Пламен Сидеров

Записки по АЛГЕБРА линейна алгебра

ВЕДИ

Глава 0. Увод

§ 01. Множества. Изображения. Принцип на математическата индукция

С $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$ ще бележим множествата съответно на естествените, целите, рационалните, реалните и комплексните числа.

Множества. Множествата ще означаваме с главни латински букви. Ако x е елемент на множеството X, ще пишем $x \in X$; в противен случай ще пишем: $x \notin X$. Често едно множество се задава като във фигурни скоби се изпишат елементите му или се даде подходящо описание на тези елементи; например $\{n \in \mathbb{N} \mid n < 5\} = \{1, 2, 3, 4\}$. Ще казваме, че множеството X се съдържа в множеството Y или, че X е подмножество на Y, ако всеки елемент на X е елемент и на Y. Ще използваме означението $X \subset Y$, ако включването е строго (т. е. има поне един елемент на Y, който не е елемент на X) и $X \subseteq Y$, ако е възможно двете множества да съвпадат. Множество, което не притежава нито един елемент ще наричаме празно множество. Без повече подробности, ще отбележим, че празното множество е единствено (ще го бележим с $\emptyset$) и се съдържа във всяко друго множество.

Нека X и Y са множества. Под *сечение* $X \cap Y$ на X и Y ще разбираме множеството, съставено от общите елементи на X и Y. Под *обединение* $X \cup Y$ на X и Y ще разбираме множеството, съставено от тези елементи, които принадлежат на поне едно от множествата X и Y. Например $\{1, 2, 3\} \cap \{2, 3, 4\} = \{2, 3\}$, $\{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\}$. По аналогичен начин се дефинират сечение и обединение на повече от две множества.

Задача 1. *Ако X, Y и Z са множества, да се докажат равенствата:*
а) $X \cup Y = Y \cup X$, $X \cap Y = Y \cap X$;
б) $(X \cup Y) \cup Z = X \cup (Y \cup Z)$;
в) $(X \cap Y) \cap Z = X \cap (Y \cap Z)$;
г) $(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z)$;
д) $(X \cap Z) \cup Z = (X \cup Z) \cap (Y \cup Z)$.

Нека X и Y са множества. Под *декартово произведение* на X и Y ще разбираме множеството $X \times Y$, състоящо се от всички наредени двойки (x, y), където $x \in X$, $y \in Y$. Под *релация* в множеството X ще разбираме всяко подмножество R на декартовото произведение $X \times X$. Ако $(x, y) \in R$, ще пишем xRy.

П р и м е р и. 1. Релацията равенство в множеството на реалните числа $\mathbb{R}$: xRy, ако $x = y$.
2. Релацията неравенство в множеството на реалните числа $\mathbb{R}$: xRy, ако $x < y$.
3. Релацията делимост в множеството на целите числа $\mathbb{Z}$: xRy, ако x дели y.
4. Релацията подобие на матрици, дефинирана в § 16: ARB, ако $A \sim B$.

Една релация R в множеството X ще наричаме:

— *рефлексивна*, ако xRx за всеки елемент $x \in X$;

— *симетрична*, ако от xRy следва yRx за всеки два елемента $x, y \in X$;

— *транзитивна*, ако от xRy и yRz следва xRz за всеки три елемента $x, y, z \in X$;

— *релация на еквивалентност*, ако е едновременно рефлексивна, симетрична и транзитивна.

Примери. 1. Релацията равенство в множеството $\mathbb{R}$ е релация на еквивалентност.

2. Релацията неравенство в множеството $\mathbb{R}$ е транзитивна, но не е рефлексивна и симетрична.

3. Релацията делимост в множеството $\mathbb{Z}$ е рефлексивна и транзитивна, но не е симетрична.

4. Релацията подобие на матрици, дефинирана в § 16, е релация на еквивалентност.

* * *

Изображения. Нека f е изображение (функция) от множеството X към (или в) множеството Y, т.е. правило, по което на всеки елемент на множеството X е съпоставен единствен елемент на множеството Y. Ще използваме означението $f : X \to Y$ или $X \xrightarrow{f} Y$. Ако на елемента $x \in X$ е съпоставен елементът $y \in Y$ (т.е. y е образът на x под действието на f), ще пишем $y = f(x)$. Под *идентитет* на множеството X ще разбираме изображението $\varepsilon : X \to X$, дефинирано чрез равенството $\varepsilon(x) = x$ за всеки елемент $x \in X$. Понякога вместо ε се използва означението id_X или само id, ако X се подразбира.

Нека $f : X \to Y$ е изображение и $U \subseteq X$. Под *ограничение* на f върху U ще разбираме изображението $f|_U : U \to Y$, такова че $f|_U(u) = f(u)$ за всеки елемент $u \in U$.

Нека $f : X \to Y$ е изображение. Ще казваме, че изображението f е:

— *инективно*, ако от $x_1 \neq x_2$ следва $f(x_1) \neq f(x_2)$ $(x_1, x_2 \in X)$;

— *върху* Y, ако всеки елемент $y \in Y$ има *праобраз* в X, т.е. елемент $x \in X$, за който $f(x) = y$;

— *биекция*, ако е едновременно инективно и върху.

Примери. 1. Да разгледаме изображението $f : \mathbb{Z} \to \mathbb{Q}$, дефинирано чрез равенството $f(n) = n$, $n \in \mathbb{Z}$. Това изображение е инективно, но не е върху.

2. Да запишем всяко рационално число като несъкратима дроб $\frac{p}{q}$, $q > 0$. Да разгледаме изображението $f : \mathbb{Q} \to \mathbb{Z}$, дефинирано чрез равенството $f\left(\frac{p}{q}\right) = p$. Това изображение е върху, но не е инективно.

3. Да означим с $2\mathbb{Z}$ множеството от всички четни числа. Да разгледаме изображението $f : \mathbb{Z} \to 2\mathbb{Z}$, дефинирано чрез равенството $f(z) = 2z$, $z \in \mathbb{Z}$. Това изображение е биекция.

Нека $f : X \to Y$ и $g : Y \to Z$ са изображения. Под *композиция* на изображенията f и g ще разбираме изображението $gf : X \to Z$, действащо по правилото $(gf)(x) = g(f(x))$, $x \in X$.

Задача 2. *Да се докаже, че ако $f : X \to Y$, $g : Y \to Z$ и $h : Z \to T$ са изображения, то $(hg)f = h(gf)$.*

Едно изображение $f : X \to Y$ ще наричаме *обратимо*, ако съществува изображение $g : Y \to X$, такова че изображенията $gf : X \to X$ и $fg : Y \to Y$ са идентитетите съответно на множествата X и Y. Изображението g ще наричаме *обратно изображение* на f и ще го бележим с f^{-1}.

Задача 3. *Да се докаже, че ако f е обратимо изображение, то f има единствено обратно изображение.*

Задача 4. *Да се докаже, че едно изображение $f : X \to Y$ е обратимо тогава и само тогава, когато е биекция. Да се докаже, че в този случай f^{-1} също е биекция.*

* * *

Принцип на математическата индукция. *Нека редицата от твърдения*

$$P_1, P_2, \dots, P_n, \dots$$

притежава следните две свойства:

а) *твърдението P_1 е вярно;*

б) *за всяко естествено число n от верността на твърдението P_n следва верността на твърдението P_{n+1}.*

Тогава всички твърдения от редицата са верни.

Ще илюстрираме принципа на математическата индукция върху следното твърдение:
За всяко естествено число n е в сила равенството

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}. \tag{1}$$

Обикновено при доказателството се извършват следните три стъпки:

1. **Основа на индукцията.** Проверяваме, че твърдението е вярно при $n = 1$. Действително, очевидно е равенството $1^2 = \dfrac{1(1+1)(2.1+1)}{6}$.

2. **Индукционно допускане (предположение).** Ако n е произволно естествено число, допускаме, че твърдението е вярно за n, т. е. че за n е в сила равенството

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

3. **Индукционна стъпка.** Въз основа на индукционното допускане доказваме, че твърдението е вярно за $n+1$, т. е. че е в сила равенството

$$1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6}.$$

Действително, използвайки индукционното допускане, последователно получаваме

$$1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{n(n+1)(2n+1) + 6(n+1)^2}{6}$$
$$= \frac{(n+1)[n(2n+1) + 6(n+1)]}{6} = \frac{(n+1)(n+2)(2n+3)}{6}.$$

Сега от принципа на математическата индукция следва, че равенството (1) е вярно за всяко естествено число n.

§ 02. Комплексни числа

Нека $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$ е полином с числови коефициенти $a_0, a_1, \dots, a_n$. Важен въпрос в математиката е дали ако коефициентите на уравнение от вида $f(x) = 0$ са в дадено числово множество M, то и корените му са в същото множество M, т. е. дали уравнението $f(x) = 0$ можем "да решим в множеството M".

Да разгледаме например множеството на целите числа $\mathbb{Z}$. Уравнението $x + 2 = 0$ е с цели коефициенти и единственият му корен $x = -2$ също е цяло число. Изобщо коренът на всяко уравнение от вида $\pm x + a = 0$ $(a \in \mathbb{Z})$ е цяло число.

Уравнението $2x - 1 = 0$ също е с цели коефициенти, но единственият му корен $x = \dfrac{1}{2}$ вече не е цяло число. Следователно в множеството $\mathbb{Z}$ не можем да решаваме дори линейни уравнения. В множеството $\mathbb{Q}$ на рационалните числа обаче, можем да решим всяко линейно уравнение.

Забележка. Множеството на рационалните числа се състои от всички реални числа от вида $\frac{a}{b}$, където $a, b \in \mathbb{Z}$ $(b \neq 0)$. То съдържа множеството $\mathbb{Z}$, тъй като всяко цяло число a се записва като $\frac{a}{1}$. При това

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} \text{ точно когато } a_1b_2 = a_2b_1; \tag{1}$$

$$\left.\begin{aligned} \frac{a_1}{b_1} + \frac{a_2}{b_2} &= \frac{a_1b_2 + a_2b_1}{b_1b_2}, \\ \frac{a_1}{b_1} \cdot \frac{a_2}{b_2} &= \frac{a_1a_2}{b_1b_2}. \end{aligned}\right\} \tag{2}$$

В частност имаме

$$\left.\begin{aligned} \frac{a_1}{1} + \frac{a_2}{1} &= \frac{a_1 + a_2}{1}, \\ \frac{a_1}{1} \cdot \frac{a_2}{1} &= \frac{a_1a_2}{1}. \end{aligned}\right\} \tag{3}$$

Вместо стандартния запис, да запишем рационалното число $\frac{a}{b}$ като наредена двойка (a, b). Тогава (1), (2) и (3) приемат вида

$$(a_1, b_1) = (a_2, b_2) \text{ точно когато } a_1b_2 = a_2b_1; \tag{1'}$$

$$\left.\begin{aligned} (a_1, b_1) + (a_2, b_2) &= (a_1b_2 + a_2b_1, b_1b_2), \\ (a_1, b_1)(a_2, b_2) &= (a_1a_2, b_1b_2) \end{aligned}\right\} \tag{2'}$$

и

$$\left.\begin{aligned} (a_1, 1) + (a_2, 1) &= (a_1 + a_2, 1), \\ (a_1, 1)(a_2, 2) &= (a_1a_2, 1). \end{aligned}\right\} \tag{3'}$$

Да допуснем сега, че познаваме целите числа, но не знаем какво е това рационално число. Да разгледаме ново множество, чиито елементи са наредени двойки (a, b) от цели числа, $b \neq 0$, като ще считаме, че се изпълнява условието (1'). Да означим това множество с $\mathbb{Q}$. Въз основа на операциите събиране и умножение на цели числа да въведем събиране и умножение на елементи на $\mathbb{Q}$ чрез формулите (2'). Тук умножението е "естествено", а събирането е донякъде "странно" (при комплексните числа, които ще въведем по-долу, ще бъде обратно). Да разгледаме подмножеството на $\mathbb{Q}$, състоящо се от всички наредени двойки цели числа от вида $(a, 1)$. Равенствата (3') (които следват от (2')) показват, че първите компоненти на две такива наредени двойки $(a_1, 1)$ и $(a_2, 1)$ се събират и умножават както целите числа a_1 и a_2, а втората им компонента 1 не се променя. Така, ако отъждествим цялото число a с наредената двойка $(a, 1)$, можем да считаме, че познатото ни числово множество $\mathbb{Z}$ се съдържа в новопостроеното множество $\mathbb{Q}$. В това по-голямо множество вече можем да решаваме всяко линейно уравнение.

Уравненията $x^2 - 2 = 0$ и $x^2 + 1 = 0$ са с цели коефициенти. Първото от тях няма корен не само в множеството $\mathbb{Z}$, но и в множеството $\mathbb{Q}$ (но има корен в множеството на реалните числа $\mathbb{R}$). Второто уравнение няма корен дори и в множеството $\mathbb{R}$.

Целта на този параграф е да разширим множеството на реалните числа $\mathbb{R}$ до ново числово множество $\mathbb{C}$ (елементите на което ще наричаме комплексни числа), в което вече можем да решаваме произволни уравнения от вида $f(x) = 0$*.

* * *

Алгебричен вид на комплексните числа. Да означим с $\mathbb{C}$ множеството от всички наредени двойки от реални числа. Нека $z_1 = (a_1, b_1)$ и $z_2 = (a_2, b_2)$ са елементи на $\mathbb{C}$. Ще казваме, че z_1 и z_2 са равни $(z_1 = z_2)$ точно когато $a_1 = a_2$ и $b_1 = b_2$. С помощта на операциите събиране и умножение на реални числа въвеждаме операциите събиране

*Фактът, че корените на всяко уравнение от вида $f(x) = 0$ с комплексни коефициенти са също комплексни числа е дълбок и доказателството му използва материал, който не се съдържа в тази книга. Ето защо само ще формулираме (и използваме) този факт, без да го доказваме.

и умножение на елементи на $\mathbb{C}$ (операциите ще бележим по същия начин) по следните правила:

$$z_1 + z_2 = (a_1 + a_2, b_1 + b_2),$$
$$z_1 z_2 = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1).$$

Директно се проверява, че така въведените операции удовлетворяват обичайните свойства на познатите ни операции събиране и умножение на реални числа. При това наредените двойки $(0,0)$ и $(1,0)$ играят ролята на нула и единица спрямо въведените операции.

Да разгледаме подмножеството на $\mathbb{C}$, състоящо се от всички наредени двойки, чиято втора компонента е равна на нула. Имаме

$$(a_1, 0) + (a_2, 0) = (a_1 + a_2, 0),$$
$$(a_1, 0)(a_2, 0) = (a_1 a_2 - 0.0, a_1.0 + a_2.0) = (a_1 a_2, 0).$$

Така сума и произведение на наредените двойки от разглеждания вид е от същия вид. При това първите компоненти на сумата и произведението на наредените двойки $(a_1, 0)$ и $(a_2, 0)$ са съответно сумата и произведението на реалните числа a_1 и a_2. Това ни позволява да отъждествяваме реалното число a с наредената двойка $(a, 0)$ от $\mathbb{C}$ и да считаме, че $\mathbb{R} \subset \mathbb{C}$. По-нататък вместо наредената двойка $(a, 0)$ ще пишем само a. Елементите на $\mathbb{C}$ ще наричаме *комплексни числа* или само числа.

Да означим с i наредената двойка $(0, 1)$. Имаме

$$i^2 = (0,1)(0,1) = (0.0 - 1.1, 0.1 + 0.1) = (-1, 0) = -1.$$

Освен това, ако $b \in \mathbb{R}$, то

$$bi = (b, 0)(0, 1) = (b.0 - 0.1, b.1 + 0.0) = (0, b).$$

Сега, ако $z = (a, b)$ е произволно комплексно число, имаме

$$z = (a, b) = (a, 0) + (0, b) = a + bi.$$

Така всяко комплексно число може да се запише във вида $z = a + bi$, $a, b \in \mathbb{R}$. Този запис ще наричаме *алгебричен вид* на комплексното число z. Реалното число a ще наричаме *реална част* на z и ще го бележим с $\operatorname{Re} z$; реалното число b ще наричаме *имагинерна част* на z и ще го бележим с $\operatorname{Im} z$. Комплексното число i ще наричаме *имагинерна единица*, а числата от вида bi, $b \in \mathbb{R}$, ще наричаме *чисто имагинерни*. Събирането на комплексни числа в алгебричен вид се извършва като съберем поотделно реалните и имагинерните им части. Умножението се извършва като формално разкрием скобите и вземем предвид, че $i^2 = -1$:

$$(a_1 + b_1 i)(a_2 + b_2 i) = a_1 a_2 + a_1 b_2 i + a_2 b_1 i + b_1 b_2 i^2 = a_1 a_2 - b_1 b_2 + (a_1 b_2 + a_2 b_1) i.$$

Нека $z = a + bi$, $a, b \in \mathbb{R}$. Числото $\overline{z} = a - bi$ ще наричаме *комплексно спрегнато* на z. Директно се проверява,че

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}.$$

Също така $z + \overline{z} = 2 \operatorname{Re} z$ и $z\overline{z} = a^2 + b^2 \in \mathbb{R}$. Под *модул* на z ще разбираме числото $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$. Модулът на едно комплексно число е неотрицателно реално число, при това $|z| = 0$ точно когато $z = 0$. Директно се проверява, че $|z_1 z_2| = |z_1|\,|z_2|$.

С помощта на операциите събиране и умножение на комплексни числа можем (както и при реални числа) да въведем обратните им операции изваждане и деление. Нека $z = a+bi$. Числото $z' = -a-bi$ е такова, че $z+z' = 0$. Ще казваме, че z' е противоположно на z и ще го бележим с $-z$. Вместо $z_1+(-z_2)$ ще пишем z_1-z_2 (изваждане на комплексни числа). Ако $z \neq 0$, да разгледаме числото $z'' = \dfrac{\overline{z}}{|z|^2}$. Имаме $z.z'' = \dfrac{z\overline{z}}{|z|^2} = 1$. Числото z'' ще наричаме обратно на z и ще го бележим с z^{-1}. Вместо $z_1 z_2^{-1}$ ще пишем $\dfrac{z_1}{z_2}$ (деление на комплексни числа). Директно се проверява, че $|z^{-1}| = |z|^{-1}$ и $\left|\dfrac{z_1}{z_2}\right| = \dfrac{|z_1|}{|z_2|}$.

* * *

Тригонометричен вид на комплексните числа. Нека в равнината е фиксирана правоъгълна координатна система с начало точката O. Тогава на всяко комплексно число $z = a+bi$ можем да съпоставим точка в равнината с координати a и b. Така получаваме биекция между множеството $\mathbb{C}$ и множеството от точките в равнината. Удобно е да отъждествяваме тези две множества (както отъждествяваме реалните числа с точките от правата). Реалните числа се изобразяват върху абсцисната ос, която ще наричаме реална, а чисто имагинерните — върху ординатната ос, която ще наричаме имагинерна. Ако едно комплексно число е означено с малка латинска буква, съответстващата му точка в равнината ще означаваме със същата главна буква.

Нека $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$. Тогава на числото $z' = z_1 + z_2$ съответства точката $Z' = (a_1+a_2, b_1+b_2)$, т.е. в сила е векторното равенство $\overrightarrow{OZ'} = \overrightarrow{OZ_1} + \overrightarrow{OZ_2}$. Прието е да се казва, че събирането на комплексни числа е съгласувано със събирането на вектори (чието начало съвпада с O).

Нека $z = a + bi$ и φ е ъгълът, който сключва векторът $\overrightarrow{OZ}$ с положителната посока на реалната ос. Да означим $|z| = r$. Сега имаме $|\overrightarrow{OZ}| = \sqrt{a^2+b^2} = |z| = r$. Също така $\cos\varphi = \dfrac{a}{r}$, т.е. $a = r\cos\varphi$ и $\sin\varphi = \dfrac{b}{r}$, т.е. $b = r\sin\varphi$. Така можем да запишем числото $z = a+bi$ във вида

$$z = r(\cos\varphi + i\sin\varphi).$$

Този запис ще наричаме *тригонометричен вид* на z. Вектора $\overrightarrow{OZ}$ ще наричаме *радиус-вектор* на комплексното число z, а ъгъла φ — *аргумент* на z и ще го бележим с $\arg z$. Така всяко комплексно число се определя от модула си (т.е. дължината на $\overrightarrow{OZ}$) и аргумента си (единствено числото 0 няма аргумент и се определя само от модула си).

З а б е л е ж к а. Аргументът на едно комплексно число z е определен с точност до число, целочислено кратно на 2π. Ако аргументът е в интервала $[0, 2\pi)$, той се нарича главен аргумент и се бележи с $\operatorname{Arg} z$. Главният аргумент еднозначно се определя от z.

Нека комплексните числа $z_k = r_k(\cos\varphi_k + i\sin\varphi_k)$ $(k = 1, 2)$ са записани в тригонометричен вид. Имаме

$$\begin{aligned} z_1z_2 &= r_1r_2(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) = \\ &= r_1r_2(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2)) = \\ &= r_1r_2(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)). \end{aligned} \tag{4}$$

Така в произведението z_1z_2 модулът е равен на произведението на модулите r_1 и r_2, а аргументът е равен на сумата на аргументите φ_1 и φ_2. Аналогично се проверява, че (при $z_2 \neq 0$)

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1-\varphi_2) + i\sin(\varphi_1-\varphi_2)).$$

От равенството (4) следва, че ако $z = r(\cos\varphi + i\sin\varphi)$, то $z^2 = r^2(\cos 2\varphi + i\sin 2\varphi)$. По индукция за всяко естествено число n получаваме равенството

$$z^n = r^n(\cos n\varphi + i\sin n\varphi). \tag{5}$$

Да означим сега $z_k = \sqrt[n]{r}\left(\cos\dfrac{\varphi+2k\pi}{n} + i\sin\dfrac{\varphi+2k\pi}{n}\right)$, $k = 0, 1, \ldots, n-1$ ($\sqrt[n]{r} \in \mathbb{R}$, $\sqrt[n]{r} \geq 0$). Имаме

$$z_k^n = r(\cos(\varphi+2k\pi) + i\sin(\varphi+2k\pi)) = r(\cos\varphi + i\sin\varphi) = z.$$

Така за всяко комплексно число z съществуват n комплексни числа $z_0, z_1, \ldots, z_{n-1}$, такива че $z_k^n = z$. Това ще записваме по следния начин:

$$\sqrt[n]{z} = \sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i\sin\frac{\varphi+2k\pi}{n}\right), \quad k = 0, 1, \ldots, n-1. \tag{6}$$

Формулите (5) и (6) се наричат *формули на Моавър*.

Пример. Нека $z = 1 = 1(\cos 0 + i\sin 0)$ и $w_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}$, $k = 0, 1, \ldots, n-1$. Имаме $w_k^n = \cos 2k\pi + i\sin 2k\pi = 1$. Числата w_k се наричат *n-ти корени на единицата*. Те са върхове на правилен n-ъгълник, вписан в единичната окръжност, един от върховете на който е числото 1.

Забележка. За решаването на едно линейно уравнение използваме само четирите аритметични действия — събиране, изваждане, умножение и деление. За да можем да решим едно квадратно уравнение, допълнително е необходимо (и достатъчно) да можем да извличаме квадратен корен. Вече видяхме, че в $\mathbf{C}$ можем да извличаме произволен n-ти корен. Това ни кара да очакваме, че в $\mathbf{C}$ ще можем да решаваме уравнения от произволна степен.

Нестрогите разсъждения в забележката се прецизират в следната теорема на Даламбер, която по-нататък ще използваме без доказателство:

Теорема. *Всеки неконстантен полином с комплексни коефициенти от степен n има n корена (броени с техните кратности), които са комплексни числа.*

§ 03. Числови полета

Нека F е подмножество на множеството на комплексните числа $\mathbb{C}$ и F съдържа поне два елемента.

Определение. *Ще казваме, че F е числово поле (или за краткост само поле), когато се изпълняват следните условия: ако z_1 и z_2 са произволни числа от F, то числата $z_1 \pm z_2$, $z_1 z_2$ и $\frac{z_1}{z_2}$ (при $z_2 \neq 0$) също са в F.*

Прието е да се казва, че множеството F е затворено относно операциите събиране, изваждане, умножение и деление.

Задача. *Да се докаже, че F е числово поле тогава и само тогава, когато се изпълняват следните условия:*

1. *ако $z_1, z_2 \in F$, то $z_1 \pm z_2$, $z_1 z_2 \in F$;*
2. *ако $z \in F$ и $z \neq 0$, то $z^{-1} \in F$.*

Примери. 1. Множествата $\mathbb{Q}$, $\mathbb{R}$ и $\mathbb{C}$ са полета. Множеството $\mathbb{Z}$ не е поле, тъй като частното на две цели числа не винаги е цяло число.

2. Да разгледаме множеството $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \,|\, a, b \in \mathbb{Q}\}$. Два елемента $a_1 + b_1\sqrt{2}$ и $a_2 + b_2\sqrt{2}$ от това множество са равни точно когато $a_1 = a_2$ и $b_1 = b_2$. Действително, нека $a_1 + b_1\sqrt{2} = a_2 + b_2\sqrt{2}$, т.е. $(b_1 - b_2)\sqrt{2} = a_2 - a_1$. Ако $b_1 - b_2 \neq 0$, то $\sqrt{2} = \frac{a_2 - a_1}{b_1 - b_2} \in \mathbb{Q}$, противоречие. Следователно $b_1 = b_2$, откъдето и $a_1 = a_2$.

Нека $z = a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ и $z \neq 0$, т.е. поне едно от числата a и b е различно от нула. Тогава $a^2 - 2b^2 \neq 0$. Действително, нека $a^2 - 2b^2 = 0$. Ако $b \neq 0$, то $2 = \frac{a^2}{b^2}$ и $\sqrt{2} = \left|\frac{a}{b}\right| \in \mathbb{Q}$, противоречие. Ако $b = 0$, то $0 = a^2 - 2b^2 = a^2$, т.е. $a = 0$, откъдето $z = 0$, отново противоречие.

Директно се проверява, че сума, разлика и произведение на елементи от $\mathbb{Q}(\sqrt{2})$ отново е в $\mathbb{Q}(\sqrt{2})$. Нека $z = a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ и $z \neq 0$. Да означим $z' = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} = \frac{1}{a^2 - 2b^2}(a - b\sqrt{2}) \in \mathbb{Q}(\sqrt{2})$. Имаме $zz' = \frac{(a + b\sqrt{2})(a - b\sqrt{2})}{a^2 - 2b^2} = \frac{a^2 - 2b^2}{a^2 - 2b^2} = 1$. Така $z' = z^{-1}$ и значи $z^{-1} \in \mathbb{Q}(\sqrt{2})$. Сега от задачата по-горе следва, че $\mathbb{Q}(\sqrt{2})$ е поле.

Твърдение. *Всяко числово поле съдържа полето на рационалните числа $\mathbb{Q}$.*

Доказателство. Нека F е поле. Тъй като F съдържа поне два елемента, то в F има ненулево число a. Тогава числата $0 = a - a$ и $1 = \frac{a}{a}$ също са във F. Оттук получаваме, че и числата $2 = 1 + 1$, $3 = 2 + 1$ и т.н. са в F, т.е. $\mathbb{N} \subseteq F$. Аналогично числата $-1 = 0 - 1$, $-2 = -1 - 1$ и т.н. са в F, т.е. $\mathbb{Z} \subseteq F$. Сега, ако a и b са произволни цели числа и $b \neq 0$, то $\frac{a}{b} \in F$. Следователно $\mathbb{Q} \subseteq F$.

§ 04. Матрици. Наредени n-орки

От училищния курс по математика знаем, че с числа можем да извършваме четирите аритметични действия — събиране, изваждане, умножение и деление. Подобни операции можем да извършваме и с някои обекти, които не са числа — например полиномите. Голяма част от съвременната алгебра се занимава с изучаването на обекти, за които са въведени подходящи операции, аналогични на действията с числа. Често свойствата на тези операции значително се различават от свойствата, които познаваме при действия с числа.

В този параграф ще разгледаме обекти, които не са числа, но са съставени от тях, а именно — *правоъгълни таблици от числа*. Всяка такава таблица ще наричаме *матрица*. Ако A е матрица с m *реда и* n *стълба* или $m \times n$ *матрица*, ще използваме означението

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

или за по-кратко $A = (a_{ij})_{m\times n}$. Числата a_{ij} се наричат *елементи* на матрицата A и имат два индекса — първият показва номера на реда, в който стои числото, а вторият — номера на стълба. Ще считаме, че елементите на матрицата принадлежат на фиксирано числово поле F (ако в началото читателят все още се плаши от думата "поле", спокойно може да си мисли, че става дума за реални числа). Множеството от всички $m \times n$ матрици с елементи от F ще означаваме с $F_{m\times n}$.

Ако $m = n$, A се нарича *квадратна матрица от ред* n. В този случай вместо $F_{m\times n}$ ще използваме означението $M_n(F)$.

Диагонала на една квадратна матрица от ред n, който започва от горния ляв ъгъл на матрицата и завършва в долния десен ъгъл (т.е. диагонала съставен от елементите $a_{11}, a_{22}, \dots, a_{nn}$) ще наричаме *главен диагонал* на A. Другия диагонал на матрицата ще наричаме *втори диагонал*.

Квадратната матрица от ред n, в която елементите по главния диагонал са единици, а всички други елементи са нули, ще наричаме *единична матрица от ред* n или само *единична матрица*, ако n се подразбира. Ще я бележим с E_n или само E.

Ако $A \in F_{m\times n}$, матрицата, която се получава като заменим всички редове на A със съответните ѝ стълбове, ще наричаме *транспонирана матрица* на A. Ще я бележим с A^t. Ясно е, че $A^t \in F_{n\times m}$ и $(A^t)^t = A$. Например, ако

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

Може да се каже, че транспонирането на квадратна матрица е завъртане на матрицата около главния ѝ диагонал или симетрия относно главния диагонал.

Сега ще въведем действията *събиране на матрици* и *умножение на матрица с число*.

Определение. *Нека* $A, B \in F_{m\times n}$ *и* $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{m\times n}$. *Нека* $\lambda, \mu \in F$.

1) *Под сума на матриците A и B ще разбираме матрицата $A + B = (a_{ij} + b_{ij})_{m\times n}$. По-подробно*

$$A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}.$$

2) *Под произведение на матрицата A с числото λ ще разбираме матрицата $\lambda A = (\lambda a_{ij})_{m\times n}$ или по-подробно*

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

3) *Матрицата $\lambda A + \mu B = (\lambda a_{ij} + \mu b_{ij})_{m\times n}$ ще наричаме линейна комбинация на матриците A и B с коефициенти съответно λ и μ.*

По аналогичен начин се дефинират сума и линейна комбинация на повече от две матрици. Ясно е, че $(A+B)^t = A^t + B^t$ и $(\lambda A)^t = \lambda A^t$.

Матрицата от $F_{m\times n}$, на която всички елементи са равни на 0 ще наричаме *нулева матрица* и ще я бележим с $\mathbf{0}_{m\times n}$ или само $\mathbf{0}$, ако m и n се подразбират.

Матрицата $(-1).A = (-a_{ij})_{m\times n}$ ще наричаме *противоположна матрица* на A и ще я бележим с $-A$. Ясно е, че $A + (-A) = \mathbf{0}$. Вместо $A + (-B)$ ще пишем $A - B$.

Задача 1. *Нека $A = (a_{ij})_{m\times n}$ е произволна матрица. Да означим с E_{ij} матрицата от $F_{m\times n}$, която има единица на (i, j)-то място и нули на всички останали места. Да се докаже, че*

$$A = \sum_{i,j} a_{ij} E_{ij} \quad (i = 1, 2, \dots, m;\ j = 1, 2, \dots, n).$$

Особено важна роля ще играят по-нататък матриците, които имат само един ред (или само един стълб). Такива матрици ще наричаме *наредени n-орки* или само *n-орки*. Най-често ще ги означаваме с малки латински букви и елементите им ще имат само един индекс; например $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$. Нулевата n-орка $(0, 0, \dots, 0)$ ще бележим с $\mathbf{o}$. Множеството от всички наредени n-орки с елементи от F ще бележим с F^n. Ако фиксираме правоъгълна координатна система в равнината, можем да отъждествяваме елементите на $\mathbb{R}^2$ с точките от равнината, т. е. с комплексните числа. Също така можем да отъждествяваме елементите на $\mathbb{R}^2$ с векторите в равнината, чието начало съвпада с началото на координатната система. Аналогично, елементите на $\mathbb{R}^3$ можем да отъждествяваме с векторите в пространството, чието начало съвпада с началото на фиксирана координатна система.

Понятието линейна комбинация ще използваме най-често за наредени n-орки.

П р и м е р. Нека $\boldsymbol{a} = (2, 1, 0)$, $\boldsymbol{b} = (-1, 3, 5)$, $\boldsymbol{c} = (8, -3, -10)$. Тогава $3\boldsymbol{a} - 2\boldsymbol{b} - \boldsymbol{c} = (3.2 - 2.(-1) - 8, 3.1 - 2.3 - (-3), 3.0 - 2.5 - (-10)) = (0, 0, 0) = \mathbf{o}$.

Важен частен случай на задача 1 е следната

Задача 2. *Нека $\boldsymbol{a} = (a_1, a_2, \dots, a_n)$ е произволна n-орка и $\boldsymbol{e}_1 = (1, 0, 0, \dots, 0)$, $\boldsymbol{e}_2 = (0, 1, 0, \dots, 0)$, $\dots$, $\boldsymbol{e}_n = (0, 0, 0, \dots, 1)$. Да се докаже, че $\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \dots + a_n\boldsymbol{e}_n$.*

§ 05. Системи линейни уравнения. Метод на Гаус

В този параграф ще изучаваме системи уравнения от първа степен с няколко неизвестни, наричани *системи линейни уравнения*. Общата теория на такива системи ще бъде развита в глава II. Тук ще разгледаме метода на Гаус за решаване на произволна система линейни уравнения, който се състои в последователно изключване на неизвестните.

Да разгледаме системата

$$\left| \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \right. . \tag{1}$$

Тази система е с m уравнения и n неизвестни. Неизвестните са означени с x_1, x_2, $\dots$, x_n. Коефициентите пред тях имат двойни индекси: a_{ij} означава коефициентът от i-тото уравнение пред j-тото неизвестно. Свободният член на i-тото уравнение е означен с b_i. Коефициентите и свободните членове принадлежат на фиксирано поле F.

Матрицата

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

ще наричаме *матрица на системата*. Да разгледаме матрицата

$$\overline{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right),$$

която се получава като присъединим към A още един стълб, съставен от свободните членове на системата (и за удобство отделен с вертикална черта). Тази матрица ще наричаме *разширена матрица на системата*.

Под *решение* на системата линейни уравнения (1) ще разбираме всяка наредена n-орка от числа $(r_1, r_2, \dots, r_n)$, която превръща всяко уравнение на (1) в тъждество след заместването на неизвестните x_i съответно с числата r_i $(i = 1, \dots, n)$.

Една система ще наричаме *несъвместима*, ако не притежава нито едно решение и *съвместима* — ако притежава поне едно решение. Една съвместима система ще наричаме *определена*, ако притежава точно едно решение и *неопределена* — ако притежава повече от едно решение (ще видим, че в последния случай системата има безбройно много решения).

Две системи с по n неизвестни ще наричаме *еквивалентни*, ако или и двете са несъвместими, или и двете са съвместими и множествата от решенията им съвпадат.

Под *елементарно преобразуване* на една система ще разбираме едно от следните преобразувания:

1) умножаване на ред от системата с число, различно от нула;
2) прибавяне на ред на системата, умножен с произволно число към друг ред.

Задача 1. *Да се докаже, че ако върху една система извършим краен брой елементарни преобразувания, получената система е еквивалентна на изходната.*

Преминаваме към метода на Гаус за решаване на произволната система линейни уравнения (1). Ако в системата има уравнение, такова че всички коефициенти пред неизвестните са равни на нула, а свободният член е различен от нула, то (заради това уравнение) системата е несъвместима. Ако в такова уравнение и свободният член е равен на нула, то се удовлетворява за всички стойности на неизвестните и като го премахнем, очевидно получаваме система, еквивалентна на дадената. Ето защо ще считаме, че във всяко уравнение на системата коефициентът пред поне едно неизвестно е различен от нула.

Нека в първото уравнение $a_{11} \neq 0$. (В действителност е възможно $a_{11} = 0$ и тогава ще се наложи да работим с коефициента пред някое друго неизвестно, но след евентуално преномериране на неизвестните, можем да считаме, че $a_{11} \neq 0$.) Ще елиминираме неизвестното x_1 от всички уравнения, освен първото. За тази цел умножаваме първото уравнение по $-\dfrac{a_{i1}}{a_{11}}$ и го прибавяме към i-тото уравнение за всяко $i = 2, 3, \dots, m$. Получаваме система от вида

$$\left|\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ b_{22}x_2 + b_{23}x_3 + \cdots + b_{2n}x_n &= b'_2 \\ b_{32}x_2 + b_{33}x_3 + \cdots + b_{3n}x_n &= b'_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ b_{m2}x_2 + b_{m3}x_3 + \cdots + b_{mn}x_n &= b'_m. \end{aligned}\right.$$

Няма да се интересуваме как числата b_{ij} и b'_i се изразяват чрез a_{ij} и b_i. Получената система е еквивалентна на изходната система (1). Както и по-горе, ако има уравнения с нулева лява част и ненулев свободен член, системата е несъвместима, а уравнения с нулева лява част и нулев свободен член можем да премахнем.

Преобразуваме получената система по аналогичен начин, без да "пипаме" първото уравнение. Считайки $b_{22} \neq 0$, умножаваме второто уравнение по $-\dfrac{b_{i2}}{b_{22}}$ и го прибавяме към i-тото за всяко $i = 3, 4, \dots, m$. Получаваме система от вида

$$\left|\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ b_{22}x_2 + b_{23}x_3 + \cdots + b_{2n}x_n &= b'_2 \\ c_{33}x_3 + \cdots + c_{3n}x_n &= b''_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ c_{t3}x_3 + \cdots + c_{tn}x_n &= b''_t. \end{aligned}\right.$$

Тази система съдържа t уравнения, $t \leq m$, тъй като е възможно да сме премахнали някои уравнения.

Кога ще приключи този процес?

Ако на някоя стъпка получим система, в която има уравнение с нулева лява част и ненулев свободен член, то тази система (а значи и изходната) е несъвместима.

В противен случай изходната система се привежда в т. нар. *трапецовиден вид*:

$$\left|\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1k}x_k + \cdots + a_{1n}x_n &= b_1 \\ b_{22}x_2 + b_{23}x_3 + \cdots + b_{2k}x_k + \cdots + b_{2n}x_n &= b'_2 \\ c_{33}x_3 + \cdots + c_{3k}x_k + \cdots + c_{3n}x_n &= b''_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ d_{kk}x_k + \cdots + d_{kn}x_n &= b_k^{(k-1)}. \end{aligned}\right. \tag{2}$$

Тук $a_{11} \neq 0$, $b_{22} \neq 0$, $c_{33} \neq 0$, $\dots$, $d_{kk} \neq 0$. При това, $k \leq n$. (На пръв поглед изглежда възможен и случаят $k > n$, т. е. накрая на системата да има няколко (повече от едно) уравнения, в които участва (с ненулев коефициент) само x_n. В действителност, това означава, че процесът просто не е доведен докрай: с първото от тези уравнения можем да елиминираме x_n от останалите.)

В този случай системата (1) *е съвместима. Тя е определена при* $k = n$ *и неопределена при* $k < n$.

Действително, при $k = n$ системата (2) приема т. нар. *триъгълен вид*:

$$\left|\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ b_{22}x_2 + b_{23}x_3 + \cdots + b_{2n}x_n &= b_2' \\ c_{33}x_3 + \cdots + c_{3n}x_n &= b_3'' \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ d_{nn}x_n &= b_n^{(n-1)}. \end{aligned}\right. \qquad (2)$$

От последното уравнение получаваме еднозначно определена стойност за x_n. Замествайки тази стойност в предпоследното уравнение, получаваме еднозначно определена стойност за x_{n-1}. Продължавайки по този начин, получаваме единствено решение на системата (3), а значи и на системата (1).

Нека $k < n$. Неизвестните x_{k+1}, $\dots$, x_n ще наричаме *свободни неизвестни*. Давайки на тези неизвестни произволни стойности, както и по-горе, последователно получаваме вече еднозначно определени стойности за x_k, x_{k-1}, $\dots$, x_2, x_1. Тъй като стойностите на свободните неизвестни можем да избираме по безбройно много начини, то системата (2), а значи и системата (1), е съвместима, но неопределена.

Да приложим казаното дотук за случая на *хомогенна система* линейни уравнения, т. е. свободните членове на които са равни на нула. Такава система винаги е съвместима, тъй като притежава *нулевото решение* $(0, 0, \dots, 0)$. Нека броят на уравненията на системата е по-малък от броя на неизвестните. Тогава системата се привежда в трапецовиден (но не триъгълен) вид. Следователно тя притежава (и то безбройно много) *ненулеви решения*, т. е. решения, в които стойността на поне едно неизвестно е различна от нула.

П р и м е р и. При практическото решаване на конкретни системи обикновено се работи не със самата система, а с разширената ѝ матрица.

1. Да се реши системата

$$\left|\begin{aligned} x_1 + 2x_2 + 5x_3 &= -9 \\ x_1 - x_2 + 3x_3 &= 2 \\ 3x_1 - 6x_2 - x_3 &= 25. \end{aligned}\right.$$

Преобразуваме разширената матрица на системата:

$$\left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 1 & -1 & 3 & 2 \\ 3 & -6 & -1 & 25 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 0 & -3 & -2 & 11 \\ 0 & -12 & -16 & 52 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 5 & -9 \\ 0 & -3 & -2 & 11 \\ 0 & 0 & -8 & 8 \end{array}\right).$$

Получената система е в триъгълен вид и следователно е определена. Единственото ѝ решение е $x_3 = -1$, $x_2 = -3$, $x_1 = 2$, т. е. наредената тройка $(2, -3, -1)$.

2. Да се реши системата

$$\left|\begin{aligned} x_1 + x_2 + 2x_3 &= 4 \\ x_1 + \phantom{x_2 +{}} x_3 &= 2 \\ x_1 + 2x_2 + 3x_3 &= 7. \end{aligned}\right.$$

Преобразуваме разширената матрица на системата:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 1 & 0 & 1 & 2 \\ 1 & 2 & 3 & 7 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & -2 \\ 0 & 1 & 1 & 3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 2 & 4 \\ 0 & -1 & -1 & -2 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Системата е несъвместима.

