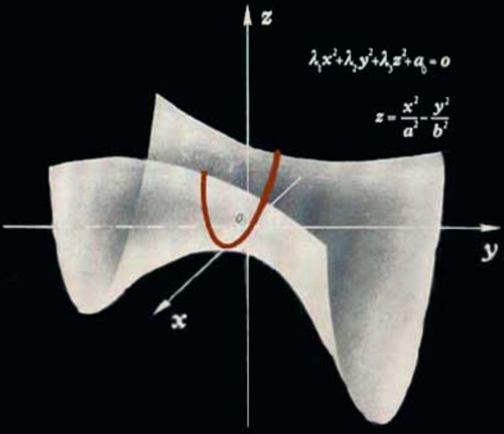
Algebra lineal

V. V. Voevodin



EDITORIAL MIR MOSCÚ

ИЗДАТЕЛЬСТВО · МИР ·

В. В. ВОЕВОДИН

ЛИНЕЙНАЯ АЛГЕБРА

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

V. V. Voevodin

Algebra lineal

Traducido al español
por el ingeniero
K. P. MEDKOV

Primera reimpresión

EDITORIAL MIR
MOSCÚ

Impreso en la URSS

Primera edición 1982

Primera reimpresión 1986

На испанском языке

INDICE

Prólogo 8

PARTE I. ESPACIOS LINEALES 9

Capítulo 1. Conjuntos, elementos, operaciones 9

§ 1. Conjuntos y elementos 9
§ 2. Operación algebraica 12
§ 3. Operación inversa 15
§ 4. Relación de equivalencia 18
§ 5. Segmentos dirigidos 21
§ 6. Adición de los segmentos dirigidos 24
§ 7. Grupos 27
§ 8. Anillos y campos 31
§ 9. Multiplicación del segmento dirigido por un número 35
§ 10. Espacios lineales 38
§ 11. Sumas finitas y productos finitos 43
§ 12. Cálculos aproximados 45

Capítulo 2. Estructura del espacio lineal 48

§ 13. Combinaciones lineales y cápsulas 48
§ 14. Dependencia lineal 50
§ 15. Sistemas equivalentes de vectores 53
§ 16. Base 57
§ 17. Ejemplos sencillos de los espacios lineales 59
§ 18. Espacios lineales de segmentos dirigidos 61
§ 19. Suma e intersección de los subespacios 64
§ 20. Suma directa de los subespacios 68
§ 21. Isomorfismo de los espacios lineales 70
§ 22. Dependencia lineal y sistemas de las ecuaciones lineales 74

Capítulo 3. Mediciones en el espacio lineal 80

§ 23. Sistemas afines de coordenadas 80
§ 24. Otros sistemas de coordenadas 85
§ 25. Problemas 88
§ 26. Producto escalar 95
§ 27. Espacio euclídeo 98
§ 28. Ortogonalidad 102
§ 29. Longitudes, ángulos, distancias 106
§ 30. Línea oblicua, perpendicular, proyección 109
§ 31. Isomorfismo euclídeo 113
§ 32. Espacio unitario 114
§ 33. Dependencia lineal y sistemas ortonormalizados 116

Capítulo 4. Volumen del sistema de vectores en un espacio lineal 118

§ 34. Productos vectorial y mixto 118
§ 35. Volumen y volumen orientado del sistema de vectores 123
§ 36. Propiedades geométricas y algebraicas del volumen 126
§ 37. Propiedades algebraicas del volumen orientado 131
§ 38. Permutaciones 133
§ 39. Existencia de un volumen orientado 136
§ 40. Determinantes 138
§ 41. Dependencia lineal y determinantes 143
§ 42. Cálculo de los determinantes 146

Capítulo 5. Línea recta y plano en el espacio lineal 148

§ 43. Ecuaciones de la línea recta y del plano 148
§ 44. Disposición conjunta 154
§ 45. Plano en el espacio lineal 158
§ 46. Recta e hiperplano 161
§ 47. Semiespacio 167
§ 48. Sistemas de ecuaciones lineales 169

Capítulo 6. Límite en el espacio lineal 175

§ 49. Espacio métrico 175
§ 50. Espacio completo 178
§ 51. Desigualdades auxiliares 181
§ 52. Espacio normalizado 183
§ 53. Convergencia en norma y convergencia coordenada 185
§ 54. Completitud de los espacios normalizados 189
§ 55. Límite y procesos de cálculo 191

PARTE II. OPERADORES LINEALES 194

Capítulo 7. Matrices y operadores lineales 194

§ 56. Operadores 194
§ 57. Espacio lineal de los operadores 198
§ 58. Anillo de los operadores 200
§ 59. Grupo de operadores regulares 202
§ 60. Matriz del operador 206
§ 61. Operaciones sobre las matrices 209
§ 62. Matrices y determinantes 214
§ 63. Paso a la otra base 217
§ 64. Matrices equivalentes y matrices semejantes 220

Capítulo 8. Polinomio característico 223

§ 65. Valores propios y vectores propios 223
§ 66. Polinomio característico 226
§ 67. Anillo de polinomios 228
§ 68. Teorema fundamental del álgebra 232
§ 69. Corolarios del teorema fundamental 237

Capítulo 9. Estructura del operador lineal 243

§ 70. Subespacios invariantes 243
§ 71. Polinomio operacional 246
§ 72. Forma triangular 249
§ 73. Suma directa de los operadores 250

§ 74. Forma de Jordan 254
§ 75. Operador conjugado 259
§ 76. Operador normal 264
§ 77. Operadores unitario y hermitiano 266
§ 78. Operadores A^*A y AA^* 270
§ 79. Descomposiciones de un operador arbitrario 273
§ 80. Operadores en el espacio real 275
§ 81. Matrices de tipo especial 279

Capítulo 10. Propiedades métricas del operador 281

§ 82. Continuidad y acotación del operador 281
§ 83. Norma del operador 283
§ 84. Normas matriciales del operador 287
§ 85. Ecuaciones operacionales 290
§ 86. Seudosoluciones y un operador seudoinverso 292
§ 87. Perturbación y regularidad del operador 296
§ 88. Solución estable de las ecuaciones 300
§ 89. La perturbación y los valores propios 308

PARTE III. FORMAS BILINEALES 310

Capítulo 11. Formas bilineales y cuadráticas 310

§ 90. Propiedades generales de las formas bilineales y cuadráticas 310
§ 91. Matrices de las formas bilineales y cuadráticas 317
§ 92. Reducción a una forma canónica 324
§ 93. Congruencia y descomposiciones matriciales 332
§ 94. Formas bilineales simétricas 338
§ 95. Hipersuperficies de segundo grado 346
§ 96. Líneas de segundo orden 352
§ 97. Superficies de segundo grado 360

Capítulo 12. Espacios bilineales métricos 366

§ 98. Matriz y determinante de Gram 366
§ 99. Subespacios regulares 373
§ 100. Ortogonalidad en las bases 377
§ 101. Operadores y formas bilineales 385
§ 102. Isomorfismo bilineal métrico 390

Capítulo 13. Formas bilineales en los procesos de cálculo 393

§ 103. Procesos de ortogonalización 393
§ 104. Ortogonalización de una sucesión de potencias 399
§ 105. Métodos de direcciones conjugadas 404
§ 106. Variantes principales 411
§ 107. Ecuaciones operacionales y seudodualidad 414

Conclusión 420

Índice alfabético de materias 423

PRÓLOGO

El presente manual reúne en sí el curso ampliado del álgebra lineal y de la geometría analítica.

Unas cuantas palabras sobre este libro. Está destinado, principalmente, para aquellos que han elegido la matemática de cálculo como una disciplina esencial de su preparación. La enseñanza concerniente a dicha especialidad está ligada en muchos aspectos con los cursos tradicionales de las matemáticas. No obstante, los cambios, tanto en la metodología de exposición como en el contenido del curso, resultan imperativos.

Ya en las primeras conferencias los estudiantes empiezan a entrar en contacto con los elementos del álgebra lineal y de la geometría analítica. Estas primeras conferencias sirven también de origen de la formación de su concepción científica. Por esta razón, la asimilación futura por los estudiantes de toda la matemática de cálculo depende muchísimo del carácter y sucesión de la exposición del curso dado.

Existe toda una serie de buenos libros referentes al álgebra lineal y a la geometría analítica. Sin embargo, su empleo directo para los fines de enseñanza resulta algo embarazoso. La causa principal de ello radica, a nuestro parecer, en que los futuros especialistas en cálculos deben saber más datos del álgebra lineal que los comunicados corrientemente en los libros de texto. Los estudiantes que se especializan en matemática de cálculo no sólo deben recibir una exposición estricta y sistemática de las bases del álgebra y geometría, sino que ya en el primer año deben tener contacto con la enorme riqueza de conocimientos que ha acumulado la matemática de cálculo.

El contacto con los problemas de cálculo permite acentuar con toda eficacia, en interés de la matemática de cálculo, todas las cuestiones de importancia que se estudian en el curso de conferencias, y establecer estrecha relación entre la teoría y los métodos numéricos del álgebra lineal. Como material de partida, para ello, sirven los hechos más sencillos de tales apartados como errores del redondeo, inestabilidad a las perturbaciones de muchos conceptos fundamentales del álgebra lineal, estabilidad de los sistemas ortonormalizados, espacios métricos y normalizados, descomposición singular, formas bilineales y su relación con los procesos de cálculo, etc.

Desde luego, la inclusión del material nuevo y suficientemente amplio es imposible sin una reconstrucción sustancial del curso tradicional. En el presente libro se hizo el intento de tal reconstrucción.

V. Voevodin

PARTE I ESPACIOS LINEALES

CAPÍTULO 1 CONJUNTOS, ELEMENTOS, OPERACIONES

§ 1. Conjuntos y elementos

En cualquier dominio de actividad siempre nos encontramos con la necesidad de considerar diversas totalidades de objetos unidos entre sí mediante un criterio común.

Así, por ejemplo, al estudiar la estructura de tal o cual mecanismo, podemos considerar la totalidad de todas las piezas del mismo. En este caso, como objeto suelto de la totalidad dada puede servir toda pieza del mecanismo, mientras que el criterio que une dichos objetos consiste en el hecho de que todos ellos pertenecen a un mecanismo bien determinado.

Hablando de la totalidad de puntos de una circunferencia en un plano, tenemos en cuenta, en esencia, los objetos —puntos del plano— unidos por la propiedad de que todos ellos son equidistantes respecto de cierto punto fijado.

Una totalidad de objetos unidos entre sí mediante un criterio común suele llamarse en las matemáticas *conjunto* y los propios objetos, *elementos* del conjunto. Al concepto de conjunto no se puede atribuir una definición rigurosa. Por supuesto, podemos decir (¡cómo lo hemos hecho!) que el conjunto es "una totalidad", "un sistema", "una clase", etc. Sin embargo, todo esto sería más bien el uso formal de la riqueza del vocabulario de la lengua.

Con el fin de determinar un concepto, es necesario, ante todo, indicar de qué modo dicho concepto está ligado con las nociones más generales. Resulta imposible hacerlo para el concepto de conjunto, puesto que en las matemáticas no existe para el conjunto un concepto más general. Por esta razón, en lugar de determinar el concepto de conjunto nos vemos obligados a ilustrarlo con unos ejemplos.

Uno de los métodos más sencillos de describir un conjunto radica en que se da la lista completa de los elementos que integran el conjunto. Por ejemplo, el conjunto de todos los libros de una biblioteca, accesibles al lector, está completamente determinado mediante las listas en los catálogos de biblioteca; el conjunto de todos los precios de mercancías está completamente determinado por la lista de precios, etc. No obstante, dicho método sólo es aplicable a los conjuntos *finitos*, es decir, a los conjuntos que contienen un número finito de elementos. Los conjuntos *infinitos*, es decir, aquellos que contienen un número infinito de elementos, no pueden ser determinados con la ayuda de una lista. ¿Cómo, por ejemplo, podríamos alistar todos los números reales?

En aquellos casos en que un conjunto no puede ser prefijado con ayuda de una lista o no es cómodo hacerlo, éste se da por indicación de la propiedad *característica*, es decir, por indicación de tal propiedad que poseen los elementos del conjunto, y sólo ellos. Por ejemplo, en los problemas en que se determinan lugares geométricos la propiedad característica del conjunto de puntos, que sirve de solución para el propio problema, no es otra cosa que una totalidad de las condiciones que han de ser satisfechas por dichos puntos de acuerdo con los requisitos del problema.

La descripción de un conjunto puede ser muy sencilla y no despertar dificultades algunas. Por ejemplo, si hablamos de un conjunto que consiste en dos números, 1 y 2, está claro que ni el número 3 ni un cuaderno ni un automóvil pueden integrar dicho conjunto. En el caso general, sin embargo, la definición de los conjuntos por medio de sus propiedades características conduce, algunas veces, a ciertas complicaciones. Existen varias razones por las cuales surgen las complicaciones citadas.

Una de ellas puede consistir en que los conceptos empleados para la descripción de un conjunto no están suficientemente determinados. Convengamos, por ejemplo, en considerar el conjunto de todos los planetas en el sistema solar. ¿De qué se trata? Se conocen nueve planetas grandes. Pero alrededor del Sol giran también más de un mil de planetas pequeños, o los asteroides. Los diámetros de algunos de estos planetas miden centenares de kilómetros y hay planetas cuyos diámetros no superan 1 km. A medida que se perfeccionen los métodos de observación irán descubriéndose unos planetas cada vez más pequeños y, al fin y al cabo, surgirá la pregunta ¿cuáles de ellos son planetas pequeños y cuáles representan meteoritos y polvo cósmico?

No siempre las dificultades en determinar la composición de un conjunto dependen sólo de las razones indicadas. Ocurre a veces que los conjuntos que a primera vista parecen estar bien determinados resultan, de hecho, determinados muy mal o incluso no están determinados del todo. Supongamos, por ejemplo, que un conjunto consta sólo de un número. Definamos este número como "*un número entero mínimo que no puede ser determinado con la ayuda de una frase que tenga menos de cien palabras rusas*". Consideraremos que se utilizan en el experimento sólo las palabras sacadas de cierto diccionario, como también sus formas gramaticales, y que, además, en el diccionario están contenidas palabras del tipo "uno", "dos", etc.

Hemos de notar que, por un lado, tal número no debe existir, pues se define por una frase, destacada más arriba con cursiva, que contiene menos de cien palabras, y por el sentido de esta frase el número no puede ser definido de la manera semejante. Pero, por otro lado, como el número de las palabras rusas utilizadas es finito, quiere decir que hay números que no pueden ser definidos por una frase que tiene menos de cien palabras y, por consiguiente, entre dichos números existe uno que es mínimo.

En la rama de las matemáticas, llamada *teoría de conjuntos*, se han acumulado muchos ejemplos de que la definición de un conjunto entraña contradicciones interiores. El estudio del problema bajo qué condiciones esto puede tener lugar, ha conducido a las investigaciones profundas en el dominio de la lógica. No obstante, dejamos al lado estas investigaciones. En lo que sigue siempre supondremos que se consideran solamente los conjuntos que están precisamente definidos sin contradicciones algunas y cuya composición no despierta ninguna duda.

Como regla, los conjuntos se designarán con las letras latinas mayúsculas A, B, . . ., y sus elementos, con las letras minúsculas a, b, ... Escribiremos $x \in A$, si el elemento x pertenece al conjunto A, y $x \notin A$, si el elemento x no pertenece al conjunto A.

A veces será introducido en la consideración el así llamado conjunto *vacío*, es decir, un conjunto que no contiene ni un solo elemento. El uso del conjunto vacío resulta cómodo en el caso cuando se desconoce de antemano si existe aunque sea un sólo elemento en la totalidad que se considera.

Ejercicios.

1. Constrúyanse los ejemplos de conjuntos finitos e infinitos. ¿Qué propiedades son para ellos características?

2. Constrúyanse los ejemplos de conjuntos cuya descripción contiene una contradicción.

3. ¿Será vacío el conjunto de raíces reales del polinomio $x^4 + 4x^3 + 7x^2 + 4x + 1$?

4. Constrúyanse los ejemplos de conjuntos cuyos elementos son también conjuntos.

5. Constrúyanse los ejemplos de conjuntos que contienen a sí mismos a título de uno de sus elementos.

§ 2. Operación algebraica

Entre toda clase de conjuntos se pueden distinguir aquellos cuyos elementos admiten que se ejecuten con ellos ciertas operaciones. Supongamos, por ejemplo, que el objeto de nuestra consideración es un conjunto de todos los números reales. En este caso para todo elemento de dicho conjunto resultan definidas tales operaciones como el cálculo del módulo del elemento, el cálculo del seno del último; para cada par de elementos están definidas las operaciones de su adición y multiplicación.

En el ejemplo considerado prestaremos una atención especial a las siguientes peculiaridades de las operaciones mencionadas. En primer lugar, el carácter determinado de todas las operaciones para *cualesquiera* elementos del conjunto dado; en segundo, *la univocidad* de todas las operaciones y, por fin, la pertenencia del resultado, que se obtiene al realizar cualquiera de las operaciones, a los elementos *del mismo* conjunto. Tal situación no siempre tiene lugar.

Una operación puede ser definida no para todos los elementos del conjunto; por ejemplo, el cálculo de un logaritmo no está definido para los números negativos. La extracción de una raíz cuadrada de los números positivos está definida, mas de manera no unívoca. No obstante, incluso si una operación está unívocamente definida para todos los elementos, el resultado de su realización puede no constituir un elemento del conjunto dado. Consideraremos una operación de división en el conjunto de los números positivos enteros. Está claro que para cualesquiera dos números de este conjunto la operación de división es realizable, pero el resultado de su ejecución no será necesariamente un número entero.

Sea dado un conjunto A que contiene al menos un solo elemento. Diremos que en el conjunto A está definida *una operación algebraica*, si se indica una ley según la cual a todo par de elementos, a y b, tomados de dicho conjunto en el orden determinado, se le pone en correspondencia de manera unívoca el tercer elemento c que también pertenece al mismo conjunto.

Esta operación puede llamarse adición y en tal caso c será la suma de los elementos a y b, lo que se designará mediante el símbolo $c = a + b$; esta misma operación puede denominarse multiplicación y entonces c, se llamará producto de los elementos a y b, lo que se designa con el símbolo $c = ab$.

En general, la terminología y la notación para la operación definida en el conjunto A no desempeñarán en adelante algún papel de importancia. Como regla, haremos uso de la notación de una suma o de un producto independientemente de cómo la operación está definida en realidad. Si, en cambio, nos hace falta subrayar ciertas propiedades generales de una operación algebraica, designaremos la operación con el símbolo $*$.

Recurriendo a unos ejemplos sencillos, veamos qué peculiaridades puede tener una operación algebraica. Supongamos que el conjunto A representa en sí una totalidad de todos los números racionales positivos. Introduzcamos para los elementos de este conjunto las operaciones ordinarias de multiplicación y división de los números y hagamos uso de la notación generalmente aceptada. No es difícil comprobar que ambas operaciones en el conjunto A son algebraicas. Sin embargo, si en la operación de multiplicación $ab = ba$ para todos los elementos de A, es decir, si la indicación del orden de los elementos no es esencial, para la operación de división, por el contrario, el orden de los elementos resulta muy esencial, pues la igualdad $a : b = b : a$ se verifica sólo en el caso en que $a = b$. De este modo, aunque la operación algebraica está definida para un par ordenado de elementos, la ordenación de los últimos no es, a veces, esencial.

Una operación algebraica se denomina *conmutativa*, si el resultado de su aplicación no depende del orden en que se eligen los elementos, es decir, si para cualesquiera dos elementos a y b del conjunto dado se verifica la igualdad $a * b = b * a$. Es evidente que entre las operaciones con los números universalmente admitidas la adición y la multiplicación son operaciones conmutativas, mientras que la sustracción y la división, operaciones no conmutativas.

Supongamos ahora que se consideran tres elementos arbitrarios a, b, c. Surge naturalmente la pregunta ¿qué sentido se atribuirá a la expresión $a * b * c$? ¿Cómo se debe aplicar una operación algebraica definida para dos elementos, a los tres elementos?

Puesto que una operación algebraica puede aplicarse solamente a un par de elementos, podemos atribuir un sentido determinado a la expresión $a * b * c$, encerrando entre paréntesis o bien dos primeros elementos, o bien dos elementos últimos. En el primer caso la expresión tomará la forma $(a * b) * c$; en el segundo caso, $a * (b * c)$. Examinemos los elementos $d = a * b$ y $e = b * c$. Debido a que ellos son elementos del conjunto inicial, las expresiones $(a * b) * c$ y $a * (b * c)$ pueden considerarse como resultado de aplicar una operación algebraica a los elementos d, c y a, e, respectivamente.

En general, los elementos $d * c$ y $a * e$ pueden resultar ser diferentes. Consideraremos de nuevo un conjunto de números racionales positivos y una operación algebraica de división de los números en este conjunto. Es fácil convencerse de que, como regla, se tiene $(a : b) : c \neq a : (b : c)$. Por ejemplo, $((3/2): 3) : (3/4) = 2/3$, pero $(3/2) : (3 : (3/4)) = 3/8$.

Una operación algebraica se denomina *asociativa*, si para cualesquiera tres elementos a, b, c del conjunto inicial se cumple la igualdad $a * (b * c) = (a * b) * c$.

La asociatividad de una operación permite hablar sobre el resultado, unívocamente definido, de aplicar la operación algebraica a los tres elementos a, b, c, con la particularidad de que por resultado se entiende cualquiera de las expresiones iguales $a * (b * c)$ y $(a * b) * c$; además la asociatividad nos permite escribir $a * b * c$ sin paréntesis.

En el caso de una operación asociativa se puede hablar también sobre la univocidad de la expresión $a_1 * a_2 * \ldots * a_n$ que contiene cualquier número finito de elementos $a_1, a_2, \ldots, a_n$. Por significado de $a_1 * a_2 \ldots * a_n$ entenderemos lo siguiente. En esta expresión distribuyamos los paréntesis de un modo arbitrario, con tal de que se pueda definirla mediante la aplicación sucesiva de una operación algebraica a los pares de elementos. Por ejemplo, para cinco elementos a_1, a_2, a_3, a_4, a_5 los paréntesis podrían ser distribuidos así: $a_1 * ((a_2 * a_3) * (a_4 * a_5))$ o bien: $((a_1 * a_2) * a_3) * (a_4 * a_5)$ o mediante muchos otros métodos.

Demostremos que para una operación asociativa el resultado de cálculo no depende de la distribución de los paréntesis. En efecto, para $n = 3$ esta afirmación se deduce de la definición de operación asociativa. Por ello suponemos $n > 3$ y consideramos que para todos los números inferiores a n nuestra afirmación ya está demostrada.

Sean dados los elementos $a_1, a_2, \ldots, a_n$ y suponemos que los paréntesis están distribuidos de una manera tal que se indica el orden en que debe realizarse la operación. Observemos que el último paso siempre será la realización de la operación con dos elementos, $a_1 * a_2 * \ldots * a_k$ y $a_{k+1} * a_{k+2} * \ldots * a_n$, para cierto k que satisface la condición $1 \leqslant k \leqslant n-1$. Puesto que ambas expresiones contienen menos que n elementos, según la hipótesis ellas se definen unívocamente y nos resta demostrar que para cualesquiera enteros y positivos k, l, $l \geqslant 1$ se verifica la igualdad

$$(a_1 * a_2 * \ldots * a_k) * (a_{k+1} * a_{k+2} * \ldots * a_n) = \\ = (a_1 * a_2 * \ldots * a_{k+l}) * (a_{k+l+1} * a_{k+l+2} * \ldots * a_n).$$

Al designar

$$a_1 * a_2 * \ldots * a_k = b,$$
$$a_{k+1} * a_{k+2} * \ldots * a_{k+l} = c,$$
$$a_{k+l+1} * a_{k+l+2} * \ldots * a_n = d,$$

obtenemos, en virtud de asociatividad de la operación, que

$$b * (c * d) = (b * c) * d,$$

y nuestra afirmación queda demostrada.

Si la operación no es solamente asociativa, sino también conmutativa, entonces la expresión $a_1 * a_2 * \ldots * a_n$ tampoco depende del orden en que se disponen los elementos. La demostración de esta afirmación queda a cargo del lector en calidad de ejercicio.

No conviene pensar que la conmutatividad y las asociatividad de una operación están ligadas entre sí de tal o cual manera. Se pueden construir las operaciones con las más diversas combinaciones de dichas propiedades. Ya hemos visto en los ejemplos de multiplicación y división de los números que una operación puede ser conmutativa y asociativa o bien no conmutativa y no asociativa. Examinemos dos ejemplos más. Supongamos que el conjunto consta de tres elementos a, b, c. Prefijemos las operaciones algebraicas por medio de las tablas:

*	a	b	c
a	a	c	b
b	c	b	a
c	b	a	c

*	a	b	c
a	a	a	a
b	b	b	b
c	c	c	c

(2.1)

y convengamos en que el primero siempre se elige un elemento según la columna, el segundo, un elemento según la fila, mientras que el resultado de la operación se tomará en el lugar donde se intersecan la fila y la columna correspondientes. En el primer caso la operación es, evidentemente, conmutativa, pero no asociativa, puesto que, por ejemplo,

$$(a * b) * c = c * c = c,$$
$$a * (b * c) = a * a = a.$$

En el segundo caso la operación no es conmutativa, sino asociativa, de lo que podemos convencernos con facilidad por una comprobación inmediata.

Ejercicios.

1. ¿Será algebraica la operación de cálculo de tg x en el conjunto de todos los números x reales?

2. Consideraremos un conjunto de números reales x que satisfacen la desigualdad $|x| \leqslant 1$. ¿Serán algebraicas en este conjunto las operaciones de multiplicación, adición, división y sustracción de los números?

3. Será conmutativa y asociativa la operación algebraica $x * y = x^2 + y$ en el conjunto de todos los números reales x, y?

4. Supongamos que un conjunto consta de un solo elemento. ¿Cómo se puede definir en este conjunto una operación algebraica?

5. Constrúyanse los ejemplos de operaciones algebraicas en un conjunto cuyos elementos son también conjuntos. ¿Serán estas operaciones conmutativas y asociativas?

§ 3. Operación inversa

Supongamos que en el conjunto A está dada cierta operación algebraica. Como ya sabemos, ésta pone en correspondencia a cualesquiera dos elementos a, b de A un elemento tercero $c = a * b$. Consideremos la totalidad C de aquellos elementos de A que pueden ser representados como resultado de haberse realizado la operación algebraica dada. Por lo visto, cualquiera que sea la operación algebraica, todos los elementos de C son, a la vez, elementos de A. Sin embargo, no es del todo obligatorio que todos los elementos de A pertenezcan a C.

Efectivamente, fijemos en el conjunto A un cierto elemento f y pongámoslo en correspondencia a todo par de elementos a, b de A. Es evidente que la correspondencia construída de tal manera es una operación algebraica y, además, conmutativa y asociativa. El conjunto C sólo contendrá un único elemento f, independientemente de la cantidad de los elementos contenidos en el conjunto A.

Por medio de una operación algebraica se determina cuáles, precisamente, elementos de A integran C. Sea esta operación tal que C coincide con A, es decir, ambos conjuntos contienen los mismos elementos. En este caso todo elemento de A puede ser representado como resultado de realización de la operación algebraica dada sobre algunos dos elementos del mismo conjunto A. Por supuesto, la representación semejante puede ser no la única. No obstante, llegamos a la conclusión de que a todo elemento de A se le puede poner en correspondencia los pares determinados de elementos de A.

Así pues, la operación algebraica inicial engendra en el conjunto A otra operación. Esta última puede no ser unívoca, puesto que a un solo elemento se le puede poner en correspondencia más que un par de elementos. Pero, incluso siendo unívoca, la operación no será algebraica, dado que está definida no para cualquier *par* de elementos, sino sólo para *un único* elemento, aunque este último puede ser arbitrario. Con relación a la operación algebraica dada sería natural atribuirle a esta nueva operación el nombre de operación "inversa". Sin embargo, en realidad por operación inversa entenderemos una noción algo diferente, más próxima al concepto de operación algebraica.

Cabe observar que la investigación de la operación "inversa" es equivalente a la investigación de aquellos elementos u, v que satisfacen la igualdad

$$u * v = b \qquad (3.1)$$

para diferentes elementos b. La investigación de la ecuación dada respecto de dos elementos u, v se reduce con facilidad a la investigación de dos ecuaciones respecto de un solo elemento. Con este fin basta fijar uno de ellos y determinar el otro, sirviéndose de la ecuación (3.1). Así pues, la investigación de la operación "inversa" es equivalente matemáticamente a la resolución de las ecuaciones

$$a * x = b, \qquad y * a = b \qquad (3.2)$$

respecto de los elementos x, y de A para diferentes elementos fijados a, b de A.

Supongamos que las ecuaciones (3.2) tienen soluciones y, además, únicas, para cualesquiera a, b. Entonces, a todo par ordenado de elementos a, b de A podemos ponerle en correspondencia los elementos x, y de A unívocamente definidos, es decir, podemos introducir *dos* operaciones algebraicas. Estas operaciones se llaman operaciones inversas *derecha* e *izquierda*, respectivamente, con relación a la operación principal. Si dichas operaciones existen, diremos que la operación principal tiene operación *inversa*. Observemos que el ejemplo considerado arriba muestra que una operación algebraica, siendo incluso conmutativa y asociativa, puede no tener operaciones inversas, ni la derecha ni la izquierda.

La existencia de la operación inversa significa en realidad que existen, en general, dos operaciones algebraicas diferentes: operaciones inversas derecha e izquierda. Por esta razón hemos de hablar sobre diferentes elementos x, y. En cambio, si una operación algebraica es conmutativa y existe para ella la operación inversa, entonces, evidentemente, $x = y$, y la operación inversa derecha coincide con la izquierda.

