версия 23.03.2019

Просьба сообщить о всех замеченных опечатках, описках, возникших вопросах, замечаниях и т.д. на адрес vvfil@mech.math.msu.su

Глава 6.

Линейные операторы. Диагонализируемые операторы.

§1. Преобразование матрицы оператора при переходе к новому базису и характеристический многочлен.

Линейное отображение $L \rightarrow L$ векторного пространства L над полем K в себя называется *линейным оператором*.

Пусть система

$$\begin{cases} y^1 = a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n \\ y^2 = a_1^2 x^1 + a_2^2 x^2 + \dots + a_n^2 x^n \\ \dots \\ y^n = a_1^n x^1 + a_2^n x^2 + \dots + a_n^n x^n \end{cases}$$

описывает рассматриваемый линейный оператор в базисе $e_1, e_2, \dots, e_n$. В матричной форме эта система записывается уравнением

$$\begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix}$$

или, более кратко, $v = Au$, где u — столбец из координат аргумента

$$\begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix}$$

в базисе $e_1, e_2, \dots, e_n$, v — столбец из координат образа

$$\begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix}$$

в базисе $e_1, e_2, \dots, e_n$,

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix}$$

— матрица оператора в базисе $e_1, e_2, \dots, e_n$.

Пусть

$$f_1 = c_1^1 e_1 + c_1^2 e_2 + \dots + c_1^n e_n,$$

$$f_2 = c_2^1 e_1 + c_2^2 e_2 + \dots + c_2^n e_n,$$

$\dots$,

$$f_n = c_n^1 e_1 + c_n^2 e_2 + \dots + c_n^n e_n$$

— другой базис пространства L .

Тогда преобразование координат векторов при переходе от базиса $e_1, e_2, \dots, e_n$ к базису $f_1, f_2, \dots, f_n$ задается системой:

$$\begin{cases} x^1 = c_1^1 y^1 + c_2^1 y^2 + \dots + c_n^1 y^n \\ x^2 = c_1^2 y^1 + c_2^2 y^2 + \dots + c_n^2 y^n \\ \dots \\ x^n = c_1^n y^1 + c_2^n y^2 + \dots + c_n^n y^n \end{cases}$$

или в матричной форме уравнением

$$\begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix} = \begin{pmatrix} c_1^1 & c_2^1 & \dots & c_n^1 \\ c_1^2 & c_2^2 & \dots & c_n^2 \\ \dots & \dots & \dots & \dots \\ c_1^n & c_2^n & \dots & c_n^n \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix}$$

или, короче, $u = Cu_1$, где u — столбец из координат вектора

$$\begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^n \end{pmatrix}$$

в базисе $e_1, e_2, \dots, e_n$, u_1 — столбец из координат того же вектора

$$\begin{pmatrix} y^1 \\ y^2 \\ \dots \\ y^n \end{pmatrix},$$

но в базисе $f_1, f_2, \dots, f_n$.

Пусть B — матрица нашего оператора в базисе $f_1, f_2, \dots, f_n$, v_1 — столбец из координат образа в базисе $f_1, f_2, \dots, f_n$, то есть, $v = Cv_1$. Тогда $v_1 = Bu_1$,

Делая подстановки в уравнение $v = Au$, получаем

$$Cv_1 = ACu_1,$$

$$v_1 = C^{-1}ACu_1.$$

Но так как матрица оператора определена однозначно, то

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

Пусть E — единичная матрица, функция $\chi(\lambda) = \det(A - \lambda E)$ является многочленом от аргумента λ . Посмотрим, что происходит с ним при переходе от базиса $e_1, e_2, \dots, e_n$ к базису $f_1, f_2, \dots, f_n$:

$$\begin{aligned} \det(B - \lambda E) &= \det(C^{-1}AC - \lambda C^{-1}EC) = \det C^{-1}(A - \lambda E)C = \\ &= \det C^{-1} \det(A - \lambda E) \det C = \det C^{-1} \det C \det(A - \lambda E) = \\ &= \det C^{-1}C \det(A - \lambda E) = \det E \det(A - \lambda E) = \det(A - \lambda E) = \chi(\lambda). \end{aligned}$$

То есть, наш многочлен не меняется при переходе к новому базису. Он называется *характеристическим многочленом* рассматриваемого оператора.

Обозначим через ε тождественное отображение $L \rightarrow L: \varepsilon(u) = u$. Заметим, что матрица $A - \lambda E$ есть ни что иное как матрица оператора $\varphi - \lambda\varepsilon$.

§2. Инвариантные подпространства.

Подпространство $L_1 \subseteq L$ пространства L называется *инвариантным* подпространством оператора φ , если $\varphi(L_1) \subseteq L_1$.

