скалярного добутку (1), можна від системи лінійно незалежних векторів $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ перейти до $\boldsymbol{H}$-ортогональної системи ненульових векторів $\boldsymbol{b}_1, ..., \boldsymbol{b}_m$, які є лінійними комбінаціями векторів $\boldsymbol{a}_1, ..., \boldsymbol{a}_m$. Загальна схема обгрунтування такої процедури аналогічна міркуванням з теореми 4.1.2.

7.3.3. Спільне перетворення двох квадратичних форм до канонічного вигляду. Звести до канонічного вигляду дві квадратичні форми за допомогою одного і того ж ортогонального перетворення змінних — це означає діагоналізувати (симетричні) матриці квадратичних форм за допомогою однієї і тієї ж ортогональної матриці. Теорема 7.3.2 вказує на те, що це можливо лише за умови переставності матриць квадратичних форм. Проте для багатьох цілей досить звести форми до канонічного вигляду за допомогою довільного (неортогонального) перетворення.

Ми розв'яжемо цю проблему, досліджуючи так звану узагальнену задачу про власні значення і власні вектори, яка полягає у відшуканні чисел λ_i (власних значень) і відповідних векторів $\boldsymbol{x}_i$ (власних векторів) з рівнянь

$$(\boldsymbol{A} - \lambda\boldsymbol{H})\boldsymbol{x} = \boldsymbol{0}, \quad \det(\boldsymbol{A} - \lambda\boldsymbol{H}) = 0. \tag{11}$$

Тут вважається, що матриці $\boldsymbol{A}$ і $\boldsymbol{H}$ є симетричними, причому $\boldsymbol{H}$ є додатно визначеною. Існує можливість звести задачу (11) до задачі типу (6.1.5) з однією симетричною матрицею. Покажемо, як це можна зробити.Частина 5) теореми 7.1.1 стверджує, що $\boldsymbol{H} = \boldsymbol{W}^T\boldsymbol{W}$, де матриця $\boldsymbol{W}$ є невиродженою. Після заміни $\boldsymbol{x} = \boldsymbol{W}^{-1}\boldsymbol{z}$ і множення зліва обох частин першого рівняння (11) на $(\boldsymbol{W}^T)^{-1}$ воно набуває вигляду:

$$(\boldsymbol{B} - \lambda\boldsymbol{E})\boldsymbol{z} = \boldsymbol{0}, \tag{12}$$

де $\boldsymbol{B} = (\boldsymbol{W}^T)^{-1}\boldsymbol{A}\boldsymbol{W}^{-1}$ — симетрична матриця. При цьому друге рівняння (11) перетворюється так:

$$0 = \det(\boldsymbol{A} - \lambda\boldsymbol{H}) = \det((\boldsymbol{W}^T)^{-1}(\boldsymbol{A} - \lambda\boldsymbol{W}^T\boldsymbol{W})\boldsymbol{W}^{-1}) =$$
$$= \det(\boldsymbol{B} - \lambda\boldsymbol{E}). \tag{13}$$

Формули (12) і (13) свідчать про те, що задача (11) зведена до задачі (6.1.5).

Повертаючись до проблеми спільного перетворення квадратичних форм, розглянемо теорему, в якій ця проблема знаходить своє вирішення.

Теорема 7.3.3. *Для квадратичних форм* $f = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ *і* $g = \boldsymbol{x}^T \boldsymbol{H}\boldsymbol{x}$*, з яких друга є додатно визначеною, існує лінійне перетворення* $\boldsymbol{x} = \boldsymbol{S}\boldsymbol{z}$ *таке, що* $f = \boldsymbol{z}^T \Lambda \boldsymbol{z}$ *і* $g = \boldsymbol{z}^T \boldsymbol{z}$, *де* Λ *є діагональною матрицею.*

Д о в е д е н н я. Власні вектори матриці $\boldsymbol{B}$ і власні вектори задачі (11) пов'язані співвідношенням $\boldsymbol{z}_i = \boldsymbol{W}\boldsymbol{x}_i$, $i = \overline{1, n}$. Власні вектори матриці $\boldsymbol{B}$ є ортонормованими, тому для власних векторів задачі (11) маємо:

$$\boldsymbol{x}_i^T \boldsymbol{H}\boldsymbol{x}_j = \boldsymbol{x}_i^T \boldsymbol{W}^T \boldsymbol{W}\boldsymbol{x}_j = \boldsymbol{z}_i^T \boldsymbol{z}_j = \delta_{ij}, \quad i,\ j = \overline{1, n}, \tag{14}$$

тобто вони є $\boldsymbol{H}$-ортонормованими. Далі за допомогою першої формули (11) і формули (14) знаходимо:

$$\boldsymbol{x}_i^T \boldsymbol{A}\boldsymbol{x}_j = \boldsymbol{x}_i^T \lambda_j \boldsymbol{H}\boldsymbol{x}_j = \lambda_j \delta_{ij}, \quad i,\ j = \overline{1, n}. \tag{15}$$

Співвідношення (15) і (14) можна написати у вигляді матричних рівностей

$$\boldsymbol{S}^T \boldsymbol{A}\boldsymbol{S} = \Lambda, \quad \boldsymbol{S}^T \boldsymbol{H}\boldsymbol{S} = \boldsymbol{E},$$

де $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ і $\boldsymbol{S} = (\boldsymbol{x}_1 \ ... \ \boldsymbol{x}_n)$. Тепер стає очевидним, що перетворення $\boldsymbol{x} = \boldsymbol{S}\boldsymbol{z}$ зводить квадратичні форми f і g до канонічного вигляду:

$$f = \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z}^T \boldsymbol{S}^T \boldsymbol{A}\boldsymbol{S}\boldsymbol{z} = \boldsymbol{z}^T \Lambda \boldsymbol{z}, \quad g = \boldsymbol{x}^T \boldsymbol{H}\boldsymbol{x} = \boldsymbol{z}^T \boldsymbol{S}^T \boldsymbol{H}\boldsymbol{S}\boldsymbol{z} = \boldsymbol{z}^T \boldsymbol{z}.$$

Зазначимо, що це перетворення є невиродженим, бо стовпці матриці $\boldsymbol{S}$, будучи $\boldsymbol{H}$-ортонормованими, є лінійно незалежними. □

П р и к л а д. Виконаємо спільне зведення до канонічного вигляду квадратичних форм

$$f = 2x_1^2 - 2x_2^2 - 3x_3^2 + 2x_1 x_3 - 10 x_2 x_3,$$

$$g = 2x_1^2 + 3x_2^2 + 2x_3^2 + 2x_1 x_3.$$

Квадратична форма g є додатно визначеною. Це випливає, наприклад, з рівності $g = x_1^2 + 3x_2^2 + x_3^2 + (x_1 + x_3)^2$. Тому спільне зведення до канонічного вигляду можливе.

Матриці квадратичних форм дорівнюють

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & -2 & -5 \\ 1 & -5 & -3 \end{pmatrix}, \quad \boldsymbol{H} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Вони не є переставними, тому для спільного перетворення квадратичних форм скористаємося теоремою 7.3.3. Характеристичне рівняння $\det(\boldsymbol{A} - \lambda\boldsymbol{H}) = 0$ є таким:

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

Воно має корені $\lambda_1 = -4,\ \lambda_2 = \lambda_3 = 1.$

Для визначення власного вектора $\boldsymbol{x}_1$, що відповідає власному значенню λ_1, маємо рівняння $(\boldsymbol{A} - \lambda_1\boldsymbol{H})\boldsymbol{x}_1 = \boldsymbol{0}$, або

$$\begin{pmatrix} 10 & 0 & 5 \\ 0 & 10 & -5 \\ 5 & -5 & 5 \end{pmatrix}\boldsymbol{x}_1 = \boldsymbol{0}.$$

Фундаментальна система розв'язків цієї однорідної системи складається з одного вектора $\boldsymbol{x}_1 = (-1/2 \quad 1/2 \quad 1)^T$. Аналогічно знаходимо власні вектори $\boldsymbol{x}_2 = (1 \quad 0 \quad 0)^T$ і $\boldsymbol{x}_3 = (0 \quad -1 \quad 1)^T$, відповідні до кратного власного значення.

Вектор $\boldsymbol{x}_1$ є $\boldsymbol{H}$-ортогональним до векторів $\boldsymbol{x}_2$ і $\boldsymbol{x}_3$:

$$\boldsymbol{x}_1^T\boldsymbol{H}\boldsymbol{x}_2 = (-1/2 \quad 1/2 \quad 1)\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0,$$

$$\boldsymbol{x}_1^T \boldsymbol{H}\boldsymbol{x}_3 = (-1/2 \quad 1/2 \quad 1)\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = 0.$$

Проте вектори $\boldsymbol{x}_2$ і $\boldsymbol{x}_3$ не є такими, оскільки

$$\boldsymbol{x}_2^T \boldsymbol{H}\boldsymbol{x}_3 = (1 \quad 0 \quad 0)\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = 1 \neq 0.$$

З подібною ситуацією ми мали справу, розглядаючи приклад у розділі 6.5. Там було вказано на необхідність застосувати ортогоналізацію Грама — Шмідта до власних векторів, відповідних до кратного власного значення. Аналогічно будемо діяти і в нашому випадку.

Застосуємо $\boldsymbol{H}$-ортогоналізацію Грама — Шмідта до векторів $\boldsymbol{x}_2$ і $\boldsymbol{x}_3$. Введемо вектори $\boldsymbol{b}_1 = \boldsymbol{x}_2$ і $\boldsymbol{b}_2 = \boldsymbol{x}_3 - \beta_{21}\boldsymbol{b}_1$, де число β_{21} знаходиться з умови $\boldsymbol{b}_2^T \boldsymbol{H}\boldsymbol{b}_1 = 0.$ Оскільки

$$0 = \boldsymbol{b}_2^T \boldsymbol{H}\boldsymbol{b}_1 = \boldsymbol{x}_3^T \boldsymbol{H}\boldsymbol{b}_1 - \beta_{21}\boldsymbol{b}_1^T \boldsymbol{H}\boldsymbol{b}_1,$$

то

$$\beta_{21} = \frac{\boldsymbol{x}_3^T \boldsymbol{H}\boldsymbol{b}_1}{\boldsymbol{b}_1^T \boldsymbol{H}\boldsymbol{b}_1} = \frac{\boldsymbol{x}_3^T \boldsymbol{H}\boldsymbol{x}_2}{\boldsymbol{x}_2^T \boldsymbol{H}\boldsymbol{x}_2}.$$

Далі знаходимо:

$$\boldsymbol{x}_2^T \boldsymbol{H}\boldsymbol{x}_2 = (1 \quad 0 \quad 0)\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 2, \quad \beta_{21} = \frac{1}{2},$$

$$\boldsymbol{b}_2 = \boldsymbol{x}_3 - \frac{1}{2}\boldsymbol{b}_1 = \boldsymbol{x}_3 - \frac{1}{2}\boldsymbol{x}_2 = \begin{pmatrix} -\frac{1}{2} & -1 & 1 \end{pmatrix}^T.$$

Вектори $\boldsymbol{x}_1$, $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ є $\boldsymbol{H}$-ортогональними. Зробимо їх $\boldsymbol{H}$-ортонормованими, розділивши на довжини $|\, \boldsymbol{x}_1 \,|_H$, $|\, \boldsymbol{b}_1 \,|_H$, $|\, \boldsymbol{b}_2 \,|_H$ відповідно. Вище ми бачили, що

$$|\, \boldsymbol{b}_1 \,|_H = \sqrt{\boldsymbol{b}_1^T \boldsymbol{H}\boldsymbol{b}_1} = \sqrt{\boldsymbol{x}_2^T \boldsymbol{H}\boldsymbol{x}_2} = \sqrt{2}.$$

Аналогічно знаходимо:

$$|\,\boldsymbol{x}_1\,|_H = \sqrt{\boldsymbol{x}_1^T \boldsymbol{H}\boldsymbol{x}_1} = \frac{3}{2}, \quad |\,\boldsymbol{b}_2\,|_H = \sqrt{\boldsymbol{b}_2^T \boldsymbol{H}\boldsymbol{b}_2} = \frac{3}{\sqrt{2}}.$$

Тепер можна написати матрицю $\boldsymbol{S}$:

$$\boldsymbol{S} = \begin{pmatrix} \dfrac{\boldsymbol{x}_1}{|\,\boldsymbol{x}_1\,|_{\boldsymbol{H}}} & \dfrac{\boldsymbol{b}_1}{|\,\boldsymbol{b}_1\,|_{\boldsymbol{H}}} & \dfrac{\boldsymbol{b}_2}{|\,\boldsymbol{b}_2\,|_{\boldsymbol{H}}} \end{pmatrix} = \begin{pmatrix} -1/3 & \sqrt{2}/2 & -\sqrt{2}/6 \\ 1/3 & 0 & -\sqrt{2}/3 \\ 2/3 & 0 & \sqrt{2}/3 \end{pmatrix}.$$

Лінійне перетворення $\boldsymbol{x} = \boldsymbol{S}\boldsymbol{z} = \boldsymbol{S}(z_1 \quad z_2 \quad z_3)^T$ зводить квадратичні форми f і g до канонічного вигляду:

$$f = \boldsymbol{z}^T \operatorname{diag}(-4,\ 1,\ 1)\boldsymbol{z} = -4z_1^2 + z_2^2 + z_3^2,$$

$$g = \boldsymbol{z}^T \boldsymbol{z} = z_1^2 + z_2^2 + z_3^2. \ \triangle$$

7.4. Екстремальні властивості квадратичних форм

Цей і наступний розділи присвячені вивченню властивостей власних значень симетричних матриць. Розглядувані при цьому методи легко узагальнюються на ермітові матриці. Більшість з одержаних у цих розділах результатів має відношення до проблеми локалізації власних значень.

7.4.1. Екстремальні властивості власних значень. Розглянемо питання про найбільше та найменше значення квадратичної форми $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ на сфері одиничного радіуса $\boldsymbol{x}^T \boldsymbol{x} = 1$ і про зв'язок цих значень з власними значеннями матриці $\boldsymbol{A}$.

Теорема 7.4.1. *Нехай λ_1, λ_n — найбільше і найменше власні значення симетричної матриці $\boldsymbol{A}$ порядку n, а $\boldsymbol{x}_1$ і $\boldsymbol{x}_n$ — відповідні до них власні вектори одиничної довжини. Тоді найбільше і найменше значення квадратичної форми $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ на сфері одиничного радіуса $\boldsymbol{x}^T \boldsymbol{x} = 1$ дорівнюють λ_1 і λ_n, причому досягаються вони на векторах $\boldsymbol{x}_1$ і $\boldsymbol{x}_n$ відповідно.*

Д о в е д е н н я. Скориставшись спектральним розкладом (6.5.18), одержимо:

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^T \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i \boldsymbol{x}_i^T \boldsymbol{x} = \sum_{i=1}^{n} \lambda_i (\boldsymbol{x}^T \boldsymbol{x}_i)^2. \tag{1}$$

Звідси випливає, що

$$\lambda_n \sum_{i=1}^{n} (\boldsymbol{x}^T \boldsymbol{x}_i)^2 \le \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \le \lambda_1 \sum_{i=1}^{n} (\boldsymbol{x}^T \boldsymbol{x}_i)^2,$$

або

$$\lambda_n \boldsymbol{x}^T \boldsymbol{x} \le \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \le \lambda_1 \boldsymbol{x}^T \boldsymbol{x}. \tag{2}$$

Якщо $\boldsymbol{x}^T \boldsymbol{x} = 1,$ то нерівності (2) набувають вигляду:

$$\lambda_n \le \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \le \lambda_1.$$

Отже, λ_1 і λ_n — найбільше і найменше значення квадратичної форми. Вони досягаються на власних векторах $\boldsymbol{x}_1$ і $\boldsymbol{x}_n$:

$$\boldsymbol{x}_1^T \boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{x}_1^T \lambda_1 \boldsymbol{x}_1 = \lambda_1, \quad \boldsymbol{x}_n^T \boldsymbol{A}\boldsymbol{x}_n = \boldsymbol{x}_n^T \lambda_n \boldsymbol{x}_n = \lambda_n . \square$$

Застосуємо теорему 7.4.1 до задачі, яка полягає в знаходженні для невід'ємно визначеної матриці $\boldsymbol{A}$ порядку n наближеного представлення у вигляді матриці рангу 1. З теореми 3.4.6 випливає, що така матриця має вигляд добутку $\boldsymbol{y}\boldsymbol{y}^T$, де $\boldsymbol{y}$ — деякий n-вимірний вектор. За міру «близькості» матриці $\boldsymbol{A}$ і шуканої матриці $\boldsymbol{y}\boldsymbol{y}^T$ ми візьмемо величину $S = \sum_{i,j=1}^{n} \{\boldsymbol{A} - \boldsymbol{y}\boldsymbol{y}^T\}_{ij}^2$. Можна сказати, що ми шукаємо такий вектор $\boldsymbol{y}$, який відповідає найменшому значенню S.

Теорема 7.4.2. *Для невід'ємно визначеної матриці* $\boldsymbol{A}$ *найменше значення величини* S *досягається при* $\boldsymbol{y} = \sqrt{\lambda}\boldsymbol{x}$, *де* λ — *найбільше власне значення матриці* $\boldsymbol{A}$ *і* $\boldsymbol{x}$ — *відповідний до нього власний вектор одиничної довжини.*

Д о в е д е н н я. Нам будуть потрібні формули для матриць $\boldsymbol{A}$, $\boldsymbol{B}$ порядку n:

$$\mathrm{tr}(\boldsymbol{A} + \boldsymbol{B}) = \mathrm{tr}\, \boldsymbol{A} + \mathrm{tr}\, \boldsymbol{B}, \tag{3}$$

$$\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA}), \tag{4}$$

$$\operatorname{tr}(\boldsymbol{AA}^T) = \sum_{i,j=1}^{n} a_{ij}^2 . \tag{5}$$

Рівність (3) є цілком очевидною. Для доведення рівності (4) скористаємося формулою (1.1.1):

$$\operatorname{tr}(\boldsymbol{AB}) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} a_{ki} b_{ik} = \sum_{i=1}^{n}\sum_{k=1}^{n} b_{ik} a_{ki} = \operatorname{tr}(\boldsymbol{BA}). \tag{6}$$

Рівність (5) можна одержати з формули (6) при $\boldsymbol{B} = \boldsymbol{A}^T$.

Почнемо доведення теореми з перетворення величини S за допомогою формули (5), а потім за допомогою формул (3) і (4):

$$S = \operatorname{tr}[(\boldsymbol{A} - \boldsymbol{yy}^T)(\boldsymbol{A} - \boldsymbol{yy}^T)^T] =$$

$$= \operatorname{tr}(\boldsymbol{A}^2) - \operatorname{tr}(\boldsymbol{yy}^T\boldsymbol{A}) - \operatorname{tr}(\boldsymbol{Ayy}^T) + \operatorname{tr}(\boldsymbol{yy}^T\boldsymbol{yy}^T) =$$

$$= \operatorname{tr}(\boldsymbol{A}^2) - 2\operatorname{tr}(\boldsymbol{y}^T\boldsymbol{Ay}) + \operatorname{tr}(\boldsymbol{y}^T\boldsymbol{yy}^T\boldsymbol{y}) .$$

Матриці $\boldsymbol{y}^T\boldsymbol{Ay}$ і $\boldsymbol{y}^T\boldsymbol{yy}^T\boldsymbol{y}$ є матрицями першого порядку, тому

$$\operatorname{tr}(\boldsymbol{y}^T\boldsymbol{Ay}) = \boldsymbol{y}^T\boldsymbol{Ay}, \qquad \operatorname{tr}(\boldsymbol{y}^T\boldsymbol{yy}^T\boldsymbol{y}) = \boldsymbol{y}^T\boldsymbol{yy}^T\boldsymbol{y} = (\boldsymbol{y}^T\boldsymbol{y})^2.$$

Отже,

$$S = \operatorname{tr}(\boldsymbol{A}^2) - 2\boldsymbol{y}^T\boldsymbol{Ay} + (\boldsymbol{y}^T\boldsymbol{y})^2.$$

Далі будемо вважати, що $\boldsymbol{y} = \sqrt{\lambda}\boldsymbol{x}$, де $\boldsymbol{x}$ — вектор одиничної довжини. При цьому

$$S = \operatorname{tr}(\boldsymbol{A}^2) - 2\lambda\boldsymbol{x}^T\boldsymbol{Ax} + \lambda^2 .$$

На підставі теореми 7.4.1 найбільше значення квадратичної форми $\boldsymbol{x}^T\boldsymbol{Ax}$ дорівнює найбільшому власному значенню μ матриці $\boldsymbol{A}$ і досягається воно на власному векторі $\boldsymbol{x}$, відповідному до власного значення μ . Отже, можна стверджувати, що

$$S \geq \operatorname{tr}(\boldsymbol{A}^2) - 2\lambda\mu + \lambda^2. \tag{7}$$

Мінімум квадратного тричлена з правої частини нерівності (7) досягається при $\lambda = \mu$.

Таким чином, матрицю $\boldsymbol{A}$ можна наближено подати у вигляді матриці $\sqrt{\lambda}\boldsymbol{x}(\sqrt{\lambda}\boldsymbol{x})^T$, де λ — найбільше власне значення матриці $\boldsymbol{A}$ і $\boldsymbol{x}$ — відповідний до нього власний вектор одиничної довжини. Зазначимо, що одержане наближення для матриці $\boldsymbol{A}$ збігається з тим доданком спектрального розкладу (6.5.18), який відповідає найбільшому власному значенню. □

7.4.2. Відношення Релея. Досі ми визначали власні значення матриці як корені її характеристичного рівняння. Покажемо, що у випадку симетричної матриці $\boldsymbol{A}$ її власні значення допускають просте тлумачення як розв'язок задачі максимізації або мінімізації функції

$$\rho(\boldsymbol{x},\ \boldsymbol{A}) = \frac{\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \tag{8}$$

при відсутності або наявності обмежень. Функція (8), яка називається *відношенням Релея*, визначена для всіх векторів $\boldsymbol{x} \in E_n$ за винятком нульового вектора. Тому в подальшому відношення Релея завжди розглядатиметься тільки для ненульових векторів.

Будемо вважати, що власні значення $\lambda_1, ..., \lambda_n$ симетричної матриці $\boldsymbol{A}$ занумеровані так, що вони не зростають:

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n, \tag{9}$$

а нумерація ортонормованих власних векторів $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ відповідає упорядкуванню (9).

Теорема 7.4.3. *Для підпростору U_{kj}, натягнутого на власні вектори $\boldsymbol{x}_k, ..., \boldsymbol{x}_j$, де $k \leq j$, виконуються співвідношення*

$$\lambda_k = \max_{\boldsymbol{x} \in U_{kj}} \rho(\boldsymbol{x},\ \boldsymbol{A}), \quad k = \overline{1,\ n}, \tag{10}$$

$$\lambda_j = \min_{\boldsymbol{x} \in U_{kj}} \rho(\boldsymbol{x}, \boldsymbol{A}), \quad j = \overline{1,\ n}. \tag{11}$$

Д о в е д е н н я. У відповідності до властивості ортонормованості власних векторів симетричної матриці вектор $\boldsymbol{x} \in U_{kj}$ повинен бути ортогональним до власних векторів $\boldsymbol{x}_1, ..., \boldsymbol{x}_{k-1}, \boldsymbol{x}_{j+1}, ..., \boldsymbol{x}_n$, тобто

$$\boldsymbol{x}^T \boldsymbol{x}_i = 0, \quad i = \overline{1,\, k-1}, \quad i = \overline{j+1,\, n}. \tag{12}$$

Тому на підставі формули (1) маємо:

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i (\boldsymbol{x}^T \boldsymbol{x}_i)^2 = \sum_{i=k}^{j} \lambda_i (\boldsymbol{x}^T \boldsymbol{x}_i)^2.$$

Звідси з урахуванням нерівностей (9) знаходимо:

$$\lambda_j \sum_{i=k}^{j} (\boldsymbol{x}^T \boldsymbol{x}_i)^2 \le \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \le \lambda_k \sum_{i=k}^{j} (\boldsymbol{x}^T \boldsymbol{x}_i)^2. \tag{13}$$

З формул (13) і (12) випливає, що

$$\lambda_j \boldsymbol{x}^T \boldsymbol{x} = \lambda_j \sum_{i=1}^{n} (\boldsymbol{x}^T \boldsymbol{x}_i)^2 \le \boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \le \lambda_k \sum_{i=1}^{n} (\boldsymbol{x}^T \boldsymbol{x}_i)^2 = \lambda_k \boldsymbol{x}^T \boldsymbol{x},$$

або

$$\lambda_j \le \rho(\boldsymbol{x},\, \boldsymbol{A}) \le \lambda_k. \tag{14}$$

Неважко перевірити, що екстремальні значення λ_j і λ_k відношення Релея досягаються на власних векторах $\boldsymbol{x}_k$ і $\boldsymbol{x}_j$.

Нерівності (14) рівносильні рівностям (10) і (11), тобто власні значення можна тлумачити як розв'язки задач на максимум і мінімум відношення Релея за додаткової умови $\boldsymbol{x} \in U_{kj}$. Якщо $k = 1$ і $j = n$, то $U_{kj} = E_n$ і задачі максимізації і мінімізації (10), (11) набувають вигляду:

$$\lambda_1 = \max \rho(\boldsymbol{x},\, \boldsymbol{A}), \quad \lambda_n = \min \rho(\boldsymbol{x},\, \boldsymbol{A}), \tag{15}$$

тобто є задачами без обмежень. Їм відповідають нерівності (14) при вказаних значеннях k і j :

$$\lambda_n \le \rho(\boldsymbol{x},\, \boldsymbol{A}) \le \lambda_1. \tag{16}$$

Їх можна також отримати безпосередньо з нерівностей (2). □

Покажемо, як можна застосувати нерівності (16) до проблеми локалізації власних значень, вибираючи конкретний вектор $\boldsymbol{x}$. Будемо вважати, наприклад, що $\boldsymbol{x}$ є вектором натурального базису $\boldsymbol{e}_i$. Тоді відношення Релея дорівнюватиме діагональному елементу a_{ii} матриці $\boldsymbol{A}$ і нерівності (16) набувають вигляду:

$$\lambda_n \le a_{ii} \le \lambda_1, \quad i = \overline{1,\, n}.$$

Відношення Релея можна розглядати для ермітової матриці $\boldsymbol{A}$ і довільних ненульових векторів $\boldsymbol{x} \in C_n$. За аналогією з формулою (8) воно має вигляд:

$$\rho(\boldsymbol{x},\, \boldsymbol{A}) = \frac{\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^* \boldsymbol{x}}. \tag{17}$$

Неважко показати, що функція (17) набуває тільки дійсних значень при комплексних векторах $\boldsymbol{x}$. Можливість застосування відношення Релея (17) до проблеми локалізації власних значень довільної матриці ілюструє

Теорема 7.4.4. *Модулі всіх власних значень будь-якої матриці* $\boldsymbol{A}$ *задовольняють нерівностям*

$$\sigma_n \le |\, \lambda_i \,| \le \sigma_1, \tag{18}$$

де σ_1, σ_n *— найбільше і найменше сингулярні числа матриці* $\boldsymbol{A}$.

Д о в е д е н н я. Зауважимо перш за все, що теорема 6.6.1 залишається в силі для комплексних матриць, якщо під матрицями $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2$ розуміти унітарні матриці.

Нехай λ_i, $\boldsymbol{x}$ — власне значення і відповідний до нього власний вектор матриці $\boldsymbol{A}$, тобто $\boldsymbol{A}\boldsymbol{x} = \lambda_i \boldsymbol{x}$. Очевидно, що

$$(\lambda_i \boldsymbol{x})^* \lambda_i \boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})^* \boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^* \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{x}.$$

З іншого боку,

$$(\lambda_i \boldsymbol{x})^* \lambda_i \boldsymbol{x} = |\, \lambda_i \,|^2 \boldsymbol{x}^* \boldsymbol{x}.$$

Тому

$$|\, \lambda_i \,|^2 = \frac{\boldsymbol{x}^* \boldsymbol{A}^* \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^* \boldsymbol{x}}. \tag{19}$$

Матриця $\boldsymbol{A}^* \boldsymbol{A}$ є ермітовою і її найбільше і найменше власні значення дорівнюють σ_1^2 і σ_n^2 відповідно. Права частина рівності (19) є відношенням Релея $\rho(\boldsymbol{x},\, \boldsymbol{A}^* \boldsymbol{A})$. На підставі нерівностей (16) маємо:

$$\sigma_n^2 \le |\lambda_i|^2 \le \sigma_1^2,$$

звідки випливають нерівності (18), справедливі для всякого власного значення λ_i. □

7.5. Теорема Куранта — Фішера

7.5.1. Доведення теореми Куранта — Фішера. Теорема 7.4.3 характеризує власні значення симетричної матриці за допомогою підпросторів, породжених її власними векторами. Недолік такого підходу полягає в тому, що власні вектори звичайно бувають невідомими. Позбутися цього недоліку можна шляхом переходу до будь-яких підпросторів евклідового простору. Як цього можна досягти, показує

Теорема 7.5.1 **(*теорема Куранта — Фішера*).** *Для власного значення λ_k симетричної матриці **A** порядку n справедливі представлення*

$$\lambda_k = \min_{V_{n-k+1}} \max_{\boldsymbol{x} \in V_{n-k+1}} \rho(\boldsymbol{x}, \boldsymbol{A}), \quad k = \overline{1, n}, \tag{1}$$

$$\lambda_k = \max_{V_k} \min_{\boldsymbol{x} \in V_k} \rho(\boldsymbol{x}, \boldsymbol{A}), \quad k = \overline{1, n}, \tag{2}$$

де через V_l позначається підпростір вимірності l і зовнішня оптимізація виконується на множині всяких підпросторів указаної вимірності.

Д о в е д е н н я. Зауважимо, що при $k = 1$ в (1) і при $k = n$ в (2) підпростори V_{n-k+1} і V_k збігаються з E_n. Для цих випадків твердження (1) і (2) збігаються з (7.4.15).

При доведенні рівності (1) разом з підпростором V_{n-k+1} розглянемо k-вимірний підпростір U_{1k}. Нехай $U = V_{n-k+1} \bigcap U_{1k}$, тоді на підставі формули Грассмана $\dim U \ge 1$.

Припустимо, що $\boldsymbol{u} \in U$ — вектор одиничної довжини. Зважаючи на те, що $\boldsymbol{u} \in V_{n-k+1}$, маємо нерівність

$$\max_{\boldsymbol{x} \in V_{n-k+1}} \rho(\boldsymbol{x}, \boldsymbol{A}) \ge \rho(\boldsymbol{u}, \boldsymbol{A}). \tag{3}$$

З іншого боку, вектор $\boldsymbol{u}$, будучи належним до U_{1k}, є ортогональним до власних векторів $\boldsymbol{x}_{k+1}, ..., \boldsymbol{x}_n$. З урахуванням цієї обставини знаходимо:

$$\rho(\boldsymbol{u}, \boldsymbol{A}) = \frac{\boldsymbol{u}^T \boldsymbol{A}\boldsymbol{u}}{\boldsymbol{u}^T \boldsymbol{u}} = \boldsymbol{u}^T \boldsymbol{A}\boldsymbol{u} = \boldsymbol{u}^T \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i \boldsymbol{x}_i^T \boldsymbol{u} = \sum_{i=1}^{n} \lambda_i (\boldsymbol{u}^T \boldsymbol{x}_i)^2 =$$

$$= \sum_{i=1}^{k} \lambda_i (\boldsymbol{u}^T \boldsymbol{x}_i)^2 \geq \lambda_k \sum_{i=1}^{k} (\boldsymbol{u}^T \boldsymbol{x}_i)^2 = \lambda_k \sum_{i=1}^{n} (\boldsymbol{u}^T \boldsymbol{x}_i)^2 = \lambda_k \boldsymbol{u}^T \boldsymbol{u} = \lambda_k. \quad (4)$$

З формул (3) і (4) випливає, що

$$\max_{\boldsymbol{x} \in V_{n-k+1}} \rho(\boldsymbol{x}, \boldsymbol{A}) \geq \lambda_k, \quad (5)$$

а тому і

$$\min_{V_{n-k+1}} \max_{\boldsymbol{x} \in V_{n-k+1}} \rho(\boldsymbol{x}, \boldsymbol{A}) \geq \lambda_k. \quad (6)$$

Візьмемо формулу (7.4.10) при фіксованому k і змінному j. Зрозуміло, що її права частина не залежить від j, тому можна вважати, що при $k = \overline{1, n}$ і $k \leq j$

$$\lambda_k = \min_{U_{kj}} \max_{\boldsymbol{x} \in U_{kj}} \rho(\boldsymbol{x}, \boldsymbol{A}).$$

Звідси при $j = n$ маємо:

$$\lambda_k = \min_{U_{kn}} \max_{\boldsymbol{x} \in U_{kn}} \rho(\boldsymbol{x}, \boldsymbol{A}). \quad (7)$$

де U_{kn} — підпростір вимірності $n - k + 1$.

Зіставляючи нерівність (6) і рівність (7), бачимо, що існує $(n - k + 1)$-вимірний підпростір U_{kn}, для якого нерівність (6) виконується як рівність. Таким чином, рівність (1) доведена.

Рівність (2) доводиться аналогічно. Наведемо відповідні міркування в скороченому вигляді.

Розглядаючи разом з підпростором V_k підпростір U_{kn}, подібно до нерівності (5) можна показати, що

$$\min_{\boldsymbol{x} \in V_k} \rho(\boldsymbol{x}, \boldsymbol{A}) \leq \lambda_k, \quad (8)$$

а тому і

$$\max_{V_k} \min_{x \in V_k} \rho(\boldsymbol{x},\ \boldsymbol{A}) \le \lambda_k. \tag{9}$$

З формули (7.4.11) випливає, що при $k = \overline{1, n}$ і $l \le k$

$$\lambda_k = \min_{x \in U_{lk}} \rho(\boldsymbol{x},\ \boldsymbol{A}).$$

Права частина цієї рівності при фіксованому k не залежить від l, тому

$$\lambda_k = \max_{U_{lk}} \min_{x \in U_{lk}} \rho(\boldsymbol{x},\ \boldsymbol{A}),$$

звідки при $l = 1$ маємо:

$$\lambda_k = \max_{U_{1k}} \min_{x \in U_{1k}} \rho(\boldsymbol{x},\ \boldsymbol{A}). \tag{10}$$

Оскільки $\dim V_k = \dim U_{1k}$, то з формул (9) і (10) випливає рівність (2). □

7.5.2. Задачі зі зв'язками. Припустимо, що розшукуються екстремальні значення відношення Релея (7.4.8) для ненульових векторів $\boldsymbol{x} \in E_n$, які задовольняють додатковій умові (системі рівнянь-зв'язків)

$$\boldsymbol{Cx} = \boldsymbol{0} \tag{11}$$

з $m \times n$-матрицею $\boldsymbol{C}$ рангу m, де $m < n$.

Задачу зі зв'язками можна звести до звичайної задачі без зв'язків для деякої симетричної матриці порядку $n - m$. Щоб пересвідчитися в цьому, введемо $n \times (n - m)$-матрицю $\boldsymbol{Q}$, стовпці якої утворюють ортонормований базис для $\mathcal{N}(\boldsymbol{C})$. Рівність (11) показує, що $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{C})$, тому $\boldsymbol{x} = \boldsymbol{Qy}$ для деякого $(n - m)$-вимірного вектора $\boldsymbol{y}$. Тепер відношення Релея дорівнює

$$\rho(\boldsymbol{x}, \boldsymbol{A}) = \frac{\boldsymbol{x}^T \boldsymbol{Ax}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{\boldsymbol{y}^T \boldsymbol{Q}^T \boldsymbol{AQy}}{\boldsymbol{y}^T \boldsymbol{Q}^T \boldsymbol{Qy}} = \frac{\boldsymbol{y}^T \boldsymbol{Q}^T \boldsymbol{AQy}}{\boldsymbol{y}^T \boldsymbol{y}} = \rho(\boldsymbol{y}, \boldsymbol{D}), \tag{12}$$

де $\boldsymbol{D} = \boldsymbol{Q}^T \boldsymbol{AQ}$ і враховано, що $\boldsymbol{Q}^T \boldsymbol{Q} = \boldsymbol{E}$.

Ми одержали задачу без зв'язків стосовно відношення Релея для симетричної матриці $\boldsymbol{D}$ порядку $n - m$. При цьому μ_i і $\boldsymbol{x}_i$ є власним значенням і відповідним власним вектором для задачі зі зв'язками тоді і тільки тоді, коли μ_i є власним значенням матриці $\boldsymbol{D}$ з власним вектором $\boldsymbol{y}_i = \boldsymbol{Q}^T \boldsymbol{x}_i$.

Нехай власні значення $\mu_1, ..., \mu_{n-m}$ задачі зі зв'язками занумеровані так, що вони не зростають:

$$\mu_1 \ge \mu_2 \ge ... \ge \mu_{n-m}, \tag{13}$$

а $\boldsymbol{y}_1, ..., \boldsymbol{y}_{n-m}$ є відповідними власними векторами матриі $\boldsymbol{D}$, що утворюють ортонормовану систему. При цьому і вектори $\boldsymbol{x}_1, ..., \boldsymbol{x}_{n-m}$ є ортонормованими, тому що

$$\boldsymbol{x}_i^T \boldsymbol{x}_j = (\boldsymbol{Q}\,\boldsymbol{y}_i)^T \boldsymbol{Q}\,\boldsymbol{y}_j = \boldsymbol{y}_i^T \boldsymbol{y}_j = \delta_{ij}, \quad i, j = \overline{1, n-m}.$$

Подібно до теореми 7.4.3 доводиться

Теорема 7.5.2. *Для підпростору $U_{kj} \subset E_n$, породженого власними векторами $\boldsymbol{x}_k, ..., \boldsymbol{x}_j$ задачі зі зв'язками виконуються співвідношення*

$$\mu_k = \max_{\boldsymbol{x} \in U_{kj}} \rho(\boldsymbol{x}, \boldsymbol{A}), \quad k = \overline{1, n-m}, \tag{14}$$

$$\mu_j = \min_{\boldsymbol{x} \in U_{kj}} \rho(\boldsymbol{x}, \boldsymbol{A}), \quad j = \overline{1, n-m}. \tag{15}$$

Д о в е д е н н я. Подібно до (7.4.10), (7.4.11) можна показати, що

$$\mu_k = \max_{\boldsymbol{y} \in W_{kj}} \rho(\boldsymbol{y}, \boldsymbol{D}), \quad k = \overline{1, n-m},$$

$$\mu_j = \min_{\boldsymbol{y} \in W_{kj}} \rho(\boldsymbol{y}, \boldsymbol{D}), \quad j = \overline{1, n-m},$$

де $W_{kj} \subset E_{n-m}$ — підпростір, натягнутий на вектори $\boldsymbol{y}_k, ..., \boldsymbol{y}_j$. За допомогою формули (12) останні рівності можна перетворити в рівності (14) і (15), де U_{kj} розглядається як підпростір простору E_n. □

Міркуваннями, на яких ґрунтується доведення теореми Куранта — Фішера, можна скористатися при дослідженні співвідношень між власними значеннями для задачі без зв'язків і задачі зі зв'язками.

***Теорема 7.5.3* (*теорема Релея*).** *Якщо* $\lambda_1, ..., \lambda_n$ *— власні значення матриці* $\boldsymbol{A}$, *занумеровані так, що вони не зростають, і* $\mu_1, ..., \mu_{n-m}$ *— власні значення (13) задачі зі зв'язками (11), то*

$$\lambda_{k+m} \leq \mu_k \leq \lambda_k, \quad k = \overline{1, n-m}. \tag{16}$$

Д о в е д е н н я. З формули (5) знаходимо:

$$\lambda_{k+m} \leq \max_{\boldsymbol{x} \in V_{n-k-m+1}} \rho(\boldsymbol{x}, \boldsymbol{A}). \tag{17}$$

З іншого боку, формула (14) показує, що

$$\mu_k = \max_{\boldsymbol{x} \in U_{k,\, n-m}} \rho(\boldsymbol{x}, \boldsymbol{A}). \tag{18}$$

Оскільки $\dim V_{n-k-m+1} = \dim U_{k, m-n}$, то праві частини формул (17) і (18) збігаються. Тому має місце ліва нерівність (16).

Для доведення правої нерівності (16) візьмемо нерівність (8) і зіставимо її з рівністю

$$\mu_k = \min_{\boldsymbol{x} \in U_{1k}} \rho(\boldsymbol{x}, \boldsymbol{A}), \tag{19}$$

яка одержується з (15). Оскільки $\dim V_k = \dim U_{1k},$ то ліва частина нерівності (8) збігається з правою частиною рівності (19). Тому права нерівність (16) виконується. ☐

7.5.3. Деякі застосування теореми Куранта — Фішера. Одне з важливих застосувань теореми Куранта — Фішера пов'язане із задачею порівняння власних значень матриць $\boldsymbol{A}$, $\boldsymbol{B}$ і $\boldsymbol{A} + \boldsymbol{B}$. Власні значення матриці $\boldsymbol{A}$ позначатимемо через $\lambda_i(\boldsymbol{A})$.

***Теорема 7.5.4* (*теорема Вейля*).** *Нехай* $\boldsymbol{A}$ *і* $\boldsymbol{B}$ *— симетричні матриці і власні значення* $\lambda_i(\boldsymbol{A}),\ \lambda_i(\boldsymbol{B}),\ \lambda_i(\boldsymbol{A} + \boldsymbol{B}),\ i = \overline{1, n},$ *занумеровані так, що вони не зростають. Тоді*

$$\lambda_k(\boldsymbol{A}) + \lambda_n(\boldsymbol{B}) \leq \lambda_k(\boldsymbol{A} + \boldsymbol{B}) \leq \lambda_k(\boldsymbol{A}) + \lambda_1(\boldsymbol{B}), \quad k = \overline{1, n}. \tag{20}$$

Д о в е д е н н я. На підставі формули (7.4.16) маємо:

$$\lambda_n(\boldsymbol{B}) \leq \rho(\boldsymbol{x}, \boldsymbol{B}) \leq \lambda_1(\boldsymbol{B}). \tag{21}$$

Скориставшись формулою (1) і лівою нерівністю (21), знаходимо:

$$\lambda_k(\boldsymbol{A}+\boldsymbol{B}) = \min_{V_{n-k+1}} \max_{\boldsymbol{x}\in V_{n-k+1}} \rho(\boldsymbol{x}, \boldsymbol{A}+\boldsymbol{B}) =$$

$$= \min_{V_{n-k+1}} \max_{\boldsymbol{x}\in V_{n-k+1}} [\rho(\boldsymbol{x}, \boldsymbol{A}) + \rho(\boldsymbol{x}, \boldsymbol{B})] \geq$$

$$\geq \min_{V_{n-k+1}} \max_{\boldsymbol{x}\in V_{n-k+1}} [\rho(\boldsymbol{x}, \boldsymbol{A}) + \lambda_n(\boldsymbol{B})] = \lambda_k(\boldsymbol{A}) + \lambda_n(\boldsymbol{B}).$$

Аналогічно, користуючись формулою (2) і правою нерівністю (21), одержимо доведення правої нерівності (20).□

П р и к л а д 1. Застосуємо теорему Вейля до матриці

$$\begin{pmatrix} 9{,}2 & 2{,}4 \\ 2{,}4 & 8{,}8 \end{pmatrix}.$$

Будемо вважати, що вона дорівнює $\boldsymbol{A} + \boldsymbol{B}$, де

$$\boldsymbol{A} = \begin{pmatrix} 8{,}2 & 2{,}4 \\ 2{,}4 & 6{,}8 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Неважко вереконатися в тому, що $\lambda_1(\boldsymbol{A}) = 10,\ \lambda_2(\boldsymbol{A}) = 5,$ $\lambda_1(\boldsymbol{B}) = 2,\ \lambda_2(\boldsymbol{B}) = 1.$ З формули (20) випливають оцінки для власних значень матриці $\boldsymbol{A} + \boldsymbol{B}$:

$$11 \leq \lambda_1(\boldsymbol{A}+\boldsymbol{B}) \leq 12,\ \ 6 \leq \lambda_2(\boldsymbol{A}+\boldsymbol{B}) \leq 7.$$

Для порівняння наведемо величини власних значень матриці $\boldsymbol{A} + \boldsymbol{B}$, знайдених як корені характеристичного многочлена:

$$\lambda_1(\boldsymbol{A}+\boldsymbol{B}) = 9 + \sqrt{5{,}8} \approx 11{,}41, \quad \lambda_2(\boldsymbol{A}+\boldsymbol{B}) = 9 - \sqrt{5{,}8} \approx 6{,}59.$$

Очевидно, що з цими значеннями узгоджуються одержані вище нерівності для власних значень матриці $\boldsymbol{A} + \boldsymbol{B}$.△

У теоремі Вейля визначені двобічні межі для власних значень суми $\boldsymbol{A} + \boldsymbol{B}$ довільних симетричних матриць $\boldsymbol{A}$ і $\boldsymbol{B}$. Ці межі можна уточнювати, якщо вважати $\boldsymbol{B}$ матрицею спеціального вигляду, наприклад, невід'ємно визначеною.

Теорема 7.5.5. *Нехай **A** і **B** — симетричні матриці, причому **B** є невід'ємно визначеною. Припустимо, що власні значення матриць **A** і **A** + **B** занумеровані так, що вони не зростають. Тоді*

$$\lambda_k(\boldsymbol{A}) \le \lambda_k(\boldsymbol{A}+\boldsymbol{B}), \quad k=\overline{1,\ n}. \tag{22}$$

Д о в е д е н н я випливає з лівої нерівності (20), якщо врахувати, що $\lambda_n(\boldsymbol{B}) = 0.$ □

Нерівності (22) показують, що власні значення симетричної матриці не зменшуються при додаванні до неї невід'ємно визначеної матриці. Можна чекати, що за певних умов, накладених на матрицю $\boldsymbol{B}$, власні значення матриці $\boldsymbol{A}$ можуть зрости як завгодно багато. Проте, якщо $\boldsymbol{B}$ є невід'ємно визначеною матрицею рангу r, це припущення може бути справедливим лише для перших r власних значень. Існують незалежні від $\boldsymbol{B}$ верхні межі для $\lambda_{r+1}(\boldsymbol{A}+\boldsymbol{B}), ..., \lambda_n(\boldsymbol{A}+\boldsymbol{B})$ і це показує наступна теорема, яку ми доведемо, не вимагаючи невід'ємної визначеності матриці $\boldsymbol{B}$.

Теорема 7.5.6. *Нехай* $\boldsymbol{A}$ *і* $\boldsymbol{B}$ *— симетричні матриці, причому* $\boldsymbol{B}$ *є матрицею рангу* $r < n$. *Припустимо, що власні значення матриць* $\boldsymbol{A}$ *і* $\boldsymbol{A} + \boldsymbol{B}$ *занумеровані так, що вони не зростають. Тоді*

$$\lambda_k(\boldsymbol{A}+\boldsymbol{B}) \le \lambda_{k-r}(\boldsymbol{A}), \quad k=\overline{r+1,\ n}. \tag{23}$$

Д о в е д е н н я. За допомогою спектрального розкладу матриці $\boldsymbol{B}$ знаходимо:

$$\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x} = \sum_{i=1}^{r}\lambda_i(\boldsymbol{B})\,(\boldsymbol{x}^T\boldsymbol{z}_i)^2,$$

де $\boldsymbol{z}_1, .., \boldsymbol{z}_r$ — ортонормовані власні вектори матриці $\boldsymbol{B}$, відповідні до ненульових власних значень $\lambda_1(\boldsymbol{B}), \dots, \lambda_r(\boldsymbol{B})$. Очевидно, що

$$\boldsymbol{x}^T(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \sum_{i=1}^{r}\lambda_i(\boldsymbol{B})(\boldsymbol{x}^T\boldsymbol{z}_i)^2. \tag{24}$$

Розглянемо задачі про знаходження власних значень для $\boldsymbol{A}$ і $\boldsymbol{A} + \boldsymbol{B}$ за наявності зв'язків $\boldsymbol{x}^T\boldsymbol{z}_i = 0, \ \ i=\overline{1,\ r}.$ Рівність (24) показує, що ці задачі збігаються і мають однакові власні значення (13). Застосувавши теорему Релея до обох задач, одержимо:

$$\begin{gathered}\lambda_{k+r}(\boldsymbol{A}) \le \mu_k \le \lambda_k(\boldsymbol{A}), \quad k=\overline{1,\ n-r},\\ \lambda_{k+r}(\boldsymbol{A}+\boldsymbol{B}) \le \mu_k \le \lambda_k(\boldsymbol{A}+\boldsymbol{B}), \quad k=\overline{1,\ n-r}.\end{gathered} \tag{25}$$

Звідси випливає нерівність

$$\lambda_{k+r}(\boldsymbol{A}+\boldsymbol{B}) \leq \lambda_k(\boldsymbol{A}), \quad k=\overline{1,\, n-r}, \tag{26}$$

рівносильна нерівності (23). □

Одержаний результат застосовується в наступній теоремі при розгляді однорангового «збурення» матриці $\boldsymbol{A}$, тобто у випадку $r(\boldsymbol{B})=1$.

Теорема 7.5.7. *Нехай* $\boldsymbol{A}$ *і* $\boldsymbol{B}$ *— симетричні матриці і власні значення матриць* $\boldsymbol{A}$, $\boldsymbol{A}+\boldsymbol{B}$ *занумеровані так, що вони не зростають. Припустимо, що* $\boldsymbol{B}$ *є матрицею рангу* 1*, єдиним ненульовим власним значенням якої є* $\lambda(\boldsymbol{B})$. *Якщо* $\lambda(\boldsymbol{B})>0$, *то*

$$\lambda_{k-1}(\boldsymbol{A}) \geq \lambda_k(\boldsymbol{A}+\boldsymbol{B}) \geq \lambda_k(\boldsymbol{A}), \quad k=\overline{2,\, n}, \tag{27}$$

а при $\lambda(\boldsymbol{B})<0$

$$\lambda_k(\boldsymbol{A}) \geq \lambda_k(\boldsymbol{A}+\boldsymbol{B}) \geq \lambda_{k+1}(\boldsymbol{A}), \quad k=\overline{1,\, n-1}. \tag{28}$$

Д о в е д е н н я. Ліва нерівність (27) є нерівністю (23) при $r=1$. За допомогою нерівностей (25) подібно до (26) можна довести, що

$$\lambda_{k+r}(\boldsymbol{A}) \leq \lambda_k(\boldsymbol{A}+\boldsymbol{B}), \quad k=\overline{1, n-r}.$$

Звідси при $r=1$ одержується права нерівність (28).

Далі доведемо праву нерівність (27) і ліву нерівність (28), урахувавши особливість матриці $\boldsymbol{B}$. Якщо $\lambda(\boldsymbol{B})>0$, то матриця $\boldsymbol{B}$ є невід'ємно визначеною, тому права нерівність (27) збігається з нерівністю (22). При $\lambda(\boldsymbol{B})<0$ матриця $\boldsymbol{B}$ є недодатно визначеною, тому ліва нерівність (28) одержується з правої нерівності (20) з урахуванням того, що для недодатно визначеної матриці $\lambda_1(\boldsymbol{B})=0$. □

П р и к л а д 2. Застосуємо теорему 7.5.7 до матриці, розглянутої в прикладі 1 після теореми Вейля. Будемо вважати, що

$$\boldsymbol{A}=\begin{pmatrix} 6{,}8 & 0 \\ 0 & 6{,}4 \end{pmatrix}, \quad \boldsymbol{B}=\begin{pmatrix} 2{,}4 & 2{,}4 \\ 2{,}4 & 2{,}4 \end{pmatrix}.$$

Тут $\lambda_1(\boldsymbol{A}) = 6{,}8,\ \lambda_2(\boldsymbol{A}) = 6{,}4$ і $\lambda(\boldsymbol{B}) > 0.$ Скориставшись формулою (27), одержимо

$$6{,}4 \le \lambda_2(\boldsymbol{A} + \boldsymbol{B}) \le 6{,}8.$$

Припустимо, що

$$\boldsymbol{A} = \begin{pmatrix} 11{,}6 & 0 \\ 0 & 11{,}2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} -2{,}4 & 2{,}4 \\ 2{,}4 & -2{,}4 \end{pmatrix},$$

тоді $\lambda_1(\boldsymbol{A}) = 11{,}6,\ \lambda_2(\boldsymbol{A}) = 11{,}2$ і $\lambda(\boldsymbol{B}) < 0.$ З формули (28) випливає, що

$$11{,}2 \le \lambda_1(\boldsymbol{A} + \boldsymbol{B}) \le 11{,}6.$$

Одержані оцінки для $\lambda_1(\boldsymbol{A} + \boldsymbol{B})$ і $\lambda_2(\boldsymbol{A} + \boldsymbol{B})$ виявилися кращими в порівнянні з обчисленими в прикладі 1. $\triangle$

Між власними значеннями симетричної матриці $\boldsymbol{A}$ і власними значеннями її підматриць, відповідних до головних мінорів, існують залежності, які можна проаналізувати за допомогою теореми Куранта — Фішера. Ми зупинимося на найпростішому випадку, коли така підматриця має порядок $n-1.$ Не порушуючи загальності, можна взяти підматрицю $\boldsymbol{A}_{n-1},$ розміщену в лівому верхньому куті матриці $\boldsymbol{A}$.

Теорема 7.5.8. *Нехай власні значення матриць $\boldsymbol{A}$ і $\boldsymbol{A}_{n-1}$ занумеровані так, що вони не зростають. Тоді власні значення матриці $\boldsymbol{A}_{n-1}$ розділяють власні значення матриці $\boldsymbol{A}$, тобто*

$$\lambda_k(\boldsymbol{A}) \ge \lambda_k(\boldsymbol{A}_{n-1}) \ge \lambda_{k+1}(\boldsymbol{A}), \quad k = \overline{1,\, n-1}. \tag{29}$$

Д о в е д е н н я. Нехай $\boldsymbol{t}_1, ..., \boldsymbol{t}_{n-1}$ — ортонормовані власні вектори матриці $\boldsymbol{A}_{n-1}$. У відповідності до формул (7.4.10) і (7.4.11) маємо:

$$\max_{\boldsymbol{t}\in U_{k,\,n-1}} \rho(\boldsymbol{t}, \boldsymbol{A}_{n-1}) = \lambda_k(\boldsymbol{A}_{n-1}) = \min_{\boldsymbol{t}\in U_{1k}} \rho(\boldsymbol{t}, \boldsymbol{A}_{n-1}), \quad k = \overline{1,\, n-1}, \tag{30}$$

де підпростори $U_{k,\,n-1}$ і U_{1k} породжуються відповідними власними векторами $\boldsymbol{t}_1, ..., \boldsymbol{t}_{n-1}$.

Зіставимо кожному $(n-1)$ - вимірному вектору $\boldsymbol{t}$ вектор $\boldsymbol{x} = \begin{pmatrix} \boldsymbol{t} \\ 0 \end{pmatrix}$ вимірності n. Неважко показати, що

$$\rho(\boldsymbol{t}, \boldsymbol{A}_{n-1}) = \rho(\boldsymbol{x}, \boldsymbol{A}) \tag{31}$$

для відповідних векторів $\boldsymbol{t}$ і $\boldsymbol{x}$. Підпосторам $U_{k,n-1}$ і U_{1k} будуть відповідати в E_n підпростори V_{n-k} і V_k тих же вимірностей.

Тепер з формул (30), (31), (5) і (8) випливають нерівності (29).□

З а у в а ж е н н я. Усі результати, одержані для відношення Релея, легко узагальнити на випадок функції $\dfrac{\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{H}\boldsymbol{x}}$, де $\boldsymbol{A}$ і $\boldsymbol{H}$ — симетричні матриці, причому $\boldsymbol{H}$ є додатно визначеною. Тут маємо справу з коренями рівняння $\det(\boldsymbol{A} - \lambda\boldsymbol{H}) = 0$ і відповідними до них $\boldsymbol{H}$-ортонормованими власними векторами.

Вправи до глави 7

7.1. Показати, що для будь-якої додатно визначеної матриці $\boldsymbol{A}$ існує представлення $\boldsymbol{A} = \boldsymbol{R}^T\boldsymbol{R}$, де $\boldsymbol{R}$ — верхня трикутна матриця.

7.2. Нехай матриця $\boldsymbol{A}$ є додатно визначеною, а матриця $\boldsymbol{C}$ — невиродженою. Довести, що матриця $\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}$ є додатно визначеною.

7.3. Сформулювати умови від'ємної визначеності симетричної матриці. Довести, зокрема, критерій Сільвестра: матриця є від'ємно визначеною тоді і тільки тоді, коли знаки провідних мінорів чергуються, причому $\det \boldsymbol{A}_1 < 0$.

7.4. Довести, що коли матриця $\boldsymbol{A}$ є додатно визначеною, то

$$r(\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C}) = r(\boldsymbol{C}).$$

7.5. Довести, що для додатно визначеної матриці $\boldsymbol{A}$ існує невироджена матриця $\boldsymbol{C}$ така, що $\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C} = \boldsymbol{E}$.

7.6. Показати, що в додатно визначеній матриці максимальний за модулем елемент знаходиться на головній діагоналі.

7.7. Показати, що при збільшенні діагонального елемента a_{ii} додатно визначеної матриці $\boldsymbol{A}$ на величину α одержується матриця $\boldsymbol{B}$ така, що $\det \boldsymbol{B} > \det \boldsymbol{A}$.

7.8. Нехай $\boldsymbol{A}$ і $\boldsymbol{B}$ — додатно визначена і довільна симетрична матриці відповідно. Довести, що $\boldsymbol{A} + t\boldsymbol{B}$ є додатно визначеною матрицею для всіх досить малих за модулем значень t.

7.9. Довести висновок 2 з теореми 7.1.4 за допомогою частини 5) теореми 7.1.1.

7.10. Показати, що міркування, наведені в зауваженні 1 з розділу 7.1, можна обґрунтувати за допомогою теореми 7.1.4, якщо за $\boldsymbol{W}$ взяти матрицю, що одержується з одиничної матриці викреслюванням вибраних належним чином стовпців.

7.11. Нехай $\boldsymbol{A}$ і $\boldsymbol{B}$ — додатно (невід'ємно) визначені матриці. Показати, що для будь-яких додатних (невід'ємних) чисел α і β матриця $\alpha\boldsymbol{A} + \beta\boldsymbol{B}$ є додатно (невід'ємно) визначеною.

7.12. Що можна стверджувати стосовно невід'ємно визначеної матриці $\boldsymbol{A}$, якщо $\operatorname{tr}\boldsymbol{A} = 0$?

7.13. Сформулювати аналог теореми 7.1.2 для недодатно визначених матриць.

7.14. Довести, що $\boldsymbol{A}$ є невід'ємно визначеною матрицею порядку n і рангу r тоді і тільки тоді, коли існує матриця $\boldsymbol{W}$ порядку n і рангу r така, що $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}$.

7.15. Для невід'ємно визначеної матриці $\boldsymbol{A}$ порядку n і рангу r довести, що існує $n \times r$-матриця $\boldsymbol{C}$ рангу r така, що $\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C} = \boldsymbol{E}_r$.

7.16. Нехай $\boldsymbol{A}$ — невід'ємно визначена матриця порядку n і рангу $r < n$. Довести, що існує дійсна невироджена матриця $\boldsymbol{C}$ така, що

$$\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{C} = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Як виглядає аналог цієї формули для симетричної матриці, у якої з r ненульових власних значень p доданих і q від'ємних?

7.17. Для невід'ємно визначеної матриці $\boldsymbol{A}$ показати, що з умови $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = 0$ випливає рівність $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$.

7.18. Для матриці $\boldsymbol{R}$ з $\boldsymbol{LR}$-розкладу симетричної матриці $\boldsymbol{A}$ довести, що кількість її нульових, додатних і від'ємних діагональних елементів збігається з кількістю нульових, додатних і від'ємних власних значень матриці $\boldsymbol{A}$.

7.19. Показати, що ранг і сигнатура квадратичної форми мають однакову парність.

7.20. Довести, що

$$\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\boldsymbol{j}^T\boldsymbol{x}\right)^2 = \boldsymbol{x}^T\left(\boldsymbol{E} - \frac{1}{n}\boldsymbol{j}\,\boldsymbol{j}^T\right)\boldsymbol{x},$$

де $\boldsymbol{j}$ — вектор з одиничними координатами і $\boldsymbol{x} = (x_1 \quad \dots \quad x_n)^T$.

7.21. Нехай симетрична матриця $\boldsymbol{A}$ поділена на чотири блоки, причому $\boldsymbol{A}_{11}$ є нульовою підматрицею першого порядку, підматриця $\boldsymbol{A}_{12}$ не є нульовою, а підматриця $\boldsymbol{A}_{22}$ є додатно визначеною. Довести, що $\det \boldsymbol{A}$ є від'ємно визначеною квадратичною формою відносно елементів підматриці $\boldsymbol{A}_{12}$.

7.22. Нехай $\boldsymbol{\Gamma}_m$ — матриця Грама системи векторів $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, …, $\boldsymbol{a}_m$. Довести, що

$$\boldsymbol{x}^T\boldsymbol{\Gamma}_m\boldsymbol{x} = \boldsymbol{y}^T\boldsymbol{y}, \quad \boldsymbol{x} = (x_1 \quad \dots \quad x_m)^T, \quad \boldsymbol{y} = \sum_{i=1}^{m} x_i\boldsymbol{a}_i.$$

Як звідси випливає теорема 5.2.1?

7.23. Довести, що подібні матриці мають однакові сліди.

7.24. Нехай $\boldsymbol{A}$ — додатно визначена матриця і $\boldsymbol{b}$ — деякий вектор. Довести, що максимальне значення функції

$$f(\boldsymbol{x}) = \frac{(\boldsymbol{b}^T\boldsymbol{x})^2}{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}$$

дорівнює $\boldsymbol{b}^T\boldsymbol{A}^{-1}\boldsymbol{b}$ і досягається воно при $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

Г Л А В А 8

ГІПЕРПОВЕРХНІ ДРУГОГО ПОРЯДКУ

У цій главі на підставі результатів, одержаних у розділі 7.2 стосовно квадратичних форм, досліджуються довільні гіперповерхні другого порядку в афінному просторі S_n та їх окремі випадки в просторах S_2 і S_3. Нижче ми побачимо, що для деяких рівняннь гіперповерхонь другого порядку не існують точки, координати яких задовольняють цим рівнянням. Такі гіперповерхні називаються уявними. Вони не вилучаються з розгляду, оскільки предметом теорії, якій присвячена ця глава, є передусім виявлення алгебраїчних властивостей рівнянь гіперповерхонь.