3. Да се реши системата

$$\left|\begin{array}{l} x_1 \qquad\quad - x_3 = 2 \\ 2x_1 + 5x_2 \qquad = 7 \\ 3x_1 + 5x_2 - x_3 = 9. \end{array}\right.$$

Преобразуваме разширената матрица на системата:

$$\left(\begin{array}{ccr|c} 1 & 0 & -1 & 2 \\ 2 & 5 & 0 & 7 \\ 3 & 5 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccr|c} 1 & 0 & -1 & 2 \\ 0 & 5 & 2 & 3 \\ 0 & 5 & 2 & 3 \end{array}\right) \to \left(\begin{array}{ccr|c} 1 & 0 & -1 & 2 \\ 0 & 5 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Системата е неопределена. За свободно неизвестно можем да изберем x_3. Нека $x_3 = p$. Тогава от второто уравнение получаваме $x_2 = \frac{3-2p}{5}$ и от първото $- x_1 = 2 + p$.Така решенията на системата са всички наредени тройки $(2 + p, \frac{3-2p}{5}, p)$, където p е произволно число.

§ 06. Пермутации

Нека n е естествено число. Да означим с Ω_n множеството на всички естествени числа от 1 до n. Всяко подреждане (разположение) на тези числа в определен ред ще наричаме *пермутация* на първите n естествени числа (или само пермутация, ако n е фиксирано). При $n = 2$ всички пермутации са две на брой: 12 и 21 (ако няма недоразумение, ще пишем числата без запетаи помежду им). При $n = 3$ пермутациите са шест: 123, 132, 213, 231, 312, 321.

Задача 1. *Да се докаже, че за всяко естествено число n броят на всички пермутации на първите n естествени числа е* $n! = 1.2.3 \ldots n$.

Нека $i_1 i_2 \ldots i_n$ е пермутация. Ще казваме, че числата i_k и i_s в тази пермутация образуват *инверсия*, ако $k < s$, но $i_k > i_s$ (т.е. в пермутацията по-голямото от двете числа е преди по-малкото). Една пермутация ще наричаме *четна*, ако броят на двойките числа в пермутацията, образуващи инверсия е четно число. В противен случай пермутацията ще наричаме *нечетна*. При $n = 3$ четните пермутации са 123 (0 инверсии), 231 (2 инверсии) и 312 (2 инверсии), а нечетните са 132 (1 инверсия), 213 (1 инверсия) и 321 (3 инверсии).

Ясно е, че ако в една пермутация разменим местата само на две числа, а останалите оставим на място, получаваме нова пермутация. Такова преобразуване ще наричаме *транспозиция*.

Лема 1. *Ако една пермутация се получава от друга с помощта на транспозиция, то двете пермутации са с различна четност.*

Д о к а з а т е л с т в о. Нека σ е първоначалната пермутация, а τ се получава от σ чрез транспозиция на числата i и j. Нека първо i и j са съседни, т.е. $\sigma = \ldots ij \ldots$, $\tau = \ldots ji \ldots$. Ясно е, че σ и τ се различават само по взаимното разположение на числата i и j. Тогава броят на инверсиите в τ или се е увеличил с единица (ако i и j не са образували инверсия в σ), или се е намалил с единица (ако са образували). Следователно σ и τ са с различна четност.

Нека сега разположението на i и j в σ е произволно и

$$\sigma = \ldots ik_1 \ldots k_t j \ldots, \qquad \tau = \ldots jk_1 \ldots k_t i \ldots .$$

Размяната на местата на i и j в σ можем да осъществим с помощта на няколко "съседни" транспозиции, а именно: първо i "прескача" последователно през $k_1, \dots, k_t$ и застава до j, после i "прескача" през j и накрая j "прескача" обратно през $k_t, \dots, k_1$ и застава на старото място на i. Общият брой на "съседните" транспозиции е $t+1+t = 2t+1$, т.е. този брой е нечетно число. Тъй като всяка "съседна" транспозиция променя четността на пермутацията, то четностите на σ и τ са различни.

Следствие 2. *При $n \geq 2$ броят на четните пермутации е равен на броя на нечетните.*

Доказателство. На всяка четна пермутация σ да съпоставим пермутация τ, която се получава, като в σ разменим местата на (например) първите две числа. Според лема 1 τ е нечетна пермутация. Лесно се проверява, че по този начин се получава взаимно еднозначно съответствие (биекция) между четните и нечетните пермутации, което доказва твърдението.

Глава I. Матрици и детерминанти

§ 1. Детерминанти

Детерминанти от втори и трети ред. Да разгледаме системата с две линейни уравнения и две неизвестни:

$$\left| \begin{array}{l} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{array} \right. \tag{1}$$

Матрицата на системата е $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Изключвайки последователно неизвестните – първо x_2, а после x_1 — получаваме системата

$$\left| \begin{array}{l} (a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - a_{12}b_2 \\ (a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - b_1a_{21}. \end{array} \right. \tag{2}$$

Нека $a_{11}a_{22} - a_{12}a_{21} \neq 0$. Тогава

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \qquad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Директната проверка показва, че получените стойности за x_1 и x_2 са решение на системата.

Забелязваме, че изключвайки от системата (1) неизвестните, освен едно от тях, получихме системата (2), в която коефициентът пред x_1 и x_2 е един и същ. (В § 3 ще видим, че това не е случайно и аналогичен факт е в сила за система с n уравнения и n неизвестни.) Този общ коефициент се изразява чрез елементите на матрицата A и се нарича *детерминанта на A* или още *детерминанта от втори ред*. За него ще използваме означението $\det A$ или $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$. При тези означения имаме

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

Така на всяка **матрица** *от втори ред съпоставяме* **число**, *наречено нейна детерминанта.*

Числителите в изразите за x_1 и x_2 също са детерминанти от втори ред и са равни съответно на $\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}$ и $\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$.

Да разгледаме сега системата с три линейни уравнения и три неизвестни:

$$\left| \begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{array} \right.$$

Матрицата на системата е

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Ако умножим първото, второто и третото уравнение съответно с $a_{22}a_{33} - a_{23}a_{32}$, $-a_{12}a_{33} + a_{13}a_{32}$ и $a_{12}a_{23} - a_{13}a_{22}$ и ги съберем, ще получим уравнението

$$\begin{aligned}(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32})x_1 \\ = b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - a_{13}a_{22}b_3 - a_{12}b_2a_{33} - b_1a_{23}a_{32}. \quad (3)\end{aligned}$$

Коефициентът пред x_1 в това уравнение се нарича *детерминанта на матрицата* A или *детерминанта от трети ред*. Използвайки аналогични означения, както при детерминанти от втори ред, имаме

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

Да означим с Δ_i детерминантата на матрицата от трети ред, която се получава от матрицата A като заменим i-ия й стълб със стълба от свободните членове на системата ($i = 1, 2, 3$). Забелязваме, че дясната страна на уравнението (3) е точно Δ_1. Ако означим $\Delta = \det A$, то (3) приема вида $\Delta x_1 = \Delta_1$. По аналогичен начин от дадената система можем да получим уравненията $\Delta x_2 = \Delta_2$ и $\Delta x_3 = \Delta_3$ (общото правило за получаването на тези уравнения ще се появи в § 3). Сега, ако $\Delta \neq 0$, намираме

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad x_3 = \frac{\Delta_3}{\Delta}.$$

Директната проверка (макар и технически по-сложна) показва, че намерените стойности за x_1, x_2 и x_3 са решение на дадената система.

Изразът за детерминанта от трети ред изглежда сложен, но практическите пресмятания се улесняват с помощта на т. нар. *правило на Сарус*. То показва кои от членовете на детерминантата се вземат със знак плюс и кои със знак минус. По-долу схематично е показано това правило

$+$  $-$ 

И така, видяхме, че и система с три линейни уравнения и три неизвестни се свежда към система от вида

$$\left| \begin{aligned} \Delta x_1 &= \Delta_1 \\ \Delta x_2 &= \Delta_2 \\ \Delta x_3 &= \Delta_3 \end{aligned} \right. ,$$

в която коефициентът Δ пред x_1, x_2 и x_3 е един и същ. Този общ коефициент нарекохме детерминанта от трети ред. Той се изразява чрез елементите на матрицата A на системата. (Ако човек разполага със значително повече търпение и време, може да провери, че аналогична ситуация е налице и за система с четири линейни уравнения и четири неизвестни.) Така стигаме до мисълта да въведем *детерминанта от n-ти ред* по аналогичен начин. Това, обаче, е практически невъзможно, тъй като с растенето на n пресмятанията се увеличават неимоверно много. Ние ще изберем друг подход: разглеждайки правилата, по които се получават изразите за детерминанта от втори и (най-вече) от трети ред, ще изведем определение (формула) за детерминанта от n-ти ред.

* * *

Детерминанти от n-ти ред. Да напомним израза за детерминантата от трети ред:

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

Какви закономерности забелязваме в този израз?

1. Всеки член от развитието на детерминантата (ако се абстрахираме от знака му) е произведение на три елемента, стоящи както в различни редове, така и в различни стълбове на матрицата A. Следователно, ако подредим множителите в едно такова произведение така, че първите им индекси да са в растящ ред $(a_{1i_1}a_{2i_2}a_{3i_3})$, то вторите им индекси образуват пермутация на числата 1, 2, 3. При това, всички възможни произведения от този вид участват в развитието на детерминантата.

2. Всяко произведение участва със знак плюс или минус. При това произведението $a_{1i_1}a_{2i_2}a_{3i_3}$ участва със знак плюс, ако $i_1i_2i_3$ е четна пермутация и със знак минус — ако е нечетна.

След тези наблюдения можем да дадем определение за детерминанта от n-ти ред. Предварително ще ни е необходимо следното означение: ако $i_1i_2\dots i_n$ е пермутация на естествените числа от 1 до n, с $[i_1i_2\dots i_n]$ ще означаваме броя на инверсиите в тази пермутация.

Определение. *Нека n е естествено число и*

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

е квадратна матрица от n-ти ред. Под детерминанта на матрицата A или детерминанта от n-ти ред ще разбираме израза

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^{[i_1i_2\dots i_n]} a_{1i_1}a_{2i_2}\dots a_{ni_n},$$

*където сумирането се извършва по всичките $n!$ на брой пермутации $i_1i_2\dots i_n$ на естествените числа от 1 до n**.

*Да разгледаме шахматна дъска с размери $n \times n$ и n топа, разположени върху нея така, че никои два не се бият. Ясно е, че тогава във всеки ред и всеки стълб ще има точно по един топ. Нека топът в първия ред е в стълб с номер i_1, топът във втория ред е в стълб с номер i_2 и т. н. Тогава произведението $a_{1i_1}a_{2i_2}\dots a_{ni_n}$ е един от членовете в развитието на детерминантата. По този начин от всевъзможните разположения на топовете върху шахматната дъска получаваме всички членове на детерминантата.

Задача 1. *Да се докаже, че ако всички числа, стоящи под главния диагонал на матрицата A са равни на нула (или, както се казва, A е триъгълна матрица), то* $\det A = a_{11}a_{22}\ldots a_{nn}$.

Ще изследваме с какъв знак участва дадено произведение от развитието на детерминантата при условие, че първите индекси на множителите му не са подредени непременно в растящ ред.

Лема 1. *Произведението $a_{k_1j_1}a_{k_2j_2}\ldots a_{k_nj_n}$ участва в развитието на детерминантата от n-ти ред със знак* $(-1)^{[k_1k_2\ldots k_n]+[j_1j_2\ldots j_n]}$.

Доказателство. Ако разменим местата на два множителя в произведението, то както в пермутацията, съставена от първите индекси, така и в пермутацията, съставена от вторите индекси се извършва по една транспозиция. Според лема 1 от § 06 и двете пермутации променят четността си, но тогава четността на сумата $[k_1k_2\ldots k_n]+[j_1j_2\ldots j_n]$ остава същата. Ясно е, че с краен брой такива размествания можем да подредим множителите така, че първите им индекси да са в растящ ред. Тогава вторите индекси образуват някаква пермутация $i_1i_2\ldots i_n$. По определение даденото произведение участва в развитието на детерминантата от n-ти ред със знак $(-1)^{[i_1i_2\ldots i_n]} = (-1)^{[12\ldots n]+[i_1i_2\ldots i_n]}$. От казаното по-горе следва, че този знак е равен на $(-1)^{[k_1k_2\ldots k_n]+[j_1j_2\ldots j_n]}$.

Теорема 2. *В сила е равенството* $\det A^t = \det A$ *(т. е. при транспониране детерминантата не се променя).*

Доказателство. Всеки член $a_{k_1j_1}a_{k_2j_2}\ldots a_{k_nj_n}$ от развитието на $\det A$ е член и от развитието на $\det A^t$ и обратно, тъй като множителите му са в различни редове и различни стълбове както на матрицата A, така и на матрицата A^t. Различието се състои само в това, че в развитието на $\det A$ първите индекси на множителите означават номерата на редовете, а вторите — номерата на стълбовете на A, докато в развитието на $\det A^t$ е обратно. Според лема 1 всяко такова произведение участва както в първото, така и във второто развитие със знак $(-1)^{[k_1k_2\ldots k_n]+[j_1j_2\ldots j_n]}$. Следователно $\det A^t = \det A$.

§ 2. Основни свойства на детерминантите

До края на параграфа ще работим с фиксирана матрица $A = (a_{ij})_{n\times n}$. За съкращаване на изложението ще използваме символиката, която следва до края на параграфа без подробни обяснения, като разчитаме, че читателят ще я тълкува по естествен начин.

Всички свойства по-долу са формулирани само за редовете на матрицата A, но поради наличието на теоремата за транспонираната детерминанта (теорема 2 от § 1) същите свойства са валидни и за стълбовете.

Да означим с $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ наредените n-орки, представляващи редовете на A, т.е. $\boldsymbol{a}_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $\boldsymbol{a}_2 = (a_{21}, a_{22}, \ldots, a_{2n})$, $\ldots$, $\boldsymbol{a}_n = (a_{n1}, a_{n2}, \ldots, a_{nn})$. Тогава матрицата A символично се записва по следния начин:

$$A = \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

Свойство 1. *Ако един от редовете на матрицата A се състои само от нули, то* $\det A = 0$.

Доказателство. Нека $a_p = 0$. От определението за детерминанта следва, че във всеки член от развитието на детерминантата участва като множител елемент на a_p, откъдето $\det A = 0$.

Свойство 2. *Нека $a_p = a'_p + a''_p$, където $a'_p = (a'_{p1}, a'_{p2}, \dots, a'_{pn})$, $a''_p = (a''_{p1}, a''_{p2}, \dots, a''_{pn})$. Тогава*

$$\det A = \det \begin{pmatrix} \vdots \\ a_p \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a'_p \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_p \\ \vdots \end{pmatrix}.$$

Доказателство. Имаме

$$\det A = \sum (-1)^{[i_1 \dots i_n]} \dots a_{pi_p} \dots = \sum (-1)^{[i_1 \dots i_n]} \dots (a'_{pi_p} + a''_{pi_p}) \dots$$

$$= \sum (-1)^{[i_1 \dots i_n]} \dots a'_{pi_p} \dots + \sum (-1)^{[i_1 \dots i_n]} \dots a''_{pi_p} \dots = \det \begin{pmatrix} \vdots \\ a'_p \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ a''_p \\ \vdots \end{pmatrix}.$$

Свойство 3. *Ако $\lambda \in F$, то*

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_p \\ \vdots \\ a_n \end{pmatrix} = \lambda \det A.$$

Доказателство. Имаме

$$\det \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_p \\ \vdots \\ a_n \end{pmatrix} = \sum (-1)^{[i_1 \dots i_n]} \dots (\lambda a_{pi_p}) \dots = \lambda \sum (-1)^{[i_1 \dots i_n]} \dots a_{pi_p} \dots = \lambda \det A.$$

Свойство 4. *Ако $a_p = a_q$ при $p \neq q$, т. е. $a_{p1} = a_{q1}$, $a_{p2} = a_{q2}$, ... , $a_{pn} = a_{qn}$, то* $\det A = 0$.

Доказателство. Нека $p < q$. На всеки член

$$a_{1i_1} \dots a_{pi_p} \dots a_{qi_q} \dots a_{ni_n}$$

от развитието на детерминантата да съпоставим члена

$$a_{1i_1} \dots a_{pi_q} \dots a_{qi_p} \dots a_{ni_n}.$$

Първият от тях участва със знак $(-1)^{[i_1 \dots i_p \dots i_q \dots i_n]}$, а вторият със знак $(-1)^{[i_1 \dots i_q \dots i_p \dots i_n]}$. Според лема 1 от § 06 тези знаци са противоположни. Освен това, от $a_{pi_p} = a_{qi_p}$ и

$a_{qi_q} = a_{pi_q}$ следва, че тези два члена са равни по абсолютна стойност. Така всички членове в развитието на детерминантата се разбиват на двойки със сума, равна на нула и следователно $\det A = 0$.

Свойство 5. *Ако* $\lambda \in F$ *и* $\boldsymbol{a}_p = \lambda \boldsymbol{a}_q$ *при* $p \neq q$ *то* $\det A = 0$.

Д о к а з а т е л с т в о. Нека $p < q$. Имаме

$$\det A = \det\begin{pmatrix} \vdots \\ \boldsymbol{a}_p \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} = \det\begin{pmatrix} \vdots \\ \lambda\boldsymbol{a}_q \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} \overset{\text{св.3}}{=} \lambda \det\begin{pmatrix} \vdots \\ \boldsymbol{a}_q \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} \overset{\text{св.4}}{=} \lambda.0 = 0.$$

Свойство 6. *Нека* $\lambda \in F$. *Ако* $p \neq q$ *и към реда* $\boldsymbol{a}_p$ *прибавим реда* $\boldsymbol{a}_q$, *умножен с* λ, *то детерминантата не се променя. По-общо, ако към даден ред прибавим произволна линейна комбинация на останалите редове, детерминантата не се променя.*

Д о к а з а т е л с т в о. Нека $p < q$. Използвайки предните свойства, последователно получаваме

$$\det\begin{pmatrix} \vdots \\ \boldsymbol{a}_p + \lambda\boldsymbol{a}_q \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} \overset{\text{св.2}}{=} \det\begin{pmatrix} \vdots \\ \boldsymbol{a}_p \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} + \det\begin{pmatrix} \vdots \\ \lambda\boldsymbol{a}_q \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} \overset{\text{св.5}}{=} \det A + 0 = \det A.$$

Обобщението се получава с неколкократно прилагане на горните разсъждения.

Свойство 7. *Ако разменим местата на два различни реда* $\boldsymbol{a}_p$ *и* $\boldsymbol{a}_q$, *детерминантата си сменя знака.*

Д о к а з а т е л с т в о. Нека $p < q$. Искаме да докажем, че

$$\det\begin{pmatrix} \vdots \\ \boldsymbol{a}_p \\ \vdots \\ \boldsymbol{a}_q \\ \vdots \end{pmatrix} = -\det\begin{pmatrix} \vdots \\ \boldsymbol{a}_q \\ \vdots \\ \boldsymbol{a}_p \\ \vdots \end{pmatrix}.$$

Имаме

$$\det\begin{pmatrix}\vdots\\ a_p\\ \vdots\\ a_q\\ \vdots\end{pmatrix} \overset{\text{св.6}}{=} \det\begin{pmatrix}\vdots\\ a_p+a_q\\ \vdots\\ a_q\\ \vdots\end{pmatrix} \overset{\text{св.6}}{=} \det\begin{pmatrix}\vdots\\ a_p+a_q\\ \vdots\\ -a_p\\ \vdots\end{pmatrix}$$

$$\overset{\text{св.6}}{=} \det\begin{pmatrix}\vdots\\ a_q\\ \vdots\\ -a_p\\ \vdots\end{pmatrix} \overset{\text{св.3}}{=} -\det\begin{pmatrix}\vdots\\ a_q\\ \vdots\\ a_p\\ \vdots\end{pmatrix}.$$

Свойство 8. *Ако един ред е линейна комбинация на останалите редове, то* $\det A = 0$.

Доказателство. Нека например $a_1 = \lambda_2 a_2 + \cdots + \lambda_n a_n$. Имаме

$$\det A = \det\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix} = \det\begin{pmatrix}\lambda_2 a_2 + \cdots + \lambda_n a_n\\ a_2\\ \vdots\\ a_n\end{pmatrix} \overset{\text{св.2}}{=} \det\begin{pmatrix}\lambda_2 a_2\\ a_2\\ \vdots\\ a_n\end{pmatrix} + \cdots + \det\begin{pmatrix}\lambda_n a_n\\ a_2\\ \vdots\\ a_n\end{pmatrix}$$

$$\overset{\text{св.5}}{=} 0 + \cdots + 0 = 0.$$

Пример. Нека L е поле и $\alpha_1, \alpha_2, \ldots, \alpha_n \in L$. Детерминантата

$$W = W(\alpha_1, \alpha_2, \ldots, \alpha_n) = \begin{vmatrix} 1 & 1 & \ldots & 1\\ \alpha_1 & \alpha_2 & \ldots & \alpha_n\\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_1^{n-1} & \alpha_2^{n-1} & \ldots & \alpha_n^{n-1}\end{vmatrix}$$

се нарича *детерминанта на Вандермонд* на елементите $\alpha_1, \alpha_2, \ldots, \alpha_n$. Ще докажем, че

$$W(\alpha_1, \alpha_2, \ldots, \alpha_n) = \prod_{n \geq i > j \geq 1} (\alpha_i - \alpha_j).$$

Умножаваме $(n-1)$-ия ред на детерминантата с α_1 и го изваждаме от n-ия, после умножаваме $(n-2)$-ия ред с α_1 и го изваждаме от $(n-1)$-ия и т.н., накрая умножаваме първия ред с α_1 и го изваждаме от втория. Получаваме

$$W = W(\alpha_1, \alpha_2, \ldots, \alpha_n) = \begin{vmatrix} 1 & 1 & \ldots & 1\\ 0 & \alpha_2 - \alpha_1 & \ldots & \alpha_n - \alpha_1\\ 0 & \alpha_2^2 - \alpha_1\alpha_2 & \ldots & \alpha_n^2 - \alpha_1\alpha_n\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \alpha_2^{n-1} - \alpha_1\alpha_2^{n-2} & \ldots & \alpha_n^{n-1} - \alpha_1\alpha_n^{n-2}\end{vmatrix}.$$

Като развием тази детерминанта по елементите на първия стълб и изнесем от стълбовете на получената детерминанта от $(n-1)$-ви ред съответно $\alpha_2-\alpha_1, \ldots, \alpha_n-\alpha_1$, получаваме

$$W(\alpha_1,\alpha_2,\ldots,\alpha_n)=(\alpha_2-\alpha_1)\ldots(\alpha_n-\alpha_1)\begin{vmatrix}1 & \ldots & 1\\ \alpha_2 & \ldots & \alpha_n\\ \alpha_2^2 & \ldots & \alpha_n^2\\ \vdots & & \vdots\\ \alpha_2^{n-2} & \ldots & \alpha_n^{n-2}\end{vmatrix}=\prod_{i=2}^{n}(\alpha_i-\alpha_1)W(\alpha_2,\ldots,\alpha_n).$$

Сега доказателството завършва с очевидна индукция по n.

§ 3. Адюнгирани количества и поддетерминанти. Формули на Крамер

Адюнгирани количества и поддетерминанти. Пресмятането на детерминанти от n-ти ред директно чрез определението за детерминанта е свързано с огромни технически трудности дори и за сравнително малки стойности на n. Съществуват обаче методи, с помощта на които детерминанта от даден ред се изразява чрез детерминанти от по-малки редове. В този параграф ще видим как детерминанта от n-ти ред се изразява чрез n детерминанти от ред $n-1$.

Нека $A=(a_{ij})_{n\times n}$ е квадратна матрица от ред n и a_{pq} е фиксиран неин елемент. Да извадим пред скоби множителя a_{pq} от всички членове в развитието на $\det A$, в които той участва. Получения израз в скобите ще означаваме с A_{pq} и ще наричаме *адюнгирано количество* на елемента a_{pq}. Като се абстрахираме от знака, всеки член от развитието на A_{pq} има вида $a_{1i_1}\ldots a_{p-1,i_{p-1}}a_{p+1,i_{p+1}}\ldots a_{ni_n}$. От определението за детерминанта следва, че в нито един такъв член не участва елемент от p-ия ред и от q-ия стълб на A, в частност числата $i_1, \ldots, i_{p-1}, i_{p+1}, \ldots, i_n$ са различни от q. Тогава броят на всички членове от развитието на A_{pq} е равен на броя на всички пермутации $i_1\ldots i_{p-1}i_{p+1}\ldots i_n$ на естествените числа от 1 до n, с изключение на числото q, т. е. равен е на $(n-1)!$.

От определенията за детерминанта и адюнгирано количество на елемент следват равенствата

$$\det A=\sum_{k=1}^{n}a_{pk}A_{pk}=a_{p1}A_{p1}+a_{p2}A_{p2}+\cdots+a_{pn}A_{pn},\tag{1}$$

$$\det A=\sum_{k=1}^{n}a_{kq}A_{kq}=a_{1q}A_{1q}+a_{2q}A_{2q}+\cdots+a_{nq}A_{nq}.\tag{2}$$

Равенствата (1) и (2) се наричат съответно *развитие на детерминантата по p-ия ред* и *развитие на детерминантата по q-ия стълб*.

Да означим с Δ_{pq} детерминантата на матрицата от $(n-1)$-ви ред, която се получава от A, като зачертаем p-ия ред и q-ия стълб на A. Имаме

$$\Delta_{pq}=\begin{vmatrix}a_{11} & \cdots & a_{1,q-1} & a_{1,q+1} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ a_{p-1,1} & \cdots & a_{p-1,q-1} & a_{p-1,q+1} & \cdots & a_{p-1,n}\\ a_{p+1,1} & \cdots & a_{p+1,q-1} & a_{p+1,q+1} & \cdots & a_{p+1,n}\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{n,q-1} & a_{n,q+1} & \cdots & a_{nn}\end{vmatrix}.$$

Тази детерминанта ще наричаме *поддетерминанта* на $\det A$ от ред $n-1$.

Теорема 1. *В сила е равенството* $A_{pq} = (-1)^{p+q}\Delta_{pq}$.

Преди да докажем теоремата, ще отбележим, че от нея и равенствата (1) и (2) получаваме

Следствие 2. *В сила са равенствата*

$$\det A = \sum_{k=1}^{n} (-1)^{p+k} a_{pk}\Delta_{pk} = (-1)^{p+1}a_{p1}\Delta_{p1} + (-1)^{p+2}a_{p2}\Delta_{p2} + \cdots + (-1)^{p+n}a_{pn}\Delta_{pn},$$

$$\det A = \sum_{k=1}^{n} (-1)^{k+q} a_{kq}\Delta_{kq} = (-1)^{1+q}a_{1q}\Delta_{1q} + (-1)^{2+q}a_{2q}\Delta_{2q} + \cdots + (-1)^{n+q}a_{nq}\Delta_{nq}.$$

Доказателство на теорема 1. От определенията за A_{pq} и Δ_{pq} следва, че (ако се абстрахираме от знака) всяко събираемо от развитието на A_{pq} участва точно веднъж в развитието на Δ_{pq}. Ще сравним знаците, с които участва всяко такова събираемо в двете развития.

Да разгледаме събираемото

$$a_{1i_1}\dots a_{p-1,i_{p-1}}a_{pq}a_{p+1,i_{p+1}}\dots a_{ni_n} \tag{3}$$

от развитието на $\det A$. От него получаваме събираемото

$$a_{1i_1}\dots a_{p-1,i_{p-1}}a_{p+1,i_{p+1}}\dots a_{ni_n} \tag{4}$$

от развитието на A_{pq}. То участва в развитието на A_{pq} със знак, с който събираемото (3) участва в развитието на $\det A$, т. е. със знак (-1) на степен

$$[i_1\dots i_{p-1}qi_{p+1}\dots i_n]. \tag{5}$$

Същото събираемо (4) участва в развитието на Δ_{pq} със знак (-1) на степен

$$[i_1\dots i_{p-1}i_{p+1}\dots i_n]. \tag{6}$$

Всяка двойка числа, образуващи инверсия в (6) образува инверсия и в (5). Ще преcметнем с колко броят на инверсиите в (5) е по-голям от броя на инверсиите в (6) заради участието на q (ще подчертаем, че числото q стои на p-то място в (5)).

Нека m от числата $i_1, \dots, i_{p-1}$ са по-големи от q и значи **образуват** инверсия с q в (5). Следователно $p-1-m$ от тях са по-малки от q (и не образуват инверсия с q). Всички естествени числа, по-малки от q са $q-1$ на брой. Тогава измежду числата $i_{p+1}, \dots, i_n$ броят на тези, които са по-малки от q, и значи **образуват** инверсия с q в (5), е равен на $q-1-(p-1-m)$. Следователно броят на новите инверсии в (5), спрямо тези в (6), е равен на

$$m + [q-1-(p-1-m)] = q-p+2m \equiv q-p \equiv q+p \pmod 2.^*$$

И така, броят на новите инверсии в (5) е число със същата четност, както числото $p+q$. Следователно, ако събираемото (4) участва в развитието на Δ_{pq} със знак ε, то същото събираемо участва в развитието на A_{p+q} със знак $(-1)^{p+q}\varepsilon$. Оттук получаваме

*Изразът $a \equiv b \pmod 2$ означава, че целите числа a и b са с еднаква четност.

равенството $A_{pq} = (-1)^{p+q}\Delta_{pq}$ (аналогично равенство е вярно дори "посъбираемно"). Теоремата е доказана.

Нека i и j са естествени числа. Да положим

$$\delta_{ij} = \begin{cases} 1, & \text{ако } i = j \\ 0, & \text{ако } i \neq j \end{cases}.$$

(Символът δ_{ij} се нарича *символ на Кронекер.*)

Твърдение 3. *В сила са равенствата*

$$\sum_{k=1}^{n} a_{pk}A_{rk} = \delta_{pr} \det A = \begin{cases} \det A, & \text{ако } p = r \\ 0, & \text{ако } p \neq r \end{cases},$$

$$\sum_{k=1}^{n} a_{kq}A_{kr} = \delta_{qr} \det A = \begin{cases} \det A & \text{ако } q = r \\ 0 & \text{ако } q \neq r \end{cases}.$$

Доказателство. Ще докажем само първото равенство. При $p = r$ получаваме развитието (1) на $\det A$ по p-ия ред. Нека $p \neq r$ (казваме, че сме развили детерминантата по елементите на p-ия ред с адюнгираните количества на елементите от r-ия ред). В този случай получаваме развитие на детерминанта, чиято матрица се получава от A, като на мястото на r-ия ред на A поставим p-ия ѝ ред (т. е. матрица с два равни реда). Според свойство 4 от § 2 тази детерминанта е равна на нула.

* * *

Формули на Крамер. Сега вече можем да се уверим, че детерминантите от n-ти ред притежават аналогични свойства на свойствата на детерминантите от втори и трети ред, формулирани в началото на § 1, а именно *детерминанта от n-ти ред е коефициентът, който се получава пред x_i* **за всяко** $i = 1, 2, \ldots, n$, *когато в (обща) система с n линейни уравнения и n неизвестни изключим всички неизвестни, освен x_i.*

Да разгледаме системата

$$\left| \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{array} \right. \tag{7}$$

Матрицата на системата е

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

Да означим $\Delta = \det A$. Детерминантата Δ ще наричаме *детерминанта на системата.* Нека Δ_i означава детерминантата на матрицата, която се получава от матрицата A като заменим i-ия стълб на A със стълба от свободните членове на системата ($i = 1, 2, \ldots, n$).

Да умножим първото уравнение на системата (7) с A_{11}, второто — с A_{21} и т. н., n-тото — с A_{n1} и да ги съберем. Получаваме уравнението

$$x_1\sum_{k=1}^{n}a_{k1}A_{k1}+x_2\sum_{k=1}^{n}a_{k2}A_{k1}+\cdots+x_n\sum_{k=1}^{n}a_{kn}A_{k1}=\sum_{k=1}^{n}b_kA_{k1}. \tag{8}$$

Според твърдение 3 коефициентът след x_1 е равен на Δ, а коефициентите след $x_2, \dots, x_n$ са равни на нула. Изразът в дясната страна на уравнението (8) е точно Δ_1. Така (8) приема вида

$$\Delta x_1 = \Delta_1.$$

Аналогично получаваме уравненията

$$\Delta x_2 = \Delta_2\,,\ \dots\,,\ \Delta x_n = \Delta_n.$$

Нека $\Delta \neq 0$. Тогава, ако системата (7) има решение, то е единствено и се получава по формулите

$$x_1 = \frac{\Delta_1}{\Delta}\,,\ x_2 = \frac{\Delta_2}{\Delta}\,,\ \dots\,,\ x_n = \frac{\Delta_n}{\Delta}. \tag{9}$$

Формулите (9) се наричат *формули на Крамер*.

Директната проверка показва, че получените стойности за x_i $(i = 1, 2, \dots, n)$ наистина са решение на системата (7). Действително, да заместим тези стойности например в първото уравнение на системата. Преобразуваме последователно лявата му страна:

$$\begin{aligned}
&a_{11}\frac{\Delta_1}{\Delta}+a_{12}\frac{\Delta_2}{\Delta}+\cdots+a_{1n}\frac{\Delta_n}{\Delta}\\
&=\frac{1}{\Delta}(a_{11}\Delta_1+a_{12}\Delta_2+\cdots+a_{1n}\Delta_n)=\frac{1}{\Delta}[a_{11}(b_1A_{11}+b_2A_{21}+\cdots+b_nA_{n1})\\
&\qquad+a_{12}(b_1A_{12}+b_2A_{22}+\cdots+b_nA_{n2})+\cdots+a_{1n}(b_1A_{1n}+b_2A_{2n}+\cdots+b_nA_{nn})]\\
&=\frac{1}{\Delta}[b_1(a_{11}A_{11}+a_{12}A_{12}+\cdots+a_{1n}A_{1n})\\
&\qquad+b_2(a_{11}A_{21}+a_{12}A_{22}+\cdots+a_{1n}A_{2n})+\cdots+b_n(a_{11}A_{n1}+a_{12}A_{n2}+\cdots+a_{1n}A_{nn})].
\end{aligned}$$

Според твърдение 3 коефициентът след b_1 в квадратните скоби е равен на Δ, а коефициентите след $b_2, \dots, b_n$ са равни на нула. Така лявата страна на първото уравнение приема вида $\frac{1}{\Delta}.b_1.\Delta = b_1$. По аналогичен начин се извършва проверката за останалите уравнения.

Следователно намерените стойности за неизвестните $x_1 = \frac{\Delta_1}{\Delta}$, $x_2 = \frac{\Delta_2}{\Delta}, \dots, x_n = \frac{\Delta_n}{\Delta}$ действително са решение на системата (7).

Така доказахме следната

Теорема 4. *Ако детерминантата на системата (7) е различна от нула, то тази система има единствено решение, което се получава по формулите на Крамер.*

§ 4. Умножение на матрици

Нека F е поле и $F_{m\times n}$ е множеството от всички $m \times n$ матрици с елементи от F. Вече познаваме действията събиране на матрици и умножение на матрица с число. (Да напомним, че с $\mathbf{0}$ означаваме нулевата матрица, а с $-A$ — противоположната матрица на матрицата A.) Тъй като тези действия се дефинират с помощта на съответните действия с числа, естествено е да очакваме, че много от свойствата на действията с числа се запазват и в този случай. Действително, нека λ и μ са произволни числа от F, а A, B и C са произволни матрици от $F_{m\times n}$. Директно от определенията се проверява, че са изпълнени следните свойства:

1) $(A+B)+C = A+(B+C)$;
2) $A + \mathbf{0} = A$;
3) $A + (-A) = \mathbf{0}$;
4) $A + B = B + A$;
5) $1.A = A$;
6) $(\lambda+\mu)A = \lambda A + \mu A$;
7) $\lambda(A+B) = \lambda A + \lambda B$;
8) $\lambda(\mu A) = (\lambda\mu)A$.

Сега ще въведем действието *умножение на матрици*, което на пръв поглед изглежда странно и се различава съществено от действието умножение на числа.*

Определение. *Нека $A = (a_{ij})_{m\times s}$, $B = (b_{ij})_{s\times n}$. Под произведение на матриците A и B ще разбираме матрицата $C = (c_{ij})_{m\times n}$, за която*

$$c_{ij} = \sum_{k=1}^{s} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj}.$$

Ще пишем $C = AB$.

Казваме, че елементът c_{ij} се получава като умножим i-ия ред на матрицата A с j-ия стълб на матрицата B. От определението е ясно, че можем да умножаваме две матрици само ако **броят на стълбовете на левия множител е равен на броя на редовете на десния множител**. В частност, ако A и B са квадратни матрици от един и същи ред, то и двете произведения AB и BA имат смисъл.

Ако $\lambda \in F$, E е единичната матрица от ред n, а A — произволна квадратна матрица от ред n, директно се проверява, че $\lambda A = (\lambda E)A = A(\lambda E)$. В частност $AE = EA = A$. Матриците от вида λE ще наричаме *скаларни матрици*.

Пример. Нека $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Тогава

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.0+1.0 & 0.0+1.1 \\ 0.0+0.0 & 0.0+0.1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = A,$$

$$BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0.0+0.0 & 0.1+0.0 \\ 0.0+1.0 & 0.1+1.0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \mathbf{0}.$$

От този пример забелязваме, че:

1. Умножението на матрици не е комутативно (в случая $AB \neq BA$).
2. Възможно е произведението на две ненулеви матрици да е равно на нулевата матрица ($B \neq \mathbf{0}$ и $A \neq \mathbf{0}$, но $BA = \mathbf{0}$).

*Едва в глава III ще стане ясно, че това определение е напълно естествено.

3. $B \neq E$, но $AB = A$.

Задача 1. *Да означим с* E_{ij} *матрицата от* $M_n(F)$, *която има единица на* (i,j)-*то място и нули на всички останали места (сравнете със задача 1 от § 04). Тези матрици се наричат* **матрични единици.** *Да се докаже, че*

$$E_{ij}E_{pq} = \delta_{jp}E_{iq} = \begin{cases} E_{iq}, & \text{ако } j = p \\ 0, & \text{ако } j \neq p. \end{cases}$$

Твърдение 1. *Нека* $\lambda \in F$, *а* A, B *и* C *са матрици, за които действията по-долу имат смисъл. Тогава са в сила следните свойства:*
а) $\lambda(AB) = (\lambda A)B = A(\lambda B)$;
б) $(A+B)C = AC + BC$; $C(A+B) = CA + CB$;
в) $(AB)C = A(BC)$.

Доказателство. Свойствата а) и б) се проверяват директно. Ще докажем само в) и то в случая, когато A, B и C са квадратни матрици от ред n, тъй като общият случай се третира аналогично (разликата е само в това, че когато A, B и C са квадратни матрици от ред n индексите се менят от 1 до n, а в общия случай се менят от 1 до "където им е позволено").

Да въведем следните означения:

$$\begin{aligned} A &= (a_{ij})_{n\times n}, \quad B = (b_{ij})_{n\times n}, \quad C = (c_{ij})_{n\times n}, \\ AB &= U = (u_{ij})_{n\times n}, \quad (AB)C = UC = S = (s_{ij})_{n\times n}, \\ BC &= V = (v_{ij})_{n\times n}, \quad A(BC) = AV = T = (t_{ij})_{n\times n}. \end{aligned}$$

Искаме да докажем, че $S = T$. Имаме

$$u_{il} = \sum_{k=1}^{n} a_{ik}b_{kl}, \qquad v_{kj} = \sum_{l=1}^{n} b_{kl}c_{lj}.$$

Сега от $S = UC$ и $T = AV$ получаваме

$$\begin{aligned} s_{ij} &= \sum_{l=1}^{n} u_{il}c_{lj} = \sum_{l=1}^{n}\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj}, \\ t_{ij} &= \sum_{k=1}^{n} a_{ik}v_{kj} = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ik}b_{kl}c_{lj}, \end{aligned}$$

т.е. $s_{ij} = t_{ij}$ $(i, j = 1, 2, \dots, n)$. Следователно $S = T$.

Твърдение 2. *Нека* A *и* B *са квадратни матрици от един и същи ред, а* A^t *и* B^t *са транспонираните им матрици. Тогава* $(AB)^t = B^tA^t$.

Доказателство. Да означим

$$\begin{aligned} A = (a_{ij})_{n\times n}, \quad B = (b_{ij})_{n\times n}, \quad AB = U = (u_{ij})_{n\times n}, \\ (AB)^t = S = (s_{ij})_{n\times n}, \quad B^tA^t = T = (t_{ij})_{n\times n}. \end{aligned}$$

Искаме да докажем, че $S = T$. Имаме

$$s_{ij} = u_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki}.$$

Елементът t_{ij} се получава като умножим i-ия **ред** на B^t, т.е. i-ия **стълб** на B, с j-ия стълб на A^t, т.е. j-ия **ред** на A, така че

$$t_{ij} = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Така $s_{ij} = t_{ij}$ $(i, j = 1, 2, \ldots, n)$. Следователно $S = T$.

§ 5. Умножение на детерминанти

Целта ни в този параграф е да докажем следната

Теорема 1. *Нека A и B са квадратни матрици от един и същи ред. Тогава е изпълнено равенството $\det(AB) = \det A.\det B$, т.е. детерминантата на произведението на две квадратни матрици от един и същи ред е равна на произведението на детерминантите им.*

За тази цел нека първо A и B са квадратни матрици не непременно от един и същи ред, например $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{k\times k}$. Да разгледаме следната помощна квадратна матрица от ред $n+k$:

$$D = \begin{pmatrix} A & 0 \\ * & B \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & 0 & 0 & \ldots & 0 \\ * & * & \ldots & * & b_{11} & b_{12} & \ldots & b_{1k} \\ * & * & \ldots & * & b_{21} & b_{22} & \ldots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & \ldots & * & b_{k1} & b_{k2} & \ldots & b_{kk} \end{pmatrix}.$$

На местата означени със $*$ стоят произволни числа.

Лема 2. *В сила е равенството* $\det D = \det A.\det B$.

Доказателство. Ще проведем индукция по n. Нека $n = 1$, т.е. $A = (a_{11})_{1\times 1}$. Тогава

$$D = \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ * & b_{11} & b_{12} & \ldots & b_{1k} \\ * & b_{21} & b_{22} & \ldots & b_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & b_{k1} & b_{k2} & \ldots & b_{kk} \end{pmatrix}.$$

Развивайки $\det D$ по първия ред, получаваме $\det D = a_{11} \det B = \det A.\det B$.