Лемма 2.1. Пусть φ и ψ линейные операторы из L в L и $\varphi - \psi = \lambda\varepsilon$ при некотором $\lambda \in K$. Тогда у операторов φ и ψ одни и те же инвариантные подпространства.

Доказательство. I. Пусть L_1 — инвариантное подпространство оператора φ .

Пусть $u \in L_1$. Тогда $\psi(u) = \varphi(u) - \lambda u$. Так как $\varphi(u) \in L_1$ и $\lambda u \in L_1$, то $\psi(u) \in L_1$. То есть, подпространство L_1 является инвариантным подпространством оператора ψ .

II То, что всякое инвариантное подпространство оператора ψ является инвариантным подпространством оператора φ вытекает из доказанного в силу симметричности положения φ и ψ в условиях леммы: $\psi - \varphi = (-\lambda)\varepsilon$. Лемма доказана.

Вид матрицы оператора, когда часть векторов базиса образует базис инвариантного подпространства.

Пусть $e_1, e_2, \dots, e_n$ — базис рассматриваемого n -мерного векторного пространства L , векторы $e_1, e_2, \dots, e_k$ — базис инвариантного подпространства $L_1 \subseteq L$, система

$$\begin{cases} y^1 = a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n \\ y^2 = a_1^2 y^1 + a_2^2 x^2 + \dots + a_n^2 x^n \\ \dots \\ y^n = a_1^n x^1 + a_2^n x^2 + \dots + a_n^n x^n \end{cases}$$

описывает рассматриваемый линейный оператор в базисе $e_1, e_2, \dots, e_n$. i -тый столбец матрицы

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix}$$

состоит из координат образа i -того базисного вектора e_i .

Образы векторов $e_1, e_2, \dots, e_k$ принадлежат инвариантному подпространству L_1 и поэтому в их разложении по базису $e_1, e_2, \dots, e_n$ коэффициенты при векторах $e_{k+1}, e_{k+2}, \dots, e_n$ равны нулю. Это означает, что в матрице

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix}$$

элементы в первых k столбцах, начиная с $(k+1)$ -ого, равны нулю, то есть, матрица имеет вид:

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{pmatrix}.$$

Вид матрицы оператора, когда пространство представлено в виде прямой суммы инвариантных подпространств. I.

Сохраняя предположения предыдущего пункта, предположим, что векторы $e_{k+1}, e_{k+2}, \dots, e_n$ составляют базис инвариантного подпространства $L_2 \subseteq L$. Тогда образы векторов $e_{k+1}, e_{k+2}, \dots,$

e_n принадлежат инвариантному подпространству L_2 и поэтому в их разложении по базису $e_1, e_2, \dots, e_n$ коэффициенты при векторах $e_1, e_2, \dots, e_k$ равны нулю. Это означает, что в матрице

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{pmatrix}$$

элементы в последних $(n - k)$ столбцах до k -ого, равны нулю, то есть, матрица имеет вид:

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & 0 & \dots & 0 \\ a_1^2 & a_2^2 & \dots & a_k^2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{pmatrix}$$

Наш оператор как бы распался на два. Один с матрицей

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k \end{pmatrix}$$

действует в пространстве L_1 . Второй с матрицей

$$\begin{pmatrix} a_{k+1}^{k+1} & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n \end{pmatrix}$$

действует в пространстве L_2 .

Вид матрицы оператора, когда пространство представлено в виде прямой суммы инвариантных подпространств. II.

Пусть пространство L представлено в виде прямой суммы $L = L_1 \oplus L_2 \oplus \dots \oplus L_k$ своих инвариантных подпространств $L_1, L_2, \dots, L_k$. Если в каждом из этих подпространств L_i мы выберем базис A_i , то объединение $A = A_1 \cup A_2 \cup \dots \cup A_k$ этих базисов есть базис пространства L . В соответствии с рассуждениями предыдущего пункта матрица рассматриваемого оператора в базисе A устроена следующим образом. Вдоль диагонали стоят последовательно матрицы ограничений нашего оператора на инвариантные подпространства $L_1, L_2, \dots, L_k$. На всех остальных позициях стоят нули. Верно и обратное: если матрица оператора имеет указанный вид, то линейные оболочки соответствующих групп базисных векторов являются инвариантными подпространствами, прямая сумма которых равна L .

Вид характеристического многочлена оператора, когда часть векторов базиса образует базис инвариантного подпространства.