8.1. Загальна теорія гіперповерхонь другого порядку

8.1.1. Центр гіперповерхні другого порядку. *Гіперповерхнею другого порядку* в афінному просторі S_n називається множина точок цього простору, координати $x_1, \ldots, x_n$ яких задовольняють рівнянню

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} + 2\boldsymbol{b}^T \boldsymbol{x} + c = 0, \tag{1}$$

де $\boldsymbol{x} = (x_1 \quad \ldots \quad x_n)^T$. Вважається, що базис в S_n є ортонормованим, а число c, вектор $\boldsymbol{b}$ і симетрична матриця $\boldsymbol{A}$ є дійсними.

Центром гіперповерхні другого порядку називають таку точку простору S_n, відносно якої всі точки гіперповерхні розміщенні симетрично парами. Розглянемо рівняння

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} + c = 0. \tag{2}$$

Якщо точка $\boldsymbol{x}$ лежить на гіперповерхні (2), то і симетрична відносно початку координат точка $-\boldsymbol{x}$ теж лежить на ній. Тому початок координат є центром гіперповерхні. У загальному випадку гіперповерхні (1) ми будемо вважати її центром таку точку, що коли розмістити в ній початок координат, то рівняння гіперповерхні набуде вигляду (2).

Паралельним перенесенням в афінному просторі S_n називається перетворення координат вигляду:

$$\boldsymbol{x} = \boldsymbol{z} + \boldsymbol{h}, \tag{3}$$

де $\boldsymbol{z} = (z_1 \ \ldots \ z_n)^T$ і $\boldsymbol{h}$ — фіксована точка простору S_n, яку називають новим початком координат. Очевидно, що при паралельному перенесенні будь-який фіксований базис не змінюється.

Зробимо паралельне перенесеня таким чином, щоб рівняння (1) перетворилося до вигляду (2). Підставимо вираз для $\boldsymbol{x}$ з (3) у рівняння (1). У результаті одержимо:

$$\boldsymbol{z}^T \boldsymbol{A}\boldsymbol{z} + 2(\boldsymbol{A}\boldsymbol{h} + \boldsymbol{b})^T \boldsymbol{z} + d = 0, \qquad d = \boldsymbol{h}^T \boldsymbol{A}\boldsymbol{h} + 2\boldsymbol{b}^T\boldsymbol{h} + c. \tag{4}$$

Неважко побачити, що рівняння (1) перетворюється в рівняння (2), коли за початок координат $\boldsymbol{h}$ узяти розв'язок системи рівнянь $\boldsymbol{A}\boldsymbol{x} = -\boldsymbol{b}$.

Якщо $\det \boldsymbol{A} \neq 0,$ то система $\boldsymbol{A}\boldsymbol{x} = -\boldsymbol{b}$ має розв'язок. Тоді гіперповерхня має єдиний центр і вона називається *центральною*. При $\det \boldsymbol{A} = 0$ система або несумісна і тоді центру нема, або вона має нескінченну множину розв'язків і тоді центрів нескінченно багато.

8.1.2. Перетин прямої лінії і гіперповерхні другого порядку. Виявимо особливості взаємного розміщення прямої лінії

$$\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1 \tag{5}$$

і гіперплощини (1). Підставляючи вираз для $\boldsymbol{x}$ з рівності (5) в рівняння (1), одержимо:

$$M\alpha_1^2 + 2N\alpha_1 + P = 0, \tag{6}$$

де

$$M = \boldsymbol{f}_1^T \boldsymbol{A}\boldsymbol{f}_1, \qquad N = (\boldsymbol{A}\boldsymbol{f}_0 + \boldsymbol{b})^T \boldsymbol{f}_1, \quad P = \boldsymbol{f}_0^T \boldsymbol{A}\boldsymbol{f}_0 + 2\boldsymbol{b}^T \boldsymbol{f}_0 + c.$$

Рівняння (6) визначає точки перетину прямої лінії і гіперповерхні. Дослідимо особливості цього рівняння.

Якщо $M \neq 0$, то рівняння (6) є квадратним. У цьому випадку є дві різні точки перетину при $N^2 - MP > 0$ і нема точок перетину, якщо $N^2 - MP < 0$. Приипустимо, що

$$N^2 - MP = 0, \tag{7}$$

але принаймні один з коефіцієнтів M, N, P є ненульовим. Маємо граничний випадок, коли дві точки перетину збігаються. Проте в цьому випадку говорять про подвійну точку перетину. Якщо в подвійній точці вектор $\boldsymbol{Ax} + \boldsymbol{b}$ не є нульовим, говорять, що в цій точці пряма лінія (3) дотикається до гіперповерхні (1).

Нехай $M = 0$ і $N \neq 0$. Тоді рівняння (6) буде рівнянням першого степеня і його розв'язок є таким:

$$\alpha_1 = -\frac{P}{2N}.$$

Якщо $M = N = 0$ і $P \neq 0$, маємо суперечливу рівність (6). При цьому пряма лінія і гіперповерхня не перетинаються.

Нарешті, при

$$M = N = P = 0$$

рівняння (6) перетворюється в тотожність. Це означає, що пряма лінія повністю належить гіперповерхні, тобто є її *прямолінійною твірною*.

Прикладом такої гіперповерхні є *конус другого порядку* — центральна гіперповерхня, у якої всі прямолінійні твірні проходять через центр. Щоб одержати рівняння конуса, знайдемо вектор $\boldsymbol{f}_1$ з рівняння (5) і підставимо його вираз у рівність $M = 0$. Тоді одержимо:

$$(\boldsymbol{x} - \boldsymbol{f}_0)^T \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{f}_0) = 0.$$

Якщо центр гіперповерхні збігається з початком координат, то конус має рівняння

$$\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} = 0.$$

8.1.3. Дотична пряма лінія і дотична гіперплощина. Подвійна точка перетину прямої лінії і гіперповерхні називається точкою дотику прямої лінії і гіперповерхні. Далі ми позначатимемо її через $\boldsymbol{x}_0$. Пряма лінія, що дотикається до гіперповерхні, називається *дотичною прямою лінією*. На підставі формули (5) вона має векторне рівняння

$$\boldsymbol{x} = \boldsymbol{x}_0 + \alpha_1 \boldsymbol{f}_1, \tag{8}$$

причому спрямовуючий вектор $\boldsymbol{f}_1$ повинен бути таким, щоб задовольнялася рівність (7). З’ясуємо, за якої умови це відбувається.

Очевидно, що $P = 0$, тому на підставі рівності (7) повинно бути $N = 0$, або

$$\boldsymbol{a}^T \boldsymbol{f}_1 = 0, \tag{9}$$

де $\boldsymbol{a} = \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{b}$. Рівність (9) показує, що спрямовуючий вектор дотичної прямої лінії знаходиться в ортогональному доповненні $U^{\perp}$ до одновимірного підпростору U, породженого ненульовим вектором $\boldsymbol{a}$. Цим ортогональним доповненням є спрямовуючий підпростір гіперплощини, перпендикулярної до $\boldsymbol{a}$ (див. приклад з розділу 5.1).

Гіперплощина, яка проходить через деяку точку $\boldsymbol{x}_0$ гіперповерхні і яка вміщує всі дотичні прямі лінії до гіперповерхні в точці $\boldsymbol{x}_0$, називається *дотичною гіперплощиною* до гіперповерхні в точці $\boldsymbol{x}_0$. Векторне рівняння дотичної гіперплощини можна одержати з векторного рівняння дотичної прямої лінії (8), вважаючи вектор $\boldsymbol{f}_1$ лінійною комбінацією базисних векторів підпростору $U^{\perp}$. Якщо при цьому помножити зліва обидві частини рівняння (8) на $\boldsymbol{a}^T$, то з урахуванням (9) одержимо рівняння

$$\boldsymbol{a}^T \boldsymbol{x} = \boldsymbol{a}^T \boldsymbol{x}_0,$$

яке після підставлення в нього виразу для вектора $\boldsymbol{a}$ перетворюється до вигляду:

$$\boldsymbol{x}_0^T \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{x} + \boldsymbol{b}^T \boldsymbol{x}_0 + c = 0. \tag{10}$$

Звідси легко одержати рівняння дотичної гіперплощини відносно координат вектора $\boldsymbol{x}$.

8.1.4. Спрощення рівняння гіперповерхні другого порядку. При спрощенні рівняння гіперповерхні другого порядку ми будемо користуватися ортогональним перетворенням, яке зводить квадратичну форму $\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x}$ до канонічного вигляду. Важливою обставиною, яка дозволяє віддати перевагу ортогональному перетворенню перед іншими невиродженими лінійними перетвореннями, є той факт, що це перетворення, як і паралельне перенесення, зберігає відстані і кути між векторами. При цьому ортонормований базис переходить в інший ортонормований базис.

Підставимо в рівняння (1) замість $\boldsymbol{x}$ його вираз у вигляді:

$$\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{y}, \tag{11}$$

де $\boldsymbol{Q} = (\boldsymbol{q}_{*1} \quad \ldots \quad \boldsymbol{q}_{*n})$ — ортогональна матриця і $\boldsymbol{y}$ — вектор з координатами $y_1, \ldots, y_n$. Розмірковуючи, як при доведенні теореми 7.2.1, одержимо:

$$\sum_{i=1}^{n} \lambda_i y_i^2 + 2\sum_{i=1}^{n} g_i y_i + c = 0. \tag{12}$$

Тут λ_i, $i = \overline{1,\ n}$, — власні значення матриці $\boldsymbol{A}$ і $g_i = \boldsymbol{b}^T \boldsymbol{q}_{*i}$, $\overline{1,\ n}$.

Далі розглянемо окремо випадки, коли $\det \boldsymbol{A} \neq 0$ і $\det \boldsymbol{A} = 0$. Нехай $\det \boldsymbol{A} \neq 0$, тоді всі власні значення λ_i відмінні від нуля. Паралельне перенесення

$$y_i = z_i - \frac{g_i}{\lambda_i}, \quad i = \overline{1,\ n},$$

зводить рівняння (12) до вигляду:

$$\sum_{i=1}^{n} \lambda_i z_i^2 = g, \tag{13}$$

де $g = -c + \sum_{i=1}^{n} g_i^2 \lambda_i^{-1}$. Зважаючи на те, що $\det \boldsymbol{A} \neq 0$, маємо центральну гіперповерхню другого порядку. Якщо матриця $\boldsymbol{A}$ є додатно визначеною і $g > 0$, то гіперповерхня (13) називається $(n-1)$-вимірним *еліпсоїдом*. Про нього йшлося в розділі 7.2.

Припустимо, що $\det \boldsymbol{A} = 0$, тобто гіперповерхня не є центральною. Позначимо через V підпростір, породжений першими r стовпцями матриці $\boldsymbol{Q}$, відповідними до ненульових власних значень $\lambda_1, ..., \lambda_r$. Тоді підпростір $V^{\perp}$ буде породжуватися $n - r$ стовпцями матриці $\boldsymbol{Q}$, відповідними до нульового власного значення матриці $\boldsymbol{A}$. На підставі теорем 5.1.2 і 3.2.3 $\boldsymbol{b} = \boldsymbol{u} - \boldsymbol{v}$, де $\boldsymbol{u} \in V, \boldsymbol{v} \in V^{\perp}$ (вектор $\boldsymbol{v}$ узятий зі знаком «–» для зручності подальших обчислень).

Виходячи із зазначених міркувань, будемо вважати, що

$$\boldsymbol{b} = \sum_{i=1}^{r} \beta_i \boldsymbol{q}_{*i} - \boldsymbol{v}.$$

Для цього треба покласти

$$\beta_i = g_i, \quad i = \overline{1,\ r}, \quad \boldsymbol{v} = \sum_{i=1}^{r} g_i \boldsymbol{q}_{*i} - \boldsymbol{b}.$$

Якщо $\boldsymbol{v} \neq \boldsymbol{0}$, то можна вважати $\boldsymbol{v} = g_0 \boldsymbol{q}_{*n}$, де $g_0 = \pm |\, \boldsymbol{v} \,|$. При $\boldsymbol{v} = \boldsymbol{0}$ теж можна вважати $\boldsymbol{v} = g_0 \boldsymbol{q}_{*n}$, але тепер $g_0 = 0$.

Отже,

$$\boldsymbol{b} = \sum_{i=1}^{r} g_i \boldsymbol{q}_{*i} - g_0 \boldsymbol{q}_{*n}$$

і далі з урахуванянням фориули (11) знаходимо:

$$\boldsymbol{b}^T \boldsymbol{x} = \boldsymbol{b}^T \boldsymbol{Q} \boldsymbol{y} = \sum_{i=1}^{r} g_i \boldsymbol{q}_{*i}^T \boldsymbol{Q} \boldsymbol{y} - g_0 \boldsymbol{q}_{*n}^T \boldsymbol{Q} \boldsymbol{y} = \sum_{i=1}^{r} g_i y_i - g_0 y_n.$$

Тепер замість рівняння (12) будемо мати:

$$\sum_{i=1}^{r} \lambda_i y_i^2 + 2\sum_{i=1}^{r} g_i y_i - 2 g_0 y_n + c = 0.$$

Виконуючи паралельне перенесення

$$y_i = z_i - \frac{g_i}{\lambda_i}, \ \ i = \overline{1,\ r}, \quad y_i = z_i, \ \ i = \overline{r+1,\ n},$$

одержимо:

$$\sum_{i=1}^{r} \lambda_i z_i^2 - 2 g_0 z_n = g, \tag{14}$$

де $g = -c + \sum_{i=1}^{r} g_i^2 \lambda_i^{-1}$. Якщо $g_0 = 0$, то це рівняння буде таким:

$$\sum_{i=1}^{r} \lambda_i z_i^2 = g. \tag{15}$$

При $g_0 \neq 0$ зробимо додаткове паралельне перенесення

$$z_i = u_i, \quad i < n, \qquad z_i = u_i - \frac{g}{2g_0}, \quad i = n.$$

Тоді з (14) одержимо рівняння

$$\sum_{i=1}^{r} \lambda_i u_i^2 - 2g_0 u_n = 0. \tag{16}$$

Зважаючи на закон інерції квадратичних форм, спрощені рівняння гіперповерхонь (13), (15), (16) неможливо перетворити одне в одного за допомогою невиродженого лінійного перетворення змінних і паралельного перенесення. При цьому різними вважаються рівняння, які неможливо перевести одне в одного множенням на –1 і зміною нумерації координат.

8.1.5. Класифікація гіперповерхонь другого порядку. Напишемо рівняння основних класів гіперповерхонь другого порядку на підставі одержаних вище спрощених рівнянь. При цьому ми будемо записувати спрощені рівняння відносно поточних координат $x_1, \ldots, x_n$. Ці рівняння називаються *канонічними рівняннями* гіперповерхонь другого порядку.

Нехай $\det \boldsymbol{A} \neq 0$, тоді маємо канонічне рівняння

$$\sum_{i=1}^{n} \lambda_i x_i^2 = g, \tag{17}$$

яке відповідає випадку (13). Якщо $\det \boldsymbol{A} = 0$ і $r(\boldsymbol{A}) = n - 1$, то маємо канонічне рівняння

$$\sum_{i=1}^{n-1} \lambda_i x_i^2 - 2g_0 x_n = 0, \tag{18}$$

яке одержується з (16) при $r = n - 1$.

Припустимо, що $\det \boldsymbol{A} = 0$ і $g_0 = 0$. У цей клас входять гіперповерхні з канонічним рівнянням типу (15), тобто

$$\sum_{i=1}^{r} \lambda_i x_i^2 = g, \quad r = \overline{1,\ n-1}. \tag{19}$$

При $\det \boldsymbol{A} = 0$ і $g_0 \neq 0$ одержуємо клас гіперповерхонь з канонічним рівнянням вигляду (16):

$$\sum_{i=1}^{r} \lambda_i x_i^2 - 2g_0 x_n = 0, \quad r = \overline{1,\ n-2}. \tag{20}$$

Канонічними рівняннями (17) — (20) вичерпуються всі можливі класи гіперповерхонь другого порядку. Випадки (17) і (18) є основними, а випадки (19) і (20) повторюють основні випадки в підпросторі меншої вимірності. Якщо мати на увазі простір S_2, то рівняння гіперповерхні зводиться до одного з таких типів:

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = g, \tag{21}$$

$$\lambda_1 x_1^2 - 2g_0 x_2 = 0, \tag{22}$$

$$\lambda_1 x_1^2 = g, \tag{23}$$

а для простору S_3 одержуються такі типи спрощених рівнянь:

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 = g, \tag{24}$$

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 - 2g_0 x_3 = 0, \tag{25}$$

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 = g, \tag{26}$$

$$\lambda_1 x_1^2 - 2g_0 x_3 = 0, \tag{27}$$

$$\lambda_1 x_1^2 = g. \tag{28}$$

Гіперповерхні в просторі S_2 називаються *лініями другого порядку*, а в просторі S_3 — *поверхнями другого порядку*.

П р и к л а д. Виконаємо спрощення рівняння гіперповерхні

$$112x_1^2 + 112x_2^2 - 96x_1x_2 - 104x_1 + 136x_2 + 31 = 0.$$

Тут

$$\boldsymbol{A} = \begin{pmatrix} 112 & -48 \\ -48 & 112 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -52 \\ 68 \end{pmatrix}, \quad c = 31.$$

Оскільки матриця $\boldsymbol{A}$ є невиродженою, то гіперповерхня є центральною. Її центром є новий початок координат $\boldsymbol{h}$, який визначається як розв'язок системи рівнянь $\boldsymbol{Ax} = -\boldsymbol{b}$. Він дорівнює $\boldsymbol{h} = (1/4 \quad -1/2)^T$. З урахуванням формул (3) і (4) маємо:

$$(z_1 \quad z_2)\begin{pmatrix} 112 & -48 \\ -48 & 112 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + d = 0, \tag{29}$$

де

$$d = (1/4 \quad -1/2)\begin{pmatrix} 112 & -48 \\ -48 & 112 \end{pmatrix}\begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix} + 2(-52 \quad 68)\begin{pmatrix} 1/4 \\ -1/2 \end{pmatrix} + 31 = -16.$$

Після спрощення рівняння (29) одержимо:

$$7z_1^2 + 7z_2^2 - 6z_1z_2 - 1 = 0.$$

Далі треба виконати ортогональне перетворення $\boldsymbol{z} = \boldsymbol{Qy}$, як у розділі 7.2. У результаті буде одержане рівняння (7.2.6). У системі координат Ox_1x_2 розглядувана гіперповерхня має вигляд, зображений на рис. 15. △

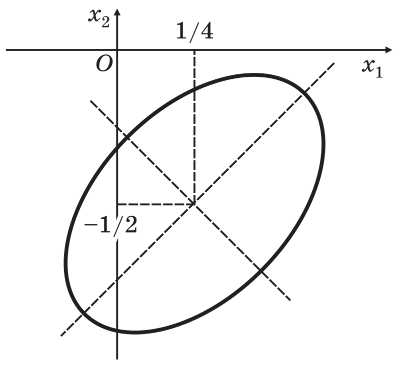

Рис. 15

8.1.6. Пара гіперплощин. Не досліджуючи геометричні властивості різних класів гіперповерхонь другого порядку (для просторів S_2 і S_3 це зроблено у двох наступних розділах), обмежимося розглядом одного простого випадку, коли ліву частину рівняння (1) можна розкласти в добуток двох дійсних лінійних множників. Геометрично це означає, що гіперповерхня (1) є *парою гіперплощин*. Для дослідження зазначеного питання введемо до розгляду симетричну матрицю

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{b}^T & c \end{pmatrix},$$

за допомогою якої рівнянню (1) можна надати вигляду:

$$\boldsymbol{r}^T \boldsymbol{B} \boldsymbol{r} = 0, \quad \boldsymbol{r} = \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix}. \tag{30}$$

Теорема 8.1.1. *Для того, щоб гіперповерхня* (1) *була парою гіперплощин, необхідно і достатньо, щоб ранг матриці* $\boldsymbol{B}$ *дорівнював* 1 *або* 2, *а індекси інерції квадратичної форми* $\boldsymbol{r}^T \boldsymbol{B} \boldsymbol{r}$ *у випадку матриці рангу* 2 *дорівнювали* 1.

Доведення. Необхідність. Нехай ліва частина рівняння (1) є добутком лівих частин рівнянь гіперповерхонь

$$\boldsymbol{c}_1^T \boldsymbol{x} + b_1 = 0, \quad \boldsymbol{c}_2^T \boldsymbol{x} + b_2 = 0.$$

Це означає, що

$$0 = 2(\boldsymbol{c}_1^T \boldsymbol{x} + b_1)(\boldsymbol{c}_2^T \boldsymbol{x} + b_2) =$$

$$= (\boldsymbol{c}_1^T \boldsymbol{x} + b_1)(\boldsymbol{c}_2^T \boldsymbol{x} + b_2) + (\boldsymbol{c}_2^T \boldsymbol{x} + b_2)(\boldsymbol{c}_1^T \boldsymbol{x} + b_1) =$$

$$= \boldsymbol{x}^T (\boldsymbol{c}_1 \boldsymbol{c}_2^T + \boldsymbol{c}_2 \boldsymbol{c}_1^T) \boldsymbol{x} + 2(b_1 \boldsymbol{c}_2^T + b_2 \boldsymbol{c}_1^T) \boldsymbol{x} + 2 b_1 b_2,$$

тобто маємо рівняння гіперповерхні другого порядку (1), для якого

$$\boldsymbol{A} = \boldsymbol{c}_1 \boldsymbol{c}_2^T + \boldsymbol{c}_2 \boldsymbol{c}_1^T, \quad \boldsymbol{b} = b_1 \boldsymbol{c}_2 + b_2 \boldsymbol{c}_1, \quad c = 2 b_1 b_2,$$

причому матриця $\boldsymbol{A}$ є симетричною.

Матрицю $\boldsymbol{B}$, рівну

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{c}_1 \boldsymbol{c}_2^T + \boldsymbol{c}_2 \boldsymbol{c}_1^T & b_1 \boldsymbol{c}_2 + b_2 \boldsymbol{c}_1 \\ b_1 \boldsymbol{c}_2^T + b_2 \boldsymbol{c}_1^T & 2 b_1 b_2 \end{pmatrix},$$

можна подати у вигляді:

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T + \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T.$$

Кожен з доданків правої частини на підставі теореми 3.4.6 є матрицею рангу 1. Тому з нерівності (3.4.12) випливає, що $r(\boldsymbol{B}) \le 2$, тобто $r(\boldsymbol{B})$ дорівнює 1 або 2.

Покажемо, що при $r(\boldsymbol{B}) = 2$ індекси інерції квадратичної форми $\boldsymbol{r}^T\boldsymbol{B}\boldsymbol{r}$ дорівнюють 1. Це означає, що серед власних значень матриці $\boldsymbol{B}$ повинно бути тільки два ненульових і що вони повинні мати протилежні знаки.

Кількість ненульових власних значень матриці $\boldsymbol{B}$ дорівнює її рангу, тому при $r(\boldsymbol{B}) = 2$ маємо тільки два ненульові власні значення ν_1, ν_2; відповідні власні вектори позначимо через $\boldsymbol{x}_1, \boldsymbol{x}_2$. Напишемо рівність (6.1.2) для власних значень $\nu_{1,2}$ і власних векторів $\boldsymbol{x}_{1,2}$:

$$\left(\nu_{1,2}\boldsymbol{E} - \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T - \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T\right)\boldsymbol{x}_{1,2} = \boldsymbol{0}. \tag{31}$$

Помноживши зліва обидві частини (31) почергово на $(\boldsymbol{c}_1^T \quad b_1)$ і на $(\boldsymbol{c}_2^T \quad b_2)$, одержимо відповідно:

$$\begin{aligned} &\nu_{1,2}\begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T \boldsymbol{x}_{1,2} - \left|\begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}\right|^2 \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T \boldsymbol{x}_{1,2} - \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T \boldsymbol{x}_{1,2} = 0, \\ &\nu_{1,2}\begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T \boldsymbol{x}_{1,2} - \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}\begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T \boldsymbol{x}_{1,2} - \left|\begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}\right|^2 \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T \boldsymbol{x}_{1,2} = 0. \end{aligned} \tag{32}$$

Маємо систему лінійних однорідних рівнянь відносно величин $(\boldsymbol{c}_1^T \quad b_1)\boldsymbol{x}_{1,2}$ і $(\boldsymbol{c}_2^T \quad b_2)\boldsymbol{x}_{1,2}$. Ці величини одночасно не обертаються в нуль, бо в протилежному випадку з формули (31) випливали б неможливі рівності $\boldsymbol{x}_{1,2} = \boldsymbol{0}$. Умова існування ненульового розв’язку системи (32) має вигляд:

$$\begin{vmatrix} \nu_{1,2} - \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix} & -\left\| \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix} \right\|^2 \\ -\left\| \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix} \right\|^2 & \nu_{1,2} - \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix} \end{vmatrix} = 0,$$

звідки випливає вираз для власних значень ν_1, ν_2:

$$\nu_{1,2} = \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix}^T \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix} \pm \left\| \begin{pmatrix} \boldsymbol{c}_1 \\ b_1 \end{pmatrix} \right\| \left\| \begin{pmatrix} \boldsymbol{c}_2 \\ b_2 \end{pmatrix} \right\|. \tag{33}$$

З відмінності від нуля власних знвчень ν_1, ν_2 і нерівності Коші — Шварца — Буняковського випливає, що перший доданок правої частини формули (33) є меншим за модулем ніж другий доданок. Тому ν_1 і ν_2 мають протилежні знаки, тобто індекси інерції дорівнюють 1.

Д о с т а т н і с т ь. Оскільки вектор $\boldsymbol{r}$ не є нульовим, то з (30) випливає, що квадратична форма $\boldsymbol{r}^T\boldsymbol{B}\boldsymbol{r}$ не є додатно визначеною. Тому вона є або невід’ємно (недодатно) визначеною, або знакозмінною. З урахуванням умов теореми існують два варіанти спектрального розкладу матриці $\boldsymbol{B}$:

$$\boldsymbol{B} = \mu_1 \boldsymbol{y}_1 \boldsymbol{y}_1^T, \quad \boldsymbol{B} = \nu_1 \boldsymbol{z}_1 \boldsymbol{z}_1^T + \nu_2 \boldsymbol{z}_2 \boldsymbol{z}_2^T. \tag{34}$$

Тут $\mu_1, \boldsymbol{y}_1$ — єдине відмінне від нуля власне значення і відповідний до нього власний вектор, ν_1, $\boldsymbol{z}_1$ і ν_2, $\boldsymbol{z}_2$ — єдині відмінні від нуля власні значення і відповідні власні вектори, причому ν_1 і ν_2 мають протилежні знаки. Нагадаємо, що власні вектори $\boldsymbol{y}_1$ і $\boldsymbol{z}_1$, $\boldsymbol{z}_2$ належать відповідним ортонормованим системам векторів, існування яких гарантується теоремою 6.5.7.

Після підставлення виразів (34) для матриці $\boldsymbol{B}$ в (30) одержимо відповідно:

$$\mu_1 (\boldsymbol{r}^T \boldsymbol{y}_1)^2 = 0, \qquad \nu_1 (\boldsymbol{r}^T \boldsymbol{z}_1)^2 + \nu_2 (\boldsymbol{r}^T \boldsymbol{z}_2)^2 = 0. \tag{35}$$

Величини $\boldsymbol{r}^T \boldsymbol{y}_1, \boldsymbol{r}^T \boldsymbol{z}_1, \boldsymbol{r}^T \boldsymbol{z}_2$ лінійно виражаються через координати вектора $\boldsymbol{x}$. Тому перше рівняння (35) є рівнянням гіперплощин, які збігаються, а друге рівняння (35) з урахуванням відмінності знаків величин ν_1 і ν_2 визначає пару різних гіперплощин. □

8.2. Лінії другого порядку

У багатьох математичних дисциплінах доводиться мати справу з лініями і поверхнями другого порядку. Розгляду їх властивостей присвячені цей і наступний розділи.

8.2.1. Властивості основних типів ліній другого порядку. Будемо вивчати лінії другого порядку, виходячи з рівнянь (8.1.21) — (8.1.23). Припустимо, що в рівнянні (8.1.21) $g \neq 0$ і числа λ_1, λ_2 мають такі ж знаки, що і g. Тоді рівняння лінії можна подати у вигляді:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1, \tag{1}$$

де $a_1 = \sqrt{g/\lambda_1}$, $a_2 = \sqrt{g/\lambda_2}$. Ця лінія називається *еліпсом* (рис. 16), а рівняння (1) — *канонічним рівнянням еліпса*.

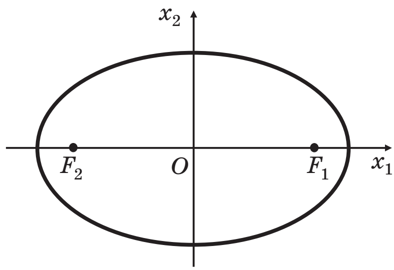

Рис. 16

Еліпс — обмежена лінія, оскільки для всіх його точок, як випливає з рівняння (1), виконуються нерівності $|x_1| \leq a_1$, $|x_2| \leq a_2$. Якщо точка $(x_1 \quad x_2)^T$ належить еліпсу, то йому належать також точки $(x_1 \quad -x_2)^T, (-x_1 \quad x_2)^T, (-x_1 \quad -x_2)^T$. Тому еліпс має дві осі симетрії — вісь Ox_1 і вісь Ox_2. Еліпс є центральною лінією другого порядку, центром еліпса є початок координат. Точки $(\pm a_1 \quad 0)^T, (0 \quad \pm a_2)^T$ називаються *вершинами* еліпса.

Звичайно вважають, що $a_1 > a_2$. Тоді величини a_1 і a_2 називаються *великою і малою напівосями* еліпса. При $a_1 = a_2$ еліпс стає *колом* з центром у початку координат і радіусом a_1.

Введемо позначення

$$c^2 = a_1^2 - a_2^2. \tag{2}$$

Точки $(\pm c \quad 0)^T$ називаються *фокусами* еліпса (точки F_1 і F_2 на рис. 16).

Теорема 8.2.1. *Сума відстаней між будь-якою точкою еліпса і його фокусами є постійною величиною, рівною* $2a_1$.

Д о в е д е н н я. Нехай ρ_1, ρ_2 — відстані між точкою $(x_1 \quad x_2)^T$ еліпса і фокусами $(c \quad 0)^T$ і $(-c \quad 0)^T$ відповідно. Очевидно, що

$$\rho_1 = \sqrt{(x_1 - c)^2 + x_2^2}.$$

За допомогою канонічного рівняння еліпса і співвідношення (2) знаходимо, що

$$\rho_1 = \sqrt{(x_1 - c)^2 + a_2^2 - \frac{a_2^2 x_1^2}{a_1^2}} = |-\varepsilon x_1 + a_1| = -\varepsilon x_1 + a_1. \tag{3}$$

Тут $\varepsilon = c/a_1$ і враховано, що $-\varepsilon x_1 + a_1 > 0$, оскільки $|x_1| \le a_1$ і $\varepsilon < 1$.

Аналогічно обчислюється відстань ρ_2:

$$\rho_2 = \sqrt{(x_1 + c)^2 + x_2^2} = \varepsilon x_1 + a_1. \tag{4}$$

Тепер з рівностей (3) і (4) випливає, що

$$\rho_1 + \rho_2 = 2a_1. \square$$

Нехай у рівнянні (8.1.21) $g \ne 0$ і числа λ_1, λ_2 мають знаки, протилежні до знаку величини g. Тоді рівняння (8.1.21) можна подати у вигляді:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = -1. \tag{5}$$

Очевидно, що координати жодної точки не задовольняють цьому рівнянню. Говорять, що (5) — рівняння *уявного* еліпса.

Якщо $g = 0$ і числа λ_1, λ_2 мають однакові знаки, то рівняння (8.1.21) зводиться до вигляду:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 0, \tag{6}$$

де $a_1 = \sqrt{|\lambda_1|^{-1}}, a_2 = \sqrt{|\lambda_2|^{-1}}$. Зрозуміло, що тільки точка $(0 \ \ 0)^T$ задовольняє рівнянню (6). Його називають рівнянням *виродженого* еліпса.

Припустимо, що $g > 0$ і числа λ_1, λ_2 мають протилежні знаки. Не порушуючи загальності, можна вважати, що $\lambda_1 > 0$, $\lambda_2 < 0$. Тоді замість еліпса будемо мати *гіперболу*. Її *канонічне рівняння* є таким:

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1, \tag{7}$$

де a_1 і a_2 обчислюються, як для еліпса.

Гіпербола (рис. 17) — необмежена лінія, вона має осі симетрії — вісь Ox_1 і вісь Ox_2. Гіпербола є центральною лінією другого порядку, центр гіперболи збігається з початком координат. Вісь Ox_1 перетинається з гіперболою в точках $(\pm a_1 \ \ 0)^T$, які називаються *вершинами* гіперболи. Інша вісь не має спільних точок з гіперболою. Величини a_1 і a_2 називаються *дійсною* і *уявною напівосями* гіперболи.

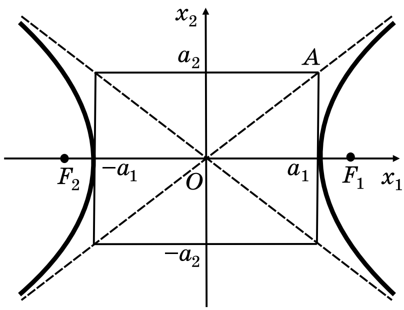

Рис. 17

Введемо позначення

$$c^2 = a_1^2 + a_2^2. \tag{8}$$

Точки $(\pm c \ \ 0)^T$ називаються *фокусами* гіперболи (точки F_1 і F_2 на рис. 17).

Теорема 8.2.2. *Модуль різниці відстаней між деякою точкою гіперболи і її фокусами є постійною величиною, рівною* $2a_1$.

Д о в е д е н н я. Позначимо через ρ_1 і ρ_2 відстанні між точкою $(x_1 \ \ x_2)^T$ гіперболи та фокусами $(c \ \ 0)^T$ і $(-c \ \ 0)^T$ відповідно.

З урахуванням канонічного рівняння гіперболи і рівності (8) нескладно показати, що

$$\rho_1 = |\varepsilon x_1 - a_1|, \qquad \rho_2 = |\varepsilon x_1 + a_1|, \tag{9}$$

де, як і для еліпса, $\varepsilon = c/a_1$.

Для всіх точок гіперболи $|x_1| \geq a_1, \varepsilon > 1.$ Тому

$$\begin{aligned} &\rho_1 = \varepsilon x_1 - a_1, \quad x_1 > 0, \qquad \rho_1 = -\varepsilon x_1 + a_1, \quad x_1 < 0, \\ &\rho_2 = \varepsilon x_1 + a_1, \quad x_1 > 0, \qquad \rho_2 = -\varepsilon x_1 - a_1, \quad x_1 < 0. \end{aligned} \tag{10}$$

Звідси випливає, що

$$|\rho_1 - \rho_2| = 2a_1. \; \square$$

Розглянемо ту частину гіперболи, для якої $x_1, x_2 \geq 0.$ Представимо рівняння (7) у вигляді:

$$x_2 = \frac{a_2}{a_1} x_1 - \frac{a_1 a_2}{x_1 + \sqrt{x_1^2 - a_1^2}}$$

і розглянемо пряму лінію

$$x_2' = \frac{a_2}{a_1} x_1.$$

За умови необмеженого зростання x_1 різниця

$$x_2' - x_2 = \frac{a_1 a_2}{x_1 + \sqrt{x_1^2 - a_1^2}}$$

прямує до нуля, залишаючись додатною. Це означає, що точки прямої лінії і гіперболи, які мають однакові значення координати x_1, зближуються, причому точки гіперболи залишаються нижче відповідних точок прямої лінії.

Аналогічну властивість мають інші частини гіперболи стосовно однієї з прямих ліній

$$x_2 = \frac{a_2}{a_1} x_1, \qquad x_2 = -\frac{a_2}{a_1} x_1,$$

які називаються *асимптотами* гіперболи (пунктирні прямі лінії на рис. 17).

Завершуючи розгляд рівняння (8.1.21), будемо вважати, що $g = 0$ і числа λ_1, λ_2 мають протилежні знаки. Тоді це рівняння можна перетворити до вигляду:

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 0, \tag{11}$$

де $a_1 = \sqrt{|\lambda_1|^{-1}}$, $a_2 = \sqrt{|\lambda_2|^{-1}}$. Очевидно, що рівняння (11) визначає *пару прямих ліній, що перетинаються*.

Звернемося до рівняння (8.1.22) і будемо замість нього розглядати аналогічне рівняння

$$\lambda_2 x_2^2 - 2g_0 x_1 = 0,$$

яке звичайно використовують на практиці. При $\lambda_2, g_0 \neq 0$ воно визначає лінію другого порядку, яка називається *параболою*. Якщо позначити g_0/λ_2 через p,то будемо мати *канонічне рівняння* параболи

$$x_2^2 = 2px_1. \tag{12}$$

На рис. 18 зображена парабола для випадку $p > 0$.

Парабола є необмеженою лінією з однією віссю симетрії — віссю Ox_1. Парабола не має центру. Точка перетину параболи з віссю симетрії називається *вершиною* параболи, а точка $(p/2 \quad 0)^T$ називається *фокусом* параболи (точка F на рис. 18).

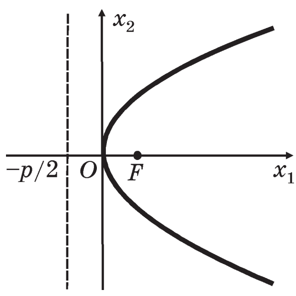

Рис. 18

Пряма лінія

$$x_1 = -\frac{p}{2}$$

називається *директрисою* параболи (пунктирна пряма лінія на рис. 18)

Теорема 8.2.3. *Відстань між деякою точкою параболи і директрисою дорівнює відстані між цією точкою і фокусом параболи.*

Д о в е д е н н я. Відстань між точкою $(x_1 \quad x_2)^T$ параболи і директрисою дорівнює

$$\rho_1 = x_1 + \frac{p}{2}, \tag{13}$$

а відстань між цією точкою і фокусом обчислюється так:

$$\rho_2 = \sqrt{(x_1 - \frac{p}{2})^2 + x_2^2} = \sqrt{(x_1 - \frac{p}{2})^2 + 2px_1} = x_1 + \frac{p}{2}, \tag{14}$$

Тут ураховано, що $x_1 \geq 0$ і $p > 0$. З формул (13) і (14) випливає рівність $\rho_1 = \rho_2$. □

Нарешті, розглянемо рівняння (8.1.23). Нехай числа λ_1 і g мають однакові знаки. При цьому рівняння (8.1.23) набуває вигляду

$$x_1^2 - a^2 = 0,$$

де $a^2 = g/\lambda_1$. Це — *пара паралельних прямих ліній.*

Якщо знаки чисел λ_1 і g є протилежними, то рівняння (8.1.23) буде таким:

$$x_1^2 + a^2 = 0,$$

де $a^2 = -g/\lambda_1$. Цьому рівнянню не задовольняють координати жодної точки. Про нього говорять як про рівняння *пари уявних прямих ліній*.

При $g = 0$ рівняння (8.1.23) рівносильне рівнянню

$$x_1^2 = 0,$$

яке визначає *пару прямих ліній, що збігаються.*

8.2.2. Директриси еліпса і гіперболи. Формули (3), (4) і (9) показують, що для всіх точок еліпса і гіперболи виконуються рівності

$$\rho_1 = \varepsilon\left|x_1 - \frac{a_1^2}{c}\right|, \qquad \rho_2 = \varepsilon\left|x_1 + \frac{a_1^2}{c}\right|. \tag{15}$$

Прямі лінії, що визначаються рівняннями

$$x_1 - \frac{a_1^2}{c} = 0, \qquad x_1 + \frac{a_1^2}{c} = 0, \tag{16}$$

називаються *директрисами* еліпса і гіперболи (пунктирні прямі лінії на рис. 19, 20). Будемо вважати, що директриса і фокус мають однакові номери, якщо вони знаходяться в одній напівплощині $x_1 > 0$ або $x_1 < 0$.

Відстані σ_1 і σ_2 між довільною точкою $(x_1 \ \ x_2)^T$ еліпса або гіперболи і директрисами (16) дорівнюють відповідно

$$\sigma_1 = \left| x_1 - \frac{a_1^2}{c} \right|, \qquad \sigma_2 = \left| x_1 + \frac{a_1^2}{c} \right|. \tag{17}$$

З формул (15) і (17) випливає, що

$$\frac{\rho_1}{\sigma_1} = \varepsilon, \qquad \frac{\rho_2}{\sigma_2} = \varepsilon.$$

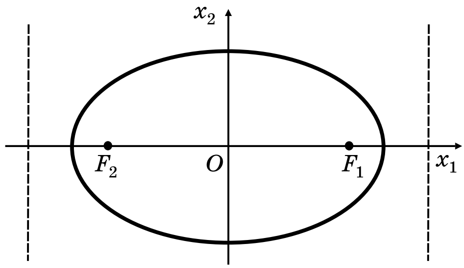

Рис. 19

Отже, відношення відстаней між будь-якою точкою еліпса або гіперболи і відповідними фокусом та директрисою є постійною величиною ε, яка називається *ексцентриситетом*. Аналогічне твердження для параболи випливає з теореми 8.2.3, якщо вважати, що для неї $\varepsilon = 1$. Нагадаємо, що для еліпса $\varepsilon < 1$, а для гіперболи $\varepsilon > 1$.

8.2.3. Дотичні прямі лінії до еліпса, гіперболи, параболи. В афінному просторі S_2 гіперплощиною є пряма лінія. Тому дотичну пряму лінію до еліпса, гіперболи і параболи можна знайти за допомогою рівняння дотичної гіперплощини (8.1.10).

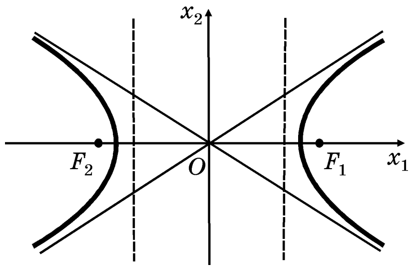

Рис. 20

Для еліпса (1) маємо:

$$\boldsymbol{A} = \begin{pmatrix} 1/a_1^2 & 0 \\ 0 & 1/a_2^2 \end{pmatrix}, \quad \boldsymbol{b} = \boldsymbol{0}, \quad c = -1 \tag{18}$$

і рівняння дотичної прямої лінії є таким:

$$\frac{x_1 x_{01}}{a_1^2} + \frac{x_2 x_{02}}{a_2^2} = 1,$$

де x_{01}, x_{02} — координати точки $\boldsymbol{x}_0$.

Аналогічно знаходимо рівняння дотичної прямої лінії для гіперболи (7):

$$\frac{x_1x_{01}}{a_1^2} - \frac{x_2x_{02}}{a_2^2} = 1.$$

У випадку параболи (12)

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} -p \\ 0 \end{pmatrix}, \quad c = 0. \tag{19}$$

Тому рівняння дотичної прямої лінії є таким:

$$x_2x_{02} = p(x_1 + x_{01}).$$

8.2.4. Оптичні властивості еліпса, гіперболи, параболи. Покажемо, що еліпс, гіпербола і парабола мають такі оптичні властивості:

1) промінь світла, що виходить з фокуса еліпса, після дзеркального відображення від еліпса проходить через інший фокус;

2)промінь світла, що виходить з фокуса гіперболи, після дзеркального відображення від гіперболи видається таким, що виходить з іншого фокуса;

3)промінь світла, що виходить з фокуса параболи, після дзеркального відображення від параболи проходить паралельно до осі симетрії параболи.

На рис. 21—23 зображені дотичні прямі лінії до еліпса, гіперболи і параболи, що проходять через точку *M* цих ліній. Оптичні властивості означають, що кути φ_1 і φ_2 повині бути рівними.

Розглянемо спочатку еліпс (рис.21). Позначимо через $\boldsymbol{c}$ вектор $(c \quad 0)^T$. Урахувавши, що $\overrightarrow{OM} = \boldsymbol{x}_0$, будемо мати:

$$\overrightarrow{MF_1} = -\boldsymbol{x}_0 + \boldsymbol{c}, \quad \overrightarrow{MF_2} = -\boldsymbol{x}_0 - \boldsymbol{c}.$$

Нехай кут φ_1 між вектором $\overrightarrow{MF_1}$ і спрямовуючим вектором $\boldsymbol{f}_1$ дотичної прямої лінії є гострим. Кут φ_2 теж є гострим, якщо вважати, що він є кутом між векторами $-\boldsymbol{f}_1$ і $\overrightarrow{MF_2}$. Ми зможемо довести рівність кутів φ_1 і φ_2, якщо покажемо, що $\cos\varphi_1 = \cos\varphi_2$. Обчислимо ці величини за допомогою формули (5.2.19):

$$\cos\varphi_1 = \frac{(-\boldsymbol{x}_0 + \boldsymbol{c})^T \boldsymbol{f}_1}{|-\boldsymbol{x}_0 + \boldsymbol{c}||\boldsymbol{f}_1|}, \qquad \cos\varphi_2 = \frac{-\boldsymbol{f}_1^T(-\boldsymbol{x}_0 - \boldsymbol{c})}{|-\boldsymbol{f}_1||-\boldsymbol{x}_0 - \boldsymbol{c}|}.$$

Вектор $\boldsymbol{f}_1$ задовольняє умові (8.1.9), яка з урахуванням (18) має вигляд:

$$\boldsymbol{x}_0^T \begin{pmatrix} 1/a_1^2 & 0 \\ 0 & 1/a_2^2 \end{pmatrix} \boldsymbol{f}_1 = 0.$$

Легко переконатися в тому, що ця рівність виконується при

$$\boldsymbol{f}_1 = (a_1^2 x_{02} \quad -a_2^2 x_{01})^T.$$

Далі з урахуванням рівності (2) маємо:

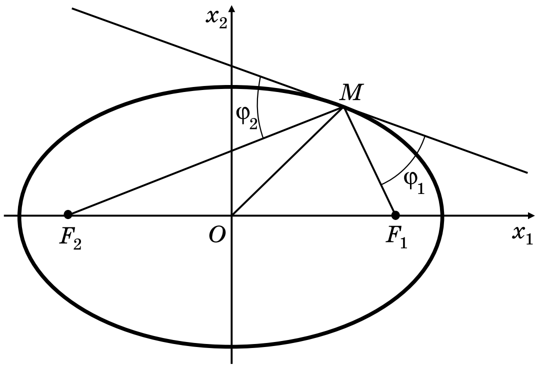

Рис. 21

$$(-\boldsymbol{x}_0 + \boldsymbol{c})^T \boldsymbol{f}_1 = (-x_{01} + c \quad -x_{02}) \begin{pmatrix} a_1^2\, x_{02} \\ -a_2^2\, x_{01} \end{pmatrix} = c a_1 x_{02} \rho_{10},$$

де ρ_{10} — величина ρ_1, обчислена для точки дотику за формулою (3). Тепер, зважаючи на те, що $|-\boldsymbol{x}_0 + \boldsymbol{c}| = \rho_{10}$, знаходимо:

$$\cos\varphi_1 = \frac{c a_1 x_{02}}{|\boldsymbol{f}_1|}.$$

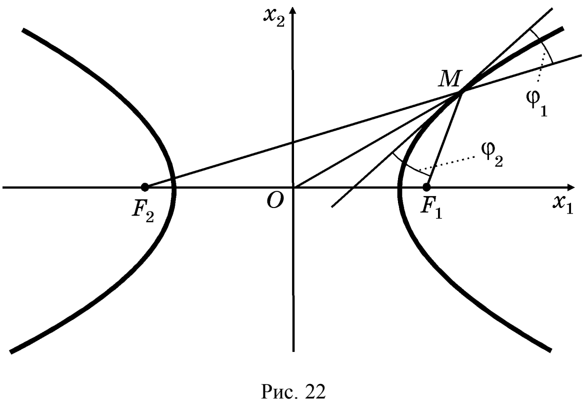

Рис. 22

Такою ж є і величина $\cos\varphi_2$, в чому неважко пересвідчитися за допомогою аналогічних обчислень з урахуванням формули (4). Отже, оптична властивість еліпса доведена.

Аналогічно виявляється оптична властивість гіперболи (7). На рис. 22 кут φ_1 між векторами $-\overrightarrow{MF_2}$ і $\boldsymbol{f}_1$ є гострим; таким же є і кут φ_2 між векторами $-\boldsymbol{f}_1$ і $\overrightarrow{MF_1}$. На підставі формули (5.2.19) маємо:

$$\cos\varphi_1 = \frac{-(-\boldsymbol{x}_0 - \boldsymbol{c})^T \boldsymbol{f}_1}{|-(-\boldsymbol{x}_0 - \boldsymbol{c})||\boldsymbol{f}_1|}, \quad \cos\varphi_2 = \frac{-\boldsymbol{f}_1^T(-\boldsymbol{x}_0 + \boldsymbol{c})}{|-\boldsymbol{f}_1||-\boldsymbol{x}_0 + \boldsymbol{c}|}.$$

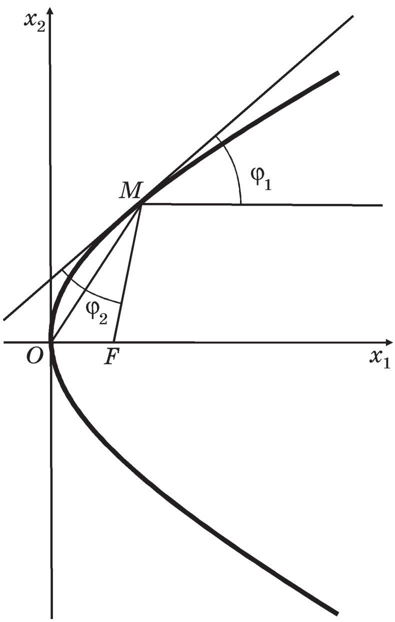

Рис. 23

Подальші нескладні обчислення показують, що $\varphi_1 = \varphi_2$. При цьому вектор $\boldsymbol{f}_1$ можна вважати рівним

$$\boldsymbol{f}_1 = (a_1^2 x_{02} \quad a_2^2 x_{01})^T$$

і треба скористатися формулами (8) і (10).

Для параболи (12) з (8.1.9) і (19) випливає така умова для вектора $\boldsymbol{f}_1$:

$$\left(\boldsymbol{x}_0^T \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + (-p \quad 0) \right) \boldsymbol{f}_1 = 0.$$

Вона виконується при

$$\boldsymbol{f}_1 = (x_{02} \quad p)^T.$$

Позначимо через $\boldsymbol{d}$ вектор $(p/2 \quad 0)^T$. Тоді у відповідності до рис. 23 на підставі формули (5.2.19) маємо:

$$\cos\varphi_1 = \frac{\boldsymbol{d}^T \boldsymbol{f}_1}{|\boldsymbol{d}||\boldsymbol{f}_1|}, \quad \cos\varphi_2 = \frac{-\boldsymbol{f}_1^T(-\boldsymbol{x}_0 + \boldsymbol{d})}{|-\boldsymbol{f}_1||-\boldsymbol{x}_0 + \boldsymbol{d}|}.$$

Легко пересвідчитися в тому, що

$$\cos\varphi_1 = \frac{x_{02}}{|\boldsymbol{f}_1|}.$$

Крім того,

$$-\boldsymbol{f}_1^T(-\boldsymbol{x}_0 + \boldsymbol{d}) = -(x_{02} \quad p)\begin{pmatrix} -x_{01} + p/2 \\ -x_{02} \end{pmatrix} = x_{02}\rho_{20},$$

де ρ_{20} — величина ρ_2, обчислена для точки дотику за формулою (14). Урахувавши, що $|-\boldsymbol{x}_0 + \boldsymbol{d}| = \rho_{20}$, знаходимо:

$$\cos\varphi_2 = \frac{x_{02}}{|\boldsymbol{f}_1|}.$$

Тепер з рівності $\cos\varphi_1 = \cos\varphi_2$ випливає, що $\varphi_1 = \varphi_2$, тобто оптична властивість параболи доведена.

8.3. Поверхні другого порядку

8.3.1. Основні типи поверхонь другого порядку. Нехай у рівнянні (8.1.24) числа $\lambda_1, \lambda_2, \lambda_3, g$ відмінні від нуля і мають однакові знаки. Стандартною заміною коефіцієнтів рівняння (8.1.24) зводиться до вигляду:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1. \qquad (1)$$

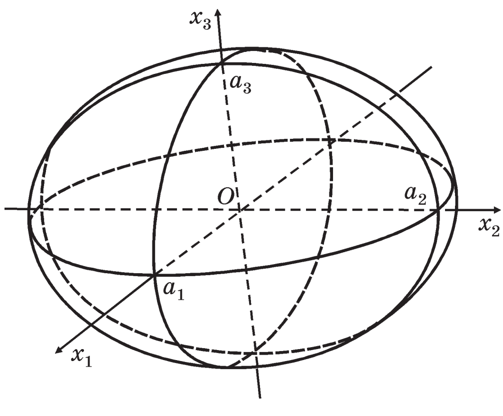

Рис. 24

Рівняння (1) визначає поверхню другого порядку, яка називається *еліпсоїдом* (рис. 24), а саме рівняння (1) називається *канонічним рівнянням* еліпсоїда. Координатні площини $x_1 = 0$, $x_2 = 0$ і $x_3 = 0$ є площинами симетрії, а початок координат — центром еліпсоїда. Числа a_1, a_2, a_3 називаються *напівосями* еліпсоїда. Еліпсоїд — обмежена поверхня, оскільки $|x_1| \le a_1, |x_2| \le a_2, |x_3| \le a_3$. Перерізом еліпсоїда і площини, паралельної до будь-якої з координатних площин, є еліпс, зокрема, вироджений або уявний.

Якщо в рівнянні (8.1.24) знаки чисел $\lambda_1, \lambda_2, \lambda_3$ є протилежними до знаку g, то замість ріняння (1) матимемо рівняння *уявного* еліпсоїда

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\frac{x_3^2}{a_3^2}=-1,$$

якому не задовольняють координати жодної точки простору S_3.

При $g=0$ і однакових знаках коефіцієнтів λ_1, λ_2, λ_3 одержується рівняння

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\frac{x_3^2}{a_3^2}=0,$$

якому задовольняє тільки початок координат. Про нього говорять як про рівняння *виродженого* еліпсоїда.

Припустимо, що знаки чисел λ_1, λ_2, g в рівнянні (8.1.24) є однаковими і протилежними до знаку λ_3. Стандартна заміна коефіцієнтів зводить це рівняння до такого вигляду:

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}-\frac{x_3^2}{a_3^2}=1. \qquad (2)$$

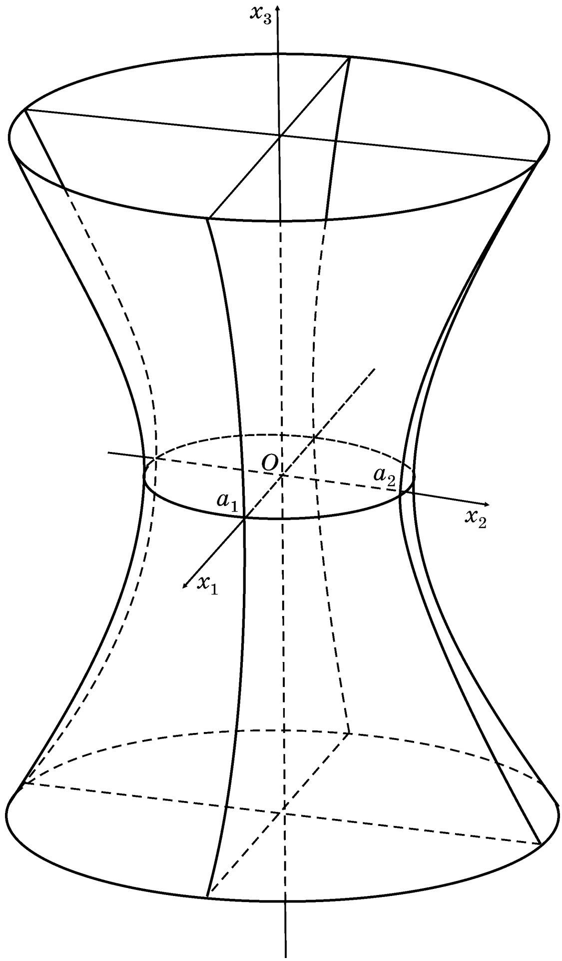

Рис. 25

Розглядувана поверхня другого порядку називається *однополим гіперболоїдом* (рис. 25). Рівняння (2) — його *канонічне рівняння*. Очевидно, що координатні площини є площинами симетрії, а початок координат — центром гіперболоїда. Перерізом однополого гіперболоїда і площини $x_3 = h$ є еліпс

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1 + \frac{h^2}{a_3^2},$$

напівосі якого необмежено збільшуються при $h \to \infty$. Неважко побачити, що перерізами однополого гіперболоїда і площин, паралельних до координатних площин $x_1 = 0$ і $x_2 = 0$, є гіперболи.

Отже, однополий гіперболоїд (2) є поверхнею, яка складається з однієї трубчастої порожнини, що необмежено розширюється в додатному і від'ємному напрямах осі Ox_3.

Нехай знаки чисел λ_1 і λ_2 є однаковими і протилежними до знаків λ_3 і g. Тоді рівняння (8. 1. 24) можна звести до вигляду:

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = -1. \qquad (3)$$

Поверхня з таким рівнянням називається *двополим гіперболоїдом* (рис. 26), а рівняння (3) — його *канонічним рівнянням*. Координатні площини є площинами симетрії, а початок координат — центром гіперболоїда. Перерізом двополого гіперболоїда і площини $x_3 = h$ є еліпс

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = -1 + \frac{h^2}{a_3^2}.$$

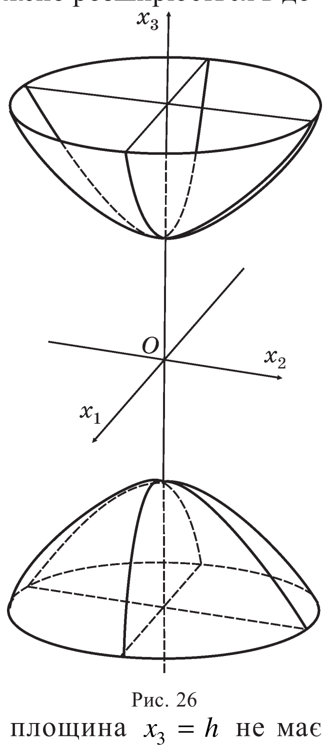

Рис. 26

При $|h| < a_3$ цей еліпс є уявним. Тому площина $x_3 = h$ не має спільних точок з двополим гіперболоїдом при $|h| < a_3$ Це означає, що поверхня складається з двох порожнин, розміщених зовні шару $|h| < a_3$. Перерізом двополого гіперболоїда і площин, паралельних до координатних площин $x_1 = 0$ і $x_2 = 0$, є гіперболи.

Припустимо, що $g = 0$ і знаки коефіцієнтів λ_1, λ_2 є протилежними до знаку λ_3. Маємо:

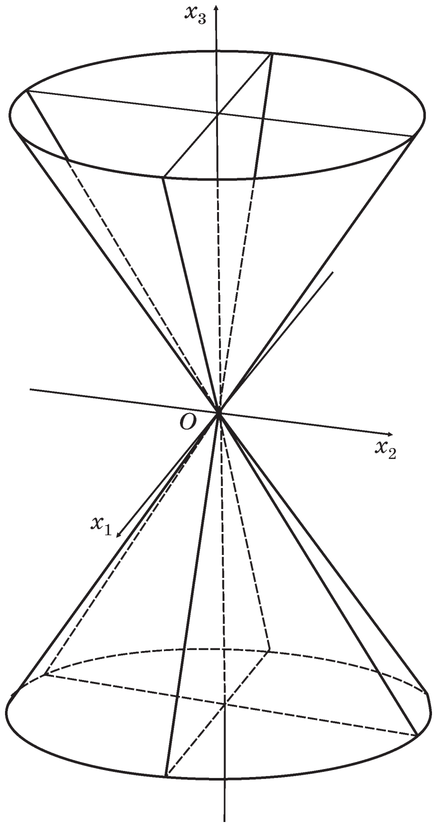

Рис. 27

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 0. \tag{4}$$

Поверхня з таким рівнянням називається *еліптичним конусом* (рис. 27), а саме рівняння (4) — його *канонічним рівнянням*. Координатні площини є площинами симетрії, а початок координат — центром конуса. Перерізом еліптичного конуса і площини, паралельної до координатної площини $x_3 = 0$, є еліпс, а перерізами еліптичного конуса і площин, паралельних до координатних площин $x_1 = 0$ і $x_2 = 0$, є гіперболи.

Звернемося до розгляду рівняння (8.1.25). Нехай числа λ_1, λ_2 і g_0 мають однакові знаки. Стандартна заміна коефіцієнтів дає рівняння

$$2x_3 = \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2}. \tag{5}$$

Поверхня, що визначається рівнянням (5), називається *еліптичним параболоїдом* (рис. 28), а рівняння (5) — його *канонічним рівнянням*. Для цієї поверхні координатні площини $x_1 = 0$ і $x_2 = 0$ є площинами симетрії. Поверхня не є центральною, вона розміщенна в напівпросторі $x_3 \geq 0$. Перерізом еліптичного параболоїда і площини $x_3 = h$ є еліпс

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}=2h,$$

напівосі якого необмежено збільшуються при $h \to +\infty$, тобто еліптичний параболоїд є нескінченною чашею. Перерізами еліптичного параболоїда і площин, паралельних до координатних площин $x_1 = 0$ і $x_2 = 0$, є параболи.

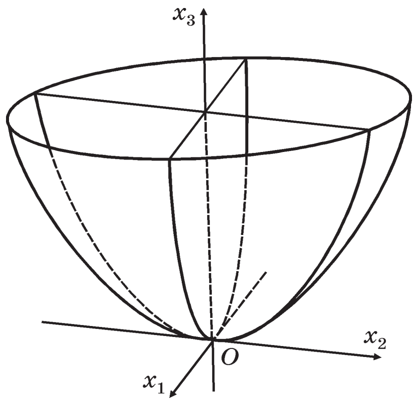

Рис. 28

Припустимо, що в рівнянні (8.1.25) числа λ_1 і λ_2 мають різні знаки, а знак величини g_0 є таким же, як у λ_1. Тоді його можна звести до рівняння

$$2x_3=\frac{x_1^2}{a_1^2}-\frac{x_2^2}{a_2^2}. \qquad (6)$$

Це рівняння визначає поверхню, яка називається *гіперболічним параболоїдом* (рис. 29), а саме рівняння (6) називається його *канонічним рівнянням*. Координатні площини $x_1 = 0$ і $x_2 = 0$ є площинами симетрії. Поверхня не є центральною. Перерізом гіперболічного параболоїда і площини, паралельної до координатної площини $x_3 = 0$, є гіпербола, а перерізами гіперболічного параболоїда і площин, паралельних до координатних площин $x_1 = 0$ і $x_2 = 0$, є параболи.