Нека $n > 1$. Да означим с $D_{11}, D_{12}, \ldots, D_{1n}$ квадратните матрици от $(n-1)$-ви ред, които се получават, като в матрицата A зачертаем съответно първия ред и първия стълб; първия ред и втория стълб; ... ; първия ред и n-ия стълб. Развивайки $\det D$ по първия ред, получаваме

$$\det D = a_{11}.\det \begin{pmatrix} D_{11} & 0 \\ * & B \end{pmatrix} - a_{12}.\det \begin{pmatrix} D_{12} & 0 \\ * & B \end{pmatrix} + \cdots + (-1)^{n+1} a_{1n}.\det \begin{pmatrix} D_{1n} & 0 \\ * & B \end{pmatrix}.$$

Индукционното предположение за матрицата $\begin{pmatrix} D_{1j} & 0 \\ * & B \end{pmatrix}$ ни дава

$$\det \begin{pmatrix} D_{1j} & 0 \\ * & B \end{pmatrix} = \det D_{1j} . \det B \qquad (j = 1, 2, \dots , n).$$

Тогава

$$\begin{aligned} \det D &= a_{11} . \det D_{11} . \det B - a_{12} . \det D_{12} . \det B + \dots + (-1)^{n+1} a_{1n} . \det D_{1n} . \det B \\ &= (a_{11} . \det D_{11} - a_{12} . \det D_{12} + \dots + (-1)^{n+1} a_{1n} . \det D_{1n}) \det B. \end{aligned}$$

Но изразът в скобите е точно развитието на $\det A$ по първия ред и значи $\det D = \det A . \det B$.

Доказателство на теорема 1. Нека $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times n}$, $AB = C = (c_{ij})_{n\times n}$ и D е следната квадратна матрица от ред $2n$:

$$D = \begin{pmatrix} A & 0 \\ -E_n & B \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}$$

Ще докажем, че $\det C = \det D$, откъдето, според лема 2, ще следва твърдението на теоремата.

Да прибавим към $(n+1)$-ия стълб на D първия стълб, умножен по b_{11}, после втория стълб, умножен по b_{21} и т. н., накрая n-ия стълб, умножен по b_{n1}. Аналогично, към $(n+2)$-ия стълб на D да прибавим линейната комбинация на първите n стълба с коефициенти съответно $b_{12}, b_{22}, \dots , b_{n2}$; … ; накрая към $2n$-ия стълб да прибавим линейната комбинация на първите n стълба с коефициенти съответно $b_{1n}, b_{2n}, \dots , b_{nn}$. Тогава, ако получената матрица е $(d_{ij})_{2n\times 2n}$, то за $j = 1, 2, \dots , n$ имаме

$$d_{1,n+j} = \sum_{k=1}^{n} a_{1k} b_{kj} = c_{1j},\ d_{2,n+j} = \sum_{k=1}^{n} a_{2k} b_{kj} = c_{2j},\ \dots ,\ d_{n,n+j} = \sum_{k=1}^{n} a_{nk} b_{kj} = c_{nj};$$
$$d_{n+1,n+j} = 0,\ d_{n+2,n+j} = 0,\ \dots ,\ d_{2n,n+j} = 0.$$

Така

$$\det D = \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & c_{11} & c_{12} & \dots & c_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} & c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & c_{n1} & c_{n2} & \dots & c_{nn} \\ -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Разменяме местата на първия и $(n+1)$-ия ред на получената матрица, после местата на втория и $(n+2)$-ия ред и т. н., накрая разменяме местата на n-ия и $2n$-ия ред. При всяка такава смяна детерминантата си сменя знака. Следователно

$$\det D = (-1)^n \det \begin{pmatrix} -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 0 & 0 & \dots & 0 \\ a_{11} & a_{12} & \dots & a_{1n} & c_{11} & c_{12} & \dots & c_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} & c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}.$$

Прилагаме лема 2 и получаваме

$$\det D = (-1)^n \det(-E_n).\det C = (-1)^n(-1)^n \det C = \det C.$$

С това теоремата е доказана.

Правилото за умножение на детерминанти, получено от равенството $\det(AB) = \det A.\det B$ се нарича правило "ред по стълб". Теоремата за транспонирана детерминанта (теорема 2 от § 1) ни дава още три равенства: $\det(AB^t) = \det A.\det B^t = \det A.\det B$, аналогично $\det(A^tB) = \det A.\det B$ и $\det(A^tB^t) = \det A.\det B$. От тези равенства получаваме още три правила за умножение на детерминанти: съответно "ред по ред", "стълб по стълб" и "стълб по ред".

§ 6. Обратна матрица

Ще казваме, че една квадратна матрица A е *неособена*, ако $\det A \neq 0$. В противен случай ще я наричаме *особена*. От теоремата за умножение на детерминанти следва, че произведението на неособени матрици също е неособена матрица.

Определение. *Една квадратна матрица A ще наричаме обратима, ако съществува квадратна матрица A' от същия ред, такава че $AA' = A'A = E$. Матрицата A' ще наричаме обратна на матрицата A.*

Ако A е обратима матрица, то тя притежава единствена обратна матрица. Действително, ако A'' е такава матрица, че $AA'' = A''A = E$, то

$$A'AA'' = A'(AA'') = A'E = A',$$
$$A'AA'' = (A'A)A'' = EA'' = A''$$

и значи $A' = A''$. Тази единствена матрица ще бележим с A^{-1}.

Ако A е обратима матрица и A^{-1} е обратната ѝ, от теоремата за умножение на детерминанти имаме $\det A.\det(A^{-1}) = \det E = 1$ и значи $\det(A^{-1}) = (\det A)^{-1}$.

Ще отбележим, че произведение на обратими матрици също е обратима матрица, при това $(AB)^{-1} = B^{-1}A^{-1}$.

Теорема 1. *Една квадратна матрица е обратима тогава и само тогава, когато е неособена.*

Доказателство. Ако A е обратима матрица, от равенството $\det A.\det(A^{-1}) = 1$ следва $\det A \neq 0$, т.е. A е неособена.

Нека сега A е неособена и $A = (a_{ij})_{n\times n}$. Да означим с X матрицата

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

където A_{ij} е адюнгираното количество на елемента a_{ij} $(i, j = 1, 2, \dots, n)$. Ще подчертаем, че адюнгираните количества на елементите от i-ия **ред** на A стоят в **i-ия стълб** на X. Нека $AX = (c_{ij})_{n\times n}$. От правилото за умножение на матрици следва, че c_{ij} е развитието на $\det A$ по i-ия ред с адюнгираните количества на елементите от j-ия ред. Знаем тогава, че $c_{ij} = \delta_{ij}\det A$ (твърдение 3 от § 3). Това означава, че $AX = \det A.E$. Нека $Y = \dfrac{1}{\det A}X$. Имаме $AY = A\left(\dfrac{1}{\det A}X\right) = \dfrac{1}{\det A}AX = \dfrac{1}{\det A}(\det A.E) = E$. Аналогично се проверява, че $YA = E$. Така Y е обратната матрица на матрицата A.

Накрая ще отбележим, че с помощта на обратна матрица могат да се решават матрични уравнения. Нека $A = (a_{ij})_{n\times n}$, $X = (x_{ij})_{n\times m}$, $B = (b_{ij})_{n\times m}$. Да разгледаме матричното уравнение $AX = B$. Ако A е обратима матрица, като умножим това уравнение отляво с A^{-1}, получаваме уравнението $X = A^{-1}B$, което е еквивалентно на горното. Аналогично можем да решаваме уравнения от вида $YA = B$. В този случай $Y = BA^{-1}$.

Да разгледаме частния случай, когато $m = 1$. Нека

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Тогава уравнението $AX = B$ е матричен запис на системата

$$\left|\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{array}\right.$$

Равенството $X = A^{-1}B$ представлява съкратен запис на формулите на Крамер.

Глава II. Линейни пространства

Всички линейни пространства във втора и трета глава (ако не е казано нещо друго) ще бъдат над фиксирано числово поле F.

§ 7. Линейни пространства

Понятието линейно пространство е типичен пример на абстрактна алгебрична структура с въведени в нея операции, подчинени на определени свойства. То е обобщение на познатите от училищния курс по математика едномерно пространство (правата), двумерно пространство (равнината) и тримерно пространство.

Определение. *Нека F е поле и V е непразно множество, чиито елементи ще наричаме вектори. Нека във V са въведени следните операции:*

I. Събиране на вектори: на всеки два вектора $\boldsymbol{a}$ и $\boldsymbol{b}$ от V е съпоставен вектор $\boldsymbol{a}+\boldsymbol{b}$ също от V, който ще наричаме сума на $\boldsymbol{a}$ и $\boldsymbol{b}$.

II. Умножение на вектор с число: на всеки вектор $\boldsymbol{a}$ от V и на всяко число λ от F е съпоставен вектор $\lambda\boldsymbol{a}$ от V.

Ще казваме, че V е линейно пространство над полето F, ако тези операции удовлетворяват следните свойства (аксиоми):

1. Събирането е асоциативно, т. е. за всеки три вектора $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ от V е в сила $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$. Този елемент ще бележим с $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$ без скоби.

2. Събирането е комутативно, т. е. за всеки два вектора $\boldsymbol{a}$ и $\boldsymbol{b}$ от V е изпълнено $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$.

3. Съществува неутрален елемент относно операцията събиране, т. е. такъв вектор $\boldsymbol{0}$ от V, че $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$ за всеки вектор $\boldsymbol{a}$ от V. По-долу ще видим, че този неутрален елемент е единствен и ще го наричаме нулев вектор.

4. За всеки вектор $\boldsymbol{a}$ от V съществува вектор $\boldsymbol{a}'$ от V, такъв че $\boldsymbol{a}+\boldsymbol{a}'=\boldsymbol{0}$. По-долу ще видим, че векторът $\boldsymbol{a}'$ еднозначно се определя от $\boldsymbol{a}$. Ще го наричаме противоположен вектор на $\boldsymbol{a}$ и ще го бележим с $-\boldsymbol{a}$.

5. За всеки вектор $\boldsymbol{a}$ от V е изпълнено $1.\boldsymbol{a}=\boldsymbol{a}$.

6. За всеки две числа λ и μ от F и за всеки вектор $\boldsymbol{a}$ от V е изпълнено $(\lambda+\mu)\boldsymbol{a}=\lambda\boldsymbol{a}+\mu\boldsymbol{a}$.

7. За всяко число λ от F и за всеки два вектора $\boldsymbol{a}$ и $\boldsymbol{b}$ от V е изпълнено $\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$.

8. За всеки две числа λ и μ от F и за всеки вектор $\boldsymbol{a}$ от V е изпълнено $\lambda(\mu\boldsymbol{a})=(\lambda\mu)\boldsymbol{a}$.

Примери на линейни пространства.

1. Да разгледаме множеството $F_{m\times n}$ от всички $m \times n$ матрици с елементи от F. От свойствата, формулирани в началото на § 4 следва, че $F_{m\times n}$ е линейно пространство над F.

2. Важен частен случай на пример 1 е множеството F^n от всички наредени n-орки с елементи от F.

При фиксирана координатна система в равнината (пространството) можем да отъждествяваме $\mathbb{R}^2$ ($\mathbb{R}^3$) с множеството от всички вектори в равнината (пространството), чието начало съвпада с началото на координатната система, разгледани с обичайните операции събиране на вектори и умножение на вектор с число. Тези множества са линейни пространства над $\mathbb{R}$.

3. Да означим с $F[x]$ множеството от всички полиноми с коефициенти от F, разгледани с обичайните операции събиране на полиноми и умножение на полином с число. От училищния курс по математика знаем, че тези операции притежават необходимите свойства, така че $F[x]$ е линейно пространство над F. Също така, множеството $F^{n+1}[x]$, състоящо се от всички полиноми с коефициенти от F и степен, ненадминаваща n, е линейно пространство над F. (Множеството от всички полиноми от степен, равна на n не е линейно пространство, тъй като сумата на два такива полинома може да е полином от по-ниска степен.)

Следствия от аксиомите.

1. Единственост на нулевия вектор.

Нека векторите $\mathbf{0}'$ и $\mathbf{0}''$ удовлетворяват третата аксиома за линейно пространство. Тогава $\mathbf{0}'+\mathbf{0}''=\mathbf{0}'$ (защото $\mathbf{0}''$ е нулев вектор) и $\mathbf{0}'+\mathbf{0}''=\mathbf{0}''$ (защото $\mathbf{0}'$ е нулев вектор). Следователно $\mathbf{0}'=\mathbf{0}''$.

2. Единственост на противоположния вектор на даден вектор.

Нека $\boldsymbol{a} \in V$ и $\boldsymbol{a}'$ и $\boldsymbol{a}''$ са противоположни вектори на $\boldsymbol{a}$. Имаме $\boldsymbol{a}'+\boldsymbol{a}+\boldsymbol{a}'' = (\boldsymbol{a}'+\boldsymbol{a})+\boldsymbol{a}'' = \mathbf{0}+\boldsymbol{a}'' = \boldsymbol{a}''$ и $\boldsymbol{a}'+\boldsymbol{a}+\boldsymbol{a}'' = \boldsymbol{a}'+(\boldsymbol{a}+\boldsymbol{a}'') = \boldsymbol{a}'+\mathbf{0} = \boldsymbol{a}'$. Следователно $\boldsymbol{a}'=\boldsymbol{a}''$.

3. За всеки вектор $\boldsymbol{a}$ от V е в сила $0.\boldsymbol{a}=\mathbf{0}$.

Имаме $\boldsymbol{a}+0.\boldsymbol{a} = 1.\boldsymbol{a}+0.\boldsymbol{a} = (1+0)\boldsymbol{a} = 1.\boldsymbol{a} = \boldsymbol{a}$, т.е. $\boldsymbol{a}+0.\boldsymbol{a}=\boldsymbol{a}$. Като прибавим към двете страни на това равенство вектора $-\boldsymbol{a}$, получаваме $0.\boldsymbol{a}=\mathbf{0}$.

4. За всяко число λ от F е в сила $\lambda\mathbf{0}=\mathbf{0}$.

В равенството $\lambda(\mu\boldsymbol{a}) = (\lambda\mu)\boldsymbol{a}$, като положим $\mu=0$ и използваме предното свойство, получаваме $\lambda.\mathbf{0} = 0.\boldsymbol{a} = \mathbf{0}$.

5. За всеки вектор $\boldsymbol{a}$ от V е изпълнено $(-1)\boldsymbol{a} = -\boldsymbol{a}$.

Имаме $\boldsymbol{a}+(-1)\boldsymbol{a} = 1.\boldsymbol{a}+(-1)\boldsymbol{a} = (1-1)\boldsymbol{a} = 0.\boldsymbol{a} = \mathbf{0}$. Следователно векторът $(-1)\boldsymbol{a}$ е противоположният вектор на $\boldsymbol{a}$.

6. Ако $\lambda \in F$, $\boldsymbol{a} \in V$ и $\lambda\boldsymbol{a}=\mathbf{0}$, то или $\lambda=0$, или $\boldsymbol{a}=\mathbf{0}$.

Действително, ако $\lambda \neq 0$, умножаваме горното равенство с λ^{-1} и получаваме $1.\boldsymbol{a}=\mathbf{0}$, т.е. $\boldsymbol{a}=\mathbf{0}$.

7. Ако $\boldsymbol{a}$ и $\boldsymbol{b}$ са вектори от V, уравнението $\boldsymbol{a}+\boldsymbol{x}=\boldsymbol{b}$ има единствено решение във V.

Очевидно $\boldsymbol{x} = (-\boldsymbol{a})+\boldsymbol{b}$ е решение на уравнението. Ако $\boldsymbol{x}'$ е вектор, за който $\boldsymbol{a}+\boldsymbol{x}'=\boldsymbol{b}$, като прибавим към двете страни на това равенство вектора $-\boldsymbol{a}$, получаваме $\boldsymbol{x}' = (-\boldsymbol{a})+\boldsymbol{b}$, т.е. $\boldsymbol{x}'=\boldsymbol{x}$.

Единственото решение на уравнението $\boldsymbol{a}+\boldsymbol{x}=\boldsymbol{b}$ по-нататък ще означаваме с $\boldsymbol{b}-\boldsymbol{a}$, т.е. вместо $\boldsymbol{b}+(-\boldsymbol{a})$ ще пишем $\boldsymbol{b}-\boldsymbol{a}$.

З а б е л е ж к а. Определението за линейно пространство изглежда дълго, а следствията от аксиомите на пръв поглед будят недоумение (едва ли човек в началото изпитва необходимост да доказва, че $(-1)\boldsymbol{a} = -\boldsymbol{a}$). В действителност, нито определението, нито следствията от аксиомите е необходимо да се помнят буквално наизуст. Достатъчно е да се знае, че въведените операции се подчиняват на естествени и "разумни" свойства и с тях може да се работи почти както с числа.

Определение. *Нека* $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ *са вектори от* V, *а* $\lambda_1, \lambda_2, \dots, \lambda_n$ *са числа от* F. *Вектора* $\lambda_1\boldsymbol{a}_1 + \lambda_2\boldsymbol{a}_2 + \dots + \lambda_n\boldsymbol{a}_n$ *ще наричаме линейна комбинация на векторите* $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ *с коефициенти съответно* $\lambda_1, \lambda_2, \dots, \lambda_n$.

Определение. *Нека* W *е непразно подмножество на линейното пространство* V. *Ще казваме, че* W *е подпространство на* V, *ако всяка линейна комбинация на вектори от* W *също принадлежи на* W.

Всяко подпространство е също линейно пространство относно въведените операции.

Задача 1. *Да се докаже, че непразното подмножество* W *на* V *е подпространство на* V *тогава и само тогава, когато се изпълняват следните две условия:*

1) *Сумата на всеки два вектора от* W *също принадлежи на* W.

2) *Произведението на всеки вектор от* W *с число от* F *принадлежи на* W.

(Казваме, че W *е затворено относно операциите събиране на вектори и умножение на вектор с число.)*

Задача 2. *Да се докаже, че всяко подпространство съдържа нулевия вектор.*

Задача 3. *Да се докаже, че сечение на произволна фамилия от подпространства също е подпространство.*

П р и м е р и   н а   п о д п р о с т р а н с т в а.

0. Очевидно множествата $\{\boldsymbol{0}\}$ и V са подпространства на V.

1. Нека V е множеството от всички вектори в равнината, чието начало съвпада с началото на фиксирана координатна система (разгледани с обичайните операции събиране на вектори и умножение на вектор с число). Това множество е линейно пространство над $\mathbb{R}$. Нека W е множеството от всички вектори от V, чийто край лежи върху фиксирана права, минаваща през началото на координатната система. Тогава W е подпространство на V.

2. Нека $V = F^n$ и k е естествено число, ненадминаващо n. Да означим с W множеството от тези наредени n-орки от F^n, чиито първи k елемента са произволни числа от F, а останалите $n-k$ са равни на нула. Директно от определението се проверява, че W е подпространство на V. Фактически можем да отъждествяваме W с F^k.

3. Линейното пространство $F^{n+1}[x]$ е подпространство на линейното пространство $F^{n+2}[x]$. Също така, за всяко естествено число n $F^{n+1}[x]$ е подпространство на $F[x]$.

Определение. *Нека* A *е произволно непразно подмножество на линейното пространство* V. *Множеството* $l(A)$, *състоящо се от всички линейни комбинации на елементи от* A *с коефициенти от* F *ще наричаме линейна обвивка на множеството* A.

Очевидно $l(V) = V$ и $l(\boldsymbol{0}) = \{\boldsymbol{0}\}$.

Задача 4. *Да се докаже, че* $l(A)$ *е подпространство на* V, *съдържащо* A *и* $l(A) = A$ *тогава и само тогава, когато* A *е подпространство на* V.

Задача 5. *Да се докаже, че* $l(A)$ *съвпада със сечението на всички подпространства на* V, *съдържащи множеството* A *(т.е.* $l(A)$ *е най-малкото (относно включване) подпространство на* V, *съдържащо* A*).*

Задача 6. *Нека* $V = F^2$ *и* $\boldsymbol{e}_1 = (1, 0)$, $\boldsymbol{e}_2 = (0, 1)$. *Да се докаже, че* $l(\boldsymbol{e}_1) = \{(\lambda, 0) \mid \lambda \in F\}$, $l(\boldsymbol{e}_1, \boldsymbol{e}_2) = F^2$.

§ 8. Линейна зависимост и независимост

Едни от най-важните понятия в линейната алгебра са понятията линейна независимост и линейна зависимост на вектори.

Определение. (!) *Нека $a_1, a_2, \dots, a_n$ са вектори от линейното пространство V над полето F. Ще казваме, че тези вектори са линейно независими над F (или още, че системата вектори $a_1, a_2, \dots, a_n$ е линейно независима над F), ако от това, че някоя линейна комбинация на векторите $a_1, a_2, \dots, a_n$ с коефициенти от F е равна на нулевия вектор следва, че всички коефициенти в тази линейна комбинация са равни на нула. За една безкрайна система вектори от V ще казваме, че е линейно независима над F, ако всяка нейна крайна подсистема е линейно независима над F.*

Ще казваме, че системата вектори $a_1, a_2, \dots, a_n$ от V е линейно зависима над F, ако съществуват числа $\lambda_1, \lambda_2, \dots, \lambda_n$ от F, не всички от които са равни на нула, но $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n = 0$.

(По-нататък, ако полето F се подразбира, ще пропускаме досадното уточнение "над F".)

Ясно е, че една система от вектори е линейно независима точно когато не е линейно зависима и обратно — една система от вектори е линейно зависима точно когато не е линейно независима.

Задача 1. *Да се докаже, че векторите $e_1 = (1,0)$ и $e_2 = (0,1)$ от F^2 са линейно независими. По-общо, векторите $e_1 = (1,0,0,\dots,0)$, $e_2 = (0,1,0,\dots,0)$, $\dots$, $e_n = (0,0,0,\dots,1)$ от F^n са линейно независими.*

Основни свойства на понятията линейна зависимост и линейна независимост на вектори.

1. *Един вектор е линейно независим тогава и само тогава, когато е ненулев.*

Това свойство следва от следствие 6 от аксиомите за линейно пространство.

2. *Всяка подсистема на линейно независима система от вектори е също линейно независима.*

Нека $A = \{a_1, a_2, \dots, a_n\}$ е линейно независима система от вектори и $B = \{a_1, a_2, \dots, a_k\}$, $k \le n$. Да допуснем, че системата B е линейно зависима и нека $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_k a_k = 0$, като например $\lambda_1 \neq 0$. Тогава $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_k a_k + 0.a_{k+1} + \dots + 0.a_n = 0$, което противоречи на линейната независимост на системата A. Следователно B е линейно независима система вектори.

От доказаното свойство следва, че всяка надсистема на линейно зависима система от вектори също е линейно зависима.

3. *Ако една система от вектори съдържа нулевия вектор или два пропорционални вектора, тя е линейно зависима.*

Действително, нека $A = \{a_1 = 0, a_2, \dots, a_n\}$. Тогава $1.a_1 + 0.a_2 + \dots + 0.a_n = 0$ и значи системата A е линейно зависима. Ако пък $a_2 = \lambda a_1$, то $\lambda a_1 - 1.a_2 + 0.a_3 + \dots + 0.a_n = 0$ и отново системата A е линейно зависима.

4. *Една система A от поне два вектора е линейно зависима тогава и само тогава, когато поне един вектор от A е линейна комбинация на останалите вектори от A.*

Нека $A = \{a_1, a_2, \dots, a_n\}$ е линейно зависима система от вектори и $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_n a_n = 0$, като например $\lambda_1 \neq 0$. Тогава $a_1 = -\dfrac{\lambda_2}{\lambda_1} a_2 - \dots - \dfrac{\lambda_n}{\lambda_1} a_n$. Обратно, ако $a_1 = \mu_2 a_2 + \dots + \mu_n a_n$, то $1.a_1 - \mu_2 a_2 - \dots - \mu_n a_n = 0$ и значи системата от вектори A е линейно зависима.

Ще завършим този параграф с т. нар. *основна лема* на линейната алгебра.

Лема 1. *Нека V е линейно пространство и са дадени две системи вектори от V:*

$$A = \{ \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \} \quad \text{и} \quad B = \{ \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_k \}.$$

Нека всеки вектор от системата B е линейна комбинация на векторите от системата A. Тогава, ако $k > n$, то векторите от системата B са линейно зависими. (Казано по-кратко, ако повече на брой вектори се изразяват линейно чрез по-малко на брой, то повечето на брой вектори са линейно зависими.)

Доказателство. Ще проведем индукция по n. Нека $n = 1$, т.е. $A = \{ \boldsymbol{a}_1 \}$. Тогава системата вектори B или съдържа нулевия вектор, или съдържа два пропорционални вектора (всички вектори от B са кратни на $\boldsymbol{a}_1$). Според свойство 3 от по-горе векторите от B са линейно зависими.

Нека $n > 1$. По условие съществуват числа λ_{ij} $(i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, n)$, такива че

$$\begin{aligned}
\lambda_{11}\boldsymbol{a}_1 + \cdots + \quad \lambda_{1,n-1}\boldsymbol{a}_{n-1} + \quad \lambda_{1n}\boldsymbol{a}_n &= \boldsymbol{b}_1,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots&\\
\lambda_{k-1,1}\boldsymbol{a}_1 + \cdots + \lambda_{k-1,n-1}\boldsymbol{a}_{n-1} + \lambda_{k-1,n}\boldsymbol{a}_n &= \boldsymbol{b}_{k-1},\\
\lambda_{k1}\boldsymbol{a}_1 + \cdots + \quad \lambda_{k,n-1}\boldsymbol{a}_{n-1} + \quad \lambda_{kn}\boldsymbol{a}_n &= \boldsymbol{b}_k.
\end{aligned}$$

Ако $\boldsymbol{b}_k = \boldsymbol{0}$, системата вектори B е линейно зависима (свойство 3). Нека $\boldsymbol{b}_k \neq \boldsymbol{0}$ и например $\lambda_{kn} \neq 0$. Елиминираме $\boldsymbol{a}_n$ от първите $k-1$ равенства като умножим последното равенство с $-\dfrac{\lambda_{in}}{\lambda_{kn}}$ и го прибавим към i-тото равенство за всяко $i = 1, 2, \ldots, k-1$. Получаваме равенствата

$$\left(\lambda_{11} - \lambda_{k1}\frac{\lambda_{1n}}{\lambda_{kn}}\right)\boldsymbol{a}_1 + \cdots + \quad \left(\lambda_{1,n-1} - \lambda_{k,n-1}\frac{\lambda_{1n}}{\lambda_{kn}}\right)\boldsymbol{a}_{n-1} = \boldsymbol{b}_1 - \frac{\lambda_{1n}}{\lambda_{kn}}\boldsymbol{b}_k,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\left(\lambda_{k-1,1} - \lambda_{k1}\frac{\lambda_{k-1,n}}{\lambda_{kn}}\right)\boldsymbol{a}_1 + \cdots + \left(\lambda_{k-1,n-1} - \lambda_{k,n-1}\frac{\lambda_{k-1,n}}{\lambda_{kn}}\right)\boldsymbol{a}_{n-1} = \boldsymbol{b}_{k-1} - \frac{\lambda_{k-1,n}}{\lambda_{kn}}\boldsymbol{b}_k.$$

Така векторите $\boldsymbol{b}_i - \dfrac{\lambda_{in}}{\lambda_{kn}}\boldsymbol{b}_k$ $(i = 1, 2, \ldots, k-1)$ са $k-1$ на брой и всеки от тях е линейна комбинация на векторите $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{n-1}$. Освен това, $k - 1 > n - 1$. Според индукционното предположение системата вектори $\boldsymbol{b}_i - \dfrac{\lambda_{in}}{\lambda_{kn}}\boldsymbol{b}_k$ $(i = 1, 2, \ldots, k-1)$ е линейно зависима, т.е. съществуват числа $\mu_1, \mu_2, \ldots, \mu_{k-1}$, не всички от които са равни на нула и такива, че

$$\mu_1\left(\boldsymbol{b}_1 - \frac{\lambda_{1n}}{\lambda_{kn}}\boldsymbol{b}_k\right) + \cdots + \mu_{k-1}\left(\boldsymbol{b}_{k-1} - \frac{\lambda_{k-1,n}}{\lambda_{kn}}\boldsymbol{b}_k\right) = \boldsymbol{0}.$$

Оттук получаваме равенството

$$\mu_1\boldsymbol{b}_1 + \cdots + \mu_{k-1}\boldsymbol{b}_{k-1} + *\boldsymbol{b}_k = \boldsymbol{0}$$

(няма да се интересуваме какъв е коефициентът пред $\boldsymbol{b}_k$). Следователно векторите $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-1}, \boldsymbol{b}_k$ са линейно зависими.

§ 9. Базис, размерност, координати

Лема 1. *Нека V е линейно пространство и $a_1, a_2, \ldots, a_s$ са линейно независими вектори от V. Ако a е вектор от V, който не принадлежи на $l(a_1, a_2, \ldots, a_s)$, то векторите $a_1, a_2, \ldots, a_s$; a продължават да бъдат линейно независими.*

Доказателство. Да допуснем, че системата от вектори $a_1, a_2, \ldots, a_s$; a е линейно зависима и нека

$$\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_s a_s + \lambda a = 0,$$

където поне един от коефициентите $\lambda_1, \lambda_2, \ldots, \lambda_s$; λ е различен от нула. Ако $\lambda = 0$, получаваме линейна зависимост на векторите $a_1, a_2, \ldots, a_s$, което е противоречие. Ако $\lambda \neq 0$, имаме $a = -\dfrac{\lambda_1}{\lambda}a_1 - \dfrac{\lambda_2}{\lambda}a_2 - \cdots - \dfrac{\lambda_s}{\lambda}a_s \in l(a_1, a_2, \ldots, a_s)$, което отново е противоречие. Следователно системата вектори $a_1, a_2, \ldots, a_s$; a е линейно независима.

Определение. *Нека V е ненулево линейно пространство над полето F и B е непразно подмножество на V. Ще казваме, че B е базис на V над F (или само базис на V, ако F се подразбира), ако:*

1) *B е линейно независима система от вектори;*

2) *всеки вектор от V е линейна комбинация на векторите от B с коефициенти от F, т. е. $V = l(B)$.*

Примери.

1. От задача 1 от § 8 и задача 2 от § 04 следва, че векторите $e_1 = (1, 0, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $e_n = (0, 0, 0, \ldots, 1)$ образуват базис на F^n над F.

2. Полиномите $1, x, x^2, \ldots, x^n$ образуват базис на $F^{n+1}[x]$. Полиномите $1, x, x^2, \ldots$ образуват базис на $F[x]$.

Определение. *Едно линейно пространство V над полето F ще наричаме крайномерно над F (или само крайномерно, ако F се подразбира), ако съществуват краен брой вектори $b_1, b_2, \ldots, b_k \in V$, такива че $V = l(b_1, b_2, \ldots, b_k)$. В противен случай ще казваме, че V е безкрайномерно над F.*

Пространствата F^n и $F^{n+1}[x]$ са крайномерни над F, а пространството $F[x]$ е безкрайномерно над F.

Теорема 2. *Всяко ненулево крайномерно пространство над полето F притежава краен базис. По-точно, ако $V = l(b_1, b_2, \ldots, b_k)$, то съществува подсистема на системата вектори $b_1, b_2, \ldots, b_k$, която е базис на V над F. Така едно ненулево линейно пространство е крайномерно тогава и само тогава, когато притежава краен базис.*

Доказателство. От $V \neq \{0\}$ следва, че поне един от векторите $b_1, b_2, \ldots, b_k$ е различен от нулевия вектор. Нека например $b_1 \neq 0$. Ако $V = l(b_1)$, то b_1 е базис на V. Нека $V \supsetneq l(b_1)$. Тогава поне един от векторите $b_2, \ldots, b_k$, например b_2, не принадлежи на $l(b_1)$ (в противен случай $V = l(b_1, b_2, \ldots, b_k) = l(b_1)$). Според лема 1 векторите b_1 и b_2 са линейно независими. Ако $V = l(b_1, b_2)$, то b_1 и b_2 образуват базис на V. В противен случай, например $b_3 \notin l(b_1, b_2)$ и тогава векторите b_1, b_2, b_3 са линейно независими. Продължавайки по същия начин, след краен брой стъпки избираме вектори $b_1, b_2, \ldots, b_n$, $n \leq k$, които са линейно независими и $V = l(b_1, b_2, \ldots, b_n)$. Следователно тези вектори образуват базис на V.

Теорема 3. *Всеки два базиса на ненулевото крайномерно пространство V над полето F съдържат равен брой вектори.*

Доказателство. Нека $a_1, a_2, \dots, a_n$ и $b_1, b_2, \dots, b_k$ са два базиса на V. Тогава всеки вектор от втория базис се изразява линейно чрез векторите от първия базис. Тъй като векторите $b_1, b_2, \dots, b_k$ са линейно независими, от основната лема (лема 1 от § 8) следва, че $k \leq n$. Аналогично $n \leq k$. Следователно $n = k$.

Определение. *Броят на векторите в кой да е базис на ненулевото крайномерно пространство V над полето F ще наричаме размерност на V над F и ще го бележим с* $\dim_F V$ *или само* $\dim V$*, ако F се подразбира. По определение размерността на нулевото пространство е равна на нула. Ако V е безкрайномерно пространство, ще пишем* $\dim V = \infty$.

От предните примери следва, че $\dim F^n = n$, $\dim F^{n+1}[x] = n+1$, $\dim F[x] = \infty$.

Теорема 4. *Нека V е линейно пространство над полето F. Тогава*

а) *V е крайномерно и $\dim V = n$ тогава и само тогава, когато във V съществуват n на брой линейно независими вектора и всеки $n+1$ на брой вектора са линейно зависими. В този случай всеки n на брой линейно независими вектора от V са базис на V;*

б) *V е безкрайномерно тогава и само тогава, когато за всяко естествено число n във V има n на брой линейно независими вектора.*

Доказателство. а) Нека V е крайномерно и $\dim V = n$. Тогава V притежава базис $a_1, a_2, \dots, a_n$, състоящ се от n на брой вектора. Нека $b_1, b_2, \dots, b_{n+1}$ е произволна система от $n+1$ на брой вектора от V. Тези вектори се изразяват линейно чрез базисните вектори $a_1, a_2, \dots, a_n$. От основната лема (лема 1 от § 8) следва, че векторите $b_1, b_2, \dots, b_{n+1}$ са линейно зависими.

Обратно, нека $a_1, a_2, \dots, a_n$ са линейно независими вектори от V и всеки $n+1$ на брой вектора от V са линейно зависими. Нека a е произволен вектор от V. Ако $a \notin l(a_1, a_2, \dots, a_n)$, според лема 1 векторите $a_1, a_2, \dots, a_n; a$ биха били линейно независими; противоречие. Следователно $a \in l(a_1, a_2, \dots, a_n)$. Така векторите $a_1, a_2, \dots, a_n$ са линейно независими и всеки вектор a от V е тяхна линейна комбинация. Следователно тези вектори са базис на V и значи V е крайномерно и $\dim V = n$.

Накрая, нека $\dim V = n$ и $b_1, b_2, \dots, b_n$ е произволна система от n на брой линейно независими вектори от V. Ако съществува вектор от V извън $l(b_1, b_2, \dots, b_n)$, прилагайки лема 1 бихме получили $n+1$ на брой линейно независими вектора във V, което противоречи на $\dim V = n$. Следователно $V = l(b_1, b_2, \dots, b_n)$ и значи тези вектори са базис на V.

б) Нека V е безкрайномерно и n е произволно естествено число. Да допуснем, че във V няма n линейно независими вектора (т. е. всеки n вектора във V са линейно зависими). Тогава от подусловие а) следва, че $\dim V < n$, противоречие.

Обратно, ако за всяко естествено число n във V има n на брой линейно независими вектора, отново от подусловие а) следва, че не е възможно V да е крайномерно пространство, т. е. V е безкрайномерно.

Твърдение 5. *Всяка линейно независима система вектори в крайномерно пространство V може да се допълни до базис на V.*

Доказателство. Нека $b_1, b_2, \dots, b_s$ са линейно независими вектори от V. Ако $V = l(b_1, b_2, \dots, b_s)$, то тези вектори са базис на V. В противен случай съществува вектор b_{s+1} от V, такъв че $b_{s+1} \notin l(b_1, b_2, \dots, b_s)$. Според лема 1 векторите $b_1, b_2, \dots, b_s, b_{s+1}$ са линейно независими. Ако $V = l(b_1, b_2, \dots, b_s, b_{s+1})$, то тези вектори са базис на V. В противен случай съществува вектор b_{s+2} от V, такъв че $b_{s+2} \notin l(b_1, b_2, \ldots, b_s, b_{s+1})$. Продължавайки по този начин (процесът не може да бъде безкраен, тъй като $\dim V < \infty$), достигаме до система вектори $b_1, \ldots, b_s, b_{s+1}, \ldots, b_n$, които са линейно независими и $V = l(b_1, \ldots, b_s, b_{s+1}, \ldots, b_n)$. Следователно тези вектори са базис на V.

Задача 1. *Нека V е крайномерно пространство и W е подпространство на V. Тогава W също е крайномерно и $\dim W \leq \dim V$, като $\dim W = \dim V$ тогава и само тогава, когато $W = V$.*

Твърдение 6. *Нека V е ненулево крайномерно пространство над полето F. Една система вектори от V е базис на V тогава и само тогава, когато всеки вектор от V се представя по единствен начин като линейна комбинация на векторите от тази система.*

Доказателство. Нека $b_1, b_2, \ldots, b_n$ е базис на V, $v \in V$ и

$$v = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n,$$
$$v = \mu_1 b_1 + \mu_2 b_2 + \cdots + \mu_n b_n.$$

Като извадим горните две равенства, получаваме

$$0 = (\lambda_1 - \mu_1) b_1 + (\lambda_2 - \mu_2) b_2 + \cdots + (\lambda_n - \mu_n) b_n.$$

От линейната независимост на векторите $b_1, b_2, \ldots, b_n$ получаваме $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$, $\ldots$, $\lambda_n = \mu_n$. Следователно векторът v се записва по единствен начин като линейна комбинация на векторите $b_1, b_2, \ldots, b_n$.

Обратно, нека векторите $b_1, b_2, \ldots, b_n$ са такива, че всеки вектор от V се представя по единствен начин като тяхна линейна комбинация. Тогава $V = l(b_1, b_2, \ldots, b_n)$ и за да докажем, че тези вектори образуват базис на V, остава да проверим, че те са линейно независими. Нека $\lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n = 0$ ($\lambda_i \in F$; $i = 1, 2, \ldots, n$). Също така, очевидно имаме $0.b_1 + 0.b_2 + \cdots + 0.b_n = 0$. Тъй като нулевият вектор се записва по единствен начин като линейна комбинация на векторите $b_1, b_2, \ldots, b_n$, то $\lambda_1 = 0$, $\lambda_2 = 0, \ldots, \lambda_n = 0$ и значи тези вектори са линейно независими.

Задача 2. *Да се докаже, че векторите $b_1, b_2, \ldots, b_n$ образуват базис на крайномерното пространство V тогава и само тогава, когато $V = l(b_1, b_2, \ldots, b_n)$ и нулевият вектор се представя по единствен начин като линейна комбинация на тези вектори.*

Твърдение 6 ни позволява да дадем следното

Определение. *Нека V е линейно пространство над полето F, $\dim V = n$ и $b_1, b_2, \ldots, b_n$ е фиксиран базис на V. Нека $v \in V$ и*

$$v = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n \quad (\lambda_i \in F,\ i = 1, 2, \ldots, n).$$

Еднозначно определените (от базиса $b_1, b_2, \ldots, b_n$) числа $\lambda_1, \lambda_2, \ldots, \lambda_n$ ще наричаме **координати на v в базиса** $b_1, b_2, \ldots, b_n$.

Задача 3. *Нека V е крайномерно пространство, $v \in V$ и $v = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k$ ($\mu_i \in F$, $v_i \in V$; $i = 1, 2, \ldots, k$). Тогава координатите на вектора v във фиксиран базис на V са линейни комбинации с коефициенти $\mu_1, \mu_2, \ldots, \mu_k$ на съответните координати на векторите $v_1, v_2, \ldots, v_k$ в същия базис.*

§ 10. Сума на подпространства

Сечение на произволна фамилия подпространства на дадено линейно пространство също е подпространство (задача 3 от § 7). Не е вярно обаче, че обединение на подпространства отново е подпространство.

Задача 1. *Да се докаже, че обединението на две подпространства отново е подпространство тогава и само тогава, когато едното подпространство се съдържа в другото.*

Ще дефинираме понятието сума на подпространства, което отново е подпространство.

Определение. *Нека $V_1, V_2, \dots, V_s$ са подпространства на линейното пространство V. Под сума $V_1 + V_2 + \dots + V_s$ на тези подпространства ще разбираме множеството от всички вектори v от V, които могат да се представят като сума $v = v_1 + v_2 + \dots + v_s$, където $v_i \in V_i$, $i = 1, 2, \dots, s$.*

Задача 2. *Да се докаже, че:*

а) $V_1 + V_2 + \dots + V_s$ *е подпространство на* V;

б) $V_1 + V_2 + \dots + V_s = l(V_1 \cup V_2 \cup \dots \cup V_s)$.

Теорема 1. *Нека V е линейно пространство и V_1 и V_2 са крайномерни подпространства на V. Тогава пространствата $V_1 + V_2$ и $V_1 \cap V_2$ също са крайномерни и*

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Доказателство. Нека $\dim V_1 = k$, $\dim V_2 = l$. Тъй като $V_1 \cap V_2$ е подпространство както на V_1, така и на V_2, от задача 1 от § 9 следва, че $V_1 \cap V_2$ е крайномерно пространство и $\dim(V_1 \cap V_2) = r \le k, l$.

Нека $a_1, \dots, a_r$ е базис на $V_1 \cap V_2$, ако $V_1 \cap V_2 \neq \{0\}$. Допълваме го (ако е необходимо) до базис $a_1, \dots, a_r; b_{r+1}, \dots, b_k$ на V_1 и до базис $a_1, \dots, a_r; c_{r+1}, \dots, c_l$ на V_2 (твърдение 5 от § 9). Ще докажем, че системата вектори

$$a_1, \dots, a_r; b_{r+1}, \dots, b_k; c_{r+1}, \dots, c_l \tag{1}$$

е базис на $V_1 + V_2$. Оттук ще следва, че $V_1 + V_2$ е крайномерно и

$$\dim(V_1 + V_2) = r + (k - r) + (l - r) = k + l - r = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

(Ако едно от двете подпространства, например V_1, е нулевото, то $V_1 \cap V_2 = \{0\}$, $V_1 + V_2 = V_2$ и твърдението на теоремата е очевидно. Ако $V_1 \neq \{0\}$ и $V_2 \neq \{0\}$, но $V_1 \cap V_2 = \{0\}$, то векторите $a_1, \dots, a_r$ просто липсват и получаваме базис на $V_1 + V_2$, който е обединение на два базиса съответно на V_1 и V_2.)