Мы начнем с замечания, относящегося к полилинейным кососимметрическим функциям. Функция $f : L \times L \times \dots \times L \rightarrow K$ является полилинейной кососимметрической функцией, если она линейна по каждому аргументу и $f(u_{\sigma(1)}, u_{\sigma(2)}, \dots, u_{\sigma(k)}) = (-1)^\sigma f(u_1, u_2, \dots, u_k)$ для любой перестановки σ множества $\{1, 2, \dots, k\}$. Примером такой функции является определитель. Он является полилинейной кососимметрической функцией от набора своих строк и от набора своих столбцов.

Отметим сразу одно свойство полилинейной кососимметрической функции. Если среди аргументов $f(u_1, u_2, \dots, u_k)$ есть два совпадающих, то $f(u_1, u_2, \dots, u_k) = 0$, ибо меняя их между собой местами, мы должны умножить значение функции на -1 . С другой стороны, при этой операции ничего не меняется, поэтому значение функции сохраняется. Одновременно это возможно тогда и только тогда, когда $f(u_1, u_2, \dots, u_k) = 0$.

Лемма 2.2. Пусть $f : L \times L \times \cdots \times L \rightarrow K$ — полилинейная кососимметрическая функция и $f(e_{i_1}, e_{i_2}, \dots, e_{i_k}) = 0$ для любой выборки базисных векторов с $i_1 < i_2 < \cdots < i_k$. Тогда функция f тождественно равна нулю.

Доказательство. I. Заметим прежде всего, что $f(e_{i_1}, e_{i_2}, \dots, e_{i_k}) = 0$ всегда (то есть, без дополнительного предположения $i_1 < i_2 < \cdots < i_k$), ибо, если среди векторов $e_{i_1}, e_{i_2}, \dots, e_{i_k}$ какой-то встречается дважды, то это следует из замечания, предшествующего формулировке леммы. Если же все векторы $e_{i_1}, e_{i_2}, \dots, e_{i_k}$ различны, то их можно переставить так, чтобы условие $i_1 < i_2 < \cdots < i_k$ выполнялось. При этом значение функции либо не меняется вовсе, либо умножается на -1 . Но так как при выполнении условия $i_1 < i_2 < \cdots < i_k$ значение функции равно нулю, то оно равно нулю и при первоначальном порядке аргументов.

II. Доказательство равенства $f(u_1, u_2, \dots, u_k) = 0$ в общем случае проведем индукцией по числу аргументов k .

При $k = 1$ про косую симметрию можно забыть и речь идет о линейном функционале. Если он равен нулю на каждом из базисных векторов, то он равен нулю тождественно и мы имеем требуемое.

Пусть утверждение доказано для $k - 1 \geq 1$. Докажем его для k . Подставим вместо u_1 любой из векторов e_i , $i = 1, 2, \dots, n$ и рассмотрим функцию $g(u_2, \dots, u_k) = f(e_i, u_2, \dots, u_k)$. Как легко видеть, функция g удовлетворяет всем нашим ограничениям и по предположению индукции $g \equiv 0$. Теперь для любых

$$u_1 = x^1 e_1 + x^2 e_2 + \cdots + x^n e_n,$$

$u_2, \dots, u_k \in L$ имеем:

$$\begin{aligned} f(u_1, u_2, \dots, u_k) &= x^1 f(e_1, u_2, \dots, u_k) + \\ &+ x^2 f(e_2, u_2, \dots, u_k) + \cdots + x^n f(e_n, u_2, \dots, u_k) = \\ &= 0 + 0 + \cdots + 0 = 0. \end{aligned}$$

Лемма доказана.

Следствие. Пусть $f, g : L \times L \times \dots \times L \rightarrow K$ — две полилинейные кососимметрические функции и $f(e_{i_1}, e_{i_2}, \dots, e_{i_k}) = g(e_{i_1}, e_{i_2}, \dots, e_{i_k}) = 0$ для любой выборки базисных векторов с $i_1 < i_2 < \dots < i_k$. Тогда функции f и g равны: $f \equiv g$.

Для получения следствия надо применить лемму 2.2 к разности $h = f - g$.

Применим теперь полученный результат к доказательству формулы

$$\begin{aligned} & \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{vmatrix} = \\ & = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k \end{vmatrix} \cdot \begin{vmatrix} a_{k+1}^{k+1} & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n \end{vmatrix}. \end{aligned}$$

Рассмотрим кососимметрические функции, заданные на пространстве $\mathbb{R}^k$ формулами

$$f(u_1, u_2, \dots, u_k) = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{vmatrix}$$

и

$$g(u_1, u_2, \dots, u_k) = \begin{vmatrix} a_1^1 & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k \end{vmatrix} \cdot \begin{vmatrix} a_{k+1}^{k+1} & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n \end{vmatrix},$$

где $u_1 = \{a_1^1, a_1^2, \dots, a_1^k\}$, $u_2 = \{a_2^1, a_2^2, \dots, a_2^k\}$, $\dots$, $u_k = \{a_k^1, a_k^2, \dots, a_k^k\} \in \mathbb{R}^k$ — аргументы наших функций, а числа a_i^j при $i = k+1, k+2, \dots, n$, $j = 1, 2, \dots, n$ фиксированы.

