САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Физический факультет
Кафедра высшей математики и математической физики

Т. А. Суслина

ЛИНЕЙНАЯ АЛГЕБРА
II СЕМЕСТР
Выпуск 2

Учебно-методическое пособие

Санкт-Петербург
2021 г.

- Рецензенты: проф., д.ф.-м.н. Ф. В. Петров;
  к.ф.-м.н. Н. Н. Сеник.
- Печатается по решению учебно-методической комиссии физического факультета СПбГУ.

Т. А. Суслина.
ЛИНЕЙНАЯ АЛГЕБРА. II СЕМЕСТР. Выпуск 2. — СПб.: СПбГУ, 2021. – 35 с.

Настоящее пособие содержит теоретический материал по курсу “Высшая алгебра”, который читается во втором семестре студентам первого курса физического факультета СПбГУ. Выпуск 2 содержит главу 2 “Линейные операторы в конечномерном линейном пространстве”. Пособие может быть полезно преподавателям, ведущим занятия по курсу “Высшая алгебра”, а также студентам, изучающим этот предмет.

Выпуск 2 является продолжением выпуска 1 учебно-методического пособия “Линейная алгебра. II семестр” (см. [3]), основанного на курсе лекций для студентов физического факультета СПбГУ. Выпуск 2 содержит главу 2 “Линейные операторы в конечномерном линейном пространстве”.

Глава 2. Линейные операторы в конечномерном линейном пространстве

§ 1. Преобразование базисов, координат и изображающих матриц линейных операторов

В главе 2 считаем, что E — *конечномерное линейное пространство над полем* K, *где* $K = \mathbb{R}$ *или* $K = \mathbb{C}$. В случае $K = \mathbb{R}$ мы называем E *вещественным линейным пространством*, в случае $K = \mathbb{C}$ — *комплексным линейным пространством.*

1.1. **Преобразование базисов.** Пусть в пространстве E заданы два базиса $\mathbf{e} = \{e_1, \dots, e_n\}$ и $\widetilde{\mathbf{e}} = \{\widetilde{e}_1, \dots, \widetilde{e}_n\}$. Разложим векторы $\widetilde{e}_k$ в базисе $\mathbf{e}$:

$$\widetilde{e}_k = \sum_{j=1}^{n} \beta_k^j e_j, \quad k = 1, \dots, n. \tag{1.1}$$

Определение 1.1. *Матрицей перехода от базиса* $\mathbf{e}$ *к базису* $\widetilde{\mathbf{e}}$ *называется матрица*

$$b = \begin{pmatrix} \beta_1^1 & \beta_2^1 & \dots & \beta_n^1 \\ \beta_1^2 & \beta_2^2 & \dots & \beta_n^2 \\ \dots & \dots & \dots & \dots \\ \beta_1^n & \beta_2^n & \dots & \beta_n^n \end{pmatrix},$$

элементы которой β_k^j *— координаты “новых” базисных векторов в “старом” базисе — определены в* (1.1).

Тогда k-й столбец матрицы b состоит из координат вектора $\widetilde{e}_k$ в "старом" базисе. Очевидно, столбцы матрицы b линейно независимы, а потому $\det b \neq 0$.

Определение 1.2. *Оператором перехода от базиса* $\mathbf{e}$ *к базису* $\widetilde{\mathbf{e}}$ *называется линейный оператор* $B : E \to E$*, такой что* $Be_k = \widetilde{e}_k$*,* $k = 1, \ldots, n$*.*

Оператор B переводит вектор $x = \sum_{k=1}^n \xi^k e_k$ в вектор

$$Bx = \sum_{k=1}^n \xi^k \widetilde{e}_k.$$

Изображающая матрица оператора B в паре базисов $\mathbf{e}$, $\mathbf{e}$ — это матрица b. Очевидно, оператор B является автоморфизмом.

Введем теперь "символические векторы" (столбцы)

$$\vec{\mathbf{e}} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}, \quad \vec{\widetilde{\mathbf{e}}} = \begin{pmatrix} \widetilde{e}_1 \\ \widetilde{e}_2 \\ \vdots \\ \widetilde{e}_n \end{pmatrix}.$$

Тогда соотношения (1.1) можно записать в матричной форме:

$$\vec{\widetilde{\mathbf{e}}} = b^t \vec{\mathbf{e}}. \tag{1.2}$$

Закон преобразования вида (1.2) (в подробной записи — (1.1)) называют *ковариантным законом.* У объектов, которые преобразуются по ковариантному закону, принято писать нижние значки: $\{e_k\}$. "Хозяйкой" законов преобразования является матрица перехода b. Ниже мы увидим, что координаты векторов преобразуются по другому закону.

1.2. **Преобразование координат векторов.** Пусть $x \in E$. Разложим вектор x в базисе $\mathbf{e}$: $x = \sum_{j=1}^n \xi^j e_j$. Теперь разложим тот же вектор x в базисе $\widetilde{\mathbf{e}}$: $x = \sum_{k=1}^n \widetilde{\xi}^k \widetilde{e}_k$. Используя

(1.1), получаем

$$x = \sum_{k=1}^{n} \widetilde{\xi}^k \Big(\sum_{j=1}^{n} \beta_k^j e_j \Big) = \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} \beta_k^j \widetilde{\xi}^k \Big) e_j.$$

Следовательно,

$$\xi^j = \sum_{k=1}^{n} \beta_k^j \widetilde{\xi}^k, \quad j = 1, \dots, n. \tag{1.3}$$

Мы получили закон преобразования координат одного и того же вектора в “старом” и “новом” базисах.

Удобно записать этот закон в матричной форме. Для этого введем векторы

$$\vec{x} = \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}, \quad \vec{\widetilde{x}} = \begin{pmatrix} \widetilde{\xi}^1 \\ \vdots \\ \widetilde{\xi}^n \end{pmatrix} \in K^n. \tag{1.4}$$

Тогда соотношения (1.3) можно записать в виде

$$\vec{x} = b\vec{\widetilde{x}}, \ \text{ или } \ \vec{\widetilde{x}} = b^{-1}\vec{x}. \tag{1.5}$$

Закон преобразования вида (1.5) (в подробной записи — (1.3)) называют *контравариантным законом.* У объектов, которые преобразуются по контравариантному закону, принято писать верхние значки: $\{\xi^k\}$.

1.3. **Преобразование изображающих матриц линейных операторов.** Для краткости будем обозначать пространство линейных операторов из E в E через $\Lambda(E)$ (прежнее обозначение $\Lambda(E, E)$).

Пусть $A \in \Lambda(E)$ и $a \in M^n$ — изображающая матрица оператора A в паре базисов $\mathbf{e}$, $\mathbf{e}$. Теперь мы будем называть *a изображающей матрицей оператора A в базисе* $\mathbf{e}$. Пусть $\widetilde{a} \in M^n$ — изображающая матрица оператора A в базисе $\widetilde{\mathbf{e}}$ (по-старому — в паре базисов $\widetilde{\mathbf{e}}$, $\widetilde{\mathbf{e}}$). Найдем связь между a и $\widetilde{a}$.

Пусть $x \in E$ и $Ax = y$. Пусть $\vec{x}$ и $\vec{y}$ — столбцы координат векторов x и y в базисе $\mathbf{e}$. Равенство $Ax = y$ перейдет в равенство $a\vec{x} = \vec{y}$.

Пусть $\vec{\widetilde{x}}$ и $\vec{\widetilde{y}}$ — столбцы координат векторов x и y в базисе $\widetilde{\mathbf{e}}$. Равенство $Ax = y$ перейдет в равенство $\widetilde{a}\vec{\widetilde{x}} = \vec{\widetilde{y}}$.

Домножая равенство $a\vec{x} = \vec{y}$ слева на b^{-1} и используя соотношения $\vec{x} = b\vec{\widetilde{x}}$ и $\vec{\widetilde{y}} = b^{-1}\vec{y}$ (см. (1.5)), получаем:

$$a\vec{x} = \vec{y} \quad \Rightarrow \quad b^{-1}a\vec{x} = b^{-1}\vec{y} \quad \Rightarrow \quad b^{-1}ab\vec{\widetilde{x}} = \vec{\widetilde{y}}.$$

Сопоставим это с равенством $\widetilde{a}\vec{\widetilde{x}} = \vec{\widetilde{y}}$. Следовательно,

$$\widetilde{a} = b^{-1}ab. \tag{1.6}$$

Вывод: *изображающая матрица $\widetilde{a}$ оператора A в базисе $\widetilde{\mathbf{e}}$ подобна изображающей матрице a этого оператора в базисе $\mathbf{e}$, а аффинитетом служит матрица b — матрица перехода от $\mathbf{e}$ к $\widetilde{\mathbf{e}}$.*

Напомним, что понятие подобия и свойства подобных матриц изучались в первом семестре; см. §8 в пособии [2].