Очевидно $V_1 + V_2$ е линейна обвивка на системата вектори (1). Остава да докажем тяхната линейна независимост. Нека

$$\underbrace{\lambda_1 a_1 + \dots + \lambda_r a_r}_{\substack{\parallel \\ a \in V_1 \cap V_2}} + \underbrace{\mu_{r+1} b_{r+1} + \dots + \mu_k b_k}_{\substack{\parallel \\ b \in V_1}} + \underbrace{\nu_{r+1} c_{r+1} + \dots + \nu_l c_l}_{\substack{\parallel \\ c \in V_2}} = 0. \tag{2}$$

Имаме $b \in V_1$ и (от горното равенство) $b = -a - c \in V_2$, следователно $b \in V_1 \cap V_2$. Тогава b се изразява линейно чрез векторите $a_1, \dots, a_r$: $b = \alpha_1 a_1 + \dots + \alpha_r a_r$. Като извадим от това равенство равенството $\boldsymbol{b} = \mu_{r+1}\boldsymbol{b}_{r+1} + \cdots + \mu_k\boldsymbol{b}_k$, получаваме $\alpha_1\boldsymbol{a}_1 + \cdots + \alpha_r\boldsymbol{a}_r - \mu_{r+1}\boldsymbol{b}_{r+1} - \cdots - \mu_k\boldsymbol{b}_k = \boldsymbol{0}$. Но векторите $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$; $\boldsymbol{b}_{r+1}, \ldots, \boldsymbol{b}_k$ са базис на V_1 и значи всички коефициенти в последното равенство, в частност $\mu_{r+1}, \ldots, \mu_k$ са равни на нула. Тогава равенството (2) приема вида

$$\lambda_1\boldsymbol{a}_1 + \cdots + \lambda_r\boldsymbol{a}_r + \nu_{r+1}\boldsymbol{c}_{r+1} + \cdots + \nu_l\boldsymbol{c}_l = \boldsymbol{0}.$$

Тъй като векторите $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$; $\boldsymbol{c}_{r+1}, \ldots, \boldsymbol{c}_l$ са базис на V_2, то и коефициентите $\lambda_1, \ldots, \lambda_r$; $\nu_{r+1}, \ldots, \nu_l$ са равни на нула. Следователно системата вектори (1) е линейно независима.

* * *

Определение. *Нека V е линейно пространство. Ще казваме, че V е директна сума на подпространствата си $V_1, V_2, \ldots, V_s$, ако всеки вектор $\boldsymbol{v}$ от V се представя по* **единствен** *начин като сума $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_s$, където $\boldsymbol{v}_i \in V_i$, $i = 1, 2, \ldots, s$. Ще използваме означението $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$.*

Твърдение 2. *Нека V е линейно пространство и V_1 и V_2 са подпространства на V. Тогава $V = V_1 \oplus V_2$ тогава и само тогава, когато са изпълнени следните две условия:*

1) $V = V_1 + V_2$;
2) $V_1 \cap V_2 = \{\boldsymbol{0}\}$.

Доказателство. Нека $V = V_1 \oplus V_2$. Тогава очевидно $V = V_1 + V_2$. Остава да докажем, че $V_1 \cap V_2 = \{\boldsymbol{0}\}$. Нека $\boldsymbol{v} \in V_1 \cap V_2$. Имаме

$$\boldsymbol{v} = \underbrace{\boldsymbol{v}}_{\in V_1} + \underbrace{\boldsymbol{0}}_{\in V_2}$$
$$\boldsymbol{v} = \underbrace{\boldsymbol{0}}_{\in V_1} + \underbrace{\boldsymbol{v}}_{\in V_2}.$$

От условието за единственост на записа на $\boldsymbol{v}$ като сума на вектори съответно от V_1 и V_2 получаваме $\boldsymbol{v} = \boldsymbol{0}$ и значи $V_1 \cap V_2 = \{\boldsymbol{0}\}$.

Обратно, нека $V = V_1 + V_2$ и $V_1 \cap V_2 = \{\boldsymbol{0}\}$. За да докажем, че $V = V_1 \oplus V_2$, трябва да проверим, че всеки вектор $\boldsymbol{v}$ от V се представя по **единствен** начин като сума на два вектора съответно от V_1 и V_2. Нека

$$\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2, \quad \boldsymbol{v} = \boldsymbol{v}_1' + \boldsymbol{v}_2' \qquad (\boldsymbol{v}_1, \boldsymbol{v}_1' \in V_1;\ \boldsymbol{v}_2, \boldsymbol{v}_2' \in V_2).$$

Като извадим тези две равенства, получаваме $\boldsymbol{0} = (\boldsymbol{v}_1 - \boldsymbol{v}_1') + (\boldsymbol{v}_2 - \boldsymbol{v}_2')$ или $\boldsymbol{v}_1' - \boldsymbol{v}_1 = \boldsymbol{v}_2 - \boldsymbol{v}_2'$. Но $\boldsymbol{v}_1' - \boldsymbol{v}_1 \in V_1$, $\boldsymbol{v}_2 - \boldsymbol{v}_2' \in V_2$ и значи $\boldsymbol{v}_1' - \boldsymbol{v}_1 = \boldsymbol{v}_2 - \boldsymbol{v}_2' \in V_1 \cap V_2$. Следователно $\boldsymbol{v}_1' - \boldsymbol{v}_1 = \boldsymbol{0} = \boldsymbol{v}_2 - \boldsymbol{v}_2'$, т.е. $\boldsymbol{v}_1 = \boldsymbol{v}_1'$ и $\boldsymbol{v}_2 = \boldsymbol{v}_2'$. Така горното представяне на $\boldsymbol{v}$ като сума на вектори съответно от V_1 и V_2 е единствено и значи $V = V_1 \oplus V_2$.

Задача 3. (*обобщение на твърдение 2*). *Нека V е линейно пространство и $V_1, V_2, \ldots, V_s$ са подпространства на V. Да се докаже, че $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$ тогава и само тогава, когато се изпълняват следните две условия:*

1) $V = V_1 + V_2 + \cdots + V_s$;
2) $V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_s) = \{\boldsymbol{0}\}$ *за всяко* $i = 1, 2, \ldots, s$.

Задача 4. *Нека V е линейно пространство и $V_1, V_2, \ldots, V_s$ са подпространства на V. Да се докаже, че $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$ тогава и само тогава, когато се изпълняват следните две условия:*

1) $V = V_1 + V_2 + \cdots + V_s$;

2) *Нулевият вектор се представя по единствен начин като сума на вектори съответно от $V_1, V_2, \ldots, V_s$.*

Твърдение 3. *Нека $e_1, e_2, \ldots, e_n$ е базис на крайномерното линейно пространство V и k е произволно естествено число, ненадминаващо n. Тогава $V = V_1 \oplus V_2$, където $V_1 = l(e_1, \ldots, e_k)$, $V_2 = l(e_{k+1}, \ldots, e_n)$. Обратно, ако $V = V_1 \oplus V_2$ и $e_1, \ldots, e_k$ и $e_{k+1}, \ldots, e_n$ са базиси съответно на V_1 и V_2, то $e_1, \ldots, e_k; e_{k+1}, \ldots, e_n$ е базис на V, в частност $\dim V = \dim V_1 + \dim V_2$ (последното следва и от теорема 1).*

Доказателство. Очевидно $V = V_1 + V_2$. От линейната независимост на векторите $e_1, \ldots, e_k; e_{k+1}, \ldots, e_n$ следва $V_1 \cap V_2 = \{0\}$ (действително, ако $v \in V_1 \cap V_2$, то $v = \alpha_1 e_1 + \cdots + \alpha_k e_k = \alpha_{k+1} e_{k+1} + \cdots + \alpha_n e_n$, откъдето $\alpha_1 e_1 + \cdots + \alpha_k e_k - \alpha_{k+1} e_{k+1} - \cdots - \alpha_n e_n = 0$ и оттук $\alpha_i = 0$ $(i = 1, 2, \ldots, n)$, значи $v = 0$). Сега от твърдение 2 получаваме $V = V_1 \oplus V_2$.

Твърдение 4. *Нека V е крайномерно линейно пространство и W е подпространство на V. Да се докаже, че съществува подпространство U на V, такова че $V = W \oplus U$.*

Доказателство. Нека $e_1, e_2, \ldots, e_k$ е базис на W. Да допълним този базис до базис $e_1, \ldots, e_k; e_{k+1}, \ldots, e_n$ на V (твърдение 5 от § 9). Да положим $U = l(e_{k+1}, \ldots, e_n)$. Тогава според твърдение 3, $V = W \oplus U$.

Задача 5. *Да се докаже, че всяко ненулево крайномерно линейно пространство V е директна сума на едномерни подпространства.*

Упътване. Ако $e_1, e_2, \ldots, e_n$ е базис на V, то $V = l(e_1) \oplus l(e_2) \oplus \cdots \oplus l(e_n)$.

§ 11. Ранг на система вектори. Ранг на матрица

В този параграф ще разгледаме метод, който ни дава възможност да определим дали една система вектори е линейно зависима или не.

Нека $A \in F_{m \times n}$ и

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

Определение. *Ако фиксираме произволни k реда и произволни k стълба на матрицата A, елементите, които стоят в пресечните им точки образуват квадратна матрица от ред k. Детерминантата на всяка такава матрица ще наричаме минор на A от ред k. С други думи, минор от ред k на A е всяка детерминанта от вида*

$$\begin{vmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_k} \\ \vdots & \ddots & \vdots \\ a_{i_k j_1} & \cdots & a_{i_k j_k} \end{vmatrix},$$

където $1 \le i_1 < i_2 < \cdots < i_k \le m$, $1 \le j_1 < j_2 < \cdots < j_k \le n$.

Определение. *Ще казваме, че матрицата A има ранг r (и ще пишем $r(A) = r$), ако A притежава различен от нула минор от ред r и всички минори от по-голям ред са равни на нула. По определение $r(0) = 0$.*

З а б е л е ж к а. Лесно се вижда, че ако всички минори от ред $r+1$ са равни на нула, то и всички минори от ред, по-голям от $r+1$ също са равни на нула.

От теоремата за транспонирана детерминанта (теорема 2 от § 1) следва, че $r(A^t) = r(A)$.

Определение. *Нека V е линейно пространство и*

$$c_1,\ c_2, \ldots,\ c_t \tag{1}$$

е система вектори от V. Ще казваме, че системата вектори (1) има ранг r (и ще пишем $r(c_1,\ c_2, \ldots,\ c_t) = r$), ако съществуват r линейно независими вектора от тази система и всеки друг вектор от системата е тяхна линейна комбинация.

Задача 1. *Да се докаже, че:*

а) *рангът на системата вектори (1) е равен на максималния брой линейно независими вектори в тази система;*

б) $r(c_1,\ c_2, \ldots,\ c_t) = \dim l(c_1,\ c_2, \ldots,\ c_t)$.

Да означим с $a_1,\ a_2, \ldots,\ a_m$ векторите редове на матрицата A, т.е. $a_1 = (a_{11}, a_{12}, \ldots, a_{1n})$, $a_2 = (a_{21}, a_{22}, \ldots, a_{2n})$, $\ldots$, $a_m = (a_{m1}, a_{m2}, \ldots, a_{mn})$. Можем да разглеждаме $a_1,\ a_2, \ldots,\ a_m$ като вектори от линейното пространство F^n. Аналогично, векторите стълбове $b_1,\ b_2, \ldots,\ b_n$ на A ще разглеждаме като вектори от F^m.

Теорема 1. *В сила са равенствата*

$$r(a_1,\ a_2, \ldots,\ a_m) = r(b_1,\ b_2, \ldots,\ b_n) = r(A).$$

Д о к а з а т е л с т в о. Поради наличието на теоремата за транспонирана детерминанта (теорема 2 от § 1), достатъчно е да докажем само равенството $r(b_1,\ b_2, \ldots,\ b_n) = r(A)$. Ако $A = \mathbf{0}$, твърдението е очевидно. Нека $r(A) = r \geq 1$. Без ограничение на общността можем да считаме, че минорът от ред r, стоящ в горния ляв ъгъл на A е различен от нула. Нека

$$D = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{pmatrix}$$

и $\Delta = \det D \neq 0$. Тогава първите r стълба на A са линейно независими. Действително, ако съществува линейна зависимост между стълбовете $b_1, \ldots, b_r$, то същата линейна зависимост ще съществува между стълбовете на матрицата D, откъдето ще следва $\det D = 0$.

Ще докажем, че всеки стълб b_l на A, $r < l \leq n$, е линейна комбинация на стълбовете $b_1, \ldots, b_r$, с което теоремата ще бъде доказана.

Нека i е произволно естествено число между 1 и m. Да разгледаме квадратната матрица D_i от $(r+1)$-ви ред, която се получава като "заградим" матрицата D с i-ия ред и l-ия стълб на A:

$$D_i = \begin{pmatrix} a_{11} & \ldots & a_{1r} & a_{1l} \\ \vdots & \ddots & \vdots & \vdots \\ a_{r1} & \ldots & a_{rr} & a_{rl} \\ a_{i1} & \ldots & a_{ir} & a_{il} \end{pmatrix}.$$

Ако $i \le r$, тази матрица има два равни реда и следователно $\det D_i = 0$. Ако $i > r$, то $\det D_i$ е минор от $(r+1)$-ви ред на A и отново $\det D_i = 0$. Да означим с $A_1, \dots, A_r; A_l$ адюнгираните количества съответно на елементите $a_{i1}, \dots, a_{ir}; a_{il}$. Очевидно $A_l = \Delta$. Развиваме $\det D_i$ по последния ред и получаваме

$$a_{i1}A_1 + \cdots + a_{ir}A_r + a_{il}\Delta = 0.$$

Тъй като $\Delta \ne 0$, то

$$a_{il} = -\frac{A_1}{\Delta}a_{i1} - \cdots - \frac{A_r}{\Delta}a_{ir}.$$

Това равенство е в сила за всяко $i = 1, 2, \dots, m$. При това, адюнгираните количества A_j $(j = 1, 2, \dots, r)$ **не зависят** от i, тъй като

$$A_j = (-1)^{(r+1)+j}\begin{vmatrix} a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1r} & a_{1l} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{r1} & \dots & a_{rj-1} & a_{rj+1} & \dots & a_{rr} & a_{rl} \end{vmatrix}.$$

Така всеки елемент a_{il} на l-ия стълб b_l е линейна комбинация на съответните елементи на стълбовете $b_1, \dots, b_r$ **с едни и същи коефициенти** за всяко $i = 1, 2, \dots, m$. Следователно b_l е линейна комбинация на $b_1, \dots, b_r$ със същите коефициенти.

Забележка. В доказателството на теоремата използвахме, че са равни на нула **не всички** минори от ред $r+1$, а само тези минори от ред $r+1$, които заграждат D. Така за да докажем, че $r(A) = r$, достатъчно е да намерим ненулев минор от ред r и да проверим, че са равни на нула само минорите от ред $r+1$, които го заграждат.

Следствие 2. *Ако A е квадратна матрица от ред n, то $\det A = 0$ тогава и само тогава, когато редовете (стълбовете) на A са линейно зависими.*

Доказателство. Едната посока на твърдението следва от свойство 4 за линейна зависимост на вектори от § 8 и свойство 8 за детерминанти от § 2.

Обратно, нека $\det A = 0$. Тогава $r(A) < n$. Според теорема 1 рангът на векторите редове (стълбове) на A също е по-малък от n и значи те са линейно зависими.

§ 12. Системи линейни уравнения. Хомогенни системи

Системи линейни уравнения. Да разгледаме системата

$$\left| \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array} \right. \qquad (1)$$

Да означим с A и $\overline{A}$ съответно матрицата и разширената матрица на системата.

Теорема 1 (теорема на Руше). *Системата (1) е съвместима тогава и само тогава, когато $r(A) = r(\overline{A})$.*

Доказателство. Да означим с $b_1, \dots, b_n$ векторите стълбове на матрицата A, а с b – стълба от свободните членове. Имаме $r(A) = r(b_1, \dots, b_n) \le r(b_1, \dots, b_n; b) = r(\overline{A})$.

При това $r(A) = r(\overline{A})$ тогава и само тогава, когато $\boldsymbol{b}$ е линейна комбинация на $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$, т.е. съществуват числа $\lambda_1, \ldots, \lambda_n$, такива че $\lambda_1 \boldsymbol{b}_1 + \cdots + \lambda_n \boldsymbol{b}_n = \boldsymbol{b}$. Но това е еквивалентно на факта, че n-орката $(\lambda_1, \ldots, \lambda_n)$ е решение на системата (1).

Забележка. Доказателството на теорема 1 звучи като тавтология (системата (1) е съвместима тогава и само тогава, когато има решение). Това, което действително представлява теоретичен интерес е съвпадението на ранга на матрицата A и ранга на системата от нейните вектори редове (стълбове).

Теоремата на Руше е критерий за съвместимост на дадена система, но не ни показва как да намираме решенията ѝ. Ще разгледаме един подход за практическото намиране на всички решения на една система (макар че поради технически трудности този подход представлява повече теоретичен интерес, за разлика от метода на Гаус от § 05).

Нека системата (1) е съвместима и $r(A) = r(\overline{A}) = r$. След евентуално разместване на уравненията и преномериране на неизвестните, можем да считаме, че минорът от ред r в горния ляв ъгъл на A (или все едно, на $\overline{A}$) е различен от нула. Тогава първите r реда на $\overline{A}$ са линейно независими и всеки друг ред е тяхна линейна комбинация. Оттук следва, че всяко решение на първите r уравнения на системата е решение и на останалите. Тогава системата (1) е еквивалентна на системата

$$\left| \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n = b_r. \end{array} \right. \tag{2}$$

Ако $r = n$, то детерминантата на системата (2) е различна от нула и тя има единствено решение, което се получава по формулите на Крамер.

Нека $r < n$. Записваме системата (2) във вида

$$\left| \begin{array}{l} a_{11}x_1 + \cdots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \cdots - a_{1n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \cdots - a_{rn}x_n. \end{array} \right. \tag{3}$$

Неизвестните $x_{r+1}, \ldots, x_n$ ще наричаме *свободни неизвестни*. Да дадем на свободните неизвестни произволни стойности $k_{r+1}, \ldots, k_n$. Получаваме система от r уравнения с r неизвестни $x_1, \ldots, x_r$, чиято детерминанта е различна от нула. Следователно тя има единствено решение $k_1, \ldots, k_r$, което се получава по формулите на Крамер. Очевидно n-орката $\mathbf{c} = (k_1, \ldots, k_r; k_{r+1}, \ldots, k_n)$ е решение на системата (3).

Задача 1. *Нека* $\mathbf{c} = (k_1, \ldots, k_r; k_{r+1}, \ldots, k_n)$ *и* $\mathbf{c}' = (k'_1, \ldots, k'_r; k'_{r+1}, \ldots, k'_n)$ *са решения на системата* (3). *Да се докаже, че ако* $k_{r+1} = k'_{r+1}, \ldots, k_n = k'_n$, *то също така* $k_1 = k'_1, \ldots, k_r = k'_r$, *т. е.* $\mathbf{c} = \mathbf{c}'$.

По описания по-горе начин можем да получим всички решения на системата (3). Нека $\mathbf{c} = (k_1, \ldots, k_r; k_{r+1}, \ldots, k_n)$ е произволно решение. Да дадем на свободните неизвестни стойности $k_{r+1}, \ldots, k_n$. Тогава за неизвестните $x_1, \ldots, x_r$ получаваме еднозначно определени стойности $k'_1, \ldots, k'_r$, т.е. получаваме решението $\mathbf{c}' = (k'_1, \ldots, k'_r; k_{r+1}, \ldots, k_n)$. Сега от задача 1 следва $\mathbf{c} = \mathbf{c}'$.

Така една съвместима система има единствено решение точно когато рангът на матрицата на системата е равен на броя на неизвестните и безбройно много решения, ако този ранг е по-малък от броя на неизвестните.

* * *

Хомогенни системи. Да приложим получените резултати за хомогенни системи линейни уравнения. Нека всички свободни членове на системата (1) са равни на нула. Тогава тя приема вида

$$\left|\begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{array}\right. \tag{$1'$}$$

Тази система винаги е съвместима, защото притежава нулевото решение. При това, ако $m < n$, то $r(A) \leq m < n$ и тогава системата има безбройно много решения, в частност има ненулево решение (този резултат получихме и в § 05).

Твърдение 2. *Нека матрицата A на системата $(1')$ е квадратна, т. е. $m = n$. В този случай системата $(1')$ има ненулево решение тогава и само тогава, когато* $\det A = 0$.

Доказателство. Ако $\det A = 0$, т. е. $r(A) < n$, системата има безбройно много решения, значи има и ненулево решение. Ако $\det A \neq 0$, т. е. $r(A) = n$, системата има единствено решение и то е нулевото.

Директно от определенията се проверява, че всяка линейна комбинация от решения на една хомогенна система също е решение на тази система. Следователно *множеството от решенията на хомогенната система $(1')$ е подпространство на линейното пространство F^n.*

Определение. *Всеки базис на пространството от решенията (ако то е ненулево) на хомогенната система $(1')$ ще наричаме фундаментална система решения на тази система. Така всяка фундаментална система решения се състои от линейно независими вектори от F^n, които са решения на хомогенната система и всяко друго решение е тяхна линейна комбинация.*

Твърдение 3. *Нека U е пространството от решенията на хомогенната система $(1')$. Тогава, ако $r(A) = r$, то $\dim U = n - r$, т. е. всяка фундаментална система решения на системата $(1')$ се състои от $n - r$ на брой решения.*

Доказателство. Ако $r(A) = 0$, т. е. $A = \mathbf{0}$, то $U = F^n$ и $\dim U = n$. Нека $A \neq \mathbf{0}$ и първите r реда на A са линейно независими. Както в първата част на параграфа, записваме системата $(1')$ във вида

$$\left|\begin{array}{l} a_{11}x_1 + \cdots + a_{1r}x_r = -a_{1,r+1}x_{r+1} - \cdots - a_{1n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = -a_{r,r+1}x_{r+1} - \cdots - a_{rn}x_n, \end{array}\right. \tag{$3'$}$$

където

$$\begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Да дадем на свободните неизвестни $x_{r+1}, \dots, x_n$ последователно следните стойности: първо $(1, 0, 0, \dots, 0)$, после $(0, 1, 0, \dots, 0)$ и т. н., накрая $(0, 0, 0, \dots, 1)$. Във всеки един от тези случаи намираме еднозначно определени стойности за неизвестните $x_1, \dots, x_r$. Така получаваме $n - r$ решения на хомогенната система $(3')$:

$$\begin{array}{l} c_1 = (k_{11}, \dots, k_{1r}; 1, 0, 0, \dots, 0), \\ c_2 = (k_{21}, \dots, k_{2r}; 0, 1, 0, \dots, 0), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ c_{n-r} = (k_{n-r,1}, \dots, k_{n-r,r}; 0, 0, 0, \dots, 1). \end{array} \tag{4}$$

Ще докажем, че системата вектори (4) е фундаментална система решения на хомогенната система (3′), с което твърдението ще бъде доказано.

Първо, системата вектори (4) е линейно независима, защото матрицата C, съставена от елементите им има ранг, равен на $n-r$ (минорът от ред $n-r$, съставен от нули и единици и стоящ в дясната част на матрицата C е равен на $1 \neq 0$). Остава да проверим, че всяко решение на (3′) е линейна комбинация на системата вектори (4). Нека $\mathbf{c} = (k_1, \dots, k_r; k_{r+1}, \dots, k_n)$ е произволно решение на системата (3′). Да разгледаме вектора $\mathbf{c}' = k_{r+1}\mathbf{c}_1 + \dots + k_n\mathbf{c}_{n-r}$. Този вектор също е решение на (3′), при това $\mathbf{c}' = (*, \dots, *; k_{r+1}, \dots, k_n)$ (няма да изписваме първите r елемента на $\mathbf{c}'$). От задача 1 следва, че $\mathbf{c} = \mathbf{c}'$ и значи $\mathbf{c}$ е линейна комбинация на векторите $\mathbf{c}_1, \dots, \mathbf{c}_{n-r}$.

Задача 2. *Да се докаже, че всяко решение $\boldsymbol{x}$ на системата (1) е от вида $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{u}$, където $\boldsymbol{x}_0$ е фиксирано решение на системата (1), а $\boldsymbol{u}$ е произволно решение на системата (1′). С други думи, всяко решение на системата (1) се записва във вида $\boldsymbol{x} = \boldsymbol{x}_0 + \lambda_1\mathbf{c}_1 + \dots + \lambda_{n-r}\mathbf{c}_{n-r}$, където $\mathbf{c}_1, \dots, \mathbf{c}_{n-r}$ е фиксирана фундаментална система решения на системата (1′).*

Вече знаем, че множеството от решенията на една хомогенна система с n неизвестни е подпространство на F^n. Вярно е и обратното:

Теорема 4. *Всяко подпространство W на линейното пространство F^n е пространството от решенията на подходяща хомогенна система линейни уравнения с n неизвестни.*

Доказателство. Нека $\dim W = n - r$, $0 \leq r \leq n$. Ще докажем, че съществува хомогенна система с n неизвестни, чиято матрица има ранг, равен на r и пространството от решенията ѝ съвпада с W.

Ако $r = n$, т.е. $\dim W = 0$ (и значи $W = \{\mathbf{o}\}$), то можем да вземем хомогенната система, чиято матрица съвпада с единичната матрица E_n от ред n. Ако пък $r = 0$, т.е. $\dim W = n$ (и значи $W = F^n$), то можем да вземем система с едно уравнение и нулеви коефициенти пред $x_1, \dots, x_n$.

Нека $0 < r < n$. Нека векторите $\boldsymbol{b}_1 = (b_{11}, \dots, b_{1n}), \dots, \boldsymbol{b}_{n-r} = (b_{n-r,1}, \dots, b_{n-r,n})$ са базис на W. Да разгледаме системата

$$\left| \begin{array}{l} b_{11}x_1 + \dots + \quad b_{1n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ b_{n-r,1}x_1 + \dots + b_{n-r,n}x_n = 0. \end{array} \right. \tag{5}$$

Тъй като векторите $\boldsymbol{b}_1, \dots, \boldsymbol{b}_{n-r}$ са линейно независими, рангът на матрицата на тази система е равен на $n-r$. Тогава пространството U от решенията ѝ е с размерност $n - (n-r) = r$. Нека векторите $\boldsymbol{a}_1 = (a_{11}, \dots, a_{1n}), \dots, \boldsymbol{a}_r = (a_{r1}, \dots, a_{rn})$ са базис на U, т.е. фундаментална система решения на системата (5). Да разгледаме системата

$$\left| \begin{array}{l} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + \dots + a_{rn}x_n = 0 \end{array} \right. \tag{6}$$

Рангът на матрицата на тази система е равен на r. Тогава пространството W' от решенията ѝ е с размерност $n-r$. Освен това, $W \subseteq W'$. Действително, векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_r$ са решения на системата (5), в частност, удовлетворяват първото ѝ уравнение. Но това означава, че векторът $\boldsymbol{b}_1$ е решение на системата (6), т.е. $\boldsymbol{b}_1 \in W'$. Аналогично $\boldsymbol{b}_2, \dots, \boldsymbol{b}_{n-r} \in W'$. Тъй като векторите $\boldsymbol{b}_1, \dots, \boldsymbol{b}_{n-r}$ са базис на W, то $W \subseteq W'$. Сега от $\dim W = n - r = \dim W'$ следва $W = W'$. Следователно пространството от решенията на хомогенната система (6) е точно W.

Глава III. Линейни изображения

§ 13. Линейни изображения

Определение. *Нека V и V' са линейни пространства и $\varphi : V \longrightarrow V'$ е изображение от V към V'. Ще казваме, че φ е линейно изображение, ако образът на произволна линейна комбинация на вектори от V е същата линейна комбинация на образите им във V', т.е. ако от $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in V$ и $\lambda_1, \ldots, \lambda_k \in F$ следва $\varphi(\lambda_1\boldsymbol{a}_1 + \cdots + \lambda_k\boldsymbol{a}_k) = \lambda_1\varphi(\boldsymbol{a}_1) + \cdots + \lambda_k\varphi(\boldsymbol{a}_k)$.*

С $\operatorname{Hom}(V, V')$ ще означаваме множеството от всички линейни изображения от V към V'. Ако V' съвпада с V, вместо линейно изображение ще използваме термина *линеен оператор* и вместо $\operatorname{Hom}(V, V)$ ще пишем само $\operatorname{Hom} V$.

Задача 1. *Да се докаже, че $\varphi : V \longrightarrow V'$ е линейно изображение тогава и само тогава, когато са изпълнени следните две условия:*

1) *ако $\boldsymbol{a}_1, \boldsymbol{a}_2 \in V$, то $\varphi(\boldsymbol{a}_1 + \boldsymbol{a}_2) = \varphi(\boldsymbol{a}_1) + \varphi(\boldsymbol{a}_2)$;*
2) *ако $\boldsymbol{a} \in V$ и $\lambda \in F$, то $\varphi(\lambda\boldsymbol{a}) = \lambda\varphi(\boldsymbol{a})$.*

Задача 2. *Нека $\varphi : V \longrightarrow V'$ е линейно изображение. Да се докаже, че:*

а) *$\varphi(0) = 0$;*
б) *за всеки вектор $\boldsymbol{a}$ от V е изпълнено $\varphi(-\boldsymbol{a}) = -\varphi(\boldsymbol{a})$;*
в) *ако $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ са линейно зависими вектори от V, то $\varphi(\boldsymbol{a}_1), \ldots, \varphi(\boldsymbol{a}_k)$ са също линейно зависими вектори от V'.*

Ще бележим с $\mathbf{0}$ нулевото изображение, т.е. изображението $\mathbf{0} : V \longrightarrow V'$, за което $\mathbf{0}(\boldsymbol{a}) = 0$ за всеки вектор $\boldsymbol{a}$ от V. С ε_V или само ε, ако V се подразбира, ще бележим идентитета на V, т.е. изображението $\varepsilon : V \longrightarrow V$, за което $\varepsilon(\boldsymbol{a}) = \boldsymbol{a}$ за всеки вектор $\boldsymbol{a}$ от V.

Примери. 1. Изображенията $\mathbf{0}$ и ε са линейни.

2. Нека m и n са естествени числа.

а) Нека $m \geq n$. Изображението $\varphi : F^m \longrightarrow F^n$, дефинирано по правилото:

$$F^m \ni (\alpha_1, \ldots, \alpha_n, \ldots, \alpha_m) \xrightarrow{\varphi} (\alpha_1, \ldots, \alpha_n) \in F^n$$

е линейно изображение.

б) Нека $m < n$. Изображението $\varphi : F^m \longrightarrow F^n$, дефинирано по правилото:

$$F^m \ni (\alpha_1, \ldots, \alpha_m) \xrightarrow{\varphi} (\alpha_1, \ldots, \alpha_m, 0, \ldots, 0) \in F^n$$

е линейно изображение.

3. Изображението $\delta : F[x] \longrightarrow F[x]$, което съпоставя на всеки полином от $F[x]$ неговата производна е линейно изображение.

Теорема 1. *Нека V и V' са линейни пространства и $\dim V = n < \infty$. Тогава за всеки базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ на V и произволни n вектора $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ от V' съществува и то единствено линейно изображение $\varphi : V \longrightarrow V'$, такова че $\varphi(\boldsymbol{e}_i) = \boldsymbol{v}_i$, $i = 1, \ldots, n$.*

Доказателство. **Съществуване.** Дефинираме изображение $\varphi : V \longrightarrow V'$ по следния начин: ако $\boldsymbol{a} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_n \boldsymbol{e}_n$ е произволен вектор от V, то $\varphi(\boldsymbol{a}) = \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n$. Изображението φ е коректно дефинирано (т.е. векторът $\varphi(\boldsymbol{a})$ от V' се определя еднозначно от вектора $\boldsymbol{a}$ от V), тъй като $\boldsymbol{a}$ се записва по **единствен** начин като линейна комбинация на базисните вектори $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Очевидно $\varphi(\boldsymbol{e}_i) = \boldsymbol{v}_i$, $i = 1, \ldots, n$. Проверката, че изображението φ е линейно е директна* и я предоставяме на читателя (за съкращаване на проверката може да се използва задача 1).

Единственост. Нека $\varphi, \psi : V \longrightarrow V'$ са линейни изображения и $\varphi(\boldsymbol{e}_i) = \boldsymbol{v}_i = \psi(\boldsymbol{e}_i)$, $i = 1, \ldots, n$. От линейността на φ и ψ следва, че ако $\boldsymbol{a} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_n \boldsymbol{e}_n$ е произволен вектор от V, то $\varphi(\boldsymbol{a}) = \lambda_1 \varphi(\boldsymbol{e}_1) + \cdots + \lambda_n \varphi(\boldsymbol{e}_n) = \lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n = \lambda_1 \psi(\boldsymbol{e}_1) + \cdots + \lambda_n \psi(\boldsymbol{e}_n) = \psi(\boldsymbol{a})$. Следователно $\varphi = \psi$.

Забележка. Единствеността в това твърдение може да бъде изказана по-свободно така: ако две линейни изображения съвпадат върху базис на V, то те изобщо съвпадат.

Определение. *Нека V и V' са линейни пространства и $\varphi : V \longrightarrow V'$ е изображение. Ще казваме, че φ е изоморфизъм, ако:*

1) *φ е линейно изображение;*
2) *φ е биекция.*

Ще казваме още, че пространствата V и V' са изоморфни и ще пишем $V \cong V'$.

Задача 3. *Да се докаже, че ако $\varphi : V \longrightarrow V'$ е изоморфизъм, то $\varphi^{-1} : V' \longrightarrow V$ е линейно изображение и значи също е изоморфизъм.*

Твърдение 2. *При изоморфизъм образ на линейно независима система вектори също е линейно независима система.*

Доказателство. Нека $\varphi : V \longrightarrow V'$ е изоморфизъм и векторите $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ от V са линейно независими. Нека $\lambda_1 \varphi(\boldsymbol{e}_1) + \cdots + \lambda_n \varphi(\boldsymbol{e}_n) = \boldsymbol{0} \in V'$. Да приложим към това равенство φ^{-1}. Получаваме $\lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_n \boldsymbol{e}_n = \boldsymbol{0} \in V$. От линейната независимост на $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ следва $\lambda_1 = \cdots = \lambda_n = 0$ и значи векторите $\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n)$ също са линейно независими.

Теорема 3. *Две крайномерни над F пространства V и V' са изоморфни тогава и само тогава, когато имат еднаква размерност.*

Доказателство. Нека $\varphi : V \longrightarrow V'$ е изоморфизъм и $\dim V = n$. Нека $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ е базис на V. Според твърдение 2 векторите $\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n)$ са линейно независими. Нека $\boldsymbol{b}$ е произволен вектор от V' и $\boldsymbol{a}$ е вектор от V, за който $\varphi(\boldsymbol{a}) = \boldsymbol{b}$. Ако $\boldsymbol{a} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_n \boldsymbol{e}_n$, то $\boldsymbol{b} = \varphi(\boldsymbol{a}) = \lambda_1 \varphi(\boldsymbol{e}_1) + \cdots + \lambda_n \varphi(\boldsymbol{e}_n)$. Така векторите $\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n)$ са линейно независими и всеки вектор от V' е тяхна линейна комбинация. Следователно тези вектори са базис на V' и значи $\dim V' = n = \dim V$.

Обратно, нека $\dim V = \dim V' = n$ и $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ и $\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n$ са базиси съответно на V и V'. Според теорема 1 съществува линейно изображение $\varphi : V \longrightarrow V'$, такова че $\varphi(\boldsymbol{e}_i) = \boldsymbol{e}'_i$, $i = 1, \ldots, n$. Ще докажем, че φ е биекция, откъдето следва $V \cong V'$.

Нека $\boldsymbol{b} = \lambda_1 \boldsymbol{e}'_1 + \cdots + \lambda_n \boldsymbol{e}'_n$ е произволен вектор от V'. Да разгледаме вектора $\boldsymbol{a} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_n \boldsymbol{e}_n$ от V. Имаме $\varphi(\boldsymbol{a}) = \lambda_1 \varphi(\boldsymbol{e}_1) + \cdots + \lambda_n \varphi(\boldsymbol{e}_n) = \lambda_1 \boldsymbol{e}'_1 + \cdots + \lambda_n \boldsymbol{e}'_n = \boldsymbol{b}$. Така всеки вектор от V' си има праобраз във V.

Нека $\boldsymbol{a}_1 = \mu_1 \boldsymbol{e}_1 + \cdots + \mu_n \boldsymbol{e}_n$ и $\boldsymbol{a}_2 = \nu_1 \boldsymbol{e}_1 + \cdots + \nu_n \boldsymbol{e}_n$ са два различни вектора от V. Тогава за поне едно i $(1 \leq i \leq n)$ имаме $\mu_i \neq \nu_i$. Оттук получаваме $\varphi(\boldsymbol{a}_1) = \mu_1 \boldsymbol{e}'_1 + \cdots + \mu_n \boldsymbol{e}'_n \neq \nu_1 \boldsymbol{e}'_1 + \cdots + \nu_n \boldsymbol{e}'_n = \varphi(\boldsymbol{a}_2)$. Следователно φ е биекция.

*т.е. използва само определенията и никакви други съображения.

Прието е доказаната теорема да се изказва още и така: за всяко естествено число n съществува единствено с точност до изоморфизъм n-мерно линейно пространство над F. При това, всяко такова пространство е изоморфно на F^n (тъй като $\dim F^n = n$). По-точно, ако $e_1, \ldots, e_n$ е фиксиран базис на V, то изображението, което съпоставя на всеки вектор от V наредената n-орка от координатите му в този базис е един изоморфизъм от V към F^n.

Ако две пространства са изоморфни, то всяко алгебрично твърдение, което е вярно за едното пространство, е вярно и за другото. В такъв смисъл двете пространства са неразличими от алгебрична гледна точка.

* * *

Нека V и V' са крайномерни линейни пространства и $\varphi \in \mathrm{Hom}(V, V')$. Нека $e_1, e_2, \ldots, e_n$ и $e'_1, e'_2, \ldots, e'_m$ са базиси съответно на V и V'. Образите на векторите $e_1, e_2, \ldots, e_n$ под действието на φ са линейни комбинации на векторите $e'_1, e'_2, \ldots, e'_m$. Нека

$$\begin{aligned} \varphi(e_1) &= a_{11}e'_1 + a_{21}e'_2 + \cdots + a_{m1}e'_m, \\ \varphi(e_2) &= a_{12}e'_1 + a_{22}e'_2 + \cdots + a_{m2}e'_m, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi(e_n) &= a_{1n}e'_1 + a_{2n}e'_2 + \cdots + a_{mn}e'_m. \end{aligned} \tag{1}$$

Ще подчертаем, че **вторият** индекс на коефициентите в записа на $\varphi(e_i)$ е равен на i.

Определение. *Матрицата*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

се нарича матрица на линейното изображение φ *във фиксираните базиси* $e_1, e_2, \ldots, e_n$ *и* $e'_1, e'_2, \ldots, e'_m$ *съответно на пространствата* V *и* V'.

Така **стълбовете** на матрицата A са съставени от координатите на $\varphi(e_1), \varphi(e_2), \ldots, \varphi(e_n)$ спрямо базиса $e'_1, e'_2, \ldots, e'_m$. Матрицата на линейно изображение зависи от избраните базиси на двете пространства. Ясно е, че на нулевото изображение $\mathbf{0}$ винаги съответства нулевата матрица, която ще бележим също с $\mathbf{0}$.

Поради важността на частния случай, когато V' съвпада с V и е фиксиран един единствен базис на V, ще се спрем на него отделно.

Нека φ е линеен оператор, действащ във V и $e_1, e_2, \ldots, e_n$ е фиксиран базис на V. Нека

$$\begin{aligned} \varphi(e_1) &= a_{11}e_1 + a_{21}e_2 + \cdots + a_{n1}e_n, \\ \varphi(e_2) &= a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \varphi(e_n) &= a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n. \end{aligned} \tag{2}$$

Определение. *Квадратната матрица*

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

ще наричаме матрица на линейния оператор φ в базиса $e_1,\ e_2,\ldots,\ e_n$.

Стълбовете на матрицата A са съставени от координатите на векторите $\varphi(e_1)$, $\varphi(e_2),\ldots,\varphi(e_n)$ спрямо базиса $e_1,\ e_2,\ldots,\ e_n$. Матрицата на линеен оператор зависи от избора на базиса. На *единичния оператор* ε (идентитета на V) съответства единичната матрица E_n (или само E, ако n е фиксирано) в кой да е базис на V.

Да резюмираме казаното дотук. Нека $e_1,\ e_2,\ldots,\ e_n$ е фиксиран базис на линейното пространство V.Тогава на всеки линеен оператор $\varphi\in\operatorname{Hom}V$ съпоставяме матрица $A\in M_n(F)$, чиито стълбове са съставени от коефициентите в съответните редове в дясната страна на равенствата (2). Обратно, нека $A=(a_{ij})_{n\times n}$ е произволна матрица от $M_n(F)$. От теорема 1 следва, че съществува, при това единствен линеен оператор $\varphi\in\operatorname{Hom}V$, такъв че координатите на векторите $\varphi(e_1),\varphi(e_2),\ldots,\varphi(e_n)$ спрямо базиса $e_1,\ e_2,\ldots,\ e_n$ се определят от равенствата (2). Тогава A е точно матрицата на φ във фиксирания базис на V.

При горните означения нека $v\in V$ и $v=\xi_1e_1+\xi_2e_2+\cdots+\xi_ne_n$, $\varphi(v)=\zeta_1e_1+\zeta_2e_2+\cdots+\zeta_ne_n$. Да означим с $\boldsymbol{\xi}$ и $\boldsymbol{\zeta}$ следните матрици стълбове:

$$\boldsymbol{\xi}=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix},\qquad \boldsymbol{\zeta}=\begin{pmatrix}\zeta_1\\ \zeta_2\\ \vdots\\ \zeta_n\end{pmatrix}.$$

Следващото твърдение показва как можем да пресмятаме координатите на вектора $\varphi(v)$ във фиксиран базис $e_1,\ e_2,\ldots,\ e_n$ на V, ако знаем координатите на v и матрицата A на оператора φ в същия базис.

Твърдение 4. *В сила е матричното равенство* $\boldsymbol{\zeta}=A\boldsymbol{\xi}$.

Д о к а з а т е л с т в о. Имаме

$$\begin{aligned}\varphi(v)&=\varphi(\xi_1e_1+\xi_2e_2+\cdots+\xi_ne_n)=\xi_1\varphi(e_1)+\xi_2\varphi(e_2)+\cdots+\xi_n\varphi(e_n)\\ &=\xi_1(a_{11}e_1+a_{21}e_2+\cdots+a_{n1}e_n)+\xi_2(a_{12}e_1+a_{22}e_2+\cdots+a_{n2}e_n)+\cdots\\ &\qquad+\xi_n(a_{1n}e_1+a_{2n}e_2+\cdots+a_{nn}e_n)\\ &=(a_{11}\xi_1+a_{12}\xi_2+\cdots+a_{1n}\xi_n)e_1+(a_{21}\xi_1+a_{22}\xi_2+\cdots+a_{2n}\xi_n)e_2+\cdots\\ &\qquad+(a_{n1}\xi_1+a_{n2}\xi_2+\cdots+a_{nn}\xi_n)e_n.\end{aligned}$$

От друга страна $\varphi(v)=\zeta_1e_1+\zeta_2e_2+\cdots+\zeta_ne_n$. Като приравним коефициентите пред съответните базисни вектори, получаваме подробен запис на матричното равенство $\boldsymbol{\zeta}=A\boldsymbol{\xi}$.