He aquí algunos de los ejemplos. Supongamos que el conjunto A representa en sí todo el eje real y la operación algebraica consiste en la multiplicación ordinaria de los números. Esta operación en el conjunto *dado* no tiene operación inversa, puesto que, por ejemplo, para $a = 0$, $b = 1$, las igualdades (3.2) no pueden tener lugar, cualesquiera que sean los números x e y. Si, en cambio, examinamos una operación de multiplicación prescrita solamente en el conjunto de números positivos, esta operación, *ahora*, ya tendrá la inversa.

En efecto, para cualesquiera números positivos a y b existen los números positivos x, y —los únicos— que satisfacen las igualdades (3.2). La operación inversa en el caso dado no es otra cosa que división de los números. El hecho de que en realidad $x = y$, no tiene en las condiciones dadas ningún interés para nosotros.

La operación de adición de los números no tiene inversa, si está definida en el conjunto de números *positivos*, puesto que, por ejemplo, las igualdades (3.2) no pueden verificarse, cualesquiera que sean los números positivos x, y, siempre que $a = 2$, $b = 1$. Si, en cambio, la operación de adición de los números está dada en *todo* el eje real, entonces la operación inversa existe y es nada más que la sustracción de números.

El ejemplo con las operaciones de multiplicación y adición de los números muestra que las operaciones directa e inversa pueden poseer unas *propiedades* más diversas. No es del todo obligatorio que de la asociatividad y la conmutatividad de una operación algebraica provenga la asociatividad y la conmutatividad de la operación inversa, incluso si ésta existe. Más aún, como ya se ha señalado más arriba, una operación algebraica conmutativa y asociativa simplemente puede no tener operaciones inversas, ni la derecha ni la izquierda.

Estos ejemplos sencillos indican una circunstancia de importancia más. Examinemos otra vez la operación de multiplicación en un conjunto de números positivos. Para tal operación las operaciones inversas derecha e izquierda coinciden y representan en sí división de los números. Puede parecer a primera vista que ahora para la operación de división de los números la operación inversa consistirá en multiplicación de los números. No obstante, esto no es del todo así.

Efectivamente, escribamos las ecuaciones correspondientes (3.2)

$$a : x = b, \qquad y : a = b.$$

Es evidente que

$$x = a : b, \qquad y = a \cdot b.$$

Por consiguiente, la operación inversa derecha para dividir los números es de nuevo la división de los números, mientras que la operación inversa izquierda es la multiplicación de los números. De este modo, una operación, inversa a la inversa, no coincide necesariamente con la operación algebraica inicial.

Ejercicios.

1. ¿Existirán las operaciones inversas derechas e izquierdas para las operaciones algebraicas definidas por las tablas (2.1)?
2. ¿Qué representan en sí las operaciones inversas derecha e izquierda para la operación algebraica $x * y = x^y$, definida en un conjunto de los números positivos x, y?
3. Demuéstrese que si las operaciones inversas derecha e izquierda coinciden, entonces la operación algebraica inicial es conmutativa.
4. Demuéstrese que si una operación algebraica tiene la inversa, entonces las operaciones inversas derecha e izquierda también tienen sus inversas. ¿Qué representan en sí estas operaciones?
5. Constrúyase un ejemplo de la operación algebraica para la cual las cuatro operaciones, inversas a las inversas, son todas coincidentes con la operación inicial.

§ 4. Relación de equivalencia

Observemos que en el examen realizado de las propiedades de una operación algebraica suponíamos implícitamente la posibilidad de que cualesquiera dos elementos del conjunto puedan ser comprobados: si son coincidentes entre sí o no lo son. Más aún, los elementos coincidentes fueron tratados con toda la facilidad sin establecer diferencia entre ellos en ninguno de los casos. Nunca suponíamos que los elementos coincidentes representan, de hecho, un solo elemento y no son los objetos diferentes. En realidad, sólo hemos aprovechado el hecho de que cierto grupo de elementos, que se llamaban iguales, en ciertas condiciones se pone de manifiesto de una manera idéntica.

Con tal situación nos encontramos con frecuencia. Al investigar las propiedades generales de los triángulos semejantes, no hacemos distinciones, en realidad, entre ningunos de los triángulos que tienen ángulos iguales. Desde el punto de vista de las propiedades que se conservan en una transformación de semejanza, los triángulos mencionados no se distinguen y podrían denominarse "iguales". Al investigar los criterios de igualdad de los triángulos, no establecemos diferencia entre aquellos que están dispuestos por diferentes lugares del plano, pero pueden ser hechos coincidentes durante su desplazamiento.

En las más diversas situaciones nos encontraremos con la necesidad de partir tal o cual conjunto en los grupos de elementos unidos entre sí según cierto criterio. Si en estas circunstancias ninguno de los elementos pertenece a los dos grupos diferentes, diremos que se trata de la partición de un conjunto en *grupos disjuntos* o en *clases*.

Los criterios según los cuales los elementos de un conjunto se parten en clases, aunque pueden ser muy distintos, sin embargo, no son del todo arbitrarios. Supongamos, por ejemplo, que desearíamos partir en clases todos los números reales incorporando los números a y b a una misma clase cuando, y sólo cuando, $b > a$. Entonces, ninguno de los números a puede caer a una clase consigo mismo, puesto que a no es mayor que el propio a. Por consiguiente, no puede existir ninguna partición en clases según el criterio dado.

Sea dado un cierto criterio. Convengamos en considerar que respecto a todo par de elementos a, b del conjunto A podemos decir que o bien el elemento a está ligado con el elemento b mediante el criterio citado o bien no está ligado con él. Si el elemento a está ligado con b, escribiremos $a \sim b$ y diremos que *a es equivalente* a b.

El análisis de los ejemplos más sencillos ya nos dicta aquellas condiciones que deben ser satisfechas por el criterio para que se puede realizar, sobre la base de éste, una partición del conjunto A en clases. A saber:

1. La reflexividad: $a \sim a$ para todo $a \in A$.
2. La simetría: si $a \sim b$, entonces $b \sim a$.
3. La transitividad: si $a \sim b$ y $b \sim c$, entonces $a \sim c$.

Un criterio que satisface estas condiciones, se denomina *relación de equivalencia*.

Demostraremos que cualquier relación de equivalencia permite partir un conjunto en clases. Efectivamente, sea K_a un grupo de elementos de A equivalentes al elemento fijado a. Debido a la propiedad de reflexividad, $a \in K_a$. Probemos que dos grupos K_a y K_b o bien coinciden o bien no tienen elementos comunes.

Supongamos que cierto elemento c pertenece a K_a y a K_b, es decir, $c \sim a$ y $c \sim b$. En virtud de la propiedad de simetría, $a \sim c$, y, debido a la propiedad de transitividad, $a \sim b$ y, por supuesto, $b \sim a$. Ahora, si $x \in K_a$, entonces $x \sim a$ y, por lo tanto, $x \sim b$, es decir, $x \in K_b$. Análogamente, si $x \in K_b$ se desprende que $x \in K_a$. De este modo, dos grupos que tienen al menos un solo elemento común, son totalmente coincidentes y hemos obtenido realmente la partición del conjunto A en clases

Desde el punto de vista del criterio en consideración, cualesquiera dos elementos pueden ser o bien equivalentes o bien no equivalentes. Nada cambiará, si *denominamos* iguales los elementos equivalentes (¡con relación al criterio dado!) y desiguales los no equivalentes (¡con relación al mismo criterio!).

Podría parecer que en este caso despreciamos el sentido de la palabra «iguales», es que ahora unos elementos, iguales según un criterio, pueden resultar ser desiguales para otro criterio. No obstante, en esto no hay nada que no sea natural. En todo problema concreto los elementos se distinguen o no se distinguen solamente respecto de aquellas propiedades que son de interés para nosotros precisamente en el problema dado. Entre tanto, en los problemas diferentes podemos revelar interés hacia diferentes propiedades de los mismos elementos.

En lo sucesivo consideraremos que siempre, cuando sea necesario, para los elementos de un conjunto debe ser definido un criterio de igualdad que permita decir que el elemento a es igual al elemento b o no es igual a éste. Si el elemento a es igual al elemento b, escribiremos $a = b$, y $a \neq b$, en el caso contrario. Supondremos también que el criterio de igualdad es la relación de equivalencia. En este caso las condiciones de reflexividad, simetría y transitividad pueden considerarse como reflexión de las propiedades más generales de una relación ordinaria de igualdad de los números.

La introducción de la relación de igualdad permite partir todo el conjunto en las clases de elementos los cuales, en virtud de una u otra razón, hemos decidido considerar iguales. Quiere decir, la diferencia entre los elementos que integran una misma clase no nos toca de ninguna manera. Por consiguiente, en todas las situaciones que han de ser consideradas en lo sucesivo, los elementos llamados iguales deben ponerse de manifiesto de un modo igual.

Al introducir la relación de igualdad de una manera *axiomática*, es decir, sin referencias a la naturaleza concreta de los elementos, convengamos en considerar que el uso del signo de igualdad sólo significa que los elementos dispuestos por ambos lados de este signo simplemente coinciden: esto es un mismo elemento. Si el signo de igualdad se emplea de la manera indicada, las propiedades de reflexividad, simetría y transitividad no requieren un acuerdo especial. Al partir el conjunto en las clases de elementos iguales, toda clase se compondrá de un único elemento.

En el caso en que la relación de igualdad se introduzca basándose sobre *la naturaleza concreta de los elementos*, puede suceder que algunas o incluso todas las clases de elementos iguales se compondrán de más de un elemento. Esto requiere que al introducir tales o cuales operaciones con los elementos, se deben imponer ciertas exigencias adicionales sobre las propias operaciones.

En efecto, según se ha acordado, los elementos iguales deben ponerse de manifiesto de una manera igual. Por esta razón, toda operación que se introduce, siendo aplicada a los elementos iguales, ha de conceder ahora los resultados iguales. En realidad nunca nos detendremos en la comprobación de que esta sugerencia se cumpla, prestando al propio lector la posibilidad de convencerse de la validez de la propiedad citada para todas las operaciones introducidas.

Ejercicios.

1. ¿Pueden partirse todos los países del globo terrestre en clases, incorporando a una misma clase dos países cuando, y sólo cuando, ellos tienen la frontera común? Si no es posible, ¿por qué?

2. Examinemos un conjunto de ciudades que tienen vías de comunicación por automóvil. Llamaremos ligadas dos ciudades A y B, si de A se puede llegar a B por carretera. ¿Se pueden partir las ciudades, según dicho criterio, en clases? Si es posible, ¿qué representan en sí las clases?

3. Llamaremos dos números complejos a y b iguales en módulo, si $|a| = |b|$. ¿Será este criterio la relación de equivalencia? ¿Qué representa en sí la partición en clases?

4. Examinemos las operaciones algebraicas de adición y multiplicación de los números complejos. ¿Cómo operan en las clases de elementos iguales en módulo?

5. Constrúyanse los ejemplos de operaciones algebraicas en el conjunto definido en el ejercicio 2. ¿Cómo estas operaciones actúan en las clases?

§ 5. Segmentos dirigidos

Los ejemplos examinados arriba pueden crear la impresión de que todos los razonamientos acerca de las operaciones sobre elementos de los conjuntos sólo se relacionan con las operaciones sobre diversos conjuntos de números. Sin embargo, esto no es así. En adelante daremos a conocer varios ejemplos de con juntos no numéricos con operaciones, pero por ahora consideraremos solamente un ejemplo el cual siempre será el objeto de nuestras referencias a lo largo de todo el curso.

Los conceptos más fundamentales de la física son tales nociones como fuerza, desplazamiento, velocidad, aceleración. Todas estas nociones se caracterizan no sólo por un número, que determina su valor, sino también por cierta dirección. Construyamos ahora un análogo geométrico de las nociones semejantes.

Sean dados en un espacio dos puntos distintos A y B. En una línea recta que pasa por los puntos estos últimos determinan de modo natural cierto segmento. Convengamos en considerar que los puntos siempre se dan en un orden determinado, por ejemplo, primero se fija A y después, B. Ahora podemos determinar la dirección en el segmento construido, a saber, la dirección del primer punto A al segundo punto B.

Un segmento, considerado junto con la dirección prefijada en él, se denomina *segmento dirigido* con el origen A y el extremo B. El segmento dirigido se llamará en otras palabras *vector* y el punto A, *punto de aplicación* del vector. Un vector con el punto de aplicación A lo denominaremos *fijado* en el punto A.

Para los segmentos dirigidos o vectores emplearemos dos formas de su designación. Si hemos de recalcar que se trata de un segmento dirigido cuyo origen se ubica en el punto A y el extremo, en el punto B, escribiremos el símbolo $\overrightarrow{AB}$. Si no nos interesa cuáles precisamente puntos de un segmento dirigido son de frontera, entonces en este caso emplearemos designaciones más sencillas, por ejemplo, las letras latinas minúsculas. En los dibujos los segmentos dirigidos se designarán por medio de las flechas, con la particularidad de que la punta de la flecha siempre se dispondrá en el punto final del segmento.

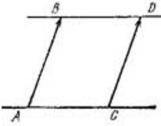


Fig. 5.1.

En todo segmento dirigido es esencial cuál de los puntos frontera constituye su origen y cuál, su extremo. Por esta razón, los segmentos dirigidos $\overrightarrow{AB}$ y $\overrightarrow{BA}$ se consideran diferentes.

Así pues, podemos construir diversos conjuntos cuyos elementos serán segmentos dirigidos. Antes de introducir las operaciones sobre elementos, llegaremos a un acuerdo de qué segmentos dirigidos se considerarán iguales.

Examinemos primero un *traslado paralelo* del segmento dirigido $\overrightarrow{AB}$ en el punto C. Supongamos que el punto C no se dispone en una recta que pasa por A y B (fig. 5.1). Tracemos una recta que pasa por los puntos A y C, después, una recta que pasa por el punto C y es paralela a la recta AB y, por fin, una recta que pasa por el punto B y es paralela a la recta AC. El punto de intersección de las dos últimas rectas se designará mediante D. El segmento dirigido $\overrightarrow{CD}$ se considerará obtenido como resultado de trasladar paralelamente el segmento $\overrightarrow{AB}$ en el punto C. Si, en cambio, el punto C se ubica en la recta que pasa por A y B, entonces el segmento $\overrightarrow{CD}$ se obtiene por el desplazamiento del segmento $\overrightarrow{AB}$ a lo largo de la recta que lo contiene hasta que el punto A coincida con el punto C.

Ahora se puede enunciar la definición de igualdad de los vectores. Dos vectores se denominan *iguales*, si se pueden hacer coincidir en el transcurso del traslado paralelo. No es difícil ver que esta definición de igualdad es la relación de equivalencia, es decir, posee las propiedades de reflexividad, simetría y transitividad.

De este modo, la totalidad de todos los vectores se divide de modo natural en unas clases de vectores iguales. Cada una de estas clases se describe sin dificultades. Se obtiene por medio del traslado paralelo de cualquiera de los vectores de la clase en todos los puntos del espacio.

Observemos que en todo punto del espacio se halla fijado un vector y sólo uno, de cada clase de los vectores iguales. Por esto, al comparar los vectores a, b se puede usar el siguiente procedimiento. Habiendo fijado un cierto punto, traslademos paralelamente en este punto los vectores a, b. Si, en este caso, ellos coinciden por completo, entonces $a = b$, y $a \neq b$, en el caso contrario.

Además de un conjunto compuesto por todos los vectores del espacio, trataremos con frecuencia unos conjuntos de otro género. En lo principal, serán los conjuntos de vectores o bien paralelos a cierta línea recta o dispuestos en ésta, o bien paralelos a cierto plano o dispuestos en éste. Los vectores de tal índole los llamaremos *colineales* o *coplanares*, respectivamente. Naturalmente en los conjuntos de vectores colineales y coplanares sigue siendo válida la definición de igualdad de los vectores enunciada más arriba.

Consideraremos también los así llamados segmentos dirigidos *nulos* cuyos orígenes coinciden con los extremos. La orientación de los vectores nulos no está definida y todos ellos se consideran, por definición, iguales. Si la indicación de los puntos frontera de un vector nulo no es obligatoria, dicho vector se designará por el símbolo $\mathbf{0}$.

También por definición, consideraremos que cualquier vector nulo es paralelo a toda recta y a todo plano. Por esta razón, siempre supondremos en lo sucesivo, si no se hacen reservas especiales, que el conjunto de vectores de un espacio, como también todo conjunto de los vectores colineales y coplanares comprende en sí el conjunto de todos los vectores nulos. Esto no se debe olvidar.

Ejercicios.

1. Demuéstrese que los vectores no nulos de un espacio pueden ser divididos en clases de vectores colineales no nulos.
2. Demuéstrese que cualquier clase de vectores colineales no nulos se la puede dividir en clases de vectores iguales no nulos.
3. Demuéstrese que toda clase de vectores iguales no nulos se incorpora íntegramente a una y sólo una clase de vectores colineales no nulos.
4. ¿Pueden dividirse los vectores no nulos de un espacio en las clases de vectores coplanares? ¿Si no se dividen, por qué?
5. Demuéstrese que todo conjunto de vectores coplanares no nulos se lo puede dividir en las clases de vectores colineales no nulos.

6. Demuéstrese que todo par de diferentes clases de los vectores colineales no nulos se incorpora íntegramente a uno y sólo un conjunto de vectores coplanares no nulos.

§ 6. Adición de los segmentos dirigidos

Como ya se ha indicado, la fuerza, el desplazamiento, la velocidad y la aceleración son imágenes de los segmentos dirigidos construidos por nosotros. Para que estos segmentos resulten útiles en la resolución de diferentes problemas físicos, se debe, al introducir las operaciones con ellos, tomar en consideración también las analogías físicas correspondientes.

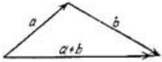


Fig. 6.1.

Está bien conocida la operación de adición de fuerzas que se realiza según la así llamada regla del paralelogramo. Conforme a esta misma regla se suman también los desplazamientos, las velocidades y las aceleraciones. De acuerdo con la terminología introducida, esta operación es algebraica, conmutativa y asociativa. Nuestra tarea inmediata consiste en construcción de una operación semejante sobre los segmentos dirigidos.

Definamos la operación de *adición de los vectores* del modo siguiente. Supongamos que es preciso sumar los vectores a y b. Traslademos paralelamente el vector b en el extremo del vector a (fig. 6.1). Entonces, se llamará *suma* $a + b$ un vector cuyo origen coincide con el origen de a y el extremo, con el extremo de b. Esta regla para realizar la operación se denomina, comúnmente, *«regla del triángulo»*.

Es evidente que la operación de adición de los vectores es algebraica. Demostremos que es conmutativa y asociativa.

Para demostrar la conmutatividad de la operación supongamos al principio que los vectores a y b no son colineales. Dispondrémoslos en el origen general O (fig. 6.2). Designaremos con las letras A y B los extremos de los vectores a y b, respectivamente y consideraremos el paralelogramo $OBCA$. De la definición de igualdad de los vectores se deduce que

$$\overrightarrow{BC} = \overrightarrow{OA} = a, \qquad \overrightarrow{AC} = \overrightarrow{OB} = b.$$

Pero, en este caso, una misma diagonal $\overrightarrow{OC}$ del paralelogramo $OBCA$ es, a la vez, $a + b$ y $b + a$. El carácter colineal de los vectores a y b es obvio.

Observemos que al mismo tiempo hemos obtenido otro método para construir la suma de los vectores. A saber, si en los vectores a, b fijados en un punto, está construido un paralelogramo, su diagonal fijada en el mismo punto será la suma $a + b$.

Con el objeto de demostrar la asociatividad de la operación de adición, apliquemos el vector a al punto arbitrario O, el vector b, al extremo del vector a y el vector c, al extremo del vector b (fig. 6.3). Designemos con A, B, C los extremos de los vectores a, b, c. Entonces

$$(a+b)+c=(\overrightarrow{OA}+\overrightarrow{AB})+\overrightarrow{BC}=\overrightarrow{OB}+\overrightarrow{BC}=\overrightarrow{OC},$$

$$a+(b+c)=\overrightarrow{OA}+(\overrightarrow{AB}+\overrightarrow{BC})=\overrightarrow{OA}+\overrightarrow{AC}=\overrightarrow{OC}.$$

Por ser transitiva la relación de igualdad de los vectores concluimos que la asociatividad de la operación también tiene lugar.

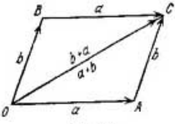


Fig. 6.2.

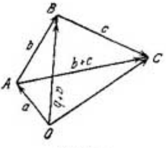


Fig. 6.3.

Las propiedades demostradas de la operación de adición de vectores permiten calcular la suma de cualquier número de vectores. Si el vector a_2 se aplica al extremo del vector a_1, el vector a_3, al extremo del vector a_2, etc. y, por último, el vector a_n, al extremo del vector a_{n-1}, entonces la suma $a_1+a_2+\ldots+a_n$ representará un vector cuyo origen coincide con el origen de a_1 y el extremo, con el de a_n. Esta regla para construir las sumas de vectores se denomina «*regla de cierre de una quebrada hasta completar un polígono*».

Planteemos ahora una pregunta sobre la existencia de la operación inversa para sumar los vectores. Como es sabido, para contestarla es preciso investigar la existencia y unicidad de la solución de la ecuación

$$a+x=b, \qquad y+a=b$$

para los vectores a, b arbitrarios. En virtud de que la operación principal es conmutativa, será suficiente, evidentemente, investigar sólo una de las ecuaciones.

Tomemos un segmento dirigido arbitrario $\overrightarrow{AB}$. Mediante una construcción geométrica elemental establecemos que siempre se verifican las correlaciones

$$\overrightarrow{AB}+\overrightarrow{BA}=0, \qquad \overrightarrow{AB}+0=\overrightarrow{AB}. \tag{6.1}$$

Por ello, la ecuación

$$\overrightarrow{AB}+x=\overrightarrow{CD}, \tag{6.2}$$

para cualesquiera vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$, tendrá, a ciencia cierta, por lo menos una solución, por ejemplo,

$$x=\overrightarrow{BA}+\overrightarrow{CD}. \tag{6.3}$$

Supongamos que la ecuación (6.2) queda satisfecha, además, por cierto otro vector z, es decir,

$$\overrightarrow{AB}+x=\overrightarrow{CD}, \qquad \overrightarrow{AB}+z=\overrightarrow{CD}.$$

Entonces, sumando $\overrightarrow{BA}$ a ambos miembros de estas igualdades, obtenemos, al tomar en consideración (6.1), que $x=\overrightarrow{BA}+\overrightarrow{CD}$, $z=\overrightarrow{BA}+\overrightarrow{CD}$ y, por consiguiente, $x=z$.

De este modo, para la operación de adición introducida por nosotros existe una operación inversa. Ella se denomina operación de *sustracción* de vectores. Si para los vectores a, b, c tiene lugar la igualdad $a+c=b$, *escribiremos simbólicamente* que $c=b-a$. El vector $b-a$, definido unívocamente por los vectores b y a, se llama *diferencia* de vectores. Un poco más tarde daremos la argumentación para tal designación y tal nombre.

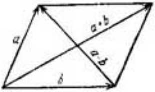


Fig. 6.4.

Es fácil indicar la regla para construir diferencia entre dos vectores dados a y b. Apliquemos estos vectores a un punto común y construyamos en ellos un paralelogramo (fig. 6.4). Ya hemos señalado anteriormente que una diagonal del paralelogramo es la suma de los vectores dados. La otra diagonal, como es fácil de ver, es la diferencia de los mismos vectores. La regla descrita para construir la suma y la diferencia de los vectores se llama, comúnmente, "*regla del paralelogramo*".

Observemos que podríamos definir la operación de adición no para el conjunto de todos los vectores de un espacio, sino sólo para uno de los conjuntos de vectores colineales o coplanares. La suma de dos vectores de cualquier conjunto de tal tipo pertenecerá de nuevo *al mismo* conjunto. Es por esto que la operación de adición de vectores queda algebraica en este caso también. Más aún, dicha operación sigue conservando todas sus propiedades y, lo que es de mayor importancia, posee, como antes, la operación inversa. La validez de esta última afirmación se desprende de la fórmula (6.3). Si los vectores

$\overrightarrow{AB}$ y $\overrightarrow{CD}$ son paralelos a cierta recta o a cierto plano, entonces es evidente que será paralelo también el vector $\overrightarrow{BA} + \overrightarrow{CD}$ o bien, que es lo mismo, el vector de la diferencia $\overrightarrow{CD} - \overrightarrow{AB}$.

Así pues, la operación de adición de los vectores es algebraica, conmutativa, asociativa y tiene operación inversa en los conjuntos de tres tipos. Éstos son: el conjunto de vectores de un espacio, el conjunto de vectores colineales y el conjunto de vectores coplanares.

Ejercicios.

1. A uno de los vértices de un cubo están aplicadas tres fuerzas iguales en valor y dirigidas a lo largo de las aristas. ¿Cuál es la dirección de la suma de estas fuerzas?
2. Sean dadas tres clases diferentes de los vectores colineales. ¿En qué caso cualquier vector del espacio puede ser representado como suma de tres vectores de estas clases?
3. A los vértices de un polígono regular están aplicadas las fuerzas iguales por su valor y dirigidas hacia el centro. ¿Cuál será la suma de estas fuerzas?
4. ¿Qué representa en sí el conjunto de sumas de los vectores tomados de dos clases diferentes de los vectores colineales?

§ 7. Grupos

Los conjuntos con una operación algebraica son en cierto sentido los más sencillos, razón por la cual parece natural empezar nuestras investigaciones precisamente por tales conjuntos. Tomaremos por axiomas las propiedades de la operación y después deduciremos de éstos los corolarios. Esto nos permite *en lo sucesivo* aplicar inmediatamente los resultados de nuestras investigaciones a todos los conjuntos, en los cuales las operaciones tienen propiedades análogas, independientemente de las peculiaridades concretas.

Se denomina *grupo* un conjunto G con una operación algebraica que es asociativa (aunque no sea necesariamente conmutativa), con la particularidad de que para dicha operación debe existir operación inversa.

Ha de notarse que la operación inversa no puede considerarse como segunda operación independiente en el grupo, puesto que ella se determina en términos de la operación principal. Según lo aceptado en la teoría de los grupos, llamemos *multiplicación* la operación prefijada en G y convengamos en emplear las notaciones correspondientes. Antes de empezar a considerar los ejemplos diferentes de grupos, deduzcamos unos corolarios más sencillos de la propia definición.

Examinemos un elemento arbitrario a del grupo G. Del hecho de que en el grupo existe la operación inversa se desprende la existencia de un único elemento e_a de tal índole que $ae_a = a$. Por consiguiente, cuando el elemento a se multiplica a la derecha por el e_a, este último desempeña el mismo papel que en la multiplicación de los números pertenece a la unidad. Ahora, sea b cualquier otro elemento del grupo. Es evidente que existe un elemento y que satisface la igualdad $ya = b$. Obtenemos pues

$$b = ya = y(ae_a) = (ya)e_a = be_a.$$

Así, e_a desempeña el papel de la unidad derecha respecto a todos los elementos del grupo G, y no sólo respecto de a. El elemento que posee la propiedad semejante debe ser único. Efectivamente, todos los elementos de esta índole satisfacen la ecuación $ax = a$, pero, conforme a la definición de operación inversa, dicha ecuación tiene una única solución. Designaremos el elemento obtenido mediante e'.

Análogamente se puede demostrar la existencia y unicidad en el grupo G del elemento e'' que satisface la igualdad $e''b = b$ para todo elemento b de G. En realidad los elementos e' y e'' coinciden, lo que se deduce de las igualdades $e''e' = e''$ y $e''e' = e'$.

De este modo hemos obtenido el primer corolario importante: en cualquier grupo G existe un elemento e y, además, *único* que satisface las igualdades

$$ae = ea = a$$

para todo a de G. Este elemento recibe el nombre de elemento *unidad* del grupo G.

De la definición de operación inversa se desprende también la existencia y unicidad, para todo elemento a, de tales elementos a' y a'' que

$$aa' = e, \qquad a''a = e.$$

Éstos se denominan *elementos inversos derecho e izquierdo*, respectivamente. Es fácil mostrar que en el caso dado ellos coinciden. En efecto, examinemos el elemento $a''aa'$ y calculémoslo empleando dos procedimientos diferentes. Se tiene

$$a''aa' = a''(aa') = a''e = a'',$$
$$a''aa' = (a''a)a' = ea' = a'.$$

Por consiguiente, $a'' = a'$. Este elemento se llama *inverso* al elemento a y se designa mediante a^{-1}.

Hemos obtenido, pues, el segundo corolario de importancia: en todo grupo G cualquier elemento a posee un *único* elemento inverso a^{-1}, para el cual

$$aa^{-1} = a^{-1}a = e. \tag{7.1}$$

Basándose en la asociatividad de la operación de grupo se puede hablar sobre la univocidad del producto de cualquier número finito de elementos del grupo prefijados (al tener en cuenta la no conmutatividad eventual de la operación de grupo) en el orden determinado. Tomando en consideración las correlaciones (7.1), no es difícil señalar la fórmula general para un elemento inverso al producto. A saber,

$$(a_1a_2 \ldots a_n)^{-1} = a_n^{-1}a_{n-1}^{-1} \ldots a_1^{-1}. \qquad (7.2)$$

De la correlación (7.1) se deduce que de elemento inverso respecto de a^{-1} sirve el propio elemento a, y el elemento inverso para la unidad será la propia unidad, es decir

$$(a^{-1})^{-1} = a, \qquad e^{-1} = e. \qquad (7.3)$$

La comprobación de que si es o no un grupo, el conjunto con una operación asociativa, se facilita mucho gracias al hecho de que en la definición de grupo el requisito de que se cumpla la operación inversa puede sustituírse por la suposición de que existen la unidad y los elementos inversos y existen, además, sólo por un lado (a la derecha, por ejemplo) sin que se suponga la unicidad de la unidad y de los elementos citados. Para precisar, resulta válido el siguiente

TEOREMA 7.1. *El conjunto G con una operación asociativa será un grupo, si en G existe al menos un elemento e que posee la propiedad de que $ae = a$ para todo a perteneciente a G, y respecto de éste todo elemento a de G tiene aunque un elemento inverso derecho a^{-1}, es decir, $aa^{-1} = e$.*

DEMOSTRACIÓN. Sea a^{-1} uno de los elementos inversos derechos para a. Se tiene

$$aa^{-1} = e = ee = eaa^{-1}.$$

Multipliquemos ambos miembros de esta igualdad a la derecha por uno de los elementos que son derechos e inversos para a^{-1}. Obtendremos que $ae = eae$, de donde se desprende la igualdad $a = ea$, puesto que e es la unidad derecha para G. De este modo, el elemento e resulta ser también la unidad izquierda para G.

