НИКОЛА ПЕТРОВ
НИКОЛА ЗЯПКОВ

ЛИНЕЙНА АЛГЕБРА И АНАЛИТИЧНА ГЕОМЕТРИЯ

част I

УНИВЕРСИТЕТСКО ИЗДАТЕЛСТВО
"ЕПИСКОП КОНСТАНТИН ПРЕСЛАВСКИ"

НИКОЛА ПЕТРОВ
НИКОЛА ЗЯПКОВ

ЛИНЕЙНА АЛГЕБРА И АНАЛИТИЧНА ГЕОМЕТРИЯ

част I

ШУМЕН
2010

ISBN 954-577-079-1

С Ъ Д Ъ Р Ж А Н И Е

П Р Е Д Г О В О Р

“Алгебрата не е нищо друго, освен написана геометрия, а геометрията е само образна алгебра.”

С.Жермен

В основата на тази книга са лекциите, които авторите от доста години четат за студентите от математическите и физическите специалности в Шуменски университет “Епископ Константин Преславски”. Тя няма други претенции, освен да бъде учебник, който би подпомогнал студентите в самостоятелната им работа за овладяване на дисциплината. В Шуменския университет курсът по линейна алгебра и аналитична геометрия се чете през първия и втория семестър, като в края на всеки семестър студентите полагат изпит. По тази причина сметнахме, че е по-рационално учебникът да се раздели на две отделни книги.

Съдържанието на първата от тях, която читателят вече държи в ръце, е подбрано по принципа “без това не може”, който предопредели неговата традиционност. Частичните изяви на авторски вкусове и предпочитания отложихме за втората част.

При изложението сме се придържали към схващането, че в учебник по математика най-абстрактните понятия трябва да се въвеждат едва тогава, когато читателят е натрупал достатъчно опит от работа в конкретни ситуации и дори сам е почувствал необходимостта от обобщение. При такъв подход, разбира се, повторенията на разсъждения и формулировки стават неизбежни, но се утешаваме с мисълта, че книгата е учебник (при това за първокурсници), а не трактат. С тези думи се мъчим донякъде да оправдаем например смущаващия факт, че абстрактните линейни пространства се въвеждат едва в гл. 4, а основните резултати за крайномерни пространства са получени в гл. 2 и гл. 3, но за конкретните координатни пространства $\mathbf{R}^n$ и $\mathbf{C}^n$. Надяваме се, че след съответната теорема за изоморфизъм чувството за неловкост отпада.

В цялата книга се предполага, че читателят е запознат с елементарните понятия от теорията на множествата и изображенията. При необходимост той би могъл да се обърне към уводните параграфи на повечето книги по математически анализ или към цитираните по-долу книги [1], [4] и [6]. Някои необходими сведения, които изпреварват системния курс по висша алгебра, са изложени в гл. 8. Добре е веднага след гл. 1 (или пък преди нея) да се прочете § 1 на гл. 7.

Задачите за упражнение, с които завършва почти всяка глава, трябва да се разглеждат като неделима част от основния текст. Като правило те доразвиват темата или излагат друга гледна точка.

Като се надяваме, че този учебник би могъл да бъде полезен и на други студенти, освен онези в Шуменския университет, бихме желали да окуражим младите читатели чрез думите на Е. Мах: “Студентът по математика често трудно се отърсва от неприятното чувство, че неговата наука ... го превъзхожда по интелигентност - едно впечатление, от което великият Ойлер признава, че изобщо не е могъл да се освободи. Това чувство може да се оправдае, като си помислим, че повечето идеи, с които боравим, често са били открити от други преди много векове. До голяма степен пред нас в науката се възправя интелектът на предшествениците ни”.

Участието на авторите е следното: глава 5, както и § 1 и § 2 от гл. 7 написа Н. Зяпков; отговорността за останалия текст е изцяло на Н. Петров.

Изказваме нашата искрена благодарност на рецензентите проф.д.м.н. Грозьо Станилов и проф. д.м.н Недялко Ненов - за вниманието, което проявиха към ръкописа, и за многобройните конструктивни препоръки. Неоценима помощ ни оказа и редакторът проф. д.м.н Керопе Чакърян, който най-акуратно прочете целия текст и бележките му ни бяха особено полезни. Изказваме му нашата сърдечна благодарност.

Авторите

ДОПЪЛНИТЕЛНА ЛИТЕРАТУРА

1. Александров, П.С., Лекции по аналитической геометрии, М., Наука, 1968.
2. Гаврилов, М., Гр. Станилов, Линейна алгебра и аналитична геометрия, С., НИ, 1991.
3. Дочев, К., Д. Димитров, Линейна алгебра, С., НИ, 1973.
4. Кострикин , А.И., Въведение в алгебрата, С., НИ, 1981.
5. Кострикин , А.И., Ю.И.Манин, Линейна алгебра и геометрия, С., НИ, 1990.
6. Тышкевич, Р.И., А.С. Феденко, Линейная алгебра и аналитическая геометрия, Минск, Вышэйшая школа, 1976.
7. Lang, S., Algèbre linéair, Paris, InterEdutions, 1976.

Г л а в а 1

ДЕТЕРМИНАНТИ

§ 1. Системи линейни уравнения - общи понятия

1. Редица важни задачи в математиката и нейните приложения се свеждат до решаване на т. нар. *системи от линейни алгебрични уравнения*. Общият вид на такава система е

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

При тези означения $x_1, x_2, \ldots, x_n$ са *неизвестни*, а уравненията са m на брой. Предполага се, че a_{ij}, b_i $(i = 1, \ldots, m;\ \ j = 1, \ldots, n)$, наричани съответно *коефициенти пред неизвестните* и *свободни членове,* са дадени числа. Удобно е коефициентите пред неизвестните да се означават с една и съща буква, снабдена с двоен индекс: в случая a_{ij} (чете се а-и-жи). Първият индекс показва номера на уравнението, а вторият - номера на неизвестното. Иначе казано, a_{ij} е коефициентът пред x_j в i -тото уравнение. Ако всички свободни членове са нули, т.е. $b_1 = b_2 = \ldots = b_m = 0$, казва се че системата е *хомогенна*.

2. Определение. Ще казваме, че наредената n-торка от числа $\left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ е *решение* на системата (1), ако след заместване на $x_1, \ldots, x_n$ съответно с $x_1^{(0)}, \ldots, x_n^{(0)}$ всяко от уравненията се превръща в числово равенство. Ако системата има решение и то е единствено, ще казваме, че тя е *определена*; ако тя има повече от едно решение, ще казваме, че е *неопределена*. Ако системата изобщо няма решение, ще казваме, че е *несъвместима*.

Скоро ще видим, че ако една система има поне две решения (неопределена), то тя има безбройно много решения. Главната ни цел ще бъде да изложим някои методи за намиране на решенията, както и да намерим необходими и достатъчни условия за съществуването на решение, т.е. да отговорим прецизно на въпроса, кога дадена система е съвместима.

3. Ясно е, че всяка система се задава еднозначно чрез коефициентите пред неизвестните и свободните членове (съвкупност от числа) и точното им местоположение. В тази връзка е удобно въвеждането на още едно понятие. Под (числова) *матрица* ще разбираме правоъгълна таблица от числа. Така например, таблицата

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

ще наричаме *матрица от коефициентите пред неизвестните* в системата (1), или накратко *матрица на системата*. Матрицата

$$\overline{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ .. & .. & .. & .. & .. \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

ще наричаме *разширена матрица* на системата. Очевидно е, че (1) се задава еднозначно чрез матрицата $\overline{A}$.

По естествен начин може да се говори за редовете и стълбовете на дадена матрица. Ако тя се състои от m реда и n стълба, ще казваме накратко, че е от *тип* $m \times n$. Ако броят на редовете е равен на броя на стълбовете, матрицата се нарича *квадратна*. Ако квадратна матрица се състои от n реда, ще казваме, че тя е от *ред* n . В квадратна матрица елементите $a_{11}, a_{22}, \dots, a_{nn}$ определят т. нар. *главен диагонал;* елементите $a_{1n},\ a_{2\ n-1}, \dots, a_{n1}$ определят *втория диагонал.* Ако в квадратна матрица всички елементи под (съответно над) главния диагонал са нули, ще казваме, че тя е *горна* (съответно *долна*) *триъгълна*. Квадратните матрици от вида

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

се наричат *диагонални* и обикновено се означават с $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. В случаите, когато от контекста е ясно какъв е типът на матрицата A, често се използва съкратено записване, при което се посочва само "общият елемент" на A, например $A = (a_{ij})$.

§ 2. Еквивалентност на системи линейни уравнения

1. Определение. Ще казваме, че две системи линейни уравнения с n неизвестни са *еквивалентни*, ако всяко решение на първата система е решение и на втората и обратно, всяко решение на втората система е решение и на първата.

С други думи две системи са еквивалентни точно тогава, когато множеството от решенията на първата система съвпада с множеството от решенията на втората. Разбира се, всеки две несъвместими системи са еквивалентни, защото множествата от решенията им са празното множество.

За да решим дадена система, ще я подменяме с друга еквивалентна, чиито решения се намират лесно. Трябва да сме сигурни обаче, че новата система е еквивалентна на изходната, защото в противен случай или ще "изгубим" решение, или ще получим и такива, които не удовлетворяват изходната система. Това налага да изучим някои основни операции (елементарни преобразувания) върху дадена система, които я трансформират в еквивалентна.

2. Твърдение. а) *Ако в една линейна система разменим местата на две уравнения, а останалите запазим непроменени, новата система е еквивалентна на изходната.*

б) *Ако умножим i -тото уравнение с някакво число и го прибавим към k -тото, $k \neq i$, то новата система, в която всички уравнения, освен k -тото са същите, а k -тото е променено по посочения начин, е еквивалентна на изходната.*

в) *Ако умножим някое от уравненията с число, различно от нула, новата система е еквивалентна на изходната.*

Д о к а з а т е л с т в о . Твърденията а) и в) са очевидни. За да докажем б), нека разглежданата система е (1), а "новата" е

$$
(2)\qquad
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\left(a_{k1}+ca_{i1}\right)x_1 + \left(a_{k2}+ca_{i2}\right)x_2 + \ldots + \left(a_{kn}+ca_{in}\right)x_n = b_k + cb_i\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m.
\end{aligned}
$$

Двете системи се различават евентуално по k-тото уравнение. Ако $\left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ е решение на (1), то удовлетворява всички уравнения на (1), а от

$$
(3)\qquad a_{i1}x_1^{(0)} + \ldots + a_{in}x_n^{(0)} = b_i,
$$

$$
(4)\qquad a_{k1}x_1^{(0)} + \ldots + a_{kn}x_n^{(0)} = b_k
$$

следва

$$
(5)\qquad \left(a_{k1}+ca_{i1}\right)x_1^{(0)} + \ldots + \left(a_{kn}+ca_{in}\right)x_n^{(0)} = b_k + cb_i,
$$

т.е. то удовлетворява всички уравнения на (2), включително и k-тото. Обратно, ако $\left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ е решение на системата (2), то разбира се удовлетворява всички уравнения на (1), освен може би k-тото. Тъй като сега (3) и (5) са верни равенства, като умножим (3) с $-c$ и го прибавим към (5), ще получим равенството (4), т.е. разглежданото решение удовлетворява и k-тото уравнение на (1).

От тези разсъждения се вижда и това, че ако едната от системите (1) и (2) е несъвместима, то другата също е несъвместима. Твърдението е доказано.

Вече е ясно, че ако върху една система приложим краен брой пъти елементарните преобразувания, описани в твърдението, ще стигнем до система, еквивалентна на дадената. Този факт непрекъснато ще имаме предвид в следващия параграф.

§ 3. Метод на Гаус за решаване на линейни системи

1. Ще изложим един алгоритъм за ефективно намиране на решенията на линейна система. Известен е като *метод на Гаус* или *метод на последователното изключване на неизвестните*.

Отново разглеждаме системата (1). Ако във всички уравнения коефициентите пред x_1 са нули, то x_1 ще приема съвършено произволни стойности и по същество решаваме система с $n-1$ неизвестни $x_2, x_3, \ldots, x_n$. По тази причина е естествено да считаме, че поне един от коефициентите пред x_1 е различен от нула. След разместване на уравненията ще предполагаме, че това е коефициентът в първото уравнение. И така, нека $a_{11} \neq 0$. Като умножим последователно първото уравнение с $a_{11}^{-1} a_{21}$ и го прибавим към второто, след това го умножим с $a_{11}^{-1} a_{31}$ и го прибавим към третото и т.н., ще получим система от вида

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\
a'_{22}x_2 + \ldots + a'_{2n}x_n &= b'_2 \\
\ldots\ldots\ldots\ldots\ldots & \\
a'_{m2}x_2 + \ldots + a'_{mn}x_n &= b'_m \ ,
\end{aligned}
\qquad (1')
$$

която е еквивалентна на изходната. Казваме, че във второто, третото и т.н. до последното уравнение сме изключили неизвестното x_1.

Ако всички коефициенти a'_{ij} , $i = 2, \ldots, m$, $j = 2, \ldots, n$, са нули, изключването на неизвестните е приключено.

Нека второто, третото и т.н. уравнения не съдържат неизвестните $x_1, x_2, \ldots, x_{k-1}$, но x_k , $1 < k \leq n$, участва с ненулев коефициент в поне едно от тези уравнения. При такава уговорка след подходящо разместване на уравненията и смяна на означенията, системата (1') може да се запише във вида

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\
a''_{2k}\,x_k + \dots + a''_{2n}\,x_n &= b''_2 \\
\dots\dots\dots\dots\dots\dots \\
a''_{mk}\,x_k + \dots + a''_{mn}\,x_n &= b''_m \ ,
\end{aligned}
$$

където $a''_{2k} \neq 0$. Сега умножаваме второто уравнение последователно с ${a''_{2k}}^{-1} a''_{3k}$ и го прибавяме към третото, после го умножаваме с ${a''_{2k}}^{-1} a''_{4k}$ и прибавяме към четвъртото и т.н. и изключваме x_k. Ще получим система от вида

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 \quad + \quad \dots \quad + \ a_{1n}x_n &= b_1 \\
a''_{2k}\,x_k \quad + \quad \dots \quad + a''_{2n}\,x_n &= b''_2 \\
a'''_{3k+1}x_{k+1} + \dots + a'''_{3n}\ x_n &= b'''_3 \\
\dots\dots\dots\dots\dots\dots \\
a'''_{mk+1}x_{k+1} + \dots + a'''_{mn}\ x_n &= b'''_m \ .
\end{aligned}
$$

Надяваме се вече е ясно, че чрез многократно прилагане на описаната процедура в общия случай ще стигнем до система от вида

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 \quad + \quad \dots \quad + a_{1n}x_n &= \overline{b}_1 \\
\overline{a}_{2k}x_k \ + \dots \ + \overline{a}_{2n}x_n &= \overline{b}_2 \\
\overline{a}_{3l}x_l + \dots + \overline{a}_{3n}x_n &= \overline{b}_3 \\
\dots\dots\dots\dots\dots\dots \\
\overline{a}_{rs}x_s + \dots + \overline{a}_{rn}x_n &= \overline{b}_r \\
0 &= \overline{b}_{r+1} \\
\dots\dots\dots\dots\dots\dots \\
0 &= \overline{b}_m ,
\end{aligned}
\tag{1''}
$$

където $1 \leq r \leq m$ и $a_{11}, \bar{a}_{2k}, \bar{a}_{3l}, \ldots, \bar{a}_{rs}$ са различни от нула. При $r > 1$ се предполага, че $1 < k < l < \ldots < s \leq n$. При $r = m$ считаме, че системата не съдържа уравнения от вида $0 = \bar{b}_i$.

За система от вида (1") ще казваме, че е *стъпаловидна*. Доказахме всъщност, че *всяка линейна система уравнения може да се приведе в стъпаловидна форма*. Изложеният метод за привеждане в стъпаловидна форма е известен като метод на Гаус.

Намирането на решенията на стъпаловидната система (1") е съвсем проста задача. Ако системата съдържа уравнение от вида $0 = \bar{b}_i$, където $\bar{b}_i \neq 0$, тя очевидно е несъвместима. Нека не съдържа такова уравнение, т.е. нека $\bar{b}_{r+1} = \ldots = \bar{b}_m = 0$. Ще я решаваме "отдолу нагоре". На неизвестните x_i с номера i, различни от $1, k, l, \ldots, s$, даваме произволни стойности λ_i, т.е. полагаме $x_i = \lambda_i$. Последното нетривиално уравнение на системата добива вида

$$\bar{a}_{rs} x_s + \bar{a}_{rs+1} \lambda_{s+1} + \ldots + \bar{a}_{rn} \lambda_n = \bar{b}_r, \qquad \bar{a}_{rs} \neq 0,$$

и от него определяме (еднозначно!) x_s. Ако $r > 1$, в уравнението с номер $r-1$ заместваме x_s с намерената стойност, а останалите неизвестни заместваме с $x_i = \lambda_i$. Ако в него x_q е неизвестното с най-малък номер, то по условие $\bar{a}_{r-1\, q} \neq 0$ и определяме (пак еднозначно!) x_q. Като продължим по този начин, ще стигнем до x_1 и всички решения на системата са намерени. Ясно е, че броят на неизвестните, на които дадохме произволни стойности, е $n - r$; наричат се *свободни неизвестни*. Останалите r на брой неизвестни, наричани *главни*, се определят еднозначно чрез свободните. При конкретните означения главни са неизвестните $x_1, x_k, x_l, \ldots, x_s$.

Преди да резюмираме във вид на твърдения получените резултати, ето един

2. Пример. Разглеждаме системата

$$\begin{aligned}
2x_1 + 3x_2 \qquad\qquad - x_5 \qquad\qquad + x_8 &= 0 \\
3x_3 - 4x_4 + x_5 \qquad\qquad &= 1 \\
x_5 + 2x_6 - x_7 \quad &= -2.
\end{aligned}$$

Тя очевидно е стъпаловидна. Главни са неизвестните x_1, x_3 и x_5, а останалите - свободни. Полагаме $x_2 = \lambda_2$, $x_4 = \lambda_4$, $x_6 = \lambda_6$, $x_7 = \lambda_7$, $x_8 = \lambda_8$ и заместваме в уравненията. От третото получаваме $x_5 = -2 - 2\lambda_6 + \lambda_3$. Заместваме x_5 във второто уравнение и намираме $x_3 = \frac{1}{3}\left(3 + 4\lambda_4 + 2\lambda_6 - \lambda_7\right)$. Накрая заместваме x_5 в първото уравнение и намираме $x_1 = -1 - \frac{3}{2}\lambda_2 - \lambda_6 + \frac{1}{2}\lambda_7 - \frac{1}{2}\lambda_8$. Общото решение зависи от 5 независими параметъра - броя на свободните неизвестни.

3. Теорема. а) *Линейна система от* m *уравнения с* n *неизвестни е съвместима тогава и само тогава, когато съответната й стъпаловидна форма не съдържа уравнения от вида* $0=\overline{b_i}$ *, където* $\overline{b_i} \neq 0$.

б) *Ако системата е съвместима и стъпаловидната форма съдържа* r *уравнения (без да броим тривиалните уравнения от вида* $0=0$*), то броят на главните неизвестни е* r *, а броят на свободните е* $n-r$.

в) *Всяка съвместима система линейни уравнения, в която уравненията са по-малко от неизвестните, е неопределена.*

г) *Всяка хомогенна система, в която уравненията са по-малко от неизвестните, е неопределена.*

д) *Една съвместима линейна система с* n *неизвестни е определена тогава и само тогава, когато броят* r *на главните неизвестни е равен на* n .

Наистина вече видяхме, че ако системата се привежда в стъпаловидна форма, в която няма уравнения от вида $0=\overline{b_i}$, където $\overline{b_i} \neq 0$, последната е съвместима, а следователно такава е и изходната. Ако пък стъпаловидната система е съвместима, тя, разбира се, не може да съдържа уравнения от вида $0=\overline{b_i}$, където $\overline{b_i} \neq 0$.

Доказателството на б) се съдържа в разсъжденията от т. 1. Твърдението в) следва веднага от б), защото ако $m<n$, то броят $n-r$ на свободните неизвестни е поне 1, тъй като винаги $r \leq m$. Като даваме на свободните неизвестни произволни стойности, можем да получим дори безбройно много решения.

Твърдението г) е частен случай на в), стига да забележим, че всяка хомогенна система (свободните членове са нули) е съвместима: едно решение е например $(0, 0, \ldots, 0)$. Често ще използваме г), затова го формулирахме отделно.

По повод на д) ще забележим отново, че ако системата е съвместима и $r<n$, то на $(n-r)$-те свободни неизвестни може да се дават произволни стойности и системата има безбройно много решения. Следователно, ако една система е определена, в сила е равенството $r=n$.

Обратно, нека си мислим, че системата е съвместима и $r=n$. Това означава, че стъпаловидната форма има вида

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1 \\ \overline{a}_{22}x_2 + \ldots + \overline{a}_{2n}x_n &= \overline{b}_2 \\ \ldots\ldots\ldots \\ \overline{a}_{nn}x_n &= \overline{b}_n, \end{aligned}$$

където всички коефициенти $a_{11}, \overline{a}_{22}, \ldots, \overline{a}_{nn}$ са ненулеви. (Уравненията от вида $0=0$ са изпуснати.) Ясно е, че системата има единствено решение, т.е. тя е определена.

От проведените разсъждения се оказа, че *произволна система линейни уравнения или няма решение, или има безбройно много решения, или има точно едно решение*.

Читателят сигурно е забелязал, че стъпаловидната форма, към която се привежда дадена система, не е еднозначно определена: при последователното изключване на неизвестни многократно сменяхме номерацията на уравненията и тези смени можеше да се правят, изобщо казано, по различни начини. Следователно разумно е да се питаме доколко числото r от теоремата (броят на главните неизвестни) зависи от стъпаловидната форма. По-късно ще видим, че то не зависи от начина, по който е получена стъпаловидната форма.

§ 4. Детерминанти от втори и трети ред. Общо определение на детерминанта

1. Методът на Гаус, изложен в предишния параграф, е прост и удобен алгоритъм за практическо намиране на решенията на произволна система от линейни уравнения. Наличието на този метод, независимо от безспорните му достойнства, не дава основание да мислим, че въпросът за решенията на произволна линейна система е намерил задоволителен отговор. Ето например две съображения.

а) Желателно е да се намерят общи формули, които да изразяват в "явен вид" неизвестните чрез коефициентите на системата. Такива формули методът на Гаус не дава.

б) В редица теоретични въпроси е достатъчно да се знае само дали системата има решения и от колко параметъра зависят те, без да е нужно тяхното конкретно намиране, т.е. желателно е да разполагаме с удобни необходими и достатъчни условия за съвместимост на системата. Засега знаем само твърдението а) на теоремата от § 3, т. 3. Дали ще се получи обаче уравнение от вида $0 = b \neq 0$, или не, можем да кажем едва след като доведем докрай цялата процедура по привеждането на системата в стъпаловидна форма.

Алгебрата отдавна е дала задоволителен отговор на повдигнатите въпроси. За целта се е наложило да бъде създаден нов и силен инструмент - теорията на детерминантите. Преди да въведем общото понятие, ще направим два традиционни експеримента.

2. Дадена е системата

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 &= b_1 \\
a_{21}x_1 + a_{22}x_2 &= b_2 .
\end{aligned} \tag{1}
$$

Ако първото и второто уравнение умножим съответно с a_{22} и $-a_{12}$ и ги съберем, а след това ги умножим съответно с $-a_{21}$ и a_{11} и отново ги съберем, ще получим системата

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

Разбира се, ако x_1 и x_2 удовлетворяват (1), то те ще удовлетворяват и (2). За да схванем правилото, по което са образувани коефициентите в системата (2), нека разгледаме матриците от втори ред

$$
\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}, \qquad \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}.
$$

Ако във всяка от тях от произведението на елементите по главния диагонал извадим произведението на елементите по втория диагонал, ще получим съответно

$$
a_{11}a_{22} - a_{12}a_{21}, \qquad b_1a_{22} - a_{12}b_2, \qquad a_{12}b_2 - b_1a_{21},
$$

т.е. коефициентите в системата (2). Следователно удобно е да се въведе следното

3. Определение. Нека A е произволна матрица от втори ред, т.е.

$$A=\left(a_{ij}\right)=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Нейна *детерминанта* ще наричаме числото $a_{11}a_{22}-a_{12}a_{21}$ и ще го бележим с някой от следните символи

$$|A|, \qquad \det\left(A\right), \qquad \det\left(a_{ij}\right), \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Ако в системата (2) $\det\left(a_{ij}\right)\neq 0$, получаваме следните формули:

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2=\frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}},$$

наричани *формули на Крамер*.

А сега да преминем към следващата по трудност задача. Разглеждаме системата

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3&=b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3&=b_2 \\ a_{31}x_1+a_{32}x_2+a_{33}x_3&=b_3. \end{aligned} \tag{3}$$

Ако първото, второто и третото уравнение умножим съответно с $a_{22}a_{33}-a_{23}a_{32}$, $-\left(a_{12}a_{33}-a_{13}a_{32}\right)$, $a_{12}a_{23}-a_{13}a_{22}$ и ги съберем, след лесни пресмятания ще получим

$$\begin{aligned} &\left(a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}+a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\right)x_1= \\ &=b_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-b_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}+b_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}. \end{aligned} \tag{4}$$

Множителите, които използвахме, са подбрани така, че след събирането на уравненията коефициентите пред x_2 и x_3 да се окажат нули. Ако читателят иска да стигне до тези множители по естествен начин, препоръчваме му да реши зад. 1 от упражненията и да приложи резултата от нея към системата

$$\begin{aligned} a_{12}y_1 + a_{22}y_2 + a_{32}y_3 &= 0 \\ a_{13}y_1 + a_{23}y_2 + a_{33}y_3 &= 0. \end{aligned}$$

Множителите по-горе са едно от решенията на тази система. По-нататък той сам ще съобрази как да получи още две уравнения, аналогични на (4), в които да участват съответно само x_2 и x_3.

Като вземем повод от (4), въвеждаме следното

4. Определение. Нека A е произволна матрица от трети ред, т.е.

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

Нейна *детерминанта* ще наричаме числото

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

Ще го бележим отново с някой от символите

$$|A|, \qquad \det\left(A\right), \qquad \det\left(a_{ij}\right), \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

За удобство на речта често ще казваме само "детерминанта от втори (или трети) ред" вместо детерминанта на матрица от втори или трети ред. Ако $A = \left(a_{11}\right)$ е матрица от тип 1×1, уславяме се под нейна детерминанта да разбираме числото a_{11}.

И така, разгледаните два частни случая навеждат на мисълта, че ще бъде уместно да се въведе следното общо

5. Определение. Нека $A = \left(a_{11}\right)$ е квадратна матрица от ред 1. Нейна детерминанта ще наричаме числото a_{11}. Нека за някое естествено число $n \geq 2$ вече сме дефинирали детерминанта от $(n-1)$-ви ред и нека

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ .. & .. & .. & .. \\ a_{i1} & a_{i2} & \dots & a_{in} \\ .. & .. & .. & .. \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

е произволна квадратна матрица от ред n. Да зачертаем в нея първия стълб и i -тия ред $(i = 1,2,\dots,n)$ и да означим с Δ_{i1} детерминантата на получената матрица от $(n-1)$-ви ред. Числото

$$a_{11}\Delta_{11} - a_{21}\Delta_{21} + \dots + \left(-1\right)^{i+1} a_{i1}\Delta_{i1} + \dots + \left(-1\right)^{n+1} a_{n1}\Delta_{n1}$$

ще наричаме *детерминанта на матрицата* A и ще го бележим с един от следните символи:

$$|A|, \qquad \det\left(A\right), \qquad \det\left(a_{ij}\right), \qquad \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ .\ . & .\ . & .\ . & .\ . \\ a_{i1} & a_{i2} & \dots & a_{in} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Въпреки че детерминантата е число, като имаме пред очи последното означение, често ще говорим за нейните редове и стълбове.

Очевидно е, че при n =2 и 3 последното определение съвпада с определенията от т. 3 и т. 4. Интуитивно е ясно, че детерминантата е дефинирана за всяко естествено число n : щом знаем какво е детерминанта от трети ред, то можем да кажем какво е детерминанта от четвърти ред, а следователно и от пети ред и т.н.

§ 5. Пермутации

1. Ще започнем със следната елементарна задача. В зала с n места предстои да седнат n посетители. По колко различни начина те могат да се разположат в залата, ако всеки има право да заема всяко свободно място?

Нека си мислим, че посетителите влизат в залата един след друг. Тогава първият може да си избере място по n начина, вторият - по $n-1$, следващият - по $n-2$ начина и т.н. Очевидно n-ят ще трябва да заеме единственото свободно място. Следователно броят на различните разполагания в залата е $n\left(n-1\right)\left(n-2\right)\ldots 1$. Ако е останало неясно защо умножихме "възможностите", може да разсъждаваме и така. Нека $f\left(n\right)$ е броят на всички възможни разполагания в залата. Влиза първият посетител и си избира, да речем, стол [1] 1. Тогава останалите ще се разположат по $f\left(n-1\right)$ начина. Ако първият посетител седне обаче на стол [1] 2, за останалите ще има отново $f\left(n-1\right)$ възможни разполагания (но вече други!). От това лесно се съобразява, че

$$f\left(n\right)=nf\left(n-1\right)$$

за всяко естествено число $n\geq 2$. Като приложим тази формула за $n-1$, $n-2$, ..., 2 и вземем предвид, че $f\left(1\right)=1$, то

$$\begin{aligned}
f\left(n\right)&=nf\left(n-1\right)\\
f\left(n-1\right)&=\left(n-1\right)f\left(n-2\right)\\
f\left(n-2\right)&=f\left(n-3\right)\\
&\ldots\ldots\ldots\ldots\\
f\left(2\right)&=2f\left(1\right)\\
f\left(1\right)&=1
\end{aligned}$$

и с почленно умножаване и съкращаване получаваме $f\left(n\right)=1.2.3\ldots n$.

Прието е произведението $1\,.\,2\,.\,3\ldots n$ да се означава с $n!$ (чете се: ен-факториел). Приема се и съглашението $0!=1$.

Уместно е да се въведе следното

2. Определение. Всяко разполагане на n елемента в някакъв ред ще наричаме *пермутация* на тези елементи.

Очевидно резултатът от т. 1 може да се формулира и така: *броят на всички пермутации на* n *елемента е равен на* $n!$.

В редица случаи се налага да сравняваме (в някакъв смисъл) доколко дадена пермутация на разглежданите елементи се различава от някоя избрана и фиксирана, наричана обикновено *главна*. Ако разглежданите n елемента са реални числа, сравняването най-често се прави спрямо пермутацията, в която те са подредени в растящ ред (от най-малкото към най-голямото число).

3. Определение. Нека $a_1, a_2, \dots, a_n$ са различни реални числа и $a, \dots, b, \dots, c, \dots, d$ е някаква тяхна пермутация. Ако $b > c$, ще казваме, че между b и c има *инверсия*. Броят на всички инверсии в пермутацията ще бележим с $\left[a, \dots, b, \dots, c, \dots, d\right]$. Ако броят на инверсиите е четно (съответно нечетно) число, ще казваме, че пермутацията е *четна* (съответно *нечетна*).

Иначе казано, между числата b и c има инверсия, ако по-голямото число предхожда по-малкото. За да преброим инверсиите в някаква пермутация $b_1, b_2, \dots, b_n$, обикновено постъпваме така: преброяваме инверсиите, които b_1 образува с числата, намиращи се вдясно от него, към техния брой прибавяме броя на инверсиите, които b_2 образува с числата, стоящи вдясно от него, след това прибавяме броя на инверсиите, които b_3 образува с числата, отстоящи вдясно от него, и т.н. Например:

$$\left[2, 5, 1, 3, 4, 7, 6\right] = 1 + 3 + 0 + 0 + 0 + 1 = 5$$

4. Забележка. Нека $i_1, i_2, \dots, i_n$ е пермутация на числата $1, 2, \dots, n$, а $a_1 < a_2 < \dots < a_n$ са реални числа. Тогава

$$\left[a_{i_1}, a_{i_2}, \dots, a_{i_n}\right] = \left[i_1, i_2, \dots, i_n\right].$$

Твърдението е очевидно, защото $a_{i_k} < a_{i_l}$ тогава и само тогава, когато $i_k < i_l$.

5. Лема. *Ако в една пермутация на числа разменим местата на две от числата, ще получим нова пермутация, която има четност, различна от четността на изходната.*

(Често казваме кратко, че ако разменим местата на две от числата, пермутацията сменя четността си; самото разместване на две от числата ще наричаме *транспозиция*.)

Доказателство. Да предположим най-напред, че числата, които разменяме, са съседни, т.е. разглеждаме пермутацията

$$\ldots, a, b, \ldots,$$

където многоточията заместват числа, които не са изписани. След размяна на a и b получаваме пермутацията

$$\ldots, b, a, \ldots$$

Очевидно, ако в изходната пермутация между a и b има (няма) инверсия, то в новата между тях вече няма (има) инверсия. Всички останали инверсии са едни и същи и в двете пермутации. Следователно

$$[\ldots, a, b, \ldots] = [\ldots, b, a, \ldots] \pm 1$$

и в разглеждания частен случай твърдението е вярно.

Да разгледаме общия случай:

$$\ldots, a, b_1, \ldots, b_k, b, \ldots$$

$$\ldots, b, b_1, \ldots, b_k, a, \ldots$$

Преходът от първата пермутация към втората може да се реализира чрез последователна размяна на съседни елементи: ще разменим местата на b и b_k, след това на b и b_{k-1} и т.н., докато b застане пред a. Дотук пермутацията ще смени четността си $k+1$ пъти. След това ще разменим a с b_1, после a с b_2 и т.н., докато стигнем до втората пермутация. При това ще настъпят още k промени на четността, или общо $(k+1)+k = 2k+1$ промени. Очевидно описаната процедура довежда до втората пермутация, като настъпват нечетен брой смени на четността. Следователно двете пермутации са с различна четност.

§ 6. Основни свойства на детерминантите

1. За да станат детерминантите истински работен инструмент, е необходимо да изучим систематично техните свойства. Целият параграф е посветен именно на тази задача.

2. Теорема. Нека $A = (a_{ij})$ *е произволна квадратна матрица от ред* n. *Тогава*

$$|A| = \sum (-1)^{[\alpha_1, \alpha_2, \ldots, \alpha_n]} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n},$$

където сумирането е по всички пермутации $\alpha_1, \alpha_2, \ldots, \alpha_n$ *на числата* $1,\ 2, \ldots, n$.

Доказателство. Ще започнем с две съображения. Твърди се, че детерминантата $|A|$ е сума на членове от вида $\pm a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n}$. В това произведение от всеки ред участва точно по един елемент, а от всеки стълб - също точно по един елемент, защото $\alpha_1, \alpha_2, \ldots, \alpha_n$ е пермутация от номерата на стълбовете. Ако множителите $a_{i\alpha_i}$ са подредени по възходящ ред на първите си индекси, знакът $\pm$ се определя от четността на пермутацията на вторите индекси.

Нека сега си мислим, че е дадена някаква квадратна матрица B от ред n, в която за разлика от обичайното първият, вторият и т.н. редове имат някакви фиктивни номера $l_1, l_2, \ldots, l_n$, където $l_1 < l_2 < \ldots < l_n$. Нека аналогично стълбовете са номерирани с $m_1, m_2, \ldots, m_n$; $m_1 < m_2 < \ldots < m_n$, а $b_{l_i m_j}$ е елементът на матрицата B, лежащ в i -тия ред и j -тия стълб при обичайната номерация. Забележката от т. 4 на § 5 показва, че ако теоремата е вярна, то

$$|B| = \sum (-1)^{[\beta_1, \beta_2, \ldots, \beta_n]} b_{l_1\beta_1} b_{l_2\beta_2} \ldots b_{l_n\beta_n},$$

където сумирането е по всички пермутации $\beta_1, \beta_2, \ldots, \beta_n$ на числата $m_1, m_2, \ldots, m_n$.

Ще докажем теоремата с индукция по n. При $n=2$ пермутациите са само 1, 2 и 2, 1, а както знаем, тогава $|A| = a_{11}a_{22} - a_{12}a_{21}$. Нека теоремата е вярна за детерминанти от ред $n-1$ и да разгледаме произволна матрица $A = (a_{ij})$ от ред n. По определение

$$|A| = a_{11}\Delta_{11} - a_{21}\Delta_{21} + \ldots + (-1)^{i+1} a_{i1}\Delta_{i1} + \ldots + (-1)^{n+1} a_{n1}\Delta_{n1},$$

където

$$\Delta_{i1} = \begin{vmatrix} a_{12} & a_{13} & \dots & a_{1n} \\ .\ . & .\ . & .\ . & .\ . \\ a_{i-12} & a_{i-13} & \dots & a_{i-1n} \\ a_{i+12} & a_{i+13} & \dots & a_{i+1n} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n2} & a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

В последната детерминанта редовете имат фиктивни номера $1, \dots, i-1, i+1, \dots, n$, а стълбовете - $2, 3, \dots, n$. По индуктивното допускане

$$\Delta_{i1} = \sum (-1)^{[\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n]} a_{1\alpha_1} \dots a_{i-1\alpha_{i-1}} a_{i+1\alpha_{i+1}} \dots a_{n\alpha_n},$$

където сумирането е по всички пермутации $\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n$ на числата $2, 3, \dots, n$. Тъй като всички числа $\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n$ са по-големи от 1, то

$$[\alpha_1, \dots, \alpha_{i-1}, 1, \alpha_{i+1}, \dots, \alpha_n] = i - 1 + [\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n].$$

Следователно

$$(-1)^{i-1} a_{i1} \Delta_{i1} = \sum (-1)^{[\alpha_1, \dots, \alpha_{i-1}, 1, \alpha_{i+1}, \dots, \alpha_n]} a_{1\alpha_1} \dots a_{i-1\alpha_{i-1}} a_{i1} a_{i+1\alpha_{i+1}} \dots a_{n\alpha_n}.$$

Уговорката за последното сумиране може да се изкаже и така: сумира се по всички пермутации на числата $1, 2, \dots, n$, в които 1 е на i -то място. È òàêà,

$$|A| = \sum_{i=1}^{n} (-1)^{i-1} a_{i1} \Delta_{i1} =$$

$$= \sum_{i=1}^{n} \sum (-1)^{[\alpha_1, \dots, \alpha_{i-1}, 1, \alpha_{i+1}, \dots, \alpha_n]} a_{1\alpha_1} \dots a_{i-1\alpha_{i-1}} a_{i1} a_{i+1\alpha_{i+1}} \dots a_{n\alpha_n}.$$

Всички пермутации на числата $1, 2, \dots, n$ (и то точно по един път!) може да се изпишат по следното правило: първо се изписват тези, в които 1 е на първо място, след това онези, в които 1 е на второ място, и т.н. Следователно последната сума може да се напише и така:

$$|A| = \sum (-1)^{[\alpha_1, \alpha_2, \ldots, \alpha_n]} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n},$$

където сумирането е по всички пермутации $\alpha_1, \alpha_2, \ldots, \alpha_n$ на числата $1, 2, \ldots, n$. С това теоремата е доказана.

Доказаната формула се нарича *формула за развиване на детерминантата*. Много често тя се взема за определение на детерминанта от n-ти ред. Тук предпочетохме друго определение (§ 4, т. 5), което, струва ни се, е по-естествено. При $n=6$ сумата съдържа 6!=720 събираеми, а при $n=7$ броят им е 5040. Поради огромния брой операции практическата полза от тази формула е незначителна. При конкретни пресмятания много по-полезни се оказват свойствата на детерминантите, които ще изложим по-долу.

В някои съвсем специални случаи пресмятането на детерминанта от n-ти ред е лесно. Оставяме на читателя да съобрази, че ако в една детерминанта всички елементи под (или над) главния диагонал са нули (такива детерминанти се наричат *триъгълни*), то тя е равна на произведението на елементите по главния диагонал.

3. Теорема. *Ако в една детерминанта умножим елементите на един ред с някакво число, то детерминантата се умножава с това число.*

Доказателство. Нека $\Delta = \det(a_{ij})$. По теоремата от т. 2

$$\Delta = \sum (-1)^{[\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n]} a_{1\alpha_1} \ldots a_{i\alpha_i} \ldots a_{n\alpha_n}.$$

Ако умножим i -тия ред на Δ с числото λ, ще получим нова детерминанта

$$\Delta^* = \sum (-1)^{[\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n]} a_{1\alpha_1} \ldots (\lambda a_{i\alpha_i}) \ldots a_{n\alpha_n} =$$
$$= \lambda \sum (-1)^{[\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n]} a_{1\alpha_1} \ldots a_{i\alpha_i} \ldots a_{n\alpha_n} = \lambda\Delta.$$

4. Теорема. *Нека в детерминантата* $\Delta = \det(a_{ij})$ *елементите на* i *-тия ред имат вида* $a_{ij} = a'_{ij} + a''_{ij}$, $j = 1, 2, \ldots, n$. *Тогава* Δ *е равна на сумата от две детерминанти* Δ' *и* Δ'', *които се различават от* Δ *само* *по това, че елементите на* i *-тия ред в първата и във втората са съответно* a'_{ij} *и* a''_{ij} , $j = 1, 2, \ldots, n$, т.е.

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ .\ . & .\ . & .\ . & .\ . \\ a'_{i1}+a''_{i1} & a'_{i2}+a''_{i2} & \ldots & a'_{in}+a''_{in} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ .\ . & .\ . & .\ . & .\ . \\ a'_{i1} & a'_{i2} & \ldots & a'_{in} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ .\ . & .\ . & .\ . & .\ . \\ a''_{i1} & a''_{i2} & \ldots & a''_{in} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

Д о к а з а т е л с т в о . Действително

$$\begin{aligned} \Delta &= \sum (-1)^{[\alpha_1,\ldots,\alpha_i,\ldots,\alpha_n]} a_{1\alpha_1} \ldots (a'_{i\alpha_i} + a''_{i\alpha_i}) \ldots a_{n\alpha_n} = \\ &= \sum (-1)^{[\alpha_1,\ldots,\alpha_i,\ldots,\alpha_n]} a_{1\alpha_1} \ldots a'_{i\alpha_i} \ldots a_{n\alpha_n} + \\ &+ \sum (-1)^{[\alpha_1,\ldots,\alpha_i,\ldots,\alpha_n]} a_{1\alpha_1} \ldots a''_{i\alpha_i} \ldots a_{n\alpha_n} = \Delta' + \Delta''. \end{aligned}$$

5. Теорема. *Ако в една детерминанта разменим местата на два реда, то тя сменя знака си.*

Д о к а з а т е л с т в о . Ще проведем индукция по реда n на детерминантата. При $n = 2$ имаме

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

а след размяна на редовете получаваме

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

Нека твърдението е вярно за детерминанти от ред $n-1$ и

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

П ъ р в и с л у ч а й: разменят се два съседни реда. За удобство ще предположим, че са разменени първият и вторият ред (общият случай е аналогичен). Нека

$$\Delta^* = \begin{vmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{11} & a_{12} & \dots & a_{1n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

е новата детерминанта, а с Δ_{i1} и Δ^*_{i1} да означим детерминантите съответно в Δ и Δ^*, които се получават чрез зачертаване на i -тия ред и първия стълб. По определение

$$\Delta = a_{11}\Delta_{11} - a_{21}\Delta_{21} + a_{31}\Delta_{31} - \dots + (-1)^{n-1} a_{n1}\Delta_{n1},$$
$$\Delta^* = a_{21}\Delta^*_{11} - a_{11}\Delta^*_{21} + a_{31}\Delta^*_{31} - \dots + (-1)^{n-1} a_{n1}\Delta^*_{n1}.$$

Да сравним Δ_{i1} и Δ^*_{i1}. Очевидно

$$\Delta^*_{11} = \Delta_{21}, \qquad \Delta^*_{21} = \Delta_{11},$$

а при $i \ge 3$ детерминантата Δ^*_{i1} се получава от Δ_{i1} чрез смяна на местата на първия и втория ред. Следователно по индуктивното допускане

$$\Delta^*_{i1} = -\Delta_{i1}, \qquad i \ge 3.$$

Като заместим тези равенства в израза за Δ^*, получаваме

$$\Delta^* = a_{21}\Delta_{21} - a_{11}\Delta_{11} - a_{31}\Delta_{31} + \dots - (-1)^{n-1} a_{n1}\Delta_{n1} = -\Delta,$$

което доказва, че при разместване на два *съседни* реда детерминантата само сменя знака си.

Втори случай: разменяме местата на два реда, които не са съседни. Ако считаме, че това са редовете с номера k и l, $k<l$ и $l-k>1$, то тази операция може да се реализира чрез последователно разместване на съседни редове. Оставяме на читателя да съобрази, че за целта се извършват $(l-k)+(l-k-1)=2(l-k)-1$ такива размествания, т.е. детерминантата ще промени знака си нечетен брой пъти (вж. доказателството на лемата в т. 5, § 5).

6. Следствия. а) *Ако детерминантата съдържа два еднакви или пропорционални реда, то тя е равна на нула.*

б) *Ако елементите на един ред умножим с някакво число и ги прибавим към съответните елементи на друг ред, то детерминантата не се променя.*

Доказателство. а) Ако детерминантата Δ съдържа два еднакви реда, нека разменим местата им: ще получим $\Delta=-\Delta$, т.е. $\Delta=0$. Ако пък редовете са пропорционални с коефициент на пропорционалност λ, то след изнасянето му пред детерминантата ще получим $\Delta=\lambda\Delta^*$, където Δ^* е детерминанта с два еднакви реда, следователно равна на нула, откъдето и $\Delta=0$.

б) Нека $\Delta=\det\left(a_{ij}\right)$ е разглежданата детерминанта. Да речем, че елементите на i -тия ред сме умножили с λ и сме ги прибавили към съответните елементи на k -тия ред, $i\neq k$. Ще получим нова детерминанта Δ^*, в която елементите на k -тия ред имат вида $a_{kj}+\lambda a_{ij}$, $j=1,2,\ldots,n$, а останалите редове са същите като в Δ. Съгласно теоремите в т. 3 и т. 4

$$\Delta^*=\Delta+\lambda\Delta^{**},$$

където в Δ^{**} i -тият ред и k -тият ред са еднакви и понеже $i\neq k$, то $\Delta^{**}=0$, което завършва доказателството.

7. Ако A е матрица от тип $m\times n$, от нея можем да образуваме нова матрица (ще я бележим с A' или A^t), в която редовете с номера 1, 2, 3 и т.н. са стълбовете на A със същите номера:

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ .. & .. & .. & .. \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \qquad A^t=\begin{pmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ a_{12} & a_{22} & \ldots & a_{m2} \\ .. & .. & .. & .. \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{pmatrix}.$$

За матрицата A^t ще казваме, че е получена от A чрез *транспониране*, или че е *транспонирана* на A. Ясно е, че A^t е от тип $m \times n$ и $(A^t)^t = A$. Ако A е квадратна матрица, операцията транспониране е симетрия спрямо главния диагонал на A. В този случай детерминантата $\left|A^t\right|$ ще наричаме *транспонирана* на A.

8. Теорема за транспонираната детерминанта. *При транспониране детерминантата не променя стойността си.*

Д о к а з а т е л с т в о. Нека $\Delta = \det\left(a_{ij}\right)$ е разглежданата детерминанта, а $\Delta^t = \det\left(a_{ji}\right)$ е транспонираната й; първият и вторият индекс на a_{ji} показват съответно номера на стълба и номера на реда на Δ^t, в които се намира a_{ji}. Като приложим теоремата от т. 2, ще получим

$$\Delta^t = \sum \left(-1\right)^{[\alpha_1, \alpha_2, \ldots, \alpha_n]} a_{\alpha_1 1} a_{\alpha_2 2} \ldots a_{\alpha_n n}.$$

Нека в произведението

(*) $$a_{\alpha_1 1} a_{\alpha_2 2} \ldots a_{\alpha_n n}$$

разместим множителите така, че първите индекси да бъдат във възходящ ред, т.е. да го запишем във вида

(**) $$a_{1\beta_1} a_{2\beta_2} \ldots a_{n\beta_n}.$$

Знакът, с който произведението (*) участва в развитието на Δ^t, се определя от четността на пермутацията $\alpha_1, \ldots, \alpha_n$, докато знакът, с който (**) участва в развитието на детерминантата Δ, се определя от четността на пермутацията $\beta_1, \ldots, \beta_n$. Ако покажем, че двете пермутации имат еднаква четност, ще следва, че всеки член от развитието на Δ^t е член от развитието на Δ, а тъй като броят на събираемите във всяко от тези развития е $n!$, ще получим равенството $\Delta^t = \Delta$.

За да достигнем до произведението (**), най-напред в (*) ще разменим местата само на два множителя, след това на още два и т.н. По-конкретно, нека в (*) сме разменили местата на два множителя и сме получили произведението

(***) $$a_{\gamma_1 \delta_1} a_{\gamma_2 \delta_2} \ldots a_{\gamma_n \delta_n}.$$

При разместване само на два множителя в пермутацията от първите индекси точно два елемента сменят местата си; същото важи и за пермутацията на вторите индекси. Съгласно лемата от § 5, т. 5 пермутациите $\alpha_1, \ldots, \alpha_n$ и $\gamma_1, \ldots, \gamma_n$ са с различна четност. По същата причина с различна четност са и пермутациите $1, \ldots, n$ и $\delta_1, \ldots, \delta_n$. Следователно числата

$$\left[\alpha_1, \ldots, \alpha_n\right] + \left[1, \ldots, n\right] = \left[\alpha_1, \ldots, \alpha_n\right]$$

и

$$\left[\gamma_1, \ldots, \gamma_n\right] + \left[\delta_1, \ldots, \delta_n\right]$$

са с еднаква четност. Като продължим тази процедура, ще получим найсетне, че същата четност има и числото

$$\left[1, \ldots, n\right] + \left[\beta_1, \ldots, \beta_n\right] = \left[\beta_1, \ldots, \beta_n\right],$$

което довършва доказателството.

9. Следствие. *Твърденията на теоремите в т. 3 - 6 са верни и за стълбовете на детерминантата.*

За да се убедим в това е достатъчно да приложим теоремата за транспонираната детерминанта.

§ 7. Адюнгирани количества и поддетерминанти

1. В определението на детерминанта от ред n (вж. § 4, т. 5) първият стълб на матрицата $A = \left(a_{ij}\right)$ играе специална роля. От друга страна, вече знаем, че ако преместим който и да е стълб на първо място, детерминантата най-много ще смени знака си. Ако сега приложим определението, ще получим "развитие на детерминантата по елементите на въпросния стълб".

Преди да уточним тези съображения, ще въведем едно понятие.

2. Определение. Нека $A = \left(a_{ij}\right)$ е квадратна матрица от ред n, в която зачертаваме i -тия ред и j -тия стълб. Детерминантата Δ_{ij} на получената матрица от ред $n-1$ ще наричаме *поддетерминанта* на елемента a_{ij}.

Сега в матрицата A да изберем произволен стълб, например j -тия, и да разменим мястото му последователно с $(j-1)$-вия, $(j-2)$ -рия и т.н., докато застане на първо място. Съгласно теорема 5 и следствие 9 от § 6 имаме

$$\left|A\right| = (-1)^{j-1} \begin{vmatrix} a_{1j} & a_{11} & \dots & a_{1j-1} & a_{1j+1} & \dots & a_{1n} \\ a_{2j} & a_{21} & \dots & a_{2j-1} & a_{2j+1} & \dots & a_{2n} \\ .. & .. & .. & .. & .. & .. & .. \\ a_{nj} & a_{n1} & \dots & a_{nj-1} & a_{nj+1} & \dots & a_{nn} \end{vmatrix}.$$

За детерминантата от дясната страна ще приложим определението. Да забележим обаче, че ако в нея зачертаем първия стълб и i -тия ред, резултатът ще бъде същият, както, ако в $\left|A\right|$ зачертаем j -тия стълб и i -тия ред, т.е. ще стигнем до поддетерминантата Δ_{ij}. Следователно

$$\left|A\right| = (-1)^{j-1}\left[a_{1j}\Delta_{1j} - a_{2j}\Delta_{2j} + \dots + (-1)^{n+1} a_{nj}\Delta_{nj}\right] =$$
$$= (-1)^{1+j} a_{1j}\Delta_{1j} + (-1)^{2+j} a_{2j}\Delta_{2j} + \dots + (-1)^{i+j} a_{ij}\Delta_{ij} + \dots + (-1)^{n+j} a_{nj}\Delta_{nj}.$$

Да означим

(1) $$A_{ij} = (-1)^{i+j} a_{ij}\Delta_{ij};$$

този израз се нарича *адюнгирано количество* на елемента a_{ij}.

Полученият резултат може да се запише във вида

(2) $$\left|A\right| = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj};$$

обикновено се казва, че това е развитие на детерминантата по адюнгираните количества на елементите от j -тия стълб. Всъщност при $j=1$ развитието (2) поставихме в основата на индуктивното определение на детерминанта.

Ще подчертаем, че адюнгираното количество A_{ij} не зависи от елементите на i -тия ред и j -тия стълб на матрицата A : при образуването на Δ_{ij} те се зачертават. В тази връзка нека $k \neq j$ и в матрицата A да подменим k -тия стълб с j -тия, а всички останали да оставим без промяна. Ще получим матрицата B с два еднакви стълба. Нейната детерминанта е равна на нула. От друга страна, да я развием по k -тия стълб и да вземем предвид, че адюнгираните количества на елементите от k -тия стълб на $\left|B\right|$ са същите както в $\left|A\right|$. Следователно

(3) $$0 = a_{1j}A_{1k} + a_{2j}A_{2k} + \ldots + a_{nj}A_{nk}, \qquad k \neq j .$$

Да резюмираме получените резултати.

3. Теорема. *Сумата от произведенията на елементите на даден стълб на детерминантата с адюнгираните им количества е равна на детерминантата. Сумата от произведенията на елементите на някой стълб със съответните адюнгирани количества на елементите от друг стълб е равна на нула, т.е.*

$$a_{1j}A_{1k} + a_{2j}A_{2k} + \ldots + a_{nj}A_{nk} = \delta_{jk}\left|A\right|,$$

където δ_{jk} е символът на Кронекер: $\delta_{jk} = 1$ при $j = k$ и $\delta_{jk} = 0$ при $j \neq k$.

Ще покажем, че аналогичен резултат е в сила и за редовете на детерминантата. За целта първо ще дадем друга интерпретация на адюнгираното количество A_{ij}. Запазваме означенията от т. 2 и фиксираме индексите i и j . В матрицата A заменяме $a_{1j}, a_{2j}, \ldots, a_{i-1j}, a_{i+1j}, \ldots, a_{nj}$ с нули. Ще получим нова матрица, която означаваме с B . Ще пресметнем $\left|B\right|$ по два начина. От една страна, тъй като в детерминантата $\left|A\right|$ адюнгираното количество A_{ij} не зависи от елементите на i -тия ред и j -тия стълб, от формула (2) получаваме, че $\left|B\right| = a_{ij}A_{ij}$. От друга страна, да си спомним, че

(4) $$\left|A\right| = \sum(-1)^{[\alpha_1, \alpha_2, \ldots, \alpha_n]} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n},$$

където сумирането е по всички пермутации на числата $1, 2, \ldots, n$. Ако в дясната страна положим $a_{1j} = \ldots = a_{i-1j} = a_{i+1j} = \ldots = a_{nj} = 0$, ще останат само събираемите, които съдържат като множител a_{ij}; тяхната сума е равна на $\left|B\right|$. Следователно за да намерим адюнгираното количество A_{ij}, е достатъчно в развитието на детерминантата като сума от $n!$ събираеми да отделим онези от тях, които съдържат множителя a_{ij}, и да го изнесем пред скоби: сумата в скобите ще бъде A_{ij}.

Като имаме предвид тази интерпретация и забележката, че събираемото $\pm\, a_{1\alpha_1} \ldots a_{n\alpha_n}$ от дясната страна на равенство (4) съдържа точно по един множител от всеки ред и от всеки стълб на детерминантата, да фиксираме i и да представим сумата в (4) като сбор на n "подсуми": в първата отделяме събираемите, съдържащи a_{i1} (това е $a_{i1}A_{i1}$), във втората отделяме събираемите, съдържащи a_{i2} (това е $a_{i2}A_{i2}$), и т.н. Следователно

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}.$$

4. Теорема. *Нека* $A = (a_{ij})$ *е квадратна матрица от ред* n. *Тогава*

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \ldots + a_{in}A_{kn} = \delta_{ik}|A|.$$

При $i = k$ твърдението току-що беше доказано. При $i \neq k$ доказателството е аналогично на разсъждението от края на т. 2. (Разгледайте детерминанта, в която i -тият и k -тият ред са еднакви.)

5. Формули на Крамер. Вече сме готови да дадем едно от първите приложения на развития дотук апарат на детерминантите. Да разгледаме произволна линейна система от n уравнения с n неизвестни:

$$\begin{aligned}
a_{11}x_1 + \ldots + a_{1k}x_k + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \ldots + a_{2k}x_k + \ldots + a_{2n}x_n &= b_2 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{n1}x_1 + \ldots + a_{nk}x_k + \ldots + a_{nn}x_n &= b_n.
\end{aligned}$$

Нека $\Delta = \det(a_{ij})$ е детерминантата на матрицата от коефициентите пред неизвестните и да допуснем, че $x_1, \ldots, x_n$ са числа, които удовлетворяват системата. Да умножим първото, второто и т.н. уравнения, съответно с адюнгираните количества A_{1k}, A_{2k} и т.н., и да съберем получените равенства. Съгласно теоремата от т. 3 ще стигнем до равенството

$$0.x_1 + \ldots + \Delta x_k + \ldots + 0.x_n = b_1A_{1k} + b_2A_{2k} + \ldots + b_nA_{nk}.$$

Пак по теоремата от т. 3 дясната страна може да се запише като детерминанта

$$\Delta_k = \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & \dots & b_2 & \dots & a_{2n} \\ . \; . & . \; . & . \; . & . \; . & . \; . \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix};$$

тя е образувана от Δ чрез подмяна на k -тия стълб със стълба от свободните членове. Следователно

$$\Delta \,.\, x_k = \Delta_k .$$

Тъй като k е произволно ($k = 1, 2, \dots, n$), получаваме формулите

(5) $$\Delta . x_1 = \Delta_1, \; \dots, \; \Delta . x_k = \Delta_k, \; \dots, \; \Delta . x_n = \Delta_n,$$

които обикновено се наричат *формули на Крамер*. Ако $\Delta \neq 0$, те може да се препишат във вида

(6) $$x_k = \frac{\Delta_k}{\Delta}, \quad k = 1, 2, \dots, n .$$

Очевидно е, че при $n = 2$ формулите (6) съвпадат с изведените в §4, т. 3.

Дотук получихме, че ако разглежданата система има решение, то трябва да удовлетворява уравненията (5). Бихме могли да разглеждаме (5) като линейна система за $x_1, \dots, x_n$, но трябва да подчертаем, че в общия случай тя не е еквивалентна на изходната. Ясно е например, че решенията на системата

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

са $(\lambda, -\lambda, 0, 0)$, където λ е произволно число. От друга страна, $\Delta = \Delta_1 = \dots = \Delta_4 = 0$, така че (5) сега има вида $0.x_1 = 0$, ..., $0.x_4 = 0$ и се удовлетворяват за всички стойности на неизвестните. Този пример показва, че има случаи, когато формулите (5) са безполезни. Толкова по-интересна е следната

6. Теорема. *Ако детерминантата Δ на линейна система от n уравнения с n неизвестни е различна от нула, то системата е* *съвместима и има единствено решение, което се дава с формулите на Крамер* (6).

Доказателство. Използваме означенията от т.5. Вече показахме, че ако системата е съвместима и $\Delta\neq 0$, то решението е единствено, защото трябва да удовлетворява (6). Нека сега $\Delta\neq 0$ и да покажем, че формулите на Крамер (6) наистина дават решение на системата. Преди да заместим в i -тото уравнение $\left(i=1,2,\ldots,n\right)$, нека напомним, че $\Delta_k=\sum_{l=1}^{n} b_l A_{lk}$, където A_{lk} е съответното адюнгирано количество в детерминантата $\Delta=\det\left(a_{lk}\right)$. Като използваме и теоремата от т. 4, получаваме

$$a_{i1}\frac{\Delta_1}{\Delta}+a_{i2}\frac{\Delta_2}{\Delta}+\ldots+a_{in}\frac{\Delta_n}{\Delta}=\sum_{k=1}^{n} a_{ik}\frac{\Delta_k}{\Delta}=\sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{n}\frac{1}{\Delta}b_l A_{lk}\right)=$$

$$=\frac{1}{\Delta}\sum_{l=1}^{n} b_l\left(\sum_{k=1}^{n} a_{ik}A_{lk}\right)=\frac{1}{\Delta}\sum_{l=1}^{n} b_l\left(\delta_{il}\Delta\right)=b_i,$$

което завършва доказателството.

От изчислителна гледна точка формулите на Крамер са твърде разточителни: пресмятането на $n+1$ детерминанти от ред n в общия случай е свързано с огромен брой събирания и умножения. В това отношение методът на Гаус е съществено по-икономичен. В контекста на въпросите, повдигнати в т. 1 на § 4, формулите на Крамер имат важна теоретична стойност, защото е вярна и обратната на теоремата от т. 6.

7. Теорема. *Ако линейна система от n уравнения с n неизвестни е съвместима и има единствено решение, то детерминантата Δ на системата е различна от нула и решението се дава от формулите на Крамер .*

Доказателство. Запазваме означенията от т. 5 и с помощта на метода на Гаус привеждаме системата в стъпаловидна форма. Съгласно теоремата от § 3, т. 3, д, стъпаловидната форма има вида

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n &= b_1\\ \overline{a}_{22}x_2+\ldots+\overline{a}_{2n}x_n &= \overline{b}_2\\ \ldots\ldots\ldots\ldots\ldots\ldots &\\ \overline{a}_{nn}x_n &= \overline{b}_n \end{aligned}\tag{7}$$

където $\Delta' = a_{11}\overline{a}_{22} \ldots \overline{a}_{nn} \neq 0$. Произведението Δ' очевидно е детерминантата на системата (7). Прилагането на метода на Гаус към изходната (дадената) система е свързано само с два вида процедури: разместване на уравнения и умножаване на уравнение с число и прибавянето му към друго уравнение. Първата от тези процедури води до разместване на редове в детерминантата $\Delta = \det\left(a_{ij}\right)$, което евентуално само сменя знака на Δ. Втората по вид процедура води до умножаване на някой ред на Δ с число и прибавянето му към друг ред. Тази операция не променя детерминантата. Следователно за детерминантата $\Delta = \det\left(a_{ij}\right)$ на дадената система и детерминантата Δ' на стъпаловидната форма имаме $\Delta = \pm\Delta'$ при подходящ избор на знака. Тъй като $\Delta' \neq 0$, то и $\Delta \neq 0$.

Взети заедно, теоремите от т. 6 и т. 7 означават, че *необходимо и достатъчно условие линейна система от n уравнения с n неизвестни да бъде определена е детерминантата ѝ да бъде различна от нула.*

§8. Детерминанта с ъгъл от нули. Умножаване на детерминанти

Пресмятането на конкретни детерминанти често пъти може да се опрости с помощта на следната

1. Теорема. Ако една детерминанта от ред $n+m \geq 2$ има вида

$$\Delta_1 = \begin{vmatrix} a_{11} & \ldots & a_{1n} & c_{11} & \ldots & c_{1m} \\ \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot \\ a_{n1} & \ldots & a_{nn} & c_{n1} & \ldots & c_{nm} \\ 0 & \ldots & 0 & b_{11} & \ldots & b_{1m} \\ \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot & \cdot\,\cdot \\ 0 & \ldots & 0 & b_{m1} & \ldots & b_{mm} \end{vmatrix},$$

или

$$\Delta_2 = \begin{vmatrix} a_{11} & \dots & a_{1n} & 0 & \dots & 0 \\ . . & . . & . . & . . & . . & . . \\ a_{n1} & \dots & a_{nn} & 0 & \dots & 0 \\ c_{11} & \dots & c_{1n} & b_{11} & \dots & b_{1m} \\ . . & . . & . . & . . & . . & . . \\ c_{m1} & \dots & c_{mn} & b_{m1} & \dots & b_{mm} \end{vmatrix},$$

то тя е равна на произведението

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ . . & . . & . . \\ a_{n1} & \dots & a_{nn} \end{vmatrix} . \begin{vmatrix} b_{11} & \dots & b_{1m} \\ . . & . . & . . \\ b_{m1} & \dots & b_{mm} \end{vmatrix}.$$

Д о к а з а т е л с т в о. За детерминанти от посочения вид ще казваме, че са *с ъгъл от нули*. Ще докажем теоремата с помощта на индукция по числото $n+m$, и то само за случая, когато нулите са в долния ляв ъгъл. Другият случай ще следва от този и от теоремата за транспонираната детерминанта (вж. §6, т.8).

При $n+m=2$, т.е. $n=m=1$, имаме

$$\Delta_1 = \begin{vmatrix} a_{11} & c_{11} \\ 0 & b_{11} \end{vmatrix} = a_{11}.b_{11}$$

и твърдението е вярно. Нека то е вярно за детерминанти от ред $n+m-1$, имащи разглеждания вид и $|A| = \det\left(a_{ij}\right)$, $|B| = \det\left(b_{kl}\right)$. Ще покажем, че е вярно и за детерминантата Δ_1 от ред $n+m$. При $n=1$ от определението на детерминанта имаме $\Delta_1 = a_{11}.|B| = |A|.|B|$. Нека $n>1$. Ако сега в Δ_1 зачертаем i -тия ред и първия стълб, $1 \le i \le n$, съответната поддетерминанта Δ_{i1} е с ъгъл от нули, има ред $n+m-1$ и за нея е приложимо индуктивното допускане. Следователно

$$\Delta_{i1} = \Delta^*_{i1}.|B|,$$

където Δ^*_{i1} е поддетерминантата в $|A|$, получена чрез зачертаване на i -тия ред и първия стълб. По определение

$$\left|A\right| = a_{11}\Delta_{11}^{*} - a_{21}\Delta_{21}^{*} + \ldots + \left(-1\right)^{n+1}\Delta_{n1}^{*}.$$

Пак по определение на детерминанта и предвид на горните равенства получаваме

$$\Delta_1 = a_{11}\Delta_{11} - a_{21}\Delta_{21} + \ldots + \left(-1\right)^{n+1}\Delta_{n1} =$$
$$= a_{11}\Delta_{11}^{*}\left|B\right| - a_{21}\Delta_{21}^{*}\left|B\right| + \ldots + \left(-1\right)^{n+1}\Delta_{n1}^{*}\left|B\right| = \left|A\right|.\left|B\right|,$$

което завършва доказателството.

2. Теорема. *Произведението*

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ .\ . & .\ . & .\ . \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} . \begin{vmatrix} b_{11} & \ldots & b_{1n} \\ .\ . & .\ . & .\ . \\ b_{n1} & \ldots & b_{nn} \end{vmatrix}$$

е равно на детерминантата

$$\begin{vmatrix} c_{11} & \ldots & c_{1n} \\ .\ . & .\ . & .\ . \\ c_{n1} & \ldots & c_{nn} \end{vmatrix},$$

където

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Д о к а з а т е л с т в о . Да разгледаме помощната детерминанта

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & 0 & 0 & \ldots & 0 \\ .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . \\ a_{n1} & a_{n2} & \ldots & a_{nn} & 0 & 0 & \ldots & 0 \\ -1 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . \\ 0 & 0 & \ldots & -1 & b_{n1} & b_{n2} & \ldots & b_{nn} \end{vmatrix}.$$

Ще я пресметнем по два начина. От една страна, по теоремата в т. 1 имаме

$$\Delta = \det\left(a_{ij}\right).\det\left(b_{kl}\right).$$

От друга страна, умножаваме редовете с номера $n+1,\ n+2,\ldots,2n$ с подходящи числа и ги прибавяме последователно към първите n реда, като се стремим на местата на елементите a_{ij} да се получат нули. Например, ако умножим $(n+1)$-вия ред до $2n$-тия ред съответно с $a_{11},\ a_{12},\ldots,a_{1n}$ и ги прибавим към първия ред, на местата на $a_{11}, a_{12},\ldots,a_{1n}$ ще застанат нули. При това на мястото на нулите в първия ред на Δ ще се появят елементи $c_{11}, c_{12},\ldots,c_{1n}$. Аналогично постъпваме с втория и т.н. до n-тия ред на Δ. Така

$$\Delta = \begin{vmatrix} 0 & 0 & \ldots & 0 & c_{11} & c_{12} & \ldots & c_{1n} \\ .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . \\ 0 & 0 & \ldots & 0 & c_{n1} & c_{n2} & \ldots & c_{nn} \\ -1 & 0 & \ldots & 0 & b_{11} & b_{12} & \ldots & b_{1n} \\ .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . & .\ . \\ 0 & 0 & \ldots & -1 & b_{n1} & b_{n2} & \ldots & b_{nn} \end{vmatrix}.$$

Не е трудно да се съобрази, че

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

Сега нека преместим стълба с номер $n+1$ на първо място, след това стълба с номер $n+2$ - на второ място и т.н. Тъй като при всяко такова преместване детерминантата сменя знака си n пъти (вж. § 6, т. 5 и т. 9), то

$$\Delta=\left(-1\right)^{n^2}.\begin{vmatrix} c_{11} & c_{12} & \dots & c_{1n} & 0 & 0 & \dots & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. \\ c_{n1} & c_{n2} & \dots & c_{nn} & 0 & 0 & \dots & 0 \\ b_{11} & b_{12} & \dots & b_{1n} & -1 & 0 & \dots & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. \\ b_{n1} & b_{n2} & \dots & b_{nn} & 0 & 0 & \dots & -1 \end{vmatrix}.$$

Като приложим теоремата от т. 1, получаваме $\Delta=\left(-1\right)^{n^2}.\det\left(c_{ij}\right).\left(-1\right)^{n}= \ =\left(-1\right)^{n(n+1)}.\det\left(c_{ij}\right).$ Числото $n\left(n+1\right)$ е четно, следователно $\Delta=\det\left(c_{ij}\right)$, а в началото видяхме, че $\Delta=\det\left(a_{ij}\right).\det\left(b_{kl}\right)$. Теоремата е доказана.