По следствию леммы 2.2 достаточно проверить совпадение значений этих функций для аргументов $u_1 = e_1$, $u_2 = e_2$, $\dots$, $u_k = e_k$, где $e_1, e_2, \dots, e_k$ — стандартный базис пространства $\mathbb{R}^k$. В этом случае

$$f(e_1, e_2, \dots, e_k) = \begin{vmatrix} 1 & 0 & \dots & 0 & a_{k+1}^1 & \dots & a_n^1 \\ 0 & 1 & \dots & 0 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{vmatrix}$$

и

$$g(e_1, e_2, \dots, e_k) = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{vmatrix} \begin{vmatrix} a_{k+1}^{k+1} & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n \end{vmatrix} =$$

$$= \begin{vmatrix} a_{k+1}^{k+1} & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n \end{vmatrix}$$

и совпадение рассматриваемых определителей следует из последовательного разложения первого определителя по первым k столбцам.

Пусть теперь $e_1, e_2, \dots, e_n$ — базис рассматриваемого n -мерного векторного пространства L , в котором действует линейный оператор φ , векторы $e_1, e_2, \dots, e_k$ — базис инвариантного относительно φ подпространства $L_1 \subseteq L$. Как мы уже знаем в таком базисе

матрица оператора φ имеет вид:

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_k^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{pmatrix}.$$

Ввиду доказанной выше формулы характеристический многочлен оператора можно написать в виде

$$\begin{aligned} \chi(\lambda) &= \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_k^1 & a_{k+1}^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 - \lambda & \dots & a_k^2 & a_{k+1}^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k - \lambda & a_{k+1}^k & \dots & a_n^k \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n - \lambda \end{vmatrix} = \\ &= \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 - \lambda & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k - \lambda \end{vmatrix} \cdot \begin{vmatrix} a_{k+1}^{k+1} - \lambda & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} - \lambda & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n - \lambda \end{vmatrix} = \\ & \chi_1(\lambda) \cdot f(\lambda), \end{aligned}$$

где

$$\chi_1(\lambda) = \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 - \lambda & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k - \lambda \end{vmatrix}$$

— характеристический многочлен ограничения оператора φ на подпространство L_1 и

$$f(\lambda) = \begin{vmatrix} a_{k+1}^{k+1} - \lambda & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} - \lambda & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n - \lambda \end{vmatrix}.$$

То есть, характеристический многочлен $\chi_1(\lambda)$ ограничения оператора φ на подпространство L_1 является множителем в характеристическом многочлене $\chi(\lambda)$ оператора φ .

Уточним сделанное наблюдение в случае, когда пространство L представлено в виде прямой суммы $L = L_1 \oplus L_2$ двух своих инвариантных подпространств.

Возьмем такой базис $e_1, e_2, \dots, e_n$ пространства L , что векторы $e_1, e_2, \dots, e_k$ составляют базис подпространства L_1 , а векторы $e_{k+1}, e_{k+2}, \dots, e_n$ составляют базис инвариантного подпространства L_2 . Как мы видели, в этом базисе матрица оператора имеет вид:

$$\begin{pmatrix} a_1^1 & a_2^1 & \dots & a_k^1 & 0 & \dots & 0 \\ a_1^2 & a_2^2 & \dots & a_k^2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n \end{pmatrix}.$$

Соответственно, для характеристического многочлена имеем:

$$\begin{aligned} \chi(\lambda) &= \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_k^1 & 0 & \dots & 0 \\ a_1^2 & a_2^2 - \lambda & \dots & a_k^2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k - \lambda & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{k+1}^{k+1} - \lambda & \dots & a_n^{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k+1}^n & \dots & a_n^n - \lambda \end{vmatrix} = \\ &= \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 - \lambda & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k - \lambda \end{vmatrix} \cdot \begin{vmatrix} a_{k+1}^{k+1} - \lambda & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} - \lambda & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n - \lambda \end{vmatrix} = \\ &= \chi_1(\lambda) \cdot \chi_2(\lambda), \end{aligned}$$

где

$$\chi_1(\lambda) = \begin{vmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_k^1 \\ a_1^2 & a_2^2 - \lambda & \dots & a_k^2 \\ \dots & \dots & \dots & \dots \\ a_1^k & a_2^k & \dots & a_k^k - \lambda \end{vmatrix}$$