Luego, si e' es una unidad arbitraria derecha y e'', una unidad arbitraria izquierda, entonces de las igualdades $e''e' = e'$ y $e''e' = e''$ proviene que $e' = e''$, es decir, toda unidad derecha equivale a cualquiera izquierda. Con esto quedan demostradas la existencia y unicidad en el conjunto G del elemento unidad el cual se designará de nuevo mediante e.

Ahora, para todo elemento inverso derecho a^{-1} se verifican las siguientes correlaciones:

$$a^{-1} = a^{-1}e = a^{-1}aa^{-1}.$$

Multipliquemos ambos miembros de esta igualdad a la derecha por uno de los elementos, derechos e inversos para a^{-1}. Llegamos a que $e = a^{-1}a$, es decir, el elemento a^{-1} es, a la vez, elemento inverso izquierdo para a. Si, ahora, $a^{-1'}$ es un elemento inverso derecho arbitrario para a y $a^{-1''}$, un elemento inverso izquierdo arbitrario, entonces de las igualdades

$$a^{-1''}aa^{-1'} = (a^{-1''}a)\,a^{-1'} = ea^{-1'} = a^{-1'},$$
$$a^{-1''}aa^{-1'} = a^{-1''}(aa^{-1'}) = a^{-1''}e = a^{-1''}$$

se deduce que $a^{-1'} = a^{-1''}$. Esto significa la existencia y unicidad, para todo elemento a de G, de un elemento inverso a^{-1}.

Ahora no es difícil de probar que el conjunto G será un grupo. En efecto, las ecuaciones $ax = b$, $ya = b$ serán, a ciencia cierta, satisfechas por los elementos

$$x = a^{-1}b, \qquad y = ba^{-1}.$$

Supongamos que existen también otras soluciones, por ejemplo, el elemento z para la primera ecuación. En este caso de las igualdades $ax = b$ y $az = b$ se desprende que $ax = az$. Al multiplicar ambos miembros a la izquierda por el elemento a^{-1}, obtenemos $x = z$. De este modo, el conjunto G es un grupo.

Un grupo se llama *conmutativo* o *abeliano*, si la operación de grupo es conmutativa. En este caso la operación, como regla, recibe el nombre de *adición* y en lugar del símbolo de multiplicación ab suele escribirse el de la suma $a + b$. La unidad de un grupo abeliano se llama elemento *nulo* y se designa mediante el símbolo 0. La operación inversa se denomina *sustracción* y el elemento inverso, *opuesto*. Se designa el último por el símbolo $-a$. Convengamos en considerar que, por definición, el símbolo de la diferencia $a - b$ significa la suma $a + (-b)$.

Si, no obstante, en virtud de tal o cual razón, la operación en un grupo conmutativo se llamará *multiplicación*, entonces la operación inversa se considerará como *división*. Los productos $a^{-1}b$ y ba^{-1}, que en el caso dado son iguales, se designarán mediante b/a y se llamarán cociente de la división de b por a.

Ejercicios.

Demuéstrese que los conjuntos a seguir son grupos abelianos. La denominación de toda operación refleja siempre no la notación, sino el contenido.

1. El conjunto: números enteros; la operación: adición de los números.
2. El conjunto: números complejos, salvo cero; la operación: multiplicación de los números.
3. El conjunto: números enteros múltiplos del número 3; la operación: sumación de los números.
4. El conjunto: números enteros racionales; la operación: multiplicación de los números.
5. El conjunto: números del tipo $a + b\sqrt{2}$, donde a, b son números racionales positivos; la operación: multiplicación de los números.
6. El conjunto: un elemento a; la operación se denomina adición y se define mediante la igualdad $a + a = a$.

7. El conjunto: números enteros 0, 1, 2, . . ., $n-1$; la operación se llama «*adición según módulo n*» y consiste en el cálculo del resto no negativo, inferior a n, de la división de una suma de dos números por el número n.

8. El conjunto: números enteros 1, 2, 3, . . ., $n-1$, donde n es un número primo; la operación se llama «*multiplicación según el módulo n*» y consiste en el cálculo del resto no negativo, inferior a n, de la división de un producto de dos números por el número n.

9. El conjunto: segmentos dirigidos colineales; la operación: sumación de los segmentos dirigidos.

10. El conjunto: segmentos dirigidos coplanares; la operación: sumación de los segmentos dirigidos.

11. El conjunto: segmentos dirigidos de un espacio; la operación: sumación de los segmentos dirigidos.

En lo que se refiere a los tres últimos ejemplos hemos de notar que el elemento nulo del grupo abeliano de segmentos dirigidos lo constituirá el segmento dirigido nulo y el segmento opuesto para $\overrightarrow{AB}$ será el segmento $\overrightarrow{BA}$. De lo demostrado más arriba proviene su unicidad. Los ejemplos de grupos no conmutativos los aduciremos más tarde.

§ 8. Anillos y campos

Examinemos un conjunto K en el que se han introducido dos operaciones. Se llamará adición una de las operaciones y la otra será multiplicación; se empleará la notación correspondiente. Supondremos que ambas operaciones están entrelazadas mediante la ley *distributiva*, es decir, para cualesquiera tres elementos a, b, c de K tienen lugar las correlaciones

$$(a+b)\,c = ac + bc, \qquad a\,(b+c) = ab + ac.$$

El conjunto K se llama *anillo*, si en él están definidas dos operaciones: adición y multiplicación, siendo asociativas ambas y ligadas entre sí por la ley distributiva, con la particularidad de que la adición es conmutativa y posee operación inversa. El anillo se denomina *conmutativo*, si la multiplicación es conmutativa, y *no conmutativo*, en el caso contrario.

Observemos que cualquier anillo es un grupo abeliano por adición. Por consiguiente, está provisto de un único elemento nulo 0. Dicho elemento tiene la propiedad de que para todo elemento a del anillo se verifica la igualdad

$$a + 0 = a.$$

La definición de elemento nulo se ha proporcionado únicamente con relación a la operación de adición. No obstante, respecto a la multiplicación este elemento también desempeña un papel singular. A saber, en todo anillo el producto de cualquier elemento por el elemento nulo es elemento nulo. Efectivamente, sea a un elemento cualquiera del anillo K, entonces

$$a\cdot 0 = a\,(0+0) = a\cdot 0 + a\cdot 0.$$

Sumando a ambos miembros de esta igualdad el elemento $-a\cdot 0$, llegamos a que $a\cdot 0 = 0$. Análogamente se demuestra que también $0\cdot a = 0$.

Al hacer uso de esta propiedad del elemento nulo podemos demostrar que en todo anillo para cualesquiera elementos a, b es válida la igualdad

$$(-a)\,b = -(ab).$$

En efecto,

$$ab + (-a)\,b = (a + (-a))\,b = 0\cdot b = 0,$$

es decir, el elemento $(-a)\,b$ es opuesto para el elemento ab. De conformidad con nuestras designaciones podemos escribirlo en la forma $-(ab)$.

Ahora es fácil probar que la ley distributiva es lícita también para la diferencia de elementos. Tenemos

$$(a - b)\,c = (a + (-b))\,c = ac + (-b)\,c = \\ = ac + (-(bc)) = ac - bc,$$

$$a\,(b - c) = a\,(b + (-c)) = ab + a\,(-c) = \\ = ab + (-(ac)) = ab - ac.$$

La ley distributiva, es decir, la regla común para abrir los paréntesis, constituye una única exigencia en la definición de anillo que liga la adición y la multiplicación. Solamente gracias a esta ley el estudio conjunto de dos operaciones citadas nos suministra más que se podría obtener al estudiarlas por separado.

Acabamos de demostrar que las operaciones algebraicas en un anillo poseen muchas propiedades habituales para las operaciones sobre los números. No conviene pensar, sin embargo, que cualquier propiedad de la adición y multiplicación de los números queda en vigor para todo anillo, aunque sea este último conmutativo. Así, por ejemplo, la multiplicación de los números posee una propiedad que es inversa a la de multiplicación por un elemento nulo. A saber si un producto de dos números es nulo, entonces por lo menos uno de los factores es igual a cero. Esta propiedad ya no es obligatoria para un anillo conmutativo, es decir, un producto de los elementos, distintos de elemento nulo, puede resultar nulo.

Los elementos no nulos cuyo producto es un elemento nulo se denomina *divisores de cero*. La presencia de tales elementos en el anillo hace la investigación mucho más complicada y no presta la oportunidad de establecer analogía profunda entre los números y los elementos del anillo conmutativo. La analogía de este género nos la proporciona el examen de los anillos privados de divisores de cero.

Supongamos que con relación a la operación de multiplicación en un anillo conmutativo hay un elemento unidad e y todo elemento no nulo a tiene elemento inverso a^{-1}. No es difícil demostrar que tanto elemento unidad como el inverso son únicos y, no obstante, lo más importante consiste en el hecho de que ahora el anillo no tiene divisores de cero. Efectivamente, sea $ab = 0$, pero $a \neq 0$. Al multiplicar ambos miembros de esta igualdad a la izquierda por el elemento a^{-1}, llegamos a que

$$a^{-1}ab = (a^{-1}a)\,b = eb = b$$

y, por supuesto, $a^{-1}0 = 0$. Por consiguiente, $b = 0$.

De la ausencia de los divisores de cero se deduce que toda igualdad se puede reducirla por un factor común no nulo. Si es que $ca = cb$ y $c \neq 0$, entonces $c\,(a - b) = 0$, de lo cual deducimos que $a - b = 0$, es decir, $a = b$.

Se denomina *campo**) un anillo conmutativo P que contiene un elemento unidad y todo elemento no nulo tiene elemento inverso.

Aprovechando la inscripción de un cociente a/b en forma de un producto ab^{-1}, podemos probar con facilidad que *en cada campo quedan en vigor todas las reglas ordinarias para operar con fracciones, teniendo en cuenta las operaciones de adición, sustracción, división y multiplicación.* A saber:

$$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad \frac{-a}{b} = -\frac{a}{b}.$$

Además, $a/b = c/d$ cuando, y sólo cuando, $ad = bc$, si, naturalmente, $b \neq 0$ y $d \neq 0$. Queda a cargo del lector comprobar la validez de estas afirmaciones en calidad de los ejercicios.

Así, desde el punto de vista de las reglas ordinarias para operar con fracciones, no hay campo que se diferencie de un conjunto de números, razón por la cual los elementos *de todo* campo los llamaremos números, siempre que, por supuesto, tal denominación no nos lleve a una ambigüedad. Como regla, el elemento nulo de cualquier campo se designará por el símbolo 0, y el elemento unidad, por el símbolo 1.

Enunciaremos ahora todos los datos comunes sobre los elementos de cualquier campo conmutativo que nos harán falta en lo sucesivo.

A. A todo par de elementos a, b le corresponde un elemento $a + b$, llamado suma de a y b, con la particularidad de que:

1) la adición es conmutativa, $a + b = b + a$,

2) la adición es asociativa, $a + (b + c) = (a + b) + c$,

3) existe un único elemento nulo 0 tal que $a + 0 = a$ para cualquier elemento a,

4) para todo elemento a existe un único elemento opuesto $-a$ tal que $a + (-a) = 0$.

---

* Al igual que en la obra de Kurosch A. C. «Curso de álgebra superior», editorial «MIR», Moscú, consideramos conveniente, para brevedad y comodidad, emplear el término «campo» como un sinónimo de «cuerpo conmutativo». (*N. del Tr.*)

B. A todo par de elementos a, b le corresponde un elemento ab, llamado producto de a y b, con la particularidad de que:

1) la multiplicación es conmutativa, $ab = ba$,

2) la multiplicación es asociativa, $a(bc) = (ab)c$,

3) existe un único elemento unidad 1 tal que $a \cdot 1 = 1 \cdot a = a$ para todo elemento a,

4) para todo elemento no nulo a existe un único elemento inverso a^{-1} tal que $aa^{-1} = a^{-1}a = 1$.

C. Las operaciones de adición y multiplicación están ligadas entre sí mediante la siguiente correlación: la multiplicación distributiva respecto de la adición, $(a + b)c = ac + bc$.

Los datos citados no pretenden constituir una independencia lógica, sino que sólo son una característica cómoda de los elementos. Las propiedades **A** describen un campo desde el punto de vista de la operación de adición y dicen que respecto a dicha operación el campo constituye un grupo abeliano. Las propiedades **B** describen un campo desde el punto de vista de la operación de multiplicación e indican que respecto a dicha operación el campo se hace un grupo abeliano, si del campo se excluye el elemento nulo. La propiedad **C** describe los lazos que unen entre sí dos operaciones.

Ejercicios.

Demuéstrese que los conjuntos 1—7 son los anillos, pero no los campos, mientras que los conjuntos 8—13 son los campos. La denominación de la operación refleja en todos los casos no la notación, sino el contenido.

1. El conjunto: números enteros; las operaciones: adición y multiplicación de los números.

2. El conjunto: números enteros múltiplos de cierto número n; las operaciones: adición y multiplicación de los números.

3. El conjunto: números reales del tipo $a + b\sqrt{2}$, donde a y b son números enteros; las operaciones: adición y multiplicación de los números.

4. El conjunto: polinomios con coeficientes reales de una variable t, incluidas las constantes; las operaciones: adición y multiplicación de los polinomios.

5. El conjunto: un solo elemento a; las operaciones se definen por las igualdades $a + a = a$ y $a \cdot a = a$.

6. El conjunto: números enteros $0, 1, 2, \ldots, n - 1$, donde n es un número compuesto; las operaciones: adición y multiplicación según el módulo n.

7. El conjunto: pares (a, b) de números enteros; las operaciones se definen mediante las fórmulas

$$(a, b) + (c, d) = (a + c, b + d); \quad (a, b) \cdot (c, d) = (ac, bd).$$

8. El conjunto: números racionales; las operaciones: adición y multiplicación de los números.

9. El conjunto: números reales; las operaciones: adición y multiplicación de los números.

10. El conjunto: números complejos; las operaciones: adición y multiplicación de los números.

11. El conjunto: números reales del tipo $a + b\sqrt{2}$, donde a y b son números racionales; las operaciones: adición y multiplicación de los números.

12. El conjunto: dos elementos a y b; las operaciones se definen por las igualdades

$$a + a = b + b = a, \qquad a + b = b + a = b,$$
$$a \cdot a = a \cdot b = b \cdot a = a, \qquad b \cdot b = b.$$

13. El conjunto: números enteros 0, 1, 2, ..., $n - 1$, donde n es un número primo; las operaciones: adición y multiplicación según el módulo n.

Llamamos la atención del lector a que en uno de los ejemplos se ha aducido un anillo con divisores de cero. ¿Cuál es el ejemplo? ¿Cuál es la forma general de los divisores de cero?

§ 9. Multiplicación del segmento dirigido por un número

Recalquemos una vez más que la operación algebraica se ha definido como operación sobre dos elementos pertenecientes a un mismo conjunto. Sin embargo, los ejemplos numerosos de la física muestran que a veces resulta razonable también considerar operaciones sobre los elementos pertenecientes a diversos conjuntos. Una de tales operaciones es dictada por los conceptos de fuerza, desplazamiento, velocidad y aceleración; considerémosla recurriendo de nuevo a los ejemplos de segmentos dirigidos.

En la física ya hace tiempo se aceptó la práctica de utilizar los segmentos. Si se dice que, por ejemplo, una fuerza se ha aumentado cinco veces, el segmento que la representa se hace "extender" cinco veces sin cambiar la dirección general. Si se habla sobre el cambio de la dirección en que actúa la fuerza, entonces en el segmento correspondiente los puntos inicial (origen) y final cambian de lugar. Partiendo de estos razonamientos, introduzcamos una operación que se denomina multiplicación del segmento dirigido por un número real.

Consideraremos al principio algunas cuestiones generales. Sea dada en un plano o en un espacio una línea recta arbitraria. Convengamos en tomar una de las direcciones en ésta por positiva y la contraria, por negativa. Una recta con la dirección definida se llamará *eje*.

Supongamos ahora que se ha dado un eje e indicado, además, el segmento graduado con ayuda del cual podemos medir cualquier otro segmento y determinar, consecuentemente, la longitud del último. A todo segmento dirigido se le atribuirá su característica numérica, la así llamada magnitud del segmento dirigido.

Se denomina *magnitud* $\{\overrightarrow{AB}\}$ del segmento dirigido $\overrightarrow{AB}$ un número, igual a la longitud del segmento $\overrightarrow{AB}$, que se toma con el signo "más", si la dirección de $\overrightarrow{AB}$ coincide con la dirección positiva del eje y con el signo "menos", si la dirección de $\overrightarrow{AB}$ coincide con la dirección negativa del eje. Las magnitudes de todos los segmentos dirigidos nulos se consideran iguales a cero, es decir,

$$\{\overrightarrow{AA}\}=0.$$

Independientemente de cuál dirección en el eje se ha aceptado por positiva, la dirección de $\overrightarrow{AB}$ es opuesta a la de $\overrightarrow{BA}$, mientras que las longitudes de los segmentos dirigidos $\overrightarrow{AB}$ y $\overrightarrow{BA}$ son iguales, por lo cual,

$$\{\overrightarrow{AB}\}=-\{\overrightarrow{BA}\}. \tag{9.1}$$

La magnitud de un segmento dirigido, a diferencia de su longitud, puede tener cualquier signo. Puesto que la longitud del segmento

Fig. 9.1.

dirigido $\overrightarrow{AB}$ es el módulo de su magnitud, entonces con el fin de designarla emplearemos el símbolo $|\overrightarrow{AB}|$. Está claro que a diferencia de (9.1)

$$|\overrightarrow{AB}|=|\overrightarrow{BA}|.$$

Supongamos que en un eje están prefijados tres puntos cualesquiera A, B, C que definen tres segmentos dirigidos $\overrightarrow{AB}$, $\overrightarrow{BC}$ y $\overrightarrow{AC}$. Cualquiera que sea la disposición de los puntos, las magnitudes de dichos segmentos dirigidos satisfacen la correlación

$$\{\overrightarrow{AB}\}+\{\overrightarrow{BC}\}=\{\overrightarrow{AC}\}. \tag{9.2}$$

En efecto, supongamos que la dirección del eje y la disposición de los puntos es tal como las muestra la fig. 9.1. En este caso será evidente que

$$|\overrightarrow{CA}|+|\overrightarrow{AB}|=|\overrightarrow{CB}|. \tag{9.3}$$

De acuerdo con la definición de magnitud del segmento dirigido y con la igualdad (9.1) tenemos

$$|\overrightarrow{CA}|=\{\overrightarrow{CA}\}=-\{\overrightarrow{AC}\}, \qquad |\overrightarrow{AB}|=\{\overrightarrow{AB}\}, \tag{9.4}$$
$$|\overrightarrow{CB}|=\{\overrightarrow{CB}\}=-\{\overrightarrow{BC}\}.$$

Por ello, de (9.3) se deduce

$$-\{\overrightarrow{AC}\}+\{\overrightarrow{AB}\}=-\{\overrightarrow{BC}\},$$

lo que coincide, en esencia, con (9.2).

En la demostración hemos usado únicamente las correlaciones (9.3) y (9.4) que sólo dependen de la disposición mutua de los puntos A, B, C en el eje, pero no dependen de si dichos puntos coinciden o no uno con el otro. Está claro que con cualquier otra disposición de los puntos la demostración se realiza del modo análogo.

La identidad (9.2) se llama *fundamental*. Desde el punto de vista de la operación de adición de los vectores dispuestos en un eje, la identidad significa que

$$\{\overrightarrow{AB}+\overrightarrow{BC}\}=\{\overrightarrow{AB}\}+\{\overrightarrow{BC}\}. \tag{9.5}$$

La magnitud de un segmento dirigido determina, salvo el traslado paralelo, el propio segmento en el eje. Mas, al tomar en consideración que los segmentos dirigidos iguales también quedan determinados, salvo el traslado paralelo, quiere decir que la magnitud de un segmento dirigido define unívocamente en el eje dado toda la totalidad de segmentos dirigidos iguales.

Supongamos ahora que se han dado el segmento dirigido $\overrightarrow{AB}$ y el número α. Se denomina *producto* $\alpha\cdot\overrightarrow{AB}$ *del segmento dirigido* $\overrightarrow{AB}$ *por un número real* α un segmento dirigido, ubicado en el eje que pasa por los puntos A, B, y cuya magnitud es igual a $\alpha\cdot\{\overrightarrow{AB}\}$. De este modo, por definición

$$\{\alpha\cdot\overrightarrow{AB}\}=\alpha\cdot\{\overrightarrow{AB}\}. \tag{9.6}$$

Para cualesquiera números α, β y cualesquiera segmentos dirigidos a, b la operación de multiplicación del segmento dirigido por un número posee las siguientes propiedades:

$$1\cdot a=a, \quad \alpha(\beta a)=(\alpha\beta)a,$$

$$(\alpha+\beta)a=\alpha a+\beta a, \qquad \alpha(a+b)=\alpha a+\alpha b.$$

Las primeras tres propiedades son bastante sencillas. Para demostrarlas basta notar que en los miembros izquierdos y derechos de las igualdades figuran los vectores dispuestos en un eje y hacer uso, luego, de las correlaciones (9.5), (9.6). Demostremos la cuarta propiedad. Supongamos, para simplificar, que $\alpha>0$. Apliquemos los vectores a, b a un punto común y construyamos sobre ellos un paralelogramo cuya diagonal será igual a $a+b$ (fig. 9.2). Al multiplicar los vectores a, b por α, la diagonal del paralelogramo resultará también multiplicada por α, en virtud de la semejanza de las figuras. Pero esto precisamente significa que $\alpha a + \alpha b = \alpha(a + b)$.

Observemos en conclusión que la introducción de la magnitud de un segmento dirigido puede interpretarse como introducción de cierta "función"

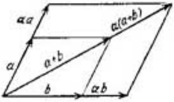


Fig. 9.2.

$$\xi = \{x\}, \tag{9.7}$$

cuyo "argumento" son los vectores x de un mismo eje, mientras que el papel de "valor" lo desempeñan los números reales ξ. En este caso

$$\{x+y\} = \{x\} + \{y\},$$
$$\{\lambda x\} = \lambda\{x\}$$

para cualesquiera vectores x e y en el eje y cualquier número λ.

Ejercicios.

1. Demuéstrese que el resultado de la operación de multiplicación por un número no depende de cómo está definida la dirección positiva en el eje.
2. Demuéstrese que el resultado de la operación de multiplicación por un número no depende de cómo está definido el segmento graduado en el eje.
3. Demuéstrese que la operación de multiplicación por un número, definida en cualquier conjunto de segmentos colineales, no hará salir de dicho conjunto.
4. Demuéstrese que la operación de multiplicación por un número, definida en cualquier conjunto de segmentos coplanares, no hará salir de dicho conjunto.
5. ¿Qué representan en sí los segmentos dirigidos opuesto y nulo desde el punto de vista de la operación de multiplicación por un número?

§ 10. Espacios lineales

La resolución de los problemas de cualquier género se reduce, al fin y al cabo, al estudio de ciertos conjuntos y, en primer lugar, al estudio de la estructura de dichos conjuntos. Para estudiar la estructura de conjuntos se emplean los más diversos métodos. Por ejemplo, partiendo de la propiedad característica que poseen los elementos (lo que se hace en los problemas referentes a la construcción de lugares geométricos) o bien partiendo de las propiedades de las operaciones, si están definidas para los elementos.

El último método parece ser de atracción especial debido a su generalidad. Efectivamente, ya hemos visto repetidamente que en los más diversos conjuntos pueden introducirse las más diversas operaciones que, no obstante, poseen propiedades iguales. Será evidente por esta razón que si en la investigación de los conjuntos se obtiene cierto resultado sólo en la base de las propiedades de una operación, este resultado tendrá lugar en todos los conjuntos, donde las operaciones poseen las mismas propiedades. En este caso la naturaleza concreta tanto de los elementos como de las operaciones sobre ellos puede ser completamente diferente.

Anteriormente hemos introducido para la consideración unos nuevos objetos matemáticos que llamamos segmentos dirigidos o vectores y hemos determinado las operaciones sobre dichos objetos. Se conoce que en realidad detrás de los vectores están los objetos físicos bien reales. Por esto, la investigación detallada de la estructura de los conjuntos representa interés por lo menos para la física.

Ya ahora conocemos tres tipos de conjuntos en los cuales las operaciones poseen las mismas propiedades. Estos conjuntos son: de vectores colineales, de vectores coplanares y de vectores en todo el espacio. A pesar de que en estos conjuntos están introducidas las mismas operaciones, tenemos derecho de esperar que la estructura de los propios conjuntos debe ser diferente.

Hay una tentación, originada por la sencillez de los conjuntos citados, que conduce al deseo de estudiarlos apoyándose sólo en las peculiaridades concretas de los elementos. Sin embargo, no podemos sino observar el hecho de que dichos conjuntos tienen mucho en común, razón por la cual parece racional comenzar su estudio partiendo de ciertas posiciones generales, abrigando esperanza, por lo menos, que *tendremos éxito en evitar repeticiones fastidiosas y monótonas* al pasar de la investigación de un conjunto a la del otro. En adición, esperamos, por supuesto, que en el caso de aparecer algún conjunto con propiedades análogas, a éste último *podríamos extender inmediatamente todos los resultados de las investigaciones realizadas.*

He aquí todos los hechos generales que se conocen acerca de los vectores que forman cualquiera de los tres conjuntos en consideración.

A. A todo par de vectores x, y le corresponde un vector $x + y$, llamado suma de x e y, con la particularidad de que:

1) la adición es conmutativa, $x + y = y + x$,

2) la adición es asociativa, $x + (y + z) = (x + y) + z$,

3) existe un único vector nulo $\mathbf{0}$ tal que $x + \mathbf{0} = x$ para cualquier vector x.

4) para todo vector x existe un único vector opuesto $-x$ tal que $x + (-x) = \mathbf{0}$.

B. A todo par α, x (donde α es un número y x, un vector) le corresponde el vector αx, llamado producto de α y x, con la particularidad de que:

1) la multiplicación por los números es asociativa, $\alpha(\beta x) = (\alpha\beta)x$,

2) $1 \cdot x = x$ para todo vector x.

C. Las operaciones de adición y multiplicación están ligadas entre sí por las siguientes correlaciones:

1) la multiplicación por un número es distributiva respecto de la adición de los vectores, $\alpha(x + y) = \alpha x + \alpha y$.

2) la multiplicación por un vector es distributiva respecto de la adición de los números, $(\alpha + \beta) x = \alpha x + \beta x$.

Al igual que en el caso de un campo, los hechos citados no pretenden ser independientes lógicamente. Las propiedades **A** describen el conjunto de vectores desde el punto de vista de la operación de adición y dicen que respecto de esta operación el conjunto representa un grupo abeliano. Las propiedades **B** describen el conjunto de vectores desde el punto de vista de la operación de multiplicación de un vector por un número. Las propiedades **C** describen los lazos existentes entre las dos operaciones.

Examinemos ahora un conjunto K y un campo P de carácter arbitrario. El conjunto K se llamará *espacio lineal* sobre el campo P, si para todos los elementos de K están definidas las operaciones de adición y multiplicación por los números de P y están cumplidos los axiomas **A**, **B**, **C**. De conformidad con esta terminología, podemos declarar que el conjunto de vectores colineales, el conjunto de vectores coplanares y el de vectores en todo el espacio son espacios lineales sobre el campo de números reales.

Los elementos de *cualquier* espacio lineal los llamaremos *vectores*, a pesar de que según su naturaleza concreta dichos elementos pueden ser bien distintos de los segmentos dirigidos. Las representaciones geométricas, relacionadas con el nombre de "vectores", nos ayudarán en aclarar y, con frecuencia, en prever los resultados necesarios, como también en buscar la interpretación geométrica de diferentes hechos la cual no siempre resulta obvia.

Los vectores del espacio lineal se designarán, como antes, mediante letras latinas minúsculas y los números, con letras minúsculas griegas. Un espacio lineal se llamará *racional*, *real* o *complejo*, según sea el campo P de numeros racionales, de números reales o de los complejos y se designará mediante **D**, **R** y **C**, respectivamente. El hecho de que en la denominación y la designación no se da ninguna referencia a los elementos del propio espacio tiene un sentido muy profundo de lo cual hablaremos mucho más tarde.

Antes de pasar a la investigación detallada de los espacios lineales, demos a conocer algunos corolarios más sencillos que provienen de la existencia de las operaciones de adición y multiplicación por un número. Conciernen, en lo esencial, a los vectores nulo y opuesto.