Розглянемо, нарешті, рівняння (8.1.26) — (8.1.28). Усі ці рівняння не залежать від однієї або двох координат. Для їх дослідження треба скористатися результатами з розділу 8.2. Поверхні, що визначаються цими рівняннями, називаються *циліндрами* (еліптичний, гіперболічний і т. д.). Очевидно, що вони вміщують прямолінійні твірні, паралельні до тієї координатної осі, яка відповідає відсутній координаті в рівнянні поверхні.

8.3.2. Прямолінійні твірні поверхні другого порядку. Однополий гіперболоїд і гіперболічний параболоїд мають властивість, пов'язану з наявністю прямолінійних твірних.

Теорема 8.3.1. *Через кожну точку однополого гіперболоїда і гіперболічного параболоїда проходять дві різні прямолінійні твірні цих поверхонь.*

Д о в е д е н н я. Напишемо рівняння однополого гіперболоїда (2) у вигляді:

$$\left(\frac{x_1}{a_1}+\frac{x_3}{a_3}\right)\left(\frac{x_1}{a_1}-\frac{x_3}{a_3}\right)=\left(1+\frac{x_2}{a_2}\right)\left(1-\frac{x_2}{a_2}\right). \qquad (7)$$

При будь-яких α і β, не рівних нулю одночасно, пара площин

$$\alpha\left(\frac{x_1}{a_1}+\frac{x_3}{a_3}\right)-\beta\left(1-\frac{x_2}{a_2}\right)=0,\quad \alpha\left(1+\frac{x_2}{a_2}\right)-\beta\left(\frac{x_1}{a_1}-\frac{x_3}{a_3}\right)=0 \qquad (8)$$

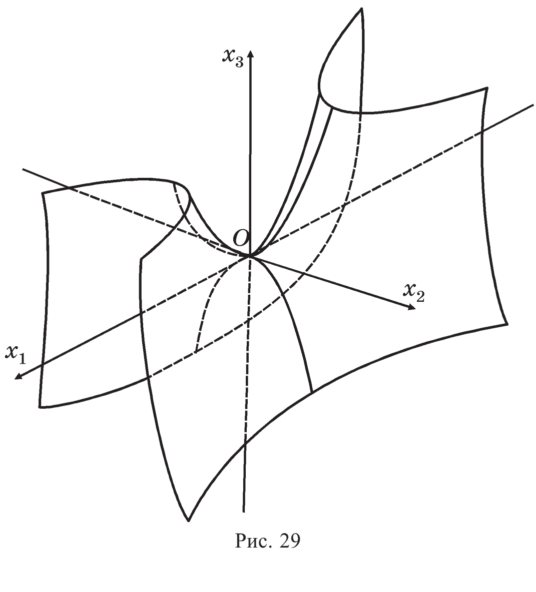

Рис. 29

визначає деяку пряму лінію, яка цілком лежить на поверхні (7). У цьому легко переконатися, виключивши α і β з рівнянь (8). Покажемо, що через кожну точку поверхні проходить одна пряма лінія з сім'ї (8)

Система рівнянь (8) відносно α, β має нульовий визначник тоді і тільки тоді, коли точка $(x_1\ x_2\ x_3)^T$ належить гіперболоїду (7). При цьому матриця системи (8) є матрицею рангу 1. Тому вектор $(\alpha\ \beta)^T$ визначається з точністю до числового множника. Звідси випливає єдиність прямої лінії (8), що проходить через кожну точку поверхні.

Аналогічно можна показати, що через кожну точку однополого гіперболоїда проходить єдина пряма лінія, яка визначається парою площин

$$\gamma\left(\frac{x_1}{a_1}+\frac{x_3}{a_3}\right)-\delta\left(1+\frac{x_2}{a_2}\right)=0,\quad \gamma\left(1-\frac{x_2}{a_2}\right)-\delta\left(\frac{x_1}{a_1}-\frac{x_3}{a_3}\right). \tag{9}$$

Очевидно, що прямі лінії (8) і (9) є різними.

Стосовно гіперболічного параболоїда (6) можна навести аналогічні міркування і довести існування двох різних сімей прямолінійних твірних, які визначаються парами площин

$$2\alpha x_3-\beta\left(\frac{x_1}{a_1}+\frac{x_2}{a_2}\right)=0,\quad \alpha\left(\frac{x_1}{a_1}-\frac{x_2}{a_2}\right)-\beta=0$$

і

$$2\gamma x_3-\delta\left(\frac{x_1}{a_1}-\frac{x_2}{a_2}\right)=0,\quad \gamma\left(\frac{x_1}{a_1}+\frac{x_2}{a_2}\right)-\delta=0.\square$$

Вправи до глави 8

8.1. Нехай $\boldsymbol{A}$ — додатно визначена матриця. Довести, що на прямій лінії (8.1.5) функція $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}+2\boldsymbol{b}^T\boldsymbol{x}$ досягає мінімального значення при

$$\alpha_1=-\frac{(\boldsymbol{A}\boldsymbol{f}_0+\boldsymbol{b})^T\boldsymbol{f}_1}{\boldsymbol{f}_1^T\boldsymbol{A}\boldsymbol{f}_1}.$$

8.2. Довести, що система рівнянь $\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}=\boldsymbol{0}$ не має розв'язків тоді і тільки тоді, коли гіперповерхня (8.1.1) зводиться до канонічного вигляду (8.1.18) або (8.1.20).

8.3. Нехай пряма лінія (8.1.5) перетинає гіперповерхню (8.1.1) у двох точках. Відрізок цієї прямої лінії, який визначається точками перетину з гіперповерхнею, називається хордою. Показати, що середина хорди визначається значенням параметра α_1 з вправи 8.1.

8.4. Показати, що коли $\boldsymbol{x}_0$ — середина хорди, то

$$\boldsymbol{x}_0 = \boldsymbol{f}_0 - \frac{\boldsymbol{f}_1^T(\boldsymbol{A}\boldsymbol{f}_0 + \boldsymbol{b})}{\boldsymbol{f}_1^T \boldsymbol{A}\boldsymbol{f}_1}\boldsymbol{f}_1.$$

Довести, що множина середин хорд належить гіперплощині

$$\boldsymbol{f}_1^T(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) = 0,$$

яка називається діаметральною гіперплощиною, спряженою до напряму $\boldsymbol{f}_1$ відносно гіперповерхні (8.1.1).

8.5. Показати, що фокуси гіперболи F_1 і F_2 (рис. 17) є точками перетину осі Ox_1 і кола радіуса $|\overrightarrow{OA}|$ з центром у початку координат.

8.6. Гіпербола з рівними напівосями називається рівносторонньою. Скласти канонічне рівняння рівносторонньої гіперболи і обчислити її ексцентриситет.

8.7. Показати, що рівняння $x_1 x_2 + c = 0$ визначає рівносторонню гіперболу, асимптотами якої є осі координат Ox_1 і Ox_2.

8.8. Що являє собою діаметральна гіперплощина, спряжена до даного напряму, для лінії другого порядку.

8.9. Написати рівняння дотичної площини для різних поверхонь другого порядку.

8.10. Що являє собою діаметральна гіперплощина, спряжена до даного напряму, для поверхні другого порядку.

Г Л А В А 9

ЖОРДАНОВА ФОРМА МАТРИЦІ

У цій главі розглядається питання про вибір спеціального перетворення подібності, при якому задана матриця зводиться до квазідіагонального вигляду, який називається жордановою формою матриці.

9.1. Жорданова форма

9.1.1. Основні визначення. Вектор $\boldsymbol{x}$ називається приєднаним вектором лінійного перетворення (матриці) $\boldsymbol{A}$, відповідним до власного значення λ, якщо для деякого цілого $m \geq 1$ виконуються співвідношення

$$(\boldsymbol{A} - \lambda\boldsymbol{E})^m \boldsymbol{x} \neq \boldsymbol{0}, \quad (\boldsymbol{A} - \lambda\boldsymbol{E})^{m+1} \boldsymbol{x} = \boldsymbol{0}.$$

Число m називається порядком приєднаного вектора $\boldsymbol{x}$. Зрозуміло, що коли $\boldsymbol{x}$ — приєднаний вектор порядку m, то вектор $(\boldsymbol{A} - \lambda\boldsymbol{E})^m \boldsymbol{x}$ є власним вектором матриці $\boldsymbol{A}$.

Далі будуть розглядатися співвідношення

$$\begin{array}{l}
\boldsymbol{A}\boldsymbol{h}_{11} = \lambda_1 \boldsymbol{h}_{11}, \quad \boldsymbol{A}\boldsymbol{h}_{1m} = \lambda_1 \boldsymbol{h}_{1m} + \boldsymbol{h}_{1,m-1}, \quad m = \overline{2,\ n_1}, \\
\boldsymbol{A}\boldsymbol{h}_{21} = \lambda_2 \boldsymbol{h}_{21}, \quad \boldsymbol{A}\boldsymbol{h}_{2m} = \lambda_2 \boldsymbol{h}_{2m} + \boldsymbol{h}_{2,m-1}, \quad m = \overline{2,\ n_2}, \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\boldsymbol{A}\boldsymbol{h}_{l1} = \lambda_l \boldsymbol{h}_{l1}, \quad \boldsymbol{A}\boldsymbol{h}_{lm} = \lambda_l \boldsymbol{h}_{lm} + \boldsymbol{h}_{l,m-1}, \quad m = \overline{2,\ n_l},
\end{array} \tag{1}$$

відносно сукупності векторів

$$\begin{array}{l} \boldsymbol{h}_{11},\ \ \boldsymbol{h}_{12},\ \ \ldots,\ \ \boldsymbol{h}_{1n_1}; \\ \boldsymbol{h}_{21},\ \ \boldsymbol{h}_{22},\ \ \ldots,\ \ \boldsymbol{h}_{2n_2}; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \boldsymbol{h}_{l1},\ \ \boldsymbol{h}_{l2},\ \ \ldots,\ \ \boldsymbol{h}_{ln_l}. \end{array} \tag{2}$$

Тут вважається, що

$$n_1 + n_2 + \ldots + n_l = n.$$

Очевидно, що вектор $\boldsymbol{h}_{11}$ є власним вектором матриці $\boldsymbol{A}$, відповідним до власного значення λ_1. Покажемо, що інші вектори першого рядка з (2) є приєднаними векторами, відповідними до власного значення λ_1. Візьмемо, наприклад, вектор $\boldsymbol{h}_{12}$ і помножимо його зліва на матрицю $\boldsymbol{A} - \lambda_1\boldsymbol{E}$. На підставі другої рівності з першої групи формул (1) будемо мати:

$$(\boldsymbol{A} - \lambda_1\boldsymbol{E})\boldsymbol{h}_{12} = \boldsymbol{h}_{11}.$$

Після подальшого множення на $\boldsymbol{A} - \lambda_1\boldsymbol{E}$ знаходимо:

$$(\boldsymbol{A} - \lambda_1\boldsymbol{E})^2\boldsymbol{h}_{12} = (\boldsymbol{A} - \lambda_1\boldsymbol{E})\boldsymbol{h}_{11} = \boldsymbol{0}.$$

Отже, $\boldsymbol{h}_{12}$ є приєднаним вектором першого порядку, відповідним до власного значення λ_1. Аналогічно можна показати, що $\boldsymbol{h}_{13}, \ldots, \boldsymbol{h}_{1n_1}$ є приєднаними векторами порядку $2, \ldots, n_1 - 1$ відповідно.

Таким чином, перший рядок у (2) складається з власного вектора і приєднаних векторів. Ця сукупність векторів називається *жордановим ланцюжком довжини* n_1, відповідним до власного значення λ_1. Інші жорданові ланцюжки у (2) мають таку ж будову: вони починаються з власного вектора, а далі йдуть приєднані вектори, розміщені в такій послідовності, що їх порядки зростають. Якщо деякому власному значенню λ відповідає s лінійно незалежних власних векторів, то будемо мати s жорданових ланцюжків, відповідних до λ. Інакше кажучи, власні значення $\lambda_1, \ldots, \lambda_l$ у співвідношеннях (1) не обов'язково є різними.

Нижче ми побачимо, що вектори (2) утворюють у n-вимірному просторі базис, який називається *канонічним базисом* відносно лінійного перетворення (матриці) $\boldsymbol{A}$. Рівності (1) показують, що базисні вектори кожного з жорданових ланцюжків переходять при лінійному перетворенні $\boldsymbol{A}$ в лінійну комбінацію векторів того ж ланцюжка. Це означає, що кожний жорданів ланцюжок векторів породжує підпростір, інваріантний відносно лінійного перетворення. Канонічний базис є об'єднанням базисів цих інваріантних підпросторів.

Жордановою клітиною $\boldsymbol{J}_k(\lambda)$ називається дводіагональна матриця порядку k, у якої головна діагональ складається з однакових елементів, рівних λ, а діагональ, прилегла до головної, складається з одиниць. Інакше кажучи,

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

Вважається, що $\boldsymbol{J}_1(\lambda) = (\lambda)$.

Жордановою формою матриці n-го порядку називається квазідіагональна матриця

$$\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_{n_1}(\lambda_1),\ \boldsymbol{J}_{n_2}(\lambda_2),\ \dots,\ \boldsymbol{J}_{n_l}(\lambda_l)), \quad \sum_{k=1}^{l} n_k = n. \tag{3}$$

Зазначимо, що коли кожна жорданова клітина $\boldsymbol{J}_{n_k}(\lambda_k)$ має порядок 1, тобто $n_k = 1$ для $k = \overline{1,\, l}$ і $l = n$, жорданова форма $\boldsymbol{J}$ є діагональною матрицею.

9.1.2. Теорема про зведення матриці до жорданової форми. При доведенні наступної теореми ми маємо на увазі, що довільне лінійне перетворення $\boldsymbol{A}$ діє в комплексному векторному просторі.

Теорема 9.1.1. *Нехай* $\boldsymbol{A}$ *— матриця* n*-го порядку. Існує базис* (*2*), *складений з власних і приєднаних векторів матриці* $\boldsymbol{A}$ *і такий, що дія лінійного перетворення* $\boldsymbol{A}$ *описується співвідношеннями* (*1*).

Перетворенням подібності з матрицею $\boldsymbol{T}$*, стовпцями якої є вектори* (*2*)*, матрицю* $\boldsymbol{A}$ *можна звести до жорданової форми*:

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{J}, \tag{4}$$

де $\boldsymbol{J}$ *визначається формулою* (*3*).

Д о в е д е н н я. У разі існування канонічного базису (2) перетворення подібності з матрицею $\boldsymbol{T}$ завдяки виконанню умов (1) зводить матрицю $\boldsymbol{A}$ до жорданової форми. У розділі 6.2 ми пересвідчилися в цьому для матриці другого порядку. Неважко перевірити це для матриці довільного порядку.

Отже, рівність (4) є справедливою при існуванні канонічного базису. Доведення його існування виконаємо за допомогою методу математичної індукції. Нехай λ — деяке власне значення матриці $\boldsymbol{A}$. Матриця $\boldsymbol{B} = \boldsymbol{A} - \lambda\boldsymbol{E}$ має ранг r, менший ніж n. Як показує рівність (4.3.8), лінійне перетворення $\boldsymbol{B}$ відображає векторний простір у підпростір $\mathscr{R}(\boldsymbol{B})$ вимірності r. Оскільки $\mathscr{R}(\boldsymbol{B})$ є інваріантним підпростором, то лінійне перетворення $\boldsymbol{B}$ відображає $\mathscr{R}(\boldsymbol{B})$ в себе. У розділі 6.1 ми бачили, що одновимірний інваріантний підпростір породжується власним вектором лінійного перетворення, тобто при $r = 1$ існує базис, про який ідеться в теоремі.

Припустимо, що теорема справедлива для підпростору $\mathscr{R}(\boldsymbol{B})$ вимірності $r < n$. Тоді в $\mathscr{R}(\boldsymbol{B})$ повинен існувати базис

$$\begin{array}{l} \boldsymbol{h}_{11},\ \boldsymbol{h}_{12},\ \ldots,\ \boldsymbol{h}_{1r_1};\\ \boldsymbol{h}_{21},\ \boldsymbol{h}_{22},\ \ldots,\ \boldsymbol{h}_{2r_2};\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \boldsymbol{h}_{p1},\ \boldsymbol{h}_{p2},\ \ldots,\ \boldsymbol{h}_{pr_p}, \end{array} \tag{5}$$

кількість векторів якого дорівнює

$$r_1 + r_2 + \ldots + r_p = r,$$

причому базис є таким, що дія лінійного перетворення $\boldsymbol{B}$ в підпросторі $\mathscr{R}(\boldsymbol{B})$ описується співвідношеннями

$$
\begin{array}{lll}
\boldsymbol{Bh}_{11} = \mu_1 \boldsymbol{h}_{11}, & \boldsymbol{Bh}_{1m} = \mu_1 \boldsymbol{h}_{1m} + \boldsymbol{h}_{1,\,m-1}, & m = \overline{2,\ r_1}, \\
\boldsymbol{Bh}_{21} = \mu_2 \boldsymbol{h}_{21}, & \boldsymbol{Bh}_{2m} = \mu_2 \boldsymbol{h}_{2m} + \boldsymbol{h}_{2,\,m-1}, & m = \overline{2,\ r_2}, \\
\dots\dots\dots & \dots\dots\dots\dots\dots\dots\dots\dots & \dots\dots\dots \\
\boldsymbol{Bh}_{p1} = \mu_p \boldsymbol{h}_{p1}, & \boldsymbol{Bh}_{pm} = \mu_p \boldsymbol{h}_{pm} + \boldsymbol{h}_{p,\,m-1}, & m = \overline{2,\ r_p}.
\end{array} \tag{6}
$$

Покажемо, як можна доповнити базис (5) до базису всього простору і домогтися виконання умов, подібних до (1).

Припустимо, що $\dim(\mathscr{N}(\boldsymbol{B}) \bigcap \mathscr{R}(\boldsymbol{B})) = q$. Зрозуміло, що кожний вектор з $\mathscr{N}(\boldsymbol{B})$ є власним вектором, відповідним до нульового власного значення матриці $\boldsymbol{B}$. Тому серед векторів (5) повинно бути q жорданових ланцюжків, відповідних до цього власного значення. Будемо вважати, що в (5) такими є перщі q ланцюжків.

Розглянемо q векторів, що знаходяться наприкінці цих ланцюжків (не виключено, що кінцевий вектор ланцюжка може збігатися з початковим). Оскільки кожен з цих векторів належить до $\mathscr{R}(\boldsymbol{B})$, то

$$
\boldsymbol{h}_{kr_k} = \boldsymbol{B}\boldsymbol{f}_k, \quad k = \overline{1,\ q}, \tag{7}
$$

для деяких векторів $\boldsymbol{f}_k$.

Підпростір $\mathscr{N}(\boldsymbol{B})$ має вимірність $n - r$. Крім свого q-вимірного перерізу з $\mathscr{R}(\boldsymbol{B})$ він повинен вміщувати $n - r - q$ базисних векторів $\boldsymbol{g}_j$, що знаходяться поза цим перерізом. Зрозуміло, що

$$
\boldsymbol{B}\boldsymbol{g}_j = \boldsymbol{0}, \quad j = \overline{1,\ n - r - q}. \tag{8}
$$

Доведемо, що вектори (5) разом з векторами

$$
\boldsymbol{f}_k, \quad k = \overline{1,\ q}, \tag{9}
$$

$$
\boldsymbol{g}_j, \quad j = \overline{1,\ n - r - q}, \tag{10}
$$

є лінійно незалежними. Для цього розглянемо рівну нульовому вектору лінійну комбінацію цих векторів:

$$
\sum_{k=1}^{p} \sum_{m=1}^{r_k} \alpha_{km} \boldsymbol{h}_{km} + \sum_{k=1}^{q} \beta_k \boldsymbol{f}_k + \sum_{j=1}^{n-r-q} \gamma_j \boldsymbol{g}_j = \boldsymbol{0}. \tag{11}
$$

Помножимо зліва цю рівність на $\boldsymbol{B}$. Скориставшись формулами (6) — (8), одержимо:

$$\sum_{k=1}^{p}\alpha_{k1}\mu_k\boldsymbol{h}_{k1}+\sum_{k=1}^{p}\sum_{m=2}^{r_k}\alpha_{km}(\mu_k\boldsymbol{h}_{km}+\boldsymbol{h}_{k,m-1})+\sum_{k=1}^{q}\beta_k\boldsymbol{h}_{kr_k}=\mathbf{0}. \qquad (12)$$

Рівність (12) являє собою рівну нульовому вектору лінійну комбінацію базисних векторів (5), тому коефіцієнти при цих векторах дорівнюють нулю. Деякі з базисних векторів (5) входять до різних сум указаної лінійної комбінації, тому зрівнювати з нулем коефіцієнти при цих векторах можна після зведення подібних членів. Це зауваження не стосується векторів, що входять до останньої суми лівої частини (12). Дійсно, в перших двох сумах коефіцієнти при векторах $\boldsymbol{h}_{kr_k}$, $k=\overline{1,\ q}$, є нульовими, оскільки $\mu_k=0$, $k=\overline{1,\ q}$, тому залишаються лише коефіцієнти β_k при векторах (9) і вони дорівнюють нулю.

Тепер з рівності (11) випливає, що

$$\sum_{k=1}^{p}\sum_{m=1}^{r_k}\alpha_{km}\boldsymbol{h}_{km}=-\sum_{j=1}^{n-r-q}\gamma_j\boldsymbol{g}_j. \qquad (13)$$

Ліва частина рівності (13) є лінійною комбінацією векторів підпростору $\mathscr{R}(\boldsymbol{B})$, а права частина належить підпростору, натягнутому на вектори (10), що знаходяться поза підпростором $\mathscr{R}(\boldsymbol{B})$. Єдиним спільним вектором указаних підпросторів є нульовий вектор. Тепер з лінійної незалежності векторів (5) і (10) випливає, що рівність (13) можлива лише коли всі коефіцієнти α_{km} і γ_j є нульовими.

Отже, рівність (11) можлива лише при нульових значеннях усіх коефіцієнтів, тобто вектори (5), (9), (10) є лінійно незалежними. Загальна їх кількість дорівнює n, тому вони утворюють базис простору.

З'ясуємо, як треба розмістити ці вектори, щоб вони відповідали умовам, подібним до співвідношень (1).

Вище ми бачили, що вектор $\boldsymbol{h}_{kr_k}$ є приєднаним вектором порядку r_k-1, тобто $\boldsymbol{B}^{r_k}\boldsymbol{h}_{kr_k}=\mathbf{0}$. Оскільки на підставі рівності (7)

$$\mathbf{0} = \boldsymbol{B}^{r_k}\boldsymbol{h}_{kr_k} = \boldsymbol{B}^{r_k+1}\boldsymbol{f}_k\,,$$

то $\boldsymbol{f}_k$ є приєднаним вектором порядку r_k. Тому цей вектор повинен розміститися у відповідному жордановому ланцюжку відразу після породжуючого вектора $\boldsymbol{h}_{kr_k}$.

Вектори (10), які на підставі формули (8) є власними векторами матриці $\boldsymbol{B}$, відповідними до нульового власного значення, можна розглядати як жорданові ланцюжки, що вміщують по одному вектору. Тому їх можна розмістити після перших q жорданових ланцюжків.

Таким чином, базисні вектори (5), (9), (10) можна записати у вигляді:

$$\begin{array}{l}
\boldsymbol{h}_{11},\quad \boldsymbol{h}_{12},\quad \ldots,\quad \boldsymbol{h}_{1r_1},\quad \boldsymbol{f}_1;\\
\boldsymbol{h}_{21},\quad \boldsymbol{h}_{22},\quad \ldots,\quad \boldsymbol{h}_{2r_2},\quad \boldsymbol{f}_2;\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\boldsymbol{h}_{q1},\quad \boldsymbol{h}_{q2},\quad \ldots,\quad \boldsymbol{h}_{qr_q},\quad \boldsymbol{f}_q;\\
\boldsymbol{g}_1;\quad \boldsymbol{g}_2;\quad \ldots;\quad \boldsymbol{g}_{n-r-q};\\
\boldsymbol{h}_{q+1,1},\quad \boldsymbol{h}_{q+1,2},\quad \ldots,\quad \boldsymbol{h}_{q+1,r_{q+1}};\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\boldsymbol{h}_{p1},\quad \boldsymbol{h}_{p2},\quad \ldots,\quad \boldsymbol{h}_{p,r_p}.
\end{array}$$

При розміщенні базисних векторів у такій послідовності дія лінійного перетворення $\boldsymbol{B}$ буде відбуватися за правилом (1).

Теорема доведена для матриці $\boldsymbol{B}$. Покажемо, що вона є справедливою і для матриці $\boldsymbol{A}$. Дійсно,

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{-1}(\boldsymbol{B}+\lambda\boldsymbol{E})\boldsymbol{T} = \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}+\lambda\boldsymbol{E}.$$

Перший доданок правої частини є жордановою формою для матриці $\boldsymbol{B}$, а $\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}+\lambda\boldsymbol{E}$ є жордановою формою для $\boldsymbol{A}$. Вона отримується за допомогою тих же жорданових ланцюжків, що і для матриці $\boldsymbol{B}$. ☐

Тепер можна розглянути умову подібності двох матриць.

Теорема 9.1.2. *Для подібності двох матриць необхідно і достатньо, щоб їх жорданові форми збігалися з точністю до розміщення жорданових клітин на головній діагоналі.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Нехай матриці $\boldsymbol{A}$ та $\boldsymbol{A}'$ є подібними і $\boldsymbol{J}$ — жорданова форма матриці $\boldsymbol{A}$. На підставі формул (4) і (4.3.7) маємо:

$$\boldsymbol{J} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}^{-1}\boldsymbol{Q}\boldsymbol{A}'\boldsymbol{Q}^{-1}\boldsymbol{T} = (\boldsymbol{Q}^{-1}\boldsymbol{T})^{-1}\boldsymbol{A}'\boldsymbol{Q}^{-1}\boldsymbol{T},$$

тобто $\boldsymbol{J}$ є жордановою формою для $\boldsymbol{A}'$.

Д о с т а т н і с т ь. Нехай

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1}, \quad \boldsymbol{A}' = \boldsymbol{T}'\boldsymbol{J}'(\boldsymbol{T}')^{-1}, \tag{14}$$

де жорданові форми $\boldsymbol{J}$ і $\boldsymbol{J}'$ складаються з одних і тих же жорданових клітин, розміщених у різному порядку. Ці матриці є подібними і перетворюючою матрицею є деяка матриця переставлень $\boldsymbol{P}$, бо

$$\boldsymbol{J}' = \boldsymbol{P}^T\boldsymbol{J}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{J}\boldsymbol{P}. \tag{15}$$

З рівностей (14) і (15) знаходимо, що

$$\boldsymbol{A}' = \boldsymbol{T}'\boldsymbol{P}^{-1}\boldsymbol{J}\boldsymbol{P}(\boldsymbol{T}')^{-1} = \boldsymbol{T}'\boldsymbol{P}^{-1}\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\boldsymbol{P}(\boldsymbol{T}')^{-1} = (\boldsymbol{T}\boldsymbol{P}(\boldsymbol{T}')^{-1})^{-1}\boldsymbol{A}\boldsymbol{T}\boldsymbol{P}(\boldsymbol{T}')^{-1}.$$

Порівняння останньої формули з рівністю (4.3.7) показує, що матриці $\boldsymbol{A}$ і $\boldsymbol{A}'$ є подібними. ☐

9.2. Властивості жорданової форми

9.2.1. Знаходження перетворюючої матриці. Доведення теореми 9.1.1 свідчить про те, що коли власні значення матриці $\boldsymbol{A}$ є дійсними, то перетворюючу матрицю $\boldsymbol{T}$ можна побудувати так, щоб вона була дійсною.

Жорданова форма матриці є єдиною з точністю до розташування жорданових клітин на діагоналі матриці $\boldsymbol{J}$. Щоб усунути цю неоднозначність, звичайно вибирається певний порядок власних значень $\lambda_1, \ldots, \lambda_n$ матриці $\boldsymbol{A}$ і тоді в матриці $\boldsymbol{J}$ спочатку йдуть жорданові клітини, що відповідають λ_1, потім ті клітини, що відповідають λ_2, і т.д. Для кожного власного значення відповідні жорданові клітини розміщуються так, щоб їх порядки не зростали. Якщо деякому власному значенню відповідають декілька жорданових клітин однакового порядку, то ці клітини нічим не відрізняються.

Доведення теореми показує, що кількість l жорданових клітин (з урахуванням повторів одних і тих же клітин) дорівнює загальній кількості лінійно незалежних власних векторів матриці. Матрицю можна діагоналізувати тоді, коли $l = n$. Кількість жорданових клітин з одним і тим же власним значенням дорівнює кількості відповідних до нього власних векторів, а сума порядків цих клітин збігається з кратністю власного значення.

Для побудови жорданової форми вказана інформація про власні значення і власні вектори є, взагалі кажучи, недостатньою. Покажемо це на прикладі. Нехай матриця восьмого порядку має власне значення λ кратності 8 і чотири власні вектори. Її жорданова форма вміщує чотири жорданові клітини. Неважко зрозуміти, що існує п'ять варіантів представлення матриці $\boldsymbol{J}$:

$$\boldsymbol{J} = \mathrm{diag}\,(\boldsymbol{J}_5(\lambda),\ \boldsymbol{J}_1(\lambda),\ \boldsymbol{J}_1(\lambda),\ \boldsymbol{J}_1(\lambda)),$$

$$\boldsymbol{J} = \mathrm{diag}\,(\boldsymbol{J}_4(\lambda),\ \boldsymbol{J}_2(\lambda),\ \boldsymbol{J}_1(\lambda),\ \boldsymbol{J}_1(\lambda)),$$

$$\boldsymbol{J} = \mathrm{diag}\,(\boldsymbol{J}_3(\lambda),\ \boldsymbol{J}_3(\lambda),\ \boldsymbol{J}_1(\lambda),\ \boldsymbol{J}_1(\lambda)),$$

$$\boldsymbol{J} = \mathrm{diag}\,(\boldsymbol{J}_3(\lambda),\ \boldsymbol{J}_2(\lambda),\ \boldsymbol{J}_2(\lambda),\ \boldsymbol{J}_1(\lambda)),$$

$$\boldsymbol{J} = \mathrm{diag}\,(\boldsymbol{J}_2(\lambda),\ \boldsymbol{J}_2(\lambda),\ \boldsymbol{J}_2(\lambda),\ \boldsymbol{J}_2(\lambda)).$$

Виявляється, що порядки жорданових клітин пов'язані з рангами деяких матриць. Щоб переконатися в цьому, розглянемо четвертий варіант матриці $\boldsymbol{J}$ з наведеного переліку.

У цьому разі матриця $\boldsymbol{J} - \lambda\boldsymbol{E}$ дорівнює

$$\boldsymbol{J} - \lambda\boldsymbol{E} = \mathrm{diag}\left(\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, (0)\right). \tag{1}$$

Обчислимо $(\boldsymbol{J} - \lambda\boldsymbol{E})^2$ і $(\boldsymbol{J} - \lambda\boldsymbol{E})^3$:

$$\begin{gathered} (\boldsymbol{J} - \lambda\boldsymbol{E})^2 = \mathrm{diag}\left(\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, (0)\right), \\ (\boldsymbol{J} - \lambda\boldsymbol{E})^3 = \mathbf{0}. \end{gathered} \tag{2}$$

Треба звернути увагу на те, як змінюються жорданові клітини $\boldsymbol{J}_k(0)$ при обчисленні їх степенів: діагональ, що складається з одиниць, зміщується паралельно до себе в напрямі правого верхнього кута і на певному кроці зникає, після чого жорданова клітина $\boldsymbol{J}_k(0)$ перетворюється в нульову підматрицю. Неважко зрозуміти, що $\boldsymbol{J}_k^s(0) = \boldsymbol{0}$ при $s \geq k$.

Рівності (1) і (2) показують, що

$$r((\boldsymbol{J} - \lambda \boldsymbol{E})^3) = 0, \quad r((\boldsymbol{J} - \lambda \boldsymbol{E})^2) = 1, \quad r(\boldsymbol{J} - \lambda \boldsymbol{E}) = 4. \tag{3}$$

Одержані результати дозволяють виявити залежності між величинами (3) і порядками жорданових клітин . Перша рівність (3) указує на те, що найбільша жорданова клітина має порядок 3. Порівняння другої рівності (3) і першої рівності (2) свідчить про те, що жорданова клітина третього порядку є єдиною. З останньої рівності (3) і формули (1) випливає, що $r(\boldsymbol{J} - \lambda \boldsymbol{E})$ дорівнює подвоєній кількості клітин третього порядку плюс кількість клітин другого порядку. У нашому випадку кількість клітин другого порядку дорівнює $4 - 2 \cdot 1 = 2$. Тепер кількість клітин першого порядку дорівнює $8 - 3 \cdot 1 - 2 \cdot 2 = 1$.

Аналогічні обчислення можна виконати у випадку жорданових клітин будь-яких порядків, якщо всі клітини відповідають одному власному значенню. Найбільша жорданова клітина має порядок n_1, рівний найменшому цілому числу такому, що $(\boldsymbol{J} - \lambda \boldsymbol{E})^{n_1} = \boldsymbol{0}$. Рангом матриці $(\boldsymbol{J} - \lambda \boldsymbol{E})^{n_1 - 1}$ є кількість клітин порядку n_1, ранг матриці $(\boldsymbol{J} - \lambda \boldsymbol{E})^{n_1 - 2}$ дорівнює подвоєній кількості клітин порядку n_1 плюс кількість клітин порядку $n_1 - 1$ і т.д. Розгляд послідовності величин $r((\boldsymbol{J} - \lambda \boldsymbol{E})^{n_1 - i})$, $i = \overline{0,\ n_1 - 1}$, дозволяє обчислити порядки всіх клітин у матриці $\boldsymbol{J}$.

Нехай $\boldsymbol{J}$ відповідає різним власним значенням. Якщо λ_1 — одне з них, то, утворюючи матриці $\boldsymbol{J} - \lambda_1 \boldsymbol{E}$, $(\boldsymbol{J} - \lambda_1 \boldsymbol{E})^2$, ..., ми зможемо перетворити в нульові підматриці лише ті клітини, які відповідають λ_1, оскільки інші клітини не обернуться в нульові підматриці через наявність ненульової головної діагоналі. Врешті решт ранги матриць $(\boldsymbol{J} - \lambda_1 \boldsymbol{E})^s$ перестануть спадати при зростанні s (звичайно,

$s \le n$). Найменше значення s, що мінімізує $r((\boldsymbol{J} - \lambda_1 \boldsymbol{E})^s)$, дорівнює порядку найбільшої жорданової клітини, відповідної до λ_1. Кількість жорданових клітин, відповідних до власного значення λ_1, і їх порядки обчислюються на підставі дослідження рангів послідовності степенів матриці $\boldsymbol{J} - \lambda_1 \boldsymbol{E}$. Аналогічно аналізуються інші власні значення.

Міркування, викладені вище стосовно матриці $\boldsymbol{J}$, залишаються в силі для будь-якої матриці $\boldsymbol{A}$, оскільки множення на невироджені матриці $\boldsymbol{T}$ і $\boldsymbol{T}^{-1}$, яке відбувається при виконанні перетворення подібності, не змінює ранги матриці і її степенів.

Отже, описана вище процедура дозволяє виявити для матриці $\boldsymbol{A}$ порядки і кількість жорданових клітин, відповідних до будь-якого власного значення. Особливості побудови жорданових ланцюжків з'ясуємо, розглянувши

П р и к л а д 1. Розглянемо матрицю п'ятого порядку

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

Її характеристичне рівняння має вигляд:

$$(\lambda + 1)^5 = 0,$$

тобто матриця $\boldsymbol{A}$ має власне значення $\lambda = -1$ п'ятої кратності. Матриця $\boldsymbol{A} + \boldsymbol{E}$ дорівнює

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

Власні вектори матриці $\boldsymbol{A}$, які задовольняють рівнянню $(\boldsymbol{A} + \boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}$, дорівнюють $\boldsymbol{x}_1 = \boldsymbol{e}_1$ і $\boldsymbol{x}_2 = \boldsymbol{e}_4$, де через $\boldsymbol{e}_i$ позначаються вектори натурального базису.

Зазначена інформація про матрицю $\boldsymbol{A}$ дозволяє стверджувати, що її жорданова форма вміщує дві жорданові клітини, проте не визначає розміри цих клітин. Дійсно, можливі два варіанти для жорданової форми:

$$\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_4(-1),\ \boldsymbol{J}_1(-1)), \quad \boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_3(-1),\ \boldsymbol{J}_2(-1)). \qquad (4)$$

Для з'ясування розмірів жорданових клітин виконаємо необхідні обчислення:

$$(\boldsymbol{A}+\boldsymbol{E})^2 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad r((\boldsymbol{A}+\boldsymbol{E})^2) = 1, \quad (\boldsymbol{A}+\boldsymbol{E})^3 = \boldsymbol{0}.$$

З останньої рівності випливає, що порядок найбільшої жорданової клітини дорівнює трьом, а передостання рівність указує на наявність однієї такої клітини. Проте і безпосередньо зрозуміло, що жорданова клітина $\boldsymbol{J}_3(-1)$ повинна бути єдиною. Отже, має місце другий варіант з (4).

Жордановій клітині $\boldsymbol{J}_3(-1)$ відповідає жорданів ланцюжок $\boldsymbol{h}_{11},\ \boldsymbol{h}_{12},\ \boldsymbol{h}_{13}$, вектори якого задовольняють рівнянням

$$(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{h}_{11} = \boldsymbol{0}, \quad (\boldsymbol{A}+\boldsymbol{E})\boldsymbol{h}_{12} = \boldsymbol{h}_{11}, \quad (\boldsymbol{A}+\boldsymbol{E})\boldsymbol{h}_{13} = \boldsymbol{h}_{12}. \qquad (5)$$

Для клітини $\boldsymbol{J}_2(-1)$ маємо вектори $\boldsymbol{h}_{21},\ \boldsymbol{h}_{22}$ і рівняння

$$(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{h}_{21} = \boldsymbol{0}, \quad (\boldsymbol{A}+\boldsymbol{E})\boldsymbol{h}_{22} = \boldsymbol{h}_{21}. \qquad (6)$$

Почнемо з рівнянь (5). Вектор $\boldsymbol{h}_{11}$, як показує перше рівняння (5), є власним вектором матриці $\boldsymbol{A}$. Він може дорівнювати $\boldsymbol{x}_1,\ \boldsymbol{x}_2$ або деякій їх лінійній комбінації. Будемо виходити з найбільш загального випадку, коли $\boldsymbol{h}_{11}$ дорівнює лінійній комбінації власних векторів $\boldsymbol{x}_1$ і $\boldsymbol{x}_2$:

$$\boldsymbol{h}_{11} = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 = (\alpha_1 \quad 0 \quad 0 \quad \alpha_2 \quad 0)^T. \qquad (7)$$

Залежність між коефіцієнтами α_1 і α_2 отримується шляхом дослідження сумісності розглядуваних нижче систем лінійних рівнянь.

Розв'яжемо друге рівняння (5). Для визначення координат вектора $\boldsymbol{h}_{12}$ маємо систему лінійних рівнянь з розширеною матрицею

$$(\boldsymbol{A}+\boldsymbol{E} \quad \boldsymbol{h}_{11}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \alpha_1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -2 & -1 & 0 & 1 & \alpha_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Перетворюючи її до ступінчастого вигляду, одержимо окремий розв'язок

$$\boldsymbol{h}_{12} = (0 \quad \alpha_1 \quad -\alpha_1 \quad 0 \quad \alpha_1+\alpha_2)^T. \tag{8}$$

Розглядувана система рівнянь виявилася сумісною і тому ніяких обмежень для α_1 і α_2 не було отримано. Продовжимо розв'язувати задачу, зосередившись на третьому рівнянні (5). Розширена матриця в цьому випадку дорівнює

$$(\boldsymbol{A}+\boldsymbol{E} \quad \boldsymbol{h}_{12}) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & \alpha_1 \\ 0 & -1 & -1 & 0 & 0 & -\alpha_1 \\ 0 & -2 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_1+\alpha_2 \end{pmatrix}. \tag{9}$$

Її п'ятий рядок указує на те, що для сумісності системи повинна виконуватися рівність $\alpha_1+\alpha_2=0$. Очевидно, що α_1 і α_2 одночасно не обертаються в нуль. Нехай $\alpha_1=1$ і $\alpha_2=-1$. Тоді на підставі формул (7) і (8) маємо:

$$\boldsymbol{h}_{11} = (1 \quad 0 \quad 0 \quad -1 \quad 0)^T, \quad \boldsymbol{h}_{12} = (0 \quad 1 \quad -1 \quad 0 \quad 0)^T.$$

Тепер ми можемо знайти вектор $\boldsymbol{h}_{13}$, перетворивши матрицю (9) до ступінчастого вигляду при вказаних вище значеннях α_1 і α_2. Окремий розв'язок системи з розширеною матрицею (9) дорівнює

$$\boldsymbol{h}_{13} = (0 \quad 0 \quad 1 \quad 0 \quad 1)^T.$$

Отже, рівняння (5) розв'язані. Звернемося до рівнянь (6). Для визначення векторів $\boldsymbol{h}_{21}$ і $\boldsymbol{h}_{22}$ скористаємося формулами (7) і (8). Як зазначалося вище, при їх одержанні не накладалися ніякі умови на α_1 і α_2. Тепер ми можемо вибрати α_1 і α_2, виходячи з єдиного обмеження, яке полягає в тому, щоб власні вектори $\boldsymbol{h}_{11}$ і $\boldsymbol{h}_{21}$ були лінійно незалежними. Зокрема, це буде в тому разі, коли $\boldsymbol{h}_{21}$ є власним вектором $\boldsymbol{x}_1$. Тоді $\alpha_1 = 1,\ \alpha_2 = 0$ і ми маємо:

$$\boldsymbol{h}_{21} = \boldsymbol{e}_1, \quad \boldsymbol{h}_{22} = (0 \quad 1 \quad -1 \quad 0 \quad 1)^T.$$

Можна було б узяти за $\boldsymbol{h}_{21}$ власний вектор $\boldsymbol{x}_2$, тоді $\alpha_1 = 0,\ \alpha_2 = 1$ і $\boldsymbol{h}_{21} = \boldsymbol{e}_4,\ \boldsymbol{h}_{22} = \boldsymbol{e}_5$.

Таким чином, усі стовпці перетворюючої матриці $\boldsymbol{T}$ знайдені і ми можемо написати її у вигляді:

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

Обчисливши $\boldsymbol{T}^{-1}$ і скориставшись формулою (9.1.4), одержимо:

$$\boldsymbol{A} = \boldsymbol{TJT}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \boldsymbol{J} \begin{pmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 2 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \end{pmatrix},$$

де

$$\boldsymbol{J} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}. \ \triangle$$

П р и к л а д 2. Побудуємо жорданову форму для матриці $\boldsymbol{A}$, яка є квадратом якої-небудь жорданової клітини $\boldsymbol{J}_k(0)$. Очевидно, що

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

Матриця $\boldsymbol{A}$ має єдине власне значення $\lambda = 0$ кратності k, якому відповідають два власні вектори, рівні векторам натурального базису $\boldsymbol{e}_1$ і $\boldsymbol{e}_2$. Тому жорданова форма $\boldsymbol{J}$ вміщує дві жорданові клітини.

Неважко переконатися в тому, що $\boldsymbol{A}^m = \boldsymbol{0}$ при $m = [(k+1)/2]$ і $\boldsymbol{A}^q \neq \boldsymbol{0}$ при $q = \overline{1,\ m-1}$. Це означає, що найвищий порядок жорданової клітини, належної до $\boldsymbol{J}$, дорівнює m. Порядок іншої жорданової клітини залежить від того, парним чи непарним є число k. Якщо $k = 2m$, то порядок другої клітини дорівнює m, тобто

$$\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_m(0),\ \boldsymbol{J}_m(0)).$$

Якщо ж $k = 2m - 1$, то порядок другої клітини повинен бути рівним $m - 1$, тобто

$$\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_m(0),\ \boldsymbol{J}_{m-1}(0)). \quad \triangle$$

9.2.2. Ортогональність векторів канонічних базисів. Наступна теорема є узагальненням теореми 6.4.1.

Теорема 9.2.1. *Довільний вектор жорданового ланцюжка матриці* $\boldsymbol{A}$*, відповідного до власного значення* λ*, є ортогональним до будь-якого вектора жорданового ланцюжка матриці* $\boldsymbol{A}^*$*, відповідного до її власного значення* $\mu \neq \bar{\lambda}$.

Д о в е д е н н я засноване на методі математичної індукції.

Нехай $\boldsymbol{x}_0$ — приєднаний вектор порядку m матриці $\boldsymbol{A}$, тоді для жорданового ланцюжка $\boldsymbol{x}_0,\ \boldsymbol{x}_1,\ \dots,\ \boldsymbol{x}_m$, написаного у зворотній послідовності в порівнянні з (9.1.2), виконуються співвідношення

$$
\begin{aligned}
\boldsymbol{x}_1 &= (\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x}_0, \\
\boldsymbol{x}_2 &= (\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x}_1, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
\boldsymbol{x}_m &= (\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x}_{m-1}, \\
(\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x}_m &= \boldsymbol{0}.
\end{aligned} \tag{10}
$$

Якщо $\boldsymbol{y}_0$ — приєднаний вектор порядку k матриці $\boldsymbol{A}^*$, то для її жорданового ланцюжка $\boldsymbol{y}_0, \ \boldsymbol{y}_1, \ \ldots, \ \boldsymbol{y}_k$ маємо аналогічні рівності

$$
\begin{aligned}
\boldsymbol{y}_1 &= (\boldsymbol{A}^* - \mu \boldsymbol{E})\boldsymbol{y}_0, \\
\boldsymbol{y}_2 &= (\boldsymbol{A}^* - \mu \boldsymbol{E})\boldsymbol{y}_1, \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
\boldsymbol{y}_k &= (\boldsymbol{A}^* - \mu \boldsymbol{E})\boldsymbol{y}_{k-1}, \\
(\boldsymbol{A}^* - \mu \boldsymbol{E})\boldsymbol{y}_k &= \boldsymbol{0}.
\end{aligned} \tag{11}
$$

Вектори $\boldsymbol{x}_m$ і $\boldsymbol{y}_k$ є власними векторами матриць $\boldsymbol{A}$ і $\boldsymbol{A}^*$ відповідно, тому на підставі теореми 6.4.1 вони є ортогональними.

Припустимо, що $\boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j} = 0$ при $i + j < l, \ l \geq 1$. Покажемо, що ця рівність виконується і при $i + j = l$. Дійсно, з урахуванням (10) і (11) при $i + j = l$ знаходимо, що

$$
\bar{\lambda}\boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j} = (\lambda\boldsymbol{x}_{m-i})^*\boldsymbol{y}_{k-j} = (\boldsymbol{A}\boldsymbol{x}_{m-i} - \boldsymbol{x}_{m-i+1})^*\boldsymbol{y}_{k-j} =
$$

$$
= (\boldsymbol{A}\boldsymbol{x}_{m-i})^*\boldsymbol{y}_{k-j} - \boldsymbol{x}_{m-i+1}^*\boldsymbol{y}_{k-j} = \boldsymbol{x}_{m-i}^*\boldsymbol{A}^*\boldsymbol{y}_{k-j} - \boldsymbol{x}_{m-i+1}^*\boldsymbol{y}_{k-j} =
$$

$$
= \boldsymbol{x}_{m-i}^*(\mu\boldsymbol{y}_{k-j} + \boldsymbol{y}_{k-j+1}) - \boldsymbol{x}_{m-i+1}^*\boldsymbol{y}_{k-j} =
$$

$$
= \mu\boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j} + \boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j+1} - \boldsymbol{x}_{m-i+1}^*\boldsymbol{y}_{k-j}.
$$

Оскільки $\boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j+1} = 0$ і $\boldsymbol{x}_{m-i+1}^*\boldsymbol{y}_{k-j} = 0$ на підставі припущення індукції, то будемо мати:

$$
\bar{\lambda}\boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j} = \mu\boldsymbol{x}_{m-i}^*\boldsymbol{y}_{k-j},
$$

або

$$(\bar{\lambda} - \mu)\boldsymbol{x}^*_{m-i}\boldsymbol{y}_{k-j} = 0.$$

Звідси, з огляду на те, що $\mu \neq \bar{\lambda}$, маємо умову ортогональності $\boldsymbol{x}^*_{m-i}\boldsymbol{y}_{k-j} = 0$. □

9.2.3. Збіжні матриці. У розділі 6.2 були визначені збіжні матриці і одержана умова збіжності для матриці простої структури. Можливість скористатися жордановою формою дозволяє узагальнити теорему 6.2.7 на випадок довільної матриці.

Теорема 9.2.2. *Для того, щоб $\boldsymbol{A}^m \to \boldsymbol{0}$ при $m \to \infty$, необхідно і достатньо, щоб усі власні значення матриці* ***A*** *були за модулем меншими ніж одиниця.*

Д о в е д е н н я. Нехай $\boldsymbol{T}$ — будь-яка невироджена матриця. Матриці $\boldsymbol{A}^m$ і $(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})^m = \boldsymbol{T}^{-1}\boldsymbol{A}^m\boldsymbol{T}$ прямують або не прямують до **0** одночасно. Тому достатньо довести теорему для жорданової форми матриці. Для того, щоб квазідіагональна матриця прямувала до **0**, необхідно і достатньо, щоб прямували до **0** всі її діагональні блоки. Отже, достатньо довести теорему для окремої жорданової клітини $\boldsymbol{J}_k(\lambda)$.

Покажемо, що елементи матриці $\boldsymbol{J}_k^m(\lambda)$ можна обчислити за формулою

$$\{\boldsymbol{J}_k^m(\lambda)\}_{ij} = C_m^{-i+j}\lambda^{m+i-j}, \tag{12}$$

де C_m^n — кількість сполучень з m елементів по n. Вона дорівнює

$$C_m^n = \frac{m!}{n!(m-n)!}, \tag{13}$$

причому вважається, що

$$C_m^n = 0, \quad n < 0, \quad n > m. \tag{14}$$

Матриця $\boldsymbol{J}_k(\lambda)$ — верхня трикутна, тому такою ж буде і $\boldsymbol{J}_k^m(\lambda)$, тобто $\{\boldsymbol{J}_k^m(\lambda)\}_{ij} = 0$ при $i > j$. Виконання цієї умови забезпечується рівністю (14). Для діагональних елементів $\{\boldsymbol{J}_k^m(\lambda)\}_{ii}$ рівність (12) виконується, бо при множенні верхніх трикутних матриць їх діагональні елементи перемножаються. Переконаємося в справедливості рівності (12) для елементів, розміщених вище головної діагоналі.

Скористаємося в цьому випадку методом математичної індукції. Рівність (12) виконується при $m = 1$ і це можна перевірити за допомогою рівностей (13) і (14). Припускаючи, що вона виконується для деякого числа $m-1$, пересвідчимося в її справедливості для m. Виходячи з рівності $\boldsymbol{J}_k^m(\lambda) = \boldsymbol{J}_k^{m-1}(\lambda)\boldsymbol{J}_k(\lambda)$, одержимо:

$$\{\boldsymbol{J}_k^m(\lambda)\}_{ij} = \{\boldsymbol{J}_k^{m-1}(\lambda)\}_{i,\,j-1} + \{\boldsymbol{J}_k^{m-1}(\lambda)\}_{ij}\lambda =$$

$$= C_{m-1}^{-i+j-1}\lambda^{m+i-j} + C_{m-1}^{-i+j}\lambda^{m+i-j} = (C_{m-1}^{-i+j-1} + C_{m-1}^{-i+j})\lambda^{m+i-j}. \qquad (15)$$

На підставі рівності (13) $C_{m-1}^{-i+j-1} + C_{m-1}^{-i+j} = C_m^{-i+j}$. Це означає, що з формули (15) випливає рівність (12).

Повернемося до питання про збіжність жорданової клітини $\boldsymbol{J}_k(\lambda)$. На підставі рівності (12) всі діагональні елементи матриці $\boldsymbol{J}_k^m(\lambda)$ дорівнюють λ^m. Для того, щоб $\boldsymbol{J}_k^m(\lambda) \to \mathbf{0}$, необхідно, щоб $\lambda^m \to 0$, тобто $|\lambda| < 1$.

Навпаки, якщо $|\lambda| < 1$, то досить показати, що при $i < j$ права частина рівності (12) прямує до нуля. Зважаючи на оцінку

$$\left|C_m^{-i+j}\lambda^{m+i-j}\right| = \left|\frac{m(m-1)...(m+i-j+1)\lambda^m}{(-i+j)!\lambda^{-i+j}}\right| \le \left|\frac{m^{-i+j}\lambda^m}{(-i+j)!\lambda^{-i+j}}\right|,$$

досить показати, що $m^{-i+j}|\lambda|^m \to 0$ при $m \to \infty$. Але це можна стверджувати на підставі правила Лопіталя. ☐

Одержана ознака збіжності $\boldsymbol{A}^m \to \mathbf{0}$ вимагає досить повних відомостей про власні значення матриці $\boldsymbol{A}$. Існує більш зручна в застосуваннях ознака, яка розглядається в розділі 12.1.

Вправи до глави 9

9.1. Довести, що матриця $\boldsymbol{T}$ залишається для матриці $\boldsymbol{A}$ перетворюючою матрицею до тієї ж жорданової форми $\boldsymbol{J}$ при тому ж розташуванні жорданових клітин на діагоналі матриці $\boldsymbol{J}$, якщо помножити всі вектори деякого жорданового ланцюжка на довільне число, відмінне від нуля.

9.2. Для жорданової клітини $\boldsymbol{J}_k(0)$ довести, що

$$\boldsymbol{J}_k^T(0)\boldsymbol{J}_k(0) = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{E}_{k-1} \end{pmatrix}.$$

9.3. Показати, що при перетворенні подібності з матрицею переставлень

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

будь-яка жорданова клітина переходить у транспоновану до неї матрицю, тобто

$$\boldsymbol{P}^{-1}\boldsymbol{J}_k(\lambda)\boldsymbol{P} = \boldsymbol{P}\boldsymbol{J}_k(\lambda)\boldsymbol{P} = \boldsymbol{J}_k^T(\lambda).$$

9.4. Довести, що жорданова форма $\boldsymbol{J}$ подібна до $\boldsymbol{J}^T$.

9.5. Довести, що будь-яка матриця $\boldsymbol{A}$ подібна до $\boldsymbol{A}^T$.

9.6.Як зміниться доведення теореми 9.1.1 у випадку, коли $\dim(\mathcal{N}(\boldsymbol{B}) \cap \mathcal{R}(\boldsymbol{B})) = 0$?

9.7. Матриця $\boldsymbol{N}$ називається *нільпотентною* матрицею *індексу* m, якщо існує ціле число $m > 1$ таке, що $\boldsymbol{N}^m = \mathbf{0}$, але $\boldsymbol{N}^{m-1} \neq \mathbf{0}$. Довести, що будь-яка жорданова клітина $\boldsymbol{J}_k(\lambda)$ є сумою скалярної і нільпотентної матриць.

9.8. Довести, що жорданову форму (9.1.3) можна подати у вигляді суми $\boldsymbol{J} = \boldsymbol{D} + \boldsymbol{N}$, де $\boldsymbol{D}$ — діагональна матриця з такою ж головною діагоналлю, як і в $\boldsymbol{J}$, а $\boldsymbol{N}$ — нільпотентна матриця, причому $\boldsymbol{N}^k = \boldsymbol{0}$, якщо k є порядком найбільшої жорданової клітини в $\boldsymbol{J}$. Показати, що матриці $\boldsymbol{D}$ і $\boldsymbol{N}$ є переставними.

9.9. Довести, що всі власні значення нільпотентної матриці дорівнюють нулю. Чи має така матриця просту структуру?

9.10. Скориставшись жордановою формою довільної матриці $\boldsymbol{A}$, довести, що $\boldsymbol{A} = \boldsymbol{A_D} + \boldsymbol{A_N}$, де $\boldsymbol{A_D}$ — матриця простої структури, а $\boldsymbol{A_N}$ — нільпотентна матриця, причому матриці $\boldsymbol{A_D}$ і $\boldsymbol{A_N}$ є переставними. Як пов'язаний індекс нільпотентності матриці $\boldsymbol{A_N}$ з виглядом жорданової форми матриці $\boldsymbol{A}$?

9.11. Нехай задана довільна матриця $\boldsymbol{A}$ і число $\varepsilon \neq 0$. Довести, що існує невироджена матриця $\boldsymbol{T} = \boldsymbol{T}(\varepsilon)$ така, що

$$\boldsymbol{T}^{-1}\boldsymbol{AT} = \operatorname{diag}(\boldsymbol{J}_{n_1}(\lambda_1, \varepsilon), \boldsymbol{J}_{n_2}(\lambda_2, \varepsilon), \dots, \boldsymbol{J}_{n_l}(\lambda_l, \varepsilon)), \quad \sum_{k=1}^{l} n_k = n,$$

де $\boldsymbol{J}_k(\lambda, \varepsilon)$ — матриця k-го порядку, рівна

$$\boldsymbol{J}_k(\lambda, \varepsilon) = \begin{pmatrix} \lambda & \varepsilon & \dots & 0 & 0 \\ 0 & \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda & \varepsilon \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Показати, що коли матриця $\boldsymbol{A}$ і число ε є дійсними, то матрицю $\boldsymbol{T}$ можна побудувати так, щоб вона була дійсною. Чи можна вважати число ε як завгодно малим за модулем?

ГЛАВА 10

СИСТЕМИ РІЗНИЦЕВИХ І ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ

У цій главі, якою відкривається друга частина книги, обговорюються застосування лінійної алгебри при дослідженні систем лінійних різницевих і диференціальних рівнянь першого порядку з постійними коефіцієнтами. Розглядаються також аналогічні різницеві і диференціальні рівняння більш високих порядків і показується, як їх можна звести до систем рівнянь першого порядку.

10.1. Системи різницевих рівнянь

10.1.1. Різницеві рівняння. На практиці часто зустрічаються задачі, в яких деяка функція $x(t)$ розглядається при послідовних значеннях аргумента t з інтервалом h, тобто мова йде про значення функції $x(t)$, $x(t+h)$, $x(t+2h)$ і т.д. Вибираючи належним чином одиницю вимірювання для t, завжди можна зробити $h=1$. У практичних задачах цей інтервал звичайно ототожнюють з періодом часу.

Рівняння $f(t,\ x(t),\dots,\ x(t+n))=0$ називається різницевим рівнянням n-го порядку. Спочатку ми розглянемо системи лінійних різницевих рівнянь першого порядку з постійними коефіцієнтами, які викликають найбільший інтерес у застосуваннях. Такі системи мають вигляд:

$$\boldsymbol{x}(t+1)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{b}\,. \qquad (1)$$

Тут $\boldsymbol{x}(t)=(x_1(t)\quad\dots\quad x_n(t))^T$ — вектор невідомих у момент часу t, причому початковий вектор $\boldsymbol{x}(0)$ вважається відомим, а матриця $\boldsymbol{A}$ і вектор $\boldsymbol{b}$ є довільними.

Знайдемо вираз для розв'язку $\boldsymbol{x}(t)$ через $\boldsymbol{x}(0)$:

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{x}(t-1) + \boldsymbol{b} = \boldsymbol{A}[\boldsymbol{A}\boldsymbol{x}(t-2) + \boldsymbol{b}] + \boldsymbol{b} = \boldsymbol{A}^2\boldsymbol{x}(t-2) +$$

$$+ (\boldsymbol{A} + \boldsymbol{E})\boldsymbol{b} = \ldots = \boldsymbol{A}^t\boldsymbol{x}(0) + (\boldsymbol{A}^{t-1} + \boldsymbol{A}^{t-2} + \ldots + \boldsymbol{E})\boldsymbol{b}. \qquad (2)$$

При $\boldsymbol{b} = \boldsymbol{0}$ ми маємо однорідну систему різницевих рівнянь $\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t)$. Перший доданок правої частини формули (2) є розв'язком однорідної системи при будь-якому $\boldsymbol{x}(0)$, тобто він є її загальним розв'язком. Вихідна система (1) є неоднорідною і формула (2) показує, що другий доданок її правої частини є окремим розв'язком цієї системи. Отже, подібно до систем лінійних алгебраїчних рівнянь загальний розв'язок неоднорідної системи різницевих рівнянь є сумою загального розв'язку відповідної однорідної системи і окремого розв'язку неоднорідної системи. Подібний факт має місце і в разі системи диференціальних рівнянь, у чому можна переконатись, ознайомившись з розділом 10.2.

Формулою (2) незручно користуватися, тому бажано надати їй іншого вигляду. Перетворимо суму

$$\boldsymbol{J} = \boldsymbol{A}^{t-1} + \boldsymbol{A}^{t-2} + \ldots + \boldsymbol{E},$$

для чого помножимо зліва обидві частини цієї рівності на $\boldsymbol{E} - \boldsymbol{A}$:

$$(\boldsymbol{E} - \boldsymbol{A})\boldsymbol{J} = \boldsymbol{E} - \boldsymbol{A}^t.$$

Якщо $(\boldsymbol{E} - \boldsymbol{A})^{-1}$ існує, то

$$\boldsymbol{J} = (\boldsymbol{E} - \boldsymbol{A})^{-1}(\boldsymbol{E} - \boldsymbol{A}^t) \qquad (3)$$

і тепер формула (2) набуває вигляду:

$$\boldsymbol{x}(t) = \boldsymbol{A}^t\boldsymbol{x}(0) + (\boldsymbol{E} - \boldsymbol{A})^{-1}(\boldsymbol{E} - \boldsymbol{A}^t)\boldsymbol{b}. \qquad (4)$$

У багатьох випадках важливою є інформація про поведінку розв'язку при $t \to \infty$. З'ясуємо, чи буде вектор $\boldsymbol{x}(t)$ незалежно від того, яке значення $\boldsymbol{x}(0)$ він мав у початковий момент, прямувати до деякого постійного вектора, коли $t \to \infty$.