Замечание. Обратно, если a — изображающая матрица оператора A в базисе $\mathbf{e}$, а матрица $\widetilde{a}$ подобна матрице a с аффинитетом b, то есть, $\widetilde{a} = b^{-1}ab$, то найдется такой базис $\widetilde{\mathbf{e}}$, в котором изображающая матрица оператора A есть $\widetilde{a}$. Новый базис можно построить по исходному с помощью формул (1.1) или (1.2).

§ 2. Определитель, след, характеристический многочлен и спектр линейных операторов

2.1. Определитель и след линейных операторов.

Определение 2.1. *Пусть $A \in \Lambda(E)$ и пусть a — изображающая матрица оператора A в каком-либо базисе $\mathbf{e}$. Определителем оператора A называется число, равное определителю матрицы a, то есть,*

$$\det A := \det a. \tag{2.1}$$

Определение корректно, поскольку при переходе к другому базису $\widetilde{\mathbf{e}}$ изображающая матрица меняется по закону (1.6): матрица $\widetilde{a}$ подобна матрице a, а у подобных матриц определители совпадают:

$$\det \widetilde{a} = \det a.$$

Свойства $\det A$

- Определитель тождественного оператора равен единице: $\det I = 1$.
  $\square$ Это свойство очевидно, поскольку изображающей матрицей оператора I в любом базисе $\mathbf{e}$ является единичная матрица. $\square$
- Пусть $A, C \in \Lambda(E)$. Тогда определитель композиции операторов AC равен произведению их определителей:

$$\det AC = \det A \cdot \det C.$$

  $\square$ Пусть a и c — изображающие матрицы операторов A и C в базисе $\mathbf{e}$. Тогда изображающей матрицей оператора AC в том же базисе является матрица ac. Остается вспомнить теорему об определителе произведения матриц (см. §6 в пособии [1]). Имеем:

$$\det AC = \det ac = \det a \cdot \det c = \det A \cdot \det C. \qquad \square$$

- Оператор $A \in \Lambda(E)$ является автоморфизмом тогда и только тогда, когда $\det A \neq 0$.
  $\square$ Мы знаем, что оператор $A \in \Lambda(E)$ является автоморфизмом тогда и только тогда, когда $\operatorname{rank} A = n$ (где $n = \dim E$). Поскольку $\operatorname{rank} A = \operatorname{rank} a$ (где a — изображающая матрица оператора A в каком-либо базисе), это равносильно тому, что $\operatorname{rank} a = n$. Последнее свойство равносильно соотношению $\det a \neq 0$. Остается учесть (2.1). $\square$

- Пусть $A \in \Lambda(E)$ — автоморфизм. Тогда

$$\det A^{-1} = \frac{1}{\det A}. \tag{2.2}$$

$\Box$ В силу предыдущего свойства, $\det A \neq 0$. Имеем:

$$1 = \det I = \det A^{-1}A = \det A^{-1} \cdot \det A. \tag{2.3}$$

Мы воспользовались равенством $A^{-1}A = I$ и свойством об определителе композиции операторов. Из (2.3) прямо вытекает (2.2). $\Box$

Определение 2.2. *Пусть $A \in \Lambda(E)$ и пусть a — изображающая матрица оператора A в каком-либо базисе* $\mathbf{e}$. *Следом оператора A называется число, равное следу матрицы a, то есть,*

$$\operatorname{Tr} A := \operatorname{Tr} a. \tag{2.4}$$

Определение корректно, поскольку при переходе к другому базису $\widetilde{\mathbf{e}}$ изображающая матрица меняется по закону (1.6): матрица $\widetilde{a}$ подобна матрице a, а у подобных матриц следы совпадают:

$$\operatorname{Tr} \widetilde{a} = \operatorname{Tr} a.$$

Свойства $\operatorname{Tr} A$

- След оператора линеен:

$$\operatorname{Tr}(\alpha A + \beta C) = \alpha \operatorname{Tr} A + \beta \operatorname{Tr} C \quad \forall A, C \in \Lambda(E), \quad \forall \alpha, \beta \in K.$$

- След произведения (композиции) операторов не зависит от порядка сомножителей:

$$\operatorname{Tr} AC = \operatorname{Tr} CA \quad \forall A, C \in \Lambda(E).$$

Докажите эти свойства самостоятельно, используя определение следа оператора (см. (2.4)) и аналогичные свойства для матриц (см. §2 в пособии [1]).

2.2. **Ориентация в конечномерном вещественном линейном пространстве.** Мы знакомы с понятием ориентации по курсу аналитической геометрии. В трехмерном пространстве (геометрических векторов) рассматриваются всевозможные упорядоченные тройки некомпланарных векторов (т.е., всевозможные базисы). На этом классе вводится отношение эквивалентности: одна тройка эквивалентна другой, если существует непрерывная деформация, переводящая первую тройку во вторую. Тогда класс всех базисов разбивается ровно на два класса эквивалентности. Выбор одного из этих классов и есть выбор ориентации в пространстве.

Сейчас мы обсудим понятие ориентации в произвольном конечномерном вещественном линейном пространстве E. Пусть $\dim E = n \geqslant 1$. Пусть $\mathbf{e} = \{e_1, \dots, e_n\}$ и $\widetilde{\mathbf{e}} = \{\widetilde{e}_1, \dots, \widetilde{e}_n\}$ — два базиса в E. Пусть $B : E \to E$ — оператор перехода от базиса $\mathbf{e}$ к $\widetilde{\mathbf{e}}$. По определению 1.2 имеем $Be_k = \widetilde{e}_k$, $k = 1, \dots, n$. Оператор B является автоморфизмом в E, а потому $\det B \neq 0$. Учитывая, что $\det B$ — вещественное число, получаем альтернативу: либо $\det B > 0$, либо $\det B < 0$.

Введем отношение эквивалентности на классе всех упорядоченных линейно независимых наборов из n векторов в E (т.е. на классе всех базисов в E).

Определение 2.3. *Говорят, что базис* $\widetilde{\mathbf{e}}$ *эквивалентен базису* $\mathbf{e}$*, если определитель оператора перехода от* $\mathbf{e}$ *к* $\widetilde{\mathbf{e}}$ *положителен*: $\det B > 0$.

Убедимся, что это действительно отношение эквивалентности.

1) *Базис* $\mathbf{e}$ *эквивалентен* $\mathbf{e}$, поскольку оператор перехода — тождественный и $\det I = 1 > 0$.

2) *Если* $\widetilde{\mathbf{e}}$ *эквивалентен* $\mathbf{e}$*, то* $\mathbf{e}$ *эквивалентен* $\widetilde{\mathbf{e}}$.

Действительно, если B — оператор перехода от $\mathbf{e}$ к $\widetilde{\mathbf{e}}$, то B^{-1} — оператор перехода от $\widetilde{\mathbf{e}}$ к $\mathbf{e}$. Из $\det B > 0$ следует, что

$$\det B^{-1} = (\det B)^{-1} > 0.$$

3) *Если $\widetilde{\mathbf{e}}$ эквивалентен $\mathbf{e}$, а $\widehat{\mathbf{e}}$ эквивалентен $\widetilde{\mathbf{e}}$, то $\widehat{\mathbf{e}}$ эквивалентен $\mathbf{e}$.*

Действительно, если B — оператор перехода от $\mathbf{e}$ к $\widetilde{\mathbf{e}}$, а $\widehat{B}$ — оператор перехода от $\widetilde{\mathbf{e}}$ к $\widehat{\mathbf{e}}$, то оператором перехода от $\mathbf{e}$ к $\widehat{\mathbf{e}}$ служит $\widehat{B}B$. Из $\det B > 0$ и $\det \widehat{B} > 0$ следует, что

$$\det \widehat{B}B = \det \widehat{B} \cdot \det B > 0.$$

Теорема 2.4. *Пусть E — конечномерное вещественное линейное пространство, $\dim E = n \geqslant 1$. Множество всех базисов в E разбивается ровно на два класса эквивалентности.*

Выбор одного из этих классов эквивалентности и есть выбор ориентации в пространстве E.