Смисълът на горната теорема е, че произведението на две детерминанти от ред n може да се представи като детерминанта от ред n. Правилото, по което са образувани елементите c_{ij}, се помни лесно: умножаваме почленно i -тия ред на първата детерминанта с j -тия стълб на втората (умножаване по правилото "ред по стълб"). Съществуват обаче още три правила. Достатъчно е да си спомним, че при транспониране детерминантата не се променя и да транспонираме която и да е от детерминантите $\det\left(a_{ij}\right)$ и $\det\left(b_{kl}\right)$. Да положим

$$c'_{ij}=\sum_{s=1}^{n}a_{si}b_{sj}\,,\quad c''_{ij}=\sum_{s=1}^{n}a_{is}b_{js}\,,\quad c'''_{ij}=\sum_{s=1}^{n}a_{si}b_{js}\,.$$

Числата c'_{ij}, c''_{ij}, c'''_{ij} са получени чрез умножаване съответно по правилата "стълб по стълб", "ред по ред", "стълб по ред". Ако във всяко от произведенията $\det\left(a_{ij}\right)^t\det\left(b_{kl}\right)$, $\det\left(a_{ij}\right)\det\left(b_{kl}\right)^t$, $\det\left(a_{ij}\right)^t\det\left(b_{kl}\right)^t$ извършим умножението по доказаното правило "ред по стълб", ще получим съответно c'_{ij}, c''_{ij}, c'''_{ij}. Следователно

$$\det\left(a_{ij}\right)\det\left(b_{kl}\right)=\det\left(c_{ij}\right)=\det\left(c'_{ij}\right)=\det\left(c''_{ij}\right)=\det\left(c'''_{ij}\right).$$

3. Пример. Нека $\Delta=\det\left(a_{ij}\right)$ е детерминанта от ред n, а A_{ij} е адюнгираното количество на a_{ij}. Детерминантата

$$\Delta^* = \begin{vmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ . \; . & . \; . & . \; . & . \; . \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{vmatrix}$$

се нарича *адюнгирана* на детерминантата Δ. Ще изразим Δ^* чрез Δ. Като си спомним теоремата от § 7, т. 4, да пресметнем $\Delta^* \Delta$ по правилото "ред по ред". Ще получим

$$\Delta \,.\, \Delta^* = \begin{vmatrix} \Delta & 0 & \dots & 0 \\ 0 & \Delta & \dots & 0 \\ . \; . & . \; . & . \; . & . \; . \\ 0 & 0 & \dots & \Delta \end{vmatrix} = \Delta^n.$$

Следователно, ако $\Delta \neq 0$, то $\Delta^* = \Delta^{n-1}$. Може да се покаже, че ако $\Delta = 0$, то и $\Delta^* = 0$ (вж. зад.5 от упражненията към тази глава). Така във всеки случай $\Delta^* = \Delta^{n-1}$.

4. Исторически бележки. Още в края на XVII в. Лайбниц (1646-1716) забелязва общо правило за изразяване на решенията на система линейни уравнения, което води до определението на детерминантите. Донякъде в емпирична форма понятието детерминанта се среща за пръв път в работите на Крамер (1704-1752) и Безу (1730-1783). Точното определение на това понятие и разкриването на неговите основни свойства дължим на Коши (1789-1857) и Якоби (1804-1851).

УПРАЖНЕНИЯ

1. Нека в хомогенната система

$$b_{11}y_1 + b_{12}y_2 + b_{13}y_3 = 0$$
$$b_{21}y_1 + b_{22}y_2 + b_{23}y_3 = 0$$

поне една от детерминантите

$$M_1 = \begin{vmatrix} b_{12} & b_{13} \\ b_{22} & b_{23} \end{vmatrix}, \qquad M_2 = \begin{vmatrix} b_{11} & b_{13} \\ b_{21} & b_{23} \end{vmatrix}, \qquad M_3 = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}$$

е различна от нула. Докажете, че всички решения на системата са

$$y_1 = M_1 t, \qquad y_2 = -M_2 t, \qquad y_3 = M_3 t,$$

където t е произволно число.

(У п ъ т в а н е . Ако например $M_3 \neq 0$, положете $y_3 = \lambda$ и определете по формулите на Крамер y_1 и y_2. Положете $\lambda = M_3 t$.)

2. Дадена е хомогенна система от $n-1$ уравнения с n неизвестни. Нека $A = (a_{ki})$ е матрицата от коефициентите пред неизвестните, $k = 1,2,\dots,n-1$; $i = 1,2,\dots,n$. Да означим с M_i детерминантата на матрицата, получена от A чрез премахване на i-тия стълб на A, $i = 1,2,\dots,n$. Докажете, че ако поне една от детерминантите $M_1, M_2, \dots, M_n$ е различна от нула, то всички решения на системата са

$$x_1 = M_1 t, \ \ x_2 = -M_2 t, \ \ \dots, \ \ x_i = (-1)^{i+1} M_i t, \dots, \ x_n = (-1)^{n+1} M_n t,$$

където t е произволно число.

3. (Детерминанта на Вандермонд) Докажете, че

$$V_n = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ .. & .. & .. & .. \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \le j < i \le n} (x_i - x_j).$$

(У п ъ т в а н е . Покажете първо, че $V_n = (x_n - x_1)(x_n - x_2) \dots (x_n - x_{n-1}) V_{n-1}$. За целта във V_n умножете $(n-1)$-вия ред с x_n и го извадете от n-тия, след това умножете $(n-2)$-рия ред с x_n и го извадете от $(n-1)$-вия и т.н.)

4. Нека функциите $f_{ij}(x)$, $i, j = 1, 2, \dots, n,$ са диференцируеми при $x = x_0$. Докажете, че функцията $\Delta(x) = \det\left(f_{ij}(x)\right)$ също е диференцируема при $x = x_0$ и

$$\Delta'(x_0)=\sum_{i=1}^{n}\begin{vmatrix} f_{11}(x_0) & f_{12}(x_0) & \dots & f_{1n}(x_0) \\ .. & .. & .. & .. \\ f'_{i1}(x_0) & f'_{i2}(x_0) & \dots & f'_{in}(x_0) \\ .. & .. & .. & .. \\ f_{n1}(x_0) & f_{n2}(x_0) & \dots & f_{nn}(x_0) \end{vmatrix}.$$

5. Нека $\Delta=\det\left(a_{ij}\right)$ е детерминанта от n -ти ред, а $\Delta^*=\det\left(A_{ij}\right)$ е адюнгираната й детерминанта. Докажете, че $\Delta^*=0$ тогава и само тогава, когато $\Delta=0$. (Упътване. Импликацията $\Delta^*=0 \Rightarrow \Delta=0$ следва от доказаното в § 8, т. 3. Обратно, нека $\Delta=0$ и допуснете, че $\Delta^*\neq 0$. Тъй като $\Delta=0$, от теоремата в § 7, т. 4 следва, че за всяко $i=1,2,\dots,n$ са в сила равенствата

$$\begin{aligned} &A_{11}a_{i1}+A_{12}a_{i2}+\dots+A_{1n}a_{in}=0\\ &A_{21}a_{i1}+A_{22}a_{i2}+\dots+A_{2n}a_{in}=0\\ &.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\\ &A_{n1}a_{i1}+A_{n2}a_{i2}+\dots+A_{nn}a_{in}=0. \end{aligned}$$

От тези равенства и допускането, че $\Delta^*\neq 0$, получете, че $a_{i1}=a_{i2}=\dots=a_{in}=0$ за всяко $i=1,2,\dots,n$. Следователно всички елементи a_{ij} в детерминантата Δ са нули, а оттук и всички елементи A_{ij} в Δ^* са нули, т.е. $\Delta^*=0$, противно на допускането.)

Г л а в а 2

КООРДИНАТНИ ЛИНЕЙНИ ПРОСТРАНСТВА

§ 1. Определение и основни свойства

1. Мотивировка. Апаратът на детерминантите не е достатъчен, за да се изследва докрай произволна система от линейни уравнения. Необходимо е да се привлекат и други идеи. Ще постъпим като шахматните анализатори: когато атаката заглъхне, връщат фигурите назад и започват анализа отново. Методът на Гаус ни даде възможност да заменим дадената система с еквивалентна, която се изследва лесно. При процедурата по тази замяна (привеждане в стъпаловидна форма) многократно повтаряхме операция от един и същ тип: умножаване на някое уравнение с число и прибавяне към друго уравнение. Тук ще моделираме тази операция в абстрактна форма, ще я изучим подробно и като резултат ще получим отговор на редица въпроси за системата, напр. кои вътрешни характеристики на системата определят колко от уравненията са следствия от останалите и следователно може да се изпуснат.

Уславяме се, че ако в множество от n елемента е посочено кой елемент е първи, кой е втори и т.н., ще говорим за *наредена* n*-торка*. За явно посочване на конкретни наредени n-торки ще пишем $\left(a_1, a_2, ..., a_n\right)$ (a_i са елементите на n-торката). Например i -тото уравнение в системата може да се интерпретира като наредена $\left(n+1\right)$-торка $\left(a_{i1}, ..., a_{in}, b_i\right)$ от числа. Това именно подсказва следното

2. Определение. Нека n е фиксирано естествено число и

$$\mathrm{R}^n = \left\{\left(a_1, a_2, ..., a_n\right) \setminus a_i \in \mathrm{R}\right\}$$

е множеството на наредените n-торки от реални числа. Ще казваме, че n-торките $a=\left(a_1, ..., a_n\right)$ и $b=\left(b_1, ..., b_n\right)$ са *равни*, ако $a_1=b_1$, $a_2=b_2, ..., a_n=b_n$. Под *сума* на две произволни n-торки $a=\left(a_1, ..., a_n\right)$ и $b=\left(b_1, ..., b_n\right)$ ще разбираме n-торката

$$a+b=\left(a_1+b_1, a_2+b_2, ..., a_n+b_n\right) \in \mathrm{R}^n .$$

Ако $\lambda \in \mathrm{R}$, под произведение на n-торката a с числото λ ще разбираме n-торката

$$\lambda a = \left(\lambda a_1, \lambda a_2, ..., \lambda a_n\right) \in \mathrm{R}^n.$$

Множеството R^n заедно с така дефинираното събиране на елементи от R^n и умножаването на елементи от R^n с реални числа ще наричаме (реално) *координатно линейно пространство* с размерност n над R. Елементите на R^n ще наричаме n*-мерни вектори* (или просто *вектори*), а реалните числа, с които умножаваме векторите - *скалари*.

Уславяме се с 0 да бележим вектора $\left(0, 0, ..., 0\right)$ и да го наричаме *нулев вектор*. Ако $a = \left(a_1, ..., a_n\right)$ е някакъв вектор, с $-a$ ще бележим вектора $\left(-a_1, ..., -a_n\right)$ и ще го наричаме *противоположен* на a. Очевидно $-a = \left(-1\right)a$. Ако a, b, c са произволни вектори от R^n, а λ, μ са произволни реални числа, от определението веднага получаваме следните свойства:

1) $a + b = b + a$ (комутативност);
2) $\left(a+b\right)+c = a+\left(b+c\right)$ (асоциативност);
3) $a + 0 = a$;
4) $a + \left(-a\right) = 0$;
5) $1.a = a$;
6) $\lambda\left(a+b\right) = \lambda a + \lambda b$;
7) $\left(\lambda+\mu\right)a = \lambda a + \mu a$;
8) $\left(\lambda\mu\right)a = \lambda\left(\mu a\right)$.

Тривиалната проверка оставяме на читателя. Опитът, който ще натрупаме в рамките на тази глава, ще ни убеди, че пресмятанията в R^n може да се извършат дори ако се абстрахираме от конкретната природа на разглежданите вектори и се ръководим само от "правилата" 1-8. Този факт ще ни даде повод за по-нататъшни обобщения.

Ако изхождаме от множеството C на комплексните числа, можем да въведем аналогично и комплексното координатно линейно пространство C^n с размерност n над C: необходимо е да се разглеждат наредените n-торки от комплексни числа, а в качеството на скалари да се използват комплексните числа. За по-голяма конкретност тук ще формулираме резултатите за R^n, макар че те ще бъдат верни и за C^n.

Ако $l_1, l_2, ..., l_k$ са вектори, а $\lambda_1, \lambda_2, ..., \lambda_k$ са числа за вектора

$$a = \lambda_1 l_1 + \lambda_2 l_2 + ... + \lambda_k l_k$$

ще казваме, че е *линейна комбинация* на векторите $l_1, l_2, ..., l_k$ с коефициенти $\lambda_1, \lambda_2, ..., \lambda_k$.

3. Определение. Ще казваме, че векторите $l_1, l_2, ..., l_k$ са *линейно зависими*, ако съществуват такива числа $\lambda_1, \lambda_2, ..., \lambda_k$, които не са едновременно нули и за които

$$\lambda_1 l_1 + \lambda_2 l_2 + ... + \lambda_k l_k = 0.$$

Ако такова равенство е възможно само при $\lambda_1 = \lambda_2 = ... = \lambda_k = 0$, ще казваме, че векторите $l_1, l_2, ..., l_k$ са *линейно независими*.

4. Няколко примера. а) Векторите $l_1 = (1, 0, 2, 3)$, $l_2 = (-2, -1, 0, 0)$, $l_3 = (8, 3, 4, 6)$ от R^4 са линейно зависими, защото $2l_1 + (-3)l_2 + (-1)l_3 = 0$.

б) Векторите

$$e_1 = (1, 0, ..., 0),$$
$$e_2 = (0, 1, ..., 0),$$
$$\dots\dots\dots$$
$$e_n = (0, 0, ..., 1)$$

от R^n са линейно независими, защото

$$\lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n = (\lambda_1, \lambda_2, ..., \lambda_n)$$

и в дясната страна ще стои нулевият вектор само ако $\lambda_1 = ... = \lambda_n = 0$. Забележително е и друго: всеки вектор $x = (x_1, ..., x_n)$ от R^n е линейна комбинация на $e_1, ..., e_n$. Действително

$$x = x_1 e_1 + ... + x_n e_n.$$

$$x = x_1 e_1 + \ldots + x_n e_n .$$

Коефициентите в тази линейна комбинация са еднозначно определени: ако

$$x = x_1' e_1 + \ldots + x_n' e_n ,$$

то от двете равенства следва

$$0 = \left(x_1 - x_1' \right) e_1 + \ldots + \left(x_n - x_n' \right) e_n ,$$

а тъй като векторите $\{e_i\}$ са линейно независими, получаваме, че $x_1 = x_1', \ldots, x_n = x_n'$.

Ще аксиоматизираме ситуацията от разгледания пример б).

5. Определение. Ще казваме, че векторите $l_1, l_2, \ldots, l_k \in \mathrm{R}^n$ образуват *базис* на пространството R^n, ако те са линейно независими и всеки вектор от R^n е тяхна линейна комбинация.

Очевидно векторите от пример б) на т. 4 образуват базис, който ще наричаме *стандартен*. Малко по-късно ще докажем, че всеки базис на R^n съдържа точно n вектора.

Налага се да направим следното терминологично уточнение. Ако $a_1, a_2, \ldots, a_k$ са дадени вектори, ще казваме, че е дадена *система* (*фамилия*) от вектори. Не се изключва възможността тя да съдържа повторения, т.е. някои от векторите с различни номера да са равни. За да подчертаем това, казваме “система”, а не “множество” или “подмножество”. Ако $\{i_1, i_2, \ldots, i_m\}$ е подмножество на множеството $\{1, 2, \ldots, k\}$, ще казваме, че векторите $a_{i_1}, a_{i_2}, \ldots, a_{i_m}$ образуват *подсистема* на дадената система от вектори.

Следващата теорема ще прилагаме многократно по-нататък.

6. Основна теорема. *Ако векторите от системата*

(A) $$a_1, a_2, \ldots, a_k$$

са линейни комбинации на векторите от системата

(B) $$b_1, b_2, \ldots, b_m$$

и $k > m$ *, то векторите от системата* (A) *са линейно зависими.*

Доказателство. По условие

$$a_1 = \lambda_{11}b_1 + \lambda_{21}b_2 + \dots + \lambda_{m1}b_m,$$
$$a_2 = \lambda_{12}b_1 + \lambda_{22}b_2 + \dots + \lambda_{m2}b_m,$$
$$\dots\dots\dots\dots\dots\dots$$
$$a_k = \lambda_{1k}b_1 + \lambda_{2k}b_2 + \dots + \lambda_{mk}b_m,$$

където λ_{ij} са числа. Да образуваме произволна линейна комбинация на $a_1, a_2, \dots, a_k$ с коефициенти $x_1, x_2, \dots, x_k$:

$$x_1a_1 + x_2a_2 + \dots + x_ka_k = \left(\lambda_{11}x_1 + \lambda_{12}x_2 + \dots + \lambda_{1k}x_k\right)b_1 +$$
$$+\left(\lambda_{21}x_1 + \lambda_{22}x_2 + \dots + \lambda_{2k}x_k\right)b_2 + \dots + \left(\lambda_{m1}x_1 + \lambda_{m2}x_2 + \dots + \lambda_{mk}x_k\right)b_m.$$

Теоремата ще бъде доказана, ако намерим такива числа $x_1, x_2, \dots, x_k$, които да не са едновременно нули и за които тази линейна комбинация да бъде равна на нулевия вектор. Това ще бъде постигнато, ако например намерим ненулево решение на хомогенната система

$$\lambda_{11}x_1 + \lambda_{12}x_2 + \dots + \lambda_{1k}x_k = 0$$
$$\dots\dots\dots\dots\dots\dots$$
$$\lambda_{m1}x_1 + \lambda_{m2}x_2 + \dots + \lambda_{mk}x_k = 0.$$

Тъй като по условие броят k на неизвестните е по-голям от броя m на уравненията, тя действително има ненулево решение (вж. гл. 1, § 3, т. 3, г). С това доказателството е завършено.

7. Следствие. *Всеки базис на* R^n *се състои точно от* n *вектора.*

Доказателство. Нека $e_1, \dots, e_n$ е стандартният базис от т. 4, б), а $l_1, \dots, l_k$ е произволен базис. По определението на базис и двете системи $\{e_i\}$, $\{l_j\}$ са линейно независими. Тъй като векторите $\{l_j\}$ са линейни комбинации на векторите $\{e_i\}$, то $k \le n$: в противен случай по основната теорема $\{l_j\}$ ще се окажат линейно зависими. От друга страна, векторите $\{e_i\}$ са линейни комбинации на базиса $\{l_j\}$. Пак по същото съображение получаваме $n \le k$. Окончателно $n = k$.

8. Следствие. *Всеки* $n+1$ *вектора от* R^n *са линейно зависими.*

Доказателство. Да изберем базис на R^n и да изразим дадените $n+1$ вектора чрез базиса. Тъй като той се състои от n вектора, твърдението следва от основната теорема.

9. Теорема. *В сила са следните твърдения.*

а) *Ако системата вектори $l_1, \ldots, l_k$ съдържа линейно зависима подсистема, то и цялата система е линейно зависима.*

б) *Всяка подсистема на линейно независима система от вектори е линейно независима.*

в) *Ако векторите $l_1, \ldots, l_k$ са линейно зависими, то поне един от тях е линейна комбинация на останалите.*

г) *Ако поне един от векторите $l_1, \ldots, l_k$* е *линейна комбинация на останалите, системата е линейно зависима.*

д) *Ако векторите $l_1, \ldots, l_k$ са линейно независими, а $l_1, \ldots, l_k, l$ са линейно зависими, то векторът l е линейна комбинация на $l_1, \ldots, l_k$.*

е) *Ако векторите $l_1, \ldots, l_k$ са линейно независими и векторът l не е тяхна линейна комбинация, системата $l_1, \ldots, l_k, l$ е линейно независима.*

Доказателство. а) Нека с точност до номерация векторите $l_1, \ldots, l_m$, $m < k$, са линейно зависими и

$$\lambda_1 l_1 + \ldots + \lambda_m l_m = 0,$$

където $\lambda_1, \ldots, \lambda_m$ не са едновременно равни на нули. Тогава и в равенството

$$\lambda_1 l_1 + \ldots + \lambda_m l_m + 0.l_{m+1} + \ldots + 0.l_k = 0$$

не всички коефициенти са нули, следователно векторите $l_1, \ldots, l_k$ са линейно зависими. Твърдението б) следва от а). За да докажем в), нека $\lambda_1 l_1 + \ldots + \lambda_{k-1} l_{k-1} + \lambda_k l_k = 0$ е линейна зависимост, в която не всички коефициенти са нули и да речем, че $\lambda_k \neq 0$. Тогава имаме

$$l_k = -\frac{\lambda_1}{\lambda_k} l_1 - \ldots - \frac{\lambda_{k-1}}{\lambda_k} l_{k-1}.$$

г) Нека например $l_k = \lambda_1 l_1 + \ldots + \lambda_{k-1} l_{k-1}$. Следователно $\lambda_1 l_1 + \ldots + \lambda_{k-1} l_{k-1} + (-1) l_k = 0$ и не всички коефициенти са нули.

д) Нека $\lambda_1 l_1 + \ldots + \lambda_k l_k + \lambda l = 0$ и $(\lambda_1, \ldots, \lambda_k, \lambda) \neq (0, \ldots, 0, 0)$. Ако $\lambda = 0$, ще се окаже, че $l_1, \ldots, l_k$ са линейно зависими, което противоречи на условието. Следователно $\lambda \neq 0$ и $l = -\frac{\lambda_1}{\lambda} l_1 - \ldots - \frac{\lambda_k}{\lambda} l_k$. Твърдението е) следва от д).

10. Координати. Нека $e_1, \ldots, e_n$ е базис на R^n (или на C^n). Тогава всеки вектор l може да се запише като линейна комбинация на базисните:

(1) $$l = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Това може да се направи по единствен начин. Доказателството е същото, както в т. 4, б). Числата $x_1, x_2, \ldots, x_n$, определени от (1), се наричат *координати на вектора* l *спрямо базиса* $e_1, e_2, \ldots, e_n$ и според доказаното при избран базис връзката "вектор-координати" е взаимно еднозначна. Тъй като координатите на наредената n-торка (вектора) l също са наредена n-торка, тази връзка засега изглежда тавтологична. Важността й ще стане ясна по-късно. Разбира се, ако сменим базиса, същият вектор l спрямо новия базис ще има, изобщо казано, други координати. Нека $l = \sum_{i=1}^{n} x_i e_i$, $m = \sum_{i=1}^{n} y_i e_i$.
Тогава

$$\lambda l + \mu m = \sum_{i=1}^{n} \left(\lambda x_i + \mu y_i \right) e_i ,$$

т.е. *координатите на линейна комбинация* $\lambda l + \mu m$ *на два вектора са същите линейни комбинации на техните координати*, а именно

$$\lambda \left(x_1, \ldots, x_n \right) + \mu \left(y_1, \ldots, y_n \right).$$

§ 2. Подпространства

1. Не е трудно да се съобрази, че ако в пространството R^n (или в C^n) изберем произволни n линейно независими вектора $e_1, e_2, \ldots, e_n$, то те са базис. Действително, ако l е произволен вектор от пространството, то съгласно следствието от § 1, т. 8 векторите $e_1, \ldots, e_n, l$ са линейно зависими, а според теоремата от § 1, т. 9 д) l е линейна комбинация на $e_1, e_2, ..., e_n$. И така, ако са налице n линейно независими вектора, линейните им комбинации запълват цялото пространство. Ясно е, че ако изберем произволно векторите $l_1, l_2, ..., l_r$ от R^n (или от C^n) и $r<n$, линейните им комбинации вече няма да запълнят цялото пространство. Това е така, защото всеки базис на R^n (или C^n) се състои от n вектора, а тук $r<n$. Въпреки това ще получим специфично подмножество, което заслужава специално изучаване.

2. Определение. Ще казваме, че непразното подмножество V на R^n е *подпространство* на линейното пространство R^n, ако заедно с всеки два вектора a, b от V то съдържа и всяка тяхна линейна комбинация $\lambda a + \mu b$.

Ясно е, че това условие може да се формулира еквивалентно и така: от $a, b \in V$ винаги да следва $a + b \in V$ и $\lambda a \in V$ за всяко число $\lambda \in \mathrm{R}$. Ще казваме накратко, че V е затворено относно операциите в R^n.

3. Примери. а) Очевидно $V = \{0\}$ (0 е нулевият вектор) е подпространство. Ще го наричаме *нулево подпространство*.

б) Нека $l_1, ..., l_k \in \mathrm{R}^n$ и V е множеството на всички техни линейни комбинации:

$$V = \left\{ \lambda_1 l_1 + ... + \lambda_k l_k \setminus \lambda_1, ..., \lambda_k \in \mathrm{R} \right\}.$$

Проверката, че V е подпространство, оставяме на читателя. Ще казваме, че V е *породено* от векторите $l_1, ..., l_k$ или че е тяхна *линейна обвивка*. Скоро ще видим, че всяко подпространство на R^n се получава по този начин.

в) Всяко подпространство V съдържа нулевия вектор, защото от $a \in V$ следва $a + (-1)a = 0 \in V$. Следователно, ако V_1 и V_2 са подпространства, то $V_1 \cap V_2$ не е празно. Покажете, че $V_1 \cap V_2$ е подпространство.

г) Ще покажем, че решенията на всяка хомогенна система от линейни уравнения с n неизвестни образуват подпространство на R^n.

Наистина, нека

$$
\begin{aligned}
&a_{11}x_1 + \ldots + a_{1n}x_n = 0 \\
&\ldots\ldots\ldots\ldots \\
&a_{i1}x_1 + \ldots + a_{in}x_n = 0 \\
&\ldots\ldots\ldots\ldots \\
&a_{m1}x_1 + \ldots + a_{mn}x_n = 0
\end{aligned}
$$

е дадена хомогенна система и $x' = \left(x_1', ..., x_n'\right)$, $x'' = \left(x_1'', ..., x_n''\right)$ са две произволни нейни решения. Имаме

$$\sum_{k=1}^{n} a_{ik}\left(\lambda x_k' + \mu x_k''\right) = \lambda \sum_{k=1}^{n} a_{ik} x_k' + \mu \sum_{k=1}^{n} a_{ik} x_k'' = \lambda.0 + \mu.0 = 0,$$

което показва, че линейната комбинация $\lambda x' + \mu x''$ на решения също е решение.

Това подпространство ще бъде изследвано подробно в § 5.

4. Базиси и размерности. Нека $V \subset \mathrm{R}^n$ е ненулево подпространство $\left(V \neq \{0\}\right)$. Ще казваме, че векторите $e_1, \ldots, e_r \in V$ образуват *базис* на подпространството V, ако те са линейно независими и всеки вектор от V е тяхна линейна комбинация.

Това определение повдига два въпроса: първо, дали във всяко ненулево подпространство съществува поне един базис и, второ, дали всеки два базиса (ако съществуват) съдържат еднакъв брой вектори.

Тъй като V е ненулево, то съдържа ненулеви вектори, а всеки ненулев вектор образува линейно независима система. Следователно V съдържа линейно независими системи от вектори. Всяка такава система ще съдържа най-малко един и най-много n вектора, защото $V \subset \mathrm{R}^n$, а в R^n всеки $n+1$ вектора са линейно зависими (вж. § 1, т. 9). Следователно във V можем да изберем поне една линейно независима система $l_1, \ldots, l_r$, в която броят r на векторите да бъде възможно най-голям. Нека сега l е произволен вектор от V и да разгледаме системата $l_1, \ldots, l_r, l$. Тя е линейно зависима поради максималния избор на r и тогава съгласно § 1, т. 9, д) векторът l е линейна комбинация на $l_1, \ldots, l_r$. Следователно векторите $l_1, \ldots, l_r$ образуват базис на V. За да докажем, че всеки два базиса на V съдържат еднакъв брой вектори, достатъчно е да повторим разсъждението от § 1, т. 7.

Да резюмираме резултатите.

5. Теорема. *Всяко ненулево подпространство* V *на* R^n *има поне един базис. Броят на векторите в който и да е базис на* V *е един и същ*; ще го бележим с $\dim V$ и ще го наричаме *размерност на подпространството* V. Ако $V=\{0\}$, ще считаме, че $\dim V=0$.

Тук символът dim е съкращение от dimension (англ.) - ‘размерност’. Фактът, че всяко ненулево подпространство има базис, може да се изтълкува и като твърдение, че всяко ненулево подпространство на L е линейна обвивка на краен брой вектори.

6. Теорема (за допълване на базиса). *Ако* $L=\mathrm{R}^n$ *или* $L=\mathrm{C}^n$, а V *е ненулево подпространство на* L, *то всеки базис на* V *може да се допълни до базис на цялото пространство* L.

Доказателство. Нека $e_1,\dots,e_k$ е базис на V. Ако $V=L$, няма какво да се доказва. Нека $V\neq L$ и да изберем по произволен начин вектор e_{k+1} от $L\setminus V$. Той не е линейна комбинация на $e_1,\dots,e_k$, защото не принадлежи на V. Съгласно теоремата от § 1, т. 9, е) векторите $e_1,\dots,e_k,e_{k+1}$ са линейно независими. Нека V_1 е тяхната линейна обвивка. Ако $V_1\neq L$, към разглежданите вектори отново присъединяваме вектор e_{k+2} от $L\setminus V_1$ и както по-горе, получаваме линейно независима система от $k+2$ вектора. Ясно е, че след краен брой стъпки тази процедура ще ни доведе до n линейно независими вектора $e_1,\dots,e_k,e_{k+1},\dots,e_n$, т.е. до базис на L.

7. Твърдение (монотонност на размерността). *За всяко подпространство* V *на* $L=\mathrm{R}^n$ *или* $L=\mathrm{C}^n$ *е изпълнено* $\dim V\le\dim L$. Равенството е в сила точно тогава, когато $V=L$.

Доказателство. Ако $V=\{0\}$, то по определение $\dim V=0$ и неравенството за размерностите е очевидно. Ако V е ненулево подпространство, неравенството следва от теоремата в т.6. Тъй като $\dim L=n$, от равенството $\dim V=\dim L$ следва, че V има базис от n вектора, а в т.1 отбелязахме, че всеки n линейно независими вектора са базис на L. Следователно всеки базис на V е базис на L и $V=L$.

§ 3. Ранг на система от вектори и ранг на матрица

1. Нека $l_1,\dots,l_m$ е произволна система вектори от R^n или от C^n и

$$V=\left\{\sum_{i=1}^{m}\lambda_i l_i \setminus \lambda_i\in\mathcal{K}\right\},\ \mathcal{K}=\mathrm{R}\ \text{ или }\ \mathrm{C}$$

е подпространството, което те пораждат. Тук ще търсим ефективни средства за пресмятане на $\dim V$ и за намиране на базиси на V .

Преди да въведем съответната терминология, ще направим някои разяснения. Ако не всички вектори l_i са нулеви, то от системата $l_1, \ldots, l_m$ можем да избираме линейно независими подсистеми. В частност, можем да изберем линейно независима подсистема, в която броят на векторите да е възможно най-голям (максимален). Нека с точност до номерация това е подсистемата $l_1, \ldots, l_r$. Поради максималния избор на r векторите $l_1, \ldots, l_r, l_i$ са линейно зависими за всяко $i = 1, 2, \ldots, m$. Според теоремата от § 1, т. 9 д) всички вектори l_i , $i = 1, 2, \ldots, m$ са линейни комбинации на $l_1, \ldots, l_r$. Следователно линейната обвивка на векторите $l_1, \ldots, l_r$ е подпространството V , те са негов базис и $\dim V = r$. Нашата задача се свежда до търсене на алгоритъм, с помощта на който от дадена система, състояща се от краен брой вектори, да изберем *максимална линейно независима подсистема*, т.е. подсистема, която да е линейно независима и броят на векторите в нея да е възможно най-голям. Ще забележим, че подсистемата $l_{i_1}, l_{i_2}, \ldots, l_{i_k}$ от линейно независими вектори е максимална линейно независима тогава и само тогава, когато за всеки вектор l_i , $i = 1, 2, \ldots, m$, от дадената система векторите $l_{i_1}, l_{i_2}, \ldots, l_{i_k}, l_i$ са линейно зависими, т.е. *ако подсистемата е линейно независима и е максимална относно включването в теоретико-множествен смисъл*. Действително, ако k е възможно най-голямо, т.е. $k = r$, твърдението току-що бе доказано. Обратно, ако за всяко $i = 1, 2, \ldots, m$ векторите $l_{i_1}, l_{i_2}, \ldots, l_{i_k}, l_i$ са линейно зависими, както по-горе се вижда, че $l_{i_1}, \ldots, l_{i_k}$ са базис на подпространството V . Според теоремата от § 2, т. 5 броят на векторите в който и да е базис на подпространството V е един и същ, следователно $k = r$.

Не е трудно да се съобрази, че всяка линейно независима подсистема може да бъде допълнена до максимална линейно независима подсистема. Действително, нека с точност до номерацията векторите $l_1, l_2, \ldots, l_k$ са линейно независими. Ако те не са максимална линейно независима подсистема, то съществува вектор l_{j_1}, $k + 1 \leq j_1 \leq m$, такъв, че векторите $l_1, l_2, \ldots, l_k, l_{j_1}$ са линейно независими. Ако те не са максимална линейно независима подсистема, ще присъединим към тях вектор l_{j_2} и т. н., докато стигнем до максимална линейно независима подсистема.

Нека по-конкретно $l_i = \left(a_{i1}, a_{i2}, ..., a_{in}\right)$, $i = 1, 2, ..., m$. Естествено е да образуваме матрицата $A = \left(a_{ij}\right)$ от тип $m \times n$; редовете ѝ са векторите l_i, $i = 1, 2, ..., m$. Търсим числа $x_1, ..., x_m$ от $\mathcal{K}$, които да не са едновременно нули и за които да бъде изпълнено векторното равенство

$$x_1 l_1 + x_2 l_2 + ... + x_m l_m = 0.$$

Записано като равенства между съответните компоненти на вектора от лявата страна и на вектора от дясната, то е еквивалентно на системата

$$\begin{array}{l} a_{11}x_1 \; + a_{21}x_2 + ... + a_{m1}x_m = 0 \\ a_{12}x_1 \; + a_{22}x_2 + ... + a_{m2}x_m = 0 \\ . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \\ a_{1n}x_1 + a_{2n}x_2 + ... + a_{mn}x_m \; = 0. \end{array}$$

Ако $m > n$, вече знаем, че системата има ненулево решение. Да допуснем за момент, че $m \leq n$ и че тя има ненулево решение. Разбира се, то удовлетворява поотделно всяко уравнение на системата, следователно както и да подберем m уравнения от системата, ще получим хомогенна система от m уравнения с m неизвестни, която има ненулево решение. Съгласно теоремата от гл. 1, § 7, т. 6 детерминантата от коефициентите пред неизвестните в такава система е равна на нула. Този резултат може да се формулира така: необходимо условие при $m \leq n$ разглежданите вектори да бъдат линейно зависими е всички детерминанти от ред m, образувани от произволни m стълба на матрицата A, да бъдат равни на нула. Оттук чисто формално следва, че ако $m \leq n$ и някоя от детерминантите, за които стана дума, е различна от нула, то векторите-редове на матрицата A са линейно независими. Този повърхностен резултат, да го наречем наблюдение, подсказва, че линейната зависимост или независимост на система вектори би могла да се формулира чрез свойства на детерминанти, образувани от различни части на матрицата A съставена от компонентите на векторите.

2. Определение. Нека $l_1, l_2, ..., l_m$ е система вектори от R^n или C^n. Максималния брой линейно независими вектори в тази система ще наричаме неин *ранг*. Ако системата се състои само от нулеви вектори, ще считаме, че рангът ѝ е числото нула.

От разсъжденията в т.1 следва, че *рангът на системата съвпада* *с размерността на подпространството, което тя поражда*. Пак от същите разсъждения се получава, че една система вектори има ранг r тогава и само тогава, когато тя съдържа r линейно независими вектора, такива, че всеки вектор от системата да е тяхна линейна комбинация.

3. Твърдение (за сравняване на рангове). *Ако векторите* $b_1, \dots, b_m$ *са линейни комбинации на векторите* $a_1, \dots, a_k$, *то рангът на системата* $b_1, \dots, b_m$ *не надминава ранга на системата* $a_1, \dots, a_k$.

Доказателство. Да означим с r_a и r_b съответно ранговете на системите $a_1, \dots, a_k$ и $b_1, \dots, b_m$, а с V_a и V_b - подпространствата, които те пораждат. При тези означения $\dim V_a = r_a$ и $\dim V_b = r_b$. От условието на твърдението следва, че V_b е подмножество на V_a. Да изберем базиси на V_a и V_b. От включването $V_b \subset V_a$ следва, че всеки от избраните r_b на брой базисни вектори на V_b е линейна комбинация на избраните базисни вектори на V_a (техният брой е r_a). Ако $r_b > r_a$, то според теоремата от § 1, т. 6 избраният базис на V_b ще се окаже линейно зависим, което противоречи на определението на базис. Следователно $r_b \leq r_a$ и твърдението е доказано.

4. Определение. Нека A е матрица от тип $m \times n$. Ранга на системата вектори-редове (съотв. вектори-стълбове) ще наричаме *ранг на матрицата по редовете* (съотв. *ранг по стълбове*).

От определението на ранг на система вектори следва в частност, че ако A е нулева матрица, рангът ѝ по редовете съвпада с ранга по стълбовете и е равен на нула.

5. Определение. Нека A е матрица от тип $m \times n$, в която по произволен начин са избрани k реда и k стълба, $1 \leq k \leq \min(m, n)$. Елементите на A, които лежат в пресечните точки на избраните редове и стълбове, образуват квадратна матрица от ред k. Нейната детерминанта ще наричаме *минор от* k*-ти ред, съдържащ се в избраните редове и стълбове*.

6. Теорема. *За произволна ненулева матрица рангът ѝ по редовете съвпада с ранга ѝ по стълбовете и е равен на най-високия ред на различен от нула минор на матрицата*.

Доказателство. Нека разглежданата матрица $A = (a_{ij})$ е от тип $m \times n$ и нека r е най-високият ред на различен от нула минор. Това означава, че всички минори от ред $r+1$ до ред $\min(m, n)$ са равни на нула. Да предположим за начало на доказателството, че различен от нула е минорът

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix}.$$

Нека $1 \leq i \leq m$, $1 \leq j \leq n$ и да разгледаме детерминантата от ред $r+1$

$$\Delta^* = \begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1j} \\ a_{21} & \dots & a_{2r} & a_{2j} \\ \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot \\ a_{r1} & \dots & a_{rr} & a_{rj} \\ a_{i1} & \dots & a_{ir} & a_{ij} \end{vmatrix}.$$

Ако $1 \leq i \leq r$ или $1 \leq j \leq r$, тя е равна на нула, защото съдържа два еднакви реда или стълба, а в останалите случаи детерминантата Δ^* е минор от ред $r+1$ на Δ и тогава е равна на нула по условие. Следователно $\Delta^* = 0$ във всички случаи. Да развием Δ^* по последния стълб:

$$a_{1j} A_{1\ r+1} + a_{2j} A_{2\ r+1} + \dots + a_{rj} A_{r\ r+1} + a_{ij} \Delta = 0$$

($A_{k\ r+1}$ са съответните адюнгирани количества; използвахме, че $A_{r+1\, r+1} = (-1)^{2r+2} \Delta = \Delta$). Тъй като $\Delta \neq 0$, получаваме равенството

$$a_{ij} = -\frac{A_{1\ r+1}}{\Delta} a_{1j} - \frac{A_{2\ r+1}}{\Delta} a_{2j} - \dots - \frac{A_{r\ r+1}}{\Delta} a_{r\,j}.$$

То е в сила за всяко $j = 1, \dots, n$, а коефициентите $A_{k\ r+1}/\Delta$, $k = 1, \dots, r$, не зависят от елементите на последния стълб на детерминантата Δ^*, т.е. не зависят от j съгласно установената връзка между адюнгирани количества и поддетерминанти. Смисълът на всичко това е, че i -тият ред, $i = 1, \ldots, m$, на матрицата A е линейна комбинация на първите r реда. Те от своя страна съдържат ненулев минор от ред r и съгласно бележката от края на т. 1 са линейно независими. Следователно рангът на A по редовете е равен на r . Като приложим този резултат за транспонираната матрица A^t , ще получим, че рангът на A по стълбове също е равен на r .

Разсъждението, което проведохме, съдържа ограничителното предположение, че разглежданият ненулев минор се съдържа в първите r реда и r стълба на матрицата A . Нека в общия случай той се съдържа в редовете с номера $m_1, \ldots, m_r$ и в стълбовете с номера $n_1, \ldots, n_r$ и по условие всички минори от ред $r+1$ (ако има такива) са равни на нула. Целта ни е с разместване на редовете и стълбовете на матрицата A "да преместим" този минор в горния ляв ъгъл на матрицата и да сведем общия случай към разгледания по-горе. При тази процедура, разбира се, ще получим някаква нова матрица A^* . Ако в A има ненулев минор от ред r и всички минори от ред $r+1$ (ако има такива) са равни на нули, то същото важи и за минорите в матрицаA^* . Това е така, защото разместването на редове и стълбове в матрицата води до евентуално разместване на редове и стълбове в съответните минори, което най-много променя знака им. Малко по-сложно е с ранговете. Трябва да сме сигурни, че ранговете на A и A^* по редовете (съотв. по стълбовете) са равни. Това не е съвсем очевидно. Вярно е наистина, че ако в A разместваме например само стълбове, то рангът по стълбовете не се променя, но се разместват компонентите на редовете, т. е. получават се други вектори-редове. Дали новата система вектори-редове има същия ранг както изходната? Отговорът е утвърдителен.

Наистина, да изоставим за момент матрицата A и да разсъждаваме по-общо. Нека във всички вектори от R^n (или от C^n) сменим по един и същ начин номерата на компонентите им (все едно във всички вектори да разместим по един и същ начин компонентите). Ще означаваме с l' вектора, в който се трансформира при тази процедура произволен вектор l . По определение два вектора l_1 и l_2 са равни тогава и само тогава, когато всички техни съответни компоненти са равни. Следователно равенствата $l_1 = l_2$ и $l_1' = l_2'$ са еквивалентни, откъдето непосредствено получаваме, че всяко векторно равенство от вида

$$l = \lambda_1 l_1 + \lambda_2 l_2 + \ldots + \lambda_k l_k$$

е еквивалентно на равенството

$$l' = \lambda_1 l'_1 + \lambda_2 l'_2 + \ldots + \lambda_k l'_k.$$

Ако сега вземем предвид и бележката от т. 2, убеждаваме се, че ако произволна система вектори $l_1, l_2, \ldots, l_m$ има ранг r, то същия ранг има и системата $l'_1, l'_2, \ldots, l'_m$.

Вече можем да забележим, че ако в матрицата A разместваме стълбовете, компонентите на всички вектори-редове се разместват по един и същ начин, следователно ранговете на матрицата по редовете и съответно по стълбовете не се променят. Същото заключение получаваме и ако разместваме редове. Следователно общият случай, свързан с разположението на ненулев минор от ред r, може наистина да се сведе до случая, разгледан в началото, което завършва доказателството на теоремата.

Доказаната теорема дава възможност да говорим просто за *ранг на матрицата* A - ще го означаваме с $r(A)$, без да уточняваме дали става дума за ранга по редовете, или по стълбовете.

Ще забележим, че ако в матрицата A всички минори от ред $r+1$ са равни на нула, то равни на нула са и минорите от ред $r+2, r+3, \ldots, \min(m,n)$. Достатъчно е да си спомним, че по определение детерминанта от ред $r+2$ се изразява чрез поддетерминанти от ред $r+1$, които в случая са минори от ред $r+1$, съдържащи се в матрицата A.

7. Следствие. *Необходимо и достатъчно условие една детерминанта да бъде равна на нула е някой неин ред* (съотв. *стълб*) *да бъде линейна комбинация на останалите редове* (съотв. *стълбове*).

Доказателство. Нека A е квадратна матрица от ред n. Ако $|A|=0$, то $r(A)<n$, което означава, че редовете й са линейно зависими, следователно някой от тях е линейна комбинация на останалите (вж. § 1, т. 8, в). Ако пък някой от редовете е линейна комбинация на останалите, то системата от редовете е линейно зависима (§ 1, т. 8, г), следователно $r(A)<n$ и минорът $|A|$ от ред n е равен на нула. Същото разсъждение важи и за стълбовете.

§ 4. Теорема на Кронекер - Капели

1. С помощта на понятието ранг на матрица вече можем да формулираме една фундаментална теорема, която дава необходимо и достатъчно условие за съвместимост на произволна система от линейни уравнения. Нека напомним, че със системата

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

асоциирахме матриците

$$
A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ .. & .. & .. & .. \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}, \qquad \overline{A} = \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ .. & .. & .. & .. \\ a_{m1} & \ldots & a_{mn} & b_m \end{pmatrix},
$$

които нарекохме съответно матрица от коефициентите пред неизвестните и разширена матрица на системата.

2. Теорема (Кронекер-Капели)[*]. *Произволна система линейни уравнения е съвместима тогава и само тогава, когато рангът на матрицата от коефициентите пред неизвестните съвпада с ранга на разширената ѝ матрица*.

Доказателство. Нека за определеност разглежданата система уравнения е (1). Ако матрицата A се състои само от нули, очевидно е, че системата е съвместима точно тогава, когато и матрицата $\overline{A}$ се състои само от нули, т. е. тогава и само тогава, когато $r\left(A\right) = r\left(\overline{A}\right) = 0$. По-нататък ще предполагаме, че A не е нулева матрица. Нека $r\left(A\right) = r\left(\overline{A}\right) = r$. По определенията на ранг на матрица и на ранг на система от вектори от векторите-стълбове на матрицата A можем да изберем r линейно независими стълба, при което всеки стълб на A ще бъде тяхна линейна комбинация. Накратко казано, избираме базис на системата от вектори-стълбове. Тъй като $r\left(\overline{A}\right) = r$, той ще бъде базис и на системата от векторите-стълбове в разширената матрица $\overline{A}$. По-специално, последният стълб на $\overline{A}$ е линейна комбинация на избраните стълбове от A, а следователно линейна комбинация на всички стълбове от

[*] Известна е и като теорема на Руше. (Бел. авт.)

A ("ненужните" стълбове може да се вземат с коефициенти нули). Това означава, че съществуват числа $\lambda_1, \dots, \lambda_n$, за които

$$
(2) \qquad \begin{aligned}
&a_{11}\lambda_1 + a_{12}\lambda_2 + \dots + a_{1n}\lambda_n = b_1, \\
&a_{21}\lambda_1 + a_{22}\lambda_2 + \dots + a_{2n}\lambda_n = b_2, \\
&\dots\dots\dots\dots\dots\dots\dots \\
&a_{m1}\lambda_1 + a_{m2}\lambda_2 + \dots + a_{mn}\lambda_n = b_m,
\end{aligned}
$$

т.е. $(\lambda_1, \dots, \lambda_n)$ е решение на системата (1). И така, ако ранговете съвпадат, (1) е съвместима.

Нека системата (1) е съвместима и $(\lambda_1, \dots, \lambda_n)$ е нейно решение, т.е. в сила са равенствата (2). Ако r е рангът на A, да изберем базис на системата от вектори-стълбове на A. Той ще бъде базис и за стълбовете на $\overline{A}$. Действително, всички стълбове на $\overline{A}$, освен последния, са стълбове и на A, следователно са линейни комбинации на избрания вече базис. От друга страна, равенствата (2) показват, че последният стълб на $\overline{A}$ е линейна комбинация на стълбовете на A, а оттук следва, че е линейна комбинация и на базисните стълбове. Следователно $r\left(\overline{A}\right) = r(A)$.

§ 5. Системи линейни уравнения. Хомогенни системи

1. Вече разполагаме с достатъчно развит апарат, за да завършим теорията на линейните системи. Ще се придържаме към означенията от § 4. Разглеждаме отново системата (1) и да предположим, че тя е съвместима. Съгласно теоремата на Кронекер-Капели рангът на матрицата $A = (a_{ij})$ съвпада с ранга на разширената матрица $\overline{A}$. Нека $r(A) = r\left(\overline{A}\right) = r$. Ако $r = 0$, т.е. ако и двете матрици са нулеви, то очевидно всяка наредена n-торка от числа е решение на системата (1). Нататък ще предполагаме, че матрицата A е ненулева. Сега можем така да изберем r линейно независими реда на $\overline{A}$, че останалите редове да бъдат техни линейни комбинации. Иначе казано, можем така да изберем r линейно независими уравнения от системата, че останалите да следват от тях и следователно, може да се изпуснат. Нека (след евентуална смяна на номерацията на уравненията) системата се редуцира на

$$
(1') \qquad
\begin{array}{l}
a_{11}x_1 \ + \ldots + a_{1r}x_r + a_{1r+1}x_{r+1} + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 \ + \ldots + a_{2r}x_r + a_{2r+1}x_{r+1} + \ldots + a_{2n}x_n = b_2 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{r1}x_1 + \ldots + a_{rr}x_r + a_{rr+1}x_{r+1} + \ldots + a_{rn}x_n = b_r,
\end{array}
$$

която е еквивалентна на дадената и освен това също е съвместима. По условие рангът на разширената матрица в (1') е r , следователно рангът на матрицата от коефициентите пред неизвестните също е r и тя съдържа ненулев минор от ред r . С точност до номерация на неизвестните нека той е

$$
\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ .\ . & .\ . & .\ . & .\ . \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{vmatrix}, \qquad \Delta \neq 0.
$$

Ако $r=n$, то съгласно теоремата от гл.1, § 7, т. 6 системата (1') е съвместима и има единствено решение. Ако $r<n$, нека считаме, че $x_{r+1}, \ldots, x_n$ са свободни неизвестни и да положим $x_{r+1} = \lambda_{r+1}, \ldots, x_n = \lambda_n$ (λ_i са произволни числа). Тъй като $\Delta \neq 0$, сега $x_1, \ldots, x_r$ (главните неизвестни) се определят еднозначно от системата

$$
(2) \qquad
\begin{array}{l}
a_{11}x_1 + \ldots + a_{1r}x_r = b_1 - a_{1r+1}\lambda_{r+1} - \ldots - a_{1n}\lambda_n \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{r1}x_1 + \ldots + a_{rr}x_r = b_r - a_{rr+1}\lambda_{r+1} - \ldots - a_{rn}\lambda_n.
\end{array}
$$

Намирането им може да стане например с формулите на Крамер.

Ако дадена система е съвместима, числото $r = r(A) = r(\overline{A})$ ще наричаме просто *ранг на системата*. От определението на ранг на матрица следва, че рангът на дадена система не надминава броя на неизвестните. Дотук доказахме

2. Твърдение. *Ако дадена система линейни уравнения с* n *неизвестни е съвместима и* r *е рангът й, то:*

а) *системата е определена тогава и само тогава, когато* $r=n$;

б) *ако* $r<n$, *решенията на системата зависят от* $r-n$ *независими параметъра (свободни неизвестни). Броят на главните неизвестни е равен на* r .

От това твърдение се вижда в частност, че както и да привеждаме системата в стъпаловидна форма, то броят на главните (съотв. свободните) неизвестни е винаги един и същ и зависи само от ранга на системата. Това не е учудващо, защото привеждането в стъпаловидна форма се извършва с помощта на елементарни преобразувания на редовете на матрицата $\overline{A}$, а те запазват ранга.

3. Хомогенни системи. В § 2, т. 3, г) показахме, че множеството от решенията на хомогенна система с n неизвестни е подпространство на R^n. Сега ще определим размерността му и ще посочим просто правило за намиране на поне един негов базис. Като запазваме означенията от § 4 и предположенията в т. 1, ще предполагаме още, че $b_1 = \ldots = b_m = 0$. Разбира се, всяка хомогенна система винаги е съвместима : $(0, 0, \ldots, 0)$ е решение. Ако $r=n$, вече знаем, че това решение е единствено. Нека $r<n$ и да приложим за (2) формулите на Крамер. Имаме

$$x_i = \frac{1}{\Delta}\begin{vmatrix} a_{11} & \ldots & -\sum\limits_{j=r+1}^{n} a_{1j}\lambda_j & \ldots & -a_{1r} \\ . & . & . & . & . \\ a_{r1} & \ldots & -\sum\limits_{j=r+1}^{n} a_{rj}\lambda_j & \ldots & -a_{rr} \end{vmatrix}, \qquad i = 1, \ldots, r.$$

Ако за всяко $i = 1, \ldots, r$ развием детерминантата по i -тия стълб, ще стигнем до изрази от вида

(2)
$$\begin{aligned} x_1 &= b_{11}\lambda_{r+1} + b_{12}\lambda_{r+2} + \ldots + b_{1n-r}\lambda_n \\ x_2 &= b_{21}\lambda_{r+1} + b_{22}\lambda_{r+2} + \ldots + b_{2n-r}\lambda_n \\ &. \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \\ x_r &= b_{r1}\lambda_{r+1} + b_{r2}\lambda_{r+2} + \ldots + b_{r\,n-r}\lambda_n , \end{aligned}$$

в които коефициентите b_{kl} зависят само от елементите на матрицата $A = \left(a_{ij}\right)$. Ако оставим $\lambda_{r+1}, \ldots., \lambda_n$ да пробягват независимо R, наредената n -торка $\left(x_1, ..., x_r, \lambda_{r+1}, ..., \lambda_n\right)$, в която x_i са пресметнати по формулите (3), ще пробягва всички решения на хомогенната система и само тях. Тъй като броят на независимите параметри е $n-r$, естествено е да очакваме, че подпространството от решенията също има размерност $n-r$. Ще покажем, че това наистина е така. За целта нека разгледаме решенията

$$
\begin{aligned}
X_1 &= \left(b_{11}, b_{21}, ..., b_{r1}, 1, 0, ..., 0\right) \\
X_2 &= \left(b_{12}, b_{22}, ..., b_{r2}, 0, 1, ..., 0\right) \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
X_{n-r} &= \left(b_{1n-r}, b_{2n-r}, ..., b_{rn-r}, 0, 0, ..., 1\right).
\end{aligned}
$$

Те са линейно независими, защото матрицата, образувана от компонентите им, е от тип $(n-r)$ и съдържа ненулевия минор

$$
\begin{vmatrix}
1 & 0 & 0 & \dots & 0 \\
0 & 1 & 0 & \dots & 0 \\
. \; . & . \; . & . \; . & . \; . & . \; . \\
0 & 0 & 0 & \dots & 1
\end{vmatrix} = 1
$$

от ред $n-r$. Ако сега $X = \left(x_1, x_2, ..., x_n\right)$ е произволно решение на системата, то като вземем предвид формулите (3), получаваме, че

$$
X = x_{r+1}X_1 + x_{r+2}X_2 + ... + x_n X_{n-r}.
$$

(Препоръчваме на читателя подробно да сравни съответните компоненти на решенията от двете страни на това равенство.) Следователно всяко решение е линейна комбинация на линейно независимите решения $X_1, \dots, X_{n-r}$, т.е. те наистина са базис на подпространството от решенията; размерността му е $n-r$. Доказахме

4. Теорема. *Ако хомогенна система линейни уравнения с* n *неизвестни има ранг* r *, то подпространството от решенията ѝ има* *размерност* $n-r$. (При $r<n$ всеки базис на това подпространство ще наричаме *фундаментална система от решения на хомогенната система*.)

Да се върнем отново на общата система линейни уравнения (1) от § 4. Ако положим $b_1 = \ldots = b_m = 0$, ще получим хомогенна система, за която ще казваме, че е съответна на дадената.

5. Теорема. *Нека* $X' = \left(x_1', \ldots, x_n'\right)$ *е произволно избрано и фиксирано решение на системата* (1) *от* § 4, *а* $X_0 = \left(x_1^{(0)}, \ldots, x_n^{(0)}\right)$ *е произволно решение на съответната хомогенна система. Тогава* $X' + X_0$ *е решение на системата (1). Всяко решение на* (1) *може да се представи във вида* $X' + X_0$ *при подходящо решение* X_0 *на съответната хомогенна система.*

Доказателство. Нека $X'' = \left(x_1'', \ldots, x_n''\right)$ *е още едно решение на системата* (1). *За всяко* $i = 1, \ldots, m$ *имаме*

$$\sum_{j=1}^{n} a_{ij} x_j' = b_i, \qquad \sum_{j=1}^{n} a_{ij} x_j'' = b_i, \qquad \sum_{j=1}^{n} a_{ij} x_j^{(0)} = 0.$$

От тези равенства получаваме

$$\sum_{j=1}^{n} a_{ij} \left(x_j' - x_j''\right) = 0, \qquad \sum_{j=1}^{n} a_{ij} \left(x_j' + x_j^{(0)}\right) = b_i.$$

Първото от тях означава, че $X' - X''$ е решение на хомогенната система, а второто - че $X' + X_0$ е решение на (1). Следователно, ако X е произволно решение на (1), то $X - X'$ е решение на хомогенната система и ако положим $X - X' = X_0$, получаваме $X = X' + X_0$.

Доказаната теорема разкрива структурата на множеството от решенията на произволна линейна система. Удобно е да се въведе специален термин.

6. Определение. Всяко непразно подмножество на R^n от вида

$$x + V = \left\{x + \upsilon \setminus \upsilon \in V\right\},$$

където $x \in \mathrm{R}^n$ е фиксиран вектор, а V е фиксирано подпространство на R^n, ще наричаме *линейно многообразие* в координатното пространство R^n. Размерността на V ще наричаме *размерност на многообразието*; V се нарича *направляващо* подпространство. Тъй като $0+V = V$, ясно е, че подпространството също е линейно многообразие.

Съгласно теоремата от т. 5 всяка съвместима система от линейни уравнения с n неизвестни "задава" линейно многообразие в R^n. Вярно е и обратното: всяко линейно многообразие може да бъде "зададено" с помощта на подходяща система линейни уравнения. Точният смисъл е следният.

7. Теорема. *Ако W е линейно многообразие в R^n, то съществува подходяща система от линейни уравнения с n неизвестни, за която множеството от решенията ѝ е точно W. Всяко подпространство на R^n съвпада с множеството от решенията на подходяща линейна хомогенна система.*

Доказателство. Най-напред ще покажем, че всяко подпространство $V \subset \mathrm{R}^n$ може да се зададе с хомогенна система линейни уравнения с n неизвестни. Ако $V = \{0\}$, достатъчно е да вземем системата $x_1=0$, $x_2=0$, ..., $x_n=0$. Ако $V = \mathrm{R}^n$, вземаме система, състояща се само от уравнението $0.x_1+0.x_2+...+0.x_n=0$. Нека по-нататък $\dim V = k$, $0<k<n$, и да изберем базис

$$
\begin{aligned}
e_1 &= \left(b_{11}, b_{12}, ..., b_{1n}\right)\\
e_2 &= \left(b_{21}, b_{22}, ..., b_{2n}\right)\\
&\dots\dots\dots\dots\dots\\
e_k &= \left(b_{k1}, b_{k2}, ..., b_{kn}\right)
\end{aligned}
$$

на V. Произволен вектор $x = \left(x_1, ..., x_n\right)$ лежи във V тогава и само тогава, когато е линейна комбинация на базисните вектори, т. е. когато

$$x = \lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_k e_k, \qquad \lambda_i \in \mathrm{R}.$$

(Когато λ_i, $i = 1, ..., k$, пробягват R, то x пробягва V.) По-подробно, имаме равенствата

$$
(4)\qquad
\begin{aligned}
&x_1 = b_{11}\lambda_1 + b_{21}\lambda_2 + \ldots + b_{k1}\lambda_k \\
&.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ . \\
&x_i = b_{1i}\lambda_1 + b_{2i}\lambda_2 + \ldots + b_{k\,i}\lambda_k \\
&.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ . \\
&x_n = b_{1n}\lambda_1 + b_{2n}\lambda_2 + \ldots + b_{kn}\lambda_k .
\end{aligned}
$$

Тъй като $e_1, \ldots, e_k$ са линейно независими, матрицата $B = \left(b_{ij}\right)$ от тип $k\ \ x\ \ n$ има ранг k, следователно съдържа ненулев минор от ред k. За нагледност и удобство ще предположим, че

$$
\begin{vmatrix} b_{11} & \ldots & b_{1k} \\ .\ . & .\ . & .\ . \\ b_{k1} & \ldots & b_{kk} \end{vmatrix} \neq 0.
$$

От първите k равенства в (4) можем да определим $\lambda_1, \ldots, \lambda_k$ (например по формулите на Крамер). Ще получим изрази от вида

$$
\begin{aligned}
&\lambda_1 = c_{11}x_1 + c_{12}x_2 + \ldots + c_{1k}x_k \\
&.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ . \\
&\lambda_k = c_{k1}x_1 + c_{k2}x_2 + \ldots + c_{kk}x_k .
\end{aligned}
$$

След заместване с тези изрази в последните $n-k$ равенства от (4) ще стигнем до уравнения от вида

$$
(5)\qquad
\begin{aligned}
&x_{k+1} = a_{11}x_1 \quad + a_{21}x_2 \quad + \ldots + a_{1k}x_k \\
&x_{k+2} = a_{21}x_1 \quad + a_{22}x_2 \quad + \ldots + a_{2k}x_k \\
&.\ . \\
&x_n \quad = a_{n-k\,1}x_1 + a_{n-k\,2}x_2 + \ldots + a_{n-k\;k}x_k .
\end{aligned}
$$

Нека $V\,'$ е подпространството от решенията на хомогенната система (5). Тя е следствие от (4), ето защо $V \subset V'$. От друга страна, броят на свободните неизвестни в (5) е k (свободни са $x_1, \ldots, x_k$), което според твърденията в т. 2 и т. 4 означава, че $\dim V\,' = k$. От включването $V \subset V'$ и $\dim V = \dim V\,'$

следва $V = V'$. Следователно подпространството V се задава с хомогенната система (5).

Нека сега $W = x' + V$ е произволно линейно многообразие в R^n, където $x' = \left(x_1', \ldots, x_n'\right) \in \mathrm{R}^n$, а $V \subset \mathrm{R}^n$ е подпространство. Задаваме V с помощта на хомогенна система от вида (5) и полагаме

$$b_i = \sum_{j=1}^{k} a_{ij} x_j' - x_{k+i}', \qquad i = 1, \ldots, n-k.$$

Твърдим, че $x' + V$ е множеството от решенията на системата

$$\begin{aligned} &a_{11}x_1 \quad + \; a_{12}x_2 \quad + \ldots + a_{1k}x_k - x_{k+1} = b_1 \\ &. \; . \\ &a_{n-k\,1}x_1 + a_{n-k\,2}x_2 + \ldots + a_{n-k\,k}x_k - x_n = b_{n-k}. \end{aligned}$$

Това следва от теоремата в т. 5, защото x' е решение на тази система, съответната й хомогенна система е (5), а подпространството от решенията й е V.

По-късно ще видим, че доказаната теорема допуска важна геометрична интерпретация: линейните многообразия са геометрични обекти от типа на прави и равнини, следователно те могат да се описват с помощта на системи линейни уравнения.

УПРАЖНЕНИЯ

1. Следните две операции ще наричаме *елементарни преобразувания на редовете* (сътв. на стълбовете) на дадена матрица: а) размяна на местата на два реда (стълба); б) умножаване на някой ред (стълб) с число и прибавянето му към друг ред (съотв. стълб). Докажете, че елементарните преобразувания, въпреки че променят матрицата, не променят ранга й.

2. Нека $A = \left(a_{ij}\right)$ е произволна матрица от тип $m \times n$. Ще казваме,че тя е *квазидиагонална*, ако $a_{ij} = 0$, винаги щом $i \neq j$. (При $m = n$ такива матрици нарекохме диагонални.) Докажете, че рангът на квазидиагонална матрица е равен на броя на ненулевите елементи, които тя съдържа.

3. Докажете, че с помощта на елементарни преобразования всяка матрица може да се трансформира в квазидиагонална. (Задачи 1-3 всъщност обосновават алгоритъм за пресмятане ранга на матрица.)

4. Нека векторите $a_1, a_2, \ldots, a_k$ са линейно независими и

$$b_i = \lambda_{i1}a_1 + \lambda_{i2}a_2 + \ldots + \lambda_{ik}a_k, \quad i = 1, 2, \ldots, m,$$

а $A = \left(\lambda_{ij}\right)$ е съответната матрица от коефициентите λ_{ij}. Докажете, че рангът на системата вектори $b_1, \ldots, b_m$ е равен на ранга на матрицата A. Ако векторите $a_1, \ldots, a_k$ не са непременно линейно независими, покажете, че рангът на системата вектори $b_1, \ldots, b_m$ не надминава ранга на матрицата A.

5. Ако някаква матрица има ранг r, покажете, че всеки минор, принадлежащ на r линейно независими реда и r линейно независими стълба, е различен от нула. (Упътване. Нека дадената матрица е A и да разгледаме частния случай, когато в нея първите r реда и първите r стълба са линейно независими, но минорът Δ от ред r, който те съдържат, е равен на нула. Нека A^* е матрицата, образувана от първите r реда на A. Тя също има ранг r. Тъй като $r\left(A\right) = r$, то всеки стълб на A е линейна комбинация на първите ѝ r стълба, следователно същото важи и за стълбовете на A^*. Тъй като $\Delta = 0$, първите r стълба на A^* са линейно зависими, следователно системата вектори-стълбове в A^* има ранг по-малък от r, което противоречи на $r\left(A^*\right) = r$. Сведете общия случай към разгледания.)

6. Нека $A = \left(a_{ij}\right)$ е квадратна матрица от ред n. Матрицата $A^* = \left(A_{ji}\right)$, където A_{ji} е адюнгираното количество на елемента a_{ij}, се нарича *адюнгирана* на матрицата A. Докажете, че: а) ако $r\left(A\right) \le n-2$, то $r\left(A^*\right) = 0$; б) ако $r\left(A\right) = n-1$, то $r\left(A^*\right) = 1$; в) ако $r\left(A\right) = n$, то $r\left(A^*\right) = n$. (Упътване. При $r\left(A\right) = n-1$, покажете, че поне едно от адюнгираните количества A_{ji} е различно от нула. Убедете се, че всеки от стълбовете на A^* е решение на хомогенната система

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = 0 \end{array}$$

Тъй като $r\left(A\right)=n-1$, подпространството от решенията на тази система има размерност 1, следователно всяко решение е пропорционално на което и да е ненулево решение. Изведете от тук, че $r\left(A^*\right)=1$.)

7. Нека L е координатното пространство R^n или C^n, а V_1 и V_2 са негови подпространства. Докажете, че $l_1+V_1=l_2+V_2$ за някакви вектори $l_1, l_2 \in L$ тогава и само тогава, когато $V_1=V_2$ и $l_1-l_2 \in V_1$.

Г л а в а 3

ЛИНЕЙНИ ИЗОБРАЖЕНИЯ И МАТРИЦИ

Навсякъде в тази глава "крайномерно линейно (векторно) пространство" означава или реалното координатно пространство R^n, или комплексното координатно пространство C^n.

§ 1. Линейни изображения

1. Нека от произволна система линейни уравнения вземем само левите им страни (вж. гл. 1, § 1, (1) и ги разгледаме като функции на $x_1, \dots, x_n$, където x_i пробягват R (или C). Имаме

$$
(1)\qquad
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= y_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= y_2 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= y_m .
\end{aligned}
$$

Разглеждани в съвкупност, равенствата (1) задават изображение $f : \mathrm{R}^n \to \mathrm{R}^m$ (съотв. $f : \mathrm{C}^n \to \mathrm{C}^m$), при което на наредената n-торка $\left(x_1, \ldots, x_n\right)$ се съпоставя наредената m-торка $\left(y_1, \ldots, y_m\right)$. Ясно е, че въпросите, дали изображението f е сюрективно, инективно или биективно, могат да се преформулират като въпроси за съответната система линейни уравнения, например дали f е сюрективно, е все едно дали както и да изберем числата $y_1 = b_1,\ y_2 = b_2,\ \ldots,\ y_m = b_m$ може да се намерят числа $x_1, \ldots, x_n$, за които да бъдат изпълнени равенствата (1). Съобразете самостоятелно, че f е инективно точно тогава, когато хомогенната система

$$
(2)\qquad
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0
\end{aligned}
$$

няма ненулево решение.