Аналізуючи розв'язок (4), неважко переконатися в тому, що для цього достатньо, щоб $\boldsymbol{A}^t \to \boldsymbol{0}$. Теореми 6.2.7 і 9.2.2 показують, що ця умова виконується, коли всі власні значення матриці $\boldsymbol{A}$ за модулем є меншими ніж одиниця. При цьому з рівності (4) знаходимо, що

$$\boldsymbol{x}(\infty) = \lim_{t\to\infty} \boldsymbol{x}(t) = (\boldsymbol{E} - \boldsymbol{A})^{-1}\boldsymbol{b}. \tag{5}$$

Якщо принаймні одне власне значення матриці $\boldsymbol{A}$ задовольняє умові $|\lambda_i| > 1$, то $\boldsymbol{x}(t)$ необмежено зростає при $t \to \infty$ (винятки можливі лише за додаткової умови для $\boldsymbol{x}(0)$). Розв'язок є обмеженим, коли всі власні значення задовольняють умові $|\lambda_i| \le 1$. Якщо власне значення λ_i від'ємне, то розв'язок зазнає коливань, причому ці коливання згасаючі, зростаючі або періодичні в залежності від того, чи модуль власного значення менше, більше або дорівнює одиниці.

10.1.2. Структура загального розв'язку однорідної системи різницевих рівнянь. З'ясуємо, як виражається загальний розв'язок однорідної системи різницевих рівнянь через власні вектори і власні значення матриці $\boldsymbol{A}$.

Якщо матриця $\boldsymbol{A}$ має просту структуру, то загальний розв'язок однорідної системи можна подати у вигляді розкладу за власними векторами цієї матриці. Дійсно, скориставшись формулою (6.2.7), одержимо:

$$\boldsymbol{A}^t\boldsymbol{x}(0) = \boldsymbol{S}\boldsymbol{\Lambda}^t\boldsymbol{S}^{-1}\boldsymbol{x}(0) = (\boldsymbol{x}_1 \ \ \dots \ \ \boldsymbol{x}_n)\,\mathrm{diag}\,(\lambda_1^t, \dots, \lambda_n^t)\,\boldsymbol{S}^{-1}\boldsymbol{x}(0) =$$

$$= (\lambda_1^t\boldsymbol{x}_1 \ \ \dots \ \ \lambda_n^t\boldsymbol{x}_n)\boldsymbol{S}^{-1}\boldsymbol{x}(0) = \sum_{i=1}^{n} c_i\lambda_i^t\boldsymbol{x}_i, \tag{6}$$

де $c_1, \dots, c_n$ — координати вектора $\boldsymbol{S}^{-1}\boldsymbol{x}(0)$.

У випадку дефектної матриці $\boldsymbol{A}$ замість розкладу розв'язку однорідної системи за власними векторами можна побудувати розклад за векторами канонічного базису. Розгляд цього питання для довільної системи рівнянь пов’язаний з громіздкими міркуваннями, тому ми обмежимося найбільш простим випадком системи двох рівнянь. При цьому ми скористаємося тими відомостями про дефектні матриці, які були одержані в розділі 6.2.

За допомогою формули (6.2.9) знаходимо:

$$\boldsymbol{A}^t\boldsymbol{x}(0) = \boldsymbol{T}\boldsymbol{J}^t\boldsymbol{T}^{-1}\boldsymbol{x}(0), \tag{7}$$

де матриця $\boldsymbol{J}^t$, як неважко переконатися за допомогою другої формули (6.2.10), дорівнює

$$\boldsymbol{J}^t = \begin{pmatrix} \lambda^t & t\lambda^{t-1} \\ 0 & \lambda^t \end{pmatrix}. \tag{8}$$

Скориставшись першою формулою (6.2.10), рівністю (8) і правилами дій над блоковими матрицями, одержимо:

$$\boldsymbol{T}\boldsymbol{J}^t = (\boldsymbol{x}_1 \quad \boldsymbol{x}_2)\begin{pmatrix} \lambda^t & t\lambda^{t-1} \\ 0 & \lambda^t \end{pmatrix} = \boldsymbol{x}_1(\lambda^t \quad t\lambda^{t-1}) + \boldsymbol{x}_2(0 \quad \lambda^t) =$$

$$= (\lambda^t\boldsymbol{x}_1 \quad t\lambda^{t-1}\boldsymbol{x}_1) + (\mathbf{0} \quad \lambda^t\boldsymbol{x}_2) = (\lambda^t\boldsymbol{x}_1 \quad t\lambda^{t-1}\boldsymbol{x}_1 + \lambda^t\boldsymbol{x}_2). \tag{9}$$

З формул (7) і (9) випливає, що

$$\begin{aligned} &\boldsymbol{A}^t\boldsymbol{x}(0) = (\lambda^t\boldsymbol{x}_1 \quad t\lambda^{t-1}\boldsymbol{x}_1 + \lambda^t\boldsymbol{x}_2)\boldsymbol{T}^{-1}\boldsymbol{x}(0) = c_1\lambda^t\boldsymbol{x}_1 + \\ &+c_2(t\lambda^{t-1}\boldsymbol{x}_1 + \lambda^t\boldsymbol{x}_2) = (c_1\lambda + tc_2)\lambda^{t-1}\boldsymbol{x}_1 + c_2\lambda^t\boldsymbol{x}_2, \end{aligned} \tag{10}$$

де c_1 і c_2 — координати вектора $\boldsymbol{T}^{-1}\boldsymbol{x}(0)$. Формула (10) дає розклад загального розв’язку однорідної системи двох різницевих рівнянь за векторами канонічного базису $\boldsymbol{x}_1, \boldsymbol{x}_2$.

10.1.3. Умова спадання розв’язку однорідної системи різницевих рівнянь. Якщо всі власні значення матриці $\boldsymbol{A}$ за модулем менші ніж одиниця, то справедлива рівність (6), з якої для однорідної системи випливає, що $\boldsymbol{x}(\infty) = \mathbf{0}$. При розв’язанні деяких задач важливо знати характер спадання розв’язку $\boldsymbol{x}(t)$ при $t \to \infty$. Може бути так, що довжини векторів $\boldsymbol{x}(t)$ спочатку зростають і лише потім, починаючи з деякого t, прямують до нуля. Умова негайного спадання $|\,\boldsymbol{x}(t)\,|$ з'ясовується в наступній теоремі.

Теорема 10.1.1. *Нехай* $\boldsymbol{x}(t)$ *— розв'язок однорідної системи різницевих рівнянь і матриця* $\boldsymbol{E}-\boldsymbol{A}^T\boldsymbol{A}$ *є додатно визначеною. Тоді* $|\,\boldsymbol{x}(t)\,|$ *негайно спадає, тобто* $|\,\boldsymbol{x}(1)\,|<|\,\boldsymbol{x}(0)\,|$ *для будь-якого початкового вектора* $\boldsymbol{x}(0)$.

Д о в е д е н н я. З додатної визначеності матриці $\boldsymbol{E}-\boldsymbol{A}^T\boldsymbol{A}$ випливає, що при $\boldsymbol{x}(t)\neq\boldsymbol{0}$ має місце нерівність

$$\boldsymbol{x}^T(t)\boldsymbol{x}(t)>\boldsymbol{x}^T(t)\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}(t)$$

для будь-якого t. Останню нерівність запишемо у вигляді:

$$|\,\boldsymbol{x}(t)\,|^2>|\,\boldsymbol{A}\boldsymbol{x}(t)\,|^2.$$

Звідси при $t=0$ маємо:

$$|\,\boldsymbol{x}(0)\,|>|\,\boldsymbol{A}\boldsymbol{x}(0)\,|=|\,\boldsymbol{x}(1)\,|,$$

що і треба довести. ☐

10.1.4. Лінійні різницеві рівняння з постійними коефіцієнтами. Многочлени Чебишова. У застосуваннях часто зустрічаються лінійні різницеві рівняння з постійними коефіцієнтами другого і більш високого порядків. Такі рівняння можуть бути зведені до системи лінійних різницевих рівнянь першого порядку. Проілюструємо це, розглядаючи *многочлени Чебишова* $T_n(x)$, які при $n=0,1,2,\ldots$ і $|\,x\,|\leq 1$ визначаються як розв'язок різницевого рівняння і початкових умов

$$\begin{gathered} y_{n+2}(x)-2xy_{n+1}(x)+y_n(x)=0, \quad n\geq 0, \\ y_0(x)=1, \quad y_1(x)=x. \end{gathered} \tag{11}$$

Напишемо різницеве рівняння разом з тотожністю:

$$\begin{gathered} y_{n+2}(x)=2xy_{n+1}(x)-y_n(x), \\ y_{n+1}(x)=y_{n+1}(x) \end{gathered}$$

і введемо до розгляду вектор $\boldsymbol{z}(n)=(y_{n+1}(x) \quad y_n(x))^T$. Тоді буде $\boldsymbol{z}(n+1)=(y_{n+2}(x) \quad y_{n+1}(x))^T$ і систему рівнянь та початкові умови можна представити у вигляді:

$$\boldsymbol{z}(n+1) = \boldsymbol{F}\boldsymbol{z}(n),\ \boldsymbol{F} = \begin{pmatrix} 2x & -1 \\ 1 & 0 \end{pmatrix},\ \boldsymbol{z}(0) = \begin{pmatrix} x \\ 1 \end{pmatrix}. \qquad (12)$$

Для розв'язання одержаної однорідної системи різницевих рівнянь необхідно знайти власні значення і власні вектори матриці $\boldsymbol{F}$. Вони є такими:

$$\lambda_1 = x + i\sqrt{1-x^2},\ \ \boldsymbol{x}_1 = (\lambda_1 \ \ 1)^T,\ \ \lambda_2 = x - i\sqrt{1-x^2},\ \ \boldsymbol{x}_2 = (\lambda_2 \ \ 1)^T.$$

Розв'язок задачі (12) визначимо за допомогою формули (6):

$$\boldsymbol{z}(n) = \boldsymbol{F}^n\boldsymbol{z}(0) = c_1\lambda_1^n\boldsymbol{x}_1 + c_2\lambda_2^n\boldsymbol{x}_2.$$

Знайдемо c_1 і c_2, обчисливши координати вектора $\boldsymbol{S}^{-1}\boldsymbol{z}(0)$:

$$\boldsymbol{S}^{-1}\boldsymbol{z}(0) = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}^{-1}\boldsymbol{z}(0) = \frac{1}{\lambda_1 - \lambda_2}\begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}\begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix},$$

звідки випливає, що $c_1 = c_2 = 1/2$.

Тепер, ураховуючи, що многочлен Чебишова $T_n(x)$ є другою координатою вектора $\boldsymbol{z}(n)$, знаходимо:

$$T_n(x) = c_1\lambda_1^n + c_2\lambda_2^n = \frac{1}{2}((x + i\sqrt{1-x^2})^n + (x - i\sqrt{1-x^2})^n). \qquad (13)$$

Будемо вважати, що $x = \cos\varphi$. Тоді $\lambda_1 = e^{i\varphi}$, $\lambda_2 = e^{-i\varphi}$ і замість (13) будемо мати:

$$T_n(x) = \cos(n \arccos x). \qquad (14)$$

10.1.5. Обчислення визначників *n*-го порядку. Метод Гаусса обчислення визначників, який розглядався в розділі 2.1, стає непридатним для цієї мети, якщо визначник має n-й порядок. Хоч загального методу для обчислення таких визначників не існує, проте найбільш успішним виявляється спосіб, заснований на зведенні проблеми до розв'язання різницевого рівняння. Приклад, що наводиться нижче, є типовим при використанні цього способу.

Нехай треба обчислити визначник n-го порядку

$$\Delta_n = \begin{vmatrix} 2\alpha & \alpha^2 & 0 & \dots & 0 & 0 \\ 1 & 2\alpha & \alpha^2 & \dots & 0 & 0 \\ 0 & 1 & 2\alpha & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 2\alpha \end{vmatrix}.$$

Розклавши його за елементами першого стовпця, одержимо:

$$\Delta_n = 2\alpha \begin{vmatrix} 2\alpha & \alpha^2 & \dots & 0 & 0 \\ 1 & 2\alpha & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 2\alpha \end{vmatrix} - \begin{vmatrix} \alpha^2 & 0 & \dots & 0 & 0 \\ 1 & 2\alpha & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 2\alpha \end{vmatrix}. \qquad (15)$$

Перший визначник з правої частини рівності (15) є визначником того ж вигляду, що і Δ_n, але на одиницю меншого порядку. Другий доданок з правої частини (15) зводиться до Δ_{n-2} після розкладення його за елементами першого рядка. Отже, рівність (15) можна написати у вигляді різницевого рівняння другого порядку

$$\Delta_n = 2\alpha\,\Delta_{n-1} - \alpha^2\Delta_{n-2}.$$

Долучимо до цього рівняння тотожність $\Delta_{n-1} = \Delta_{n-1}$. Тоді, позначаючи вектор $(\Delta_n \quad \Delta_{n-1})^T$ через $\boldsymbol{x}(n+1)$, одержимо систему різницевих рівнянь

$$\boldsymbol{x}(n+1) = \boldsymbol{F}\boldsymbol{x}(n), \quad \boldsymbol{F} = \begin{pmatrix} 2\alpha & -\alpha^2 \\ 1 & 0 \end{pmatrix}. \qquad (16)$$

Щоб одержати початковий вектор для системи різницевих рівнянь (16), обчислимо Δ_2 і Δ_3:

$$\Delta_2 = \begin{vmatrix} 2\alpha & \alpha^2 \\ 1 & 2\alpha \end{vmatrix} = 3\alpha^2, \quad \Delta_3 = \begin{vmatrix} 2\alpha & \alpha^2 & 0 \\ 1 & 2\alpha & \alpha^2 \\ 0 & 1 & 2\alpha \end{vmatrix} = 4\alpha^3.$$

Початковим вектором є $\boldsymbol{x}(4) = (4\alpha^3 \quad 3\alpha^2)^T$. Далі маємо:

$$\boldsymbol{x}(n+1) = \boldsymbol{F}\boldsymbol{x}(n) = \boldsymbol{F}^2\boldsymbol{x}(n-1) = \boldsymbol{F}^3\boldsymbol{x}(n-2) = \ldots = \boldsymbol{F}^{n-3}\boldsymbol{x}(4). \quad (17)$$

Матриця $\boldsymbol{F}$ є дефектною, оскільки двократному власному значенню, рівному α, відповідає один власний вектор $\boldsymbol{x}_1 = (\alpha \quad 1)^T$. Другий вектор канонічного базису знаходиться як окремий розв'язок рівняння (6.2.8) і дорівнює $\boldsymbol{x}_2 = (1 \quad 0)^T$. Матриця $\boldsymbol{T}$ з (6.2.10) і обернена до неї є такими:

$$\boldsymbol{T} = \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{T}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -\alpha \end{pmatrix}.$$

Тепер з формул (17) і (10) випливає, що

$$\boldsymbol{x}(n+1) = (c_1\alpha + (n-3)c_2)\alpha^{n-4}\boldsymbol{x}_1 + c_2\alpha^{n-3}\boldsymbol{x}_2, \quad (18)$$

де c_1 і c_2 — координати вектора $\boldsymbol{T}^{-1}\boldsymbol{x}(4)$. Оскільки

$$\boldsymbol{T}^{-1}\boldsymbol{x}(4) = \begin{pmatrix} 0 & 1 \\ 1 & -\alpha \end{pmatrix}\begin{pmatrix} 4\alpha^3 \\ 3\alpha^2 \end{pmatrix} = \begin{pmatrix} 3\alpha^2 \\ \alpha^3 \end{pmatrix},$$

то $c_1 = 3\alpha^2, c_2 = \alpha^3$. Підставляючи ці величини і вирази для векторів $\boldsymbol{x}_1, \boldsymbol{x}_2$ у рівність (18), одержимо:

$$\boldsymbol{x}(n+1) = ((n+1)\alpha^n \quad n\alpha^{n-1})^T.$$

Шуканий визначник є першою координатою вектора $\boldsymbol{x}(n+1)$, тобто $\Delta_n = (n+1)\alpha^n$.

З а у в а ж е н н я. Матриця $\boldsymbol{F}$ у формулах (12) і (16) є матрицею Фробеніуса (6.1.6) для многочленів $\lambda^2 - 2x\lambda + 1$ і $\lambda^2 - 2\alpha\lambda + \alpha^2$ відповідно. Цією обставиною можна скористатися при знаходженні власних векторів. Необхідна для цього інформація щодо матриці Фробеніуса наводиться в розділі 10.2.Очевидно, що $\boldsymbol{F}$ буде матрицею Фробеніуса для деякого многочлена при розгляді різницевого рівняння довільного порядку.

10.2. Системи диференціальних рівнянь

10.2.1. Однорідна система диференціальних рівнянь і матрична експонента. Розглянемо систему лінійних однорідних диференціальних рівнянь з постійними коефіцієнтами

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^{n} a_{ij} x_j(t), \quad i = \overline{1,\ n}, \tag{1}$$

при початкових умовах $x_i(0) = x_{0i}$, $i = \overline{1,\ n}$. Якщо $\boldsymbol{A}$ — матриця з елементами a_{ij}, $\boldsymbol{x}(t) = (x_1(t)\ \ ...\ \ x_n(t))^T$ і $\boldsymbol{x}_0 = (x_{01}\ ...\ x_{0n})^T$, то систему і початкові умови можна записати у вигляді:

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\boldsymbol{x}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0. \tag{2}$$

Тут і в подальшому під похідною будь-якої матриці $\boldsymbol{X}$ ми розуміємо матрицю, яка одержується з $\boldsymbol{X}$ диференціюванням її елементів. Зокрема,

$$\frac{d\boldsymbol{x}(t)}{dt} = \left(\frac{dx_1(t)}{dt} \quad \cdots \quad \frac{dx_n(t)}{dt} \right)^T.$$

Нехай матриці $\boldsymbol{A}$ відповідає жорданова форма (9.1.3). Нагадаємо, що ця форма стає діагональною матрицею, якщо $\boldsymbol{A}$ має просту структуру. Виходячи з рівності (9.1.4), покажемо, що розв'язок задачі (2) має вигляд:

$$\boldsymbol{x}(t) = \boldsymbol{T} e^{\boldsymbol{J}t} \boldsymbol{T}^{-1} \boldsymbol{x}_0. \tag{3}$$

Тут $e^{\boldsymbol{J}t}$ — квазідіагональна матриця

$$e^{\boldsymbol{J}t} = \mathrm{diag}(\boldsymbol{I}_{n_1}(\lambda_1 t), \boldsymbol{I}_{n_2}(\lambda_2 t), ..., \boldsymbol{I}_{n_l}(\lambda_l t)), \quad \sum_{k=1}^{l} n_k = n, \tag{4}$$

кожний діагональний блок якої відповідає діагональному блоку матриці $\boldsymbol{J}$, причому

$$\boldsymbol{I}_k(\lambda t) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & (t^2/2!)e^{\lambda t} & ... & (t^{k-1}/(k-1)!)e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & ... & (t^{k-2}/(k-2)!)e^{\lambda t} \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & e^{\lambda t} \end{pmatrix}. \tag{5}$$

Очевидно, що розв'язок (3) задовольняє початковій умові $\boldsymbol{x}(0) = \boldsymbol{x}_0$. Обчислимо похідну від матриці $e^{\boldsymbol{J}t}$. Оскільки

$$\frac{d}{dt}e^{\boldsymbol{J}t} = \operatorname{diag}\left(\frac{d}{dt}\boldsymbol{I}_{n_1}(\lambda_1 t), \frac{d}{dt}\boldsymbol{I}_{n_2}(\lambda_2 t), ..., \frac{d}{dt}\boldsymbol{I}_{n_l}(\lambda_l t)\right), \tag{6}$$

то досить знати, як обчислювати похідну від підматриці $\boldsymbol{I}_k(\lambda t)$. Очевидно, що

$$\frac{d}{dt}\boldsymbol{I}_k(\lambda t) = \begin{pmatrix} \lambda e^{\lambda t} & \lambda t e^{\lambda t} & (\lambda t^2/2!)e^{\lambda t} & ... & (\lambda t^{k-1}/(k-1)!)e^{\lambda t} \\ 0 & \lambda e^{\lambda t} & \lambda t e^{\lambda t} & ... & (\lambda t^{k-2}/(k-2)!)e^{\lambda t} \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & \lambda e^{\lambda t} \end{pmatrix} +$$

$$+\begin{pmatrix} 0 & e^{\lambda t} & t e^{\lambda t} & ... & (t^{k-2}/(k-2)!)e^{\lambda t} \\ 0 & 0 & e^{\lambda t} & ... & (t^{k-3}/(k-3)!)e^{\lambda t} \\ ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 \end{pmatrix}.$$

Неважко переконатися в тому, що перший доданок правої частини останньої формули дорівнює $\lambda \boldsymbol{I}_k(\lambda t)$, а другий можна подати у вигляді добутку $\boldsymbol{J}_k(0)\boldsymbol{I}_k(\lambda t)$ (або $\boldsymbol{I}_k(\lambda t)\boldsymbol{J}_k(0)$). Тепер з урахуванням рівності $\boldsymbol{J}_k(\lambda) = \lambda \boldsymbol{E}_k + \boldsymbol{J}_k(0)$ маємо:

$$\frac{d}{dt}\boldsymbol{I}_k(\lambda t) = \boldsymbol{J}_k(\lambda)\boldsymbol{I}_k(\lambda t). \tag{7}$$

Далі за допомогою формул (3), (6), (9.1.3) і (4) знаходимо:

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{T}\frac{d}{dt}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1}\boldsymbol{x}_0 =$$

$$= \boldsymbol{T}\operatorname{diag}(\boldsymbol{J}_{n_1}(\lambda_1)\boldsymbol{I}_{n_1}(\lambda_1 t), \boldsymbol{J}_{n_2}(\lambda_2)\boldsymbol{I}_{n_2}(\lambda_2 t), ..., \boldsymbol{J}_{n_l}(\lambda_l)\boldsymbol{I}_{n_l}(\lambda_l t))\boldsymbol{T}^{-1}\boldsymbol{x}_0 =$$

$$= \boldsymbol{T}\boldsymbol{J}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1}\boldsymbol{x}_0. \tag{8}$$

Формулі (9.1.4) можна надати вигляд рівності $\boldsymbol{TJ} = \boldsymbol{AT}$, тому з формул (8) і (3) випливає, що

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\boldsymbol{T}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1}\boldsymbol{x}_0 = \boldsymbol{A}\boldsymbol{x}(t),$$

тобто маємо рівняння (2).

Таким чином, формула (3) визначає розв'язок задачі (2) Якщо матриця $\boldsymbol{A}$ має просту структуру, то замість (3) одержимо з урахуванням рівності (6.2.6):

$$\boldsymbol{x}(t) = \boldsymbol{S}e^{\Lambda t}\boldsymbol{S}^{-1}\boldsymbol{x}_0, \quad e^{\Lambda t} = \operatorname{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, ..., e^{\lambda_n t}). \tag{9}$$

Матриця $\boldsymbol{T}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1}$ (або $\boldsymbol{S}e^{\Lambda t}\boldsymbol{S}^{-1}$) називається *матричною експонентою* і позначається через $e^{\boldsymbol{A}t}$. Підставою для надання їй такої назви послужила та обставина, що в разі однієї невідомої задача (2) має розв'язок у вигляді експоненти $x(t) = e^{At}x_0$.

10.2.2. Умова спадання розв'язку однорідної системи диференціальних рівнянь. З формул (3) — (5), а також (9) випливає, що $|\boldsymbol{x}(t)| \to 0$ при необмеженому збільшенні t за умови $\operatorname{Re}\lambda_i < 0$ для всіх i. Якщо для всіх власних значень $\operatorname{Re}\lambda_i \le 0$, то розв'язок обмежений при $t \to \infty$. Нарешті, розв'язок необмежено зростає при $t \to \infty$, коли $\operatorname{Re}\lambda_i > 0$ принаймні для одного власного значення (за винятком випадків з вельми спеціальними початковими умовами).

Як і в разі однорідної системи різницевих рівнянь, розв'язок системи диференціальних рівнянь може короткостроково збільшуватися перш ніж почне спадати. Умова спадання довжини розв'язку з моменту $t = 0$ з'ясовується в наступній теоремі.

Теорема 10.2.1. *Нехай $\boldsymbol{x}(t)$ — розв'язок однорідної системи диференціальних рівнянь (2) і матриця $\boldsymbol{A} + \boldsymbol{A}^T$ є від'ємно визначеною. Тоді $|\boldsymbol{x}(t)|$ спадає з моменту $t = 0$, а при $\boldsymbol{A} + \boldsymbol{A}^T = \boldsymbol{0}$ величина $|\boldsymbol{x}(t)|$ залишається постійною.*

Д о в е д е н н я. Неважко переконатися в тому, що

$$\frac{d}{dt}\boldsymbol{x}^T(t)\boldsymbol{y}(t) = \left(\frac{d\boldsymbol{x}(t)}{dt}\right)^T \boldsymbol{y}(t) + \boldsymbol{x}^T(t)\frac{d\boldsymbol{y}(t)}{dt}.$$

Звідси для розв'язку рівняння $\dfrac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\boldsymbol{x}(t)$ маємо:

$$\frac{d}{dt}\ |\boldsymbol{x}(t)|^2 = \frac{d}{dt}(\boldsymbol{x}^T(t)\boldsymbol{x}(t)) = \left(\frac{d\boldsymbol{x}(t)}{dt}\right)^T \boldsymbol{x}(t) + \boldsymbol{x}^T(t)\frac{d\boldsymbol{x}(t)}{dt} =$$

$$= \boldsymbol{x}^T(t)(\boldsymbol{A}^T + \boldsymbol{A})\boldsymbol{x}(t).$$

Оскільки $\boldsymbol{x}^T(t)(\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{x}(t) < 0$, то довжина вектора $\boldsymbol{x}(t)$ спадає в кожний момент часу. При $\boldsymbol{A} + \boldsymbol{A}^T = \boldsymbol{0}$ довжина $|\boldsymbol{x}(t)|$ залишається постійною. □

10.2.3. Неоднорідні системи диференціальних рівнянь. Покажемо, що неоднорідна система лінійних диференціальних рівнянь з початковими умовами

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{f}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0 \tag{10}$$

має розв'язок

$$\boldsymbol{x}(t) = \int_0^t e^{\boldsymbol{A}(t-s)}\boldsymbol{f}(s)\,ds + e^{\boldsymbol{A}t}\boldsymbol{x}_0. \tag{11}$$

Вище ми бачили, що матриці $\boldsymbol{J}_k(0)$ і $\boldsymbol{I}_k(\lambda t)$ є переставними. Як наслідок маємо переставність матриць $\boldsymbol{J}_k(\lambda)$ і $\boldsymbol{I}_k(\lambda t)$, а тому і матриць $\boldsymbol{J}$ та $e^{\boldsymbol{J}t}$. Далі з урахуванням формули (8) знаходимо, що

$$\frac{d}{dt}e^{\boldsymbol{A}t} = \boldsymbol{T}\frac{d}{dt}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1} = \boldsymbol{T}\boldsymbol{J}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1} =$$

$$= \boldsymbol{T}e^{\boldsymbol{J}t}\boldsymbol{J}\boldsymbol{T}^{-1} = \boldsymbol{T}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1} = e^{\boldsymbol{A}t}\boldsymbol{A}.$$

Тепер за допомогою рівняння (10) одержуємо:

$$\frac{d}{dt}(e^{-\boldsymbol{A}t}\boldsymbol{x}(t)) = e^{-\boldsymbol{A}t}\left(\frac{d\boldsymbol{x}(t)}{dt} - \boldsymbol{A}\,\boldsymbol{x}(t)\right) = e^{-\boldsymbol{A}t}\boldsymbol{f}(t),$$

звідки

$$e^{-\boldsymbol{A}t}\boldsymbol{x}(t) = \int_0^t e^{-\boldsymbol{A}s}\boldsymbol{f}(s)ds + \boldsymbol{c}.$$

З останньої формули, в якій треба покласти $\boldsymbol{c} = \boldsymbol{x}_0$, випливає рівність (11). При $\boldsymbol{f}(s) = \boldsymbol{0}$ з неї одержується розв'язок однорідної системи (2).

10.2.4. Лінійні диференціальні рівняння з постійними коефіцієнтами. Диференціальні рівняння (1) є рівняннями першого порядку. Розглянемо лінійне диференціальне рівняння n-го порядку з постійними коефіцієнтами

$$\frac{d^n y(t)}{dt^n} + a_{n-1}\frac{d^{n-1}y(t)}{dt^{n-1}} + \ldots + a_1\frac{dy(t)}{dt} + a_0 y(t) = f(t). \qquad (12)$$

Будемо шукати його розв'язок при початкових умовах

$$y(0) = y_0, \quad \frac{dy(0)}{dt} = y_0', \quad \ldots, \quad \frac{d^{n-1}y(0)}{dt^{n-1}} = y_0^{(n-1)}. \qquad (13)$$

Задачу (12), (13) можна звести до системи диференціальних рівнянь (10) подібно до того, як це було зроблено для різницевого рівняння другого порядку в попередньому розділі. Для цього введемо позначення

$$y(t) = x_1(t),$$

$$\frac{dy(t)}{dt} = x_2(t),$$

$$\cdots\cdots\cdots\cdots$$

$$\frac{d^{n-2}y(t)}{dt^{n-2}} = x_{n-1}(t),$$

$$\frac{d^{n-1}y(t)}{dt^{n-1}} = x_n(t).$$

Далі напишемо рівняння (12), долучивши до нього тотожності:

$$\frac{dx_n(t)}{dt} = -a_{n-1}x_n(t) - a_{n-2}x_{n-1}(t) - \ldots - a_1 x_2(t) - a_0 x_1(t) + f(t),$$

$$\frac{dx_{n-1}(t)}{dt} = x_n(t),$$

$$\cdots\cdots\cdots\cdots$$

$$\frac{dx_2(t)}{dt} = x_3(t),$$

$$\frac{dx_1(t)}{dt} = x_2(t).$$

Тепер систему диференціальних рівнянь і початкові умови запишемо у вигляді:

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{F}\boldsymbol{x}(t) + \boldsymbol{f}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \qquad (14)$$

де

$$\boldsymbol{x}(t) = (x_n(t) \quad x_{n-1}(t) \quad \dots \quad x_1(t))^T, \quad \boldsymbol{x}_0 = (y_0^{(n-1)} \quad y_0^{(n-2)} \quad \dots \quad y_0)^T,$$

$$\boldsymbol{f}(t) = (f(t) \quad 0 \quad \dots \quad 0)^T,$$

а матриця $\boldsymbol{F}$ є матрицею Фробеніуса (6.1.6) для многочлена

$$\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \dots + a_1\lambda + a_0,$$

коефіцієнти якого збігаються з коефіцієнтами рівняння (12).

Візьмемо розв'язок задачі (14) у вигляді (11), після чого шуканий розв'язок $y(t)$ задачі (12), (13) буде визначатися останньою координатою вектора $\boldsymbol{x}(t)$.

При розв'язанні задачі (12), (13) треба враховувати, що власному значенню λ матриці Фробеніуса $\boldsymbol{F}$ відповідає власний вектор $(\lambda^{n-1} \quad \lambda^{n-2} \quad \dots \quad \lambda \quad 1)^T$. У цьому неважко переконатися за допомогою безпосередньої перевірки. Можна довести, що вказаний власний вектор є єдиним для власного значення λ незалежно від його кратності.

Вправи до глави 10

10.1. Нехай треба знайти функцію $y(t)$, яка задовольняє різницевому рівнянню і початковим умовам

$$y(t+n) + a_{n-1}y(t+n-1) + \dots + a_1 y(t+1) + a_0 y(t) = a,$$

$$y(0) = y_0, \quad y(1) = y_1, \quad \dots, \quad y(n-1) = y_{n-1}.$$

Звести цю задачу до розв'язання системи (10.1.1) з початковою умовою $\boldsymbol{x}(0) = \boldsymbol{x}_0$. Якими є матиця $\boldsymbol{A}$ і вектори $\boldsymbol{x}(t)$, $\boldsymbol{x}_0$, $\boldsymbol{b}$?

10.2. Не користуючись жордановою формою, одержати вираз для матричної експонети $e^{\boldsymbol{A}t}$ у випадку , коли матриця $\boldsymbol{A}$ має просту структуру.

ГЛАВА 11

ПСЕВДООБЕРНЕННЯ МАТРИЦЬ І МЕТОД НАЙМЕНШИХ КВАДРАТІВ

Несумісні системи лінійних алгебраїчних рівнянь часто виникають на практиці і тому їх треба вміти розв'язувати. Якщо рівнянь більше ніж невідомих, існує можливість знайти розв'язок з частини рівнянь, проігнорувавши інші рівняння. Проте такий спосіб важко обгрунтувати, якщо всі рівняння мають одне і те ж «походження».

Замість такого розв'язування одних рівнянь і допущення великих похибок в інших доцільно вибрати вектор $\boldsymbol{x}$ так, щоб мінімізувати середню похибку для всіх рівнянь системи $\boldsymbol{Ax} = \boldsymbol{b}$. Існують різні способи для вибору такого усереднення. Найбільш слушний спосіб пов'язаний з уточненням того, що розуміють під розв'язком системи $\boldsymbol{Ax} = \boldsymbol{b}$.

Для довільного вектора $\boldsymbol{x}$ розглянемо вектор $\boldsymbol{Ax} - \boldsymbol{b}$, який називається відхилом вектора $\boldsymbol{x}$. Для того, щоб $\boldsymbol{x}$ був розв'язком системи, необхідно і достатньо, щоб його відхил дорівнював нульовому вектору. У свою чергу , для того, щоб відхил дорівнював нульовому вектору, необхідно і достатньо, щоб дорівнювала нулю його довжина. Таким чином, усі розв'язки системи $\boldsymbol{Ax} = \boldsymbol{b}$ і тільки вони задовольняють рівності $|\boldsymbol{Ax} - \boldsymbol{b}|^2 = 0$.

Оскільки нульові значення довжини відхилу є найменшими, то розв'язання системи можна розглядати як задачу знаходження таких векторів $\boldsymbol{x}$, для яких досягає свого найменшого значення функція $|\boldsymbol{Ax} - \boldsymbol{b}|^2$. Знаходження векторів, які мінімізують цю функцію, має сенс і в тому разі, коли розв'язки системи не існують.

Псевдорозв'язком системи $\boldsymbol{Ax} = \boldsymbol{b}$ називається будь-який вектор $\boldsymbol{x}$, для якого функція $|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2$ досягає свого найменшого значення. Псевдорозв'язок найменшої довжини називається *нормальним* псевдорозв'язком. Метод розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$, пов'язаний з мінімізацією функції $|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2$, називається *методом найменших квадратів*.

Далі буде показано, як можна одержати нормальний псевдорозв'язок, а також дослідити всі розв'язки, одержувані методом найменших квадратів, проте спочатку мова буде йти про так звану псевдообернену матрицю, тісно пов'язану із задачею знаходження псевдорозв'язків системи рівнянь. Застосування псевдообернених матриць дозволяє досягти більшої загальності при побудові моделей багатовимірного статистичного аналізу і кращого розуміння геометричних міркувань, на яких грунтується метод найменших квадратів.

11.1. Псевдообернена матриця

11.1.1. Визначення псевдооберненої матриці. Має місце

Теорема 11.1.1. *Для довільної матриці A існує матриця* $\boldsymbol{A}^+$, *яка визначається формулами*

$$\boldsymbol{A}^+ = \lim_{\varepsilon\to 0}(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}\boldsymbol{A}^T, \tag{1}$$

$$\boldsymbol{A}^+ = \lim_{\varepsilon\to 0}\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E})^{-1}. \tag{2}$$

Д о в е д е н н я. Покажемо, що матриця $\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E}$ є невиродженою. Дійсно, квадратична форма

$$\boldsymbol{x}^T(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})\boldsymbol{x} = |\,\boldsymbol{Ax}\,|^2 + \varepsilon^2\,|\,\boldsymbol{x}\,|^2$$

є додатною і може дорівнювати нулю лише при $\boldsymbol{x} = \boldsymbol{0}$, тобто матриця $\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E}$ є додатно визначеною і тому невиродженою. Аналогічно доводиться невиродженість матриці $\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E}$.

Рівність обох виразів для $\boldsymbol{A}^+$, якщо вони існують, є наслідком очевидної рівності

$$\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E}) = (\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})\boldsymbol{A}^T,$$

яку можна перетворити до вигляду:

$$(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}\boldsymbol{A}^T = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E})^{-1}.$$

Доведемо, що границі у формулах (1) і (2) існують. З теореми 6.5.7 випливає, що

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^T, \tag{3}$$

де $\boldsymbol{Q}$ — ортогональна і Λ — діагональна матриці, тому з формули (1) знаходимо:

$$\boldsymbol{A}^+ = \lim_{\varepsilon\to 0}\boldsymbol{Q}(\Lambda + \varepsilon^2\boldsymbol{E})^{-1}\boldsymbol{Q}^T\boldsymbol{A}^T. \tag{4}$$

Якщо матриця Λ не має нульових діагональних елементів λ_i, то очевидно, що границя в правій частині рівності (4) існує. Якщо ж серед діагональних елементів матриці Λ є нульові елементи, то їм відповідають елементи матриці $(\Lambda + \varepsilon^2\boldsymbol{E})^{-1}$, рівні ε^{-2}, і тому границя в правій частині рівності (4) може існувати лише тоді, коли множник $\boldsymbol{Q}^T\boldsymbol{A}^T$ має відповідні нульові рядки. Але саме такий випадок має місце, бо рівність $\Lambda = \boldsymbol{Q}^T\boldsymbol{A}^T(\boldsymbol{Q}^T\boldsymbol{A}^T)^T$, яка є наслідком формули (3), показує, що коли i-й рядок Λ складається з нулів, то матриця $\boldsymbol{Q}^T\boldsymbol{A}^T$ також має нульовий i-й рядок.

Аналогічно доводиться існування границі в рівності (2). □

Матриця $\boldsymbol{A}^+$, визначена в теоремі 11.1.1, називається *псевдооберненою* матрицею для матриці $\boldsymbol{A}$. Підставою для надання їй такої назви є та обставина, що коли матриця $\boldsymbol{A}$ квадратна і невироджена, то $\boldsymbol{A}^+ = \boldsymbol{A}^{-1}$. З рівностей (1) і (2) випливає, що для $m \times n$-матриці псевдообернена матриця має розміри $n \times m$.

Рівність (1) і формула (5.1.8) показують, що

$$\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T, \tag{5}$$

коли стовпці матриці $\boldsymbol{A}$ лінійно незалежні, а з рівності (2) і формули (5.1.8) випливає, що

$$\boldsymbol{A}^+ = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1} \tag{6}$$

при лінійно незалежних рядках матриці $\boldsymbol{A}$.

Неважко переконатися в тому, що коли матриця $\boldsymbol{A}$ діагональна, то $\boldsymbol{A}^+$ теж є діагональною і її діагональні елементи дорівнюють a_{ii}^{-1}, якщо $a_{ii} \neq 0$, і нулю, якщо $a_{ii} = 0$. Цей результат є узагальненням правила (2.2.3) обернення невиродженої матриці.

Операції псевдообернення і транспонування матриці є переставними. Дійсно, за допомогою рівностей (1) і (2) знаходимо, що

$$(\boldsymbol{A}^T)^+ = \lim_{\varepsilon\to 0}(\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E})^{-1}\boldsymbol{A} = \lim_{\varepsilon\to 0}[\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E})^{-1}]^T = (\boldsymbol{A}^+)^T,$$

тобто маємо рівність

$$(\boldsymbol{A}^T)^+ = (\boldsymbol{A}^+)^T. \tag{7}$$

Зазначимо також, що коли $\boldsymbol{A}$ є нульовою $m \times n$-матрицею, то $\boldsymbol{A}^+$ є нульовою матрицею розмірів $n \times m$.

11.1.2. Властивості псевдооберненої матриці. Вище були розглянуті найпростіші властивості псевдооберненої матриці, які безпосередньо випливають з її визначення. Далі буде йти мова про менш очевидні властивості, при доведенні яких знадобляться рівності

$$\boldsymbol{A}^+\boldsymbol{A} = \lim_{\varepsilon\to 0}[\boldsymbol{E} - \varepsilon^2(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}], \tag{8}$$

$$\boldsymbol{A}\boldsymbol{A}^+ = \lim_{\varepsilon\to 0}[\boldsymbol{E} - \varepsilon^2(\boldsymbol{A}\boldsymbol{A}^T + \varepsilon^2\boldsymbol{E})^{-1}]. \tag{9}$$

Рівність (8) доводится за допомогою рівності (1):

$$\boldsymbol{A}^+\boldsymbol{A} = \lim_{\varepsilon\to 0}(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}\boldsymbol{A}^T\boldsymbol{A} = \lim_{\varepsilon\to 0}(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E} -$$

$$- \varepsilon^2\boldsymbol{E}) = \lim_{\varepsilon\to 0}[\boldsymbol{E} - \varepsilon^2(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}].$$

Аналогічно можна довести рівність (9) за допомогою рівності (2).

Теорема 11.1.2. *Для довільної матриці* $\boldsymbol{A}$ *виконуються рівності*

$$\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T, \quad \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^T. \tag{10}$$

Д о в е д е н н я. З урахуванням рівностей (8) і (1) маємо:

$$\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^T = \lim_{\varepsilon\to 0}[\boldsymbol{A}^T - \varepsilon^2(\boldsymbol{A}^T\boldsymbol{A} + \varepsilon^2\boldsymbol{E})^{-1}\boldsymbol{A}^T] = \boldsymbol{A}^T.$$

Аналогічно доводиться друга рівність (10) за допомогою рівностей (9) і (2). □

Доведемо теорему, яка характеризує псевдообернену матрицю як єдиний розв'язок сукупності матричних рівнянь.

***Теорема 11.1.3* (*теорема Пенроуза*).** *Для довільної матриці* $\boldsymbol{A}$ *матриця* $\boldsymbol{X}$ *дорівнює* $\boldsymbol{A}^{+}$ *тоді і тільки тоді, коли*

$$\boldsymbol{XA} = (\boldsymbol{XA})^{T}, \qquad \boldsymbol{AX} = (\boldsymbol{AX})^{T}, \tag{11}$$

$$\boldsymbol{AXA} = \boldsymbol{A}, \tag{12}$$

$$\boldsymbol{XAX} = \boldsymbol{X}. \tag{13}$$

Д о в е д е н н я. Н е о б х і д н і с т ь. Доведемо, що матриця $\boldsymbol{X} = \boldsymbol{A}^{+}$ задовольняє рівнянням (11) — (13). Інакше кажучи, покажемо, що

$$\boldsymbol{A}^{+}\boldsymbol{A} = (\boldsymbol{A}^{+}\boldsymbol{A})^{T}, \qquad \boldsymbol{AA}^{+} = (\boldsymbol{AA}^{+})^{T}, \tag{14}$$

$$\boldsymbol{AA}^{+}\boldsymbol{A} = \boldsymbol{A}, \tag{15}$$

$$\boldsymbol{A}^{+}\boldsymbol{AA}^{+} = \boldsymbol{A}^{+}. \tag{16}$$

З рівностей (8) і (9) випливає симетричність матриць $\boldsymbol{A}^{+}\boldsymbol{A}$ і $\boldsymbol{AA}^{+}$, тобто рівності (14) виконуються.

Рівність(15) можна довести, скориставшись рівностями (8) і (2):

$$\boldsymbol{AA}^{+}\boldsymbol{A} = \lim_{\varepsilon\to 0}[\boldsymbol{A} - \varepsilon^{2}\boldsymbol{A}(\boldsymbol{A}^{T}\boldsymbol{A} + \varepsilon^{2}\boldsymbol{E})^{-1}] = \boldsymbol{A}.$$

Рівність (16) доводиться за допомогою рівності (1) і другої рівності (10):

$$\boldsymbol{A}^{+}\boldsymbol{AA}^{+} = \lim_{\varepsilon\to 0}(\boldsymbol{A}^{T}\boldsymbol{A} + \varepsilon^{2}\boldsymbol{E})^{-1}\boldsymbol{A}^{T}\boldsymbol{AA}^{+} = \lim_{\varepsilon\to 0}(\boldsymbol{A}^{T}\boldsymbol{A} + \varepsilon^{2}\boldsymbol{E})^{-1}\boldsymbol{A}^{T} = \boldsymbol{A}^{+}.$$

Д о с т а т н і с т ь. Припустивши, що матриця $\boldsymbol{X}$ задовольняє рівнянням (11) — (13), доведемо рівність $\boldsymbol{X} = \boldsymbol{A}^{+}$.

З рівності (12) і першої рівності (11) випливає, що

$$\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{XA})^{T} = \boldsymbol{AA}^{T}\boldsymbol{X}^{T}. \tag{17}$$

Унаслідок рівності (17) і перших рівностей (10) і (11) маємо:

$$\boldsymbol{A}^{+}\boldsymbol{A} = \boldsymbol{A}^{+}\boldsymbol{AA}^{T}\boldsymbol{X}^{T} = \boldsymbol{A}^{T}\boldsymbol{X}^{T} = (\boldsymbol{XA})^{T} = \boldsymbol{XA}. \tag{18}$$

На підставі рівності (13) і другої рівності (11) маємо:

$$\boldsymbol{X}^{T} = (\boldsymbol{XAX})^{T} = (\boldsymbol{AX})^{T}\boldsymbol{X}^{T} = \boldsymbol{AXX}^{T}. \tag{19}$$

Помноживши зліва обидві частини рівності (19) на $\boldsymbol{AA}^+$, одержимо з урахуванням рівності (15):

$$\boldsymbol{AA}^+\boldsymbol{X}^T = \boldsymbol{AA}^+\boldsymbol{AXX}^T = \boldsymbol{AXX}^T. \tag{20}$$

На підставі формул (19) і (20) маємо:

$$\boldsymbol{X}^T = \boldsymbol{AA}^+\boldsymbol{X}^T,$$

звідки з урахуванням другої рівності (14) знаходимо, що

$$\boldsymbol{X} = \boldsymbol{X}(\boldsymbol{AA}^+)^T = \boldsymbol{XAA}^+. \tag{21}$$

Скориставшись формулами (18) і (16), з (21) одержимо:

$$\boldsymbol{X} = \boldsymbol{A}^+\boldsymbol{AA}^+ = \boldsymbol{A}^+. \ \square$$

Теорема Пенроуза дуже корисна при доведенні матричних тотожностей. Наприклад, якщо припустити, що деякий вираз збігається з псевдоoберненою матрицею для даної матриці, то зручний шлях доведення цього припущення полягає в перевірці співвідношень (11) — (13). Зокрема, таким способом неважко довести, що

$$(\boldsymbol{A}^+)^+ = \boldsymbol{A}, \tag{22}$$

$$(\boldsymbol{A}^T\boldsymbol{A})^+ = \boldsymbol{A}^+(\boldsymbol{A}^T)^+, \quad (\boldsymbol{AA}^T)^+ = (\boldsymbol{A}^T)^+\boldsymbol{A}^+, \tag{23}$$

$$(\boldsymbol{Q}_1\boldsymbol{AQ}_2)^+ = \boldsymbol{Q}_2^T\boldsymbol{A}^+\boldsymbol{Q}_1^T, \tag{24}$$

де $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2$ — ортогональні матриці.

Тепер ми доведемо теорему, яка показує, що псевдообернення довільної матриці можна звести до псевдообернення симетричної матриці.

Теорема 11.1.4. *Для довільної матриці* $\boldsymbol{A}$ *виконуються рівності*

$$\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{A}^T, \quad \boldsymbol{A}^+ = \boldsymbol{A}^T(\boldsymbol{AA}^T)^+. \tag{25}$$

Д о в е д е н н я. Скориставшись першою рівністю (23), другою рівністю (10), а потім першою рівністю (10) і рівністю (16), одержимо:

$$(\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{A}^T = \boldsymbol{A}^+(\boldsymbol{A}^T)^+\boldsymbol{A}^T = \boldsymbol{A}^+((\boldsymbol{A}^T)^+\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{A}^+ = \boldsymbol{A}^+\boldsymbol{AA}^+ = \boldsymbol{A}^+.$$

Отже, перша рівність (25) виконується. Аналогічно можна довести другу рівність (25). $\square$

З'ясуємо залежності між рангами матриць $\boldsymbol{A}$, $\boldsymbol{A}^+$, $\boldsymbol{A}^+\boldsymbol{A}$, і $\boldsymbol{A}\boldsymbol{A}^+$.

Теорема 11.1.5. *Для будь-якої матриці* $\boldsymbol{A}$ *виконуються рівності*

$$r(\boldsymbol{A}) = r(\boldsymbol{A}^+) = r(\boldsymbol{A}^+\boldsymbol{A}) = r(\boldsymbol{A}\boldsymbol{A}^+). \tag{26}$$

Д о в е д е н н я. З теореми 3.4.4 випливає, що

$$r(\boldsymbol{A}^+\boldsymbol{A}) \le r(\boldsymbol{A}). \tag{27}$$

З іншого боку, на підставі рівності (15) і тієї ж теореми маємо:

$$r(\boldsymbol{A}) = r(\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A}) \le r(\boldsymbol{A}^+\boldsymbol{A}). \tag{28}$$

Співвідношення (27) і (28) доводять, що

$$r(\boldsymbol{A}) = r(\boldsymbol{A}^+\boldsymbol{A}).$$

Аналогічно можна довести рівність $r(\boldsymbol{A}) = r(\boldsymbol{A}\boldsymbol{A}^+)$. Після цього залишається показати, що $r(\boldsymbol{A}^+) = r(\boldsymbol{A}^+\boldsymbol{A})$ (або $r(\boldsymbol{A}^+) = r(\boldsymbol{A}\boldsymbol{A}^+)$) і це можна зробити за допомогою рівності (16). □

Наприкінці розглянемо умови невід'ємної визначеності блокових матриць. З математичної статистики відомо, що невід'ємно визначеними є коваріаційні матриці, тому наступна теорема разом з теоремою 7.1.5 являє інтерес при з'ясуванні властивостей блоків цих матриць.

Теорема 11.1.6. *Нехай* $\boldsymbol{A}$ *— квадратна матриця*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^T & \boldsymbol{A}_{22} \end{pmatrix},$$

де $\boldsymbol{A}_{11}$ *і* $\boldsymbol{A}_{22}$ *— симетричні підматриці. Для невід'ємної визначеності матриці* $\boldsymbol{A}$ *необхідно і достатньо, щоб*:

1) *матриці* $\boldsymbol{A}_{11}$ *і* $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^+\boldsymbol{A}_{12}$ *були невід'ємно визначеними*;

2) $\boldsymbol{A}_{11}\boldsymbol{A}_{11}^+\boldsymbol{A}_{12} = \boldsymbol{A}_{12}$.

Д о в е д е н н я. Н е о б х і д н і с т ь. Якщо $\boldsymbol{A}$ — невід'ємно визначена матриця, то $\boldsymbol{A} = \boldsymbol{W}^T\boldsymbol{W}$. За допомогою блокового представлення $\boldsymbol{W} = (\boldsymbol{X} \quad \boldsymbol{Y})$ матрицю $\boldsymbol{A}$ можна подати у вигляді:

$$\boldsymbol{A} = (\boldsymbol{X} \quad \boldsymbol{Y})^T(\boldsymbol{X} \quad \boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{X}^T\boldsymbol{X} & \boldsymbol{X}^T\boldsymbol{Y} \\ \boldsymbol{Y}^T\boldsymbol{X} & \boldsymbol{Y}^T\boldsymbol{Y} \end{pmatrix}.$$

Звідси випливає, що $\boldsymbol{A}_{11} = \boldsymbol{X}^T\boldsymbol{X}$ є невід'ємно визначеною матрицею і $\boldsymbol{A}_{12} = \boldsymbol{X}^T\boldsymbol{Y}$. Крім того,

$$\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{+} = \boldsymbol{X}^T\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{+} = \boldsymbol{X}^T\boldsymbol{X}\boldsymbol{X}^{+}(\boldsymbol{X}^T)^{+} = \boldsymbol{X}^T(\boldsymbol{X}^T)^{+} = (\boldsymbol{X}^{+}\boldsymbol{X})^T = \boldsymbol{X}^{+}\boldsymbol{X}$$

і тому

$$(\boldsymbol{A}_{11}\boldsymbol{A}_{11}^{+})\boldsymbol{A}_{12} = \boldsymbol{X}^{+}\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{Y} = \boldsymbol{A}_{12}\,.$$

Для матриці $\boldsymbol{U} = \boldsymbol{Y} - \boldsymbol{X}\boldsymbol{A}_{11}^{+}\boldsymbol{A}_{12}$ виконується рівність

$$\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^{+}\boldsymbol{A}_{12}\,,$$

тобто матриця $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^{+}\boldsymbol{A}_{12}$ є невід'ємно визначеною.

Д о с т а т н і с т ь. Якщо матриці $\boldsymbol{A}_{11}$ і $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^{+}\boldsymbol{A}_{12}$ є невід'ємно визначеними, то існують квадратні корені з цих матриць. Нехай

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{0} \end{pmatrix}\boldsymbol{A}_{11}^{1/2}, \qquad \boldsymbol{Y} = \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{0} \end{pmatrix}\boldsymbol{A}_{11}^{1/2}\boldsymbol{A}_{11}^{+}\boldsymbol{A}_{12} + \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{E} \end{pmatrix}(\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^T\boldsymbol{A}_{11}^{+}\boldsymbol{A}_{12})^{1/2}.$$

Тоді

$$(\boldsymbol{X} \quad \boldsymbol{Y})^T(\boldsymbol{X} \quad \boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^T & \boldsymbol{A}_{22} \end{pmatrix} = \boldsymbol{A}\,,$$

тобто матриця $\boldsymbol{A}$ — невід'ємно визначена. ☐

Аналогічно можна дослідити випадок, коли невід'ємно визначеними є матриці $\boldsymbol{A}_{22}$ і $\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{+}\boldsymbol{A}_{12}^T$.

11.2. Псевдообернення добутку матриць

11.2.1. Зв'язок між $(\boldsymbol{AB})^{+}$ і $\boldsymbol{B}^{+}\boldsymbol{A}^{+}$. Для квадратних невироджених матриць $\boldsymbol{A}$ і $\boldsymbol{B}$ має місце рівність $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. Вище ми одержали для окремих випадків рівності (11.1.23) і (11.1.24) аналогічного типу для псевдообернених матриць. Проте, взагалі кажучи, рівність $(\boldsymbol{AB})^{+} = \boldsymbol{B}^{+}\boldsymbol{A}^{+}$ не виконується, в чому неважко переконатися на такому прикладі:

$$\boldsymbol{A} = (1 \quad 1), \qquad \boldsymbol{B} = (1 \quad 0)^T, \qquad (\boldsymbol{AB})^{+} = (1), \qquad \boldsymbol{B}^{+}\boldsymbol{A}^{+} = (1/2).$$

Питання про псевдообернення добутку матриць певною мірою розв'язується наступною теоремою.

Теорема 11.2.1. *Рівність*

$$(\boldsymbol{AB})^{+} = \boldsymbol{B}^{+}\boldsymbol{A}^{+} \tag{1}$$

має місце тоді і тільки тоді, коли виконуються умови

$$\boldsymbol{A}^{+}\boldsymbol{AB}(\boldsymbol{AB})^{T} = \boldsymbol{B}(\boldsymbol{AB})^{T}, \tag{2}$$

$$(\boldsymbol{AB})^{T}\boldsymbol{ABB}^{+} = (\boldsymbol{AB})^{T}\boldsymbol{A}. \tag{3}$$

Д о в е д е н н я. Обмежимося доведенням достатності, скориставшись теоремою Пенроуза. Для цього покажемо, що матриця $\boldsymbol{X} = \boldsymbol{B}^{+}\boldsymbol{A}^{+}$ задовольняє рівнянням

$$\boldsymbol{XAB} = (\boldsymbol{XAB})^{T}, \quad \boldsymbol{ABX} = (\boldsymbol{ABX})^{T}, \tag{4}$$

$$\boldsymbol{ABXAB} = \boldsymbol{AB}, \tag{5}$$

$$\boldsymbol{XABX} = \boldsymbol{X}. \tag{6}$$

Помножимо зліва обидві частини рівності (2) на $\boldsymbol{B}^{+}$. На підставі першої рівності (11.1.10) будемо мати:

$$\boldsymbol{B}^{+}\boldsymbol{A}^{+}\boldsymbol{AB}(\boldsymbol{AB})^{T} = (\boldsymbol{AB})^{T}. \tag{7}$$

Аналогічно за допомогою рівності (3) і другої рівності (11.1.10) знаходимо:

$$(\boldsymbol{AB})^{T}\boldsymbol{ABB}^{+}\boldsymbol{A}^{+} = (\boldsymbol{AB})^{T}. \tag{8}$$

Помножимо справа обидві частини рівності (7) на $((\boldsymbol{AB})^{T})^{+}$ і скористаємося другою формулою (11.1.10) при перетворенні лівої частини одержаної рівності і формулою (11.1.7) та першою формулою (1.1.14) при спрощенні її правої частини. У результаті одержимо, що

$$\boldsymbol{B}^{+}\boldsymbol{A}^{+}\boldsymbol{AB} = (\boldsymbol{AB})^{+}\boldsymbol{AB}. \tag{9}$$

Помноживши зліва обидві частини рівності (8) на $((\boldsymbol{AB})^{T})^{+}$, подібно до формули (9) одержимо:

$$\boldsymbol{ABB}^{+}\boldsymbol{A}^{+} = \boldsymbol{AB}(\boldsymbol{AB})^{+}. \qquad (10)$$

Праві частини рівностей (9) та (10) є симетричними матрицями, тому матриця $\boldsymbol{B}^{+}\boldsymbol{A}^{+}$ задовольняє рівнянням (4). Якщо обидві частини рівності (9) помножити зліва на $\boldsymbol{AB}$, то з урахуванням формули (11.1.15) одержимо рівність

$$\boldsymbol{ABB}^{+}\boldsymbol{A}^{+}\boldsymbol{AB} = \boldsymbol{AB}, \qquad (11)$$

яка означає, що $\boldsymbol{B}^{+}\boldsymbol{A}^{+}$ задовольняє рівнянню (5).

Нам залишається довести, що матриця $\boldsymbol{B}^{+}\boldsymbol{A}^{+}$ задовольняє рівнянню (6). Ця частина доведення є найбільш складною і вона пов'язана з розглядом матричного рівняння

$$\boldsymbol{B}^{+}\boldsymbol{A}^{+}\boldsymbol{ABY} = \boldsymbol{B}^{+}\boldsymbol{A}^{+}. \qquad (12)$$

Покажемо, що матриця $\boldsymbol{Y}$ існує, тобто рівняння (12) є розв'язуваним.

На підставі міркувань, використаних у розділі 4.3 при визначенні вимірності образу матриці, можна стверджувати, що коли $r(\boldsymbol{G}) = r(\boldsymbol{GH})$, то $\mathscr{R}(\boldsymbol{G}) = \mathscr{R}(\boldsymbol{GH})$ і тому рівняння $\boldsymbol{GHY} = \boldsymbol{G}$ має розв'язок. У нашому випадку $\boldsymbol{G} = \boldsymbol{B}^{+}\boldsymbol{A}^{+}$ і $\boldsymbol{H} = \boldsymbol{AB}$. Тому треба довести рівність

$$r(\boldsymbol{B}^{+}\boldsymbol{A}^{+}) = r(\boldsymbol{B}^{+}\boldsymbol{A}^{+}\boldsymbol{AB}). \qquad (13)$$

За допомогою рівностей (11.1.16) і (11.1.14) знаходимо, що

$$\boldsymbol{B}^{+}\boldsymbol{A}^{+} = (\boldsymbol{B}^{+}\boldsymbol{BB}^{+})(\boldsymbol{A}^{+}\boldsymbol{AA}^{+}) = \boldsymbol{B}^{+}(\boldsymbol{B}^{+})^{T}\boldsymbol{B}^{T}\boldsymbol{A}^{T}(\boldsymbol{A}^{+})^{T}\boldsymbol{A}^{+},$$

звідки на підставі теореми 3.4.4 маємо:

$$r(\boldsymbol{B}^{+}\boldsymbol{A}^{+}) \le r(\boldsymbol{B}^{T}\boldsymbol{A}^{T}) = r(\boldsymbol{AB}). \qquad (14)$$

З іншого боку, за допомогою рівностей (11.1.26) і (9) та теореми 3.4.4 знаходимо, що

$$r(\boldsymbol{AB}) = r((\boldsymbol{AB})^{+}\boldsymbol{AB}) = r(\boldsymbol{B}^{+}\boldsymbol{A}^{+}\boldsymbol{AB}) \le r(\boldsymbol{B}^{+}\boldsymbol{A}^{+}). \qquad (15)$$

З формул (14) і (15) випливає рівність (13).

Отже, матричне рівняння (12) є розв'язуваним. Помноживши зліва обидві частини рівняння (12) на $\boldsymbol{AB}$, за допомогою рівності (11) одержимо:

$$\boldsymbol{ABY} = \boldsymbol{ABB}^{+}\boldsymbol{A}^{+}. \tag{16}$$

З рівнянь (12) і (16) знаходимо, що

$$\boldsymbol{B}^{+}\boldsymbol{A}^{+}\boldsymbol{ABB}^{+}\boldsymbol{A}^{+} = \boldsymbol{B}^{+}\boldsymbol{A}^{+},$$

тобто матриця $\boldsymbol{B}^{+}\boldsymbol{A}^{+}$ задовольняє рівнянню (6).

Висновок. *Якщо* $\boldsymbol{AB} = \boldsymbol{0}$, *то* $\boldsymbol{B}^{+}\boldsymbol{A}^{+} = \boldsymbol{0}$.

Д о в е д е н н я. При $\boldsymbol{AB} = \boldsymbol{0}$ умови (2) і (3) виконані. □

11.2.2. Обчислення псевдооберненої матриці в окремих випадках. Теорема 11.2.1 дає можливість обчислити псевдообернену матрицю в деяких важливих випадках. Зокрема, неважко показати, що

$$(\boldsymbol{A}^{+}\boldsymbol{A})^{+} = \boldsymbol{A}^{+}\boldsymbol{A}, \quad (\boldsymbol{AA}^{+})^{+} = \boldsymbol{AA}^{+}, \tag{17}$$

$$(\boldsymbol{AQ})^{+} = \boldsymbol{Q}^{T}\boldsymbol{A}^{+}, \quad (\boldsymbol{QA})^{+} = \boldsymbol{A}^{+}\boldsymbol{Q}^{T}, \tag{18}$$

де $\boldsymbol{Q}$ — ортогональна матриця. Зазначимо принагідно, що формули (18) можна також одержати з рівності (11.1.24).