Доказательство. 1) Сначала докажем, что классов эквивалентности не меньше двух. Фиксируем базис $\mathbf{e} = \{e_1, \dots, e_n\}$ и построим базис $\mathbf{e}_* = \{-e_1, e_2, \dots, e_n\}$, то есть, $e_{*1} = -e_1$, $e_{*k} = e_k$ при $k = 2, \dots, n$. Пусть B_* — оператор перехода от $\mathbf{e}$ к $\mathbf{e}_*$, а b_* — соответствующая матрица перехода. Тогда

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

Имеем: $\det B_* = \det b_* = -1 < 0$, а потому базис $\mathbf{e}_*$ не эквивалентен базису $\mathbf{e}$. Значит, они принадлежат разным классам эквивалентности. Стало быть, классов эквивалентности не меньше двух.

2) Теперь покажем, что классов эквивалентности ровно два. Фиксируем базис $\mathbf{e}$, построим $\mathbf{e}_*$ как в пункте 1. Будем

“сортировать” остальные базисы. Пусть $\widetilde{\mathbf{e}}$ — произвольный базис и B — оператор перехода от $\mathbf{e}$ к $\widetilde{\mathbf{e}}$.

Если $\det B > 0$, то $\widetilde{\mathbf{e}}$ эквивалентен $\mathbf{e}$.

Если же $\det B < 0$, то $\widetilde{\mathbf{e}}$ эквивалентен $\mathbf{e}_*$. Действительно,

$$Be_k = \widetilde{e}_k, \quad B_* e_k = e_{*k} \quad \Rightarrow \quad B_* B^{-1} \widetilde{e}_k = e_{*k}.$$

Значит, оператором перехода от $\widetilde{\mathbf{e}}$ к $\mathbf{e}_*$ служит $B_* B^{-1}$. Имеем:

$$\det B_* B^{-1} = \frac{\det B_*}{\det B} = -\frac{1}{\det B} > 0.$$

Мы показали, что каждый базис $\widetilde{\mathbf{e}}$ попадает либо в первый класс эквивалентности (содержащий $\mathbf{e}$), либо во второй (содержащий $\mathbf{e}_*$). Значит, классов эквивалентности ровно два. □

Замечание. В комплексном конечномерном линейном пространстве нет понятия ориентации. Наличие ориентации — чисто вещественный эффект.

2.3. **Характеристический многочлен и спектр линейного оператора.**

Определение 2.5. *Пусть E — конечномерное линейное пространство над полем K, где $K = \mathbb{R}$ или $K = \mathbb{C}$. Пусть $A \in \Lambda(E)$. Характеристическим многочленом оператора A называется многочлен*

$$d_A(\lambda) = \det(A - \lambda I).$$

Равносильное определение: характеристическим многочленом оператора A называется характеристический многочлен его изображающей матрицы a (в каком-либо базисе). При переходе к другому базису изображающая матрица меняется по закону подобия (см. (1.6)), а характеристический многочлен является инвариантом подобия. Поэтому такое определение корректно.

Вспомним разложение характеристического многочлена (см. §8 в пособии [2]):

$$d_A(\lambda) = \delta_n \lambda^n + \delta_{n-1}\lambda^{n-1} + \dots + \delta_1\lambda + \delta_0, \qquad (2.5)$$

где $\delta_n = (-1)^n$, $\delta_{n-1} = (-1)^{n-1}\operatorname{Tr} A$, $\delta_0 = \det A$.

Замечание. В вещественном случае необходимо сделать комментарий. Оператор $A-\lambda I$ имеет смысл лишь при $\lambda \in \mathbb{R}$. Тогда многочлен $d_A(\lambda)$ исходно определен при $\lambda \in \mathbb{R}$. Но после того, как мы записали его в виде (2.5), нет проблем распространить этот многочлен на все значения $\lambda \in \mathbb{C}$. В случае второго определения можно сразу считать, что $\lambda \in \mathbb{C}$.

Определение 2.6. *Собственными значениями (собственными числами) оператора A называются корни характеристического многочлена $d_A(\lambda)$. Совокупность собственных значений называется спектром оператора A.*

Если $K = \mathbb{R}$, то $d_A(\lambda)$ — многочлен с вещественными коэффициентами, но его корни могут оказаться комплексными числами с $\operatorname{Re}\lambda \neq 0$.

Есть два (основных) способа нумерации собственных значений. *Первый способ*: пусть $\lambda_1, \dots, \lambda_n$ — все собственные значения, причем кратные корни многочлена $d_A(\lambda)$ повторяются столько раз, какова их кратность. Тогда разложение характеристического многочлена на множители имеет вид

$$d_A(\lambda) = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n).$$

Второй способ: пусть $\mu_1, \dots, \mu_p$ — все различные собственные значения, а $\sigma_1, \dots, \sigma_p$ — их кратности. Разумеется, $\sigma_1 + \dots + \sigma_p = n$. Тогда разложение характеристического многочлена на множители имеет вид

$$d_A(\lambda) = (-1)^n(\lambda - \mu_1)^{\sigma_1}\cdots(\lambda - \mu_p)^{\sigma_p}.$$

Определение 2.7. *Алгебраической кратностью собственного значения μ_j оператора A называется кратность корня*

μ_j характеристического многочлена $d_A(\lambda)$, то есть число σ_j.

Определение 2.8. *Если $\sigma_j = 1$, то собственное значение μ_j оператора A называется простым. Если $\sigma_j > 1$, то собственное значение μ_j оператора A называется кратным.*

В следующем параграфе будет введено понятие другой кратности — геометрической кратности собственного значения.

§ 3. Собственные элементы и собственные подпространства линейных операторов

3.1. **Собственные элементы. Геометрическая кратность собственных значений.**

Предложение 3.1. *Пусть E — n-мерное линейное пространство над полем K, где $K = \mathbb{R}$ или $K = \mathbb{C}$, причем $n \geqslant 1$. Пусть $A \in \Lambda(E)$ и μ — собственное значение оператора A. В случае $K = \mathbb{R}$ дополнительно предположим, что $\mu \in \mathbb{R}$. Тогда существует элемент $f \in E$, $f \neq \mathbf{0}$, такой что $Af = \mu f$.*

Доказательство. Число μ является корнем характеристического многочлена, а потому

$$d_A(\mu) = \det(A - \mu I) = 0.$$

Следовательно, $\operatorname{rank}(A - \mu I) < n$. В силу соотношения

$$\operatorname{rank}(A - \mu I) + \dim \operatorname{Ker}(A - \mu I) = \dim E = n,$$

получаем $\dim \operatorname{Ker}(A - \mu I) > 0$. Это означает, что

$$\operatorname{Ker}(A - \mu I) \neq \{\mathbf{0}\}.$$

Тогда найдется элемент $\mathbf{0} \neq f \in \operatorname{Ker}(A - \mu I)$. Имеем:

$$(A - \mu I)f = \mathbf{0} \;\Leftrightarrow\; Af = \mu f. \qquad \square$$

Определение 3.2. *В условиях предложения* 3.1 *элемент* $f \neq \mathbf{0}$, *такой что* $Af = \mu f$, *называется собственным элементом (или собственным вектором) оператора* A, *отвечающим собственному значению* μ.

Замечание. 1°. Если E — вещественное пространство, а собственное значение $\mu \notin \mathbb{R}$, то никакого собственного вектора, отвечающего μ, нет. Причина в том, что умножение на комплексное число $\mu \notin \mathbb{R}$ в пространстве E не имеет смысла.
2°. Если f — собственный вектор оператора A, отвечающий μ, то любой ненулевой вектор вида cf (где $c \in K$) тоже является собственным вектором, отвечающим μ.

Определение 3.3. *В условиях предложения* 3.1 *собственным подпространством оператора* A, *отвечающим собственному значению* μ, *называется подпространство* $F_\mu := \mathrm{Ker}(A - \mu I)$.

Собственное подпространство F_μ состоит из всех собственных векторов, отвечающих μ, и из нулевого вектора.