En *todo* espacio lineal para cualquier elemento x se verifica la igualdad

$$0 \cdot x = 0,$$

donde en el segundo miembro 0 significa el vector nulo y en el primer miembro 0 es el número nulo. Para demostrar esta correlación examinemos el elemento $0\cdot x + x$. Tenemos

$$0\cdot x + x = 0\cdot x + 1\cdot x = (0 + 1)\, x = 1\cdot x = x.$$

Por consiguiente,

$$x = 0\cdot x + x.$$

Al sumar a ambos miembros de la igualdad el elemento $-x$, hallamos

$$0 = x + (-x) = (0\cdot x + x) + (-x) =$$
$$= 0\cdot x + (x + (-x)) = 0\cdot x + 0 = 0\cdot x.$$

Ahora ya es fácil indicar la expresión *explícita* para el elemento opuesto $-x$ en términos del propio elemento x. A saber,

$$-x = (-1)\, x.$$

La validez de esta fórmula se deduce de las correlaciones sencillas

$$x + (-1)\, x = 1\cdot x + (-1)\, x = (1 - 1)\, x = 0\cdot x = 0.$$

Esto, a su vez, permite demostrar la validez de las correlaciones

$$-(\alpha x) = (-\alpha)\, x = \alpha\, (-x),$$

puesto que

$$-(\alpha x) = (-1)\,(\alpha x) = (-\alpha)\, x = \alpha\, ((-1)\, x) = \alpha\, (-x)$$

Recordemos que por definición de operación de sustracción, $x - y = x + (-y)$ para cualesquiera vectores x e y. La expresión explícita para el vector opuesto permite señalar que las leyes distributivas son legítimas también para *la diferencia*. En efecto, cualesquiera que sean los números α, β y los vectores x, y, tendremos

$$(\alpha - \beta)\, x = \alpha x + (-\beta)\, x = \alpha x + (-(\beta x)) = \alpha x - \beta x,$$
$$\alpha\, (x - y) = \alpha\, (x + (-1)\, y) = \alpha x + (-\alpha)\, y =$$
$$= \alpha x + (-(\alpha y)) = \alpha x - \alpha y.$$

De aquí proviene, en particular, que para *todo* número

$$\alpha\cdot 0 = 0,$$

puesto que

$$\alpha\cdot 0 = \alpha\, (x - x) = \alpha x - \alpha x = \alpha x + (-\alpha x) = 0.$$

Y, por fin, el último corolario. Si para algún número α y el vector x se verifica la correlación

$$\alpha x = 0, \tag{10.1}$$

entonces, o *bien* $\alpha = 0$ o *bien* $x = 0$. Efectivamente, si la igualdad (10.1) se realiza, puede haber una de las dos posibilidades: o bien $\alpha = 0$ o bien $\alpha \neq 0$. El caso $\alpha = 0$ confirma nuestra afirmación. Sea, ahora, $\alpha \neq 0$. En este caso

$$x = 1 \cdot x = \left(\frac{1}{\alpha} \cdot \alpha\right) x = \frac{1}{\alpha} (\alpha x) = \frac{1}{\alpha} \cdot 0 = 0.$$

De aquí se desprende, en particular, que en todo espacio lineal cualquier igualdad puede *ser reducida* formalmente por un factor común no nulo, no importa que dicho factor sea un número o un vector. Efectivamente, si es que $\alpha x = \beta x$ y $x \neq 0$, tenemos $(\alpha - \beta) x = 0$, y entonces $\alpha - \beta = 0$, es decir, $\alpha = \beta$. Si, en cambio, $\alpha x = \alpha y$ y $\alpha \neq 0$, se tiene $\alpha (x - y) = 0$ y entonces $x - y = 0$, es decir, $x = y$.

Así pues, *desde el punto de vista de las operaciones de multiplicación, adición y sustracción tienen lugar formalmente todas las reglas de transformaciones equivalentes para las expresiones algebraicas.* En lo que sigue estas reglas ya no serán objeto de especificaciones especiales.

Así damos por terminado el primer conocimiento con los espacios lineales. Sólo nos resta observar que no es por casualidad que hemos escrito en la forma única las propiedades de operaciones en el campo y en el espacio lineal. Existen rasgos de la semejanza sorprendente (y de la diferencia, en igual medida) entre los axiomas de un campo y de un espacio lineal sobre el campo. El lector ha de reflexionar en esta circunstancia.

Ejercicios.

Demuéstrese que los conjuntos a seguir son espacios lineales. La denominación de la operación refleja siempre no la notación, sino el contenido.

1. El campo: números reales; el conjunto: números reales; la adición: adición de los números reales; la multiplicación por un número: multiplicación de un número real por un número real.

2. El campo: números reales; el conjunto: números complejos; la adición: adición de los números complejos; la multiplicación por un número: multiplicación de un número complejo por un número real.

3. El campo: números racionales; el conjunto: números reales; la adición: adición de los números reales; la multiplicación por un número: multiplicación de un número real por un número racional.

4. El campo: número cualquiera; el conjunto: un solo vector a; la adición: adición se determina por la regla $a + a = a$; la multiplicación por un número: multiplicación del vector a por un número cualquiera α se determina mediante la regla $\alpha a = a$.

5. El campo: números reales; el conjunto: polinomios con coeficientes reales de una variable t, incluidas las constantes; la adición: adición de los polinomios; la multiplicación por un número: multiplicación del polinomio por un número real.

6. El campo: números racionales; el conjunto: números de la forma $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{5}$, donde a, b, c y d son unos números racionales; la adición: adición de los números de la forma indicada; la multiplicación por un número: multiplicación del número de la forma indicada por un número racional.

7. El campo: campo cualquiera; el conjunto: el mismo campo; la adición: adición de los elementos (¡de los vectores!) del campo; la multiplicación por un número: multiplicación del elemento (¡del vector!) del campo por un elemento (¡un número!) del campo.

§ 11. Sumas finitas y productos finitos

Los campos y los espacios lineales serán conjuntos principales que trataremos en lo sucesivo. En estos conjuntos están introducidas dos operaciones: la adición y la multiplicación. Si se realizan un gran número de operaciones sobre los elementos, aparecen expresiones que contienen en cantidad considerable tanto sumandos como factores. Para asegurar la comodidad de su inscripción introduzcamos símbolos correspondientes. Supondremos, además, que la adición y la multiplicación son operaciones *conmutativas* y *asociativas*.

Sea dado un número finito de los elementos, no necesariamente distintos. Convengamos en considerar que todos los elementos están enumerados de cierta manera y se les han atribuido unos números que varían seguidamente a partir del número k hasta algún número p. Los elementos se designarán mediante una letra, indicando el número. El propio número, el cual se llamará en adelante *índice*, puede ocupar en la designación un lugar arbitrario. El índice puede disponerse al lado de la letra encerrado entre paréntesis, abajo cerca de la letra, arriba cerca de la letra, etc. Esto no tiene ninguna importancia. Con mayor frecuencia lo escribiremos al lado de la letra, abajo, a la derecha.

Una suma de elementos $a_k, a_{k+1}, \ldots, a_p$ se designará mediante el símbolo de la forma siguiente

$$a_k + a_{k+1} + \ldots a_p = \sum_{i=k}^{p} a_i. \tag{11.1}$$

El índice i en esta fórmula se denomina *índice de adición*. Nada variará, por supuesto, si lo designamos con cualquier otra letra. A veces, bajo el signo de la suma se indicará de manera explícita aquella totalidad de los índices según los cuales se realiza la adición. Por ejemplo, la suma en consideración podría ser escrita de la forma:

$$a_k + a_{k+1} + \ldots + a_p = \sum_{k \leqslant i \leqslant p} a_i.$$

Es evidente que si todo elemento a_i es igual al producto del elemento b_i y elemento α, donde α no depende del índice de adición i, entonces

$$\sum_{i=k}^{p} \alpha b_i = \sum_{i=k}^{p} b_i,$$

es decir, el factor que no depende del índice de adición puede sacarse fuera del signo de la suma.

Supongamos ahora que los elementos están marcados con dos índices cada uno de los cuales varía independientemente. Aceptemos para estos elementos una designación general a_{ij} y sea, por ejemplo, $k \leqslant i \leqslant p$; $m \leqslant j \leqslant n$. Dispongamos los elementos en forma de una tabla rectangular

$$\begin{array}{llll} a_{km} & a_{k,\,m+1} & \ldots & a_{kn}, \\ a_{k+1,\,m} & a_{k+1,\,m+1} & \ldots & a_{k+1,\,n}, \\ \ldots & \ldots & \ldots & \ldots \\ a_{pm} & a_{p,\,m+1} & \ldots & a_{pn}. \end{array}$$

Está claro que independientemente del orden en que se realiza la sumación, el resultado será el mismo. Por ello, al tomar en consideración la designación introducida para la suma, tenemos

$$(a_{km}+a_{k,\,m+1}+\ldots+a_{kn})+(a_{k+1,\,m}+a_{k+1,\,m+1}+\ldots$$
$$\ldots+a_{k+1,\,n})+\ldots+(a_{pm}+a_{p,\,m+1}+\ldots+a_{pn})=$$
$$=\sum_{j=m}^{n} a_{kj}+\sum_{j=m}^{n} a_{k+1,\,j}+\ldots+\sum_{j=m}^{n} a_{pj}=\sum_{i=k}^{p}\Big(\sum_{j=m}^{n} a_{ij}\Big).$$

Por otro lado, esta misma suma es igual a

$$(a_{km}+a_{k+1,\,m}+\ldots+a_{pm})+(a_{k,\,m+1}+$$
$$+a_{k+1,\,m+1}+\ldots+a_{p,\,m+1})+\ldots+(a_{kn}+a_{k+1,\,n}+\ldots+a_{pn})=$$
$$=\sum_{i=k}^{p} a_{im}+\sum_{i=k}^{p} a_{i,\,m+1}+\ldots+\sum_{i=k}^{p} a_{in}=\sum_{j=m}^{n}\Big(\sum_{i=k}^{p} a_{ij}\Big).$$

Por consiguiente,

$$\sum_{i=k}^{p}\Big(\sum_{j=m}^{n} a_{ij}\Big)=\sum_{j=m}^{n}\Big(\sum_{i=k}^{p} a_{ij}\Big).$$

Si convenimos en realizar la adición siempre de manera sucesiva según los índices de las sumas dispuestas de derecha a la izquierda, los paréntesis pueden ser omitidos y, en definitiva, obtenemos

$$\sum_{i=k}^{p}\sum_{j=m}^{n} a_{ij}=\sum_{j=m}^{n}\sum_{i=k}^{p} a_{ij}.$$

Quiere decir que al sumar según dos índices se puede cambiar el orden de la adición. Si, por ejemplo, $a_{ij}=\alpha_i b_{ij}$, donde α_i no depende del índice j, entonces

$$\sum_{i=k}^{p}\sum_{j=m}^{n} \alpha_i b_{ij}=\sum_{i=k}^{p} \alpha_i \sum_{j=m}^{n} b_{ij}.$$

Los resultados análogos tienen lugar para las sumas realizadas según cualquier número finito de índices.

El producto de los elementos $a_k, a_{k+1}, \ldots, a_p$ se designará mediante el símbolo de tal forma:

$$a_k a_{k+1} \ldots a_p = \prod_{i=k}^{p} a_i.$$

Luego, si $a_i = \alpha b_i$, entonces

$$\prod_{i=k}^{p} \alpha b_i = \alpha^{p-k+1} \prod_{i=k}^{p} b_i.$$

Igual que en el caso de sumación, al calcular un producto según dos índices, se puede cambiar el orden de cálculo del producto, es decir,

$$\prod_{i=k}^{p} \prod_{j=m}^{n} a_{ij} = \prod_{j=m}^{n} \prod_{i=k}^{p} a_{ij}.$$

Todos estos hechos se demuestran según el mismo esquema que se usa en el caso de sumación de los números.

Ejercicios.

Calcúlense las siguientes expresiones:

$$\sum_{i=1}^{n} 1, \quad \sum_{i=1}^{n} i, \quad \sum_{i=1}^{n} i^2, \quad \sum_{i=1}^{n} 8(i-1),$$

$$\sum_{r=1}^{n} \sum_{j=1}^{m} rj, \qquad \sum_{i=1}^{n} \sum_{j=1}^{m} (i+5j),$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{k=1}^{p} (2j-i)^2 k,$$

$$\prod_{i=1}^{n} 2, \quad \prod_{p=1}^{n} 10^p, \quad \prod_{i=1}^{n} \prod_{j=1}^{n} 2^{i-j}, \quad \prod_{i=1}^{n} \prod_{j=1}^{m} \prod_{k=1}^{p} 2^{i+j+k}.$$

§ 12. Cálculos aproximados

Los conjuntos examinados son de amplio uso en las más diversas investigaciones teóricas. Para obtener el resultado en dichas investigaciones hemos de realizar casi siempre ciertas operaciones sobre los elementos de los conjuntos. Muy a menudo surge la necesidad en la realización de cálculos con elementos de los campos numéricos. Desearíamos atraer la atención a *una peculiaridad* muy importante *en la realización práctica* de los cálculos semejantes.

Supongamos, para concretar, que tenemos un campo de números reales. Supongamos, además, que todo número se representa como una fracción decimal infinita. Ni un ser humano ni una computadora más moderna pueden operar con las fracciones infinitas. Por esta razón, en la práctica cada fracción de tal índole se sustituye por una fracción decimal finita, próxima a la primera, o por un número racional conveniente.

Así pues, un número real exacto es sustituido por un número aproximado. En las investigaciones teóricas, donde se sobreentiende la prefijación exacta de los números, tal o cual expresión se sustituye bastante a menudo por otra expresión, igual a la primera, aunque escrita, quizás, en una forma distinta. Por supuesto, en este caso tal sustitución no puede despertar objeciones ni tampoco las dudas. En cambio, si deseamos calcular una expresión con ayuda de unos números prefijados aproximadamente, ya no es indeferente para nosotros en qué forma se ha dado la expresión.

Consideremos un ejemplo sencillo. Es fácil comprobar que en el caso de prefijar exactamente el número $\sqrt{2}$ se tiene

$$(\sqrt{2}-1)^6 = 99 - 70\sqrt{2}. \tag{12.1}$$

Puesto que $\sqrt{2} = 1{,}4142\ldots$, los números $7/5 = 1{,}4$ y $17/12 = 1{,}4166\ldots$ pueden considerarse valores aproximados para $\sqrt{2}$. Pero, al sustituir 7/5 en los miembros izquierdo y derecho de (12.1), obtenemos 0,00509 . . . y 1,0, respectivamente. Para 17/12 tenemos 0,00523 . . . y —0,1666 . . . Los resultados de la sustitución son muy distintos uno del otro y no se ve de una tirada cuál de ellos es más próximo al valor exacto. Esto nos muestra *la precaución con la que se deben tratar los números aproximados.*

Nos hemos detenido sólo en una fuente de aparición de los números aproximados que es el redondeo de los números dados con toda la exactitud. En la realidad existen también varias otras fuentes. Por ejemplo, los datos iniciales para los cálculos se obtienen con frecuencia de un experimento, mientras que cualquier experimento puede proporcionar resultados de exactitud limitada. Ya las operaciones más simples, como son la multiplicación y división, pueden conducir al aumento de la cantidad de órdenes en las fracciones. Por ello, nos veremos obligados a despreciar una parte de los órdenes en los resultados de los cálculos intermedios, es decir, aquí también ciertos números se sustituyen por los aproximados, etc.

La investigación detallada de las operaciones con los números aproximados sale de los márgenes de este curso. No obstante, la diferencia entre los cálculos teóricos y prácticos será con frecuencia el objeto de nuestra discusión. La necesidad en tal discusión está provocada por lo que *los cálculos teóricos no pueden ser realizados, como regla, en la forma precisa.*

Ejercicios.

1. ¿Cuál fracción decimal finita es necesaria para aproximar $\sqrt{2}$ para que en los resultados del cálculo de los miembros izquierdo y derecho en (12.1) coincidan los primeros seis órdenes?

2. Supongamos que el resultado de realización de toda operación sobre dos números reales se redondea de acuerdo con cualquier regla conocida hasta t órdenes después de la coma. ¿Queden en vigor las propiedades de conmutatividad y asociatividad de las operaciones?

3. ¿Se cumplirán en las condiciones del ejercicio antecedente las leyes distributivas?

4. ¿A qué deducción llega Ud., si en los ejercicios 2 y 3 se han obtenido respuestas negativas?

CAPÍTULO 2 ESTRUCTURA DEL ESPACIO LINEAL

§ 13. Combinaciones lineales y cápsulas

Supongamos que en el espacio lineal K, definido sobre un campo P, se ha elegido un número finito de vectores arbitrarios $e_1, e_2, \ldots, e_n$ que no son obligatoriamente distintos. Llamaremos estos vectores *sistema de vectores*. Un sistema de vectores se denominará *subsistema* del segundo sistema, si el primer sistema sólo contiene ciertos vectores del segundo y no contiene ningunos otros vectores.

Sobre los vectores del sistema dado y los vectores obtenidos de los primeros se realizarán las operaciones de adición y multiplicación por unos números. Está claro que todo vector x del tipo

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n, \tag{13.1}$$

donde $\alpha_1, \alpha_2, \ldots, \alpha_n$ son unos números del campo P, se obtiene de los vectores del sistema dado $e_1, e_2, \ldots, e_n$ con ayuda de las operaciones citadas. Más aún, cualquiera que sea el orden en que se realicen estas operaciones, obtendremos solamente los vectores del tipo (13.1).

Respecto del vector x de (13.1) suele decirse que *se expresa linealmente* en términos de los vectores $e_1, e_2, \ldots, e_n$. El segundo miembro de (13.1) se denomina *combinación lineal* de estos vectores y los números $\alpha_1, \alpha_2, \ldots, \alpha_n$ son *coeficientes de la combinación lineal*.

Fijemos el sistema de vectores $e_1, e_2, \ldots, e_n$ y dejemos que los coeficientes de las combinaciones lineales tomen cualesquiera valores del campo P. En este caso quedará definido cierto conjunto de vectores de K. Este conjunto lleva el nombre de *cápsula lineal* de los vectores $e_1, e_2, \ldots, e_n$ y se designa con $L(e_1, e_2, \ldots, e_n)$.

Nuestro interés hacia las cápsulas lineales se debe a dos circunstancias. En primer lugar, toda cápsula lineal tiene una estructura muy simple: representa en sí una totalidad de todas las combinaciones lineales de vectores del sistema dado. Segundo, la cápsula lineal de cualquier sistema de vectores de todo espacio lineal es de por sí un espacio lineal.

Efectivamente, el cumplimiento de todos los axiomas de un espacio lineal es casi obvio. Sólo los axiomas referentes al vector nulo y al opuesto necesitan, quizás, algunas explicaciones. El vector nulo pertenece a ciencia cierta a cualquier cápsula lineal y corresponde a los valores nulos de los coeficientes de la combinación lineal, es decir,

$$0 = 0 \cdot e_1 + 0 \cdot e_2 + \ldots + 0 \cdot e_n.$$

El vector, opuesto para el vector (13.1), tendrá la forma

$$-x = (-\alpha_1)\, e_1 + (-\alpha_2)\, e_2 + \ldots + (-\alpha_n)\, e_n.$$

La unicidad de los vectores nulo y opuesto proviene de su unicidad como vectores pertenecientes al espacio lineal K.

Observemos que la cápsula lineal de los vectores $e_1, e_2, \ldots, e_n$ es un espacio lineal "*mínimo*" que contiene dichos vectores. En efecto, la propia cápsula lineal sólo consiste de las combinaciones lineales de los vectores $e_1, e_2, \ldots, e_n$, mientras que todo espacio lineal, en el que están contenidos los vectores $e_1, e_2, \ldots, e_n$, ha de contener también todas sus combinaciones lineales.

Así pues, todo espacio lineal contiene, en el caso general, un conjunto infinito de otros espacios lineales representados por las cápsulas lineales. Ahora surgen unas preguntas:

¿En qué condiciones las cápsulas lineales de dos sistemas *diferentes* de vectores consisten de *los mismos* vectores del espacio inicial?

¿Qué número *mínimo* de vectores define una misma cápsula lineal?

¿Será el espacio lineal inicial una cápsula lineal de algunos de *sus* vectores?

Las respuestas a estas preguntas y otras las daremos en el futuro más próximo. Con este fin emplearemos en gran escala la noción de combinación lineal y, en particular, la propiedad de su *transitividad*. A saber, si un cierto vector z es una combinación lineal de los vectores $x_1, x_2, \ldots, x_r$, y cada uno de ellos, a su turno, es una combinación lineal de los vectores $y_1, y_2, \ldots, y_s$, entonces el vector z también puede ser representado como combinación lineal de los vectores $y_1, y_2, \ldots, y_s$. Demostremos esta propiedad. Sea

$$z = \sum_{i=1}^{r} \beta_i x_i \tag{13.2}$$

y, además, para todos los números i, $1 \leqslant i \leqslant r$,

$$x_i = \sum_{j=1}^{s} \gamma_{ij} y_j,$$

donde β_i y γ_{ij} son unos números arbitrarios del campo P.

Al sustituir la expresión para x_i en el segundo miembro de (13.2) y al hacer uso de las propiedades correspondientes de las sumas finitas, obtenemos

$$z=\sum_{i=1}^{r}\beta_i x_i=\sum_{i=1}^{r}\beta_i\sum_{j=1}^{s}\gamma_{ij}y_j=\sum_{i=1}^{r}\sum_{j=1}^{s}\beta_i\gamma_{ij}y_j=$$

$$=\sum_{j=1}^{s}\sum_{i=1}^{r}\beta_i\gamma_{ij}y_j=\sum_{j=1}^{s}\Big(\sum_{i=1}^{r}\beta_i\gamma_{ij}\Big)y_j=\sum_{j=1}^{s}\xi_j y_j,$$

donde los coeficientes ξ_j significan las siguientes expresiones:

$$\xi_j=\sum_{i=1}^{r}\beta_i\gamma_{ij}.$$

Por consiguiente, la transitividad de la noción de combinación lineal realmente tiene lugar.

Ejercicios.

1. ¿Qué son en un espacio de segmentos dirigidos las cápsulas lineales de los sistemas de uno, dos, tres y de mayor número de segmentos dirigidos?
2. Examinemos un espacio lineal de polinomios que dependen de la variable t y están definidos sobre el campo de números reales. ¿Qué representa en sí la cápsula lineal del sistema de vectores t^2+1, t^2+t, 1?
3. ¿En qué espacio todas las cápsulas lineales coinciden con el mismo espacio?
4. Demuéstrese que un espacio lineal de todos los segmentos dirigidos no puede ser cápsula lineal de dos segmentos dirigidos cualesquiera.

§ 14. Dependencia lineal

Consideraremos otra vez los vectores arbitrarios $e_1, e_2, \ldots, e_n$ en un espacio lineal. Puede ocurrir que uno de ellos se expresa linealmente en términos de los demás. Sea, por ejemplo, el vector e_1. Entonces, cada uno de los vectores $e_1, e_2, \ldots, e_n$ se expresa linealmente en términos de $e_2, e_3, \ldots, e_n$. Por esta razón cualquier combinación lineal de los vectores $e_1, e_2, \ldots, e_n$ es también una combinación lineal de los vectores $e_2, e_3, \ldots, e_n$. Por consiguiente, las cápsulas lineales de los vectores $e_1, e_2, \ldots, e_n$ y $e_2, e_3, \ldots, e_n$ coinciden.

Supongamos luego que entre los vectores $e_2, e_3, \ldots, e_n$ hay un vector, por ejemplo, e_2 que también se expresa linealmente en términos de los vectores restantes. Al repetir nuestros razonamientos, llegamos a la conclusión de que ahora cualquier combinación lineal de los vectores $e_1, e_2, \ldots, e_n$ es también una combinación lineal de los vectores $e_3, e_4, \ldots, e_n$. Continuando este proceso, pasaremos, al fin y al cabo, del sistema $e_1, e_2, \ldots, e_n$ a un sistema de vectores del cual ya *no podremos* excluir ni uno de los vectores. La cápsula lineal del nuevo sistema de vectores *coincide*, evidentemente, con la cápsula lineal de los vectores $e_1, e_2, \ldots, e_n$. Además, podemos decir que si entre $e_1, e_2, \ldots, e_n$ hubo aunque un solo vector no nulo, el nuevo *sistema de vectores o bien consiste solamente en un vector no nulo o bien ninguno de sus vectores se expresa linealmente en términos de los vectores restantes.*

Tal sistema de vectores se llama *linealmente independiente.*

Si un sistema de vectores no es linealmente independiente, se denomina *linealmente dependiente.* En particular, según se infiere de la definición, un sistema compuesto sólo por un vector nulo será linealmente dependiente. La dependencia y la independencia lineales constituyen las propiedades del sistema de vectores. Sin embargo, muy a menudo los adjetivos correspondientes se aplican también a los mismos vectores, razón por la cual en lugar del "sistema linealmente independiente de vectores" diremos, a veces, "sistema de los vectores linealmente independientes", etc.

Desde el punto de vista de las nociones introducidas los razonamientos realizados significan que se ha demostrado el

LEMA 14.1. *Si entre los vectores $e_1, e_2, \ldots, e_n$ no todos son nulos y si dicho sistema es linealmente dependiente, entonces en éste puede existir un subsistema linealmente independiente de vectores, en cuyos términos es linealmente expresado cualquiera de los vectores $e_1, e_2, \ldots, e_n$.*

Una circunstancia, inesperada a primera vista, determina si un sistema de vectores $e_1, e_2, \ldots, e_n$ es linealmente dependiente o linealmente independiente. Ya se ha observado que el vector nulo pertenece a la cápsula lineal y es representado a ciencia cierta por la combinación lineal (13.1) con valores nulos de los coeficientes. A pesar de esto, puede expresarse linealmente en términos de los vectores $e_1, e_2, \ldots, e_n$ de otro método, es decir, determinarse por otro surtido de los coeficientes de la combinación lineal. La independencia lineal de los vectores $e_1, e_2, \ldots, e_n$ está estrechamente vinculada con la unicidad de la representación del elemento nulo en términos de los vectores citados. A saber, es válido el

TEOREMA 14.1. *Un sistema de los vectores $e_1, e_2, \ldots, e_n$ es linealmente independiente cuando, y sólo cuando, de la igualdad*

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = 0 \tag{14.1}$$

se deduce que todos los coeficientes de la combinación lineal son nulos.

DEMOSTRACIÓN. Sea $n = 1$. Si $e_1 \neq 0$, entonces, según se ha mostrado antes, de la correlación $\alpha_1 e_1 = 0$ debe provenir que $\alpha_1 = 0$. En cambio, si de la igualdad $\alpha_1 e_1 = 0$ se desprende que el coeficiente α_1 es nulo, entonces, evidentemente, e_1 no puede ser nulo.

Examinemos ahora el caso en que $n \geqslant 2$. Sea el sistema de vectores linealmente independiente. Supongamos que la igualdad (14.1) resulta lícita con cierto surtido de los coeficientes, entre los cuales existe aunque uno que sea distinto de cero. Sea, por ejemplo, $\alpha_1 \neq 0$.

Entonces, de (14.1) obtenemos

$$e_1 = \left(-\frac{\alpha_2}{\alpha_1}\right) e_2 + \left(-\frac{\alpha_3}{\alpha_1}\right) e_3 + \ldots + \left(-\frac{\alpha_n}{\alpha_1}\right) e_n,$$

es decir, el vector e_1 se expresa linealmente en términos de los demás vectores del sistema. Esto contradice a la condición de independencia lineal del sistema, por lo cual no puede haber coeficientes no nulos entre aquellos que satisfacen (14.1).

Si de la igualdad (14.1) se deduce que todos los coeficientes son nulos, el sistema de vectores no puede ser linealmente dependiente. En efecto, supondremos lo contrario y sea, por ejemplo, un vector e_1 que se expresa linealmente en términos de los demás, es decir,

$$e_1 = \beta_2 e_2 + \beta_3 e_3 + \ldots + \beta_n e_n.$$

En este caso la igualdad (14.1) será satisfecha a ciencia cierta por los coeficientes $\alpha_1 = -1$, $\alpha_2 = \beta_2$, ..., $\alpha_n = \beta_n$, entre los cuales por lo menos uno no es nulo. El teorema queda demostrado.

Este teorema es de tan amplio uso en diversas investigaciones que muy a menudo su afirmación se considera simplemente como definición de independencia lineal.

He aquí dos propiedades sencillas de los sistemas de vectores vinculadas con la dependencia lineal.

LEMA 14.2. *Si algunos de los vectores $e_1, e_2, \ldots, e_n$ son linealmente dependientes, todo el sistema $e_1, e_2, \ldots, e_n$ será linealmente dependiente.*

DEMOSTRACIÓN Sin restringir la generalidad podemos considerar que los primeros vectores $e_1, e_2, \ldots, e_k$ son linealmente dependientes. Por consiguiente, existen tales números $\alpha_1, \alpha_2, \ldots, \alpha_k$, entre los cuales hay distintos de cero, que

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k = 0.$$

De aquí fluye la legitimidad de la igualdad

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k + 0 \cdot e_{k+1} + \ldots + 0 \cdot e_n = 0.$$

Mas, esta igualdad significa dependencia lineal de los vectores $e_1, e_2, \ldots, e_n$, puesto que entre los números $\alpha_1, \alpha_2, \ldots, \alpha_k$, $0, \ldots, 0$ hay algunos que no son nulos.

LEMA 14.3. *Si entre los vectores $e_1, e_2, \ldots, e_n$ hay aunque un solo vector nulo, todo el sistema $e_1, e_2, \ldots, e_n$ será linealmente dependiente.*

DEMOSTRACIÓN. En efecto, un sistema compuesto por un vector nulo es linealmente dependiente. Por ello, de la propiedad que acabamos de demostrar se infiere dependencia lineal de todo el sistema.

El teorema que sigue representa un resultado más importante referente a la dependencia lineal.

TEOREMA 14.2. *Los vectores $e_1, e_2, \ldots, e_n$ son linealmente dependientes cuando, y sólo cuando, o bien $e_1 = \mathbf{0}$ o bien cierto vector e_k, $2 \leqslant k \leqslant n$, sea una combinación lineal de los vectores antecedentes.*

DEMOSTRACIÓN. Supongamos que los vectores $e_1, e_2, \ldots, e_n$ son linealmente dependientes. Entonces en la igualdad (14.1) no todos los coeficientes son nulos. Sea α_k el último coeficiente no nulo. Si $k = 1$, quiere decir $e_1 = \mathbf{0}$. Sea ahora $k > 1$. En este caso de la igualdad

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_k e_k = \mathbf{0}$$

encontramos que

$$e_k = \left(-\frac{\alpha_1}{\alpha_k}\right) e_1 + \left(-\frac{\alpha_2}{\alpha_k}\right) e_2 + \ldots + \left(-\frac{\alpha_{k-1}}{\alpha_k}\right) e_{k-1}.$$

Con esto queda demostrada la necesidad de la afirmación enunciada en el teorema. La suficiencia es evidente, puesto que tanto el caso cuando $e_1 = 0$, como el caso en que e_k se expresa linealmente en términos de los vectores anteriores, significa la dependencia lineal de los primeros vectores en el sistema $e_1, e_2, \ldots, e_n$. De aquí se deduce la dependencia lineal de todo el sistema de vectores.