За допомогою рівностей (11.1.5) і (11.1.6) легко довести, що для скелетного розкладу $\boldsymbol{A} = \boldsymbol{BC}$ виконуються умови теореми 11.2.1, тому

$$\boldsymbol{A}^{+} = \boldsymbol{C}^{T}(\boldsymbol{CC}^{T})^{-1}(\boldsymbol{B}^{T}\boldsymbol{B})^{-1}\boldsymbol{B}^{T} = \boldsymbol{C}^{T}(\boldsymbol{B}^{T}\boldsymbol{AC}^{T})^{-1}\boldsymbol{B}^{T}. \tag{19}$$

Якщо $m \times n$-матриця $\boldsymbol{A}$ має ранг $r = \min(m,\ n)$, то її псевдообернення спрощується. Нехай $r = m$, тоді, як зазначалося в розділі 3.4, скелетний розклад представляється у вигляді добутку $\boldsymbol{A} = \boldsymbol{EA}$. При цьому рівність (19) переходить у формулу (11.1.6). Якщо $r = n$, то скелетний розклад визначається рівністю $\boldsymbol{A} = \boldsymbol{AE}$ і тоді (19) записується у вигляді формули (11.1.5).

П р и к л а д. Знайдемо $\boldsymbol{A}^{+}$ для матриці

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

Перетворивши матрицю $\boldsymbol{A}$ до ступінчастого вигляду, одержимо:

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

Оскільки $r(\boldsymbol{A}) = 2 < 3$, то для псевдообернення треба мати скелетний розклад. Він є таким:

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

Скористаємося рівністю (19). Послідовно знаходимо:

$$\boldsymbol{B}^T \boldsymbol{A} \boldsymbol{C}^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix},$$

$$(\boldsymbol{B}^T \boldsymbol{A} \boldsymbol{C}^T)^{-1} = \frac{1}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

$$\boldsymbol{A}^+ = \boldsymbol{C}^T (\boldsymbol{B}^T \boldsymbol{A} \boldsymbol{C}^T)^{-1} \boldsymbol{B}^T =$$

$$= \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \\ -1 & -2 & 1 \end{pmatrix}. \triangle$$

Розглянемо більш складний приклад псевдообернення добутку. Доведемо, що

$$(\boldsymbol{V}^{-1} \boldsymbol{H})^+ = \boldsymbol{H}^+ \boldsymbol{V} (\boldsymbol{E} - (\boldsymbol{\Gamma} \boldsymbol{V})^+ \boldsymbol{V}), \tag{20}$$

де $\boldsymbol{H}$ і $\boldsymbol{V}$ — довільна прямокутна та невироджена матриці відповідно і $\boldsymbol{\Gamma} = \boldsymbol{E} - \boldsymbol{H}\boldsymbol{H}^+$. Для цього введемо позначення:

$$\boldsymbol{A} = \boldsymbol{V}^{-1} \boldsymbol{H}, \quad \boldsymbol{X} = \boldsymbol{H}^+ \boldsymbol{V} (\boldsymbol{E} - (\boldsymbol{\Gamma} \boldsymbol{V})^+ \boldsymbol{V})$$

і перевіримо виконання умов теореми Пенроуза.

Попередньо зауважимо, що з очевидної рівності $\boldsymbol{H}^{+}\boldsymbol{\Gamma}\boldsymbol{V} = \boldsymbol{0}$ на підставі висновку з теореми 11.2.1 маємо:

$$(\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{H} = \boldsymbol{0}. \tag{21}$$

Скориставшись формулою (21), одержимо, що

$$\begin{aligned}(\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{V} &= (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{E}\boldsymbol{V} = (\boldsymbol{\Gamma}\boldsymbol{V})^{+}(\boldsymbol{\Gamma} + \boldsymbol{H}\boldsymbol{H}^{+})\boldsymbol{V} = \\ &= (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{\Gamma}\boldsymbol{V} + (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{H}\boldsymbol{H}^{+}\boldsymbol{V} = (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{\Gamma}\boldsymbol{V}.\end{aligned} \tag{22}$$

Повертаючись до рівності (20), обчислимо $\boldsymbol{XA}$ і $\boldsymbol{AX}$. Скориставшись формулою (21), будемо мати:

$$\boldsymbol{XA} = \boldsymbol{H}^{+}\boldsymbol{V}(\boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{V})\boldsymbol{V}^{-1}\boldsymbol{H} = \boldsymbol{H}^{+}(\boldsymbol{E} - \boldsymbol{V}(\boldsymbol{\Gamma}\boldsymbol{V})^{+})\boldsymbol{H} = \boldsymbol{H}^{+}\boldsymbol{H}. \tag{23}$$

Далі з урахуванням формул (22) і (11.1.15) знаходимо:

$$\boldsymbol{AX} = \boldsymbol{V}^{-1}\boldsymbol{H}\boldsymbol{H}^{+}\boldsymbol{V}(\boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{V}) = \boldsymbol{V}^{-1}(\boldsymbol{E} - \boldsymbol{\Gamma})\boldsymbol{V}(\boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{\Gamma}\boldsymbol{V}) =$$

$$= \boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{\Gamma}\boldsymbol{V} - \boldsymbol{V}^{-1}\boldsymbol{\Gamma}\boldsymbol{V} + \boldsymbol{V}^{-1}\boldsymbol{\Gamma}\boldsymbol{V}(\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{\Gamma}\boldsymbol{V} = \boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{\Gamma}\boldsymbol{V}.$$

Очевидно, що матриці $\boldsymbol{XA}$ і $\boldsymbol{AX}$ симетричні, тобто матриця $\boldsymbol{X}$ задовольняє умовам (11.1.11). Обчислимо $\boldsymbol{AXA}$ і $\boldsymbol{XAX}$, скориставшись формулою (23):

$$\boldsymbol{AXA} = \boldsymbol{A}\boldsymbol{H}^{+}\boldsymbol{H} = \boldsymbol{V}^{-1}\boldsymbol{H}\boldsymbol{H}^{+}\boldsymbol{H} = \boldsymbol{V}^{-1}\boldsymbol{H} = \boldsymbol{A},$$

$$\boldsymbol{XAX} = \boldsymbol{H}^{+}\boldsymbol{H}\boldsymbol{H}^{+}\boldsymbol{V}(\boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{V}) = \boldsymbol{H}^{+}\boldsymbol{V}(\boldsymbol{E} - (\boldsymbol{\Gamma}\boldsymbol{V})^{+}\boldsymbol{V}) = \boldsymbol{X}.$$

Отже, $\boldsymbol{X}$ задовольняє рівнянням (11.1.12) і (11.1.13), тому можна зробити висновок про справедливість рівності (20).

11.3. Проекційні матриці

11.3.1. Проекції вектора на образ і ядро матриці. Для визначення нормального псевдорозв'язку і одержання загальних результатів при користуванні методом найменших квадратів нам будуть потрібні властивості псевдооберненої матриці, пов'язані з ортогональним проектуванням на образ і ядро матриці. Далі для стислості викладення ми будемо називати ортогональну проекцію просто проекцією.

Теорема 11.3.1. *Для будь-якого вектора* $\boldsymbol{x}$ *вектор* $\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x}$ *є проекцією* $\boldsymbol{x}$ *на* $\mathcal{R}(\boldsymbol{A}^{T})$, *а вектор* $(\boldsymbol{E}-\boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{x}$ *є проекцією* $\boldsymbol{x}$ *на* $\mathcal{N}(\boldsymbol{A})$. *Для будь-якого вектора* $\boldsymbol{x}$ *вектор* $\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{x}$ *— проекція* $\boldsymbol{x}$ *на* $\mathcal{R}(\boldsymbol{A})$, *а* $(\boldsymbol{E}-\boldsymbol{A}\boldsymbol{A}^{+})\boldsymbol{x}$ *— проекція* $\boldsymbol{x}$ *на* $\mathcal{N}(\boldsymbol{A}^{T})$.

Д о в е д е н н я. З формули (11.1.1) випливає, що для будь-якого вектора $\boldsymbol{x}$ має місце рівність

$$\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x}=\lim_{\varepsilon\to 0}(\boldsymbol{A}^{T}\boldsymbol{A}+\varepsilon^{2}\boldsymbol{E})^{-1}\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{x}.$$

У відповідності до теореми 5.1.4 з урахуванням симетричності матриці $\boldsymbol{A}^{T}\boldsymbol{A}$ маємо:

$$\boldsymbol{x}=\boldsymbol{u}+\boldsymbol{v},\quad \boldsymbol{u}\in\mathcal{R}(\boldsymbol{A}^{T}\boldsymbol{A}),\quad \boldsymbol{v}\in\mathcal{N}(\boldsymbol{A}^{T}\boldsymbol{A}),$$

причому $\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{u}$. Оскільки $\boldsymbol{u}\in\mathcal{R}(\boldsymbol{A}^{T}\boldsymbol{A})$, то $\boldsymbol{u}=\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{y}$ для деякого вектора $\boldsymbol{y}$ і тому з урахуванням рівності (11.1.3) маємо:

$$(\boldsymbol{A}^{T}\boldsymbol{A}+\varepsilon^{2}\boldsymbol{E})^{-1}\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{A}^{T}\boldsymbol{A}+\varepsilon^{2}\boldsymbol{E})^{-1}(\boldsymbol{A}^{T}\boldsymbol{A})^{2}\boldsymbol{y}=\boldsymbol{Q}(\Lambda+\varepsilon^{2}\boldsymbol{E})^{-1}\Lambda^{2}\boldsymbol{Q}^{T}\boldsymbol{y},$$

звідки після граничного переходу над кожним елементом матриці випливає, що

$$\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x}=\lim_{\varepsilon\to 0}(\boldsymbol{A}^{T}\boldsymbol{A}+\varepsilon^{2}\boldsymbol{E})^{-1}\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{Q}\Lambda\boldsymbol{Q}^{T}\boldsymbol{y}=\boldsymbol{A}^{T}\boldsymbol{A}\boldsymbol{y}=\boldsymbol{u}.$$

Це означає на підставі другої формули (5.1.7), що $\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x}$ є проекцією $\boldsymbol{x}$ на $\mathcal{R}(\boldsymbol{A}^{T})$, що і треба довести.

Якщо $\boldsymbol{u}$ — проекція $\boldsymbol{x}$ на $\mathcal{R}(\boldsymbol{A}^{T})$, то на підставі теореми 5.1.4 вектор $\boldsymbol{x}-\boldsymbol{u}=(\boldsymbol{E}-\boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{x}$ є проекцією $\boldsymbol{x}$ на $\mathcal{N}(\boldsymbol{A})$.

Друга частина теореми одержується з першої шляхом заміни $\boldsymbol{A}$ на $\boldsymbol{A}^{T}$. Це випливає з рівності

$$\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{x}=(\boldsymbol{A}^{T})^{+}\boldsymbol{A}^{T}\boldsymbol{x},$$

яка доводиться за допомогою другої формули (11.1.14).

Висновок 1. *Якщо* A *— довільна ненульова матриця, то* $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ *тоді і тільки тоді, коли* $\boldsymbol{x}=(\boldsymbol{E}-\boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{y}$ *для деякого вектора* $\boldsymbol{y}$.

Д о в е д е н н я. Оскільки $\boldsymbol{x} \in \mathcal{N}(\boldsymbol{A})$, то $\boldsymbol{x}$ можна представити як проекцію деякого вектора $\boldsymbol{y}$ на $\mathcal{N}(\boldsymbol{A})$, тобто $\boldsymbol{x} = (\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{y}$.

З іншого боку, $\boldsymbol{Ax} = \boldsymbol{A}(\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{y} = \boldsymbol{0}$.

Висновок 2. *Якщо A — довільна матриця, то* $\boldsymbol{x} \in \mathcal{R}(\boldsymbol{A})$ *тоді і тільки тоді, коли знайдеться вектор* $\boldsymbol{u}$ *такий, що* $\boldsymbol{x} = \boldsymbol{AA}^{+}\boldsymbol{u}$.

Д о в е д е н н я. Оскільки $\boldsymbol{x} \in \mathcal{R}(\boldsymbol{A})$, то $\boldsymbol{x}$ можна представити у вигляді проекції деякого вектора $\boldsymbol{u}$ на $\mathcal{R}(\boldsymbol{A})$, тобто $\boldsymbol{x} = \boldsymbol{AA}^{+}\boldsymbol{u}$.

З іншого боку, рівність $\boldsymbol{x} = \boldsymbol{AA}^{+}\boldsymbol{u}$ показує, що $\boldsymbol{x} \in \mathcal{R}(\boldsymbol{A})$. □

П р и к л а д. Повернемося до задачі, розглянутої в прикладі 1 з розділу 5.2. Там обчислювалася проекція вектора $\boldsymbol{a}$ на підпростір, породжений векторами $\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \boldsymbol{u}_3$. Можна вважати, що мова йде про проекцію $\boldsymbol{a}$ на множину стовпців матриці $\boldsymbol{A} = (\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{u}_3)$ або на множину рядків матриці $\boldsymbol{A}^{T}$. Для обчислення цієї проекції ми можемо скористатися теоремою 11.3.1. Оскільки в прикладі з розділу 11.2 ми виконали псевдообернення матриці, рядки якої складаються з координат векторів $\boldsymbol{u}_1,\ \boldsymbol{u}_2,\ \boldsymbol{u}_3$, то можна скористатися одержаним результатом і шукати проекцію вектора $\boldsymbol{a}$ на вказаний підпростір як проекцію $\boldsymbol{a}$ на $\mathcal{R}(\boldsymbol{A}^{T})$. У відповідності до теореми 11.3.1 шукана проекція дорівнює

$$\boldsymbol{u} = \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{a} = \frac{1}{6}\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \\ -1 & -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ -2 \end{pmatrix},$$

тобто маємо той же результат, що і в прикладі 1 з розділу 5.2. △

11.3.2. Властивості проекційних матриць. Матриця, яка не змінюється при множенні на саму себе, називається *ідемпотентною*. Симетрична ідемпотентна матриця називається *проекційною*. Неважко переконатися за допомогою формул (11.1.14) і (11.1.16) в тому, що матриці $\boldsymbol{A}^{+}\boldsymbol{A}$, $\boldsymbol{AA}^{+}$, $\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A}$, $\boldsymbol{E} - \boldsymbol{AA}^{+}$, які зустрічалися вище при розгляді проекцій вектора на образ і ядро матриці, є проекційними. Розглянемо докладніше властивості проекційних матриць. Спочатку буде доведена допоміжна

Теорема 11.3.2. *Симетрична матриця* $\boldsymbol{P}$ *порядку n є ідемпотентною матрицею рангу r тоді і тільки тоді, коли r її власних значень дорівнюють одиниці, а інші* $n - r$ *власні значення дорівнюють нулю.*

Д о в е д е н н я. Н е о б х і д н і с т ь. Оскільки $\boldsymbol{P}^2 = \boldsymbol{P}$, то з рівності $\boldsymbol{Px} = \lambda\boldsymbol{x}$ випливає, що

$$\lambda\boldsymbol{x}^T\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{Px} = \boldsymbol{x}^T\boldsymbol{P}^2\boldsymbol{x} = (\boldsymbol{Px})^T\boldsymbol{Px} = \lambda^2\boldsymbol{x}^T\boldsymbol{x}$$

і тому $\lambda(\lambda - 1) = 0$.

Отже, власні значення матриці $\boldsymbol{P}$ дорівнюють або одиниці, або нулю. На підставі теореми 6.5.7. і висновку з теореми 3.4.4 r дорівнює кількості ненульових власних значень матриці $\boldsymbol{P}$, тобто r власних значень дорівнюють одиниці і $(n - r)$ — нулю.

Д о с т а т н і с т ь. Без обмеження загальності можна вважати, що одиничними є перші r власних значень матриці. Тому на підставі теореми 6.5.7 існує ортогональна матриця $\boldsymbol{Q}$ така, що

$$\boldsymbol{P} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^T, \qquad \Lambda = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Отже, $r(\boldsymbol{P}) = r$ і

$$\boldsymbol{P}^2 = \boldsymbol{Q}\Lambda^2\boldsymbol{Q}^T = \boldsymbol{Q}\Lambda\boldsymbol{Q}^T = \boldsymbol{P}. \ \square$$

Теорема 11.3.3. *Проекційна матриця має такі властивості*:

1) *слід проекційної матриці дорівнює її рангу*;

2) *проекційна матриця є невід'ємно визначеною*;

3) *для проекційної матриці* $\boldsymbol{P}$ *виконується рівність*

$$\boldsymbol{P}^+ = \boldsymbol{P}; \tag{1}$$

4) *для проекційної матриці* $\boldsymbol{P}$ *і довільної матриці* $\boldsymbol{A}$ *виконуються рівності*

$$(\boldsymbol{PA})^+ = (\boldsymbol{PA})^+\boldsymbol{P}, \qquad (\boldsymbol{AP})^+ = \boldsymbol{P}(\boldsymbol{AP})^+. \tag{2}$$

Д о в е д е н н я. 1) Формула (6.2.2) показує, що слід матриці дорівнює сумі її власних значень. Ця сума на підставі теореми 11.3.2 дорівнює $r(\boldsymbol{P})$. Отже, $\operatorname{tr}\boldsymbol{P} = r(\boldsymbol{P})$.

2) Очевидно, що

$$\boldsymbol{x}^T \boldsymbol{P}\boldsymbol{x} = \boldsymbol{x}^T \boldsymbol{P}^2 \boldsymbol{x} = (\boldsymbol{P}\boldsymbol{x})^T \boldsymbol{P}\boldsymbol{x} \geq 0,$$

тобто матриця $\boldsymbol{P}$ є невід'ємно визначеною.

3) Рівність (1) легко доводиться за допомогою теореми Пенроуза.

4) Для доведення першої формули (2) представимо матрицю $\boldsymbol{PA}$ у вигляді добутку $\boldsymbol{P}(\boldsymbol{PA})$ і застосуємо теорему 11.2.1 і рівність (1):

$$(\boldsymbol{PA})^+ = (\boldsymbol{P}(\boldsymbol{PA}))^+ = (\boldsymbol{PA})^+ \boldsymbol{P}^+ = (\boldsymbol{PA})^+ \boldsymbol{P}.$$

Аналогічно доводиться друга рівність (2). □

Наприкінці покажемо, що для $m \times n$ -матриці $\boldsymbol{A}$ мають місце співвідношення:

$$r(\boldsymbol{E} - \boldsymbol{A}\boldsymbol{A}^+) = m - r(\boldsymbol{A}), \quad r(\boldsymbol{E} - \boldsymbol{A}^+\boldsymbol{A}) = n - r(\boldsymbol{A}). \tag{3}$$

Дійсно, на підставі частини 1) теореми 11.3.3 і формули (11.1.26) маємо:

$$r(\boldsymbol{E} - \boldsymbol{A}\boldsymbol{A}^+) = \operatorname{tr} \boldsymbol{E} - \operatorname{tr}(\boldsymbol{A}\boldsymbol{A}^+) = m - r(\boldsymbol{A}\boldsymbol{A}^+) = m - r(\boldsymbol{A}).$$

Аналогічно доводиться друга рівність (3).

11.4. Метод найменших квадратів

11.4.1. Нормальний псевдорозв'язок. Нагадаємо, що нормальним псевдорозв'язком системи $\boldsymbol{Ax} = \boldsymbol{b}$ називається вектор з мінімальною довжиною серед усіх векторів, що мінімізують функцію $|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2$. Питання про нормальний псевдорозв'язок вирішується наступною теоремою.

Теорема 11.4.1. *Нормальний псевдорозв'язок системи* $\boldsymbol{Ax} = \boldsymbol{b}$ *дорівнює*

$$\boldsymbol{x}_0 = \boldsymbol{A}^+ \boldsymbol{b}. \tag{1}$$

Д о в е д е н н я. У відповідності до методу найменших квадратів ми шукаємо деякий вектор $\boldsymbol{x}_0$, що задовольняє нерівності

$$|\boldsymbol{Ax}_0 - \boldsymbol{b}|^2 \le |\boldsymbol{Ax} - \boldsymbol{b}|^2 \tag{2}$$

для будь-якого вектора $\boldsymbol{x}$. Інакше кажучи, треба знайти вектор $\boldsymbol{Ax}_0 \in \mathcal{R}(\boldsymbol{A})$ такий, щоб відстань між ним і $\boldsymbol{b}$ була мінімальною. Як зазначалося в розділі 5.2, це буде тоді, коли $\boldsymbol{Ax}_0$ є проекцією $\boldsymbol{b}$ на $\mathcal{R}(\boldsymbol{A})$. На підставі теореми 11.3.1 вказана проекція дорівнює $\boldsymbol{AA}^+\boldsymbol{b}$. Отже,

$$\boldsymbol{Ax}_0 = \boldsymbol{AA}^+\boldsymbol{b}. \tag{3}$$

Зіставлення формул (1) і (3) свідчить про те, що за вектор $\boldsymbol{x}_0$, про який ідеться вище в доведенні теореми, можна взяти нормальний псевдорозв'язок.

Покажемо, що $\boldsymbol{x}_0$ має найменшу можливу довжину серед векторів, що мінімізують функцію $|\boldsymbol{Ax} - \boldsymbol{b}|^2$. Очевидно, що треба порівнювати $\boldsymbol{x}_0$ з тими векторами $\boldsymbol{x}$, для яких нерівність (2) виконується як рівність. Ця умова, як буде показано нижче, рівносильна виконанню рівності

$$\boldsymbol{Ax} = \boldsymbol{Ax}_0. \tag{4}$$

Наведемо міркування, які дозволяють обґрунтувати умову (4).

На підставі теореми про ортогональний розклад можна вважати, що $\boldsymbol{b} = \boldsymbol{u} + \boldsymbol{v}$, де $\boldsymbol{u}$ і $\boldsymbol{v}$ — проекції вектора $\boldsymbol{b}$ на підпростори $\mathcal{R}(\boldsymbol{A})$ і $\mathcal{R}^{\perp}(\boldsymbol{A})$ відповідно. Оскільки $\boldsymbol{Ax} \in \mathcal{R}(\boldsymbol{A})$, то $\boldsymbol{Ax} - \boldsymbol{u} \in \mathcal{R}(\boldsymbol{A})$. Крім того, $\boldsymbol{v} \perp \boldsymbol{Ax} - \boldsymbol{u}$. На підставі теореми Піфагора маємо:

$$|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2 = |\,\boldsymbol{Ax} - \boldsymbol{u} - \boldsymbol{v}\,|^2 = |\,\boldsymbol{Ax} - \boldsymbol{u}\,|^2 + |\,\boldsymbol{v}\,|^2 \ge |\,\boldsymbol{v}\,|^2. \tag{5}$$

Вище ми бачили, що $\boldsymbol{u} = \boldsymbol{AA}^+\boldsymbol{b} = \boldsymbol{Ax}_0$, тому

$$\boldsymbol{v} = \boldsymbol{b} - \boldsymbol{Ax}_0. \tag{6}$$

З (5) і (6) випливає нерівність (2) і при цьому легко побачити, що вона виконується як рівність тоді і тільки тоді, коли $\boldsymbol{Ax} = \boldsymbol{u}$, тобто за умови (4).

Скориставшись рівністю (4), далі одержимо:

$$\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x}_0 = \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{A}^{+}\boldsymbol{b} = \boldsymbol{A}^{+}\boldsymbol{b} = \boldsymbol{x}_0\,,$$

тому

$$\boldsymbol{x} = \boldsymbol{x}_0 + (\boldsymbol{x} - \boldsymbol{x}_0) = \boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x} + (\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{x}\,.$$

Доданки правої частини останньої формули задовольняють умові $(\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x})^T(\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{x} = 0$. Це означає, що $\boldsymbol{A}^{+}\boldsymbol{A}\boldsymbol{x} \perp (\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{x}$, або $\boldsymbol{x}_0 \perp \boldsymbol{x} - \boldsymbol{x}_0$. Тепер на підставі теореми Піфагора

$$|\,\boldsymbol{x}\,|^2 = |\,\boldsymbol{x}_0\,|^2 + |\,\boldsymbol{x} - \boldsymbol{x}_0\,|^2 \geq |\,\boldsymbol{x}_0\,|^2$$

з рівністю тоді і тільки тоді, коли $\boldsymbol{x} = \boldsymbol{x}_0$. Отже, $\boldsymbol{x}_0$ має найменшу можливу довжину. ☐

Таким чином, розв’язуючи систему $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ методом найменших квадратів, ми повинні знайти нормальний псевдорозв’язок (1).

Припустимо, що матриця $\boldsymbol{A}$ має розміри $m \times n$, де $m > n$, і ранг n. Це відповідає випадку, коли кількість рівнянь системи $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ перевищує кількість невідомих. Така система звичайно буває несумісною. Її нормальний псевдорозв’язок $\boldsymbol{x}_0 = \boldsymbol{A}^{+}\boldsymbol{b}$ на підставі формули (11.1.5) дорівнює $\boldsymbol{x}_0 = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$. Очевидно, що $\boldsymbol{x}_0$ є єдиним розв’язком так званих нормальних рівнянь $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$, з якими мають справу в математичній статистиці. Проте їх використання далеко не завжди є найбільш ефективним або точним методом розв’язання задач математичної статистики. Про це буде йти мова в розділі 12.2.

11.4.2. Дослідження розв’язків, одержуваних методом найменших квадратів. Сформулюємо мовою псевдообернених матриць основні результати, що стосуються розв’язання довільних систем лінійних рівнянь методом найменших квадратів. Спочатку виявимо структуру множини псевдорозв’язків системи рівнянь $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

Теорема 11.4.2. *Вектор* $\boldsymbol{x}$ *мінімізує функцію* $|\,\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\,|^2$ *тоді і тільки тоді, коли він має вигляд*:

$$\boldsymbol{x} = \boldsymbol{A}^{+}\boldsymbol{b} + (\boldsymbol{E} - \boldsymbol{A}^{+}\boldsymbol{A})\boldsymbol{y} \tag{7}$$

для деякого вектора $\boldsymbol{y}$.

Д о в е д е н н я. Н е о б х і д н і с т ь. При доведенні теореми 11.4.1 було з'ясовано, що вектор $\boldsymbol{x}$ мінімізує $|\,\boldsymbol{Ax}-\boldsymbol{b}\,|^2$ тоді і тільки тоді, коли $\boldsymbol{Ax}=\boldsymbol{Ax}_0$. Оскільки при цьому

$$\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{A}^+\boldsymbol{b})=\boldsymbol{Ax}_0-\boldsymbol{AA}^+\boldsymbol{b}=\boldsymbol{0},$$

то на підставі висновку 1 з теореми 11.3.1 можна стверджувати, що $\boldsymbol{x}-\boldsymbol{A}^+\boldsymbol{b}=(\boldsymbol{E}-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{y}$ для деякого вектора $\boldsymbol{y}$. Отже, маємо рівність (7).

Д о с т а т н і с т ь. Оскільки з рівності (7) випливає, що

$$\boldsymbol{Ax}=\boldsymbol{AA}^+\boldsymbol{b}+\boldsymbol{A}(\boldsymbol{E}-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{y}=\boldsymbol{AA}^+\boldsymbol{b}=\boldsymbol{Ax}_0,$$

то, як зазначалося вище, $\boldsymbol{x}$ мінімізує $|\,\boldsymbol{Ax}-\boldsymbol{b}\,|^2$. □

Теорема 11.4.2 дозволяє з'ясувати геометричний зміст множини псевдорозв'язків системи рівнянь $\boldsymbol{Ax}=\boldsymbol{b}$. Другий доданок правої частини рівності (7) у відповідності до теореми 11.3.1 є довільним вектором, належним до підпростору $\mathcal{N}(\boldsymbol{A})$ Тому формула (7) є векторним рівнянням площини, спрямовуючим підпростором якої є $\mathcal{N}(\boldsymbol{A})$, а початковою точкою — нормальний псевдорозв'язок $\boldsymbol{A}^+\boldsymbol{b}$. На підставі висновку з теореми 3.3.1 початковою точкою площини можна вважати будь-яку її точку. Це означає, що множина псевдорозв'язків системи $\boldsymbol{Ax}=\boldsymbol{b}$ утворює площину, спрямовуючим підпростором якої є $\mathcal{N}(\boldsymbol{A})$, а початковою точкою — будь-який псевдорозв'язок.

Доданки правої частини рівності (7) є ортогональними, бо

$$(\boldsymbol{A}^+\boldsymbol{b})^T(\boldsymbol{E}-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{y}=\boldsymbol{b}^T(\boldsymbol{A}^T)^+(\boldsymbol{E}-\boldsymbol{A}^T(\boldsymbol{A}^T)^+)\boldsymbol{y}=0.$$

Це означає, що нормальний псевдорозв'язок є нормальним вектором площини (7). У відповідності до виявлених у розділі 5.2 властивостей нормального вектора площини можна стверджувати, що нормальний псевдорозв'язок є єдиним псевдорозв'язком, ортогональним до $\mathcal{N}(\boldsymbol{A})$ і таким, що має найменшу довжину серед усіх псевдорозв'язків. Останнє твердження було обґрунтоване в теоремі 11.4.1 за допомогою інших геометричних міркувань.

У наступній теоремі йдеться про умову розв'язуваності системи лінійних рівнянь.

Теорема 11.4.3. *Система рівнянь* $\boldsymbol{Ax} = \boldsymbol{b}$ *має розв'язок тоді і тільки тоді, коли*

$$\boldsymbol{AA^{+}b} = \boldsymbol{b}\,. \tag{8}$$

Д о в е д е н н я. Н е о б х і д н і с т ь. Підставивши $\boldsymbol{x}$ з рівності (7) у функцію $|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2$, одержимо її мінімальне значення $|\,\boldsymbol{AA^{+}b} - \boldsymbol{b}\,|^2$. Оскільки система рівнянь $\boldsymbol{Ax} = \boldsymbol{b}$ має розв'язок, то це мінімальне значення дорівнює нулю. Отже, маємо умову (8).

Д о с т а т н і с т ь. Якщо виконується рівність (8), то рівняння $\boldsymbol{Ax} = \boldsymbol{b}$ задовольняється при підставленні в нього $\boldsymbol{x}$ з формули (7). □

Дві останні теореми є важливим знаряддям дослідження загальних систем лінійних рівнянь. Нижче наводиться приклад застосування теореми 11.4.2 при дослідженні методом найменших квадратів задачі з обмеженнями.

11.4.3. Метод найменших квадратів для задачі з обмеженнями. Наступні дві теореми є узагальненням теорем 11.4.2 і 11.4.1.

Теорема 11.4.4. *Вектор* $\boldsymbol{x}$ *мінімізує функцію* $|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2$ *при обмеженні у вигляді сумісної системи рівнянь*

$$\boldsymbol{Fx} = \boldsymbol{f} \tag{9}$$

тоді і тільки тоді, коли

$$\boldsymbol{x} = \boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{G}^{+}\boldsymbol{g} + \boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{y} \tag{10}$$

для деякого вектора $\boldsymbol{y}$*, де*

$$\boldsymbol{G} = \boldsymbol{A\Gamma}_1, \quad \boldsymbol{g} = \boldsymbol{b} - \boldsymbol{AF}^{+}\boldsymbol{f}\,, \quad \boldsymbol{\Gamma}_1 = \boldsymbol{E} - \boldsymbol{F}^{+}\boldsymbol{F}\,, \quad \boldsymbol{\Gamma}_2 = \boldsymbol{E} - \boldsymbol{G}^{+}\boldsymbol{G}\,. \tag{11}$$

Д о в е д е н н я. Нехай γ — множина розв'язків системи рівнянь (9). На підставі теореми 11.4.2 маємо:

$$\boldsymbol{x} = \boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{\Gamma}_1\boldsymbol{z} \tag{12}$$

для деякого вектора $\boldsymbol{z}$.

Підставивши $\boldsymbol{x}$ у функцію $|\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2$, одержимо:

$$\min_{\boldsymbol{x}\in\gamma} |\,\boldsymbol{Ax} - \boldsymbol{b}\,|^2 = \min_{\boldsymbol{z}} |\,\boldsymbol{Gz} - \boldsymbol{g}\,|^2\,.$$

Останній мінімум на підставі теореми 11.4.2 досягається тоді і тільки тоді, коли

$$\boldsymbol{z} = \boldsymbol{G}^{+}\boldsymbol{g} + \boldsymbol{\Gamma}_2\boldsymbol{y} \tag{13}$$

для деякого вектора $\boldsymbol{y}$. З рівностей (12) і (13) випливає, що $\boldsymbol{x}$ мінімізує $|\boldsymbol{Ax} - \boldsymbol{b}|^2$ при обмеженні (9) тоді і тільки тоді, коли

$$\boldsymbol{x} = \boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{\Gamma}_1(\boldsymbol{G}^{+}\boldsymbol{g} + \boldsymbol{\Gamma}_2\boldsymbol{y}). \tag{14}$$

Перетворимо цю формулу, скориставшись другою рівністю (11.1.25) і першою рівністю (11). Ураховуючи, що матриця $\boldsymbol{\Gamma}_1$ є проекційною, одержимо:

$$\begin{aligned}\boldsymbol{\Gamma}_1\boldsymbol{G}^{+} &= \boldsymbol{\Gamma}_1\boldsymbol{G}^{T}(\boldsymbol{G}\boldsymbol{G}^{T})^{+} = \boldsymbol{\Gamma}_1^2\boldsymbol{A}^{T}(\boldsymbol{G}\boldsymbol{G}^{T})^{+} = \\ &= \boldsymbol{\Gamma}_1\boldsymbol{A}^{T}(\boldsymbol{G}\boldsymbol{G}^{T})^{+} = \boldsymbol{G}^{T}(\boldsymbol{G}\boldsymbol{G}^{T})^{+} = \boldsymbol{G}^{+}.\end{aligned} \tag{15}$$

За допомогою формули (15) рівність (14) перетворюється в рівність (10). □

Теорема 11.4.5. *Вектор з мінімальною довжиною серед усіх векторів, що мінімізують* $|\boldsymbol{Ax} - \boldsymbol{b}|^2$ *при обмеженні (9), дорівнює*

$$\boldsymbol{x}_0 = \boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{G}^{+}\boldsymbol{g}.$$

Д о в е д е н н я. Покажемо, що третій доданок правої частини рівності (10) є ортогональним до першого і другого доданків.

З урахуванням рівності $\boldsymbol{\Gamma}_1\boldsymbol{F}^{+} = \boldsymbol{0}$ маємо:

$$(\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{y})^{T}\boldsymbol{F}^{+}\boldsymbol{f} = \boldsymbol{y}^{T}\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}_1\boldsymbol{F}^{+}\boldsymbol{f} = 0,$$

звідки випливає, що $\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{y} \perp \boldsymbol{F}^{+}\boldsymbol{f}$. Беручи до уваги формулу (15) і рівність $\boldsymbol{\Gamma}_2\boldsymbol{G}^{+} = \boldsymbol{0}$, одержимо:

$$(\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{y})^{T}\boldsymbol{G}^{+}\boldsymbol{g} = \boldsymbol{y}^{T}\boldsymbol{\Gamma}_2\boldsymbol{\Gamma}_1\boldsymbol{G}^{+}\boldsymbol{g} = \boldsymbol{y}^{T}\boldsymbol{\Gamma}_2\boldsymbol{G}^{+}\boldsymbol{g} = 0.$$

Це означає, що $\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{y} \perp \boldsymbol{G}^{+}\boldsymbol{g}$.

Тепер за теоремою Піфагора з рівності (10) знаходимо:

$$|\boldsymbol{x}_0|^2 = |\boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{G}^{+}\boldsymbol{g}|^2 + |\boldsymbol{\Gamma}_1\boldsymbol{\Gamma}_2\boldsymbol{y}|^2 \geq |\boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{G}^{+}\boldsymbol{g}|^2,$$

причому строга нерівність має місце при $\boldsymbol{x}_0 \neq \boldsymbol{F}^{+}\boldsymbol{f} + \boldsymbol{G}^{+}\boldsymbol{g}$. □

11.4.4. Двокроковий метод найменших квадратів. На початку глави було зазначено, що розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$ можна розглядати як задачу мінімізації функції $|\boldsymbol{Ax} - \boldsymbol{b}|^2$. Якщо система є сумісною, то ця функція обертається в нуль при підставленні в неї нормального псевдорозв'язку (1), а для несумісної системи величина $\varepsilon_1 = |\boldsymbol{Ax}_0 - \boldsymbol{b}|^2$ відмінна від нуля. При надто великій величині ε_1 можна домогтися її зменшення, дослідивши більш складну задачу.

Припустимо, що є можливість залучити додаткову інформацію про вихідну задачу, збільшивши кількість стовпців матриці $\boldsymbol{A}$ при незмінному векторі $\boldsymbol{b}$. У такому разі нам доведеться мати справу з системою рівнянь

$$(\boldsymbol{A} \quad \boldsymbol{B})\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \boldsymbol{b},$$

або

$$\boldsymbol{Ax} + \boldsymbol{By} = \boldsymbol{b}. \tag{16}$$

Зрозуміло, що можна наново знайти нормальний псевдорозв'язок ускладненої задачі, проте менш трудомістким є наближене обчислення цього псевдорозв'язку, якщо на першому кроці знайти нормальний псевдорозв'язок для $\boldsymbol{x}$, вважаючи $\boldsymbol{y}$ відомим, а на другому кроці скласти і розв'язати рівняння відносно $\boldsymbol{y}$. Нижче ми знайдемо наближений нормальний псевдорозв'язок $\boldsymbol{x}_*$, $\boldsymbol{y}_*$ системи (16) і покажемо, що за певних умов величина $\varepsilon_{12} = |\boldsymbol{Ax}_* + \boldsymbol{By}_* - \boldsymbol{b}|^2$ є меншою ніж ε_1.

Теорема 11.4.6. *Для нормального псевдорозв'язку* $\boldsymbol{x}_*$, $\boldsymbol{y}_*$ *системи* (*16*) *мають місце наближені формули*

$$\boldsymbol{x}_* = \boldsymbol{A}^+(\boldsymbol{E} - \boldsymbol{B}(\boldsymbol{\Gamma B})^+)\boldsymbol{b}, \tag{17}$$

$$\boldsymbol{y}_* = (\boldsymbol{\Gamma B})^+\boldsymbol{b}, \tag{18}$$

де $\boldsymbol{\Gamma} = \boldsymbol{E} - \boldsymbol{AA}^+$, *а величини* ε_1 і ε_{12} *пов'язані співвідношенням*

$$\varepsilon_{12} = \varepsilon_1 - |\boldsymbol{\Gamma B}(\boldsymbol{\Gamma B})^+\boldsymbol{b}|^2. \tag{19}$$

Д о в е д е н н я. З рівності $\boldsymbol{Ax} = \boldsymbol{b} - \boldsymbol{By}$ і теореми 11.4.1 випливає, що

$$\boldsymbol{x}_* = \boldsymbol{A}^+(\boldsymbol{b} - \boldsymbol{B}\boldsymbol{y}). \tag{20}$$

Замінимо цим виразом вектор $\boldsymbol{x}$ у рівності (16). Тоді отримаємо рівняння відносно $\boldsymbol{y}$:

$$\boldsymbol{\Gamma}\boldsymbol{B}\boldsymbol{y} = \boldsymbol{\Gamma}\boldsymbol{b},$$

з якого знаходимо $\boldsymbol{y}_*$:

$$\boldsymbol{y}_* = (\boldsymbol{\Gamma}\boldsymbol{B})^+\boldsymbol{\Gamma}\boldsymbol{b}.$$

На підставі частини 4) теореми 11.3.3 з останньої рівності випливає рівність (18), а після підставлення у формулу (20) виразу для $\boldsymbol{y}_*$ одержимо рівність (17).

Для доведення рівності (19) будемо виходити з тотожності

$$\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}_* + \boldsymbol{B}\boldsymbol{y}_* - \boldsymbol{b} - \boldsymbol{\Gamma}\boldsymbol{B}\boldsymbol{y}_*, \tag{21}$$

яка легко перевіряється за допомогою рівностей (1), (17) і (18). Покажемо, що останній доданок правої частини рівності (21) є ортогональним до суми інших її доданків.

Оскільки $\boldsymbol{\Gamma}\boldsymbol{A} = \boldsymbol{0}$ і $\boldsymbol{\Gamma}$ є проекційною матрицею, то з урахуванням формули (18) і другої рівності (11.1.10) одержимо, що

$$(\boldsymbol{\Gamma}\boldsymbol{B}\boldsymbol{y}_*)^T(\boldsymbol{A}\boldsymbol{x}_* + \boldsymbol{B}\boldsymbol{y}_* - \boldsymbol{b}) = \boldsymbol{y}_*^T\boldsymbol{B}^T\boldsymbol{\Gamma}(\boldsymbol{B}\boldsymbol{y}_* - \boldsymbol{b}) =$$

$$= \boldsymbol{y}_*^T(\boldsymbol{B}^T\boldsymbol{\Gamma}^2\boldsymbol{B}(\boldsymbol{\Gamma}\boldsymbol{B})^+ - \boldsymbol{B}^T\boldsymbol{\Gamma})\boldsymbol{b} = \boldsymbol{y}_*^T((\boldsymbol{\Gamma}\boldsymbol{B})^T\boldsymbol{\Gamma}\boldsymbol{B}(\boldsymbol{\Gamma}\boldsymbol{B})^+ - (\boldsymbol{\Gamma}\boldsymbol{B})^T)\boldsymbol{b} = 0,$$

тобто $\boldsymbol{\Gamma}\boldsymbol{B}\boldsymbol{y}_* \perp \boldsymbol{A}\boldsymbol{x}_* + \boldsymbol{B}\boldsymbol{y}_* - \boldsymbol{b}$.

На підставі теореми Піфагора з рівності (21) знаходимо:

$$|\,\boldsymbol{A}\boldsymbol{x}_0 - \boldsymbol{b}\,|^2 = |\,\boldsymbol{A}\boldsymbol{x}_* + \boldsymbol{B}\boldsymbol{y}_* - \boldsymbol{b}\,|^2 + |\,\boldsymbol{\Gamma}\boldsymbol{B}\boldsymbol{y}_*\,|^2,$$

звідки випливає рівність (19). □

Другий доданок у правій частині формули (19) дорівнює нулю при $\boldsymbol{\Gamma}\boldsymbol{B} = \boldsymbol{0}$. Покажемо, що ця рівність має місце в тому випадку, коли стовпці матриці $\boldsymbol{B}$ є лінійно залежними від стовпців матриці $\boldsymbol{A}$.

Дійсно, рівність $\boldsymbol{\Gamma}\boldsymbol{A} = \boldsymbol{0}$ показує, що $\boldsymbol{\Gamma}\boldsymbol{a}_{*j} = \boldsymbol{0}$ для кожного стовпця $\boldsymbol{a}_{*j}$ матриці $\boldsymbol{A}$. Звідси випливає, що $\boldsymbol{\Gamma}\boldsymbol{b}_{*j} = \boldsymbol{0}$ для будь-якого стовпця $\boldsymbol{b}_{*j}$ матриці $\boldsymbol{B}$, який є лінійною комбінацією стовпців матриці $\boldsymbol{A}$. Тому $\boldsymbol{\Gamma}\boldsymbol{B} = \boldsymbol{0}$, якщо всі стовпці матриці $\boldsymbol{B}$ лінійно залежать від стовпців матриці $\boldsymbol{A}$.

Отже, величина ε_1 не буде зменшена при вказаній особливості матриці $\boldsymbol{B}$. Тому цей випадок не являє інтересу при користуванні двокроковим методом найменших квадратів. Цей висновок є зрозумілим у світлі сказаного вище про представлення матрицею $\boldsymbol{B}$ додаткової по відношенню до матриці $\boldsymbol{A}$ інформації.

Слід зауважити, що другий доданок правої частини рівності (19) може обернутися в нуль навіть при $\mathbf{\Gamma}\boldsymbol{B} \neq \mathbf{0}$, якщо вектор $\boldsymbol{b}$ виявиться ненульовим розв'язком системи

$$\mathbf{\Gamma}\boldsymbol{B}(\mathbf{\Gamma}\boldsymbol{B})^{+}\boldsymbol{z} = \mathbf{0}. \tag{22}$$

Ця можливість випливає з теореми 3.5.1.

11.5. Середньоквадратичне дискретне наближення

Метод найменших квадратів — широко відомий метод обчислень, який використовується не тільки при розв'язанні систем лінійних рівнянь.У цьому розділі мова йде про застосування методу найменших квадратів до однієї задачі теорії найближень функцій.

11.5.1. Формулювання задачі середньоквадратичного наближення. Функції $u_1(x), \ldots, u_m(x)$ називаються лінійно залежними на деякій множині *X*, якщо існують такі числа $\alpha_1, \ldots, \alpha_m$, не всі рівні нулю, що

$$\sum_{k=1}^{m} \alpha_k u_k(x) = 0$$

для всіх точок множини *X*. Якщо ця рівність виконується лише тоді, коли всі числа $\alpha_1, \ldots, \alpha_m$ дорівнюють нулю, функції називаються лінійно незалежними на *X*.

Нехай неперервну функцію $f(x)$ треба наближено представити на *X* лінійною комбінацією лінійно незалежних неперервних функцій $u_1(x), \ldots, u_m(x)$. Позначимо шукану лінійну комбінацію через $u(x)$:

$$u(x) = \sum_{k=1}^{m} z_k u_k(x), \tag{1}$$

де $z_1, \ldots, z_m$ — деякі числа. Функції вигляду (1) називають узагальненими многочленами. Задача про середньоквадратичне наближення функцій полягає в тому, що за рахунок належного вибору коефіцієнтів $z_1, \ldots, z_m$ треба замінити функцію $f(x)$ узагальненим многочленом $u(x)$ так, щоб відхилення $u(x)$ від $f(x)$ на множині X було мінімальним.

Коли множина X складається з точок $x_1, \ldots, x_n$ числової осі, які вважаються різними, і відхилення $u(x)$ від $f(x)$ визначається величиною

$$\Delta = \left(\sum_{l=1}^{n} [\, f(x_l) - u(x_l)]^2 \right)^{1/2}, \tag{2}$$

маємо справу з так званою задачею середньоквадратичного дискретного наближення.

Права частина формули (2) є довжиною вектора $\boldsymbol{u} - \boldsymbol{f}$, де

$$\boldsymbol{u} = (u(x_1) \ \ \ldots \ \ u(x_n))^T, \quad \boldsymbol{f} = (f(x_1) \ \ \ldots \ \ f(x_n))^T.$$

На підставі рівності (1) вектор $\boldsymbol{u}$ можна подати у вигляді:

$$\boldsymbol{u} = \sum_{k=1}^{m} z_k \boldsymbol{u}_{*k} = \boldsymbol{Uz},$$

де $\boldsymbol{z} = (z_1 \ \ \ldots \ \ z_m)^T$ і $\boldsymbol{u}_{*k}$ — стовпці матриці

$$\boldsymbol{U} = \begin{pmatrix} u_1(x_1) & u_2(x_1) & \ldots & u_m(x_1) \\ u_1(x_2) & u_2(x_2) & \ldots & u_m(x_2) \\ \ldots & \ldots & \ldots & \ldots \\ u_1(x_n) & u_2(x_n) & \ldots & u_m(x_n) \end{pmatrix}. \tag{3}$$

Отже, $\Delta = |\boldsymbol{Uz} - \boldsymbol{f}|$ і тому задача зводиться до мінімізації величини $|\boldsymbol{Uz} - \boldsymbol{f}|^2$. Це питання розглядалося в розділі 11.4.Там було показано,що сформульована задача розв’язується шляхом відшукання нормального псевдорозв’язку.

Далі ми обмежимося найпростішим випадком, коли функції $u_1(x), \ldots, u_m(x)$ задовольняють умові, щоб стовпці матриці $\boldsymbol{U}$ були ортонормованими.Тоді на підставі формул (11.4.1) і (11.1.5) будемо мати вираз для нормального псевдорозв'язку:

$$z_0 = \boldsymbol{U}^+ \boldsymbol{f} = (\boldsymbol{U}^T \boldsymbol{U})^{-1} \boldsymbol{U}^T \boldsymbol{f} = \boldsymbol{U}^T \boldsymbol{f}. \tag{4}$$

При цьому похибка середньоквадратичного дискретного наближення, рівна мінімальному значенню величини Δ, є такою:

$$\Delta_{\min} = \left|(\boldsymbol{U}\boldsymbol{U}^T - \boldsymbol{E})\boldsymbol{f}\right|.$$

Задача середньоквадратичного дискретного наближення розглядається за умови, що $n > m$. Припустимо, що кількість точок x_l збігається з кількістю функцій $u_k(x)$, тобто $n = m$.Тоді матриця $\boldsymbol{U}$ стає ортогональною і тому $\Delta_{\min} = 0$. Це означає, що мінімальне відхилення $u(x)$ від $f(x)$ є нульовим.Розглядувану задачу в цьому випадку називають задачею інтерполяції, а точки x_1, ..., x_n — вузлами інтерполяції.

11.5.2.Середньоквадратичне дискретне наближення функцій тригонометричними многочленами. Розглянемо важливий для застосувань приклад середньоквадратичного дискретного наближення функцій тригонометричними многочленами. Припустимо, що наближення функції $f(x)$ тригонометричним многочленом розшукується на сукупності рівновіддалених точок

$$x_l = (l-1)\alpha, \quad l = \overline{1,\ n}, \quad \alpha = \frac{2\pi}{n} \tag{5}$$

відрізка $[0,\ 2\pi]$ числової осі. Візьмемо матрицю $\boldsymbol{U}$ у вигляді:

$$\boldsymbol{U} = \sqrt{\frac{2}{n}} \begin{pmatrix} 1/\sqrt{2} & \cos x_1 & \sin x_1 & \ldots & \cos mx_1 & \sin mx_1 \\ 1/\sqrt{2} & \cos x_2 & \sin x_2 & \ldots & \cos mx_2 & \sin mx_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1/\sqrt{2} & \cos x_n & \sin x_n & \ldots & \cos mx_n & \sin mx_n \end{pmatrix}, \tag{6}$$

де m — деяке число,що задовольняє умові $2m+1 < n$. Пересвідчимося в ортонормованості стовпців матриці $\boldsymbol{U}$, після чого напишемо відповідний до цієї матриці узагальнений многочлен.

Підсумовуючи геометричну прогресію, з урахуванням(5) одержимо:

$$\sum_{l=1}^{n} e^{ikx_l} = \sum_{l=1}^{n} e^{ik(l-1)\alpha} = \frac{e^{ikn\alpha}-1}{e^{ik\alpha}-1} = 0. \tag{7}$$

За допомогою (7) і формули Ейлера легко довести для стовпців матриці $\boldsymbol{U}$ рівності

$$\begin{gathered} \boldsymbol{u}_{*1}^T \boldsymbol{u}_{*2k} = 0, \quad \boldsymbol{u}_{*1}^T \boldsymbol{u}_{*2j+1} = 0, \quad \boldsymbol{u}_{*2k}^T \boldsymbol{u}_{*2j+1} = 0, \\ \boldsymbol{u}_{*2k}^T \boldsymbol{u}_{*2j} = 0, \quad \boldsymbol{u}_{*2k+1}^T \boldsymbol{u}_{*2j+1} = 0, \end{gathered} \tag{8}$$

де $k, j = \overline{1,\ m}$, причому у двох останніх рівностях (8) вважається, що $k \neq j$.

Далі будемо обчислювати інші скалярні добутки стовпців матриці $\boldsymbol{U}$. Очевидно, що

$$\boldsymbol{u}_{*1}^T \boldsymbol{u}_{*1} = 1. \tag{9}$$

З урахуванням формули (7) маємо:

$$\boldsymbol{u}_{*2k}^T \boldsymbol{u}_{*2k} =$$

$$= \frac{2}{n}\sum_{l=1}^{n}\cos^2 kx_l = \frac{1}{n}\sum_{l=1}^{n}\left(1+\frac{1}{2}(e^{2ikx_l}+e^{-2ikx_l})\right) = 1, \quad k = \overline{1,\ m}. \tag{10}$$

Аналогічно доводиться, що

$$\boldsymbol{u}_{*2j+1}^T \boldsymbol{u}_{*2j+1} = 1, \quad j = \overline{1,\ m}. \tag{11}$$

Рівності (8) — (11) показують, що стовпці матриці $\boldsymbol{U}$ є ортонормованими. Їй відповідає тригонометричний многочлен

$$u(x) = \sqrt{\frac{2}{n}}\left(\frac{z_0}{\sqrt{2}} + \sum_{k=1}^{m}(z_k' \cos kx + z_k'' \sin kx)\right), \tag{12}$$

коефіцієнти якого одержуються на підставі формул (4) і (6):

$$z_0 = \frac{1}{\sqrt{n}}\sum_{l=1}^{n} f(x_l),$$

$$z_k' = \sqrt{\frac{2}{n}} \sum_{l=1}^{n} f(x_l) \cos k x_l, \quad k = \overline{1,\ m},$$

$$z_k'' = \sqrt{\frac{2}{n}} \sum_{l=1}^{n} f(x_l) \sin k x_l, \quad k = \overline{1,\ m}.$$

Ці формули носять назву *формул Бесселя*. Підставляючи знайдені коефіцієнти у формулу (12), одержимо вираз для шуканого тригонометричного многочлена. Зазначимо, що вигляд тригонометричного многочлена може бути відмінним від (12) і його можна вибрати багатьма способами.

11.6. Обчислення псевдооберненої матриці і нормального псевдорозв'язку

Псевдообернення матриці є досить серйозною проблемою в порівнянні з обчисленням оберненої матриці. З багатьох існуючих способів знаходження псевдооберненої матриці ми розглянемо тут два способи.

11.6.1. Псевдообернення за допомогою скелетного розкладу. Якщо відомий скелетний розклад матриці $\boldsymbol{A}$, то $\boldsymbol{A}^+$ обчислюється за допомогою формули (11.2.19). Розглянемо спосіб побудови скелетного розкладу, відмінний від використаного в розділі 3.4. Він заснований на так званій другій формі сингулярного розкладу матриці, яка вводиться нижче.

Для $m \times n$ -матриці $\boldsymbol{A}$ рангу r сингулярний розклад виглядає так:

$$\boldsymbol{A} = \boldsymbol{Q}_1 \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{Q}_2 = \boldsymbol{Q}_1 \begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix} \boldsymbol{\Sigma} (\boldsymbol{E}_r \quad \boldsymbol{0}) \boldsymbol{Q}_2,$$

де $\boldsymbol{Q}_1, \boldsymbol{Q}_2$ — ортогональні матриці і $\boldsymbol{\Sigma}$ — діагональна матриця, діагональними елементами якої є ненульові сингулярні числа. Матриця $\boldsymbol{U}_1 = \boldsymbol{Q}_1 \begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix}$ складається з перших r стовпців матриці $\boldsymbol{Q}_1$, а матриця $\boldsymbol{U}_2 = (\boldsymbol{E}_r \quad \boldsymbol{0})\boldsymbol{Q}_2$ — з перших r рядків матриці $\boldsymbol{Q}_2$. Отже, ми маємо розклад

$$\boldsymbol{A} = \boldsymbol{U}_1 \boldsymbol{\Sigma} \boldsymbol{U}_2, \tag{1}$$

який іноді називають *другою формою сингулярного розкладу*. На відміну від загальновживаної першої форми в розклад входять матриці менших розмірів. Очевидно, що для їх знаходження треба обчислити тільки координати тих векторів сингулярних базисів, які відповідають ненульовим сингулярним числам.

На підставі рівності (1) можна одержати скелетний розклад, об'єднавши множники, наприклад, таким способом:

$$\boldsymbol{A} = \boldsymbol{U}_1(\boldsymbol{\Sigma} \boldsymbol{U}_2).$$

Тепер можна застосовувати теорему 11.2.1 і написати $\boldsymbol{A}^+$ у вигляді:

$$\boldsymbol{A}^+ = (\boldsymbol{\Sigma} \boldsymbol{U}_2)^+ \boldsymbol{U}_1^+.$$

Оскільки стовпці матриці $\boldsymbol{U}_1$ ортонормовані, то на підставі формули (11.1.5) маємо рівність $\boldsymbol{U}_1^+ = \boldsymbol{U}_1^T$.

Добуток $\boldsymbol{\Sigma} \boldsymbol{U}_2$ є скелетним розкладом, тому при його псевдооберненні знову застосуємо теорему 11.2.1, а потім рівність (11.1.6). У результаті одержимо:

$$(\boldsymbol{\Sigma} \boldsymbol{U}_2)^+ = \boldsymbol{U}_2^+ \boldsymbol{\Sigma}^+ = \boldsymbol{U}_2^T \boldsymbol{\Sigma}^{-1}.$$

Отже, маємо рівність

$$\boldsymbol{A}^+ = \boldsymbol{U}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{U}_1^T, \tag{2}$$

яка є аналогом рівності (11.1.24).

Зазначимо наприкінці, що коли відомий сингулярний розклад матриці $\boldsymbol{A}$, то $\boldsymbol{A}^+$ легко визначається за допомогою рівності (11.1.24). Проте вдаватися до обчислення сингулярного розкладу не завжди доцільно через його трудомісткість при чисельній реалізації. Саме тому являють інтерес інші методи псевдообернення, зокрема, засновані на використанні скелетного розкладу.

П р и к л а д 1. За допомогою формули (2) обчислимо $\boldsymbol{A}^{+}$ для матриці $\boldsymbol{A}$, розглянутої в прикладі з розділу 11.2. Знайдемо спочатку сингулярні числа матриці $\boldsymbol{A}$, обчисливши власні значення матриці $\boldsymbol{A}\boldsymbol{A}^{T}$. Ця матриця дорівнює

$$\boldsymbol{A}\boldsymbol{A}^{T}=\begin{pmatrix}1&1&0&0\\0&0&1&-1\\1&1&-1&1\end{pmatrix}\begin{pmatrix}1&0&1\\1&0&1\\0&1&-1\\0&-1&1\end{pmatrix}=\begin{pmatrix}2&0&2\\0&2&-2\\2&-2&4\end{pmatrix},$$

а її власними значеннями є числа 0, 2, 6. Сингулярні числа дорівнюють $\sigma_1=\sqrt{6},\ \sigma_2=\sqrt{2},\ \sigma_3=0.$ Ненульовим сингулярним числам відповідають ортонормовані власні вектори матриці $\boldsymbol{A}\boldsymbol{A}^{T}$

$$\boldsymbol{e}_1=\frac{1}{\sqrt{6}}(1\quad -1\quad 2)^{T},\quad \boldsymbol{e}_2=\frac{1}{\sqrt{2}}(1\quad 1\quad 0)^{T}.$$

За допомогою другої формули (6.6.16) можна знайти ортонормовані власні вектори матриці $\boldsymbol{A}^{T}\boldsymbol{A}$, відповідні до σ_1 і σ_2:

$$\boldsymbol{f}_1=\frac{1}{\sigma_1}\boldsymbol{A}^{T}\boldsymbol{e}_1=\frac{1}{\sqrt{6}}\begin{pmatrix}1&0&1\\1&0&1\\0&1&-1\\0&-1&1\end{pmatrix}\frac{1}{\sqrt{6}}\begin{pmatrix}1\\-1\\2\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1\\1\\-1\\1\end{pmatrix},$$

$$\boldsymbol{f}_2=\frac{1}{\sigma_2}\boldsymbol{A}^{T}\boldsymbol{e}_2=\frac{1}{\sqrt{2}}\begin{pmatrix}1&0&1\\1&0&1\\0&1&-1\\0&-1&1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}.$$

Матриці $\boldsymbol{U}_1$ і $\boldsymbol{U}_2$ з формули (1) дорівнюють

$$\boldsymbol{U}_1=(\boldsymbol{e}_1\quad \boldsymbol{e}_2)=\begin{pmatrix}1/\sqrt{6}&1/\sqrt{2}\\-1/\sqrt{6}&1/\sqrt{2}\\2/\sqrt{6}&0\end{pmatrix},\quad \boldsymbol{U}_2=\begin{pmatrix}\boldsymbol{f}_1^{T}\\\boldsymbol{f}_2^{T}\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1&1&-1&1\\1&1&1&-1\end{pmatrix}.$$

Тепер на підставі формули (2) одержується матриця

$$\boldsymbol{A}^{+} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1/\sqrt{6} & 0 \\ 0 & 1/\sqrt{2} \end{pmatrix}\begin{pmatrix} 1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} =$$

$$= \frac{1}{6}\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 2 & -1 \\ -1 & -2 & 1 \end{pmatrix},$$

яка збігається з матрицею, отриманою в розділі 11.2. $\triangle$

11.6.2. Метод Гревілля. Розглянемо псевдообернення блокової матриці у тому випадку, коли один з двох блоків є рядком.

Теорема 11.6.1. *Нехай*

$$\boldsymbol{A}_{n+1} = \begin{pmatrix} \boldsymbol{A}_n \\ \boldsymbol{a}_{n+1}^T \end{pmatrix},$$

де $\boldsymbol{a}_{n+1}$ *— матриця-стовпець. Тоді*

$$\boldsymbol{A}_{n+1}^{+} = ((\boldsymbol{E} - \boldsymbol{l}_{n+1}\boldsymbol{a}_{n+1}^T)\boldsymbol{A}_n^{+} \quad \boldsymbol{l}_{n+1}),$$

де

$$\boldsymbol{l}_{n+1} = \begin{cases} \alpha^{-1}(\boldsymbol{E} - \boldsymbol{A}_n^{+}\boldsymbol{A}_n)\boldsymbol{a}_{n+1}, & (\boldsymbol{E} - \boldsymbol{A}_n^{+}\boldsymbol{A}_n)\boldsymbol{a}_{n+1} \neq \boldsymbol{0}, \\ \beta^{-1}\boldsymbol{A}_n^{+}(\boldsymbol{A}_n^T)^{+}\boldsymbol{a}_{n+1}, & (\boldsymbol{E} - \boldsymbol{A}_n^{+}\boldsymbol{A}_n)\boldsymbol{a}_{n+1} = \boldsymbol{0}, \end{cases}$$

$$\alpha = \boldsymbol{a}_{n+1}^T(\boldsymbol{E} - \boldsymbol{A}_n^{+}\boldsymbol{A}_n)\boldsymbol{a}_{n+1}, \qquad \beta = 1 + \boldsymbol{a}_{n+1}^T\boldsymbol{A}_n^{+}(\boldsymbol{A}_n^T)^{+}\boldsymbol{a}_{n+1}.$$

Д о в е д е н н я полягає в перевірці виконання умов теореми Пенроуза. Нехай $(\boldsymbol{E} - \boldsymbol{A}_n^{+}\boldsymbol{A}_n)\boldsymbol{a}_{n+1} \neq \boldsymbol{0}$, тоді з урахуванням очевидної рівності $\boldsymbol{A}_n\boldsymbol{l}_{n+1} = \boldsymbol{0}$ маємо:

$$\boldsymbol{A}_{n+1}\boldsymbol{A}_{n+1}^{+} = \begin{pmatrix} \boldsymbol{A}_n\boldsymbol{A}_n^{+} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{pmatrix}, \tag{3}$$

$$\boldsymbol{A}_{n+1}^{+}\boldsymbol{A}_{n+1} = \boldsymbol{A}_n^{+}\boldsymbol{A}_n + \boldsymbol{l}_{n+1}\boldsymbol{a}_{n+1}^T(\boldsymbol{E} - \boldsymbol{A}_n^{+}\boldsymbol{A}_n). \tag{4}$$

Легко переконатися в тому, що матриця $\boldsymbol{A}_{n+1}\boldsymbol{A}_{n+1}^{+}$ є симетричною. Симетричність матриці $\boldsymbol{A}_{n+1}^{+}\boldsymbol{A}_{n+1}$ стає очевидною, якщо в правій частині рівності (4) замінити $\boldsymbol{l}_{n+1}$ його виразом з умови теореми. Перевірка двох інших умов теореми Пенроуза не викликає труднощів, якщо скористатися рівністю (3).