Разглежданото изображение f е зададено по много специален начин, което влече и някои негови специални свойства. Лесно е да се провери например, че $f\left(l_1+l_2\right)=f\left(l_1\right)+f\left(l_2\right)$ и $f\left(\lambda l\right)=\lambda f\left(l\right)$ каквито и да са векторите l_1, l_2, l от R^n (или от C^n) и скаларът λ. Както ще видим по-късно, тези две свойства се оказват характеристични за изображението, дефинирано с равенствата (1). Засега ще ги постулираме в следното

2. Определение. Нека $L=\mathrm{R}^n$ или C^n, а $M=\mathrm{R}^m$ (съотв. C^m). Ще казваме, че изображението $f: L\to M$ е *линейно изображение на линейното пространство* L *в линейното пространство* M, ако:

а) $f\left(l_1+l_2\right)=f\left(l_1\right)+f\left(l_2\right)$ за всички $l_1, l_2 \in L$;

б) $f\left(\lambda l\right)=\lambda f\left(l\right)$ за всеки вектор $l\in L$ и за всяко $\lambda\in\mathrm{R}$ (или C).

Всяко линейно изображение на линейното пространство L в себе си ще наричаме *линеен оператор, дефиниран в* L.

3. Примери и коментари. а) Изображението f, зададено с равенствата (1) и разгледано в т. 1, е линейно.

б) Нека изображението $f: L\to M$ е дефинирано с равенството $f\left(l\right)=0$ за всяко $l\in L$. То очевидно е линейно. Нарича се *нулево изображение*.

в) Нека изображението $f: L\to L$ е дефинирано с равенството $f\left(l\right)=l$ за всяко $l\in L$. То също е линейно. Това е така нареченият *единичен оператор* (*идентитет*).

г) Нека $l=\left(x_1, \ldots, x_i, \ldots, x_n\right)$ и f_i е изображението на L в себе си, определено с равенството $f_i\left(l\right)=\left(0, \ldots, x_i, \ldots, 0\right)$ за всеки вектор $l\in L$ при фиксирано i. Този линеен оператор е частен случай на така наречените *оператори на проектиране* (*проектори*).

Ще забележим, че $f\left(0\right)=0$ за всяко линейно изображение f. Това следва от равенството $f\left(0\right)=f\left(0+0\right)=f\left(0\right)+f\left(0\right)$.

А сега да се опитаме да построим възможно най-общ пример на линейно изображение. Нека за целта изберем произволен базис $e_1, \ldots, e_n$ на L, а f е произволно линейно изображение на L в M. Тъй като всеки вектор $l\in L$ се представя еднозначно във вида

$$l=x_1e_1+x_2e_2+\ldots+x_ne_n,$$

то

(3) $$f(l) = x_1 f(e_1) + x_2 f(e_2) + \ldots + x_n f(e_n).$$

Следователно, за да "изчисляваме" образа $f(l)$ на даден вектор l, е достатъчно да знаем образите $f(e_1), \ldots, f(e_n)$ на базисните вектори. Обратно, ако $e_1, \ldots, e_n$ е базис на L, как трябва да се изберат векторите $f(e_1), \ldots, f(e_n)$ от M, че изображението $f: L \to M$, зададено с формулата (3), да се окаже линейно? Следващото твърдение показва, че те може да се избират без каквито и да било ограничения.

4. Твърдение. *Нека* $e_1, \ldots, e_n$ *е базис на* L ($= \mathrm{R}^n$ или C^n), а $g_1, \ldots, g_n \in M$ ($= \mathrm{R}^m$ съотв. C^m). *Съществува, и то единствено, линейно изображение* $f: L \to M$, *за което* $f(e_i) = g_i$, $i = 1, \ldots, n$.

Доказателство. Нека f съществува. Предишното разсъждение показва, че f ще се задава с формулата (3) при $f(e_i) = g_i$, което доказва единствеността.

Да дефинираме изображението $f: L \to M$ с равенството

$$f(x_1 e_1 + \ldots + x_n e_n) = x_1 g_1 + \ldots + x_n g_n$$

и да покажем, че то е линейно изображение. Наистина, ако $l = \sum_{i=1}^{n} x_i e_i$, то

$$f(\lambda l) = f\left(\sum_{i=1}^{n} \lambda x_i e_i\right) = \sum_{i=1}^{n} \lambda x_i g_i = \lambda \sum_{i=1}^{n} x_i g_i = \lambda f(l).$$

Нека $l_1 = \sum_{i=1}^{n} x_i' e_i$, $\quad l_2 = \sum_{i=1}^{n} x_i'' e_i$. Тогава

$$f(l_1 + l_2) = f\left(\sum_{i=1}^{n} (x_i' + x_i'') e_i\right) = \sum_{i=1}^{n} (x_i' + x_i'') g_i = \sum_{i=1}^{n} x_i' g_i + \sum_{i=1}^{n} x_i'' g_i = \\ = f(l_1) + f(l_2).$$

С това съществуването на исканото линейно изображение е доказано.

5. Матрица на линейно изображение. Нека отново $L = R^n$ или C^n, а $M = \mathrm{R}^m$ (съотв. C^m). Предполагаме, че $e_1, \ldots, e_n$ и $e_1^*, \ldots, e_m^*$ са произволно избрани базиси съответно на L и M, а $f: L \to M$ е линейно изображение на L в M. Тъй като $f\left(e_i\right)$, $i = 1, \dots, n$, са вектори от M, те се изразяват чрез разглеждания базис на M, т.е.

$$
\begin{aligned}
& f\left(e_1\right) = a_{11}e_1^* + a_{21}e_2^* + \ldots + a_{m1}e_m^* \\
& f\left(e_2\right) = a_{12}e_1^* + a_{22}e_2^* + \ldots + a_{m2}e_m^* \\
& \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& f\left(e_n\right) = a_{1n}e_1^* + a_{2n}e_2^* + \ldots + a_{mn}e_m^*.
\end{aligned} \tag{4}
$$

Тъй като векторите $f\left(e_i\right)$ се определят еднозначно чрез координатите си $\left(a_{1i}, a_{2i}, \ldots, a_{mi}\right)$ спрямо базиса $e_1^*, \dots, e_m^*$, а според твърдението в т.4 f се определя еднозначно от $f\left(e_i\right)$, $i = 1, \dots, n$, излиза, че *линейното изображение f се определя еднозначно от матрицата*

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

Тя, разбира се, зависи както от избрания базис на L, така и от избрания базис на M. Нарича се *матрица на линейното изображение f спрямо избраните базиси на L и M*. Специално подчертаваме, че *стълбовете на A се състоят от координатите на векторите* $f\left(e_1\right), \dots, f\left(e_n\right)$ спрямо избрания базис на пространството M. Матрицата A има тип $m \times n$, т.е. $\dim M \times \dim L$. Ако $f: L \to L$ е линеен оператор, в качеството на $e_1^*, \dots, e_n^*$ задължително се избира същият базис $e_1, \dots, e_n$ и се говори за *матрица на оператора f спрямо избрания базис на L*. В този случай A е квадратна матрица от ред n.

Нека $l = x_1e_1 + x_2e_2 + \ldots + x_ne_n \in L$ и

$$f\left(l\right) = y_1e_1^* + y_2e_2^* + \ldots + y_me_m^*.$$

Като заместим равенствата (4) в (3) и сравним съответните коефициенти от лявата и дясната страна (тъй като координатите на всеки вектор са еднозначно определени!), ще стигнем до равенствата (1). Следователно в *термините на координатите всяко линейно изображение на едно крайномерно векторно пространство в друго крайномерно пространство се задава с равенства от вида* (1), *където* $A=\left(a_{ij}\right)$ е матрицата на изображението спрямо избраните базиси.

§ 2. Ядро и образ на линейно изображение

1. Определение. Нека $L=\mathrm{R}^n$, $M=\mathrm{R}^m$, или $L=\mathrm{C}^n$ и $M=\mathrm{C}^m$, а $f:L\to M$ е линейно изображение. Множеството

$$\mathrm{Ker} f=\left\{l\in L \mid f\left(l\right)=0\right\}$$

ще наричаме *ядро* на линейното изображение f; *образ* на f ще наричаме множеството

$$\mathrm{Im} f=\left\{f\left(l\right) \mid l\in L\right\}.$$

Тъй като винаги $f\left(0\right)=0$, то $\mathrm{Ker}\, f$ и $\mathrm{Im}\, f$ съдържат поне нулевия вектор съответно на L и M. Ще покажем, че *ядрото и образът са подпространства съответно на* L и M. Наистина, нека $l_1,l_2\in\mathrm{Ker}\, f$. Тогава $f\left(\lambda l_1+\mu l_2\right)=\lambda f\left(l_1\right)+\mu f\left(l_2\right)=\lambda.0+\mu.0=0$, т.е. $\lambda.l_1+\mu.l_2\in\mathrm{Ker}\, f$, което означава, че ядрото е подпространство. Доказателството за образа оставяме за упражнение.

Ако $\mathrm{Ker} f=\left\{0\right\}$, ще казваме, че ядрото е *тривиално*.

2. Твърдение. *Равенството* $f\left(l_1\right)=f\left(l_2\right)$, *където* $l_1,l_2\in L$, *е в сила тогава и само тогава, когато* $l_1-l_2\in\mathrm{Ker}\, f$. *Линейното изображение* f *е инективно точно тогава, когато неговото ядро* $\mathrm{Ker}\, f$ *е тривиално*.

Доказателство. Действително, равенството $f\left(l_1\right)=f\left(l_2\right)$ е еквивалентно на $f\left(l_1-l_2\right)=0$, а последното е в сила точно тогава, когато $l_1-l_2\in\mathrm{Ker}\, f$. Останалото вече е очевидно.

3. Определение. Размерността на ядрото $\operatorname{Ker} f$ на линейното изображение f ще наричаме *дефект* на изображението, а размерността на образа му $\operatorname{Im} f$ ще наричаме негов *ранг*.

4. Теорема. *При означенията от т.* 1 *нека* $f: L \to M$ *е линейно изображение. Тогава*

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim L.$$

Доказателство. Нека $\mathrm{k} = \dim \operatorname{Ker} f$. Ако $\mathrm{k} > 0$, избираме базис $e_1, e_2, \ldots, e_k$ на подпространството $\operatorname{Ker} f$. Съгласно теоремата от гл. 2, § 2, т. 6 той може да бъде допълнен до базис $e_1, \ldots, e_k, \ldots, e_n$ на цялото пространство L ($n = \dim L$). Ще покажем, че $f(e_{k+1}), \ldots, f(e_n)$ образуват базис на подпространството $\operatorname{Im} f$, с което теоремата ще бъде доказана. Действително, нека $l \in L$ и $l = x_1 e_1 + \ldots + x_k e_k + x_{k+1} e_{k+1} + \ldots + x_n e_n$. Тъй като $f(e_i) = 0$, $i = 1, \ldots, k$, то

$$f(l) = x_{k+1} f(e_{k+1}) + \ldots + x_n f(e_n),$$

което показва, че векторите $f(e_{k+1}), \ldots, f(e_n)$ наистина пораждат $\operatorname{Im} f$. Остава да се убедим, че те са линейно независими. Наистина, ако

$$\sum_{i=k+1}^{n} \lambda_i f(e_i) = 0, \qquad \text{то} \qquad f\left(\sum_{i=k+1}^{n} \lambda_i e_i\right) = 0$$

и следователно векторът $l_0 = \lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n$ принадлежи на ядрото $\operatorname{Ker} f$. Тогава l_0 се изразява чрез базиса на $\operatorname{Ker} f$: $l_0 = \mu_1 e_1 + \ldots + \mu_k e_k$. Сравнявайки двете представяния на l_0 (предвид линейната независимост на $e_1, \ldots, e_n$), получаваме, че $\mu_1 = \ldots = \mu_k = \lambda_{k+1} = \ldots = \lambda_n = 0$.

Проведеното разсъждение лесно се приспособява и за случая $\mathrm{k} = 0$, т.е. когато ядрото $\operatorname{Ker} f$ е тривиално. Сега базисът на ядрото е празното множество от вектори и направо избираме базис $e_1, \ldots, e_n$ на цялото пространство L, след което проверяваме, че векторите $f(e_1), \ldots, f(e_n)$ са базис на подпространството $\operatorname{Im} f$. Детайлите пропускаме.

Ще изложим още едно доказателство на теоремата. За целта избираме базиси $e_1, \ldots, e_n$ и $e_1^*, \ldots, e_m^*$ съответно на L и M и нека

$A = (a_{ij})$ е матрицата на линейното изображение f спрямо тези базиси. Системата вектори $f(e_1), ..., f(e_n)$ поражда $\operatorname{Im} f$ и $\dim \operatorname{Im} f$ съвпада с ранга на тази система. От друга страна, стълбовете на матрицата A са координатите на векторите $f(e_1), ..., f(e_n)$. Следователно рангът на въпросната система вектори съвпада с ранга $r(A)$ на матрицата A. И така, рангът $\dim \operatorname{Im} f$ на изображението f е равен на ранга $r(A)$ на матрицата му. Нека $r(A) = A$. Ядрото на изображението f се състои от векторите с координати $(x_1, ..., x_n)$, удовлетворяващи хомогенната система (2) от § 1, т. 1. Съгласно теоремата от гл. 2, § 5, т. 4 подпространството от решенията й има размерност $n-r$. Така $\dim \operatorname{Ker} f = n - r$. Очевидно $\dim\operatorname{Ker} f + \dim\operatorname{Im} f = (n-r) + r = = \dim L$.

Всъщност доказахме нещо повече:

5. Твърдение. *Нека* $L = \mathrm{R}^n$, $M = \mathrm{R}^m$, *или* $L = \mathrm{C}^n$ *и* $M = \mathrm{C}^m$, а $f: L \to M$ *е линейно изображение. Нека* A *е матрицата на* f *спрямо произволно избрани базиси на* L *и* M. *Тогава рангът на линейното изображение* f *е равен на ранга* $r(A)$ *на матрицата му. Ако този ранг е* r*, то дефектът на* f *е* $n - r = \dim L - r(A)$.

С правилото, по което се сменя матрицата на линейното изображение при смяна на базисите, ще се запознаем по-късно. Още сега ще забележим обаче, че рангът й е инвариантен при смяна на базисите, защото съвпада с ранга на изображението, който пък не зависи от никакви базиси.

§ 3. Действия с линейни изображения и операции с матрици

Навсякъде в този параграф $\mathcal{K}$ ще означава или множеството R на реалните числа, или множеството C на комплексните числа, а $L = \mathcal{K}^n$ и $M = \mathcal{K}^m$ са координатни линейни пространства.

1. Сума на линейни изображения и умножаване на линейно изображение с число. Нека $f: L \to M$ и $g: L \to M$ са линейни изображения. Дефинираме ново изображение $f + g: L \to M$ с правилото

$(f+g)(l)=f(l)+g(l)$ за всяко $l\in L$. Ще го наричаме *сума* на f и g. От определението веднага следва, че

$$(f+g)(l_1+l_2)=f(l_1+l_2)+g(l_1+l_2)=f(l_1)+f(l_2)+g(l_1)+g(l_2)=$$
$$=(f+g)(l_1)+(f+g)(l_2),$$

$$(f+g)(\lambda l)=f(\lambda l)+g(\lambda l)=\lambda f(l)+\lambda g(l)=\lambda\big(f(l)+g(l)\big)=\lambda(f+g)(l),$$

което означава, че *сумата на линейни изображения също е линейно изображение*.

По-нататък за скалара $\lambda\in\mathcal{K}$ и линейното изображение $f:L\to M$ дефинираме изображение $\lambda f:L\to M$ с правилото $(\lambda f)(l)=\lambda f(l)$, което ще наричаме *произведение на линейното изображение* f *с числото* λ. Като пропускаме тривиалната проверка, ще подчертаем, че λf *също е линейно изображение*.

Да запазим означенията от § 1, т. 5. Там видяхме, че ако в L и M изберем базиси, то на всяко линейно изображение може да се съпостави еднозначно определена матрица. Нека сега $A=(a_{ij})$ и $B=(b_{ij})$ са матриците на f и g спрямо избраните базиси. И двете матрици са от един и същ тип: $m\times n$. Не е трудно да намерим матриците на изображенията λf и $f+g$.

Тъй като $f(e_j)=a_{1j}e_1^*+a_{2j}e_2^*+\ldots+a_{mj}e_m^*$, то $\lambda f(e_j)=\lambda a_{1j}e_1^*+\lambda a_{2j}e_2^*+\ldots+\lambda a_{mj}e_m^*$, което означава, че елементите на j-тия стълб в матрицата на изображението λf са елементите на j-тия стълб на A (матрицата на f), умножени с λ. Понеже $j=1,2,\ldots,n$, получаваме, че матрицата на λf се получава след като всички елементи в матрицата на f се умножат с λ. За изображението $f+g$ имаме

$$(f+g)(e_j)=f(e_j)+g(e_j)=\left(a_{1j}e_1^*+a_{2j}e_2^*+\ldots+a_{mj}e_m^*\right)+\left(b_{1j}e_1^*+b_{2j}e_2^*+\ldots+b_{mj}e_m^*\right)=$$

$$=\left(a_{1j}+b_{1j}\right)e_1^*+\left(a_{2j}+b_{2j}\right)e_2^*+\ldots+\left(a_{mj}+b_{mj}\right)e_m^*,$$

което показва, че неговата матрица $C = \left(c_{ij}\right)$ се получава от матриците на f и g, като се съберат съответните им елементи, т.е. $c_{ij} = a_{ij} + b_{ij}$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

Горните разсъждения мотивират следното

2. Определение. Нека A и B са произволни матрици от еднакъв тип $m \times n$,

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

Тяхна *сума* (ще я бележим с $A + B$) ще наричаме матрицата

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

Под *произведение на матрицата A с числото λ* (означение: λA) ще разбираме матрицата

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

Ясно е, че *ако линейните изображения f и g имат матрици съответно A и B, то матрицата на $f + g$ е $A + B$, а на λf е λA.*

Почти очевидно е, че за всеки три матрици A, B, C от един и същи тип $m \times n$ и за произволни числа λ и μ са в сила следните свойства:

1) $A + B = B + A$ (комутативност);

2) $\left(A+B\right)+C=A+\left(B+C\right)$ (асоциативност);

3) да означим с 0 матрицата от тип $m\times n$, в която всички елементи са нули (нулевата матрица); тогава $A+0=A$ за всяка матрица A от типа $m\times n$;

4) нека $-A=\left(-1\right)A$; тогава $A+\left(-A\right)=0$;

5) $1.A=A$ за всяка матрица A;

6) $\lambda\left(A+B\right)=\lambda A+\lambda B$;

7) $\left(\lambda+\mu\right)A=\lambda A+\mu A$;

8) $\left(\lambda\mu\right)A=\lambda\left(\mu A\right)$.

Проверката пропускаме. Засега ще отбележим само, че 1) - 8) са същите "правила за смятане", както в R^n (или в C^n).

3. Композиция на линейни изображения и произведение на матрици. Нека $L_1=\mathcal{K}^p$, $L_2=\mathcal{K}^n$ и $L_3=\mathcal{K}^m$ са три координатни линейни пространства и $g\colon L_1\to L_2$, $f\colon L_2\to L_3$ са линейни изображения. Малко по-нагледно:

$$L_1\xrightarrow{\;g\;}L_2\xrightarrow{\;f\;}L_3.$$

Изображението $f\circ g\colon L_1\to L_3$, дефинирано с равенството

$$\left(f\circ g\right)\left(l\right)=f\left(g\left(l\right)\right)\qquad \text{за всяко } l\in L_1,$$

се нарича *композиция на изображенията* f *и* g. Проверката, че *композицията на линейни изображения е линейно изображение*, е чисто техническа:

$$f\circ g\left(l+l'\right)=f\left(g\left(l+l'\right)\right)=f\left(g\left(l\right)+g\left(l'\right)\right)=f\left(g\left(l\right)\right)+f\left(g\left(l'\right)\right)=$$
$$=f\circ g\left(l\right)+f\circ g\left(l'\right),$$
$$f\circ g\left(\lambda l\right)=f\left(g\left(\lambda l\right)\right)=f\left(\lambda g\left(l\right)\right)=\lambda f\left(g\left(l\right)\right)=\lambda f\circ g\left(l\right).$$

По-нататък нека $e_1,\dots,e_p$; $e'_1,\dots,e'_n$ и $e''_1,\dots,e''_m$ са базиси съответно на L_1, L_2 и L_3. Нека $A=\left(a_{ik}\right)$ е матрицата на $f\colon L_2\to L_3$, а $B=\left(b_{kj}\right)$ - матрицата на $g\colon L_1\to L_2$ спрямо избраните базиси. Не е трудно да намерим матрицата C на $f\circ g$. Ще напомним, че според § 1, т. 5 A, B, C са съответно от тип $m\times n$, $n\times p$, $m\times p$. Имаме

$$f\circ g\left(e_j\right)=f\left(g\left(e_j\right)\right)=f\left(\sum_{k=1}^{n}b_{kj}e'_k\right)=\sum_{k=1}^{n}b_{kj}f\left(e'_k\right)=\sum_{k=1}^{n}b_{kj}\left(\sum_{i=1}^{m}a_{ik}e''_i\right)=$$

$$=\sum_{i=1}^{m}\sum_{k=1}^{n}a_{ik}b_{kj}e''_i,$$

което след полагането

$$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}\,;\qquad i=1,\dots,m,\ j=1,\dots,p,$$

добива вида

$$f\circ g\left(e_j\right)=\sum_{i=1}^{m}c_{ij}e''_i.$$

Последното равенство показва, че матрицата на $f\circ g$ е $C=\left(c_{ij}\right)$. Този резултат подсказва следното

4. Определение. Нека

$$A=\begin{pmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ .. & .. & .. & ..\\ a_{m1} & a_{m2} & \dots & a_{mn}\end{pmatrix},\quad B=\begin{pmatrix}b_{11} & b_{12} & \dots & b_{1p}\\ b_{21} & b_{22} & \dots & b_{2p}\\ .. & .. & .. & ..\\ b_{n1} & b_{n2} & \dots & b_{np}\end{pmatrix}$$

са произволни матрици от тип съответно $m\times n$ и $n\times p$. Тяхно *произведение* (означение: AB) ще наричаме матрицата

$$AB=\begin{pmatrix}c_{11} & c_{12} & \dots & c_{1p}\\ c_{21} & c_{22} & \dots & c_{2p}\\ .. & .. & .. & ..\\ c_{m1} & c_{m2} & \dots & c_{mp}\end{pmatrix}$$

от тип $m\times p$, където

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\dots+a_{in}b_{nj},\qquad i=1,\dots,m,\quad j=1,\dots,p.$$

Последната формула се помни лесно: с елементите на i -тия ред на A умножаваме съответните елементи на j-тия стълб на B и после събираме произведенията; накратко: "умножаваме ред по стълб". Обръщаме внимание, че *произведението* AB *е дефинирано само ако броят на стълбовете на* A *съвпада с броя на редовете на* B. В този смисъл е възможно AB да е дефинирано, но BA да не е. Но дори когато AB и BA са едновременно дефинирани, то изобщо казано $AB \neq BA$. Например

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

При така въведеното умножение имаме, че *матрицата на композицията на две линейни изображения е произведение на техните матрици*.

Ще отбележим основните свойства на операцията умножение на матрици:

а) *Умножението на матрици е асоциативно*, т.е. $(AB)C = A(BC)$. По-прецизно казано, ако произведението $(AB)C$ е дефинирано, то $A(BC)$ също е дефинирано и те са равни. Същата бележка важи и за следващите по-долу свойства.

Действително, за да бъде дефинирано произведението $(AB)C$ е необходимо и достатъчно матриците $A = (a_{ik})$, $B = (b_{kl})$, $C = (c_{lj})$ да бъдат от тип съответно $m \times n$, $n \times p$, $p \times q$, което показва, че тогава произведението $A(BC)$ също ще бъде дефинирано и всяка от матриците $(AB)C$ и $A(BC)$ ще бъде от тип $m \times q$. Елементите от i -тия ред в матрицата AB имат вида

$$a_{i1}b_{1l} + a_{i2}b_{2l} + \ldots + a_{in}b_{nl} = \sum_{k=1}^{n} a_{ik}b_{kl}, \qquad l = 1, 2, \ldots, p.$$

Умножавайки i -тия ред на матрицата AB с j-тия стълб на C, за елемента d_{ij} в матрицата $D = (d_{ij}) = (AB)C$ получаваме

$$d_{ij} = \sum_{l=1}^{p}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj}.$$

Като сменим реда на сумирането имаме

$$d_{ij} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{p} b_{kl} c_{lj} \right).$$

Тук сумата в скобите (сумата по l) е всъщност елементът с координати (k, j) в матрицата BC, а цялата дясна страна е произведението на i -тия ред на матрицата A с j-тия стълб на матрицата BC. Следователно $(AB)C = A(BC)$.

Читателят би могъл да докаже същото и без пресмятания, стига да попълни детайлите в следното разсъждение. От елементарната теория на множествата знаем, че за всеки три изображения f, g, h на множества, за които композицията $(f \circ g) \circ h$ е дефинирана, $(f \circ g) \circ h = f \circ (g \circ h)$ (композицията на изображения е асоциативна). По-нататък нека A, B, C са произволни матрици, за които произведението $(AB)C$ е дефинирано. Избираме координатни линейни пространства L_1, L_2, L_3, L_4 с подходящи размерности и дефинираме линейни изображения $f: L_3 \to L_4$, $g: L_2 \to L_3$, $h: L_1 \to L_2$ така, че матриците им да бъдат съответно A, B, C (съгласно §1, т.4 и т.5 това може да се направи). Тогава матрицата на $(f \circ g) \circ h$ ще бъде $(AB)C$, а на $f \circ (g \circ h)$ - съответно $A(BC)$. Тъй като равните линейни изображения имат равни матрици, то $(AB)C = A(BC)$.

б) *Събирането и умножението са свързани с дистрибутивните закони*:

$$A(B+C) = AB + BC, \quad (A+B)C = AC + BC.$$

Доказателството е елементарна проверка и го пропускаме.

в) За всеки две матрици A, B, за които произведението AB е дефинирано, и за всяко число λ е в сила

$$\lambda(AB) = (\lambda A)B = A(\lambda B).$$

И тук проверката е непосредствена.

г) Нека E_n е квадратна матрица от ред n, в която всички елементи по главния диагонал са единици, а елементите извън него са нули. Директно се проверява, че за всяка квадратна матрица A от ред n е в сила

$$AE_n = E_n A = A\,.$$

Матрицата E_n ще наричаме *единична матрица от ред* n. Ако няма опасност от недоразумения, ще изпускаме индекса n.

д) Ако A, B са квадратни матрици от ред n, то

$$\det\left(AB\right) = \det A . \det B\,.$$

Това равенство е непосредствено следствие от определението за умножение на матрици и от теоремата за умножаване на детерминанти (гл. 1, § 8, т. 2).

е) $\left(AB\right)^t = B^t A^t$ (A^t е транспонираната матрица на A). Наистина $\left(j,i\right)$-тият елемент на матрицата $\left(AB\right)^t$ е

$$\sum_k a_{ik} b_{kj} = \sum_k b_{kj} a_{ik}\,.$$

Дясната страна в това равенство е $\left(j,i\right)$-тият елемент на матрицата $B^t A^t$.

5. Обратна матрица. Нека A е квадратна матрица от ред n. Търсим матрица X, за която да бъде изпълнено $AX = XA = E$ (единичната матрица от ред n). Очевидно, ако матрицата X съществува, то тя също трябва да бъде квадратна от ред n. Ще покажем, че ако X съществува, то тя е единствена и $\det A \neq 0$, т.е. A е неособена. Действително, нека матриците X_1 и X_2 удовлетворяват условията $AX_1 = X_1 A = E$, $AX_2 = X_2 A = E$. Тъй като умножението на матрици е асоциативно, имаме $\left(X_1 A\right) X_2 = X_1 \left(A X_2\right)$, т.е. $EX_2 = X_1 E$ и $X_1 = X_2$. Освен това от равенството $AX = E$ имаме $\det A . \det X = 1$, откъдето $\det A \neq 0$. Ще покажем, че последното условие е достатъчно за съществуването на матрицата X. Нека

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

$\det A \neq 0$ и да образуваме матрицата

$$A^{-1} = \frac{1}{\det A} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & a_{22} & \dots & A_{n2} \\ .. & .. & .. & .. \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}, \tag{5}$$

където A_{ij} са адюнгираните количества на елементите a_{ij} в детерминантата на A. Непосредствено се проверява, че $AA^{-1} = A^{-1}A = E$ - достатъчно е да се използват формулите от гл. 1, § 7, т. 3 и т. 4.

Ако за матрицата A съществува матрица X, за която $AX = XA = E$, ще казваме, че A е *обратима матрица*, а единствената матрица X с това свойство ще наричаме *обратна матрица* на A. Нашите разсъждения показват, че квадратната матрица A е обратима тогава и само тогава, когато A е неособена; A^{-1} се дава с формулата (5).

От определението на обратна матрица се вижда, че ако X е обратна на матрицата A, то A е обратна на X, т.е.

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

Оставяме на читателя да съобрази, че *произведението на обратими матрици е обратима матрица*, и ако $A_1, A_2, \dots, A_k$ са обратими матрици от ред n, то

$$\left(A_1A_2\dots A_k\right)^{-1} = A_k^{-1}A_{k-1}^{-1}\dots A_1^{-1} .$$

Понятието обратна матрица е естествено свързано със следната задача. Нека $\mathcal{K} = \mathrm{R}$ или $\mathcal{K} = \mathrm{C}$, а $L = \mathcal{K}^n$ е съответното координатно линейно пространство и $f : L \to L$ е линеен оператор. Кога операторът f е биективно изображение?

Да предположим, че операторът $f : L \to L$ е биективно изображение и нека $f^{-1}: L \to L$ е обратното изображение. По определение на обратно изображение имаме $f \circ f^{-1}\left(l\right) = f^{-1} \circ f\left(l\right) = l$ за всеки вектор $l \in L$. Ще забележим, че f^{-1} също е линеен оператор. Действително, нека $l_1', l_2' \in L$.

Тъй като f е сюрективно изображение, то $l_1' = f(l_1)$, $l_2' = f(l_2)$ за подходящи вектори $l_1, l_2 \in L$. Очевидно $f^{-1}(l_1') = l_1$, $f^{-1}(l_2') = l_2$. Имаме

$$f^{-1}(l_1' + l_2') = f^{-1}(f(l_1) + f(l_2)) = f^{-1}(f(l_1 + l_2)) =$$
$$= f^{-1} \circ f(l_1 + l_2) = l_1 + l_2 = f^{-1}(l_1') + f^{-1}(l_2').$$

Аналогично

$$f^{-1}(\lambda l_1') = f^{-1}(\lambda f(l_1)) = f^{-1}(f(\lambda l_1)) = f^{-1} \circ f(\lambda l_1) = \lambda l_1 = \lambda f^{-1}(l_1').$$

Да изберем базис на L и нека A и X са матрици съответно на f и f^{-1}. Имаме равенствата $f \circ f^{-1} = f^{-1} \circ f = \mathsf{E}$, където E е единичният оператор (вж. § 1, т. 3, в). Матрицата на E е единичната матрица E (убедете се самостоятелно). Тъй като матрицата на композицията на две линейни изображения е произведението на техните матрици, от последните равенства получаваме, че $AX = XA = E$. Следователно, ако операторът $f: L \to L$ е биективно изображение, то матрицата му A е обратима. Матрицата на f^{-1} е A^{-1}.

Обратно, нека $f: L \to L$ е произволен линеен оператор и да предположим, че спрямо избран базис матрицата му A е обратима. Дефинираме линеен оператор $g: L \to L$ така, че матрицата му спрямо вече избрания базис да бъде A^{-1}. Равенствата $AA^{-1} = A^{-1}A = E$ означават, че $f \circ g = g \circ f = \varepsilon$ - единичният оператор в L, т.е. f е обратимо (биективно) изображение и $f^{-1} = g$.

Окончателно получихме, че *операторът* $f: L \to L$ *е биективно изображение тогава и само тогава, когато матрицата му е обратима*,

6. Ранг на произведение на матрици. Нека $A = (a_{ij})$, $B = (b_{rs})$ са две матрици, за които произведението AB е дефинирано. Какво може да се каже за ранга на матрицата $C = AB$? Нека A е от тип $m \times n$, B е от тип $n \times p$ и $C = (c_{ij})$; C е от тип $m \times p$. Знаем, че

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}, \quad 1 \le i \le m, 1 \le j \le p.$$

Да фиксираме i и да меним j от 1 до p. Тогава равенствата показват, че i-тият ред на матрицата C е линейна комбинация на редовете на матрицата B. Тъй като i е произволно, всеки ред на C е линейна комбинация на редовете на B. Според твърдението в т. 3, § 3, гл. 2 рангът на системата вектори-редове в C не надминава ранга на системата вектори-редове на B, т.е.

$$r(AB) \leq r(B).$$

Сега да фиксираме j, а да меним i. Горните равенства показват, че стълбовете на C са линейни комбинации на стълбовете на A и по цитираното твърдение имаме

$$r(AB) \leq r(A).$$

Да предположим например, че матрицата A е неособена. Тогава A^{-1} съществува и $B = A^{-1}.AB$. Според доказаното

$$r(B) = r(A^{-1}.AB) \leq r(AB) \leq r(B),$$

следователно $r(AB) = r(B)$. Аналогично се доказва, че ако B е неособена, то $r(AB) = r(A)$. И така, доказахме

7. Теорема. *Ако A и B са матрици, за които произведението AB е дефинирано, то рангът на AB не надминава ранга на който и да е от множителите, т.е.*

$$r(AB) \leq \min\big(r(A), r(B)\big).$$

Ако някоя от матриците A, B е неособена, то рангът на произведението AB е равен на ранга на другата матрица.

8. Действието на линейно изображение чрез координати. В § 1, т. 5 отбелязахме, че ако $f: L \to M$ е произволно линейно изображение на крайномерното линейно пространство L в крайномерното линейно пространство M, а $e_1, \ldots, e_n$ и $e_1^*, \ldots, e_m^*$ са базиси съответно на L и M, то произволен вектор

$$l = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

се изобразява във вектор

$$f(l) = y_1e_1^* + y_2e_2^* + \ldots + y_me_m^*,$$

където

(1)
$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &. \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n, \end{aligned}$$

а $A = \left(a_{ij}\right)$ е матрицата на f спрямо избраните базиси. С помощта на матрици последните равенства може да се запишат съвсем кратко и удобно. За целта нека

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

са матриците-стълбове от координатите съответно на l и $f\left(l\right)$. По определение две матрици са равни точно тогава, когато са от един и същ тип и всички съответни елементи са равни. Като си спомним правилото за умножаване на матрици, равенствата (1) може да се запишат като едно матрично равенство

(2)
$$Y = AX$$

То се помни лесно: ако матрицата на линейното изображение f умножим отдясно със стълба от координатите на вектора l, ще получим стълба от координатите на образа му $f\left(l\right)$.

§ 4. Смяна на базиса

1. Смяна на координатите на вектор при смяна на базиса. Нека отново L е крайномерно линейно пространство, в което са избрани два базиса: $e_1, e_2, \ldots, e_n$ и $e_1', e_2', \ldots, e_n'$. За удобство на речта първия от тях ще наричаме "стар базис", а втория - "нов". За произволен вектор $l \in L$ имаме

(1)
$$l = \sum_{i=1}^{n} x_i e_i = \sum_{j=1}^{n} x_j' e_j' .$$

Тук ще потърсим връзката между “старите” координати x_i и “новите” x'_j.

От определението за базис следва, че съществуват еднозначно определени числа α_{ij}, за които

(2)
$$\begin{aligned} e'_1 &= \alpha_{11}e_1 + \alpha_{21}e_2 + \ldots + \alpha_{n1}e_n \\ e'_2 &= \alpha_{12}e_1 + \alpha_{22}e_2 + \ldots + \alpha_{n2}e_n \\ &. \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \\ e'_n &= \alpha_{1n}e_1 + \alpha_{2n}e_2 + \ldots + \alpha_{nn}e_n. \end{aligned}$$

Матрицата

$$T = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ .. & .. & .. & .. \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{pmatrix}$$

се нарича *матрица на прехода* от стария базис към новия. *Тя е неособена*. Наистина да допуснем, че $\det(T) = 0$. Това означава, че рангът на матрицата T е по-малък от n, следователно стълбовете ѝ са линейно зависими, т.е. съществуват числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, които не са едновременно равни на нула и за които

$$\lambda_1(\alpha_{11}, \alpha_{21}, \ldots, \alpha_{n1}) + \lambda_2(\alpha_{12}, \alpha_{22}, \ldots, \alpha_{n2}) + \ldots + \lambda_n(\alpha_{1n}, \alpha_{2n}, \ldots, \alpha_{nn}) = (0, 0, \ldots, 0)$$

От това равенство и равенствата (2) лесно следва, че

$$\lambda_1 e'_1 + \lambda_2 e'_2 + \ldots + \lambda_n e'_n = 0.$$

Получихме, че векторите $e'_1, e'_2, \ldots, e'_n$ са линейно зависими, противно на условието, че те са базис.

Тъй като матрицата T е неособена, то тя е обратима: нейна обратна ще бъде матрицата на прехода от новия към стария базис. Полезно е да се запомни, че *стълбовете на T са старите координати на новите базисни вектори.*

Като заместим равенствата (2) в (1), получаваме

$$\sum_{i=1}^{n} x_i e_i = \sum_{j=1}^{n} x'_j e'_j = \sum_{j=1}^{n} x'_j \left(\sum_{i=1}^{n} \alpha_{ij} e_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_{ij} x'_j \right) e_i \ .$$

Сравнявайки последователно коефициентите пред $e_1, e_2, \ldots, e_n$ в най-лявата и най-дясната страна, стигаме до формулите

(3)
$$\begin{aligned} x_1 &= \alpha_{11} x'_1 + \alpha_{12} x'_2 + \ldots + \alpha_{1n} x'_n \\ x_2 &= \alpha_{21} x'_1 + \alpha_{22} x'_2 + \ldots + \alpha_{2n} x'_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= \alpha_{n1} x'_1 + \alpha_{n2} x'_2 + \ldots + \alpha_{nn} x'_n . \end{aligned}$$

Ако положим

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix},$$

в матрична форма те се записват като

(4)
$$X = TX' ,$$

т.е. старите координати се получават, ако умножим отдясно матрицата на прехода с матрицата стълб, съставена от новите координати.

2. Смяна на матрицата на линеен оператор при смяна на базиса. Запазваме означенията от предишната точка. Нека $f : L \to L$ е линеен оператор, A е матрицата му спрямо стария базис, а B - спрямо новия. Ще търсим връзка между матриците A и B. Ако l е произволен вектор, то съгласно § 4, т. 8 имаме

(5)
$$Y = AX \ , \quad Y' = BX' ,$$

където X, Y са матриците-стълбове от старите координати съответно на l и $f(l)$, а X', Y' са съответно матриците-стълбове от новите им координати. Според формулата (4) за смяна на координатите имаме

$$X = TX' , \quad Y = TY' ,$$

следователно

$$TBX' = ATX' \quad \text{или} \quad \left(TB - AT\right)X' = 0 .$$

Тъй като векторът l е произволен, последното равенство означава, че линейният оператор с матрица $TB - AT$ (спрямо новия базис) изобразява всеки вектор в нулевия, т.е. това е нулевият оператор. Следователно $TB - AT = 0$, откъдето веднага се получава, че

$$B = T^{-1}AT . \tag{6}$$

С това нашата задача е решена: връзката между матрицата A на оператора f спрямо стария базис и матрицата му B спрямо новия се дава с (6), където T е матрицата на прехода от стария базис към новия.

Във връзка с този резултат ще направим някои бележки.

а) *Подобни матрици*. Нека A и B са произволни реални (съответно комплексни) квадратни матрици от ред *n*. Ще казваме, че те са *подобни*, ако съществува неособена реална (съответно комплексна) матрица T, за която да бъде изпълнено равенството (6). Означение: $A \sim B$. Нека $\mathcal{K} = \mathrm{R}$ или C и $M_n\left(\mathcal{K}\right)$ е множеството на квадратните матрици от ред n с елементи от $\mathcal{K}$. Не е трудно да се убедим, че релацията $\sim$ е релация на еквивалентност в множеството $M_n\left(\mathcal{K}\right)$. Наистина $A \sim A$ (релацията е рефлексивна), защото $A = E^{-1}AE$, където E е единичната матрица от ред *n*. Нека $A \sim B$, т.е. $B = T^{-1}AT$. Тъй като $(T^{-1})^{-1} = T$, веднага се получава, че $(T^{-1})^{-1}AT^{-1} = B$, което означава, че $B \sim A$ (релацията е симетрична). Най-сетне нека $A \sim B$, $B \sim C$. Тогава $B = T_1^{-1}AT_1$, $C = T_2^{-1}AT_2$, откъдето получаваме, че $C = \left(T_1T_2\right)^{-1}A\left(T_1T_2\right)$, което означава, че $A \sim C$ (релацията е транзитивна).

Следователно релацията "подобие на матрици" разбива множеството $M_n\left(\mathcal{K}\right)$ на класове на еквивалентност (класове подобни матрици).

Равенството (6) означава, че матриците на даден линеен оператор спрямо всеки два базиса са подобни и подобието се реализира от матрицата на прехода от единия базис към другия. От втората част на теоремата от § 3, т. 7 веднага следва, че подобните матрици имат равни рангове. По-долу ще посочим и други характеристики на квадратните матрици, които се запазват при подобие на матрици.

б) *Следа на матрица и следа на линеен оператор*. Следа на квадратната матрица $A=\left(a_{ij}\right)$ (означение $\mathrm{tr}A$; англ. trace - ‘следа’) се дефинира като

$$\mathrm{tr}A=\sum_{i=1}^{n}a_{ii}\ .$$

Лесно се вижда, че за всеки две квадратни матрици от ред n

$$\mathrm{tr}AB=\mathrm{tr}BA\,.$$

Наистина

$$\mathrm{tr}AB=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}b_{ji}\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n}b_{ji}a_{ij}\right)=\mathrm{tr}BA\ .$$

Очевидно

$$\mathrm{tr}\left(A+B\right)=\mathrm{tr}A+\mathrm{tr}B\ ,\quad \mathrm{tr}\lambda A=\lambda\mathrm{tr}A\,,\ (\ \lambda\in\mathcal{K}\ ).$$

От предходното свойство веднага следва, че подобните матрици имат равни следи:

$$\mathrm{tr}\left(T^{-1}.AT\right)=\mathrm{tr}\left(AT.T^{-1}\right)=\mathrm{tr}A\,.$$

Нека сега $f:L\to L$ е линеен оператор в крайномерно пространство. *Следа* $\mathrm{tr}\ f$ *на оператора* f се нарича следата на матрицата му A спрямо произволно избран базис:

$$\mathrm{tr}f=\mathrm{tr}A\,.$$

Това определение е коректно, защото, ако сменим базиса, матрицата A съгласно (6) се подменя с подобна на нея матрица, чиято следа е същата. В този смисъл следата на оператора f е негов *инвариант*, т.е. тя зависи от f и не зависи от избора на такива помощни неща като базиси и матрици, с помощта на които я дефинирахме.

в) *Детерминанта на линеен оператор*. Ще покажем, че подобните матрици имат равни детерминанти. За целта най-напред ще забележим, че ако T е обратима матрица, то $\det T^{-1}=\left(\det T\right)^{-1}$. Действително, от равенството $TT^{-1} = E$ имаме $\det T.\det T^{-1} = \det E = 1$. Следователно

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

Това обосновава коректността на следното определение: *детерминанта на оператора* f (означение $\det f$) ще наричаме детерминантата на матрицата му A спрямо който и да е базис, т.е. $\det f = \det A$. Тя също е инвариант на оператора.

г) *Характеристичен полином на матрица и на линеен оператор.* Нека $A = \left(a_{ij}\right)$ е произволна квадратна матрица от ред n, а t е независима променлива. Да разгледаме детерминантата

$$\det\left(tE - A\right)=\begin{vmatrix} t-a_{11} & -a_{12} & \dots & -a_{1\mathrm{n}} \\ -a_{21} & t-a_{22} & \dots & -a_{2n} \\ .. & .. & .. & .. \\ -a_{n1} & -a_{n2} & \dots & t-a_{nn} \end{vmatrix}.$$

Ако я развием, очевидно ще получим полином на t, който ще означим с $\varphi(t)$:

$$\varphi_A\left(t\right) = a_0t^n + a_1t^{n-1} + \dots + a_{n-k}t^k + \dots + a_{n-1}t + a_n$$

Той се нарича *характеристичен полином на матрицата* A. За да определим коефициента a_n, е достатъчно да положим $t = 0$. Имаме $\varphi_A\left(0\right) = a_n$, т. е. $a_n = \det\left(-A\right) = \left(-1\right)^n \det A$.

За да определим a_0 и a_1, ще използваме, че детерминантата е сума от произведения на нейни елементи, в които от всеки ред и всеки стълб участва точно по един елемент и произведението се взема с подходящ знак. Произведението

$$\left(t - a_{11}\right)\left(t - a_{22}\right)\dots\left(t - a_{nn}\right)$$

участва в развитието на детерминантата и очевидно е равно на $t^n - \left(a_{11} + a_{22} + \dots + a_{nn}\right)t^{n-1}$ +членове от по-ниска степен. Всички останали произведения, които участват в развитието на детерминантата, са от степен най-много $n-2$ по отношение на t, защото, ако такова произведение съдържа множител a_{ij} ($i \neq j$), то няма да съдържа $t-a_{ii}$ и $t-a_{jj}$, т.е. степента му ще е най-много $n-2$. Следователно $a_0=1$ и $a_1=-\left(a_{11}+a_{22}+...+a_{nn}\right)=-\mathrm{tr}A$.

Пресмятането на останалите коефициенти в характеристичния полином е по-сложно. Удобно е да се въведе подходящо понятие. Ако в квадратна матрица от ред n някакъв минор от ред m ($m \leq n$) принадлежи на редове с номера $i_1, i_2, ..., i_m$ и на стълбове със същите номера, ще казваме, че той е *главен минор* от ред m. Може да се докаже, че коефициентът a_{n-k} пред t^k в характеристичния полином на квадратна матрица A от ред n е равен на сумата на всички главни минори от ред $n-k$, умножена с $\left(-1\right)^{n-k}$. Тук доказахме само, че

$$\varphi_A\left(t\right)=t^n-\left(a_{11}+a_{22}+...+a_{nn}\right)t^{n-1}+...+\left(-1\right)^n \det A.$$

Ще покажем, че подобните матрици имат равни характеристични полиноми. Наистина, нека $B=T^{-1}AT$. Тогава

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

Корените на полинома $\varphi_A\left(t\right)$ се наричат *характеристични корени на матрицата* A.

Ако f е линеен оператор и A е матрицата му спрямо някакъв базис, характеристичният полином на A се нарича и *характеристичен полином на оператора.* Той зависи само от оператора и не зависи от избора на базиса: ако сменим базиса, матрицата A ще се смени с подобна на A и тя ще има същия характеристичен полином.

УПРАЖНЕНИЯ

1. Нека A е произволна матрица, която с помощта на произволни хоризонтални и вертикални прави разделяме на части, които също ще разглеждаме като матрици и ще ги наричаме *клетки* или *блокове* на A; в такъв случай се казва, че A е клетъчна матрица. Малко по-нагледно:

$$A=\left(\begin{array}{c:c:c:c} A_{11} & A_{12} & \dots & A_{1q} \\ \hdashline & & & \\ \hdashline A_{p1} & A_{p2} & \dots & A_{pq} \end{array}\right);$$

тук $A_{\alpha\beta}$ са клетките. В конкретния случай се казва, че A има *клетъчен тип* $p\times q$.

а) Нека матриците A и B са от един и същ тип и са разбити на клетки по един и същ начин, т.е.

$$A=\begin{pmatrix} A_{11} & ... & A_{1q} \\ .. & .. & .. \\ A_{p1} & ... & A_{pq} \end{pmatrix}, \quad B=\begin{pmatrix} B_{11} & ... & B_{1q} \\ .. & .. & .. \\ B_{p1} & ... & B_{pq} \end{pmatrix}$$

и всеки две клетки $A_{\alpha\beta}$ и $B_{\alpha\beta}$ са от един същ тип. Докажете, че

$$A+B=\begin{pmatrix} A_{11}+B_{11} & ... & A_{1q}+B_{1q} \\ .. & .. & .. \\ A_{p1}+B_{p1} & ... & A_{pq}+B_{pq} \end{pmatrix}, \quad \lambda A=\begin{pmatrix} \lambda A_{11} & ... & \lambda A_{1q} \\ .. & .. & .. \\ \lambda A_{p1} & ... & \lambda A_{pq} \end{pmatrix}.$$

б) Нека матрицата A е от тип $m\times n$, а матрицата B е от тип $n\times k$. Нека те са разбити на клетки по следния начин:

$$A=\begin{pmatrix} A_{11} & ... & A_{1q} \\ .. & .. & .. \\ A_{p1} & ... & A_{pq} \end{pmatrix}, \quad B=\begin{pmatrix} B_{11} & ... & B_{1r} \\ .. & .. & .. \\ B_{q1} & ... & B_{qr} \end{pmatrix},$$

т.е. първата е от клетъчен тип $p\times q$, а втората - от тип $q\times r$. Нека освен това броят на стълбовете във всяка клетка A_{ij} е равен на броя на редовете във всяка клетка B_{jl}. Докажете, че

$$AB = \begin{pmatrix} C_{11} & \dots & C_{1r} \\ .. & .. & .. \\ C_{p1} & \dots & C_{pr} \end{pmatrix},$$

където за клетките C_{il} имаме

$$C_{il} = \sum_{j=1}^{q} A_{ij} B_{jl} .$$

2. Нека E_{ij} е матрица от тип $m \times n$, в която на пресечната точка на i -тия ред и j -тия стълб стои единица, а останалите елементи са нули. (Това са така наречените *матрични единици*. В означението E_{ij} липсва указание за типа на матрицата. Обикновено той се подразбира.) Дефинираме следните квадратни матрици от ред n, наричани *елементарни матрици*:

$$F_{st} = E - E_{ss} - E_{tt} + E_{st} + E_{ts}, \quad s \neq t;$$
$$F_{st}\left(\lambda\right) = E + \lambda E_{st}, \quad s \neq t;$$
$$F_{s}\left(\lambda\right) = E + \left(\lambda - 1\right) E_{ss}, \quad \lambda \neq 0.$$

Тук E е единичната матрица от ред n.

а) Изпишете елементарните матрици в явен вид (като таблици от числа).

б) Покажете, че елементарните матрици са обратими и

$$F_{st}^{-1} = F_{st}, \quad F_{st}\left(\lambda\right)^{-1} = F_{st}\left(-\lambda\right), \quad F_{s}\left(\lambda\right)^{-1} = F_{s}\left(\lambda^{-1}\right).$$

в) Нека $A = \left(a_{ij}\right)$ е квадратна матрица от ред n. Убедете се, че матриците $F_{st}A$, $F_{st}\left(\lambda\right)A$ и $F_{s}\left(\lambda\right)A$ се получават от A съответно чрез: размяна на местата на s -тия ред и t -тия ред; умножаване на t -тия ред с λ и прибавяне към s -тия ред; умножаване на s -тия ред с λ.

г) Как следва да се интерпретират матриците $AF_{s,t}$, $AF_{s,t}\left(\lambda\right)$ и $AF_{s}\left(\lambda\right)$?

д) Докажете, че ако A е квадратна матрица от ред n, то съществуват елементарни матрици $P_1, \dots, P_k$ и $Q_1, \dots, Q_m$ такива, че

$$P_k P_{k-1} \dots P_1 A Q_1 Q_2 \dots Q_m = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & 0 & \\ & & & 1 & & & \\ & 0 & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

Колко е броят на единиците в последната матрица? (Отговор: той е равен на ранга на A.)

е) Ако A е неособена матрица, покажете, че в дясната страна на равенството от д) стои единичната матрица E. Изведете оттук, че

$$A^{-1} = Q_1 Q_2 \dots Q_m P_k P_{k-1} \dots P_1.$$

ж) Докажете, че ако A е неособена квадратна матрица, то с помощта на елементарни преобразувания само на редовете (или само на стълбовете) матрицата A може да се трансформира в единичната матрица E. Като използвате в) и г), изведете, че съществуват елементарни матрици $P_1, P_2, \dots, P_r$ и $Q_1, Q_2, \dots, Q_s$, за които

$$P_r P_{r-1} \dots P_1 A = E, \quad A Q_1 Q_2 \dots Q_s = E.$$

Следователно $A^{-1} = P_r P_{r-1} \dots P_1 = Q_1 Q_2 \dots Q_s$.

По-нататък обосновете следния алгоритъм: нека с последователно прилагане на елементарни преобразувания само върху редовете на матрицата A я трансформираме в единичната матрица E; ако същите преобразувания в същата последователност приложим върху редовете на единичната матрица, ще получим обратната матрица A^{-1}. Формулирайте аналогичен алгоритъм за стълбовете. (У п ъ т в а н е . Използвайте резултата от в) и равенството $A^{-1} = P_r P_{r-1} \dots P_1 . E$.)

З а б е л е ж к а . С формулирания алгоритъм намирането на матрицата A^{-1} се сведе до елементарни преобразувания на матрицата A и на единичната матрица. С това се заобикаля пресмятането на n^2 детерминанти от ред $n-1$ - адюнгираните количества A_{ij}, както и пресмятането на $\det A$.

3. Обобщете твърдението д) от предната задача за произволна (не непременно квадратна) матрица.

4. Докажете, че всяка матрица с ранг 1 може да се представи като произведение на матрица-стълб и матрица-ред.

5. Докажете, че произведението на две горни триъгълни матрици е горна триъгълна и че обратната на горна триъгълна също е горна триъгълна матрица.

6. Нека $M_n(\mathcal{K})$ е множеството на всички квадратни матрици от ред n с елементи от $\mathcal{K}$, където $\mathcal{K} = \mathrm{R}$ или C. Докажете, че ако за матрицата $X \in M_n(\mathcal{K})$ равенството $AX = XA$ е изпълнено за всяка матрица $A \in M_n(\mathcal{K})$, то X е *скаларна* матрица, т.е. $X = \lambda E$ за някое $\lambda \in \mathcal{K}$. (Упътване. Използвайте, че в частност $E_{ij}X = XE_{ij}$, където E_{ij} са матричните единици от зад. 2.)

7. Нека I е непразно подмножество на $M_n(\mathcal{K})$ със следните свойства: а) ако $X, Y \in I$, то матриците $X + Y$ и $X - Y$ също принадлежат на I; б) ако $X \in I$, то матриците AX и XA принадлежат на I за всяка матрица $A \in M_n(\mathcal{K})$. Докажете, че или I се състои само от нулевата матрица, или $I = M_n(\mathcal{K})$.

Глава 4

ЛИНЕЙНИ ПРОСТРАНСТВА

§ 1. Определения

1. Ако погледнем съдържанието на предишните две глави в малко по-едър план, ще забележим, че при изучаването там на наглед различни математически обекти е налице известна аналогия и дори повторяемост. Ще поясним малко по-подробно. Работихме с множеството R^n на наредените n-торки от реални числа. Въведохме операции събиране на елементи от R^n и умножаване на n-торки с реални числа. Основните свойства на двете операции до голяма степен напомнят операциите с числа (вж. гл. 2, § 1, т. 2). Абсолютно същото важи и за C^n (там обаче умножавахме n-торките с числа от C). По-късно при $L = \mathrm{R}^n$ и $M = \mathrm{R}^m$ (или при $L = \mathrm{C}^n$ и $M = \mathrm{C}^m$) разглеждахме всевъзможните линейни изображения $f: L \to M$. Събирахме линейни изображения и ги умножавахме с числа - реални (съотв. комплексни). Най-сетне събирахме матрици и ги умножавахме с числа. И в трите случая основните свойства на операциите се оказаха едни и същи. Ако се абстрахираме от конкретната природа на обектите, с които работихме, бихме могли да кажем, че и в трите ситуации е налице някакво множество L (от наредени n-торки от числа, от линейни изображения, от матрици), в което по някакъв начин са дефинирани операции събиране и умножаване с числа. Интуитивно е ясно, че трите случая са разновидности на по-обща ситуация, която тук се стремим да дефинираме аксиоматично. За целта се налага да се изчистят несъществените детайли, да се обобщят някои понятия и да се открои логическият скелет.

2. Определение. Ще казваме, че в множеството L е въведена *бинарна алгебрична операция* $*$, ако е зададено правило, което на всяка наредена двойка (l_1, l_2) от елементи на L съпоставя еднозначно определен елемент $l_1 * l_2$, принадлежащ на L. (С други думи зададено е изображение $*: L \times L \to L$.)

Например, ако $L = \mathbf{N}$ е множеството на естествените (целите положителни) числа, то обичайното събиране $a+b$ на числа е бинарна алгебрична операция в $\mathbf{N}$, докато изваждането $a-b$ не е, тъй като $a-b$ невинаги е естествено число. (Условието $l_1 * l_2 \in L$, т.е. L да бъде *затворено относно операцията* $*$, е съществена част от определението!)

Ако $L = Z$ (множеството на целите числа), обичайното събиране $a+b$ и изваждането $a-b$ са бинарни алгебрични операции в $\mathcal{Z}$. Операцията събиране, която въведохме в R^n (съответно в C^n), също е бинарна.

По-особен е например случаят, когато умножавахме матрици с числа и резултатът беше матрица. Тук не може да говорим за бинарна алгебрична операция, защото са налице две множества: от една страна на матриците, а от друга страна, на числата. Ще се наложи да употребим термина *външна операция*. В глави 2 и 3 в качеството на множество $\mathcal{K}$ от числа, с които умножаваме, избирахме или R, или C. Практиката е показала, че за нуждите на обобщението, към което се стремим, е удобно и достатъчно да се предположи само, че $\mathcal{K}$ е *числово поле* (вж. гл. 7, § 2) и да не се налагат други ограничения.

Без да се впускаме в повече анализи, въвеждаме следното

3. Определение. Нека L е дадено непразно множество, а $\mathcal{K}$ е дадено числово поле. Нека освен това:

а) в L е зададена бинарна алгебрична операция, която ще наричаме събиране $l_1 + l_2$;

б) е зададена и външна операция $(\lambda, l) \mapsto \lambda l$, т.е. правило, което на всяко число λ от $\mathcal{K}$ и на всеки елемент l от L съпоставя еднозначно определен елемент от L, който ще бележим с λl.

Ще казваме, че L е *линейно* (*векторно*) *пространство над полето* $\mathcal{K}$, ако са изпълнени следните условия (аксиоми на линейно пространство):

1) $l_1 + l_2 = l_2 + l_1$ за всеки два елемента l_1, l_2 от L (*комутативност на събирането*);

2) $(l_1 + l_2) + l_3 = l_1 + (l_2 + l_3)$ за всеки три елемента l_1, l_2, l_3 от L (*асоциативност на събирането*);

3) в L съществува елемент, който ще бележим с 0, за който $l + 0 = l$ за всеки $l \in L$ (в L *съществува нулев елемент*);

4) за всеки $l \in L$ съществува елемент от L, който ще бележим с $-l$, за който $l + (-l) = 0$. (Елементът $-l$ ще наричаме *противоположен на* l.);

5) $1l = l$ за всеки $l \in L$;

6) $\lambda(l_1 + l_2) = \lambda l_1 + \lambda l_2$ за всички $\lambda \in \mathcal{K}$, $l_1, l_2 \in L$;

7) $(\lambda + \mu) l = \lambda l + \mu l$ за всички $\lambda, \mu \in \mathcal{K}$, $l \in L$;

8) $(\lambda\mu) l = \lambda(\mu l)$ за всички $\lambda, \mu \in \mathcal{K}$, $l \in L$.

Елементите на L ще наричаме *вектори*, а елементите на полето $\mathcal{K}$ - *скалари*.

Обикновено векторите ще бележим с малки латински букви, а скаларите - с малки гръцки букви.

4. Примери. а) Нека $L = \mathrm{R}^n$ е реалното координатно линейно просранство, разгледано в гл. 2, а $\mathcal{K} = \mathrm{R}$. То очевидно е линейно пространство над R. Същото важи и за комплексното координатно просранство $L = \mathrm{C}^n$, при $\mathcal{K} = \mathrm{C}$ - то е линейно пространство над C.

б) Ще обобщим пример а). Нека $\mathcal{K}$ е произволно числово поле, а $L = \mathcal{K}^n$ е множеството на всички наредени n-торки с елементи от $\mathcal{K}$. Събирането на n-торки и умножаването им с числа от $\mathcal{K}$ дефинираме както за R^n и C^n. Почти очевидно е, че $\mathcal{K}^n$ е линейно пространство над $\mathcal{K}$. При $n = 1$ получаваме, че всяко поле $\mathcal{K}$ може да се разглежда като линейно пространство над себе си.

в) Нека $L = \mathcal{K}$ е произволно числово поле, а $\mathbf{Q}$ е полето на рационалните числа. Събирането в l е обичайната сума на числа. Външната операция е умножаването на числата от $\mathcal{K}$ с рационални числа. Тъй като всяко числово поле съдържа $\mathbf{Q}$, то $\mathcal{K}$ е линейно пространство над $\mathbf{Q}$.

г) Нека $\mathcal{K}$ е числово поле и $M_{m\mathrm{x}n}(\mathcal{K})$ означава множеството на всички матрици от тип $m \mathrm{x}\, n$ с коефициенти от $\mathcal{K}$. Лесно е да се съобрази, че $M_{m\mathrm{x}n}(\mathcal{K})$, разглеждано заедно с операциите събиране на матрици и умножаване на матрица с число, е линейно пространство над $\mathcal{K}$ (вж. гл. 3, § 4, т. 2). Всъщност трябва да се съобрази само, че операциитте не извеждат вън от разглежданото множество, а това следва от определението на числово поле.

д) Нека $\mathcal{K}$ е числово поле, $L = \mathcal{K}^n$, $M = \mathcal{K}^m$. Нека $\mathcal{L}(L, M)$ е множеството на всички линейни изображения $f : L \to M$. В гл. 3, § 4, т. 1 въведохме операциите сума на линейни изображения и произведение на линейно изображение с число и показахме, че и в двата случая резултатът от съответната операция е отново линейно изображение на L в M. Очевидно е, че нулевото изображение $0 : L \to M$, което на всеки вектор от L съпоставя нулевия вектор от M, е линейно и че $f + 0 = f$ за всяко линейно изображение $f \in \mathcal{L}(L, M)$. Ако $-f = (-1)f$, то разбира се $f + (-f) = 0$. Вече е ясно, че $\mathcal{L}(L, M)$ е линейно пространство над $\mathcal{K}$.

е) Нека $[a, b]$ е затворен интервал от R, а L е множеството на всички функции, които са дефинирани и непрекъснати във всички точки на $[a, b]$, а стойностите им са реални числа. Сума на функциите f и g се дефинира по обичайния начин: $\left(f+g\right)\left(x\right)=f\left(x\right)+g\left(x\right)$. Нека $\lambda \in \mathrm{R}$, $f \in L$. Дефинираме: $\left(\lambda f\right)\left(x\right)=\lambda f\left(x\right)$. Тъй като сума на непрекъснати функции е непрекъсната и произведението на непрекъсната функция с число е непрекъсната функция, то L е затворено относно въведените операции. Ако означим с 0 функцията, която е тъждествено нула в $\left[a,b\right]$, става ясно, че L е линейно пространство над R.

ж) Нека L е множество, състоящо се само от един елемент, който ще означим с 0, т.е. $L=\left\{0\right\}$, а $\mathcal{K}$ е числово поле. В L дефинираме събиране и външна операция с правилата: $0+0=0$, $\lambda 0=0$, за всяко $\lambda \in \mathcal{K}$. Осемте аксиоми на линейно пространство очевидно са изпълнени, следователно L е линейно пространство над $\mathcal{K}$, състоящо се само от нулевия елемент. Пространството L ще наричаме *нулево линейно пространство*. Всички линейни пространства, които съдържат повече от един вектор, ще наричаме *ненулеви линейни пространства*.

Понятието линейно (векторно) пространство е едно от фундаменталните в цялата математика. Оказва се, че е възможно да се развие богата по съдържание теория на абстрактните линейни пространства и резултатите й след това да се прилагат направо към всеки конкретен случай. Разбира се, ако извършваме пресмятания в някакво абстрактно линейно пространство, единствените позволени правила за смятане са аксиомите 1) - 8) от определението и следствията от тях. Те са твърде естествени и помненето им не е трудно.

5. Някои следствия от аксиомите. а) *Нулевият елемент* 0 *от аксиома* 3) *е единствен*. Наистина, нека $0'$ е още един елемент от L, за който $l+0'=l$ за всеки $l \in L$. Тогава в частност $0+0'=0$, а $0'+0=0'$, защото 0 е нулев. Според 1) имаме $0+0'=0'+0$, следователно $0'=0$. В определението се условихме, че елементите на линейното пространство L ще наричаме вектори. В унисон с това единствения елемент 0 от L, удовлетворяващ 3), ще наричаме *нулев вектор*.

б) *За всеки вектор* $l \in L$ *неговият противоположен е единствен*. (Съгласно аксиома 4) поне един противоположен на вектора l съществува.) Действително, нека векторите l' и l'' са противоположни на l. Съгласно аксиома 2) имаме

$$\left(l'+l\right)+l''=l'+\left(l+l''\right)$$

Като използваме аксиомите 1) и 3), лявата страна е равна на $0+l''=l''$, защото l' е противоположен на вектора l; по същата причина дясната страна пък е равна на $l' + 0 = l'$. Следователно $l'' = l'$ и означението $-l$ за противоположния вектор на l не крие опасност от недоразумения.

в) *За всеки вектор* $l \in L$ е в сила $0l = 0$. (Тук 0 означава числото нула вляво и нулевия вектор вдясно.) Наистина, като използваме аксиома 7), имаме

$$0l = \left(0+0\right)l = 0l + 0l .$$

Ако към двете страни на последното равенство прибавим $-\left(0l\right)$ и използваме аксиома 2), ще получим, че $0l = 0$.

г) *За всеки вектор* l *е в сила* $-l = \left(-1\right)l$. Действително, като приложим аксиоми 5) и 7) и твърдението б), имаме

$$0 = 0l = \left(1-1\right)l = 1l + \left(-1\right)l = l + \left(-1\right)l ,$$

откъдето следва, че $\left(-1\right)l$ е противоположният вектор на l .

д) Според определението по-горе може да се събират само два вектора. Ако l_1, l_2, ..., l_n е наредена n-торка от вектори, $n \geq 3$, тяхната сума може да се образува по различни начини като с помощта на скоби, без да се нарушава редът на векторите, се посочи последователността, в която ще се събират по два вектора. Например, при $n = 4$:

$$\left(\left(l_1 + l_2\right) + l_3\right) + l_4, \quad \left(l_1 + \left(l_2 + l_3\right)\right) + l_4, \quad l_1 + \left(l_2 + \left(l_3 + l_4\right)\right),$$

$$l_1 + \left(\left(l_2 + l_3\right) + l_4\right), \quad \left(l_1 + l_2\right) + \left(l_3 + l_4\right).$$

Като се използва аксиома 2), е лесно да се забележи, че посочените суми са равни. В сила е следното общо твърдение: *резултатът от последователното прилагане на операцията събиране върху* n *вектора,* $n \geq 3$*, не зависи от разполагането на скобите*. При $n = 3$ твърдението е просто аксиома 2). Педантичното доказателство за произволно n няма да излагаме. То може да се извърши с помощта на индукция по n.

Следствие д) позволява да пишем суми от вида $l_1 + l_2 + \ldots + l_n$ без употреба на скоби.

Преди да пристъпим към въвеждането на редица допълнителни понятия, като линейна независимост, линейна зависимост, базиси, размерности, координати и т.н., длъжни сме да направим едно принципно разяснение. В гл. 2 и гл. 3 работихме с две конкретни линейни пространства: реалното координатно и комплексното координатно. Споменатите понятия там бяха въведени, като мотивите произтичаха от конкретни математически задачи. Скромният опит, който там натрупахме, потвърждава съдържателността на тези понятия. Естествено е да ги въведем и в произволно линейно пространство, като считаме, че допълнителна мотивировка вече не е нужна. Следва да признаем и още нещо: когато там формулирахме определения и твърдения, “тайно” имахме предвид обобщенията. Не бива да учудва, че сега пренасянето на определения и на почти всички твърдения ще е буквално.

6. Линейна независимост и линейна зависимост. Нека L е линейно пространство над (числово) поле $\mathcal{K}$. Ще казваме, че векторите $l_1, l_2, \ldots, l_k$ от L са *линейно независими*, ако векторът

$$\lambda_1 l_1 + \lambda_2 l_2 + \ldots + \lambda_k l_k, \quad \lambda_i \in \mathcal{K},$$

е нулевият единствено когато $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$.