Определение 3.4. *В условиях предложения* 3.1 *геометрической кратностью* τ_μ *собственного значения* μ *называется размерность собственного подпространства*:

$$\tau_\mu := \dim F_\mu.$$

Геометрическая кратность показывает, сколько линейно независимых собственных векторов отвечает данному собственному значению.

Примеры

- Пусть $E = \mathbb{R}^2$, и пусть $A_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ — оператор поворота на угол $\varphi \in [0, \pi]$ (против часовой стрелки). Рассмотрим стандартный базис

$$\mathbf{e} = \{\vec{e}_1, \vec{e}_2\}, \quad \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Изображающей матрицей оператора A_φ в базисе $\mathbf{e}$ служит матрица поворота

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

Вычислим характеристический многочлен

$$d(\lambda) = \begin{vmatrix} \cos\varphi - \lambda & -\sin\varphi \\ \sin\varphi & \cos\varphi - \lambda \end{vmatrix} = (\cos\varphi - \lambda)^2 + \sin^2\varphi.$$

Найдем собственные значения λ_1, λ_2 — корни многочлена $d(\lambda)$:

$$\lambda_1 = \cos\varphi + i\sin\varphi = e^{i\varphi}, \quad \lambda_2 = \cos\varphi - i\sin\varphi = e^{-i\varphi}.$$

Если $0 < \varphi < \pi$, то $\lambda_1, \lambda_2 \notin \mathbb{R}$, а потому оператор A_φ не имеет никаких собственных векторов. Это ясно и из геометрического смысла: если бы нашелся ненулевой вектор f такой, что $A_\varphi f = \mu f$, то вектор $A_\varphi f$ был бы коллинеарен f (в терминах аналитической геометрии). Но таких векторов не существует, поскольку оператор A_φ поворачивает векторы на угол φ.

Если $\varphi = 0$, то $A_0 = I$ — тождественный оператор, он имеет кратное собственное значение $\lambda_1 = \lambda_2 = 1$, соответствующее собственное подпространство совпадает с $\mathbb{R}^2$.

Если $\varphi = \pi$, то $A_\pi = -I$, он имеет кратное собственное значение $\lambda_1 = \lambda_2 = -1$, соответствующее собственное подпространство совпадает с $\mathbb{R}^2$.

- **Упражнение.** Пусть $E = \mathbb{C}^2$ и $B_\varphi : \mathbb{C}^2 \to \mathbb{C}^2$ — оператор умножения на матрицу a_φ, где $0 < \varphi < \pi$. Собственные значения оператора B_φ — это собственные значения матрицы a_φ. Они уже найдены: $\lambda_1 = e^{i\varphi}$, $\lambda_2 = e^{-i\varphi}$. Найдите собственные векторы.

Теорема 3.5. *Пусть E — n-мерное пространство над полем K, где $K = \mathbb{R}$ или $K = \mathbb{C}$, причем $n \geqslant 1$. Пусть*

$A \in \Lambda(E)$ и μ — собственное значение оператора A. В случае $K = \mathbb{R}$ предполагаем, что $\mu \in \mathbb{R}$. Тогда геометрическая кратность τ_μ собственного значения μ не превосходит его алгебраической кратности σ_μ:

$$\tau_\mu \leqslant \sigma_\mu.$$

Доказательство. Рассмотрим собственное подпространство $F_\mu = \operatorname{Ker}(A - \mu I)$. Геометрическую кратность собственного значения μ обозначим для краткости $\tau := \tau_\mu = \dim F_\mu$, а алгебраическую кратность обозначим $\sigma := \sigma_\mu$.

Выберем какой-либо базис $f_1, \ldots, f_\tau$ в F_μ. Дополним линейно независимый набор $f_1, \ldots, f_\tau$ до базиса в пространстве E:

$$f_1, \ldots, f_\tau; z_{\tau+1}, \ldots, z_n. \tag{3.1}$$

Пусть $a \in M^n$ — изображающая матрица оператора A в этом базисе. Векторы f_j — собственные, а потому $Af_j = \mu f_j$, $j = 1, \ldots, \tau$. Запишем разложение векторов Az_k в базисе (3.1):

$$Az_k = \sum_{j=1}^{\tau} \alpha_k^j f_j + \sum_{l=\tau+1}^{n} \alpha_k^l z_l, \quad k = \tau + 1, \ldots, n.$$

Тогда изображающая матрица a имеет вид

$$a = \begin{pmatrix} \mu & 0 & \ldots & 0 & \alpha_{\tau+1}^1 & \ldots & \alpha_n^1 \\ 0 & \mu & \ldots & 0 & \alpha_{\tau+1}^2 & \ldots & \alpha_n^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mu & \alpha_{\tau+1}^\tau & \ldots & \alpha_n^\tau \\ 0 & 0 & \ldots & 0 & \alpha_{\tau+1}^{\tau+1} & \ldots & \alpha_n^{\tau+1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \alpha_{\tau+1}^n & \ldots & \alpha_n^n \end{pmatrix}.$$

Обозначим через $b \in M^{n-\tau}$ правый нижний блок этой матрицы. Вычислим характеристический многочлен:

$$d_a(\lambda) = \begin{vmatrix} \mu-\lambda & 0 & \dots & 0 & \alpha_{\tau+1}^1 & \dots & \alpha_n^1 \\ 0 & \mu-\lambda & \dots & 0 & \alpha_{\tau+1}^2 & \dots & \alpha_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mu-\lambda & \alpha_{\tau+1}^\tau & \dots & \alpha_n^\tau \\ 0 & 0 & \dots & 0 & \alpha_{\tau+1}^{\tau+1}-\lambda & \dots & \alpha_n^{\tau+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \alpha_{\tau+1}^n & \dots & \alpha_n^n-\lambda \end{vmatrix}.$$

Получаем

$$d_a(\lambda) = (\mu-\lambda)^\tau d_b(\lambda).$$

Следовательно, число μ является корнем характеристического многочлена кратности σ не меньшей, чем τ. Тем самым, доказано, что $\tau \leqslant \sigma$. □

Поясним, что $\sigma = \tau$, если μ не является корнем многочлена $d_b(\lambda)$, и $\sigma > \tau$, если μ является корнем многочлена $d_b(\lambda)$.

В силу предложения 3.1 в условиях теоремы выполнено $\tau_\mu \geqslant 1$. С другой стороны, очевидно, что $\sigma_\mu \leqslant n$. Мы получаем следствие.

Следствие 3.6. 1°. *В условиях теоремы* 3.5 *выполнены неравенства*

$$1 \leqslant \tau_\mu \leqslant \sigma_\mu \leqslant n.$$

2°. *Если μ — простое собственное значение, то $\tau_\mu = \sigma_\mu = 1$.*

Примеры

- Пусть $A = \mathbf{0}_{E\to E}$ — нулевой оператор в n-мерном пространстве E. Его изображающая матрица в любом базисе — это нулевая матрица, характеристический многочлен равен

$$d(\lambda) = \begin{vmatrix} -\lambda & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & -\lambda \end{vmatrix} = (-\lambda)^n.$$

Собственное значение одно: $\mu = 0$; его алгебраическая кратность максимальна: $\sigma = n$. Любой ненулевой вектор $f \in E$ является собственным, так как $Af = 0 \cdot f$. Очевидно, собственное подпространство совпадает с E, а геометрическая кратность равна $\tau = n$. В этом примере $\tau = \sigma = n$.

- Пусть $A = I : E \to E$ — тождественный оператор в n-мерном пространстве E. Его изображающая матрица в любом базисе — это единичная матрица, характеристический многочлен равен

$$d(\lambda) = \begin{vmatrix} 1-\lambda & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 1-\lambda \end{vmatrix} = (1-\lambda)^n.$$

Собственное значение одно: $\mu = 1$; его алгебраическая кратность максимальна: $\sigma = n$. Любой ненулевой вектор $f \in E$ является собственным, так как $Af = 1 \cdot f$. Очевидно, собственное подпространство совпадает с E, а геометрическая кратность равна $\tau = n$. В этом примере $\tau = \sigma = n$.