При $(\boldsymbol{E} - \boldsymbol{A}_n^+ \boldsymbol{A}_n)\boldsymbol{a}_{n+1} = \boldsymbol{0}$, або $\boldsymbol{a}_{n+1}^T(\boldsymbol{E} - \boldsymbol{A}_n^+ \boldsymbol{A}_n) = \boldsymbol{0}$, маємо:

$$\boldsymbol{A}_{n+1}\boldsymbol{A}_{n+1}^+ = \begin{pmatrix} \boldsymbol{A}_n\boldsymbol{A}_n^+ - \beta^{-1}(\boldsymbol{A}_n^T)^+\boldsymbol{a}_{n+1}\boldsymbol{a}_{n+1}^T\boldsymbol{A}_n^+ & \beta^{-1}(\boldsymbol{A}_n^T)^+\boldsymbol{a}_{n+1} \\ \beta^{-1}\boldsymbol{a}_{n+1}^T\boldsymbol{A}_n^+ & \beta^{-1}\boldsymbol{a}_{n+1}^T\boldsymbol{A}_n^+(\boldsymbol{A}_n^T)^+\boldsymbol{a}_{n+1} \end{pmatrix}, \quad (5)$$

$$\boldsymbol{A}_{n+1}^+\boldsymbol{A}_{n+1} = \boldsymbol{A}_n^+\boldsymbol{A}_n\,, \quad (6)$$

причому рівність (5) одержана з урахуванням формули

$$\boldsymbol{A}_n\boldsymbol{A}_n^+(\boldsymbol{A}_n^T)^+ = (\boldsymbol{A}_n^T)^+,$$

яка є наслідком другої рівності (11.1.25) і першої рівності (11.1.23).

Цілком очевидно, що матриці $\boldsymbol{A}_{n+1}\boldsymbol{A}_{n+1}^+$ і $\boldsymbol{A}_{n+1}^+\boldsymbol{A}_{n+1}$, які визначаються формулами (5) і (6), є симетричними. Дві інші умови теореми Пенроуза легко перевіряються за допомогою рівності (6). ☐

Випадок, коли блоком є стовпець, можна одержати транспонуванням матриць $\boldsymbol{A}_{n+1}$ і $\boldsymbol{A}_{n+1}^+$.

На теоремі 11.6.1 ґрунтується *метод Гревілля* послідовного знаходження псевдооберненої матриці. Він полягає в тому, що спочатку обчислюється $\boldsymbol{A}_1^+$ для матриці-рядка $\boldsymbol{A}_1$, а потім послідовно знаходяться $\boldsymbol{A}_2^+$, $\boldsymbol{A}_3^+$ і т.д. Цей метод є особливо зручним у тому випадку, коли на протязі деякого часу інформація щодо розв'язуваної проблеми поповнюється і розрахунки, виконані для матриці $\boldsymbol{A}_n$, треба повторити з матрицею $\boldsymbol{A}_{n+1}$, одержаною з $\boldsymbol{A}_n$ добавленням рядка (стовпця).

11.6.3. Обчислення нормального псевдорозв'язку. Якщо відома псевдообернена матриця, то нормальний псевдорозв'язок можна знайти за допомогою рівності (11.4.1). Проте не обов'язково обчислювати псевдообернену матрицю при розшукуванні нормального псевдорозв'язку. Більш прийнятним для розв'язання цієї проблеми є спосіб, заснований на використанні рівності (2).

Нагадаємо, що стовпцями матриць $\boldsymbol{U}_1$ і $\boldsymbol{U}_2^T$ є вектори сингулярних базисів, відповідні до ненульових сингулярних числел. Скориставшись позначеннями розділу 6.6 і розмірковуючи, як при доведенні рівності (6.4.4), одержимо вираз для нормального псевдорозв'язку:

$$\boldsymbol{x}_0 = \boldsymbol{A}^{+}\boldsymbol{b} = \boldsymbol{U}_2^T\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1^T\boldsymbol{b} =$$

$$= (\boldsymbol{f}_1 \quad \ldots \quad \boldsymbol{f}_r)\,\mathrm{diag}(\sigma_1^{-1}, \ldots, \sigma_r^{-1})\begin{pmatrix} \boldsymbol{e}_1^T \\ \ldots \\ \boldsymbol{e}_r^T \end{pmatrix}\boldsymbol{b} = \sum_{i=1}^{r}\sigma_i^{-1}\boldsymbol{f}_i\boldsymbol{e}_i^T\boldsymbol{b}\,.$$

Звідси на підставі формул (6.6.16), які визначають зв'язки між векторами різних сингулярних базисів, маємо:

$$\boldsymbol{x}_0 = \sum_{i=1}^{r}\frac{\boldsymbol{A}^T\boldsymbol{e}_i}{\sigma_i^2}\boldsymbol{e}_i^T\boldsymbol{b} = \sum_{i=1}^{r}\frac{\boldsymbol{b}^T\boldsymbol{e}_i}{\sigma_i^2}\boldsymbol{A}^T\boldsymbol{e}_i\,, \tag{7}$$

$$\boldsymbol{x}_0 = \sum_{i=1}^{r}\frac{\boldsymbol{f}_i\boldsymbol{f}_i^T\boldsymbol{A}^T\boldsymbol{b}}{\sigma_i^2} = \sum_{i=1}^{r}\frac{\boldsymbol{b}^T\boldsymbol{A}\boldsymbol{f}_i}{\sigma_i^2}\boldsymbol{f}_i. \tag{8}$$

Формули (7) і (8) дають вирази для нормального псевдорозв'язку через вектори одного з двох сингулярних базисів.

П р и к л а д 2. Знайдемо нормальний псевдорозв'язок системи рівнянь

$$\begin{aligned} x_1 + x_2 \phantom{{}+x_3} &= 2, \\ x_2 + x_3 &= 3. \end{aligned}$$

Скористаємося формулою (7). Очевидно, що

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{A}\boldsymbol{A}^T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Власні значення і ортонормовані власні вектори матриці $\boldsymbol{A}\boldsymbol{A}^T$ є такими:

$$\lambda_1 = 3, \quad \boldsymbol{e}_1 = \frac{1}{\sqrt{2}}(1 \quad 1)^T, \quad \lambda_2 = 1, \quad \boldsymbol{e}_2 = \frac{1}{\sqrt{2}}(-1 \quad 1)^T.$$

Сингулярні числа матриці $\boldsymbol{A}$ дорівнюють $\sigma_1 = \sqrt{3},\ \sigma_2 = 1.$ Далі знаходимо:

$$\frac{\boldsymbol{b}^T\boldsymbol{e}_1}{\sigma_1^2} = \frac{5}{3\sqrt{2}}, \quad \frac{\boldsymbol{b}^T\boldsymbol{e}_2}{\sigma_2^2} = \frac{1}{\sqrt{2}},$$

$$\boldsymbol{A}^T\boldsymbol{e}_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},$$

$$\boldsymbol{A}^T\boldsymbol{e}_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Тепер за допомогою формули (7) знаходимо нормальний псевдорозв'язок:

$$\boldsymbol{x}_0 = \frac{\boldsymbol{b}^T\boldsymbol{e}_1}{\sigma_1^2}\boldsymbol{A}^T\boldsymbol{e}_1 + \frac{\boldsymbol{b}^T\boldsymbol{e}_2}{\sigma_2^2}\boldsymbol{A}^T\boldsymbol{e}_2 =$$

$$= \frac{5}{3\sqrt{2}} \cdot \frac{1}{\sqrt{2}}(1 \quad 2 \quad 1)^T + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}(-1 \quad 0 \quad 1)^T = \begin{pmatrix} \frac{1}{3} & \frac{5}{3} & \frac{4}{3} \end{pmatrix}^T. \ \triangle$$

Вправи до глави 11

11.1. Довести, що $(\alpha\boldsymbol{A})^+ = \alpha^{-1}\boldsymbol{A}^+$, де α — число, не рівне нулю.

11.2. Нехай $\boldsymbol{A} = \boldsymbol{a}$, де $\boldsymbol{a}$ — вектор. Довести, що $\boldsymbol{A}^+ = \alpha\boldsymbol{a}^T$. Чому дорівнює α?

11.3. Довести, що

$$(\boldsymbol{A} \quad \boldsymbol{0})^+ = \begin{pmatrix} \boldsymbol{A}^+ \\ \boldsymbol{0}^T \end{pmatrix}.$$

11.4. Довести, що

$$r(\boldsymbol{AB}) = r(\boldsymbol{A}^+\boldsymbol{AB}) = r(\boldsymbol{ABB}^+).$$

11.5. Довести, що коли $\boldsymbol{M}$ і $\boldsymbol{N}$ — невироджені матриці і $\boldsymbol{A}$ — довільна матриця, то

$$(\boldsymbol{MA})^{+}\boldsymbol{MA} = \boldsymbol{A}^{+}\boldsymbol{A}, \quad \boldsymbol{AN}(\boldsymbol{AN})^{+} = \boldsymbol{AA}^{+}.$$

11.6. Довести, що коли $\boldsymbol{M}$ — невироджена і $\boldsymbol{Q}$ — ортогональна матриця, то для довільної матриці $\boldsymbol{A}$ виконується рівність

$$\boldsymbol{A}^{+} = \boldsymbol{Q}(\boldsymbol{MA}^{T}\boldsymbol{AQ})^{+}\boldsymbol{MA}^{T}.$$

11.7. Обґрунтувати можливість застосування теореми 11.2.1 при доведенні рівностей (11.3.2).

11.8. Нехай система $\boldsymbol{Ax} = \boldsymbol{b}$ є сумісною. Дослідити задачу про мінімум квадратичної форми $\boldsymbol{x}^{T}\boldsymbol{x}$ на множині векторів, що задовольняють умові $\boldsymbol{Ax} = \boldsymbol{b}$. Порівняти з вправою 5.1.

11.9. Нехай $\boldsymbol{A}$ — додатно визначена матриця порядку n і система рівнянь $\boldsymbol{Bx} = \boldsymbol{c}$, де $\boldsymbol{B}$ — матриця розмірів $m \times n$, є сумісною. Довести, що мінімальне значення квадратичної форми $f = \boldsymbol{x}^{T}\boldsymbol{Ax}$ при обмеженні $\boldsymbol{Bx} = \boldsymbol{c}$ дорівнює $f_{\min} = \boldsymbol{c}^{T}\boldsymbol{S}^{+}\boldsymbol{c}$, де $\boldsymbol{S} = \boldsymbol{BA}^{-1}\boldsymbol{B}^{T}$, причому воно досягається при $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{B}^{T}\boldsymbol{S}^{+}\boldsymbol{c}$.

11.10. Що треба розуміти під множиною γ в теоремі 11.4.4, якщо система (11.4.9) є несумісною?

11.11. Довести, що система (11.4.22) має ненульові розв'язки.

11.12. Побудувати середньоквадратичне дискретне наближення функції $f(x)$ за допомогою узагальненого многочлена

$$u(x) = \sqrt{\frac{2}{n}}\left(\frac{z_0}{\sqrt{2}}T_0(x) + \sum_{k=1}^{m} z_k T_k(x)\right),$$

де $T_k(x)$, $k = \overline{0,\ m}$, — многочлени Чебишова. Вважати, що побудова наближення виконується на сукупності точок $x_l = t_{nl}$, $l = \overline{1,\ n}$, $n > m$, де t_{nl} — корені многочлена $T_n(x)$ на відрізку $[-1, 1]$.

ГЛАВА 12

МАТРИЧНІ ОБЧИСЛЕННЯ

Розв'язання систем лінійних алгебраїчних рівнянь і розв'язання алгебраїчної проблеми власних значень є основними обчислювальними задачами лінійної алгебри. З багатьох застосовуваних на практиці методів у цій главі вивчаються: 1) метод збурень і найпростіші ітераційні методи для розв'язання систем лінійних рівнянь; 2) ***QR***-алгоритм разом зі степеневим методом для розв'язання алгебраїчної проблеми власних значень. Один з розділів глави присвячений аналізу труднощів, які виникають при користуванні обчислювальними методами, і деяким шляхам їх подолання.

12.1. Норми векторів і матриць

12.1.1. Векторні норми. У багатьох задачах виникає необхідність порівнювати вектори. Наприклад, в евклідовому просторі вектори можна порівнювати за довжиною. Узагальненням поняття довжини є норма вектора.

Дійсний або комплексний векторний простір називається *нормованим* простором, якщо виконані такі умови.

A. Існує правило, за яким кожному вектору $\boldsymbol{a}$ з указаного простору зіставляється дійсне число $\|\boldsymbol{a}\|$, що називається *нормою* цього вектора.

B. Указане правило підпорядковане аксіомам:

1) $\|\boldsymbol{a}\| > 0$, якщо $\boldsymbol{a} \neq \mathbf{0}$, і $\|\mathbf{0}\| = 0$;

2) $\|\alpha\boldsymbol{a}\| = |\alpha| \, \|\boldsymbol{a}\|$ для будь-якого числа α;

3) $\|\boldsymbol{a} + \boldsymbol{b}\| \le \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$ («нерівність трикутника»).

З нерівності трикутника можна одержати одне корисне співвідношення. Очевидно, що для будь-яких векторів $\boldsymbol{a}$ і $\boldsymbol{b}$ маємо:

$$\| \boldsymbol{a} \| = \| \boldsymbol{a} - \boldsymbol{b} + \boldsymbol{b} \| \leq \| \boldsymbol{a} - \boldsymbol{b} \| + \| \boldsymbol{b} \|,$$

звідки

$$\| \boldsymbol{a} \| - \| \boldsymbol{b} \| \leq \| \boldsymbol{a} - \boldsymbol{b} \|.$$

Міняючи місцями $\boldsymbol{a}$ і $\boldsymbol{b}$, одержимо:

$$\| \boldsymbol{b} \| - \| \boldsymbol{a} \| \leq \| \boldsymbol{a} - \boldsymbol{b} \|.$$

Тому

$$\big| \| \boldsymbol{a} \| - \| \boldsymbol{b} \| \big| \leq \| \boldsymbol{a} - \boldsymbol{b} \|. \tag{1}$$

Найбільш уживані норми вектора виражаються через його координати $a_1, \ldots, a_n$ за допомогою формул:

$$\| \boldsymbol{a} \|_1 = \sum_{i=1}^{n} | a_i |, \quad \| \boldsymbol{a} \|_E = \left(\sum_{i=1}^{n} | a_i |^2 \right)^{1/2}, \quad \| \boldsymbol{a} \|_\infty = \max_{1 \leq i \leq n} | a_i |.$$

Норма $\| * \|_E$ називається *евклідовою*. Очевидно, що $\| \boldsymbol{a} \|_E = | \boldsymbol{a} |$.

Неважко перевірити, що норми $\| * \|_1$ і $\| * \|_\infty$ задовольняють усім аксіомам норми. Для евклідової норми виконання перших двох аксіом є очевидним. Покажемо, що третя аксіома також виконується. Дійсно, на підставі нерівності Коші — Шварца — Буняковського маємо:

$$\| \boldsymbol{a} + \boldsymbol{b} \|^2 = (\boldsymbol{a} + \boldsymbol{b})^T (\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{a}^T \boldsymbol{a} + 2\boldsymbol{a}^T \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{b} \leq$$

$$\leq \| \boldsymbol{a} \|^2 + 2\| \boldsymbol{a} \| \| \boldsymbol{b} \| + \| \boldsymbol{b} \|^2 = (\| \boldsymbol{a} \| + \| \boldsymbol{b} \|)^2,$$

звідки випливає нерівність трикутника.

Нехай вектори послідовності

$$\boldsymbol{x}_1, \ \boldsymbol{x}_2, \ \ldots, \ \boldsymbol{x}_m, \ \ldots \tag{2}$$

належать нормованому простору. Вектор $\boldsymbol{x}_0$ цього простору називається границею послідовності, якщо

$$\lim_{m \to \infty} \| \boldsymbol{x}_m - \boldsymbol{x}_0 \| = 0 . \tag{3}$$

Кажуть, що послідовність (2) збігається до $\boldsymbol{x}_0$ за нормою $\| * \|$.

Наявність збіжності в нормованому просторі дозволяє говорити про неперервність функцій. Функція $f(\boldsymbol{x})$ називається неперервною в точці $\boldsymbol{x}_0$, якщо із збіжності до $\boldsymbol{x}_0$ послідовності (2) за деякою нормою випливає збіжність за тією ж нормою до $f(\boldsymbol{x}_0)$ послідовності

$$f(\boldsymbol{x}_1),\ f(\boldsymbol{x}_2),\ \ldots,\ f(\boldsymbol{x}_m),\ \ldots\ .$$

Теорема 12.1.1. *У нормованому просторі норма є неперервною функцією.*

Д о в е д е н н я. Треба показати, що із збіжності до $\boldsymbol{x}_0$ послідовності (2) повинна випливати збіжність до $\|\boldsymbol{x}_0\|$ послідовності

$$\|\boldsymbol{x}_1\|,\ \|\boldsymbol{x}_2\|,\ \ldots,\ \|\boldsymbol{x}_m\|,\ \ldots .$$

Нерівність (1) показує, що

$$\big|\|\boldsymbol{x}_m\| - \|\boldsymbol{x}_0\|\big| \le \|\boldsymbol{x}_m - \boldsymbol{x}_0\|.$$

З урахуванням формули (3) маємо:

$$\lim_{m\to\infty}\big|\|\boldsymbol{x}_m\| - \|\boldsymbol{x}_0\|\big| = 0,$$

що і треба було довести. ☐

12.1.2. Еквівалентність норм. Норми $\|*\|_\varphi$ і $\|*\|_\psi$ називаються *еквівалентними*, якщо існують додатні числа α і β такі, що

$$\alpha\|*\|_\varphi \le \|*\|_\psi \le \beta\|*\|_\varphi . \tag{4}$$

Очевидно, що кожна норма еквівалентна самій собі. З формули (4) випливає, що

$$\beta^{-1}\|*\|_\psi \le \|*\|_\varphi \le \alpha^{-1}\|*\|_\psi ,$$

тобто властивість еквівалентності норм є симетричною.

Припустимо, що

$$\alpha_1\|*\|_\varphi \le \|*\|_\psi \le \beta_1\|*\|_\varphi, \quad \alpha_2\|*\|_\psi \le \|*\|_\chi \le \beta_2\|*\|_\psi .$$

Тоді, як неважко побачити,

$$\alpha_1\alpha_2\|*\|_\varphi \le \|*\|_\chi \le \beta_1\beta_2\|*\|_\varphi .$$

Це означає, що коли еквівалентними є норми $\| * \|_\varphi$ і $\| * \|_\psi$, а також $\| * \|_\psi$ і $\| * \|_\chi$, то еквівалентними є також норми $\| * \|_\varphi$ і $\| * \|_\chi$.

Як приклад розглянемо деякі з нерівностей (4) для векторних норм $\| * \|_1$, $\| * \|_E$ і $\| * \|_\infty$. Ці нерівності, які дозволяють визначити величини α і β, є такими:

$$\frac{1}{\sqrt{n}} \| * \|_1 \le \| * \|_E \le \| * \|_1,$$

$$\| * \|_\infty \le \| * \|_E \le n \| * \|_\infty,$$

$$\| * \|_\infty \le \| * \|_1 \le n \| * \|_\infty .$$

Обмежимося обгрунтуванням лівої половини першої пари нерівностей (інші нерівності є очивидними). Нехай $\boldsymbol{j}$ — вектор з одиничними координатами і $\boldsymbol{t}$ — вектор, координатами якого є модулі координат вектора $\boldsymbol{x}$. Скориставшись нерівністю Коші — Шварца — Буняковського, одержимо:

$$\| \boldsymbol{x} \|_1 = \boldsymbol{j}^T \boldsymbol{t} \le | \boldsymbol{j} \| \boldsymbol{t} | = \sqrt{n} \| \boldsymbol{x} \|_E,$$

що і треба довести.

Теорема 12.1.2. *Послідовність* (2) *векторів нормованого простору збігається за нормою* $\| * \|_\psi$ *тоді і тільки тоді, коли вона збігається за будь-якою еквівалентною нормою* $\| * \|_\varphi$.

Д о в е д е н н я. На підставі формули (4) маємо:

$$\alpha \| \boldsymbol{x}_m - \boldsymbol{x}_0 \|_\varphi \le \| \boldsymbol{x}_m - \boldsymbol{x}_0 \|_\psi \le \beta \| \boldsymbol{x}_m - \boldsymbol{x}_0 \|_\varphi .$$

Ці нерівності показують, що $\| \boldsymbol{x}_m - \boldsymbol{x}_0 \|_\psi \to 0$ при $m \to \infty$ за умови $\| \boldsymbol{x}_m - \boldsymbol{x}_0 \|_\varphi \to 0$.

Обернене твердження випливає з доведеного, якщо врахувати симетричність властивості еквівалентності норм. □

Теорема 12.1.3. *Для будь-якої норми* $\| * \|_\varphi$ *існують додатні числа* γ, δ *такі, що кожний вектор* $\boldsymbol{x}$, *для якого* $\| \boldsymbol{x} \|_E = 1$, *задовольняє нерівностям*

$$\gamma \le \| \boldsymbol{x} \|_\varphi \le \delta. \tag{5}$$

Д о в е д е н н я. Множина векторів $\boldsymbol{x}$, для яких $\| \boldsymbol{x} \|_E = 1$, є обмеженою і замкненою, а функція $\| \boldsymbol{x} \|_\varphi$ є неперервною на цій множині за теоремою 12.1.1. На підставі теореми Вейєрштрасса для функції $\| \boldsymbol{x} \|_\varphi$ знайдуться числа γ і δ, для яких мають місце нерівності (5).

Оскільки множина $\| \boldsymbol{x} \|_E = 1$ не вміщує вектор $\boldsymbol{x} = \boldsymbol{0}$, то функція $\| \boldsymbol{x} \|_\varphi$ на множині $\| \boldsymbol{x} \|_E = 1$ має лише додатні значення. Тому γ, $\delta > 0$. □

Тепер ми можемо довести теорему, яка часто дозволяє звести дослідження властивостей векторних норм до вивчення однієї з них.

Теорема 12.1.4. *У нормованому просторі всі векторні норми є еквівалентними.*

Д о в е д е н н я. Будемо спочатку розглядати норму $\| * \|_\varphi$ і евклідову норму $\| * \|_E$. Довільному вектору $\boldsymbol{x} \ne \boldsymbol{0}$ зіставимо вектор $\boldsymbol{x}_*$ такий, що $\boldsymbol{x} = \| \boldsymbol{x} \|_E \, \boldsymbol{x}_*$. Тоді на підставі аксіоми 2) норми буде

$$\| \boldsymbol{x} \|_\varphi = \| \boldsymbol{x} \|_E \, \| \boldsymbol{x}_* \|_\varphi .$$

За теоремою 12.1.3 маємо нерівності

$$\gamma \le \| \boldsymbol{x}_* \|_\varphi \le \delta.$$

Після множення цих нерівностей на $\| \boldsymbol{x} \|_E$ одержимо:

$$\gamma \| \boldsymbol{x} \|_E \le \| \boldsymbol{x} \|_\varphi \le \delta \| \boldsymbol{x} \|_E .$$

Ці нерівності, очевидно, задовольняються для $\boldsymbol{x} = \boldsymbol{0}$. Вони збігаються з нерівностями (4), тому норми $\| * \|_\varphi$ і $\| * \|_E$ є еквівалентними.

Аналогічно можна довести еквівалентність будь-якої іншої норми $\| * \|_\psi$ і евклідової норми. Це означає, що еквівалентними є норми $\| * \|_\varphi$ і $\| * \|_\psi$. □

Оскільки всі норми є еквівалентними, то інколи буває однаково, якою нормою скористатися. Проте складність обчислень часто суттєво залежить від того, з якою нормою ми маємо справу. Вибір найбільш придатної для конкретного випадку норми може істотно спростити міркування.

12.1.3. Збіжність за нормою і координатна збіжність. У нормованому просторі крім збіжності за нормою можна ввести інше поняття збіжності. Позначимо координати вектора $\boldsymbol{x}_m$ з послідовності (2) через $x_{m1}, \ldots, x_{mn}$, а координати граничного вектора $\boldsymbol{x}_0$ — через $x_{01}, \ldots, x_{0n}$. Якщо

$$\lim_{m\to\infty} x_{mi} = x_{0i}, \quad i = \overline{1,\ n}.$$

то говорять, що має місце координатна збіжність послідовності (2) до вектора $\boldsymbol{x}_0$. Поняття координатної збіжності є цілком природним: якщо два вектора є «близькими», то можна припустити, що повинні бути «близькими» їх координати.

Теорема 12.1.5. *Координатна збіжність послідовності векторів має місце тоді і тільки тоді, коли послідовність збігається за будь-якою векторною нормою.*

Д о в е д е н н я. Якщо має місце координатна збіжність, то

$$\lim_{m\to\infty} |\, x_{mi} - x_{0i} \,| = 0, \quad i = \overline{1,\ n},$$

або

$$\lim_{m\to\infty} \max_{1\le i\le n} |\, x_{mi} - x_{0i} \,| = 0.$$

Це означає збіжність послідовності (2) за нормою $\|*\|_\infty$. Обернене твердження теореми випливає з еквівалентності векторних норм і теореми 12.1.2. □

12.1.4. Норми матриць. Подібно до векторної норми може бути введена величина, яка називається нормою матриці. Для матриці $\boldsymbol{A}$ її *норма* визначається як дійсне число $\|\boldsymbol{A}\|$, яке задовольняє аксіомам:

1) $\|\boldsymbol{A}\| > 0$, якщо $\boldsymbol{A} \ne \boldsymbol{0}$, і $\|\boldsymbol{0}\| = 0$;

2) $\|\alpha\boldsymbol{A}\| = |\alpha|\,\|\boldsymbol{A}\|$ для будь-якого числа α;

3) $\|\boldsymbol{A} + \boldsymbol{B}\| \le \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$;

4) $\|\boldsymbol{AB}\| \le \|\boldsymbol{A}\|\|\boldsymbol{B}\|$.

Ці аксіоми стосуються довільної комплексної матриці. Нижче ми будемо розглядати норми дійсних квадратних матриць. Одержані при цьому результати легко поширюються на комплексний випадок.

Оскільки в більшості задач, де мова йде про норму векторів, доводиться розглядати і норму матриць, то доцільно вводити її таким чином, щоб вона була розумно пов'язана з нормою векторів. Говорять, що норма матриці *узгоджена* з нормою векторів, якщо для будь-якої матриці $\boldsymbol{A}$ і будь-якого вектора $\boldsymbol{x}$ виконується нерівність

$$\| \boldsymbol{Ax} \| \le \| \boldsymbol{A} \| \| \boldsymbol{x} \|. \qquad (6)$$

Можна довести, що всяка норма матриць узгоджена з якими-небудь нормами векторів. Перевагу узгоджених норм ілюструє

Теорема 12.1.6. *Модулі власних значень матриці не перевищують будь-яку її узгоджену норму.*

Д о в е д е н н я. Нехай λ — власне значення і $\boldsymbol{x}$ — відповідний до нього власний вектор. Тоді $\boldsymbol{Ax} = \lambda \boldsymbol{x}$ і

$$|\lambda| \| \boldsymbol{x} \| = \| \lambda \boldsymbol{x} \| = \| \boldsymbol{Ax} \| \le \| \boldsymbol{A} \| \| \boldsymbol{x} \|,$$

звідки випливає, що $|\lambda| \le \| \boldsymbol{A} \|$. □

Дослідимо спосіб отримання норми матриці, *підпорядкованої* деякій векторній нормі. Мова йде про норму

$$\| \boldsymbol{A} \| = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\| \boldsymbol{Ax} \|}{\| \boldsymbol{x} \|}, \qquad (7)$$

яка є максимальним відношенням норми перетвореного вектора $\boldsymbol{x}$ при лінійному перетворенні $\boldsymbol{A}$ до норми вектора $\boldsymbol{x}$.

Оскільки заміна $\boldsymbol{x}$ на $\alpha \boldsymbol{x}$ не змінює величину $\| \boldsymbol{Ax} \| \| \boldsymbol{x} \|^{-1}$, то можна припустити, що $\| \boldsymbol{x} \| = 1$, тоді замість рівності (7) будемо мати:

$$\| \boldsymbol{A} \| = \max_{\| \boldsymbol{x} \| = 1} \| \boldsymbol{Ax} \|. \qquad (8)$$

Рівність (8) показує, що для всякої норми матриць, підпорядкованої якій-небудь векторній нормі, $\| \boldsymbol{E} \| = 1$.

Унаслідок неперервності векторної норми і обмеженості та замкненості множини векторів $\| \boldsymbol{x} \| = 1$ можна стверджувати на підставі теореми Вейєрштрасса, що максимум у правій частині рівності (8) досягається, тобто знайдеться вектор $\boldsymbol{x}_0$ такий, що $\| \boldsymbol{x}_0 \| = 1$ і $\| \boldsymbol{A}\boldsymbol{x}_0 \| = \| \boldsymbol{A} \|$.

Теорема 12.1.7. *Для підпорядкованої норми* (8) *виконуються аксіоми норми і умова узгодженості* (6).

Д о в е д е н н я. 1) Очевидно, що $\| \boldsymbol{A} \| \geq 0$. Якщо $\max\limits_{\|\boldsymbol{x}\|=1} \| \boldsymbol{A}\boldsymbol{x} \| = 0$, то $\| \boldsymbol{A}\boldsymbol{x} \| = 0$ при $\| \boldsymbol{x} \| = 1$. Унаслідок довільності $\boldsymbol{x}$ матриця $\boldsymbol{A}$ є нульовою. Отже, аксіома 1) норми матриці виконується.

2) Використавши аксіому 2) для векторної норми, одержимо:

$$\| \alpha \boldsymbol{A} \| = \max_{\|\boldsymbol{x}\|=1} \| \alpha \boldsymbol{A}\boldsymbol{x} \| = | \alpha | \max_{\|\boldsymbol{x}\|=1} \| \boldsymbol{A}\boldsymbol{x} \| = | \alpha | \| \boldsymbol{A} \|.$$

3) Для матриці $\boldsymbol{A} + \boldsymbol{B}$ з урахуванням аксіоми 3) для векторної норми маємо:

$$\| \boldsymbol{A} + \boldsymbol{B} \| = \max_{\|\boldsymbol{x}\|=1} \| \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{x} \| \leq \max_{\|\boldsymbol{x}\|=1} (\| \boldsymbol{A}\boldsymbol{x} \| + \| \boldsymbol{B}\boldsymbol{x} \|) \leq$$

$$\leq \max_{\|\boldsymbol{x}\|=1} \| \boldsymbol{A}\boldsymbol{x} \| + \max_{\|\boldsymbol{x}\|=1} \| \boldsymbol{B}\boldsymbol{x} \| = \| \boldsymbol{A} \| + \| \boldsymbol{B} \|.$$

Перш ніж перевірити виконання аксіоми 4), доведемо нерівність (6) для норми (8). Для будь-якого $\boldsymbol{x} \neq \boldsymbol{0}$ визначимо вектор $\boldsymbol{y} = \boldsymbol{x} \| \boldsymbol{x} \|^{-1}$, норма якого дорівнює одиниці. Тоді

$$\| \boldsymbol{A}\boldsymbol{x} \| = \| \boldsymbol{x} \| \| \boldsymbol{A}\boldsymbol{y} \| \leq \| \boldsymbol{x} \| \max_{\|\boldsymbol{y}\|=1} \| \boldsymbol{A}\boldsymbol{y} \| = \| \boldsymbol{A} \| \| \boldsymbol{x} \|$$

і це є умова узгодженості.

4) Перевіримо виконання аксіоми 4). Ураховуючи узгодженість норми, знаходимо:

$$\| \boldsymbol{A}\boldsymbol{B} \| = \max_{\|\boldsymbol{x}\|=1} \| \boldsymbol{A}(\boldsymbol{B}\boldsymbol{x}) \| \leq \max_{\|\boldsymbol{x}\|=1} (\| \boldsymbol{A} \| \| \boldsymbol{B}\boldsymbol{x} \|) =$$

$$= \| \boldsymbol{A} \| \max_{\|\boldsymbol{x}\|=1} \| \boldsymbol{B}\boldsymbol{x} \| = \| \boldsymbol{A} \| \| \boldsymbol{B} \|.$$

Теорема доведена. □

З'ясуємо, якими є норми матриці, підпорядковані векторним нормам $\| * \|_1$ і $\| * \|_\infty$. Нехай $\| \boldsymbol{x} \|_1 = 1$. Тоді

$$\| \boldsymbol{Ax} \|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \sum_{i=1}^{n} \sum_{j=1}^{n} | a_{ij} | \, | x_j | = \sum_{j=1}^{n} | x_j | \sum_{i=1}^{n} | a_{ij} | \le$$

$$\le \left(\max_{1 \le j \le n} \sum_{i=1}^{n} | a_{ij} | \right) \sum_{j=1}^{n} | x_j | = \max_{1 \le j \le n} \sum_{i=1}^{n} | a_{ij} |,$$

тобто

$$\| \boldsymbol{A} \|_1 = \max_{\| \boldsymbol{x} \|_1 = 1} \| \boldsymbol{Ax} \|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} | a_{ij} |. \tag{9}$$

При $\| \boldsymbol{x} \|_\infty = 1$ маємо:

$$\| \boldsymbol{Ax} \|_\infty = \max_{1 \le i \le n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \max_{1 \le i \le n} \sum_{j=1}^{n} | a_{ij} | \, | x_j | \le$$

$$\le \max_{1 \le j \le n} | x_j | \left(\max_{1 \le i \le n} \sum_{j=1}^{n} | a_{ij} | \right) = \max_{1 \le i \le n} \sum_{j=1}^{n} | a_{ij} |,$$

тобто

$$\| \boldsymbol{A} \|_\infty = \max_{\| \boldsymbol{x} \|_\infty = 1} \| \boldsymbol{Ax} \|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} | a_{ij} |. \tag{10}$$

Можна показати, що максимуми в правих частинах формул (9) і (10) досягаються. Інакше кажучи, існує вектор $\boldsymbol{x}_0$ такий, що $\| \boldsymbol{x}_0 \|_1 = 1$ і $\| \boldsymbol{Ax}_0 \|_1 = \| \boldsymbol{A} \|_1$; аналогічною є ситуація і для норми $\| \boldsymbol{A} \|_\infty$. На розгляді цих питань ми не зупиняємося.

Якщо векторна норма є евклідовою, то відповідна підпорядкована норма матриці називається *спектральною* нормою і позначається символом $\| * \|_2$.

Теорема 12.1.8. *Спектральна норма матриці дорівнює її найбільшому сингулярному числу.*

Д о в е д е н н я. За допомогою теореми 7.4.1 обчислюємо квадрат спектральної норми:

$$\| A \|_2^2 = \max_{|x|=1} | Ax |^2 = \max_{|x|=1} (Ax)^T Ax = \max_{|x|=1} x^T A^T Ax = \sigma_1^2,$$

де σ_1 — найбільше сингулярне число матриці A. Отже,

$$\| A \|_2 = \sigma_1 . \ \square \tag{11}$$

Аналогічно можна показати, що

$$\| A^{-1} \|_2 = \sigma_n^{-1} , \tag{12}$$

де σ_n — найменше сингулярне число.

Величина σ_1^2 дорівнює максимальному власному значенню матриці $A^T A$ (або AA^T), тобто $\sigma_1^2 = \lambda_{\max}(A^T A)$. Тому з урахуванням теореми 6.2.3 маємо:

$$\| A \|_2^2 = \lambda_{\max}(A^T A) = [\lambda_{\max}((A^T A)^2)]^{1/2} =$$

$$= [\lambda_{\max}((A^T A)^T (A^T A))]^{1/2} = (\| A^T A \|_2^2)^{1/2} = \| A^T A \|_2 . \tag{13}$$

Рівність

$$\| A^{-1} \|_2^2 = \| (A^T A)^{-1} \|_2 \tag{14}$$

можна довести аналогічно.

Норми матриць (9) — (11) належать до найбільш уживаних норм. Такою ж є і *евклідова* норма

$$\| A \|_E = \left(\sum_{i,\ j=1}^{n} |a_{ij}|^2 \right)^{1/2} .$$

Очевидно, що вона задовольняє першим трьом аксіомам норми, оскільки її можна розглядати як норму вектора в n^2-вимірному арифметичному просторі. Перевіримо виконання аксіоми 4). На підставі нерівності Коші — Шварца — Буняковського маємо:

$$\left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2 \le \sum_{k=1}^{n} |a_{ik}|^2 \sum_{k=1}^{n} |b_{kj}|^2, \quad i, j = \overline{1,\ n}.$$

Підсумовуючи ці нерівності по всім i та j, одержимо необхідну нерівність

$$\| AB \|_E \le \| A \|_E \| B \|_E.$$

Звідси, зокрема, випливає, що евклідова норма матриці узгоджена з векторною евклідовою нормою:

$$\| \boldsymbol{Ax} \|_E \le \| \boldsymbol{A} \|_E \, \| \boldsymbol{x} \|_E.$$

Не існує такої норми в арифметичному просторі, якій була б підпорядкована евклідова норма матриці. Дійсно, для підпорядкованої норми $\| \boldsymbol{E} \| = 1$, проте $\| \boldsymbol{E} \|_E = \sqrt{n} \ne 1$.

Для евклідової норми на підставі формул (7.4.5) і (7.4.4) знаходимо:

$$\| \boldsymbol{A} \|_E = \sqrt{\operatorname{tr}(\boldsymbol{A}\boldsymbol{A}^T)} = \sqrt{\operatorname{tr}(\boldsymbol{A}^T\boldsymbol{A})}. \tag{15}$$

Має місце також

Теорема 12.1.9. *Евклідова норма матриці дорівнює квадратному кореню із суми квадратів її сингулярних чисел.*

Д о в е д е н н я випливає з формул (15), (6.6.14) і (7.4.4):

$$\| \boldsymbol{A} \|_E = \sqrt{\operatorname{tr}(\boldsymbol{Q}_1 \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T \boldsymbol{Q}_1^T)} = \sqrt{\operatorname{tr}(\boldsymbol{Q}_1^T \boldsymbol{Q}_1 \boldsymbol{\Sigma}\boldsymbol{\Sigma}^T)} = \sqrt{\operatorname{tr}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T)} = \sqrt{\sum_{i=1}^{n} \sigma_i^2}. \ \square$$

12.1.5. Збіжність за нормою і елементна збіжність. Вище обговорювалося питання щодо зв'язку між збіжністю за нормою і координатною збіжністю послідовності векторів. У разі послідовності матриць збіжність, подібна до координатної збіжності, розглядалася в розділах 6.2 і 9.2. Ми будемо називати її елементною збіжністю. Для дослідження збіжності за нормою послідовності матриць нам знадобиться виконання лише перших трьох аксіом норми. Норму, яка задовольняє аксіомам 1) — 3) (але не обов'язково 4)), називають *узагальненою* нормою матриці. Усі розглянуті вище норми матриці є узагальненими. Проте існують узагальнені норми матриці, які не є матричними нормами.

Як приклад розглянемо величину $\max\limits_{i,j} |a_{ij}|$. Неважко побачити, що вона задовольняє першим трьом аксіомам норми. Щоб показати, що аксіома 4) не задовольняється, покладемо $\boldsymbol{A} = \boldsymbol{B} = \boldsymbol{J}$, де $\boldsymbol{J}$ — матриця з одиничними елементами. Оскільки $\boldsymbol{J}^2 = n\boldsymbol{J}$, то

$$\max_{i,j} |\{\boldsymbol{AB}\}_{ij}| = n \max_{i,j} |\{\boldsymbol{J}\}_{ij}| = n.$$

З іншого боку,

$$\max_{i,j} |\{\boldsymbol{A}\}_{ij}| = \max_{i,j} |\{\boldsymbol{B}\}_{ij}| = 1.$$

Тепер можна побачити, що аксіома 4) не виконується.

Розглянемо теорему, яка є аналогом теореми 12.1.5 для матриць.

Теорема 12.1.10. *Елементна збіжність послідовності матриць має місце тоді і тільки тоді, коли послідовність збігається за будь-якою узагальненою матричною нормою.*

Д о в е д е н н я. Перші три аксіоми матричної норми, яким задовольняє узагальнена матрична норма, є цілком аналогічними аксіомам векторної норми. Тому, розглядаючи матриці як n^2-вимірні вектори, ми доведемо дану теорему шляхом посилання на теорему 12.1.5. □

Тепер ми можемо повернутися до збіжних матриць, які були розглянуті в розділах 6.2 і 9.2, і вказати більш зручну в багатьох випадках ознаку збіжності $\boldsymbol{A}^m \to \boldsymbol{0}$ ніж у теоремах 6.2.7 і 9.2.2.

Теорема 12.1.11. *Для того, щоб $\boldsymbol{A}^m \to \boldsymbol{0}$ при $m \to \infty$, достатньо, щоб деяка норма матриці $\boldsymbol{A}$ була меншою ніж одиниця.*

Д о в е д е н н я. За допомогою аксіоми 4) норми матриці одержимо:

$$\| \boldsymbol{A}^m \| \le \| \boldsymbol{A} \| \| \boldsymbol{A}^{m-1} \| \le \| \boldsymbol{A} \|^2 \| \boldsymbol{A}^{m-2} \| \le \ldots \le \| \boldsymbol{A} \|^m.$$

Якщо $\| \boldsymbol{A} \| < 1$, то $\| \boldsymbol{A}^m \| \to 0$ при $m \to \infty$. На підставі теореми 12.1.10 це означає, що $\boldsymbol{A}^m \to \boldsymbol{0}$. □

12.1.6. Нескінченні ряди векторів і матриць. У подальшому нам доведеться мати справу з нескінченними рядами векторів і матриць. Під вектором $\sum_{k=0}^{\infty} \boldsymbol{u}_k$, де $\boldsymbol{u}_k = (u_{k1} \ \ldots \ u_{kn})^T$, розуміють вектор, i-та координата якого є числовим рядом $\sum_{k=0}^{\infty} u_{ki}$. Тому збіжність векторного ряду рівносильна одночасній збіжності числових рядів $\sum_{k=0}^{\infty} u_{ki}$, $i = \overline{1, n}$, тобто маємо справу з координатною збіжністю векторного ряду. На підставі теореми 12.1.5 координатна збіжність векторного ряду рівносильна його збіжності за будь-якою векторною нормою. Інакше кажучи, векторний ряд $\sum_{k=0}^{\infty} \boldsymbol{u}_k$ збігається тоді і тільки тоді, коли збіжним є ряд $\sum_{k=0}^{\infty} \|\boldsymbol{u}_k\|$.

Подібно до цього стосовно матричного ряду $\sum_{k=0}^{\infty} \boldsymbol{U}_k$ можна розглядати елементну збіжність або збіжність за деякою узагальненою нормою у випадку квадратних матриць $\boldsymbol{U}_k$. Як приклад розглянемо питання про збіжність матричного ряду, який є сумою нескінченної геометричної прогресії.

Теорема 12.1.12. *Для того, щоб матричний ряд*

$$\boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + \ldots + \boldsymbol{A}^k + \ldots \tag{16}$$

збігався, необхідно і достатньо, щоб матриця $\boldsymbol{A}$ *була збіжною. При цьому сума ряду дорівнює* $(\boldsymbol{E} - \boldsymbol{A})^{-1}$.

Доведення. Необхідність випливає з необхідної умови збіжності числових рядів.

Достатність. Нехай матриця $\boldsymbol{A}$ є збіжною. На підставі теорем 6.2.7 і 9.2.2 власні значення матриці $\boldsymbol{A}$ за модулем є меншими ніж одиниця. Тому матриця $\boldsymbol{E} - \boldsymbol{A}$ є невиродженою. Розмірковуючи, як при виведенні формули (10.1.3), одержимо:

$$\boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + \ldots + \boldsymbol{A}^k + \ldots = (\boldsymbol{E} - \boldsymbol{A})^{-1}.$$

Висновок. *Для того, щоб матричний ряд* (*16*) *збігався, необхідно і достатньо, щоб* $|\lambda_1| < 1$, *де* λ_1 *— найбільше за модулем власне значення матриці* $\boldsymbol{A}$.

Доведення випливає з теорем 6.2.7 і 9.2.2.□

Теорема 12.1.12 дозволяє оцінити швидкість збіжності ряду (16).

Теорема 12.1.13. *Якщо для деякої матричної норми* $\| \boldsymbol{A} \| < 1$, *то*

$$\| (\boldsymbol{E} - \boldsymbol{A})^{-1} - (\boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + ... + \boldsymbol{A}^m) \| \le \frac{\| \boldsymbol{A} \|^{m+1}}{1 - \| \boldsymbol{A} \|}.$$

Д о в е д е н н я. Очевидно, що

$$(\boldsymbol{E} - \boldsymbol{A})^{-1} - (\boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + ... + \boldsymbol{A}^m) = \boldsymbol{A}^{m+1} + \boldsymbol{A}^{m+2} + ...$$

Звідси, підсумовуючи нескінченну геометричну прогресію, одержимо:

$$\| (\boldsymbol{E} - \boldsymbol{A})^{-1} - (\boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + ... + \boldsymbol{A}^m) \| \le$$

$$\le \| \boldsymbol{A} \|^{m+1} + \| \boldsymbol{A} \|^{m+2} + ... = \frac{\| \boldsymbol{A} \|^{m+1}}{1 - \| \boldsymbol{A} \|}. \ \square$$

12.2. Зумовленість матриць

12.2.1. Число зумовленості матриці. При чисельному розв’язанні систем лінійних рівнянь навіть точними методами виникає декілька джерел неточності розв’язку. Одним з них є необхідність округлення чисел у процесі обчислень. Друге джерело з’являється тоді, коли система $\boldsymbol{Ax} = \boldsymbol{b}$ виникає з практичної задачі і тому неточність вихідних даних веде до того, що коефіцієнти і праві частини системи відомі лише наближено. Це може породжувати великі зміни в розв’язку, тобто може мати місце нестійкість при розв’язанні задачі.

Для оцінки похибок розв’язку, породжуваних указаними обставинами, припустимо, що вихідні дані, тобто $\boldsymbol{A}$ і $\boldsymbol{b}$, певною мірою невизначені і ми хочемо знати, як це відбивається на розв’язку $\boldsymbol{x}$. Спочатку будемо вважати, що матриця $\boldsymbol{A}$ відома точно, а вектор $\boldsymbol{b}$ — з деякою невизначеністю, яка характеризується «збуренням» $\Delta\boldsymbol{b}$. З’ясуємо, наскільки великим може бути «збурення» $\Delta\boldsymbol{x}$ розв’язку $\boldsymbol{x}$. Ми будемо виходити з рівняння $\boldsymbol{Ax} = \boldsymbol{b}$, в яке підставлені збурені вектори $\boldsymbol{x}$ і $\boldsymbol{b}$:

$$\boldsymbol{A}(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{b} + \Delta\boldsymbol{b}. \tag{1}$$

У подальших оцінках, пов'язаних з проблемою розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$, ми будемо користуватися евклідовою нормою векторів і підпорядкованою їй спектральною нормою матриць. З рівності (1) знаходимо, що $\boldsymbol{A}\Delta\boldsymbol{x} = \Delta\boldsymbol{b}$, звідки $\Delta\boldsymbol{x} = \boldsymbol{A}^{-1}\Delta\boldsymbol{b}$. Тепер на підставі нерівності (12.1.6) маємо:

$$\| \Delta\boldsymbol{x} \| \le \| \boldsymbol{A}^{-1} \| \| \Delta\boldsymbol{b} \|. \tag{2}$$

З іншого боку, оскільки $\boldsymbol{Ax} = \boldsymbol{b}$, аналогічно одержуємо:

$$\| \boldsymbol{b} \| \le \| \boldsymbol{A} \| \| \boldsymbol{x} \|. \tag{3}$$

У співвідношеннях (2) і (3) для деяких $\Delta\boldsymbol{b}$ і $\boldsymbol{x}$ мають місце рівності.

Помноживши (2) і (3), одержимо:

$$\| \Delta\boldsymbol{x} \| \| \boldsymbol{b} \| \le \| \boldsymbol{A} \| \| \boldsymbol{A}^{-1} \| \| \boldsymbol{x} \| \| \Delta\boldsymbol{b} \|,$$

звідки

$$\frac{\| \Delta\boldsymbol{x} \|}{\| \boldsymbol{x} \|} \le c(\boldsymbol{A}) \frac{\| \Delta\boldsymbol{b} \|}{\| \boldsymbol{b} \|}, \tag{4}$$

де

$$c(\boldsymbol{A}) = \| \boldsymbol{A} \| \| \boldsymbol{A}^{-1} \|. \tag{5}$$

Величина $c(\boldsymbol{A})$ називається *числом зумовленості* невиродженої матриці $\boldsymbol{A}$. З рівностей (12.1.11) і (12.1.12) маємо:

$$c(\boldsymbol{A}) = \sigma_1 \sigma_n^{-1} \ge 1.$$

З визначення $c(\boldsymbol{A})$ випливає, що

$$c(\alpha\boldsymbol{A}) = c(\boldsymbol{A}), \tag{6}$$

де $\alpha \ne 0$ — деяке число, а внаслідок рівностей (12.1.13) і (12.1.14) можна стверджувати, що

$$c(\boldsymbol{A}^T \boldsymbol{A}) = [c(\boldsymbol{A})]^2. \tag{7}$$

Повертаючись до нерівності (4), можна тлумачити величину $\| \Delta\boldsymbol{b} \| \| \boldsymbol{b} \|^{-1}$ як міру відносної невизначеності вектора $\boldsymbol{b}$. Якщо, наприклад, координати $\boldsymbol{b}$ відомі до п'яти значущих десяткових цифр, то $\|\Delta\boldsymbol{b}\|\|\boldsymbol{b}\|^{-1}$ дорівнює приблизно 10^{-4} або 10^{-5}. Аналогічно $\|\Delta\boldsymbol{x}\|\|\boldsymbol{x}\|^{-1}$ можна тлумачити як відносну похибку розв'язку $\boldsymbol{x}$, зумовлену невизначеністю вектора $\boldsymbol{b}$. Нерівність (4) свідчить про те, що $c(\boldsymbol{A})$ обмежує зверху відношення відносної похибки розв'язку $\boldsymbol{x}$ до відносної невизначеності вектора $\boldsymbol{b}$. Якщо величина $c(\boldsymbol{A})$ є малою, то матриця $\boldsymbol{A}$ називається *добре зумовленою* відносно задачі розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$. При досить великій величині $c(\boldsymbol{A})$ матриця $\boldsymbol{A}$ є *погано зумовленою* відносно тієї ж задачі. Чим більшим є число $c(\boldsymbol{A})$, тим гірше зумовлена матриця $\boldsymbol{A}$; звичайно величина $c(\boldsymbol{A}) \approx 10^3 — 10^4$ вже означає погану зумовленість.

Прикладом добре зумовлених матриць є ортогональні матриці, стосовно яких на підставі рівності (7) можна стверджувати, що їх число зумовленості дорівнює одиниці. Настільки ж добре зумовленими є матриці вигляду $\alpha\boldsymbol{Q}$, де $\alpha \neq 0$ — число і $\boldsymbol{Q}$ — ортогональна матриця. Останнє твердження є наслідком рівності (6).

Формула (7) пояснює, чому нормальні рівняння $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$ важко розв'язувати: число зумовленості матриці $\boldsymbol{A}^T\boldsymbol{A}$ є квадратом числа зумовленості $\boldsymbol{A}$, тобто, формуючи $\boldsymbol{A}^T\boldsymbol{A}$, можна одержати систему з погано зумовленою матрицею. Значно краще (виключаючи лише дуже малі задачі) скористатися при знаходженні нормального псевдорозв'язку результатами, одержаними в розділі 11.6.

Зробимо декілька зауважень.

З а у в а ж е н н я 1. Норма і число зумовленості матриці в практичних задачах звичайно не обчислюються, а тільки оцінюються, бо нема потреби і можливості розв'язувати задачу про власні значення для знаходження σ_1 і σ_n.

З а у в а ж е н н я 2. Число зумовленості має різні значення при використанні різних норм матриць. Проте ці значення звичайно мають порівнянні величини.

З а у в а ж е н н я 3. Погана або добра зумовленість матриці $\boldsymbol{A}$ відносно задачі розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$ ніяк не пов'язана зі стійкістю проблеми власних значень для тієї ж матриці (про це мова буде йти наприкінці розділу).

П р и к л а д 1. Нехай

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0{,}99 \\ 0{,}99 & 0{,}98 \end{pmatrix}.$$

Власними значеннями матриці $\boldsymbol{A}$ є $\lambda_1 \approx 1{,}98005$ і $\lambda_2 \approx -0{,}00005$. Оскільки матриця $\boldsymbol{A}$ симетрична, то $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^2$ і на підставі теореми 6.2.3 її сингулярні числа дорівнюють $|\lambda_1|$ і $|\lambda_2|$, тобто $\sigma_1 \approx 1{,}98005$, $\sigma_2 \approx 0{,}00005$. Тому $c(\boldsymbol{A}) \approx 39601$. Таким чином, матриця $\boldsymbol{A}$ погано зумовлена і тому при розв'язанні системи $\boldsymbol{Ax} = \boldsymbol{b}$ можна чекати серйозних неприємностей.

Розглянемо систему

$$\begin{aligned} x_1 + 0{,}99x_2 &= 1{,}99, \\ 0{,}99x_1 + 0{,}98x_2 &= 1{,}97, \end{aligned}$$

точним розв'язком якої є $x_1 = x_2 = 1$. Проте система

$$\begin{aligned} x_1 + 0{,}99x_2 &= 1{,}9902, \\ 0{,}99x_1 + 0{,}98x_2 &= 1{,}9697 \end{aligned}$$

має розв'язок $x_1 = -3{,}93$, $x_2 = 5{,}98$, тобто наслідком невизначеності $\Delta\boldsymbol{b} = (0{,}0002 \quad -0{,}0003)^T$ є похибка $\Delta\boldsymbol{x} = (-4{,}93 \quad 4{,}98)^T$. Далі знаходимо:

$$\|\Delta\boldsymbol{x}\| \approx 5\sqrt{2}, \qquad \|\boldsymbol{x}\| = \sqrt{2}, \qquad \frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \approx 5,$$

$$\|\Delta\boldsymbol{b}\| \approx 10^{-4}\sqrt{13}, \qquad \|\boldsymbol{b}\| \approx 2\sqrt{2}, \qquad \frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|} \approx \frac{10^{-4}\sqrt{13}}{2\sqrt{2}}.$$

Отже,

$$\frac{\|\Delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \approx 39223\frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|},$$

що узгоджується зі знайденим вище значенням $c(\boldsymbol{A})$.

Таким чином, якщо координати вектора $\boldsymbol{b}$ відомі з невизначеністю 0,0003, то вектор $\boldsymbol{x}$ може бути знайдений з відносною похибкою порядку 5. Це є наслідком поганої зумовленості матриці $\boldsymbol{A}$. $\triangle$

Цей приклад свідчить про те, що система з погано зумовленою матрицею є мало стійкою, тобто її розв'язок сильно змінюється при малій зміні правих частин. Однією з причин нестійкості може бути мализна визначника матриці системи, бо на підставі рівності (5) число зумовленості залежить від $\| \boldsymbol{A}^{-1} \|$, а ця величина може бути великою, якщо $|\det \boldsymbol{A}|$ малий. У розглянутому прикладі саме такий випадок має місце. Проте не існує прямого зв'язку між малим значенням $|\det \boldsymbol{A}|$ і поганою зумовленістю матриці $\boldsymbol{A}$. Це можна побачити на такому прикладі. Нехай $n = 100$, $\sigma_1 = \ldots = \sigma_n = 10^{-1}$, тоді $|\det \boldsymbol{A}| = 10^{-100}$ і в той же час $c(\boldsymbol{A}) = 1$, тобто матриця добре зумовлена, хоч її визначник є дуже малим. Отже, значення $c(\boldsymbol{A})$ є більш важливим критерієм важкості розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$ ніж малі значення $|\det \boldsymbol{A}|$ або великий порядок $\boldsymbol{A}$. Погана зумовленість системи виявляється при використанні будь-якого методу її розв'язання.

Покажемо тепер, що величина $c(\boldsymbol{A})$ характеризує важкість розв'язання системи $\boldsymbol{Ax} = \boldsymbol{b}$ і тоді, коли $\boldsymbol{A}$ має невизначеність, а вектор $\boldsymbol{b}$ заданий точно. У цьому випадку маємо рівність

$$(\boldsymbol{A} + \Delta\boldsymbol{A})(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{b}. \tag{8}$$

Будемо вважати, що як матриця $\boldsymbol{A}$, так і матриця $\boldsymbol{A} + \Delta\boldsymbol{A}$ є невиродженими. Остання вимога, як можна показати, виконується при досить малій величині $\| \Delta\boldsymbol{A} \|$.

З рівності (8) випливає, що

$$\boldsymbol{A}\Delta\boldsymbol{x} + \Delta\boldsymbol{A}(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{0},$$

звідки

$$\Delta\boldsymbol{x} = -\boldsymbol{A}^{-1}\Delta\boldsymbol{A}(\boldsymbol{x} + \Delta\boldsymbol{x}).$$

Переходячи в останній рівності до норми, одержимо, що

$$\| \Delta\boldsymbol{x} \| \le \| \boldsymbol{A}^{-1} \| \| \Delta\boldsymbol{A} \| \| \boldsymbol{x} + \Delta\boldsymbol{x} \|$$

і тому

$$\frac{\|\Delta \boldsymbol{x}\|}{\|\boldsymbol{x}+\Delta \boldsymbol{x}\|} \leq c(\boldsymbol{A}) \frac{\|\Delta \boldsymbol{A}\|}{\|\boldsymbol{A}\|}.$$

Отже, відносна невизначеність $\boldsymbol{x}$ обмежена відносною невизначеністю матриці $\boldsymbol{A}$, помноженою на її число зумовленості.

У загальному випадку, коли має місце невизначеність як елементів матриці системи, так і її правих частин, число зумовленості теж відіграє вирішальну роль при оцінюванні відносної невизначеності розв'язку системи. Не зупиняючись на цьому випадку, обмежимося натомість розглядом прикладу.

П р и к л а д 2. Розв'яжемо систему рівнянь

$$\begin{aligned} \sqrt{2}x_1 + \quad x_2 &= \sqrt{3}, \\ 2x_1 + \sqrt{2}x_2 &= \sqrt{6}, \end{aligned} \tag{9}$$

обчислюючи її коефіцієнти і праві частини спочатку з точністю до двох, а потім до трьох значущих цифр.

У першому випадку маємо:

$$\begin{aligned} 1{,}4x_1 + \quad x_2 &= 1{,}7, \\ 2x_1 + 1{,}4x_2 &= 2{,}4 \end{aligned} \tag{10}$$

і розв'язок цієї системи $\boldsymbol{x} = (1/2 \quad 1)^T$. У другому випадку система має вигляд:

$$\begin{aligned} 1{,}41x_1 + \quad x_2 &= 1{,}73, \\ 2x_1 + 1{,}41x_2 &= 2{,}45 \end{aligned} \tag{11}$$

і її розв'язок дорівнює $\boldsymbol{x} \approx (0{,}90 \quad 0{,}46)^T$.

Порівняння обох розв'язків показує, що ми маємо справу з нестійкістю при розв'язанні системи рівнянь. Немає підстав сподіватися, що підвищення точності обчислень допоможе уникнути цієї неприємності. Справа в тому, що матриця вихідної системи (9) вироджена і система має нескінченну множину розв'язків, а матриці систем (10) і (11) невироджені і системи мають єдині розв'язки. $\triangle$

Отже, постає природне питання про те, яке практичне значення мають такі поняття, як сумісна чи несумісна система, лінійна залежність, базис, ранг матриці і т. д. З цього питання починаються ті відмінності, які відрізняють математику «точну» від математики «наближеної». Відповідь на це питання потребує глибокого розуміння особливостей розв'язуваних задач і запровадження таких запобіжних заходів, які допомогли б уникнути ситуацій, подібних до розглянутих у прикладах, або принаймні пом'якшити їх вплив. Нижче ми звертаємо увагу на два таких заходи: частковий вибір провідного елемента в методі Гаусса і використання ортонормованих систем векторів. Перш ніж перейти до цих питань, зробимо

Зауваження 4. Досі ми не аналізували похибки округлення, а лише обговорювали невизначеність розв'язку систем лінійних рівнянь, обумовлену невизначеністю вихідних даних задачі. Докладні дослідження показують, що похибки округлень можна включити до невизначеності даних, тобто повна похибка, обумовлена похибками округлень, визначається числом $c(\boldsymbol{A})$.

12.2.2. Метод виключення з частковим вибором провідного елемента. Нестійкість при розв'язуванні систем лінійних рівнянь може виникнути навіть при добре зумовленій матриці системи, якщо метод розв'язання має певні вади. Як приклад розглянемо систему рівнянь $\boldsymbol{Ax} = \boldsymbol{b}$, де

$$\boldsymbol{A} = \begin{pmatrix} 0{,}0001 & 1 \\ 1 & 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Оскільки сингулярні числа матриці $\boldsymbol{A}$ дорівнюють $\sigma_1 \approx 1{,}618$, $\sigma_2 \approx 0{,}618$, то $c(\boldsymbol{A}) \approx 2{,}6$ і тому матриця $\boldsymbol{A}$ є добре зумовленою. Розв'яжемо систему методом виключення, округлюючи в процесі обчислень до трьох значущих цифр. Елементарні перетворення розширеної матриці мають вигляд:

$$\begin{pmatrix} 0{,}0001 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 10000 & 10000 \\ 1 & 1 & 2 \end{pmatrix} \sim$$

$$\sim\begin{pmatrix}1 & 10000 & 10000\\ 0 & -10000 & -10000\end{pmatrix}\sim\begin{pmatrix}1 & 10000 & 10000\\ 0 & 1 & 1\end{pmatrix}.$$

Останній матриці відповідає система

$$\begin{aligned}x_1 + 10000x_2 &= 10000,\\ x_2 &= 1,\end{aligned}$$

розв'язок якої дорівнює $\boldsymbol{x} = (0 \quad 1)^T$.