Ако съществуват числа $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathcal{K}$, които не са едновременно нули и за които

$$\lambda_1 l_1 + \lambda_2 l_2 + \ldots + \lambda_k l_k = 0,$$

ще казваме, че векторите са *линейно зависими*.

За всяка система от краен брой вектори е в сила теоремата от т. 8, § 1, гл. 2. Препоръчваме на читателя да провери, че нейното доказателство никъде не използва конкретната природа на векторите и на числовото поле.

Нека L е произволно линейно пространство над поле $\mathcal{K}$, $l_1, l_2, \ldots, l_s$ са вектори от L, а $\lambda_1, \lambda_2, \ldots, \lambda_s$ са скалари от $\mathcal{K}$. Векторът $l = \lambda_1 l_1 + \lambda_2 l_2 + \ldots + \lambda_s l_s$ се нарича *линейна комбинация* на векторите l_i с коефициенти λ_i. В сила е следната

7. Теорема. *Ако всеки от векторите $a_1, a_2, \ldots, a_k \in L$ е линейна комбинация на векторите $b_1, b_2, \ldots, b_m \in L$ и $k > m$, то векторите $a_1, a_2, \ldots, a_k$ са линейно зависими.*

За координатните векторни пространства R^n и C^n тази теорема бе доказана в гл. 2, § 1, т. 6. Доказателството никъде не използва, че векторите са наредени n-торки от числа, следователно може да се пренесе буквално и за произволно линейно пространство. За да разсеем тук и по-нататък евентуални съмнения, налага се да направим следната *забележка*. Въпросното доказателство бе получено като следствие от факта, че ако в хомогенна система от линейни уравнения неизвестните са повече от уравненията, то системата има ненулево решение; фактът пък бе получен като просто наблюдение при решаването на системи по метода на Гаус. За да не отвличаме излишно вниманието на читателя, в теорията на системите от линейни уравнения, развита в гл.1 и гл.2, общо взето не налагахме ограничения за числовото поле, на което принадлежат коефициентите на системата или компонентите на решенията. Нека сега си мислим, че е дадена произволна линейна хомогенна система, в която коефициентите пред неизвестните принадлежат на някакво фиксирано числово поле $\mathcal{K}$. Както видяхме, при нейното решаване и изследване върху коефициентите се прилагат само аритметичните операции събиране, изваждане, умножение и деление, които не извеждат вън от полето $\mathcal{K}$. Разбира се, понякога избирахме свободни неизвестни (параметри), но нищо не пречи да ги изберем само от $\mathcal{K}$. Следователно споменатата теорема за хомогенни системи би могла да се прецизира и така: ако в хомогенна линейна система с коефициенти от числовото поле $\mathcal{K}$ броят на неизвестните е по-голям от броя на уравненията, то системата има ненулево решение, чиито компоненти са числа от $\mathcal{K}$. Това уточнение е необходимо, защото за педантичното доказателство на теорема 7 по-горе е необходимо всички използвани скалари да принадлежат на $\mathcal{K}$. След този коментар остава само да препишем буквално доказателството от гл. 2.

§ 2. Базиси и размерности

1. Определение. Нека L е ненулево линейно пространство над поле $\mathcal{K}$. Ще казваме, че наредената n-торка от вектори $e_1, e_2, \dots, e_n \in L$ е *базис* на L, ако: а) векторите $e_1, e_2, \dots, e_n$ са линейно независими; б) всеки вектор от L е тяхна линейна комбинация. Ако L е нулевото пространство, негов *базис* ще наричаме празното множество от вектори. Ако едно линейно пространство има базис (от краен брой вектори), ще казваме, че то е *крайномерно линейно пространство*. Ако не съществува базис (състоящ се от краен брой вектори), ще казваме, че пространството е *безкрайномерно*.

Подчертаваме, че всеки базис се разглежда като наредена система от вектори. Ако разменим местата им, отново ще получим базис, но той ще бъде вече друг.

Условията а) и б) са еквивалентни на изискването всеки вектор от ненулевото линейно пространство L да се представя, и то еднозначно, като линейна комбинация от базисните вектори. Наистина, нека $e_1, e_2, \dots, e_n$ е базис на L. Тогава от равенството

$$x_1 e_1 + x_2 e_2 + \dots + x_n e_n = x_1' e_1 + x_2' e_2 + \dots + x_n' e_n, \quad x_i, x_i' \in \mathcal{K}$$

следва

$$\left(x_1 - x_1'\right) e_1 + \left(x_2 - x_2'\right) e_2 + \dots + \left(x_n - x_n'\right) e_n = 0,$$

а тъй като базисните вектори са линейно независими, последното е възможно единствено когато $x_1 - x_1' = x_2 - x_2' = \ldots = x_n - x_n' = 0$, т.е. $x_1 = x'$, $x_2 = x_2'$, ..., $x_n = x_n'$. Следователно всеки вектор от L се представя, и то еднозначно, като линейна комбинация на базисните вектори.

Обратно, нека всеки вектор от L се представя по единствен начин като линейна комбинация на някакви фиксирани вектори $e_1, e_2, \ldots, e_n \in L$. В частност това важи и за нулевия вектор. За него разбира се имаме $0e_1 + 0e_2 + \ldots + 0e_n = 0$. От единствеността на това представяне следва, че равенството

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0, \quad \lambda_i \in \mathcal{K},$$

е възможно само когато $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$, т.е. векторите $e_1, e_2, \ldots, e_n$ са линейно независими и образуват базис.

2. Твърдение. *Ако L е крайномерно линейно пространство, то броят на векторите в които и да са два негови базиса е един и същ. Ако този брой е n, то всеки $n+1$ вектора от L са линейно зависими.*

Доказателство. Ако $L = \{0\}$ е нулево пространство, то по определение негов базис е празното множество от вектори. В случая базисът е единствен и броят на елементите му е равен на нула. Нулевият вектор 0 е единственият вектор от L, а тъй като $1.0 = 0$, той е линейно зависим.

Нека L е ненулево линейно пространство, а $e_1, \ldots, e_k$ и $e_1', \ldots, e_m'$ са два негови базиса. Това означава в частност, че всеки вектор e_i е линейна комбинация на векторите $e_1', \ldots, e_m'$. Ако допуснем за момент, че $k > m$, то според теоремата в т.7 на предишния параграф векторите $e_1, \ldots, e_k$ са линейно зависими, противно на условието, че те са базис. Следователно $k \le m$. Аналогично получаваме $m \le k$. Така $k = m$. Втората част на твърдението също следва от цитираната теорема, защото всеки $n+1$ вектора от L са линейни комбинации на векторите от даден базис, броят на които е n.

Доказаното твърдение обосновава коректността на следното

3. Определение. Ако L е крайномерно линейно пространство, броят на векторите в който и да е негов базис ще наричаме *размерност* на L и ще го бележим с $\dim L$.

Според това определение размерността на нулевото пространство е равна на нула.

Не всяко линейно пространство е крайномерно. Например, нека $L = \mathrm{R}\left[x\right]$ е множеството на всички полиноми на променливата x с реални коефициенти. Сумата на полиноми и произведението на полином с число, дефинирани по традиционния начин, дават възможност $\mathrm{R}\left[x\right]$ да се разглежда като линейно пространство над R. По определение полиномът $f\left(x\right) = a_0 + a_1 x + \ldots + a_k x^k$ е нулевият полином тогава и само тогава, когато $a_0 = a_1 = \ldots = a_k = 0$. Това означава, че "векторите" $1, x, \ldots, x^k$ от L са линейно независими за всяко естествено число k. Ако това пространство е крайномерно, ще стигнем до противоречие с втората част на твърдението в т. 2. Следователно L не е крайномерно пространство.

Понятието базис може да се разшири и за безкрайномерни линейни пространства. В рамките на този елементарен курс няма да го използваме, затова само ще скицираме идеите. Предварително ще напомним, че понятието "система вектори $l_1, l_2, \ldots, l_k$" не изключва възможността някои от векторите да са равни. Празното множество от вектори също е "система от вектори"; според определението в т. 6, § 1 тя е линейно независима.

По-нататък нека L е произволно линейно пространство над поле $\mathcal{K}$. За подмножеството S на L ще казваме, че е *линейно независимо*, ако всяка крайна система различни вектори от S е линейно независима. (Подмножеството S би могло да бъде и безкрайно множество.) Ако $S \subset L$ и S съдържа поне една крайна система от различни линейно зависими вектори, ще казваме, че S е *линейно зависимо* множество. По-нататък за подмножеството S на L ще казваме, че е *максимално линейно независимо* подмножество, ако: а) S е линейно независимо; б) за всеки вектор $l \in L \setminus S$ множеството $S \cup \left\{l\right\}$ е линейно зависимо. Всяко максимално линейно независимо подмножество на L ще наричаме *базис* на линейното пространство L.

За крайномерни линейни пространства това определение е еквивалентно на определението на базис от т.1. Ако L е нулевото линейно пространство, няма какво да се доказва, защото единственото линейно независимо подмножество на L е празното множество от вектори. Нека $\dim L = n > 0$ и $e_1, e_2, \ldots, e_n$ е базис на пространството L в смисъл на определението от т. 1. Според твърдението от т.2 всеки $n+1$ вектора от L са линейно зависими, следователно $S = \left\{e_1, e_2, \ldots, e_n\right\}$ е максимално линейно независимо подмножество на пространството L. Обратно, нека S е някакво максимално линейно независимо подмножество на L. Пак по твърдението от т. 2 броят на елементите на S не надминава n. Нека

$S=\{l_1,l_2,...,l_k\}$, $k\le n$ и l е произволен вектор от L. Ако l принадлежи на S, то $l=l_i$ за някое $i=1,...,n$. В този случай векторът l очевидно е линейна комбинация на векторите от S. Ако векторът l не принадлежи на S, то векторите $l_1,...,l_k,l$ са линейно зависими, защото S е максимално линейно независимо подмножество на пространството L. Тогава съществува линейна зависимост

$$\lambda_1 l_1+\lambda_2 l_2+...+\lambda_k l_k+\lambda l=0\ ,\ \lambda_i\ ,\ \lambda\in\mathrm{K}\ ,$$

в която не всички коефициенти са равни на нула. Тъй като $l_1,...,l_k$ са линейно независими, имаме $\lambda\ne 0$ и векторът l може да се запише като линейна комбинация на векторите $l_1,...,l_k$ от S. Следователно подмножеството S е базис на линейното пространство L в смисъл на определението от т.1.

Ако линейното пространство L не е крайномерно, цялата трудност е да се установи, че то съдържа максимални линейно независими подмножества. За целта е необходимо да се използва трансфинитна индукция, или еквивалентно - лемата на Цорн. Без да се занимаваме с това, нека си мислим, че S е максимално линейно независимо подмножество на линейното пространство L. С разсъждения както в предишния абзац може да се докаже, че за всеки ненулев вектор $l\in L$ съществуват еднозначно определени вектори $l_1,l_2,...,l_k$, принадлежащи на S и зависещи от l, и ненулеви скалари $\lambda_1,\lambda_2,...,\lambda_k$, принадлежащи на полето $\mathcal{K}$, за които

$$l=\lambda_1 l_1+\lambda_2 l_2+...+\lambda_k l_k$$

4. Координати. Ако L е крайномерно линейно пространство над поле $\mathcal{K}$ и $e_1,e_2,...,e_n$ е негов базис, то всеки вектор l - се записва във вида

(2) $$l=x_i e_1+x_2 e_2+...+x_n e_n\ ,\ x_i\in\mathcal{K}$$

В т. 1 видяхме, че за всеки вектор това записване е еднозначно определено. Числата от наредената n-торка $(x_1,x_2,...,x_n)$ се наричат *координати на вектора спрямо l базиса $e_1,e_2,...,e_n$*.

Ако векторите l и m имат координати съответно $\left(x_1, x_2, \dots, x_n\right)$ и $\left(y_1, y_2, \dots, y_n\right)$, непосредствено се вижда, че линейната комбинация $\lambda x + \mu m$ има координати

$$\left(\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, \dots, \lambda x_n + \mu y_n\right) .$$

В гл. 3, § 4, т. 1 за пространствата R^n и C^n намерихме правилото, по което се сменят координатите на вектор при смяна на базиса. Това правило се пренася буквално и за абстрактни крайномерни линейни пространства - достатъчно е да се прегледа отново текстът на разсъждението в цитираната точка. В него конкретната природа на векторите не се използва. По-конкретно в сила е следният резултат. Нека L е произволно крайномерно линейно пространство, а $e_1, e_2, \dots, e_n$ и $e_1', e_2', \dots, e_n'$ са два негови базиса. Да наречем първия базис "стар базис", а втория - "нов базис". Нека

$$e_j' = \sum_{i=1}^{n} \alpha_{ij} e_i \ , \ \alpha_{ij} \in \mathcal{K} \ , \quad j = 1, 2, \dots, n \ .$$

Квадратната матрица $T = \left(d_{ij}\right)$ от ред n ще наричаме *матрица на прехода от стария базис към новия*. Тя е неособена. Ако l е произволен вектор, а X и X' означават матрици-стълбове, състоящи се съответно от координатите на вектора l спрямо стария базис и спрямо новия базис, то

$$X = TX' \ .$$

5. Подпространства. Нека L е линейно пространство над поле $\mathcal{K}$, а M е непразно подмножество на L. Ще казваме, че M е *подпространство* на линейното пространство L, ако заедно с всеки два вектора l , $m \in M$, подмножеството M съдържа и всяка тяхна линейна комбинация $\lambda l + \mu m \quad \left(\lambda, \mu \in \mathcal{K}\right)$.

Ще покажем, че *непразното подмножеството $M \subset L$ е подпространство на линейното пространство L тогава и само тогава, когато M е линейно пространство относно същите операции, дефинирани в L* .

Наистина, нека M е подпространство и l, m са вектори от M. Тогава сумата им $l + m$ също принадлежи на M, защото $l + m = 1l + 1m$, а по определение на подпространство дясната страна принадлежи на M. Ако λ е произволен скалар от полето $\mathcal{K}$, пак по определението на подпространство имаме $\lambda l = \lambda l + 0l \in M$. Получихме, че M е затворено относно операциите събиране на вектори и умножаване на вектор с число. Нулевият вектор 0 принадлежи на M, защото $0 = 0l$. За всеки вектор l от M противоположният му вектор $-l$ също е в M, защото $-l = (-1)l$. Останалите аксиоми на линейно пространство, като комутативност на събирането на вектори, асоциативност на събирането и т.н., са изпълнени в M, защото са изпълнени в линейното пространство L, а M е подмножество на L .С това доказахме, че всяко подпространство на линейното пространство L е линейно пространство относно операциите, дефинирани в L.

Обратно, нека M е непразно подмножество на линейното пространство L и нека M е линейно пространство относно операциите, дефинирани в L. Това означава в частност, че за всеки два вектора l, m от M и за всеки два скалара λ, μ от $\mathcal{K}$ векторите $l + m$, λl, μm се съдържат в линейното пространство M, а следователно в M се съдържа и векторът $\lambda l + \mu m$, т.е. M е подпространство.

Във всяко линейно пространство L подмножеството $M = \{0\}$, състоящо се само от нулевия вектор, очевидно е подпространство - така нареченото *нулево подпространство*. Подпространство на L е и самото линейно пространство L. Тези две подпространства обикновено се наричат *несобствени* подпространства.

Най-общ пример на подпространство се строи по следния начин. В линейното пространство L избираме непразно подмножество S (крайно или безкрайно) и нека $l(S)$ е множеството на всички линейни комбинации на краен брой вектори от S. Очевидно $l(S)$ удовлетворява условията на определението и следователно е подпространство на L, съдържащо S. То се нарича *линейна обвивка* на S или *подпространство, породено от* S. В тази терминология условието S да бъде базис на L може да се изкаже така: S е линейно независимо подмножество и линейната му обвивка съвпада с L.

Линейната обвивка на S може да се интерпретира като минималното (относно включването на множества) подпространство на L, съдържащо S. Ще покажем по-точно, че *линейната обвивка* $l(S)$ *е сечението на всички подпространства на* L*, съдържащи* S. Действително, тъй като линейната обвивка $l(S)$ е подпространство на L, съдържащо S, то сечението на всички подпространства, съдържащи S, е подмножество на $l(S)$. От друга страна, всяко подпространство M, съдържащо S, съдържа в частност всички линейни комбинации на краен брой вектори от S, т.е. имаме включването $l(S) \subset M$. Оттук следва, че $l(S)$ е подмножество на сечението на всички подпространства, съдържащи S, което завършва доказателството.

Ако линейната обвивка $l(S)$ е крайномерно линейно пространство, размерността му се нарича ранг на множеството S. В гл. 2, § 2 и § 3 доста подробно се занимавахме с частния случай, когато пространството L е R^n, а S е крайно множество от вектори. Там определението на ранг беше друго, но беше показано, че рангът е равен на размерността на подпространството $l(S)$.

Ако L е крайномерно линейно пространство, то всяко негово подпространство M също е крайномерно и $\dim M \le \dim L$. *Равенството е налице тогава и само тогава, когато $M = L$.*

Доказателството е абсолютно същото както на следствието от т.7, § 2, гл. 2. Пак оттам се пренася и важната

6.Теорема (за допълване на базиса). *Ако L е крайномерно линейно пространство, а M е негово подпространство, то всеки базис на M може да се допълни до базис на цялото пространство L.*

§ 3. Линейни изображения

Всички понятия и резултати от гл. 3 се пренасят буквално и за абстрактни линейни пространства.

1. Определение. Нека L и M са линейни пространства над едно и също поле $\mathcal{K}$. Ще казваме, че изображението $f: L \to M$ е *линейно*, ако:

а) $f(l_1 + l_2) = f(l_1) + f(l_2)$ за всички $l_1, l_2 \in L$;

б) $f(\lambda l) = \lambda f(l)$ за всеки вектор $l \in L$ и за всеки скалар $\lambda \in \mathcal{K}$.

Всяко линейно изображение $f: L \to M$ на линейното пространство L в себе си ще наричаме *линеен оператор*, дефиниран в L.

От определението следва, че всяко линейно изображение f изобразява нулевия вектор на L в нулевия вектор на M. Това е така, защото $f(0) = f(0+0) = f(0) + f(0)$. Пак от определението следва, че за всяка линейна комбинация $x_1 l_1 + x_2 l_2 + \ldots + x_k l_k$, $l_i \in L$, $x_i \in \mathcal{K}$ имаме

$$f(x_1 l_1 + x_2 l_2 + \ldots + x_k l_k) = x_1 f(l_1) + x_2 f(l_2) + \ldots + x_k f(l_k).$$

В гл. 3 бяха дадени достатъчно много примери на линейни изображения в случая, когато L и M са координатни линейни пространства. Ето още един.

Пример. Нека $L = \mathrm{R}[x]$ е линейното пространство на всички полиноми на променливата x с реални коефициенти ($\mathcal{K} = \mathrm{R}$). Дефинираме изображение $d\colon L \to L$ с правилото

$$d\big(f(x)\big) = f'(x),$$

където $f'(x)$ е производната на полинома $f(x)$. Тъй като $\big(f(x)+g(x)\big)' = f'(x)+g'(x)$ и $\big(\lambda f(x)\big)' = \lambda f'(x)$, ясно е, че става дума за линеен оператор; наричат го *оператор на диференциране*. Този пример показва, че в безкрайномерни пространства L описанието на линейните оператори, дефинирани в L, може да бъде много по-заплетено, отколкото в гл. 3.

Видяхме, че съществуват разнообразни примери (модели) на линейни пространства. Принципен въпрос е кои от примерите ще считаме за съществено различни и кои - не. Достатъчно е да решим кога две линейни пространства L и M ще бъдат разглеждани като несъществено различни. Същината на проблема е в операциите, т.е. може ли да се намери съответствие (аналогия) между операциите в L и операциите в M и, разбира се, между елементите на L и M да се установи взаимно еднозначно съответствие, което да "отъждествява" елементите на L с елементите на M. Оказва се, че най-разумното решение на проблема дава следното

2. Определение. Нека L и M са линейни пространства над едно и също поле $\mathcal{K}$. За изображението $f\colon L \to M$ ще казваме, че е *изоморфизъм*, ако:

а) f е линейно изображение;

б) f е биективно.

Ако между L и M съществува поне един изоморфизъм, ще казваме, че те са *изоморфни* и ще пишем $L \cong M$.

В това определение ролите на пространствата L и M са несиметрични. Ще покажем, че обратното изображение $f^{-1}\colon M \to L$ (то съществува, защото f е биективно) също е линейно, а тъй като то също е биективно, ще получим, че f^{-1} е изоморфизъм.

Действително, нека $m_1, m_2 \in M$. Тъй като f е сюрективно, то $m_1 = f(l_1)$, $m_2 = f(l_2)$ при подходящи вектори $l_1, l_2 \in L$. Имаме

$$f^{-1}(m_1+m_2) = f^{-1}\big(f(l_1)+f(l_2)\big) = l_1 + l_2 = f^{-1}(m_2) + f^{-1}(m_2).$$

Аналогично

$$f^{-1}(\lambda m_1) = f^{-1}(\lambda f(l_1)) = f^{-1}(f(\lambda l_1)) = \lambda l_1 = \lambda f^{-1}(m_1).$$

Това обосновава коректността на горното определение, защото виждаме, че ако линейното пространство L е изоморфно на M, то M е изоморфно на L. Ако $f:L \to L'$ и $g:L' \to L''$ са изоморфизми, композицията $g \circ f:L \to L''$ е биективно изображение, защото е композиция на биективни. Неговата линейност се проверява лесно (вж. гл. 3, § 4, т. 3). Това показва, че пространствата L и L'' също са изоморфни. Разбира се, всяко линейно пространство е изоморфно със себе си: достатъчно е да положим $f(l) = l$ за всеки вектор $l \in L$.

Тези бележки показват, че релацията изоморфизъм, която въведохме в съвкупността на всички линейни пространства над едно и също поле $\mathcal{K}$, е релация на еквивалентност. Тя разбива съвкупността на класове на еквивалентност, състоящи се от изоморфни пространства. Принципният момент е, че по-нататък ще разглеждаме изоморфните пространства като несъществено различни; нещо повече, всяко твърдение, доказано за конкретно пространство L, е валидно и за всяко пространство, изоморфно на L. Разбира се, става дума само за твърдения, които се формулират в термините на операциите събиране на вектори и умножаване на вектори с числа от $\mathcal{K}$.

3. Теорема. *Две крайномерни пространства L и M над едно и също поле $\mathcal{K}$ са изоморфни тогава и само тогава, когато имат равни размерности. В частност всяко n-мерно линейно пространство над поле $\mathcal{K}$ е изоморфно на координатното пространство $\mathcal{K}^n$.*

Доказателство. Нека $f:L \to M$ е изоморфизъм и $e_1,\dots,e_n$ е базис на L. Ще покажем, че n-те вектора $f(e_1),\dots,f(e_n)$ са базис на M. Тъй като f е сюрективно, то всеки вектор $l^* \in M$ има вида

$$l^* = f(l) = f\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i f(e_i),$$

т.е. линейната обвивка на векторите $f(e_1),\dots,f(e_n)$ е пространството M. Ще покажем, че те са линейно независими. Действително, нека $\lambda_1 f(e_1) + \dots + \lambda_n f(e_n) = 0$, т.е. $f(\lambda_1 e_1 + \dots + \lambda_n e_n) = 0$. Тъй като f е инективно и $f(0) = 0$, получаваме, че

$$\lambda_1 e_1 + \ldots + \lambda_n e_n = 0.$$

Тъй като $e_1, \ldots, e_n$ са базис на пространството L, то $\lambda_1 = \ldots = \lambda_n = 0$. Следователно векторите $f(e_1), \ldots, f(e_n)$ са базис на M и $\dim M = \dim L$.

Обратно, нека $\dim L = \dim M = n$, а $e_1, \ldots, e_n$ и $e_1^*, \ldots, e_n^*$ са базиси съответно на L и M. Дефинираме изображение $f : L \to M$ с правилото: ако $l = x_1 e_1 + \ldots + x_n e_n$, то $f(l) = x_1 e_1^* + \ldots + x_n e_n^*$. Определението е коректно, защото всеки вектор $l \in L$ се изразява еднозначно като линейна комбинация на базисните вектори. Проверката, че f е линейно изображение, вече правихме (гл. 3, § 1, т. 4). Сюрективността на f е очевидна. Очевидно е също, че $f(l) = 0$ точно когато $l = 0$, защото $e_1^*, \ldots, e_n^*$ са базис на пространството M. Следователно, ако $f(l_1) = f(l_2)$, то $0 = f(l_1) - f(l_2) = f(l_1 - l_2)$ и $l_1 - l_2 = 0$, т.е. изображението f е инективно. С това доказахме, че f е изоморфизъм. Теоремата е доказана.

Нека $M = \mathcal{K}^n$ е n-мерното координатно линейно пространство и $e_1^*, \ldots, e_n^*$ е неговият стандартен базис. При горните означения имаме

$$f(x_1 e_1 + \ldots + x_n e_n) = x_1 e_1^* + \ldots + x_n e_n^* = (x_1, \ldots, x_n),$$

т.е. на всеки вектор съпоставяме наредената n-торка от неговите координати спрямо избрания базис. Така дефинираният изоморфизъм обикновено се нарича *координатен*. Той лежи в основата на така наречения *координатен метод*, който най-широко се използва в геометрията.

Доказаната теорема показва, че $\mathcal{K}^n$ е универсален модел на крайномерно пространство над поле $\mathcal{K}$. В този смисъл можем ясно да кажем, че в гл. 2 отчасти, а в гл. 3 - изцяло, изучавахме теорията на абстрактните крайномерни линейни пространства, като за удобство там използвахме конкретен модел - координатното пространство $\mathcal{K}^n$. За да не отвличаме вниманието на читателя, предполагахме, че $\mathcal{K} = \mathrm{R}$ или C. Бегло препрочитане ще ни убеди, че всички резултати остават в сила и ако $\mathcal{K}$ е произволно числово поле.

Понятията ядро и образ на линейното изображение $f:L\to M$ се въвеждат и за произволни линейни пространства (включително и безкрайномерни) над поле $\mathcal{K}$:

$$\mathrm{Ker} f=\{l\in L\mid f(l)=0\},\quad \mathrm{Im}\, f=\{f(l)\mid l\in L\}.$$

Доказателството, че *ядрото* $\mathrm{Ker} f$ и *образът* $\mathrm{Im} f$ са подпространства съответно на L и M е познато. В него размерността не играе никаква роля. Напомняме обаче (гл. 3, § 2), че в крайномерния случай

$$\dim\,\mathrm{Ker} f+\dim\,\mathrm{Im} f=\dim L.$$

4. Линейни изображения и матрици. Нека отново L и M са крайномерни линейни пространства над числово поле $\mathcal{K}$. С $\mathcal{L}(L,M)$ означихме множеството на всички линейни изображения на L в M (вж. § 1, т. 4, д). Както видяхме, относно операциите сума на линейни изображения и произведение на линейно изображение с число то е линейно пространство над $\mathcal{K}$. В гл. 3 показахме как при избрани базиси съответно на L и M на всяко линейно изображение $f:L\to M$ може да се съпостави еднозначно определена матрица A от тип $m\times n$ с елементи от $\mathcal{K}$, т.е. $A\in M_{m\times n}(\mathcal{K})$. Напомняме, че $m=\dim M$, $n=\dim L$.

Иначе казано, имаме изображението

$$\varphi:\mathcal{L}(L,M)\to M_{m\times n}(\mathcal{K}),$$

при което $\varphi(f)=A$. Ще покажем, че то е биективно. За целта нека $e_1,\dots,e_n$ и $e_1^*,\dots,e_n^*$ са съответните избрани базиси на пространствата L и M. Както знаем, стълбовете на матрицата на изображението f са координатите на векторите $f(e_1),\dots,f(e_n)$ спрямо избрания базис на пространството M. От твърдението в т. 4, § 1, гл. 3 следва, че ако f и g са различни линейни изображения на L в M, то съществува номер i, за който $f(e_i)\neq g(e_i)$. Това означава, че i-тият стълб на матрицата $\varphi(f)$ е различен от i-тия стълб на матрицата $\varphi(g)$, т.е. $\varphi(f)\neq\varphi(g)$ и изображението φ е инективно. Ще покажем, че то е и сюрективно. За целта нека A е произволна матрица от $M_{m\times n}(\mathcal{K})$ и да означим с $g_1,g_2,\dots,g_n$ онези вектори от M, чиито координати са съответно първия, втория и т.н. стълбове на матрицата A. Пак според твърдението в т. 4, § 1, гл. 3 съществува единствено линейно изображение $f:L\to M$, за което $f(e_i)=g_i$, $i=1,2,\dots,n$. Матрицата на това изображение е очевидно матрицата A, т.е. $\varphi\left(f\right)=A$ и φ е сюрективно.

Ако $g\colon L\to M$ е още едно линейно изображение с матрица B, то както знаем $\varphi\left(f+g\right)=A+B=\varphi\left(f\right)+\varphi\left(g\right)$ и $\varphi\left(\lambda f\right)=\lambda\varphi\left(f\right)$ ($\lambda\in\mathcal{K}$). От абстрактната гледна точка, на която застанахме, можем просто да кажем, че φ е изоморфизъм на линейни пространства, т.е. $\mathcal{L}(L,M)\cong M_{m\times n}(\mathcal{K})$. В много случаи този изоморфизъм е особено полезен, защото дава възможност операциите с линейни изображения да се сведат до операции с матрици. Ползата от това не бива да се надценява, защото изоморфизмът φ зависи от случайния избор на базисите в L и M, което води до известни неудобства.

§ 4. Сума на подпространства

1. Нека M_1 и M_2 са подпространства на линейно пространство L над поле $\mathcal{K}$. Без значение е дали L е крайномерно или не. С M_1+M_2 ще означим множеството на всички вектори l от L, които се представят във вида $l=l_1+l_2$, където, $l_1\in M_1$, $l_2\in M_2$, т.е.

$$M_1+M_2=\left\{l_1+l_2 \mid l_1\in M_1, l_2\in M_2\right\}$$

Ще проверим, че множеството M_1+M_2 е подпространство на L. Наистина, за всеки два вектора $l=l_1+l_2$, $m=l_1'+l_2'$, за които $l_1,l_1'\in M_1$, $l_2,l_2'\in M_2$, и числа $\lambda,\mu\in\mathcal{K}$, имаме

$$\lambda l+\mu m=\lambda\left(l_1+l_2\right)+\mu\left(l_1'+l_2'\right)=\left(\lambda l_1+\mu l_1'\right)+\left(\lambda l_2+\mu l_2'\right),$$

Тази линейна комбинация е в M_1+M_2, защото двете събираеми в най-дясната страна са съответно в подпространствата M_1 и M_2.

По-нататък за подпространството M_1+M_2 ще казваме, че е *сума на подпространствата* M_1 *и* M_2. По-общо, ако $M_1,M_2,\ldots,M_k$ са подпространства на L, тяхната *сума* се дефинира като

$$M_1+M_2+\ldots+M_k=\left\{\sum_{i=1}^{k} l_i \mid i=1,2,\ldots,k\right\}.$$

Рутинната проверка, че $M_1 + M_2 + \ldots + M_k$ е подпространство на L, пропускаме. (При $k = 2$ вече я направихме.)

2. Теорема. *Нека* L *е крайномерно линейно пространство, а* M_1 *и* M_2 са негови подпространства. Тогава

$$\dim\left(M_1 + M_2\right) = \dim M_1 + \dim M_2 - \dim\left(M_1 \cap M_2\right).$$

Доказателство. Ако $M_1 \subset M_2$, то $M_1 + M_2 = M_2$, $M_1 \cap M_2 = M_1$ и твърдението очевидно е вярно. Същото важи и ако $M_2 \subset M_1$.

По-нататък ще предполагаме, че нито едно от включванията $M_1 \subset M_2$ и $M_2 \subset M_1$ не е налице и полагаме $M = M_1 \cap M_2$. Тъй като сечението на подпространства е подпространство, то M е подпространство на L. За определеност предполагаме, че M не е нулевото подпространство и нека $e_1, \ldots, e_k$ е базис на M. Тъй като M е подпространство на M_1, можем да допълним базиса му до базис $e_1, \ldots, e_k, e'_{k+1}, \ldots, e'_r$ на пространството M_1. (При нашите предположения $M \neq M_1$ и следователно $r > k$.) Тъй като M е подпространство и на пространството M_2, $M \neq M_2$, отново допълваме базиса му до базис $e_1, \ldots, e_k, e''_{k+1}, \ldots, e''_s$ на M_2. При тези означения $\dim M_1 = r$, $\dim M_2 = s$, $\dim M = k$. Ако покажем, че векторите

(1) $$e_1, \ldots, e_k, e'_{k+1}, \ldots, e'_r, e''_{k+1}, \ldots, e''_s$$

(броят им е $r + s - k$) са базис на $M_1 + M_2$, с това ще бъде доказана и теоремата. Ясно е, че всеки вектор от $M_1 + M_2$ е линейна комбинация на векторите от системата (1), защото тя съдържа базиси както на M_1, така и на M_2. Остава да се убедим, че (1) е линейно независима система. Нека

(2) $$\sum_{i=1}^{k} \alpha_i e_i + \sum_{j=k+1}^{r} \beta_j e'_j + \sum_{t=k+1}^{s} \gamma_t e''_t = 0$$

с някакви числа α_i, β_j, γ_t от полето $\mathcal{K}$. Да препишем (2) във вида

(3) $$\sum_{i=1}^{k} \alpha_i e_i + \sum_{j=k+1}^{r} \beta_j e'_j = -\sum_{t=k+1}^{s} \gamma_t e''_t .$$

Очевидно векторът от лявата страна на (3) е в M_1, а векторът от дясната страна - в M_2. И така в лявата страна на (3) стои вектор от $M_1 \cap M_2 = M$. Следователно

$$\sum_{i=1}^{k} \alpha_i e_i + \sum_{j=k+1}^{r} \beta_j e'_j = \sum_{i=1}^{k} \delta_i e_i$$

с някакви числа δ_i от $\mathcal{K}$. Тъй като всеки вектор от M_1 се записва еднозначно чрез базиса на M_1, от последното равенство следва, че $\alpha_i = \delta_i$, $i = 1,\dots,k$ и $\beta_j = 0$, $j = k+1,\dots,r$. Тъй като всеки вектор от M_2 се записва еднозначно чрез базиса на M_2, сега от (3) следва, че $\alpha_i = 0$, $i = 1,\dots,k$ и $\gamma_t = 0$, $t = k+1,\dots,s$. С това показахме, че равенството (2) е възможно само ако всички коефициенти α_i, β_j, γ_t са равни на нула, т.е. системата (1) е базис на пространството $M_1 + M_2$.

Ако сечението $M = M_1 \cap M_2$ е нулевото подпространство, разсъжденията са аналогични с единствената разлика, че базис на M е празното множество от вектори.

З а б е л е ж к и . а) Заключението на теоремата остава в сила и ако L е произволно (не непременно крайномерно) линейно пространство, а M_1 и M_2 са негови крайномерни подпространства. Доказателството е същото.

б) Нека L е крайномерно линейно пространство , а M_1 и M_2 са негови подпространства. Тъй като $M_1 + M_2$ е подпространство на L, то според теоремата от края на т. 5, § 2 винаги имаме неравенството $\dim\left(M_1 + M_2\right) \le \dim L$. Като използваме току -що доказаната теорема, получаваме неравенството

$$\dim M_1 + \dim M_2 - \dim L \le \dim\left(M_1 \cap M_2\right).$$

От него следва например, че в тримерното пространство L сечението на всеки две двумерни подпространства има размерност поне единица, че в четиримерно пространство сечението на всеки две тримерни подпространства има размерност поне две и т.н.

в) Ако сечението $M = M_1 \cap M_2$ не е нулевото подпространство, произволен вектор $l \in M_1 + M_2$ може по различни начини да се запише във вида $l = l_1 + l_2$, $l_i \in M_i$. Действително, ако l' е произволен ненулев вектор от M, то $l'_1 = l_1 + l'$ и $l'_2 = l_2 - l'$ са вектори съответно от M_1 и M_2, различни са от съответните l_1 и l_2 и $l = l_1 + l_2 = l'_1 + l'_2$

3. Директни суми. Нека L е произволно линейно пространство (без ограничения за размерността му), а $M_1, M_2, \ldots, M_k$ са негови подпространства. Ще казваме, че сумата

$$M_1 + M_2 + \ldots + M_k$$

е *директна*, ако всеки вектор l от нея се представя еднозначно във вида

(4) $$l = l_1 + l_2 + \ldots + l_k, \quad l_i \in M_i, \quad i = 1, 2, \ldots, k,$$

т.е. ако от равенствата (4) и

(5) $$l = l_1' + l_2' + \ldots + l_k', \quad l_i' \in M_i, \quad i = 1, 2, \ldots, k$$

винаги следва $l_1 = l_1'$, $l_2 = l_2'$, ..., $l_k = l_k'$. Ако сумата е директна, ще пишем

$$M_1 \oplus M_2 \oplus \ldots \oplus M_k \quad \text{или} \quad \bigoplus_{i=1}^{k} M_i .$$

Според забележката по-горе необходимо условие сумата $M_1 + M_2$ да е директна е $M_1 \cap M_2 = \{0\}$. Ще покажем, че това условие е и достатъчно. Наистина, нека $l = l_1 + l_2 = l_1' + l_2'$, където $l_i, l_i' \in M_i$, $i = 1, 2$. От това равенство следва, че $l_1 - l_1' = l_2' - l_2$. Лявата страна е вектор от M_1, а дясната - от M_2, т.е. $l_1 - l_1' \in M_1 \cap M_2 = \{0\}$. Следователно $l_1 = l_1'$, $l_2 = l_2'$ и представянето е еднозначно.

Ще забележим допълнително, че ако M_1 и M_2 са крайномерни подпространства, от теоремата в т. 2 следва, че $\dim\left(M_1 + M_2\right) = \dim M_1 + \dim M_2$ тогава и само тогава, когато $\dim\left(M_1 \cap M_2\right) = 0$, т.е. когато $M_1 \cap M_2 = \{0\}$.

С това доказахме

4. Твърдение. *Сумата* $M_1 + M_2$ *е директна тогава и само тогава, когато* $M_1 \cap M_2 = \{0\}$. *Ако подпространствата* M_1 *и* M_2 *са крайномерни, сумата им* $M_1 + M_2$ *е директна тогава и само тогава, когато* $\dim\left(M_1 + M_2\right) = \dim M_1 + \dim M_2$.

Твърдението може да се обобщи и за повече от две подпространства, а именно: сумата $M_1 + M_2 + \ldots + M_k$ *е директна тогава и само тогава, когато сечението на всяко от подпространствата* M_i *със сумата на останалите е нулевото подпространство, т.е.*

$$M_i \cap \sum_{\substack{j=1 \\ j \neq i}}^{k} M_j = \{0\}, \quad i = 1, 2, \ldots, k$$

Доказателството оставяме за упражнение.

Примери. а) Нека L е n-мерно линейно пространство и $e_1, e_2, \ldots, e_n$ е някакъв негов базис. За всяко $i = 1, 2, \ldots, n$ да означим с M_i едномерното подпространство, породено от вектора e_i, т.е.

$$M_i = \{\lambda_i e_i \mid \lambda_i \in K\}.$$

Непосредствено от определението на базис следва, че

$$L = M_1 \oplus M_2 \oplus \ldots \oplus M_n,$$

т.е. *всяко* n*-мерно пространство може да се представи като директна сума на* n *едномерни подпространства.*

б) Нека отново L е крайномерно линейно пространство и M е негово подпространство. Ако $M = \{0\}$, полагаме $M' = L$, а ако $M = L$ - полагаме $M' = \{0\}$. Нека M е собствено подпространство. Да изберем базис $e_1, \ldots, e_k$ на M и да го допълним до базис $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ на L. (Това може да се направи съгласно теоремата за допълване на базиса от § 2, т. 6.) Да означим с M' подпространството, породено от векторите $e_{k+1}, \ldots, e_n$. Те са линейно независими и от определението на линейна обвивка следва, че са базис на M'. Почти очевидно е, че във всички случаи имаме

$$L = M \oplus M'.$$

(Всеки вектор $l \in L$ се представя еднозначно във вида $l = (\lambda_1 e_1 + \ldots + \lambda_k e_k) + (\lambda_{k+1} e_{k+1} + \ldots + \lambda_n e_n)$; "излишните" скоби поставихме, за да разграничим двете събираеми съответно от M и M'.) Подпространството M' се нарича *директно допълнение* на M.

За подпространството M на L се казва, че е *директно събираемо на* L, ако съществува подпространство M', за което $L = M \oplus M'$. Всъщност показахме, че в крайномерния случай *всяко подпространство е директно събираемо.* Същото е вярно и за безкрайномерни пространства. За да се установи този факт, е необходимо да се докаже с помощта на лемата на Цорн, че теоремата за допълване на базиса е вярна и в този случай.

Ще подчертаем, че директното допълнение M' на подпространството M не е еднозначно определено - причината е, че допълването на базиса на M до базис на L може да се направи по различни начини. Възможността да "разглобим" пространството L на директна сума от две подпространства с по-ниски размерности често е удобна при разсъждения с индукция по $\dim L$.

5. Проектори. С всяко директно разлагане $L = M_1 \oplus M_2$ се свързват два линейни оператора $p_i\colon L \to L$, $i = 1, 2$, дефинирани по следния начин: ако $l = l_1 + l_2$, $l_1 \in M_1$, $l_2 \in M_2$, то

$$p_1(l) = l_1, \quad p_2(l) = l_2.$$

Изображенията p_i, наречени *проектори*, са коректно дефинирани, защото за всеки вектор l компонентите му l_1 и l_2 са еднозначно определени. Твърдението, че p_i са линейни оператори, е очевидно. Ще забележим, че

$$p_i \circ p_i(l) = p_i(p_i(l)) = p_i(l_i) = l_i = p_i(l),$$

$$p_1 \circ p_2(l) = p_1(p_2(l)) = p_1(l_2) = 0, \quad p_2 \circ p_1(l) = p_2(l_1) = 0$$

$$(p_1 + p_2)(l) = p_1(l) + p_2(l) = l_1 + l_2 = l,$$

за всеки вектор $l \in L$, което означава, че

$$p_i^2 = p_i \circ p_i = p_i, \quad p_1 p_2 = p_2 p_1 = 0, \quad p_1 + p_2 = \mathcal{E}$$

(С $\mathcal{E}$ означаваме единичния оператор: $\mathcal{E}(l) = l$ за всеки вектор $l \in L$.)

Това подсказва следното по-общо

6. Определение. Ще казваме, че линейният оператор $p\colon L \to L$ е *проектор* (*оператор на проектиране*), ако $p^2 = p \circ p = p$.

Следващото твърдение показва, че с всеки проектор p може да се асоциира съответно директно разлагане на L.

7. Твърдение. *Нека $p\colon L \to L$ е проектор и нека*

$$M = \operatorname{Im} p = \left\{ p\left(l\right) \mid l \in L \right\},$$

$$M' = \operatorname{Im}(\mathcal{E} - p) = \{l - p(l) \mid l \in L\}.$$

Тогава $L = M \oplus M'$.

Доказателство. В гл. 3, § 2 вече стана дума, че образът на линейно изображение е подпространство, следователно M и M' са подпространства на L. Ще покажем, че сечението им е нулевото подпространство. Наистина, всеки вектор от $M \cap M'$ има вида

(5) $$p\left(l\right) = l' - p\left(l'\right)$$

при подходящи $l, l' \in L$. Лявата страна означава, че векторът е от M, а дясната - че е от M'. Ако към двете страни на (5) приложим оператора p и отчетем, че $p^2 = p$, ще получим

$$p\left(l\right) = p\left(l'\right) - p\left(l'\right) = 0,$$

което показва, че $M \cap M' = \left\{0\right\}$. Според твърдението в т. 4 сумата $M + M'$ е директна. Очевидното равенство $l = p\left(l\right) + \left(l - p\left(l\right)\right)$ за всеки вектор l от L показва, че $M \oplus M' = L$.

Обобщение на това твърдение е дадено в упражненията.

8. Външни директни суми. Сумата на подпространства може да се разглежда като конструкция, с помощта на която от няколко подпространства на едно и също пространство построихме ново подпространство - сумата им. Аналогична конструкция е възможна и ако $L_1, \ldots, L_k$ са произволни линейни пространства над едно и също поле $\mathcal{K}$, без да се иска да са подпространства на едно и също пространство.

И така, нека $L = L_1 \times L_2 \times \ldots \times L_k$ е декартовото произведение на множествата $L_1, L_2, \ldots, L_k$, т.е. L се състои от всички наредени k-орки $(l_1, l_2, \ldots, l_k)$, в които $l_i \in L_i$, $i = 1, 2, \ldots, k$. Както обичайно, ще считаме, че две такива k-орки са равни точно тогава, когато всичките им съответни компоненти са равни. Дефинираме:

$$\left(l_1, \ldots, l_k\right) + \left(l'_1, \ldots, l'_k\right) = \left(l_1 + l'_1, \ldots, l_k + l'_k\right),$$

$$\lambda\left(l_1, \ldots, l_k\right) = \left(\lambda l_1, \ldots, \lambda l_k\right), \quad \lambda \in \mathcal{K}.$$

Лесно се проверява, че относно така дефинираните операции L е линейно пространство над $\mathcal{K}$. Проверката по нищо не се отличава от онази, която правихме в координатното пространство и тук я пропускаме. Така дефинираното линейно пространство $L = L_1 \times L_2 \times \ldots \times L_k$ се нарича *външна директна сума на пространствата* $L_1, L_2, \ldots, L_k$. Терминологията се мотивира от следното съображение. Нека за всяко $i = 1, 2, \ldots, k$ положим

$$\overline{L_i} = \left\{ \left(0, \ldots, l_i, \ldots, 0\right) \mid l_i \in L_i \right\},$$

където векторът l_i е на i -то място. Очевидно $\overline{L_i}$ е подпространство на L и

$$L = \overline{L_1} \oplus \overline{L_2} \oplus \ldots \oplus \overline{L_k}.$$

Да дефинираме изображение $f_i : L_i \to \overline{L_i}$ с правилото $f_i\left(l_i\right) = \left(0, \ldots, l_i, \ldots, 0\right)$. То очевидно е изоморфизъм, който отъждествява всеки вектор $l_i \in L_i$ с вектора $\left(0, \ldots, l_i, \ldots, 0\right)$ от $\overline{L_i}$. Благодарение на този изоморфизъм бихме могли да не правим разлика между пространствата L_i и $\overline{L_i}$ и да считаме просто, че $L_i \subset L$. В този смисъл понятията външно и вътрешно произведение съвпадат.

§ 5. Факторпространства

1. Нека L е произволно линейно пространство над поле $\mathcal{K}$, а M е негово подпространство. Целта ни е да изложим една друга конструкция, с помощта на която от L и M се строи трето линейно пространство.

В множеството L въвеждаме двучленна релация $\sim$ с помощта на следното определение: ще считаме, че два вектора l_1 и l_2 са *еквивалентни по модул подпространството* M и ще пишем $l_1 \sim l_2$ точно тогава, когато $l_1 - l_2 \in M$. Ще проверим, че $\sim$ е релация на еквивалентност. Наистина, $l \sim l$ за всеки вектор $l \in L$, защото $l - l = 0 \in M$, т.е. релацията е рефлексивна. Тя е симетрична, защото, ако $l_1 \sim l_2$, то $l_1 - l_2 \in M$, следователно $l_2 - l_1 = -\left(l_1 - l_2\right)$ също е в M (M е подпространство!), т.е. $l_2 \sim l_1$. Най-сетне нека $l_1 \sim l_2$ и $l_2 \sim l_3$. Тогава

$l_1 - l_3 = (l_1 - l_2) + (l_2 - l_3) \in M$, защото двете събираеми $l_1 - l_2$ и $l_2 - l_3$ по условие са в подпространството M. Следователно $l_1 \sim l_3$ (релацията е транзитивна).

Въведената релация на еквивалентност има допълнително и следните две важни свойства, които се проверяват непосредствено:

а) ако $l_1 \sim l_2$ и $l_1' \sim l_2'$, то $l_1 + l_1' \sim l_2 + l_2'$;

б) ако $l_1 \sim l_2$, то $\lambda l_1 \sim \lambda l_2$ за всеки скалар $\lambda \in \mathcal{K}$.

Свойствата а) и б) обикновено се резюмират с думите, че *релацията на еквивалентност ~ е съгласувана с операциите* в линейното пространство L.

Както знаем, всяка релация на еквивалентност разбива множеството, в което е въведена, на две по две непресичащи се подмножества, наречени класове на еквивалентност. Напомняме, че два вектора l_1 и l_2 са в един и същи клас точно тогава, когато са еквивалентни. За всеки $l \in L$ с $\overline{l}$ ще означаваме класа на еквивалентност, съдържащ l; елементът l се нарича *представител* на класа. Всеки клас се определя еднозначно от който и да е свой представител, т.е. $\overline{l} = \overline{l'}$ точно тогава, когато $l \sim l'$. Означаваме

$$L / M = \left\{ \overline{l} \mid l \in L \right\},$$

т.е. елементите на множеството L / M са различните класове на еквивалентност. Класът $\overline{l}$ може да бъде описан по-нагледно и като множеството

$$\overline{l} = l + M = \left\{ l + m \mid m \in M \right\}$$

Действително, $l \sim l + m$ за всяко $m \in M$, което означава, че всеки вектор от $l + M$ е еквивалентен на l. Обратно, ако $l_1 \sim l$, то $l_1 - l \in M$, т.е. $l_1 = l + m$ за някой вектор $m \in M$. Следователно $\overline{l} = l + M$.

Ще отбележим, че $l + M = l' + M$ тогава и само тогава, когато $l - l' \in M$. Това е така, защото $\overline{l} = \overline{l'}$ точно тогава, когато $l \sim l'$.

Ще превърнем множеството L / M в линейно пространство над K, дефинирайки операции с равенствата

(1) $$\overline{l_1} + \overline{l_2} = \overline{l_1 + l_2}, \quad \lambda \overline{l} = \overline{\lambda l}, \quad . \lambda \in \mathcal{K}.$$

Засега това определение изглежда съмнително: сумата на два класа би трябвало да зависи само от самите класове, а не от случайния избор на представителите им, както е заложено в (1). Иначе казано, ако в (1) сменим представителите на класовете, дали резултатът от събирането и умножаването със скалар ще бъдат същите класове? Ще проверим, че определенията наистина са коректни. Действително, нека $\bar{l}_1 = \bar{l}_1'$, $\bar{l}_2 = \bar{l}_2'$. Всяко от двете равенства е равносилно на $l_1 \sim l_1'$ и $l_2 \sim l_2'$. Съгласно споменатото по-горе свойство а) имаме $l_1 + l_2 \sim l_1' + l_2'$, т.е. $\overline{l_1 + l_2} = \overline{l_1' + l_2'}$. Следователно събирането е коректно дефинирано. Като използваме б), от $\bar{l} = \bar{l}'$, което е равносилно на $l \sim l'$, получаваме $\lambda l \sim \lambda l'$ или $\lambda\bar{l} = \overline{\lambda l'}$. Това показва, че умножението със скалар също е коректно дефинирано.

Педантичната проверка, че в L/M са в сила осемте аксиоми за линейно пространство, ще оставим за упражнение. Ще се ограничим само с някои пояснения. Нулевият вектор в L/M е класът $\bar{0} = 0 + M = M$, където 0 е нулевият вектор на L. Това е така, защото $\bar{l} + \bar{0} = \overline{l + 0} = \bar{l}$. Противоположният на вектора $\bar{l}$ е $\overline{-l}$.

2. Определение. Множеството L/M заедно с операциите, дефинирани с (1), е линейно пространство над $\mathcal{K}$; ще го наричаме *факторпространство на* L *по подпространството* M.

З а б е л е ж к а . Релацията на еквивалентност, която въведохме в L, зависи от подпространството M. Естествен е въпросът, дали не би могло да се действа по по-обща схема, като си послужим с произволна релация на еквивалентност ~. Тя разбира се ще разбие L на класове на еквивалентност и нека $L/\sim$ е факторможеството. Въвеждайки операции в $L/\sim$, искаме не просто да го превърнем в линейно пространство, а освен това то да наследи операциите от L, т.е. стигаме до определенията (1). За да докажем тяхната коректност, ще се окаже, че е необходимо релацията да бъде съгласувана с операциите в L. Но както показва зад. 11 от упражненията, тогава тя задължително е "еквивалентност по модул някакво подпространство M ".

Факторпространствата са особено полезен инструмент в линейната алгебра. В следващите точки ще направим някои приложения.

3. Канонично изображение $L \to L/M$. При означенията от т. 1 дефинираме изображение $\varphi: L \to L/M$ с равенството

$$\varphi(l) = \bar{l}, \quad l \in L.$$

То е линейно, защото

$$\varphi(l_1+l_2)=\overline{l_1+l_2}=\overline{l_1}+\overline{l_2}=\varphi(l_1)+\varphi(l_2),$$

$$\varphi(\lambda l)=\overline{\lambda l}=\lambda\overline{l}=\lambda\varphi(l).$$

Очевидно е, че φ е сюрективно изображение.

Изображението φ се нарича *канонично линейно изображение* на пространството L върху факторпространството L/M. Вместо линейно изображение често се използва и терминът *хомоморфизъм*.

Ядрото на φ е подпространството M. Действително, равенството $\varphi(l)=\overline{l}=\overline{0}$, т.е. $l+M=M$, е изпълнено точно тогава, когато $l\in M$. Следователно $\operatorname{Ker}\varphi=M$.

Дотук конструкцията и останалите разсъждения не зависеха от размерността. Сега да предположим, че L е крайномерно и нека $\dim L=n$. Тъй като в L всеки $n+1$ вектора са линейно зависими, образите им под действие на линейното изображение φ също са линейно зависими. Но φ е сюрективно, следователно във факторпространството L/M всеки $n+1$ вектора са линейно зависими, т.е. то е крайномерно. Съгласно теоремата от т. 4, § 2, гл. 3

$$\dim\operatorname{Im}\varphi+\dim\operatorname{Ker}\varphi=\dim L.$$

Тук $\operatorname{Im}\varphi=L/M$, $\operatorname{Ker}\varphi=M$ и от горното равенство следва, че

$$\dim(L/M)=\dim L-\dim M$$

С това доказахме

4.Твърдение. *Ако линейното пространство L е крайномерно, то факторпространството L/M също е крайномерно и*

$$\dim(L/M)=\dim L-\dim M$$

Често пъти това елементарно твърдение открива път за доказателства на различни теореми с помощта на индукция по $\dim L$. Механизмът е прост: ако твърдението е проверено за подпространства с по-малка размерност и ако за L/M при $M\neq\{0\}$ са приложими условията на съответното твърдение, то за L/M е в сила индуктивното допускане, защото $\dim(L/M) < \dim L$. Оттук често пъти лесно се прави изводът, че твърдението е вярно и за L.

Размерността на L/M обикновено се нарича *коразмерност* на подпространството M.

Ще докажем следната теорема

5. Теорема за хомоморфизмите. *Нека L и L' са линейни пространства над едно и също поле* K*, а $f{:}L \to L'$ е сюрективно линейно изображение. Тогава пространствата $L/\mathrm{Ker}f$ и L' са изоморфни. По-точно съществува, и то единствен, изоморфизъм $g{:}L/\mathrm{Ker}f \to L'$, за който $f = g \circ \varphi$, където $\varphi{:}L \to L/\mathrm{Ker}f$ е каноничният хомоморфизъм.* Нагледно:

$$\begin{tikzcd} L \arrow[r, "\varphi"] \arrow[dr, "f"'] & L/\mathrm{Ker}f \arrow[d, "g"] \\ & L' \end{tikzcd}$$

(За горната диаграма се казва, че е комутативна, т.е. резултатът от "движението"от L към L' не зависи от това по посока на кои стрелки ще се "движим". Точният смисъл е, че $f = g \circ \varphi$.)

Доказателство. Запазваме означенията от т. 1, т. 2 и т. 3. За краткост полагаме $\mathrm{Ker}f = M$; образа $f(l) \in L'$ на вектора $l \in L$ ще означаваме с l'. Ако исканият изоморфизъм g наистина съществува, би трябвало $f(l) = l' = g\varphi(l) = g(\bar{l})$, т.е. $g(\bar{l}) = l'$. От това равенство е ясно, че ако g съществува, той е единствен. Но то подсказва и как да дефинираме g. За целта нека $l' \in L'$. Тъй като изображението f е сюрективно, то съществува вектор $l \in L$, за който $f(l) = l'$. Ако l_1 и l_2 са два вектора от L, за които $f(l_1) = f(l_2)$, то $f(l_1 - l_2) = 0$, т.е. $l_1 - l_2 \in \mathrm{Ker}f = M$. Това означава, че $\bar{l}_1 = \bar{l}_2$. Иначе казано, ако векторът $l' \in L'$ е фиксиран, множеството на векторите $l \in L$, за които $f(l) = l'$, е единствен клас на еквивалентност $\bar{l} \in L/M$. Сега дефинираме: $g(\bar{l}) = l'$, ако $f(l) = l'$. Вече сме сигурни, че това определение е коректно и $f = g \circ \varphi$. Остава да се убедим, че изображението g е линейно и биективно. По самото си определение то е сюрективно. Вече стана дума, че за всеки вектор $l' \in L'$ съществува единствен $\bar{l} \in L / M$ такъв, че $g\left(\bar{l}\right) = l'$, т.е. g е инективно. Нека $f\left(l_1\right) = l'_1$, $f\left(l_2\right) = l'_2$. Тъй като f е линейно, то $f\left(\lambda l_1 + \mu l_2\right) = \lambda l'_1 + \mu l'_2$. Следователно

$$g\left(\lambda \bar{l}_1 + \mu \bar{l}_2\right) = g\left(\overline{\lambda l_1 + \mu l_2}\right) = \lambda l'_1 + \mu l'_2 = \lambda g\left(\bar{l}_1\right) + \mu g\left(\bar{l}_2\right),$$

което доказва линейността на g.

На пръв поглед чисто формална, доказаната теорема има доста дълбок смисъл: вместо да изучаваме образите L' на L при всевъзможните сюрективни хомоморфизми на L в различни линейни пространства, достатъчно е да изучаваме само всевъзможните факторпространства L / M (когато меним подпространството M).

6. Базиси. Нека линейното пространство L е крайномерно и, както до сега, M е негово подпространство. Ако $e_1, \ldots, e_n$ е произволен базис на L ($n = \dim L$) и $M \neq \{0\}$, съответните вектори $\bar{e}_1, \ldots, \bar{e}_n$ в L / M са линейно зависими. Това е ясно, защото според твърдението в т. 4 размерността на факторпространството е по-малка от n. Същото може да се съобрази и непосредствено: нека m е ненулев вектор от M и $m = \lambda_1 e_1 + \ldots + \lambda_n e_n$. Не всички скалари λ_i са нули, защото $m \neq 0$. Равенството се трансформира в $\bar{0} = \lambda_i \bar{e}_1 + \ldots + \lambda_n \bar{e}_n$, което показва, че $\bar{e}_1, \ldots, \bar{e}_n$ са линейно зависими. Въпреки това, линейната им обвивка е цялото факторпространство L / M, защото L се състои от всички линейни комбинации на $e_1, \ldots, e_n$. Ако от системата вектори $\bar{e}_1, \ldots, \bar{e}_n$ изберем линейно независима подсистема, в която броят на векторите да бъде възможно най-голям (максимална линейно независима подсистема), ще получим базис на L / M.

Ако $M = \{0\}$, пространствата L и L / M са изоморфни (даже съвпадат), защото сега $l + M = \bar{l} = l$ и всеки базис на L е базис и на факторпространството L / M.

Ако $M = L$, факторпространството L / L е нулевото пространство и негов базис е празното множество от вектори.

Нека M е собствено подпространство на линийното пространство L, т.е. $\{0\} \neq M \neq L$. Удобно е да се избере базис $e_1, \ldots, e_k$ на M и след това да се допълни до базис $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ на пространството L. Тогава $\bar{e}_1 = \ldots = \bar{e}_k = \bar{0}$, защото $e_1, \ldots, e_k$ са в M. Твърдим, че векторите

$\overline{e}_{k+1},\dots,\overline{e}_n$ са базис на факторпространството. Според казаното по-горе, достатъчно е да се убедим, че те са линейно независими. Нека

$$\lambda_{k+1}\overline{e}_{k+1}+\dots+\lambda_n\overline{e}_n=\overline{0},\quad \lambda_i\in\mathsf{K}$$

Това означава, че векторът $m=\lambda_{k+1}e_{k+1}+\dots+\lambda_n e_n$ принадлежи на M, следователно $m=\mu_1 e_1+\dots+\mu_k e_k$, $\mu_j\in\mathsf{K}$ и

$$\lambda_{k+1}e_{k+1}+\dots+\lambda_n e_n=\mu_1 e_1+\dots+\mu_k e_k.$$

Тъй като векторите $e_1,\dots,e_n$ са базис на L, от последното равенство следва, че всички скалари λ_i и μ_j са нули.

Забележка. Тъй като $k=\dim M$, от това разсъждение още веднъж следва, че $\dim\left(L/M\right)=\dim L-\dim M$. Нещо повече, ако $f:L\to L'$ е произволно линейно изображение на L в произволно пространство L' (и двете трябва да бъдат над едно и също поле!), то $f:L\to \operatorname{Im} f\subset L'$ е сюрективно и според теоремата за хомоморфизмите пространствата $\operatorname{Im} f$ и $L/\operatorname{Ker} f$ са изоморфни. Следователно $\dim \operatorname{Im} f=\dim(L/\operatorname{Ker} f)=\dim L-\dim \operatorname{Ker} f$ и отново получихме друга позната теорема.

Да се занимаем и с обратната задача: как по произволен базис $\overline{e}_{k+1},\dots,\overline{e}_n$ на факторпространството L/M да построим поне един базис на L? За целта от всеки клас $\overline{e}_i=e_i+M$, $i=k+1,\dots,n$, по произволен начин избираме по един вектор; това може да бъде и векторът e_i. Да изберем и базис $e_1,\dots,e_k$ на подпространството M. Твърдим, че векторите $e_1,\dots,e_k,e_{k+1},\dots,e_n$ са базис на L. Тъй като броят им е равен на $\dim L$, достатъчно е да докажем, че те са линейно независими. Наистина, от

$$\lambda_1 e_1+\dots+\lambda_k e_k+\lambda_{k+1}e_{k+1}+\dots+\lambda_n e_n=0$$

следва $\lambda_{k+1}\overline{e}_{k+1}+\dots+\lambda_n\overline{e}_n=\overline{o}$. Тъй като $\overline{e}_{k+1},\dots,\overline{e}_n$ са базис на L/M, то $\lambda_{k+1}=\dots=\lambda_n=0$. Тогава от предходното равенство следва, че $\lambda_1=\dots=\lambda_k=0$, защото $e_1,\dots,e_k$ са базис на M.

7. Линейни оператори и факторпространства. Нека $f:L\to L$ е линеен оператор, дефиниран в линейното пространство L, а M е подпространство на L. Ще казваме, че подпространството M е *инвариантно относно оператора* f, ако за всеки вектор $m \in M$ векторът $f(m)$ също принадлежи на M.

Образът на подпространството M под действие на оператора f ще означаваме с $f(M)$, т.е.

$$f(M) = \{ f(m) \mid m \in M \}.$$

При това означение подпространството M е инвариантно относно оператора f точно когато $f(M) \subset M$. Очевидно е, че тривиалните подпространства $\{0\}$ и L са инвариантни относно всеки оператор.

Да предположим, че подпространството M е инвариантно относно оператора f. Дефинираме изображение $\bar{f}: L/M \to L/M$ с правилото: за всеки вектор $\bar{l} = M + l$ от факторпространството $\bar{f}(\bar{l}) = M + f(l)$, т.е. $\bar{f}(\bar{l}) = \overline{f(l)}$. Ще проверим, че определението е коректно, т.е. че $\bar{f}(\bar{l})$ зависи само от класа $\bar{l}$, а не от конкретния представител l. Действително, нека $\bar{l} = \bar{l}'$, което означава, че $l = l' + m$ за някакъв вектор $m \in M$. Тогава $f(l) = f(l') + f(m)$ и $f(m) \in M$, заради инвариантността на M. Следователно $\overline{f(l)} = \overline{f(l')}$ и $\bar{f}(\bar{l}) = \bar{f}(\bar{l}')$. Така дефинираното изображение $\bar{f}$ всъщност е линеен оператор - линейността му следва непосредствено от линейността на оператора f.

Сега да предположим, че пространството L е крайномерно, да изберем негов базис $e_1, \ldots, e_n$, но така, че $e_1, e_2, \ldots, e_k$ да образуват базис на M и нека

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

е матрицата на f спрямо този базис. По определение

$$f(e_i) = a_{1i}e_1 + \ldots + a_{ki}e_k + a_{k+1i}e_{k+1} + \ldots + a_{ni}e_n.$$

Като вземем предвид, че $\overline{e}_1 = \ldots = \overline{e}_k = \overline{0}$, последното равенство се трансформира в

(3) $$\overline{f(e_i)} = \overline{f}\left(\overline{e}_i\right) = a_{k+1i}\overline{e}_{k+1} + \ldots + a_{ni}\overline{e}_n\,, \quad i = k+1, \ldots, n.$$

В т. 6 видяхме, че $\overline{e}_{k+1}, \ldots, \overline{e}_n$ са базис на факторпространството L/M и равенствата (3) означават, че матрицата $\overline{A}$ на оператора $\overline{f}$ спрямо този базис е

$$\overline{A} = \begin{pmatrix} a_{k+1k+1} & \ldots & a_{k+1n} \\ \ldots. & .. & .. \\ a_{nk+1} & \ldots & a_{nn} \end{pmatrix},$$

т.е. тя се получава от A с премахване на първите k реда и k стълба. Подчертаваме, че базисът $e_1, \ldots, e_n$ е все пак специално подбран.

§ 6. Собствени вектори и собствени стойности на линеен оператор

1. Определение. Нека L е линейно пространство над поле $\mathcal{K}$, а $f: L \to L$ е линеен оператор, дефиниран в L. За ненулевия вектор $g \in L$ ($g \neq 0$) ще казваме, че е *собствен вектор* на оператора f, ако съществува скалар $\lambda \in \mathcal{K}$, за който

(1) $$f(g) = \lambda g.$$

Скаларът λ се нарича *собствена стойност* на оператора.

Може да се случи при фиксирано λ равенството (1) да е изпълнено за различни ненулеви вектори. За тях се казва, че *съответстват на собствената стойност* λ. Лесно е да се съобрази, че всички *те заедно с нулевия вектор образуват подпространство на пространството* L.

Интересът към собствените вектори се мотивира от много причини, които засега не можем да изложим. Става дума за задачи, които в приложенията на линейната алгебра възникват по естествен начин, например в квантовата механика. Ще отбележим един много частен пример.

В крайномерно пространство матрицата на линеен оператор спрямо някакъв базис е диагонална тогава и само тогава, когато базисът се състои от собствени вектори на оператора. По диагонала й стоят собствени стойности на оператора.

Твърдението е тривиално и следва направо от определенията на собствен вектор и на матрица на линеен оператор спрямо даден базис. Разбира се, не за всеки оператор може да се намери базис от собствени вектори. Ако за някой оператор такъв базис съществува, операторът се нарича *диагонализуем*.

2. Нека пространството L над полето $\mathcal{K}$ има размерност n, нека $e_1, e_2, \dots, e_n$ е негов произволен базис и $A = \left(a_{ij}\right)$ е матрица на оператора f спрямо този базис. Считаме, че търсеният собствен вектор g има координати $x_1, x_2, \dots, x_n$. Тогава равенството (1) се записва във вида $AX = \lambda X$ (вж. гл. 3, § 3, т. 8) или по-подробно като система уравнения с неизвестни $x_1, x_2, \dots, x_n$ и λ:

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \lambda x_1 \\ &a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = \lambda x_2 \\ &\dots\dots\dots\dots\dots\dots\dots \\ &a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = \lambda x_n. \end{aligned} \tag{2}$$

Тя очевидно е еквивалентна на системата

$$\begin{aligned} &\left(a_{11} - \lambda\right)x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ &a_{21}x_1 + \left(a_{22} - \lambda\right)x_2 + \dots + a_{2n}x_n = 0 \\ &\dots\dots\dots\dots\dots\dots \\ &a_{n1}x_1 + a_{n2}x_2 + \dots + \left(a_{nn} - \lambda\right)x_n = 0. \end{aligned} \tag{3}$$

Преди да анализираме системата (3), ще подчертаем следното. Всички коефициенти a_{ij} в системата (3) са от полето $\mathcal{K}$, защото са елементи на матрицата на линеен оператор, дефиниран в линейно пространство L над поле $\mathcal{K}$. По определение на собствена стойност скаларът λ трябва да принадлежи на полето $\mathcal{K}$. Тъй като $g \in L$, координатите му $x_1, x_2, \dots, x_n$ също трябва да принадлежат на $\mathcal{K}$.