Ejercicios.

1. Demuéstrese que si algún vector de un espacio lineal se representa de modo único en forma de una combinación lineal de los vectores $e_1, e_2, \ldots, e_n$, entonces el sistema citado de vectores es linealmente independiente.
2. Demuéstrese que si el sistema de vectores $e_1, e_2, \ldots, e_n$ es linealmente independiente, entonces cualquier vector de la cápsula lineal de dichos vectores se representa de modo único en forma de su combinación lineal.
3. Demuéstrese que el sistema de vectores $e_1, e_2, \ldots, e_n$ es linealmente dependiente cuando, y sólo cuando, o bien $e_n = \mathbf{0}$, o bien cierto vector e_k, $1 \leqslant k \leqslant n-1$, es una combinación lineal de los vectores posteriores.
4. Consideraremos un espacio lineal de polinomios que dependen de la variable t y están definidos sobre un campo de números reales. Demuéstrese que el sistema de vectores $1, t, t^2, \ldots, t^n$ es linealmente independiente para todo n.
5. Demuéstrese que un sistema de dos segmentos dirigidos no colineales es linealmente independiente.

§ 15. Sistemas equivalentes de vectores

Examinaremos dos sistemas de vectores de un espacio lineal K. Supongamos que las cápsulas lineales de estos sistemas coinciden y constituyen cierto conjunto L. Al conjunto L le pertenece a ciencia cierta cualquiera de los vectores de ambos sistemas y, además, cada uno de los vectores de L puede ser representado, en el caso dado, en forma de la combinación lineal tanto de los vectores de un sistema como de los vectores del otro. Por consiguiente:

Dos sistemas de vectores poseen la propiedad de que cualquier vector de cada sistema se expresa linealmente en términos de los vectores del otro sistema.

Los sistemas de tal tipo se llaman *equivalentes.*

De lo dicho se desprende que si las cápsulas lineales de dos sistemas de vectores coinciden, entonces los sistemas son equivalentes. Sean dados, ahora, cualesquiera dos sistemas equivalentes. En virtud de que la noción de combinación lineal es transitiva, toda combinación lineal de los vectores de un sistema puede representarse como combinación lineal de los vectores del otro sistema, es decir, las cápsulas lineales de ambos sistemas coinciden. Así pues, es válido el

LEMA 15.1. *Para que las cápsulas lineales de dos sistemas de vectores coincidan, es necesario y suficiente que dichos sistemas sean equivalentes.*

Hemos de notar que el concepto de equivalencia de dos sistemas de vectores es *una relación de equivalencia*. La reflexividad es obvia, puesto que todo sistema es equivalente a sí mismo; la simetría fluye de la definición de sistemas equivalentes y la transitividad de este concepto proviene de la transitividad de la noción de combinación lineal. Por esta razón *el conjunto de todos los sistemas de vectores de un espacio lineal se puede dividir en clases de los sistemas equivalentes.* Es importante subrayar que a todos los sistemas de cada clase les corresponde *una y sólo una* cápsula lineal.

Sobre la proporción del número de vectores en los sistemas equivalentes nada puede decirse en el caso general. En cambio, si de los dos sistemas equivalentes aunque uno es linealmente independiente, podemos hacer deducciones definitivas acerca de la cantidad de los vectores. Estas deducciones se fundamentan en el

TEOREMA 15.1. *Si cada uno de los vectores de un sistema linealmente independiente* $e_1, e_2, \ldots, e_n$ se expresa linealmente en términos de los vectores $y_1, y_2, \ldots, y_m$, entonces $n \leqslant m$.

DEMOSTRACIÓN. Por hipótesis del teorema, el vector e_n se expresa linealmente en términos de los vectores $y_1, y_2, \ldots, y_m$, por lo cual el sistema

$$e_n, y_1, y_2, \ldots, y_m \tag{15.1}$$

es linealmente dependiente. El vector e_n no es nulo, razón por la cual, de acuerdo con el teorema 14.2, cierto vector y_l de (15.1) es una combinación lineal de los vectores anteriores. Al eliminar este vector, obtendremos el sistema:

$$e_n, y_1, \ldots, y_{l-1}, y_{l+1}, \ldots, y_m. \tag{15.2}$$

Ahora, haciendo uso de la transitividad de la noción de combinación lineal, es fácil mostrar que cada uno de los vectores $e_1, e_2, \ldots, e_n$ se expresa linealmente en términos de los vectores (15.2).

Adjuntemos a los vectores (15.2) por la izquierda el vector e_{n-1}. De nuevo llegamos a que el sistema

$$e_{n-1},\ e_n,\ y_1,\ \ldots,\ y_{l-1},\ y_{l+1},\ \ldots,\ y_m \tag{15.3}$$

es linealmente dependiente. El vector e_{n-1} no es nulo, por lo cual, de acuerdo con el teorema 14.2, uno de los vectores restantes (15.3) es una combinación lineal de los vectores anteriores. De este vector no puede servir e_n, puesto que en tal caso sería linealmente dependiente el sistema de dos vectores e_{n-1}, e_n y, por consiguiente, todo el sistema $e_1, e_2, \ldots, e_n$. De este modo, cierto vector y_j de (15.3) se expresa linealmente en términos de los anteriores. Si lo eliminamos, obtendremos de nuevo un sistema por cuyo intermedio se expresa linealmente cada uno de los vectores $e_1, e_2, \ldots, e_n$.

Continuando este proceso advertimos que los vectores $y_1, y_2, \ldots, y_m$ no pueden ser agotados antes de que se hayan adjuntado todos los vectores $e_1, e_2, \ldots, e_n$. En el caso contrario resulta que cada uno de los vectores $e_1, e_2, \ldots, e_n$ se expresa linealmente en términos de una parte de los vectores de este mismo sistema, es decir, todo el sistema ha de ser linealmente dependiente. Puesto que esto contradice a la hipótesis del teorema, de aquí se deduce que $n \leqslant m$.

Consideraremos unos corolarios de este teorema. Sean dados dos sistemas equivalentes de vectores linealmente independientes. De acuerdo con el teorema demostrado, cada uno de dichos sistemas contiene no más vectores que el otro. Por consiguiente:

Los sistemas equivalentes linealmente independientes se componen de un mismo número de vectores.

Luego, tomemos n vectores arbitrarios, construyamos en ellos una cápsula lineal y elijamos en ésta $n + 1$ vectores cualesquiera. Como el número de estos vectores es mayor que el de vectores dados, ellos no pueden ser linealmente independientes. Por esto:

Cualesquiera $n + 1$ vectores de la cápsula lineal de un sistema de n vectores son linealmente dependientes.

La afirmación del lema 14.1 significa en términos de los sistemas equivalentes que cualquiera que sea el sistema de los vectores que no son nulos simultáneamente, existe en él un subsistema linealmente independiente que es equivalente al sistema. Este subsistema lleva el nombre de *base* del sistema inicial.

Por supuesto, todo sistema puede tener más que una base. Todas las bases de los sistemas equivalentes son de por sí sistemas equivalentes. Del primer corolario del teorema 15.1 se infiere que constan de un mismo número de vectores. Este número es la característica de todos los sistemas equivalentes y se llama *rango*. Por definición, el rango de los sistemas de vectores nulos se considera igual a cero.

Examinemos ahora dos sistemas linealmente independientes que se componen de un mismo número de vectores. Sustituyamos un vector cualquiera del primer sistema por cierto vector del segundo sistema. A continuación, en el sistema obtenido sustituyamos otra vez uno de los vectores del primer sistema por alguno de los vectores restantes del segundo sistema, etc. El proceso de sustituciones se realizará hasta que el primer sistema sea sustituido por el segundo sistema. Si la sustitución se realiza de un modo arbitrario, los sistemas intermedios pueden resultar linealmente dependientes. No obstante, queda válido el

TEOREMA 15.2. *El proceso de la sustitución sucesiva puede realizarse de una manera tal que todos los sistemas intemedios serán lineamente independientes.*

DEMOSTRACIÓN. Sean dados dos sistemas linealmente independientes de los vectores $y_1, y_2, \ldots, y_n$ y $z_1, z_2, \ldots, z_n$. Supongamos que se han realizado k pasos del proceso, donde $k \geqslant 0$. Sin restringir la generalidad consideraremos que todos los vectores $y_1, \ldots, y_k$ están sustituidos por los vectores $z_1, \ldots, z_k$ y todos los sistemas obtenidos, incluido el sistema

$$z_1, \ldots, z_k, \; y_{k+1}, \ldots, y_n,$$

son linealmente independientes. Ésta suposición tiene lugar a ciencia cierta cuando $k = 0$.

Supongamos luego que durante la sustitución del vector y_{k+1} por cualquiera de los vectores $z_{k+1}, \ldots, z_n$ todos los sistemas

$$z_1, \ldots, z_k, z_i, y_{k+2}, \ldots, y_n$$

son linealmente dependientes para $i = k + 1, \ldots, n$. Dado que el sistema

$$z_1, \ldots, z_k, y_{k+2}, \ldots, y_n \qquad (15.4)$$

es linealmente independiente, resulta que los vectores z_i, para $i = k + 1, \ldots, n$, se expresan de modo lineal en términos de este sistema. Pero, a través del mismo se expresan linealmente, además, los vectores z_i cuando $i = 1, 2, \ldots, k$. Por consiguiente, a través del sistema (15.4) deben expresarse linealmente todos los vectores $z_1, \ldots, z_n$. Esto no es posible en virtud del teorema 15.1. Por esta razón, el proceso de sustitución enunciado en el teorema realmente tiene lugar.

Ejercicios.

Demuéstrese que las transformaciones del sistema de vectores a seguir, que se llaman *elementales*, conducen a un sistema equivalente.

1. La adjunción al sistema de vectores de cualquier combinación lineal de estos vectores.
2. La eliminación del sistema de vectores de cualquier vector que es una combinación lineal de los vectores restantes.
3. La multiplicación de cualquier vector de un sistema por un número distinto de cero.
4. La adición a cualquier vector de un sistema de cualquier combinación lineal de los vectores restantes.
5. La permutación de dos vectores.

§ 16. Base

Sea dado un espacio lineal arbitrario que se compone *no sólo de un vector nulo*. En tal espacio se tiene a ciencia cierta aunque un vector no nulo y, por lo tanto, existe un sistema linealmente independiente, por lo menos, de un vector. Por consiguiente, son posibles dos casos: o bien existe un sistema linealmente independiente que contiene un número de vectores tan grande como se quiera o bien existe un sistema linealmente independiente que contiene el número máximo de vectores. En el primer caso el espacio lineal se llama *de dimensión infinita* y en el segundo caso, *de dimensión finita*.

A excepción de unos ejemplos episódicos, *nuestra atención será dirigida a lo largo de todo el curso exclusivamente a los espacios de dimensión finita*. En particular, un espacio lineal de dimensión finita lo constituirá cualquier cápsula lineal construida en el número finito de vectores de un espacio arbitrario (no necesariamente de dimensión finita).

Así pues, supongamos que en el espacio lineal de dimensión finita K los vectores $e_1, e_2, \ldots, e_n$ constituyen un sistema linealmente independiente con el número máximo de vectores. Esto significa que para cualquier vector x de K el sistema $e_1, e_2, \ldots, e_n, x$ será linealmente dependiente. De conformidad con el teorema 14.2, el vector x se expresa linealmente en términos de $e_1, e_2, \ldots, e_n$. Como el vector x es arbitrario, mientras que los vectores $e_1, e_2, \ldots, e_n$ son fijados, podemos decir que:

Cualquier espacio lineal de dimensión finita es una cápsula lineal del número finito de sus vectores.

Ahora al investigar espacios lineales de dimensión finita, podemos hacer uso de toda clase de información referente a las cápsulas lineales y los sistemas equivalentes de vectores. Introduzcamos la siguiente definición.

Un sistema linealmente independiente de los vectores en términos de los cuales se expresa linealmente todo vector del espacio, se llama *base del espacio*.

La noción de base está ligada aquí con un sistema linealmente independiente que contiene el número máximo de vectores. No obstante, es evidente que todas las bases de un mismo espacio lineal de dimensión finita representan sistemas equivalentes linealmente independientes. Como sabemos, tales sistemas contienen un número igual de vectores. Por consiguiente, el número de los vectores de una base es la característica del espacio lineal de dimensión finita. Este número recibe el nombre de *dimensión* del espacio lineal K y se designa $\dim K$. Si $\dim K = n$, el propio espacio K se llama n-dimensional. Está claro que:

En un espacio lineal n-dimensional todo sistema linealmente inde-

pendiente de n vectores forma una base, mientras que todo sistema de $n+1$ vectores es linealmente dependiente.

Observemos que en todos los razonamientos anteriores suponíamos que el espacio lineal consiste no sólo en un vector nulo. Un espacio, que contiene sólo un vector nulo, no posee base en nuestro sentido y, rigiéndonos por la definición, consideraremos que su dimensión es igual a cero.

La base es de importancia enorme en el estudio de los espacios lineales de dimensión finita y en nuestras investigaciones siempre vamos a utilizarla. La base permite describir con facilidad la estructura de cualquier espacio lineal definido sobre un campo arbitrario P. Además, con ayuda de la base se puede construir *un aparato eficaz que reduce la realización de las operaciones sobre los elementos del espacio a las operaciones correspondientes sobre los números del campo P.*

Como ya hemos mostrado más arriba, todo vector x del espacio lineal K puede ser representado en forma de una combinación lineal

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n, \qquad (16.1)$$

donde, $\alpha_1, \alpha_2, \ldots, \alpha_n$ son ciertos números de P y $e_1, e_2, \ldots, e_n$ la base de K. La combinación lineal (16.1) se denomina *descomposición del vector x según la base* y los propios números $\alpha_1, \alpha_2, \ldots, \alpha_n$ se llaman *coordenadas del vector x respecto de dicha base.* El hecho de que el vector x se ha dado mediante sus coordenadas $\alpha_1, \alpha_2, \ldots, \alpha_n$ se notará del modo siguiente:

$$x = (\alpha_1, \alpha_2, \ldots, \alpha_n).$$

En lo que sigue no se indica, por regla general, a qué base se refieren las coordenadas dadas, siempre que esto no lleve a la ambigüedad.

Es fácil mostrar que para todo vector x de K su *descomposición según la base es única.* Esto se demuestra por un procedimiento empleado muy a menudo en la resolución de los problemas referentes a la dependencia lineal. Supongamos que existe otra descomposición:

$$x = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n. \qquad (16.2)$$

Al restar término a término (16.2) de (16.1), obtenemos

$$(\alpha_1 - \beta_1) e_1 + (\alpha_2 - \beta_2) e_2 + \ldots + (\alpha_n - \beta_n) e_n = 0.$$

En virtud de que los vectores $e_1, e_2, \ldots, e_n$ son linealmente independientes, de aquí se deduce que todos los coeficientes de la combinación lineal son nulos y que, por consiguiente, las descomposiciones (16.1) y (16.2) coinciden.

De este modo, con la base fijada del espacio lineal K todo vector de K se define unívocamente por la totalidad de sus coordenadas respecto de esta base.

Supongamos ahora que cualesquiera dos vectores x e y de K vienen dados por sus coordenadas respecto de una misma base e_1,

$e_2, \ldots, e_n$, es decir,

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n,$$
$$y = \gamma_1 e_1 + \gamma_2 e_2 + \ldots + \gamma_n e_n,$$

entonces

$$x + y = (\alpha_1 + \gamma_1) e_1 + (\alpha_2 + \gamma_2) e_2 + \ldots + (\alpha_n + \gamma_n) e_n.$$

Luego, para cualquier número λ del campo P tenemos

$$\lambda x = (\lambda\alpha_1) e_1 + (\lambda\alpha_2) e_2 + \ldots + (\lambda\alpha_n) e_n.$$

De aquí proviene que *al adicionar dos vectores de un espacio lineal, sus coordenadas respecto de cualquier base se suman y al multiplicar un vector por un número, todas sus coordenadas se multiplican por dicho número.*

Ejercicios.

1. Demuéstrese que el rango de un sistema de vectores coincide con la dimensión de su cápsula lineal.
2. Demuéstrese que los sistemas equivalentes de vectores tienen un mismo rango.
3. Demuéstrese que si la cápsula lineal L_1 está construida en los vectores de la cápsula lineal L_2, entonces $\dim L_1 \leqslant \dim L_2$.
4. Demuéstrese que si la cápsula lineal L_1 está construida en los vectores de la cápsula lineal L_2 y si $\dim L_1 = \dim L_2$, entonces las propias cápsulas lineales coinciden.
5. Demuéstrese que un espacio lineal de polinomios con coeficientes reales dado sobre un campo de números reales es de dimensión finita.

§ 17. Ejemplos sencillos de los espacios lineales

Los conceptos fundamentales de dependencia lineal y de base se pueden ilustrar con unos ejemplos muy sencillos pero aleccionadores, si tomamos a título de espacios lineales los conjuntos de números con operaciones ordinarias de adición y multiplicación. La validez de los axiomas del espacio lineal para los conjuntos de este tipo es completamente obvia, por lo cual no nos detendremos en su comprobación. Los elementos del espacio se llamarán, como antes, *vectores*.

Consideraremos un espacio lineal *complejo* que representa en sí un grupo de adición de todos los números complejos con multiplicación sobre un campo de números complejos. Está claro que cualquier número z_1, distinto de cero, representa en sí un vector linealmente independiente. Sin embargo, cualesquiera dos vectores no nulos z_1 y z_2 ya son siempre linealmente dependientes. Para demostrarlo basta hallar tales dos números complejos α_1 y α_2, distintos de cero simultáneamente, que sea $\alpha_1 z_1 + \alpha_2 z_2 = 0$. Mas, esta igualdad puede realizarse, evidentemente, cuando $\alpha_1 = -z_2$, $\alpha_2 = z_1$. Por consiguiente, el espacio lineal considerado es unidimensional.

Resulta ser algo diferente un espacio lineal *real* que representa un grupo de adición de todos los números complejos con multiplicación sobre un campo de números reales. A título de coeficientes de las combinaciones lineales pueden emplearse, en este caso, solamente los números reales, razón por la cual este espacio lineal no puede ser unidimensional. En efecto, no existen números reales α_1 y α_2, distintos de cero simultáneamente, para los cuales la combinación lineal $\alpha_1 z_1 + \alpha_2 z_2$ se reduciría a cero, por ejemplo, para $z_1 = 1$, $z_2 = i$. Sea al cargo del lector demostrar, a título de ejercicio, que el espacio lineal dado es bidimensional.

Es importante recalcar que aunque ambos espacios examinados se componen de *los mismos* elementos, *se diferencian* uno del otro en principio. Ahora ya está claro que un espacio lineal real, que representa un grupo de adición de todos los números reales con multiplicación sobre un campo de números reales, es unidimensional. Consideraremos luego un espacio lineal racional que representa en sí un grupo de adición de todos los números reales con multiplicación sobre un campo de números racionales.

Trataremos de construir, como antes, un sistema que contenga el número máximo de los vectores linealmente independientes r_1, r_2, r_3, ... Está claro que se puede tomar, por ejemplo, $r_1 = 1$. Puesto que a título de coeficientes de las combinaciones lineales α_1, α_2, α_3, ... se admiten sólo los números racionales, es natural que mediante un número del tipo $\alpha_1 \cdot 1$ no es posible representar, por ejemplo, $\sqrt{2}$. Por lo tanto, el espacio no puede ser unidimensional. Por eso, en calidad del segundo vector, linealmente independiente con la unidad, podemos tomar precisamente $\sqrt{2}$. Sin embargo, mediante un número del tipo $\alpha_1 \cdot 1 + \alpha_2 \sqrt{2}$ no se puede representar, por ejemplo, $\sqrt[4]{2}$.

Efectivamente, supongamos que para ciertos α_1 y α_2 racionales se verifica la igualdad $\sqrt[4]{2} = \alpha_1 + \alpha_2 \sqrt{2}$. Elevando al cuadrado ambos miembros, obtendremos

$$\sqrt{2} = (\alpha_1^2 + 2\alpha_2^2) + 2\alpha_1\alpha_2 \sqrt{2}$$

o bien

$$\frac{2(1-2\alpha_1\alpha_2)}{\alpha_1^2+2\alpha_2^2} = \sqrt{2}.$$

Esto es imposible, puesto que en el primer miembro figura un número racional y en el segundo, irracional.

For lo tanto, el espacio en consideración tampoco puede ser bidimensional. Entonces, ¿cuál será? Por sorprendente que sea, es *de dimensión infinita*. Sin embargo, la demostración de esta afirmación sale de los márgenes de nuestro curso.

La atención especial que se ha prestado a los ejemplos de espacios lineales de dimensión pequeña se explica por el hecho de que con ayuda de tales espacios podemos construir los espacios lineales de cualquier dimensión. Mas, hablaremos de esto más tarde.

Ejercicios.

1. ¿Qué dimensión tiene un espacio lineal de números racionales dado sobre un campo de números racionales?
2. Constrúyanse los sistemas linealmente independientes de vectores en un espacio de los números complejos dado sobre un campo de números racionales.
3. ¿Será espacio lineal un grupo de adición de los números racionales dado sobre un campo de números reales? Si no ¿por qué?

§ 18. Espacios lineales de segmentos dirigidos

Ya se ha señalado anteriormente que los conjuntos de segmentos dirigidos colineales, de segmentos dirigidos coplanares y de segmentos dirigidos en todo el espacio forman dos espacios lineales sobre un campo de números reales. Nuestra tarea inmediata consiste en revelar su dimensión y construir la base.

LEMA 18.1. *La condición necesaria y suficiente de la dependencia lineal de dos vectores es su carácter colineal.*

DEMOSTRACIÓN. Hemos de notar que la afirmación del lema es evidente, si entre dos vectores se tiene aunque uno nulo. Por eso supondremos que ambos vectores no son nulos.

Sean a y b dos vectores linealmente dependientes. En este caso existen los números α y β tales que

$$\alpha a+\beta b=0.$$

Puesto que, de acuerdo con la suposición, $a\neq 0$, $b\neq 0$, entonces $\alpha\neq 0$, $\beta\neq 0$, y por eso

$$b=\left(-\frac{\alpha}{\beta}\right)a.$$

Por consiguiente, según la definición de operación de la multiplicación de un segmento dirigido por un número, los vectores a y b son colineales.

Supongamos ahora que los vectores a y b son colineales. Apliquémoslos a un punto común O. Estos vectores se dispodrán en cierta recta la cual transformaremos en un eje con una dirección determinada. Los vectores a y b son no nulos, razón por la cual existe un número real λ tal que la magnitud del segmento dirigido a es igual al producto de la magnitud del segmento dirigido b por el número λ, es decir, $\{a\}=\lambda\{b\}$. Mas, de acuerdo con la definición de operación de multiplicación de un segmento dirigido por un número, esto quiere decir que $a=\lambda b$. Así pues, los vectores a y b son linealmente dependientes.

Del lema demostrado se desprende que el espacio lineal de segmentos dirigidos colineales es *unidimensional* y de su base puede servir *cualquier vector no nulo.*

El lema 18.1. permite deducir un corolario muy importante. A saber, si los vectores a, b son colineales y $a \neq 0$, entonces existe tal número λ que $b = \lambda a$. En efecto, estos vectores son linealmente dependientes, es decir, para ciertos números α, β que no son nulos simultáneamente, se tiene $\alpha a + \beta b = 0$. Si suponemos que $\beta = 0$, de aquí se deduce que $\alpha = 0$. Por consiguiente, $\beta \neq 0$ y a título de número λ se puede tomar $\lambda = (-\alpha)/\beta$.

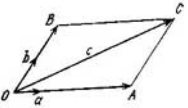


Fig. 18.1.

LEMA 18.2. *La condición necesaria y suficiente de la dependencia lineal de tres vectores es su carácter coplanar.*

DEMOSTRACIÓN. Sin restringir la generalidad supondremos que ningún par de los tres vectores indicados es colineal, puesto que en el caso contrario la afirmación del lema se deduce inmediatamente del lema 18.1.

Así pues, sean a, b, c tres vectores linealmente dependientes. Por consiguiente, existen tales números reales α, β, γ, entre los cuales no todos son nulos, que

$$\alpha a + \beta b + \gamma c = 0.$$

Si, por ejemplo, $\gamma \neq 0$, entonces de esta igualdad hallamos

$$c = \left(-\frac{\alpha}{\gamma}\right) a + \left(-\frac{\beta}{\gamma}\right) b.$$

Apliquemos los vectores a, b, c a un punto común O. En este caso de la última igualdad se desprende que el vector c es igual a la diagonal del paralelogramo construido en los vectores $(-\alpha/\gamma)\, a$ y $(-\beta/\gamma)\, b$. Esto significa que siendo trasladados paralelamente en el punto común, los vectores a, b, c resultan ser dispuestos en un mismo plano y, por consiguiente, son coplanares.

Supongamos ahora que los vectores a, b, c son coplanares. Traslademoslos en un plano y apliquemos al punto común O (fig. 18.1). Por los extremos del vector c tracemos las rectas paralelas a los vectores a y b, y consideraremos el paralelogramo $OACB$. Los vectores a, $\overrightarrow{OA}$ y b, $\overrightarrow{OB}$ son colineales por construcción y no nulos, por lo cual existen tales números λ, μ que

$$\overrightarrow{OA} = \lambda a, \quad \overrightarrow{OB} = \mu b.$$

Pero, $\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB}$, por consiguiente $c = \lambda a + \mu b$, o bien

$$\lambda a + \mu b + (-1)\, c = 0.$$

Puesto que los números λ, μ, -1 son a ciencia cierta diferentes de cero, la última igualdad significa dependencia lineal de los vectores a, b, c.

Ahora podemos resolver la cuestión sobre la dimensión de un espacio lineal de segmentos dirigidos coplanares. De acuerdo con el lema que acabamos de demostrar, la dimensión de este espacio debe ser menos de tres. Pero cualesquiera dos segmentos dirigidos no colineales de dicho espacio son linealmente independientes. Por eso, el espacio lineal de segmentos dirigidos coplanares es un espacio *bidimensional* y su base pueden constituir *cualesquiera dos vectores no colineales.*

LEMA 18.3. *Cualesquiera cuatro vectores son linealmente dependientes.*

DEMOSTRACION. Sin restringir la generalidad supondremos que ninguna terna de cuatro vectores es coplanar, puesto que en el caso contrario la afirmación del lema fluye inmediatamente del lema 18.2. Apliquemos los vectores a, b, c, d al òrigen común O y por el extremo D del vector d tracemos los planos paralelos a los que se definen por los pares de vectores b, c; a, c y a, b, respectivamente (fig. 18.2). De la regla del paralelogramo para la adición de vectores proviene que

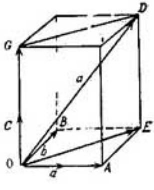


Fig. 18.2.

$$\overrightarrow{OD}=\overrightarrow{OC}+\overrightarrow{OE},\quad \overrightarrow{OE}=\overrightarrow{OA}+\overrightarrow{OB},$$

por ello

$$\overrightarrow{OD}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}. \tag{18.1}$$

Los vectores a, $\overrightarrow{OA}$ y también b, $\overrightarrow{OB}$ y c, $\overrightarrow{OC}$ son colineales según su construcción, con la particularidad de que a, b, c son no nulos. Por consiguiente, existen tales números λ, μ, ν que

$$\overrightarrow{OA}=\lambda a,\quad \overrightarrow{OB}=\mu b,\quad \overrightarrow{OC}=\nu c.$$

Teniendo presente (18.1), esto nos da la correlación

$$d=\lambda a+\mu b+\nu c,$$

de donde se desprende la dependencia lineal de los vectores a, b, c, d.

Del lema demostrado concluimos que la dimensión de un espacio lineal de todos los segmentos dirigidos debe ser inferior a cuatro. Pero no puede ser menos que tres, puesto que, de conformidad con el lema 18.2, cualesquiera tres segmentos dirigidos no coplanares son linealmente independientes. Por eso el espacio lineal de todos los segmentos dirigidos es *tridimensional* y de su base pueden servir *cualesquiera tres vectores no coplanares.*

Los espacios lineales considerados no son muy ilustrativos desde el punto de vista geométrico, puesto que se admite la existencia en dichos espacios de un número infinito de los vectores iguales entre sí. Se harán mucho más ilustrativos, si en toda clase de los vectores iguales fijamos un representante y por palabra "vector" entendemos siempre un segmento dirigido elegido sólo de la totalidad de estos representantes.

Uno de los métodos más cómodos para la fijación consiste en el examen de los conjuntos de vectores dirigidos sujetados en cierto punto O. En tal caso en lugar de un espacio lineal de segmentos dirigidos colineales obtenemos un espacio de segmentos dirigidos que están sujetados en el punto O y se ubican en la recta que pasa por este punto; en lugar de un espacio lineal de segmentos dirigidos coplanares obtenemos un espacio de segmentos dirigidos sujetados al punto O y ubicados en el plano que pasa por este punto; y, por fin, en lugar de un espacio lineal de todos los segmentos dirigidos, un espacio de segmentos dirigidos sujetados en el punto O.

En adelante trataremos, principalmente, solamente los vectores sujetados. Los espacios lineales correspondientes se designarán mediante V_1, V_2 y V_3, donde el índice abajo significa la dimensión. Un espacio lineal compuesto de un solo segmento dirigido nulo, lo designaremos mediante V_0.

La introducción de estos espacios permite establecer una correspondencia biunívoca entre los puntos y los segmentos dirigidos. Para ello es suficiente a todo vector ponerle en correspondencia su punto final. Teniendo presente tal *interpretación geométrica*, los elementos de un espacio lineal abstracto también se llamarán, a veces, *puntos* en lugar de denominarlos vectores.

Ejercicios.

Indíquese en los espacios V_1, V_2, V_3 el sentido geométrico de tales conceptos como:

1. La cápsula lineal.
2. La dependencia e independencia lineales.
3. Los sistemas equivalentes de vectores.
4. Las transformaciones equivalentes elementales del sistema de vectores.
5. El rango del sistema de vectores.