— характеристический многочлен ограничения оператора φ на подпространство L_1

$$\chi_2(\lambda) = \begin{vmatrix} a_{k+1}^{k+1} - \lambda & a_{k+2}^{k+1} & \dots & a_n^{k+1} \\ a_{k+1}^{k+2} & a_{k+2}^{k+2} - \lambda & \dots & a_n^{k+2} \\ \dots & \dots & \dots & \dots \\ a_{k+1}^n & a_{k+2}^n & \dots & a_n^n - \lambda \end{vmatrix}$$

— характеристический многочлен ограничения оператора φ на подпространство L_2 . То есть, в этом случае *характеристический многочлен $\chi(\lambda)$ оператора φ распадается в произведение характеристических многочленов $\chi_1(\lambda)$ и $\chi_2(\lambda)$ ограничений оператора φ на подпространства L_1 и L_2 , соответственно.*

Пусть пространство L представлено в виде прямой суммы $L = L_1 \oplus L_2 \oplus \dots \oplus L_k$ своих инвариантных подпространств $L_1, L_2, \dots, L_k$. Для удобства ссылок на предыдущее рассуждение заметим, что

$$L = L_1 \oplus (L_2 \oplus (L_3 \oplus \dots \oplus (L_{k-1} \oplus L_k) \dots)).$$

Как мы видели, если в каждом из этих подпространств L_i мы выберем базис A_i , то объединение $A = A_1 \cup A_2 \cup \dots \cup A_k$ этих базисов есть базис пространства L , а матрица рассматриваемого оператора в базисе A устроена следующим образом. Вдоль диагонали стоят последовательно матрицы ограничений нашего оператора на инвариантные подпространства $L_1, L_2, \dots, L_k$. На всех остальных позициях стоят нули. Повторяя $(k-2)$ раза предыдущее рассуждение, получаем что, *характеристический многочлен $\chi(\lambda)$ оператора φ распадается в произведение характеристических многочленов $\chi_1(\lambda), \chi_2(\lambda), \dots, \chi_k(\lambda)$ ограничений оператора φ на подпространства $L_1, L_2, \dots, L_k$, соответственно.*

§3. Одномерные и двумерные инвариантные подпространства и собственные векторы линейного оператора.

Вектор $u \in L$ называется *собственным вектором* оператора $\varphi : L \rightarrow L$, если найдется такое число $\lambda \in K$, что $\varphi(u) = \lambda u$. Нулевой вектор $\vec{0}$ удовлетворяет этому условию при любом λ . В связи с этим, когда речь идет о поиске собственных векторов, дополнительно предполагают, не оговаривая это специально, что речь идет о ненулевых собственных векторах. Но говоря о пространствах собственных векторов, нулевой вектор в них включают.

Существование вектора u , для которого $\varphi(u) = \lambda u$ равносильно тому, что оператор $\varphi - \lambda E$ имеет ненулевое ядро. В свою очередь это имеет место тогда и только тогда, когда определитель матрицы $A - \lambda E$ оператора $\varphi - \lambda E$ равен нулю, то есть, тогда и только тогда, когда $\chi(\lambda) = 0$, то есть, тогда и только тогда, когда λ является корнем характеристического многочлена оператора φ . Корни характеристического многочлена оператора φ называются также *собственными значениями* оператора φ .

Если $\varphi(u) = \lambda u$, то вектор u называется *собственным вектором*, соответствующим собственному значению λ , а λ называется *собственным значением*, соответствующим собственному вектору u .

Лемма 3.1. Пусть L_1 — одномерное инвариантное подпространство. Тогда любой вектор из L_1 является собственным.

Доказательство. Пусть u — ненулевой вектор подпространства L_1 . В силу одномерности пространства L_1 вектор u составляет его базис. Поэтому образ v вектора u получается умножением вектора u на некоторое число, что и дает требуемое. Лемма доказана.

Лемма 3.2. Пусть для оператора $\varphi : L \rightarrow L$ и $\lambda \in K$ $M_\lambda = \{u : u \in L, \varphi(u) = \lambda u\}$. Тогда множество M_λ является инвариантным подпространством.

Доказательство. Множество M_λ совпадает с ядром оператора $\varphi - \lambda E$ и доказываемое вытекает из леммы 2.1. Лемма доказана.

Лемма 3.3. Пусть в обозначениях леммы 3.2 числа $\lambda_1, \lambda_2, \dots, \lambda_i \in K$ различны и $u_k \in M_{\lambda_k} \setminus \{\vec{0}\}$ для $k = 1, 2, \dots, i$. Тогда

векторы $u_1, u_2, \dots, u_i$ линейно независимы.

Доказательство поведем индукцией по i .

При $i = 1$ мы имеем один ненулевой вектор и он составляет линейно независимую систему.