По условие търсеният собствен вектор g трябва да бъде ненулев, следователно координатите му $x_1, x_2, \dots, x_n$ не са едновременно равни на нула и са ненулево решение на хомогенната система (3). Но както знаем, хомогенна система от n линейни уравнения с n неизвестни има ненулево решение точно тогава, когато детерминантата от коефициентите пред неизвестните е равна на нула. Следователно

(4) $$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ . \; . \; . & . \; . \; . & . \; . & . \; . \; . \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0.$$

В гл.3, § 4, т. 2 въведохме полинома $\varphi_A = \det(tE\text{-}A)$ и го нарекохме характеристичен полином на матрицата A; корените му нарекохме характеристични корени на матрицата. Тъй като $\det(tE\text{-}A) = (-1)^n \det(A - tE)$, равенството (4) показва, че *собствената стойност λ е характеристичен корен на матрицата A, принадлежащ на полето $\mathcal{K}$*. Вярно е и обратното: *всеки характеристичен корен λ на матрицата A, принадлежащ на полето $\mathcal{K}$, е собствена стойност на оператора f*. Наистина при това условие всички коефициенти ред неизвестните в системата (3) са в полето $\mathcal{K}$ и детерминантата от тези коефициенти е равна на нула. Следователно системата има ненулево решение. Както коментирахме в края на § 2, можем да изберем даже такова ненулево решение $x'_1,\dots,x'_n$, на което всичките компоненти x'_i да принадлежат на полето $\mathcal{K}$. Да означим с g (ненулевия) вектор с координати $x'_1,\dots,x'_n$. Сега от равенствата (2) получаваме, че за конкретния вектор $g \in L$ и за число $\lambda \in \mathsf{K}$ е изпълнено равенството (1), т.е. g е собствен вектор, съответстващ на собствената стойност λ.

От казаното дотук следва, че условието (1) е еквивалентно на следните две аналитични условия: а) λ е характеристичен корен на матрицата A, принадлежащ на полето $\mathcal{K}$; б) координатите $x_1, x_2, \dots, x_n$ на вектора g принадлежат на полето $\mathcal{K}$, и са ненулево решение на системата (3).

Как да се намерят всички решения на хомогенна система, видяхме в гл. 2. По-деликатно е условието а). Съгласно теоремата на Д'Аламбер (вж. гл. 8, § 3) всеки полином с числови коефициенти и степен поне единица има толкова корени (те са изобщо комплексни числа), колкото е степента му, стига всеки корен да се брои толкова пъти, колкото е кратността му. В качеството на собствени стойности в равенството (1) може да се използват обаче само онези корени на характеристичния полином $\varphi_A(t)$ (той е с коефициенти от $\mathcal{K}$), които принадлежат на полето $\mathcal{K}$. Възможно е такива корени изобщо да няма. Ако $\mathcal{K} = \mathrm{C}$, проблемът се решава от теоремата на Д'Аламбер: всички корени на характеристичния полином са в C. Но ако например $\mathcal{K} = \mathrm{R}$, възниква следната достатъчно трудна обща задача: при какви условия полином с реални коефициенти има реални корени и как да се определи броят на различните реални корени, ако има такива? Отговора отлагаме за курса по висша алгебра.

Във всеки случай от горните разсъждения получаваме следния алгоритъм за намиране на собствените стойности и собствените вектори (ако има такива) на линеен оператор f, дефиниран в крайномерно пространство над поле $\mathcal{K}$: а) спрямо някакъв базис изписваме матрицата A на оператора; б) намираме характеристичния полином $\varphi_A(t) = \det(tE\text{-}A)$; в) намираме всички корени $\lambda_1, \ldots, \lambda_k$ на $\varphi_A(t)$, които принадлежат на полето $\mathcal{K}$ - те именно са собствените стойности на оператора; г) за всяко $i = 1, 2, \ldots, k$ заместваме в системата (3) λ с λ_i и я решаваме - ненулевите й решения (и само те) са координатите на собствените вектори, съответстващи на собствената стойност λ_i.

Ще подчертаем, че *ако характеристичният полином има поне един корен в полето* $\mathcal{K}$, *операторът има поне един собствен вектор*.

3. Теорема. *Нека* $g_1, g_2, \ldots, g_k$ *са собствени вектори на линеен оператор* f, *нека* $f(g_i) = \lambda_i g_i$, $i = 1, 2, \ldots, k$, *и собствените стойности* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *са различни. Тогава векторите* $g_1, g_2, \ldots, g_k$ *са линейно независими.*

Доказателство. Ще разсъждаваме с индукция по k. При $k = 1$ твърдението е очевидно, защото по определение всеки собствен вектор е ненулев. Нека твърдението е вярно за $k-1$ вектора и да допуснем, че

$$\alpha_1 g_1 + \alpha_2 g_2 + \ldots + \alpha_k g_k = 0.$$

Към това равенство прилагаме оператора f и получаваме

$$\alpha_1 \lambda_1 g_1 + \alpha_2 \lambda_2 g_2 + \ldots + \alpha_k \lambda_k g_k = 0.$$

От двете равенства непосредствено следва, че

$$\alpha_1 (\lambda_1 - \lambda_k) g_1 + \alpha_2 (\lambda_2 - \lambda_k) g_2 + \ldots + \alpha_{k-1} (\lambda_{k-1} - \lambda_k) g_{k-1} = 0.$$

Според индуктивното допускане тук всички коефициенти са нули, а тъй като по условие $\lambda_i - \lambda_k \neq 0$, $i = 1,2,\ldots,k-1$, то $\alpha_1 = \alpha_2 = \ldots = \alpha_{k-1} = 0$, а следователно и $\alpha_k = 0$, което завършва доказателството.

4. Спектър. Множеството от собствените стойности на линейния оператор f се нарича негов *спектър*. Ако f е дефиниран в крайномерно пространство L над поле $\mathcal{K}$, спектърът се състои от онези корени на характеристичния полином на f, които се съдържат в $\mathcal{K}$. Ако въпросните корени имат кратност единица (прости корени), спектърът се нарича *прост*.

5. Теорема. *Ако линейният оператор f е дефиниран в пространство с размерност n, спектърът му е прост и се състои от n собствени стойности, то съществува базис на пространството, състоящ се от собствени вектори* (*операторът е диагонализуем*).

Твърдението е непосредствено следствие от теоремата в т. 3: на n-те различни собствени стойности $\lambda_1, \lambda_2, \ldots, \lambda_n$ съответстват собствени вектори $g_1, g_2, \ldots, g_n$, които са линейно независими. Тъй като пространството е n-мерно, те са негов базис.

За крайномерни пространства над полето C на комплексните числа твърдението може да се изкаже и така: *ако характеристичният полином на оператора няма кратни корени, операторът е диагонализуем*. Ще отбележим, че обратното не е вярно: достатъчно е да се разгледа единичният оператор $\mathcal{E}$, дефиниран с равенството $\mathcal{E}(l) = l$ за всеки вектор l от пространството L. Спектърът му се състои само от числото 1, при $n > 1$ не е прост, но операторът е диагонализуем - спрямо всеки базис на L матрицата му е единичната матрица $\mathcal{E}$ от ред n.

§ 7. Комплексификация на реално линейно пространство

1. Постановка на задачата. Нека L е реално линейно пространство ($\mathcal{K} = \mathrm{R}$). Искаме да построим линейно пространство L^{C} над полето C на комплексните числа и да са удовлетворени следните условия : а) векторите от L да са вектори и от L^{C}, т.е. $L \subset L^{\mathrm{C}}$; б) в L^{C} правилата за събиране на вектори и за умножаване на вектори с комплексни числа да са такива, че когато се прилагат в частност за вектори от L и за реални числа, да съвпадат със съответните правила, вече въведени в реалното пространство L.

Мотиви да се търси такова комплексно пространство, което да съдържа L като подмножество (не непременно като подпространство), т.е.

да се разшири полето на скаларите, може да се намерят в предишния параграф. Там видяхме, че ако реалното пространство L е крайномерно, собствените стойности на оператора $f: L \to L$ са реалните корени на характеристичния му полином. И точно тук възниква дисхармония: за някои оператори характеристичните им полиноми или изобщо нямат реални корени, или пък само част от корените им са реални. Ако полето на скаларите е C, тази дисхармония изчезва.

2. Ще дадем едно от възможните решения на поставената задача. Нека

$$L^{\mathrm{C}} = \left\{ \left(l_1, l_2 \right) \mid l_1, l_2 \in L \right\},$$

т.е. L^{C} е множеството на всички наредени двойки вектори от L. Както обикновено, $\left(l_1, l_2 \right) = \left(l_1', l_2' \right)$ точно тогава, когато $l_1 = l_1'$, $l_2 = l_2'$. В L^{C} въвеждаме операция събиране с правилото

$$\left(l_1, l_2 \right) + \left(l_1', l_2' \right) = \left(l_1 + l_1', l_2 + l_2' \right). \tag{1}$$

Нейната комутативност и асоциативност следват веднага от комутативността и асоциативността на събирането в L. Ролята на нулев елемент играе двойката $\left(0, 0 \right)$, където 0 е нулевият вектор на L. Очевидно $\left(l_1, l_2 \right) + \left(-l_1, -l_2 \right) = \left(0, 0 \right)$, следователно всеки елемент от L^{C} има противоположен елемент.

Да дефинираме умножение на елементите на L^{C} с комплексни числа. Ако a е реално число, нека

$$a\left(l_1, l_2 \right) = \left(al_1, al_2 \right), \tag{2}$$

а ако i е имагинерната единица ($i^2 = -1$), нека

$$i\left(l_1, l_2 \right) = \left(-l_2, l_1 \right). \tag{3}$$

Всяко комплексно число се записва еднозначно във вида $a + bi$, $a, b \in R$. Ако искаме да превърнем L^{C} в линейно пространство над C, трябва да е изпълнено

$$\left(a+bi\right)\left(l_1,l_2\right)=a\left(l_1,l_2\right)+b\left[i\left(l_1,l_2\right)\right]=\left(al_1,al_2\right)+b\left(-l_2,l_1\right)=$$
$$=\left(al_1,al_2\right)+\left(-bl_2,bl_1\right)=\left(al_1-bl_2,bl_1+al_2\right).$$

Това ни задължава да дефинираме

(4) $$\left(a+bi\right)\left(l_1,l_2\right)=\left(al_1-bl_2,bl_1+al_2\right).$$

Част от формалната проверка, че операциите, дефинирани с (1) и (4), превръщат L^{C} в линейно пространство над C, направихме по-горе. Останалата част оставяме за упражнение.

Всеки вектор $l\in L$ ще отъждествяваме с двойката $\left(l,0\right)$ и ще пишем $l=\left(l,0\right)$. В този смисъл ще считаме, че $L\subset L^{\mathrm{C}}$. Очевидно $l+l'=\left(l+l',0\right)=\left(l,0\right)+\left(l',0\right)$ и $al=\left(al,0\right)=a\left(l,0\right)$, $a\in\mathrm{R}$. Всичко това показва, че условията а) и б), поставени в т. 1, са удовлетворени.

3. Определение. Построеното в т. 2 комплексно линейно пространство L^{C} ще наричаме *комплекцификация* на реалното пространство L.

Ще преминем към по-удобни означения. Тъй като $i\left(l,0\right)=\left(0,l\right)$ и $l=\left(l,0\right)$, ще пишем

$$il=\left(0,l\right).$$

Следователно

(5) $$\left(l_1,l_2\right)=\left(l_1,0\right)+\left(0,l_2\right)=l_1+il_2.$$

Като се отказваме от наредените двойки, ще считаме, че всеки вектор от L^{C} се записва еднозначно във вида l_1+il_2; $l_1,l_2\in L$. Можем да не помним и равенството (4): ако в израза $\left(a+bi\right)\left(l_1+il_2\right)$ разкрием формално скобите и отчетем, че $i^2=-1$, ще получим дясната страна на (4).

Въпреки че L е подмножество на L^{C} и е затворено относно събирането, не бива да се мисли, че L е подпространство на L^{C}. Действително, ако $l\in L$ и $l\neq 0$, векторът il вече не е в L.

Ще дадем един абсолютно елементарен пример. Множеството $L=\mathrm{R}$ на реалните числа е линейно пространство над полето R. То е едномерно реално пространство. Според нашата конструкция елементите на комплексификацията L^{C} се записват еднозначно във вида $a+ib$; $a,b\in\mathrm{R}$. Очевидно е, че $L^{\mathrm{C}}=\mathrm{C}$ и пространството L^{C} също е едномерно линейно пространство, но над полето C.

Ще докажем следното

4. Твърдение. *Ако* L *е крайномерно реално линейно пространство, то всеки базис на* L *(над* R*) е базис (над* C*) и на комплексификацията му* L^{C}*. В частност* $\dim_{\mathrm{R}} L = \dim_{\mathrm{C}} L^{\mathrm{C}}$.

(Поставихме долните индекси R и C, за да подчертаем, че L е пространство над R, а L^{C} - над C. Този начин на означаване е общоприет и се използва винаги, когато има опасност от недоразумение.)

Доказателство. Нека $e_1,\dots,e_n$ е базис на реалното пространство L (R - базис). Всеки вектор от L^{C} се записва еднозначно като l_1+il_2; $l_1,l_2\in L$. Нека $l_1=a_1e_1+\dots+a_ne_n$, $l_2=b_1e_1+\dots+b_ne_n$; реалните числа a_j, b_j са еднозначно определени, защото векторите $e_1,\dots,e_n$ са R - базис на L. Следователно еднозначно е и записването

$$l_1+il_2=\left(a_1+b_1i\right)e_1+\dots+\left(a_n+b_ni\right)e_n.$$

Така $e_1,\dots,e_n$ е базис на L^{C} над C (вж. § 2, т. 1), което доказва твърдението.

5. Комплексификация на линеен оператор. Нека в реалното линейно пространство L е дефиниран линеен оператор $f:L\to L$. Дефинираме изображение $f^{\mathrm{C}}:L^{\mathrm{C}}\to L^{\mathrm{C}}$ с правилото

$$f^{\mathrm{C}}\left(l_1+il_2\right)=f\left(l_1\right)+if\left(l_2\right).$$

Тривиално е да се провери, че f^{C} е линеен оператор в комплексификацията L^{C}; той се нарича *комплексификация* на (реалния) оператор f.

Да разгледаме частния случай, когато реалното пространство L е крайномерно и $e_1,\dots,e_n$ е някакъв негов базис. Ако A е матрицата на f спрямо избрания базис, то както знаем, стълбовете ѝ са координатите на векторите $f\left(e_j\right)$, $j=1,\dots,n$. Както видяхме, $e_1,\dots,e_n$ е базис (над C) и на L^{C}.

Тъй като $f^{\mathrm{C}}\left(e_j\right)=f\left(e_j\right)$, ясно е, че матрицата на f^{C} спрямо разглеждания базис е същата матрица A.

Просто, но ефективно приложение на разглежданията от този параграф е следното

6. Твърдение. *Ако L е ненулево крайномерно реално линейно пространство и f е линеен оператор в L, то винаги съществува едномерно или двумерно подпространство M, което е инвариантно относно оператора, т.е. $f\left(M\right)\subset M$.*

Доказателство. Нека $\varphi\left(t\right)$ е характеристичният полином на оператора f. Както знаем, той е с реални коефициенти. Ако λ е реален корен на $\varphi\left(t\right)$, то според резултатите в § 6 съществува собствен вектор $g\in L$ ($g\neq 0$), за който $f\left(g\right)=\lambda g$. Едномерното подпространство M на L с базис вектора g очевидно е инвариантно относно оператора f.

Нека $\lambda=a+bi$ е комплексен корен на характеристичния полином $\varphi\left(t\right)$. Разглеждаме оператора $f^{\mathrm{C}}:L^{\mathrm{C}}\to L^{\mathrm{C}}$. Според резултатите в т. 5 неговият характеристичен полином също е $\varphi\left(t\right)$. Както по-горе, съществува ненулев собствен вектор $g=l_1+il_2$, за който $f^{\mathrm{C}}\left(g\right)=\lambda g$. Разбира се, $l_1,l_2\in L$. Последното равенство, преписано по-подробно, дава

$$f\left(l_1\right)+if\left(l_2\right)=f^{\mathrm{C}}\left(l_1+il_2\right)=\left(a+bi\right)\left(l_1+il_2\right)=\left(al_1-bl_2\right)+i\left(bl_1+al_2\right).$$

Следователно

$$f\left(l_1\right)=al_1-bl_2\,,\qquad f\left(l_2\right)=bl_1+al_2\,. \tag{6}$$

Да означим с M подпространството на L, породено от векторите l_1 и l_2. Поне един от тях е ненулев, защото l_1+il_2 е ненулев вектор от L^{C}. Следователно размерността на M е най-малко 1 и най-много 2. Включването $f\left(M\right)\subset M$ следва от (6).

§ 8. Триъгълен вид на линеен оператор. Теорема на Хамилтън-Кейли

Тук ще направим някои приложения на изложеното в предишните няколко параграфа. Необходимо е читателят предварително да се запознае с § 3 на гл. 8. Резултатите оттам ще използваме съвсем свободно, без дори да припомняме формулировките.

1. Нека L е крайномерно линейно пространство над поле $\mathcal{K}$. Ще казваме, че линейният оператор $f\colon L \to L$ е в *триъгълен вид* или че е *приведен в триъгълен вид*, ако е намерен базис на L, спрямо който матрицата A на оператора има триъгълен вид:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ .. & .. & .. & .. \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Удобно е това условие да се формулира в термините на подпространства на L. За целта нека $\dim L = n$ и $e_1, e_2, \dots, e_n$ е базис, спрямо който операторът е в триъгълен вид. Нека L_i е подпространството, породено от $e_1, e_2, \dots, e_i$, $i = 1,2,\dots,n$. Очевидно

(1) $$L_1 \subset L_2 \subset \dots \subset L_i \subset \dots \subset L_n \subset L \quad \text{и} \quad \dim L_i = i.$$

Веригата (1) от подпространства обикновено се нарича *максимално уплътнен флаг* или просто *флаг*. От матрицата A може “да се прочете”, че

(2) $$f(e_i) = a_{1i}e_1 + a_{2i}e_2 + \dots + a_{ii}e_i \in L_i, \quad i = 1,2,\dots,n.$$

От това лесно следва, че $f(L_i) \subset L_i$, $i = 1,2,\dots,n$. Ще казваме кратко, че *флагът* (1) *е инвариантен относно оператора* f. И така, ако спрямо някакъв базис операторът е в триъгълен вид, то от този базис може да се конструира флаг, който е инвариантен относно оператора.

Обратно, ако съществува флаг от вида (1), който е инвариантен относно оператора f, то f може да се приведе в триъгълен вид. За целта е достатъчно да изберем базис $e_1, e_2, \dots, e_n$ на L, съгласуван с флага, т.е.

такъв, че векторите $e_1, e_2, \ldots, e_i$ да са базис на подпространството L_i, $i = 1,2,\ldots,n$. Това винаги може да се направи: за едномерното пространство L_1 избираме базис e_1, допълваме го до базис e_1, e_2 на подпространството L_2, след това допълваме до базис e_1, e_2, e_3 на L_3 и т.н. Тъй като $f(L_i) \subset L_i$, то по-специално $f(e_i) \in L_i$, което означава, че векторът $f(e_i)$ трябва да има вида (2) при подходящи скалари $a_{ji} \in \mathcal{K}$. Записано подробно за всяко $i = 1,2,\ldots,n$ това ни дава, че матрицата A на оператора е триъгълна спрямо построения базис. Резюмираме:

2. Твърдение. *В крайномерно пространство L операторът $f: L \to L$ се привежда в триъгълен вид тогава и само тогава, когато съществува флаг от вида* (1), *който е инвариантен относно оператора.*

Ще забележим, че ако операторът f се привежда в триъгълен вид и A е съответната триъгълна матрица от т. 1, то $a_{11}, \ldots, a_{nn}$ са нейните характеристични корени. Тъй като всички елементи a_{ij} в матрицата A принадлежат на полето $\mathcal{K}$, излиза, че всички корени на характеристичния полином на оператора f (вж. гл. 3, § 4, т. 2) принадлежат на полето $\mathcal{K}$. Този факт е необходимо условие операторът да се привежда в триъгълен вид. По-късно ще докажем, че това условие е и достатъчно. Това донякъде обяснява защо в следващата теорема се иска полето да бъде $\mathbb{C}$.

3. Теорема. *Нека L е ненулево крайномерно линейно пространство над полето на комплексните числа. Тогава всеки линеен оператор, дефиниран в L, може да се приведе в триъгълен вид.*

Доказателство. Ще разсъждаваме с индукция по $n = \dim L$. Ако $n = 1$, твърдението е очевидно. Предполагаме, че то е вярно за комплексни пространства с размерност $n-1$ и нека L е пространство с размерност n. Разглеждаме оператор $f: L \to L$. По теоремата на Д'Аламбер (гл. 7, § 3) всички корени на характеристичния му полином са в полето $\mathbb{C}$. Нека λ е един от корените и $e_1 \neq 0$ е съответен собствен вектор: $f(e_1) = \lambda e_1$. Да означим с L' едномерното подпространство с базис вектора e_1, а с L'' - произволно директно допълнение на L' (вж. § 4, т. 4), т.е.

$$L = L' \oplus L''. \tag{3}$$

Разбира се, $\dim L'' = n - \dim L' = n-1$, но подпространството L'' изобщо не е длъжно да бъде инвариантно относно оператора f (включването

$f\left(L''\right) \subset L''$ може да не е в сила). По тази причина ограничението на f върху подпространството L'' може и да не е линеен оператор, дефиниран в L'', и не можем да приложим индуктивното допускане. Тази трудност ще заобиколим по следния начин. Изхождайки от разлагането (3), дефинираме проекторите $p_1 : L \to L'$ и $p_2 : L \to L''$ (вж. § 4, т. 5). По определение $p_1\left(l' + l''\right) = l'$ и $p_2\left(l' + l''\right) = l''$ за всеки два вектора $l' \in L'$, $l'' \in L''$. При това $p_1 + p_2 = \mathcal{E}$, $f = \left(p_1 + p_2\right) \circ f = p_1 \circ f + p_2 \circ f$. Сега подпространството L'' е инвариантно относно оператора $p_2 \circ f$. Наистина $p_2 \circ f\left(L''\right) \subset p_2\left(L\right) \subset L''$.

Разглеждаме действието на $p_2 \circ f$ в L''. Става дума за линеен оператор, дефиниран в комплексно пространство с размерност $n-1$. Според индуктивното допускане в L'' съществува флаг

(4) $$L_1'' \subset L_2'' \subset \ldots \subset L_{n-1}'' = L'', \quad \dim L_i'' = i,$$

който е инвариантен относно $p_2 \circ f$. Вече лесно можем да построим в L флаг, който е инвариантен относно f. За целта полагаме

$$L_1 = L', \quad L_i = L_{i-1}'' + L', \quad i = 2,3,\ldots,n$$

Не е трудно да се убедим, че

(5) $$L' = L_1 \subset L_2 \subset \ldots \subset L_n = L$$

наистина е флаг, т.е. че dim. Деиствително, нека $e_2, e_3, \ldots, e_n$ е базис, съгласуван с флага (4), т.е. $e_2, \ldots, e_i$ е базис на L_{i-1}'', $i = 2,3,\ldots,n$. Тъй като e_1 е базис на L' и $L_i = L_{i-1}' \oplus L'$, то $e_1, e_2, \ldots, e_i$ е базис на L_i.

Ще проверим, че флагът (5) е инвариантен относно оператора f, т.е. че $f\left(L_i\right) \subset L_i$, $i = 1,2,\ldots,n$. При $i = 1$ подпространството L_1 се състои от векторите αe_1, $\alpha \in \mathrm{C}$, и $f\left(\alpha e_1\right) = \alpha \lambda e_1 \in L_1$. Нека $i \geq 2$. Сега

(6) $$f\left(L_i\right) = \left(p_1 + p_2\right) \circ f\left(L_i\right) = p_1 \circ f\left(L_i\right) + p_2 \circ f\left(L_i\right).$$

Включването $p_1 \circ f\left(L_i\right) \subset L'$ е очевидно, защото проекторът p_1 изобразява L в L'. За да се убедим, че $p_2 \circ f\left(L_i\right) \subset L_{i-1}''$, ще забележим, че $p_2 \circ f\left(L'\right) \subset p_2\left(L'\right) = \left\{0\right\}$ (проекторът p_2 действа върху L' като нулевия оператор), а включването $p_2 \circ f\left(L''_{i-1}\right) \subset L''_{i-1}$ е изпълнено по условие за флага (4). И така, $f\left(L_i\right) \subset L_i$ за $i = 1,2,\dots,n$. Окончателно доказахме, че (5) е флаг, инвариантен относно оператора f. Базисът $e_1, e_2, \dots, e_n$ е съгласуван с флага и спрямо него матрицата на оператора е триъгълна.

Второ доказателство. Ще използваме техниката на факторпространствата. Запазваме означенията от предишното доказателство. Както по-горе, строим едномерното подпространство L', породено от собствения вектор e_1, $f\left(e_1\right) = \lambda e_1$. То е инвариантно относно f. Във факторпространството L / L' разглеждаме линейния оператор $\bar{f}: L / L' \to L / L'$, индуциран от $f: L \to L$. Тъй като $\dim\left(L / L'\right) = n - 1$, за $\bar{f}$ е приложимо индуктивното допускане. Следователно $\bar{f}$ се привежда в триъгълен вид и нека $\bar{e}_2, \dots, \bar{e}_n$ е съответен базис на L / L', спрямо който матрицата $\bar{A}$ на оператора $\bar{f}$ е триъгълна. Както видяхме в § 5, т. 6, векторите $e_1, e_2, \dots, e_n$ образуват базис на L. Ще покажем, че матрицата A на оператора f спрямо този базис е триъгълна. Тъй като $f\left(e_1\right) = \lambda e_1$, първият стълб на A е n-торката $\left(\lambda, 0, \dots, 0\right)$. Както видяхме в § 5, т. 7, ако в A премахнем първия ред и първия стълб, ще получим матрицата $\bar{A}$ на оператора $\bar{f}$. Следователно A има вида

$$A = \left(\begin{array}{c|ccc} \lambda & a_{12} & \dots & a_{1n} \\ \hline 0 & & & \\ \vdots & & \bar{A} & \\ 0 & & & \end{array}\right)$$

и е горна триъгълна матрица, защото $\bar{A}$ е горна триъгълна. Това доказва теоремата.

От това доказателство може да се извлече и нещо повече. Нека $\varphi(t)$ е характеристичният полином на оператора f, а $\bar{\varphi}(t)$ - на оператора $\bar{f}$. Очевидно

$$\varphi(t) = \det\left(tE_n - A\right) = \left(t - \lambda\right)\det\left(tE_{n-1} - \bar{A}\right) = \left(t - \lambda\right)\bar{\varphi}(t).$$

Тази проста връзка между двата характеристични полинома показва например, че ако всички корени на полинома $\varphi(t)$ принадлежат на някакво числово поле $\mathcal{K}$, то и всички корени на $\bar{\varphi}(t)$ са в това поле. Буквално повтаряне на доказателството, заедно с тази забележка за корените, ще ни даде следната малко по-прецизна

4. Теорема. *Нека L е крайномерно линейно пространство над поле $\mathcal{K}$. Ако всички характеристични корени на оператора $f : L \to L$ са в полето $\mathcal{K}$, то f може да се приведе в триъгълен вид.*

5. Следствие. *Нека A е квадратна матрица от ред n с коефициенти от поле $\mathcal{K}$; накратко: $A \in M_n(\mathcal{K})$. Ако всички характеристични корени на матрицата A са в полето $\mathcal{K}$, то съществува такава неособена матрица $C \in M_n(\mathcal{K})$, че матрицата $C^{-1}AC$ да е горна триъгълна.*

Следствието всъщност е преформулировка на теоремата на матричен език. За да разберем това, достатъчно е да разгледаме координатното пространство $L = \mathcal{K}^{\mathrm{n}}$, да изберем базис в него (например стандартния) и да дефинираме оператор $f : L \to L$ така, че матрицата му спрямо избрания базис да е A. Според теоремата в т. 4 съществува базис, спрямо който матрицата B на оператора е горна триъгълна. Връзката между двете матрици е $B = C^{-1}AC$ при подходяща неособена матрица $C \in M_n(\mathcal{K})$ (вж. гл. 3, § 4, т. 2).

Ако $\mathcal{K} = \mathrm{C}$, теоремата на Д'Аламбер гарантира, че всички характеристични корени на матрицата A са в C и получаваме

6. Следствие. *Ако $\mathcal{K} = \mathrm{C}$, за всяка квадратна матрица $A \in M_n(\mathrm{C})$ съществува неособена матрица $C \in M_n(\mathrm{C})$, за която $C^{-1}AC$ е горна триъгълна.*

7. Някои приложения. Ако A е квадратна матрица от ред n и k е естествено число, полагаме $A^k = A.A \dots A$ (k множителя). Приемаме, че $A^0 = E$. Ако A е обратима, означаваме $A^{-k} = (A^{-1})^k$. Ще покажем, че *ако $\lambda_1, \lambda_2, \dots, \lambda_n$ са всичките характеристични корени на матрицата A, то $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ са всичките характеристични корени на матрицата A^k, k - естествено число. Ако матрицата A е обратима, същото е вярно и за отрицателни цели числа k.*

За целта прилагаме следствието на предишната точка и нека $B = C^{-1}AC$ е горна триъгълна матрица. Тъй като подобните матрици имат едни и същи характеристични корени, а в триъгълна матрица те са просто елементите по главния диагонал, B има вида

$$B = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & * & \\ 0 & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

Със * са означени елементите над диагонала. От правилото за умножаване на матрици непосредствено следва, че произведението на горни триъгълни матрици е горна триъгълна, а по диагонала стоят произведенията на съответните диагонални елементи. Следователно

$$B^k = \begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & * & \\ 0 & & \ddots & \\ & & & \lambda_n^k \end{pmatrix}.$$

Очевидно е, че характеристичните корени на B^k са $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$. От друга страна, $B^k = C^{-1}AC \ldots C^{-1}AC = C^{-1}A^kC$, което показва, че B^k и A^k са подобни, следователно имат едни и същи характеристични корени.

Ще забележим, че квадратната матрица A е обратима (неособена) точно тогава, когато всичките ѝ характеристични корени са различни от нула. Това е така, защото подобните матрици имат равни детерминанти и при горните означения $\det A = \det B = \lambda_1.\lambda_2 \ldots \lambda_n$. Не е трудно да се съобрази, че ако A е обратима, то матрицата B^{-1} има вида

$$B^{-1} = \begin{pmatrix} \lambda_1^{-1} & & & \\ & \lambda_2^{-1} & * & \\ & & \ddots & \\ & & & \lambda_n^{-1} \end{pmatrix}.$$

Сега от $C^{-1}AC = B$ следва $C^{-1}A^{-1}C = B^{-1}$, което показва, че характеристичните корени на матрицата A^{-1} са $\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}$. Нека k е отрицателно цяло число. Тъй като $A^k = \left(A^{-1}\right)^k$ и $-k > 0$, характеристичните корени на матрицата A^k са $\left(\lambda_i^{-1}\right)^{-k} = \lambda_i^k$, $k = 1,2,\dots,n$.

б) Квадратната матрица A от ред n се нарича *нилпотентна,* ако съществува естествено число k, за което $A^k = 0$. Ще покажем, че A е *нилпотентна тогава и само тогава, когато всичките й характеристични корени са равни на нула, т.е. характеристичният й полином е* t^n.

Действително, нека A е нилпотентна и $\lambda_1, \lambda_2, \dots, \lambda_n$ са характеристичните й корени. Нека $A^k = 0$. Характеристичните корени на нулевата матрица са равни на нула и според а) имаме $\lambda_1^k = \dots = \lambda_n^k = 0$, т.е. $\lambda_1 = \dots = \lambda_n = 0$.

Обратно, нека $\lambda_1 = \dots = \lambda_n = 0$. Привеждаме A в триъгълен вид:

$$C^{-1}AC = B = \begin{pmatrix} 0 & & & \\ & 0 & * & \\ 0 & & \ddots & \\ & & & 0 \end{pmatrix}.$$

Не е трудно да се провери директно, че $B^n = 0$. Тъй като $B^n = \left(C^{-1}AC\right)^n = C^{-1}A^nC$, то $C^{-1}A^nC = 0$ и като умножим отляво с C, а отдясно с C^{-1}, ще получим $A^n = 0$. Пътем видяхме и следното: ако квадратната матрица A от ред n е нилпотентна, то $A^n = 0$.

7. Нека $\varphi(t) = a_m t^m + a_{m-1}t^{m-1} + \dots + a_0$ е произволен полином на променливата t с коефициенти от поле $\mathcal{K}$, а $f: L \to L$ е линеен оператор, дефиниран в линейно пространство L над същото поле. В гл. 3 дефинирахме сума на линеен оператор с число от $\mathcal{K}$, както и композиция $f \circ g$ на операторите f и g. Във всички случаи резултатът от операциите бе линеен оператор, дефиниран в L. Ако k е естествено число, полагаме $f^k = f \circ f \circ \dots \circ f$ (в дясната страна стоят k “множителя”) и $f^0 = \mathcal{E}$, където $\mathcal{E}$ е единичният оператор. Тогава изразът

$$\varphi(f) = a_m f^m + \ldots + a_1 f + a_0 \varepsilon$$

е коректно дефиниран и $\varphi(f)$ е линеен оператор, дефиниран в L. (Заместихме променливата t с оператора f.) Изискването коефициентите на полинома $\varphi(t)$ да бъдат от полето $\mathcal{K}$ е съществено: можем да умножаваме векторите от L само с числа от $\mathcal{K}$ и нищо повече.

Аналогично, ако A е квадратна матрица от ред n с елементи от поле $\mathcal{K}$, а $\mathcal{E}$ е единичната матрица от ред n, коректно дефиниран е и изразът

$$\varphi(A) = a_m A^m + \ldots + a_1 A + a_0 \mathcal{E}$$

и $\varphi(A)$ е квадратна матрица от ред n с елементи от полето $\mathcal{K}$.

Ако $\dim L = n$, линейното пространство $\mathcal{L}(L, L)$ (вж. § 3, т. 4) е крайномерно (то е изоморфно с $M_n(\mathcal{K})$), следователно съществува естествено число m, за което операторите $\mathcal{E}, f, f^2, \ldots, f^m$ са линейно зависими, т.е. съществуват числа $a_i \in \mathcal{K}$, които не са едновременно равни на нула и за които

$$. \ a_0 \mathcal{E} + a_1 f + \ldots + a_m f^m$$

С други думи съществува ненулев полином $\varphi(t)$ с коефициенти от полето $\mathcal{K}$, за който $\varphi(f) = 0$ (нулевият оператор). Обикновено се казва, че *операторът* f *анулира полинома* $\varphi(t)$. Току-що показахме, че за всеки оператор в крайномерно линейно пространство такъв полином съществува. Между ненулевите полиноми с коефициенти от полето $\mathcal{K}$, които се анулират от конкретен оператор, избираме полином с възможно най-ниска степен. Ако го умножим с подходяща ненулева константа, можем да считаме, че старшият му коефициент е единица. Такъв полином ще наричаме *минимален полином на оператора* f.

Ще забележим, че ако $\mu(t)$ е минимален полином на оператора f, а $\varphi(t)$ е произволен полином с коефициенти от $\mathcal{K}$, който се анулира от f, то $\mu(t)$ дели $\varphi(t)$. Наистина, да разделим с остатък:

$$\varphi(t)=\mu(t)q(t)+r(t),$$

където $q(t)$, $r(t)$ са полиноми с коефициенти от $\mathcal{K}$ и $r(t)=0$ или степента на $r(t)$ е по-ниска от степента на $\mu(t)$. Ако $r(t)$ е нулевият полином, няма какво да се доказва. В противен случай, тъй като $\varphi(f)=\mu(f)=0$, получаваме $r(f)=0$, което противоречи на минималния избор на степента на $\mu(t)$.

Оттук следва по-специално, че минималният полином на оператора f е еднозначно определен: ако $\mu_1(t)$ е още един минимален полином, то $\mu(t)$ дели $\mu_1(t)$, а $\mu_1(t)$ дели $\mu(t)$, т.е.

$$\mu_1(t)=\mu(t)q(t), \qquad \mu(t)=\mu_1(t)q_1(t).$$

От двете равенства получаваме, че $q(t)$ и $q_1(t)$ са константи. Тъй като старшите коефициенти на $\mu(t)$ и $\mu_1(t)$ са единици, то $q(t)=q_1(t)=1$ и $\mu_1(t)=\mu(t)$.

8. Теорема на Хамилтън - Кейли. *Нека L е крайномерно линейно пространство над поле $\mathcal{K}$, а $f:L\to L$ е линеен оператор, дефиниран в L. Ако всички корени на характеристичния полином $\varphi(t)$* на *f принадлежат на полето $\mathcal{K}$, то*

$$\varphi(f)=0$$

При $\mathcal{K}=\mathrm{C}$ *заключението е вярно за всеки оператор f.*

Доказателство. Ще докажем теоремата с индукция по $n=\dim L$. При $n=1$, ако e_1 е базис на L, то $f(e_1)=\lambda e_1$ за подходящо $\lambda\in\mathcal{K}$ и характеристичният полином на оператора f е $t-\lambda$. Тъй като $(f-\lambda\mathcal{E})(e_1)=0$ (нулевия вектор), то $f-\lambda\mathcal{E}$ е нулевият оператор и твърдението е вярно.

По-нататък нека теоремата е вярна за всички линейни оператори, удовлетворяващи условията й и дефинирани в пространства с размерност $n-1$. Нека $\dim L=n$. Съгласно теоремата в т. 4 операторът f може да се приведе в триъгълен вид. Нека

$$L_1\subset L_2\subset\ldots\subset L_n=L, \quad \dim L_i=i,$$

е съответният флаг от т. 1. По условие $f\left(L_i\right) \subset L_i$, $i = 1,\ldots,n$. Ако $e_1,\ldots,e_n$ е базис на L, съгласуван с флага, матрицата A на f спрямо този базис има вида

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ .. & .. & .. & .. \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}.$$

Следователно за характеристичния полином $\varphi\left(t\right)$ имаме

$$\varphi\left(t\right) = \left(t - a_{11}\right)\left(t - a_{22}\right)\ldots\left(t - a_{nn}\right).$$

Разглеждаме ограничението $\bar{f}$ на оператора f върху подпространството L_{n-1}, т.е. $\bar{f}\left(l'\right) = f\left(l'\right)$ за всеки вектор $l' \in L_{n-1}$. Тъй като $f\left(L_{n-1}\right) \subset L_{n-1}$, ясно е, че $\bar{f}$ е линеен оператор в пространството L_{n-1}, на което векторите $e_1,\ldots,e_{n-1}$ са базис. Тъй като

(7) $$f\left(e_i\right) = a_{1i}e_1 + \ldots + a_{ii}e_i, \qquad i = 1,\ldots,n,$$

матрицата $\overline{A}$ на оператора $\bar{f}$ спрямо този базис се получава от A чрез премахване на последния ред и на последния стълб. Следователно характеристичният полином $\bar{\varphi}\left(t\right)$ на оператора $\bar{f}$ (т. е. характеристичният полином на $\overline{A}$) е

$$\bar{\varphi}\left(t\right) = \left(t - a_{11}\right)\left(t - a_{22}\right)\ldots\left(t - a_{n-1n-1}\right).$$

Ще се отклоним за малко със следната забележка: ако $h_i\left(t\right)$, $i = 1, 2, 3, 4$ са полиноми с коефициенти от полето $\mathcal{K}$ и

$$h_3\left(t\right) = h_1\left(t\right) + h_2\left(t\right), \quad h_4\left(t\right) = h_1\left(t\right)h_2\left(t\right),$$

то от тези равенства следват равенствата

$$h_3(f) = h_1(f) + h_2(f), \qquad h_4(f) = h_1(f)h_2(f)$$

за всеки линеен оператор f. Ако A е квадратна матрица от ред n, имаме аналогично

$$h_3(A) = h_1(A) + h_2(A), \qquad h_4(A) = h_1(A)h_2(A).$$

Това се получава лесно от равенствата

$$af^i + bf^i = (a+b)f^i,\ af^i \circ bf^j = abf^{i+j}$$

за всеки две цели неотрицателни числа i, j и за всички a,b от полето $\mathcal{K}$ и аналогично от равенствата

$$aA^i + bA^i = (a+b)A^i,\ aA^i \circ bA^j = abA^{i+j}$$

Детайлите оставяме за упражнение.

И така, за оператора $\bar{f}$ в $(n-1)$-мерно пространство L_{n-1} е приложимо индуктивното допускане, следователно

$$\bar{\varphi}(\bar{f}) = (\bar{f} - a_{11}\mathcal{E})\dots(\bar{f} - a_{n-1n-1}\mathcal{E}) = 0$$

(нулевият оператор в L_{n-1}). Всъщност получихме, че

$$(f - a_{11}\mathcal{E})\dots(f - a_{n-1n-1}\mathcal{E})(l') = 0$$

за всеки вектор $l' \in L_{n-1}$. От това равенство веднага ще следва, че

$$\varphi(f) = (f - a_{11}\mathcal{E})\dots(f - a_{n-1n-1}\mathcal{E})(f - a_{nn}\mathcal{E}) = 0,$$

стига да покажем, че за всеки вектор $l \in L$ е в сила $(f - a_{nn}\mathcal{E})(l) \in L_{n-1}$.

Действително, всеки вектор $l \in L$ може да се запише във вида $l = ae_n + l'$, където $l' \in L_{n-1}$. От (7) следва по-специално, че

$$(f - a_{nn}\mathcal{E})(e_n) = a_{1n}e_1 + \dots + a_{n-1n}e_{n-1} \in L_{n-1}.$$

По условие $f\left(l'\right)\in L_{n-1}$, следователно

$$\left(f-a_{nn}\mathcal{E}\right)\left(l\right)=a\left(f-a_{nn}\mathcal{E}\right)\left(e_n\right)+f\left(l'\right)-a_{nn}l'\in L_{n-1},$$

което завършва доказателството.

9. Следствие. *Нека* A *е квадратна матрица от ред* n *с числови елементи, а* $\varphi(t)$ *е нейният характеристичен полином. Тогава* $\varphi\left(A\right)=0$.

Доказателство. В координатното пространство $L=\mathrm{C}^{\mathrm{n}}$ избираме базис и дефинираме линеен оператор f така, че матрицата му спрямо избрания базис да бъде A. Характеристичният полином на f също е $\varphi(t)$. По теоремата на Хамилтън-Кейли имаме $\varphi\left(f\right)=0$. Тъй като матрицата на сума от линейни оператори е сумата от матриците им, а на произведение - произведението от матриците им, ясно е, че матрицата на оператора $\varphi\left(f\right)$ е $\varphi\left(A\right)$. Тъй като $\varphi\left(f\right)$ е нулевият оператор, то $\varphi\left(A\right)=0$.

УПРАЖНЕНИЯ

1. Докажете, че в определението на линейно пространство аксиомата за комутативност на събирането е следствие от останалите аксиоми. (Упътване. Използвайте, че $\left(1+1\right)\left(l_1+l_2\right)=1.\left(l_1+l_2\right)+1.\left(l_1+l_2\right)$ и $\left(1+1\right)\left(l_1+l_2\right)=\left(1+1\right)l_1+\left(1+1\right)l_2$.)

2. Ако $\mathcal{K}$ е числово поле, покажете, че $\mathcal{K}$ може да се разглежда като едномерно линейно пространство над $\mathcal{K}$. Опишете всички базиси на това поле.

3. Ако L и M са крайномерни линейни пространства над поле $\mathcal{K}$, покажете, че $\dim\mathcal{L}\left(L,M\right)=\dim L.\dim M$.

4. В следващите примери операциите са обичайните сума на матрици и умножаване на матрица с число; $\mathcal{K}$ е фиксирано числово поле. Покажете, че посочените множества са линейни пространства над $\mathcal{K}$ и определете размерностите им:

а) $M_n(\mathcal{K})$ - множеството на квадратните матрици от ред n с елементи от $\mathcal{K}$;

б) $M_1 = \left\{ A \in M_n(\mathcal{K}) \mid A = A^t \right\}$ - множеството на *симетричните* матрици от ред n с елементи от $\mathcal{K}$ (Матрицата A се нарича *симетрична*, ако $A = A^t$; A^t е транспонираната на A.);

в) $M_2 = \left\{ A \in M_n(\mathcal{K}) \mid A = -A^t \right\}$ - множеството на *антисиметричните* матрици от ред n;

г) множеството на горните триъгълни матрици от ред n с елементи от $\mathcal{K}$;

д) множеството на онези матрици от $M_n(\mathcal{K})$, чиято следа е равна на нула.

(Отговор: а) n^2; б) $\dfrac{n(n+1)}{2}$; в) $\dfrac{n(n-1)}{2}$; г) $\dfrac{n(n+1)}{2}$; д) n^2-1.)

5. Докажете, че за всяко естествено число n функциите $\sin x, \sin 2x, \dots, \sin nx$ са линейно независими над R в интервала $[0,2\pi]$. Колко е размерността на линейното пространство над R, състоящо се от реалните и непрекъснатите функции, дефинирани в интервала $[0,2\pi]$?

6. Нека L е линейно пространство над поле $\mathcal{K}$ и $L^* = \mathcal{L}(L,\mathcal{K})$ - множеството на линейните функции, дефинирани в L и със стойности в $\mathcal{K}$. Относно операциите сума на линейни изображения и умножаване на линейно изображение с число L^* е линейно пространство, което се нарича *дуално* на пространството L. Да предположим, че L е крайномерно.

а) Ако $e_1,\dots,e_n$ е базис на L, дефинираме функциите e^i с правилото: ако $l = x_1e_1 + \dots + x_ie_i + \dots + x_ne_n$, то $e^i(l) = x_i$, $i = 1,\dots,n$. Докажете, че функциите $e^1,\dots,e^n$ са линейни и образуват базис на L^*. (Той се нарича *дуален на базиса* $e_1,\dots,e_n$. По-специално имаме $e^i(e_i) = 1$, $e^i(e_j) = 0$ при $i \neq j$.)

б) Нека $f \in L^*$, $f \neq 0$. Полагаме

$$M_f = \left\{ l \in L \mid f(l) = 0 \right\}.$$

Докажете, че M_f е подпространство на L и $\dim M_f = \dim L - 1$.

в) Докажете, че всяко подпространство M на L с размерност $\dim L - 1$ може да се зададе както в б) при подходяща функция $f \in L^*$, т.е. $M = M_f$.

г) Ако $M_f = M_g$, $f \neq 0$, $g \neq 0$, докажете, че $f = \lambda g$ за някое $\lambda \in \mathcal{K}$, $\lambda \neq 0$.

д) Ако $f \in L^*$, $f \neq 0$ и λ е произволно, но фиксирано число от $\mathcal{K}$, докажете, че решенията на уравнението $f(l) = \lambda$ са точно векторите от линейното многообразие $l_1 + M_f$, където векторът l_1 е едно от решенията на уравнението.

7. Докажете, че $M_n(\mathcal{K}) = M_1 \oplus M_2$, където M_1 и M_2 са подпространствата, състоящи се съответно от симетричните и антисиметричните матрици.

8. Ако M_i, $i = 1, 2, 3$, са подпространства на линейното пространство L, докажете, че $(M_1 + M_2) + M_3 = M_1 + (M_2 + M_3)$.

9. Нека L е линейно пространство, а $p_i : L \to L$, $i = 1, 2, \ldots, k$, са проектори, т.е. $p_i^2 = p_i$. Нека освен това са изпълнени следните условия: а) $p_1 + p_2 + \ldots + p_k = \mathcal{E}$; б) $p_i p_j = 0$ при $i \neq j$. Ако $L_i = \mathrm{Im} p_i$, докажете, че $L = L_1 \oplus \ldots \oplus L_k$.

10. Нека линейното пространство L е директна сума на свои подпространства, т.е. $L = L_1 \oplus \ldots \oplus L_k$. Докажете, че обединението от базис на L_1, базис на L_2 и т.н. базис на L_k е базис на L. Ако L е крайномерно, то $\dim L = \sum_{i=1}^{k} \dim L_i$.

11. Нека в линейното пространство L е въведена релация на еквивалентност ~, съгласувана с операциите, т.е. от $l_1 \sim l_2$ и $l_1' \sim l_2'$ винаги следва, че $l_1 + l_1' \sim l_2 + l_2'$ и $\lambda l_1 \sim \lambda l_2$ за всяко λ от полето на скаларите. Да означим с M класа на еквивалентност, съдържащ нулевия вектор. Докажете, че M е подпространство на L и $l_1 \sim l_2$ тогава и само тогава, когато $l_1 - l_2 \in M$.

12. Нека M и N са подпространства на линейното пространство L. Докажете, че факторпространствата $(M + N)/N$ и $M/(M \cap N)$ са изоморфни. (У п ъ т в а н е . Проверете, че изображението, при което класът $(m + n) + N$ се изобразява в $m + (M \cap N)$, е изоморфизъм, $m \in M$,

$n \in N$.) В частност, ако сумата $M+N$ е директна, то факторпространството $\left(M+N\right)/N$ е изоморфно на M.

13. Нека L е крайномерно линейно пространство над C, а f и g са линейни оператори в L, които комутират, т.е. $fg = gf$. Докажете, че:

а) съществува поне един ненулев вектор, който е собствен и за двата оператора;

б) операторите може едновременно да се приведат в триъгълен вид, т.е. съществува базис на L, спрямо който и двата оператора са в триъгълен вид.

Упътване. Ако λ е собствена стойност на оператора f, нека $L_\lambda = \left\{ l \in L \mid f\left(l\right) = \lambda l \right\}$. Докажете, че L_λ е ненулево подпространство на L, което е инвариантно относно оператора g. Всеки собствен вектор на оператора g, принадлежащ на подпространството L_λ, е собствен вектор и на оператора f. По-нататък използвайте индукция по $\dim L$. Ако l_0 е собствен вектор на операторите f и g, а M е подпространството на L, породено от вектора l_0, разгледайте факторпространството L/M и използвайте резултатите от § 5, т. 6 и т. 7.

14. Нека A е квадратна матрица от ред n, на която всички характеристични корени са равни на единица. Нека $N = A - E_n$. Докажете, че матрицата N е нилпотентна и $N^n = 0$. Докажете, че A е обратима и

$$A^{-1} = \left(E_n + N\right)^{-1} = E_n - N + N^2 - \ldots + \left(-1\right)^{n-1} N^{n-1}.$$

(Упътване. Приведете матрицата A в триъгълен вид.)

Глава 5

ПРОСТРАНСТВОТО НА СВОБОДНИТЕ ВЕКТОРИ

Въпросите, разглеждани в тази глава, традиционно се считат за чисто геометрични. Тук обаче няма да изграждаме строго геометрията. Напротив, ще предполагаме, че читателят е запознат от училищния курс с основните обекти и понятия на евклидовата геометрия. Нашата цел е да я "алгебризираме", т.е. на геометричните обекти и понятия да съпоставим обекти и понятия от линейната алгебра, и то така, че съпоставянето да бъде по възможност взаимно еднозначно. По този път геометричните проблеми ще се превърнат в аналитични, т.е. в проблеми на линейната алгебра. Аксиоматичното изграждане на геометрията ще обсъдим бегло във втората част на курса.

§ 1. Свободен вектор. Афинни операции със свободни вектори

1. Отсечка AB, на която точката A е обявена за начало, а точката B за край, ще наричаме *насочена отсечка* и ще я означаваме с $\overrightarrow{AB}$. Ако B съвпада с A, ще казваме, че насочената отсечка е *нулева*.

За две ненулеви насочени отсечки $\overrightarrow{AB}$ и $\overrightarrow{CD}$ ще казваме, че са *колинеарни*, ако правите AB и CD са успоредни или съвпадат и ще пишем $\overrightarrow{AB} \| \overrightarrow{CD}$. Ако ненулевата отсечка $\overrightarrow{AB}$ е успоредна на някаква права или лежи върху нея, ще казваме, че правата и насочената отсечка са *колинеарни*. Ще считаме, че всяка нулева отсечка е колинеарна на всяка отсечка и на всяка права.

Напомняме, че права с избрана върху нея посока се нарича *ос*. Ненулевата насочена отсечка $\overrightarrow{AB}$ по естествен начин превръща правата AB в ос: за посока върху правата се избира посоката на насочената отсечка; тази ос ще бележим с $AB^{\rightarrow}$. Ако ненулевите отсечки $\overrightarrow{AB}$ и $\overrightarrow{CD}$ са колинеарни, ще казваме, че те имат *еднакви* или *противоположни посоки* в зависимост от това, дали осите $AB^{\rightarrow}$ и $CD^{\rightarrow}$ имат еднакви или противоположни посоки. Подчертаваме, че посоките на две отсечки се сравняват само ако те са ненулеви и колинеарни.

2. Определение. Две ненулеви насочени отсечки $\overrightarrow{AB}$ и $\overrightarrow{CD}$ наричаме *равни* и пишем $\overrightarrow{AB} = \overrightarrow{CD}$, ако: а) дължините им са равни; б) колинеарни са; в) посоките им са еднакви. Считаме, че всеки две нулеви насочени отсечки са равни.

От определението непосредствено следват свойствата:

а) $\overrightarrow{AB} = \overrightarrow{AB}$ (рефлексивност);

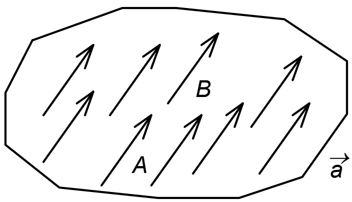
б) ако $\overrightarrow{AB} = \overrightarrow{CD}$, то $\overrightarrow{CD} = \overrightarrow{AB}$ (симетричност);

в) ако $\overrightarrow{AB} = \overrightarrow{CD}$ и $\overrightarrow{CD} = \overrightarrow{EF}$, то $\overrightarrow{AB} = \overrightarrow{EF}$ (транзитивност).

Следователно релацията "равенство на насочени отсечки" е релация на еквивалентност в множеството на всички насочени отсечки; тя го разбива на непресичащи се класове на еквивалентност.

3. Определение. Всеки клас на еквивалентност от равни насочени отсечки ще наричаме *свободен вектор*.

Ако е дадена насочената отсечка $\overrightarrow{AB}$, множеството на всички насочени отсечки, равни на $\overrightarrow{AB}$, е свободен вектор $\vec{a}$, определен от $\overrightarrow{AB}$. Тези насочени отсечки наричаме *представители на свободния вектор* и ще пишем $\overrightarrow{AB} \in \vec{a}$. Образно свободния вектор $\vec{a}$ можем да си представяме по следния начин (черт.1):



черт.1

Дължина на свободния вектор $\vec{a}$ (означение: $|\vec{a}|$) ще наричаме дължината на който и да е негов представител, *посока* на свободния вектор $\vec{a}$ ще наричаме посоката на който и да е негов представител.

Свободните вектори ще означаваме с малки латински букви със стрелка над тях. За краткост вместо *свободен вектор* по-нататък ще казваме *вектор*.

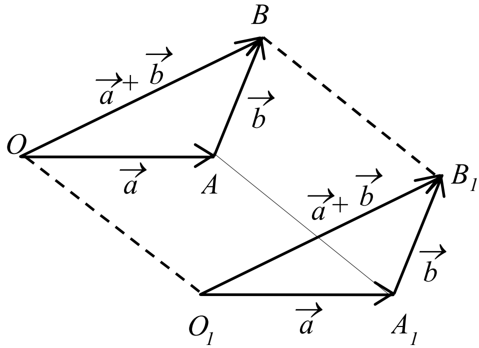
Множеството от всички нулеви насочени отсечки ще наричаме *нулев вектор* и ще го бележим с $\vec{0}$.

Ако векторът $\vec{a}$ е определен от представителя си $\overrightarrow{AB}$, то вектора, определен от $\overrightarrow{BA}$, означаваме с $-\vec{a}$ и го наричаме *противоположен* на $\vec{a}$.

Два вектора $\vec{a}$ и $\vec{b}$ наричаме *колинеарни*, ако който и да е представител на $\vec{a}$ е колинеарен на представител на $\vec{b}$. В частност, ако построим представители с общо начало за двата вектора, то те ще лежат на една права. Считаме, че *нулевият вектор е колинеарен на всеки вектор*.

4. Определение. *Сума* (*сбор*) на два вектора $\vec{a}$ и $\vec{b}$ наричаме вектора $\vec{c} = \vec{a} + \vec{b}$, определен по следния начин. Нека O е произволна точка, $\overrightarrow{OA} \in \vec{a}$ и $\overrightarrow{AB} \in \vec{b}$; тогава $\overrightarrow{OB} \in \vec{c}$, т.е. насочената отсечка $\overrightarrow{OB}$ е представител на $\vec{c}$.

Ще докажем, че даденото определение е коректно, т.е. че не зависи от избора на точката O и от представителите на $\vec{a}$ и $\vec{b}$. Нека O_1 е друга точка и $\overrightarrow{O_1A_1} \in \vec{a}$, $\overrightarrow{A_1B_1} \in \vec{b}$; трябва да докажем, че $\overrightarrow{O_1B_1} \in \vec{c}$, т.е. $\overrightarrow{OB} = \overrightarrow{O_1B_1}$ (черт.2).



черт. 2

По построение $\overrightarrow{OA} = \overrightarrow{O_1A_1}$ (представители на $\vec{a}$). Това означава, че четириъгълникът OO_1AA_1 е успоредник, откъдето следва, че $\overrightarrow{OO_1} = \overrightarrow{AA_1}$. Аналогично от $\overrightarrow{AB} = \overrightarrow{A_1B_1}$ следва $\overrightarrow{AA_1} = \overrightarrow{BB_1}$. Тогава от транзитивността на релацията "равенство на насочени отсечки" получаваме, че $\overrightarrow{OO_1} = \overrightarrow{BB_1}$, т.е. четириъгълникът OO_1B_1B е успоредник. Това означава, че $\overrightarrow{OB} = \overrightarrow{O_1B_1}$.

Да построим успоредника $OABC$, където $\overrightarrow{OA} \in \vec{a}$, $\overrightarrow{OC} \in \vec{b}$, като векторите са неколинеарни (черт.3). Тъй като $\overrightarrow{OC} = \overrightarrow{AB}$, то $\overrightarrow{AB} \in \vec{b}$ и тогава диагоналът $\overrightarrow{OB}$ е представител на $\vec{a} + \vec{b}$. Това геометрично построение на сумата (сбора) $\vec{a} + \vec{b}$ на два вектора обикновено се нарича *правило на успоредника*.

5. Свойства на операцията събиране на вектори. Ще докажем, че:

1) $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$ (асоциативност);

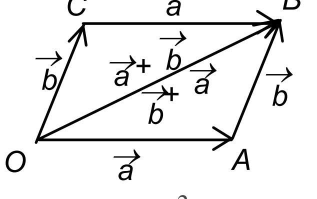
2) $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ (комутативност);

3) $\vec{a} + \vec{0} = \vec{a}$ за всеки свободен вектор $\vec{a}$;

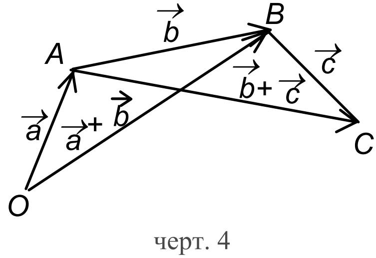
4) $\vec{a} + (-\vec{a}) = \vec{0}$ за всеки свободен вектор $\vec{a}$.

Доказателство. 1) Нека $\overrightarrow{OA} \in \vec{a}$, $\overrightarrow{AB} \in \vec{b}$, $\overrightarrow{BC} \in \vec{c}$ (черт.4). Съгласно определението в т.4 $\overrightarrow{OB} \in \vec{a} + \vec{b}$ и $\overrightarrow{AC} \in \vec{b} + \vec{c}$. Тогава $\overrightarrow{OC} \in (\vec{a} + \vec{b}) + \vec{c}$ и $\overrightarrow{OC} \in \vec{a} + (\vec{b} + \vec{c})$, което означава, че $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$.

Свойствата 3) и 4) следват непосредствено от определението за сума, а черт.3 достатъчно добре илюстрира свойство 2).



черт. 3

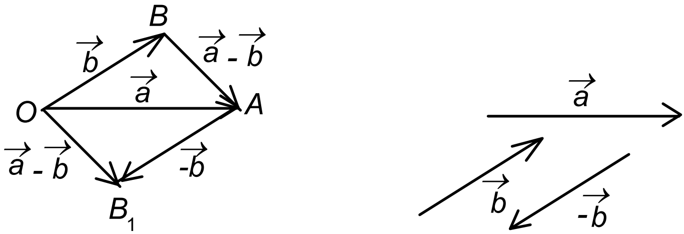


черт. 4

6. Определение. *Разлика* на векторите $\vec{a}$ и $\vec{b}$ (означение: $\vec{a}-\vec{b}$) се нарича сумата на векторите $\vec{a}$ и $-\vec{b}$, т.е.

$$\vec{a}-\vec{b}=\vec{a}+(-\vec{b})$$

Геометричното построяване на разликата се вижда от черт.5.



черт. 5

На черт.3 (правилото на успоредника) разликата $\vec{a}-\vec{b}$ е векторът, определен от насочената отсечка $\vec{CA}$. От черт.5 се вижда, че разликата $\vec{a}-\vec{b}$ има представители $\vec{OB_1}$ и $\vec{BA}$, защото четириъгълникът OB_1AB е успоредник.

7. Умножаване на вектор с реално число. *Произведение на вектора* $\vec{a}$ *с реалното число* λ наричаме вектора (ще го бележим с $\lambda\vec{a}$), дефиниран по следния начин: 1) $|\lambda\vec{a}|=|\lambda||\vec{a}|$; 2) ако $\lambda>0$, векторите $\vec{a}$ и $\lambda\vec{a}$ са колинеарни и имат еднакви посоки; 3) ако $\lambda<0$, векторите $\vec{a}$ и $\lambda\vec{a}$ са колинеарни и имат противоположни посоки.

От 1) следва, че $\lambda\vec{a}=\vec{0}$ точно когато $\lambda=0$ или $\vec{a}=\vec{0}$. Пак от определението веднага следва, че $(-1)\vec{a}=-\vec{a}$.

8. Операцията умножение на вектор с число има следните свойства:

1) $1.\vec{a}=\vec{a}$;

2) $\quad (\lambda\mu)\vec{a} = \lambda(\mu\vec{a})$;

3) $\quad (\lambda+\mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$;

4) $\quad \lambda(\vec{a}+\vec{b}) = \lambda\vec{a} + \lambda\vec{b}$.

Тук $\vec{a}$ и $\vec{b}$ са произволни вектори, а λ и μ - произволни реални числа.

Доказателство. 1) е очевидно. Ще докажем останалите свойства.

2) Векторите $\lambda(\mu\vec{a})$ и $(\lambda\mu)\vec{a}$ са колинеарни и имат равни дължини, защото

$$|\lambda(\mu\vec{a})| = |\lambda||\mu\vec{a}| = |\lambda||\mu||\vec{a}| = |\lambda\mu||\vec{a}|.$$

Остава да докажем, че те имат еднакви посоки. Ако $\lambda > 0$, $\mu > 0$, то $\lambda\mu > 0$ и всеки от векторите $\mu\vec{a}$, $\lambda(\mu\vec{a})$, $(\lambda\mu)\vec{a}$ има посоката на $\vec{a}$.

Ако поне едно от числата λ и μ е нула, то $\lambda(\mu\vec{a}) = \vec{0}$ и $(\lambda\mu)\vec{a} = \vec{0}$.

Ако $\lambda > 0$ и $\mu < 0$, то $\lambda\mu < 0$ и трите колинеарни вектора $\mu\vec{a}$, $\lambda(\mu\vec{a})$ и $(\lambda\mu)\vec{a}$ имат посоки, противоположни на посоката на $\vec{a}$. Следователно посоките на $\lambda(\mu\vec{a})$ и $(\lambda\mu)\vec{a}$ съвпадат. Останалите възможности за знаците на λ и μ се разглеждат аналогично.

3) Ако $\lambda = 0$ или $\mu = 0$, или $\vec{a} = \vec{0}$, равенството е тривиално.

Векторите $(\lambda+\mu)\vec{a}$ и $\lambda\vec{a} + \mu\vec{a}$ са колинеарни, защото всеки от тях е колинеарен на $\vec{a}$.

Нека λ и μ имат еднакви знаци. В такъв случай $\lambda\vec{a}$ и $\mu\vec{a}$ имат еднакви посоки и затова

$$|\lambda\vec{a}+\mu\vec{a}|=|\lambda\vec{a}|+|\mu\vec{a}|.$$

От друга страна

$$|\lambda\vec{a}|+|\mu\vec{a}|=|\lambda||\vec{a}|+|\mu||\vec{a}|=(|\lambda|+|\mu|)|\vec{a}|=|\lambda+\mu||\vec{a}|=|(\lambda+\mu)\vec{a}|,$$

т.е. векторите $(\lambda+\mu)\vec{a}$ и $\lambda\vec{a}+\mu\vec{a}$ имат равни дължини, колинеарни са и посоките им съвпадат. Следователно те са равни.

Да разгледаме другия случай, когато λ и μ имат различни знаци и нека за определеност $|\lambda|>|\mu|$. Тогава $\lambda\vec{a}$ и $\mu\vec{a}$ са колинеарни и имат противоположни посоки.

$$|\lambda\vec{a}+\mu\vec{a}|=|\lambda\vec{a}|-|\mu\vec{a}| \qquad (|\lambda\vec{a}|>|\mu\vec{a}|)$$

От друга страна

$$|(\lambda+\mu)\vec{a}|=|\lambda+\mu||\vec{a}|=(|\lambda|-|\mu|)|\vec{a}|=|\lambda||\vec{a}|-|\mu||\vec{a}|=|\lambda\vec{a}|-|\mu\vec{a}|.$$

Остава да докажем, че векторите $(\lambda+\mu)\vec{a}$ и $\lambda\vec{a}+\mu\vec{a}$ имат еднакви посоки. Тъй като $|\lambda|>|\mu|$, посоките им съвпадат с посоката на $\lambda\vec{a}$, т.е. те имат еднакви посоки.

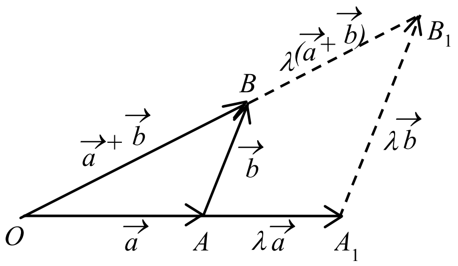
Накрая, ако $\lambda=-\mu$, то $(\lambda+\mu)\vec{a}=\vec{0}$ и $\lambda\vec{a}+\mu\vec{a}=\vec{0}$, което окончателно доказва исканото равенство.

4) Ако някой от векторите $\vec{a}$, $\vec{b}$ е нулевият или $\lambda=0$, равенството е очевидно. Ако $\vec{a}$, $\vec{b}$ са колинеарни и ненулеви, то $\vec{b}=\mu\vec{a}$, където $\mu=\dfrac{|\vec{b}|}{|\vec{a}|}$, ако $\vec{a}$ и $\vec{b}$ имат еднакви посоки и $\mu=-\dfrac{|\vec{b}|}{|\vec{a}|}$, ако имат противоположни посоки. Наистина

$$|\mu \vec{a}|=|\mu||\vec{a}|=\frac{|\vec{b}|}{|\vec{a}|}|\vec{a}|=|\vec{b}|.$$

В такъв случай

$$\lambda \vec{a}+\lambda \vec{b}=\lambda \vec{a}+\lambda(\mu \vec{a})=\lambda \vec{a}+(\lambda\mu)\vec{a}=(\lambda+\lambda\mu)\vec{a}=$$
$$=\left[\lambda(1+\mu)\right]\vec{a}=\lambda[(1+\mu)\vec{a}]=\lambda(\vec{a}+\mu \vec{a})=\lambda(\vec{a}+\vec{b}).$$



черт.6

Нека $\vec{a}$ и $\vec{b}$ са неколинеарни и за определеност $\lambda>0$ $(\lambda>1)$ (черт.6).

Разглеждаме триъгълниците OAB и OA_1B_1. Тъй като $\vec{b}$ и $\lambda\vec{b}$ са колинеарни, триъгълниците са подобни, защото $\angle OAB=\angle OA_1B_1$ и

$$\frac{OA}{OA_1}=\frac{|\vec{a}|}{|\lambda \vec{a}|}=\frac{1}{\lambda}=\frac{|\vec{b}|}{|\lambda \vec{b}|}=\frac{AB}{A_1B_1}.$$

От подобието на триъгълниците следва, че $\angle AOB=\angle A_1OB_1$, т.е. точките B и B_1 лежат на една права, минаваща през O. От черт.6 е ясно, че $\vec{OB}\in(\vec{a}+\vec{b})$, $\vec{OB_1}\in(\lambda\vec{a}+\lambda\vec{b})$. Пак от подобието на триъгълниците имаме

$$\frac{|\vec{OB}|}{|\vec{OB_1}|}=\frac{|\vec{OA}|}{|\vec{OA_1}|}=\frac{1}{\lambda}.$$

Тъй като $\lambda > 0$, от тук следва, че $\overrightarrow{OB_1} = \lambda \overrightarrow{OB}$, т.е. $\lambda(\vec{a} + \vec{b}) = \lambda \vec{a} + \lambda \vec{b}$. Случаят $\lambda < 0$ се разглежда аналогично.