- Пусть $E = \Omega_{n-1}$ — пространство многочленов степени не выше $n-1$ с вещественными коэффициентами (это n-мерное пространство над полем $\mathbb{R}$). Пусть $D : E \to E$ — оператор дифференцирования. Рассмотрим базис $e_1, \dots, e_n$, где

$$e_1(t) = 1,\ e_2(t) = t,\ e_3(t) = \frac{t^2}{2},\ \dots,\ e_n(t) = \frac{t^{n-1}}{(n-1)!}.$$

Изображающая матрица оператора D в этом базисе имеет вид

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

Характеристический многочлен равен

$$d(\lambda) = \begin{vmatrix} -\lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & -\lambda & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -\lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & -\lambda \end{vmatrix} = (-\lambda)^n.$$

Собственное значение одно: $\mu = 0$; его алгебраическая кратность максимальна: $\sigma = n$.

Собственный элемент — это ненулевой многочлен $P \in \Omega_{n-1}$ такой, что $DP = \mathbf{0}$, то есть, $P'(t) = 0$ тождественно по t. Следовательно, $P(t) = \mathrm{Const}$. Таким образом, собственное подпространство F состоит из констант, то есть, $F = \Omega_0$, а геометрическая кратность равна единице $\tau = 1$.

Если $n > 1$ в этом примере, то геометрическая кратность меньше алгебраической: $1 = \tau < \sigma = n$.

3.2. **Прямая сумма собственных подпространств линейного оператора.** Пусть E — n-мерное линейное пространство над полем K, где $K = \mathbb{R}$ или $K = \mathbb{C}$, причем $n \geqslant 1$. Пусть $A \in \Lambda(E)$. В комплексном случае рассмотрим все различные собственные значения оператора A, а в вещественном случае рассмотрим все различные *вещественные* собственные значения оператора A. Обозначим набор таких собственных значений

$$\mu_1, \ldots, \mu_p.$$

(В вещественном случае этот набор может быть пустым; см. пример про оператор поворота в $\mathbb{R}^2$.)

Пусть $F_1, \ldots, F_p$ — соответствующие собственные подпространства, т.е. $F_k = \operatorname{Ker}(A - \mu_k I)$.

Теорема 3.7. *При сделанных предположениях линейная сумма* $F_A := F_1 + \cdots + F_p$ *является прямой суммой*:

$$F_A = F_1 \dotplus F_2 \dotplus \cdots \dotplus F_p.$$

Доказательство. Требуется проверить, что для любого вектора $x \in F_A$ представление в виде $x = x_1 + \cdots + x_p$, где $x_j \in F_j$ $(j = 1, \ldots, p)$ единственно.

Предположим, что

$$x = x_1 + \cdots + x_p, \quad x = \widetilde{x}_1 + \cdots + \widetilde{x}_p, \quad x_j, \widetilde{x}_j \in F_j,\ j = 1, \ldots, p.$$

Требуется доказать, что $x_j = \widetilde{x}_j$, $j = 1, \ldots, p$.

Обозначим $f_j := x_j - \widetilde{x}_j$, $j = 1, \ldots, p$. Имеем

$$f_1 + \cdots + f_p = \mathbf{0}, \quad f_j \in F_j, \quad j = 1, \ldots, p. \tag{3.2}$$

Теперь требуется доказать, что $f_j = \mathbf{0}$, $j = 1, \ldots, p$.

Применим к равенству (3.2) оператор $(A - \mu_1 I)$ и учтем, что $Af_j = \mu_j f_j$, $j = 1, \ldots, p$. Получаем:

$$(\mu_2 - \mu_1) f_2 + (\mu_3 - \mu_1) f_3 + \cdots + (\mu_p - \mu_1) f_p = \mathbf{0}. \tag{3.3}$$

Далее, применяя к равенству (3.3) оператор $(A - \mu_2 I)$, приходим к

$$(\mu_3 - \mu_1)(\mu_3 - \mu_2) f_3 + \cdots + (\mu_p - \mu_1)(\mu_p - \mu_2) f_p = \mathbf{0}. \tag{3.4}$$

Каждый раз число слагаемых уменьшается на единицу. Будем продолжать эту процедуру. Предпоследнее равенство имеет вид

$$(\mu_{p-1} - \mu_1) \cdots (\mu_{p-1} - \mu_{p-2}) f_{p-1} + (\mu_p - \mu_1) \cdots (\mu_p - \mu_{p-2}) f_p = \mathbf{0}. \tag{3.5}$$

Наконец, последнее равенство содержит всего одно слагаемое:

$$(\mu_p - \mu_1) \cdots (\mu_p - \mu_{p-1}) f_p = \mathbf{0}. \tag{3.6}$$

Множитель $(\mu_p - \mu_1) \cdots (\mu_p - \mu_{p-1})$ отличен от нуля, поскольку все числа $\mu_1, \ldots, \mu_p$ различны. Поэтому из (3.6) вытекает, что $f_p = \mathbf{0}$. Подставляя $f_p = \mathbf{0}$ в (3.5), получаем $f_{p-1} = \mathbf{0}$. И так далее. В итоге мы убеждаемся, что $f_1 = \cdots = f_p = \mathbf{0}$. $\square$

3.3. **Критерий существования собственного базиса. Диагонализуемость линейных операторов.**

Теорема 3.8. 1°. *Пусть E — конечномерное линейное пространство над полем $\mathbb{C}$, причем $n = \dim E \geqslant 1$. Пусть $A \in \Lambda(E)$. В E существует базис из собственных элементов оператора A тогда и только тогда, когда для всех различных собственных значений оператора A совпадают геометрические и алгебраические кратности.*
2°. *Пусть E — конечномерное линейное пространство над полем $\mathbb{R}$, причем $n = \dim E \geqslant 1$. Пусть $A \in \Lambda(E)$. В E существует базис из собственных элементов оператора A тогда и только тогда, когда все различные собственные значения оператора A вещественны и для них совпадают геометрические и алгебраические кратности.*

Доказательство. 1°. В условиях первого пункта теоремы, пусть $\mu_1, \ldots, \mu_p$ — все различные собственные значения оператора A. Пусть σ_j и τ_j — алгебраическая и геометрическая кратности собственного значения μ_j. Отметим, что $\sigma_1 + \cdots + \sigma_p = n$.

Рассмотрим собственные подпространства

$$F_j = \operatorname{Ker}(A - \mu_j I), \quad \tau_j = \dim F_j.$$

В силу теоремы 3.7 линейная сумма $F_A := F_1 + \cdots + F_p$ — это прямая сумма:

$$F_A = F_1 \dot{+} \cdots \dot{+} F_p, \tag{3.7}$$

а тогда

$$\dim F_A = \tau_1 + \cdots + \tau_p.$$

Отметим, что подпространство (3.7) состоит из всевозможных собственных элементов оператора A (и из нулевого вектора $\mathbf{0}$). Поэтому существование в E базиса из собственных элементов (кратко — собственного базиса) оператора A равносильно тому, что $F_A = E$. Последнее равенство равносильно тому, что $\dim F_A = \dim E = n$. Таким образом, существование собственного базиса равносильно соотношению

$$\tau_1 + \dots + \tau_p = n. \tag{3.8}$$

В силу теоремы 3.5 выполнены неравенства $\tau_j \leqslant \sigma_j$ при $j = 1, \dots, p$, а потому

$$\tau_1 + \dots + \tau_p \leqslant \sigma_1 + \dots + \sigma_p = n. \tag{3.9}$$

Достаточность. Если $\tau_j = \sigma_j$ при всех $j = 1, \dots, p$, то неравенство (3.9) превратится в равенство, а тогда выполнено (3.8), что равносильно существованию собственного базиса.

Необходимость. Пусть существует собственный базис. Тогда выполнено (3.8). Это возможно только в том случае, когда $\tau_j = \sigma_j$ при всех $j = 1, \dots, p$. Действительно, если хотя бы для одного j выполнено $\tau_j < \sigma_j$, то неравенство (3.9) будет строгим: $\tau_1 + \dots + \tau_p < \sigma_1 + \dots + \sigma_p = n$.

2°. В условиях второго пункта надо рассмотреть все различные *вещественные* собственные значения $\mu_1, \dots, \mu_p$ оператора A и провести аналогичные рассуждения. При этом, если у оператора A существует хотя бы одно невещественное собственное значение, то окажется, что $\sigma_1 + \dots + \sigma_p < n$. Тогда заведомо $\dim F_A = \tau_1 + \dots + \tau_p < n$, и собственного базиса не существует. Поэтому условие, чтобы все собственные значения были вещественными, необходимо. Если оно выполнено, то по-прежнему $\sigma_1 + \dots + \sigma_p = n$ и надо просто повторить рассуждения из первого пункта. □

Определение 3.9. *Оператор $A \in \Lambda(E)$ называется диагонализуемым, если в E существует базис, в котором изображающая матрица оператора A диагональна.*

Предложение 3.10. *Пусть E — конечномерное линейное пространство над полем $\mathbb{C}$ или $\mathbb{R}$, причем $n = \dim E \geqslant 1$. В E существует собственный базис оператора $A \in \Lambda(E)$ тогда и только тогда, когда оператор A диагонализуем.*

Доказательство. Утверждение практически очевидно.