§ 14. Действия с линейни изображения

Определение. *Нека V и V' са линейни пространства, φ, $\psi \in \operatorname{Hom}(V, V')$ и $\lambda \in F$. Под сума на линейните изображения φ и ψ ще разбираме изображението $\varphi + \psi : V \longrightarrow V'$, действащо по следния начин: ако $\boldsymbol{a} \in V$, то $(\varphi + \psi)(\boldsymbol{a}) = \varphi(\boldsymbol{a}) + \psi(\boldsymbol{a}) \in V'$. Под произведение на линейното изображение φ с числото λ ще разбираме изображението $\lambda\varphi : V \longrightarrow V'$, действащо по следния начин: ако $\boldsymbol{a} \in V$, то $(\lambda\varphi)(\boldsymbol{a}) = \lambda\varphi(\boldsymbol{a}) \in V'$.*

Задача 1. а) *Да се докаже, че изображенията $\varphi + \psi$ и $\lambda\varphi$ са линейни, т. е. принадлежат на $\operatorname{Hom}(V, V')$.*

б) *Да се докаже, че с така въведените операции събиране на линейни изображения и умножение на линейно изображение с число множеството $\operatorname{Hom}(V, V')$ се превръща в линейно пространство над F.*

Теорема 1. *Нека V и V' са крайномерни линейни пространства и $\dim V = n$, $\dim V' = m$. Тогава линейните пространства $\operatorname{Hom}(V, V')$ и $F_{m\times n}$ са изоморфни. В частност изоморфни са линейните пространства $\operatorname{Hom} V$ и $M_n(F)$.*

Д о к а з а т е л с т в о. Нека $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ и $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_m$ са базиси съответно на V и V'. Нека $\Phi : \operatorname{Hom}(V, V') \longrightarrow F_{m\times n}$ е изображение, което съпоставя на всяко линейно изображение $\varphi \in \operatorname{Hom}(V, V')$ матрицата му $A \in F_{m\times n}$ в тези базиси. Ще докажем, че Φ е изоморфизъм на линейни пространства.

Нека $\varphi, \psi \in \operatorname{Hom}(V, V')$ и $\varphi \neq \psi$. Тогава за поне едно число i $(1 \le i \le n)$ имаме $\varphi(\boldsymbol{e}_i) \neq \psi(\boldsymbol{e}_i)$ (в противен случай, според теорема 1 от § 13, бихме имали $\varphi = \psi$). Следователно матриците на φ и ψ се различават поне по i-ия си стълб.

Нека сега $A = (a_{ij})_{m\times n}$ е произволна матрица от $F_{m\times n}$. Според теорема 1 от § 13 съществува (единствено) линейно изображение $\varphi \in \operatorname{Hom}(V, V')$, което се задава чрез равенствата (1). Тогава A е матрицата на φ във фиксираните базиси. Така доказахме, че Φ е биекция.

Проверката за линейност на Φ е директна и я предоставяме на читателя.

Определение. *Нека V, V' и V'' са линейни пространства и $\varphi \in \operatorname{Hom}(V, V')$, $\psi \in \operatorname{Hom}(V', V'')$. Под произведение $\psi\varphi : V \longrightarrow V''$ на φ и ψ ще разбираме тяхната композиция (първо прилагаме φ към даден вектор $\boldsymbol{a}$ от V и после към получения вектор от V' прилагаме ψ, т. е. $(\psi\varphi)(\boldsymbol{a}) = \psi(\varphi(\boldsymbol{a}))$).*

Директно се проверява, че произведение на линейни изображения също е линейно изображение, т. е. $\psi\varphi \in \operatorname{Hom}(V, V'')$.

От теорема 1 виждаме, че понятията сума на линейни изображения и произведение на линейно изображение с число са съгласувани със съответните операции за матриците им. Сега ще видим, че същото е вярно и за произведение на линейни изображения. За краткост, ако на линейното изображение φ съответства матрицата A (и базисите се подразбират), ще пишем $\varphi \longrightarrow A$.

Твърдение 2. *Нека линейните пространства V, V' и V'' са крайномерни и $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\dots$, $\boldsymbol{e}_n$; $\boldsymbol{e}'_1, \boldsymbol{e}'_2, \dots, \boldsymbol{e}'_m$ и $\boldsymbol{e}''_1, \boldsymbol{e}''_2, \dots, \boldsymbol{e}''_k$ са фиксирани базиси съответно на V, V' и V''. Нека $\varphi \in \operatorname{Hom}(V, V')$, $\psi \in \operatorname{Hom}(V', V'')$. Тогава, ако $\varphi \longrightarrow A \in F_{m\times n}$, $\psi \longrightarrow B \in F_{k\times m}$, то $\psi\varphi \longrightarrow BA \in F_{k\times n}$.*

Д о к а з а т е л с т в о. Нека $A = (a_{ij})_{m\times n}$, $B = (b_{ij})_{k\times m}$. Имаме

$$
\begin{aligned}
\varphi(\boldsymbol{e}_1) &= a_{11}\boldsymbol{e}'_1 + a_{21}\boldsymbol{e}'_2 + \dots + a_{m1}\boldsymbol{e}'_m, & \psi(\boldsymbol{e}'_1) &= b_{11}\boldsymbol{e}''_1 + b_{21}\boldsymbol{e}''_2 + \dots + b_{k1}\boldsymbol{e}''_k,\\
\varphi(\boldsymbol{e}_2) &= a_{12}\boldsymbol{e}'_1 + a_{22}\boldsymbol{e}'_2 + \dots + a_{m2}\boldsymbol{e}'_m, & \psi(\boldsymbol{e}'_2) &= b_{12}\boldsymbol{e}''_1 + b_{22}\boldsymbol{e}''_2 + \dots + b_{k2}\boldsymbol{e}''_k,\\
&\dots\dots\dots\dots\dots\dots & &\dots\dots\dots\dots\dots\dots\\
\varphi(\boldsymbol{e}_n) &= a_{1n}\boldsymbol{e}'_1 + a_{2n}\boldsymbol{e}'_2 + \dots + a_{mn}\boldsymbol{e}'_m; & \psi(\boldsymbol{e}'_m) &= b_{1m}\boldsymbol{e}''_1 + b_{2m}\boldsymbol{e}''_2 + \dots + b_{km}\boldsymbol{e}''_k;
\end{aligned}
$$

Сега

$$(\psi\varphi)(e_1) = \psi(\varphi(e_1)) = \psi(a_{11}e'_1 + a_{21}e'_2 + \cdots + a_{m1}e'_m) = a_{11}\psi(e'_1) + a_{21}\psi(e'_2) + \cdots + a_{m1}\psi(e'_m)$$
$$= a_{11}(b_{11}e''_1 + b_{21}e''_2 + \cdots + b_{k1}e''_k) + a_{21}(b_{12}e''_1 + b_{22}e''_2 + \cdots + b_{k2}e''_k) + \cdots$$
$$+ a_{m1}(b_{1m}e''_1 + b_{2m}e''_2 + \cdots + b_{km}e''_k)$$
$$= (b_{11}a_{11} + b_{12}a_{21} + \cdots + b_{1m}a_{m1})e''_1 + (b_{21}a_{11} + b_{22}a_{21} + \cdots + b_{2m}a_{m1})e''_2 + \cdots$$
$$+ (b_{k1}a_{11} + b_{k2}a_{21} + \cdots + b_{km}a_{m1})e''_k.$$

Забелязваме, че коефициентът пред e''_i е точно i-ият ред на матрицата B, умножен по първия стълб на матрицата A. Следователно първият стълб на матрицата на $\psi\varphi$ съвпада с първия стълб на матрицата BA. Аналогично се проверява и за останалите стълбове.

З а б е л е ж к а. Едва сега става ясно, че действието умножение на матрици, което дефинирахме в § 4 е естествено и всъщност произлиза от действието умножение на линейни изображения.

§ 15. Ранг и дефект на линейно изображение. Обратим линеен оператор

Ранг и дефект на линейно изображение. В първата част на параграфа V и V' ще бъдат фиксирани линейни пространства, а $\varphi : V \longrightarrow V'$ — линейно изображение.

Определение. *Под ядро на* φ *ще разбираме множеството* $\operatorname{Ker}\varphi = \{v \in V | \varphi(v) = 0\}$. *Под образ на* φ *ще разбираме множеството* $\operatorname{Im}\varphi = \{v' \in V' | \exists v \in V : \varphi(v) = v'\}$.

Директно се проверява, че $\operatorname{Ker}\varphi$ е подпространство на V, а $\operatorname{Im}\varphi$ е подпространство на V'.

Задача 1. *Да се докаже, че следните условия са еквивалентни:*

1) $\operatorname{Ker}\varphi = \{0\}$;

2) *ако* $v_1, v_2 \in V$ *и* $v_1 \neq v_2$, *то* $\varphi(v_1) \neq \varphi(v_2)$.

Определение. *Под ранг на* φ *ще разбираме числото* $r(\varphi) = \dim \operatorname{Im}\varphi$. *Под дефект на* φ *ще разбираме числото* $d(\varphi) = \dim \operatorname{Ker}\varphi$.

Твърдение 1. *Ако* V *и* V' *са крайномерни пространства и* A *е матрицата на* φ *в кои да са два базиса съответно на* V *и* V', *то* $r(\varphi) = r(A)$.

Д о к а з а т е л с т в о. Нека $e_1, \ldots, e_n$ и $e'_1, \ldots, e'_m$ са базиси съответно на V и V' и A е матрицата на φ в тези базиси. Знаем, че рангът на A е равен на ранга на системата от нейните вектори стълбове. Но стълбовете на A са съставени от координатите на $\varphi(e_1), \ldots, \varphi(e_n)$ (в базиса $e'_1, \ldots, e'_m$). Следователно $r(A) = r(\varphi(e_1), \ldots, \varphi(e_n)) = \dim l(\varphi(e_1), \ldots, \varphi(e_n))$.

От друга страна, тъй като всеки вектор от V е линейна комбинация на $e_1, \ldots, e_n$, то всеки вектор от $\operatorname{Im}\varphi$ е линейна комбинация на $\varphi(e_1), \ldots, \varphi(e_n)$, откъдето следва $\operatorname{Im}\varphi = l(\varphi(e_1), \ldots, \varphi(e_n))$. Тогава $r(\varphi) = \dim \operatorname{Im}\varphi = \dim l(\varphi(e_1), \ldots, \varphi(e_n))$. Следователно $r(\varphi) = r(A)$.

Теорема 2 (теорема за ранга и дефекта). *Ако* $\dim V = n < \infty$, *то* $r(\varphi) + d(\varphi) = n$.

Д о к а з а т е л с т в о. Нека $d(\varphi) = d$ и $e_1, \ldots, e_d$ е базис на $\operatorname{Ker}\varphi$. Да допълним тези линейно независими вектори до базис $e_1, \ldots, e_d, e_{d+1}, \ldots, e_n$ на V. Ще докажем, че векторите $\varphi(e_{d+1}), \dots, \varphi(e_n)$ образуват базис на $\operatorname{Im}\varphi$. Оттук ще следва, че $r(\varphi) = n-d$ и значи $r(\varphi) + d(\varphi) = n$.

(Ако $\operatorname{Im}\varphi = \{\mathbf{0}\}$, т. е. $\varphi = \mathbf{0}$, то $\operatorname{Ker}\varphi = V$ и значи $d(\varphi) = n$. Ако $\operatorname{Ker}\varphi = \{\mathbf{0}\}$, то векторите $e_1, \dots, e_d$ в следващите разсъждения просто липсват.)

Нека v' е произволен вектор от $\operatorname{Im}\varphi$ и v е вектор от V, за който $v' = \varphi(v)$. Нека $v = \lambda_1 e_1 + \dots + \lambda_d e_d + \lambda_{d+1} e_{d+1} + \dots + \lambda_n e_n$. Като вземем предвид, че векторите $e_1, \dots, e_d$ са от $\operatorname{Ker}\varphi$, получаваме $v' = \varphi(v) = \lambda_{d+1}\varphi(e_{d+1}) + \dots + \lambda_n\varphi(e_n)$. Така всеки вектор от $\operatorname{Im}\varphi$ е линейна комбинация на векторите $\varphi(e_{d+1}), \dots, \varphi(e_n)$.

Нека $\mu_{d+1}\varphi(e_{d+1}) + \dots + \mu_n\varphi(e_n) = \mathbf{0}$, следователно $\varphi(\mu_{d+1}e_{d+1} + \dots + \mu_n e_n) = \mathbf{0}$. Тогава векторът $\mu_{d+1}e_{d+1} + \dots + \mu_n e_n$ е от $\operatorname{Ker}\varphi$ и значи се изразява линейно чрез векторите $e_1, \dots, e_d$. Нека $\mu_{d+1}e_{d+1} + \dots + \mu_n e_n = \mu_1 e_1 + \dots + \mu_d e_d$, т. е. $-\mu_1 e_1 - \dots - \mu_d e_d + \mu_{d+1}e_{d+1} + \dots + \mu_n e_n = \mathbf{0}$. Тогава всички коефициенти, в частност и $\mu_{d+1}, \dots, \mu_n$ са равни на нула. Следователно векторите $\varphi(e_{d+1}), \dots, \varphi(e_n)$ са линейно независими и значи образуват базис на $\operatorname{Im}\varphi$.

* * *

Обратим линеен оператор. До края на параграфа V ще бъде фиксирано линейно пространство и $\dim V = n < \infty$.

За един линеен оператор $\varphi : V \longrightarrow V$ ще казваме, че е *обратим*, ако притежава обратен линеен оператор, т. е. такъв оператор $\varphi' : V \longrightarrow V$, че $\varphi\varphi' = \varphi'\varphi = \varepsilon$. Както и при матрици се доказва, че φ' еднозначно се определя чрез φ; ще го бележим с φ^{-1}. Всеки обратим линеен оператор е биекция и значи е изоморфизъм.

Твърдение 3. *Нека φ е обратим линеен оператор във V. Ако в даден базис на V на φ съответства матрица A, то A е обратима матрица и в същия базис на φ^{-1} съответства матрицата A^{-1}.*

Доказателство. Нека в дадения базис на φ^{-1} съответства матрица A'. Имаме $\varphi\varphi^{-1} = \varphi^{-1}\varphi = \varepsilon$. От твърдение 2 от § 14 получаваме $AA' = A'A = E$. Следователно A е обратима матрица и $A^{-1} = A'$.

Твърдение 4. *Нека $\varphi \in \operatorname{Hom} V$. Следните условия са еквивалентни:*

1) *φ е обратим линеен оператор;*
2) *$\operatorname{Ker}\varphi = \{\mathbf{0}\}$ (т. е. $d(\varphi) = 0$);*
3) *$\operatorname{Im}\varphi = V$ (т. е. $r(\varphi) = n$);*
4) *φ преобразува кой да е базис на V също в базис;*
5) *матрицата на φ във всеки базис на V е обратима.*

Доказателство. 1) $\Rightarrow$ 2). Нека φ е обратим и $v \in \operatorname{Ker}\varphi$. Тогава $v = (\varphi^{-1}\varphi)(v) = \varphi^{-1}(\varphi(v)) = \varphi^{-1}(\mathbf{0}) = \mathbf{0}$.

2) $\Rightarrow$ 3). Нека $\operatorname{Ker}\varphi = \{\mathbf{0}\}$, т. е. $d(\varphi) = 0$. От равенството $r(\varphi) + d(\varphi) = n$ следва $r(\varphi) = n$, т. е. $\operatorname{Im}\varphi = V$.

3) $\Rightarrow$ 4). Нека $\operatorname{Im}\varphi = V$ и $e_1, \dots, e_n$ е базис на V. Имаме $V = \operatorname{Im}\varphi = l(\varphi(e_1), \dots, \varphi(e_n))$. Освен това векторите $\varphi(e_1), \dots, \varphi(e_n)$ са линейно независими (в противен случай $\dim V < n$) и значи са базис на V.

4) $\Rightarrow$ 5). Нека $e_1, \dots, e_n$ е произволен базис на V и A е матрицата на φ в този базис. По условие векторите $\varphi(e_1), \dots, \varphi(e_n)$ също са базис на V. Рангът на матрицата A е равен на ранга на системата от нейните вектори стълбове, които са координатите на $\varphi(e_1), \dots, \varphi(e_n)$ (в базиса $e_1, \dots, e_n$). Така $r(A) = \dim l(\varphi(e_1), \dots, \varphi(e_n)) = \dim V = n$, което означава, че $\det A \neq 0$. Следователно A е обратима матрица.

5) $\Rightarrow$ 1). Нека $e_1, \dots, e_n$ е фиксиран базис на V и на φ съответства матрица A в този базис. По условие A е обратима матрица. Нека при зададения базис на матрицата A^{-1} съответства линеен оператор φ'. Тогава на операторите $\varphi\varphi'$ и $\varphi'\varphi$ съответстват матриците $AA^{-1} = E$ и $A^{-1}A = E$. Следователно $\varphi\varphi' = \varphi'\varphi = \varepsilon$, т.е. φ е обратим линеен оператор (и $\varphi^{-1} = \varphi'$).

З а б е л е ж к а. Ако V е безкрайномерно пространство, от 2), както и от 3) не следва обратимост на φ.

Задача 2. а) *Нека $\varphi, \tau \in \operatorname{Hom} V$ и τ е обратим оператор. Да се докаже, че* $r(\varphi) = r(\varphi\tau) = r(\tau\varphi)$.

б) *Нека $A, T \in M_n(F)$ и T е обратима матрица. Да се докаже, че* $r(A) = r(AT) = r(TA)$.

§ 16. Смяна на базиса

Нека V е крайномерно пространство. В този параграф ще видим как се сменят координатите на вектор и матрицата на линеен оператор при смяна на базиса на V.

Нека

$$e_1,\ e_2, \dots,\ e_n \tag{1}$$

и

$$f_1,\ f_2, \dots,\ f_n \tag{2}$$

са два базиса на V. Всеки вектор от втория базис се изразява линейно чрез векторите от първия базис. Нека

$$\begin{aligned} f_1 &= \tau_{11}e_1 + \tau_{21}e_2 + \cdots + \tau_{n1}e_n, \\ f_2 &= \tau_{12}e_1 + \tau_{22}e_2 + \cdots + \tau_{n2}e_n, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ f_n &= \tau_{1n}e_1 + \tau_{2n}e_2 + \cdots + \tau_{nn}e_n. \end{aligned} \tag{3}$$

Определение. *Матрицата*

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

ще наричаме матрица на прехода от базиса (1) *към базиса* (2).

Стълбовете на матрицата T са съставени от координатите на векторите от втория базис спрямо първия базис. Тогава стълбовете на матрицата T са линейно независими и значи $r(T) = n$. Следователно T е обратима матрица. Обратно, нека $T = (\tau_{ij})_{n\times n}$ е обратима матрица и $e_1, \dots, e_n$ е базис на V. Тогава векторите $f_1, \dots, f_n$, определени с равенствата (3) са линейно независими и значи образуват базис на V. Така всяка обратима матрица T е матрица на прехода от произволен базис $e_1, \dots, e_n$ на V към базиса $f_1, \dots, f_n$, определен с равенствата (3).

Нека $\tau : V \longrightarrow V$ е линеен оператор, който има матрица T **в базиса (1)**. От определението за матрица на линеен оператор следва $\tau(e_i) = f_i$ $(i = 1, 2, \ldots, n)$. Тъй като T е обратима матрица, то τ е обратим оператор и $\tau^{-1}(f_i) = e_i$ $(i = 1, 2, \ldots, n)$. Матрицата T^{-1} е матрица на прехода от базиса (2) към базиса (1).

При горните означения нека $v \in V$, $v = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$ и $v = \eta_1 f_1 + \eta_2 f_2 + \cdots + \eta_n f_n$. Да означим с ξ и η следните матрици стълбове:

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}.$$

Твърдение 1. *В сила е матричното равенство* $\xi = T\eta$.

Д о к а з а т е л с т в о. Имаме

$$\begin{aligned} v &= \eta_1 f_1 + \eta_2 f_2 + \cdots + \eta_n f_n \\ &= \eta_1(\tau_{11}e_1 + \tau_{21}e_2 + \cdots + \tau_{n1}e_n) + \eta_2(\tau_{12}e_1 + \tau_{22}e_2 + \cdots + \tau_{n2}e_n) + \cdots \\ &\qquad + \eta_n(\tau_{1n}e_1 + \tau_{2n}e_2 + \cdots + \tau_{nn}e_n) \\ &= (\tau_{11}\eta_1 + \tau_{12}\eta_2 + \cdots + \tau_{1n}\eta_n)e_1 + (\tau_{21}\eta_1 + \tau_{22}\eta_2 + \cdots + \tau_{2n}\eta_n)e_2 + \cdots \\ &\qquad + (\tau_{n1}\eta_1 + \tau_{n2}\eta_2 + \cdots + \tau_{nn}\eta_n)e_n. \end{aligned}$$

От друга страна, $v = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$. Като приравним коефициентите пред съответните базисни вектори, получаваме подробно разписано матричното равенство $\xi = T\eta$.

Твърдение 2. *Нека* $\varphi \in \operatorname{Hom} V$ *и на* φ *съответства матрица* $A = (a_{ij})_{n\times n}$ *в базиса* (1) *и матрица* $B = (b_{ij})_{n\times n}$ *в базиса* (2). *Тогава* $B = T^{-1}AT$.

Д о к а з а т е л с т в о. Нека τ е линейният оператор, който има матрица T **в базиса (1)**. Тъй като B е матрицата на φ **в базиса (2)**, то $\varphi(f_i) = b_{1i}f_1 + \cdots + b_{ni}f_n$ $(i = 1, \ldots, n)$. Прилагаме оператора τ^{-1} към това равенство, като вземем предвид, че $\tau^{-1}(f_i) = e_i$. Получаваме $\tau^{-1}\varphi(f_i) = b_{1i}e_1 + \cdots + b_{ni}e_n$. Заместваме в лявата част на последното равенство f_i с $\tau(e_i)$ и получаваме $\tau^{-1}\varphi\tau(e_i) = b_{1i}e_1 + \cdots + b_{ni}e_n$ $(i = 1, \ldots, n)$. Тези равенства означават, че операторът $\tau^{-1}\varphi\tau$ има матрица B **в базиса (1)**.

От друга страна, операторите φ и τ имат матрици съответно A и T **в базиса (1)**. Тогава операторът $\tau^{-1}\varphi\tau$ има матрица $T^{-1}AT$ също **в базиса (1)**. Следователно $B = T^{-1}AT$.

* * *

Определение. *Нека* $A, B \in M_n(F)$. *Ще казваме, че матриците* A *и* B *са подобни, ако съществува обратима матрица* $T \in M_n(F)$, *такава че* $B = T^{-1}AT$. *В този случай ще пишем* $A \sim B$.

З а б е л е ж к а. Ще отбележим, че релацията "подобие" на матрици е релация на еквивалентност. В частност, няма значение дали ще пишем $A \sim B$ или $B \sim A$.

От твърдение 2 е ясно, че матриците на един и същи оператор в различни базиси на V са подобни. Вярно е и обратното:

Твърдение 3. *Нека $A, B \in M_n(F)$ и $A \sim B$. Тогава съществува линеен оператор $\varphi \in \operatorname{Hom} V$ и два базиса на V, такива че матриците на φ в тези два базиса са съответно A и B.*

Доказателство. Нека $B = T^{-1}AT$ и $e_1, \dots, e_n$ е произволен базис на V. Нека φ е линейният оператор, който има матрица A в базиса $e_1, \dots, e_n$, а τ е линейният оператор, който има матрица T в същия базис. Операторът τ е обратим (тъй като матрицата T е обратима) и значи векторите $f_1 = \tau(e_1), \dots, f_n = \tau(e_n)$ също образуват базис на V. При това, T е матрицата на прехода от първия към втория базис. Според твърдение 2 матрицата на оператора φ в базиса $f_1, \dots, f_n$ е точно B.

Твърдение 4. *Подобните матрици имат:*

а) *равни детерминанти;*

б) *еднакви рангове.*

Доказателство. Нека $A \sim B$ и $B = T^{-1}AT$.

а) Използвайки теоремата за умножение на детерминанти (теорема 3 от § 5), последователно получаваме

$$\det B = \det(T^{-1}AT) = \det(T^{-1}).\det(A).\det T = (\det T)^{-1}.\det A.\det T = \det A.$$

б) Според твърдение 3 A и B са матрици на един и същи оператор φ в подходящи базиси на V. Сега според твърдение 1 от § 15 имаме $r(A) = r(\varphi) = r(B)$.

§ 17. Собствени вектори и собствени стойности. Инвариантни подпространства

Собствени вектори и собствени стойности. Нека $A = (a_{ij})_{n\times n}$ е матрица от $M_n(F)$. Да разгледаме матрицата

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

Детерминантата на тази матрица е полином на x от степен n с коефициенти от F. Действително, произведението на елементите, стоящи по главния диагонал е полином на x от степен n и старши коефициент, равен на $(-1)^n$. Всеки друг член от развитието на детерминантата не съдържа поне два елемента от главния диагонал и следователно е полином на x от степен, ненадминаваща $n-2$.

Задача 1. *Да се докаже, че коефициентът пред $(n-1)$-вата степен и свободният член на полинома $\det(A - xE)$ са равни съответно на $(-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})$ и $\det A$.*

Определение. *Полинома $\det(A - xE)$ ще наричаме характеристичен полином на матрицата A и ще го бележим с $f_A(x)$. Корените на този полином ще наричаме характеристични корени на A.*

Задача 2. *Да се докаже, че ако матрицата A е триъгълна (с нулеви елементи под главния диагонал), то характеристичните корени на A са елементите, стоящи по главния диагонал на A.*

Твърдение 1. *Подобните матрици имат едни и същи характеристични полиноми.*

Доказателство. Нека $A \sim B$ и $B = T^{-1}AT$. Прилагайки теоремата за умножение на детерминанти, последователно получаваме

$$f_B(x) = \det(B - xE) = \det(T^{-1}AT - xE) = \det(T^{-1}(A - xE)T)$$
$$= \det(T^{-1}).\det(A - xE).\det T = (\det T)^{-1}.\det(A - xE).\det T = \det(A - xE) = f_A(x).$$

Нека V е линейно пространство, $\dim V = n < \infty$ и $\varphi \in \operatorname{Hom} V$. Матриците на φ в различни базиси на V са подобни и следователно имат едни и същи характеристични полиноми. Това ни дава възможност да говорим за характеристичен полином на линеен оператор.

Определение. *Под характеристичен полином на линейния оператор φ ще разбираме характеристичния полином на матрицата на φ в кой да е базис на V. Корените на този полином ще наричаме характеристични корени на φ.*

Определение. *Нека V е линейно пространство и $\varphi \in \operatorname{Hom} V$. Един вектор $\boldsymbol{v}$ от V ще наричаме собствен вектор на φ, ако:*

1) $\boldsymbol{v} \neq 0$;

2) *съществува число λ от F, такова че $\varphi(\boldsymbol{v}) = \lambda \boldsymbol{v}$. Числото λ ще наричаме собствена стойност на φ. Ще казваме, че $\boldsymbol{v}$ е собствен вектор, съответстващ на λ, както и че λ е собствена стойност, съответстваща на $\boldsymbol{v}$.*

Задача 3. *Нека V е линейно пространство, $\varphi \in \operatorname{Hom} V$ и λ е собствена стойност на φ. Да се докаже, че множеството U от всички собствени вектори на φ, съответстващи на λ, заедно с нулевия вектор (т.е. $U = \{\boldsymbol{v} \in V \mid \varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}\}$) е подпространство на V.*

Теорема 2. *Нека V е крайномерно пространство над полето F и $\varphi \in \operatorname{Hom} V$. Тогава характеристичните корени на φ, лежащи в полето F, и само те, са всичките собствени стойности на φ.*

Доказателство. Нека $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ е базис на V и $A = (a_{ij})_{n\times n}$ е матрицата на φ в този базис. Нека $\lambda \in F$ е собствена стойност на φ и $\boldsymbol{v} = \xi_1\boldsymbol{e}_1 + \xi_2\boldsymbol{e}_2 + \cdots + \xi_n\boldsymbol{e}_n$ е собствен вектор на φ, съответстващ на λ. Тогава векторното равенство $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$ е еквивалентно на системата от равенства

$$\left|\begin{array}{l} a_{11}\xi_1 + a_{12}\xi_2 + \cdots + a_{1n}\xi_n = \lambda\xi_1 \\ a_{21}\xi_1 + a_{22}\xi_2 + \cdots + a_{2n}\xi_n = \lambda\xi_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \cdots + a_{nn}\xi_n = \lambda\xi_n \end{array}\right.$$

или все едно на

$$\left|\begin{array}{lll} (a_{11} - \lambda)\xi_1 + & a_{12}\xi_2 + \cdots + & a_{1n}\xi_n = 0 \\ a_{21}\xi_1 + & (a_{22} - \lambda)\xi_2 + \cdots + & a_{2n}\xi_n = 0 \\ \dots\dots\dots & \dots\dots\dots & \dots\dots\dots \\ a_{n1}\xi_1 + & a_{n2}\xi_2 + \cdots + & (a_{nn} - \lambda)\xi_n = 0. \end{array}\right. \tag{1}$$

По определение векторът $\boldsymbol{v}$ е ненулев. Така ненулевата n-орка $(\xi_1, \xi_2, \ldots, \xi_n)$ е решение на хомогенна система с матрица $A - \lambda E$. Следователно $\det(A - \lambda E) = 0$, т.е. $f_A(\lambda) = 0$. Така λ е характеристичен корен (лежащ в F) на линейния оператор φ.

Обратно, нека λ е характеристичен корен на φ, лежащ в F. Тогава $\det(A - \lambda E) = 0$, откъдето следва, че хомогенната система с матрица $A - \lambda E$ има ненулево решение

$(\xi_1, \xi_2, \dots, \xi_n)$, т. е. изпълняват се равенствата (1). Тогава за ненулевия вектор $\boldsymbol{v} = \xi_1 \boldsymbol{e}_1 + \xi_2 \boldsymbol{e}_2 + \dots + \xi_n \boldsymbol{e}_n$ е в сила $\varphi(\boldsymbol{v}) = \lambda \boldsymbol{v}$, т. е. $\boldsymbol{v}$ е собствен вектор на φ, а λ е съответстващата му собствена стойност.

Забележка. Някои от корените на характеристичния полином на φ може и да не са в полето F, но според теоремата на Даламбер (в края на параграф 02) всички корени са комплексни числа. Така, ако V е линейно пространство над $\mathbb{C}$, всички характеристични корени на φ са собствени стойности на φ. Ако обаче полето е друго, например $\mathbb{R}$, то φ може изобщо да няма собствени стойности.

Теорема 3. *Нека V е линейно пространство и $\varphi \in \operatorname{Hom} V$. Ако $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ са собствени вектори на φ, съответстващи на различни собствени стойности, то $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ са линейно независими вектори.*

Доказателство. Нека $\lambda_1, \dots, \lambda_k$ са собствените стойности на φ, съответстващи на $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$. Ще проведем индукция по k. При $k = 1$ векторът $\boldsymbol{v}_1$ по определение е ненулев и значи е линейно независим.

Нека $k > 1$ и

$$\mu_1 \boldsymbol{v}_1 + \dots + \mu_{k-1} \boldsymbol{v}_{k-1} + \mu_k \boldsymbol{v}_k = \boldsymbol{0}. \tag{2}$$

Като приложим φ към това равенство и вземем предвид, че $\varphi(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i$ $(i = 1, \dots, k)$, получаваме равенството

$$\mu_1 \lambda_1 \boldsymbol{v}_1 + \dots + \mu_{k-1} \lambda_{k-1} \boldsymbol{v}_{k-1} + \mu_k \lambda_k \boldsymbol{v}_k = \boldsymbol{0}. \tag{3}$$

Да умножим равенството (2) с λ_k и да го извадим от равенството (3). Получаваме $\mu_1(\lambda_1 - \lambda_k)\boldsymbol{v}_1 + \dots + \mu_{k-1}(\lambda_{k-1} - \lambda_k)\boldsymbol{v}_{k-1} = \boldsymbol{0}$. Според индукционното предположение всички коефициенти в това равенство са равни на нула и тъй като $\lambda_i \neq \lambda_k$ за $i = 1, \dots, k-1$, получаваме $\mu_1 = \dots = \mu_{k-1} = 0$. Сега от равенството (2) (тъй като $\boldsymbol{v}_k \neq \boldsymbol{0}$) получаваме и $\mu_k = 0$. Следователно векторите $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ са линейно независими.

Следствие 4. *Нека V е линейно пространство, $\dim V = n < \infty$ и $\varphi \in \operatorname{Hom} V$. Ако φ има n различни собствени стойности $\lambda_1, \dots, \lambda_n$, то съществува базис на V, в който матрицата D на φ е диагонална, по-точно*

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Доказателство. Нека $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ са собствени вектори на φ, съответстващи на $\lambda_1, \dots, \lambda_n$. Според теорема 3 тези вектори са линейно независими и тъй като са n на брой, те образуват базис на V. Сега остава да си припомним определението за матрица на линеен оператор в даден базис. Така търсеният базис е $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$.

Забележка. Всъщност само от определенията следва, че матрицата на един линеен оператор φ в даден базис е диагонална тогава и само тогава, когато базисните вектори са собствени вектори на φ (съответстващи на не непременно различни собствени стойности). В следващия параграф ще видим, че ако не поставяме ограничения за характеристичните корени на φ (освен те да са в основното поле F), съществува базис, в който матрицата на φ е "почти диагонална".

Следствие 5. *Нека A е квадратна матрица от ред n, която има n различни характеристични корена. Тогава A е подобна на диагонална матрица. (В този случай казваме, че A се привежда в диагонален вид.)*

Д о к а з а т е л с т в о. Нека $V = \mathbb{C}^n$ и $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ е фиксиран базис на V. Тогава A е матрицата в този базис на линеен оператор φ, който има n различни собствени стойности. Според следствие 4, съществува базис $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ на V, в който матрицата D на φ е диагонална и, както знаем, $A \sim D$.

* * *

Инвариантни подпространства. Нека V е линейно пространство, U е подпространство на V и $\varphi \in \operatorname{Hom} V$. Ще казваме, че подпространството U е *инвариантно относно* φ или φ*-инвариантно*, ако за всеки вектор $\boldsymbol{u} \in U$ имаме $\varphi(\boldsymbol{u}) \in U$.

Ако U е инвариантно подпространство относно φ, можем да разглеждаме действието на φ само върху U. Този линеен оператор в линейното пространство U ще наричаме *ограничение на* φ *върху* U и ще бележим с $\varphi_{|_U}$, т.е. $\varphi_{|_U}(\boldsymbol{u}) = \varphi(\boldsymbol{u})$ за $\boldsymbol{u} \in U$.

П р и м е р и. 1. Подпространствата $\{\boldsymbol{0}\}$ и V са инвариантни относно действието на всеки оператор от $\operatorname{Hom} V$.

2. Подпространствата $\operatorname{Im}\varphi$ и $\operatorname{Ker}\varphi$ са инвариантни относно φ.

3. Нека λ е собствена стойност на φ.

а) Подпространството $U = \{\boldsymbol{v} \in V \mid \varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}\}$ от задача 3 е инвариантно относно φ (в него φ действа като хомотетия).

б) Нека $\boldsymbol{v}$ е един собствен вектор, съответстващ на λ и $U = l(\boldsymbol{v})$. Тогава U е едномерно φ-инвариантно подпространство на V.

От пример 3.б) виждаме, че ако един линеен оператор притежава собствена стойност, то той има едномерно инвариантно подпространство (вярно е и обратното). В частност, всеки оператор в крайномерно ненулево пространство над $\mathbb{C}$ има едномерно инвариантно подпространство. Следващата теорема изяснява положението за линейни пространства над $\mathbb{R}$.

Теорема 6. *Нека V е крайномерно ненулево линейно пространство над $\mathbb{R}$ и $\varphi \in \operatorname{Hom} V$. Тогава φ има едномерно или двумерно инвариантно подпространство.*

Д о к а з а т е л с т в о. Нека $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ е базис на V и φ има матрица A в този базис. Знаем, че изображението, съпоставящо на всеки вектор от V наредената n-орка от координатите му във фиксирания базис е изоморфизъм от V към $\mathbb{R}^n$. Всъщност, можем да отъждествяваме векторите от V с наредените n-орки реални числа, съставени от техните координати. Така (допускайки известна волност) просто ще считаме, че $V = \mathbb{R}^n$. При това отъждествяване имаме $\boldsymbol{e}_1 = (1, 0, 0, \dots, 0)$, $\boldsymbol{e}_2 = (0, 1, 0, \dots, 0)$, …, $\boldsymbol{e}_n = (0, 0, 0, \dots, 1)$. По този начин V става **подмножество** на линейното пространство $\mathbb{C}^n$, като векторите $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n$ са базис на $\mathbb{C}^n$ над $\mathbb{C}$. Да означим с $\overline{\varphi}$ линейния оператор в $\mathbb{C}^n$, който има матрица A в този базис. Имаме $\overline{\varphi}(\boldsymbol{v}) = \varphi(\boldsymbol{v})$ за всеки вектор $\boldsymbol{v} \in V$. Според теоремата на Даламбер (в края на параграф 02) $\overline{\varphi}$ има собствена стойност $\lambda \in \mathbb{C}$. Нека $\lambda = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$. Нека $\boldsymbol{c} = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{C}^n$ е собствен вектор на $\overline{\varphi}$, съответстващ на λ и $\xi_k = \eta_k + i\zeta_k$ ($\eta_k, \zeta_k \in \mathbb{R}$, $k = 1, 2, \dots, n$). Да означим $\boldsymbol{a} = (\eta_1, \eta_2, \dots, \eta_n)$, $\boldsymbol{b} = (\zeta_1, \zeta_2, \dots, \zeta_n) \in V$. Тогава $\boldsymbol{c} = \boldsymbol{a} + i\boldsymbol{b}$. Прилагайки $\overline{\varphi}$ към това равенство, получаваме

$$\overline{\varphi}(\boldsymbol{c}) = \overline{\varphi}(\boldsymbol{a}) + i\overline{\varphi}(\boldsymbol{b}) = \varphi(\boldsymbol{a}) + i\varphi(\boldsymbol{b}). \tag{4}$$

От друга страна имаме

$$\overline{\varphi}(\boldsymbol{c}) = \lambda\boldsymbol{c} = (\alpha + i\beta)(\boldsymbol{a} + i\boldsymbol{b}) = (\alpha\boldsymbol{a} - \beta\boldsymbol{b}) + i(\beta\boldsymbol{a} + \alpha\boldsymbol{b}). \tag{5}$$

Приравнявайки реалните и имагинерните части в десните страни на равенствата (4) и (5), получаваме

$$\begin{aligned}\varphi(\boldsymbol{a}) &= \alpha\boldsymbol{a} - \beta\boldsymbol{b},\\ \varphi(\boldsymbol{b}) &= \beta\boldsymbol{a} + \alpha\boldsymbol{b}.\end{aligned} \tag{6}$$

Нека $U = l(\boldsymbol{a}, \boldsymbol{b})$ (U е подпространство на V). Равенствата (6) показват, че U е инвариантно относно φ. При това, $\dim U \leq 2$ и $U \neq \{\boldsymbol{0}\}$. С това теоремата е доказана.

Задача 4. *Да се докаже, че ако числото λ от доказателството на теорема 6 не е реално, то $\dim U = 2$.*

§ 18. Жорданова нормална форма: съществуване

В следващите два параграфа V ще бъде крайномерно линейно пространство над F, а φ — линеен оператор от $\operatorname{Hom} V$. Ще напомним, че ако U е φ-инвариантно подпространство на V, с $\varphi_{|_U}$ бележим ограничението на φ върху U.

Преди да формулираме основния резултат, ще направим няколко бележки.

1. Знаем, че подпространствата $\operatorname{Ker}\varphi$ и $\operatorname{Im}\varphi$ са φ-инвариантни. Вярно е също и че за всяко число $\lambda \in F$ и за всяко естествено число m подпространствата $\operatorname{Ker}(\varphi - \lambda\varepsilon)^m$ и $\operatorname{Im}(\varphi - \lambda\varepsilon)^m$ също са φ-инвариантни. Ще направим проверка само за $\operatorname{Ker}(\varphi - \lambda\varepsilon)^m$. Нека $\boldsymbol{v} \in \operatorname{Ker}(\varphi - \lambda\varepsilon)^m$, т.е. $(\varphi - \lambda\varepsilon)^m(\boldsymbol{v}) = \boldsymbol{0}$. Операторът $\lambda\varepsilon$ комутира с всеки оператор и значи $\varphi - \lambda\varepsilon$ комутира с φ. Тогава $(\varphi - \lambda\varepsilon)^m(\varphi(\boldsymbol{v})) = \varphi((\varphi - \lambda\varepsilon)^m(\boldsymbol{v})) = \varphi(\boldsymbol{0}) = \boldsymbol{0}$, т.е. $\varphi(\boldsymbol{v}) \in \operatorname{Ker}(\varphi - \lambda\varepsilon)^m$.

2. Нека пространството V е директна сума на подпространствата си U_1 и U_2, които са φ-инвариантни и $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ е базис на U_1, а $\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ е базис на U_2. Тогава векторите $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k; \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ са базис на V. Нека A_1 е матрицата на $\varphi_{|_{U_1}}$ в базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$, а A_2 е матрицата на $\varphi_{|_{U_2}}$ в базиса $\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$. Векторите $\varphi(\boldsymbol{e}_1), \dots, \varphi(\boldsymbol{e}_k)$ са в U_1 и значи са линейна комбинация на векторите $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$. Аналогично векторите $\varphi(\boldsymbol{e}_{k+1}), \dots, \varphi(\boldsymbol{e}_n)$ са линейна комбинация на $\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$. Сега от определението за матрица на линеен оператор следва, че матрицата A на φ в базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k; \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ е клетъчно диагонална, по-точно $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$. Аналогично твърдение е вярно и ако V е сума на повече от две подпространства.