Пусть утверждение доказано для $(i - 1)$. Докажем его для i .

Рассмотрим линейную комбинацию наших векторов, равную нулю:

$$\mu_1 u_1 + \mu_2 u_2 + \dots + \mu_i u_i = \vec{0}.$$

Тогда

$$\begin{aligned} \varphi(\mu_1 u_1 + \mu_2 u_2 + \dots + \mu_i u_i) &= \varphi(\vec{0}), \\ \mu_1 \varphi(u_1) + \mu_2 \varphi(u_2) + \dots + \mu_i \varphi(u_i) &= \vec{0}, \\ \mu_1 \lambda_1 u_1 + \mu_2 \lambda_2 u_2 + \dots + \mu_i \lambda_i u_i &= \vec{0}. \end{aligned}$$

Вычитая из последнего равенства первое, умноженное на λ_i , получаем:

$$\mu_1(\lambda_1 - \lambda_i)u_1 + \mu_2(\lambda_2 - \lambda_i)u_2 + \dots + \mu_{i-1}(\lambda_{i-1} - \lambda_i)u_{i-1} = \vec{0}.$$

Так как числа $\lambda_1, \lambda_2, \dots, \lambda_i \in K$ различны, то каждый из векторов $v_k = (\lambda_k - \lambda_i)u_k$, $k = 1, 2, \dots, i - 1$ является ненулевым вектором подпространства M_{λ_k} . При этом

$$\mu_1 v_1 + \mu_2 v_2 + \dots + \mu_{i-1} v_{i-1} = \vec{0}.$$

В силу предположения индукции $\mu_1 = \mu_2 = \dots = \mu_{i-1} = 0$. Подставляя это в первое равенство, получаем

$$\mu_i u_i = \vec{0},$$

а так как $u_i \neq \vec{0}$, то $\mu_i = 0$, то есть, исходная линейная комбинация является тривиальной, что и дает требуемое. Лемма доказана.

Лемма 3.4. В обозначениях леммы 3.3 сумма $M_{\lambda_1} + M_{\lambda_2} + \dots + M_{\lambda_i}$ является прямой.

Доказательство. Нам надо проверить, что при любом $k = 1, 2, \dots, i$ пересечение подпространства L_k с суммой всех остальных состоит только из нулевого вектора $\vec{0}$. Но если это не так, то какой-то ненулевой вектор u подпространства L_k представим в

виде суммы векторов из L_j , $j = 1, 2, \dots, i$, $j \neq k$, что противоречит лемме 3.3. Лемма доказана.

Поиск собственных векторов.

Процедура поиска собственных векторов вытекает из замечаний начала параграфа.

Каждый вектор соответствует какому-нибудь собственному значению. Поэтому первый шаг состоит в нахождении собственных значений оператора φ , то есть, корней его характеристического многочлена.

Пусть корни $\lambda_1, \lambda_2, \dots, \lambda_i$ найдены. Берем один из них λ_k .

Собственный вектор, соответствующий λ_k , описывается уравнением $\varphi(u) = \lambda_k u$ или уравнением $(\varphi - \lambda_k \varepsilon)(u) = \vec{0}$. Находим ядро оператора $\varphi - \lambda_k \varepsilon$, то есть, все собственные векторы, соответствующие собственному значению λ_k .

На языке систем уравнений эта процедура выглядит следующим образом.

Пусть система

$$\begin{cases} y^1 = a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n \\ y^2 = a_1^2 y^1 + a_2^2 x^2 + \dots + a_n^2 x^n \\ \dots \\ y^n = a_1^n x^1 + a_2^n x^2 + \dots + a_n^n x^n \end{cases}$$

задает оператор φ . Вектор $u = \{x^1, x^2, \dots, x^n\}$ является собственным вектором, соответствующим собственному значению λ , тогда и только тогда, когда

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \dots + a_n^1 x^n = \lambda x^1 \\ a_1^2 y^1 + a_2^2 x^2 + \dots + a_n^2 x^n = \lambda x^2 \\ \dots \\ a_1^n x^1 + a_2^n x^2 + \dots + a_n^n x^n = \lambda x^n \end{cases}$$

или

$$\begin{cases} (a_1^1 - \lambda)x^1 + a_2^1x^2 + \dots + a_n^1x^n = 0 \\ a_1^2y^1 + (a_2^2 - \lambda)x^2 + \dots + a_n^2x^n = 0 \\ \dots \\ a_1^n x^1 + a_2^n x^2 + \dots + (a_n^n - \lambda)x^n = 0. \end{cases}$$

Эта система задает в рассматриваемой системе координат подпространство M_λ .

Существование одномерных инвариантных подпространств.

Пусть λ — корень характеристического многочлена рассматриваемого оператора.