Необходимость. Пусть в пространстве E существует базис $\mathbf{f} = \{f_1, \ldots, f_n\}$ из собственных векторов оператора A. Тогда $Af_j = \lambda_j f_j$, где λ_j — соответствующие собственные значения. Очевидно, изображающая матрица a оператора A в базисе $\mathbf{f}$ имеет вид

$$a = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}. \tag{3.10}$$

Матрица a диагональна, а значит, оператор A диагонализуем.

Достаточность. Пусть оператор диагонализуем, то есть, существует базис $\mathbf{f} = \{f_1, \ldots, f_n\}$, в котором изображающая матрица a оператора A диагональна, то есть a имеет вид (3.10). Поскольку столбцы изображающей матрицы состоят из координат векторов Af_k, $k = 1, \ldots, n$, получаем

$$Af_k = \lambda_k f_k, \quad k = 1, \ldots, n.$$

Таким образом, числа λ_k, стоящие на диагонали матрицы a, — это собственные значения, а базисные векторы f_k — это собственные векторы оператора A. Значит, базис $\mathbf{f}$ является собственным базисом оператора A. $\square$

Из теоремы 3.8 и предложения 3.10 вытекает следствие.

Следствие 3.11. 1°. *Пусть E — конечномерное линейное пространство над полем $\mathbb{C}$, причем $n = \dim E \geqslant 1$. Оператор $A \in \Lambda(E)$ диагонализуем тогда и только тогда, когда для всех различных собственных значений оператора A совпадают геометрические и алгебраические кратности.*
2°. *Пусть E — конечномерное линейное пространство над полем $\mathbb{R}$, причем $n = \dim E \geqslant 1$. Оператор $A \in \Lambda(E)$ диагонализуем тогда и только тогда, когда все различные собственные значения оператора A вещественны и для них совпадают геометрические и алгебраические кратности.*

Вспомним определение диагонализуемой матрицы.

Определение 3.12. *Матрица a называется диагонализуемой, если она подобна диагональной матрице $\widetilde{a}$, то есть существует неособая матрица b, такая что $\widetilde{a} = b^{-1}ab$.*

Теперь мы обсудим, как связаны свойства диагонализуемости оператора и диагонализуемости его изображающей матрицы.

Предложение 3.13. 1°. *Пусть E — конечномерное линейное пространство над полем $\mathbb{C}$, $n = \dim E \geqslant 1$. Оператор $A \in \Lambda(E)$ диагонализуем тогда и только тогда, когда его изображающая матрица a (в каком-либо базисе) диагонализуема.*
2°. *Пусть E — конечномерное линейное пространство над полем $\mathbb{R}$, причем $n = \dim E \geqslant 1$. Оператор $A \in \Lambda(E)$ диагонализуем тогда и только тогда, когда его изображающая матрица a (в каком-либо базисе) диагонализуема в классе вещественных матриц.*

Доказательство. 1°. Пусть $\mathbf{e} = \{e_1, \ldots, e_n\}$ — некоторый базис в E. Пусть a — изображающая матрица оператора A в базисе $\mathbf{e}$.

Необходимость. Пусть оператор A диагонализуем, то есть, в E существует базис $\mathbf{f} = \{f_1, \ldots, f_n\}$, в котором изображающая матрица $\widetilde{a}$ оператора A диагональна:

$$\widetilde{a} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}. \tag{3.11}$$

Обозначим через b матрицу перехода от исходного базиса $\mathbf{e}$ к базису $\mathbf{f}$. Вспомним закон преобразования изображающих матриц: $\widetilde{a} = b^{-1}ab$. Мы убедились, что матрица a диагонализуема.

Достаточность. Пусть матрица a диагонализуема, то есть, существует неособая матрица b такая, что $\widetilde{a} = b^{-1}ab$. Тогда матрица $\widetilde{a}$ является изображающей матрицей оператора A в базисе $\mathbf{f}$, который связан с базисом $\mathbf{e}$ матрицей перехода b (см. ковариантный закон преобразования). Таким образом, оператор A диагонализуем.

Утверждение 2° устанавливается аналогично. $\square$

Как следствие, мы установим критерий диагонализуемости матриц. Следующее утверждение формулируется в матричных (а не операторных) терминах.

Предложение 3.14. 1°. *Матрица $a \in M^n$ с комплексными элементами диагонализуема тогда и только тогда, когда существуют числа $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ и линейно независимые векторы $\vec{f}_1, \ldots, \vec{f}_n \in \mathbb{C}^n$, такие что $a\vec{f}_j = \lambda_j \vec{f}_j$, $j = 1, \ldots, n$. При этом выполнено $\widetilde{a} = b^{-1}ab$, где диагональная матрица $\widetilde{a}$ имеет вид* (3.11), *а матрица b составлена из столбцов $\vec{f}_1, \ldots, \vec{f}_n$.*
2°. *Матрица $a \in M^n$ с вещественными элементами диагонализуема в классе вещественных матриц тогда и только*

тогда, когда существуют числа $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ *и линейно независимые векторы* $\vec{f}_1, \ldots, \vec{f}_n \in \mathbb{R}^n$, *такие что* $a\vec{f}_j = \lambda_j \vec{f}_j$ *при всех* $j = 1, \ldots, n$. *При этом выполнено* $\widetilde{a} = b^{-1}ab$, *где диагональная матрица* $\widetilde{a}$ *имеет вид* (3.11), *а матрица* b *составлена из столбцов* $\vec{f}_1, \ldots, \vec{f}_n$.

Доказательство. 1°. Рассмотрим оператор

$$A = \widehat{a} : \mathbb{C}^n \to \mathbb{C}^n,$$

т.е. оператор умножения на матрицу a. Обозначим через $\mathbf{e} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ стандартный базис в $\mathbb{C}^n$. Изображающей матрицей оператора A в базисе $\mathbf{e}$ является сама матрица a.

В силу предложения 3.13 матрица a диагонализуема тогда и только тогда, когда оператор A диагонализуем. Как следует из предложения 3.10, диагонализуемость оператора A равносильна существованию собственного базиса $\mathbf{f} = \{\vec{f}_1, \ldots, \vec{f}_n\}$ в $\mathbb{C}^n$. При этом $a\vec{f}_j = \lambda_j \vec{f}_j$, $j = 1, \ldots, n$, где λ_j — соответствующие собственные значения. Очевидно, матрица перехода b от стандартного базиса $\mathbf{e}$ к базису $\mathbf{f}$ состоит из столбцов $\vec{f}_1, \ldots, \vec{f}_n$. Изображающая матрица $\widetilde{a}$ оператора A в собственном базисе $\mathbf{f}$ диагональна (имеет вид (3.11)), и выполнено $\widetilde{a} = b^{-1}ab$.

2°. Второе утверждение доказывается аналогично с помощью рассмотрения оператора $A = \widehat{a} : \mathbb{R}^n \to \mathbb{R}^n$. □

§ 4. Функции от операторов

Материал этого параграфа излагается обзорно (многие факты сообщаются без доказательства или доказательство поручается читателю). Мы обсудим различные подходы к определению функций от линейных операторов в конечномерном линейном пространстве E над полем K, где $K = \mathbb{C}$ или $K = \mathbb{R}$.

4.1. **Степени A^m, $m \in \mathbb{Z}_+$.**

Определение 4.1. *Пусть $A \in \Lambda(E)$ и пусть $m \in \mathbb{Z}_+$. Определение оператора A^m индуктивно:*

$$A^0 := I,\ A^1 := A,\ A^2 := A \cdot A,\ A^3 := A^2 \cdot A,\ \dots,$$
$$A^m := A^{m-1} \cdot A,\ \dots$$

Проверьте самостоятельно, что A^m и A^k коммутируют, причем справедливо равенство

$$A^m A^k = A^k A^m = A^{m+k}, \quad m, k \in \mathbb{Z}_+.$$

(Для проверки надо использовать ассоциативность умножения операторов.)