3. Нека U е φ-инвариантно подпространство и $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ е базис на U. Да допълним този базис до базис $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k; \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ на V. Нека A_1 е матрицата на $\varphi_{|_U}$ в базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$. Както в предната бележка се вижда, че матрицата A на φ в базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k; \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ е клетъчно триъгълна, по-точно $A = \begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}$ ($A_1 \in M_k(F)$, $A_2 \in M_{n-k}(F)$, $A_3 \in F_{k\times(n-k)}$). Оттук получаваме $A - xE_n = \begin{pmatrix} A_1 - xE_k & A_3 \\ 0 & A_2 - xE_{n-k} \end{pmatrix}$. От лема 2 от § 5 (след съответна модификация) следва $\det(A - xE_n) = \det(A_1 - xE_k).\det(A_2 - xE_{n-k})$, т.е. $f_A(x) = f_{A_1}(x) f_{A_2}(x)$. В частност, всички характеристични корени на оператора $\varphi_{|_U} \in \operatorname{Hom} U$ са характеристични корени и на $\varphi \in \operatorname{Hom} V$.

Определение. *Под жорданова клетка от ред n, съответстваща на числото λ ще* *разбираме квадратната матрица от ред n*

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

Примери.

$$J_1(2) = (2),\ J_2(2) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},\ J_3(2) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},\ J_4(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Определение. *Под жорданова матрица от ред n ще разбираме всяка квадратна матрица от ред n от вида*

$$J = \begin{pmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & 0 & & J_{n_k}(\lambda_k) \end{pmatrix},$$

където $J_{n_i}(\lambda_i)$ *са жорданови клетки от ред* n_i*, съответстващи на* λ_i $(i = 1, 2, \dots, k)$. *Числата* $n_1, n_2, \dots, n_k$*, както и* $\lambda_1, \lambda_2, \dots, \lambda_k$ *не са непременно различни. Очевидно* $n_1 + n_2 + \cdots + n_k = n$.

Ясно е, че всяка жорданова клетка е жорданова матрица (но не и обратното). Също така, всяка жорданова матрица е триъгълна и характеристичните ѝ корени са числата, стоящи по главния ѝ диагонал. По-точно, характеристичният полином на J е полиномът $f_J(x) = (\lambda_1 - x)^{n_1}(\lambda_2 - x)^{n_2} \dots (\lambda_k - x)^{n_k}$ (напомняме, че числата λ_i не са непременно различни).

Примери. Всяка диагонална матрица е жорданова матрица (съставена от жорданови клетки от ред 1). Други примери са матриците

$$\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 3 \end{array}\right),\ \left(\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array}\right),\ \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 \end{array}\right)$$

Определение. *Под жорданов базис за линейния оператор* φ *ще разбираме всеки базис на* V*, в който матрицата на* φ *е жорданова.*

Нека $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ е жорданов базис за линейния оператор φ и матрицата на φ в този базис е само една жорданова клетка, съответстваща на числото 0, т.е. матрицата $J_n(0)$. От определението за матрица на линеен оператор следва, че $\varphi(\boldsymbol{v}_1) = \boldsymbol{0}$, $\varphi(\boldsymbol{v}_2) = \boldsymbol{v}_1$,

$\dots, \varphi(\boldsymbol{v}_n) = \boldsymbol{v}_{n-1}$. Схематично действието на φ върху базиса ще записваме по следния начин:

$$\varphi: \quad \begin{array}{c} \mathbf{0} \\ \uparrow \\ \boldsymbol{v}_1 \\ \uparrow \\ \boldsymbol{v}_2 \\ \uparrow \\ \vdots \\ \uparrow \\ \boldsymbol{v}_n \end{array} .$$

Нека сега матрицата J на φ в базиса $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ се състои от k жорданови клетки $J_{n_1}(0), \dots, J_{n_k}(0)$, съответстващи на числото 0. Тогава базисните вектори могат да се разбият на k групи: $\boldsymbol{v}_1^1, \dots, \boldsymbol{v}_1^{n_1}; \dots; \boldsymbol{v}_k^1, \dots, \boldsymbol{v}_k^{n_k}$, съответстващи на жордановите клетки $J_{n_1}(0), \dots, J_{n_k}(0)$ и върху всяка група φ действа по описания по-горе начин. Схематично действието на φ върху базиса ще записваме в следната таблица:

$$\varphi: \quad \begin{array}{ccc} \mathbf{0} & & \mathbf{0} \\ \uparrow & & \uparrow \\ \boldsymbol{v}_1^1 & & \boldsymbol{v}_k^1 \\ \uparrow & \dots & \uparrow \\ \vdots & & \vdots \\ \uparrow & & \uparrow \\ \boldsymbol{v}_1^{n_1} & & \boldsymbol{v}_k^{n_k} \end{array} . \tag{1}$$

Вярно е и обратното: ако можем да разбием векторите в един базис на V на няколко групи, върху които φ действа по показания в таблица (1) начин, то този базис е жорданов базис за оператора φ.

Сега вече можем да формулираме основния резултат.

Теорема 1 (Жордан). *Нека V е линейно пространство над F, $\dim V = n < \infty$, $\varphi \in \operatorname{Hom} V$ и всички характеристични корени на φ са в F. Тогава са в сила следните две твърдения:*

I. **Съществуване**. *Съществува жорданов базис за φ.*

II. **Единственост**. *Ако J и J' са матриците на φ в два различни жорданови базиса, то J и J' се състоят от едни и същи жорданови клетки, евентуално разположени в различен ред.*

Като преминем от операторен на матричен език, получаваме.

Следствие 2. *Ако всички характеристични корени на една матрица $A \in M_n(F)$ са в F, то A е подобна на жорданова матрица. Две жорданови матрици са подобни тогава и само тогава, когато се състоят от едни и същи жорданови клетки, евентуално разположени в различен ред.* .

Прието е да се казва, че матрицата A се привежда в *жорданова нормална форма*.

Прецизирането на доказателството на следствие 2 предоставяме на читателя.

* * *

Доказателство на теоремата на Жордан (съществуване). Ще започнем със следната лема на Фитинг, която представлява и самостоятелен интерес.

Лема 3. *Нека V е крайномерно пространство и $\varphi \in \operatorname{Hom} V$. Тогава съществува естествено число m, такова че $V = \operatorname{Ker}\varphi^m \oplus \operatorname{Im}\varphi^m$.*

Доказателство. Очевидни са включванията $V \supseteq \operatorname{Im}\varphi \supseteq \operatorname{Im}\varphi^2 \supseteq \dots$. Тъй като V е крайномерно пространство, само краен брой от тези включвания могат да са строги. Следователно съществува естествено число m, такова че $\operatorname{Im}\varphi^m = \operatorname{Im}\varphi^{m+1} = \dots$. В частност $\operatorname{Im}\varphi^m = \operatorname{Im}\varphi^{2m}$.

Нека $\boldsymbol{v}$ е произволен вектор от V. Имаме $\varphi^m(\boldsymbol{v}) \in \operatorname{Im}\varphi^m$, а значи и $\varphi^m(\boldsymbol{v}) \in \operatorname{Im}\varphi^{2m}$. Следователно съществува вектор $\boldsymbol{w} \in V$, такъв че $\varphi^m(\boldsymbol{v}) = \varphi^{2m}(\boldsymbol{w})$ или $\varphi^m(\boldsymbol{v} - \varphi^m(\boldsymbol{w})) = \boldsymbol{0}$. Това означава, че $\boldsymbol{v} - \varphi^m(\boldsymbol{w}) \in \operatorname{Ker}\varphi^m$. Очевидно е равенството $\boldsymbol{v} = (\boldsymbol{v} - \varphi^m(\boldsymbol{w})) + \varphi^m(\boldsymbol{w})$. При това $\boldsymbol{v} - \varphi^m(\boldsymbol{w}) \in \operatorname{Ker}\varphi^m$, а $\varphi^m(\boldsymbol{w}) \in \operatorname{Im}\varphi^m$. Така получихме $V = \operatorname{Ker}\varphi^m + \operatorname{Im}\varphi^m$. За да докажем, че тази сума е директна, остава да проверим, че $\operatorname{Ker}\varphi^m \cap \operatorname{Im}\varphi^m = \{\boldsymbol{0}\}$. Имаме (теорема 1 от § 10)

$$\dim V = \dim(\operatorname{Ker}\varphi^m + \operatorname{Im}\varphi^m) = \dim\operatorname{Ker}\varphi^m + \dim\operatorname{Im}\varphi^m - \dim(\operatorname{Ker}\varphi^m \cap \operatorname{Im}\varphi^m).$$

От друга страна теоремата за ранга и дефекта ни дава $\dim V = \dim\operatorname{Ker}\varphi^m + \dim\operatorname{Im}\varphi^m$. Като сравним десните страни на двете равенства, получаваме $\dim(\operatorname{Ker}\varphi^m \cap \operatorname{Im}\varphi^m) = 0$, т.е. $\operatorname{Ker}\varphi^m \cap \operatorname{Im}\varphi^m = \{\boldsymbol{0}\}$.

Доказателството на теоремата на Жордан (съществуване) ще разбием на три стъпки.

* * *

Стъпка 1. Редукция към случая на линеен оператор с една единствена собствена стойност.

Нека $\lambda_1, \dots, \lambda_k$ са всичките различни собствени стойности на φ $(1 \le k \le n)$ и нека $k > 1$. Прилагайки лемата на Фитинг за оператора $\varphi - \lambda_1\varepsilon$, получаваме $V = V_1 \oplus V'$, където $V_1 = \operatorname{Ker}(\varphi - \lambda_1\varepsilon)^m$, $V' = \operatorname{Im}(\varphi - \lambda_1\varepsilon)^m$. Пространствата V_1 и V' са φ-инвариантни (бележка 1 в началото на параграфа).

Лема 4. а) *Всички собствени вектори на φ, съответстващи на λ_1 са във V_1 и λ_1 е единствената собствена стойност на оператора $\varphi_{|V_1} \in \operatorname{Hom} V_1$.*

б) *Всички собствени вектори на φ, съответстващи на $\lambda_2, \dots, \lambda_k$ са във V' и (както следва от а)) λ_1 не е собствена стойност на оператора $\varphi_{|V'} \in \operatorname{Hom} V'$.*

Доказателство. а) Нека $\boldsymbol{v}_1 \in V$ е собствен вектор на φ, съответстващ на λ_1. Равенството $\varphi(\boldsymbol{v}_1) = \lambda_1\boldsymbol{v}_1$ записваме във вида $(\varphi - \lambda_1\varepsilon)(\boldsymbol{v}_1) = \boldsymbol{0}$. Оттук получаваме $(\varphi - \lambda_1\varepsilon)^m(\boldsymbol{v}_1) = \boldsymbol{0}$, което означава, че $\boldsymbol{v}_1 \in \operatorname{Ker}(\varphi - \lambda_1\varepsilon)^m = V_1$.

Нека λ е произволна собствена стойност на оператора $\varphi_{|V_1}$ и $\boldsymbol{v} \in V_1$ е собствен вектор, съответстващ на λ. От равенството $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$ изваждаме равенството $\lambda_1\varepsilon(\boldsymbol{v}) = \lambda_1\boldsymbol{v}$ и получаваме $(\varphi - \lambda_1\varepsilon)(\boldsymbol{v}) = (\lambda - \lambda_1)\boldsymbol{v}$. Прилагаме към това равенство оператора $\varphi - \lambda_1\varepsilon$: $(\varphi - \lambda_1\varepsilon)^2(\boldsymbol{v}) = (\lambda - \lambda_1)(\varphi - \lambda_1\varepsilon)(\boldsymbol{v}) = (\lambda - \lambda_1)^2\boldsymbol{v}$. Към полученото равенство отново прилагаме оператора $\varphi - \lambda_1\varepsilon$ и т.н. докато получим равенството $(\varphi - \lambda_1\varepsilon)^m(\boldsymbol{v}) = (\lambda - \lambda_1)^m\boldsymbol{v}$. Но $\boldsymbol{v} \in V_1$, т.е. $(\varphi - \lambda_1\varepsilon)^m(v) = \boldsymbol{0}$. Следователно и $(\lambda - \lambda_1)^m\boldsymbol{v} = \boldsymbol{0}$. Тъй като $\boldsymbol{v} \neq \boldsymbol{0}$, то $\lambda = \lambda_1$ и значи λ_1 е единствена собствена стойност на оператора $\varphi_{|V_1}$.

б) Нека $\boldsymbol{v}_2 \in V$ е собствен вектор на φ, съответстващ на λ_2. Както в доказателството на а) (замествайки λ с λ_2) получаваме $(\varphi - \lambda_1\varepsilon)^m(\boldsymbol{v}_2) = (\lambda_2 - \lambda_1)^m\boldsymbol{v}_2$. Но $(\varphi - \lambda_1\varepsilon)^m(\boldsymbol{v}_2) \in \operatorname{Im}(\varphi - \lambda_1\varepsilon)^m = V'$, следователно $(\lambda_2 - \lambda_1)^m\boldsymbol{v}_2 \in V'$. Тъй като $(\lambda_2 - \lambda_1)^m \neq 0$, то и $\boldsymbol{v}_2 \in V'$. Така всеки собствен вектор на φ, съответстващ на λ_2 е във V'. Същите са разсъжденията и за останалите собствени стойности на φ. Лемата е доказана.

Прилагайки лемата на Фитинг за пространството V', получаваме $V' = V_2 \oplus V''$, където $V_2 = \operatorname{Ker}(\varphi - \lambda_2\varepsilon)^s$, $V'' = \operatorname{Im}(\varphi - \lambda_2\varepsilon)^s$ за подходящо естествено число s. Тогава

$V = V_1 \oplus V_2 \oplus V''$. После прилагаме лемата на Фитинг за V'' и т. н. Така след краен брой стъпки достигаме до следното твърдение:

Пространството V е директна сума $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, където $V_i = \operatorname{Ker}(\varphi - \lambda_i \varepsilon)^{m_i}$ за подходящи естествени числа m_i $(i = 1, \ldots, k)$. При това, пространствата V_i са φ-инвариантни и за всяко $i = 1, \ldots, k$ операторът $\varphi_{|V_i} \in \operatorname{Hom} V_i$ има една единствена собствена стойност, равна на λ_i.

Да допуснем сега, че сме намерили жорданов базис за всеки един от операторите $\varphi_{|V_i} \in \operatorname{Hom} V_i$, т. е. че матрицата на $\varphi_{|V_i}$ в този базис е жорданова $(i = 1, \ldots, k)$. От бележка 2 в началото на параграфа следва, че матрицата на φ в базиса на V, съставен от тези базиси на $V_1, \ldots, V_k$ е също жорданова. Следователно достатъчно е да докажем теоремата на Жордан (съществуване) за оператор с една единствена собствена стойност.

* * *

Стъпка 2. Редукция към случая на оператор с единствена собствена стойност, равна на нула.

Вече ще считаме, че самият оператор $\varphi \in \operatorname{Hom} V$ има единствена собствена стойност λ. (В този случай характеристичният полином на φ е $(\lambda - x)^n$.) Да означим $\psi = \varphi - \lambda\varepsilon$ или все едно $\varphi = \psi + \lambda\varepsilon$. Нека $\boldsymbol{v}$ е собствен вектор на φ. От равенството $\varphi(\boldsymbol{v}) = \lambda\boldsymbol{v}$ следва $(\varphi - \lambda\varepsilon)(\boldsymbol{v}) = \boldsymbol{0}$, т. е. $\psi(\boldsymbol{v}) = \boldsymbol{0} = 0.\boldsymbol{v}$. Така числото 0 е собствена стойност на оператора ψ (и $\boldsymbol{v}$ е собствен вектор).

Нека μ е произволна собствена стойност на ψ и $\boldsymbol{w}$ е собствен вектор на ψ, съответстващ на μ. Имаме $\psi(\boldsymbol{w}) = \mu\boldsymbol{w}$, откъдето $\varphi(\boldsymbol{w}) = (\psi + \lambda\varepsilon)(\boldsymbol{w}) = (\mu + \lambda)\boldsymbol{w}$. Така $\boldsymbol{w}$ е собствен вектор на φ, съответстващ на числото $\mu + \lambda$. Но φ има единствена собствена стойност, равна на λ и значи $\mu = 0$. Следователно операторът ψ има единствена собствена стойност, равна на нула.

Нека съществува жорданов базис за ψ и J_0 е жордановата матрица на ψ в този базис. От равенството $\varphi = \psi + \lambda\varepsilon$ следва, че матрицата на φ в същия базис е $J = J_0 + \lambda E$, която очевидно също е жорданова. Така, ако намерим жорданов базис за оператора $\psi = \varphi - \lambda\varepsilon$, ще сме намерили жорданов базис и за φ. Следователно достатъчно е да докажем теоремата на Жордан (съществуване) за оператор с единствена собствена стойност, равна на нула.

* * *

Стъпка 3. Доказателство на теоремата на Жордан (съществуване) за оператор с единствена собствена стойност, равна на нула.

Ще считаме, че операторът $\varphi \in \operatorname{Hom} V$ има единствена собствена стойност, равна на нула.

Ще проведем индукция по $n = \dim V$. При $n = 1$ няма какво да доказваме. Нека $n > 1$ и $\boldsymbol{v}$ е собствен вектор на φ. Имаме $\varphi(\boldsymbol{v}) = 0.\boldsymbol{v} = \boldsymbol{0}$. Следователно $\operatorname{Ker}\varphi \neq \{\boldsymbol{0}\}$. От теоремата за ранга и дефекта следва, че $\dim \operatorname{Im}\varphi < n$. Пространството $\operatorname{Im}\varphi$ е φ-инвариантно и от бележка 3 в началото на параграфа следва, че характеристичните корени на $\varphi|_{\operatorname{Im}\varphi} \in \operatorname{Hom} \operatorname{Im}\varphi$ са характеристични корени и на $\varphi \in \operatorname{Hom} V$. Следователно $\varphi|_{\operatorname{Im}\varphi}$ има единствена собствена стойност, равна на нула. Сега можем да приложим индукционното предположение за оператора $\varphi|_{\operatorname{Im}\varphi}$: за него съществува жорданов базис в пространството $\operatorname{Im}\varphi$. (Ако $\operatorname{Im}\varphi = \{\boldsymbol{0}\}$, т. е. $\operatorname{Ker}\varphi = V$, то φ е нулевият оператор и няма какво да доказваме.) Нека върху този базис операторът $\varphi|_{\operatorname{Im}\varphi}$ действа, както е показано в следната таблица:

$$
\varphi|_{\operatorname{Im}\varphi}:\quad
\begin{array}{ccc}
0 & & 0\\
\uparrow & & \uparrow\\
v_1^1 & & v_k^1\\
\uparrow & \dots & \uparrow\\
\vdots & & \vdots\\
\uparrow & & \uparrow\\
v_1^{n_1} & & v_k^{n_k}
\end{array}
\tag{2}
$$

Ще допълним този базис до жорданов базис за оператора $\varphi \in \operatorname{Hom} V$.

Очевидно векторите $v_1^1, \dots, v_k^1$ са от $\operatorname{Ker}\varphi$ (може да се докаже, че те са базис на пространството $\operatorname{Ker}\varphi \cap \operatorname{Im}\varphi$). Първо допълваме (ако е необходимо) тези вектори до базис $v_1^1, \dots, v_k^1; v_{k+1}^1, \dots, v_t^1$ на пространството $\operatorname{Ker}\varphi$. По-нататък, тъй като всички вектори от таблицата (2) са от $\operatorname{Im}\varphi$, то съществуват вектори $v_1^{n_1+1}, \dots, v_k^{n_k+1} \in V$, такива че $\varphi(v_1^{n_1+1}) = v_1^{n_1}, \dots, \varphi(v_k^{n_k+1}) = v_k^{n_k}$. Присъединяваме новите вектори към таблица (2) и получаваме таблицата

$$
\varphi:\quad
\begin{array}{ccc|ccc}
0 & & 0 & 0 & & 0\\
\uparrow & & \uparrow & & & \\
v_1^1 & & v_k^1 & v_{k+1}^1 & \dots & v_t^1\\
\uparrow & & \uparrow & & & \\
\vdots & \dots & \vdots & & & \\
\uparrow & & \uparrow & & & \\
v_1^{n_1} & & v_k^{n_k} & & & \\
\hdashline
\uparrow & & \uparrow & & & \\
v_1^{n_1+1} & & v_k^{n_k+1} & & &
\end{array}
\tag{3}
$$

Ще докажем, че векторите v_i^j от таблицата (3) са базис на V. Тогава те ще бъдат жорданов базис за φ.

Нека броят на векторите v_i^j в таблицата (3) е равен на n'. Тогава числото n' е сума на следните три числа:

1) броят на векторите v_i^j от таблицата (2), който е равен на $\dim \operatorname{Im}\varphi$;

2) броят на векторите $v_{k+1}^1, \dots, v_t^1$, който е равен на $t - k = \dim \operatorname{Ker}\varphi - k$;

3) броят на векторите $v_1^{n_1+1}, \dots, v_k^{n_k+1}$, който е равен на k.

Следователно $n' = \dim \operatorname{Im}\varphi + (\dim \operatorname{Ker}\varphi - k) + k = \dim \operatorname{Im}\varphi + \dim \operatorname{Ker}\varphi$. От теоремата за ранга и дефекта следва $n' = \dim V$.

Остава да докажем, че векторите v_i^j от таблицата (3) са линейно независими. Нека

$$
\sum_{i=1}^{k} \sum_{j=2}^{n_i+1} \mu_i^j v_i^j + \sum_{l=1}^{t} \mu_l^1 v_l^1 = 0 \tag{4}
$$

(векторите от първия ред на таблицата са в отделна сума). Прилагаме оператора φ към това равенство. При действието на φ векторите от всеки ред на таблицата (3) се "придвижват" с един ред нагоре по същия стълб, т.е. $\varphi(v_i^j) = v_i^{j-1}$ за $i = 1, \ldots, k$; $j = 2, \ldots, n_i + 1$ и $\varphi(v_l^1) = 0$ за $l = 1, \ldots, t$. Следователно получаваме

$$\sum_{i=1}^{k}\sum_{j=2}^{n_i+1}\mu_i^j v_i^{j-1} = 0.$$

Но в това равенство присъстват само вектори от таблицата (2), които са част от базиса на $\operatorname{Im}\varphi$ и значи са линейно независими. Следователно $\mu_i^j = 0$ за $i = 1, \ldots, k$; $j = 2, \ldots, n_i + 1$. Така равенството (4) приема вида $\sum_{l=1}^{t}\mu_l^1 v_l^1 = 0$. Но векторите $v_1^1, \ldots, v_t^1$ са базис на $\operatorname{Ker}\varphi$ и значи са линейно независими. Следователно и $\mu_l^1 = 0$, $l = 1, \ldots, t$. Така векторите v_i^j от таблицата (3) са линейно независими и следователно образуват базис на V.

С това теоремата на Жордан (съществуване) е доказана.

§ 19. Жорданова нормална форма: единственост

Ще използваме означенията от предния параграф. Ще напомним, че с $r(\varphi)$ и $d(\varphi)$ бележим съответно ранга и дефекта на линейния оператор φ.

Нека $\lambda_1, \ldots, \lambda_k$ са различните характеристични корени на φ и J е матрицата на φ в един жорданов базис. Да означим с $N_p(\lambda_i)$ броя на клетките от ред p, съответстващи на λ_i $(i = 1, \ldots, k)$. Матрицата J еднозначно (с точност до реда на клетките) се определя от:

1) числата $\lambda_1, \ldots, \lambda_k$;

2) числата $N_p(\lambda_i)$ $(i = 1, \ldots, k;\ p = 1, 2, \ldots\ )$.

Както знаем, характеристичните корени не зависят от избора на базиса. За числата $N_p(\lambda_i)$ ще докажем равенството

$$N_p(\lambda_i) = r\left[(\varphi - \lambda_i\varepsilon)^{p-1}\right] - 2r[(\varphi - \lambda_i\varepsilon)^p] + r\left[(\varphi - \lambda_i\varepsilon)^{p+1}\right] \qquad (1)$$

(считаме, че $r[(\varphi - \lambda_i\varepsilon)^0] = n$). Дясната страна на това равенство също не зависи от избора на базиса. С това единствеността в теоремата на Жордан ще бъде доказана.

Да означим $\psi_i = \varphi - \lambda_i\varepsilon$. Като използваме теоремата за ранга и дефекта, равенството (1) приема вида

$$N_p(\lambda_i) = \left[n - d(\psi_i^{p-1})\right] - 2\left[n - d(\psi_i^p)\right] + \left[n - d(\psi_i^{p+1})\right] = 2d(\psi_i^p) - d(\psi_i^{p-1}) - d(\psi_i^{p+1}). \quad (2)$$

Ще докажем равенството (1) записано в този вид, при това, без ограничение на общността, ще считаме, че $i = 1$.

Както и в предния параграф, ще сведем доказателството към случая на оператор с единствена собствена стойност, равна на нула. След евентуално разместване на жордановите клетки, можем да считаме, че на първите s места $(s \geq 1)$ по главния диагонал на матрицата J стои характеристичният корен λ_1, а на останалите места стоят $\lambda_2, \ldots, \lambda_k$. Нека базисът, в който операторът φ има матрица J е $v_1, \ldots, v_s;\ v_{s+1}, \ldots, v_n$. Да означим $V_1 = l(v_1, \ldots, v_s)$, $V' = l(v_{s+1}, \ldots, v_n)$. Имаме $V = V_1 \oplus V'$. От начина, по който действа φ върху базисните вектори следва, че подпространствата V_1 и V' са φ-инвариантни.

(В действителност ще се окаже, че това са точно подпространствата, означени по същия начин в стъпка 1 от доказателството на теоремата на Жордан (съществуване) в предния параграф.) Тези подпространства са инвариантни и относно действието на оператора $\psi_1 = \varphi - \lambda_1\varepsilon$. Матрицата на ψ_1 във фиксирания базис е $J_1 = J - \lambda_1 E$, която също е жорданова, има клетъчната структура на J и първите s елемента по главния ѝ диагонал са равни на нула, а останалите са ненулеви. Следователно операторът $\psi_1|_{V_1} \in \operatorname{Hom} V_1$ има единствена собствена стойност, равна на нула и числото нула не е собствена стойност на оператора $\psi_1|_{V'} \in \operatorname{Hom} V'$. В частност $\operatorname{Ker}(\psi_1|_{V'}) = \{\mathbf{0}\}$.

Нека $\boldsymbol{v} \in \operatorname{Ker}\psi_1$ и $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}'$, където $\boldsymbol{v}_1 \in V_1$, $\boldsymbol{v}' \in V'$. Имаме $\mathbf{0} = \psi_1(\boldsymbol{v}) = \psi_1(\boldsymbol{v}_1) + \psi_1(\boldsymbol{v}')$. При това $\psi_1(\boldsymbol{v}_1) \in V_1$, $\psi_1(\boldsymbol{v}') \in V'$. Тъй като сумата $V = V_1 \oplus V'$ е директна, то $\psi_1(\boldsymbol{v}') = \mathbf{0}$ (както и $\psi_1(\boldsymbol{v}_1) = \mathbf{0}$). Но от $\operatorname{Ker}(\psi_1|_{V'}) = \{\mathbf{0}\}$ следва $\boldsymbol{v}' = \mathbf{0}$ и значи $\boldsymbol{v} = \boldsymbol{v}_1 \in V_1$. Така $\operatorname{Ker}\psi_1 \subseteq V_1$, откъдето $\operatorname{Ker}\psi_1 = \operatorname{Ker}\psi_1 \cap V_1 = \operatorname{Ker}(\psi_{1|_{V_1}})$. Следователно достатъчно е да докажем равенството (2) (при $i = 1$) за пространството V_1 и за оператора $\psi_{1|_{V_1}}$.

Без да ограничаваме общността, ще считаме, че V просто съвпада с V_1. Така ψ_1 има единствена собствена стойност, равна на нула. Както знаем от предния параграф, векторите от фиксирания базис могат да се разбият на групи, върху които ψ_1 действа по показания в следната таблица начин:

$$\psi_1 : \quad \begin{array}{ccc} \mathbf{0} & & \mathbf{0} \\ \uparrow & & \uparrow \\ \boldsymbol{v}_1^1 & & \boldsymbol{v}_t^1 \\ \uparrow & \dots & \uparrow \\ \vdots & & \vdots \\ \uparrow & & \uparrow \\ \boldsymbol{v}_1^{n_1} & & \boldsymbol{v}_t^{n_t}. \end{array} \qquad (3)$$

Да означим с N_p^0 броя на стълбовете с дължина p в таблицата (3). От начина на действие на ψ_1 върху базисните вектори следва, че N_p^0 е точно броят на жордановите клетки от ред p в матрицата на ψ_1. Следователно трябва да докажем равенството

$$N_p^0 = 2d(\psi_1^p) - d(\psi_1^{p-1}) - d(\psi_1^{p+1}). \qquad (2')$$

Лема 1. *Векторите* $\boldsymbol{v}_i^j$*, стоящи в първите* p *реда на таблицата* (3) *(т. е. горният индекс на които не надминава* p*) образуват базис на пространството* $\operatorname{Ker}\psi_1^p$.

Доказателство. Нека $\boldsymbol{v}_i^j$ е произволен вектор от j-ия ред на таблицата (3). От начина на действие на ψ_1 е ясно, че $\psi_1^j(\boldsymbol{v}_i^j) = \mathbf{0}$, а значи и $\psi_1^r(\boldsymbol{v}_i^j) = \mathbf{0}$ за $r \geq j$. В същото време, ако $r < j$, то $\psi_1^r(\boldsymbol{v}_i^j) = \boldsymbol{v}_i^{j-r} \neq \mathbf{0}$.

От казаното следва първо, че векторите от първите p реда на таблицата са от $\operatorname{Ker}\psi_1^p$. Освен това те са линейно независими, тъй като са част от базис. Остава да докажем, че всеки вектор от $\operatorname{Ker}\psi_1^p$ е тяхна линейна комбинация.

Нека $\boldsymbol{v} \in \operatorname{Ker}\psi_1^p$ и

$$\boldsymbol{v} = \sum_{i=1}^{t}\sum_{j=1}^{n_i} \mu_i^j \boldsymbol{v}_i^j. \qquad (4)$$

Да допуснем, че в това равенство участва с ненулев коефициент вектор $\boldsymbol{v}_s^m$, за който $m > p$. Като приложим ψ_1^p към равенството (4), в дясната страна ще получим линейна комбинация на векторите $\boldsymbol{v}_i^j$, в която ще участва с ненулев коефициент векторът $\boldsymbol{v}_s^{m-p}$.

Следователно дясната страна ще бъде различна от нулевия вектор. От друга страна $\psi_1^p(v) = \mathbf{0}$, което е противоречие. Така всички горни индекси на векторите с ненулеви коефициенти в дясната страна на равенството (4) не надминават p, което означава, че v е линейна комбинация на векторите от първите p реда на таблицата (3).

Следствие 2. *Броят на стълбовете в таблицата* (3) *с дължина, по-голяма или равна на p е равен на $d(\psi_1^p) - d(\psi_1^{p-1})$.*

Доказателство. Един стълб има дължина поне p тогава и само тогава, когато е "успял да достигне" до p-ия ред, т. е. има свой "представител" вектор в p-ия ред. Така броят на стълбовете с дължина, по-голяма или равна на p съвпада с броя на векторите от p-ия ред на таблицата. Но той се получава като от броя векторите от първите p реда на таблицата извадим броя на векторите от първите $p-1$ реда. От лема 1 непосредствено следва, че тази разлика е равна на $d(\psi_1^p) - d(\psi_1^{p-1})$.

Интересуващото ни число N_p^0 се получава като от броя на стълбовете с дължина поне p извадим броя на стълбовете с дължина, по-голяма от p. Казано по друг начин, числото N_p^0 е равно на броя на стълбовете с дължина, по-голяма или равна на p минус броя на стълбовете с дължина, по-голяма или равна на $p+1$. Сега от следствие 2 получаваме

$$N_p^0 = \left[d(\psi_1^p) - d(\psi_1^{p-1})\right] - \left[d(\psi_1^{p+1}) - d(\psi_1^p)\right] = 2d(\psi_1^p) - d(\psi_1^{p-1}) - d(\psi_1^{p+1}),$$

което е точно равенството (2').

С това единствеността в теоремата на Жордан е доказана.

§ 20. Теорема на Хамилтън — Кейли

Нека $f(x) = a_0x^k + \dots + a_{k-1}x + a_k \in F[x]$ и $A \in M_n(F)$. Под стойност на полинома $f(x)$ при $x = A$ ще разбираме матрицата $f(A) = a_0A^k + \dots + a_{k-1}A + a_kE \in M_n(F)$. Ще казваме, че матрицата A е корен на полинома $f(x)$, ако $f(A) = \mathbf{0} \in M_n(F)$.

Теорема 1 (Хамилтън — Кейли). *Всяка матрица $A \in M_n(F)$ е корен на характеристичния си полином $f_A(x)$.*

Преди да докажем теоремата, ще направим няколко бележки.

1. Нека $A, B, T \in M_n(F)$, като T е обратима матрица и $\lambda \in F$. Тогава

$$T^{-1}(A+B)T = T^{-1}AT + T^{-1}BT,$$
$$T^{-1}(AB)T = (T^{-1}AT)(T^{-1}BT),$$
$$T^{-1}(\lambda A)T = \lambda T^{-1}AT.$$

Оттук следва, че ако $f(x) \in F[x]$, то $f(T^{-1}AT) = T^{-1}f(A)T$. Като умножим това равенство отляво с T, а отдясно с T^{-1}, получаваме $f(A) = Tf(T^{-1}AT)T^{-1}$. От двете равенства следва, че $f(A) = \mathbf{0}$ тогава и само тогава, когато $f(T^{-1}AT) = \mathbf{0}$.

2. Нека матрицата на оператора φ във фиксиран базис $v_1, \dots, v_n$ на линейното пространство V е жорданова клетка от ред n, съответстваща на числото 0, т. е. матрицата $J_n(0)$. От начина на действие на φ върху базисните вектори следва, че $\varphi^n(v_i) = \mathbf{0}$, $i = 1, \dots, n$. Следователно φ^n е нулевият оператор. Тогава $J_n(0)^n = \mathbf{0}$ и значи $J_n(0)^r = \mathbf{0}$ при $r \geq n$.

3. Нека

$$J = \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ 0 & & & J_t \end{pmatrix}$$

е (не непременно жорданова) клетъчно диагонална матрица. От правилата за действия с матрици следва, че ако $f(x) \in F[x]$, то

$$f(J) = \begin{pmatrix} f(J_1) & & 0 \\ & f(J_2) & \\ & & \ddots \\ 0 & & & f(J_t) \end{pmatrix}.$$

В частност, ако $f(J_i) = \mathbf{0}$ $(i = 1, 2, \dots, t)$, то и $f(J) = \mathbf{0}$.

Доказателство на теоремата на Хамилтън — Кейли. Тъй като $A \in M_n(F)$ и $F \subseteq \mathbb{C}$, то $A \in M_n(\mathbb{C})$. Характеристичните корени на A са комплексни числа, следователно можем да приложим следствие 2 от § 18. Нека J е жордановата нормална форма за матрицата A $(J \in M_n(\mathbb{C}))$. Матриците A и J са подобни и имат еднакви характеристични полиноми. От бележка 1 следва, че е достатъчно да докажем теоремата за матрицата J.

Нека $\lambda_1, \dots, \lambda_k$ са различните характеристични корени на J и числото λ_i се среща n_i пъти по главния диагонал на J $(i = 1, \dots, k)$. Тогава характеристичният полином на J е

$$f_J(x) = \det(J - xE) = (\lambda_1 - x)^{n_1} \dots (\lambda_k - x)^{n_k}.$$

Нека I е жорданова клетка на J, съответстваща (например) на λ_1. Тогава матрицата I е от ред най-много n_1 и $I - \lambda_1 E$ е жорданова клетка, съответстваща на числото нула. От бележка 2 следва, че $(I - \lambda_1 E)^{n_1} = \mathbf{0}$, а значи и $(\lambda_1 E - I)^{n_1} = \mathbf{0}$. Тогава

$$f_J(I) = (\lambda_1 E - I)^{n_1} \dots (\lambda_k E - I)^{n_k} = \mathbf{0}.$$

(Тук с една и съща буква E означаваме единични матрици от различни редове.)

Горното разсъждение е в сила за всяка жорданова клетка на матрицата J. Сега от бележка 3 следва $f_J(J) = \mathbf{0}$. Теоремата е доказана.

Глава IV. Евклидови пространства

В четвърта глава (освен в последния параграф) всички линейни пространства ще бъдат над полето реалните числа $\mathbb{R}$.

§ 21. Евклидови пространства

Определение. *Нека V е линейно пространство над $\mathbb{R}$. Ще казваме, че във V е въведено скаларно произведение, ако на всеки два вектора $\boldsymbol{a}, \boldsymbol{b} \in V$ е съпоставено реално число, което ще означаваме с $(\boldsymbol{a}, \boldsymbol{b})$, като се удовлетворяват следните свойства (аксиоми) (тук $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ са произволни вектори от V, а λ е произволно реално число):*

I. $(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a})$.
II. $(\boldsymbol{a} + \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{c}) + (\boldsymbol{b}, \boldsymbol{c})$.
III. $(\lambda \boldsymbol{a}, \boldsymbol{b}) = \lambda(\boldsymbol{a}, \boldsymbol{b})$.
IV. $(\boldsymbol{a}, \boldsymbol{a}) \geq 0$ *като* $(\boldsymbol{a}, \boldsymbol{a}) = 0$ *точно когато* $\boldsymbol{a} = \boldsymbol{0}$.

Ако в едно пространство е въведено скаларно произведение, ще казваме, че това пространство е евклидово.

Примери. 1. Знаем, че можем да интерпретираме елементите на пространствата $\mathbb{R}^2$ и $\mathbb{R}^3$ като вектори, чието начало съвпада с началото на фиксирана координатна система. Скаларното произведение на вектори, дефинирано в курса по геометрия превръща тези пространства в евклидови.

Полезно е, когато се формулират твърдения, отнасящи се за евклидови пространства, човек да си представя съответната геометрична ситуация в тримерното евклидово пространство $\mathbb{R}^3$.

2. Нека $V = \mathbb{R}^n$ и $\boldsymbol{a} = (\xi_1, \xi_2, \dots, \xi_n)$, $\boldsymbol{b} = (\eta_1, \eta_2, \dots, \eta_n) \in V$. Полагаме $(\boldsymbol{a}, \boldsymbol{b}) = \xi_1\eta_1 + \xi_2\eta_2 + \dots + \xi_n\eta_n \in \mathbb{R}$. Непосредствено се проверява, че така дефинираната операция е скаларно произведение.

Ако V е произволно n-мерно пространство над $\mathbb{R}$, при фиксиран базис то се "превръща" в $\mathbb{R}^n$. Така всяко крайномерно пространство може да се направи евклидово.

3. Нека a и b са реални числа, $a < b$ и $V = C[a, b]$ е множеството от всички непрекъснати в интервала $[a, b]$ функции, приемащи реални стойности. С обичайните операции събиране на функции и умножение на функция с реално число V се превръща в линейно пространство над $\mathbb{R}$. Нека $f(x)$, $g(x) \in V$. Полагаме $(f(x), g(x)) = \int_a^b f(x)g(x)\,dx \in \mathbb{R}$. От анализа знаем, че така дефинираната операция удовлетворява аксиомите за скаларно произведение.

До края на параграфа V ще бъде евклидово пространство.

От аксиома III при $\lambda = 0$ следва равенството $(\boldsymbol{0}, \boldsymbol{b}) = 0$ за всеки вектор $\boldsymbol{b} \in V$. Също така, като приложим неколкократно аксиомите I, II и III, получаваме, че за произволни две линейни комбинации $\sum_i \lambda_i \boldsymbol{a}_i$ и $\sum_j \mu_j \boldsymbol{b}_j$ е в сила равенството

$$\Big(\sum_i \lambda_i \boldsymbol{a}_i, \sum_j \mu_j \boldsymbol{b}_j\Big) = \sum_{i,j} \lambda_i \mu_j (\boldsymbol{a}_i, \boldsymbol{b}_j).$$

Следващите няколко определения обобщават добре известни геометрични понятия, свързани с векторите в равнината и тримерното евклидово пространство.

Под *дължина* на вектора $\boldsymbol{a}$ ще разбираме числото $|\boldsymbol{a}| = \sqrt{(\boldsymbol{a}, \boldsymbol{a})}$. От аксиома IV следва, че $|\boldsymbol{a}|$ е неотрицателно реално число и $|\boldsymbol{a}| = 0$ точно когато $\boldsymbol{a} = \boldsymbol{0}$. Ако $\lambda \in \mathbb{R}$, имаме

$$|\lambda \boldsymbol{a}| = \sqrt{(\lambda \boldsymbol{a}, \lambda \boldsymbol{a})} = \sqrt{\lambda^2 (\boldsymbol{a}, \boldsymbol{a})} = |\lambda||\boldsymbol{a}|.$$

Ще казваме, че векторът $\boldsymbol{e}$ е *единичен*, ако $|\boldsymbol{e}| = 1$. От равенството $|\lambda \boldsymbol{a}| = |\lambda||\boldsymbol{a}|$ непосредствено следва, че ако $\boldsymbol{a}$ е произволен ненулев вектор, то векторът $\boldsymbol{a}_0 = \dfrac{1}{|\boldsymbol{a}|}\boldsymbol{a}$ е единичен. Прието е да се казва, че сме *нормирали* вектора $\boldsymbol{a}$.

Ще казваме, че два вектора $\boldsymbol{a}$ и $\boldsymbol{b}$ са ортогонални, ако $(\boldsymbol{a}, \boldsymbol{b}) = 0$. Нулевият вектор е ортогонален на всеки вектор от V. Обратно, ако един вектор е ортогонален на всеки вектор от V, то той (бидейки ортогонален на себе си) е нулевият вектор. Един базис на V ще наричаме *ортогонален*, ако базисните вектори са два по два ортогонални и *ортонормиран*, ако освен това са с дължина единица. Така един базис $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ е ортонормиран, ако $(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij} = \begin{cases} 1 & \text{при } i = j \\ 0 & \text{при } i \neq j \end{cases}$.

От всеки ортогонален базис, след нормиране на базисните вектори, се получава ортонормиран базис.

Твърдение 1 (теорема на Питагор). *Ако векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са два по два ортогонални, то $|\boldsymbol{a}_1 + \dots + \boldsymbol{a}_k|^2 = |\boldsymbol{a}_1|^2 + \dots + |\boldsymbol{a}_k|^2$.*

Доказателство. Вземайки предвид, че $(\boldsymbol{a}_i, \boldsymbol{a}_j) = 0$ при $i \neq j$, получаваме

$$|\boldsymbol{a}_1 + \dots + \boldsymbol{a}_k|^2 = (\boldsymbol{a}_1 + \dots + \boldsymbol{a}_k, \boldsymbol{a}_1 + \dots + \boldsymbol{a}_k) = (\boldsymbol{a}_1, \boldsymbol{a}_1) + \dots + (\boldsymbol{a}_k, \boldsymbol{a}_k) = |\boldsymbol{a}_1|^2 + \dots + |\boldsymbol{a}_k|^2.$$

Твърдение 2. *Всяка система $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ от два по два ортогонални ненулеви вектори е линейно независима.*

Доказателство. Нека $\lambda_1 \boldsymbol{a}_1 + \dots + \lambda_k \boldsymbol{a}_k = \boldsymbol{0}$. Тъй като векторите са два по два ортогонални, умножавайки това равенство скаларно по $\boldsymbol{a}_i$ $(1 \leq i \leq k)$, получаваме $\lambda_i (\boldsymbol{a}_i, \boldsymbol{a}_i) = 0$. Но $(\boldsymbol{a}_i, \boldsymbol{a}_i) \neq 0$, следователно $\lambda_i = 0$. Така всички коефициенти в горната линейна комбинация са равни на нула и значи векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са линейно независими.