Очевидно, що це помилковий розв'язок. Отже, незважаючи на добру зумовленість матриці $\boldsymbol{A}$, система виявляється нестійкою, бо похибки округлень призвели до великих похибок у розв'язку. Неважко зрозуміти, що причиною ускладнень є мале значення першого провідного елемента.

На відміну від системи з погано зумовленою матрицею, коли нестійкість виявляється при всякому методі розв'язання, в даному випадку можна уникнути нестійкості, якщо порівнювати кожний провідний елемент з усіма іншими можливими провідними елементами того ж самого стовпця, а потім переставити рядки матриці таким чином, щоб найбільший з них за модулем став новим провідним елементом. Це так званий метод виключення з частковим вибором провідного елемента. У нашому випадку треба поміняти місцями перший і другий рядки. Отже,

$$\begin{pmatrix}0{,}0001 & 1 & 1\\ 1 & 1 & 2\end{pmatrix}\sim\begin{pmatrix}1 & 1 & 2\\ 0{,}0001 & 1 & 1\end{pmatrix}\sim\begin{pmatrix}1 & 1 & 2\\ 0 & 1 & 1\end{pmatrix},$$

звідки знаходимо, що $\boldsymbol{x} = (1 \quad 1)^T$. Легко переконатися в тому, що цей розв'язок є близьким до справжнього розв'язку.

Метод виключення з частковим вибором провідного елемента відрізняється від більш надійного методу повного вибору провідного елемента, коли за провідний беруть елемент, максимальний за модулем у всій матриці рівнянь, що залишилася після виконання деякої кількості виключень. Проте через складність цього методу до нього удаються у виключних випадках, бо часткового вибору провідного елемента, як правило, буває достатньо.

12.2.3. Лінійна залежність і ортонормовані базиси. Матриця системи (9) має ранг, рівний одиниці, а після наближеного обчислення її елементів він став дорівнювати двом. Це означає, що лінійна залежність векторів (у даному випадку стовпців матриці) може бути порушена при малій зміні цих векторів. Це явище призводить до значних ускладнень при використанні поняття базису в практичних задачах. Проте це відбувається не завжди, як свідчить наступна теорема, яку ми наведемо без доведення.

Теорема 12.2.1. *Нехай* $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ *— ортонормований базис. Якщо для системи векторів* $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n$ *виконується нерівність*

$$\sum_{i=1}^{n} |\boldsymbol{\varepsilon}_i|^2 < 1,$$

то система $\boldsymbol{e}_1 + \boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{e}_n + \boldsymbol{\varepsilon}_n$ *є лінійно незалежною.* □

З теореми випливає, що при малих змінах ортонормованих систем векторів зберігається їх лінійна незалежність. Ця їх особливість пояснює широке використання ортонормованих базисів і ортогональних перетворень при побудові обчислювальних алгоритмів.

12.2.4. Зумовленість і власні значення. Нехай алгебраїчна проблема власних значень розв’язується у випадку, коли матриця $\boldsymbol{A}$ задана з похибкою $\Delta\boldsymbol{A}$. Стосовно матриці $\boldsymbol{A}$ будемо припускати, що вона має різні власні значення $\lambda_1, \ldots, \lambda_n$. Якщо $\boldsymbol{S}$ — її діагоналізуюча матриця, то

$$\boldsymbol{S}^{-1}(\boldsymbol{A} + \Delta\boldsymbol{A})\boldsymbol{S} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) + \boldsymbol{S}^{-1}\Delta\boldsymbol{A}\boldsymbol{S}.$$

На підставі теореми 6.2.5 матриця $\boldsymbol{S}^{-1}(\boldsymbol{A} + \Delta\boldsymbol{A})\boldsymbol{S}$ має ті ж власні значення, що і матриця $\boldsymbol{A} + \Delta\boldsymbol{A}$. З теореми Гершгоріна, застосованої до рядків матриці $\boldsymbol{S}^{-1}(\boldsymbol{A} + \Delta\boldsymbol{A})\boldsymbol{S}$, випливає, що

$$|\lambda - \lambda_i - \{\boldsymbol{S}^{-1}\Delta\boldsymbol{A}\boldsymbol{S}\}_{ii}| \le \sum_{\substack{j=1 \\ j \ne i}}^{n} |\{\boldsymbol{S}^{-1}\Delta\boldsymbol{A}\boldsymbol{S}\}_{ij}|, \quad i = \overline{1,\ n}, \qquad (12)$$

де λ є довільним власним значенням матриці $\boldsymbol{A} + \Delta\boldsymbol{A}$. Очевидно, що

$$|\lambda-\lambda_i| = |\lambda-\lambda_i-\{S^{-1}\Delta AS\}_{ii}+\{S^{-1}\Delta AS\}_{ii}| \le$$

$$\le |\lambda-\lambda_i-\{S^{-1}\Delta AS\}_{ii}|+|\{S^{-1}\Delta AS\}_{ii}|, \quad i=\overline{1,\,n}. \qquad (13)$$

З нерівностей (13) і (12) знаходимо, що

$$|\lambda-\lambda_i| \le \sum_{j=1}^{n} |\{S^{-1}\Delta AS\}_{ij}|, \quad i=\overline{1,\,n},$$

або

$$|\lambda-\lambda_i| \le \| S^{-1}\Delta AS \|, \quad i=\overline{1,\,n},$$

де через $\|*\|$ позначена норма $\|*\|_\infty$. На підставі аксіоми 4) норми матриці маємо остаточно:

$$|\lambda-\lambda_i| \le \| S^{-1}\| \|\Delta A\| \|S\| = c(S)\|\Delta A\|, \quad i=\overline{1,\,n},$$

тобто похибка визначення власних значень залежить від числа зумовленості діагоналізуючої матриці. Зрозуміло, що найкращим буде той випадок, коли матриця A нормальна, бо діагоналізуюча матриця є унітарною і тому $c(S)=1$.

12.3. Збурення розв'язків лінійних алгебраїчних систем

У розділі 12.2 аналізувався вплив малих збурень вихідних даних при реалізації чисельних методів розв'язання основних задач лінійної алгебри. Нижче мова буде йти про одержання виразів для розв'язків збурених систем лінійних рівнянь у вигляді рядів за степенями малого «параметра збурення» ε. Розглядаються найпростіші варіанти можливих задач про збурення. У разі необхідності читач може за аналогією з розглянутими зразками сформулювати задачу і побудувати її розв'язок для більш складних випадків.

12.3.1. Випадок сумісності незбуреної системи. Будемо виходити з системи (12.2.8), яку подамо у вигляді:

$$(A_0+\varepsilon A_1)x=b. \qquad (1)$$

Тут $\boldsymbol{A}_0$ і $\boldsymbol{A}_1$ — дійсні матриці, ε — деякий додатний числовий параметр. При малих значеннях ε можна вважати, що система (1) одержана збуренням системи

$$\boldsymbol{A}_0\boldsymbol{x} = \boldsymbol{b}. \tag{2}$$

Припустимо, що при малих ε виконується умова

$$\det(\boldsymbol{A}_0 + \varepsilon\boldsymbol{A}_1) \neq 0. \tag{3}$$

Тоді рівняння (1) має розв'язок $\boldsymbol{x} = \boldsymbol{x}(\varepsilon)$ і наша задача буде полягати в знаходженні розкладу $\boldsymbol{x}(\varepsilon)$ за степенями ε і з'ясуванні його зв'язків з розв'язком рівняння (2) в різних випадках.

Зовсім простим є випадок, коли $\det \boldsymbol{A}_0 \neq 0$ і при досить малих значеннях ε виконується умова (3). Тоді розв'язок $\boldsymbol{x}(\varepsilon)$ рівняння (1) дорівнює

$$\boldsymbol{x}(\varepsilon) = (\boldsymbol{A}_0 + \varepsilon\boldsymbol{A}_1)^{-1}\boldsymbol{b} = (\boldsymbol{E} + \varepsilon\boldsymbol{A}_0^{-1}\boldsymbol{A}_1)^{-1}\boldsymbol{A}_0^{-1}\boldsymbol{b}$$

і його при необхідності можна подати у вигляді векторного ряду

$$\boldsymbol{x}(\varepsilon) = \sum_{k=0}^{\infty} \varepsilon^k (-1)^k (\boldsymbol{A}_0^{-1}\boldsymbol{A}_1)^k \boldsymbol{A}_0^{-1}\boldsymbol{b}, \tag{4}$$

якщо скористатися теоремою 12.1.12.

Цей ряд збігається, якщо збіжним є ряд

$$\sum_{k=0}^{\infty} \varepsilon^k \, \| \boldsymbol{A}_0^{-1}\boldsymbol{A}_1 \|^k \, \| \boldsymbol{A}_0^{-1}\boldsymbol{b} \|.$$

Тут і в подальшому $\| * \|$ — яка-небудь векторна норма і узгоджена з нею матрична норма. Останній ряд збігається, як геометрична прогресія, за умови, що $\varepsilon \, \| \boldsymbol{A}_0^{-1}\boldsymbol{A}_1 \| < 1.$ З іншого боку, на підставі висновку з теореми 12.1.12 ряд (4) збігається при $\varepsilon \, | \lambda_1 | < 1$, де λ_1 – найбільше за модулем власне значення матриці $\boldsymbol{A}_0^{-1}\boldsymbol{A}_1$. Теорема 12.1.6 показує, що першій нерівності для ε відповідає менший інтервал збіжності ряду (4). Проте ця нерівність більш придатна для практичного застосування, бо принаймні деякі з норм матриці $\boldsymbol{A}_0^{-1}\boldsymbol{A}_1$ обчислюються просто на відміну від величини $| \lambda_1 |$.

Похибка при знаходженні розв'язку $\boldsymbol{x}(\varepsilon)$ за допомогою ряду (4) обчислюється на підставі теореми 12.1.13.

12.3.2. Випадок несумісності незбуреної системи. Нехай $\det \boldsymbol{A}_0 = 0$, тоді матриця $\boldsymbol{A}_0$ має власне значення $\lambda = 0$, яке ми будемо вважати простим. Йому відповідає єдиний власний вектор $\boldsymbol{h}_{11}$, тобто однорідне рівняння

$$\boldsymbol{A}_0 \boldsymbol{x} = \boldsymbol{0} \tag{5}$$

має фундаментальну систему розв'язків, яка складається з одного вектора $\boldsymbol{h}_{11}$. На підставі теореми 3.5.1 можна стверджувати, що $r(\boldsymbol{A}_0) = n - 1$. Оскільки $r(\boldsymbol{A}_0) = r(\boldsymbol{A}_0^T)$, то спряжене однорідне рівняння

$$\boldsymbol{A}_0^T \boldsymbol{y} = \boldsymbol{0} \tag{6}$$

теж має лише один (з точністю до числового множника) розв'язок $\boldsymbol{y}_0$.

Далі для простоти будемо вважати, що $\det \boldsymbol{A}_1 \neq 0$. Нижче ми зустрінемося з рівнянням

$$\boldsymbol{B}\boldsymbol{x} = \boldsymbol{0}, \tag{7}$$

де

$$\boldsymbol{B} = \boldsymbol{A}_1^{-1} \boldsymbol{A}_0.$$

На підставі висновку з теореми 3.4.4 $r(\boldsymbol{B}) = r(\boldsymbol{A}_0)$, тому рівняння (7) має, як і рівняння (5), фундаментальну систему розв'язків, що складається з одного вектора, яким є $\boldsymbol{h}_{11}$. Він є власним вектором матриці $\boldsymbol{B}$, відповідним до її власного значення $\lambda = 0$. На відміну від матриць $\boldsymbol{A}_0$ і $\boldsymbol{A}_0^T$ це власне значення матриці $\boldsymbol{B}$ не обов'язково є простим. Нижче ми окремо розглянемо випадки, коли $\lambda = 0$ є простим і кратним власним значенням.

Почнемо з більш простого випадку, коли власне значення $\lambda = 0$ матриці $\boldsymbol{B}$ має кратність 1. Розклад за степенями ε розв'язку $\boldsymbol{x} = \boldsymbol{x}(\varepsilon)$ рівняння (1) будемо шукати у вигляді:

$$\boldsymbol{x}(\varepsilon) = \frac{\theta \boldsymbol{h}_{11}}{\varepsilon} + \sum_{k=0}^{\infty} \varepsilon^k (\boldsymbol{x}_k + \theta_k \boldsymbol{h}_{11}). \tag{8}$$

Для знаходження невідомих векторів $\boldsymbol{x}_k$ і величин θ, θ_k підставимо (8) в рівняння (1):

$$(\boldsymbol{A}_0 + \varepsilon\boldsymbol{A}_1)\left(\frac{\theta\boldsymbol{h}_{11}}{\varepsilon} + (\boldsymbol{x}_0 + \theta_0\boldsymbol{h}_{11}) + \ldots + \varepsilon^k(\boldsymbol{x}_k + \theta_k\boldsymbol{h}_{11}) + \ldots\right) = \boldsymbol{b}. \quad (9)$$

Порівняємо в рівності (9) коефіцієнти при однакових степенях ε. Спочатку зробимо це для степенів ε^{-1} і ε^0:

$$\theta\boldsymbol{A}_0\boldsymbol{h}_{11} = \boldsymbol{0}, \quad (10)$$

$$\boldsymbol{A}_0(\boldsymbol{x}_0 + \theta_0\boldsymbol{h}_{11}) + \theta\boldsymbol{A}_1\boldsymbol{h}_{11} = \boldsymbol{b}. \quad (11)$$

Рівність (10) виконується при будь-якому θ, оскільки $\boldsymbol{A}_0\boldsymbol{h}_{11} = \boldsymbol{0}$. На тій же підставі рівняння (11) можна спростити:

$$\boldsymbol{A}_0\boldsymbol{x}_0 = \boldsymbol{b} - \theta\boldsymbol{A}_1\boldsymbol{h}_{11}. \quad (12)$$

У відповідності до теореми Фредгольма неоднорідне рівняння (12) має ров'язок тоді і тільки тоді, коли його права частина є ортогональною до $\boldsymbol{y}_0$:

$$(\boldsymbol{b} - \theta\boldsymbol{A}_1\boldsymbol{h}_{11})^T\boldsymbol{y}_0 = 0. \quad (13)$$

Покажемо, що величина

$$a = \boldsymbol{h}_{11}^T\boldsymbol{A}_1^T\boldsymbol{y}_0 \quad (14)$$

не може дорівнювати нулю.

Припустимо, що $\boldsymbol{A}_1\boldsymbol{h}_{11} \perp \boldsymbol{y}_0$. Тоді рівняння

$$\boldsymbol{A}_0\boldsymbol{z} = \boldsymbol{A}_1\boldsymbol{h}_{11} \quad (15)$$

є розв'язуваним на підставі теореми Фредгольма. З (15) випливає, що

$$\boldsymbol{B}\boldsymbol{z} = \boldsymbol{h}_{11}. \quad (16)$$

Оскільки $\boldsymbol{A}_0\boldsymbol{h}_{11} = \boldsymbol{0}$, то

$$\boldsymbol{B}\boldsymbol{h}_{11} = \boldsymbol{0}. \quad (17)$$

Рівності (16) і (17) показують, що $\boldsymbol{z}$ належить інваріантному підпростору, натягнутому на власний вектор $\boldsymbol{h}_{11}$ матриці $\boldsymbol{B}$, відповідний до $\lambda = 0$. Тому

$$z = \alpha \boldsymbol{h}_{11},$$

де α — деяке число. Звідси після множення на матрицю $\boldsymbol{B}$ знаходимо за допомогою рівності (17):

$$\boldsymbol{B}z = \boldsymbol{0}.$$

Ця рівність знаходиться в суперечності з (16), тому не може бути $\boldsymbol{A}_1\boldsymbol{h}_{11} \perp \boldsymbol{y}_0$. Це означає, що $a \neq 0$.

Тепер з рівності (13) знаходимо:

$$\theta = a^{-1}\boldsymbol{b}^T \boldsymbol{y}_0.$$

Величина θ обертається в нуль при $\boldsymbol{b}^T \boldsymbol{y}_0 = 0$, тобто при розв'язуваності рівняння (2). На підставі теореми Фредгольма θ буде ненульовою величиною лише в разі нерозв'язуваності рівняння (2). Отже, лише за цієї умови розклад для $\boldsymbol{x}(\varepsilon)$ буде вміщувати член з ε^{-1}.

Знаючи θ, можна знайти $\boldsymbol{x}_0$ як окремий розв'язок рівняння (12). При цьому вектор $\boldsymbol{x}_0$ можна вибрати ортогональним до власного вектора $\boldsymbol{h}_{11}$.

Таким чином, ми обчислили θ і $\boldsymbol{x}_0$. Далі знайдемо θ_k при $k \geq 0$ і $\boldsymbol{x}_k$ при $k > 0$. З рівності (9) випливає, що

$$\boldsymbol{A}_0(\boldsymbol{x}_k + \theta_k \boldsymbol{h}_{11}) + \boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \theta_{k-1}\boldsymbol{h}_{11}) = \boldsymbol{0},$$

або

$$\boldsymbol{A}_0\boldsymbol{x}_k = -\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \theta_{k-1}\boldsymbol{h}_{11}). \tag{18}$$

Умова розв'язуваності цього рівняння має вигляд:

$$(\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \theta_{k-1}\boldsymbol{h}_{11}))^T \boldsymbol{y}_0 = 0,$$

звідки

$$\theta_{k-1} = -a^{-1}\boldsymbol{x}_{k-1}^T \boldsymbol{A}_1^T \boldsymbol{y}_0. \tag{19}$$

Тепер $\boldsymbol{x}_k$ визначається як окремий розв'язок рівняння (18), ортогональний до $\boldsymbol{h}_{11}$. Отже, маємо процедуру послідовного знаходження величин θ_k і векторів $\boldsymbol{x}_k$. Проаналізуємо більш докладно структуру загального розв'язку рівняння $\boldsymbol{A}_0 z = \boldsymbol{b}_0$, до якого зводяться рівняння (12) і (18).

Нехай V — одновимірний підпростір, натягнутий на вектор $\boldsymbol{h}_{11}$, і U — підпростір, натягнутий на інші вектори канонічного базису відносно лінійного перетворення $\boldsymbol{A}_0$. Оскільки $E_n = U \oplus V$, то можна вважати, що загальний розв'язок $\boldsymbol{z}_0 + \zeta \boldsymbol{h}_{11}$ рівняння $\boldsymbol{A}_0 \boldsymbol{z} = \boldsymbol{b}_0$ є сумою векторів $\boldsymbol{z}_0 \in U$ і $\zeta \boldsymbol{h}_{11} \in V$, причому окремий розв'язок системи $\boldsymbol{A}_0 \boldsymbol{z} = \boldsymbol{b}_0$ можна вибрати ортогональним до вектора $\boldsymbol{h}_{11}$.

Підставимо загальний розв'язок $\boldsymbol{z}_0 + \zeta \boldsymbol{h}_{11}$ у рівняння $\boldsymbol{A}_0 \boldsymbol{z} = \boldsymbol{b}_0$. Тоді одержимо, що $\boldsymbol{A}_0 \boldsymbol{z}_0 = \boldsymbol{b}_0$. На підставі теореми 9.2.1 $\boldsymbol{y}_0 \perp U$, а з умови розв'язуваності рівняння $\boldsymbol{A}_0 \boldsymbol{z} = \boldsymbol{b}_0$ випливає, що $\boldsymbol{b}_0 \perp \boldsymbol{y}_0$. Тому $\boldsymbol{b}_0 \in U$, тобто $\boldsymbol{A}_0 \boldsymbol{z}_0 \in U$.

Таким чином, $\boldsymbol{z}_0, \boldsymbol{A}_0 \boldsymbol{z}_0 \in U$, тому $\boldsymbol{A}_0$ є лінійним перетворенням, що діє з U в U. Покажемо, що воно є невиродженим. Припустимо, що для деякого вектора $\boldsymbol{z} \in U$ існує два прообрази $\boldsymbol{u}, \boldsymbol{u}' \in U$, тобто $\boldsymbol{z} = \boldsymbol{A}_0 \boldsymbol{u}, \ \boldsymbol{z} = \boldsymbol{A}_0 \boldsymbol{u}'.$ Очевидно, що $\boldsymbol{u} - \boldsymbol{u}' \in U$. З іншого боку, $\boldsymbol{A}_0(\boldsymbol{u} - \boldsymbol{u}') = \boldsymbol{0},$ тому $\boldsymbol{u} - \boldsymbol{u}' \in V$. Єдиним вектором, спільним для U і V, є нульовий вектор, тому $\boldsymbol{u} = \boldsymbol{u}'.$ Отже, кожний вектор з U має єдиний прообраз і тому перетворення $\boldsymbol{A}_0$ є невиродженим. Нехай $\boldsymbol{C}$ — обернене до нього перетворення, тоді $\boldsymbol{z}_0 = \boldsymbol{C} \boldsymbol{b}_0$. Нижче ми скористаємося одержаним результатом при дослідженні збіжності ряду (8).

Спочатку оцінімо $|\theta_{k-1}|$ і $\|\boldsymbol{x}_k\|$. З рівності (19) на підставі нерівності Коші — Шварца — Буняковського і еквівалентності векторних норм знаходимо, що

$$|\theta_{k-1}| \le |-a^{-1}| \, \|\boldsymbol{x}_{k-1}\|_E \, \|\boldsymbol{A}_1^T \boldsymbol{y}_0\|_E = C_1 \|\boldsymbol{x}_{k-1}\|_E \le$$
$$\le C_1 C_2 \|\boldsymbol{x}_{k-1}\| = C_3 \|\boldsymbol{x}_{k-1}\|. \tag{20}$$

Тут і в подальшому ми позначаємо через C_i деякі незалежні від k постійні величини.

Як зазначалося вище, окремий розв'язок рівняння (18) має вигляд:

$$\boldsymbol{x}_k = -\boldsymbol{C} \boldsymbol{A}_1 (\boldsymbol{x}_{k-1} + \theta_{k-1} \boldsymbol{h}_{11}). \tag{21}$$

Скориставшись аксіомами норми векторів і матриць, з рівності (21) одержимо:

$$\| \boldsymbol{x}_k \| \le \| \boldsymbol{C}\boldsymbol{A}_1 \| \, \| \boldsymbol{x}_{k-1} + \theta_{k-1}\boldsymbol{h}_{11} \| = C_4 \| \boldsymbol{x}_{k-1} + \theta_{k-1}\boldsymbol{h}_{11} \|. \qquad (22)$$

З урахуванням оцінки (20) далі маємо:

$$\| \boldsymbol{x}_{k-1} + \theta_{k-1}\boldsymbol{h}_{11} \| \le \| \boldsymbol{x}_{k-1} \| + \| \theta_{k-1}\boldsymbol{h}_{11} \| = \| \boldsymbol{x}_{k-1} \| + C_5 |\theta_{k-1}| \le$$
$$\le \| \boldsymbol{x}_{k-1} \| + C_5 C_3 \| \boldsymbol{x}_{k-1} \| = C_6 \| \boldsymbol{x}_{k-1} \|. \qquad (23)$$

З (22) і (23) випливає, що

$$\| \boldsymbol{x}_k \| \le C_4 C_6 \| \boldsymbol{x}_{k-1} \| = C_7 \| \boldsymbol{x}_{k-1} \|.$$

Послідовно застосовуючи останню оцінку, одержимо, що

$$\| \boldsymbol{x}_k \| \le C_7 \| \boldsymbol{x}_{k-1} \| \le C_7^2 \| \boldsymbol{x}_{k-2} \| \le \ldots \le C_7^k \| \boldsymbol{x}_0 \| = C_7^k C_8. \quad (24)$$

Тепер можна оцінити коефіцієнт загального члена ряду (8). З урахуванням оцінок (23) і (24) знаходимо:

$$\| \boldsymbol{x}_k + \theta_k \boldsymbol{h}_{11} \| \le C_6 \| \boldsymbol{x}_k \| \le C_6 C_7^k C_8 = C_7^k C_9.$$

Ця оцінка показує, що ряд (8) збігається за нормою $\| * \|$, як геометрична прогресія для $\varepsilon < C_7^{-1}$.

Більш докладне дослідження збіжності ряду (8) з оцінкою величин C_i можна виконати для конкретних матриць $\boldsymbol{A}_0$ і $\boldsymbol{A}_1$. На цьому питанні ми не зупиняємося.

П р и к л а д. Нехай треба знайти розв'язок рівняння

$$\begin{pmatrix} 1+\varepsilon & -2 \\ -1 & 2+\varepsilon \end{pmatrix} \boldsymbol{x}(\varepsilon) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

у вигляді розкладу за степенями малого параметра ε. Будемо обчислювати координати вектора $\boldsymbol{x}(\varepsilon)$ в першому наближенні, тобто з точністю до членів порядку ε^1.

Розглядуване рівняння є рівнянням (1) при

$$\boldsymbol{A}_0 = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix}, \quad \boldsymbol{A}_1 = \boldsymbol{E}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Матриця $\boldsymbol{A}_0$ має просте власне значення $\lambda = 0$ і відповідний до нього власний вектор $\boldsymbol{h}_{11} = (2 \quad 1)^T$. Матриця $\boldsymbol{A}_0^T$ теж має просте власне значення $\lambda = 0$, якому відповідає власний вектор $\boldsymbol{y}_0 = (1 \quad 1)^T$. Для матриці $\boldsymbol{B}$, яка збігається з $\boldsymbol{A}_0$, власне значення $\lambda = 0$ є простим, тому для $\boldsymbol{x}(\varepsilon)$ існує розклад (8). Обчислимо коефіцієнти цього розкладу.

Величини a і θ дорівнюють

$$a = \boldsymbol{h}_{11}^T \boldsymbol{A}_1^T \boldsymbol{y}_0 = 3, \quad \theta = a^{-1} \boldsymbol{b}^T \boldsymbol{y}_0 = 1.$$

Вектор $\boldsymbol{x}_0 = (x_{01} \quad x_{02})^T$ знаходимо як окремий розв'язок рівняння (12), ортогональний до $\boldsymbol{h}_{11}$. Рівняння (12) має вигляд:

$$\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_{01} \\ x_{02} \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Розв'язуючи його методом виключення, знаходимо загальний розв'язок $(2x_{02} - 1 \quad x_{02})^T$, де x_{02} відіграє роль параметра, який може набути будь-яких числових значень. Обчислимо його, врахувавши, що $\boldsymbol{x}_0 \perp \boldsymbol{h}_{11}$. Тоді одержимо окремий розв'язок

$$\boldsymbol{x}_0 = \begin{pmatrix} -\frac{1}{5} & \frac{2}{5} \end{pmatrix}^T.$$

Далі обчислимо θ_0:

$$\theta_0 = -a^{-1} \boldsymbol{x}_0^T \boldsymbol{A}_1^T \boldsymbol{y}_0 = -\frac{1}{15},$$

а потім $\boldsymbol{x}_1 = (x_{11} \quad x_{12})^T$ як окремий розв'язок рівняння (18), ортогональний до $\boldsymbol{h}_{11}$. Це рівняння є таким:

$$\begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} 1/3 \\ -1/3 \end{pmatrix},$$

а його загальний розв'язок дорівнює $(2x_{12} + 1/3 \quad x_{12})^T$. Окремим розв'язком, ортогональним до $\boldsymbol{h}_{11}$, є вектор

$$\boldsymbol{x}_1 = \begin{pmatrix} \frac{1}{15} & -\frac{2}{15} \end{pmatrix}^T.$$

Тепер можна знайти θ_1:

$$\theta_1 = -a^{-1}\boldsymbol{x}_1^T \boldsymbol{A}_1^T \boldsymbol{y}_0 = \frac{1}{45}.$$

Розклад для $\boldsymbol{x}(\varepsilon)$ з указаною вище точністю має вигляд:

$$\boldsymbol{x}(\varepsilon) = \frac{\theta \boldsymbol{h}_{11}}{\varepsilon} + \boldsymbol{x}_0 + \theta_0 \boldsymbol{h}_{11} + \varepsilon(\boldsymbol{x}_1 + \theta_1 \boldsymbol{h}_{11}) + \ldots .$$

Підставивши сюди знайдені вище числа θ, θ_0, θ_1 і вектори $\boldsymbol{h}_{11}$, $\boldsymbol{x}_0$, $\boldsymbol{x}_1$, одержимо:

$$\boldsymbol{x}(\varepsilon) = \frac{1}{\varepsilon}\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -1/3 \\ 1/3 \end{pmatrix} + \varepsilon \begin{pmatrix} 1/9 \\ -1/9 \end{pmatrix} + \ldots . \; \triangle$$

12.3.3. Випадок несумісності незбуреної системи (продовження). Припустимо, що $\lambda = 0$ є кратним власним значенням матриці $\boldsymbol{B}$. Основні властивості розглядуваної задачі про збурення можна виявити, розглянувши два окремих випадки, коли кратність власного значення дорівнює 3 і йому відповідає: 1) один жорданів ланцюжок довжини 3; 2) два жорданових ланцюжки довжини 2 і 1.

У першому випадку матриця $\boldsymbol{B}$ має один власний вектор $\boldsymbol{h}_{11}$, який є одночасно власним вектором матриці $\boldsymbol{A}_0$ для того ж власного значення. При цьому рівняння (6) теж має один (з точністю до числового множника) розв'язок $\boldsymbol{y}_0$. Шукаємо розв'язок рівняння (1) у вигляді ряду

$$\boldsymbol{x}(\varepsilon) = \frac{\mu_{01}\boldsymbol{h}_{11}}{\varepsilon^3} + \frac{\mu_{11}\boldsymbol{h}_{11} + \mu_{12}\boldsymbol{h}_{12}}{\varepsilon^2} + \frac{\mu_{21}\boldsymbol{h}_{11} + \mu_{22}\boldsymbol{h}_{12} + \mu_{23}\boldsymbol{h}_{13}}{\varepsilon} + $$
$$+ \sum_{k=0}^{\infty} \varepsilon^k (\boldsymbol{x}_k + \nu_{1k}\boldsymbol{h}_{11} + \nu_{2k}\boldsymbol{h}_{12} + \nu_{3k}\boldsymbol{h}_{13}). \quad (25)$$

Тут $\boldsymbol{h}_{11}$, $\boldsymbol{h}_{12}$, $\boldsymbol{h}_{13}$ — вектори жорданового ланцюжка, відповідного до нульового власного значення матриці $\boldsymbol{B}$. На підставі формул (9.1.1) можна написати співвідношення для цих векторів:

$$\boldsymbol{B}\boldsymbol{h}_{11} = \boldsymbol{0}, \quad \boldsymbol{B}\boldsymbol{h}_{12} = \boldsymbol{h}_{11}, \quad \boldsymbol{B}\boldsymbol{h}_{13} = \boldsymbol{h}_{12}. \quad (26)$$

Підставимо ряд (25) у рівняння (1) і порівняємо коефіцієнти при $\varepsilon^{-3}, \varepsilon^{-2}, \varepsilon^{-1}$. Тоді одержимо:

$$\mu_{01}\boldsymbol{A}_0\boldsymbol{h}_{11} = \boldsymbol{0},$$

$$\boldsymbol{A}_0(\mu_{11}\boldsymbol{h}_{11} + \mu_{12}\boldsymbol{h}_{12}) + \mu_{01}\boldsymbol{A}_1\boldsymbol{h}_{11} = \boldsymbol{0},$$

$$\boldsymbol{A}_0(\mu_{21}\boldsymbol{h}_{11} + \mu_{22}\boldsymbol{h}_{12} + \mu_{23}\boldsymbol{h}_{13}) + \boldsymbol{A}_1(\mu_{11}\boldsymbol{h}_{11} + \mu_{12}\boldsymbol{h}_{12}) = \boldsymbol{0}.$$

Якщо помножити зліва обидві частини цих рівностей на $\boldsymbol{A}_1^{-1}$, то будемо мати:

$$\mu_{01}\boldsymbol{B}\boldsymbol{h}_{11} = \boldsymbol{0},$$

$$\boldsymbol{B}(\mu_{11}\boldsymbol{h}_{11} + \mu_{12}\boldsymbol{h}_{12}) + \mu_{01}\boldsymbol{h}_{11} = \boldsymbol{0},$$

$$\boldsymbol{B}(\mu_{21}\boldsymbol{h}_{11} + \mu_{22}\boldsymbol{h}_{12} + \mu_{23}\boldsymbol{h}_{13}) + \mu_{11}\boldsymbol{h}_{11} + \mu_{12}\boldsymbol{h}_{12} = \boldsymbol{0}.$$

Формули (26) показують, що останні рівності виконуються при будь-якому μ_{01} і за умов

$$\mu_{12} + \mu_{01} = 0, \quad \mu_{22} + \mu_{11} = 0, \quad \mu_{23} + \mu_{12} = 0. \tag{27}$$

Порівнюючи коефіцієнти при ε^0 після підставлення ряду (25) в рівняння (1), одержимо:

$$\boldsymbol{A}_0(\boldsymbol{x}_0 + \nu_{10}\boldsymbol{h}_{11} + \nu_{20}\boldsymbol{h}_{12} + \nu_{30}\boldsymbol{h}_{13}) + \boldsymbol{A}_1(\mu_{21}\boldsymbol{h}_{11} + \mu_{22}\boldsymbol{h}_{12} + \mu_{23}\boldsymbol{h}_{13}) = \boldsymbol{b}.$$

Ця рівність за умов

$$\nu_{20} + \mu_{21} = 0, \quad \nu_{30} + \mu_{22} = 0 \tag{28}$$

і при врахуванні співвідношень (26) набуває вигляду:

$$\boldsymbol{A}_0\boldsymbol{x}_0 = \boldsymbol{b} - \mu_{23}\boldsymbol{A}_1\boldsymbol{h}_{13}. \tag{29}$$

На підставі теореми Фредгольма рівняння (29) є розв'язуваним при

$$(\boldsymbol{b} - \mu_{23}\boldsymbol{A}_1\boldsymbol{h}_{13})^T \boldsymbol{y}_0 = 0. \tag{30}$$

Скористаємося цим рівнянням для знаходження величини μ_{23}. Для цього покажемо, що величина $c = \boldsymbol{h}_{13}^T\boldsymbol{A}_1^T\boldsymbol{y}_0$ не є рівною нулю.

Припустимо, що $\boldsymbol{A}_1\boldsymbol{h}_{13} \perp \boldsymbol{y}_0$. Тоді на підставі теореми Фредгольма повинне бути розв'язуваним рівняння

$$\boldsymbol{A}_0\boldsymbol{z} = \boldsymbol{A}_1\boldsymbol{h}_{13},$$

або

$$\boldsymbol{B}\boldsymbol{z} = \boldsymbol{h}_{13}. \tag{31}$$

Ця рівність разом з рівностями (26) показує, що вектор $\boldsymbol{z}$ належить інваріантному підпростору, натягнутому на вектори жорданового ланцюжка $\boldsymbol{h}_{11}, \boldsymbol{h}_{12}, \boldsymbol{h}_{13}$. Тому

$$\boldsymbol{z} = \alpha\boldsymbol{h}_{11} + \beta\boldsymbol{h}_{12} + \gamma\boldsymbol{h}_{13},$$

де α, β, γ – деякі числа. Звідси з урахуванням рівностей (26) маємо:

$$\boldsymbol{B}\boldsymbol{z} = \beta\boldsymbol{h}_{11} + \gamma\boldsymbol{h}_{12}. \tag{32}$$

З формул (31) і (32) випливає рівність $\boldsymbol{h}_{13} = \beta\boldsymbol{h}_{11} + \gamma\boldsymbol{h}_{12}$, яка є неможливою через лінійну незалежність векторів жорданового ланцюжка. Це означає, що не може бути $\boldsymbol{A}_1\boldsymbol{h}_{13} \perp \boldsymbol{y}_0$. Тому $c \neq 0$.

Тепер з рівності (30) випливає, що

$$\mu_{23} = c^{-1}\boldsymbol{b}^T\boldsymbol{y}_0. \tag{33}$$

Після підставлення одержаного значення μ_{23} в рівняння (29) знайдемо $\boldsymbol{x}_0$ як окремий розв'язок, вважаючи $\boldsymbol{x}_0 \perp \boldsymbol{h}_{11}$.

Продовжимо порівнювати коефіцієнти при ε^k після підставлення ряду (25) в рівняння (1). При $k \geq 1$ одержимо:

$$\boldsymbol{A}_0(\boldsymbol{x}_k + \nu_{1k}\boldsymbol{h}_{11} + \nu_{2k}\boldsymbol{h}_{12} + \nu_{3k}\boldsymbol{h}_{13}) +$$

$$+\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \nu_{1,k-1}\boldsymbol{h}_{11} + \nu_{2,k-1}\boldsymbol{h}_{12} + \nu_{3,k-1}\boldsymbol{h}_{13}) = \boldsymbol{0}.$$

Звідси, вважаючи справедливими рівності

$$\nu_{2k} + \nu_{1,k-1} = 0, \quad \nu_{3k} + \nu_{2,k-1} = 0, \quad k \geq 1, \tag{34}$$

з урахуванням (26) знаходимо:

$$\boldsymbol{A}_0\boldsymbol{x}_k = -\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \nu_{3,k-1}\boldsymbol{h}_{13}). \tag{35}$$

Обчислимо $\nu_{3,\,k-1}$, скориставшись умовою розв'язуваності рівняння (35), яка має вигляд:

$$(\boldsymbol{x}_{k-1} + \nu_{3,\,k-1}\boldsymbol{h}_{13})^T \boldsymbol{A}_1^T \boldsymbol{y}_0 = 0.$$

Звідси випливає, що

$$\nu_{3,\,k-1} = -c^{-1}\boldsymbol{x}_{k-1}^T \boldsymbol{A}_1^T \boldsymbol{y}_0, \quad k \geq 1. \tag{36}$$

Далі після підставлення величини $\nu_{3,\,k-1}$ і раніше знайденого вектора $\boldsymbol{x}_{k-1}$ в рівняння (35) знаходимо $\boldsymbol{x}_k$ як окремий розв'язок цього рівняння, ортогональний до $\boldsymbol{h}_{11}$.

Описана процедура дозволяє послідовно знаходити всі вектори $\boldsymbol{x}_k$ і коефіцієнти при векторах жорданового ланцюжка в ряді (25). Напишемо цей ряд, спростивши його за допомогою одержаних вище співвідношень (27), (28) і (34), які можна подати у вигляді:

$$\mu_{01} = \mu_{23}, \ \ \mu_{11} = \nu_{30}, \ \ \mu_{12} = -\mu_{23}, \ \ \mu_{21} = \nu_{31}, \ \ \mu_{22} = -\nu_{30},$$
$$\nu_{1k} = \nu_{3,\,k+2}, \ \ \nu_{2k} = -\nu_{3,\,k+1}, \quad k \geq 0.$$

Маємо вираз для розв'язку:

$$\boldsymbol{x}(\varepsilon) = \frac{\mu_{23}\boldsymbol{h}_{11}}{\varepsilon^3} + \frac{\nu_{30}\boldsymbol{h}_{11} - \mu_{23}\boldsymbol{h}_{12}}{\varepsilon^2} + \frac{\nu_{31}\boldsymbol{h}_{11} - \nu_{30}\boldsymbol{h}_{12} + \mu_{23}\boldsymbol{h}_{13}}{\varepsilon} +$$
$$+\sum_{k=0}^{\infty} \varepsilon^k (\boldsymbol{x}_k + \nu_{3,\,k+2}\boldsymbol{h}_{11} - \nu_{3,\,k+1}\boldsymbol{h}_{12} + \nu_{3k}\boldsymbol{h}_{13}). \tag{37}$$

Нагадаємо, що величини μ_{23} і ν_{3k} обчислюються за формулами (33) і (36).

Розглянемо випадок, коли власному значенню $\lambda = 0$ матриці $\boldsymbol{B}$ відповідає два жорданових ланцюжки довжини 2 і 1. Шукаємо розв'язок рівняння (1) у вигляді суми двох рядів, відповідних до жорданових ланцюжків:

$$\boldsymbol{x}(\varepsilon) = \frac{\sigma_{01}\boldsymbol{h}_{11}}{\varepsilon^2} + \frac{\sigma_{11}\boldsymbol{h}_{11} + \sigma_{12}\boldsymbol{h}_{12}}{\varepsilon} + \frac{\tau\boldsymbol{h}_{21}}{\varepsilon} +$$

$$+\sum_{k=0}^{\infty} \varepsilon^k (\boldsymbol{x}_k + \chi_{1k}\boldsymbol{h}_{11} + \chi_{2k}\boldsymbol{h}_{12} + \tau_k \boldsymbol{h}_{21}). \qquad (38)$$

Тут $\boldsymbol{h}_{11}, \boldsymbol{h}_{12}$ і $\boldsymbol{h}_{21}$ — вектори жорданових ланцюжків, які на підставі формул (9.1.1) задовольняють рівностям

$$\boldsymbol{B}\boldsymbol{h}_{11} = \boldsymbol{0}, \quad \boldsymbol{B}\boldsymbol{h}_{12} = \boldsymbol{h}_{11}, \quad \boldsymbol{B}\boldsymbol{h}_{21} = \boldsymbol{0}. \qquad (39)$$

Підставимо ряд (38) у рівняння (1) і порівняємо коефіцієнти при ε^{-2} і ε^{-1}. Тоді одержимо:

$$\sigma_{01}\boldsymbol{A}_0\boldsymbol{h}_{11} = \boldsymbol{0},$$

$$\boldsymbol{A}_0(\sigma_{11}\boldsymbol{h}_{11} + \sigma_{12}\boldsymbol{h}_{12} + \tau\boldsymbol{h}_{21}) + \sigma_{01}\boldsymbol{A}_1\boldsymbol{h}_{11} = \boldsymbol{0}.$$

Ці рівності виконуються завдяки співвідношенням (39) і за умови

$$\sigma_{12} + \sigma_{01} = 0. \qquad (40)$$

Порівняємо коефіцієнти при ε^0:

$$\boldsymbol{A}_0(\boldsymbol{x}_0 + \chi_{10}\boldsymbol{h}_{11} + \chi_{20}\boldsymbol{h}_{12} + \tau_0\boldsymbol{h}_{21}) + \boldsymbol{A}_1(\sigma_{11}\boldsymbol{h}_{11} + \sigma_{12}\boldsymbol{h}_{12} + \tau\boldsymbol{h}_{21}) = \boldsymbol{b}.$$

Звідси, вважаючи, що

$$\chi_{20} + \sigma_{11} = 0, \qquad (41)$$

одержимо з урахуванням формул (39):

$$\boldsymbol{A}_0\boldsymbol{x}_0 = \boldsymbol{b} - \boldsymbol{A}_1(\sigma_{12}\boldsymbol{h}_{12} + \tau\boldsymbol{h}_{21}). \qquad (42)$$

Кількість власних векторів матриць $\boldsymbol{A}_0$ і $\boldsymbol{B}$, відповідних до $\lambda = 0$, визначається величиною $r(\boldsymbol{A}_0) = r(\boldsymbol{B})$. Такою ж є кількість фундаментальних розв'язків системи (6). Якщо ці розв'язки позначені через $\boldsymbol{y}_1$ і $\boldsymbol{y}_2$, то на підставі теореми Фредгольма рівняння (42) є розв'язуваним при

$$(\boldsymbol{b} - \boldsymbol{A}_1(\sigma_{12}\boldsymbol{h}_{12} + \tau\boldsymbol{h}_{21}))^T \boldsymbol{y}_i = 0, \quad i = \overline{1,\ 2}. \qquad (43)$$

З формули (43) знаходимо систему лінійних рівнянь відносно величин σ_{12} і τ

$$\begin{pmatrix} \boldsymbol{h}_{12}^T \boldsymbol{A}_1^T \boldsymbol{y}_1 & \boldsymbol{h}_{21}^T \boldsymbol{A}_1^T \boldsymbol{y}_1 \\ \boldsymbol{h}_{12}^T \boldsymbol{A}_1^T \boldsymbol{y}_2 & \boldsymbol{h}_{21}^T \boldsymbol{A}_1^T \boldsymbol{y}_2 \end{pmatrix} \begin{pmatrix} \sigma_{12} \\ \tau \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}^T \boldsymbol{y}_1 \\ \boldsymbol{b}^T \boldsymbol{y}_2 \end{pmatrix}. \tag{44}$$

Покажемо, що визначник цієї системи рівнянь не дорівнює нулю. Для цього розглянемо рівну нульовому вектору лінійну комбінацію стовпців визначника:

$$\alpha \begin{pmatrix} \boldsymbol{h}_{12}^T \boldsymbol{A}_1^T \boldsymbol{y}_1 \\ \boldsymbol{h}_{12}^T \boldsymbol{A}_1^T \boldsymbol{y}_2 \end{pmatrix} + \beta \begin{pmatrix} \boldsymbol{h}_{21}^T \boldsymbol{A}_1^T \boldsymbol{y}_1 \\ \boldsymbol{h}_{21}^T \boldsymbol{A}_1^T \boldsymbol{y}_2 \end{pmatrix} = \boldsymbol{0}.$$

Звідси знаходимо, що

$$(\alpha \boldsymbol{h}_{12} + \beta \boldsymbol{h}_{21})^T \boldsymbol{A}_1^T \boldsymbol{y}_i = 0, \quad i = \overline{1,\ 2}.$$

Це означає, що рівняння

$$\boldsymbol{A}_0 \boldsymbol{z} = \boldsymbol{A}_1 (\alpha \boldsymbol{h}_{12} + \beta \boldsymbol{h}_{21}) \tag{45}$$

є розв’язуваним на підставі теореми Фредгольма.

З рівності (45) знаходимо:

$$\boldsymbol{B}\boldsymbol{z} = \alpha \boldsymbol{h}_{12} + \beta \boldsymbol{h}_{21}. \tag{46}$$

Рівності (46) і (39) вказують на те, що вектор $\boldsymbol{z}$ належить інваріантному підпростору, натягнутому на вектори $\boldsymbol{h}_{11}, \boldsymbol{h}_{12}, \boldsymbol{h}_{21}$. Тому

$$\boldsymbol{z} = \gamma \boldsymbol{h}_{11} + \delta \boldsymbol{h}_{12} + \zeta \boldsymbol{h}_{21},$$

де γ, δ, ζ — деякі числа. Звідси з урахуванням рівностей (39) випливає, що

$$\boldsymbol{B}\boldsymbol{z} = \delta \boldsymbol{h}_{11}. \tag{47}$$

Порівнюючи формули (46) і (47), одержимо:

$$\alpha \boldsymbol{h}_{12} + \beta \boldsymbol{h}_{21} = \delta \boldsymbol{h}_{11}.$$

Через лінійну незалежність векторів $\boldsymbol{h}_{11}, \boldsymbol{h}_{12}, \boldsymbol{h}_{21}$ остання рівність можлива лише при $\alpha = \beta = \delta = 0$. Це означає, у свою чергу, що лінійно незалежними є стовпці визначника системи рівнянь (44). На підставі теореми 3.4.2 визначник є відмінним від нуля, тому система (44) має розв'язок. Після обчислення σ_{12} і τ можна знайти окремий розв'язок рівняння (42), ортогональний до власних векторів $\boldsymbol{h}_{11}$ і $\boldsymbol{h}_{21}$.

Продовжимо порівнювати коефіцієнти при ε^k після підставлення ряду (38) в рівняння (1). При $k \geq 1$ маємо:

$$\boldsymbol{A}_0(\boldsymbol{x}_k + \chi_{1k}\boldsymbol{h}_{11} + \chi_{2k}\boldsymbol{h}_{12} + \tau_k\boldsymbol{h}_{21}) +$$

$$+\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \chi_{1,\,k-1}\boldsymbol{h}_{11} + \chi_{2,\,k-1}\boldsymbol{h}_{12} + \tau_{k-1}\boldsymbol{h}_{21}) = \boldsymbol{0}.$$

Звідси, вважаючи, що

$$\chi_{2k} + \chi_{1,\,k-1} = 0, \quad k \geq 1, \tag{48}$$

з урахуванням (39) будемо мати:

$$\boldsymbol{A}_0\boldsymbol{x}_k = -\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \chi_{2,\,k-1}\boldsymbol{h}_{12} + \tau_{k-1}\boldsymbol{h}_{21}). \tag{49}$$

Це рівняння є розв'язуваним при

$$(\boldsymbol{A}_1(\boldsymbol{x}_{k-1} + \chi_{2,\,k-1}\boldsymbol{h}_{12} + \tau_{k-1}\boldsymbol{h}_{21}))^T \boldsymbol{y}_i = 0, \quad i = \overline{1,\,2}. \tag{50}$$

Остання формула визначає систему рівнянь відносно $\chi_{2,\,k-1}$ і τ_{k-1}, подібну до (44). Обчислимо ці величини і підставимо їх і вектор $\boldsymbol{x}_{k-1}$ у рівняння (49). Після цього вектор $\boldsymbol{x}_k$ знаходиться як окремий розв'язок рівняння (49), ортогональний до $\boldsymbol{h}_{11}$ і $\boldsymbol{h}_{21}$.

Напишемо вираз для ряду (38) з урахуванням рівностей (40), (41) і (48), які візьмемо у вигляді:

$$\sigma_{01} = -\sigma_{12}, \quad \sigma_{11} = -\chi_{20}, \quad \chi_{1k} = -\chi_{2,\,k+1}, \quad k \geq 0.$$

Маємо:

$$\boldsymbol{x}(\varepsilon) = -\frac{\sigma_{12}\boldsymbol{h}_{11}}{\varepsilon^2} + \frac{-\chi_{20}\boldsymbol{h}_{11} + \sigma_{12}\boldsymbol{h}_{12}}{\varepsilon} + \frac{\tau\boldsymbol{h}_{21}}{\varepsilon} +$$

$$+\sum_{k=0}^{\infty}\varepsilon^k(\boldsymbol{x}_k - \chi_{2,\,k+1}\boldsymbol{h}_{11} + \chi_{2k}\boldsymbol{h}_{12} + \tau_k\boldsymbol{h}_{21}). \tag{51}$$

Нагадаємо, що пари чисел σ_{12} і τ, а також χ_{2k} і τ_k при $k \geq 0$ визначаються з систем рівнянь (44) і (50).

Наприкінці зазначимо, що збіжнісить рядів (37) і (51) можна дослідити подібно до ряду (8).

12.4. Ітераційні методи для лінійних систем

Ітераційні методи дають розв'язок системи лінійних рівнянь як границю нескінченного обчислювального процесу, який дозволяє за вже знайденим наближенням до розв'язку побудувати наступне, більш точне. Важлива перевага збіжних ітераційних методів полягає в тому, що помилка в деякому наближенні виправляється в подальших обчисленнях, і таке виправлення звичайно вимагає лише кількох зайвих кроків одноманітних обчислень. Ця обставина є особливо цінною при обчисленнях за допомогою комп'ютерів.

З багатьох існуючих ітераційних методів для лінійних систем ми розглянемо два найбільш простих: метод простої ітерації і метод Гаусса — Зейделя.

12.4.1. Метод простої ітерації. Найпростішим ітераційним методом розв'язання систем лінійних рівнянь є *метод простої ітерації*, який застосовується до системи вигляду:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}. \tag{1}$$

Припустимо, що відомий наближений розв'язок $\boldsymbol{x}^{(0)}$ системи. Усі наступні наближення визначимо за правилом:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{A}\boldsymbol{x}^{(k)} + \boldsymbol{b}, \qquad k = 0, 1, 2, \ldots . \tag{2}$$

Якщо послідовність $\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots$ збігається до деякого вектора, то він буде розв'язком системи. Це твердження стає очевидним, якщо в (2) зробити граничний перехід при $k \to \infty$. Умову збіжності ітераційної послідовності $\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots$ дає

Теорема 12.4.1. *Для збіжності послідовних наближень достатньо, щоб усі власні значення матриці* ***A*** *були за модулем меншими ніж одиниця*:

$$|\lambda_i| < 1, \qquad i = \overline{1, n}. \tag{3}$$

Д о в е д е н н я. Послідовні наближення $\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots$ пов'язані між собою різницевим рівнянням (2), тому на підставі результатів, одержаних у розділі 10.1, можна стверджувати, що умова (3) є достатньою для збіжності цієї послідовності.

Висновок. *Для збіжності послідовних наближень достатньо, щоб яка-небудь норма матриці* $\boldsymbol{A}$ *була меншою ніж одиниця.*

Д о в е д е н н я базується на теоремі 12.1.6. □

З а у в а ж е н н я 1. Можна довести, що умова (3) є необхідною, якщо вимагати, щоб послідовні наближення збігалися до $(\boldsymbol{E} - \boldsymbol{A})^{-1}\boldsymbol{b}$ при будь-якому початковому наближенні $\boldsymbol{x}^{(0)}$.

З а у в а ж е н н я 2. Ітераційний метод можна застосувати до будь-якої системи $\boldsymbol{Ax} = \boldsymbol{b}$ після «розщеплення» матриці $\boldsymbol{A}$. Якщо вважати, що $\boldsymbol{A} = \boldsymbol{S} - \boldsymbol{T}$, то система $\boldsymbol{Ax} = \boldsymbol{b}$ рівносильна системі $\boldsymbol{Sx} = \boldsymbol{Tx} + \boldsymbol{b}$. Звичайно, $\boldsymbol{S}$ повинна бути простою (наприклад, діагональною або трикутною) і невиродженою матрицею. Тоді $\boldsymbol{x} = \boldsymbol{S}^{-1}\boldsymbol{Tx} + \boldsymbol{S}^{-1}\boldsymbol{b}$, тобто ми маємо систему вигляду (1).

12.4.2. Метод Гаусса — Зейделя. У методі простої ітерації наступне наближення $\boldsymbol{x}^{(k+1)}$ обчислюється на підставі попереднього шляхом підставлення в праву частину рівняння (1) всіх координат вектора $\boldsymbol{x}^{(k)}$. Це означає, що поки не обчислене повністю наближення $\boldsymbol{x}^{(k+1)}$, треба зберігати в пам'яті комп'ютера всі координати вектора $\boldsymbol{x}^{(k)}$. Ця обставина є джерелом практичних труднощів при розв'язанні систем великого порядку. Тому в *методі Гаусса —Зейделя* при обчисленні $(k+1)$-го наближення невідомої x_i враховуються вже обчислені раніше $(k+1)$-і наближення невідомих x_1, …, x_{i-1}. Інакше кажучи, координати $x_1^{(k+1)}$, …, $x_n^{(k+1)}$ вектора $\boldsymbol{x}^{(k+1)}$ обчислюються за формулами ($k = 0, 1, 2, \ldots$):

$$x_1^{(k+1)} = \sum_{j=1}^{n} a_{1j}x_j^{(k)} + b_1,$$

$$x_2^{(k+1)} = a_{21}x_1^{(k+1)} + \sum_{j=2}^{n} a_{2j}x_j^{(k)} + b_2,$$

$$x_3^{(k+1)} = \sum_{j=1}^{2} a_{3j}x_j^{(k+1)} + \sum_{j=3}^{n} a_{3j}x_j^{(k)} + b_3,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x_n^{(k+1)} = \sum_{j=1}^{n-1} a_{nj}x_j^{(k+1)} + a_{nn}x_n^{(k)} + b_n.$$

З'ясуємо умову збіжності методу Гаусса — Зейделя. Подамо матрицю $\boldsymbol{A}$ у вигляді суми $\boldsymbol{A} = \boldsymbol{F} + \boldsymbol{G}$, де матриця $\boldsymbol{F}$ отримується з $\boldsymbol{A}$ шляхом заміни нулями діагональних елементів і всіх елементів, розміщених вище головної діагоналі. Тепер останню систему запишемо у вигляді:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{F}\boldsymbol{x}^{(k+1)} + \boldsymbol{G}\boldsymbol{x}^{(k)} + \boldsymbol{b}. \tag{4}$$

Теорема 12.4.2. *Для збіжності послідовних наближень, одержуваних методом Гаусса — Зейделя, достатньо, щоб усі корені рівняння*

$$\det(\lambda(\boldsymbol{E} - \boldsymbol{F}) - \boldsymbol{G}) = 0 \tag{5}$$

були за модулем меншими ніж одиниця.

Д о в е д е н н я. Матриця $\boldsymbol{E} - \boldsymbol{F}$ є нижньою трикутною і має одиниці на головній діагоналі. Оскільки така матриця є невиродженою, то з рівності (4) знаходимо, що

$$\boldsymbol{x}^{(k+1)} = (\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{G}\boldsymbol{x}^{(k)} + (\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{b}.$$

Отже, метод Гаусса — Зейделя рівносильний методу простої ітерації, застосованому до системи

$$\boldsymbol{x} = (\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{G}\boldsymbol{x} + (\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{b}.$$

Це дає можливість скористатися теоремою 12.4.1, на підставі якої можна стверджувати, що для збіжності методу Гаусса — Зейделя достатньо, щоб власні значення матриці $(\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{G}$, тобто корені рівняння

$$\det(\lambda\boldsymbol{E} - (\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{G}) = 0,$$

були за модулем меншими ніж одиниця.

Перетворимо це рівняння, помноживши обидві його частини на $\det(\boldsymbol{E} - \boldsymbol{F}) \neq 0$:

$$0 = \det(\boldsymbol{E} - \boldsymbol{F}) \cdot \det(\lambda\boldsymbol{E} - (\boldsymbol{E} - \boldsymbol{F})^{-1}\boldsymbol{G}) = \det(\lambda(\boldsymbol{E} - \boldsymbol{F}) - \boldsymbol{G}).$$

Отже, ми одержали рівняння (5). □

Наведемо більш просту достатню ознаку збіжності.

Теорема 12.4.3. *Для збіжності послідовних наближень, одержуваних методом Гаусса — Зейделя, достатньо, щоб виконувалась одна з умов*:

$$\| \boldsymbol{A} \|_1 < 1, \tag{6}$$

$$\| \boldsymbol{A} \|_\infty < 1. \tag{7}$$

Д о в е д е н н я. Покажемо, що за умови (6) всі корені рівняння (5) будуть за модулем меншими ніж одиниця. Припустимо, що це не так, тобто $|\lambda| \geq 1$. Оцінимо суму модулів позадіагональних елементів *j*-го стовпця матриці $\lambda(\boldsymbol{E} - \boldsymbol{F}) - \boldsymbol{G}$. Урахувавши формулу (12.1.9), будемо мати:

$$|\lambda| \sum_{i=j+1}^{n} |a_{ij}| + \sum_{i=1}^{j-1} |a_{ij}| \leq |\lambda| \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| = |\lambda| \left(\sum_{i=1}^{n} |a_{ij}| - |a_{jj}| \right) <$$

$$< |\lambda| (1 - |a_{jj}|) = |\lambda| - |\lambda| |a_{jj}| \leq |\lambda| - |a_{jj}| \leq |\lambda - a_{jj}|.$$

Звідси випливає, що матриця $\lambda(\boldsymbol{E} - \boldsymbol{F}) - \boldsymbol{G}$ є матрицею Адамара при $|\lambda| \geq 1$. На підставі теореми Адамара ця матриця є невиродженою, тобто λ не може бути коренем рівняння (5) при $|\lambda| \geq 1$. Тому всі корені цього рівняння за модулем менші ніж одиниця.

У випадку виконання нерівності (7) теорема доводиться аналогічно, якщо скористатися формулою (12.1.10). □

12.5. *QR*-алгоритм

12.5.1. *QR*-алгоритм знаходження власних значень. Загальна схема розв'язання алгебраїчної проблеми власних значень, описана в розділі 6.1, є теоретичною основою при розгляді вказаної проблеми. Вона придатна для практичного застосування лише в окремих простих випадках, наприклад, для матриць другого порядку. Відшукання коренів характеристичного многочлена в загальному випадку пов'язане з подоланням настільки значних обчислювальних труднощів, що більш прийнятним виявився альтернативний шлях, про який ішлося після доведення теореми 6.2.5. Він пов'язаний зі зведенням матриці за допомогою перетворень подібності до трикутного вигляду. Зокрема, на цьому побудований розглядуваний нижче метод, відомий як $\boldsymbol{QR}$ - алгоритм.

$\boldsymbol{QR}$-*алгоритм* застосовується до матриці, яка попередньо зведена до форми Хессенберга, як це описано в розділі 4.4. Кожний крок $\boldsymbol{QR}$-алгоритму полягає в переході від матриці $\boldsymbol{A}_k$ до матриці $\boldsymbol{A}_{k+1}$, де $k = 0, 1, 2, \ldots$ і $\boldsymbol{A}_0 = \boldsymbol{A}$, за такими правилами:

1) за допомогою перетворень відбиття одержується $\boldsymbol{QR}$-розклад матриці $\boldsymbol{A}_k$:

$$\boldsymbol{A}_k = \boldsymbol{Q}_k \boldsymbol{R}_k; \tag{1}$$

2) матриця $\boldsymbol{A}_{k+1}$ одержується шляхом переставлення у $\boldsymbol{QR}$-розкладі матриці $\boldsymbol{A}_k$ множників:

$$\boldsymbol{A}_{k+1} = \boldsymbol{R}_k \boldsymbol{Q}_k. \tag{2}$$

Після виконання цих операцій маємо:

$$\boldsymbol{Q}_k^T \boldsymbol{A}_k \boldsymbol{Q}_k = \boldsymbol{Q}_k^T (\boldsymbol{Q}_k \boldsymbol{R}_k) \boldsymbol{Q}_k = \boldsymbol{A}_{k+1},$$

або

$$\boldsymbol{A}_{k+1} = (\boldsymbol{Q}_0 \boldsymbol{Q}_1 \ldots \boldsymbol{Q}_k)^T \boldsymbol{A} \boldsymbol{Q}_0 \boldsymbol{Q}_1 \ldots \boldsymbol{Q}_k. \tag{3}$$

На підставі теореми 6.2.5 матриці $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots$ мають ті ж власні значення, що і матриця $\boldsymbol{A}$.

Суттєва особливість $\boldsymbol{QR}$-алгоритму полягає в тому, що трудомісткість одного його кроку знижується на порядок, якщо матрицю попередньо звести до форми Хессенберга. Важливість цієї обставини стає зрозумілою при врахуванні того, що ітерації $\boldsymbol{QR}$-алгоритму, як доводиться в наступній теоремі, зберігають форму Хессенберга.