Операциите събиране на свободни вектори и умножаване на свободен вектор с реално число обикновено се наричат *афинни*. Доказаните в т.5 и т.8 осем свойства на тези операции показват, че *свободните вектори образуват реално линейно (векторно) пространство*.

§ 2. Координати на вектор и на точка

1. Понятията и резултатите от гл. 4 се пренасят буквално и за линейното пространство на свободните вектори. Главната ни цел тук е да изтъкнем геометричния характер на тези понятия. Нов обект ще бъдат само точките и техните координати.

2. Теорема. *Два свободни вектора са линейно зависими тогава и само тогава, когато са колинеарни.*

Д о к а з а т е л с т в о . Нека векторите $\vec{a}$ и $\vec{b}$ са линейно зависими. Това означава, че съществуват числа λ_1 и λ_2, които не са едновременно нули и за които

$$\lambda_1 \vec{a} + \lambda_2 \vec{b} = \vec{0}.$$

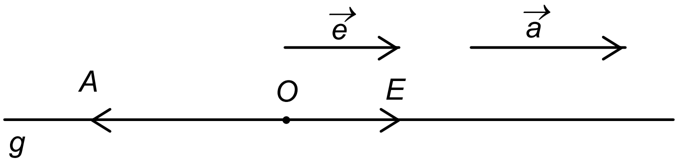
Нека за определеност $\lambda_1 \neq 0$. Тогава

$$\vec{a} = -\frac{\lambda_2}{\lambda_1} \vec{b}$$

и съгласно § 1, т. 7 $\vec{a}$ е колинеарен на $\vec{b}$.

Обратно, нека векторите $\vec{a}$ и $\vec{b}$ са колинеарни. Ако и двата вектора са нулеви, те очевидно са линейно зависими. Нека $\vec{a} \neq \vec{0}$. В такъв случай $\vec{b} = \mu \vec{a}$, където $\mu = \frac{|\vec{b}|}{|\vec{a}|}$, ако векторите $\vec{a}$ и $\vec{b}$ са еднопосочни и $\mu = -\frac{|\vec{b}|}{|\vec{a}|}$, ако са с противоположни посоки. Очевидно $\mu \vec{a} + (-1)\vec{b} = \vec{0}$, т.е. векторите са линейно зависими.

3. Нека g е дадена права. Избираме върху нея точка O и ненулев вектор $\vec{e}$ с представител $\vec{OE}$, където $E \in g$ (черт.7).



черт.7

Нека $\vec{a}$ е произволен вектор, колинеарен на правата g. Тъй като той е колинеарен на $\vec{e}$, от доказателството на теоремата в т.2 се вижда, че съществува еднозначно определено реално число x, за което

$$\vec{a} = x\,\vec{e}\,.$$

Числото x е нула, ако $\vec{a} = \vec{0}$, x е положително, ако $\vec{a}$ и $\vec{e}$ имат еднакви посоки и отрицателно - ако имат противоположни посоки. Обратно, ако оставим x да пробяга реалните числа, то $\vec{a} = x\,\vec{e}$ ще пробяга множеството на всички вектори, колинеарни на правата g.

Съвкупността от точката O и ненулев вектор $\vec{e}$ (означение $O\vec{e}$) се нарича *афинна координатна система върху правата* g. Точката O се нарича *начало* на координатната система. Еднозначно определеното реално число x, съответстващо на вектора $\vec{a}$, се нарича негова *афинна координата спрямо координатната система* $O\vec{e}$.

Читателят сигурно е забелязал, че координатата x на вектора $\vec{a}$ изобщо не зависи от избора на началото O. В терминологията на гл.4 ситуацията изглежда така: в едномерното пространство на свободните вектори, колинеарни с дадена права g, избрахме в качеството на базис произволен ненулев вектор $\vec{e}$ и коефициента x от равенството $\vec{a} = x\,\vec{e}$ нарекохме координата на вектора $\vec{a}$.

Ролята на началото O е съществена при въвеждането на координати на точките. Нека A е произволна точка от правата g. Векторът с представител $\overrightarrow{OA}$ се нарича *радиус-вектор на точката* A. Координатата на радиус-вектора на A се нарича *координата на точката* A спрямо афинната координатна система $O\vec{e}$, т.е.

$$\overrightarrow{OA} = x_A \overrightarrow{OE} \qquad\qquad (x_A \text{ - координата на точката } A).$$

Ако $|\vec{e}|=1$, координатната система $O\vec{e}$ се нарича *нормирана*. Сега от $|\vec{e}|=1$ следва

$$|\vec{a}|=|x\,\vec{e}|=|x||\vec{e}|=|x|,$$

т.е. при *нормирана координатна система върху права дължината на вектор, колинеарен на правата, е равна на абсолютната стойност на неговата координата*.

4. Определение. Ако представители на няколко вектора лежат в една равнина (еквивалентно: векторите са успоредни на поне една равнина), ще казваме, че векторите са *компланарни*.

5. Теорема. *Три вектора са линейно зависими тогава и само тогава, когато са компланарни*.

Доказателство. а) Нека векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ са линейно зависими. Следователно съществуват числата λ_1, λ_2, λ_3, които не са едновременно нули така, че да е изпълнено равенството

$$\lambda_1\vec{a}+\lambda_2\vec{b}+\lambda_3\vec{c}=\vec{0}.$$

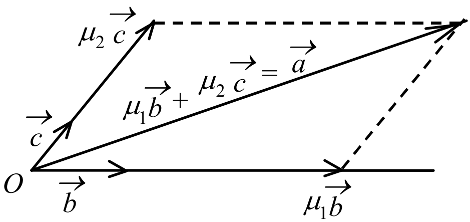
Да допуснем, че $\lambda_1 \neq 0$. Тогава

$$\vec{a}=(-\frac{\lambda_2}{\lambda_1})\vec{b}+(-\frac{\lambda_3}{\lambda_2})\vec{c},$$

т.е. стигаме до равенство от вида

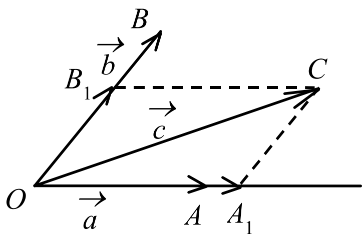
$$\vec{a}=\mu_1\vec{b}+\mu_2\vec{c}.$$

Ако $\vec{b}$ и $\vec{c}$ са колинеарни, то векторът $\vec{a}$ е колинеарен с тях, т.е. $\vec{a}$, $\vec{b}$, $\vec{c}$ са колинеарни и следователно компланарни. С помощта на правилото за събиране на вектори лесно се съобразява, че ако векторите $\vec{b}$ и $\vec{c}$ не са колинеарни, то $\vec{b}$, $\vec{c}$ и всяка тяхна линейна комбинация са компланарни. Ситуацията е илюстрирана на черт.8 при $\mu_1 > 0$ и $\mu_2 > 0$.



черт.8

б) Обратно, нека $\vec{a}$, $\vec{b}$, $\vec{c}$ са компланарни. Ако два от тях, например $\vec{a}$ и $\vec{b}$, са колинеарни, то съгласно теоремата от т.2 те са линейно зависими, т.е. съществуват числа λ и μ, които не са едновременно нули и $\lambda \vec{a} + \mu \vec{b} = \vec{0}$, откъдето $\lambda \vec{a} + \mu \vec{b} + 0.\vec{c} = \vec{0}$, т.е. трите вектора са линейно зависими. Тъй като по определение нулевият вектор е колинеарен на всеки вектор, остава да разгледаме случая, когато и трите вектора са ненулеви и никои два от тях не са колинеарни. За целта в една равнина построяваме представители $\vec{OA} \in \vec{a}$, $\vec{OB} \in \vec{b}$, $\vec{OC} \in \vec{c}$ (черт.9). Построяваме успоредника OA_1CB_1 така, че $\vec{OC}$ да бъде диагонал, а A_1 и B_1 да лежат съответно върху правите OA и OB. Тъй като $\vec{a}$ и $\vec{b}$ са ненулеви вектори, то $\vec{OA_1} = \lambda \vec{a}$ и $\vec{OB_1} = \mu \vec{b}$. Но



черт.9

$$\vec{OC} = \vec{OA_1} + \vec{OB_1},$$

т.е. $\vec{c} = \lambda \vec{a} + \mu \vec{b}$ или все едно $\lambda \vec{a} + \mu \vec{b} + (-1).\vec{c} = 0$. Теоремата е доказана. От доказателството получаваме следните важни

6. Следствия. а) *В пространството на свободните вектори всеки три некомпланарни вектора са линейно независими.*

б) А*ко два вектора* $\vec{a}$, $\vec{b}$ *са неколинеарни и са компланарни с дадена равнина* ε, *то всеки вектор* $\vec{c}$, *който е компланарен с* ε, *се представя, и то еднозначно, като линейна комбинация на векторите* $\vec{a}$ *и* $\vec{b}$.

Тъй като а) е очевидно, ще докажем б). Съществуването на представяне $\vec{c} = \lambda \vec{a} + \mu \vec{b}$ беше доказано по-горе. Нека $\vec{c} = \lambda_1 \vec{a} + \mu_1 \vec{b}$ е още едно такова представяне. От двете равенства следва

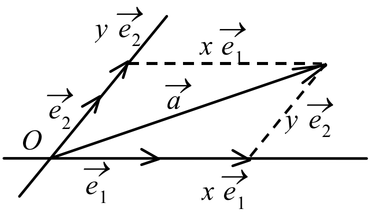
$$\vec{0} = (\lambda - \lambda_1)\vec{a} + (\mu - \mu_1)\vec{b},$$

а тъй като $\vec{a}$ *и* $\vec{b}$ са линейно независими, то $\lambda - \lambda_1 = 0$, $\mu - \mu_1 = 0$, т.е. $\lambda = \lambda_1$, $\mu = \mu_1$ и представянето е единствено.

7. Афинна координатна система в равнината. Ако в дадена равнина ε са избрани точка O и двойка *неколинеарни* вектори $\vec{e}_1$ и $\vec{e}_2$, ще казваме, че е зададена *афинна координатна система* $O\vec{e}_1\vec{e}_2$ *в равнината* ε *с начало* O *и базисни вектори* $\vec{e}_1$, $\vec{e}_2$.

В т.6 видяхме, че за произволен вектор $\vec{a}$, компланарен с дадената равнина, е изпълнено

$$\vec{a} = x\vec{e}_1 + y\vec{e}_2 .$$



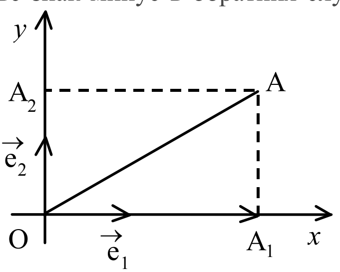
черт.10

Еднозначно определените числа x и y се наричат *афинни координати на вектора* $\vec{a}$ *спрямо координатна система* $O\vec{e}_1\vec{e}_2$. Нагледно това е показано на черт.10, разположен изцяло в равнината ε. Представителите на векторите са взети с начало точката O. Тя се нарича *начало* на координатната система. Ще подчертаем, че координатите на вектора $\vec{a}$ не зависят от избора на началото $\vec{a}$: ако O' е произволна точка в равнината ε, векторът $\vec{a}$ спрямо координатната система $O'\vec{e}_1\vec{e}_2$ (базисните вектори $\vec{e}_1$ и $\vec{e}_2$ са същите!) има същите координати.

Дотук разглежданията в т.6 и т.7 може да се интерпретират и така: *свободните вектори, компланарни с дадена равнина, образуват двумерно линейно пространство; всяка двойка неколинеарни вектори от това пространство е негов базис*.

Ролята на началото O е съществена при въвеждане на координати на точките от равнината. Нека A е произволна точка от равнината ε; векторът с представител $\vec{OA}$ се нарича *радиус-вектор на точката* A. Неговите координати (т.е. координатите на $\vec{OA}$) се наричат *координати на точката* A спрямо дадената афинна координатна система.

Ако базисните вектори $\vec{e}_1$ и $\vec{e}_2$ имат дължини единица ($|\vec{e}_1|=|\vec{e}_2|=1$) и са ортогонални помежду си, координатната система се нарича *ортонормирана (декартова)*. Тя е добре позната от училищния курс по математика. Оста $O\vec{e}_1$ (правата през O, колинеарна на $\vec{e}_1$ и с посока, съвпадаща с посоката на $\vec{e}_1$) се нарича *абсцисна*, а оста $O\vec{e}_2$ - *ординатна*. Традиционното означение на тези оси е съответно Ox и Oy. Ще напомним, че *алгебрична мярка* на насочената отсечка $\vec{AB}$, лежаща върху някаква ос (означение: $\overline{AB}$), се нарича дължината на отсечката, взета със знак плюс, ако посоката на отсечката съвпада с посоката на оста, и взета със знак минус в обратния случай. Декартовите координати на точката A имат прост геометричен смисъл (черт. 11). Нека A_1 и A_2 са ортогоналните проекции на точката A съответно върху абсцисната и ординатната ос. Първата координата (абсцисата) на A е алгебричната мярка на насочената отсечка $\vec{OA}_1$, а втората координата (ординатата) –



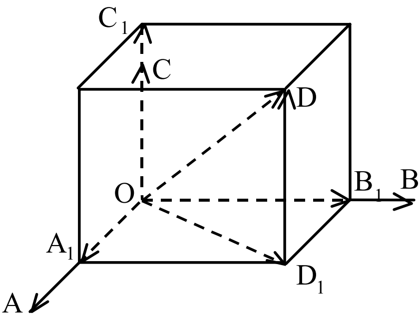
черт.11

алгебричната мярка на $\overrightarrow{OA_2}$.

8. Теорема. *Всеки четири свободни вектора са линейно зависими.*

Д о к а з а т е л с т в о . Нека векторите са $\vec{a}$, $\vec{b}$, $\vec{c}$, $\vec{d}$. Ако три от тях, например $\vec{a}$, $\vec{b}$, $\vec{c}$, са компланарни, то те са линейно зависими. Съществува ненулева тройка от числа λ, μ, ν, за които

$$\lambda \vec{a} + \mu \vec{b} + \nu \vec{c} = \vec{0},$$



черт.12

следователно

$$\lambda \vec{a} + \mu \vec{b} + \nu \vec{c} + 0.\vec{d} = \vec{0},$$

т.е. четирите вектора са линейно зависими.

По-нататък нека всеки три от дадените вектори $\vec{a}$, $\vec{b}$, $\vec{c}$, $\vec{d}$ са некомпланарни, т.е. линейно независими. С начало точката O построяваме съответно техните представители $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$, $\overrightarrow{OD}$ (черт.12).

Построяваме паралелепипед, в който точките O и D да бъдат срещуположни върхове, а трите ръба през върха O да лежат върху правите, OB и OC . За целта е достатъчно освен дадените равнини, определени от OAB , OAC , OBC , да построим още три, минаващи през D и съответно успоредни на всяка една от споменатите три. В означенията от чертежа имаме

$$\overrightarrow{OD} = \overrightarrow{OC_1} + \overrightarrow{OD_1}, \qquad \overrightarrow{OD_1} = \overrightarrow{OA_1} + \overrightarrow{OB_1},$$

откъдето

$$\overrightarrow{OD} = \overrightarrow{OA_1} + \overrightarrow{OB_1} + \overrightarrow{OC_1}.$$

Тъй като разглежданите вектори са в частност ненулеви, съществуват еднозначно определени числа λ, μ, ν, за които

$$\overrightarrow{OA_1} = \lambda \overrightarrow{OA}, \qquad \overrightarrow{OB_1} = \mu \overrightarrow{OB}, \qquad \overrightarrow{OC_1} = \nu \overrightarrow{OC}$$

Следователно

$$\overrightarrow{OD} = \lambda\overrightarrow{OA} + \mu\overrightarrow{OB} + \nu\overrightarrow{OC},$$

или все едно $\vec{d} = \lambda\vec{a} + \mu\vec{b} + \nu\vec{c}$, $\lambda\vec{a} + \mu\vec{b} + \nu\vec{c} + (-1)\vec{d} = \vec{0}$, т.е. векторите са отново линейно зависими.

От т. 6, а) и от доказателството, което току-що изложихме, следва важно

9. Твърдение. *Линейното пространство на свободните вектори е тримерно. Ако са дадени три некомпланарни вектора, то всеки вектор се представя еднозначно като тяхна линейна комбинация.*

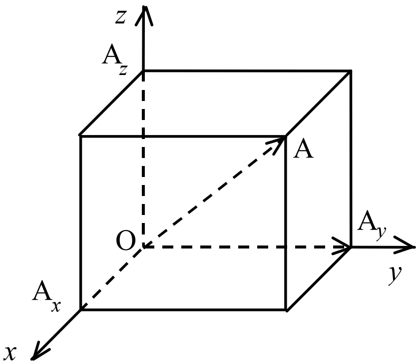
10. Афинна координатна система в пространството. Ако в пространството е избрана точка O и тройка *некомпланарни* свободни вектори $\vec{e_1}$, $\vec{e_2}$, $\vec{e_3}$, ще казваме, че е зададена *афинна координатна система* $O\vec{e_1}\vec{e_2}\vec{e_3}$ с начало O и базисни вектори $\vec{e_1}$, $\vec{e_2}$, $\vec{e_3}$. Вече видяхме, че ако $\vec{a}$ е произволен вектор, то съществуват числа x, y, z, за които

$$\vec{a} = x\vec{e_1} + y\vec{e_2} + z\vec{e_3}.$$

Те са еднозначно определени - факт, който е непосредствено следствие от линейната независимост на $\vec{e_1}$, $\vec{e_2}$, $\vec{e_3}$. Наредената тройка числа (x, y, z) се наричат *афинни координати на вектора* $\vec{a}$ спрямо дадената координатна система. Макар и негласно, за базисните вектори винаги се предполага, че е казано кой от тях е първи, кой е втори и кой е трети. (Счита се например, че $O\vec{e_2}\vec{e_1}\vec{e_3}$ е друга координатна система.) Базисните вектори $\vec{e_1}$, $\vec{e_2}$, $\vec{e_3}$ очевидно имат координати съответно $(1, 0, 0)$, $(0,1,0)$, $(0,0,1)$.

Ако A е произволна точка, векторът с представител $\overrightarrow{OA}$ се нарича *радиус-вектор на точката* A. Координатите на радиус-вектора $\overrightarrow{OA}$ се наричат *афинни координати на точката* A спрямо дадената координатна система.

Ако читателят иска да вникне в геометричния смисъл на координатите, уместно е да се върне към черт.12 и да си мисли, че базисни са векторите $\vec{a}$, $\vec{b}$, $\vec{c}$. Построенията обяснихме в т.8; координатите на вектора $\vec{d}$ (или все едно на точката D) са (λ, μ, ν).



черт.13

Ако базисните вектори са с дължини единица и са два по два ортогонални (перпендикулярни), координатната система се нарича *ортонормирана* (*декартова*). Осите, определени от $O\vec{e_1}$, $O\vec{e_2}$ и $O\vec{e_3}$, се наричат съответно *абсцисна, ординатна* и *апликатна*. Често ги означават съответно с Ox, Oy, Oz (черт.13). Равнините, определени от които и да са две от координатните оси, се наричат *координатни равнини* и се използват традиционни означения: осите Ox и Oy определят равнината Oxy, равнината Oyz се определя от осите Oy и Oz, равнината Oxz се определя от осите Ox и Oz. Геометричният смисъл на координатите спрямо декартова координатна система е особено прост: точката A се проектира ортогонално върху координатните оси и нека проекциите ѝ върху Ox, Oy, Oz са съответно A_x, A_y, A_z; алгебричните мерки на насочените отсечки $\overrightarrow{OA_x}$, $\overrightarrow{OA_y}$, $\overrightarrow{OA_z}$ наричаме съответно *абсциса, ордината* и *апликата* на A. Разбира се, проектирането може да се реализира, като се прекарват през A равнини, успоредни на координатните, и после се търсят прободните им точки с осите.

11. Теорема. *Координатите на вектор, който е линейна комбинация на няколко вектора, са също линейни комбинации на съответните координати на векторите.*

Д о к а з а т е л с т в о . Нека спрямо дадена афинна координатна система $O\vec{e_2}\vec{e_1}\vec{e_3}$ векторите $\vec{a_1}, \vec{a_2} \ldots, \vec{a_k}$ имат съответно координати (x_i, y_i, z_i), $i = 1, 2, \ldots, k$, и нека

$$\vec{a} = \lambda_1 \vec{a_1} + \lambda_2 \vec{a_2} + \ldots + \lambda_k \vec{a_k}.$$

Ако координатите на $\vec{a}$ са (x, y, z), тъй като $\vec{a_i} = x_i \vec{e_1} + y_i \vec{e_2} + z_i \vec{e_3}$, то

$$x\vec{e_1} + y\vec{e_2} + z\vec{e_3} = \sum_{i=1}^{k} \lambda_i (x_i \vec{e_1} + y_i \vec{e_2} + z_i \vec{e_3}) =$$

$$= (\sum_{i=1}^{k} \lambda_i x_i)\vec{e_1} + (\sum_{i=1}^{k} \lambda_i y_i)\vec{e_2} + (\sum_{i=1}^{k} \lambda_i z_i)\vec{e_3}.$$

Следователно

$$x = \sum_{i=1}^{k} \lambda_i x_i, \qquad y = \sum_{i=1}^{k} \lambda_i y_i, \qquad z = \sum_{i=1}^{k} \lambda_i z_i.$$

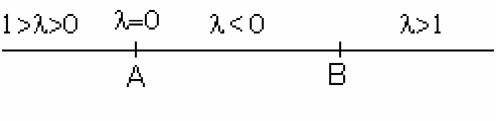
От доказаната теорема и от резултата от т.2 лесно следва, че *два вектора са колинеарни тогава и само тогава, когато съответните им координати са пропорционални*.

12. Координати на вектор, зададен с наредена двойка точки. Нека точките A и B спрямо афинна координатна система $O\vec{e_2}\vec{e_1}\vec{e_3}$ имат координати съответно (x_1, y_1, z_1) и (x_2, y_2, z_2). Тогава *векторът, определен от насочената отсечка* $\vec{AB}$, *има координати* $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$ ("координатите на края минус координатите на началото"). Този полезен факт следва непосредствено от равенството $\vec{AB} = \vec{OB} - \vec{OA}$ и от теоремата в т.11.

13. Деление на отсечка в дадено отношение. Просто отношение на три точки. Нека A, B, C е наредена тройка точки, лежащи върху една права, и нека $A \neq B$. Тогава векторите $\vec{AC}$ и $\vec{BC}$ са колинеарни и съществува число λ, такова, че е в сила равенството

$$\vec{AC} = \lambda \vec{BC}$$

и казваме, че точката C *дели насочената отсечка* $\vec{AB}$ *в отношение* λ. Числото λ се нарича *просто отношение* на точките A, B, C (взети в този ред). Обикновено се използва означението $\lambda = (ABC)$.



черт. 14

Ако втората и третата точка съвпадат $\left(B=C\right)$, отсечката $\overrightarrow{BC}$ е нулева, $\overrightarrow{AC}$ е ненулева и простото отношение $\left(ABC\right)$ очевидно не е дефинирано. Във всички останали случаи отсечката $\overrightarrow{BC}$ е ненулева и λ е еднозначно определено, тъй като векторите с представители $\overrightarrow{AC}$ и $\overrightarrow{BC}$ са колинеарни. Оставяме на читателя да съобрази самостоятелно, че ако точката C пробяга всички точки от правата с изключение на точката B, то простото отношение $\lambda=\left(ABC\right)$ пробягва всички реални числа с изключение на числото 1. Ако се съобрази почти очевидният факт, че щом $C\neq C'$, то $\left(ABC\right)\neq\left(ABC'\right)$, тогава лесно се получава, че с помощта на простото отношение може да се построи биекция между множеството на реалните числа без числото 1 и множеството на всички точки от правата без точката B. По-точно: за всяко $\lambda\in\mathrm{R}$, $\lambda\neq 1$, съществува, и то еднозначно определена точка C от правата, за която $\left(ABC\right)=\lambda$.

А сега да изберем произволна координатна система $O\vec{e}_1\vec{e}_2\vec{e}_3$ и нека спрямо нея точките A, B, C имат координати съответно $\left(x_1,y_1,z_1\right)$, $\left(x_2,y_2,z_2\right)$ и $\left(x_3,y_3,z_3\right)$. От векторните равенства

$$\overrightarrow{AC}=\overrightarrow{OC}-\overrightarrow{OA}, \qquad \overrightarrow{BC}=\overrightarrow{OC}-\overrightarrow{OB}$$

получаваме $\overrightarrow{OC}-\overrightarrow{OA}=\lambda(\overrightarrow{OC}-\overrightarrow{OB})$, откъдето

$$\overrightarrow{OC}=\frac{1}{1-\lambda}\overrightarrow{OA}-\frac{\lambda}{1-\lambda}\overrightarrow{OB}.$$

Координатите на радиус-векторите $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$ са координатите на точките A, B, C и според теоремата в т.11 имаме

$$x = \frac{x_1 - \lambda x_2}{1 - \lambda}, \qquad y = \frac{y_1 - \lambda y_2}{1 - \lambda}, \qquad z = \frac{z_1 - \lambda z_2}{1 - \lambda}.$$

Ако точката C е средата на $\overrightarrow{AB}$, то очевидно $\lambda = -1$ и за координатите на C получаваме

$$x = \frac{1}{2}\left(x_1 + x_2\right), \qquad y = \frac{1}{2}\left(y_1 + y_2\right), \qquad z = \frac{1}{2}\left(z_1 + z_2\right).$$

Ако в горните формули оставим λ да клони към 1, то x, y, z ще клонят към $\pm\infty$ (всичко зависи от това, дали λ приближава единицата отляво или отдясно). Изглежда съблазнително към правата да се присъедини някаква нова, фиктивна точка, да я наречем "безкрайна" и да я означим с U, за която да имаме по определение $\left(A, B, U\right) = 1$. Тогава с помощта на простото отношение ще се реализира биекция между множеството $\mathbb{R}$ на реалните числа и множеството от точките на "разширената" права. Граничният преход, за който стана дума, подсказва, че когато C се отдалечава от B до безкрайност наляво, ще отиде в U, но същото ще стане (или би трябвало да стане), ако C се отдалечава от B "безкрайно надясно". Излиза, че "разширената права" повече прилича на окръжност (двата "края" $+\infty$ и $-\infty$ се "връзват" в точката U). През миналия век математиците поставили здрава логическа база под тази фантазия (и не само под нея) и построили красива и хармонична теория, която днес наричаме проективна геометрия.

13. Аналитично условие за компланарност на три вектора. Нека координатната система $O\vec{e_1}\vec{e_2}\vec{e_3}$ е произволна. Ще покажем, че *векторите* $\vec{a}$, $\vec{b}$, $\vec{c}$ *съответно с координати* $\left(x_1, y_1, z_1\right)$, $\left(x_2, y_2, z_2\right)$ и $\left(x_3, y_3, z_3\right)$ *са компланарни тогава и само тогава, когато*

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = 0.$$

Действително, както видяхме в § 2, т.5, компланарност е геометричен еквивалент на линейна зависимост. Следователно $\vec{a}$, $\vec{b}$, $\vec{c}$ са компланарни точно тогава, когато съществуват λ, μ, ν, които не са едновременно нули и за които

$$\lambda\vec{a} + \mu\vec{b} + \nu\vec{c} = \vec{0}.$$

Тогава според т.11 за съответните координати имаме

$$\lambda x_1 + \mu x_2 + \nu x_3 = 0$$

$$\lambda y_1 + \mu y_2 + \nu y_3 = 0$$

$$\lambda z_1 + \mu z_2 + \nu z_3 = 0.$$

Ако разгледаме трите равенства като хомогенна система от три уравнения с неизвестни λ, μ, ν, то тя има ненулево решение тогава и само тогава, когато детерминантата от коефициентите пред неизвестните е равна на нула, а това е точно изписаната по-горе детерминанта (по-точно транспонираната й).

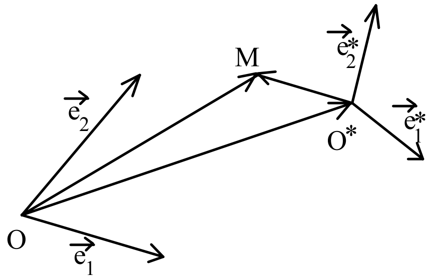
Ето още едно разсъждение, което води до същия резултат, но се основава на фактите от гл. 4, § 3. Както видяхме там, при изоморфизъм свойството на няколко вектора да бъдат линейно зависими (съответно независими) се запазва. Нека L е пространството на свободните вектори. То има размерност 3, следователно е изоморфно на R^3. При избрания вече базис на L да разгледаме координатния изоморфизъм $x\vec{e_1} + y\vec{e_2} + z\vec{e_3} \to$ $\to (x, y, z)$ (гл. 4, § 3, т. 3). Образи на $\vec{a}$, $\vec{b}$, $\vec{c}$ са съответните им тройки от координати (елементи на R^3). Да съставим матрица, в която редовете са x_i, y_i, z_i, $i = 1, 2, 3$. Векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ са линейно зависими точно тогава, когато редовете на тази матрица са линейно зависими, а това е еквивалентно на условието детерминантата й да бъде равна на нула.

§ 3. Смяна на координатите

1. Смяна на координатите в равнината. Да разгледаме в равнината две афинни координатни системи: $O\vec{e_1}\vec{e_2}$ и $O^*\vec{e_1^*}\vec{e_2^*}$. За удобство първата от тях ще наричаме "стара", а втората - "нова" (черт.15); същите прилагателни ще употребяваме и за координатите.
Нека освен това

$$\vec{e_1^*} = \alpha_{11}\vec{e_1} + \alpha_{21}\vec{e_2}, \qquad \vec{e_2^*} = \alpha_{12}\vec{e_1} + \alpha_{22}\vec{e_2}.$$

Както видяхме, множеството на векторите, компланарни с равнината $O\vec{e_1}\vec{e_2}$, е двумерно векторно пространство, а $\vec{e_1}$, $\vec{e_2}$ и $\vec{e_1^*}$, $\vec{e_2^*}$ са два



черт. 15

негови базиса. Следователно, ако при $n=2$ приложим резултатите от гл. 3, § 5, т. 1 за произволен вектор

$$\vec{a} = x\vec{e}_1 + y\vec{e}_2 = x'\vec{e}_1^* + y'\vec{e}_2^*$$

получаваме

$$(1) \qquad \begin{aligned} x &= \alpha_{11}x' + \alpha_{12}y' \\ y &= \alpha_{21}x' + \alpha_{22}y'. \end{aligned}$$

(Тук $x_1 = x$, $x_2 = y$ и $x_1' = x'$, $x_2' = y'$. Препоръчваме на читателя за упражнения да получи (1) самостоятелно.) Ако означим с $T = (\alpha_{ij})$ матрицата на прехода от стария базис към новия,

$$T = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix},$$

равенството (1) се записва матрично като

$$(2) \qquad \begin{pmatrix} x \\ y \end{pmatrix} = T \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

(Ако матрицата-стълб от новите координати умножим отляво с матрицата на прехода, ще получим матрицата-стълб от старите координати.) Разбира се матрицата T е неособена, защото векторите $\vec{e}_1^*$, $\vec{e}_2^*$ са линейно независими (неколинеарни). Следователно от (2) след умножаване отляво с T^{-1} ще получим

$$(3) \qquad \begin{pmatrix} x' \\ y' \end{pmatrix} = T^{-1} \begin{pmatrix} x \\ y \end{pmatrix}.$$

И така, формулите (1) или (2) дават правилото, по което се сменят координатите на векторите. По-интересно е как се сменят координатите на точките. Нека

$$\overrightarrow{OO^*} = a\overrightarrow{e_1} + b\overrightarrow{e_2},$$

т.е. новото начало има стари координати (a, b). Ако старите координати на произволна точка M са (x, y), а новите са (x', y'), от равенството

$$\overrightarrow{O^*M} = \overrightarrow{OM} - \overrightarrow{OO^*}$$

следва, че векторът $\overrightarrow{O^*M}$ има стари координати $(x-a, y-b)$; новите му координати са (x', y'). (Напомняме, че координатите на точка са координатите на нейния радиус-вектор.) Като приложим за вектора $\overrightarrow{O^*M}$ формулите (1) и пренесем $-a$ и $-b$ вдясно, получаваме

$$(4) \qquad \begin{aligned} x &= \alpha_{11}x' + \alpha_{12}y' + a \\ y &= \alpha_{21}x' + \alpha_{22}y' + b. \end{aligned}$$

Това всъщност е правилото, по което се сменят координатите на точките. В матрична форма то има вида

$$(5) \qquad \begin{pmatrix} x \\ y \end{pmatrix} = T \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

За афинните координатни системи $O\overrightarrow{e_1}\overrightarrow{e_2}$ и $O^*\overrightarrow{e_1^*}\overrightarrow{e_2^*}$ (базисните вектори са едни и същи; различно е само началото) се казва, че втората е получена от първата чрез *транслация* (успоредно пренасяне) с вектор $\overrightarrow{OO^*}$. В този частен случай формулите (4) добиват вида

$$x = x' + a$$

$$y = y' + b.$$

Тук ще направим едно необходимо отклонение на тема *ъгли*. В елементарната геометрия *ъгъл* се дефинира като геометрична фигура, определена от два лъча с общо начало. Говори се за *вътрешност* на ъгъла и за негова *мярка* (*големина*). Ако лъчите са означени с p и q, мярката на ъгъла, измерван с радиани, ще бележим с $(p,q)_e$; буквата e следва да напомня, че става дума за ъгъл в смисъла на елементарната геометрия (*елементарно-геометричен ъгъл*). Напомняме, че винаги $0 \le (p,q)_e \le \pi$; $(p,q)_e = 0$ единствено, когато лъчите p и q съвпадат. Тук изобщо не се държи сметка кой от лъчите p и q е първи и кой - втори.

Ще казваме, че два лъча с общо начало определят *насочен* (*ориентиран*) ъгъл, ако е посочено кой от лъчите е първи и кой - втори. Ако лъчът p е първи, а q - втори, ще пишем $\angle(p,q)$. Мярка на насочен ъгъл се въвежда малко по-сложно. За целта ще напомним, че в равнината може да се изберат точно две посоки на въртене. По традиция посоката на въртене, обратна на посоката на въртене на часовниковата стрелка, наричаме положителна, а другата посока - отрицателна. За ориентирания ъгъл $\angle(p,q)$ ще казваме, че е *положително ориентиран*, ако посоката на въртене на лъча p към q в равнината, определена от p и q, при което се описва вътрешността на елементарно-геометричния ъгъл, е положителна; в противен случай ще казваме, че ъгълът е *отрицателно ориентиран*. На ориентирания ъгъл $\angle(p,q)$ съпоставяме число (p,q), което наричаме *мярка* на $\angle(p,q)$ и определяме по следния начин:

$$(p,q) = \varepsilon (p,q)_\ell + k.2\pi,$$

където $\varepsilon = 1$, ако ъгълът е положително ориентиран и $\varepsilon = -1$ в противоположния случай; $k = 0, \pm 1, \pm 2, \dots$ е произволно цяло число. Ще подчертаем, че мярката на насочен ъгъл не е определена еднозначно, а само с точност до целочислено кратно на 2π. Предимствата на тази наглед неестествена дефиниция стават ясни от практиката.

От определението непосредствено следва, че

(6) $$(p,q) = -(q,p) + k.2\pi.$$

За упражнение на читателя оставяме следната *релация на Шал*: ако три лъча a, b, c са с общо начало и лежат в една равнина, то

(7) $$\left(a,b\right)+\left(b,c\right)+\left(c,a\right)=k.2\pi\ .$$

Нека $\vec{a}$ и $\vec{b}$ са два ненулева свободни вектора и C е произволна точка, като $\vec{CA}$ и $\vec{CB}$ са съответните представители на векторите. Под насочен ъгъл между векторите $\vec{a}$ и $\vec{b}$ разбираме насочения ъгъл $\angle\left(CA^{\rightarrow},CB^{\rightarrow}\right)$, т.е. $\angle\left(CA^{\rightarrow},CB^{\rightarrow}\right)=\angle(\vec{a},\vec{b})$.

А сега да се върнем към смяната на координатите в частния случай, когато и двете координатни системи $O\vec{e_1}\vec{e_2}$ и $O^*\vec{e_1^*}\vec{e_2^*}$ са ортонормирани (декартови). За положителна посока на въртене в равнината (ориентираме равнината) избираме онази, при която $\angle(\vec{e_1},\vec{e_2})$ е положително ориентиран. Нека при така избраната ориентация $\angle(\vec{e_1},\vec{e_1^*})=\varphi$. От геометрични съображения е ясно, че координатите на единичния вектор $\vec{e_1^*}$ спрямо координатната система $O\vec{e_1}\vec{e_2}$ са $\left(\cos\varphi,\sin\varphi\right)$, т.е.

$$\alpha_{11}=cos\varphi,\qquad \alpha_{21}=sin\varphi.$$

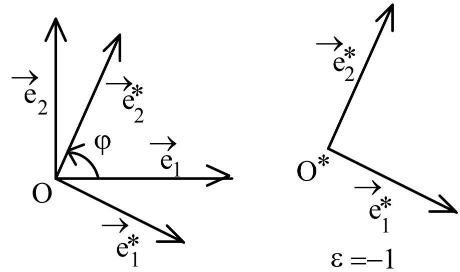
Съгласно релацията на Шал имаме $(\vec{e_1},\vec{e_2^*})+(\vec{e_2^*},\vec{e_1^*})+(\vec{e_1^*},\vec{e_1})=k.2\pi$, което заедно с (6) дава

(8) $$(\vec{e_1},\vec{e_2^*})=(\vec{e_1},\vec{e_1^*})+(\vec{e_1^*},\vec{e_2^*})+k.2\pi=\varphi+\varepsilon\frac{\pi}{2}+k.2\pi\ .$$

Следователно координатите на единичния вектор $\vec{e_2^*}$ спрямо $O\vec{e_1}\vec{e_2}$ са:

$$\alpha_{12}=cos(\vec{e_1},\vec{e_2^*})=cos(\varphi+\varepsilon\frac{\pi}{2})=-\varepsilon\, sin\varphi,$$

$$\alpha_{22}=sin(\vec{e_1},\vec{e_2^*})=sin(\varphi+\varepsilon\frac{\pi}{2})=\varepsilon\, cos\varphi.$$



черт.16

И така, в разглеждания частен случай матрицата T добива вида

$$T = \begin{pmatrix} \cos\varphi & -\varepsilon\sin\varphi \\ \sin\varphi & \varepsilon\cos\varphi \end{pmatrix},$$

а формулите (4) се конкретизират като

$$\text{(9)} \qquad \begin{aligned} x &= x'\cos\varphi - \varepsilon y'\sin\varphi + a \\ y &= x'\sin\varphi + \varepsilon y'\cos\varphi + b, \end{aligned}$$

където $\varepsilon = 1$, ако двете координатни системи са еднакво ориентирани и $\varepsilon = -1$ в противоположен случай. Очевидно

$$\det T = \varepsilon, \qquad TT^{t} = T^{t}T = E \text{, т.е.} \qquad T^{-1} = T^{t},$$

следователно $\det T = \det T^{-1} = \varepsilon$. Получихме следния резултат: *две декартови координатни системи в равнината са еднакво ориентирани точно тогава, когато* $\det T = +1$ и противоположно ориентирани точно тогава, когато $\det T = -1$.

2. Смяна на координатите в пространството. Случаят е напълно аналогичен на този в равнината, когато координатните системи са произволни афинни. Нека $O\vec{e}_1\vec{e}_2\vec{e}_3$ и $O^*\vec{e}_1^{\,*}\vec{e}_2^{\,*}\vec{e}_3^{\,*}$ са две афинни координатни системи в пространството и нека спрямо първата от тях началото O^* има координати (a, b, c), а $\vec{e}_1^{\,*}$, $\vec{e}_2^{\,*}$, $\vec{e}_3^{\,*}$ - съответно $(\alpha_{11}, \alpha_{21}, \alpha_{31})$, $(\alpha_{12}, \alpha_{22}, \alpha_{32})$, $(\alpha_{13}, \alpha_{23}, \alpha_{33})$. Ако произволна точка M в пространството има координати (x, y, z) спрямо $O\vec{e}_1\vec{e}_2\vec{e}_3$ и (x', y', z') спрямо $O^*\vec{e}_1^{\,*}\vec{e}_2^{\,*}\vec{e}_3^{\,*}$, то

$$\text{(10)} \qquad \begin{aligned} x &= \alpha_{11}x' + \alpha_{12}y' + \alpha_{13}z' + a \\ y &= \alpha_{21}x' + \alpha_{22}y' + \alpha_{23}z' + b \\ z &= \alpha_{31}x' + \alpha_{32}y' + \alpha_{33}z' + c \end{aligned}$$

Матрицата $T = \left(\alpha_{ij}\right)_1^3$ е неособена, тъй като векторите $\vec{e}_1^{\,*}$, $\vec{e}_2^{\,*}$, $\vec{e}_3^{\,*}$ са линейно независими. Ако втората координатна система е получена от първата чрез *транслация*, т.е. ако $\vec{e}_i^{\,*} = \vec{e}_i$, $i = 1, 2, 3$, то $\alpha_{ii} = 1$, $\alpha_{ij} = 0$ при $i \neq j$, където $i, j = 1, 2, 3$ и формулите (10) добиват вида

$$x = x' + a$$

11) $$y = y' + b$$

$$z = z' + c.$$

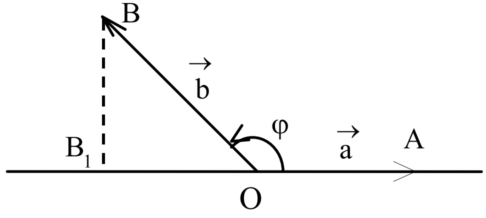
§ 4. Скаларно произведение на свободни вектори. Дължини и разстояния

1. Определение. *Скаларно произведение* на ненулевите вектори $\vec{a}$ и $\vec{b}$ ще наричаме числото (означение: $\vec{a}\,\vec{b}$), равно на произведението от дължините на векторите и косинуса на елементарно-геометричния ъгъл между тях, т.е.

(1) $$\vec{a}\,\vec{b} = |\vec{a}||\vec{b}|\cos(\vec{a}, \vec{b})_e.$$

Ако някой от векторите $\vec{a}$, $\vec{b}$ е нулевият, ще считаме, че $\vec{a}\,\vec{b} = 0$.

2. Нека векторите $\vec{a}$ и $\vec{b}$ не са колинеарни. Да изберем произволна точка O и да построим техни представители с начало O (черт.17), като $\overrightarrow{OA} \in \vec{a}$ и $\overrightarrow{OB} \in \vec{b}$. Проектираме ортогонално (перпендикулярно) точката B върху оста $OA^{\rightarrow}$. Означаваме с φ ъгъла между векторите $\vec{a}$ и $\vec{b}$, а с B_1 - проекцията на точката B. За алгебричната мярка $\overline{OB_1}$ върху оста $OA^{\rightarrow}$ очевидно имаме



черт.17

(2) $$\overline{OB_1} = |\vec{b}| \cos\varphi .$$

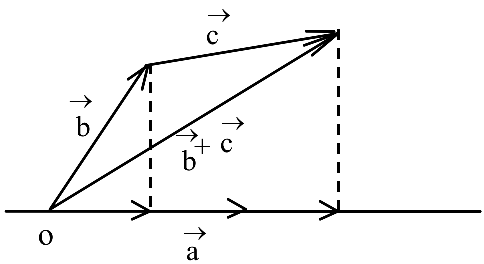
Формулата (2) е в сила и при $\varphi < \dfrac{\pi}{2}$ и при $\varphi = \dfrac{\pi}{2}$. Лесно е да се съобрази, че свободният вектор, определен от насочената отсечка $\overrightarrow{OB_1}$, зависи само от векторите $\vec{a}$ и $\vec{b}$, а не от конкретно избраните техни представители; по тази причина ще го бележим с $np_{\vec{a}}\vec{b}$ и ще го наричаме *ортогонална проекция на вектора* $\vec{b}$ *върху вектора* $\vec{a}$. Впрочем, горната конструкция е валидна дори и векторите да са колинеарни, но ненулеви. Ако $\vec{b}$ е нулев вектор, ще считаме, че ортогоналната проекция е нулевият вектор. При тези уговорки ще пишем

$$\overline{OB_1} = \overline{np_{\vec{a}}\vec{b}} .$$

Като вземем предвид (2), определението (1) може да се запише и като

(3) $$\vec{a}\,\vec{b} = |\vec{a}|\,\overline{np_{\vec{a}}\vec{b}} , \qquad \text{при } \vec{a} \neq \vec{0} ,$$

т.е. скаларното произведение на векторите $\vec{a}$ и $\vec{b}$ *е равно на дължината на* $\vec{a}$, *умножена с алгебричната мярка на проекцията на* $\vec{b}$ *върху оста, определена от* $\vec{a}$.



черт.18

От геометрични съображения (черт.18) лесно следват зависимостите

(4) $$\text{пр}_{\vec{a}}(\vec{b}+\vec{c}) = \text{пр}_{\vec{a}}\vec{b} + \text{пр}_{\vec{a}}\vec{c}, \qquad \overline{\text{пр}_{\vec{a}}(\vec{b}+\vec{c})} = \overline{\text{пр}_{\vec{a}}\vec{b}} + \overline{\text{пр}_{\vec{a}}\vec{c}}.$$

Разбира се, равенствата (4) са валидни и за произволен краен брой събираеми.

3. Свойства на скаларното произведение.

а) $\vec{a}\,\vec{b} = \vec{b}\,\vec{a}$;

б) $\vec{a}(\vec{b}+\vec{c}) = \vec{a}\,\vec{b} + \vec{a}\,\vec{c}$;

в) $(\lambda\vec{a})\vec{b} = \lambda(\vec{a}\,\vec{b})$;

г) $\vec{a}\,\vec{a} = |\vec{a}|^2 \geq 0$, като равенство имаме точно когато $\vec{a} = \vec{0}$.

Тук $\vec{a}$, $\vec{b}$, $\vec{c}$ са произволни вектори, а λ е произволно реално число.

Д о к а з а т е л с т в о . Свойство а) е очевидно, защото елементарно-геометричният ъгъл между $\vec{a}$ и $\vec{b}$ е същият, както между $\vec{b}$ и $\vec{a}$. Свойство б) лесно следва от (3) и (4) при $\vec{a} \neq \vec{0}$:

$$\vec{a}(\vec{b}+\vec{c}) = |\vec{a}|\overline{\text{пр}_{\vec{a}}(\vec{b}+\vec{c})} = |\vec{a}|(\overline{\text{пр}_{\vec{a}}\vec{b}} + \overline{\text{пр}_{\vec{a}}\vec{c}}) = |\vec{a}|\overline{\text{пр}_{\vec{a}}\vec{b}} + |\vec{a}|\overline{\text{пр}_{\vec{a}}\vec{c}} =$$
$$= \vec{a}\,\vec{b} + \vec{a}\,\vec{c}.$$

При $\vec{a} = \vec{0}$, равенството б) е очевидно.

За да докажем в), ще разгледаме два случая. Нека $\lambda > 0$; тогава векторите $\vec{a}$ и $\lambda\vec{a}$ са колинеарни и имат еднакви посоки. Â òàêúâ ñëó÷àé $\varphi = (\vec{a},\vec{b})_e = (\lambda\vec{a},\vec{b})_e$. Tîãàâà

$$(\lambda\vec{a})\vec{b} = |\lambda\vec{a}||\vec{b}|cos\varphi = |\lambda||\vec{a}||\vec{b}|cos\varphi = \lambda|\vec{a}||\vec{b}|cos\varphi = \lambda(\vec{a}\,\vec{b}).$$

Ако $\lambda < 0$, векторите $\vec{a}$ и $\lambda\vec{a}$ са колинеарни и имат противоположни посоки и за ъглите имаме

$$(\vec{a},\vec{b})_e = \varphi, \qquad (\lambda\vec{a},\vec{b})_e = \pi - \varphi.$$

Следователно

$$(\lambda \vec{a})\vec{b} = |\lambda \vec{a}||\vec{b}|\cos(\pi-\varphi) = |\lambda||\vec{a}||\vec{b}|(-\cos\varphi) =$$

$$= -\lambda|\vec{a}||\vec{b}|(-\cos\varphi) = \lambda(\vec{a}\vec{b}).$$

Свойство в) е очевидно в случаите: $\lambda = 0$; $\vec{a} = \vec{0}$; $\vec{b} = \vec{0}$;.

Свойство г) следва веднага от определението, тъй като ъгълът между ненулевите вектори $\vec{a}$ и $\vec{a}$ има мярка нула, а $cos 0 = 1$. За удобство въвеждаме означението $\vec{a}^2 = \vec{a}\vec{a}$ и ще говорим за скаларен квадрат на вектора $\vec{a}$. Ясно е, че

$$|\vec{a}| = \sqrt{\vec{a}\vec{a}} = \sqrt{\vec{a}^2}.$$

Ще отбележим някои непосредствени следствия от доказаните свойства.

1) От а) и в) веднага се получава, че $\vec{a}(\lambda\vec{b}) = \lambda(\vec{a}\vec{b})$.

2) От б) и в) лесно следва обобщението:

$$(5)\qquad \begin{aligned}&(\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \ldots + \lambda_k\vec{a}_k)(\mu_1\vec{b}_1 + \mu_2\vec{b}_2 + \ldots + \mu_l\vec{b}_l) =\\ &= \sum_{i=1}^{k}\sum_{j=1}^{l}\lambda_i\mu_j(\vec{a}_i\vec{b}_j).\end{aligned}$$

3) *Два ненулеви вектора са ортогонални тогава и само тогава, когато скаларното им произведение е равно на нула*. Наистина от определението на скаларно произведение се вижда, че за ненулеви вектори $\vec{a}$, $\vec{b}$ то е равно на нула точно тогава, когато $cos\varphi = 0$, което е еквивалентно на условието ъгълът между векторите да бъде прав. Удобно е да се счита, че *нулевият вектор е ортогонален на всеки вектор*.

4. Скаларно произведение чрез координати. Нека е избрана декартова координатна система $O\vec{e}_1\vec{e}_2\vec{e}_3$. Тъй като базисните вектори имат дължини 1 и са два по два ортогонални, то

$$(6)\qquad \vec{e}_i\vec{e}_i = 1,\ \vec{e}_i\vec{e}_j = 0, \qquad i \neq j, \quad i, j = 1, 2, 3.$$

Да предположим, че спрямо избраната вече координатна система векторите $\vec{a}$ и $\vec{b}$ имат съответно координати (x_1, y_1, z_1) и (x_2, y_2, z_2),

$$\vec{a} = x_1\vec{e_1} + y_1\vec{e_2} + z_1\vec{e_3}, \qquad \vec{b} = x_2\vec{e_1} + y_2\vec{e_2} + z_2\vec{e_3}.$$

Тогава, като приложим (5) и вземем предвид равенствата (6), получаваме

(7) $$\vec{a}\vec{b} = x_1x_2 + y_1y_2 + z_1z_2.$$

Оттук за скаларния квадрат на вектора $\vec{a}$ имаме

$$\vec{a}^2 = \vec{a}\vec{a} = x_1^2 + y_1^2 + z_1^2,$$

а за дължината му получаваме формулата

(8) $$|\vec{a}| = \sqrt{\vec{a}^2} = \sqrt{x_1^2 + y_1^2 + z_1^2}.$$

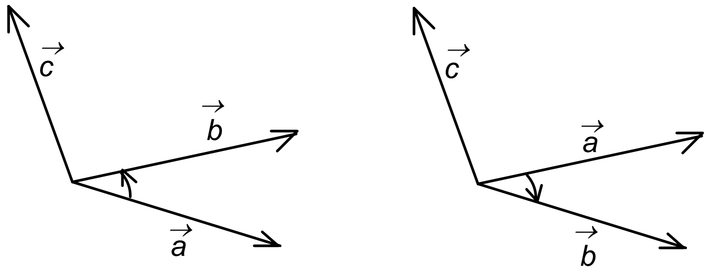
Формулите (7) и (8) се използват много често. Специално подчертаваме, че те са в сила само ако координатната система е декартова.

5. Разстояние между две точки. Нека точките M_1 и M_2 имат декартови координати съответно (x_1, y_1, z_1) и (x_2, y_2, z_2). Разстоянието между точките M_1 и M_2 е равно на дължината на свободния вектор, определен от насочената отсечка $\overrightarrow{M_1M_2}$. Ако го означим с $\mathrm{d}(\mathrm{M}_1, M_2)$, то $\mathrm{d}(\mathrm{M}_1, M_2) = |\overrightarrow{M_1M_2}|$. Насочената отсечка $\overrightarrow{M_1M_2}$ има координати $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$, а съгласно (8) за дължината му получаваме

$$|\overrightarrow{M_1M_2}| = \mathrm{d}(M_1, M_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

§ 5. Векторно произведение на свободни вектори

1. Определение. Нека $\vec{a}$, $\vec{b}$, $\vec{c}$ е наредена тройка *некомпланарни* свободни вектори. За наредената тройка $\vec{a}$, $\vec{b}$, $\vec{c}$ ще казваме, че е *дясна*, ако посоката на въртене на $\vec{a}$ към $\vec{b}$, при която се описва елементарно-геометричният ъгъл между векторите $\vec{a}$ и $\vec{b}$, наблюдавана от края на вектора $\vec{c}$, е обратна на посоката на въртене на часовниковата стрелка. Ако тази посока е по часовниковата стрелка, ще казваме, че тройката е *лява* (черт.19).



черт.19

2. Определение. *Векторно произведение* на векторите $\vec{a}$ и $\vec{b}$ наричаме вектор (ще го бележим с $\vec{a} \wedge \vec{b}$), определен от следните три условия:

а) $|\vec{a} \wedge \vec{b}| = |\vec{a}||\vec{b}| sin\varphi$, където φ е мярката на елементарно-геометричния ъгъл между $\vec{a}$ и $\vec{b}$;

б) векторът $\vec{a} \wedge \vec{b}$ е ортогонален на всеки от векторите $\vec{a}$ и $\vec{b}$;

в) наредената тройка вектори $\vec{a}$, $\vec{b}$, $\vec{a} \wedge \vec{b}$ е дясна.

В това определение се предполага, че вектори $\vec{a}$ и $\vec{b}$ са неколинеарни. Ако вектори $\vec{a}$ и $\vec{b}$ са колинеарни, то $\vec{a} \wedge \vec{b} = \vec{0}$ по определение.

3. Свойства на векторното произведение:

а) $\vec{a} \wedge \vec{b} = -(\vec{b} \wedge \vec{a})$;

б) $(\lambda \vec{a}) \wedge \vec{b} = \lambda(\vec{a} \wedge \vec{b})$; $\quad \vec{a} \wedge \lambda \vec{b} = \lambda(\vec{a} \wedge \vec{b})$

в) $(\vec{a} + \vec{b}) \wedge \vec{c} = \vec{a} \wedge \vec{c} + \vec{b} \wedge \vec{c}$; $\quad \vec{a} \wedge (\vec{b} + \vec{c}) = \vec{a} \wedge \vec{b} + \vec{a} \wedge \vec{c}$.

за произволни свободни вектори $\vec{a}$, $\vec{b}$, $\vec{c}$ и за всяко реално число λ.

Д о к а з а т е л с т в о . а) Ако векторите $\vec{a}$ и $\vec{b}$ са колинеарни, и двете страни са равни на нулевия вектор. Нека те не са колинеарни. Тъй като $\vec{a}\wedge\vec{b}$ и $\vec{b}\wedge\vec{a}$ са ортогонални на $\vec{a}$ и $\vec{b}$, те са колинеарни, а по определението имат равни дължини. Тройките $\vec{a}$, $\vec{b}$, $\vec{a}\wedge\vec{b}$ и $\vec{b}$, $\vec{a}$, $\vec{b}\wedge\vec{a}$ са десни. В такъв случай тройката $\vec{a}$, $\vec{b}$, $\vec{b}\wedge\vec{a}$ е лява. Следователно векторите $\vec{a}\wedge\vec{b}$ и $\vec{b}\wedge\vec{a}$ имат противоположни посоки и получаваме а).

б) Ако векторите $\vec{a}$ и $\vec{b}$ са колинеарни или пък $\lambda=0$, равенствата са очевидни. Нека $\vec{a}$ и $\vec{b}$ не са колинеарни и $\lambda\neq 0$. Ще докажем само първото равенство, защото второто следва от него с прилагане на а). В зависимост от знака на λ за ъглите имаме $\varphi=(\vec{a},\vec{b})_e=(\lambda\vec{a},\vec{b})$ при $\lambda>0$ и $(\lambda\vec{a},\vec{b})=\pi-\varphi$ при $\lambda<0$. Дължините на векторите $(\lambda\vec{a})\wedge\vec{b}$ и $\lambda(\vec{a}\wedge\vec{b})$ са равни, защото

$$|(\lambda\vec{a})\wedge\vec{b}|=|\lambda\vec{a}||\vec{b}|sin\varphi=|\lambda||\vec{a}||\vec{b}|sin\varphi=|\lambda(\vec{a}\wedge\vec{b})|,$$

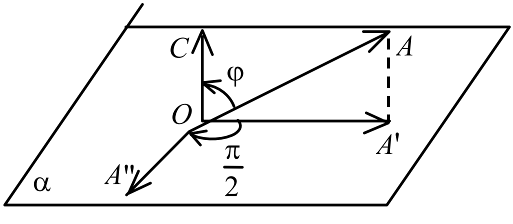
ако $\lambda>0$ и

$$|(\lambda\vec{a})\wedge\vec{b}|=|\lambda\vec{a}||\vec{b}|\sin(\pi-\varphi)=|\lambda||\vec{a}||\vec{b}|\sin\varphi=|\lambda(\vec{a}\wedge\vec{b})|,$$

ако $\lambda<0$. Те са колинеарни, защото всеки от тях е ортогонален на векторите $\vec{a}$ и $\vec{b}$. Ако $\lambda>0$, посоките на $(\lambda\vec{a})\wedge\vec{b}$ и $\lambda(\vec{a}\wedge\vec{b})$ съвпадат с посоката на $\vec{a}\wedge\vec{b}$. Ако $\lambda<0$, векторите $(\lambda\vec{a})\wedge\vec{b}$ и $\lambda(\vec{a}\wedge\vec{b})$ имат еднакви посоки, защото всеки от тях има посока, противоположна на посоката на $\vec{a}\wedge\vec{b}$. С това исканото равенство е доказано.

в) Ще докажем само равенството

(1) $$(\vec{a}+\vec{b})\wedge\vec{c}=\vec{a}\wedge\vec{c}+\vec{b}\wedge\vec{c}.$$



черт.20

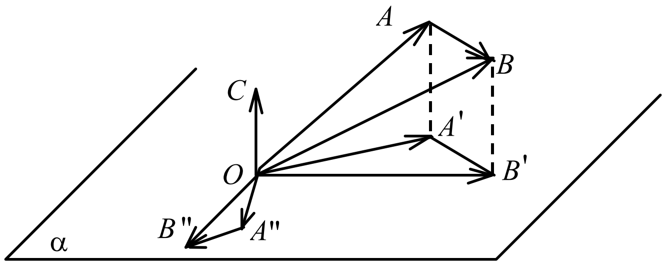
Второто равенство в) следва веднага от него с прилагане на а). Разбира се, (1) е очевидно, ако някой от векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ е нулевият. По-нататък ще предполагаме, че те не са нулеви. Най-напред ще отбележим, че ако векторът $\vec{c}$ има дължина единица, $|\vec{c}|=1$, векторното произведение $\vec{a}\wedge\vec{c}$ има проста геометрична конструкция (черт.20). Нека $\vec{OA}\in\vec{a}$, $\vec{OC}\in\vec{c}$ и елементарно-геометричният ъгъл между $\vec{a}$ и $\vec{c}$ е φ. През точката O построяваме равнина α, перпендикулярна на $\vec{OC}$. Да проектираме ортогонално A върху равнината α и нека A' е проекцията на точката A. Гледано от точка C, да завъртим $\vec{OA'}$ в равнината α около точката O на ъгъл $\dfrac{\pi}{2}$ по посока на часовниковата стрелка. Нека A'' е образът на точката A' при това въртене. Ясно е, че $|\vec{OA''}|=|\vec{OA'}|$ и освен това

$$|\vec{OA''}|=|\vec{OA'}|=|\vec{OA}|cos(\frac{\pi}{2}-\varphi)=|\vec{a}||\vec{c}|sin\varphi.$$

Векторите, чиито представители са съответно $\vec{OA}$, $\vec{OC}$ и $\vec{OA''}$, образуват дясна тройка и $\vec{OA''}$ е ортогонален на $\vec{OC}$ и на $\vec{OA}$, защото е перпендикулярен на равнината, определена от точките O, A, A'. Следователно

$$\vec{OA''}\in\vec{a}\wedge\vec{c}.$$

Сега ще докажем (1), като най-напред ще предполагаме, че $|\vec{c}|=1$. Построението е направено на черт.21:



черт.21

На него $\vec{OC} \in \vec{c}$, $\vec{OA} \in \vec{a}$, $\vec{AB} \in \vec{b}$, $\vec{OB} \in \vec{a}+\vec{b}$. Изпълнено е

$$\vec{OB'} = \vec{OA'} + \vec{A'B'}, \qquad \vec{OB''} = \vec{OA''} + \vec{A''B''}.$$

Както по-горе, след завъртането получаваме

$$\vec{OB''} \in (\vec{a}+\vec{b}) \wedge \vec{c}, \qquad \vec{OA''} \in \vec{a} \wedge \vec{c}, \qquad \vec{A''B''} \in \vec{b} \wedge \vec{c}.$$

Следователно получаваме (1). Общият случай се получава лесно от частния с прилагане на б). Нека $|\vec{c}| = \lambda$; тъй като векторът $\lambda^{-1}\vec{c}$ има дължина 1, то

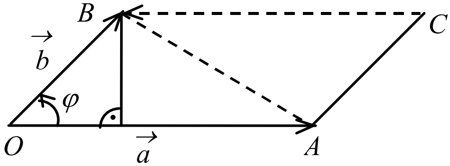
$$(\vec{a}+\vec{b}) \wedge \vec{c} = \lambda\lambda^{-1}[(\vec{a}+\vec{b}) \wedge \vec{c}] = \lambda[(\vec{a}+\vec{b}) \wedge (\lambda^{-1}\vec{c})] =$$

$$= \lambda[\vec{a} \wedge (\lambda^{-1}\vec{c}) + \vec{b} \wedge (\lambda^{-1}\vec{c})] = \vec{a} \wedge \vec{c} + \vec{b} \wedge \vec{c}.$$

Доказаните свойства дават възможност да умножаваме векторно линейни комбинации на свободни вектори:

$$(\lambda_1 \vec{a_1} + \lambda_2 \vec{a_2} + \ldots + \lambda_k \vec{a_k}) \wedge (\mu_1 \vec{b_1} + \mu_2 \vec{b_2} + \ldots + \mu_l \vec{b_l}) =$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{l} \lambda_i \mu_j (\vec{a_i} \wedge \vec{b_j}).$$



черт.22

4. Твърдение. *Ако векторите $\vec{a}$ и $\vec{b}$ не са колинеарни, то дължината $|\vec{a}\wedge\vec{b}|$ на векторното им произведение е равна на лицето на успоредника, построен върху тези вектори.*

Д о к а з а т е л с т в о. Достатъчно е да се погледне черт.22. За височината h на успоредника $OACB$ имаме

$$h = |\vec{b}|\sin\varphi.$$

Следователно за лицето му S получаваме

$$S = |\vec{a}|h = |\vec{a}||\vec{b}|\sin\varphi = |\vec{a}\wedge\vec{b}|$$

Полезно е да отбележим за приложенията, че за лицето S_Δ на триъгълника $\vec{a}$, построен върху векторите $\vec{a}$ *и* $\vec{b}$, очевидно имаме

$$S_\Delta = \frac{1}{2}|\vec{a}\wedge\vec{b}|.$$

5. Векторното произведение чрез координати. Нека е дадена ортонормирана (декартова) координатна система $O\vec{e}_1\vec{e}_2\vec{e}_3$. Допълнително ще предположим, че наредената тройка вектори $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ е *дясна*. Ще казваме кратко, че системата е *дясна декартова*. С помощта на определението на векторно произведение и на неговите свойства лесно се съобразява, че

$$\vec{e}_i\wedge\vec{e}_i = \vec{0}, \qquad i = 1, 2, 3;$$
$$\vec{e}_1\wedge\vec{e}_2 = \vec{e}_3, \qquad \vec{e}_2\wedge\vec{e}_3 = \vec{e}_1, \qquad \vec{e}_3\wedge\vec{e}_1 = \vec{e}_2,$$
$$\vec{e}_2\wedge\vec{e}_1 = -\vec{e}_3, \qquad \vec{e}_3\wedge\vec{e}_2 = -\vec{e}_1, \qquad \vec{e}_1\wedge\vec{e}_3 = -\vec{e}_2.$$

Нека векторите $\vec{a}$ *и* $\vec{b}$ имат координати съответно (x_1, y_1, z_1) и (x_2, y_2, z_2), т.е.

$$\vec{a} = x_1\vec{e_1} + y_1\vec{e_2} + z_1\vec{e_3}, \qquad \vec{b} = x_2\vec{e_1} + y_2\vec{e_2} + z_2\vec{e_3}$$

Като се използват тъждеството от края на т.3 и равенствата от тази точка, с непосрведствено пресмятане, което оставяме на читателя, се вижда, че

$$\vec{a}\wedge\vec{b} = \left(y_1z_2 - y_2z_1\right)\vec{e_1} + \left(z_1x_2 - z_2x_1\right)\vec{e_2} + \left(x_1y_2 - x_2y_1\right)\vec{e_3}.$$

По определение коефициентът пред $\vec{e_i}$ е i -тата координата на $\vec{a}\wedge\vec{b}$. Тъй като зависимостта между координатите на векторите и координатите на тяхното векторно произведение се налага да се запомни, ще препишем (2) във вида

$$\vec{a}\wedge\vec{b} = \begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}\vec{e_1} - \begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}\vec{e_2} + \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}\vec{e_3}. \tag{3}$$

§ 6. Смесено произведение на три свободни вектора

1. Определение. *Смесено произведение* $\vec{a}\,\vec{b}\,\vec{c}$ на векторите $\vec{a}$, $\vec{b}$, $\vec{c}$, взети в този ред, наричаме скаларното произведение на векторите $\vec{a}\wedge\vec{b}$ и $\vec{c}$, т.е.

$$\vec{a}\,\vec{b}\,\vec{c} = (\vec{a}\wedge\vec{b})\,\vec{c}.$$

2. Смесеното произведение $\vec{a}\,\vec{b}\,\vec{c}$ е равно на нула тогава и само тогава, когато векторите $\vec{a}$, $\vec{b}$, и $\vec{c}$ са компланарни.

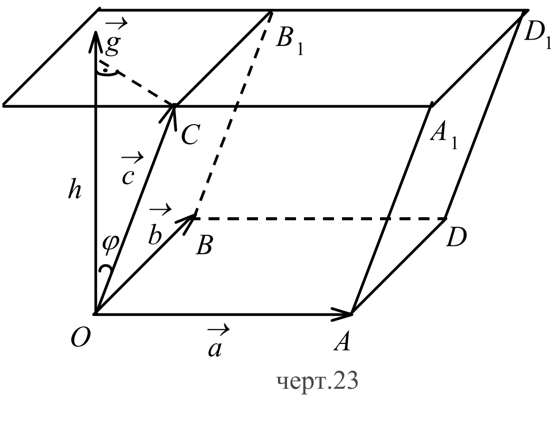
Доказателство. Нека $\vec{a}$, $\vec{b}$, и $\vec{c}$ са компланарни с равнината α. Векторът $\vec{a}\wedge\vec{b}$ е перпендикулярен на α (или е нулевият вектор). Това следва от определението за векторно произведение. Следователно

$$\vec{a}\,\vec{b}\,\vec{c} = (\vec{a}\wedge\vec{b})\,\vec{c} = 0$$

(следствие 3, т. 3, § 4, гл. 5).

Обратно, нека $\vec{a}\,\vec{b}\,\vec{c}=\vec{0}$. Ако $\vec{a}\wedge\vec{b}=\vec{0}$, то $\vec{a}$ и $\vec{b}$ са колинеарни и следователно $\vec{a}$, $\vec{b}$, и $\vec{c}$ са компланарни. Ако $\vec{a}\wedge\vec{b}\neq\vec{0}$, то векторът $\vec{a}\wedge\vec{b}$ е перпендикулярен на $\vec{c}$. Тъй като $\vec{a}\wedge\vec{b}$ е перпендикулярен на $\vec{a}$ и на $\vec{b}$ (по определение), то $\vec{c}$ е успореден на равнина, определена от представителите на $\vec{a}$ и $\vec{b}$. Следователно векторите $\vec{a}$, $\vec{b}$ и $\vec{c}$ са компланарни.