Ще отбележим, че ако векторите $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ в n-мерното евклидово пространство V удовлетворяват равенствата $(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}$ $(i, j = 1, \dots, n)$, то те образуват ортонормиран базис на V. Действително, тези вектори са два по два ортогонални и с дължина единица и според предното твърдение са линейно независими, т. е. образуват базис на V.

Твърдение 3. а) (Ортогонализация по метода на Грам — Шмид). *Нека $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са линейно независими вектори. Тогава съществуват вектори $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$, които са два по два ортогонални и за всяко $i = 1, \dots, k$ е в сила $l(\boldsymbol{a}_1, \dots, \boldsymbol{a}_i) = l(\boldsymbol{e}_1, \dots, \boldsymbol{e}_i)$. По-точно, векторът $\boldsymbol{e}_i$ е линейна комбинация от вида $\boldsymbol{e}_i = \boldsymbol{a}_i + \lambda_1 \boldsymbol{e}_1 + \dots + \lambda_{i-1} \boldsymbol{e}_{i-1}$. При това, ако първите няколко вектора $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ $(m \leq k)$ са два по два ортогонални, то $\boldsymbol{e}_1 = \boldsymbol{a}_1, \dots, \boldsymbol{e}_m = \boldsymbol{a}_m$.*

б) *Нека $\dim V = n < \infty$. Тогава всяка система $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ от два по два ортогонални ненулеви вектори от V може да се допълни до ортогонален базис $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k; \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ на V. В частност (след нормиране на векторите $\boldsymbol{e}_i$) V притежава ортонормиран базис.*

Доказателство. а) Ще построим векторите $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ на k стъпки. При всяка стъпка ще бъде изпълнено $l(\boldsymbol{a}_1, \dots, \boldsymbol{a}_i) = l(\boldsymbol{e}_1, \dots, \boldsymbol{e}_i)$, $i = 1, \dots, k$. Тъй като векторите $a_1, \dots, a_k$ са линейно независими, то $i = \dim l(a_1, \dots, a_i) = \dim l(e_1, \dots, e_i)$. Следователно векторите $e_1, \dots, e_k$ също ще бъдат линейно независими и в частност, ненулеви. Така за всяко $i = 1, \dots, k$ ще имаме $(e_i, e_i) \neq 0$.

Нека $e_1 = a_1$. Да положим $e_2 = a_2 + \lambda_1 e_1$. Ще подберем числото λ_1 така, че векторът e_2 да е ортогонален на вектора e_1, т. е.

$$0 = (e_2, e_1) = (a_2 + \lambda_1 e_1, e_1) = (a_2, e_1) + \lambda_1 (e_1, e_1).$$

Оттук намираме $\lambda_1 = -\dfrac{(a_2, e_1)}{(e_1, e_1)}$. Очевидно векторите e_1 и e_2 са линейни комбинации на векторите a_1 и a_2 и обратно. Следователно $l(a_1, a_2) = l(e_1, e_2)$. При това, ако $(a_2, a_1) = 0$, то $\lambda_1 = 0$ и значи $e_2 = a_2$.

Да допуснем, че вече сме построили векторите $e_1, \dots, e_{i-1}$, като $l(a_1, \dots, a_{i-1}) = l(e_1, \dots, e_{i-1})$ и ако векторите $a_1, \dots, a_{i-1}$ са били два по два ортогонални, то $e_1 = a_1$, $\dots, e_{i-1} = a_{i-1}$. Ще търсим вектора e_i във вида

$$e_i = a_i + \lambda_1 e_1 + \dots + \lambda_{i-1} e_{i-1}. \tag{1}$$

Коефициентите $\lambda_1, \dots, \lambda_{i-1}$ ще определим така, че векторът e_i да е ортогонален на векторите $e_1, \dots, e_{i-1}$. Нека $1 \leq t \leq i-1$. Да умножим равенството (1) скаларно с вектора e_t. Като вземем предвид, че векторите $e_1, \dots, e_{i-1}$ са два по два ортогонални, получаваме

$$0 = (e_i, e_t) = (a_i + \lambda_1 e_1 + \dots + \lambda_t e_t + \dots + \lambda_{i-1} e_{i-1}, e_t) = (a_i, e_t) + \lambda_t (e_t, e_t).$$

Оттук намираме $\lambda_t = -\dfrac{(a_i, e_t)}{(e_t, e_t)}$. По допускане $l(a_1, \dots, a_{i-1}) = l(e_1, \dots, e_{i-1})$ и освен това $e_i = a_i + \lambda_1 e_1 + \dots + \lambda_{i-1} e_{i-1} \in l(e_1, \dots, e_{i-1}, a_i) = l(a_1, \dots, a_{i-1}, a_i)$; $a_i = e_i - \lambda_1 e_1 - \dots - \lambda_{i-1} e_{i-1} \in l(e_1, \dots, e_{i-1}, e_i)$. Следователно $l(a_1, \dots, a_i) = l(e_1, \dots, e_i)$. При това, ако векторите $a_1, \dots, a_i$ са два по два ортогонални, то по допускане $e_t = a_t$, откъдето $\lambda_t = 0$ за $t = 1, \dots, i-1$ и значи $e_i = a_i$.

Продължавайки този процес, построяваме търсената ортогонална система вектори $e_1, \dots, e_k$.

б) Според твърдение 2 векторите $e_1, \dots, e_k$ са линейно независими. Допълваме ги до базис $e_1, \dots, e_k; a_{k+1}, \dots, a_n$ на V. Необходимия ортогонален базис ще получим като ортогонализираме тези вектори по метода на Грам — Шмид. Тъй като векторите $e_1, \dots, e_k$ са два по два ортогонални, от а) следва, че на първите k стъпки ортогонализацията ще работи "на празен ход", т. е. първите k вектора в намерения базис ще бъдат отново векторите $e_1, \dots, e_k$. Твърдението е доказано.

Нека $e_1, \dots, e_n$ е ортонормиран базис на V и $a = \xi_1 e_1 + \xi_2 e_2 + \dots + \xi_n e_n$, $b = \eta_1 e_1 + \eta_2 e_2 + \dots + \eta_n e_n$. Като вземем предвид, че $(e_i, e_j) = \delta_{ij}$, получаваме

$$(a, b) = \xi_1 \eta_1 + \xi_2 \eta_2 + \dots + \xi_n \eta_n. \tag{2}$$

Нещо повече, *един базис $e_1, \dots, e_n$ на V е ортонормиран тогава и само тогава, когато скаларното произведение се задава чрез равенството* (2).

Тъй като всяко крайномерно пространство притежава ортонормиран базис, горните разсъждения показват, че за всяко естествено число n съществува в някакъв смисъл единствено n-мерно евклидово пространство. Това наблюдение се прецизира в теорема 4 по-долу.

Определение. *Нека V и V' са евклидови пространства. Ще казваме, че изображението $\varphi : V \longrightarrow V'$ е изоморфизъм на евклидови пространства, ако:*

1) *φ е изоморфизъм на линейни пространства;*

2) *за всеки два вектора $\boldsymbol{a}$, $\boldsymbol{b} \in V$ е изпълнено $(\boldsymbol{a}, \boldsymbol{b}) = (\varphi(\boldsymbol{a}), \varphi(\boldsymbol{b}))$.*

Ако съществува такова изображение, ще казваме, че V и V' са изоморфни като евклидови пространства и ще използваме познатото означение $V \cong V'$.

Теорема 4. *Нека V и V' са крайномерни евклидови пространства. Тогава $V \cong V'$ точно когато $\dim V = \dim V'$. Така всяко n-мерно евклидово пространство е изоморфно на $\mathbb{R}^n$.*

Доказателство. Знаем, че ако $V \cong V'$, то $\dim V = \dim V'$.

Обратно, нека $\dim V = \dim V' = n$ и $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ и $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n$ са ортонормирани базиси съответно на V и V'. Знаем, че съществува (единствен) изоморфизъм на линейни пространства $\varphi : V \longrightarrow V'$, за който $\varphi(\boldsymbol{e}_i) = \boldsymbol{e}'_i$, $i = 1, \dots, n$. Нека $\boldsymbol{a} = \sum\limits_{i=1}^{n} \xi_i \boldsymbol{e}_i$, $\boldsymbol{b} = \sum\limits_{i=1}^{n} \eta_i \boldsymbol{e}_i$ са произволни вектори от V. Тогава $\varphi(\boldsymbol{a}) = \sum\limits_{i=1}^{n} \xi_i \boldsymbol{e}'_i$, $\varphi(\boldsymbol{b}) = \sum\limits_{i=1}^{n} \eta_i \boldsymbol{e}'_i$. Тъй като и двата базиса са ортонормирани, то

$$(\boldsymbol{a}, \boldsymbol{b}) = \left(\sum_{i=1}^{n} \xi_i \boldsymbol{e}_i, \sum_{i=1}^{n} \eta_i \boldsymbol{e}_i \right) = \sum_{i=1}^{n} \xi_i \eta_i = \left(\sum_{i=1}^{n} \xi_i \boldsymbol{e}'_i, \sum_{i=1}^{n} \eta_i \boldsymbol{e}'_i \right) = (\varphi(\boldsymbol{a}), \varphi(\boldsymbol{b})).$$

Следователно φ е изоморфизъм на евклидови пространства.

* * *

Нека U е подпространство на V. Ако един вектор $\boldsymbol{a} \in V$ е ортогонален на всички вектори от U, за краткост ще пишем $(\boldsymbol{a}, U) = 0$. Ако всеки вектор от едно подпространство U_1 е ортогонален на всеки вектор от друго подпространство U_2, ще пишем $(U_1, U_2) = 0$.

Определение. *Под ортогонално допълнение на подпространството U на V ще разбираме множеството $U^{\perp} = \{ \boldsymbol{a} \in V \mid (\boldsymbol{a}, U) = 0 \}$.*

Директно от определението се проверява, че $U^{\perp}$ също е подпространство на V.

Твърдение 5. *Нека V е крайномерно евклидово пространство и U е подпространство на V. Тогава $V = U \oplus U^{\perp}$.*

Доказателство. Нека $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ е ортонормиран базис на U. Да го допълним до ортонормиран базис $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k; \boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ на V и да означим $U' = l(\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n)$. Знаем, че $V = U \oplus U'$. Ще докажем, че $U^{\perp} = U'$, с което твърдението ще бъде доказано.

Тъй като $(\boldsymbol{e}_i, \boldsymbol{e}_t) = 0$ за $i = 1, \dots, k$; $t = k+1, \dots, n$, то $(U, U') = 0$, т.е. $U' \subseteq U^{\perp}$. Нека $\boldsymbol{u}$ е произволен вектор от $U^{\perp}$ и $\boldsymbol{u} = \lambda_1 \boldsymbol{e}_1 + \dots + \lambda_k \boldsymbol{e}_k + \lambda_{k+1} \boldsymbol{e}_{k+1} + \dots + \lambda_n \boldsymbol{e}_n$. Имаме $(\boldsymbol{e}_i, \boldsymbol{u}) = 0$ за $i = 1, \dots, k$. От друга страна, тъй като базисът е ортонормиран, то $(\boldsymbol{e}_i, \boldsymbol{u}) = (\boldsymbol{e}_i, \lambda_i \boldsymbol{e}_i) = \lambda_i$. Така $\lambda_i = 0$ за $i = 1, \dots, k$ и значи $\boldsymbol{u} = \lambda_{k+1} \boldsymbol{e}_{k+1} + \dots + \lambda_n \boldsymbol{e}_n \in U'$. Следователно $U^{\perp} \subseteq U'$, откъдето $U^{\perp} = U'$.

Твърдение 6. *Нека V е крайномерно евклидово пространство и U е подпространство на V. Тогава $(U^{\perp})^{\perp} = U$.*

Доказателство. Ще ползваме означенията от доказателството на твърдение 5. Векторите $\boldsymbol{e}_{k+1}, \dots, \boldsymbol{e}_n$ са ортонормиран базис на $U^{\perp}$, а векторите $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ ги допълват до ортонормиран базис на V. Повтаряйки същите разсъждения, както в доказателството на твърдение 5, но за подпространството $U^{\perp}$, получаваме $(U^{\perp})^{\perp} = l(\boldsymbol{e}_1, \dots, \boldsymbol{e}_k)$. От друга страна $l(\boldsymbol{e}_1, \dots, \boldsymbol{e}_k) = U$, следователно $(U^{\perp})^{\perp} = U$.

З а б е л е ж к а. Ако V е безкрайномерно евклидово пространство и U е произволно подпространство на V, в общия случай не е вярно, че $V = U \oplus U^{\perp}$ и $(U^{\perp})^{\perp} = U$.

Твърдение 7. *Нека V е крайномерно евклидово пространство, U е подпространство на V и $\boldsymbol{a}$ е произволен вектор от V. Тогава съществуват единствени вектори $\boldsymbol{a}_0 \in U$ и $\boldsymbol{h} \in U^{\perp}$, такива че $\boldsymbol{a} = \boldsymbol{a}_0 + \boldsymbol{h}$.*

Д о к а з а т е л с т в о. Твърдението следва непосредствено от равенството $V = U \oplus U^{\perp}$.

Вектора $\boldsymbol{a}_0$ в горното твърдение ще наричаме *ортогонална проекция* на $\boldsymbol{a}$ в U, а $\boldsymbol{h}$ — *перпендикуляр, спуснат от* $\boldsymbol{a}$ *към* U.

Твърдение 8. *При означенията от предното твърдение, ако $\boldsymbol{u}$ е произволен вектор от U, различен от $\boldsymbol{a}_0$, то $|\boldsymbol{a} - \boldsymbol{a}_0| < |\boldsymbol{a} - \boldsymbol{u}|$.*

Д о к а з а т е л с т в о. Имаме $\boldsymbol{a}-\boldsymbol{a}_0 = \boldsymbol{h} \in U^{\perp}$ и $\boldsymbol{a}_0-\boldsymbol{u} \in U$, така че векторите $\boldsymbol{a}-\boldsymbol{a}_0$ и $\boldsymbol{a}_0-\boldsymbol{u}$ са ортогонални. От твърдение 1 следва $|\boldsymbol{a}-\boldsymbol{a}_0|^2+|\boldsymbol{a}_0-\boldsymbol{u}|^2 = |(\boldsymbol{a}-\boldsymbol{a}_0)+(\boldsymbol{a}_0-\boldsymbol{u})|^2 = |\boldsymbol{a} - \boldsymbol{u}|^2$. Тъй като $\boldsymbol{a}_0 - \boldsymbol{u} \neq \boldsymbol{0}$, то $|\boldsymbol{a}_0 - \boldsymbol{u}|^2 > 0$, откъдето $|\boldsymbol{a} - \boldsymbol{a}_0|^2 < |\boldsymbol{a} - \boldsymbol{u}|^2$, т.е. $|\boldsymbol{a} - \boldsymbol{a}_0| < |\boldsymbol{a} - \boldsymbol{u}|$.

Горното твърдение обобщава факта, че перпендикулярът спуснат от точка до права (равнина) е най-късото разстояние от точката до правата (равнината).

§ 22. Детерминанта на Грам

Определение. *Нека V е евклидово пространство и $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k$ са произволни вектори от V. Детерминантата*

$$\Gamma(\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k) = \begin{vmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_1) & (\boldsymbol{a}_1, \boldsymbol{a}_2) & \dots & (\boldsymbol{a}_1, \boldsymbol{a}_k) \\ (\boldsymbol{a}_2, \boldsymbol{a}_1) & (\boldsymbol{a}_2, \boldsymbol{a}_2) & \dots & (\boldsymbol{a}_2, \boldsymbol{a}_k) \\ \vdots & \vdots & \ddots & \vdots \\ (\boldsymbol{a}_k, \boldsymbol{a}_1) & (\boldsymbol{a}_k, \boldsymbol{a}_2) & \dots & (\boldsymbol{a}_k, \boldsymbol{a}_k) \end{vmatrix}$$

ще наричаме детерминанта на Грам на векторите $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_k$.

Теорема 1. *Нека V е евклидово пространство и $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k \in V$. Тогава детерминантата на Грам на векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ е неотрицателно реално число и е равна на нула точно когато векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са линейно зависими.*

Д о к а з а т е л с т в о. Нека векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са линейно зависими и $\lambda_1\boldsymbol{a}_1 + \dots + \lambda_k\boldsymbol{a}_k = \boldsymbol{0}$, като $(\lambda_1, \dots, \lambda_k) \neq (0, \dots, 0)$. Умножавайки това равенство скаларно последователно с $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$, получаваме хомогенната система

$$\left| \begin{array}{l} \lambda_1(\boldsymbol{a}_1, \boldsymbol{a}_1) + \dots + \lambda_k(\boldsymbol{a}_1, \boldsymbol{a}_k) = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \lambda_1(\boldsymbol{a}_k, \boldsymbol{a}_1) + \dots + \lambda_k(\boldsymbol{a}_k, \boldsymbol{a}_k) = 0. \end{array} \right.$$

Тази система има ненулево решение $\lambda_1, \dots, \lambda_k$ и следователно детерминантата ѝ е равна на нула. От друга страна тази детерминанта е точно $\Gamma(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$. Така, ако векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са линейно зависими, то $\Gamma(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k) = 0$.

Нека сега векторите $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ са линейно независими. Ще докажем, че $\Gamma(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k) > 0$. Нека $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ е ортонормиран базис на пространството $l(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k)$ и

$$\boldsymbol{a}_i = \lambda_{i1}\boldsymbol{e}_1 + \dots + \lambda_{ik}\boldsymbol{e}_k$$

$(i = 1, \ldots, k)$. Тъй като базисът $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ е ортонормиран, то

$$(\boldsymbol{a}_i, \boldsymbol{a}_j) = \lambda_{i1}\lambda_{j1} + \cdots + \lambda_{ik}\lambda_{jk}. \tag{1}$$

От линейната независимост на векторите $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ следва линейна независимост на редовете на детерминантата

$$\Delta = \begin{vmatrix} \lambda_{11} & \ldots & \lambda_{1k} \\ \vdots & \ddots & \vdots \\ \lambda_{k1} & \ldots & \lambda_{kk} \end{vmatrix}.$$

Следователно $\Delta \neq 0$ и значи $\Delta^2 > 0$. От друга страна, използвайки равенството (1) и правилото "ред по ред" за умножение на детерминанти, получаваме $\Delta^2 = \Gamma(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k)$. Следователно $\Gamma(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k) > 0$.

Следствие 2 (Неравенство на Коши — Буняковски). *Нека V е евклидово пространство и $\boldsymbol{a}, \boldsymbol{b} \in V$. Тогава е в сила неравенството $|(\boldsymbol{a}, \boldsymbol{b})| \leq |\boldsymbol{a}|\,|\boldsymbol{b}|$. При това, равенство се достига точно когато векторите $\boldsymbol{a}$ и $\boldsymbol{b}$ са линейно зависими* (колинеарни).

Доказателство. Имаме

$$\begin{aligned} \Gamma(\boldsymbol{a}, \boldsymbol{b}) &= \begin{vmatrix} (\boldsymbol{a}, \boldsymbol{a}) & (\boldsymbol{a}, \boldsymbol{b}) \\ (\boldsymbol{b}, \boldsymbol{a}) & (\boldsymbol{b}, \boldsymbol{b}) \end{vmatrix} = (\boldsymbol{a}, \boldsymbol{a})(\boldsymbol{b}, \boldsymbol{b}) - (\boldsymbol{a}, \boldsymbol{b})^2 \\ &= |\boldsymbol{a}|^2|\boldsymbol{b}|^2 - (\boldsymbol{a}, \boldsymbol{b})^2. \end{aligned}$$

Сега от неравенството $0 \leq \Gamma(\boldsymbol{a}, \boldsymbol{b})$ следва $(\boldsymbol{a}, \boldsymbol{b})^2 \leq |\boldsymbol{a}|^2|\boldsymbol{b}|^2$, т. е. $|(\boldsymbol{a}, \boldsymbol{b})| \leq |\boldsymbol{a}|\,|\boldsymbol{b}|$.

Забележка. 1 (Координатна форма на неравенството на Коши — Буняковски). Нека $V = \mathbb{R}^n$ и $\boldsymbol{a} = (\xi_1, \ldots, \xi_n)$, $\boldsymbol{b} = (\eta_1, \ldots, \eta_n)$. Тогава неравенството на Коши — Буняковски (записано във вида $(\boldsymbol{a}, \boldsymbol{b})^2 \leq |\boldsymbol{a}|^2|\boldsymbol{b}|^2$) приема вида

$$(\xi_1\eta_1 + \cdots + \xi_n\eta_n)^2 \leq (\xi_1^2 + \cdots + \xi_n^2)(\eta_1^2 + \cdots + \eta_n^2).$$

2 (Интегрална форма на неравенството на Коши — Буняковски). Нека $V = C[a, b]$ е евклидовото пространство от пример 3 от предния параграф. Тогава неравенството на Коши — Буняковски приема вида

$$\left(\int_a^b f(x)g(x)\,dx \right)^2 \leq \int_a^b f^2(x)\,dx \cdot \int_a^b g^2(x)\,dx.$$

Нека V е евклидово пространство и $\boldsymbol{a}$ и $\boldsymbol{b}$ са ненулеви вектори от V. Неравенството на Коши — Буняковски можем да запишем във вида

$$-|\boldsymbol{a}|\,|\boldsymbol{b}| \leq (\boldsymbol{a}, \boldsymbol{b}) \leq |\boldsymbol{a}|\,|\boldsymbol{b}|,$$

откъдето получаваме

$$-1 \leq \frac{(\boldsymbol{a}, \boldsymbol{b})}{|\boldsymbol{a}|\,|\boldsymbol{b}|} \leq 1.$$

Следователно съществува единствен ъгъл $\alpha \in [0, \pi]$, такъв че

$$\cos\alpha = \frac{(\boldsymbol{a}, \boldsymbol{b})}{|\boldsymbol{a}|\,|\boldsymbol{b}|}.$$

Този ъгъл ще наричаме *ъгъл между векторите $\boldsymbol{a}$ и $\boldsymbol{b}$*.

Така виждаме, че понятието скаларно произведение на вектори в произволно евклидово пространство съответства на познатото ни от курса по геометрия понятие: *скаларното произведение на два ненулеви вектора е равно на произведението на дължините им по косинуса на ъгъла между тях.*

Задача 1. *Нека a_1, a_2 и a_3 са вектори от евклидовото пространство $\mathbb{R}^3$. Да се докаже, че:*

а) *детерминантата $\Gamma(a_1, a_2)$ е равна на квадрата на лицето на успоредника, построен върху векторите a_1 и a_2;*

б) *детерминантата $\Gamma(a_1, a_2, a_3)$ е равна на квадрата на обема на паралелепипеда, построен върху векторите a_1, a_2 и a_3.*

Твърдение 3 (неравенство на триъгълника). *Нека V е евклидово пространство и a, $b \in V$. Тогава е в сила неравенството $|a + b| \leq |a| + |b|$.*

Доказателство. Използвайки неравенството на Коши — Буняковски, последователно получаваме

$$\begin{aligned}|a+b|^2 &= (a+b, a+b) = (a,a) + (b,b) + 2(a,b)\\ &\leq |a|^2 + |b|^2 + 2|a|\,|b| = (|a| + |b|)^2,\end{aligned}$$

откъдето $|a + b| \leq |a| + |b|$.

С непосредствена индукция неравенството на триъгълника се обобщава за произволен брой вектори: ако a_1, … , a_k са вектори от евклидовото пространство V, то $|a_1 + \cdots + a_k| \leq |a_1| + \cdots + |a_k|$.

Забележка. Ако $V = \mathbb{R}^n$, неравенството на триъгълника може да се изкаже в координатна форма (аналогично на неравенството на Коши — Буняковски) по следния начин: нека $\xi_1, \ldots, \xi_n$ и $\eta_1, \ldots, \eta_n$ са n-орки от реални числа. Тогава е в сила неравенството

$$\sqrt{\sum_{i=1}^{n}(\xi_i + \eta_i)^2} \leq \sqrt{\sum_{i=1}^{n}\xi_i^2} + \sqrt{\sum_{i=1}^{n}\eta_i^2}.$$

При това, равенство се достига точно когато векторите $(\xi_1, \ldots, \xi_n)$ и $(\eta_1, \ldots, \eta_n)$ са "еднопосочно колинеарни", т.е. когато или единият от двата вектора е нулевият, или съществува положително реално число k, такова че $\eta_i = k\xi_i$, $i = 1, \ldots, n$.

Следствие 4. *Нека a и b са вектори от евклидовото пространство V. Тогава е в сила неравенството $|a - b| \geq |a| - |b|$.*

Доказателство. От неравенството на триъгълника получаваме

$$|a| = |b + (a - b)| \leq |b| + |a - b|,$$

откъдето $|a - b| \geq |a| - |b|$.

§ 23. Ортогонални оператори

Определение. *Една матрица $A \in M_n(\mathbb{R})$ ще наричаме ортогонална, ако $AA^t = E$.*

Така една матрица A е ортогонална точно когато е обратима и $A^{-1} = A^t$. Тъй като матриците A^{-1} и A комутират, условието за ортогоналност на A се изразява и с равенството $A^tA = E$.

Нека $A = (a_{ij}) \in M_n(\mathbb{R})$ е ортогонална матрица. От правилото за умножение на матрици следва, че матричното равенство $A^t A = E$ е еквивалентно на n^2 на брой числови равенства $\sum\limits_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij},\ i, j = 1, \dots , n$. Тези равенства означават, че стълбовете на A, разгледани като вектори в n-мерното евклидово пространство $\mathbb{R}^n$, образуват ортонормиран базис на $\mathbb{R}^n$. Аналогично (разглеждайки равенството $A\,A^t = E$) получаваме, че A е ортогонална матрица точно когато векторите-редове на A образуват ортонормиран базис на $\mathbb{R}^n$.

Твърдение 1. *Във всяко крайномерно евклидово пространство V матрицата на прехода от един ортонормиран базис на V към друг ортонормиран базис на V е ортогонална.*

Доказателство. Нека $\dim V = n$, $\boldsymbol{e}_1, \dots , \boldsymbol{e}_n$ и $\boldsymbol{f}_1, \dots , \boldsymbol{f}_n$ са ортонормирани базиси на V и $T = (\tau_{ij})$ е матрицата на прехода от първия към втория базис. Използвайки, че и двата базиса са ортонормирани, последователно получаваме

$$\delta_{ij} = (\boldsymbol{f}_i, \boldsymbol{f}_j) = \Big(\sum_{k=1}^{n} \tau_{ki}\boldsymbol{e}_k, \sum_{k=1}^{n} \tau_{kj}\boldsymbol{e}_k\Big) = \sum_{k=1}^{n} \tau_{ki}\tau_{kj} \quad (i, j = 1, \dots, n).$$

Тези равенства означават, че е изпълнено матричното равенство $T^tT = E$, т. е.T е ортогонална матрица.

Свойства на ортогоналните матрици.

1. Ако A е ортогонална матрица, то $\det A = \pm 1$.

Действително имаме

$$(\det A)^2 = \det A.\det A = \det A.\det A^t = \det(AA^t) = \det E = 1.$$

Следователно $\det A = \pm 1$.

2. Ако A е ортогонална матрица, то A^{-1} също е ортогонална матрица.

Имаме $A^{-1}A = E$ и след транспониране получаваме $A^t(A^{-1})^t = E$. Но $A^t = A^{-1}$ и значи $A^{-1}(A^{-1})^t = E$. Следователно A^{-1} е ортогонална матрица.

3. Ако A и B са ортогонални матрици от един и същи ред, то AB също е ортогонална матрица.

Имаме

$$(AB)^t = B^tA^t = B^{-1}A^{-1} = (AB)^{-1},$$

което означава, че AB е ортогонална матрица.

До края на параграфа V ще бъде n-мерно евклидово пространство.

Определение. *Нека $\varphi \in \operatorname{Hom} V$. Ще казваме, че φ е ортогонален оператор, ако за всеки два вектора $\boldsymbol{u}$ и $\boldsymbol{v}$ от V е изпълнено $(\varphi(\boldsymbol{u}), \varphi(\boldsymbol{v})) = (\boldsymbol{u}, \boldsymbol{v})$.*

Прието е да се казва, че φ запазва скаларното произведение, в частност запазва дължините на векторите и големините на ъглите между тях.

Нека $\boldsymbol{e}_1, \dots , \boldsymbol{e}_n$ е базис на V. Тъй като всеки вектор от V е линейна комбинация на векторите $\boldsymbol{e}_1, \dots , \boldsymbol{e}_n$ и φ е линеен оператор, то φ е ортогонален оператор тогава и само тогава, когато запазва скаларното произведение само на двойките базисни вектори, т. е. когато $(\varphi(\boldsymbol{e}_i), \varphi(\boldsymbol{e}_j)) = (\boldsymbol{e}_i, \boldsymbol{e}_j)$ $(i, j = 1, \dots , n)$.

Задача 1. *Нека $\varphi \in \operatorname{Hom} V$. Да се докаже, че φ е ортогонален оператор тогава и само тогава, когато за всеки вектор $\boldsymbol{u}$ от V е изпълнено $(\varphi(\boldsymbol{u}), \varphi(\boldsymbol{u})) = (\boldsymbol{u}, \boldsymbol{u})$, т. е. когато φ запазва само дължините на векторите.*

У п ъ т в а н е. Нека φ запазва дължините на векторите. За да докажете, че φ е ортогонален, използвайте равенството $(\varphi(\boldsymbol{u}+\boldsymbol{v}), \varphi(\boldsymbol{u}+\boldsymbol{v})) = (\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v})$.

Твърдение 2. *Нека $e_1, \ldots, e_n$ е ортонормиран базис на V, $\varphi \in \operatorname{Hom} V$ и A е матрицата на φ в този базис. Тогава φ е ортогонален оператор точно когато A е ортогонална матрица. В частност, всеки ортогонален оператор е обратим.*

Д о к а з а т е л с т в о. Нека $A = (a_{ij})_{n\times n}$. Имаме $\varphi(e_i) = \sum\limits_{k=1}^{n} a_{ki}e_k$ $(i = 1, \ldots, n)$. Използвайки, че базисът $e_1, \ldots, e_n$ е ортонормиран, получаваме

$$(\varphi(e_i), \varphi(e_j)) = \Big(\sum_{k=1}^{n} a_{ki}e_k, \sum_{k=1}^{n} a_{kj}e_k\Big) = \sum_{k=1}^{n} a_{ki}a_{kj}.$$

От друга страна $(e_i, e_j) = \delta_{ij}$. Следователно $(\varphi(e_i), \varphi(e_j)) = (e_i, e_j)$ точно когато $\sum\limits_{k=1}^{n} a_{ki}a_{kj} = \delta_{ij}$ $(i, j = 1, \ldots, n)$. Това означава, че φ е ортогонален оператор точно когато A е ортогонална матрица.

Задача 2. *Да се докаже, че φ е ортогонален оператор точно когато преобразува ортонормиран базис на V в ортонормиран базис на V.*

Твърдение 3. *Собствените стойности на ортогонален оператор са равни на ± 1.*

Д о к а з а т е л с т в о. Нека φ е ортогонален оператор във V, λ е собствена стойност на φ и $\boldsymbol{v}$ е съответстващ ѝ собствен вектор на φ. Имаме $(\varphi(\boldsymbol{v}), \varphi(\boldsymbol{v})) = (\boldsymbol{v}, \boldsymbol{v})$, откъдето последователно получаваме $(\lambda\boldsymbol{v}, \lambda\boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{v})$, $\lambda^2(\boldsymbol{v}, \boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{v})$, $(\lambda^2 - 1)(\boldsymbol{v}, \boldsymbol{v}) = 0$. Тъй като $(\boldsymbol{v}, \boldsymbol{v}) \neq 0$, то $\lambda^2 = 1$, значи $\lambda = \pm 1$.

Ще отбележим без доказателство, че всички характеристични корени на ортогонален оператор (а значи и на ортогонална матрица) по модул са равни на 1 (вижте и следствие 2 от § 27).

Твърдение 4. *Всеки два собствени вектора, съответстващи на различни собствени стойности на ортогонален оператор, са ортогонални помежду си.*

Д о к а з а т е л с т в о. Нека φ е ортогонален оператор, λ_1 и λ_2 са различни собствени стойности на φ и $\boldsymbol{v}_1$ и $\boldsymbol{v}_2$ са съответстващи им собствени вектори на φ. Имаме $(\varphi(\boldsymbol{v}_1), \varphi(\boldsymbol{v}_2)) = (\boldsymbol{v}_1, \boldsymbol{v}_2)$, откъдето последователно получаваме

$$(\lambda_1\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2) = (\boldsymbol{v}_1, \boldsymbol{v}_2), \quad \lambda_1\lambda_2(\boldsymbol{v}_1, \boldsymbol{v}_2) = (\boldsymbol{v}_1, \boldsymbol{v}_2), \quad (\lambda_1\lambda_2 - 1)(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0.$$

Тъй като $\lambda_{1,2} = \pm 1$, но $\lambda_1 \neq \lambda_2$, то $\lambda_1\lambda_2 = -1$ и значи $\lambda_1\lambda_2 - 1 = -2 \neq 0$. Следователно $(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$.

Теорема 5. *Нека V е евклидово пространство и φ е ортогонален оператор във V. Тогава съществува ортонормиран базис на V, в който матрицата D на φ е клетъчно диагонална:*

$$D = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_k \end{pmatrix},$$

и всяка една от матриците D_i $(i = 1, \ldots, k)$ е от вида

$$D_i = (1), \quad D_i = (-1) \quad \text{или} \quad D_i = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

където α е подходящ ъгъл, който не е целочислено кратен на π (ако ъгълът α е целочислено кратен на π, то $\sin\alpha = 0$ и последната клетка се разбива на две клетки от първите два вида).

Д о к а з а т е л с т в о. Ще използваме неколкократно, че (според твърдение 3) собствените стойности на φ са равни на ± 1.

Ще проведем индукция по $n = \dim V$. Нека $n = 1$ и $\boldsymbol{v}_1$ е базис на V, като $|\boldsymbol{v}_1| = 1$. Тогава $\varphi(\boldsymbol{v}_1) \in V = l(\boldsymbol{v}_1)$, т. е. $\varphi(\boldsymbol{v}_1) = \lambda \boldsymbol{v}_1$ ($\lambda \in \mathbb{R}$). Следователно λ е собствена стойност на φ и значи $\lambda = \pm 1$. Тогава матрицата на φ в базиса $\boldsymbol{v}_1$ на V е $D = (\pm 1)$.

Нека $n = 2$. Нека първо φ има собствена стойност λ (т. е. реален характеристичен корен) и $\boldsymbol{v}'_1$ е собствен вектор на φ, съответстващ на λ. Заменяйки $\boldsymbol{v}'_1$ с вектора $\boldsymbol{v}_1 = \dfrac{1}{|\boldsymbol{v}'_1|}\boldsymbol{v}'_1$, получаваме собствен вектор на φ, съответстващ на λ, който е с дължина 1. Допълваме вектора $\boldsymbol{v}_1$ до ортонормиран базис $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ на V. Имаме $V = l(\boldsymbol{v}_1) \oplus l(\boldsymbol{v}_2)$; при това $l(\boldsymbol{v}_1)^\perp = l(\boldsymbol{v}_2)$. Използвайки, че $\lambda = \pm 1$ и че φ е ортогонален оператор, последователно получаваме.

$$(\boldsymbol{v}_1, \varphi(\boldsymbol{v}_2)) = \pm(\lambda \boldsymbol{v}_1, \varphi(\boldsymbol{v}_2)) = \pm(\varphi(\boldsymbol{v}_1), \varphi(\boldsymbol{v}_2)) = \pm(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0.$$

Това означава, че $\varphi(\boldsymbol{v}_2) \in l(\boldsymbol{v}_1)^\perp$. Но $l(\boldsymbol{v}_1)^\perp = l(\boldsymbol{v}_2)$ и значи $\varphi(\boldsymbol{v}_2) = \mu \boldsymbol{v}_2$ ($\mu \in \mathbb{R}$). Тогава μ е собствена стойност на φ, т. е. $\mu = \pm 1$. Сега от равенствата $\varphi(\boldsymbol{v}_1) = \lambda \boldsymbol{v}_1$ и $\varphi(\boldsymbol{v}_2) = \mu \boldsymbol{v}_2$ следва, че матрицата D на φ в ортонормирания базис $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ на V се състои от две клетки от ред 1, всяка от които е (1) или (-1).

Нека сега φ няма собствени стойности, т. е. няма реални характеристични корени. Нека $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ е ортонормиран базис на V и $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ е матрицата на φ в този базис. Характеристичният полином на A е

$$f_A(x) = \begin{vmatrix} a - x & b \\ c & d - x \end{vmatrix} = x^2 - (a+d)x + (ad - bc).$$

Тъй като A е ортогонална матрица, то $\det A = ad - bc = \pm 1$. От друга страна полиномът $f_A(x)$ няма реални корени и значи дискриминантата му е отрицателна, т. е. $(a+d)^2 - 4(ad - bc) < 0$. Следователно $\det A = +1$. Сега имаме $A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ и $A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. От равенството $A^t = A^{-1}$ получаваме $d = a$ и $b = -c$. Така матрицата A има вида $A = \begin{pmatrix} a & -c \\ c & a \end{pmatrix}$ и освен това $\det A = a^2 + c^2 = 1$. Тогава съществува ъгъл α, такъв че $\cos\alpha = a$, $\sin\alpha = c$. (Тъй като характеристичните корени на A не са реални, то $c \neq 0$ и $\alpha \neq l\pi$ ($l \in \mathbb{Z}$).) Следователно матрицата на φ в (произволен) ортонормиран базис на V е

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

за подходящ ъгъл α.

Нека $n > 2$. Според теорема 6 от § 17 операторът φ има едномерно или двумерно инвариантно подпространство U и φ е ортогонален оператор на U. От доказаното по-горе следва, че съществува ортонормиран базис на U, в който матрицата на оператора $\varphi_{|U}$ е от търсения вид. Нека $U^\perp$ е ортогоналното допълнение на U. Ще докажем, че подпространството $U^\perp$ е φ-инвариантно. Нека $\boldsymbol{v} \in U^\perp$, т. е. $(\boldsymbol{v}, U) = 0$ и $\boldsymbol{u}$ е произволен вектор от U. Операторът $\varphi_{|U}$ е обратим и значи съществува вектор $\boldsymbol{u}'$ от U, такъв че $\varphi(\boldsymbol{u}') = \boldsymbol{u}$. Сега имаме

$$(\varphi(\boldsymbol{v}), \boldsymbol{u}) = (\varphi(\boldsymbol{v}), \varphi(\boldsymbol{u}')) = (\boldsymbol{v}, \boldsymbol{u}') = 0.$$

Това означава, че $(\varphi(\boldsymbol{v}), U) = 0$, т.е. $\varphi(\boldsymbol{v}) \in U^{\perp}$. Следователно подпространството $U^{\perp}$ е φ-инвариантно и φ е ортогонален оператор на $U^{\perp}$.

Знаем, че $V = U \oplus U^{\perp}$ и $\dim U = 1$ или 2. Тогава $\dim U^{\perp} \leq n-1$. Според индукционното предположение съществува ортонормиран базис на $U^{\perp}$, в който матрицата на оператора $\varphi_{|_{U^{\perp}}}$ е от търсения вид. Да разгледаме базиса на V, съставен от ортонормираните базиси на U и $U^{\perp}$, в които матриците съответно на операторите $\varphi_{|_U}$ и $\varphi_{|_{U^{\perp}}}$ са от търсения вид. Тъй като $(U, U^{\perp}) = 0$, този базис на V също е ортонормиран. Сега от бележка 2 в началото на § 18 следва, че матрицата на φ в така намерения ортонормиран базис на V също е от посочения в условието на теоремата вид.

З а б е л е ж к а. От доказаната теорема следва, че цялото пространство V се представя като директна сума на едномерни и | или двумерни φ-инвариантни подпространства. Ако едно такова подпространство е едномерно (права), φ действа в него като идентитет или симетрия относно точка. Ако подпространството е двумерно (равнина), φ действа в него като ротация на ъгъл α около точка.

Следствие 6. *За всяка ортогонална матрица A съществува ортогонална матрица T, такава че матрицата $D = T^{-1}AT$ е от посочения в условието на теорема* 5 *вид.*

Д о к а з а т е л с т в о. Нека $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ е ортонормиран базис на пространството V и φ е линейният оператор във V с матрица A в този базис. Тогава φ е ортогонален оператор (твърдение 2). Според теорема 5 съществува ортонормиран базис $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ на V, в който матрицата D на φ е от посочения в тази теорема вид. Сега, ако T е матрицата на прехода от първия към втория базис, то T е ортогонална матрица (твърдение 1) и $D = T^{-1}AT$.

§ 24. Симетрични оператори

Определение. *Една матрица $A \in M_n(\mathbb{R})$ ще наричаме симетрична, ако $A^t = A$.*

Нека $A = (a_{ij})_{n \times n}$. Тогава матричното равенство $A^t = A$ е еквивалентно на n^2 на брой числови равенства: $a_{ji} = a_{ij}$ $(i, j = 1, \ldots, n)$. Това означава, че елементите на A, симетрични относно главния диагонал са равни.

С в о й с т в а на с и м е т р и ч н и т е м а т р и ц и.

1. Симетричните матрици образуват подпространство на линейното пространство $M_n(\mathbb{R})$.

Това свойство следва директно от определенията.

2. Ако A е обратима симетрична матрица, то A^{-1} също е симетрична матрица.

Нека $C = A^{-1}$. Имаме $AC = E$, откъдето последователно получаваме

$$(AC)^t = E, \quad C^t A^t = E, \quad C^t A = E, \quad C^t = A^{-1}, \text{ т.е. } C^t = C.$$

Следователно C е симетрична матрица.

3. Ако A и B са симетрични матрици и $AB = BA$, то AB също е симетрична матрица.

Имаме $(AB)^t = B^t A^t = BA = AB$. Следователно AB е симетрична матрица.

До края на параграфа V ще бъде n-мерно евклидово пространство.