§ 19. Suma e intersección de los subespacios

La introducción de las cápsulas lineales ha mostrado que todo espacio lineal contiene un conjunto infinito de otros espacios lineales. El significado de estos espacios no se limita sólo a las cuestiones examinadas anteriormente.

Las cápsulas lineales se han definido con ayuda de la indicación inmediata de su *estructura*. Podríamos emplear también otro acceso, al definir los espacios "menores" a través de *las propiedades* de los vectores. Supongamos que en el espacio lineal K está definido un conjunto de vectores L. Si, al realizar las mismas operaciones que en el espacio K, el conjunto L es por sí mismo un espacio lineal, llamaremos a L *subespacio* lineal, subrayando en la denominación el hecho de que el subespacio consta de los vectores de cierto espacio. Por lo visto, el subespacio mínimo es aquel que se compone sólo de un vector nulo. Tal subespacio se denominará *nulo* y se designará mediante el símbolo 0. El subespacio máximo es el espacio K. Estos dos subespacios se llaman *triviales* y los demás, *no triviales*. Es evidente que todo subespacio junto con cada par de sus elementos x, y contiene también todas sus combinaciones lineales $\alpha x + \beta y$. Lo recíproco es también cierto. A saber:

Si el conjunto de vectores L de un espacio lineal K, junto con cada par de sus elementos x, y contiene también *todas* sus combinaciones lineales $\alpha x + \beta y$, será un subespacio.

Efectivamente, de todos los axiomas para el espacio lineal es preciso comprobar sólo aquellos que se refieren a los vectores nulo y opuesto. La validez de los axiomas restantes es obvia. Tomemos $\alpha = 0$, $\beta = 0$. De conformidad con los corolarios que provienen de las propiedades de las operaciones para los vectores del espacio K, llegamos a que $0 \cdot x + 0 \cdot y = 0$, es decir, el vector nulo pertenece al conjunto L. Tomemos ahora $\alpha = -1$, $\beta = 0$. Se tiene $(-1)\, x + + 0 \cdot y = (-1)\, x$, por lo cual en L junto con todo vector x figura también un vector opuesto a x. Así pues, el conjunto L es un subespacio.

La presencia de la base permite enunciar la afirmación de que en todo espacio de dimensión finita cualquier subespacio será una cápsula lineal. Por esto, en los espacios lineales de dimensión finita la cápsula lineal constituye un método más general para definir los subespacios lineales. En el espacio de dimensión infinita no es así. No obstante, no locabe olvidar que existe muchísimo en común entre los conceptos y los hechos en los espacios de dimensión finita y los análogos correspondientes en los espacios de dimensión infinita. En nuestro deseo de subrayar esta circunstancia, aun en los espacios de dimensión finita utilizaremos con mayor frecuencia el término *subespacio lineal* en vez de *cápsula lineal*.

Sea K un espacio n-dimensional. Al igual que en el mismo espacio K, en cualquier subespacio suyo L puede construirse la base. Si en el espacio K la base elegida es $e_1, e_2, \ldots, e_n$, en el caso general los vectores básicos del subespacio L no pueden elegirse directamente del número de los vectores $e_1, e_2, \ldots, e_n$ aunque sea por aquella razón que ni uno de ellos puede figurar en dicho subespacio. Sin embargo, resulta válido en cierto sentido el siguiente lema inverso.

LEMA 19.1. *Si en un subespacio L de dimensión s se ha elegido una base arbitraria $t_1, \ldots, t_s$, entonces en el espacio K de n-ésima* *dimensión se pueden elegir los vectores $t_{s+1}, \ldots, t_n$ de una manera tal que el sistema de los vectores $t_1, \ldots, t_s, t_{s+1}, \ldots, t_n$ será la base en todo K.*

DEMOSTRACIÓN. Examinemos sólo aquellos sistemas linealmente independientes de vectores en K que contienen los vectores $t_1, \ldots, t_s$. Está claro que entre estos sistemas hay un sistema $t_1, \ldots, t_s, t_{s+1}, \ldots, t_p$ que tiene el número máximo de vectores. Mas, en este caso cualquiera que sea el vector x de K, el sistema $t_1, \ldots, t_p, x$ debe ser linealmente dependiente. Por consiguiente, el vector x debe expresarse linealmente en términos de los vectores $t_1, \ldots, t_p$. Esto significa que los vectores $t_1, \ldots, t_s, t_{s+1}, \ldots \ldots, t_p$ forman la base en K y $p = n$.

Consideraremos otra vez el espacio lineal arbitrario K. Este espacio engendra el conjunto de todos los subespacios suyos, el cual se designará mediante U. En el conjunto U se pueden definir dos operaciones algebraicas que, a base de unos subespacios, permiten construir otros.

Se denomina *suma* $L_1 + L_2$ de los subespacios lineales L_1, L_2 un conjunto de todos los vectores del tipo $z = x + y$, donde $x \in L_1$, $y \in L_2$.

Se denomina *intersección* $L_1 \cap L_2$ de los subespacios lineales L_1, L_2 un conjunto de todos los vectores pertenecientes simultáneamente tanto a L_1 como a L_2.

Observemos que la suma de los subespacios, como también la intersección de ellos, siempre son conjuntos no vacíos, ya que les pertenece a ciencia cierta el vector nulo del espacio K. Demostremos que estos conjuntos son subespacios.

En efecto, tomemos dos vectores arbitrarios z_1, z_2 de la suma $L_1 + L_2$. Esto significa que $z_1 = x_1 + y_1$, $z_2 = x_2 + y_2$, donde $x_1, x_2 \in L_1$ e $y_1, y_2 \in L_2$. Examinaremos ahora una combinación lineal arbitraria $\alpha z_1 + \beta z_2$. Tenemos $\alpha z_1 + \beta z_2 = (\alpha x_1 + \beta x_2) = (\alpha y_1 + \beta y_2)$. Puesto que $\alpha x_1 + \beta x_2 \in L_1$ y $\alpha y_1 + \beta y_2 \in L_2$, entonces $\alpha z_1 + \beta z_2 \in L_1 + L_2$. Por consiguiente, $L_1 + L_2$ es un subespacio. Sea, ahora, $z_1, z_2 \in L_1 \cap L_2$, es decir, $z_1, z_2 \in L_1$ y $z_1, z_2 \in L_2$. Está claro que $\alpha z_1 + \beta z_2 \in L_1$ y $\alpha z_1 + \beta z_2 \in L_2$, es decir, $\alpha z_1 + \beta z_2 \in L_1 \cap L_2$. Por consiguiente, $L_1 \cap L_2$ es también un subespacio.

De este modo, las operaciones de adición de los subespacios y de su intersección son algebraicas. Estas operaciones son, evidentemente, conmutativas y asociativas. Además, para todo subespacio L de K se verifica

$$L + 0 = L, \quad L \cap K = L.$$

Las leyes distributivas, que ligan ambas operaciones, están ausentes.

Como se nota con facilidad ya en los ejemplos más sencillos, la dimensión de la suma de dos subespacios arbitrarios depende no sólo de la dimensión de los propios subespacios, sino también de cuán grande es su parte común. Resulta válido el

TEOREMA 19.1. *Para cualesquiera dos subespacios L_1, L_2 de dimensión finita tiene lugar la igualdad*

$$\dim (L_1 \cap L_2) + \dim (L_1 + L_2) = \dim L_1 + \dim L_2. \quad (19.1)$$

DEMOSTRACIÓN. Designaremos las dimensiones de los subespacios L_1, L_2, $L_1 \cap L_2$ con r_1, r_2, m, respectivamente. Elijamos en la intersección $L_1 \cap L_2$ una base cualquiera $c_1, \ldots, c_m$. Los vectores son linealmente independientes y se disponen en L_1. Conforme al lema 19.1, en L_1 existen tales vectores $a_1, \ldots, a_k$ que el sistema $a_1, \ldots, a_k, c_1, \ldots, c_m$ será la base en L_1. Por analogía, en el subespacio L_2 existen tales vectores $b_1, b_2, \ldots, b_p$, que el sistema $b_1, \ldots, b_p, c_1, \ldots, c_m$ será la base en L_2. En esto caso

$$r_1 = k + m, \qquad r_2 = p + m.$$

Si demostramos que el sistema de vectores

$$a_1, \ldots, a_k, \quad c_1, \ldots, c_m, \quad b_1, \ldots, b_p \quad (19.2)$$

es la base del subespacio $L_1 + L_2$, entonces la afirmación del teorema tiene lugar, puesto que

$$m + (k + m + p) = (k + m) + (p + m).$$

Todo vector de los subespacios L_1, L_2 se expresa linealmente en términos de los vectores de su base y con mayor razón, linealmente en términos de los vectores (19.2). Por eso en términos de estos vectores también se expresará linealmente cualquier vector de la suma $L_1 + L_2$. Resta mostrar que el sistema (19.2) es linealmente independiente. Sea

$$\alpha_1 a_1 + \ldots + \alpha_k a_k + \gamma_1 c_1 + \ldots + \gamma_m c_m + \beta_1 b_1 + \ldots$$
$$\ldots + \beta_p b_p = 0. \quad (19.3)$$

Denotemos

$$b = \beta_1 b_1 + \ldots + \beta_p b_p. \quad (19.4)$$

Está claro que $b \in L_2$. Pero de (19.3) se deduce que $b \in L_1$. Por consiguiente, $b \in L_1 \cap L_2$, es decir,

$$b = \nu_1 c_1 + \ldots + \nu_m c_m \quad (19.5)$$

para ciertos números $\nu_1, \ldots, \nu_m$. Al comparar (19.4), (19.5) obtenemos

$$\beta_1 b_1 + \ldots + \beta_p b_p + (-\nu_1) c_1 + \ldots + (-\nu_m) c_m = 0.$$

El sistema de vectores $b_1, \ldots, b_p, c_1, \ldots, c_m$ es linealmente independiente por construcción y por ello

$$\beta_1 = \ldots = \beta_p = \nu_1 = \ldots = \nu_m = 0.$$

En virtud de la independencia lineal del sistema de vectores $a_1, \ldots, a_k, c_1, \ldots, c_m$, de (19.3) se desprende ahora que

$$\alpha_1 = \ldots = \alpha_k = \gamma_1 = \ldots = \gamma_m = 0.$$

El teorema queda demostrado.

Ejercicios.

1. Establézcase, en el ejemplo del espacio lineal V_3, el significado geométrico de las operaciones de la suma e intersección de subespacios.
2. ¿Qué es la suma de los subespacios V_1 y V_2?
3. ¿Qué es la intersección de los subespacios V_1 y V_2?
4. Demuéstrese que la dimensión de una intersección de cualquier número de subespacios no es superior a la mínima de las dimensiones de estos subespacios.
5. Demuéstrese que la dimensión de una suma de cualquier número de subespacios no es inferior a la máxima de las dimensiones de estos subespacios.

§ 20. Suma directa de los subespacios

Sean dados los subespacios $L_1, L_2, \ldots, L_m$ de cierto espacio lineal. Por definición de la operación de adición, todo vector x, perteneciente a la suma

$$K = L_1 + L_2 + \ldots + L_m, \tag{20.1}$$

puede ser representado en la forma

$$x = x_1 + x_2 + \ldots + x_m, \tag{20.2}$$

donde $x_i \in L_i$ para i cualquiera. En el caso general esta representación no será única. En cambio, si todo vector de K admite una única representación (20.2), entonces la suma (20.1) se llama *suma directa* y se designa del modo siguiente:

$$K = L_1 \dotplus L_2 \dotplus \ldots \dotplus L_m. \tag{20.3}$$

Las sumas directas poseen varias propiedades especiales. No obstante, para nosotros serán de interés no tanto dichas propiedades como *los rasgos comunes en la descomposición* (20.2) *y la descomposición según la base*. Supongamos que cierto espacio K puede descomponerse en suma directa (20.3) de sus subespacios $L_1, L_2, \ldots, L_m$. Entonces, debido a la unicidad de la descomposición (20.2), el sistema de los subespacios $L_1, L_2, \ldots, L_m$ puede considerarse como cierta "base generalizada" del espacio K y la descomposición (20.2), como descomposición según la "base generalizada". Tal interpretación de la suma directa es especialmente útil al estudiar los espacios lineales de gran dimensión, puesto que en estos espacios hemos de estudiar, como regla, no todos los componentes en la descomposición según la base, sino sólo una pequeña parte de ellos. El uso de la suma directa hace posible evitar tanto las descomposiciones voluminosas como las investigaciones de los pormenores superfluos.

Sea K un espacio lineal n-dimensional. Tomemos su base arbitraria $e_1, e_2, \ldots, e_n$ y construyamos una totalidad de cápsulas lineales $L_1 = L_1(e_1)$, $L_2 = L_2(e_2)$, $\ldots$, $L_n = L_n(e_n)$. Es evidente en este caso que K será la suma directa de estos n subespacios unidimensionales. Pero el espacio K puede ser representado por diferentes medios en forma de una suma directa de los subespacios que tengan otra dimensión. El fundamento para tal representación lo constituye el

TEOREMA 20.1. *Para que el espacio K sea una suma directa de sus subespacios $L_1, \ldots, L_m$, es necesario y suficiente que la reunión de las bases de estos subespacios constituya la base de todo el espacio.*

DEMOSTRACIÓN. Supongamos que K es una suma directa de los subespacios $L_1, \ldots, L_m$ y los vectores $e_1, \ldots, e_{s_1}; \ldots$ $\ldots; e_{s_{m-1}+1}, \ldots, e_{s_m}$ constituyen las bases de dichos subespacios. Entonces para cada vector x de K tiene lugar la descomposición (20.2). Al representar cada uno de los vectores x_i en forma de una descomposición según la base del subespacio correspondiente L_i, obtenemos que

$$x = \alpha_1 e_1 + \ldots + \alpha_{s_1} e_{s_1} + \ldots + \alpha_{s_{m-1}+1} e_{s_{m-1}+1} + \ldots + \alpha_{s_m} e_{s_m} \qquad (20.4)$$

para ciertos números $\alpha_1, \ldots, \alpha_{s_m}$.

De este modo, todo vector de K se puede representar en forma de una combinación lineal de los vectores $e_1, \ldots, e_{s_m}$. Para que se pueda afirmar que estos vectores constituyen la base del espacio K, nos resta demostrar su independencia lineal. Examinaremos la igualdad

$$\beta_1 e_1 + \ldots + \beta_{s_1} e_{s_1} + \ldots + \beta_{s_{m-1}+1} e_{s_{m-1}+1} + \ldots + \beta_{s_m} e_{s_m} = 0 \qquad (20.5)$$

con los coeficientes numéricos $\beta_1, \ldots, \beta_{s_m}$ y designaremos

$$\begin{gathered} \beta_1 e_1 + \ldots + \beta_{s_1} e_{s_1} = y_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \beta_{s_{m-1}+1} e_{s_{m-1}+1} + \ldots + \beta_{s_m} e_{s_m} = y_m. \end{gathered} \qquad (20.6)$$

Es obvio que $y_i \in L_i$ y de (20.5) proviene la siguiente igualdad:

$$0 = y_1 + \ldots + y_m.$$

Todos los subespacios contienen el vector nulo, por lo cual se verifica a ciencia cierta la correlación

$$0 = 0 + \ldots + 0.$$

Debido a la unicidad de la descomposición del vector nulo de K según los subespacios $L_1, \ldots, L_m$ concluimos que

$$y_1 = \ldots = y_m = 0.$$

De aquí se desprende que todos los coeficientes de las combinaciones lineales (20.6) son iguales a cero, es decir, los vectores $e_1, \ldots, e_{s_m}$ son linealmente independientes.

Supongamos ahora que los vectores $e_1, \ldots, e_{s_1}; \ldots; e_{s_{m-1}+1}, \ldots, e_{s_m}$, que constituyen las bases de los subespacios $L_1, \ldots, L_m$, forman la base de K. En este caso para todo vector x de K tiene lugar una única descomposición (20.4). Al denotar

$$\begin{aligned} \alpha_1 e_1 + \ldots + \alpha_{s_1} e_{s_1} &= x_1, \\ \ldots\ldots\ldots\ldots &\ldots\ldots \\ \alpha_{s_{m-1}+1} e_{s_{m-1}+1} + \ldots + \alpha_{s_m} e_{s_m} &= x_m, \end{aligned} \qquad (20.7)$$

resulta que para x existe al menos una descomposición (20.2). Todo vector x_i de (20.7) es una combinación lineal de los vectores básicos L_i. En virtud de la unicidad de la descomposición (20.4) para el vector x, llegamos a la conclusión de que la descomposición (20.2) para él es también única. El teorema queda demostrado.

Ejercicios.

1. ¿En qué condiciones el espacio V_3 será la suma directa de sus subespacios V_1 y V_2?
2. ¿En qué condiciones el espacio V_3 será la suma directa de sus dos subespacios del tipo V_1?
3. ¿Podrá ser V_3 una suma directa de sus dos subespacios del tipo V_2? Si no ¿por qué?
4. Demuéstrese que para que la suma (20.1) sea directa, es necesario y suficiente que la descomposición (20.2) sea única para el vector nulo.
5. Demuéstrese que para que la suma (20.1) sea directa, es necesario y suficiente que la intersección de cada uno de los subespacios L_i, $i = 1, \ldots, m$, con la suma de los demás contenga sólo el vector nulo.

§ 21. Isomorfismo de los espacios lineales

Consideraremos el conjunto de todos los espacios lineales definidos sobre un mismo campo P. Será natural preguntar en qué consiste la semejanza y la diferencia entre todos estos espacios.

Todo espacio lineal contiene en su descripción dos facetas esencialmente diferentes. En primer lugar, el espacio lineal es una totalidad de objetos concretos que se denominan vectores. En segundo lugar, sobre estos objetos concretos están definidas las operaciones de adición y multiplicación por un número que poseen ciertas propiedades. Por esto, el interés puede dirigirse o bien hacia la naturaleza de los vectores y las propiedades de éstos o bien hacia las propiedades de las operaciones indicadas independientemente de la naturaleza de los elementos.

La naturaleza de los vectores nos interesaba sólo en el estudio de los segmentos dirigidos y únicamente en la medida que fue necesaria para introducir operaciones y establecer las propiedades de las mismas. A continuación, la investigación posterior de los segmentos dirigidos se apoyaba exclusivamente en las propiedades de las operaciones. Del modo análogo procederemos también en cada caso concreto. Por ello, se considerará que dos espacios de una misma estructura respecto a las operaciones de adición y multiplicación por un número poseen propiedades iguales o son *isomorfos*. Con más precisión:

Dos espacios lineales dados sobre un mismo campo se llaman isomorfos, si se puede establecer tal correspondencia biunívoca entre sus vectores que a la suma de cualesquiera dos vectores del primer espacio le corresponda la suma de los vectores correspondientes del segundo espacio y al producto de un número por un vector del primer espacio le corresponda el producto de este mismo número por el vector correspondiente del segundo espacio.

Sean K y K' los espacios isomorfos. El hecho de que a todo vector x de K se le ha puesto en correspondencia un vector determinado x' de K' puede entenderse como introducción de cierta "función"

$$x' = \omega(x), \tag{21.1}$$

cuyo "argumento" es el vector x del espacio K y el "valor" es el vector x' del espacio K'. Ambas propiedades de esta función se pueden escribir del modo siguiente. Para cualesquiera x, y de K y para todo número λ

$$\begin{gathered}\omega(x+y) = \omega(x) + \omega(y),\\ \omega(\lambda x) = \lambda\omega(x).\end{gathered} \tag{21.2}$$

El carácter biunívoco de la correspondencia entre K y K' significa que a cualesquiera argumentos diferentes de la función (21.1) les corresponden diferentes valores, es decir, si

$$x \neq y, \tag{21.3}$$

entonces

$$\omega(x) \neq \omega(y). \tag{21.4}$$

Por consiguiente, de la igualdad o desigualdad de los valores de la función proviene, correspondientemente, la igualdad o desigualdad de los argumentos.

Los espacios isomorfos tienen mucho en común. En particular, al vector nulo le corresponde el vector nulo, pues

$$\omega(0) = \omega(0\cdot x) = 0\cdot\omega(x) = 0\cdot x' = 0'.$$

Sin embargo, el corolario más importante consiste en que a un sistema linealmente independiente de vectores le corresponde de nuevo un sistema linealmente independiente.

Efectivamente, sean $x_1, x_2, \ldots, x_n$ unos vectores linealmente independientes. Consideraremos una combinación lineal $\alpha_1 \omega(x_1) + \alpha_2\omega(x_2) + \ldots + \alpha_n\omega(x_n)$ y harémosla igual a cero. Debido a las propiedades de la correspondencia isomorfa tenemos

$$0' = \alpha_1\omega(x_1) + \alpha_2\omega(x_2) + \ldots + \alpha_n\omega(x_n) = \\ = \omega(\alpha_1x_1 + \alpha_2x_2 + \ldots + \alpha_nx_n) = \omega(0),$$

de donde se desprende que

$$\alpha_1x_1 + \alpha_2x_2 + \ldots + \alpha_nx_n = 0.$$

Como los vectores $x_1, x_2, \ldots, x_n$ son linealmente independientes, todos los coeficientes deben ser nulos.

El corolario demostrado permite afirmar que si dos espacios lineales de dimensión finita son isomorfos, tienen la misma dimensión. La afirmación recíproca es también válida. A saber, tiene lugar el

TEOREMA 21.1 *Dos espacios lineales cualesquiera que tienen una misma dimensión finita y están dados sobre un mismo campo son isomorfos.*

DEMOSTRACIÓN Sean K y K' dos espacios lineales de dimensión n. Elijamos una base $e_1, e_2, \ldots, e_n$ en el espacio K y una base $e'_1, e'_2, \ldots, e'_n$ en el espacio K'. Construyamos un isomorfismo ω del modo siguiente, haciendo uso de los sistemas indicados de vectores. A todo vector

$$x = \alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ne_n$$

del espacio K le pondremos en correspondencia el vector

$$\omega(x) = \alpha_1e'_1 + \alpha_2e'_2 + \ldots + \alpha_ne'_n$$

del espacio K'. La correspondencia establecida será biunívoca, puesto que la descomposición según la base es única.

Elijamos ahora dos vectores arbitrarios x e y de K y un número cualquiera λ y supongamos que

$$x = \alpha_1e_1 + \alpha_2e_2 + \ldots + \alpha_ne_n,$$
$$y = \beta_1e_1 + \beta_2e_2 + \ldots + \beta_ne_n.$$

Se tiene

$$\omega(x+y) = \omega((\alpha_1+\beta_1)e_1 + (\alpha_2+\beta_2)e_2 + \ldots + (\alpha_n+\beta_n)e_n) = \\ = (\alpha_1+\beta_1)e'_1 + (\alpha_2+\beta_2)e'_2 + \ldots + (\alpha_n+\beta_n)e'_n = \\ = (\alpha_1e'_1 + \alpha_2e'_2 + + \alpha_ne'_n) + (\beta_1e'_1 + \beta_2e'_2 + \ldots + \beta_ne'_n) = \omega(x) + \omega(y),$$

$$\omega(\lambda x) = \omega((\lambda\alpha_1)e_1 + (\lambda\alpha_2)e_2 + \ldots + (\lambda\alpha_n)e_n) = \\ = (\lambda\alpha_1)e'_1 + (\lambda\alpha_2)e'_2 + \ldots + (\lambda\alpha_n)e'_n = \\ = \lambda(\alpha_1e'_1 + \alpha_2e'_2 + \ldots + \alpha_ne'_n) = \lambda\omega(x).$$

Las igualdades obtenidas demuestran la validez de la afirmación del teorema.

La importancia de este teorema es muy grande. Precisamente este teorema permite ahora hablar con toda la certeza de que desde el punto de vista de todos los corolarios procedentes de los axiomas cualesquiera dos espacios lineales que tienen una misma dimensión y están dados sobre un mismo campo son indistinguibles. Por consiguiente, podríamos construir un espacio lineal n-dimensional sobre el campo dado y, al investigar sólo dicho espacio, aclarar las leyes propias a todos los espacios de dimension finita.

Sea dado un campo P. Consideremos un conjunto cuyos elementos son toda clase de surtidos ordenados de n números $\alpha_1, \alpha_2, \ldots, \alpha_n$ del campo P. Si x es un elemento de este conjunto, escribiremos

$$x = (\alpha_1, \alpha_2, \ldots, \alpha_n). \tag{21.5}$$

Las operaciones de adición y multiplicación por el número λ del campo P determinemos del modo siguiente:

$$\begin{aligned}(\alpha_1, \alpha_2, \ldots, \alpha_n) + (\beta_1, \beta_2, \ldots, \beta_n) &= \\ &= (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n), \\ \lambda(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n).\end{aligned} \tag{21.6}$$

Es fácil comprobar que los axiomas del espacio lineal están cumplidos. En particular, el vector nulo se determina por el surtido de ceros, es decir,

$$\mathbf{0} = (0, 0, \ldots, 0),$$

y el vector opuesto para el vector (21.5) será

$$-x = (-\alpha_1, -\alpha_2, \ldots, -\alpha_n).$$

Este espacio es n-dimensional y una de sus bases se indica con facilidad inmediatamente. A saber,

$$\begin{aligned} e_1 &= (1, 0, \ldots, 0, 0), \\ e_2 &= (0, 1, \ldots, 0, 0), \\ &\cdots\cdots\cdots \\ e_n &= (0, 0, \ldots, 0, 1). \end{aligned} \tag{21.7}$$

Puesto que para el elemento (21.5) tiene lugar la descomposición

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n,$$

los números $\alpha_1, \alpha_2, \ldots, \alpha_n$ se llamarán *coordenadas* del vector x.

El espacio de tal tipo lo denominaremos *aritmético* y lo designaremos mediante P_n subrayando con esto su relación con el campo P. En el caso de que P sea un campo de los números complejos, reales o racionales, tales espacios n-dimensionales se designarán con $\mathbf{C}_n$, $\mathbf{R}_n$ o $\mathbf{D}_n$, respectivamente.

Ahora puede parecer que no hay ninguna necesidad en el estudio de los espacios lineales n-dimensionales arbitrarios. En efecto, sabemos que desde el punto de vista de los corolarios procedentes de los axiomas, los espacios lineales isomorfos son indistinguibles, razón por la cual siempre podemos estudiar con todo éxito solamente, por ejemplo, P_n. No obstante, los razonamientos generales permiten poner en evidencia las propiedades más importantes de los espacios lineales, es decir, aquellas que *no dependen de los sistemas básicos* o, en otras palabras, son *invariantes en los isomorfismos*.

Al estudiar solamente los espacios P_n, siempre seríamos atados a *una base concreta*, por lo cual no podríamos ver con facilidad la invariación de unas u otras deducciones. Con todo eso, se debe observar que las propiedades particulares del espacio P_n no se mezclen con las propiedades generales de los espacios lineales. No es tan fácil de hacerlo ni mucho menos.

Para concluir, indiquemos una circunstancia más. Por analogía con el espacio P_n consideraremos el espacio P_∞ cuyos elementos son toda clase de surtidos infinitos ordenados de los números $\alpha_1, \alpha_2, \ldots$ del campo P. El elemento x de este conjunto lo designaremos por analogía con (21.5)

$$x = (\alpha_1, \alpha_2, \ldots)$$

y por analogía con (21.6) introduciremos las operaciones sobre los elementos.

El espacio P_∞ ya será de dimensión infinita. Si suponemos que los espacios de dimensión infinita son isomorfos respecto del espacio P_∞, no será difícil comprender que los espacios de dimensión infinita y los de dimensión finita han de tener mucho en común. De eso no se debe olvidar.

Ejercicios.

1. Constrúyase una correspondencia isomorfa entre el espacio V_1 y un espacio de números reales dado sobre un campo de números reales.
2. Constrúyase una correspondencia isomorfa entre el espacio V_2 y un espacio de números complejos dado sobre un campo de números reales.
3. Demuéstrese que en los espacios isomorfos los sistemas equivalentes de vectores se transforman en sistemas equivalentes.
4. Demuéstrese que en los espacios isomorfos una intersección de subespacios se transforma en otra intersección de subespacios.
5. Demuéstrese que en los espacios isomorfos una suma directa de subespacios se transforma en otra suma directa de subespacios.

§ 22. Dependencia lineal y sistemas de ecuaciones lineales

La investigación de varias cuestiones, ligadas de uno u otro modo con la dependencia lineal, se reduce a la resolución del problema siguiente.

Sean dados un sistema de vectores $a_1, a_2, \ldots, a_m$ y el vector b. Se necesita establecer si el vector b es una combinación lineal del sistema dado de vectores y, además, hallar los coeficientes de la combinación lineal.

Si el vector b es una combinación lineal de los vectores $a_1, a_2, \ldots, a_m$, entonces existen tales números $z_1, z_2, \ldots, z_m$ que

$$z_1a_1 + z_2a_2 + \ldots + z_ma_m = b. \tag{22.1}$$

Por consiguiente, el problema planteado se reduce a la investigación de la ecuación vectorial (22.1) respecto de los números $z_1, z_2, \ldots, z_m$.

Supongamos que los vectores se han dado por sus coordenadas respecto de cierta base $e_1, e_2, \ldots, e_k$, es decir,

$$\begin{aligned} a_1 &= (a_{11}, a_{21}, \ldots, a_{k1}), \\ a_2 &= (a_{12}, a_{22}, \ldots, a_{k2}), \\ &\ldots\ldots\ldots\ldots \\ a_m &= (a_{1m}, a_{2m}, \ldots, a_{km}), \\ b &= (b_1, b_2, \ldots, b_k). \end{aligned}$$

Al igualar las coordenadas correspondientes de los vectores en los miembros primero y segundo de la ecuación (22.1), obtenemos

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 + \ldots + a_{1m}z_m &= b_1, \\ a_{21}z_1 + a_{22}z_2 + \ldots + a_{2m}z_m &= b_2, \\ &\ldots\ldots\ldots\ldots \\ a_{k1}z_1 + a_{k2}z_2 + \ldots + a_{km}z_m &= b_k. \end{aligned} \tag{22.2}$$

Este sistema de ecuaciones que refleja la inscripción de la ecuación (22.1) en coordenadas se llama *sistema de las ecuaciones algebraicas lineales*. Los números $b_1, b_2, \ldots, b_k$ se denominan *miembros derechos* y $z_1, z_2, \ldots, z_m$, *incógnitas* del sistema de ecuaciones. La totalidad ordenada de los valores de las incógnitas que satisface cada una de las ecuaciones (22.2) recibe el nombre de *solución* del sistema. Si el sistema de ecuaciones algebraicas lineales tiene al menos una sola solución, se llama *compatible*; en el caso contrario el sistema se denomina *incompatible*.