3. Нека векторите $\vec{a}$, $\vec{b}$ и $\vec{c}$ не са компланарни, образуват дясна тройка и нека $\vec{g}=\vec{a}\wedge\vec{b}$. Да построим паралелепипед върху дадените вектори, т.е. паралелепипед, на който три ръба през общ връх да бъдат представители съответно на $\vec{a}$, $\vec{b}$ и $\vec{c}$ (черт. 23). От § 5, т. 4 знаем, че $|\vec{g}|=S$ е лицето на успоредника, построен върху представителите на векторите $\vec{a}$ и $\vec{b}$, т.е. лицето на основата на паралелепипеда. Векторът $\vec{g}$ е ортогонален на равнината α, определена от представителите на $\vec{a}$ и $\vec{b}$, а тройката $\vec{a}$, $\vec{b}$, $\vec{g}$ е дясна. Тъй като тройката $\vec{a}$, $\vec{b}$, $\vec{c}$ е дясна, то векторите $\vec{g}$ и $\vec{c}$ са разположени от една и съща страна спрямо равнината α. Тъй като векторът $\vec{g}$ е ортогонален на α, ъгълът φ между $\vec{g}$



черт.23

и $\vec{c}$ е остър и $\cos\varphi > 0$. За дължината на височината на паралелепипеда получаваме

$$h = |\vec{c}|\cos\varphi.$$

Следователно за обема му V имаме

$$V = Sh = |\vec{a} \wedge \vec{b}||\vec{c}|\cos\varphi = (\vec{a} \wedge \vec{b})\,\vec{c} = \vec{a}\,\vec{b}\,\vec{c}.$$

Получихме, че *ако векторите* $\vec{a}$, $\vec{b}$, $\vec{c}$ *не са компланарни и образуват дясна тройка, то смесеното им произведение е равно на обема на паралелепипеда, построен върху тях*.

Нека сега векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ отново са некомпланарни, но образуват лява тройка. В този случай $\vec{g}$ и $\vec{c}$ са от различни страни на равнината α и тъй като $\vec{g}$ е ортогонален на α, то ъгълът φ между $\vec{g}$ и $\vec{c}$ е тъп, т.е. $\cos\varphi < 0$. За височината h на паралелепипеда имаме

$$h = -|\vec{c}|\cos\varphi,$$

а за обема му V получаваме

$$V = Sh = |\vec{a} \wedge \vec{b}|(-|\vec{c}|\cos\varphi) = -\vec{a}\,\vec{b}\,\vec{c}.$$

3. Нека векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ са некомпланарни и да построим паралелепипед върху тях (виж т.2). Обемът на този паралелепипед, взет със знак "+", ако $\vec{a}$, $\vec{b}$, $\vec{c}$ образуват дясна тройка, и със знак "-", ако $\vec{a}$, $\vec{b}$, $\vec{c}$ образуват лява тройка, наричаме *ориентиран обем* на паралелепипеда, построен върху векторите $\vec{a}$, $\vec{b}$, $\vec{c}$.

В т.2 доказахме, че ако векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ са некомпланарни, то смесеното им произведение е равно на ориентирания обем на паралелепипеда, построен върху тези вектори.

4. Свойства на смесеното произведение. В сила са:

а) $(\vec{a'} + \vec{a''})\,\vec{b}\,\vec{c} = \vec{a'}\,\vec{b}\,\vec{c} + \vec{a''}\,\vec{b}\,\vec{c}$,

$$\vec{a}(\vec{b'} + \vec{b''})\,\vec{c} = \vec{a}\,\vec{b'}\,\vec{c} + \vec{a}\,\vec{b''}\,\vec{c},$$

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

б) $(\lambda\,\vec{a})\,\vec{b}\,\vec{c} = \vec{a}(\lambda\,\vec{b})\,\vec{c} = \vec{a}\,\vec{b}(\lambda\,\vec{c}) = \lambda(\vec{a}\,\vec{b}\,\vec{c})$ *за всяко реално число* λ;

в) *ако в смесеното произведение разменим местата на два вектора, то сменя знака си, т.е.*

$$\vec{a}\,\vec{b}\,\vec{c} = -(\vec{b}\,\vec{a}\,\vec{c}) = -(\vec{a}\,\vec{c}\,\vec{b}) = -(\vec{c}\,\vec{b}\,\vec{a}) = \vec{c}\,\vec{a}\,\vec{b} = \vec{b}\,\vec{c}\,\vec{a}.$$

По-общо в сила е

г) *Ако* $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ *са произволни свободни вектори, а* (i,j,k) *е пермутация на числата* 1, 2, 3*, то*

$$\vec{e}_i\,\vec{e}_j\,\vec{e}_k = (-1)^{[i,j,k]}\,\vec{e}_1\,\vec{e}_2\,\vec{e}_3 ,$$

където $[i,j,k]$ е *броят на инверсиите в пермутацията* (виж гл. 1, § 5).

Доказателство. Твърденията а) и б) следват лесно от съответните свойства на векторното и на скаларното произведение. За илюстрация ще докажем само по едно от исканите равенства:

$$(\vec{a}'+\vec{a}'')\,\vec{b}\,\vec{c} = [(\vec{a}'+\vec{a}'')\wedge\vec{b}]\,\vec{c} = [\vec{a}'\wedge\vec{b}+\vec{a}''\wedge\vec{b}]\,\vec{c} =$$

$$= (\vec{a}'\wedge\vec{b})\,\vec{c} + (\vec{a}''\wedge\vec{b})\,\vec{c} = \vec{a}'\,\vec{b}\,\vec{c} + \vec{a}''\,\vec{b}\,\vec{c} ;$$

$$(\lambda\,\vec{a})\,\vec{b}\,\vec{c} = [(\lambda\,\vec{a})\wedge\vec{b}]\,\vec{c} = [\lambda(\vec{a}\wedge\vec{b})]\,\vec{c} = \lambda[(\vec{a}\wedge\vec{b})\,\vec{c}] =$$

$$= \lambda(\vec{a}\,\vec{b}\,\vec{c}).$$

в) Ако векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ са компланарни, всички разглеждани смесени произведения са равни на нула и няма какво повече да се доказва. Нека те не са компланарни. Както видяхме в т.2, смесеното произведение на три вектора, независимо в какъв ред ще ги вземем, с точност до знак съвпада с обема V на паралелепипеда, построен върху тях, т.е. всяко от шестте смесени произведения, които могат да се образуват от три вектора, е равно на V или на $-V$. Не е трудно да се съобрази, че ако в една наредена тройка некомпланарни вектори разменим местата на два от тях, то тройката сменя ориентацията си, т.е. ако изходната е била дясна (лява), то новата ще бъде лява (дясна). Следователно, ако $\vec{a}\,\vec{b}\,\vec{c} = V$, то $\vec{a} \;\wedge\; \vec{b}$ и т.н.

г) Всъщност става дума за компактно записване на равенствата от в). Достатъчно е да се изпишат шестте пермутации на числата 1, 2, 3 и да се приложи в) при $\vec{a} = \vec{e}_1$, $\vec{b} = \vec{e}_2$, $\vec{c} = \vec{e}_3$.

Ако интерпретираме смесеното произведение като функция f на три векторни променливи: $f(\vec{a}, \vec{b}, \vec{c}) = \vec{a}\,\vec{b}\,\vec{c}$, то свойствата а) и б) изразяват, че тя е линейна функция на всяка от трите променливи, щом другите две са фиксирани. За да бъдем по-ясни, нека например фиксираме $\vec{b}$ и $\vec{c}$ и да меним $\vec{a}$; получаваме функция $f_1(\vec{a}) = \vec{a}\,\vec{b}\,\vec{c}$ на една променлива и за нея, съгласно а) и б), имаме

$$f_1(\vec{a'} + \vec{a''}) = f_1(\vec{a'}) + f_1(\vec{a''}), \qquad f_1(\lambda\,\vec{a}) = \lambda f_1(\vec{a}).$$

5. Изразяване на смесеното произведение чрез координатите на векторите. Нека спрямо произволна афинна координатна система $O\vec{e}_1\vec{e}_2\vec{e}_3$ имаме

$$\vec{a} = a_{11}\,\vec{e}_1 + a_{12}\,\vec{e}_2 + a_{13}\,\vec{e}_3, \qquad \vec{b} = a_{21}\,\vec{e}_1 + a_{22}\,\vec{e}_2 + a_{23}\,\vec{e}_3$$

$$\vec{c} = a_{31}\,\vec{e}_1 + a_{32}\,\vec{e}_2 + a_{33}\,\vec{e}_3$$

(Защо за координатите на векторите избрахме нетрадиционно означаване с a_{ij}, ще стане ясно след малко.) Тогава от свойствата а) и б) в предишната точка имаме:

$$\vec{a}\,\vec{b}\,\vec{c} = \left(\sum_{i=1}^{3} a_{1i}\,\vec{e}_i\right)\left(\sum_{j=1}^{3} a_{2j}\,\vec{e}_j\right)\left(\sum_{k=1}^{3} a_{3k}\,\vec{e}_k\right) = \sum_{i,j,k=1}^{3} a_{1i}a_{2j}a_{3k}\,\vec{e}_i\,\vec{e}_j\,\vec{e}_k$$

Ако между числата i, j, k има равни, векторите $\vec{e}_i, \vec{e}_j, \vec{e}_k$ са компланарни и $\vec{e}_i\,\vec{e}_j\,\vec{e}_k = 0$. Следователно, достатъчно е да сумираме само по всички *пермутации* (i, j, k) на числата 1, 2, 3. Тази уговорка за сумирането и г) от т. 3 ни дават:

(1) $$\vec{a}\,\vec{b}\,\vec{c} = \Big(\sum_{(i,j,k)}^{3} (-1)^{[i,j,k]} a_{1i}a_{2j}a_{3k}\Big)\vec{e_i}\,\vec{e_j}\,\vec{e_k}$$

или

(2) $$\vec{a}\,\vec{b}\,\vec{c} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \vec{e_i}\,\vec{e_j}\,\vec{e_k}\,.$$

Да предположим, че координатната система е *дясна декартова*. Тогава $\vec{e_1} \wedge \vec{e_2} = \vec{e_3}$ и $\vec{e_1}\,\vec{e_2}\,\vec{e_3} = 1$. Следователно (2) ще добие вида

(3) $$\vec{a}\,\vec{b}\,\vec{c} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Следва да се запомни, че в първия ред на детерминантата стоят координатите на първия от трите вектора, във втория ред - координатите на втория вектор и в третия ред - координатите на третия вектор.

6. Коментар. Бихме могли съществено да съкратим изложението с помощта на следната схема. Веднага след определението на смесено произведение да въведем за векторите $\vec{a}$, $\vec{b}$, $\vec{c}$ координати (x_i, y_i, z_i), съответно $i = 1, 2, 3$, спрямо някаква дясна декартова координатна система. От § 5, т. 5 знаем, че в този случай векторното произведение $\vec{a} \wedge \vec{b}$ има координати

$$\left(\begin{vmatrix} y_1 & z_1 \\ y_2 & z_2 \end{vmatrix}, -\begin{vmatrix} x_1 & z_1 \\ x_2 & z_2 \end{vmatrix}, \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \right).$$

Тогава с препращане към формулата (7) за скаларното произведение от § 4, т. 4 и с позоваване на формулата за развиване на детерминантата по адюнгираните количества (в случая на третия ред) щяхме да получим равенството

$$(4) \qquad \vec{a}\,\vec{b}\,\vec{c} = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

(всъщност (3) в други означения). Веднага след това бихме могли да декларираме, че свойствата а) - в) от т.3 са очевидни, защото съгласно (4) те са друга редакция на вече познатите ни свойства на детерминантите. Причината да предпочетем по-дългото чисто геометрично изложение е в изкушението да покажем, че до понятието детерминанта от трети ред водят и потребности на геометрията. Ако читателят прегледа внимателно още веднъж цялата глава, мислено зачертае изреченията, в които участва думата "детерминанта", с изненада ще открие, че почти нищо съдържателно не е изчезнало. По-сложно ще стане обаче, когато трябва да се възприеме "грапавата" формула (1). Ако все още не знаехме що е детерминанта, то (1) би ни принудило да я въведем поне в разглеждания частен случай (детерминанта от трети ред), а твърденията а) - в) от т. 3 ни поднасят наготово нейните основни свойства.

УПРАЖНЕНИЯ

1. (*Релация на Шал*) Докажете, че за всеки три точки A, B, C, лежащи върху ос, е изпълнено

$$\overline{AB} + \overline{BC} + \overline{CA} = 0,$$

където с $\overline{AB}$ е означена алгебричната мярка на насочената отсечка $\vec{AB}$.

2. Нека $\vec{a}$, $\vec{b}$, $\vec{c}$ са три произволни свободни вектора. Произведението $(\vec{a} \wedge \vec{b}) \wedge \vec{c}$ се нарича *двойно векторно произведение*. Разглежда се и двойното векторно произведение $\vec{a} \wedge (\vec{b} \wedge \vec{c})$. Докажете равенствата

$$(\vec{a} \wedge \vec{b}) \wedge \vec{c} = (\vec{a}\,\vec{c}).\vec{b} - (\vec{b}\,\vec{c}).\vec{a},$$

$$\vec{a} \wedge (\vec{b} \wedge \vec{c}) = (\vec{a}\,\vec{c}).\vec{b} - (\vec{a}\,\vec{b}).\vec{c}.$$

3. Докажете *тъждеството на Якоби*:

$$(\vec{a}\wedge\vec{b})\wedge\vec{c}+(\vec{b}\wedge\vec{c})\wedge\vec{a}+(\vec{c}\wedge\vec{a})\wedge\vec{b}=\vec{0},$$

където $\vec{a}$, $\vec{b}$, $\vec{c}$ са произволни свободни вектори.

4. Докажете, че за всеки четири вектора $\vec{a}$, $\vec{b}$, $\vec{c}$ и $\vec{d}$ е изпълнено:

$$(\vec{a}\wedge\vec{b})(\vec{c}\wedge\vec{d})=\begin{vmatrix}\vec{a}\vec{c} & \vec{a}\vec{d}\\ \vec{b}\vec{c} & \vec{b}\vec{d}\end{vmatrix}.$$

5. Нека $O\vec{e_1}\vec{e_2}\vec{e_3}$ и $O\vec{e_1^*}\vec{e_2^*}\vec{e_3^*}$ са еднакво (противоположно) ориентирани афинни координатни системи в пространството. Докажете, че матрицата на прехода от базиса $\vec{e_1}$, $\vec{e_2}$, $\vec{e_3}$ към базиса $\vec{e_1^*}$, $\vec{e_2^*}$, $\vec{e_3^*}$ има положителна (отрицателна) детерминанта. Докажете аналогичното твърдение в равнината.

Това упражнение показва, че понятието *еднакво* (*противоположно*) *ориентирани* координатни системи може да се въведе чисто аналитично. А именно: ще казваме, че афинните координатни системи $O\vec{e_1}\vec{e_2}\vec{e_3}$ и $O\vec{e_1^*}\vec{e_2^*}\vec{e_3^*}$ са еднакво (противоположно) ориентирани тогава и само тогава, когато детерминантата на матрицата на прехода от единия към другия базис е положително (отрицателно) число. Същото определение може да се въведе и за афинни координатни системи в равнината.

Г л а в а 6

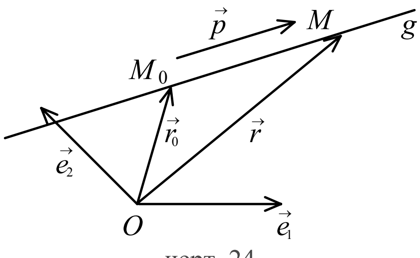
ПРАВИ И РАВНИНИ

С помощта на координатния метод продължаваме да алгебризираме класическата евклидова геометрия. На основните геометрични фигури - правите и равнините - се съпоставят алгебрични обекти - уравнения от първа степен. Показва се как с помощта на уравненията може да се изразяват някои от най-важните геометрични понятия и величини, като взаимни разполагания, ъгли и разстояния между основните геометрични фигури. В повечето от случаите координатната система е произволна афинна. При разглеждане на метрични задачи[*], свързани с ъгли и разстояния, тя обикновено се предполага декартова.

За краткост и удобство често вместо изрази като "точка M с координати (x, y)" ще пишем кратко "точка $M(x, y)$", а вместо "вектор $\vec{p}$ с координати (a,b)" ще използваме и "вектор $\vec{p}(a,b)$". Понякога ще си позволяваме и волността с един и същ символ да означаваме както дадена насочена отсечка, така и свободния вектор, който тя определя. Точният смисъл, който влагаме, ще бъде ясен от контекста (вж. например равенство (1) по-долу).

§ 1. Уравнения на права в равнината

1. Параметрични уравнения на права в равнина. Нека в равнина ε е дадена права g. Очевидно е, че правата се определя еднозначно със задаването на точка M_0, лежаща върху g, и ненулев вектор $\vec{p}$, колинеарен на правата (черт. 24). Произволна точка M от ε ще лежи върху правата g точно тогава, когато насочената отсечка $\overrightarrow{MM_0}$ е колинеарна на $\vec{p}$ или, еквивалентно, точно когато



черт. 24

---

[*] Задачите, свързани с измерване на разстояния, измерване на ъгли, определяне на лица и обеми обикновено, се наричат метрични задачи.

(1) $$\overrightarrow{MM}_0 = \lambda\vec{p}$$

за някакво реално число λ. Да изберем в ε произволна афинна координатна система $O\vec{e}_1\vec{e}_2$ и да въведем координати за елементите, определящи правата. Именно, нека $M_0\left(x_0, y_0\right)$, $\vec{p}\left(a,b\right)$ и $M\left(x,y\right)$, а $\vec{r}_0$ и $\vec{r}$ са съответно радиус-векторите на точките M_0 и M. В тези означения (1) се записва като $\vec{r} - \vec{r}_0 = \lambda\vec{p}$ или

(2) $$\vec{r} = \vec{r}_0 + \lambda\vec{p}.$$

Смисълът на това равенство е, че *точка M с радиус-вектор $\vec{r}$ лежи върху правата g точно тогава, когато е изпълнено съотношението* (2) *за някакво реално число λ*. Уравнението (2) се нарича *векторно параметрично уравнение на правата g*. В еквивалентна форма то може да бъде преписано като двойка равенства за съответните координати:

(3) $$g\colon \begin{aligned} x &= x_0 + \lambda a \\ y &= y_0 + \lambda b. \end{aligned}$$

Равенствата (3) се наричат *скаларни параметрични уравнения на правата g*.

Ще подчертаем, че тъй като векторът $\vec{p}$ е ненулев, числата a, b не са едновременно нули. Ясно е и обратното: ако изберем произволни реални числа x_0, y_0, a, b, за които $|a| + |b| \neq 0$, множеството от точките с координати $\left(x, y\right)$, определени от (3), където λ пробягва реалните числа, е еднозначно определена права g, минаваща през точката $M_0\left(x_0, y_0\right)$ и колинеарна на вектора $\vec{p}\left(a,b\right)$.

С тези елементарни разсъждения на геометричния обект права съпоставихме алгебричния обект уравнение (2) (уравнения (3)) и, обратно: на посочените алгебрични обекти съответства геометричният обект права.

2. Общо уравнение на права в равнината. Скаларните параметрични уравнения (3) означават, че двата стълба в матрицата

$$\begin{pmatrix} x - x_0 & a \\ y - y_0 & b \end{pmatrix}$$

са линейно зависими, което е еквивалентно на условието минорът от втори ред (детерминантата й) да е равен на нула, т.е.

$$b\left(x-x_0\right)-a\left(y-y_0\right)=0.$$

След полагането $A=b$, $B=-a$, $C=-bx_0+ay_0$, последното уравнение се преписва като

$$Ax+By+C=0. \tag{4}$$

Тъй като (4) е еквивалентно на (3), то се нарича *общо уравнение на правата* g. Ще подчертаем, че числата A, B не са едновременно нули и че векторът $\vec{p}\left(-B,A\right)$ е колинеарен на правата g с уравнение (4).

3. Теорема. *Нека в равнината* ε *е избрана афинна координатна система* $O\vec{e}_1\vec{e}_2$. *Тогава за всяка права* g от ε *съществува уравнение от вида* (4), $|A|+|B|\neq 0$ (*общо уравнение на правата*), *в смисъл, че точка* $M\left(x,y\right)$ *от* ε *лежи върху правата точно тогава, когато координатите й удовлетворяват* (4). *Обратно: на всяко уравнение* (4), *в което* $|A|+|B|\neq 0$, *съответства единствена права* g, *за която* (4) *е нейно общо уравнение*.

Доказателство. В едната посока твърдението вече е доказано. Нека сега е дадено уравнение (4). Ще възстановим правата. Тъй като A, B не са едновременно нули, поне едно решение на (4) съществува. Нека $\left(x_0,y_0\right)$ е едно конкретно решение, а $\left(x,y\right)$ - произволно, т.е.

$$Ax_0+By_0+C=0,$$

$$Ax+By+C=0.$$

Следователно

$$A\left(x-x_0\right)+B\left(y-y_0\right)=0.$$

Тъй като A,B не са едновременно нули, всички решения на последното уравнение се дават с формулите

$$x-x_0=-\lambda B, \qquad y-y_0=\lambda A,$$

където λ пробягва реалните числа. Но те са всъщност скаларни параметрични уравнения на единствената права g, минаваща през точката $M_0\left(x_0, y_0\right)$ и колинеарна на вектора $\vec{p}\left(-B, A\right)$. Ако бяхме избрали друго конкретно решение $\left(x_1, y_1\right)$ на (4), щяхме да получим права, която минава както през $M_1\left(x_1, y_1\right)$, така и през $M_0\left(x_0, y_0\right)$ и е колинеарна на вектора $\vec{p}$, т.е. същата права g.

Очевидно е, че ако умножим (4) с произволно ненулево число, ще получим ново уравнение, което задава същата права. Малко по-късно ще покажем и обратното: ако две уравнения от вида (4) задават една и съща права, то те са пропорционални.

4. Твърдение. *Векторът* $\vec{q}\left(\alpha, \beta\right)$ *е колинеарен на правата* g*, зададена с уравнение* (4)*, тогава и само тогава, когато*

$$A\alpha + B\beta = 0. \tag{5}$$

Ако координатната система е декартова, векторът $\vec{n}\left(A, B\right)$ *е ортогонален на правата* g.

Д о к а з а т е л с т в о . Тъй като A и B не са едновременно нули, (5) е еквивалентно на

$$\alpha = -\lambda B, \qquad \beta = \lambda A, \qquad \left(\lambda \in \mathrm{R}\right),$$

което означава, че векторът $\vec{q}\left(\alpha, \beta\right)$ е колинеарен на вектора $\vec{p}\left(-B, A\right)$, за който вече видяхме, че е колинеарен на правата g.

Ако координатната система е декартова, за скаларното произведение $\vec{p}\vec{n}$ имаме

$$\vec{p}\vec{n} = \left(-B\right)A + AB = 0,$$

т.е. $\vec{n}$ е ортогонален на $\vec{p}$, а следователно и на правата g.

Ще отбележим, че при декартова координатна система общото уравнение (4) може да се запише и в термините на скаларни произведения на радиус-вектори. Наистина, нека $M_0\left(x_0, y_0\right)$ е някаква точка от правата g, $\vec{r}_0$ е нейният радиус-вектор, а $M\left(x, y\right)$ е произволна точка с радиус-вектор $\vec{r}$. Имаме

$$C = -Ax_0 - By_0 = -\vec{n}\vec{r}_0, \qquad Ax + By = \vec{n}\vec{r}$$

и уравнението (4) се записва като

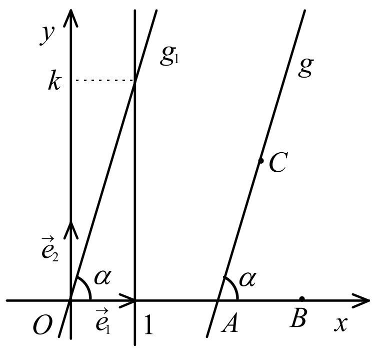
(6) $$\vec{n}\vec{r} = \vec{n}\vec{r}_0.$$

Следователно правата g може да се зададе и като множеството на точките от равнината с радиус-вектори r, удовлетворяващи (6), където $\vec{n}$ е ненулев вектор, ортогонален на g, а r_0 е радиус-вектор на произволна, но фиксирана точка от правата.

Полезно е да се отбележи най-сетне, че ако в общото уравнение (4) коефициентът A е равен на нула, то правата е успоредна на оста $O\vec{e}_1$. Ако пък $B = 0$, правата е успоредна на оста $O\vec{e}_2$. Това следва веднага от твърдението в т. 4, но и само по себе си е очевидно: ако например уравнението е $By + C = 0$, ясно е, че правата се състои от всички точки с координати $(x, -\frac{C}{B})$, където x е произволно. Очевидно е, че тези точки лежат на права, успоредна на $O\vec{e}_1$.

5. Декартово уравнение на права. Да предположим, че координатната система е декартова и g е права с общо уравнение (4), в което $B \neq 0$. Геометричният смисъл на това условие е, че правата не е успоредна на ординатната ос. Ако решим (4) спрямо y, ще получим уравнение от вида:

(7) $$g: \quad y = kx + l,$$



черт. 25

което се нарича *декартово уравнение на правата*; коефициентът k се нарича *ъглов коефициент на правата*. Нека правата g пресича абсцисната ос в точка A (черт. 25). Върху абсцисната ос избираме точка B така, че абсцисата ѝ да бъде по-голяма от абсцисата на A. Върху правата g избираме точка C така, че ординатата ѝ да бъде положителна. Елементарно-геометричния ъгъл $\alpha = \angle BAC$ ще наричаме *ъгъл, който правата* g *сключва с положителната посока на* *абсцисната ос*. Това понятие може да се разшири и за прави, които са успоредни на абсцисната ос или съвпадат с нея - ще считаме, че ъгълът, който те сключват с положителната й посока, е равен на нула. Ще отбележим, че ъгълът, за който стана дума, зависи и от ориентацията на ординатната ос. Ако на черт. 25 сменим посоката на ординатната ос и запазим посоката на абсцисната ос, ъгълът ще стане $\pi - \alpha$.

Ще покажем, че ако правата g не е успоредна на ординатната ос, то ъгловият коефициент k в декартовото й уравнение (7) е равен на тангеса на ъгъла, който правата сключва с положителната посока на абсцисната ос, т.е.

$$k = \operatorname{tg}\alpha .$$

За целта да разгледаме правата g_1 с уравнение

$$g_1: \quad y = kx .$$

Тя е успоредна на g (защото векторът $\vec{n}(k,-1)$ е ортогонален и на двете прави) и пресича правата с уравнение $x = 1$ (оста на тангесите) в точка с ордината k. Следователно $\operatorname{tg}\alpha = k$.

6. Уравнение на права през две точки. Нека отново координатната система $O\vec{e}_1\vec{e}_2$ е произволна, а $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ са две различни точки. Ще напишем общо уравнение на правата g, определена от двете точки. За целта нека $M(x, y)$ е произволна точка. Тя ще лежи върху правата g точно тогава, когато векторите $\overrightarrow{M_1M_2}$ и $\overrightarrow{M_1M}$ са колинеарни, или все едно, когато са линейно зависими. Те имат координати съответно $(x_2 - x_1, y_2 - y_1)$ и $(x - x_1, y - y_1)$ и условието за линейна зависимост е

(8)
$$\begin{vmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{vmatrix} = 0.$$

Следователно точка с координати (x, y) лежи върху правата g тогава и само тогава, когато координатите й удовлетворяват уравнението (8), което се нарича *уравнение на права през две точки*. Ако развием детерминантата, можем да го запишем във вида

(8′) $$\left(y_2-y_1\right)\left(x-x_1\right)-\left(x_2-x_1\right)\left(y-y_1\right)=0.$$

Коефициентите y_2-y_1 и $-\left(x_2-x_1\right)$ съответно пред x и y не са едновременно равни на нула, защото точките M_1 и M_2 са различни. Получихме, че уравнението (8) е уравнение от вида (4). От свойствата на детерминантите непосредствено следва, че (8) може да се запише и като

(9) $$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix}=0.$$

7. Пропорционалност на наредени n-торки от числа. Нека $a=\left(a_1,a_2,...,a_n\right)$, $b=\left(b_1,b_2,...,b_n\right)$ са две наредени n-торки от числа и $b\neq\left(0,0,...,0\right)$. Ако съществува число λ, за което $a=\lambda b$, т.е.

(10) $$a_1=\lambda b_1,\quad a_2=\lambda b_2,\quad . \ . \ . \ ,\quad a_n=\lambda b_n,$$

ще казваме, че n-торката a е пропорционална на n-торката b и ще пишем

(11) $$\frac{a_1}{b_1}=\frac{a_2}{b_2}=\ldots\frac{a_n}{b_n},$$

дори и в случаите, когато някой от знаменателите b_i е равен на нула. Ако за някое $i=1,2,\ldots,n$ имаме $b_i=0$, то от $a_i=\lambda b_i$ следва $a_i=0$. Следователно в пропорцията (11) $\frac{a_j}{b_j}$, $j=1,2,\ldots,n$, или е числото λ от равенствата (10), или е символът $\frac{0}{0}$. Последният изразява само това, че $a_j=b_j=0$ и нищо повече. Уславяме се, че пропорцията (11) е само друго записване на равенствата (10) и няма никакъв друг смисъл. Лесно е да се съобрази, че при ненулева n-торка b равенствата (10) са налице тогава и само тогава, когато $a_ib_j=a_jb_i$ за всички $i,j=1,2,\ldots,n$.

Като се имат предвид тези бележки, уравнението (8) на права през две точки се записва във вида

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}. \tag{12}$$

§ 2. Взаимно положение на две и на три прави в равнината. Сноп прави

1. Отново ще предполагаме, че координатната система $O\vec{e}_1\vec{e}_2$ е произволна. Ако в равнината са дадени две прави g_1 и g_2, възможностите за тяхното взаимно положение са добре известни: или съвпадат, или са успоредни, или се пресичат (имат единствена обща точка). Ако правите са зададени с общи уравнения

$$\begin{aligned} g_1:&\quad A_1x + B_1y + C_1 = 0, \\ g_2:&\quad A_2x + B_2y + C_2 = 0, \end{aligned} \tag{1}$$

ще потърсим аналитичните условия, които описват трите посочени възможности. Фактически задачата за общите точки на двете прави е задача за решенията на системата (1) от две линейни уравнения с две неизвестни. Тя се решава лесно с помощта на теоремата на Кронекер-Капели (гл. 2, § 4). Образуваме матриците

$$A = \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix}, \qquad \overline{A} = \begin{pmatrix} A_1 & B_1 & -C_1 \\ A_2 & B_2 & -C_2 \end{pmatrix}.$$

И двете имат ранг поне единица, защото $|A_1| + |B_1| \neq 0$, $|A_2| + |B_2| \neq 0$. Налице са следните възможности:

а) $r(A) = r(\overline{A}) = 1$. Следователно

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2},$$

защото всички минори от втори ред са нули. Очевидно всеки от двата реда в $\overline{A}$ е пропорционален на другия; иначе казано, двете уравнения в (1) са пропорционални в смисъл, че ако умножим едното уравнение с подходящо число, ще получим другото. Следователно *правите g_1 и g_2 съвпадат*.

б) $r(A) = 1$, $\quad r(A) = 2$. Това условие е еквивалентно на

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}\neq\frac{C_1}{C_2},$$

защото минорът от втори ред в A е нула, а $\overline{A}$ съдържа ненулев минор от ред 2. Системата (1) няма решение, т.е. *правите са успоредни*.

в) $r\left(A\right)=r\left(\overline{A}\right)=2$ или, еквивалентно,

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}\neq 0.$$

В този случай системата (1) има единствено решение (получава се по формулите на Крамер), т.е. *правите се пресичат*.

Условията а) - в) за ранговете изчерпват всички възможности и взаимно се изключват, следователно за правите g_1 и g_2 с уравнения (1) е в сила следното

2.Твърдение. а) *правите g_1 и g_2 съвпадат тогава и само тогава, когато*

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2};$$

б) *правите g_1 и g_2 са успоредни тогава и само тогава, когато*

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}\neq\frac{C_1}{C_2};$$

в) *правите g_1 и g_2 се пресичат тогава и само тогава, когато*

$$\frac{A_1}{A_2}\neq\frac{B_1}{B_2}..$$

До този резултат може да се стигне и със следното чисто геометрично разсъждение. За съжаление то не се поддава на обобщения (например за равнини). Знаем, че векторите $\vec{p}_1\left(-B_1,A_1\right)$ и $\vec{p}_2\left(-B_2,A_2\right)$ са колинеарни съответно на правите g_1 и g_2. Ако векторите $\vec{p}_1$ и $\vec{p}_2$ не са колинеарни, правите, разбира се, ще се пресичат и обратно. Чрез координатите условието за неколинеарност се изразява аналитично както във в). Нека по-нататък $\vec{p}_1$ и $\vec{p}_2$ са колинеарни, т.е. $\vec{p}_1=\lambda\vec{p}_2$. Тогава има две възможности: C_1 и λC_2 са или равни, или не. Ясно е, че в първия случай правите съвпадат, а във втория са успоредни. Предположенията, които направихме, се записват аналитично както в а) и б).

В т. 2 на предишния параграф при избрана координатна система на всяка права съпоставихме нейно общо уравнение. Остана впечатлението, че при това съпоставяне има доста много произвол. Доказаното твърдение дава в частност точен отговор за степента на този произвол: *при избрана координатна система общото уравнение на всяка права е определено еднозначно с точност до пропорционалност*.

3. Задачата за намиране на аналитични еквиваленти на различните взаимни разположения на три прави в равнината, зададени с общи уравнения, не предлага принципни трудности. Тъй като възможностите за взаимните разположения сега са значително повече, изчерпателното ѝ решение би се превърнало в отегчително упражнение върху прилагане на теоремата на Кронекер-Капели и позовавания на случаите за две прави. За илюстрация ще разгледаме само два случая, втория от които ще обособим в отделна точка.

Пример. Три прави са зададени с общи уравнения. Кои са аналитичните условия, еквивалентни на изискването правите да бъдат в *общо положение*, т.е. всеки две да се пресичат, без и трите да минават през обща точка.

Нека трите прави са g_1, g_2, g_3 и

(2) $$g_i : A_i x + B_i y + C_i = 0, \qquad i = 1,2,3,$$

са техните общи уравнения. Нека

$$A = \begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \end{pmatrix}, \qquad \overline{A} = \begin{pmatrix} A_1 & B_1 & -C_1 \\ A_2 & B_2 & -C_2 \\ A_3 & B_3 & -C_3 \end{pmatrix}.$$

За да се пресичат всеки две от правите, необходимо и достатъчно е всеки от трите минора от втори ред в матрицата A да бъде ненулев (вж. твърдението в т. 2). В частност това означава и $r\left(A\right) = 2$. От друга страна, за да няма решение системата от трите уравнения (2) (правите да не минават през обща точка), необходимо и достатъчно е $r\left(\overline{A}\right) = 3$, т.е. $\det\overline{A} \neq 0$. Окончателно, трите прави са в общо положение тогава и само тогава, когато трите минора от втори ред в матрицата A са различни от нула (никои два реда в A не са пропорционални) и $\det\overline{A} \neq 0$.

4. Сноп прави в равнината. Множеството на всички прави в равнината, минаващи през фиксирана точка M_0, се нарича *сноп прави*, а точката M_0 се нарича *център* на снопа.

Ако точката M_0 има координати $\left(x_0, y_0\right)$, то произволна права с общо уравнение

$$Ax + By + C = 0$$

ще минава през точката M_0 точно тогава, когато

$$A_0 x + B_0 y + C_0 = 0.$$

От последното равенство определяме C, заместваме го в общото уравнение и получаваме, че всяка права от снопа ще има уравнение от вида

$$A\left(x - x_0\right) + B\left(y - y_0\right) = 0, \qquad |A| + |B| \neq 0.$$

Ако меним A и B, ще получим всички прави от снопа. Често пъти е по-удобно снопът да се зададе чрез две негови фиксирани различни прави g_1 и g_2. Възниква въпросът: при какви условия произволна права g_3 ще минава през пресечната точка на g_1 и g_2? Запазваме означенията от края на предишната точка. Тъй като g_1 и g_2 са различни и се пресичат, то

(3)
$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0.$$

Следователно матрицата A има ранг 2. За да имат трите прави обща точка, необходимо и достатъчно е $\overline{A}$ също да има ранг 2. Първите два реда в $\overline{A}$ са линейно независими поради (3), следователно рангът й ще бъде 2 точно тогава, когато третият й ред е линейна комбинация на първите два. С други думи, уравнението на g_3 може да се запише като линейна комбинация на уравненията на g_1 и g_2:

(4)
$$g_3 \colon \lambda\left(A_1 x + B_1 y + C_1\right) + \mu\left(A_2 x + B_2 y + C_2\right) = 0.$$

Разбира се, λ и μ не са едновременно нули, защото (4) е уравнение на права.

Обратно, нека λ и μ са произволни реални числа, които не са едновременно нули. Тогава в уравнението (4) коефициентите пред x и y не са едновременно нули (те са $\lambda A_1 + \mu A_2$ и $\lambda B_1 + \mu B_2$) заради условието

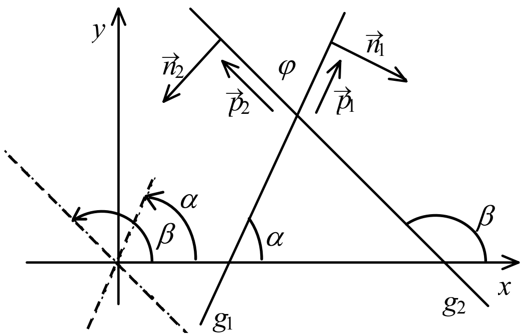
(3), т.е. (4) е уравнение на права. Тя очевидно минава през пресечната точка на правите g_1 и g_2. С това доказахме, че *ако правите*

$$g_1 : A_1x + B_1y + C_1 = 0,$$
$$g_2 : A_2x + B_2y + C_2 = 0$$

се пресичат, то уравнението на всяка права от снопа, определен от g_1 *и* g_2*, има вида* (4) *при подходяща ненулева двойка реални числа* (λ, μ). *Обратно, за всяка ненулева двойка* (λ, μ) *уравнението* (4) *задава права от снопа.*

§ 3. Ъгъл между две прави

1. Две пресичащи се прави винаги определят две двойки връхни ъгли; ако в едната двойка ъглите имат мярка φ, а в другата двойка - мярка ψ, то винаги $\varphi + \psi = \pi$. Тук имаме предвид елементарно геометричните ъгли φ и ψ, т.е. $0 < \varphi < \pi$, $0 < \psi < \pi$. *Който и да е от ъглите* φ *и* ψ ще наричаме *ъгъл между правите*. Това определение разширяваме и за успоредни или съвпадащи прави: ще считаме, че ъгълът между тях е 0 или π $(\varphi = 0,\ \psi = \pi)$.



черт. 26

Ако двете прави са зададени аналитично чрез някакъв вид уравнения (параметрични, общи, декартови), проблемът за определяне на ъгъла между тях е чисто технически и не изисква привличане на нови идеи. Задачата е метрична и, естествено, в *целия параграф* ще предполагаме, че *координатните ситеми са декартови*.

2. Нека правите g_1 и g_2 са зададени с векторни параметрични уравнения

$$g_1: \ \vec{r} = \vec{r}_1 + \lambda\vec{p}_1, \qquad\qquad g_2: \ \vec{r} = \vec{r}_2 + \mu\vec{p}_2 .$$

Ненулевите вектори $\vec{p}_1$ и $\vec{p}_2$ са колинеарни съответно на правите g_1 и g_2 и естествено е да приемем, че ъгълът между правите е $\varphi = \angle\left(\vec{p}_1, \vec{p}_2\right)_e$ (черт. 26). От определенията на скаларно и векторно произведение на свободни вектори знаем, че

$$\text{(1)} \qquad \cos\varphi = \frac{\vec{p}_1\vec{p}_2}{|\vec{p}_1||\vec{p}_2|}, \qquad \sin\varphi = \frac{|\vec{p}_1 \wedge \vec{p}_2|}{|\vec{p}_1||\vec{p}_2|}$$

От първото равенство и от условието $0 \le \varphi \le \pi$ ъгълът φ се определя еднозначно. Ако правите не са ортогонални ($\vec{p}_1\vec{p}_2 \neq 0$), от (1) следва

$$\text{(2)} \qquad \operatorname{tg}\varphi = \frac{|\vec{p}_1 \wedge \vec{p}_2|}{\vec{p}_1\vec{p}_2}.$$

Въвеждаме координати и изразяваме чрез тях десните страни на (1). Нека $\vec{p}_1\left(a_1, b_1\right)$, $\vec{p}_2\left(a_2, b_2\right)$. За да изразим векторното произведение, декартовата координатна система $O\vec{e}_1\vec{e}_2$ допълваме временно до дясна декартова координатна система $O\vec{e}_1\vec{e}_2\vec{e}_3$ на цялото пространство. Сега $\vec{p}_i\left(a_i, b_i, 0\right)$, $i = 1,2$, и

$$\vec{p}_1\vec{p}_2 = a_1a_2 + b_1b_2, \qquad \vec{p}_1 \wedge \vec{p}_2\left(0, 0, a_1b_2 - a_2b_1\right).$$

Следователно

$$\text{(3)} \qquad \cos\varphi = \frac{a_1a_2 + b_1b_2}{\sqrt{a_1^2 + b_1^2}.\sqrt{a_2^2 + b_2^2}}, \qquad \sin\varphi = \frac{|a_1b_2 - a_2b_1|}{\sqrt{a_1^2 + b_1^2}.\sqrt{a_2^2 + b_2^2}},$$

$$\text{(4)} \qquad \operatorname{tg}\varphi = \frac{|a_1b_2 - a_2b_1|}{a_1a_2 + b_1b_2}.$$

Едва ли е нужно да се запаметяват формулите (3) и (4) - те просто са “координатно” записване на (1) и (2). Ще отбележим още, че вместо (4) може да се използва и формула, в която знакът за абсолютна стойност е изпуснат. Става дума за равенството

(5) $$\operatorname{tg}\psi = \frac{a_1 b_2 - a_2 b_1}{a_1 a_2 + b_1 b_2},$$

където ψ е един от двата ъгъла, които правите определят. Това следва веднага от (4) и от тъждеството $\operatorname{tg}\psi = \operatorname{tg}\left(\pi\text{-}\varphi\right) = -\operatorname{tg}\varphi$.

3. Нека правите са зададени с общи уравнения:

$$g_1 : A_1 x + B_1 y + C_1 = 0, \qquad g_2 : A_2 x + B_2 y + C_2 = 0.$$

Както знаем, векторът $\vec{n}_i\left(A_i, B_i\right)$ е ортогонален на правата g_i, $i = 1{,}2$. Сега за ъгъл φ между правите можем да приемем елементарно-геометричния ъгъл между векторите $\vec{n}_1$ и $\vec{n}_2$, т.е. $\varphi = \angle\left(\vec{n}_1, \vec{n}_2\right)_e$.

$$\cos\varphi = \frac{\vec{n}_1 \vec{n}_2}{\left|\vec{n}_1\right|\left|\vec{n}_2\right|}, \qquad \sin\varphi = \frac{\left|\vec{n}_1 \wedge \vec{n}_2\right|}{\left|\vec{n}_1\right|\left|\vec{n}_2\right|}.$$

Записани чрез координатите, тези равенства изглеждат точно както (3): достатъчно е малките букви a_i, b_i да се заменят с главни A_i, B_i, $i = 1{,}2$. Ще отбележим, че ако се работи с различни уравнения (параметрични или общи) на едни и същи прави, ъгълът φ, пресметнат по горните формули, може да се окаже различен. Това не бива да учудва - самата дефиниция на ъгъл между две прави съдържа двузначност: ако φ е едната стойност на ъгъла, $0 \le \varphi \le \pi$, то $\pi - \varphi$ е другата стойност. Лесно е да се съобрази, че ако меним уравненията, горните формули винаги ще дават една от двете стойности.

4. Нека правите са зададени с декартови уравнения:

$$g_1 : y = k_1 x + l_1, \qquad g_2 : y = k_2 x + l_2.$$

За нормалните вектори $\vec{n}_1\left(k_1, -1\right)$ и $\vec{n}_2\left(k_2, -1\right)$, сключващи ъгъл φ, ще приложим резултата от т. 3. Сега

$$\vec{n}_1\vec{n}_2 = k_1k_2 + 1, \qquad\qquad |\vec{n}_1 \wedge \vec{n}_2| = |k_1 - k_2|$$

(относно второто равенство вж. съответното пресмятане за $\vec{p}_1 \wedge \vec{p}_2$ в т. 2).

Правите са ортогонални точно тогава, когато $\vec{n}_1\vec{n}_2 = 0$ или, все едно, *когато*

$$k_1k_2 + 1 = 0.$$

Нека правите не са ортогонални. Тогава

$$\operatorname{tg}\varphi = \frac{|\vec{n}_1 \wedge \vec{n}_2|}{\vec{n}_1\vec{n}_2} = \frac{|k_1 - k_2|}{1 + k_1k_2}. \tag{6}$$

Очевидно е, че *правите са успоредни или съвпадат точно тогава,* когато $\operatorname{tg}\varphi = 0$, т.е. *когато*

$$k_1 = k_2.$$

От друга страна, знаем, че ако α и β са съответно ъглите, които правите g_1 и g_2 сключват с положителната посока на абсцисната ос (вж. § 1, т. 5 и черт. 26), то $k_1 = \operatorname{tg}\alpha$ и $k_2 = \operatorname{tg}\beta$. За ъгъл между правите можем да считаме който и да е от ъглите $|\alpha - \beta|$ и $\pi - |\alpha - \beta|$. След заместване в известното тъждество

$$\operatorname{tg}(\alpha - \beta) = \frac{\operatorname{tg}\alpha - \operatorname{tg}\beta}{1 + \operatorname{tg}\alpha\operatorname{tg}\beta}$$

получаваме

$$\operatorname{tg}(\alpha - \beta) = \frac{k_1 - k_2}{1 + k_1k_2}. \tag{7}$$

От последното равенство винаги може да се определи $|\alpha - \beta|$.

§ 4. Уравнения на равнина

1. Параметрични уравнения на равнина. Един от възможните начини равнина ε да бъде зададена чрез нейни елементи е следният: да се зададе точка M_0, лежаща в ε, и двойка неколинеарни вектори $\vec{p}_1$ и

$\vec{p}$, компланарни с равнината (черт. 27). Очевидно е, че произволна точка M ще лежи в равнината точно тогава, когато векторите $\vec{p}_1$, $\vec{p}_2$ и $\overrightarrow{M_0M}$ са компланарни. Тъй като $\vec{p}_1$ и $\vec{p}_2$ са линейно независими (неколинеарни), геометричното условие за компланарност е еквивалентно на аналитичното условие $\overrightarrow{MM_0}$ да бъде линейна комбинация на $\vec{p}_1$ и $\vec{p}_2$, т.е.

$$\overrightarrow{M_0M} = \lambda\vec{p}_1 + \mu\vec{p}_2 . \tag{1}$$

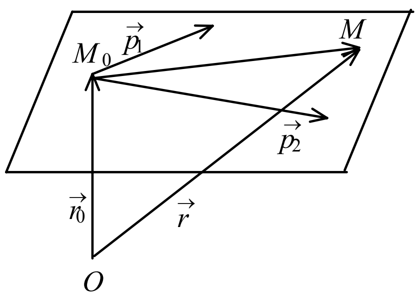
И така, точка M лежи в ε точно тогава, когато е изпълнено (1) за някакви реални числа λ и μ.

Избираме произволна афинна координатна система $O\vec{e}_1\vec{e}_2\vec{e}_3$ и въвеждаме радиус-вектори на точките и координати:

$$\vec{r}_0 = \overrightarrow{OM}_0\left(x_0, y_0, z_0\right),\ \vec{r} = \overrightarrow{OM}\left(x, y, z\right),\ \vec{p}_1\left(a_1, b_1, c_1\right),\ \vec{p}_2\left(a_2, b_2, c_2\right).$$

Тъй като $\overrightarrow{M_0M} = \vec{r} - \vec{r}_0$, съотношението (1) се преписва като

$$\vec{r} - \vec{r}_0 = \lambda\vec{p}_1 + \mu\vec{p}_2, \tag{2}$$



черт. 27

което се нарича *векторно параметрично уравнение* на равнината ε; λ и μ се наричат параметри. Отново подчертаваме: *точка с радиус-вектор $\vec{r}$ лежи в равнината ε точно тогава, когато $\vec{r}$ удовлетворява уравнението* (2) *при подходящи реални числа λ, μ*. Векторното равенство (2) може да се запише по еквивалентен начин като тройка скаларни уравнения за съответните координати:

$$\begin{aligned} x - x_0 &= \lambda a_1 + \mu a_2 \\ y - y_0 &= \lambda b_1 + \mu b_2 \\ z - z_0 &= \lambda c_1 + \mu c_2 . \end{aligned} \tag{3}$$

Уравненията (3) се наричат *скаларни параметрични уравнения на равнината* ε.

2. Общо уравнение на равнина. Във връзка с (3) нека разгледаме матрицата

$$Q = \begin{pmatrix} x - x_0 & a_1 & a_2 \\ y - y_0 & b_1 & b_2 \\ z - z_0 & c_1 & c_2 \end{pmatrix}.$$

Да предположим, че са изпълнени равенствата (3). Те означават, че първият стълб в Q е линейна комбинация на втория и третия. Следователно $\det Q = 0$, т.е.

$$A\left(x - x_0\right) + B\left(y - y_0\right) + C\left(z - z_0\right) = 0, \tag{4}$$

където

$$A = b_1c_2 - b_2c_1, \quad B = -\left(a_1c_2 - a_2c_1\right), \quad C = a_1b_2 - a_2b_1.$$

Тъй като векторите $\vec{p}_1$ и $\vec{p}_2$ не са колинеарни, то поне едно от числата A, B, C е различно от нула. Ако положим още

$$D = -Ax_0 - By_0 - Cz_0,$$

то (4) ще добие вида

$$Ax + By + Cz + D = 0, \qquad |A| + |B| + |C| \neq 0. \tag{5}$$

И така, от (3) следва (5).

Обратно, нека е изпълнено условието (5) за точка с координати $\left(x, y, z\right)$, т.е. детерминантата на разглежданата матрица е равна на нула. Това условие, заедно с факта, че вторият и третият ѝ стълб съдържат ненулев минор от втори ред, означава, че рангът на матрицата е 2 и че първият стълб е линейна комбинация на останалите два, т.е. изпълнени са равенствата (3) при подходящи λ и μ. Окончателно получихме, че условията (3) и (5) са еквивалентни.

Уравнението (5) се нарича *общо уравнение на равнината* ε.

Дотук отново получихме факт, който е част от същността на така наречения координатен метод, водещ началото си от Декарт (1596 - 1650 г.)

и Ферма (1601 - 1665 г.), а именно: на геометричния обект равнина ε беше съпоставен алгебричният обект тройка скаларни параметрични уравнения от вида (3) или еквивалентният му алгебричен обект - уравнение (5) от първа степен с три неизвестни. В случая всички уравнения са с реални коефициенти. Сега ще покажем, че е налице съответствие и в обратната посока, т.е. на всяко уравнение от вида (5), за което $|A|+|B|+|C|\neq 0$, съответства и то единствена равнина, състояща се от всички точки с координати (x,y,z), удовлетворяващи (5).

Нека например $B\neq 0$. Да изберем точка $M_0(x_0,y_0,z_0)$, за която

(6) $$Ax_0+By_0+Cz_0+D=0,$$

и нека (x,y,z) е произволно решение на (5). Следователно

(7) $$A(x-x_0)+B(y-y_0)+C(z-z_0)=0.$$

За да намерим всички решения на (5), е достатъчно да намерим всички решения $(x-x_0,y-y_0,z-z_0)$ на хомогенното уравнение (7). Две линейно независими решения на (7) са $(0,-C,B)$ и $(-B,A,0)$ и всяко решение е тяхна линейна комбинация (вж. гл.2, § 5):

(8) $$\begin{aligned} x-x_0&=\lambda.0+\mu.(-B)\\ y-y_0&=\lambda.(-C)+\mu A\\ z-z_0&=\lambda B+\mu.0. \end{aligned}$$

Както видяхме, параметричните уравнения (8) задават равнина, минаваща през точката M_0 и съдържаща представители на неколинеарните вектори $\vec{p}_1(0,-C,B)$, $\vec{p}_2(-B,A,0)$. Съгласно проведените по-горе разсъждения, уравненията (8) са еквивалентни на единственото уравнение

$$\begin{vmatrix} x-x_0 & 0 & -B\\ y-y_0 & -C & A\\ z-z_0 & B & 0 \end{vmatrix}=0,$$

което след развиване на детерминантата по първия стълб добива вида

$$-AB\left(x-x_0\right)-B^2\left(y-y_0\right)-BC\left(z-z_0\right)=0.$$

Ако съкратим на $-B\neq 0$ и съберем полученото уравнение с равенството (6), ще получим уравнението (5). Тъй като (5) и (8) са еквивалентни, показахме, че множеството от решенията на уравнението (5) (точките с координати $\left(x,y,z\right)$, удовлетворяващи (5)) е точно множеството на точките от равнината, зададена параметрично с уравненията (8). Да резюмираме получения резултат.

3. Теорема. *Нека в пространството е избрана произволна афинна координатна система. Тогава всяка равнина* ε *се описва с някакво уравнение от вида* (5) *в смисъл, че* ε *се състои от всички точки с координати* $\left(x,y,z\right)$, *удовлетворяващи* (5). (Ще казваме, че (5) е *общо уравнение на равнината.*) *Обратно, множеството от точките* $M\left(x,y,z\right)$, *чиито координати удовлетворяват дадено уравнение от вида* (5), *е някаква равнина*.

4. Твърдение. *Векторът* $\vec{q}\left(\alpha,\beta,\gamma\right)$ *е компланарен с равнината* ε, *зададена с общо уравнение* (5) *тогава и само тогава, когато*

(9) $$A\alpha+B\beta+C\gamma=0.$$

Ако координатната система е декартова, векторът $\vec{n}\left(A,B,C\right)$ *е ортогонален на равнината*.

Д о к а з а т е л с т в о. Нека насочената отсечка $\overrightarrow{M_1M_2}$ е представител на вектора $\vec{q}$ и $M_i\left(x_i,y_i,z_i\right)$, $i=1,2$. Както знаем, $\alpha=x_2-x_1$, $\beta=y_2-y_1$, $\gamma=z_2-z_1$. Винаги можем да изберем отсечката така, че началото й M_1 да лежи в равнината ε или, все едно, да бъде изпълнено

(10) $$Ax_1+By_1+Cz_1+D=0.$$

При това условие векторът $\vec{q}=\overrightarrow{M_1M_2}$ ще бъде компланарен на равнината точно тогава, когато точката M_2 лежи в равнината или, еквивалентно, когато

(11) $$Ax_2+By_2+Cz_2+D=0.$$

Ако е налице (11) и от него извадим равенството (10), ще получим

$$A\left(x_2 - x_1\right) + B\left(y_2 - y_1\right) + C\left(z_2 - z_1\right) = 0, \tag{12}$$

т.е. (9). Обратно, от (12) и (10) следва (11), което означава, че точката M_2 лежи в равнината ε, т.е. векторът $\vec{q}$ е компланарен с нея.

Ако координатната система е декартова, лявата страна на (9) е всъщност скаларното произведение $\vec{n}\vec{q}$. Фактът, че то е равно на нула, означава, че векторът $\vec{n}\left(A, B, C\right)$ е ортогонален на всеки вектор $\vec{q}$, копланарен с равнината, т.е. $\vec{n}$ е ортогонален на равнината.

5. Общо уравнение на равнина през три точки. Нека M_1, M_2, M_3 са три точки, които не лежат на една права, и $M_i\left(x_i, y_i, z_i\right)$, $i = 1,2,3$, спрямо някаква афинна координатна система. Целта ни е да напишем общо уравнение на равнината ε, определена от трите точки. Точка $M\left(x, y, z\right)$ лежи в равнината тогава и само тогава, когато векторите $\overrightarrow{M_1M}$, $\overrightarrow{M_1M_2}$, $\overrightarrow{M_1M_3}$ са компланарни, което е еквивалентно на условието детерминантата от координатите им да бъде равна на нула, т.е.

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0. \tag{13}$$

Вторият и третият й ред съдържат поне един ненулев минор от втори ред, защото векторите $\overrightarrow{M_1M_2}$ и $\overrightarrow{M_1M_3}$ не са колинеарни. Ако развием детерминантата по първия ред, ще получим уравнение от вида (5), в което коефициентите пред неизвестните не са едновременно равни на нула. Следователно (13) е общо уравнение на някаква равнина. Тя, разбира се, минава през точките M_i, $i = 1,2,3$, защото координатите им очевидно удовлетворяват уравнението.

Като се използват само свойствата на детерминантите, лесно е да се съобрази, че общото уравнение (13) на равнина през три точки може да се напише и във вида

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0.$$

§ 5. Взаимно положение на две и на три равнини

1. Навсякъде в този параграф ще предполагаме, че координатната система е произволна афинна. Основен инструмент за изследването ще бъде (както в § 2) теоремата на Кронекер - Капели. И така, нека равнините ε_1 и ε_2 са зададени с общи уравнения:

$$\begin{aligned} &\varepsilon_1 : A_1x + B_1y + C_1z + D_1 = 0, \\ &\varepsilon_2 : A_2x + B_2y + C_2z + D_2 = 0. \end{aligned} \tag{1}$$

Образуваме матриците

$$A = \begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix}, \qquad \overline{A} = \begin{pmatrix} A_1 & B_1 & C_1 - D_1 \\ A_2 & B_2 & C_2 - D_2 \end{pmatrix}.$$

И двете имат ранг поне единица, защото всеки от редовете на A по условие съдържа ненулев елемент. Следователно $1 \le r(A) \le r(\overline{A}) \le \ \le r(A)+1$ и за ранговете са налице следните възможности:

а) $r(A) = r(\overline{A}) = 1$. Това означава, че всички минори от втори ред са нули, следователно

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = \frac{D_1}{D_2}.$$

Получихме, че уравненията в (1) са пропорционални, т.е. равнините съвпадат.

б) $r(A) = 1$, $r(\overline{A}) = 2$. Това означава, че минорите от втори ред в матрицата A са нули, а $\overline{A}$ съдържа ненулев минор от втори ред. Еквивалентно:

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}\neq\frac{D_1}{D_2}.$$

По теоремата на Кронекер-Капели системата (1) няма решение, т.е. равнините нямат общи точки (успоредни са).

в) $r\left(A\right)=2$, а следователно $r\left(\overline{A}\right)=2$. Това означава, че матрицата A съдържа ненулев минор от втори ред или, еквивалентно, че двата й реда не са пропорционални. По теоремата на Кронекер-Капели сега системата (1) е съвместима и решенията й, съгласно твърдението в гл.2, § 5, т.2, зависят от един независим параметър. В разглежданата ситуация равнините са различни, защото, ако съвпадат, системата (1) би се редуцирала само на едното уравнение и решенията биха зависили от два независими параметъра. Следователно при условието в) равнините са различни и се пресичат в права.

Тъй като условията а) - в) за ранговете взаимно се изключват, ясно е, че за равнините ε_1 и ε_2 с общи уравнения (1) е в сила следното

2. Твърдение. а) *равнините съвпадат тогава и само тогава, когато*

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}=\frac{D_1}{D_2};$$

б) *равнините са успоредни тогава и само тогава, когато*

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}\neq\frac{D_1}{D_2};$$

в) *равнините се пресичат* (*в права*) *тогава и само тогава, когато поне една от детерминантите*

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}, \qquad \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \qquad \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}$$

е различна от нула.

От този резултат се вижда по-специално, че спрямо фиксирана координатна система *общото уравнение на дадена равнина е определено не съвсем еднозначно, а само с точност до пропорционалност*.

3. Взаимно положение на три равнини. Сноп равнини. Нека трите равнини са зададени с общи уравнения

$$\begin{aligned} &\varepsilon_1 : A_1 x + B_1 y + C_1 z + D_1 = 0 \\ (2) \qquad &\varepsilon_2 : A_2 x + B_2 y + C_2 z + D_2 = 0 \\ &\varepsilon_3 : A_3 x + B_3 y + C_3 z + D_3 = 0. \end{aligned}$$

Образуваме матриците

$$A = \begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{pmatrix}, \qquad \overline{A} = \begin{pmatrix} A_1 & B_1 & C_1 & -D_1 \\ A_2 & B_2 & C_2 & -D_2 \\ A_3 & B_3 & C_3 & -D_3 \end{pmatrix}.$$

Тъй като уравненията (2) са уравнения на равнини, всеки ред на A съдържа поне един ненулев елемент. Следователно $1 \le r\left(A\right)$. Понеже $1 \le r\left(A\right) \le r\left(\overline{A}\right) \le r\left(A\right)+1$, то възможностите за ранговете на матриците A и $\overline{A}$ са следните:

а) $r\left(A\right) = r\left(\overline{A}\right) = 1$. Това означава, че всеки два реда на матрицата $\overline{A}$ са пропорционални и от твърдението в т. 2 получаваме, че трите равнини съвпадат.

б) $r\left(A\right) = 1$, $r\left(\overline{A}\right) = 2$. Според първото от тези условия всеки два реда на матрицата A са пропорционални, което съгласно т. 2 означава, че всеки две от равнините или съвпадат, или са успоредни. Тъй като $r\left(\overline{A}\right) = 2$, да предположим например, че първите два реда на $\overline{A}$ съдържат ненулев минор от втори ред. Това гарантира, че ε_1 и ε_2 са успоредни. За ε_3 остават двете посочени възможности: или е успоредна на ε_1 и ε_2, или съвпада с някоя от тях - всичко зависи от това, дали в $\overline{A}$ има и други ненулеви минори от втори ред, образувани с участието на третия ред на $\overline{A}$, и кои са те; например, ако всеки два реда на $\overline{A}$ съдържат такива минори, трите равнини са различни и успоредни.

в) $r\left(A\right) = 2$, $r\left(\overline{A}\right) = 3$. Сега трите равнини със сигурност са различни, защото в противен случай съгласно т. 2 матрицата $\overline{A}$ би съдържала два пропорционални реда и рангът ѝ би бил най-много 2. Тъй като ранговете на A и $\overline{A}$ са различни, то според теоремата на Кронекер-Капели системата уравнения (2) няма решение, т.е. трите равнини нямат обща точка. По условие $r(A)=2$, следователно поне два от редовете на A съдържат ненулев минор от втори ред; геометричният смисъл е, че съответните две равнини се пресичат в права. Тогава за третата равнина остават две възможности: или е успоредна на една от споменатите две, или пресича другата, или пресича всяка от другите две. В последния случай трите пресечни прави са успоредни, защото трите равнини нямат обща точка. Може да се покаже с примери, че и двете възможности се реализират. Намирането на допълнителни аналитични условия (освен $r(A)=2$, $r(\overline{A})=3$), които характеризират всяка от двете геометрични конфигурации, пропускаме.

г) $r(A)=r(\overline{A})=2$. Това условие означава, че два от редовете на A са линейно независими, а третият е тяхна линейна комбинация. Да предположим за определеност, че първите два реда на A са линейно независими, следователно такива са и първите два реда на $\overline{A}$; от една страна, това означава, че равнините ε_1 и ε_2 се пресичат в права (според т. 2), а от друга страна, че съществуват реални числа λ и μ, за които $\lambda A_1+\mu A_2=A_3,\ \lambda B_1+\mu B_2=B_3,\ \lambda C_1+\mu C_2=C_3,\quad D_1+\mu D_2=D_3$.
Следователно уравнението на третата равнина може да се запише във вида

$$(3)\ \varepsilon_3{:}\lambda\left(A_1x+B_1y+C_1z+D_1\right)+\mu\left(A_2x+B_2y+C_2z+D_2\right)=0,\ |\lambda|+|\mu|\neq 0$$

Оттук вече е очевидно, че всяка обща точка на равнините ε_1 и ε_2 лежи в ε_3, защото координатите ѝ удовлетворяват (3). Накратко казано, две от равнините се пресичат в права, а третата минава през пресечницата им. Тази конфигурация се появява в редица геометрични разглеждания, ето защо ще я анализираме в малко по-общ план.

Множеството от всички равнини, които минават през фиксирана права, се нарича *сноп равнини*; правата се нарича *ос* на снопа. Ясно е, че всеки две различни равнини от снопа го определят еднозначно. Като запазваме означенията от тази точка, нека ε_1 и ε_2 са две различни равнини от снопа. Да образуваме уравнението (3), като изберем λ и μ по произволен начин, стига да не са едновременно равни на нула. В (3) коефициентите пред x,y,z са съответно $\lambda A_1+\mu A_2$, $\lambda B_1+\mu B_2$,

$\lambda C_1 + \mu C_2$. Ако допуснем за момент, че те едновременно са нули , ще се окаже, че и трите детерминанти

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}, \qquad \begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \qquad \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}$$

са равни на нула, което според т. 2 означава, че равнините ε_1 и ε_2 са успоредни или се сливат, противно на условието. Следователно (3) наистина е уравнение на някаква равнина. Тя очевидно принадлежи на снопа, защото минава през оста му (точките с координати (x, y, z), удовлетворяващи уравненията на ε_1 и ε_2, удовлетворяват и (3)). Ще покажем, че *ако в* (3) *меним* λ *и* μ, $|\lambda| + |\mu| \neq 0$, *ще получим уравненията на всички равнини от снопа.* Действително, нека ε_3 е произволна равнина от снопа и за ε_1, ε_2 и ε_3 да разгледаме съвместно уравненията им (2). По условие рангът на матрицата A е поне две, защото ε_1 и ε_2 са различни и се пресичат в права. Ако $r(A) = 3$, то $\det A \neq 0$ и системата от трите линейни уравнения (2) има единствено решение, т.е. трите равнини биха имали единствена обща точка, противно на допускането, че са равнини от снопа. Следователно $r(A) = 2$. Ако допуснем, че $r(\overline{A}) = 3$, то по теоремата на Кронекер-Капели системата (2) няма решение и трите равнини няма да имат обща точка. Окончателно получихме, че $r(A) = r(\overline{A}) = 2$ и, както видяхме по-горе, уравнението на ε_3 може да се запише във вида (3).

д) $r(A) = r(\overline{A}) = 3$. Сега системата (2) има единствено решение (защото $\det A \neq 0$), т.е. трите равнини имат единствена обща точка и изглеждат както равнините, определени от три стени на тетраедър.

§ 6. Уравнения на права в пространството

1. Параметрични уравнения на права в пространството. Отново, както в § 1, т. 1, ще предполагаме, че права g в пространството е зададена чрез точка M_0, лежаща върху g, и ненулев вектор $\vec{p}$, колинеарен на правата. Нека спрямо избрана афинна координатна система точката M_0 има радиус-вектор $\vec{r}_0$ с координати (x_0, y_0, z_0), векторът $\vec{p}$ има координати (a,b,c) и нека $M(x,y,z)$ е произволна точка с радиус-вектор $\vec{r}$. Като повторим буквално разсъждението от § 1, т. 1, стигаме до извода, че точката M лежи върху правата g точно тогава, когато е изпълнено равенството

$$\vec{r} = \vec{r}_0 + \lambda\vec{p} \tag{1}$$

за някакво реално число λ. Това равенство отново се нарича *векторно параметрично уравнение на правата* g. От него изобщо не личи дали правата е в пространството, или е в някаква равнина. Различието се появява, ако заменим векторното уравнение (1) със скаларни уравнения за съответните координати. Сега те са три и отново се наричат *скаларни параметрични уравнения на правата* g:

$$\begin{aligned} x &= x_0 + \lambda a \\ y &= y_0 + \lambda b \\ z &= z_0 + \lambda c. \end{aligned} \tag{2}$$

2. *Всяка права в пространството може да се зададе и като пресечница на две различни равнини*. Нека спрямо произволна афинна координатна система равнините имат уравнения

$$\begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0 \\ A_2x + B_2y + C_2z + D_2 &= 0, \end{aligned} \tag{3}$$

които разглеждаме като система. Както знаем от § 5, т. 2, разглежданите равнини се пресичат в права точно тогава, когато матрицата

$$A = \begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix}$$

има ранг 2. Следователно всеки път, когато пишем система от вида (3) с допълнителното изискване горната матрица да има ранг 2, всъщност задаваме права.