Если $a \in M^n$ — изображающая матрица оператора A в некотором базисе $\mathbf{e}$, то изображающей матрицей оператора A^m в том же базисе служит a^m. Это свойство следует из того, что композиции (произведению) операторов отвечает произведение их изображающих матриц.

4.2. **Многочлены от A.**

Определение 4.2. *Пусть $P(t)$ — многочлен вида*

$$P(t) = p_0 + p_1 t + p_2 t^2 + \dots + p_m t^m,$$

где $p_j \in \mathbb{C}$ в случае $K = \mathbb{C}$ и $p_j \in \mathbb{R}$ в случае $K = \mathbb{R}$. Пусть $A \in \Lambda(E)$. Оператор $P(A)$ определен формулой

$$P(A) = p_0 I + p_1 A + p_2 A^2 + \dots + p_m A^m.$$

Проверьте, что два многочлена $P(A)$ и $Q(A)$ коммутируют и справедливо равенство

$$P(A)Q(A) = Q(A)P(A) = (PQ)(A),$$

где многочлен PQ — это произведение $(PQ)(t) = P(t)Q(t)$.

Если $a \in M^n$ — изображающая матрица оператора A в некотором базисе $\mathbf{e}$, то изображающей матрицей оператора $P(A)$ в том же базисе служит $P(a)$.

Предложение 4.3. *Пусть* $A \in \Lambda(E)$ *и пусть*

$$d_A(t) = \delta_0 + \delta_1 t + \cdots + \delta_n t^n$$

есть характеристический многочлен оператора A. *Тогда выполнено тождество Кэли*

$$d_A(A) = \delta_0 I + \delta_1 A + \cdots + \delta_n A^n = \mathbf{0}_{E \to E}.$$

Доказательство. Пусть $a \in M^n$ — изображающая матрица оператора A в некотором базисе $\mathbf{e}$. Тогда характеристический многочлен оператора A совпадает с характеристическим многочленом матрицы a: $d_A(t) = d_a(t)$.

Для любого многочлена P изображающей матрицей оператора $P(A)$ является матрица $P(a)$, поэтому для оператора $d_A(A) = d_a(A)$ изображающей матрицей является $d_a(a)$. В силу тождества Кэли матрица $d_a(a)$ равна нулевой матрице (см. §8 в пособии [2]). Следовательно, $d_A(A)$ равен нулевому оператору. □

4.3. **Степени A^{-m}, $m \in \mathbb{N}$.**

Определение 4.4. *Пусть* $A \in \Lambda(E)$, *причем* $\det A \neq 0$. *Тогда существует обратный оператор* A^{-1}. *Пусть* $m \in \mathbb{N}$. *По определению,*

$$A^{-m} := (A^{-1})^m.$$

Проверьте, что в случае $\det A \neq 0$ любые целые степени A^k и A^j коммутируют:

$$A^k A^j = A^j A^k = A^{k+j}, \quad k, j \in \mathbb{Z}.$$

4.4. **Дробно-рациональные функции от операторов.** В этом пункте ограничимся (более простым) случаем пространства E над полем $\mathbb{C}$. Пусть задана дробно-рациональная функция

$$R(t) = \frac{P(t)}{Q(t)},$$

где $P(t)$ и $Q(t)$ — некоторые многочлены, причем $\deg P = m$, $\deg Q = k \geqslant 1$.

Выделим целую часть рациональной дроби, то есть, представим дробь $R(t)$ в виде

$$R(t) = S(t) + \frac{P_1(t)}{Q(t)},$$

где $S(t)$, $P_1(t)$ — некоторые многочлены, причем $S(t) = 0$, если $m < k$, и $\deg S = m - k$, если $m \geqslant k$, а $\deg P_1 \leqslant k - 1$.

Пусть $t_1, \ldots, t_q$ — все различные корни многочлена $Q(t)$, а $j_1, \ldots, j_q$ — их кратности ($j_1 + \cdots + j_q = k$). Тогда правильную дробь $\frac{P_1(t)}{Q(t)}$ можно разложить в сумму простых дробей, и мы получим представление

$$R(t) = S(t) + \frac{\alpha_1^{(1)}}{t - t_1} + \frac{\alpha_2^{(1)}}{(t - t_1)^2} + \cdots + \frac{\alpha_{j_1}^{(1)}}{(t - t_1)^{j_1}} + \cdots + \\ + \frac{\alpha_1^{(q)}}{t - t_q} + \frac{\alpha_2^{(q)}}{(t - t_q)^2} + \cdots + \frac{\alpha_{j_q}^{(q)}}{(t - t_q)^{j_q}}. \tag{4.1}$$

Пусть $A \in \Lambda(E)$. Предположим, что числа $t_1, \ldots, t_q$ не являются собственными значениями оператора A. Тогда корректно определены операторы $(A - t_s I)^{-m}$, $s = 1, \ldots, q$, при $m \in \mathbb{N}$.

Определение 4.5. *Пусть дробно-рациональная функция $R(t)$ допускает представление* (4.1). *Пусть $A \in \Lambda(E)$. Предположим, что числа $t_1, \ldots, t_q$ не являются собственными значениями оператора A. По определению,*

$$R(A) = S(A) + \\ + \alpha_1^{(1)}(A - t_1 I)^{-1} + \alpha_2^{(1)}(A - t_1 I)^{-2} + \cdots + \alpha_{j_1}^{(1)}(A - t_1 I)^{-j_1} + \\ + \cdots + \alpha_1^{(q)}(A - t_q I)^{-1} + \alpha_2^{(q)}(A - t_q I)^{-2} + \cdots + \alpha_{j_q}^{(q)}(A - t_q I)^{-j_q}.$$

Если $a \in M^n$ — изображающая матрица оператора A в некотором базисе $\mathbf{e}$, то изображающей матрицей оператора $R(A)$ в том же базисе служит $R(a)$.

Замечание. Пусть E — пространство над полем $\mathbb{R}$, и многочлены $P(t)$ и $Q(t)$ имеют вещественные коэффициенты. Тогда данное выше определение годится только в том случае, когда все корни $t_1, \ldots, t_q$ многочлена $Q(t)$ вещественны. Если это не так, то надо модифицировать определение, используя разложение дроби $R(t)$ с участием простых дробей третьего и четвертого типов, имеющих вид $\frac{\alpha+\beta t}{(t^2+\gamma t+\delta)^m}$ (где $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, а дробь в знаменателе имеет отрицательный дискриминант).

4.5. **Аналитические функции от операторов.** Пусть функция $f(t)$ допускает разложение в сходящийся степенной ряд (ряд Тейлора)

$$f(t) = \sum_{k=0}^{\infty} f_k t^k, \qquad f_k = \frac{f^{(k)}(0)}{k!}. \tag{4.2}$$

Определение 4.6. *Пусть $f(t)$ раскладывается в ряд* (4.2)*. Пусть $A \in \Lambda(E)$. По определению,*

$$f(A) = \sum_{k=0}^{\infty} f_k A^k.$$

Это определение требует объяснения, как понимается сходимость ряда из операторов. Один из способов объяснения дается через переход к изображающим матрицам. Пусть a — изображающая матрица оператора A в каком-либо базисе $\mathbf{e}$. Эквивалентное определение: оператор $f(A)$ — это оператор с изображающей матрицей

$$f(a) = \sum_{k=0}^{\infty} f_k a^k$$

в базисе $\mathbf{e}$. Сходимость матричного ряда понимается поэлементно: должны сходиться числовые ряды

$$[f(a)]_{lj} = \sum_{k=0}^{\infty} f_k [a^k]_{lj}, \quad l, j = 1, \ldots, n.$$

Пример. Экспонента от оператора задана рядом

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

4.6. **Функции от диагонализуемых операторов.** Пусть оператор $A \in \Lambda(E)$ диагонализуем. Тогда существует базис $\mathbf{f} = \{f_1, \ldots, f_n\}$ из собственных векторов оператора A. При этом $Af_j = \lambda_j f_j$, $j = 1, \ldots, n$, где λ_j — соответствующие соственные значения оператора A. Изображающая матрица a оператора A в базисе $\mathbf{f}$ диагональна:

$$a = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Пусть $\Phi(t)$ — некоторая функция (в вещественном случае это функция от вещественного аргумента с вещественными значениями, в комплексном случае это функция от комплексного аргумента с комплексными значениями). Предполагается, что числа $\lambda_1, \ldots, \lambda_n$ принадлежат области определения функции $\Phi(t)$. Вспомним определение функции от диагональной матрицы:

$$\Phi(a) = \begin{pmatrix} \Phi(\lambda_1) & 0 & \ldots & 0 \\ 0 & \Phi(\lambda_2) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \Phi(\lambda_n) \end{pmatrix}. \qquad (4.3)$$