По сказанному выше есть по крайней мере один ненулевой вектор u , для которого $\varphi(u) = \lambda u$.

В комплексном случае, то есть, $K = \mathbb{C}$, корень характеристического многочлена всегда есть. Соответственно, всегда есть ненулевой собственный вектор. Соответственно, всегда есть одномерное инвариантное подпространство.

В действительном случае, то есть, $K = \mathbb{R}$, корня характеристического многочлена может не быть. Тогда не будет и ненулевых собственных векторов. Если же корень характеристического многочлена есть, будет и соответствующий ненулевой собственный вектор.

Существование одномерных или двумерных инвариантных подпространств в действительном случае.

Как уже было объяснено ранее, задача сводится к нахождению ненулевых решение одной из систем

$$\begin{cases} a_1^1x^1 + a_2^1x^2 + \dots + a_n^1x^n = \lambda x^1 \\ a_1^2y^1 + a_2^2x^2 + \dots + a_n^2x^n = \lambda x^2 \\ \dots \\ a_1^n x^1 + a_2^n x^2 + \dots + a_n^n x^n = \lambda x^n. \end{cases}$$

Если есть действительное собственное значение λ , то положительный ответ был получен выше. Пусть действительных корней характеристического уравнения нет. Но комплексный корень $\lambda = \mu + \nu i$ всегда есть. Считая, что мы рассматриваем отображение с матрицей

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \dots & & & \\ a_1^n & a_2^n & \dots & a_n^n \end{pmatrix}$$

в пространстве $\mathbb{C}^n$, в соответствии с замечаниями предыдущего рассуждения получаем существование комплексного ненулевого решения нашей системы, то есть, решения вида $u = v + iw$, где векторы v и w имеют только действительные координаты в рассматриваемом базисе:

$$Au = \lambda u,$$

$$A(v + iw) = (\mu + \nu i)(v + iw),$$

$$Av + iAw = (\mu v - \nu w) + i(\mu w + \nu v).$$

Это числовое равенство, записанное в матричной форме, должно выполняться отдельно для действительной и для мнимой частей, поэтому

$$\begin{cases} Av = \mu v - \nu w \\ Aw = \nu v + \mu w. \end{cases}$$

Это равенство использует только действительные числа и оно выполнено для матрицы нашего оператора в действительном случае. Это означает, что для векторов v_1 и w_1 , столбцы координат обозначены в этой системе через v и w , соответственно, справедлива система

$$\begin{cases} \varphi(v_1) = \mu\varphi(v_1) - \nu\varphi(w_1) \\ \varphi(w_1) = \nu\varphi(v_1) + \mu\varphi(w_1). \end{cases}$$

Отсюда следует, что образы векторов v_1 и w_1 лежат в их линейной оболочке $L(v_1, w_1)$ и, далее, что образ любой линейной комбинации векторов v_1 и w_1 лежит в их линейной оболочке $L(v_1, w_1)$ и, следовательно,

эта линейная оболочка $L(v_1, w_1)$ является инвариантным подпространством рассматриваемого оператора. Ее размерность не больше двух.

§4. Диагонализируемый оператор.

Оператор $\varphi : L \rightarrow L$ называется *диагонализируемым*, если в пространстве L есть базис из собственных векторов оператора φ . Название объясняется тем, что в этом базисе матрица оператора имеет диагональный вид:

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

где λ_i — собственное значение, соответствующее i -тому базисному вектору.

Простой признак диагонализируемости оператора состоит в следующем. Пусть характеристический многочлен оператора обладает n различными корнями. Тогда оператор диагонализируем, ибо каждому корню мы можем сопоставить по крайней мере один собственный вектор. Их окажется n штук. По лемме 3.3 они линейно независимы. n линейно независимых векторов в n -мерном пространстве составляют базис и это дает требуемое.

В общем случае ситуация требует анализа. Пусть $\lambda_i, i = 1, 2, \dots, k$ — все собственные числа оператора φ . В этом перечислении они попарно различны.

Как мы видели подпространство M_{λ_i} собственных векторов оператора φ , соответствующих собственному значению λ , инвариантно. Соответственно, характеристический многочлен $\chi(\lambda)$ оператора φ разлагается в произведение

$$\chi_i(\lambda) \cdot f(\lambda),$$

где $\chi_i(\lambda)$ — характеристический многочлен ограничения оператора φ на подпространство M_{λ_i} . Так как все векторы подпространства M_{λ_i} являются собственными и соответствуют собственному значению

λ_i , то $\chi_i(\lambda) = (\lambda_i - \lambda)^{l_i}$, где l_i — размерность подпространства M_{λ_i} . Для характеристического многочлена $\chi(\lambda)$ оператора φ это означает, что кратность m_i корня λ_i не меньше размерности l_i подпространства M_{λ_i} : $m_i \geq l_i$.