При конкретните приложения умението да се преминава свободно от параметричното задаване (2) на конкретна права g към задаването ѝ (3) като ос на сноп равнини и обратно е доста полезно. Нека правата g е зададена параметрично с уравненията (2). Тъй като векторът $\vec{p}(a,b,c)$ е ненулев, нека например $b \neq 0$. От първото и второто уравнение изключваме λ, а след това правим същото с второто и третото. Получаваме системата

$$
(4) \qquad \begin{aligned} b(x-x_0)-a(y-y_0)&=0 \\ c(y-y_0)-b(z-z_0)&=0. \end{aligned}
$$

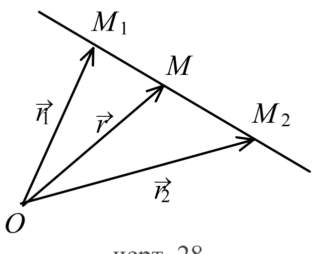
Всяко от уравненията в нея е уравнение на равнина, защото $b \neq 0$. След прилагане на споменатия по-горе критерий лесно се вижда, че двете равнини се пресичат в права, т.е. (4) задават същата права, която задават и (2).

Обратно, нека правата g е зададена с уравненията (3). Всеки вектор, компланарен и с двете равнини, е колинеарен на g. За да намерим такъв ненулев вектор $\vec{p}(a,b,c)$, съгласно § 4, т. 4 е необходимо и достатъчно да решим системата

$$
(5) \qquad \begin{aligned} A_1a+B_1b+C_1c&=0 \\ A_2a+B_2b+C_2c&=0. \end{aligned}
$$

Съгласно зад. 1 от упражненията към глава 1 всяко нейно решение има вида

$$
(6) \qquad a=\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} t, \quad b=-\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix} t, \quad c=\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} t .
$$



черт. 28

За да напишем уравненията (2), остава да намерим поне едно конкретно решение (x_0, y_0, z_0) на системата (3) (точка от правата).

Ще отбележим, че ако координатната система е дясна декартова, решенията на системата (5) може да се напишат въз основа само на геометричната нагледност. Както знаем, сега всеки от векторите $\vec{n}_i\left(A_i, B_i, C_i\right)$, $i=1{,}2$, е ортогонален на съответната равнина, а векторът $\vec{n}_1 \wedge \vec{n}_2$ е колинеарен на пресечницата им. Разбира се, $\vec{n}_1 \wedge \vec{n}_2 \neq \vec{0}$, защото равнините не са успоредни. Следователно всеки вектор, колинеарен на g, има вида $t\left(\vec{n}_1 \wedge \vec{n}_2\right)$, а точно това е записано в (6), но в термините на координатите.

3. Параметрични уравнения на отсечка. Нека M_1 и M_2 са две различни точки в пространството и нека спрямо афинна координатна система с начало O те имат радиус-вектори съответно $\vec{r}_1$ и $\vec{r}_2$. Търсим аналитично описание за точките от отсечката M_1M_2. Както видяхме, точка M с радиус-вектор $\vec{r}$ лежи върху *правата*, определена от M_1 и M_2, точно тогава, когато

(7) $$\overrightarrow{M_1M} = \lambda \overrightarrow{M_1M_2}$$

за някакво реално число λ. Лесно се съобразява (черт. 28), че M ще бъде точка от отсечката M_1M_2 точно тогава, когато $0 \le \lambda \le 1$. Да изразим (7) чрез радиус-векторите на точките:

$$\vec{p}_2\left(-B, A, 0\right)$$

или

(8) $$\vec{r} = \left(1-\lambda\right)\vec{r}_1 + \lambda\vec{r}_2, \qquad \lambda \qquad .$$

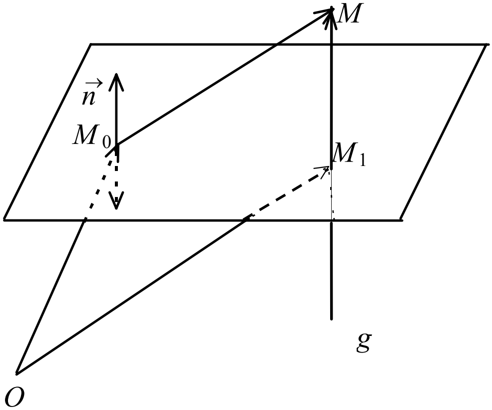
Равенството (8) заедно с условието за λ се нарича *векторно параметрично уравнение на отсечката* g . Ако положим $\alpha = 1-\lambda$, $\beta = \lambda$, (8) може да се препише във вида

(9) $$\vec{r} = \alpha\,\vec{r}_1 + \beta\,\vec{r}_2, \quad \alpha \ge 0, \quad \beta \ge 0, \quad \alpha+\beta = 1.$$

Предимство на (9) е, че в него $\vec{r}_1$ и $\vec{r}_2$ участват симетрично. При необходимост читателят би могъл да замени векторните уравнения (8) и (9) със скаларни за съответните координати. Тук няма да ги изписваме.

§ 7. Разстояние от точка до равнина и от точка до права

Навсякъде в този параграф ще предполагаме, че всички разглеждани координатни системи са декартови.



черт. 29

1. Нека ε и M са дадени съответно равнина и точка в пространството. През точката M (черт. 29) прекарваме права g, перпендикулярна на ε. Нека тя пробожда равнината в точка M_1. Както знаем, дължината на отсечката $M\quad{}_1$ се нарича *разстояние от точката M до равнината ε*; ще го бележим с $d\left(M,\varepsilon\right)$. Да предположим, че M_0 е някаква (произволна, но фиксирана) точка от равнината, а $\vec{n}$ е вектор с дължина единица, който е ортогонален на равнината. Ясно е, че ортогоналната проекция на отсечката M_0M върху правата g е отсечката M_1M. С помощта на вектора $\vec{n}$ (той е колинеарен на g) превръщаме g в ос. От свойствата на скаларното произведение знаем, че

$$(1) \qquad \vec{n}\,\overrightarrow{M_0M} = |\vec{n}|\overline{np_{\vec{n}}\,\overrightarrow{M_0M}} = 1.\overline{M_1M} = \overline{M_1M}\,,$$

където в най-дясната страна на последното равенство стои алгебричната мярка на $\overrightarrow{M_1M}$ върху оста g, определена от $\vec{n}$. Вече е ясно, че

(2) $$d\left(M,\varepsilon\right) = \pm(\vec{n}.\overrightarrow{M_0M}),$$

където знакът е плюс, ако векторите $\vec{n}$ и $\overrightarrow{M_0M}$ са еднопосочни, и минус - в противен случай. Числото

(3) $$\delta\left(M,\varepsilon\right) = \vec{n}.\overrightarrow{M_0M}$$

ще наричаме *насочено* (*ориентирано*) *разстояние* от точката M до равнината ε. Подчертаваме, че то зависи от избора на вектора $\vec{n}$. Той обаче не внася голям произвол: единствените вектори с дължина 1 и ортогонални на ε са $\vec{n}$ и $-\vec{n}$. Ако си представяме, че представител на $\vec{n}$ е нанесен с начало, лежащо в равнината ε, то за точките M, които лежат в онова полупространство, в което е краят на този представител, имаме $\delta\left(M,\varepsilon\right) > 0$, за точките от другото полупространство $\delta\left(M,\varepsilon\right) < 0$ и $\delta\left(M,\varepsilon\right) = 0$ за точките от ε.

По-нататък да изберем някаква декартова координатна система и нека спрямо нея $M\left(x,y,z\right)$, $M_0\left(x_0,y_0,z_0\right)$ и равнината ε е зададена с общо уравнение

(4) $$\varepsilon\colon Ax + By + Cz + D = 0.$$

Както видяхме (вж. § 4, т. 4), сега векторът с координати $\left(A,B,C\right)$ е ортогонален на равнината ε. Нуждаем се обаче от вектор, който освен това да има дължина единица.

В тази връзка се използва следната терминологията: нека спрямо декартова координатна система дадена равнина има общо уравнение (4). Ще казваме, че то е *нормално*, ако векторът с координати $\left(A,B,C\right)$ има дължина единица, т.е. ако $A^2 + B^2 + C^2 = 1$. От разсъжденията по-горе се вижда, че в качеството на A, B, C в нормалното уравнение може да се използват или координатите на $\vec{n}$ (за който стана дума), или на $-\vec{n}$. Тъй като всеки две уравнения на ε спрямо една и съща координатна система са пропорционални, ясно е, че всяка равнина (спрямо фиксирана декартова координатна система) има точно две нормални уравнения и ако (4) е общо уравнение на равнината, то

(5) $$\frac{\pm 1}{\sqrt{A^2+B^2+C^2}}\left(Ax+By+Cz+D\right)=0$$

са двете нейни нормални уравнения. (Две са заради избора на знака.)

Да предположим, че (4) е нормално уравнение на равнината и да изразим чрез координатите дясната страна на (3). Сега $\vec{n}\left(A,B,C\right)$ и $\overrightarrow{M_0M}\left(x-x_0,y-y_0,z-z_0\right)$; следователно

(6) $$\delta\left(M,\varepsilon\right)=A\left(x-x_0\right)+B\left(y-y_0\right)+C\left(z-z_0\right).$$

Тъй като точката M_0 е от равнината, то

(7) $$Ax_0+By_0+Cz_0+D=0.$$

От (6) и (7) получаваме окончателно, че

(8) $$\delta\left(M,\varepsilon\right)=Ax+By+Cz+D.$$

С това доказахме следното

2. Твърдение. *Ориентираното разстояние от точка до равнина е равно на числото, което се получава, след като в лявата страна на нейно нормално уравнение заместим променливите с координатите на точката.*

3. Проблемът за аналитичното изразяване на разстоянието от точка до права в равнината се решава буквално както в т. 1, ето защо ще спестим обясненията.

Нека в равнината е избрана декартова координатна система, дадена е права g с общо уравнение

(9) $$g\colon Ax+By+C=0,$$

$M\left(x,y\right)$ е произволна точка от равнината, а $M_0\left(x_0,y_0\right)$ е някаква точка от правата. От M спускаме перпендикуляр към правата и нека M_1 е петата му. Дължината на отсечката M_1M ще наричаме *разстояние от точката* M до правата g и ще пишем $d\left(M,g\right)$. Нека освен това векторът $\vec{n}$ с дължина единица е ортогонален на правата g. Отново, както в т.1, въвеждаме насоченото разстояние

$$\delta(M,g)=\vec{n}\,\overrightarrow{M_0M}\ ;\quad d(M,g)=\left|\delta(M,g)\right|.$$

Знаем, че при декартова координатна система векторът с координати (A,B) е ортогонален на правата с общо уравнение (9). Ако той има дължина единица, т.е. ако $A^2+B^2=1$, ще казваме, че уравнението (9) е *нормално*. Буквално както за равнина, се вижда, че всяка права спрямо зададена декартова координатна система има точно две нормални уравнения и те се получават от (9) след умножаване с $\pm 1/\sqrt{A^2+B^2}$.

По-нататък ще предполагаме, че уравнението (9) е нормално и нека $\vec{n}(A,B)$. Векторът $\overrightarrow{M_0M}$ има координати $(x-x_0,y-y_0)$, следователно

$$\delta(M,g)=\vec{n}\,\overrightarrow{M_0M}=A(x-x_0)+B(y-y_0).$$

Точката M_0 по условие лежи върху правата, което означава, че

$$Ax_0+By_0+C=0,$$

откъдето за ориентираното разстояние получаваме окончателно

$$\delta(M,g)=Ax+By+C\ . \tag{10}$$

Твърдението, което се съдържа във формулата (10), е напълно аналогично на доказаното в т. 2 и няма да го формулираме отделно.

§ 8. Знак на линейния полином $Ax+By+Cz+D$

1. Изпъкнали множества. Нека X е множество от точки. Ще казваме, че то е *изпъкнало*, ако за всеки две точки от X отсечката, която ги съединява, лежи изцяло в X . Ще считаме, че празното множество също е изпъкнало.

Според това определение всяка отсечка, права, равнина и цялото пространство очевидно са изпъкнали множества.

Понятието "изпъкнало множество" не бива да се смесва с понятието "изпъкнал многоъгълник". Например според традиционното за евклидовата геометрия схващане триъгълникът

ABC е множеството на точките от отсечките AB , BC , CA (страни на триъгълника), където A, B, C са три точки, които не лежат на една права. В този смисъл триъгълникът не е изпъкнало множество. Изпъкнало множество ще получим, ако към точките от страните присъединим и точките, които лежат "вътре" в триъгълника.

Не е трудно да се докаже, че сечението на произволна фамилия от изпъкнали множества също е изпъкнало множество. Оттук лесно следва, че ако Y е произволно множество от точки, то съществува единствено изпъкнало множество X , удовлетворяващо следните условия: а) $X \supset Y$; б) ако X' е изпъкнало и $X' \supset Y$, то $X' \supset X$. За целта в качеството на X е достатъчно да се вземе сечението на всички изпъкнали множества, които съдържат Y . (Такива множества съществуват, например цялото пространство.) Множеството X с горните условия се нарича *изпъкнала обвивка на* Y . Формалните доказателства оставяме на читателя.

2. Полупространства и полуравнини. Нека ε е фиксирана равнина. Ще казваме, че две точки M_1 и M_2, които не лежат в ε , *са в едно и също отворено полупространство*, или че са от една и съща страна на равнината, ако отсечката M_1M_2 не пресича равнината. Ако със средствата на училищната геометрия се опитаме да докажем строго, че множеството от точките, които са вън от равнината, се разбива на точно две отворени полупространства, които са изпъкнали множества и сечението им е празно, ще срещнем неочаквани затруднения. За целта ще използваме някои от средствата на аналитичната геометрия, с които вече разполагаме.

И така, избираме произволна афинна координатна система в пространството и нека

(1) $$\varepsilon : Ax + By + Cz + D = 0$$

е общо уравнение на равнината ε спрямо избраната координатна система. За точка M с координати (x, y, z) полагаме

(2) $$f(M) \equiv f(x, y, z) = Ax + By + Cz + D .$$

Ясно е, че множеството на точките, за които $f(M) = 0$, е равнината ε .

Ще докажем следното твърдение:

Нека точките M_1 и M_2 не лежат в равнината ε . Те са в едно и също отворено полупространство, определено от ε , тогава и само *тогава, когато числата* $f(M_1)$ *и* $f(M_2)$ *са с еднакви знаци. Ако точките* M_1 *и* M_2 *са в едно и също полупространство, то и всички точки от отсечката* M_1M_2 *са в това отворено полупространство.*

Наистина, нека M е произволна точка от отсечката M_1M_2. Нейните координати се изразяват чрез координатите на M_1 и M_2 по формулите (8) от § 6, т. 3. Ако считаме, че M_i има радиус-вектор $\vec{r}_i$ с координати (x_i, y_i, z_i), $i = 1,2$, то с леко пресмятане получаваме, че

$$(3) \qquad f(M) = (1-\lambda) f(M_1) + \lambda f(M_2), \qquad 0 \le \lambda \le 1.$$

Тъй като точките M_1 и M_2 не лежат в равнината ε, то $f(M_1) \neq 0$ и $f(M_2) \neq 0$. Ако $f(M_1)$ и $f(M_2)$ са с еднакви знаци, от (3) е очевидно, че $f(M)$ ще има същия знак и в частност $f(M) \neq 0$. Следователно никоя точка от отсечката M_1M_2 не лежи в равнината ε.

Обратно, нека отсечката M_1M_2 не пресича равнината, но да допуснем за момент, че $f(M_1)$ и $f(M_2)$ са с противоположни знаци. Тогава дясната страна на (3) се анулира при

$$\lambda_0 = \frac{-f(M_1)}{f(M_2) - f(M_1)}, \qquad 0 < \lambda_0 < 1.$$

Ако M_0 е съответната точка от отсечката, която се получава при $\lambda = \lambda_0$, от (3) следва, че $f(M_0) = 0$, т.е. отсечката пресича равнината, противно на допускането. Следователно $f(M_1)$ и $f(M_2)$ са с еднакви знаци. Втората част на твърдението следва от (3): ако $f(M_1)$ и $f(M_2)$ са с еднакви знаци, то $f(M)$ има същия знак, което завършва доказателството.

След всичко това да означим

$$X^+ = \{M \mid f(M) > 0\}, \quad X^- = \{M \mid f(M) < 0\}.$$

Тъй като коефициентите A, B, C в уравнението (1) не са едновременно нули, то множествата X^+ и X^- не са празни, но сечението им е празното множество. Според доказаното твърдение те са изпъкнали множества.

Празни са и сеченията им с равнината ε. Очевидно цялото пространство (като множество от точки) е обединението $X^{+} \bigcup \varepsilon \bigcup X^{-}$; X^{+} и X^{-} са съответстващите на нагледните ни представи отворени полупространства, определени от равнината. Множествата $X^{+} \bigcup \varepsilon$ и $X^{-} \bigcup \varepsilon$ обикновено се наричат *затворени полупространства*.

Аналогично стоят нещата, ако във фиксирана равнина ε е дадена права g. Ако изберем по произволен начин афинна координатна система, правата ще има общо уравнение от вида

(4) $$g : Ax + By + C = 0.$$

Точките от равнината, които не лежат на правата, се разбиват на две отворени полуравнини X^{+} и X^{-}, които са изпъкнали множества; аналитично те се характеризират така:

$$X^{+} = \left\{ M\left(x, y\right) \mid Ax + By + C > 0 \right\},$$

$$X^{-} = \left\{ M\left(x, y\right) \mid Ax + By + C < 0 \right\}.$$

Тъй като доказателството буквално повтаря проведеното по-горе, няма да го излагаме.

3. Знак на полинома $Ax + By + Cz + D$. В т. 2 дадохме аналитична интерпретация на геометричната ситуация. Това подсказва, че можем да геометризираме следната аналитична задача. Нека

(5) $$f\left(x, y, z\right) = Ax + By + Cz + D$$

е произволен полином от първа степен на три променливи с реални коефициенти. Това означава по-специално, че A, B, C не са едновременно нули. Интересуваме се от решенията на неравенството

(6) $$f\left(x, y, z\right) > 0.$$

Въвеждаме афинна координатна система и интерпретираме $\left(x, y, z\right)$ като координати на точка M. Вече е очевидно, че множеството от решенията на (6) може да се интерпретира като отворено полупространство, определено от равнина ε с уравнение $f\left(x, y, z\right) = 0$.

Полученият резултат изглежда толкова елементарен, че би могло да го наречем наблюдение. Оказва се обаче, че е възможно геометричните понятия да се обобщят, вместо полинома (5) да се вземе полином $f\left(x_1,...,x_n\right)$ от първа степен на n променливи с реални коефициенти. Решенията на уравнението $f\left(x_1,...,x_n\right)=0$ се интерпретират като хиперравнина в n-мерно афинно пространство, а решенията на неравенството $f>0$ - като отворено полупространство. Подобен геометричен "нагледˮ в n-мерно пространство се оказва много удобен в дисциплината, наречена *линейно програмиране*. В рамките на тази книга няма да се занимаваме с обобщенията, за които споменахме. Ще отбележим в заключение, че резултатите в този параграф са подсказани от разглежданията в § 7. Там обаче работихме само с декартови координатни системи.

У П Р А Ж Н Е Н И Я

Във всички следващи задачи се предполага, ако не е казано друго, че координатната система е афинна.

1. В равнината е дадена права g, която не минава през началото на координатната система и пресича координатните оси в точките $M_1\left(a,0\right)$ и $M_2\left(0,b\right)$. Докажете, че уравнението на правата g може да се напише във вида

$$\frac{x}{a}+\frac{y}{b}=1 .$$

(Нарича се *отрезово уравнение* на правата.)

2. (Уравнение на права през точка и перпендикулярна на даден вектор) В равнината е избрана декартова координатна система и са дадени точка $M_0\left(x_0,y_0\right)$ и ненулев вектор $\vec{n}\left(A,B\right)$. Докажете, че правата, която минава през точката M_0 и е перпендикулярна на вектора $\vec{n}$, има уравнение

$$A\left(x-x_0\right)+B\left(y-y_0\right)=0 .$$

3. (Уравнение на права през точка и перпендикулярна на дадена права) В равнината са дадени права g, която спрямо декартова координатна система има уравнение $Ax+By+C=0$, и точка

$M_0\left(x_0, y_0\right)$. Докажете, че правата, която минава през точката M_0 и е перпендикулярна на правата g, има уравнение

$$B\left(x-x_0\right)-A\left(y-y_0\right)=0.$$

4. (Уравнение на права през точка и успоредна на дадена права) Дадена е права g с уравнение $Ax+By+C=0$ и точка $M_0\left(x_0, y_0\right)$. Докажете, че правата, която минава през точката M_0 и е успоредна на правата g, има уравнение

$$A\left(x-x_0\right)+B\left(y-y_0\right)=0.$$

5. В равнината са дадени права g, която спрямо декартова координатна система има уравнение $Ax+By+C=0$, и две различни точки $M_1\left(x_1, y_1\right)$ и $M_2\left(x_2, y_2\right)$. Докажете, че точките M_1 и M_2 са разположени симетрично спрямо правата g тогава и само тогава, когато

$$Ax_1+By_1+C=-\left(Ax_2+By_2+C\right), \quad B\left(x_2-x_1\right)-A\left(y_2-y_1\right)=0.$$

Докажете, че симетралата на отсечката M_1M_2 има уравнение

$$\left(x_2-x_1\right)(x-\frac{x_1+x_2}{2})+\left(y_2-y_1\right)(y-\frac{y_1+x_2}{2})=0.$$

6. В равнината са дадени две прави g_1 и g_2, които се пресичат, и спрямо декартова координатна система имат уравнения

$$g_1\colon Ax_1+By_1+C_1=0, \quad g_2\colon Ax_2+By_2+C_2=0.$$

Докажете, че ъглополовящите на ъглите, които правите g_1 и g_2 сключват, имат уравнения

$$\frac{1}{\sqrt{A_1^2+B_1^2}}\left(A_1x+B_1y+C_1\right)=\frac{\varepsilon}{\sqrt{A_2^2+B_2^2}}\left(A_2x+B_2y+C_2\right), \quad \varepsilon=\pm 1.$$

7. В равнината са дадени три точки $M_i\left(x_i, y_i\right)$, $i=1,2,3$.Докажете, че следните условия са еквивалентни:

а) трите точки са колинеарни (лежат на една права);

б) в сила е равенството

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}=0;$$

в) съществуват реални числа $\lambda_1, \lambda_2, \lambda_3$, които не са едновременно равни на нула и за които:

$$\lambda_1+\lambda_2+\lambda_3=0, \quad \lambda_1x_1+\lambda_2x_2+\lambda_3x_3=0, \quad \lambda_1y_1+\lambda_2y_2+\lambda_3y_3=0.$$

(У п ъ т в а н е . Използвайте, че ако една детерминанта е равна на нула, то редовете ѝ са линейно зависими.)

8. В равнината са дадени три точки $M_i\left(x_i, y_i\right)$, $i=1,2,3$, които не лежат на една права. Докажете, че за всяка точка $M\left(x, y\right)$ съществуват еднозначно определени числа m_1, m_2, m_3, за които:

$$m_1+m_2+m_3=1, \quad x=m_1x_1+m_2x_2+m_3x_3, \quad y=m_1y_1+m_2y_2+m_3y_3.$$

З а б е л е ж к а . Числата m_1, m_2, m_3 се наричат *барицентрични координати* на точката M спрямо триъгълника $M_1M_2M_3$. Терминологията е продиктувана от следното съображение от физиката. Нека m_1, m_2, m_3 са произволни положителни реални числа, за които $m_1+m_2+m_3=1$. Ако във върховете на триъгълника $M_1M_2M_3$ поставим материални точки с маси съответно m_1, m_2 и m_3, то центърът на тежестта на системата от трите материални точки е точка M, чиито координати се пресмятат по горните формули. Ще подчертаем, че във физиката масата на материална точка винаги е положително число, а барицентричните координати могат да бъдат и отрицателни числа.

9. Нека точките M_1, M_2, M_3 не лежат на една права и спрямо дадена афинна координатна система (в равнината или в пространството) имат радиус-вектори съответно $\vec{r}_1, \vec{r}_2, \vec{r}_3$. Докажете, че точка M с радиус-вектор $\vec{r}$ лежи вътре в триъгълника $M_1M_2M_3$ или върху страните му точно тогава, когато съществуват неотрицателни реални числа $\lambda_1, \lambda_2, \lambda_3$, за които:

$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \qquad \vec{r} = \lambda_1\vec{r}_1 + \lambda_2\vec{r}_2 + \lambda_3\vec{r}_3.$$

Кога точката M лежи върху някоя от страните на триъгълника $M_1M_2M_3$? Обобщете твърдението за тетраедър, т.е. за четири точки M_i, $i = 1,2,3,4$, които не лежат в една равнина (не са компланарни). Като се върнете към задача 8, покажете, че точката M лежи върху страните на триъгълника $M_1M_2M_3$ или вътре в него тогава и само тогава, когато барицентричните й координати m_1, m_2, m_3 са неотрицателни числа.

(Упътване. Използвайте параметричните уравнения на отсечка от § 6 т. 3 и следния факт: точката M е вътре в триъгълника $M_1M_2M_3$ тогава и само тогава, когато е вътрешна точка на някаква отсечка M_1M', където точката M' е вътрешна точка на отсечката M_2M_3.)

10. Дадена равнина ε не минава през началото на координатната система и пресича координатните оси съответно в точки $M_1\left(a, 0, 0\right)$, $M_2\left(0, b, 0\right)$ и $M_3\left(0, 0, c\right)$. Докажете, че уравнението на равнината ε може да се напише във вида

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1.$$

(Нарича се *отрезово уравнение* на равнината.)

11. Дадена е равнина ε с уравнение $Ax + By + Cz + D = 0$. Намерете необходими и достатъчни условия равнината ε да има някое от следните свойства: а) да съвпада с някоя от координатните равнини; б) да бъде успоредна на някоя от координатните равнини; в) да бъде успоредна на някоя от координатните оси; г) да минава през някоя от координатните оси.

12. (Уравнение на равнина през точка и перпендикулярна на дадена равнина) Нека ε е равнина, която спрямо дадена декартова координатна система има уравнение $Ax + By + Cz + D = 0$, и $M_0\left(x_0, y_0, z_0\right)$ е дадена точка. Да се докаже, че всяка равнина ε_1, която минава през точка M_0 и е перпендикулярна на равнината ε, има уравнение

$$\alpha\left(x-x_0\right)+\beta\left(y-y_0\right)+\gamma\left(z-z_0\right)=0,$$

където $A\alpha + B\beta + C\gamma = 0$ и $|\alpha|+|\beta|+|\gamma| \neq 0$.

13. (Уравнение на равнина през точка и успоредна на дадена равнина) Дадена е равнина ε с уравнение $Ax + By + Cz + D = 0$ и точка $M_0\left(x_0, y_0, z_0\right)$, която не лежи в равнината ε. Докажете, че равнината, която минава през точката M_0 и е успоредна на равнината ε, има уравнение

$$A\left(x-x_0\right)+B\left(y-y_0\right)+C\left(z-z_0\right)=0.$$

14. Дадени са две равнини ε_1 и ε_2, които се пресичат, и спрямо декартова координатна система имат уравнения

$$\varepsilon_1 : A_1x + B_1y + C_1z + D_1 = 0, \qquad \varepsilon_2 : A_2x + B_2y + C_2z + D_2 = 0.$$

Докажете, че ъглополовящите равнини на двустенните ъгли, които равнините сключват, имат уравнения

$$\frac{1}{\sqrt{A_1^2+B_1^2+C_1^2}}\left(A_1x+B_1y+C_1z+D_1\right)=\frac{\varepsilon}{\sqrt{A_2^2+B_2^2+C_2^2}}\left(A_2x+B_2y+C_2z+D_2\right)$$

$$, \varepsilon = \pm 1.$$

15. В пространството са дадени четири точки $M_i\left(x_i, y_i, z_i\right)$, $i = 1, 2, 3, 4$. Докажете, че следните условия са еквивалентни:

а) точките са компланарни (лежат в една равнина);

б) в сила е равенството

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0;$$

в) съществуват реални числа λ_i, $i = 1, 2, 3, 4$, които не са едновременно равни на нула и за които:

$$\sum_{i=1}^{4} \lambda_i = 0, \quad \sum_{i=1}^{4} \lambda_i x_i = 0, \quad \sum_{i=1}^{4} \lambda_i y_i = 0, \quad \sum_{i=1}^{4} \lambda_i z_i = 0.$$

16. В пространството са дадени четири точки $M_i\left(x_i, y_i, z_i\right)$, $i = 1, 2, 3, 4$, които не са компланарни. Да се докаже, че за всяка точка $M\left(x, y, z\right)$ съществуват еднозначно определени числа m_i, $i = 1, 2, 3, 4$, за които:

$$\sum_{i=1}^{4} m_i = 1, \quad x = \sum_{i=1}^{4} m_i x_i, \quad y = \sum_{i=1}^{4} m_i y_i, \quad z = \sum_{i=1}^{4} m_i z_i.$$

(Числата m_i се наричат *барицентрични координати* на точката M спрямо тетраедъра $M_1 M_2 M_3 M_4$.)

17. Докажете, че всяка права в пространството може да бъде зададена чрез двойка уравнения от поне един от следните три вида:

а) $\begin{matrix} x = az + p \\ y = bz + q \end{matrix}$; б) $\begin{matrix} x = ay + p \\ z = by + q \end{matrix}$; в) $\begin{matrix} y = ax + p \\ z = bx + q \end{matrix}$

(наричат се *канонични уравнения* на правата). Ако например б) са канонични уравнения на правата, докажете, че $x = a\lambda + p$, $y = \lambda$, $z = b\lambda + q$ са нейни параметрични уравнения.

(У п ъ т в а н е. Ако правата е зададена като пресечница на две равнини с уравнения $A_1 x + B_1 y + C_1 z + D_1 = 0$ и $A_2 x + B_2 y + C_2 z + D_2 = 0$, използвайте резултата от § 5, т. 2 и покажете, че системата от уравненията на двете равнини винаги може да се реши спрямо някои две от неизвестните x, y, z.)

18. В пространството са дадени две различни точки $M_1\left(x_1, y_1, z_1\right)$ и $M_2\left(x_2, y_2, z_2\right)$. Докажете, че правата, която минава през двете точки, има параметрични уравнения

$$x = x_1 + \lambda\left(x_2 - x_1\right), \quad y = y_1 + \lambda\left(y_2 - y_1\right), \quad z = z_1 + \lambda\left(z_2 - z_1\right).$$

19. В пространството са дадени равнина ε и точка M_0. Спрямо декартова координатна система равнината има уравнение $Ax + By + Cz + D = 0$, а точката има координати $\left(x_0, y_0, z_0\right)$. Докажете, че правата, която минава през точката M_0 и е перпендикулярна на равнината ε, има параметрични уравнения

$$x = x_0 + \lambda A, \qquad y = y_0 + \lambda B, \qquad z = z_0 + \lambda C.$$

20. Дадена е равнината $\varepsilon: Ax + By + Cz + D = 0$ и две точки $M_1\left(x_1, y_1, z_1\right)$ и $M_2\left(x_2, y_2, z_2\right)$. Намерете необходими и достатъчни условия равнината да пресича отсечката M_1M_2.

Глава 7

НЕОБХОДИМИ СВЕДЕНИЯ ОТ АЛГЕБРАТА

§ 1. Комплексни числа

1. В училищния курс по математика се разглеждат последователно множествата $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ съответно на естествените, целите, рационалните и реалните числа.

Когато се решава квадратно уравнение с реални коефициенти и отрицателна детерминанта, се оказва, че то няма реални корени. Този факт (а и много други) показва, че е необходимо да се разшири множеството на реалните числа така, че в новото числово множество операциите събиране и умножение по възможност да запазят познатите си свойства и освен това да може да се извлича квадратен корен от отрицателно реално число.

2. Нека

$$\mathbf{C} = \{(a,b) \mid a,b \in \mathbf{R}\}$$

е множеството на всички наредени двойки от реални числа. Ще считаме, че $(a,b) = (c,d)$ тогава и само тогава, когато $a = c$ и $b = d$. В множеството $\mathbf{C}$ дефинираме операции събиране и умножение с равенствата:

(1)
$$(a,b) + (c,d) = (a+c, b+d),$$

(2)
$$(a,b)(c,d) = (ac - bd, ad + bc).$$

Непосредствено се проверява, че ако α, β, γ са три произволни елемента на $\mathbf{C}$, то са налице следните свойства:

1) $\alpha + \beta = \beta + \alpha$ - комутативен закон при събиране;

2) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ - асоциативен закон при събиране;

3) съществува еднозначно определен елемент $\omega \in \mathbf{C}$, за който $\alpha + \omega = \alpha$ за всеки елемент $\alpha \in \mathbf{C}$. Елементът $\omega = (0,0)$ наричаме *нулев елемент* на $\mathbf{C}$ или просто *нула*.

4) За всеки елемент $\alpha \in \mathbf{C}$ съществува елемент $-\alpha \in \mathbf{C}$, така че

$$\alpha + (-\alpha) = \omega.$$

Този елемент е еднозначно определен и ако $\alpha = (a,b)$, то $-\alpha = (-a,-b)$.

5) $\alpha\beta = \beta\alpha$;

6) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$;

7) $(\alpha+\beta)\gamma = \alpha\gamma + \beta\gamma$;

8) $\alpha\omega = \omega$, където ω е нулевият елемент на $\mathbf{C}$.

Да разгледаме подмножеството на $\mathbf{C}$, което се състои от всички наредени двойки, чиято втора компонента е нула. Изпълнено е

$$(a,0)+(b,0)=(a+b,0),$$

$$(a,0)(b,0)=(ab-0.0, a.0+b.0)=(ab,0).$$

Това означава, че сумата и произведението на наредените двойки от разглеждания вид е от същия вид и първите компоненти на сумата и произведението на $(a,0)$ с $(b,0)$ са съответно сумата и произведението на реалните числа a и b.

Поради това, ще отъждествяваме реалното число a с наредената двойка $(a,0)$ от $\mathbf{C}$ (т.е. $a=(a,0)$) и можем да считаме, че $\mathbf{R} \subset \mathbf{C}$.

3. Определение. Множеството $\mathbf{C}$ на наредените двойки реални числа с дефинираните операции събиране (1) и умножение (2) ще наричаме *множество на комплексните числа*, а всеки елемент на $\mathbf{C}$ ще наричаме *комплексно число*.

4. Да означим с i наредената двойка $(0,1)$. Изпълнено е

$$i^2 = i.i = (0,1)(0,1) = (0.0-1.1, 0.1+1.0) = (-1,0) = -1.$$

Ако $b \in \mathbf{R}$, то

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

Тогава за произволно комплексно число $\alpha = (a,b)$ имаме:

$$\alpha = (a,b) = (a,0)+(0,b) = a+bi.$$

По такъв начин всяко комплексно число може да се запише по единствен начин във вида $a+bi$, $a,b \in \mathbf{R}$. Този запис ще наричаме *алгебричен вид* на комплексното число α. Реалните числа a и b ще наричаме съответно *реална* и *имагинерна* част на α и ще ги означаваме с $a=\operatorname{Re}\alpha$, $b=\operatorname{Im}\alpha$. Комплексното число $i=(0,1)$ ще наричаме *имагинерна единица*, а числата от вида bi, $b\in\mathbf{R}$ - *чисто имагинерни*. Тъй като $i^2=-1$, комплексното число i е решение на уравнението $x^2+1=0$ и $i=\sqrt{-1}$. Това позволява да намираме квадратен корен от произволно отрицателно число. Например: $\sqrt{-25}=\sqrt{25(-1)}=\pm5\sqrt{-1}=\ =\pm5i$, $\sqrt{-7}=\pm\mathrm{i}\sqrt{7}$.

Събирането на комплексни числа в алгебричен вид се извършва, като съберем съответно реалните и имагинерните им части, т.е.

$$(a+bi)+(c+di)=(a,b)+(c,d)=(a+c,b+d)=(a+c)+(b+d)i$$

Умножението се извършва, като формално разкрием скобите и вземем пред вид, че $i^2=-1$, т.е.

$$(a+bi)(c+di)=ac+adi+bci+bdi^2=(ac-bd)+(ad+bc)i=(ac-bd,ad+bc)$$

5. Нека $\alpha=a+bi$ е произволно комплексно число. Числото $\overline{\alpha}=a-bi$ се нарича *комплексно спрегнато* на α. Ясно е, че $\operatorname{Re}\alpha=\operatorname{Re}\overline{\alpha}$ и $\operatorname{Im}\overline{\alpha}=-\operatorname{Im}\alpha$. Непосредствено се проверяват следните равенства

$$\overline{\alpha+\beta}=\overline{\alpha}+\overline{\beta},\qquad\qquad \overline{\alpha\beta}=\overline{\alpha}.\overline{\beta}.$$

С метода на пълната математическа индукция може да се докаже, че

$$\overline{\alpha_1+\alpha_2+...+\alpha_k}=\overline{\alpha}_1+\overline{\alpha}_2+...+\overline{\alpha}_k,$$
$$\overline{\alpha_1\alpha_2...\alpha_k}=\overline{\alpha}_1.\overline{\alpha}_2...\overline{\alpha}_k,$$

където k е произволно естествено число и $\alpha_i\in\mathbf{C}$. Изпълнени са също така и равенствата:

$$\alpha+\overline{\alpha}=2\operatorname{Re}\alpha,\qquad\qquad \alpha.\overline{\alpha}=a^2+b^2\in\mathbf{R},$$

т.е. сумата и произведението на две комплексно спрегнати числа са реални числа. Очевидно е, че числото α е реално тогава и само тогава, когато $\alpha=\overline{\alpha}$.

6. Определение. Числото $|\alpha| = \sqrt{\alpha\overline{\alpha}} = \sqrt{a^2+b^2}$ ще наричаме *модул* на комплексното число α.

Модулът на едно комплексно число α е неотрицателно реално число и $|\alpha| = 0$ точно тогава, когато $\alpha = 0$. Ясно е, че ако α е реално, то модулът му съвпада с абсолютната стойност на реалното число α.

Непосредствено се проверява, а след това се обобщава, че

$$|\alpha\beta| = |\alpha||\beta|, \qquad |\alpha_1\alpha_2\ldots\alpha_k| = |\alpha_1||\alpha_2|\ldots|\alpha_k|.$$

В частност

$$|\alpha^n| = |\alpha|^n,\ n \in \mathrm{Z}.$$

7. С помощта на операциите събиране и умножение на комплексни числа можем да въведем обратните им операции изваждане и деление, като числата ще бъдат представени в алгебричен вид. Нека $\alpha = a + bi$ и $\beta = c + di$. От т. 2 е ясно, че $-\beta = -c - di$. Тогава под *разлика* на числата α и β ще разбираме сумата $\alpha + (-\beta)$ и ще пишем $\alpha - \beta$, т.е.

$$\alpha - \beta = \alpha + (-\beta) = (a + bi) + (-c - di) = (a - c) + (b - d)i.$$

Ако $\alpha \neq 0$, да разгледаме числото $\dfrac{\overline{\alpha}}{|\alpha|^2} = \dfrac{1}{a^2+b^2}\overline{\alpha}$.

Тогава

$$\alpha.\frac{\overline{\alpha}}{a^2+b^2} = \frac{\alpha.\overline{\alpha}}{a^2+b^2} = \frac{a^2+b^2}{a^2+b^2} = 1.$$

Числото $\dfrac{1}{a^2+b^2}\overline{\alpha} = \dfrac{1}{|\alpha|^2}\overline{\alpha}$ ще наричаме *обратно* (*реципрочно*) на α и ще го бележим с α^{-1}. В такъв случай под частно на комплексните числа β и α, $\alpha \neq 0$, ще разбираме числото $\beta\alpha^{-1}$ и ще пишем β / α:

$$\beta\alpha^{-1} = \frac{\beta}{\alpha}.$$

Ако $\alpha = a + bi$, $\beta = c + di$, то

$$\frac{\beta}{\alpha} = \beta\alpha^{-1} = \frac{1}{|\alpha|^2}\beta\overline{\alpha} = \frac{1}{a^2+b^2}\Big[(c+di)(a-bi)\Big] =$$

$$= \frac{1}{a^2+b^2}\Big[(ca+bd)+(ad-bc)\Big] = \frac{ac+bd}{a^2+b^2} + \frac{ad-bc}{a^2+b^2}i .$$

Същият резултат се получава, ако числителят и знаменателят на дробта β / α, $\alpha \neq 0$, се умножат с комплексно спрегнатото число на знаменателя, т.е

$$\frac{\beta}{\alpha} = \frac{\beta\overline{\alpha}}{\alpha.\overline{\alpha}} = \frac{\beta\overline{\alpha}}{|\alpha|^2} = \frac{ac+bd}{a^2+b^2} + \frac{ad-bc}{a^2+b^2}i.$$

Непосредствено се проверява, че

$$\left|\alpha^{-1}\right| = |\alpha|^{-1}, \qquad \left|\frac{\alpha}{\beta}\right| = \frac{|\alpha|}{|\beta|}.$$

8. Нека в равнината въведем правоъгълна координатна система Oxy . Тогава на всяко комплексно число $\alpha = (a,b) = a+bi$ съответства единствена точка от равнината с координати a и b . Обратно, на всяка точка от равнината с абсциса a и ордината b съответства комплексното число $\alpha = a+bi$. Така получаваме взаимно еднозначно съответствие между множеството на комплексните числа и точките в равнината. Блгодарение на това съответствие е удобно да отъждествим тези две множества, т.е. да отъждествим всяко комплексно число α със съответната му точка от равнината и ще я означаваме със същата буква α. (По подобен начин отъждествявахме реалните числа с точките от правата.) При това съответствие реалните числа се изобразяват върху абсцисната ос, която ще наричаме *реална*, а чисто имагинерните - върху ординатната ос, която ще наричаме *имагинерна*.

Нека $\alpha = a+bi$, $\beta = c+di$. Тогава на числото $\gamma = \alpha+\beta = (a+c)+(b+d)i$ съответства точката $\gamma(a+c,b+d)$, т.е. изпълнено е векторното равенство

$$\overrightarrow{O\alpha} + \overrightarrow{O\beta} = \overrightarrow{O\gamma} .$$

Последното показва, че събирането на комплексни числа е съгласувано със събирането на вектори. Вектора $\overrightarrow{O\alpha}$ ще наричаме *радиус-вектор* на комплексното число α.

Нека $\alpha = a + bi$ и φ е ъгълът, който векторът $\overrightarrow{O\alpha}$ сключва с положителната посока на реалната ос Ox. Нека $|\alpha| = r = \sqrt{a^2 + b^2} =$ $=|\overrightarrow{O\alpha}|$. Ясно е, че

$$\cos\varphi = \frac{a}{r}, \qquad \sin\varphi = \frac{b}{r},$$

$$a = r\cos\varphi, \qquad b = r\sin\varphi.$$

Тогава

$$\alpha = a + bi = r\cos\varphi + ir\sin\varphi = r(\cos\varphi + i\sin\varphi).$$

Най-дясната страна в последните равенства ще наричаме *тригонометричен вид* на α. Ъгълът φ ще наричаме *аргумент* на α и ще го бележим с $\arg\alpha$. Удобно е да се счита, че аргументът е определен само с точност до кратно на 2π, т.е.

$$\arg\alpha = \varphi + 2k\pi, \qquad k = 0, \pm 1, \pm 2, \ldots$$

Така всяко комплексно число се определя еднозначно от модула си (дължината на радиус-вектора си) и аргумента си. Единствено числото 0 се определя само от модула си; ще считаме, че негов аргумент не е дефиниран.

Изобщо в сила е следното свойство: комплексните числа $\alpha = r(\cos\varphi + i\sin\varphi)$ и $\beta = r_1(\cos\varphi_1 + i\sin\varphi_1)$ са равни тогава и само тогава, когато $r = r_1$ и $\varphi = \varphi_1 + 2k\pi$, т.е. $\varphi - \varphi_1 = 2k\pi$.

9. Нека $\alpha = r(\cos\varphi + i\sin\varphi)$ и $\beta = r_1(\cos\varphi_1 + i\sin\varphi_1)$. Тогава

$$\alpha\beta = rr_1(\cos\varphi + i\sin\varphi)(\cos\varphi_1 + i\sin\varphi_1) =$$

$$= rr_1\left[(\cos\varphi\cos\varphi_1 - \sin\varphi\sin\varphi_1) + i(\sin\varphi\cos\varphi_1 + \cos\varphi\sin\varphi_1)\right] =$$

$$= rr_1\left[\cos(\varphi + \varphi_1) + i\sin(\varphi + \varphi_1)\right],$$

т.е. модулът на произведението е равен на произведението от модулите, а аргументът е равен на сумата от аргументите, или в друг запис

$$|\alpha\beta| = |\alpha||\beta|, \qquad \arg(\alpha\beta) = \arg\alpha + \arg\beta .$$

Като се извърши делението по метода от т. 7, за частното α / β получаваме

$$\frac{\alpha}{\beta} = \frac{r}{r_1}\left[\cos(\varphi - \varphi_1) + i\sin(\varphi - \varphi_1)\right],$$

$$\left|\frac{\alpha}{\beta}\right| = \frac{|\alpha|}{|\beta|}, \quad \arg\left(\frac{\alpha}{\beta}\right) = \arg\alpha - \arg\beta .$$

При горните означения имаме $\alpha^2 = r^2(\cos 2\varphi + i\sin 2\varphi)$. С индукция лесно получаваме равенството

$$\alpha^n = r^n(\cos n\varphi + i\sin n\varphi), \qquad n = 1, 2, 3, \ldots,$$

известно като *формула на Моавър за степенуване* на комплексни числа.

10. Теорема. *Ако* α *е различно от нула комплексно число, то съществуват точно* n *различни комплексни числа* β*, за които* $\beta^n = \alpha$*. Всяко от тях ще бележим с* $\sqrt[n]{\alpha}$ *и ще го наричаме* n*-ти корен на* α*. Ако* $\alpha = r(\cos\varphi + i\sin\varphi)$*, всички стойности на* $\sqrt[n]{\alpha}$ *се дават по формулата*

$$\beta_k = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i\sin\frac{\varphi + 2k\pi}{n}\right),$$

където $k = 0, 1, \ldots, n-1$, а $\sqrt[n]{r}$ *е аритметичен корен от положителното число* r.

Доказателство. Нека $\beta = r_1(\cos\psi + i\sin\psi)$ е такова комплексно число, че $\beta^n = \alpha$. Следователно

$$r_1^n(\cos n\psi + i\sin n\psi) = r(\cos\varphi + i\sin\varphi).$$

Според свойството от т. 8 следва, че

$$r = r_1^n, \qquad n\psi - \varphi = 2s\pi, \quad s \in \mathbf{Z},$$

откъдето

$$r_1 = \sqrt[n]{r}, \quad \psi = \frac{\varphi + 2s\pi}{n}.$$

Да разделим s на n с остатък: $s = nq + k$, $0 \le k < n$. Тогава

$$\psi = \frac{\varphi + 2s\pi}{n} = \frac{\varphi + 2k\pi}{n} + 2q\pi, \quad q \in \mathbf{Z}.$$

В такъв случай

$$\beta = \sqrt[n]{r}\left[\cos\left(\frac{\varphi + 2k\pi}{n} + 2q\pi\right) + i\sin\left(\frac{\varphi + 2k\pi}{n} + 2q\pi\right)\right] =$$

$$= \sqrt[n]{r}\left[\cos\frac{\varphi + 2k\pi}{n} + \sin\frac{\varphi + 2k\pi}{n}\right] = \beta_k.$$

И така, всяко число β, за което $\beta^n = \alpha$, съвпада с някое от числата β_k. От формулата на Моавър за степенуване веднага следва, че $\beta_k^n = \alpha$ при $k = 0, 1, \ldots, n-1$. Най-сетне числата β_k са различни, защото разликата на аргументите на β_k и β_l е $2\pi(k-l)/n$; тя не е кратна на 2π, тъй като $0 < k - l < n$ и $k - l$ не се дели на n. С това теоремата е доказана.

Формулата за намиране на n-ти корен на комплексно число е известна като *формула на Моавър за кореннуване*.

С л е д с т в и е. *Всички n-ти корени на* 1 *са комплексните числа*

$$\varepsilon_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \qquad k = 0,1,\ldots,n-1.$$

Геометрично те са върховете на правилен n-ъгълник, вписан в окръжност с център нулата и радиус 1, като един от върховете е числото 1.

Ще казваме, че един n-ти корен на единицата е *примитивен*, ако той не е корен на единицата от по-ниска степен, т.е. $\varepsilon^n = 1$, но $\varepsilon^m \neq 1$ за всяко m, за което $0 < m < n$. Например $\varepsilon_1 = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$ е примитивен n-ти корен на единицата. Не е трудно да се докаже, че ε_k е примитивен тогава и само тогава, когато числото k е взаимно просто с n.

УПРАЖНЕНИЯ

1. Докажете неравенствата

а) $|\alpha + \beta| \leq |\alpha| + |\beta|, \qquad |\alpha| - |\beta| \leq |\alpha + \beta|$;

б) $|\alpha_1 + \alpha_2 + ... + \alpha_n| \leq |\alpha_1| + |\alpha_2| + ... + |\alpha_n|$,

където $\alpha, \beta, \alpha_1, ..., \alpha_n \in \mathbf{C}$.

2. Докажете, че ако β е една стойност на n-я корен от комплексното число α, то всички стойности на корена са $\varepsilon_k\beta$, $k = 0,1,\ldots,n-1$, където ε_k са n-тите корени на единицата.

3. Докажете, че ако ε е примитивен n-ти корен на единицата, то неговите степени $\varepsilon^0 = 1, \varepsilon^1, \varepsilon^2, \ldots, \varepsilon^{n-1}$ са всичките стойности на $\sqrt[n]{1}$.

4. Докажете, че ако ε е примитивен n-ти корен на единицата, равенството $\varepsilon^m = 1$, $m \in \mathbf{Z}$, е изпълнено точно тогава, когато n дели m.

§ 2. Числови полета

1. В множеството $\mathbf{Z}$ на целите числа операциите събиране, изваждане и умножение са винаги изпълними, докато операцията деление на число, различно от нула, не винаги е изпълнима - резултатът може да не е цяло число. В множествата $\mathbf{Q}$, $\mathbf{R}$ и $\mathbf{C}$, съдържащи множеството $\mathbf{Z}$, са изпълними всички споменати операции, включително и делението на число, различно от нула. Съществуват и други множества от числа, в които четирите операции са изпълними.

2. Да разгледаме множеството

$$\mathbf{Q}\left(\sqrt{3}\right)=\left\{a+b\sqrt{3} \mid a,b\in\mathbf{Q}\right\}.$$

Две числа $a_1 + b_1\sqrt{3}$ и $a_2 + b_2\sqrt{3}$ от това множество са равни тогава и само тогава, когато $a_1 = a_2$ и $b_1 = b_2$. Действително, нека $a_1 + b_1\sqrt{3} = a_2 + b_2\sqrt{3}$ и да допуснем за момент, че $b_1 \neq b_2$. Тогава

$$\sqrt{3} = \frac{a_1 - a_2}{b_2 - b_1}$$

и излиза, че $\sqrt{3}$ е рационално число, а знаем, че това не е така. Следователно $b_1 = b_2$, откъдето и $a_1 = a_2$.

Непосредствено се проверява, че сума, разлика и произведение на числа от $\mathbf{Q}\left(\sqrt{3}\right)$ отново е в $\mathbf{Q}\left(\sqrt{3}\right)$. Да видим как стои въпросът с делението на число, различно от нула. Нека $\alpha = a + b\sqrt{3} \neq 0$, т.е. поне едно от числата a,b не е нула, и нека $\beta = c + d\sqrt{3}$ е число от $\mathbf{Q}\left(\sqrt{3}\right)$. Тогава

$$\frac{\beta}{\alpha} = \frac{c + d\sqrt{3}}{a + b\sqrt{3}} = \frac{ac - 3db}{a^2 - 3b^2} + \frac{ad - bc}{a^2 - 3b^2}\sqrt{3}\ .$$

За да сме сигурни, че частното β / α принадлежи на $\mathbf{Q}\left(\sqrt{3}\right)$, трябва да се убедим, че $a^2 - 3b^2 \neq 0$. Да допуснем, че $a^2 - 3b^2 = 0$, но $b \neq 0$. Тогава $3 = a^2 / b^2$ и $\sqrt{3} = |a / b|$ се оказва рационално число, тъй като модулът на рационално число, е рационално. Това обаче е невъзможно. Ако $b = 0$, от $a^2 - 3b^2 = 0$ следва $a = 0$, противно на допускането, че $a + b\sqrt{3} \neq 0$. Следователно частното β / α при $\alpha \neq 0$ винаги принадлежи на $\mathbf{Q}\left(\sqrt{3}\right)$.

Горните свойства на операциите в $\mathbf{Q}\left(\sqrt{3}\right)$ остават в сила и за всяко множество от вида $\mathbf{Q}\left(\sqrt{\alpha}\right)$, където α е цяло и не е точен квадрат на цяло число.

Това дава основание всички тези множества да разглеждаме като примери на едно по-общо понятие, наречено числово поле.

3. Определение. Числовото множество $\mathcal{K} \subset \mathbf{C}$, съдържащо поне два различни елемента, наричаме *числово поле*, ако са изпълнени следните условия: ако $\alpha, \beta \in \mathcal{K}$, то $\alpha \pm \beta$, $\alpha\beta$ и α / β (при $\beta \neq 0$) също принадлежат на $\mathcal{K}$.

Ще докажем следното

4. Твърдение. Всяко числово поле съдържа полето на рационалните числа.

Доказателство. Нека $\mathcal{K}$ е произволно числово поле. Тъй като $\mathcal{K}$ съдържа поне две различни числа, то $\mathcal{K}$ съдържа поне едно число, различно от нула. Нека $a \neq 0$, $a \in \mathcal{K}$. Тогава от определението следва, че $a - a = 0 \in \mathcal{K}$ и $a / a = 1 \in \mathcal{K}$. Но тогава 1+1=2, 2+1=3 и т.н. също са в $\mathcal{K}$, т.е. $\mathcal{K}$ съдържа множеството $\mathbf{N}$ на естествените числа. Разликите 0-1=-1, 0-2=-2 и т.н. също са в $\mathcal{K}$, т.е. $\mathbf{Z} \subset \mathcal{K}$. Най-сетне, ако $a, b \in \mathbf{Z}$ и $b \neq 0$, то $a / b \in \mathcal{K}$, което показва, че всички рационални числа принадлежат на $\mathcal{K}$.

§ 3. Полиноми

1. Определение. Нека $\mathcal{K}$ е числово поле, а f е изображение на $\mathcal{K}$ в $\mathcal{K}$. Ще казваме, че f е *полиномно изображение, дефинирано в* $\mathcal{K}$, или *полином на една променлива с коефициенти от* $\mathcal{K}$, ако съществуват такива елементи $a_0, a_1, \dots, a_n \in \mathcal{K}$, че

$$f\left(x\right) = a_0 + a_1 x + \dots + a_n x^n$$

за всяко x от полето $\mathcal{K}$. Числата $a_0, a_1, \dots, a_n$ се наричат *коефициенти* на полинома f.

Ако $a_n \neq 0$, то a_n се нарича *старши коефициент* на полинома, а n - степен на полинома; записва се $\deg f = n$. За коефициента a_0 се казва, че е *свободен член*. Ако коефициентите $a_1, a_2, \dots, a_n$ са равни на нула, полиномът се нарича *константа* и се пише $f\left(x\right) = a_0$. Ако всички коефициенти на полинома са равни на нула, казва се, че той е *нулевият полином* и се пише

$f=0$ или $f(x)=0$ (второто означение не бива да се смесва с някакво "уравнение"). На нулевия полином приписваме степен $-\infty$, като се договаряме, че символът $-\infty$ (той не е число!) има следните свойства:

(1) $$-\infty+(-\infty)=-\infty, \qquad -\infty+l=-\infty, \qquad -\infty< l$$

за всяко цяло число l.

При записването на даден полином членовете му $a_i x^i$, $i=0,1,\dots,n$, може да се подреждат или по растящите степени на x, или по намаляващите. Прието е множеството на всички полиноми на една променлива x и с коефициенти от полето $\mathcal{K}$ да се бележи с $\mathcal{K}[x]$

2. Операции с полиноми. Нека

$$f(x)=a_0+a_1x+\dots+a_nx^n,$$

$$g(x)=b_0+b_1x+\dots+b_mx^m$$

са два полинома от $\mathcal{K}[x]$. Ако например $n>m$, то можем да запишем $g(x)$ като

$$g(x)=b_0+b_1x+\dots+b_mx^m+0.x^{m+1}+\dots+0.x^n$$

и да считаме, че $b_{m+1}=\dots=b_n=0$. При тази уговорка полиномът

$$(f+g)(x)=(a_0+b_0)+(a_1+b_1)x+\dots+(a_n+b_n)x^n$$

е добре дефиниран полином с коефициенти от $\mathcal{K}$ и се нарича *сума* на полиномите f и g. Обикновено вместо $(f+g)(x)$ се употребява означението $f(x)+g(x)$.

Нека λ е произволно число от полето $\mathcal{K}$. Дефинираме полинома

$$(\lambda f)(x)=\lambda f(x)=(\lambda a_0)+(\lambda a_1)x+\dots+(\lambda a_n)x^n.$$

Той се нарича *произведение* на полинома $f(x)$ и числото λ.

Лесно е да се провери, че множеството $\mathcal{K}[x]$ заедно с дефинираните операции събиране на полиноми и умножаване на полиноми с числа от $\mathcal{K}$ е линейно пространство над полето $\mathcal{K}$.

Полиномът

$$(fg)(x) = f(x)g(x) = c_0 + c_1 x + \ldots + c_{m+n} x^{m+n},$$

където

$$c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0, \qquad k = 0, 1, \ldots, m+n,$$

се нарича *произведение* на полиномите $f(x)$ и $g(x)$. Ясно е, че неговите коефициенти принадлежат на същото поле, на което принадлежат коефициентите на полиномите $f(x)$ и $g(x)$.

Не е трудно да се докаже, че

(2) $$\deg(fg) = \deg f + \deg g$$

за всеки два полинома $f(x)$ и $g(x)$. Наистина, ако поне един от двата полинома е нулевият, то fg е нулевият полином и горното равенство е едно от договорените равенства (1). Ако и двата полинома не са нулеви и $\deg f = n$, $\deg g = m$, т.е. $a_n \neq 0$, $b_m \neq 0$, то $c_{m+n} = a_n b_m$ е различно от нула число, следователно $\deg fg = n + m = \deg f + \deg g$.

3. Твърдение. *Нека* $f(x)$ *и* $\mathrm{g}(\mathrm{x})$ *са два полинома от* $\mathcal{K}[x]$, *за които* $\mathrm{f}(\mathrm{x}) = \mathrm{g}(\mathrm{x})$ *за всяко* x *от числовото поле* $\mathcal{K}$ *и нека*

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n,$$

$$g(x) = b_0 + b_1 x + \ldots + b_n x^n.$$

Тогава $a_i = b_i, \quad i = 0, 1, \ldots, n$.

Доказателство. От равенството $f(0) = g(0)$ следва, че $a_0 = b_0$. Да допуснем, че твърдението не е вярно и нека k е най-голямото естествено число, за което $a_k \neq b_k$. Тогава от $f(x) - g(x) = 0$ следва равенството

(3) $$0 = (a_1 - b_1)x + \ldots + (a_{k-1} - b_{k-1})x^{k-1} + (a_k - b_k)x^k,$$

което е изпълнено за всяко $x \in \mathcal{K}$. Ако $k=1$, получаваме $0=\left(a_1-b_1\right)x$, а тъй като x може да приема и стойности, различни от нула, то $a_1-b_1=0$, противно на допускането. Следователно $k \geq 2$. От (3) при $x \neq 0$ следва

(4) $$a_k-b_k=\frac{b_1-a_1}{x^{k-1}}+\frac{b_2-a_2}{x^{k-2}}+\ldots+\frac{b_{k-1}-a_{k-1}}{x}.$$

Нека $d=\left|a_k-b_k\right|$, $d>0$. Всяко числово поле $\mathcal{K}$ съдържа полето Q на рационалните числа, следователно в $\mathcal{K}$ съществува достатъчно голямо число x_0 (например рационално), за което

$$\left|\frac{b_i-a_i}{x_0^{k-i}}\right|<\frac{d}{k-1}, \qquad i=1,2,\ldots,k-1.$$

Тогава от (4) следва

$$d=\left|a_k-b_k\right|=\left|\sum_{i=1}^{k-1}\frac{b_i-a_i}{x_0^{k-i}}\right| \leq \sum_{i=1}^{k-1}\left|\frac{b_i-a_i}{x_0^{k-i}}\right|<\frac{d}{k-1}.(k-1)=d,$$

т.е. $d<d$. Полученото противоречие доказва твърдението.

4. Теорема (за деление с остатък). *Нека $f(x)$ и $g(x)$ са полиноми с коефициенти от числово поле $\mathcal{K}$ и $g(x)$ не е нулевият полином. Тогава съществуват еднозначно определени полиноми $q(x)$ и $r(x)$ с коефициенти от $\mathcal{K}$, за които:*

а) $f(x)=g(x)q(x)+r(x)$;

б) $\deg r<\deg g$.

Доказателство. Като предполагаме, че полиномите $q(x)$ и $r(x)$ съществуват, ще покажем, че те са единствените, които удовлетворяват условията а) и б). Действително, нека $q_1(x)$ и $r_1(x)$ е още една двойка полиноми от $\mathcal{K}[x]$, за които

$$f(x) = g(x)q_1(x) + r_1(x), \quad \deg r_1 < \deg g.$$

Тогава

$$0 = g(x)\big(q(x) - q_1(x)\big) + r(x) - r_1(x)$$

(в лявата страна стои нулевият полином) и

(5)
$$g(x)\big(q(x) - q_1(x)\big) = r_1(x) - r(x).$$

Ако $q - q_1$ не е нулевият полином, то степента му е неотрицателно число и според (2) степента на полинома в лявата страна на равенството (5) е по-голяма или равна на $\deg g$, докато степента на полинома в дясната страна на (5) е по-малка от $\deg g$. Това противоречи на твърдението от т. 3 (равните полиноми в частност имат равни степени), следователно $q(x) = q_1(x)$ и $r(x) = r_1(x)$.

Сега ще докажем, че исканите полиноми q и r съществуват. За целта нека

$$f(x) = a_n x^n + \ldots + a_0,$$

$$g(x) = b_m x^m + \ldots + b_0,$$

където $b_m \neq 0$. Ще докажем твърдението с индукция по n. Ако $\deg f < m$, полагаме $q = 0$, $r = f$ и заключенията а), б) от теоремата са налице. Нека $n \geq m$ и да допуснем, че заключението на теоремата в частта съществуване е вярно за всички полиноми от степен по-малка от n. Тогава полиномът

$$f_1(x) = f(x) - a_n b_m^{-1} x^{n-m} g(x)$$

очевидно е с коефициенти от полето $\mathcal{K}$, степента му е по-малка от n и според индуктивното допускане $f_1(x) = q_1(x)g(x) + r(x)$ за подходящи полиноми $q_1(x)$ и $r(x)$, като $\deg r < \deg g$. Следователно

$$f(x) = a_n b_m^{-1} x^{n-m} g(x) + f_1(x) =$$

$$= a_n b_m^{-1} x^{n-m} g(x) + q_1(x)g(x) + r(x) =$$

$$= \left(a_n b_m^{-1} x^{n-m} g(x) + q_1(x) \right) g(x) + r(x),$$

с което изразихме $f(x)$ в искания вид и завършихме доказателството.

Полиномът $q(x)$ от теоремата се нарича *непълно частно* (или просто *частно*) от делението на $f(x)$ на $g(x)$, а $r(x)$ се нарича *остатък* от делението. Ако остатъкът е нулевият полином, казва се че полиномът g дели полинома f.

Тук няма да обясняваме общия алгоритъм за намиране на полиномите $q(x)$ и $r(x)$ - той много напомня известния от училище алгоритъм за деление на естествени числа. Ще се ограничим само със специалния случай, когато $g(x) = x - c$ - тогава делението може да се извърши по схема, известна като *правило* на Хорнер, което е по-удобно от общия алгоритъм.

И така, като запазваме горните означения нека $g(x) = x - c$ и

$$f(x) = q(x)(x-c) + r. \tag{6}$$

Тъй като $\deg r < \deg g = 1$, то полиномът r е константа. Ясно е, че ако $\deg f = n \geq 1$, то $\deg q = n-1$. Следователно

$$f(x) = \left(c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \ldots + c_0 \right)(x-c) + r,$$

където c_i, $i = 0, 1, \ldots, n-1$ са коефициентите на търсения полином $q(x)$. Като използваме твърдението от т. 3 и сравним коефициентите пред равните степени на x в двете страни на последното равенство, получаваме:

$$\begin{aligned}
a_n &= c_{n-1} \\
a_{n-1} &= c_{n-2} - cc_{n-1} \\
a_{n-2} &= c_{n-3} - cc_{n-2} \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
a_1 &= c_0 - cc_1 \\
a_0 &= r - cc_0,
\end{aligned}$$

откъдето очевидно $c_{n-1} = a_n$, $c_i = a_{i+1} + cc_{i+1}$, $i = n-2, n-3, \ldots, 0$, $r = a_0 + ca_0$. Удобно е коефициентите на полиномите да се подредят в следната таблица:

	a_n	a_{n-1}	a_{n-2}	$\cdots$	a_1	a_0
c	a_n	c_{n-2}	c_{n-3}	$\cdots$	c_0	r

като в нея $a_n = c_{n-1}$. Сега правилото за пресмятане на коефициента c_i при $i < n-1$ става съвсем нагледно: той е равен на произведението на c и на предхождащия го коефициент c_{i+1} плюс коефициента a_{i+1}, който стои над c_i; същото правило важи и за остатъка r.

Ако в равенството (6) положим $x = c$, ще получим $f(c) = r$, т.е. остатъкът r при делението на полинома $f(x)$ на полинома $x - c$ е равен на стойността на $f(x)$ при $x = c$. Ако за някое комплексно число c имаме $f(c) = 0$, казва се, че c е *корен* или *нула* на полинома $f(x)$. От казаното веднага получаваме следната проста, но полезна

5. Теорема (Безу). *Комплексното число c е корен на полинома $f(x)$ (с числови коефициенти) тогава и само тогава, когато съществува полином $q(x)$, за който*

$$f(x) = (x-c)q(x).$$

Доказателството на следващото важно твърдение се излага в курсовете по висша алгебра.

6. Теорема (Д'Аламбер). *За всеки полином с числови коефициенти, който не е константа, съществува поне едно комплексно число , което е корен на полинома.*

Ще изведем едно следствие от теоремата, което е полезно за приложенията. За целта нека $f(x)$ е произволен полином с числови коефициенти (можем да считаме, че $f(x) \in \mathbf{C}[x]$) и нека $\deg f = n \geq 1$. Нека $\alpha_1 \in \mathbf{C}$ е корен на полинома. Тогава по теоремата на Безу имаме

$$f(x) = (x-\alpha_1)q_1(x),$$

където $q_1(x)$ е полином с числови коефициенти. Ако $n=1$, то $q_1(x)$ е константа. Нека $n>1$. Тогава полиномът $q_1(x)$ има степен $n-1$ и не е константа. По теоремата на Д'Аламбер съществува комплексно число α_2, което е корен на $q_1(x)$ и

$$q_1(x)=(x-\alpha_2)q_2(x).$$

Повтаряйки това разсъждение, ще получим

$$q_2(x)=(x-\alpha_3)q_3(x)$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$q_{n-1}(x)=(x-\alpha_n)q_n(x),$$

където комплексните числа $\alpha_2,\dots,\alpha_n$ са корени съответно на полиномите $q_1(x),\dots,q_{n-1}(x)$, а полиномът $q_n(x)$ е константата a. Оттук лесно следва, че

(7) $$f(x)=a(x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n).$$

Старшият коефициент на полинома в дясната страна е a и според твърдението в т. 3 той е равен на старшия коефициент на полинома $f(x)$. За равенствата (7) се казва, че полиномът $f(x)$ е разложен на линейни множители, а $\alpha_1,\alpha_2,\dots,\alpha_n$ са корените му. Нека при подходяща номерация на корените предположим, че $\alpha_1,\alpha_2,\dots,\alpha_r$, $1\le r\le n$, са различните корени на полинома. Тогава равенството (7) може да се запише и като

(8) $$f(x)=a(x-\alpha_1)^{k_1}(x-\alpha_2)^{k_2}\dots(x-\alpha_r)^{k_r}.$$

Показателят k_i се нарича *кратност* на корена α_i. Ще подчертаем, че $k_1+k_2+\dots+k_r=n=\deg f$. С това доказахме

7. Следствие. *В полето на комплексните числа всеки полином $f(x)$ с числови коефициенти и положителна степен има разлагане от вида* (7) (*разлага се на линейни множители*). *Броят на корените на полинома е равен на степента му, ако всеки корен се брои толкова пъти, колкото е кратността му.*

Азбучен указател

доц. Никола Т. Петров
доц. Никола П. Зяпков

Българска
Първо издание 2000
Допечатка 2004
Допечатка 2010

Редактор: **проф. д-р. Керопе Чакърян**

Рецензенти: **проф.д.м.н. Грозьо Станилов**
проф. д.м.н. Недялко Ненов

Художник на корицата: **Деян Евтимов**

Формат 60/84/16
Печатни коли 16,25
Тираж 500
ISBN 954-577-079-1