Определение 4.7. *При сделанных предположениях оператор $\Phi(A)$ — это оператор в E, изображающая матрица которого в базисе* $\mathbf{e}$ *есть $\Phi(a)$.*

Собственными значениями оператора $\Phi(A)$ являются числа $\Phi(\lambda_1), \ldots, \Phi(\lambda_n)$, а собственные векторы прежние:

$$\Phi(A)f_j = \Phi(\lambda_j)f_j, \quad j = 1, \ldots, n.$$

Замечания. 1) Аналитическую функцию $f(A)$ можно определить для любого оператора A, но аналитичность f — очень сильное ограничение. Зато для диагонализуемых операторов можно определить любую функцию $\Phi(A)$ (с подходящей областью определения функции Φ).

2) Разные определения согласованы между собой. Если A диагонализуем, то определение рациональной функции $R(A)$ через разложение (4.1) и через диагонализацию дают один и тот же оператор $R(A)$. Аналогично определение аналитической функции $f(A)$ через степенной ряд и через диагонализацию дают один и тот же оператор $f(A)$.

4.7. **Применение к системам линейных дифференциальных уравнений первого порядка.** Функции от матриц применяются, например, в теории дифференциальных уравнений. Рассмотрим систему линейных дифференциальных уравнений первого порядка:

$$\begin{cases} \dfrac{dx_1(t)}{dt} = \alpha_{11}x_1(t) + \alpha_{12}x_2(t) + \cdots + \alpha_{1n}x_n(t), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dx_n(t)}{dt} = \alpha_{n1}x_1(t) + \alpha_{n2}x_2(t) + \cdots + \alpha_{nn}x_n(t). \end{cases} \tag{4.4}$$

Здесь $x_1(t), \ldots, x_n(t)$ — функции от $t \in \mathbb{R}$ с комплексными значениями, а $\alpha_{lj} \in \mathbb{C}$ — постоянные коэффициенты. Ищем решение системы, удовлетворяющее начальным условиям

$$x_1(0) = x_{01},\ x_2(0) = x_{02},\ \ldots,\ x_n(0) = x_{0n}. \tag{4.5}$$

Здесь начальные данные $x_{0j} \in \mathbb{C}$ заданы. В теории дифференциальных уравнений задачу (4.4), (4.5) называют *задачей Коши.*

Удобно записать эту систему в матричном виде

$$\frac{d\vec{x}(t)}{dt} = a\vec{x}(t), \quad \vec{x}(0) = \vec{x}_0, \tag{4.6}$$

где $\vec{x}(t)$ — столбец неизвестных функций

$$\vec{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix},$$

a — матрица коэффициентов:

$$a = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{nn} \end{pmatrix},$$

а $\vec{x}_0$ — столбец начальных данных:

$$\vec{x}_0 = \begin{pmatrix} x_{01} \\ \vdots \\ x_{0n} \end{pmatrix} \in \mathbb{C}^n.$$

В курсе дифференциальных уравнений устанавливается теорема о существовании и единственности решения задачи Коши. Мы убедимся, что это решение можно записать с помощью матричной экспоненты.

Предложение 4.8. *Решение задачи* (4.6) *дается формулой*

$$\vec{x}(t) = e^{ta}\vec{x}_0, \quad t \in \mathbb{R},$$

где e^{ta} *— экспонента от матрицы* ta.

Доказательство. Запишем матричную экспоненту в виде суммы ряда

$$e^{ta} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k a^k.$$

Продифференцируем по t:

$$\frac{de^{ta}}{dt} = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{dt^k}{dt} a^k = \sum_{k=1}^{\infty} \frac{1}{k!} k t^{k-1} a^k.$$

(В теории степенных функциональных рядов устанавливается законность перемены порядка дифференцирования и суммирования.) Преобразуем правую часть:

$$\frac{de^{ta}}{dt} = a \sum_{k=1}^{\infty} \frac{1}{(k-1)!} t^{k-1} a^{k-1} = a \sum_{m=0}^{\infty} \frac{1}{m!} t^m a^m = a e^{ta}.$$

Применяя получившееся равенство к вектору $\vec{x_0}$, получаем

$$\frac{d(e^{ta}\vec{x}_0)}{dt} = a(e^{ta}\vec{x}_0).$$

Поскольку $e^{0\cdot a}$ — это единичная матрица, то

$$(e^{ta}\vec{x}_0)|_{t=0} = \vec{x}_0.$$

Тем самым вектор-функция $\vec{x}(t) = e^{ta}\vec{x}_0$ является решением задачи Коши (4.6). $\square$

Рассмотрим случай, когда матрица a диагонализуема. В силу предложения 3.14 найдутся числа $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ и базис $\{\vec{f}_1, \ldots, \vec{f}_n\}$ в $\mathbb{C}^n$, такие что $a\vec{f}_j = \lambda_j \vec{f}_j$, $j = 1, \ldots, n$. При этом выполнено $\widetilde{a} = b^{-1}ab$, где матрица $\widetilde{a}$ диагональна и имеет вид

$$\widetilde{a} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

а матрица b составлена из столбцов $\vec{f}_1, \ldots, \vec{f}_n$.

Выполним подстановку (замену неизвестной вектор-функции)

$$\vec{x}(t) = b\vec{y}(t) \;\Leftrightarrow\; \vec{y}(t) = b^{-1}\vec{x}(t).$$

Домножая уравнение $\frac{d\vec{x}(t)}{dt} = a\vec{x}(t)$ на матрицу b^{-1}, получаем уравнение для $\vec{y}(t)$:

$$\frac{d\vec{y}(t)}{dt} = (b^{-1}ab)\vec{y}(t) \;\Leftrightarrow\; \frac{d\vec{y}(t)}{dt} = \widetilde{a}\vec{y}(t).$$

Домножая начальное условие $\vec{x}(0) = \vec{x}_0$ на матрицу b^{-1} и вводя обозначение $\vec{y}_0 := b^{-1}\vec{x}_0$, получаем

$$\vec{y}(0) = \vec{y}_0 \equiv \begin{pmatrix} y_{01} \\ \vdots \\ y_{0n} \end{pmatrix}.$$

Мы приходим к задаче Коши

$$\frac{d\vec{y}(t)}{dt} = \widetilde{a}\vec{y}(t), \quad \vec{y}(0) = \vec{y}_0.$$

За счет того, что матрица $\widetilde{a}$ диагональна, эта задача расщепляется на отдельные задачи для каждой скалярной функции $y_j(t)$:

$$\frac{dy_j(t)}{dt} = \lambda_j y_j(t), \quad y_j(0) = y_{0j}, \quad j = 1, \ldots, n.$$

Скалярные задачи Коши легко решаются:

$$y_j(t) = e^{\lambda_j t} y_{0j}, \quad j = 1, \ldots, n.$$

Замечание. Если матрица a не диагонализуема, то такого расщепления не происходит. В общем случае можно вычислять матричную экспоненту e^{ta}, используя приведение матрицы a к жордановой форме. (Жорданову форму мы будем изучать в конце курса алгебры.) Тогда у задачи Коши существуют неэкспоненциальные решения.

Список литературы

[1] Бирман М. Ш., Суслина Т. А., *Линейная алгебра, выпуск* 1. *Матрицы. Определители.* Учебно-методическое пособие. — СПб.: СПбГУ, 1999. – 41 с.

[2] Бирман М. Ш., Суслина Т. А., Фаддеев М. М., *Линейная алгебра, выпуск* 2. *Матрицы. Системы линейных алгебраических уравнений.* Учебно-методическое пособие. — СПб.: СПбГУ, 1999. – 49 с.

[3] Суслина Т. А., *Линейная алгебра. II семестр, выпуск* 1. Учебно-методическое пособие. — СПб.: СПбГУ, 2021. – 76 с.