De este modo, la respuesta a la pregunta sobre si es o no el vector b una combinación lineal de los vectores $a_1, a_2, \ldots, a_m$ se determina por la compatibilidad o no compatibilidad del sistema (22.2). Si el sistema es compatible, cualquier solución de él proporciona los coeficientes de la descomposición del vector b según el sistema de vectores $a_1, a_2, \ldots, a_m$.

Dos sistemas de ecuaciones algebraicas lineales respecto de las mismas incógnitas se llaman *equivalentes*, si toda solución de un sistema es solución para el otro sistema o ambos sistemas son incompatibles.

El método general para solucionar los sistemas puede fundamentarse sobre la transformación sucesiva del sistema inicial (22.2) en tal sistema equivalente, para el cual la solución se halla de manera suficientemente simple. Demos a conocer ahora uno de los métodos llamado *método de eliminación* o *método de Gauss.*

El proceso de resolución incluye, en el caso general, no más que $k-1$ etapas. Con el fin de distinguir los coeficientes de las incógnitas y los miembros derechos, que se obtienen en el transcurso de la transformación en diferentes etapas, marcarémoslos con un índice adicional que se pone por arriba y significa el número de la etapa realizada. De conformidad con esta observación, el sistema inicial (22.2) tendrá la forma

$$\begin{aligned} a_{11}^{(0)}z_1+a_{12}^{(0)}z_2+\ldots+a_{1m}^{(0)}z_m&=b_1^{(0)},\\ a_{21}^{(0)}z_1+a_{22}^{(0)}z_2+\ldots+a_{2m}^{(0)}z_m&=b_2^{(0)},\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{k1}^{(0)}z_1+a_{k2}^{(0)}z_2+\ldots+a_{km}^{(0)}z_m&=b_k^{(0)}. \end{aligned} \tag{22.3}$$

Examinemos la primera ecuación. Si todos los coeficientes de las incógnitas y el miembro derecho son iguales a cero, entonces cualquier totalidad de los números $z_1, z_2, \ldots, z_m$ satisfará esta ecuación. Por consiguiente, al excluir de la consideración la primera ecuación, obtendremos un sistema equivalente. Puede suceder que todos los coeficientes de las incógnitas en la primera ecuación son iguales a cero, mientras que el miembro derecho no es nulo. En este caso ninguna totalidad de los números $z_1, z_2, \ldots, z_m$ puede satisfacer tal ecuación. Entonces, el sistema será incompatible y su investigación se ha dado por terminada.

Supongamos que entre los coeficientes de las incógnitas en la primera ecuación se tiene aunque uno diferente de cero. Sin restringir la generalidad, se puede considerar que $a_{11}^{(0)}\neq 0$, puesto que en el caso contrario se lo puede conseguir por permutación de las incógnitas. Llamemos al elemento $a_{11}^{(0)}$ elemento *rector*. Expresemos la incógnita z_1 de la primera ecuación en términos de las incógnitas restantes y del miembro derecho y a continuación sustituyamos la expresión obtenida para z_1 en todas las ecuaciones, salvo la primera. Reduciendo en todas las ecuaciones los términos semejantes obtendremos un sistema nuevo

$$\begin{aligned} a_{11}^{(0)}z_1+a_{12}^{(0)}z_2+\ldots+a_{1m}^{(0)}z_m&=b_1^{(0)},\\ a_{22}^{(1)}z_2+\ldots+a_{2m}^{(1)}z_m&=b_2^{(1)},\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{k2}^{(1)}z_2+\ldots+a_{km}^{(1)}z_m&=b_k^{(1)}. \end{aligned} \tag{22.4}$$

Los coeficientes de este sistema están ligados con los del sistema anterior mediante las siguientes correlaciones:

$$a_{ij}^{(1)} = a_{ij}^{(0)} - a_{1j}^{(0)} \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$$

$$b_i^{(1)} = b_i^{(0)} - b_1^{(0)} \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$$

para cualesquiera i, j.

Los sistemas (22.3) y (22.4) son equivalentes. En efecto, supongamos que el sistema (22.3) es compatible. Entonces, cualquier solución $z_1, z_2, \ldots, z_m$ del mismo convierte todas las ecuaciones del sistema (22.3) en las identidades. Reiterando de nuevo con cualquiera de las soluciones el proceso de eliminación, nos convencemos de que la solución mencionada es también solución del sistema (22.4). Supongamos luego que cierta solución del sistema (22.4) no es la solución del sistema (22.3). Satisface a ciencia cierta la primera ecuación de (22.3). Supongamos que dicha solución no satisface una ecuación de número $i \geqslant 2$. Entonces, repitiendo otra vez el proceso de eliminación, llegamos a que la solución elegida no debe satisfacer la i-ésima ecuación del sistema (22.4). Pero esto contradice a la hipótesis. Ahora está claro que si uno de los sistemas es incompatible, el otro sistema será también incompatible.

Hemos descrito solamente la primera etapa de la transformación del sistema. Todas las etapas restantes se realizan de la manera análoga. En la segunda etapa eliminemos la incógnita z_2 de todas las ecuaciones a excepción de las dos primeras; en la tercera etapa, la incógnita z_3 de todas las ecuaciones a excepción de las tres primeras, etc. Si, en el transcurso de las transformaciones, no encontramos las ecuaciones, donde todos los coeficientes de las incógnitas son iguales a cero, entonces acabadas $k-1$ etapas, llegaremos al sistema:

$$\begin{array}{rl}
a_{11}^{(0)}z_1 + a_{12}^{(0)}z_2 + \ldots + a_{1k}^{(0)}z_k & + a_{1,k+1}^{(0)}z_{k+1} + \ldots + a_{1m}^{(0)}z_m = b_1^{(0)}, \\
a_{22}^{(1)}z_2 + \ldots + a_{2k}^{(1)}z_k & + a_{2,k+1}^{(1)}z_{k+1} + \ldots + a_{2m}^{(1)}z_m = b_2^{(1)}, \\
\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{kk}^{(k-1)}z_k & + a_{k,k+1}^{(k-1)}z_{k+1} + \ldots + a_{km}^{(k-1)}z_m = b_k^{(k-1)},
\end{array} \qquad (22.5)$$

que es equivalente al sistema (22.4). Si en el transcurso de la transformación encontramos unas ecuaciones que se satisfacen idénticamente, el sistema (22.5) se compondrá de un número menor de ecuaciones.

Las incógnitas $z_{k+1}, \ldots, z_n$ se denominan *indeterminadas*. Evidentemente, cualesquiera que sean los valores que se atribuyan a estas incógnitas, todas las demás incógnitas del sistema (22.5), a partir de z_k, también pueden determinarse sucesivamente. Los coeficientes $a_{11}^{(0)}, a_{22}^{(1)}, \ldots, a_{kk}^{(k-1)}$, por los cuales tendremos que dividir, son elementos rectores en las etapas separadas, razón por la cual todos ellos son distintos de cero.

Así pues, desde el punto de vista *teórico*, el concepto de dependencia lineal se ha investigado en un grado suficientemente completo. No obstante, en el aspecto *práctico* puede conducir a unas dificultades muy serias. Por ejemplo, consideraremos en el espacio $\mathbf{R}_k$ un sistema de vectores

$$\begin{aligned} a_1 &= (1, & -2, && 0, \ldots, 0, && 0), \\ a_2 &= (0, & 1, && -2, \ldots, 0, && 0), \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k-1} &= (0, & 0, && 0, \ldots, 1, && -2), \\ a_k &= (-2^{-(k-1)}, & 0, && 0, \ldots, 0, && 1). \end{aligned} \tag{22.6}$$

Es linealmente dependiente, puesto que

$$2^{-k}a_1 + 2^{-(k-1)}a_2 + \ldots + 2^{-1}a_k = 0.$$

Observemos que $2^{-(k-1)} < 10^{-12}$ para $k > 40$, por lo cual al realizar los cálculos prácticos, surge, naturalmente, el deseo de menospreciar el valor de la coordenada tan pequeño. Además, todos los números, como regla, suelen conocerse con inexactitud y casi siempre contienen errores mucho más grandes. Pero, si incluso conociéramos las coordenadas con toda la exactitud, ya los cálculos iniciales con ellas llevarían a los resultados inexactos, si los cálculos se realizaban de manera aproximada. Se debe añadir que la mayoría de las máquinas computadoras modernas no percibe los números tan pequeños como es $2^{-(k-1)}$ para $k > 64$, y opera con ellos como si fueran ceros. Por esta razón, en realidad, en lugar del sistema de vectores (22.6) en la práctica podemos tropezar con un sistema del tipo:

$$\begin{aligned} a_1 &= (1, & -2, && 0, \ldots, 0, && 0), \\ a_2 &= (0, & 1, && -2, \ldots, 0, && 0), \\ & \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k-1} &= (0, & 0, && 0, \ldots, 1, && -2), \\ a_k &= (0, & 0, && 0, \ldots, 0, && 1). \end{aligned} \tag{22.7}$$

Mas, este sistema es linealmente independiente.

De este modo, las alteraciones pequeñas en coordenadas de los vectores pueden llevar a que bajo las condiciones en que las propias coordenadas se prefijen de modo aproximado y los cálculos con ellas sean también aproximados, un sistema linealmente dependiente puede hacerse linealmente independiente y, viceversa, un sistema linealmente independiente, linealmente dependiente. Pero en este caso será natural preguntar ¿qué sentido práctico tienen los conceptos de dependencia lineal, rango, base, sistemas compatible y no compatible y, en general, todo lo que hemos investigado hasta ahora? Para esta pregunta no existe respuesta sencilla, puesto que está asociada con una comprensión profunda de los problemas que se resuelven. *Esta pregunta da origen a las diferencias que destinguen la matemática "precisa" de la matemática "aproximada".*

Ejercicios.

1. Demuéstrese que si el sistema (22.2) es compatible, tendrá una única solución cuando, y sólo cuando, el sistema de vectores $a_1, a_2, \ldots, a_m$ sea linealmente independiente.

2. Demuéstrese que si el sistema de vectores $a_1, a_2, \ldots, a_m$ tiene el rango r, el sistema (22.5) se compondrá de r ecuaciones.

3. Como soluciones del sistema consideraremos los vectores del espacio P_m. Supongamos que $b = 0$ y el sistema de vectores $a_1, a_2, \ldots, a_m$ tiene el rango r. Demuéstrese que el conjunto de todas las soluciones del sistema (22.2) forma en este caso un subespacio $(m - r)$-dimensional del espacio P_m.

4. Hállense todas las soluciones del sistema de ecuaciones algebraicas lineales

$$\begin{aligned}\sqrt{2}\, z_1 + 1\cdot z_2 &= \sqrt{3},\\ 2\cdot z_1 + \sqrt{2}\, z_2 &= \sqrt{6}.\end{aligned}$$

Resuélvase este mismo sistema al profijar $\sqrt{2}$, $\sqrt{3}$, $\sqrt{6}$ con la exactitud diferente. Compárense los resultados entre sí.

5. Establézcase la relación existente entre el método de Gauss y las transformaciones elementales de un sistema de vectores.

CAPÍTULO 3 MEDICIONES EN EL ESPACIO LINEAL

§ 23. Sistemas afines de coordenadas

Hay gran cantidad de los problemas científicos y técnicos que requieren descripción exacta de la posición en el espacio de diferentes objetos geométricos tales como el punto, la figura, la línea, la superficie, etc. Cuando se trata de un objeto complejo, resulta muy importante conocer no sólo la característica general de la posición, como, por ejemplo, la ubicación del centro de gravedad, sino también la posición de todo punto del objeto.

A título de ejemplo recordemos que la predicción de los eclipses lunares y solares es posible gracias a que se conoce la posición de los cuerpos celestes en cualquier momento de tiempo. La transmisión de las imágenes de televisión a grandes distancias puede realizarse porque la posición de todo punto de la imagen que se transmite está bien definida.

Evidentemente, es necesario proporcionar un método que describe la posición de un solo punto, pues todo objeto geométrico puede ser definido como cierta totalidad de puntos. En este caso, obviamente conviene considerar independientemente la posición de un punto en una recta, en un plano o en un espacio, puesto que la descripción espacial de un objeto es lejos de ser siempre oportuna. Por ejemplo, una foto puede considerarse, a ciencia cierta, sólo en un plano, mientras el movimiento de un punto material, siendo ausentes las fuerzas que actúan contra él, sólo en una línea recta.

Una de las descripciones más difundidas de la posición de un punto está basada sobre una idea muy simple. Ya hemos notado que entre todos los puntos y los segmentos dirigidos sujetados puede establecerse una correspondencia biunívoca. Por ello, la descripción de la posición de un punto puede ser sustituida por la descripción de la posición del segmento dirigido correspondiente. Esta última se determina plenamente por las coordenadas del segmento respecto de cualquier base, es decir, por ciertos surtidos ordenados de números. Por consiguiente, la posición de un punto debe determinarse también por los surtidos ordenados de números. Pasamos, ahora, a la investigación de esta idea.

Sea dada una línea recta. Fijemos en ésta un punto arbitrario O y consideremos un espacio lineal V_1 de vectores dispuestos en la recta dada y sujetados al punto O. Elijamos en el citado espacio un vector básico a. Transformemos luego la recta en un eje, al definir en la misma una dirección de modo tal que la magnitud del segmento a sea positiva (fig. 23.1).

Un eje con el punto O y el vector básico a, definidos en el eje, forma *un sistema afín de coordenadas en la línea recta*. El punto O

Fig. 23.1.

se llama *origen del sistema de coordenadas* y la longitud del vector a, *unidad de escala*.

La posición de cualquier punto M en la recta se determina unívocamente por la posición del vector $\overrightarrow{OM}$. Los vectores a, $\overrightarrow{OM}$ son colineales y $a \neq 0$, a consecuencia de lo cual, de acuerdo con el corolario del lema 18.1, existe tal número real α que

$$\overrightarrow{OM} = \alpha a. \tag{23.1}$$

Este número se denomina *coordenada afín* del punto M en la recta. El hecho de que el punto M tiene la coordenada α se denota por el símbolo $M(\alpha)$.

Hemos de notar que, siendo fijado un sistema afín de coordenadas en la recta, la correlación (23.1) define unívocamente la coordenada afín α de cualquier punto M en la recta. Evidentemente, lo recíproco es también cierto. A saber, todo número α determina unívocamente, mediante la correlación (23.1), un cierto punto M de la línea recta. De este modo, siendo fijado un sistema afín de coordenadas, *existe una correspondencia biunívoca entre todos los números reales y los puntos de la recta.*

La definición de los puntos por medio de sus coordenadas permite calcular las magnitudes de los segmentos dirigidos y las distancias entre los puntos. Sean dados los puntos $M_1(\alpha_1)$ y $M_2(\alpha_2)$. Tenemos

$$\{\overrightarrow{M_1M_2}\} = \{\overrightarrow{OM}_2 - \overrightarrow{OM}_1\} = \{\alpha_2 a - \alpha_1 a\} =$$
$$= \{(\alpha_2 - \alpha_1) a\} = (\alpha_2 - \alpha_1)\{a\} = (\alpha_2 - \alpha_1)\,|a|. \tag{23.2}$$

Al designar mediante $\rho(M_1, M_2)$ la distancia entre los puntos M_1 y M_2, tenemos

$$\rho(M_1, M_2) = |\{\overrightarrow{M_1M_2}\}| = |\alpha_2 - \alpha_1|\,|a|. \tag{23.3}$$

Las fórmulas se hacen particularmente sencillas, si la longitud del vector básico es igual a la unidad. En este caso

$$\{\overrightarrow{M_1M_2}\} = \alpha_2 - \alpha_1, \qquad \rho(M_1, M_2) = |\alpha_2 - \alpha_1|. \tag{23.4}$$

Supongamos ahora que se ha dado un cierto plano. Fijemos en el mismo un punto arbitrario O y consideremos el espacio lineal V_2 de los vectores dispuestos en el plano dado y sujetados al punto O. Elijamos en el espacio un par cualquiera de vectores básicos a, b. A las líneas rectas que contienen dichos vectores atribuyámosles direcciones determinadas de un modo tal que las magnitudes de los segmentos a, b sean positivas (fig. 23.2).

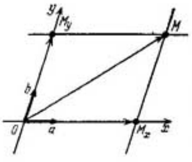


Fig. 23.2.

En un plano dos ejes que se cortan en el punto O y llevan los vectores básicos a, b dados forman *un sistema afín de coordenadas en el plano*. El eje que contiene el primer vector básico se llama eje Ox o eje de *abscisas*; el eje que contiene el segundo vector básico se denomina eje Oy o eje de *ordenadas*. La posición de cualquier punto M en el plano se determina también unívocamente por el vector $\overrightarrow{OM}$ y para éste, a su vez, existe la única descomposición del tipo

$$\overrightarrow{OM} = \alpha a + \beta b. \tag{23.5}$$

Los números reales α, β se llaman de nuevo *coordenadas afines* del punto M. La primera coordenada se llama *abscisa* y la segunda, *ordenada* de M. El hecho de que el punto M tiene las coordenadas α, β se denota mediante el símbolo $M(\alpha, \beta)$.

En los ejes de coordenadas Ox, Oy existen los únicos puntos M_x, M_y tales que

$$\overrightarrow{OM} = \overrightarrow{OM_x} + \overrightarrow{OM_y}. \tag{23.6}$$

Dichos puntos se disponen en la intersección de los ejes de coordenadas con las rectas que son paralelas a los ejes y pasan por el punto M. Se denominan *proyecciones afines* del punto M sobre los ejes de coordenadas. Los vectores $\overrightarrow{OM_x}$, $\overrightarrow{OM_y}$ se llaman *proyecciones afines del vector* $\overrightarrow{OM}$. Debido a la unicidad de las descomposiciones (23.5), (23.6) llegamos a la conclusión de que

$$\overrightarrow{OM_x} = \alpha a, \qquad \overrightarrow{OM_y} = \beta b. \tag{23.7}$$

De este modo, si el punto M tiene las coordenadas $M(\alpha, \beta)$ los puntos M_x, M_y, que pertenecen al plano, tienen las coordenadas $M_x(\alpha, 0)$, $M_y(0, \beta)$. Además, si

$$\overrightarrow{OM} = (\alpha, \beta),$$

entonces

$$\overrightarrow{OM_x} = (\alpha, 0), \qquad \overrightarrow{OM_y} = (0, \beta).$$

Todo vector básico forma en su eje un sistema propio de coordenadas. Por ello, los puntos M_x, M_y se pueden considerar también como puntos de los ejes Ox, Oy, dados en estos sistemas propios de coordenadas. Sin embargo, de (23.7) proviene que la coordenada del punto M_x en el eje Ox es igual a la abscisa del punto M. Análogamente, la coordenada del punto M_y en el eje Oy es igual a la ordenada del punto M. Siendo estas afirmaciones del todo evidentes, son de gran importancia, puesto que permiten emplear las fórmulas (23.2)—(23.4).

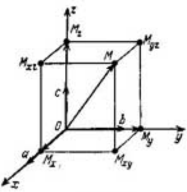


Fig. 23.3.

El par ordenado de números α, β determina unívocamente un punto. En efecto, las correlaciones (23.7) permiten construir unívocamente las proyecciones afines de un punto, las cuales, a su vez, determinan unívocamente, un punto del plano. Por consiguiente, siendo fijado un sistema afín de coordenadas, *existe una correspondencia biunívoca entre todos los pares ordenados de números reales y puntos del plano.*

De modo análogo se introduce el sistema afín de coordenadas en un espacio. Fijemos un punto O y consideremos el espacio lineal V_3 de vectores sujetados al punto O. Elijamos en este espacio una terna de los vectores básicos a, b, c. Atribuyamos a las rectas que contienen dichos vectores las direcciones determinadas de modo tal que las magnitudes de los segmentos a, b, c sean positivas (fig. 23.3).

Tres ejes en el espacio que se cortan en el punto O y llevan definidos en sí los vectores básicos a, b, c, forman *un sistema afín de coordenadas en el espacio.* El eje que contiene el primer vector básico se llama eje Ox o eje de abscisas, el eje con el segundo vector básico es Oy o eje de ordenadas, el tercer eje se denomina Oz o eje de z-

coordenadas. Los ejes de coordenadas elegidos dos a dos determinan los así llamados *planos de coordenadas* que se denominarán Oxy, Oyz y Oxz.

La posición de cualquier punto M de un espacio se determina, como antes, por el vector $\overrightarrow{OM}$, para el cual existe una única descomposición

$$\overrightarrow{OM} = \alpha a + \beta b + \gamma c.$$

Los números reales α, β, γ se llaman *coordenadas afines* del punto M en el espacio. La primera coordenada se llama *abscisa*, la segunda, *ordenada* y la coordenada tercera, *z-coordenada* del punto M. El hecho de que el punto M tiene las coordenadas α, β, γ se indicará mediante el símbolo $M(\alpha, \beta, \gamma)$.

Tracemos por el punto M del plano unos planos paralelos a los de coordenadas. Los puntos de intersección de estos planos con los ejes de coordenadas Ox, Oy, Oz se designarán con M_x, M_y, M_z y se denominarán *proyecciones afines del punto M sobre los ejes de coordenadas.* La intersección de los planos de coordenadas con los pares de planos que pasan por el punto M, determina los puntos M_{yz}, M_{xz}, M_{xy} que llevarán los nombres de *proyecciones afines del punto M sobre los planos de coordenadas.* Respectivamente, los vectores $\overrightarrow{OM}_{yz}$, $\overrightarrow{OM}_x$, etc. se denominarán *proyecciones afines del vector* $\overrightarrow{OM}$. Es evidente que

$$\overrightarrow{OM} = \overrightarrow{OM}_x + \overrightarrow{OM}_y + \overrightarrow{OM}_z,$$

$$\overrightarrow{OM}_{yz} = \overrightarrow{OM}_y + \overrightarrow{OM}_z,$$

$$\overrightarrow{OM}_{xz} = \overrightarrow{OM}_x + \overrightarrow{OM}_z,$$

$$\overrightarrow{OM}_{xy} = \overrightarrow{OM}_x + \overrightarrow{OM}_y.$$

Al igual que en el caso de un plano, si el punto M tiene las coordenadas

$$M(\alpha, \beta, \gamma),$$

las proyecciones afines de este punto tendrán por coordenadas:

$$M_x(\alpha, 0, 0), \quad M_y(0, \beta, 0), \quad M_z(0, 0, \gamma), \tag{23.8}$$
$$M_{yz}(0, \beta, \gamma), \quad M_{xz}(\alpha, 0, \gamma), \quad M_{xy}(\alpha, \beta, 0).$$

Por analogía, si

$$\overrightarrow{OM} = (\alpha, \beta, \gamma),$$

se tiene

$$\overrightarrow{OM}_x=(\alpha,\ 0,\ 0),\quad \overrightarrow{OM}_y=(0,\ \beta,\ 0),\quad \overrightarrow{OM}_z=(0,\ 0,\ \gamma),$$

$$\overrightarrow{OM}_{yz}=(0,\ \beta,\ \gamma),\quad \overrightarrow{OM}_{xz}=(\alpha,\ 0,\ \gamma),\quad \overrightarrow{OM}_{xy}=(\alpha,\ \beta,\ 0).$$

Esta vez también, todo vector básico y todo par de vectores básicos forman sistemas afines propios en los ejes de coordenadas y en los planos de coordenadas. Las coordenadas de los puntos en estos sistemas coinciden nuevamente con las coordenadas afines de los mismos, considerados como puntos del espacio. Ahora bien, siendo fijado un sistema afín de coordenadas, *existe una correspondencia biunívoca entre todas las ternas ordenadas de números reales y los puntos del espacio.*

Entre los sistemas afines de coordenadas en una recta, en un plano y en un espacio, son de mayor uso los llamados *sistemas rectangulares cartesianos de coordenadas.* Estos últimos se caracterizan por el hecho de que todos los vectores básicos tienen una longitud igual a la unidad y los ejes de coordenadas, tanto en el caso de un plano como en el de un espacio, son perpendiculares dos a dos. Los vectores básicos en el sistema cartesiano de coordenadas se designan, comúnmente, por las letras i, j, k. En lo sucesivo haremos uso, como regla, sólo de los sistemas que acabamos de mencionar.

Ejercicios.

1. ¿Cuál de los puntos A (α), B $(-\alpha)$ se halla más a la derecha en el eje de coordenadas dibujado en la fig. 23.1?
2. ¿Qué representa en sí el lugar geométrico de los puntos M $(\alpha,\ \beta,\ \gamma)$ para los cuales las proyecciones afines M_{xy} tienen por coordenadas M_{xy} $(-3,\ 2,\ 0)$?
3. ¿Dependen las coordenadas de los puntos del modo de elegir la dirección en los ejes de coordenadas?
4. ¿Cómo varían las coordenadas de los puntos con el cambio de longitud de los vectores básicos?
5. ¿Qué coordenadas tiene el centro de un paralelepípedo, si el origen de coordenadas coincide con uno de sus vértices y los vectores básicos, con las aristas?

§ 24. Otros sistemas de coordenadas

Los sistemas de coordenadas empleados en las matemáticas permiten prefijar, mediante números, la posición de cualquier punto de un espacio, de un plano o de una recta. Esto hace posibles toda clase de cálculos sobre las coordenadas y, lo que es muy importante, permite utilizar las computadoras modernas no sólo para diversos cálculos numéricos, sino también para la resolución de problemas geométricos, la investigación de cualesquiera objetos geométricos y correlaciones. Además de los sistemas afines de coordenadas examinados anteriormente se emplean con frecuencia otros sistemas.

Sistema polar de coordenadas. Elijamos en un plano una recta y fijemos en la misma el sistema cartesiano de coordenadas. El origen O de este sistema se llamará *polo* y el eje de coordenadas, *eje polar*. Consideraremos en lo sucesivo que el segmento de escala del sistema de coordenadas en la recta se emplea para medir longitudes de los segmentos cualesquiera en el plano. Examinemos un punto arbitrario M del plano. Es evidente que su posición estará completamente definida, si se prefijan la distancia ρ entre los puntos M, O y el ángulo φ que se forma al hacer girar el rayo Ox alrededor del punto O en *el sentido contrahorario* hasta que su dirección coincida con la del segmento $\overrightarrow{OM}$ (fig. 24.1).

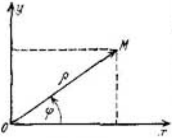


Fig. 24.1.

Se llaman *coordenadas polares* del punto M en un plano los números ρ y φ. El número ρ se llama *radio polar*, el número φ, *ángulo polar*. Comúnmente se supone que

$$0 \leqslant \varphi < +\infty, \ 0 \leqslant \varphi < 2\pi. \qquad (24.1)$$

Si el punto M coincide con el polo O, el ángulo polar se considera *indeterminado*.

Con cualquier sistema polar de coordenadas se enlaza de un modo natural cierto sistema rectangular cartesiano de coordenadas. En este sistema el origen de coordenadas coincide con el polo, el eje de abscisas, con el eje polar y el eje de ordenadas se obtiene girando el eje polar alrededor del punto O a un ángulo $\pi/2$.

Denotemos las coordenadas del punto M en el sistema rectangular cartesiano de coordenadas Oxy mediante α, β. Son evidentes las fórmulas

$$\alpha = \rho \cos \varphi, \quad \beta = \rho \operatorname{sen} \varphi.$$

De aquí obtenemos también las correlaciones inversas

$$\rho^2 = \alpha^2 + \beta^2, \ \cos \varphi = \frac{\alpha}{+(\alpha^2+\beta^2)^{1/2}},$$

$$\operatorname{sen} \varphi = \frac{\beta}{+(\alpha^2+\beta^2)^{1/2}}.$$

Estas fórmulas permiten calcular las coordenadas polares de un punto partiendo de sus coordenadas cartesianas, y viceversa.

Coordenadas cilíndricas. Elijamos en el espacio un plano cualquiera π y fijemos en éste un sistema polar de coordenadas. Por el polo O tracemos el eje Oz perpendicular al plano π (fig. 24.2). Convengamos, nuevamente, en considerar que para medir las longitudes de todos los segmentos en el espacio se utiliza un mismo segmento de escala. Introduzcamos en el plano π un sistema rectangular cartesiano de coordenadas que corresponda al sistema polar. Junto con el eje Oz el sistema introducido formará un sistema cartesiano de coordenadas en el espacio.

Consideraremos las proyecciones M_z y M_{xy} del punto M sobre el eje Oz y el plano Oxy. El punto M_{xy}, como punto del plano π,

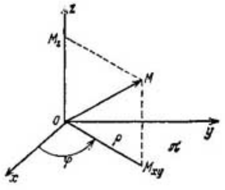


Fig. 24.2.

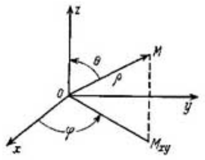


Fig. 24.3.

tiene por coordenadas polares ρ, φ. El punto M_z, que pertenece al eje Oz, tiene la coordenada z.