Кратность m_i корня λ_i называется его *алгебраической кратностью*, размерность l_i подпространства M_{λ_i} называется *геометрической кратностью* корня λ_i . Мы установили, что *геометрическая кратность корня λ_i не больше его алгебраической кратности*.

Допустим теперь, что оператор диагонализирован. Тогда, как легко видеть, $m_i = l_i$,

$$\dim M_{\lambda_1} + \dim M_{\lambda_2} + \cdots + \dim M_{\lambda_k} = n,$$

и

$$M_{\lambda_1} \oplus M_{\lambda_2} \oplus \cdots \oplus M_{\lambda_k} = L.$$

Допустим теперь, что $m_i = l_i$. Так как сумма кратностей корней равна степени многочлена (то есть, $m_1 + m_2 + \cdots + m_k = n$), то $l_1 + l_2 + \cdots + l_k = n$. (В действительном случае мы должны оговорить, что все корни действительны.) При этом, как мы видели, сумма $M_{\lambda_1} + M_{\lambda_2} + \cdots + M_{\lambda_k}$ является прямой. Так как $l_1 + l_2 + \cdots + l_k = n$, то есть,

$$\dim M_{\lambda_1} + \dim M_{\lambda_2} + \cdots + \dim M_{\lambda_k} = n,$$

то

$$M_{\lambda_1} \oplus M_{\lambda_2} \oplus \cdots \oplus M_{\lambda_k} = L.$$

Если теперь в каждом из этих подпространств M_{λ_i} мы выберем базис A_i , то объединение $A = A_1 \cup A_2 \cup \cdots \cup A_k$ этих базисов есть базис пространства L .

В соответствии с рассуждениями предыдущего пункта матрица рассматриваемого оператора в базисе A устроена следующим образом. Вдоль диагонали стоят последовательно матрицы ограничений нашего оператора на инвариантные подпространства $L_1, L_2, \dots, L_k$, имеющие диагональный вид. Соответственно, и матрица нашего оператора имеет диагональный вид.

Подведем итог. *Оператор диагонализуем*

1) *тогда и только тогда, когда все корни характеристического многочлена принадлежат полю K* (в комплексном случае это выполняется автоматически) и кратность каждого корня λ_i равна размерности соответствующего подпространства M_{λ_i} ,

2) *тогда и только тогда, когда*

$$M_{\lambda_1} \oplus M_{\lambda_2} \oplus \cdots \oplus M_{\lambda_k} = L.$$

3) *тогда и только тогда, когда*

$$\dim M_{\lambda_1} + \dim M_{\lambda_2} + \cdots + \dim M_{\lambda_k} = n.$$

§5. Связь невырожденного оператора с переходом к новым координатам.

Система

$$\begin{cases} x^1 = c_1^1 y^1 + c_2^1 y^2 + \cdots + c_n^1 y^n \\ x^2 = c_1^2 y^1 + c_2^2 y^2 + \cdots + c_n^2 y^n \\ \cdots \\ x^n = c_1^n y^1 + c_2^n y^2 + \cdots + c_n^n y^n \end{cases}$$

с невырожденной матрицей появлялась у нас дважды.

В первом случае это была система преобразования координат при переходе от базиса A : $e_1, e_2, \dots, e_n$ к базису B в пространстве L :

$$f_1 = c_1^1 e_1 + c_2^1 e_2 + \cdots + c_n^1 e_n,$$

$$f_2 = c_1^2 e_1 + c_2^2 e_2 + \cdots + c_n^2 e_n,$$

$\cdots,$

$$f_n = c_1^n e_1 + c_2^n e_2 + \cdots + c_n^n e_n.$$

Во втором случае речь шла о линейном операторе $\varphi : L \rightarrow L$, при котором $\varphi(e_i) = f_i$, $i = 1, 2, \dots, n$.

Что делает оператор φ ?

Вектору u пространства L , представимому в виде линейной комбинации

$$u = y^1 e_1 + y^2 e_2 + \cdots + y^n e_n$$

векторов базиса A сопоставляется вектор

$$v = \varphi(u) = y^1 f_1 + y^2 f_2 + \cdots + y^n f_n,$$

имеющий такое же разложение, но уже по базису B .

Далее вступает в дело наша система, позволяющая перейти от координат $\{y^1, y^2, \dots, y^n\}$ вектора v в базисе B к его координатам $\{x^1, x^2, \dots, x^n\}$ в основном базисе A .

В силу указанной связи свойства невырожденных линейных операторов и преобразований координат в некоторых отношениях тождественны.