Теорема 12.5.1. *Майже трикутна форма матриці зберігається при застосуванні до неї* $\boldsymbol{QR}$ *- алгоритму.*

Д о в е д е н н я. Нехай матриця $\boldsymbol{A}_k$ має форму Хесенберга. Її можна перетворити у верхню трикутну матрицю $\boldsymbol{R}_k$ шляхом множення зліва на матриці обертання $\boldsymbol{T}_{12}(\varphi_1), \boldsymbol{T}_{23}(\varphi_2), \ldots, \boldsymbol{T}_{n-1,\,n}(\varphi_{n-1})$:

$$\boldsymbol{R}_k = \boldsymbol{T}_{n-1,\,n}(\varphi_{n-1}) \ldots \boldsymbol{T}_{23}(\varphi_2) \boldsymbol{T}_{12}(\varphi_1) \boldsymbol{A}_k.$$

Звідси випливає, що

$$\boldsymbol{A}_k = \boldsymbol{T}_{12}^T(\varphi_1)\boldsymbol{T}_{23}^T(\varphi_2)\ldots\boldsymbol{T}_{n-1,n}^T(\varphi_{n-1})\boldsymbol{R}_k.$$

Тепер

$$\boldsymbol{A}_{k+1} = \boldsymbol{R}_k\boldsymbol{T}_{12}^T(\varphi_1)\boldsymbol{T}_{23}^T(\varphi_2)\ldots\boldsymbol{T}_{n-1,n}^T(\varphi_{n-1}).$$

Множення справа матриці $\boldsymbol{R}_k$ на матриці обертання може змінити в $\boldsymbol{R}_k$ тільки діагональні блоки другого порядку і елементи, розміщені вище цих блоків. При цьому трикутна матриця $\boldsymbol{R}_k$ може перетворитися тільки в майже трикутну. □

Особливо простим є симетричний випадок, бо ітерації ***QR***-алгоритму зберігають не тільки форму Хессенберга, але і симетричність матриць $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots$, залишаючи їх тридіагональними. Це доводить

Теорема 12.5.2. *Симетричність матриці зберігається при застосуванні до неї **QR**-алгоритму.*

Д о в е д е н н я. Нехай $\boldsymbol{A}_k$ — симетрична матриця і $\boldsymbol{A}_k = \boldsymbol{Q}_k\boldsymbol{R}_k$. Тоді

$$\boldsymbol{A}_k = \boldsymbol{A}_k^T = \boldsymbol{R}_k^T\boldsymbol{Q}_k^T$$

і тому

$$\boldsymbol{R}_k^T = \boldsymbol{A}_k\boldsymbol{Q}_k = \boldsymbol{Q}_k\boldsymbol{R}_k\boldsymbol{Q}_k.$$

Тепер $\boldsymbol{A}_{k+1} = \boldsymbol{R}_k\boldsymbol{Q}_k$ і , отже,

$$\boldsymbol{A}_{k+1}^T = \boldsymbol{Q}_k^T\boldsymbol{R}_k^T = \boldsymbol{Q}_k^T\boldsymbol{Q}_k\boldsymbol{R}_k\boldsymbol{Q}_k = \boldsymbol{R}_k\boldsymbol{Q}_k = \boldsymbol{A}_{k+1}. \ \square$$

З а у в а ж е н н я. За допомогою перетворень Хаусхолдера можна одержати ***QL***-розклад матриці у вигляді добутку ортогональної матриці ***Q*** і нижньої трикутної матриці ***L***. Заснований на цьому розкладі ***QL***-алгоритм у теоретичному сенсі майже не відрізняється від ***QR***-алгоритму. Вибір одного з них залежить від особливостей конкретної задачі.

12.5.2. Збіжність *QR*-алгоритму. Для з'ясування механізму дії ***QR***-алгоритму дослідимо будову матриць $\boldsymbol{A}_k$ при великих *k*. Обмежимося випадком, коли $\boldsymbol{A}$ є невиродженою матрицею простої структури, яка має форму Хессенберга. Будемо вважати ненульовими всі піддіагональні елементи такої матриці (випадок, коли один з її піддіагональних елементів дорівнює нулю, розглядається нижче). Покажемо, що при цьому матриця простої структури буде мати різні власні значення.

Нехай λ_i — деяке власне значення матриці $\boldsymbol{A}$ порядку n, тоді мінор матриці $\lambda_i \boldsymbol{E} - \boldsymbol{A}$, розміщений у перших $n-1$ стовпцях і останніх $n-1$ рядках, є ненульовим. Це означає, що $r(\lambda_i \boldsymbol{E} - \boldsymbol{A}) = n-1$. Припустимо, що λ_i є кратним характеристичним коренем. Тоді буде $r(\lambda_i \boldsymbol{E} - \boldsymbol{A}) \le n-2$, що суперечить попередній рівності. Отже, власні значення є різними. Ми будемо припускати, що їх модулі теж різні.

Таким чином, існує матриця $\boldsymbol{S}$ така, що

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}, \quad \boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n), \tag{4}$$

причому вважається, що власні значення $\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n$ задовольняють умовам:

$$|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|. \tag{5}$$

Далі нам знадобиться представлення матриці $\boldsymbol{A}^{k+1}$ у вигляді:

$$\boldsymbol{A}^{k+1} = \boldsymbol{P}_{k+1}\boldsymbol{U}_{k+1}, \tag{6}$$

де

$$\boldsymbol{P}_{k+1} = \boldsymbol{Q}_0 \ldots \boldsymbol{Q}_{k-1}\boldsymbol{Q}_k, \quad \boldsymbol{U}_{k+1} = \boldsymbol{R}_k\boldsymbol{R}_{k-1}\ldots\boldsymbol{R}_0 \tag{7}$$

і матриці $\boldsymbol{Q}_0, \ldots, \boldsymbol{Q}_k$, $\boldsymbol{R}_0, \ldots, \boldsymbol{R}_k$ є множниками $\boldsymbol{QR}$-розкладів (1). Рівність (6) легко доводиться, якщо врахувати наступну формулу, яка є наслідком рівностей (7), (1) і (3):

$$\boldsymbol{P}_{k+1}\boldsymbol{U}_{k+1} = \boldsymbol{P}_k\boldsymbol{Q}_k\boldsymbol{R}_k\boldsymbol{U}_k = \boldsymbol{P}_k\boldsymbol{A}_k\boldsymbol{U}_k = \boldsymbol{P}_k\boldsymbol{P}_k^T\boldsymbol{A}\boldsymbol{P}_k\boldsymbol{U}_k = \boldsymbol{A}\boldsymbol{P}_k\boldsymbol{U}_k.$$

Припустимо, що існує $\boldsymbol{LR}$-розклад матриці $\boldsymbol{S}^{-1}$. Це означає, що $\boldsymbol{S}^{-1} = \boldsymbol{LR}$, де $\boldsymbol{L}$ — нижня трикутна матриця з одиницями на головній діагоналі і $\boldsymbol{R}$ — верхня трикутна матриця. З формул (4) і (6) випливає, що

$$\boldsymbol{S}\Lambda^{k+1}\boldsymbol{L}\boldsymbol{R} = \boldsymbol{P}_{k+1}\boldsymbol{U}_{k+1},$$

звідки

$$\boldsymbol{P}_{k+1}^T\boldsymbol{S}\Lambda^{k+1}\boldsymbol{L}\Lambda^{-k-1} = \boldsymbol{U}_{k+1}\boldsymbol{R}^{-1}\Lambda^{-k-1}.$$

Позначимо кожну з частин цієї рівності через $\boldsymbol{N}_k$. Очевидно, що $\boldsymbol{N}_k$ є верхньою трикутною матрицею. Тепер рівність (3) можна подати у необхідному для подальшого аналізу вигляді:

$$\boldsymbol{A}_{k+1} = \boldsymbol{N}_k\Lambda^{k+1}\boldsymbol{L}^{-1}\Lambda\boldsymbol{L}\Lambda^{-k-1}\boldsymbol{N}_k^{-1}. \tag{8}$$

Матриця $\boldsymbol{B} = \boldsymbol{L}^{-1}\Lambda\boldsymbol{L}$ є нижньою трикутною і її діагональні елементи дорівнюють $\lambda_1, \ldots, \lambda_n$. Якщо b_{ij} — елементи цієї матриці, то елементи матриці $\boldsymbol{C}_k = \Lambda^{k+1}\boldsymbol{B}\Lambda^{-k-1}$ дорівнюють

$$\{\boldsymbol{C}_k\}_{ij} = b_{ij}\left(\frac{\lambda_i}{\lambda_j}\right)^{k+1}. \tag{9}$$

Оскільки матриця $\boldsymbol{B}$ є нижньою трикутною, то $b_{ij} = 0$ при $i < j$. Якщо $i > j$, то елементи матриці $\boldsymbol{C}_k$ не є нульовими, проте вони прямують до нуля при $k \to \infty$. Це випливає з формул (9) і (5). Отже, при досить великих k матрицю $\boldsymbol{C}_k$ можна наближено вважати діагональною, а її діагональні елементи — рівними власним значенням матриці $\boldsymbol{A}$.

Формула (8) показує, що при обмеженості матриць $\boldsymbol{N}_k$ і $\boldsymbol{N}_k^{-1}$ матриця $\boldsymbol{A}_{k+1} = \boldsymbol{N}_k\boldsymbol{C}_k\boldsymbol{N}_k^{-1}$ прямує до верхньої трикутної матриці, діагональні елементи якої є власними значеннями матриці $\boldsymbol{A}$. Таким чином, залишається показати, що при великих значеннях k матриці $\boldsymbol{N}_k$ і $\boldsymbol{N}_k^{-1}$ є обмеженими. Покажемо, як це випливає з рівностей

$$\boldsymbol{N}_k = \boldsymbol{P}_{k+1}^T\boldsymbol{S}\Lambda^{k+1}\boldsymbol{L}\Lambda^{-k-1}, \qquad \boldsymbol{N}_k^{-1} = \Lambda^{k+1}\boldsymbol{L}^{-1}\Lambda^{-k-1}\boldsymbol{S}^{-1}\boldsymbol{P}_{k+1}.$$

Дійсно, матриці $\boldsymbol{S}$ і $\boldsymbol{S}^{-1}$ є постійними, а $\boldsymbol{P}_{k+1}$ і $\boldsymbol{P}_{k+1}^T$ обмежені як ортогональні матриці. Що стосується матриць $\Lambda^{k+1}\boldsymbol{L}\Lambda^{-k-1}$ і

$\Lambda^{k+1}L^{-1}\Lambda^{-k-1}$ то, як показує рівність (9), їх елементи не перевищують за модулем відповідні елементи матриць L і L^{-1}.

Таким чином, при виконанні умов (5) матриці A_{k+1} наближаються до верхньої трикутної матриці. Можна показати, що при послабленні обмежень (5) матриці A_{k+1} наближаються до форми, відповідної до розкладу Шура, тобто одержується квазітрикутна матриця з діагональними підматрицями першого і другого порядків.

12.5.3. Прискорення збіжності *QR*-алгоритму. Укажемо на два способи поліпшення збіжності ***QR***-алгоритму. Один з них пов'язаний зі зменшенням розмірів розв'язуваної задачі. Припустимо, що після застосування кількох кроків ***QR***-алгоритму до майже трикутної матриці один з піддіагональних елементів стає настільки малим, що його можна вважати нульовим. Тоді ми будемо мати справу з квазітрикутною матрицею з двома діагональними блоками, які, у свою чергу, є майже трикутними матрицями. Всі власні значення вихідної матриці будуть одержані, якщо ми знайдемо власні значення діагональних блоків, тобто задача «розщеплюється» на дві задачі меншого розміру.

Розглянемо другу можливість удосконалення ***QR***-алгоритму. Рівності (1) і (2) можна замінити співвідношеннями

$$A_k - \sigma_k E = Q_k R_k, \quad A_{k+1} = R_k Q_k + \sigma_k E, \tag{10}$$

де $\sigma_0, \sigma_1, \sigma_2, \ldots$ — деякі числа. Такий перехід не змінює власних значень, бо матриця A_{k+1} подібна до A_k:

$$Q_k^T A_k Q_k = Q_k^T (Q_k R_k + \sigma_k E) Q_k = A_{k+1}.$$

Алгоритм, заснований на рівностях (10), називається ***QR**-алгоритмом зі зсувом*. Зсув визначається величиною σ_k. Найбільше прискорення збіжності ***QR***-алгоритму досягається тоді, коли зсув є близьким до власного значення матриці A.

Існують модифікації ***QR***-алгоритму зі зсувом, пристосовані до різних окремих випадків , зокрема, до знаходження комплексно спряжених власних значень дійсної матриці.

П р и к л а д 1. Для матриці

$$\boldsymbol{A} = \begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix}$$

власні значення якої дорівнюють

$$\lambda_1 = \frac{5+\sqrt{13}}{2} \approx 4{,}30, \quad \lambda_2 = \frac{5-\sqrt{13}}{2} \approx 0{,}70,$$

виконаємо один крок ***QR***-алгоритму без зсуву і зі зсувом, рівним елементу a_{22}.

QR-алгоритм застосовують до матриці у формі Хессенберга. Саме такою є будь-яка матриця другого порядку, тому до неї можна застосувати ***QR***-алгоритм.

Спочатку треба одержати ***QR***-розклад матриці ***A***. Зробимо це за допомогою перетворення обертання. Помножимо зліва матрицю ***A*** на матрицю обертання

$$\boldsymbol{T}_{12}(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

і виберемо найменший додатний кут φ, виходячи з умов

$$\{\boldsymbol{T}_{12}(\varphi)\boldsymbol{A}\}_{21} = 0, \quad \{\boldsymbol{T}_{12}(\varphi)\boldsymbol{A}\}_{11} > 0.$$

Очевидно, що $\operatorname{tg}\varphi = 1/4$, $\sin\varphi = 1/\sqrt{17}$, $\cos\varphi = 4/\sqrt{17}$, тому

$$\boldsymbol{T}_{12}(\varphi) = \frac{1}{\sqrt{17}}\begin{pmatrix} 4 & -1 \\ 1 & 4 \end{pmatrix}.$$

Знаходимо верхню трикутну матрицю ***R***:

$$\boldsymbol{R} = \boldsymbol{T}_{12}(\varphi)\boldsymbol{A} = \frac{1}{\sqrt{17}}\begin{pmatrix} 4 & -1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 4 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{\sqrt{17}}\begin{pmatrix} 17 & -5 \\ 0 & 3 \end{pmatrix}.$$

Звідси одержується ***QR***-розклад матриці ***A***:

$$\boldsymbol{A}_0 = \boldsymbol{A} = \boldsymbol{T}_{12}^T(\varphi)\boldsymbol{R} = \frac{1}{\sqrt{17}}\begin{pmatrix} 4 & 1 \\ -1 & 4 \end{pmatrix}\frac{1}{\sqrt{17}}\begin{pmatrix} 17 & -5 \\ 0 & 3 \end{pmatrix} = \boldsymbol{Q}_0\boldsymbol{R}_0.$$

У результаті виконання першого кроку ***QR***-алгоритму знаходимо матрицю $\boldsymbol{A}_1$:

$$\boldsymbol{A}_1 = \boldsymbol{R}_0\boldsymbol{Q}_0 = \frac{1}{\sqrt{17}}\begin{pmatrix} 17 & -5 \\ 0 & 3 \end{pmatrix}\frac{1}{\sqrt{17}}\begin{pmatrix} 4 & 1 \\ -1 & 4 \end{pmatrix} =$$

$$= \begin{pmatrix} 73/17 & -3/17 \\ -3/17 & 12/17 \end{pmatrix} \approx \begin{pmatrix} 4,29 & -0,18 \\ -0,18 & 0,71 \end{pmatrix}. \qquad (11)$$

Нехай зсув σ_0 дорівнює елементу a_{22}, тобто $\sigma_0 = 1$. Для матриці

$$\boldsymbol{A}_0 - \sigma_0\boldsymbol{E} = \begin{pmatrix} 3 & -1 \\ -1 & 0 \end{pmatrix}$$

знаходимо матрицю обертання

$$\boldsymbol{T}_{12}(\psi) = \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix},$$

де, як неважко побачити, $\operatorname{tg}\psi = 1/3,\ \sin\psi = 1/\sqrt{10},\ \cos\psi = 3/\sqrt{10}.$

Далі маємо:

$$\boldsymbol{T}_{12}(\psi)(\boldsymbol{A}_0 - \sigma_0\boldsymbol{E}) = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -1 & 0 \end{pmatrix} = \frac{1}{\sqrt{10}}\begin{pmatrix} 10 & -3 \\ 0 & -1 \end{pmatrix} = \boldsymbol{R},$$

звідки випливає, що

$$\boldsymbol{A}_0 - \sigma_0\boldsymbol{E} = \boldsymbol{T}_{12}^T(\psi)\boldsymbol{R} = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}\frac{1}{\sqrt{10}}\begin{pmatrix} 10 & -3 \\ 0 & -1 \end{pmatrix} = \boldsymbol{Q}_0\boldsymbol{R}_0.$$

Нарешті знаходимо, що

$$\boldsymbol{A}_1 = \boldsymbol{R}_0\boldsymbol{Q}_0 + \sigma_0\boldsymbol{E} =$$

$$= \frac{1}{\sqrt{10}}\begin{pmatrix} 10 & -3 \\ 0 & -1 \end{pmatrix}\frac{1}{\sqrt{10}}\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4,3 & 0,1 \\ 0,1 & 0,7 \end{pmatrix}. \qquad (12)$$

Кінцевою метою ***QR***-алгоритму є одержання матриці $\boldsymbol{A}_k$, в якій лівий нижній елемент є дуже малим. Тоді матрицю можна наближено вважати верхньою трикутною і її власні значення будуть збігатися з елементами головної діагоналі. Звичайно, один крок ***QR***-алгоритму не може привести до такої матриці, проте порівняння формул (11) і (12) указує на те, що ***QR***-алгоритм зі зсувом дає краще наближення до граничної матриці ніж алгоритм без зсуву. △

12.5.4. Степеневі ітерації. ***QR***-алгоритм знаходження власних значень звичайно поєднують з так званим оберненим степеневим методом зі зсувом для знаходження власних векторів.

Звичайний *степеневий метод* реалізують за такою схемою. Вибирається початковий вектор $\boldsymbol{u}_0$, а потім будується послідовність $\boldsymbol{u}_1 = \boldsymbol{A}\boldsymbol{u}_0, \boldsymbol{u}_2 = \boldsymbol{A}\boldsymbol{u}_1$ і в загальному випадку $\boldsymbol{u}_{k+1} = \boldsymbol{A}\boldsymbol{u}_k$. Якщо припустити, що матриця $\boldsymbol{A}$ має повний набір власних векторів $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ відповідних до власних значень $\lambda_1, \ldots, \lambda_n$, то вектор $\boldsymbol{u}_k$ як розв'язок різницевого рівняння на підставі формули (10.1.6) має вигляд:

$$\boldsymbol{u}_k = \sum_{i=1}^{n} c_i \lambda_i^k \boldsymbol{x}_i ,$$

де $c_1, \ldots, c_n$ — деякі числа.

Припустимо, що власні значення занумеровані так, що їх модулі не спадають, кратність λ_n дорівнює одиниці і нема іншого власного значення з таким же модулем. Отже, нехай

$$|\lambda_1| \le |\lambda_2| \le \ldots \le |\lambda_{n-1}| < |\lambda_n|.$$

Тоді, якщо початковий вектор $\boldsymbol{u}_0$ вміщує доданок $c_n \boldsymbol{x}_n$, цей доданок поступово стає домінуючим, тобто $\boldsymbol{u}_k \approx c_n \lambda_n^k \boldsymbol{x}_n$. Отже, $\boldsymbol{u}_k$ прямує при $k \to \infty$ до власного вектора матриці $\boldsymbol{A}$.

В оберненому степеневому методі працюють з матрицею $\boldsymbol{A}^{-1}$ замість $\boldsymbol{A}$, тобто мають справу з ітераціями $\boldsymbol{v}_{k+1} = \boldsymbol{A}^{-1}\boldsymbol{v}_k$.

Проте найкращим є *обернений степеневий метод зі зсувом*, якому відповідають ітерації $\boldsymbol{w}_{k+1} = (\boldsymbol{A} - \sigma\boldsymbol{E})^{-1}\boldsymbol{w}_k$. Як випливає з теорем 6.2.2 і 6.2.3, матриця $(\boldsymbol{A} - \sigma\boldsymbol{E})^{-1}$ має ті ж власні вектори, що і вихідна матриця $\boldsymbol{A}$. При цьому різницеве рівняння має розв'язок

$$\boldsymbol{w}_k = \sum_{i=1}^{n} \frac{c_i \boldsymbol{x}_i}{(\lambda_i - \sigma)^k}.$$

Якщо σ є близьким до λ_i, то знаменник *i*-го доданка цієї суми буде близьким до нуля і, отже, знадобиться лише один або два кроки, щоб зробити *i*-й доданок повністю домінуючим. Коли λ_i вже обчислене за допомогою іншого методу (наприклад, ***QR***-алгоритму), то за σ можна взяти це обчислене значення. Якщо для λ_i ще не одержане наближене значення, то обернений степеневий метод зі зсувом повинен сам вибирати σ або навіть σ_k, якщо ми хочемо змінювати зсув на кожному кроці. У симетричному випадку за σ_k доцільно взяти відношення Релея:

$$\sigma_k = \rho(\boldsymbol{w}_k) = \frac{\boldsymbol{w}_k^T \boldsymbol{A} \boldsymbol{w}_k}{\boldsymbol{w}_k^T \boldsymbol{w}_k}.$$

Обернений степеневий метод з постійним і особливо зі змінним зсувом — найбільш ефективний метод знаходження власних векторів. Збіжність при гарному виборі σ_k настільки швидка, що метод придатний і без указаних вище обмежень на власні значення.

П р и к л а д 2. Для матриці $\boldsymbol{A}$, розглянутої вище в прикладі 1, знайдемо власні вектори оберненим степеневим методом зі зсувом, використовуючи наближені значення власних чисел, одержаних за допомогою ***QR***-алгоритму. За початковий вектор візьмемо $\boldsymbol{w}_0 = (3 \quad -1)^T$.

Формула (12) показує, що наближені значення власних чисел дорівнюють 4,3 і 0,7. У першому випадку $\sigma = 4{,}3$, тому

$$\boldsymbol{A} - \sigma\boldsymbol{E} = \begin{pmatrix} -0{,}3 & -1 \\ -1 & -3{,}3 \end{pmatrix}, \quad (\boldsymbol{A} - \sigma\boldsymbol{E})^{-1} = \begin{pmatrix} 330 & -100 \\ -100 & 30 \end{pmatrix}.$$

Власний вектор наближено дорівнює

$$\boldsymbol{w}_1 = (\boldsymbol{A} - \sigma\boldsymbol{E})^{-1}\boldsymbol{w}_0 = \begin{pmatrix} 330 & -100 \\ -100 & 30 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1090 \\ -330 \end{pmatrix} \approx 165\begin{pmatrix} 6{,}61 \\ -2 \end{pmatrix}. \tag{13}$$

Для другого власного вектора при $\sigma = 0{,}7$ маємо:

$$\boldsymbol{A}-\sigma\boldsymbol{E}=\begin{pmatrix}3,3 & -1\\ -1 & 0,3\end{pmatrix},\quad (\boldsymbol{A}-\sigma\boldsymbol{E})^{-1}=\begin{pmatrix}-30 & -100\\ -100 & -330\end{pmatrix}.$$

Тепер

$$\boldsymbol{w}_1=(\boldsymbol{A}-\sigma\boldsymbol{E})^{-1}\boldsymbol{w}_0=\begin{pmatrix}-30 & -100\\ -100 & -330\end{pmatrix}\begin{pmatrix}3\\ -1\end{pmatrix}=$$

$$=\begin{pmatrix}10\\ 30\end{pmatrix}\approx -15\begin{pmatrix}-0,67\\ -2\end{pmatrix}. \tag{14}$$

Порівняємо власні вектори, знайдені наближено, з їх точними значеннями, які для власних чисел

$$\lambda_1=\frac{5+\sqrt{13}}{2},\quad \lambda_2=\frac{5-\sqrt{13}}{2},$$

є такими

$$\boldsymbol{x}_1=(3+\sqrt{13}\quad -2)^T\approx(6,61\quad -2)^T, \tag{15}$$

$$\boldsymbol{x}_2=(3-\sqrt{13}\quad -2)^T\approx(-0,61\quad -2)^T. \tag{16}$$

Формули (15) і (13) показують, що після першого кроку оберненого степеневого методу зі зсувом одержується досить точне наближення для власного вектора $\boldsymbol{x}_1$. Одночасно з формул (16) і (14) випливає, що наближення для $\boldsymbol{x}_2$ є лише відносно прийнятним за точністю. Виявлена відмінність обумовлена тим, що вже початковий вектор $\boldsymbol{w}_0$ є кращим наближенням для $\boldsymbol{x}_1$ ніж для $\boldsymbol{x}_2$.

Візьмемо інший початковий вектор $\boldsymbol{w}_0=(1\quad 3)^T$, який є кращим наближенням для $\boldsymbol{x}_2$ ніж для $\boldsymbol{x}_1$. У цьому разі замість (13) і (14) будемо мати відповідно:

$$\boldsymbol{w}_1=\begin{pmatrix}330 & -100\\ -100 & 30\end{pmatrix}\begin{pmatrix}1\\ 3\end{pmatrix}=\begin{pmatrix}30\\ -10\end{pmatrix}=5\begin{pmatrix}6\\ -2\end{pmatrix},$$

$$\boldsymbol{w}_1=\begin{pmatrix}-30 & -100\\ -100 & -330\end{pmatrix}\begin{pmatrix}1\\ 3\end{pmatrix}=\begin{pmatrix}-330\\ -1090\end{pmatrix}\approx 545\begin{pmatrix}-0,61\\ -2\end{pmatrix}.$$

Тут ми бачимо, що наближення для $\boldsymbol{x}_1$ є відносно непоганим, а наближення для $\boldsymbol{x}_2$ — вельми точним.

Отже, точність наближення збільшується при вдалому виборі початкового вектора $\boldsymbol{w}_0$. $\triangle$

12.5.5. *QR*-алгоритм знаходження сингулярного розкладу матриці. Теореми 6.6.1 — 6.6.3 вказують найбільш натуральний шлях побудови сингулярного розкладу матриці $\boldsymbol{A}$. Він пов'язаний із знаходженням власних значень матриці $\boldsymbol{A}\boldsymbol{A}^T$ і власних векторів матриць $\boldsymbol{A}\boldsymbol{A}^T$ і $\boldsymbol{A}^T\boldsymbol{A}$. Проте практично цей шлях є мало прийнятним. Це обумовлено тим, що при знаходженні сингулярного розкладу одним з найважливіших є питання про наявність нульового сингулярного числа і його кратність. При визначенні сингулярних чисел як квадратних коренів з власних значень матриці $\boldsymbol{A}\boldsymbol{A}^T$ виникає утруднення, оскільки квадрати малих сингулярних чисел менше відрізняються від нуля ніж самі ці числа, а при обчисленні квадратних коренів з дуже малих чисел відносна похибка виявляється великою. Якщо матриця $\boldsymbol{A}\boldsymbol{A}^T$ все ж використовується, то її елементи повинні обчислюватися з дуже великою точністю.

На практиці застосовується інший метод, заснований на використанні дводіагональної форми матриці. Ми опишемо його для квадратної матриці. Він є придатним і для $m\times n$-матриці, коли $m > n$, а при $m < n$ можна цей метод застосувати до транспонованої матриці. За допомогою перетворень Хаусхолдера матриця $\boldsymbol{A}$ зводиться до дводіагональної форми $\boldsymbol{A}_0$, а потім ця матриця перетворюється до діагональної форми за допомогою ітераційного процесу. Перехід від $\boldsymbol{A}_k$ до $\boldsymbol{A}_{k+1}$ відбувається шляхом множення матриці $\boldsymbol{A}_k$ на послідовність матриць обертання:

$$\boldsymbol{A}_{k+1} = \boldsymbol{T}_{n-1,\,n}(\varphi_n)\,\ldots\,\boldsymbol{T}_{23}(\varphi_3)\boldsymbol{T}_{12}(\varphi_2)\boldsymbol{A}_k\boldsymbol{T}_{12}(\psi_2)\boldsymbol{T}_{23}(\psi_3)\,\ldots\,\boldsymbol{T}_{n-1,\,n}(\psi_n). \quad (17)$$

Нехай кут ψ_2 є довільним, а наступні значення кута виберемо так, щоб матриця $\boldsymbol{A}_{k+1}$ мала ту ж форму, що і $\boldsymbol{A}_k$, тобто була дводіагональною. Це означає, що множення на $\boldsymbol{T}_{12}(\psi_2)$ не обертає в нуль ні одного елемента матриці $\boldsymbol{A}_k$ і робить ненульовим елемент $\{\boldsymbol{A}_k\}_{21}$. Нехай $\boldsymbol{T}_{12}(\varphi_2)$ обертає в нуль $\{\boldsymbol{A}_k\boldsymbol{T}_{12}(\psi_2)\}_{21}$ і робить ненульовим елемент $\{\boldsymbol{A}_k\boldsymbol{T}_{12}(\psi_2)\}_{13}$, потім множник $\boldsymbol{T}_{23}(\psi_3)$ обертає в нуль $\{\boldsymbol{T}_{12}(\varphi_2)\boldsymbol{A}_k\boldsymbol{T}_{12}(\psi_2)\}_{13}$ і робить ненульовим елемент $\{\boldsymbol{T}_{12}(\varphi_2)\boldsymbol{A}_k\boldsymbol{T}_{12}(\psi_2)\}_{32}$ і т. д. Нарешті, $\boldsymbol{T}_{n-1,\,n}(\varphi_n)$ обертає в нуль елемент з n-го рядка і ($n-1$)-го стовпця і не добавляє ненульових елементів.

Рівність (17) можна записати так:

$$\boldsymbol{A}_{k+1} = \boldsymbol{Q}_1\boldsymbol{A}_k\boldsymbol{Q}_2\,, \tag{18}$$

де через $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2$ позначені ортогональні матриці, які є добутками відповідних матриць обертання. Для матриці $\boldsymbol{M}_k = \boldsymbol{A}_k^T\boldsymbol{A}_k$, яка є тридіагональною, на підставі рівності (18) маємо:

$$\boldsymbol{M}_{k+1} = \boldsymbol{A}_{k+1}^T\boldsymbol{A}_{k+1} = \boldsymbol{Q}_2^T\boldsymbol{M}_k\boldsymbol{Q}_2\,. \tag{19}$$

Можна довести, що перехід від матриці $\boldsymbol{M}_k$ до матриці $\boldsymbol{M}_{k+1}$, який визначається за допомогою формули (19), відповідає одному кроку ***QR***-алгоритму зі зсувом і цей перехід можна забезпечити належним вибором кута ψ_2.

Таким чином, кожний крок алгоритму, зберігаючи дводіагональний вигляд матриці, зменшує модулі її позадіагональних елементів. Зрештою вони стають порівнянними з похибками обчислень і їх можна покласти рівними нулю, після чого одержується сингулярний розклад матриці.

Повне обговорення деталей застосування ***QR***-алгоритму до проблеми обчислення сингулярного розкладу виходить за межі цієї книги. Укажемо лише на ту обставину, що кут φ_2 можна знаходити, використовуючи тільки матрицю $\boldsymbol{A}_k$ і не обчислюючи матрицю $\boldsymbol{M}_k$.

Описана процедура застосовується до дводіагональної матриці з ненульовими елементами головної і прилеглої до неї діагоналей. Якщо принаймні один з цих елементів є нульовим, то задачу знаходження сингулярного розкладу можна спростити завдяки тому, що дводіагональна матриця зводиться до квазідіагональної. Це цілком очевидно у випадку, коли нульовим є елемент діагоналі, прилеглої до головної. Якщо ж нульовим є елемент головної діагоналі, нехай це буде i-й, то множенням зліва на матриці обертання i-й рядок можна зробити нульовим, після чого дводіагональна матриця стає квазідіагональною. Покажемо це на прикладі матриці п’ятого порядку у випадку, коли $i = 2$. Послідовність матриць, яка вміщує $\boldsymbol{A}_0$ і одержані з неї матриці, виглядає так:

$$\boldsymbol{A}_0 = \begin{pmatrix} * & * & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{pmatrix}, \quad \boldsymbol{T}_{23}\boldsymbol{A}_0 = \begin{pmatrix} * & * & 0 & 0 & 0 \\ 0 & 0 & 0 & + & 0 \\ 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{pmatrix},$$

$$\boldsymbol{T}_{24}\boldsymbol{T}_{23}\boldsymbol{A}_0 = \begin{pmatrix} * & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & + \\ 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{pmatrix}, \quad \boldsymbol{T}_{25}\boldsymbol{T}_{24}\boldsymbol{T}_{23}\boldsymbol{A}_0 = \begin{pmatrix} * & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \end{pmatrix}.$$

Остання матриця є квазідіагональною.

Аналізуючи особливості сингулярних базисів квазідіагональних матриць за допомогою теореми 7.3.1, неважко переконатися в тому, що ортогональні матриці з сингулярного розкладу таких матриць повинні бути також квазідіагональними. Тому задача знаходження цього розкладу «розщеплюється» на менші задачі, відповідні до діагональних підматриць.

12.5.6. Застосування сингулярного розкладу. Знання сингулярного розкладу дозволяє визначити ранг матриці, який збігається з кількістю її ненульових сингулярних чисел.

Матриця $\boldsymbol{A}$ розмірів $m \times n$, де $m \ge n$, називається матрицею повного рангу, якщо $r(\boldsymbol{A}) = n$, або матрицею недостатнього рангу, якщо $r(\boldsymbol{A}) < n$. Оскільки ранг матриці завжди повинен бути цілим числом, він обов’язково буде розривною функцією від елементів матриці. Як завгодно малі зміни в матриці недостатнього рангу можуть перетворити її в матрицю повного рангу, як ми бачили це при розгляді прикладу 2 з розділу 12.2. На практиці ми маємо справу з ефективним рангом — кількістю сингулярних чисел, більших деякої величини, яка відображає точність даних. Це також розривна функція, але її розриви менш численні і неприємні ніж у теоретичного рангу. Значною перевагою використання сингулярного розкладу для визначення рангу матриці є те, що рішення про нехтування треба виносити лише стосовно малих чисел (малих сингулярних чисел), а не векторів чи систем векторів.

Нехай треба розв'язати систему $\boldsymbol{Ax} = \boldsymbol{b}$ за допомогою методу найменших квадратів, скориставшись для цього сингулярним розкладом матриці $\boldsymbol{A}$. Як зазначено вище, при його обчисленні ми повинні вирішити, які з малих сингулярних чисел можна вважати нульовими. Для кожної конкретної задачі, виходячи з точності вихідних даних і оцінки похибки округлень, вибирають деяке число ε для знаходження другої форми сингулярного розкладу: якщо обчислене значення сингулярного числа є меншим ніж ε, то його вважають рівним нулю.

Визнаючи деяке додатне сингулярне число рівним нулю, ми тим самим вважаємо деякий стовпець матриці $\boldsymbol{A}$ лінійно залежним від інших стовпців. Це означає, що отримана на підставі формули (11.6.1) матриця $\boldsymbol{A}_\varepsilon$ в порівнянні з $\boldsymbol{A}$ несе меншу інформацію про вихідну задачу, проте має більш компактний вигляд ніж $\boldsymbol{A}$. Вибираючи ε відносно великим, ми робимо задачу більш простою і одночасно менш змістовною. Тому збільшувати ε треба обачливо, щоб утримати в матриці $\boldsymbol{A}_\varepsilon$ достатньо інформації про вихідну задачу.

Вправи до глави 12

12.1. Нехай найбільше значення в правій частині формули (12.1.9) досягається при $j = k$. Показати, що за вектор $\boldsymbol{x}_0$, для якого досягається це значення, можна взяти вектор натурального базису $\boldsymbol{e}_k$.

Припустимо, що найбільше значення в правій частині формули (12.1.10) досягається при $i = l$. Показати, що це відбувається для вектора $\boldsymbol{x}_0$ з координатами $x_{0m} = |a_{lm}| a_{lm}^{-1}$ при $a_{lm} \neq 0$ і $x_{0m} = 1$ при $a_{lm} = 0$.

12.2. Довести, що евклідова і спектральна норми не змінюються при множенні матриці справа і зліва на будь-які ортогональні матриці.

12.3. Показати, що для будь-якого власного значення λ невиродженої матриці $\boldsymbol{A}$ має місце оцінка

$$\frac{1}{\|\boldsymbol{A}^{-1}\|} \leq |\lambda| \leq \|\boldsymbol{A}\|.$$

12.4. Довести для числа зумовленості, що

$$c(\boldsymbol{A}) = c(\boldsymbol{A}^{-1}), \qquad c(\boldsymbol{AB}) \leq c(\boldsymbol{A})c(\boldsymbol{B}).$$

12.5. При розгляді прикладу 2 з розділу 12.2 було з’ясовано, що малі зміни координат векторів можуть призвести до порушення їх лінійної залежності або незалежності. Які висновки можна звідси зробити стосовно задачі про перетин гіперплощин?

12.6. Якими є джерела нестійкості при виконанні обчислень:
1) скелетного розкладу матриці за схемою, викладеною в розділі 3.4;
3) жорданової форми матриці за схемою, викладеною в розділі 9.2;
2) псевдооберненої матриці за методом Гревілля?

12.7. Проаналізувати стійкість проблеми власних значень, замінивши умови (12.2.12) для рядків аналогічними умовами для стовпців матриці.

12.8. Для системи рівнянь $\boldsymbol{Ax} = \boldsymbol{b}$, поданої у вигляді:

$$\boldsymbol{x} = \left(\boldsymbol{E} - \frac{2\boldsymbol{A}}{\|\boldsymbol{A}\|}\right)\boldsymbol{x} + \frac{2\boldsymbol{b}}{\|\boldsymbol{A}\|},$$

довести збіжність методу простої ітерації за умови додатної визначеності матриці $\boldsymbol{A}$.

12.9. Показати, що обернений степеневий метод пристосований для знаходження власного вектора, відповідного до найменшого за модулем власного значення.

ВКАЗІВКИ ДО ВПРАВ

1.1. Узяти за $\boldsymbol{x}$ стовпець $\boldsymbol{e}_{*i}$ матриці $\boldsymbol{E}$. Тоді $\boldsymbol{Ax} = \boldsymbol{a}_{*i} = \boldsymbol{0}$, де $\boldsymbol{a}_{*i}$ — стовпець матриці $\boldsymbol{A}$.

1.3. Для матриць $\boldsymbol{L}_i(\lambda)$ і $\boldsymbol{L}_i(\lambda\,|\,j)$ вектори $\boldsymbol{u}$ і $\boldsymbol{v}$ є стовпцями матриці $\boldsymbol{E}$, а для $\boldsymbol{L}_{ij}$ ці вектори є різницями стовпців $\boldsymbol{E}$.

1.4. Визначник $\det(\boldsymbol{A}^{-1})$ дорівнює одиниці. Скористатися формулою (1.3.2).

2.3. $\boldsymbol{L}^{-1}(\boldsymbol{A} \quad \boldsymbol{E}) = (\boldsymbol{L}^{-1}\boldsymbol{A} \quad \boldsymbol{L}^{-1}) = (\boldsymbol{R} \quad \boldsymbol{L}^{-1})$.

2.4. $\boldsymbol{A}^{-1}(\boldsymbol{A} \quad \boldsymbol{B}) = (\boldsymbol{E} \quad \boldsymbol{A}^{-1}\boldsymbol{B})$; $\quad \begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \boldsymbol{B}^{-1} = \begin{pmatrix} \boldsymbol{AB}^{-1} \\ \boldsymbol{E} \end{pmatrix}$.

2.5. Скористатися блоковим представленням множників $\boldsymbol{LDU}$-розкладу і застосувати теорему 2.2.1.

2.6. Навести належні приклади.

2.7. Урахувати, що $\boldsymbol{AP}^T = (\boldsymbol{PA}^T)^T$, і скористатися міркуваннями щодо рядків матриці $\boldsymbol{PA}$.

3.1. Скористатися визначенням лінійної залежності (незалежності) і теоремою 3.1.1.

3.5, 3.6. Скористатися теоремою 3.4.2.

3.7. Можливість представлення у вигляді суми r матриць випливає з теореми про базисний мінор, а неможливість представлення у вигляді суми меншої кількості матриць випливає з теореми 3.4.7.

3.8. $\boldsymbol{A} = \boldsymbol{L}^{-1}\,\boldsymbol{\Delta}\,\boldsymbol{M}^{-1} = \left(\boldsymbol{L}^{-1}\begin{pmatrix} \boldsymbol{E}_r \\ \boldsymbol{0} \end{pmatrix}\right)\left((\boldsymbol{E}_r \quad \boldsymbol{0})\boldsymbol{M}^{-1}\right)$.

3.9. При доведенні достатності скористатися теоремою 3.4.3.

3.10. Скористатися теоремою Кронекера — Капеллі.

3.12. Показати, що пряма лінія паралельна до площини і скористатися теоремою 3.6.2.

3.13. З'ясувати, яким є взаємне розміщення спрямовуючих підпросторів прямих ліній і (ненульового) вектора, що з'єднує їх початкові точки.

3.14. Скористатися теоремою 3.5.4.

5.1. Визначник Грама векторів $\boldsymbol{a}_{1*}^T, \ldots, \boldsymbol{a}_{m*}^T, \boldsymbol{x}$ є невід'ємною величиною:

$$G(\boldsymbol{a}_{1*}^T, \ldots, \boldsymbol{a}_{m*}^T, \boldsymbol{x}) \geq 0.$$

Цю нерівність можна записати так:

$$\boldsymbol{x}^T \boldsymbol{x} \det \boldsymbol{\Gamma}_m + \begin{vmatrix} & & & \boldsymbol{a}_{1*}\boldsymbol{x} \\ & \boldsymbol{\Gamma}_m & & \cdots \\ & & & \boldsymbol{a}_{m*}\boldsymbol{x} \\ \boldsymbol{x}^T\boldsymbol{a}_{1*}^T & \ldots & \boldsymbol{x}^T\boldsymbol{a}_{m*}^T & 0 \end{vmatrix} \geq 0,$$

де $\boldsymbol{\Gamma}_m$ — матриця Грама векторів $\boldsymbol{a}_{1*}^T, \ldots, \boldsymbol{a}_{m*}^T$. З урахуванням умови $\boldsymbol{Ax} = \boldsymbol{b}$ остання нерівність набуває вигляду:

$$\boldsymbol{x}^T \boldsymbol{x} \geq -\frac{1}{\det \boldsymbol{\Gamma}_m} \begin{vmatrix} \boldsymbol{\Gamma}_m & \boldsymbol{b} \\ \boldsymbol{b}^T & 0 \end{vmatrix}.$$

Тут права частина нерівності є мінімальним значенням $|\,\boldsymbol{x}\,|^2$ і досягається воно, якщо вибрати $\boldsymbol{x}$ рівним лінійній комбінації векторів $\boldsymbol{a}_{1*}^T, \ldots, \boldsymbol{a}_{m*}^T$. Саме в цьому випадку нерівність $G(\boldsymbol{a}_{1*}^T, \ldots, \boldsymbol{a}_{m*}^T, \boldsymbol{x}) \geq 0$ перетворюється в рівність. Отже, $\boldsymbol{x} = \sum_{j=1}^{m} c_j \boldsymbol{a}_{j*}^T$. Після підставлення $\boldsymbol{x}$ у рівняння $\boldsymbol{Ax} = \boldsymbol{b}$ одержується система рівнянь $\sum_{j=1}^{m} \boldsymbol{a}_{i*}\boldsymbol{a}_{j*}^T c_j = b_i$, $i = \overline{1,\ m}$.

5.2. Нехай $\boldsymbol{x} = \boldsymbol{f}_0 + \alpha_1 \boldsymbol{f}_1$ — векторне рівняння шуканої прямої лінії в просторі S_3. Тут $\boldsymbol{f}_0,\ \boldsymbol{f}_1$ — відомий і шуканий вектори відповідно. Якщо задані прямі лінії визначаються векторними рівняннями $\boldsymbol{x} = \boldsymbol{g}_0 + \beta_1 \boldsymbol{g}_1$ і $\boldsymbol{x} = \boldsymbol{h}_0 + \gamma_1 \boldsymbol{h}_1$, то умови перпендикулярності мають вигляд:

$$\boldsymbol{g}_1^T \boldsymbol{f}_1 = 0, \quad \boldsymbol{h}_1^T \boldsymbol{f}_1 = 0.$$

Маємо однорідну систему двох рівнянь відносно трьох невідомих координат вектора $\boldsymbol{f}_1$. Оскільки вектори $\boldsymbol{g}_1$ і $\boldsymbol{h}_1$ не є колінеарними, то фундаментальна система розв'язків складається з одного вектора, який є шуканим вектором $\boldsymbol{f}_1$.

6.1. Послідовно розглянути випадки рівностей для z_1 і z_2: 1) $z_1 = 0$; 2) $z_2 = 0$; 3) $\arg z_1 = \arg z_2$; 4) $\arg z_1 = \arg z_2 + \pi$.

6.3. За допомогою елементарних перетворень з рядками першого діагонального блоку звести визначник до такого вигляду, коли цей блок стає трикутним.

6.5. Нульове власне значення має кратність $n-1$. Йому відповідають власні вектори, що знаходяться з рівняння $\boldsymbol{b}^T \boldsymbol{x} = 0$. Власному вектору $\boldsymbol{a}$ відповідає власне значення $\boldsymbol{b}^T \boldsymbol{a}$.

6.10. Застосувати висновок з теореми 3.4.4 до матриці $\boldsymbol{S\Lambda S}^{-1}$.

6.11. Скористатися теоремою 6.2.5.

6.14. Побудувати відповідну матрицю Фробеніуса.

6.15. Матриця $\boldsymbol{A} = \boldsymbol{B}\operatorname{diag}(d_1,\ \ldots,\ d_n)$ є матрицею Адамара.

6.16. Нехай i-та нерівність є строгою. Зменшивши d_i настільки мало, щоб ця нерівність не порушилася, можна інші умови перетворити в строгі нерівності.

6.17. У рівності $\boldsymbol{A}(\boldsymbol{u} + \boldsymbol{v}i) = (\mu + \nu i)(\boldsymbol{u} + \boldsymbol{v}i)$ порівняти дійсні і уявні частини. Потім помножити зліва на $\boldsymbol{v}^T$ і $\boldsymbol{u}^T$ відповідно, транспонувати одну з рівностей і взяти їх різницю, після чого дійти необхідних висновків. 2) Рівності $\boldsymbol{Ax} = \lambda \boldsymbol{x}$, $\boldsymbol{Ay} = \mu \boldsymbol{y}$ помножити зліва на

$\boldsymbol{y}^T$ і $\boldsymbol{x}^T$ відповідно, одну з них транспонувати і взяти їх різницю, після чого дійти необхідного висновку.

6.18. Скористатися вправою 6.17 і теоремою 6.2.6.

6.21. $\boldsymbol{Hu} = (\boldsymbol{E} - 2\boldsymbol{uu}^T)\boldsymbol{u} = -\boldsymbol{u}; \quad \boldsymbol{Hv} = (\boldsymbol{E} - 2\boldsymbol{uu}^T)\boldsymbol{v} = \boldsymbol{v}\,(\boldsymbol{v} \perp \boldsymbol{u})$.

6.25. Перевірити виконання рівності

$$(\alpha\boldsymbol{A} + \beta\boldsymbol{E})(\bar{\alpha}\boldsymbol{A}^* + \bar{\beta}\boldsymbol{E}) = (\bar{\alpha}\boldsymbol{A}^* + \bar{\beta}\boldsymbol{E})(\alpha\boldsymbol{A} + \beta\boldsymbol{E}).$$

6.26. Якщо $(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$, то з урахуванням вправи 6.25 будемо мати:

$$0 = ((\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x})^*(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{x}^*(\lambda\boldsymbol{E} - \boldsymbol{A})^*(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} =$$

$$= \boldsymbol{x}^*(\lambda\boldsymbol{E} - \boldsymbol{A})(\lambda\boldsymbol{E} - \boldsymbol{A})^*\boldsymbol{x} = ((\bar{\lambda}\boldsymbol{E} - \boldsymbol{A}^*)\boldsymbol{x})^*(\bar{\lambda}\boldsymbol{E} - \boldsymbol{A}^*)\boldsymbol{x},$$

тобто $(\bar{\lambda}\boldsymbol{E} - \boldsymbol{A}^*)\boldsymbol{x} = \boldsymbol{0}$.

6.28. Рівності $\boldsymbol{Ax} = \lambda\boldsymbol{x}$, $\boldsymbol{A}^*\boldsymbol{x} = \mu\boldsymbol{x}$ помножити зліва на $\boldsymbol{x}^*$, до однієї з них застосувати операцію спряження і взяти їх різницю, після чого дійти необхідного висновку.

7.1. $\boldsymbol{A} = \boldsymbol{U}^T\boldsymbol{DU} = (\boldsymbol{D}^{1/2}\boldsymbol{U})^T\boldsymbol{D}^{1/2}\boldsymbol{U} = \boldsymbol{R}^T\boldsymbol{R}$.

7.2. $\boldsymbol{C}^T\boldsymbol{AC} = \boldsymbol{C}^T\boldsymbol{W}^T\boldsymbol{WC} = (\boldsymbol{WC})^T\boldsymbol{WC}$.

7.4. $r(\boldsymbol{C}^T\boldsymbol{AC}) = r(\boldsymbol{C}^T\boldsymbol{W}^T\boldsymbol{WC}) = r((\boldsymbol{WC})^T\boldsymbol{WC}) = r(\boldsymbol{WC}) = r(\boldsymbol{C})$.

7.5. $\boldsymbol{C}^T\boldsymbol{AC} = \boldsymbol{C}^T\boldsymbol{Q}\Lambda\boldsymbol{Q}^T\boldsymbol{C} = (\Lambda^{1/2}\boldsymbol{Q}^T\boldsymbol{C})^T\Lambda^{1/2}\boldsymbol{Q}^T\boldsymbol{C}$, тому $\Lambda^{1/2}\boldsymbol{Q}^T\boldsymbol{C} = \boldsymbol{E}$.

7.6. Скориставшись зауваженням 1 з розділу 7.1, дослідити знак мінора другого порядку, у правому верхньому куті якого розміщений максимальний за модулем елемент.

7.7. Розкласти $\det\boldsymbol{B}$ за елементами i-го рядка або i-го стовпця.

7.8. При $t = 0$ всі провідні мінори матриці $\boldsymbol{A} + t\boldsymbol{B}$ є додатними. Оскільки провідні мінори цієї матриці є неперервними функціями від t, вони залишаються додатними при $|t| < \delta$, де δ — деяке досить мале число.

7.9. $\boldsymbol{B}^{-1} = (\boldsymbol{W}^T\boldsymbol{W})^{-1} = \boldsymbol{W}^{-1}(\boldsymbol{W}^{-1})^T = \boldsymbol{X}^T\boldsymbol{X}$, де $\boldsymbol{X} = (\boldsymbol{W}^{-1})^T$ — невироджена матриця.

7.11. Скористатися ознакою 1) з теорем 7.1.1 і 7.1.2.

7.12. Скористатися теоремою 6.2.1.

7.14. Скористатися першою формулою (7.1.2) і формулою (5.1.8).

7.15. З рівності $\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{Q} = \mathrm{diag}(\Lambda_r, \boldsymbol{0})$, де $\Lambda_r = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$, випливає, що $\boldsymbol{Q}_r^T\boldsymbol{A}\boldsymbol{Q}_r = \Lambda_r$ де $\boldsymbol{Q}_r$ — матриця, утворена першими r стовпцями матриці $\boldsymbol{Q}$. Далі покласти $\boldsymbol{Q}_r = \boldsymbol{C}\Lambda_r^{1/2}$.

7.17. З частини 5) теореми 7.1.2 випливає, що $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{B}^T\boldsymbol{B}$, де $\boldsymbol{B} = \boldsymbol{W}\boldsymbol{x}$. Оскільки $\boldsymbol{B}^T\boldsymbol{B} = \boldsymbol{0}$, то для кожного стовпця $\boldsymbol{b}_{*i}$ матриці $\boldsymbol{B}$ виконується рівність $\boldsymbol{b}_{*i}^T\boldsymbol{b}_{*i} = 0$. Це означає, що всі стовпці матриці $\boldsymbol{B}$ є нульовими. Тому $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{W}^T\boldsymbol{B} = \boldsymbol{0}$.

7.18. Скористатися законом інерції квадратичних форм.

7.21. Скористатися третьою формулою (1.4.6).

7.22. $$\boldsymbol{x}^T\boldsymbol{\Gamma}_m\boldsymbol{x} = \left(\left(\sum_{i=1}^m x_i\boldsymbol{a}_i^T\boldsymbol{a}_1\right)\left(\sum_{i=1}^m x_i\boldsymbol{a}_i^T\boldsymbol{a}_2\right)\ldots\left(\sum_{i=1}^m x_i\boldsymbol{a}_i^T\boldsymbol{a}_m\right)\right)\boldsymbol{x} =$$

$$= \left(\sum_{i=1}^m x_i\boldsymbol{a}_i^T\right)\left(\sum_{j=1}^m x_j\boldsymbol{a}_j\right) = \boldsymbol{y}^T\boldsymbol{y}.$$

7.23. $\mathrm{tr}\,(\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}) = \mathrm{tr}\,(\boldsymbol{A}\boldsymbol{Q}\boldsymbol{Q}^{-1}) = \mathrm{tr}\,\boldsymbol{A}$.

7.24. Якщо вважати, що $\boldsymbol{x} = (\boldsymbol{A}^{1/2})^{-1}\boldsymbol{y}$, то функція $f(\boldsymbol{x})$ перетворюється у відношення Релея для квадратичної форми $\boldsymbol{y}^T\boldsymbol{a}\boldsymbol{a}^T\boldsymbol{y}$, де $\boldsymbol{a} = (\boldsymbol{A}^{1/2})^{-1}\boldsymbol{b}$. Оскільки $r(\boldsymbol{a}\boldsymbol{a}^T) = 1$, то матриця $\boldsymbol{a}\boldsymbol{a}^T$ має нульове власне значення кратності $n - 1$. Ще одне власне значення, обчислене за допомогою формули (6.2.2), дорівнює

$$\mathrm{tr}\,(\boldsymbol{a}\boldsymbol{a}^T) = \mathrm{tr}\,(\boldsymbol{a}^T\boldsymbol{a}) = \boldsymbol{a}^T\boldsymbol{a}.$$

Вона є найбільшим власним значенням матриці $\boldsymbol{a}\boldsymbol{a}^T$, тому

$$\max f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{a} = \boldsymbol{b}^T\boldsymbol{A}^{-1}\boldsymbol{b}.$$

Рівняння $(\lambda \boldsymbol{E} - \boldsymbol{a}\boldsymbol{a}^T)\boldsymbol{y} = \mathbf{0}$, яке визначає власний вектор $\boldsymbol{y}$, показує, що вектори $\boldsymbol{y}$ і $\boldsymbol{a}$ є колінеарними, тому можна вважати, що $\boldsymbol{y} = (\boldsymbol{A}^{1/2})^{-1}\boldsymbol{b}$ і, отже, $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$.

9.11. Нехай матриця $\boldsymbol{T}_0$ є такою, що $\boldsymbol{T}_0^{-1}\boldsymbol{A}\boldsymbol{T}_0 = \boldsymbol{J}$. Візьмемо матрицю $\boldsymbol{D} = \operatorname{diag}(1, \varepsilon, \ldots, \varepsilon^{n-1})$ і помножимо рівність $\boldsymbol{T}_0^{-1}\boldsymbol{A}\boldsymbol{T}_0 = \boldsymbol{J}$ зліва на $\boldsymbol{D}^{-1}$ і справ на $\boldsymbol{D}$. Тоді $\boldsymbol{D}^{-1}\boldsymbol{J}\boldsymbol{D}$ буде квазідіагональною матрицею необхідного вигляду. Перетворююча матриця дорівнює $\boldsymbol{T} = \boldsymbol{T}_0\boldsymbol{D}$.

11.1, 11.2. Скористатися формулою (11.1.1) або (11.1.2).

11.3. Скористатися теоремою Пенроуза.

11.4. На підставі формули(11.1.15) і теореми 3.4.4

$$r(\boldsymbol{A}\boldsymbol{B}) = r(\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{B}) \le r(\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{B}).$$

З іншого боку, $r(\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{B}) \le r(\boldsymbol{A}\boldsymbol{B})$. Отже, $r(\boldsymbol{A}\boldsymbol{B}) = r(\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{B})$.

11.5. Перша рівність доводиться за допомогою формул (11.2.20) і (11.2.21), а друга одержується з першої після транспонування обох її частин.

11.6. $\boldsymbol{Q}(\boldsymbol{M}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Q})^+\boldsymbol{M}\boldsymbol{A}^T = \boldsymbol{Q}\boldsymbol{Q}^T(\boldsymbol{M}\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{M}\boldsymbol{A}^T = (\boldsymbol{M}\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{M}\boldsymbol{A}^T =$

$= (\boldsymbol{M}\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{M}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^+\boldsymbol{A}^T = \boldsymbol{A}^+$. При доведенні врахована перша формула з вправи 11.5.

11.8. Не зменшуючи загальності, можна вважати, що матриця $\boldsymbol{A}$ має лінійно незалежні рядки. Скориставшись псевдорозв'язком (11.4.1), одержимо з урахуванням формули (11.1.6):

$$\boldsymbol{x}_0 = \boldsymbol{A}^+\boldsymbol{b} = \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{b}.$$

Нехай $(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{b} = \boldsymbol{c}$, тоді $\boldsymbol{x}_0 = \boldsymbol{A}^T\boldsymbol{c} = \sum_{j=1}^{m} c_j \boldsymbol{a}_{j*}$. Оскільки $\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{c} = \boldsymbol{b}$, то координати вектора $\boldsymbol{c}$ задовольняють системі рівнянь

$$\begin{pmatrix} \boldsymbol{a}_{1*}\boldsymbol{a}_{1*}^T & \dots & \boldsymbol{a}_{1*}\boldsymbol{a}_{m*}^T \\ \dots & \dots & \dots \\ \boldsymbol{a}_{m*}\boldsymbol{a}_{1*}^T & \dots & \boldsymbol{a}_{m*}\boldsymbol{a}_{m*}^T \end{pmatrix} \begin{pmatrix} c_1 \\ \dots \\ c_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \dots \\ b_m \end{pmatrix}.$$

Отже, маємо розв'язок вправи 5.1.

11.9. Задача зводиться до знаходження вектора $\boldsymbol{y} = \boldsymbol{A}^{1/2}\boldsymbol{x}$ найменшої довжини з сумісної системи рівнянь $\boldsymbol{B}(\boldsymbol{A}^{1/2})^{-1}\boldsymbol{y} = \boldsymbol{c}$. На підставі теореми 11.4.1 $\boldsymbol{y} = (\boldsymbol{B}(\boldsymbol{A}^{1/2})^{-1})^{+}\boldsymbol{c}$, звідки $\boldsymbol{x} = (\boldsymbol{A}^{1/2})^{-1}(\boldsymbol{B}(\boldsymbol{A}^{1/2})^{-1})^{+}\boldsymbol{c}$. З урахуванням другої рівності (11.1.25) $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{B}^T\boldsymbol{S}^{+}\boldsymbol{c}$, де $\boldsymbol{S} = \boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}^T$. Величина $f_{\min}$ знаходиться після підставлення $\boldsymbol{x}$ у квадратичну форму і врахування формули (11.1.16).

11.11. Порівняти порядок і ранг матриці системи. Скористатися першою рівністю (11.3.3).

11.12. Матриця (11.5.3) є такою:

$$\boldsymbol{U} = \sqrt{\frac{2}{n}} \begin{pmatrix} 1/\sqrt{2}\, T_0(x_1) & T_1(x_1) & \dots & T_m(x_1) \\ 1/\sqrt{2}\, T_0(x_2) & T_1(x_2) & \dots & T_m(x_2) \\ \dots & \dots & \dots & \dots \\ 1/\sqrt{2}\, T_0(x_n) & T_1(x_n) & \dots & T_m(x_n) \end{pmatrix}.$$

Тут $x_l = t_{nl} = \cos\left(l - \dfrac{1}{2}\right)\beta, \quad l = \overline{1, n}, \quad \beta = \dfrac{\pi}{n}$. Ортонормованість стовпців матриці $\boldsymbol{U}$ доводиться за допомогою рівності (10.1.14) і формули Ейлера. Коефіцієнти узагальненого многочлена на підставі формули (11.5.4) дорівнюють

$$z_0 = \frac{1}{\sqrt{n}} \sum_{l=1}^{m} f(x_l), \quad z_k = \sqrt{\frac{2}{n}} \sum_{l=1}^{m} f(x_l) T_k(x_l), \quad k = \overline{1, m}.$$

12.2. Якщо $\boldsymbol{A}$ замінити матрицею $\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2$, де $\boldsymbol{Q}_1$ і $\boldsymbol{Q}_2$ — ортогональні матриці, то замість матриці $\boldsymbol{A}\boldsymbol{A}^T$ будемо мати $\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{Q}_1^T$. Ця матриця на підставі теореми 6.2.5 має ті ж власні числа, що і $\boldsymbol{A}\boldsymbol{A}^T$.

Це означає незмінність сингулярних чисел матриці $\boldsymbol{A}$ і, отже, незмінність спектральної норми. Незмінність евклідової норми випливає з теореми 12.1.9.

12.3. Скористатися теоремою 12.1.6, урахувавши при цьому рівність $\boldsymbol{A}^{-1}\boldsymbol{x} = \lambda^{-1}\boldsymbol{x}$.

12.4. Скористатися формулою (12.2.5) і аксіомою 4) норми матриці.

12.8. Скориставшись теоремою 6.2.2, довести, що власні значення матриці $\boldsymbol{E} - 2\| \boldsymbol{A} \|^{-1}\boldsymbol{A}$ розміщені в інтервалі $(-1,\ 1)$.

12.9. Порівняти зі степеневим методом.

ПРЕДМЕТНИЙ ПОКАЖЧИК

Навчальне видання

ТРАВКІН
Юрій Іванович

ЛІНІЙНА АЛГЕБРА І АНАЛІТИЧНА ГЕОМЕТРІЯ

Навчальний посібник

Видано в авторській редакції

Комп'ютерна верстка *В. Ю. Куплевацького*

Підписано до друку 2.04.2009.
Формат 60x84/16. Гарнітура «Таймс».
Друк офсетний. Папір офсетний.
Ум.-друк. арк. 26,12. Обл.-вид. арк. 27,2.
Наклад 300 прим. Зам. № 09-23.

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців і розповсюджувачів
видавничої продукції ДК № 1002 від 31.07.2002 р.

Видання і друк ТОВ «Майдан»
61002, Харків, вул. Чернишевська, 59.
Тел.: (8-057) 700-37-30