Определение. *Нека $\varphi \in \operatorname{Hom} V$. Ще казваме, че φ е симетричен оператор, ако за всеки два вектора $\boldsymbol{u}$ и $\boldsymbol{v}$ от V е изпълнено $(\varphi(\boldsymbol{u}), \boldsymbol{v}) = (\boldsymbol{u}, \varphi(\boldsymbol{v}))$.*

Нека $e_1, \dots, e_n$ е произволен базис на V. Тъй като всеки вектор на V е линейна комбинация на векторите $e_1, \dots, e_n$ и φ е линеен оператор, то φ е симетричен оператор тогава и само тогава, когато само за двойките базисни вектори е изпълнено $(\varphi(e_i), e_j) = (e_i, \varphi(e_j))$ $(i, j = 1, \dots, n)$.

Твърдение 1. *Нека $e_1, \dots, e_n$ е ортонормиран базис на V, $\varphi \in \operatorname{Hom} V$ и A е матрицата на φ в този базис. Тогава φ е симетричен оператор точно когато A е симетрична матрица.*

Доказателство. Нека $A = (a_{ij})_{n \times n}$. Имаме $\varphi(e_i) = \sum\limits_{k=1}^{n} a_{ki} e_k$ $(i = 1, \dots, n)$. Използвайки, че базисът $e_1, \dots, e_n$ е ортонормиран, получаваме

$$(\varphi(e_i), e_j) = \left(\sum_{k=1}^{n} a_{ki} e_k, e_j\right) = (a_{ji} e_j, e_j) = a_{ji}(e_j, e_j) = a_{ji},$$

$$(e_i, \varphi(e_j)) = \left(e_i, \sum_{k=1}^{n} a_{kj} e_k\right) = (e_i, a_{ij} e_i) = a_{ij}(e_i, e_i) = a_{ij}.$$

От тези равенства следва, че φ е симетричен оператор, т.е. $(\varphi(e_i), e_j) = (e_i, \varphi(e_j))$, точно когато $a_{ji} = a_{ij}$ $(i, j = 1, \dots, n)$, т.е. когато A е симетрична матрица.

Твърдение 2. *Характеристичните корени на симетрична матрица са реални числа.*

Доказателство. Нека $A = (a_{ij})_{n \times n}$ е симетрична матрица и λ е (изобщо казано комплексен) характеристичен корен на A. Ще докажем, че λ е реално число.

Имаме $f_A(\lambda) = \det(A - \lambda E) = 0$. Следователно хомогенната система с матрица $A - \lambda E$ има ненулево решение $(\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{C}^n$, т.е. в сила са равенствата

$$\left|\begin{array}{l} (a_{11} - \lambda)\xi_1 + \quad a_{12}\xi_2 \quad + \dots + \quad a_{1n}\xi_n \quad = 0 \\ \quad a_{21}\xi_1 \quad + (a_{22} - \lambda)\xi_2 + \dots + \quad a_{2n}\xi_n \quad = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \quad a_{n1}\xi_1 \quad + \quad a_{n2}\xi_2 \quad + \dots + (a_{nn} - \lambda)\xi_n = 0 \end{array}\right.$$

или

$$\left|\begin{array}{l} a_{11}\xi_1 + a_{12}\xi_2 + \dots + a_{1n}\xi_n = \lambda\xi_1 \\ a_{21}\xi_1 + a_{22}\xi_2 + \dots + a_{2n}\xi_n = \lambda\xi_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}\xi_1 + a_{n2}\xi_2 + \dots + a_{nn}\xi_n = \lambda\xi_n \end{array}\right.$$

Като умножим първото равенство с $\overline{\xi}_1$, второто с $\overline{\xi}_2$ и т.н., n-тото с $\overline{\xi}_n$ и ги съберем, получаваме

$$\sum_{i,j=1}^{n} a_{ij}\xi_j\overline{\xi}_i = \lambda \sum_{i=1}^{n} \xi_i\overline{\xi}_i = \lambda \sum_{i=1}^{n} |\xi_i|^2.$$

Да означим $u = \sum\limits_{i,j=1}^{n} a_{ij}\xi_j\overline{\xi}_i$, $v = \sum\limits_{i=1}^{n} |\xi_i|^2$. Числото v е реално и различно от 0 (по-точно, $v > 0$). Като използваме, че A е симетрична матрица, т.е. $a_{ij} = a_{ji}$, получаваме

$$\overline{u} = \sum_{i,j=1}^{n} \overline{a}_{ij}\overline{\xi}_j\overline{\overline{\xi}}_i = \sum_{i,j=1}^{n} a_{ij}\overline{\xi}_j\xi_i = \sum_{i,j=1}^{n} a_{ji}\overline{\xi}_j\xi_i.$$

Остава да "разчетем", че последната сума е равна точно на u. Тогава $\overline{u} = u$ и значи u е реално число, а следователно и $\lambda = \dfrac{u}{v}$ е реално число.

Преминавайки от матричен на операторен език, получаваме

Следствие 3. *Характеристичните корени на симетричен оператор са реални числа (и значи всички характеристични корени на оператора са негови собствени стойности).*

Твърдение 4. *Всеки два собствени вектора, съответстващи на различни собствени стойности на симетричен оператор, са ортогонални помежду си.*

Д о к а з а т е л с т в о. Нека φ е симетричен оператор, λ_1 и λ_2 са различни собствени стойности на φ и $\boldsymbol{v}_1$ и $\boldsymbol{v}_2$ са съответстващи им собствени вектори на φ. Имаме $(\varphi(\boldsymbol{v}_1), \boldsymbol{v}_2) = (\boldsymbol{v}_1, \varphi(\boldsymbol{v}_2))$, откъдето $(\lambda_1\boldsymbol{v}_1, \boldsymbol{v}_2) = (\boldsymbol{v}_1, \lambda_2\boldsymbol{v}_2)$, $\lambda_1(\boldsymbol{v}_1, \boldsymbol{v}_2) = \lambda_2(\boldsymbol{v}_1, \boldsymbol{v}_2)$, $(\lambda_1 - \lambda_2)(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$. Тъй като $\lambda_1 \neq \lambda_2$, то $(\boldsymbol{v}_1, \boldsymbol{v}_2) = 0$.

Теорема 5. *Нека φ е симетричен оператор във V. Тогава съществува ортонормиран базис на V, в който матрицата на φ е диагонална (по главния ѝ диагонал стоят собствените стойности на φ, а базисните вектори са собствени вектори на φ).*

Д о к а з а т е л с т в о. Трябва да докажем, че пространството V притежава ортонормиран базис $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$, състоящ се от собствени вектори на φ, т. е. $\varphi(\boldsymbol{v}_1) = \lambda_1\boldsymbol{v}_1, \dots, \varphi(\boldsymbol{v}_n) = \lambda_n\boldsymbol{v}_n$ $(\lambda_1, \dots, \lambda_n \in \mathbb{R})$. Тогава матрицата D на φ в този базис е

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Ще проведем индукция по $n = \dim V$. При $n = 1$ твърдението е очевидно. Нека $n > 1$ и да допуснем, че теоремата е доказана за по-малки стойности на n. Нека λ_1 е собствена стойност на φ и $\boldsymbol{v}_1'$ е собствен вектор на φ, съответстващ на λ_1 (напомняме, че характеристичните корени на φ са реални числа и следователно φ притежава собствена стойност). Заменяйки $\boldsymbol{v}_1'$ с вектора $\boldsymbol{v}_1 = \dfrac{1}{|\boldsymbol{v}_1'|}\boldsymbol{v}_1'$, получаваме собствен вектор на φ, съответстващ на λ_1, който е с дължина 1. Да означим $U = \ell(\boldsymbol{v}_1)$ и нека $U^\perp$ е ортогоналното допълнение на U. Ясно е, че $U^\perp = \{\boldsymbol{v} \in V \,|\, (\boldsymbol{v}, \boldsymbol{v}_1) = 0\}$.

Ще докажем, че подпространството $U^\perp$ е φ-инвариантно, т. е. ако един вектор $\boldsymbol{v}$ принадлежи на $U^\perp$, то $\varphi(\boldsymbol{v})$ също принадлежи на $U^\perp$. Нека $\boldsymbol{v} \in U^\perp$, т. е. $(\boldsymbol{v}, \boldsymbol{v}_1) = 0$. Използвайки, че φ е симетричен оператор, получаваме

$$(\varphi(\boldsymbol{v}), \boldsymbol{v}_1) = (\boldsymbol{v}, \varphi(\boldsymbol{v}_1)) = (\boldsymbol{v}, \lambda_1\boldsymbol{v}_1) = \lambda_1(\boldsymbol{v}, \boldsymbol{v}_1) = 0.$$

Следователно $\varphi(\boldsymbol{v}) \in U^\perp$.

И така, φ съпоставя на всеки вектор от $U^\perp$ вектор, който също лежи в $U^\perp$, което означава, че φ е (симетричен) оператор, действащ в пространството $U^\perp$. Знаем, че $V = U \oplus U^\perp$ и освен това $\dim U = 1$. Следователно $\dim U^\perp = n - 1$ и според индукционното предположение $U^\perp$ притежава ортонормиран базис $\boldsymbol{v}_2, \dots, \boldsymbol{v}_n$, състоящ се от собствени вектори на φ, т. е. $\varphi(\boldsymbol{v}_2) = \lambda_2\boldsymbol{v}_2, \dots, \varphi(\boldsymbol{v}_n) = \lambda_n\boldsymbol{v}_n$ $(\lambda_2, \dots, \lambda_n \in \mathbb{R})$. Освен това векторите $\boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ са от $U^\perp$ и значи са ортогонални на вектора $\boldsymbol{v}_1$. Тогава $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ е базис на V, който е ортонормиран и $\varphi(\boldsymbol{v}_i) = \lambda_i\boldsymbol{v}_i$ за $i = 1, 2, \dots, n$. Следователно матрицата D на φ в този базис е диагонална и по главния ѝ диагонал стоят числата $\lambda_1, \lambda_2, \dots, \lambda_n$.

Следствие 6. *За всяка симетрична матрица A съществува ортогонална матрица T, такава че матрицата $D = T^{-1}AT$ е диагонална (и значи всяка симетрична матрица е подобна на диагонална).*

Доказателство. Нека $e_1, \dots, e_n$ е ортонормиран базис на пространството V и φ е линейният оператор във V с матрица A в този базис. Тогава φ е симетричен оператор (твърдение 1). Според теорема 5 съществува ортонормиран базис $v_1, \dots, v_n$ на V, в който матрицата D на φ е диагонална. Сега, ако T е матрицата на прехода от първия към втория базис, то T е ортогонална матрица (твърдение 1 от § 23) и $D = T^{-1}AT$.

§ 25. Билинейни и квадратични форми

Определение. *Нека V е линейно пространство над $\mathbb{R}$ и φ е изображение, което на всяка наредена двойка вектори u и v от V съпоставя реално число $\varphi(u, v)$. Ще казваме, че φ е билинейно изображение (форма), ако е линейно по двата си аргумента, т. е. ако се изпълняват следните условия:*

$$1.\ \varphi(u_1 + u_2, v) = \varphi(u_1, v) + \varphi(u_2, v),$$
$$\varphi(\lambda u, v) = \lambda\varphi(u, v);$$
$$2.\ \varphi(u, v_1 + v_2) = \varphi(u, v_1) + \varphi(u, v_2),$$
$$\varphi(u, \lambda v) = \lambda\varphi(u, v)$$

(u, v, u_1, u_2, v_1, $v_2 \in V$, $\lambda \in \mathbb{R}$). *Ако освен това за всеки два вектора u и v от V е изпълнено $\varphi(u, v) = \varphi(v, u)$, ще казваме, че φ е симетрична билинейна форма.*

Скаларното произведение в евклидово пространство е пример за симетрична билинейна форма.

Нека $\dim V = n < \infty$ и $e_1, \dots, e_n$ е базис на V. Нека $u = \sum\limits_{i=1}^{n} \xi_i e_i$, $v = \sum\limits_{j=1}^{n} \eta_j e_j$. Както при скаларно произведение, и тук получаваме $\varphi(u, v) = \sum\limits_{i,j=1}^{n} \xi_i \eta_j \varphi(e_i, e_j)$. Да означим $\varphi(e_i, e_j) = a_{ij}$ $(i, j = 1, \dots, n)$. Тогава

$$\varphi(u, v) = \sum_{i,j=1}^{n} a_{ij} \xi_i \eta_j. \tag{1}$$

Определение. *Матрицата $A = (a_{ij})_{n \times n}$ ще наричаме матрица на билинейната форма φ в базиса $e_1, \dots, e_n$.*

При фиксиран базис всяка билинейна форма определя матрица от $M_n(\mathbb{R})$. Обратно, нека $A = (a_{ij})_{n \times n}$ е матрица от $M_n(\mathbb{R})$. Директно се проверява, че изображението φ, дефинирано с равенството (1) е билинейно. При това $\varphi(e_i, e_j) = a_{ij}$ $(i, j = 1, \dots, n)$, т.е. матрицата на φ в базиса $e_1, \dots, e_n$ е точно A.

Твърдение 1. *Една билинейна форма е симетрична тогава и само тогава, когато матрицата ѝ във всеки базис е симетрична.*

Доказателството е непосредствено и го предоставяме на читателя.

Твърдение 2. *Нека $e_1, \dots, e_n$ и $f_1, \dots, f_n$ са два базиса на V и T е матрицата на прехода от първия към втория базис. Нека φ е билинейна форма и матриците ѝ в първия и втория базис са съответно A и B. Тогава е в сила равенството $B = T^t A T$.*

Доказателство. Нека $A = (a_{ij})_{n\times n}$, $B = (b_{ij})_{n\times n}$, $T = (\tau_{ij})_{n\times n}$. Имаме

$$b_{ij} = \varphi(\boldsymbol{f}_i, \boldsymbol{f}_j) = \varphi\left(\sum_{k=1}^{n} \tau_{ki}\boldsymbol{e}_k, \sum_{s=1}^{n} \tau_{sj}\boldsymbol{e}_s\right) = \sum_{k,s=1}^{n} \tau_{ki}\tau_{sj}\varphi(\boldsymbol{e}_k, \boldsymbol{e}_s)$$
$$= \sum_{k,s=1}^{n} \tau_{ki}\tau_{sj}a_{ks} = \sum_{s=1}^{n}\left(\sum_{k=1}^{n} \tau_{ki}a_{ks}\right)\tau_{sj}.$$

Сумата в скобите в дясната страна на равенството представлява i-ия стълб на T, т. е. i-ия ред на T^t, умножен по s-ия стълб на A, т. е. равна е на елемента, който стои на (i, s)-то място в матрицата T^tA. Тогава цялата сума представлява i-ия ред на T^tA, умножен по j-ия стълб на T, т. е. равна е на елемента, който стои на (i, j)-то място в матрицата T^tAT $(i, j = 1, \ldots, n)$. Следователно $B = T^tAT$.

Определение. *Нека φ е симетрична билинейна форма във V. Изображението $\tilde{\varphi} : V \longrightarrow \mathbb{R}$, което на всеки вектор $\boldsymbol{u}$ от V съпоставя числото $\tilde{\varphi}(\boldsymbol{u}) = \varphi(\boldsymbol{u}, \boldsymbol{u})$ ще наричаме квадратична форма. Ще казваме, че формите φ и $\tilde{\varphi}$ са асоциирани.*

Директно се проверява, че е изпълнено равенството

$$\varphi(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{2}\left[\varphi(\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) - \varphi(\boldsymbol{u}, \boldsymbol{u}) - \varphi(\boldsymbol{v}, \boldsymbol{v})\right] = \frac{1}{2}\left[\tilde{\varphi}(\boldsymbol{u}+\boldsymbol{v}) - \tilde{\varphi}(\boldsymbol{u}) - \tilde{\varphi}(\boldsymbol{v})\right]$$

(при проверката се използва, че $\varphi(\boldsymbol{u}, \boldsymbol{v}) = \varphi(\boldsymbol{v}, \boldsymbol{u})$). В дясната страна на равенството стоят стойности само на квадратичната форма $\tilde{\varphi}$. Следователно *симетричната билинейна форма φ еднозначно се определя от асоциираната си квадратична форма* (това оправдава разглеждането само на квадратични форми, получени от симетрични билинейни форми).

При въведените означения, ако $\boldsymbol{u} = \sum\limits_{i=1}^{n} \xi_i\boldsymbol{e}_i$, имаме $\tilde{\varphi}(\boldsymbol{u}) = \varphi(\boldsymbol{u}, \boldsymbol{u}) = \sum\limits_{i,j=1}^{n} a_{ij}\xi_i\xi_j$. Симетричната матрица $A = (a_{ij})_{n\times n}$ ще наричаме *матрица на квадратичната форма* $\tilde{\varphi}$ в базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Така матриците на φ и $\tilde{\varphi}$ съвпадат.

Забележка. Ако се абстрахираме от линейното пространство V и от фиксирания базис, можем да разглеждаме квадратичната форма като полином на $\xi_1, \ldots, \xi_n$ с коефициенти a_{ij}.

Нека елементите извън главния диагонал на матрицата A са равни на нула (т. е. $a_{ij} = 0$ при $i \neq j$). Да означим за краткост $a_{ii} = \lambda_i$ $(i = 1, \ldots, n)$. Тогава квадратичната форма приема вида $\tilde{\varphi}(\boldsymbol{u}) = \lambda_1\xi_1^2 + \cdots + \lambda_n\xi_n^2$. Този вид на квадратичната форма ще наричаме *каноничен вид*.

Теорема 3. *За всяка квадратична форма $\tilde{\varphi}$ в крайномерно пространство V съществува базис на V, в който квадратичната форма е в каноничен вид. При това, ако във V е въведено скаларно произведение, този базис може да се избере ортонормиран.*

Доказателство. Нека $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ е базис на V и $A = (a_{ij})_{n\times n}$ е матрицата на квадратичната форма $\tilde{\varphi}$ в този базис. Тъй като A е симетрична матрица, съществува ортогонална матрица T, такава че матрицата $D = T^{-1}AT$ е диагонална (следствие 6 от § 24):

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Матрицата T е неособена и следователно е матрица на прехода от базиса $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ към друг базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. Освен това $T^{-1} = T^t$ и значи $D = T^tAT$. Тогава матрицата на квадратичната форма в базиса $e_1, \dots, e_n$ е точно D и значи $\tilde{\varphi}(u) = \lambda_1\xi_1^2 + \dots + \lambda_n\xi_n^2$ $(u = \sum_{i=1}^{n} \xi_i e_i)$.

Накрая, ако V е евклидово пространство, избираме базиса $f_1, \dots, f_n$ ортонормиран. Тъй като T е ортогонална матрица, базисът $e_1, \dots, e_n$ също ще бъде ортонормиран. Теоремата е доказана.

Забележка. Привеждането на една квадратична форма в каноничен вид чрез ортогонална трансформация се нарича *привеждане към главни оси*. Съществуват и неортогонални трансформации, които привеждат квадратичната форма в каноничен вид.

Нека $V = \mathbb{R}^2$ и $c \in \mathbb{R}$. Уравнението (спрямо ξ_1 и ξ_2) $\tilde{\varphi}(u) = \sum_{i,j=1}^{2} a_{ij}\xi_i\xi_j = c$ задава крива от втора степен. Геометричният смисъл на привеждането към главни оси е, че съществува подходяща (ортонормирана) координатна система, в която уравнението на кривата е в каноничен вид. Аналогичен е геометричният смисъл (за повърхнини от втора степен) в тримерното евклидово пространство.

Нека матрицата D на квадратичната форма $\tilde{\varphi}$ в даден базис е диагонална и в този базис $\tilde{\varphi}(u) = \lambda_1\xi_1^2 + \dots + \lambda_r\xi_r^2$, като $\lambda_i \neq 0$ $(i = 1, \dots, r)$, $r \leq n$. Броят r на ненулевите коефициенти λ_i е равен на ранга на матрицата D. Ако T е обратима матрица, то $r(T^t DT) = r(D)$. Следователно рангът на матрицата на квадратичната форма във всеки базис е един и същ. Този ранг ще наричаме *ранг на квадратичната форма* и ще го бележим с $r(\tilde{\varphi})$. Така броят на нулевите коефициенти във всеки каноничен вид на $\tilde{\varphi}$ е един и същ и е равен на $n - r(\tilde{\varphi})$.

Нека сега $\tilde{\varphi}(u) = \lambda_1\xi_1^2 + \dots + \lambda_p\xi_p^2 - \lambda_{p+1}\xi_{p+1}^2 - \dots - \lambda_r\xi_r^2$,като $\lambda_i > 0$ за $i = 1, \dots, r$. Полагаме $e_i' = \dfrac{1}{\sqrt{\lambda_i}} e_i$ за $i = 1, \dots, r$ и $e_j' = e_j$ за $j = r+1, \dots, n$. Тогава в базиса $e_1', \dots, e_n'$ квадратичната форма приема така наречения *нормален вид*: $\tilde{\varphi}(u) = \eta_1^2 + \dots + \eta_p^2 - \eta_{p+1}^2 - \dots - \eta_r^2$ $(u = \sum_{i=1}^{n} \eta_i e_i')$.

Теорема 4 (закон за инерцията). *Броят на положителните, броят на отрицателните и броят на нулевите коефициенти във всеки каноничен вид на една квадратична форма е един и същ (т. е. нормалният вид на квадратичната форма е единствен).*

Доказателство. Нека $\tilde{\varphi}$ е квадратична форма във V. Вече знаем, че броят на нулевите коефициенти във всеки каноничен вид на $\tilde{\varphi}$ е един и същ.

Нека $e_1, \dots, e_n$ и $f_1, \dots, f_n$ са два базиса на V, в които квадратичната форма е в нормален вид и двата нормални вида са съответно

$$\tilde{\varphi}(u) = \xi_1^2 + \dots + \xi_p^2 - \xi_{p+1}^2 - \dots - \xi_r^2 \quad \left(u = \sum_{i=1}^{n} \xi_i e_i\right)$$

и

$$\tilde{\varphi}(u) = \eta_1^2 + \dots + \eta_{p'}^2 - \eta_{p'+1}^2 - \dots - \eta_r^2 \quad \left(u = \sum_{i=1}^{n} \eta_i f_i\right).$$

Трябва да докажем, че $p = p'$.

Да допуснем, че $p \neq p'$ и нека например $p > p'$. Да означим $V_1 = l(e_1, \dots, e_p)$ $(\dim V_1 = p > 0$, тъй като $p > p' \geq 0)$ и $V_2 = l(f_{p'+1}, \dots, f_r, \dots, f_n)$ $(\dim V_2 = n - p' > 0$, тъй като $n \geq p > p')$. Имаме (теорема 1 от § 10)

$$\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - \dim(V_1 + V_2).$$

Тъй като $\dim(V_1 + V_2) \leq n$, то

$$\dim(V_1 \cap V_2) = p + (n - p') - \dim(V_1 + V_2) \geq p + (n - p') - n = p - p' > 0.$$

Следователно $V_1 \cap V_2 \neq \{\mathbf{0}\}$. Нека $\mathbf{0} \neq \boldsymbol{u} \in V_1 \cap V_2$ и $\boldsymbol{u} = \sum_{i=1}^{p} \xi_i \boldsymbol{e}_i = \sum_{i=p'+1}^{n} \eta_i \boldsymbol{f}_i$. От $\boldsymbol{u} \neq \mathbf{0}$ следва, че поне един от коефициентите $\xi_1, \ldots, \xi_p$ е различен от нула. Сега имаме

$$\tilde{\varphi}(\boldsymbol{u}) = \xi_1^2 + \cdots + \xi_p^2 > 0$$

и

$$\tilde{\varphi}(\boldsymbol{u}) = -\eta_{p'+1}^2 - \cdots - \eta_r^2 \leq 0,$$

което е противоречие. С това теоремата е доказана.

Числото p се нарича *положителен индекс на инерция*, числото $q = r - p$ — *отрицателен индекс на инерция*, а числото $\sigma = p - q$ — *сигнатура* на квадратичната форма.

§ 26. Положително дефинитни квадратични форми

Определение. *Нека V е линейно пространство над $\mathbb{R}$ и $\tilde{\varphi}$ е квадратична форма във V. Ще казваме, че $\tilde{\varphi}$ е положително дефинитна квадратична форма, ако за всеки вектор $\boldsymbol{u}$ от V е изпълнено $\tilde{\varphi}(\boldsymbol{u}) \geq 0$, като $\tilde{\varphi}(\boldsymbol{u}) = 0$ точно когато $\boldsymbol{u} = \mathbf{0}$.*

Нека във V е въведено скаларно произведение $(\boldsymbol{u}, \boldsymbol{v})$. Да положим $\varphi(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$. Тогава φ е симетрична билинейна форма, чиято асоциирана квадратична форма $\tilde{\varphi}$ е положително дефинитна. Обратно, нека $\tilde{\varphi}$ е положително дефинитна квадратична форма и φ е (единствената) симетрична билинейна форма, асоциирана с $\tilde{\varphi}$. Тогава във V можем да въведем скаларно произведение като положим $(\boldsymbol{u}, \boldsymbol{v}) = \varphi(\boldsymbol{u}, \boldsymbol{v})$.

Твърдение 1. *Една квадратична форма в n-мерно пространство е положително дефинитна тогава и само тогава, когато тя има каноничен вид $\lambda_1\xi_1^2 + \cdots + \lambda_n\xi_n^2$ и $\lambda_i > 0$ за $i = 1, \ldots, n$, т. е. когато нормалният вид на формата е $\eta_1^2 + \cdots + \eta_n^2$.*

Доказателството на това твърдение предоставяме на читателя.

Твърдение 2 (едновременна канонизация на двойка форми). *Нека V е крайномерно пространство и $\tilde{\varphi}$ и $\tilde{\psi}$ са две квадратични форми във V, поне една от които е положително дефинитна. Тогава съществува базис на V, в който и двете форми са в каноничен вид.*

Д о к а з а т е л с т в о. Нека $\tilde{\psi}$ е положително дефинитна и ψ е асоциираната ѝ симетрична билинейна форма. Въвеждаме във V скаларно произведение като положим $(\boldsymbol{u}, \boldsymbol{v}) = \psi(\boldsymbol{u}, \boldsymbol{v})$. Според твърдение 3 от предния параграф съществува ортонормиран (спрямо това скаларно произведение) базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ на V, в който квадратичната форма φ е в каноничен вид. От друга страна (тъй като базисът $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ е ортонормиран), ако $\boldsymbol{u} = \sum_{i=1}^{n} \xi_i \boldsymbol{e}_i$, то $\tilde{\psi}(\boldsymbol{u}) = \psi(\boldsymbol{u}, \boldsymbol{u}) = (\boldsymbol{u}, \boldsymbol{u}) = \xi_1^2 + \cdots + \xi_n^2$, така че и квадратичната форма $\tilde{\psi}$ е в каноничен (дори нормален) вид в този базис. Твърдението е доказано.

Нека $A = (a_{ij})_{n \times n}$ е квадратна матрица от ред n. Да положим $\Delta_0 = 1$ и

$$\Delta_k = \begin{vmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \ldots & a_{kk} \end{vmatrix}$$

за $k = 1, \ldots, n$. Ясно е ,че $\Delta_1 = a_{11}$ и $\Delta_n = \det A$. Детерминантите $\Delta_1, \ldots, \Delta_n$ ще наричаме *главни минори* на матрицата A.

Твърдение 3 (метод на Якоби). *Нека* $\dim V = n < \infty$ *и* $e_1, \ldots, e_n$ *е базис на* V. *Нека* $\tilde{\varphi}$ *е квадратична форма във* V *и* $A = (a_{ij})_{n \times n}$ *е матрицата ѝ в този базис. Тогава, ако главните минори на матрицата* A *са различни от нула, то съществува базис* $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ *на* V, *в който квадратичната форма има вида*

$$\tilde{\varphi}(\boldsymbol{u}) = \frac{\Delta_0}{\Delta_1}\xi_1^2 + \frac{\Delta_1}{\Delta_2}\xi_2^2 + \cdots + \frac{\Delta_{n-1}}{\Delta_n}\xi_n^2, \qquad \boldsymbol{u} = \sum_{i=1}^{n} \xi_i \boldsymbol{f}_i \in V.$$

Д о к а з а т е л с т в о. Нека $1 \leq j \leq n$. Да разгледаме системата

$$\left| \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1j}x_j = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2j}x_j = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jj}x_j = 1. \end{array} \right. \tag{1}$$

По-кратко k-тото уравнение на системата може да се запише във вида

$$\sum_{s=1}^{j} a_{ks}x_s = \delta_{kj} = \begin{cases} 0 & \text{при } k < j \\ 1 & \text{при } k = j \end{cases}$$

$(k = 1, \ldots, j)$. Детерминантата на системата е Δ_j и по условие $\Delta_j \neq 0$. Следователно тя има единствено решение $(x_{j1}, x_{j2}, \ldots, x_{jj})$. Замествайки в k-тото уравнение неизвестните $x_1, \ldots, x_j$ съответно с $x_{j1}, \ldots, x_{jj}$, получаваме равенствата

$$\sum_{s=1}^{j} a_{ks}x_{js} = \delta_{kj} \tag{2}$$

$(k = 1, \ldots, j)$. От вида на свободните членове на системата (1) и от формулите на Крамер получаваме $x_{jj} = \dfrac{\Delta_{j-1}}{\Delta_j}$ (формулата е вярна и при $j = 1$). В частност $x_{jj} \neq 0$ $(j = 1, \ldots, n)$.

Да разгледаме матрицата

$$T = \begin{pmatrix} x_{11} & x_{21} & \ldots & x_{n1} \\ & x_{22} & \ldots & x_{n2} \\ & & \ddots & \vdots \\ 0 & & & x_{nn} \end{pmatrix}$$

Матрицата T е триъгълна и в j-ия ѝ стълб $(j = 1, \ldots, n)$ стоят елементите $(x_{j1}, \ldots, x_{jj}, 0, \ldots, 0)$, т.е. първите j елемента са решението на системата (1) (ще подчертаем, че **вторият индекс** на елементите от i-ия **ред** на матрицата T е равен на i). Имаме $\det T = x_{11}x_{22}\ldots x_{nn} \neq 0$ (по-точно $\det T = \dfrac{\Delta_0}{\Delta_1} \cdot \dfrac{\Delta_1}{\Delta_2} \cdots \dfrac{\Delta_{n-1}}{\Delta_n} = \dfrac{\Delta_0}{\Delta_n}$). Тъй като T е обратима матрица, тя е матрица на прехода от базиса $e_1, \ldots, e_n$ към базис $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$, като $\boldsymbol{f}_j = x_{j1}e_1 + x_{j2}e_2 + \cdots + x_{jj}e_j$ $(j = 1, \ldots, n)$. Нека $B = (b_{ij})_{n \times n}$ е матрицата на квадратичната форма $\tilde{\varphi}$ в новия базис и φ е асоциираната ѝ симетрична билинейна форма. Тогава за $1 \leq i \leq j \leq n$ имаме

$$\begin{aligned}
b_{ij} = \varphi(\boldsymbol{f}_i, \boldsymbol{f}_j) &= \varphi\left(\sum_{k=1}^{i} x_{ik}\boldsymbol{e}_k, \sum_{s=1}^{j} x_{js}\boldsymbol{e}_s\right) \\
&= \sum_{k=1}^{i}\sum_{s=1}^{j} x_{ik}x_{js}\varphi(\boldsymbol{e}_k, \boldsymbol{e}_s) = \sum_{k=1}^{i}\sum_{s=1}^{j} x_{ik}x_{js}a_{ks} \qquad (3) \\
&= \sum_{k=1}^{i} x_{ik}\left(\sum_{s=1}^{j} a_{ks}x_{js}\right).
\end{aligned}$$

Сумата в скобите в дясната страна на това равенство е точно лявата страна на равенството (2). Следователно равенството (3) приема вида

$$b_{ij} = \sum_{k=1}^{i} x_{ik}\delta_{kj}. \qquad (4)$$

Сега при $i < j$ имаме $k \leq i < j$, откъдето $\delta_{kj} = 0$ и значи $b_{ij} = 0$. Тъй като матрицата B е симетрична, то $b_{ij} = 0$ и за $i > j$. При $i = j$ равенството (4) приема вида $b_{jj} = \sum_{k=1}^{j} x_{jk}\delta_{kj} = x_{jj}\delta_{jj} = x_{jj} = \frac{\Delta_{j-1}}{\Delta_j}$. Така намираме

$$B = \begin{pmatrix} \dfrac{\Delta_0}{\Delta_1} & & & 0 \\ & \dfrac{\Delta_1}{\Delta_2} & & \\ & & \ddots & \\ 0 & & & \dfrac{\Delta_{n-1}}{\Delta_n} \end{pmatrix}.$$

Следователно в базиса $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ квадратичната форма $\tilde{\varphi}$ има търсеният каноничен вид

$$\tilde{\varphi}(\boldsymbol{u}) = \frac{\Delta_0}{\Delta_1}\xi_1^2 + \frac{\Delta_1}{\Delta_2}\xi_2^2 + \cdots + \frac{\Delta_{n-1}}{\Delta_n}\xi_n^2.$$

Теорема 4 (критерий на Силвестър). *Една квадратична форма в крайномерно пространство V е положително дефинитна тогава и само тогава, когато главните минори на матрицата ѝ във всеки базис на V са положителни числа.*

Д о к а з а т е л с т в о. Нека $\tilde{\varphi}$ е квадратична форма във V и главните минори на матрицата ѝ във всеки базис на V са положителни. Тогава тя има каноничен вид, посочен в твърдение 3 и според твърдение 1 $\tilde{\varphi}$ е положително дефинитна.

Обратно, нека $\tilde{\varphi}$ е положително дефинитна квадратична форма и φ е асоциираната ѝ симетрична билинейна форма. Нека $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ е произволен базис на V и $A = (a_{ij})_{n\times n}$ е матрицата на $\tilde{\varphi}$ в този базис. Въвеждаме във V скаларно произведение като положим $(\boldsymbol{u}, \boldsymbol{v}) = \varphi(\boldsymbol{u}, \boldsymbol{v})$. Имаме $a_{ij} = \varphi(\boldsymbol{e}_i, \boldsymbol{e}_j) = (\boldsymbol{e}_i, \boldsymbol{e}_j)$ $(i, j = 1, \ldots, n)$. Следователно за главните минори Δ_k $(k = 1, \ldots, n)$ на A имаме

$$\Delta_k = \begin{vmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \ldots & a_{kk} \end{vmatrix} = \begin{vmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & \ldots & (\boldsymbol{e}_1, \boldsymbol{e}_k) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{e}_k, \boldsymbol{e}_1) & \ldots & (\boldsymbol{e}_k, \boldsymbol{e}_k) \end{vmatrix} = \Gamma(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k).$$

Тъй като векторите $e_1, \dots, e_k$ са линейно независими, то $\Gamma(e_1, \dots, e_k) > 0$ ($k = 1, \dots, n$) (теорема 1 от § 22), т.е. $\Delta_k > 0$, $k = 1, \dots, n$. С това теоремата е доказана.

§ 27. Унитарни пространства

Понятието скаларно произведение може да се въведе и за линейни пространства над полето на комплексните числа $\mathbb{C}$. Такива пространства се наричат унитарни. Целта на този параграф е да скицира как се пренасят понятията и резултатите, изложени в §§ 21-24, от евклидови за унитарни пространства.

Определение. *Нека V е линейно пространство над $\mathbb{C}$. Ще казваме, че във V е въведено скаларно произведение, ако на всеки два вектора $a, b \in V$ е съпоставено комплексно число, което ще означаваме с (a, b), като се удовлетворяват следните свойства (аксиоми) (тук a, b, c са произволни вектори от V, а λ е произволно комплексно число):*

I. $(a, b) = \overline{(b, a)}$.
II. $(a + b, c) = (a, c) + (b, c)$.
III. $(\lambda a, b) = \lambda(a, b)$.
IV. $(a, a) \in \mathbb{R}$, $(a, a) \geq 0$ *като* $(a, a) = 0$ *точно когато* $a = 0$.

Ако в едно пространство е въведено скаларно произведение, ще казваме, че това пространство е унитарно.

Пример. Нека $V = \mathbb{C}^n$ и $a = (\xi_1, \xi_2, \dots, \xi_n)$, $b = (\eta_1, \eta_2, \dots, \eta_n) \in V$. Полагаме $(a, b) = \xi_1\overline{\eta_1} + \xi_2\overline{\eta_2} + \dots + \xi_n\overline{\eta_n} \in \mathbb{C}$. Непосредствено се проверява, че така дефинираната операция е скаларно произведение.

Ако V е произволно n-мерно пространство над $\mathbb{C}$, при фиксиран базис то се "превръща" в $\mathbb{C}^n$. Така всяко крайномерно пространство може да се направи унитарно.

Непосредствено от аксиомите се извеждат следните следствия (формулирани за произволни вектори от V и произволни скалари от $\mathbb{C}$).

1. $(0, b) = 0$.
2. $(a, b + c) = (a, b) + (a, c)$.
3. $(a, \lambda b) = \overline{\lambda}(a, b)$.
4. За произволни две линейни комбинации $\sum_i \lambda_i a_i$ и $\sum_j \mu_j b_j$ е в сила

$$\Big(\sum_i \lambda_i a_i, \sum_j \mu_j b_j\Big) = \sum_{i,j} \lambda_i \overline{\mu_j}(a_i, b_j).$$

По-нататък всички определения и твърдения от § 21 и § 22, отнасящи се за евклидови пространства, остават в сила (без промени във формулировките) и за унитарни пространства. При това промени в доказателствата практически също няма, освен появата на комплексното спрягане в част от пресмятанията.

По-долу ще скицираме (често без доказателства или само с упътване) как се пренасят за унитарни пространства определенията и твърденията от § 23 и § 24.

До края на параграфа V ще бъде n-мерно унитарно пространство.

* * *

Унитарни оператори. Нека $A \in M_n(\mathbb{C})$. С $\overline{A}$ ще означаваме комплексно спрегнатата матрица на A, т. е. матрицата, която се получава от A като заменим всички елементи на A с комплексно спрегнатите им. Директно се проверява, че $\overline{\det A} = \det \overline{A}$ и (ако A е обратима) $\overline{A^{-1}} = \overline{A}^{-1}$.

Определение. *Една матрица $A \in M_n(\mathbb{C})$ ще наричаме унитарна, ако $A\overline{A}^t = E$.*

Всяка ортогонална матрица е унитарна.

Както и за евклидови пространства, се съобразява, че A е унитарна матрица точно когато векторите-редове (векторите-стълбове) на A образуват ортонормиран базис на пространството $\mathbb{C}^n$. Също така, матрицата на прехода от ортонормиран базис на V към ортонормиран базис на V е унитарна.

Определение. *За един оператор $\varphi \in \operatorname{Hom} V$ ще казваме, че е унитарен, ако за всеки два вектора $\boldsymbol{u}$ и $\boldsymbol{v}$ от V е изпълнено $(\varphi(\boldsymbol{u}), \varphi(\boldsymbol{v})) = (\boldsymbol{u}, \boldsymbol{v})$.*

Както и при евклидови пространства, един оператор е унитарен точно когато матрицата му във всеки ортонормиран базис на V е унитарна.

Твърдение 1. *Собствените стойности на унитарен оператор по модул са равни на 1.*

Следствие 2. *Характеристичните корени на унитарна (а значи и на ортогонална) матрица по модул са равни на 1.*

Твърдение 3. *Всеки два собствени вектора, съответстващи на различни собствени стойности на унитарен оператор, са ортогонални помежду си.*

Теорема 4. *Нека V е унитарно пространство и φ е унитарен оператор във V. Тогава съществува ортонормиран базис на V, в който матрицата на φ е диагонална (по главния ѝ диагонал стоят собствените стойности на φ, а базисните вектори са собствени вектори на φ).*

Доказателството на тази теорема в идейно отношение не се различава от доказателството на аналогичната теорема за симетричен оператор (теорема 5 от § 24).

Както следствие 6 от § 23, и тук получаваме

Следствие 5. *За всяка унитарна матрица A съществува унитарна матрица T, такава че матрицата $D = T^{-1}AT$ е диагонална.*

* * *

Ермитови оператори.

Определение. *Една матрица $A \in M_n(\mathbb{C})$ ще наричаме ермитова, ако $\overline{A}^t = A$.*

Всяка симетрична матрица е ермитова.

Ако $A = (a_{ij})_{n\times n}$, то A е ермитова матрица точно когато $a_{ji} = \overline{a_{ij}}$ ($i, j = 1, \dots, n$), т. е. когато елементите на A, симетрични относно главния диагонал, са комплексно спрегнати.

Определение. *За един оператор $\varphi \in \operatorname{Hom} V$ ще казваме, че е ермитов, ако за всеки два вектора $\boldsymbol{u}$ и $\boldsymbol{v}$ от V е изпълнено $(\varphi(\boldsymbol{u}), \boldsymbol{v}) = (\boldsymbol{u}, \varphi(\boldsymbol{v}))$.*

Както и при евклидови пространства, един оператор е ермитов точно когато матрицата му във всеки ортонормиран базис на V е ермитова.

Твърдение 6. *Собствените стойности на ермитов оператор са реални числа.*

Д о к а з а т е л с т в о. Нека φ е ермитов оператор, λ е собствена стойност на φ и $\boldsymbol{v}$ е съответстващ ѝ собствен вектор на φ. Имаме $(\varphi(\boldsymbol{v}), \boldsymbol{v}) = (\boldsymbol{v}, \varphi(\boldsymbol{v}))$, откъдето последователно получаваме

$$(\lambda v, v) = (v, \lambda v), \quad \lambda(v, v) = \overline{\lambda}(v, v), \quad (\lambda - \overline{\lambda})(v, v) = 0.$$

Тъй като $(v, v) \neq 0$, то $\lambda = \overline{\lambda}$ и значи λ е реално число.

Преминавайки от операторен на матричен език, получаваме

Следствие 7. *Характеристичните корени на ермитова (а значи и на симетрична) матрица са реални числа.*

Твърдение 8. *Всеки два собствени вектора, съответстващи на различни собствени стойности на ермитов оператор, са ортогонални помежду си.*

Теорема 9. *Нека φ е ермитов оператор във V. Тогава съществува ортонормиран базис на V, в който матрицата на φ е диагонална (по главния ѝ диагонал стоят собствените стойности на φ, а базисните вектори са собствени вектори на φ).*

Доказателството на тази теорема не се различава от доказателството на съответната теорема за симетричен оператор (теорема 5 от § 24). По същия начин получаваме и аналог на следствие 6 от § 24, а именно:

Следствие 10. *За всяка ермитова матрица A съществува унитарна матрица T, такава че матрицата $D = T^{-1}AT$ е диагонална (и значи всяка ермитова матрица е подобна на диагонална).*

ЛИТЕРАТУРА

1. К. Дочев, Д. Димитров, *Линейна алгебра*, София, 1973.
2. А. Г. Курош, *Курс высшей алгебры*, Москва, 1965.
3. И. М. Гельфанд, *Лекции по линейной алгебре*, Москва, 1971.

С Ъ Д Ъ Р Ж А Н И Е