Se llaman *coordenadas cilíndricas* del punto M en el espacio tres números ρ, φ, z. En este caso se supone también que

$$0 \leqslant \rho < +\infty, \qquad 0 \leqslant \varphi < 2\pi.$$

El ángulo φ para los puntos del eje Oz *no está definido.*

La relación entre las coordenadas cartesianas en el sistema $Oxyz$ y las coordenadas cilíndricas se determina por medio de las correlaciones

$$x = \rho \cos \varphi, \qquad y = \rho \operatorname{sen} \varphi, \qquad z = z.$$

Coordenadas esféricas. Consideremos en un espacio el sistema rectangular cartesiano de coordenadas $Oxyz$ y el correspondiente sistema polar de coordenadas en el plano Oxy (fig. 24.3). Sea M un punto cualquiera del espacio distinto de O, y sea M_{xy} la proyección de M sobre el plano Oxy. Designemos con ρ la distancia del punto M al punto O y con θ, el ángulo formado por el vector $\overrightarrow{OM}$ y el vector básico del eje Oz. Sea, por fin, φ el ángulo polar de la proyección M_{xy}.

Se llaman *coordenadas esféricas* del punto M en el espacio tres números ρ, φ, θ. El número ρ es *el radio*, el número φ, *la longitud* y el número θ, *la latitud*. En este caso se supone que

$$0 \leqslant \rho < +\infty, \qquad 0 \leqslant \varphi < 2\pi, \qquad 0 \leqslant \theta \leqslant \pi.$$

La longitud *no está definida* para todos los puntos del eje Oz, mientras que la latitud no está definida para el punto O.

La relación existente entre las coordenadas cartesianas en el sistema $Oxyz$ y las coordenadas esféricas se determina por las correlaciones

$$x = \rho \operatorname{sen} \theta \cos \varphi, \qquad y = \rho \operatorname{sen} \theta \operatorname{sen} \varphi, \qquad z = \rho \cos \theta.$$

Ejercicios.

1. Constrúyase una línea tal que las coordenadas de sus puntos en el sistema polar satisfagan la correlación $\rho = \cos 3\varphi$.
2. Constrúyase una línea tal que las coordenadas de sus puntos en el sistema cilíndrico satisfagan las correlaciones $\rho = \varphi^{-1}$, $z = \varphi$.
3. Constrúyase una superficie tal que las coordenadas de sus puntos en el sistema esférico de coordenadas satisfagan las correlaciones

$$0 \leqslant \rho \leqslant 1, \qquad \varphi = \pi/2, \qquad 0 \leqslant \theta \leqslant \pi/2.$$

§ 25. Problemas

Consideremos varios problemas sencillos referentes a la aplicación de los sistemas rectangulares cartesianos de coordenadas. Para concretar, examinemos dichos problemas en el espacio. Los problemas análogos en un plano difieren de éstos sólo en pequeños detalles insignificantes. Convengamos en considerar siempre que se tiene un sistema fijado de coordenadas cuyo origen es el punto O y los vectores básicos son i, j, k.

Coordenadas de un vector. Supongamos que en el espacio están dados dos puntos M_1 $(\alpha_1, \beta_1, \gamma_1)$ y M_2 $(\alpha_2, \beta_2, \gamma_2)$. Estos puntos determinan el vector $\overrightarrow{M_1M_2}$, que tiene ciertas coordenadas respecto de la base i, j, k. Establezcamos la relación que existe entre las coordenadas del vector M_1M_2 y las de los puntos $\overrightarrow{M_1, M_2}$. Tenemos

$$\overrightarrow{M_1M_2} = \overrightarrow{OM_2} - \overrightarrow{OM_1}.$$

Luego, por definición de coordenadas afines de los puntos M_1, M_2,

$$\overrightarrow{OM_1} = \alpha_1 i + \beta_1 j + \gamma_1 k, \quad \overrightarrow{OM_2} = \alpha_2 i + \beta_2 j + \gamma_2 k.$$

Por ello, de aquí se deduce que

$$M_1M_2 = (\alpha_2 - \alpha_1)\, i + (\beta_2 - \beta_1)\, j + (\gamma_2 - \gamma_1)\, k,$$

o bien, de acuerdo con las designaciones aceptadas,

$$\overrightarrow{M_1M_2}=(\alpha_2-\alpha_1,\ \beta_2-\beta_1,\ \gamma_2-\gamma_1). \qquad (25.1)$$

Coordenadas de las proyecciones de un vector. Consideremos, otra vez, en el espacio el segmento dirigido $\overrightarrow{M_1M_2}$. Al proyectar los puntos M_1 y M_2 sobre el mismo plano de coordenadas o el mismo eje de coordenadas, obtendremos un nuevo segmento dirigido. Éste se llama *proyección coordenada* del *vector* $\overrightarrow{M_1M_2}$.

Todo vector de un espacio tiene seis proyecciones coordenadas: tres proyecciones sobre los ejes de coordenadas y tres sobre los planos de coordenadas. Son fáciles de calcular las coordenadas de las proyecciones en la base i, j, k según las coordenadas de los puntos M_1 $(\alpha_1, \beta_1, \gamma_1)$, M_2 $(\alpha_2, \beta_2, \gamma_2)$. Con este fin es suficiente hacer uso de las fórmulas (23.8), (25.1).

Supongamos, por ejemplo, que deseamos calcular las coordenadas de la proyección $\overrightarrow{M_{1x}M_{2x}}$. Teniendo presente que los puntos M_{1x} y M_{2x} tienen por coordenadas

$$M_{1x}\ (\alpha_1,\ 0,\ 0), \qquad M_{2x}\ (\alpha_2,\ 0,\ 0),$$

hallamos que

$$\overrightarrow{M_{1x}M_{2x}}=(\alpha_2-\alpha_1,\ 0,\ 0). \qquad (25.2)$$

Por analogía,

$$\overrightarrow{M_{1xz}M_{2xz}}=(\alpha_2-\alpha_1,\ 0,\ \gamma_2-\gamma_1),$$

y así para el resto de las proyecciones.

Al comparar la primera de las fórmulas (23.4) con las fórmulas del tipo (25.2), llegamos a que

$$\{\overrightarrow{M_{1x}M_{2x}}\}=\alpha_2-\alpha_1, \quad \{\overrightarrow{M_{1y}M_{2y}}\}=\beta_2-\beta_1, \quad \{\overrightarrow{M_{1z}M_{2z}}\}=\gamma_2-\gamma_1.$$

Por esto la magnitud de las proyecciones del vector sobre los ejes de coordenadas coinciden con las coordenadas de este vector.

La segunda de las fórmulas (23.4) permite calcular, a partir de las coordenadas de los puntos M_1 y M_2, las longitudes de las proyecciones del vector $\overrightarrow{M_1M_2}$ sobre los ejes de coordenadas. A saber,

$$|\overrightarrow{M_{1x}M_{2x}}|=|\alpha_2-\alpha_1|, \quad |\overrightarrow{M_{1y}M_{2y}}|=|\beta_2-\beta_1|, \quad |\overrightarrow{M_{1z}M_{2z}}|=|\gamma_2-\gamma_1|.$$

Longitud de un vector. Deduzcamos la fórmula para calcular la longitud de un vector en el espacio. Es obvio que la longitud $|\overrightarrow{M_1M_2}|$ del vector $\overrightarrow{M_1M_2}$ es igual a la distancia $\rho\ (M_1, M_2)$ entre los puntos

M_1, M_2 y equivale también a la longitud de la diagonal de un paralelepípedo rectangular cuyas caras son paralelas a los planos de coordenadas y pasan por los puntos M_1 y M_2 (fig. 25.1). La longitud de cualquier arista del paralelepípedo es igual a la de la proyección del

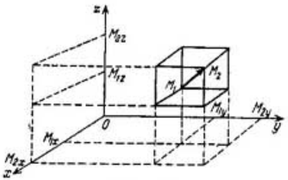


Fig. 25.1.

vector $\overrightarrow{M_1M_2}$ sobre el eje de coordenadas paralelo a la arista. Por ende, haciendo uso del teorema de Pitágoras, obtenemos

$$|\overrightarrow{M_1M_2}| = (|\overrightarrow{M_{1x}M_{2x}}|^2 + |\overrightarrow{M_{1y}M_{2y}}|^2 + |\overrightarrow{M_{1z}M_{2z}}|^2)^{1/2}.$$

Ahora, si los puntos M_1, M_2 están definidos por sus coordenadas M_1 $(\alpha_1, \beta_1, \gamma_1)$ y M_2 $(\alpha_2, \beta_2, \gamma_2)$, entonces

$$\rho(M_1, M_2) = ((\alpha_2 - \alpha_1)^2 + (\beta_2 - \beta_1)^2 + (\gamma_2 - \gamma_1)^2)^{1/2}. \qquad (25.3)$$

Si el vector $\overrightarrow{M_1M_2}$ está definido por las coordenadas x, y, z respecto de la base i, j, k, entonces

$$|\overrightarrow{M_1M_2}| = (x^2 + y^2 + z^2)^{1/2}. \qquad (25.4)$$

Una fórma análoga la tienen las fórmulas también en el caso de un plano. Si los puntos M_1 (α_1, β_1), M_2 (α_2, β_2) o el vector $\overrightarrow{M_1M_2} = (x, y)$ vienen dados mediante sus coordenadas, entonces

$$\rho(M_1, M_2) = ((\alpha_2 - \alpha_1)^2 + (\beta_2 - \beta_1)^2)^{1/2}, \quad |\overrightarrow{M_1M_2}| = (x^2 + y^2)^{1/2}.$$

Ángulo formado por vectores. Examinemos los vectores no nulos a, b en el espacio. Apliquémoslos al punto O. Designemos con π el plano que pasa por el punto O y contiene ambos vectores. Se llama *ángulo formado por dos vectores* a, b el ángulo *mínimo* que se obtiene girando alrededor del punto O uno de los vectores en el plano π para que su dirección coincida con la del otro vector. Si al menos uno de los vectores es nulo, el ángulo *no está definido*. Nuestra tarea consistirá en calcular el coseno del ángulo entre dos vectores, a partir de las coordenadas de dichos vectores. Para el coseno admitamos la designación $\cos\{a, b\}$.

Denotaremos mediante A, B los extremos de los vectores a, b en el plano π. Evidentemente, el ángulo formado por los vectores a, b no es otra cosa que el ángulo AOB del triángulo AOB cuyos lados constituyen los vectores a, b y $b-a$ (fig. 25.2).

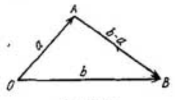


Fig. 25.2.

Supongamos que los vectores a, b están dados mediante sus coordenadas

$$a = (x_1, y_1, z_1), \qquad b = (x_2, y_2, z_2).$$

En este caso

$$b - a = (x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

Según se sabe de la geometría elemental, el cuadrado de la longitud de un lado del triángulo es igual a la suma de los cuadrados de las longitudes de los otros dos lados menos el producto duplicado de las longitudes de estos lados por el coseno del ángulo entre ellos. Por esta razón

$$|b-a|^2 = |a|^2 + |b|^2 - 2\,|a|\cdot|b|\cos\{a, b\}$$

o bien, al tomar en consideración la fórmula (25.4),

$$(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2 = x_1^2+y_1^2+z_1^2+$$
$$+x_2^2+y_2^2+z_2^2-2\,(x_1^2+y_1^2+z_1^2)^{1/2}\,(x_2^2+y_2^2+z_2^2)^{1/2}\cos\{a, b\}.$$

Realizando ciertas transformaciones elementales hallamos

$$\cos\{a, b\} = \frac{x_1x_2+y_1y_2+z_1z_2}{(x_1^2+y_1^2+z_1^2)^{1/2}\,(x_2^2+y_2^2+z_2^2)^{1/2}} \tag{25.5}$$

Los cambios en la fórmula para el caso de un plano son obvios.

División de un segmento en la razón dada. Supongamos que en un espacio se han dado una recta y dos puntos distintos en la misma M_1 y M_2. Elijamos en dicha recta la dirección positiva. En el eje obtenido los puntos M_1 y M_2 determinan un segmento dirigido $\overrightarrow{M_1M_2}$. Sea M un punto cualquiera del eje, distinto de M_2. El número

$$\lambda = \frac{\{\overrightarrow{M_1M}\}}{\{\overrightarrow{MM_2}\}} \tag{25.6}$$

se llama *razón en que el punto M divide el segmento dirigido* $\overrightarrow{M_1M_2}$.

Con el cambio de dirección en el eje, los números $\{\overrightarrow{M_1M}\}$ y $\{\overrightarrow{MM_2}\}$ cambian de signo simultáneamente. Por lo tanto, la razón (25.6) no depende de la dirección positiva elegida en el eje. Luego, con el cambio de la escala de longitudes de los segmentos en el eje, los números $\{\overrightarrow{M_1M}\}$ y $\{\overrightarrow{MM_2}\}$ se multiplican por el mismo número. Por lo tanto, la razón (25.6) no depende de la unidad elegida para medir longitudes. De aquí se desprende que la razón (25.6) *no depende de cómo se elige en el eje el sistema de coordenadas.*

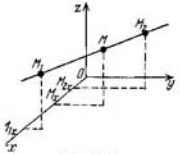


Fig. 25.3.

El problema consiste en calcular las coordenadas del punto M, que divide el segmento $\overrightarrow{M_1M_2}$ en la razón λ, si se conocen las coordenadas de los puntos M_1, M_2 y el número λ, con la particularidad de que $\lambda \neq -1$. Así pues, supongamos que están dados M_1 $(\alpha_1, \beta_1, \gamma_1)$, M_2 $(\alpha_2, \beta_2, \gamma_2)$ y se desconoce M (α, β, γ). Proyectemos estos puntos sobre los ejes de coordenadas, por ejemplo, sobre el eje Ox (fig. 25.3). De los razonamientos de semejanza se ve que el punto M_x divide el segmento dirigido $\overrightarrow{M_{1x}M_{2x}}$ también en la razón λ. Por ello,

$$\lambda = \frac{\{\overrightarrow{M_{1x}M_x}\}}{\{\overrightarrow{M_xM_{2x}}\}}. \tag{25.7}$$

De conformidad con la fórmula (23.4), $\{\overrightarrow{M_{1x}M_x}\} = \alpha - \alpha_1$, $\{\overrightarrow{M_xM_{2x}}\} = \alpha_2 - \alpha$. Ahora, teniendo en cuenta la expresión (25.7), hallamos que $\alpha = (\alpha_1 + \lambda\alpha_2)/(1 + \lambda)$. De modo análogo se calculan las coordenadas β y γ. Así pues,

$$\alpha = \frac{\alpha_1 + \lambda\alpha_2}{1+\lambda}, \qquad \beta = \frac{\beta_1 + \lambda\beta_2}{1+\lambda}, \qquad \gamma = \frac{\gamma_1 + \lambda\gamma_2}{1+\lambda}.$$

Observemos que $\lambda > 0$, si el punto M se halla dentro del segmento $\overrightarrow{M_1M_2}$; $\lambda < 0$, si el punto M se encuentra fuera del segmento $\overrightarrow{M_1M_2}$, y $\lambda = 0$, si el punto M coincide con M_1. Cuando el punto M se desplaza desde el punto M_1 hasta el punto M_2 (excluyendo el punto M_2), la razón λ toma primero el valor nulo y luego, de la manera sucesiva, todos los valores positivos posibles siempre crecientes. Si el punto M se desplaza desde el punto M_1 en dirección positiva del eje (véase la fig. 25.3), la razón λ tomará primero el valor cero y a continuación valores negativos siempre decrecientes, aproximándose tan cerca como se quiera al valor de $\lambda = -1$, pero quedando en todo momento mayor que dicho valor. Si el punto M se desplaza en dirección negativa desde el punto M_2, la razón λ toma todos los valores negativos posibles en el orden de crecimiento, quedando siempre inferior a $\lambda = -1$.

De este modo, entre todos los números reales y los puntos de la recta se podría establecer una correspondencia biunívoca, si en la recta hubiera un punto M que divida el segmento $\overrightarrow{M_1M_2}$ en la razón $\lambda = -1$ y si al punto M, coincidente con M_2, se le pudiera poner en correspondencia algún número. Este problema se resuelve comúnmente al completar la recta con un "punto" complementario convencional y los "números", con un "número" complementario convencional. Tal punto se denomina "infinito" y el número, "infinitamente grande".

Proyecciones ortogonales del vector. Sean dados en el espacio cierto eje u y un segmento dirigido $\overrightarrow{AB}$. Tracemos por los puntos A, B unos planos perpendiculares al eje u (fig. 25.4). La intersección de estos planos con el eje determina los puntos A_u, B_u, de los cuales A_u se ubica en el mismo plano con A, mientras que B_u se halla en el mismo plano con B. El segmento dirigido $\overrightarrow{A_uB_u}$ se llama *proyección ortogonal del segmento* $\overrightarrow{AB}$ *sobre el eje* u. Para designarlo se utiliza el siguiente símbolo:

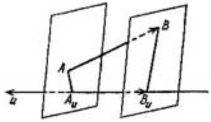


Fig. 25.4.

$$\overrightarrow{A_uB_u} = \mathrm{pr}_u\, \overrightarrow{AB}.$$

Con el eje u fijado, todo vector x del espacio define unívocamente su proyección ortogonal x'. Se puede considerar, por ende, que tenemos una "función" dada

$$x' = \mathrm{pr}_u\, x, \tag{25.8}$$

cuyo "argumento" puede constituir cualquier vector del espacio y el "valor", un vector en el eje u. Demostraremos ahora que dicha función posee las siguientes propiedades:

$$\left.\begin{aligned} \mathrm{pr}_u\,(x+y) &= \mathrm{pr}_u\, x + \mathrm{pr}_u\, y, \\ \mathrm{pr}_u\,(\lambda x) &= \lambda\, \mathrm{pr}_u\, x, \end{aligned}\right\} \tag{25.9}$$

válidas para cualesquiera vectores x e y, como también para todo número λ.

En efecto, fijemos un sistema rectangular cartesiano de coordenadas en el que el eje u coincida con el eje coordenado de abscisas. Supongamos que en este sistema

$$x = (\alpha_1,\ \beta_1,\ \gamma_1),$$
$$y = (\alpha_2,\ \beta_2,\ \gamma_2),$$

entonces

$$x+y=(\alpha_1+\alpha_2,\ \beta_1+\beta_2,\ \gamma_1+\gamma_2),$$
$$\lambda x=(\lambda\alpha_1,\ \lambda\beta_1,\ \lambda\gamma_1).$$

En el sistema de coordenadas elegido la proyección ortogonal del vector sobre el eje u coincide con su proyección coordenada sobre el eje de abscisas. Como se ha indicado antes, la proyección de cualquier vector sobre el eje de abscisas tiene la primera coordenada coincidente con la primera coordenada del propio vector, mientras que las coordenadas restantes son nulas. Por eso

$$\begin{aligned}
&\mathrm{pr}_u (x+y)=(\alpha_1+\alpha_2,\ 0,\ 0),\\
&\mathrm{pr}_u (\lambda x)=(\lambda\alpha_1,\ 0,\ 0),\\
&\mathrm{pr}_u\, x=(\alpha_1,\ 0,\ 0),\\
&\mathrm{pr}_u\, y=(\alpha_2,\ 0,\ 0).
\end{aligned} \qquad (25.10)$$

De acuerdo con la regla de adición de vectores y multiplicación de éstos por un número, de las últimas dos igualdades (25.10) concluímos que

$$\mathrm{pr}_u\, x+\mathrm{pr}_u\, y=(\alpha_1+\alpha_2,\ 0,\ 0),$$
$$\lambda\, \mathrm{pr}_u\, x=(\lambda\alpha_1,\ 0,\ 0).$$

Comparando los segundos miembros de las igualdades obtenidas con los miembros correspondientes de las primeras dos (25.10), nos convencemos de que ambas propiedades (25.9) son justas.

Sean dados en el espacio un plano π y un segmento dirigido $\overrightarrow{AB}$. Al trazar perpendiculares de los puntos A y B sobre el plano π, obtendremos en el plano dado dos puntos A_π y B_π que determinan el segmento dirigido $\overrightarrow{A_\pi B_\pi}$. Este segmento lleva el nombre de *proyección ortogonal del segmento dirigido AB sobre el plano π*. Para designarlo se utiliza la misma notación, es decir,

$$\overrightarrow{A_\pi B_\pi}=\mathrm{pr}_\pi\, \overrightarrow{AB}.$$

Por supuesto, para las proyecciones ortogonales sobre un mismo plano tienen lugar correlaciones análogas a (25.9). Para demostar esta afirmación, se puede fijar un sistema rectangular cartesiano de coordenadas en el que el plano π sea un plano coordenado y otra vez, hacer uso de las propiedades correspondientes de las proyecciones sobre un plano de coordenadas.

Hemos considerado las proyecciones ortogonales de los vectores en el espacio. Indudablemente, la analogía completa tiene lugar también para los vectores en un plano.

Ejercicios.

1. Dos vectores no nulos están dados mediante sus coordenadas cartesianas. ¿En qué casos serán perpendiculares entre sí?
2. Hállense las coordenadas del centro de gravedad de tres puntos materiales, si se conocen sus coordenadas cartesianas y sus masas.
3. Hállese el área de un triángulo, si se conocen las coordenadas cartesianas de sus tres vértices.
4. En un espacio están dados los vectores no nulos x, a, b, c, con la particularidad de que a, b, c son perpendiculares dos a dos. Demuéstrese que

$$\cos^2\{x,\ a\}+\cos^2\{x,\ b\}+\cos^2\{x,\ c\}=1.$$

5. Designemos con π un plano coordenado cualquiera y con u, un eje coordenado cualquiera en el plano π. Demuéstrese que para cualquier vector x se verifica

$$\mathrm{pr}_u\,(\mathrm{pr}_\pi\, x)=\mathrm{pr}_u\, x.$$

§ 26. Producto escalar

El uso de los segmentos dirigidos para expresar fuerzas y desplazamientos conduce a un concepto muy importante de producto escalar de vectores.

Del curso de la física sabemos que si un vector a representa una fuerza tal, que el punto de su aplicación se desplaza desde el origen del vector b al extremo del mismo, el trabajo ω de esta fuerza se determinará por la igualdad

$$\omega=|a|\,|b|\cos\{a,\ b\}. \tag{26.1}$$

El segundo miembro de esta igualdad se llama *producto escalar de los vectores* a, b. Para su designación se ha aceptado el símbolo (a, b). De este modo,

$$(a,\ b)=|a|\,|b|\cos\{a,\ b\}. \tag{26.2}$$

Hablando en rigor, la definición aducida de producto escalar se refiere sólo a los vectores no nulos a, b, puesto que solamente para vectores de tal tipo está definido el ángulo. Sin embargo, tomando en consideración la preimagen del producto escalar, es fácil comprender cúal ha de ser la definición adicional en el caso en que siquiera uno de los vectores sea igual a cero. Si, bien una fuerza, bien un desplazamiento, se da mediante un vector nulo, el trabajo que se cumple es nulo. Por esta razón consideraremos que $(a, b)=0$, siempre que al menos uno de los vectores a, b sea igual a cero.

De la fórmula (26.2) provienen ciertas propiedades geométricas del producto escalar. Por ejemplo, el ángulo, formado por dos vectores no nulos, será agudo (obtuso) cuando, y sólo cuando, el producto escalar de estos vectores es positivo (negativo).

Si el ángulo formado por los vectores es recto o al menos uno de los vectores es nulo, el producto escalar será igual a cero. Los vectores de este tipo se llamarán *ortogonales*.

Se denominarán vectores *ortonormalizados* aquellos vectores ortogonales cuya longitud es igual a la unidad. En particular, los vectores básicos i, j, k de un sistema rectangular cartesiano de coordenadas son ortonormalizados. De la fórmula (26,2) se deduce que

$$\begin{aligned} &(i, i)=1, \quad (i, j)=0, \quad (i, k)=0, \\ &(j, i)=0, \quad (j, j)=1, \quad (j, k)=0, \\ &(k, i)=0, \quad (k, j)=0, \quad (k, k)=1. \end{aligned} \tag{26.3}$$

Examinemos los vectores no nulos a, b. Por el vector a tracemos el eje u atribuyéndole una dirección tal que la magnitud del vector a sea positiva. Será evidente que

$$\{\mathrm{pr}_u b\} = |b| \cos\{a, b\}.$$

La proyección del vector b sobre un eje construido de este modo, la llamaremos *proyección del vector b sobre el vector a* y la indicaremos mediante el símbolo $\mathrm{pr}_a b$. Naturalmente, la proyección de un vector sobre otro conserva las propiedades (25.9). En las designaciones nuevas

$$(a, b) = |a| \{\mathrm{pr}_a b\} = |b| \{\mathrm{pr}_b a\}. \tag{26.4}$$

Estas fórmulas permiten establecer propiedades algebraicas muy importantes del producto escalar. A saber, para cualesquiera vectores a, b, c y todo número real α se verifican las correlaciones:

$$\begin{aligned} &1)\ (a, b) = (b, a), \\ &2)\ (\alpha a, b) = \alpha (a, b), \\ &3)\ (a + b, c) = (a, c) + (b, c), \\ &4)\ (a, a) > 0 \text{ cuando } a \neq 0; \ (0, 0) = 0. \end{aligned} \tag{26.5}$$

Hemos de notar que las correlaciones (26.5) se cumplen, a ciencia cierta, si por lo menos uno de los vectores es nulo. En el caso general la validez de las propiedades 1,4 se desprende directamente de la fórmula (26.2). En cuanto a las propiedades 2,3, haremos uso de las fórmulas (26.4) y las propiedades de proyecciones, para establecerlas. Tenemos

$$(\alpha a, b) = |b| \{\mathrm{pr}_b (\alpha a)\} = |b| \{\alpha \cdot \mathrm{pr}_b a\} = \alpha |b| \{\mathrm{pr}_b a\} = \alpha (a, b),$$
$$(a + b, c) = |c| \{\mathrm{pr}_c (a + b)\} = |c| \{\mathrm{pr}_c a + \mathrm{pr}_c b\} =$$
$$= |c| \{\mathrm{pr}_c a\} + |c| \{\mathrm{pr}_c b\} = (a, c) + (b, c).$$

Las propiedades 2,3 sólo están ligadas con el primer factor del producto escalar. Las propiedades análogas tienen lugar también respecto del segundo factor. En efecto,

$$(a, \alpha b) = (\alpha b, a) = \alpha (b, a) = \alpha (a, b),$$
$$(a, b + c) = (b + c, a) = (b, a) + (c, a) = (a, b) + (a, c).$$

Además, en virtud de la igualdad $a - b = a + (-1)\, b$, serán lícitas también las siguientes correlaciones:

$$(a - b,\ c) = (a,\ c) - (b,\ c),$$
$$(a,\ b - c) = (a,\ b) - (a,\ c)$$

puesto que

$$(a - b,\ c) = (a + (-1)\, b,\ c) = (a,\ c) + ((-1)\, b,\ c) =$$
$$= (a,\ c) + (-1)\,(b,\ c) = (a,\ c) - (b,\ c).$$

TEOREMA 26.1. *Si dos vectores a, b están dados mediante sus coordenadas rectangulares cartesianas, entonces el producto escalar de estos vectores es igual a la suma de los productos, realizados dos a dos, de las coordenadas correspondientes.*

DEMOSTRACION. Supongamos, para concretar, que los vectores están dados en un espacio, es decir, $a = (x_1,\ y_1,\ z_1)$, $b = (x_2,\ y_2,\ z_2)$. Puesto que

$$a = x_1 i + y_1 j + z_1 k,$$
$$b = x_2 i + y_2 j + z_2 k,$$

entonces, al efectuar las transformaciones algebraicas del producto escalar, obtenemos

$$(a,\ b) = x_1x_2\,(i,\ i) + x_1y_2\,(i,\ j) + x_1z_2\,(i,\ k) +$$
$$+ y_1x_2\,(j,\ i) + y_1y_2\,(j,\ j) + y_1z_2\,(j,\ k) +$$
$$+ z_1x_2\,(k,\ i) + z_1y_2\,(k,\ j) + z_1z_2\,(k,\ k).$$

Ahora, de acuerdo con (26.3) se tiene

$$(a,\ b) = x_1x_2 + y_1y_2 + z_1z_2, \tag{26.6}$$

y el teorema queda demostrado.

La fórmula (26.6) permite escribir las expresiones, recién obtenidas, (25.4), (25.5) para la longitud del vector y el ángulo formado por los vectores en términos de un producto escalar. A saber,

$$|a| = (a,\ a)^{1/2}, \tag{26.7}$$
$$\cos\{a,\ b\} = \frac{(a,\ b)}{|a| \cdot |b|}$$

Puede parecer que estas fórmulas son triviales, puesto que se desprenden directamente de (26.2) sin referencia alguna a las fórmulas (25.4), (25.5). No obstante, sin darnos prisa en hacer tal deducción, veamos una circunstancia de importancia particular.

Ha de ser notado que toda nuestra investigación se desarrollaba, de hecho, en tres etapas. En primer lugar, partiendo de la fórmula (26.2), hemos demostrado la validez de las propiedades (26.5). A continuación, basándonos solamente en estas propiedades y en el carácter ortonormal de los vectores básicos del sistema de coordenadas, deducimos la fórmula (26.6). Y, por fin, haciendo uso de las fórmulas (25.4), (25.5), que se han obtenido sin referencias al concepto de producto escalar de los vectores, llegamos a las fórmulas (26.7).

Al partir de lo dicho, podríamos ahora introducir el producto escalar no por definición de su forma explícita, sino *de un modo axiomático*, como cierta función numérica definida para todo par de vectores, exigiendo en tal caso que las propiedades (26.5) se cumplan sin falta. Entonces, para cualesquiera sistemas de coordenadas, donde los vectores básicos están ortonormalizados en el sentido de producto escalar axiomático, seguirá siendo válida la correlación (26.6). Por consiguiente, teniendo presente un modelo del sistema rectangular cartesiano de coordenadas, podríamos considerar axiomáticamente que las longitudes de los vectores y los ángulos entre ellos se calculan según las fórmulas (26.7). Desde luego, tendríamos que convencernos, en tal caso, de que las longitudes y los ángulos que se han introducido de modo semejante, poseen las propiedades necesarias.

Ejercicios.

1. Están dados dos vectores a y b. ¿Bajo qué condiciones impuestas al número α los vectores a y $b + \alpha a$ son ortogonales? ¿Cuál es la interpretación geométrica de este problema?
2. El vector a se ha dado en el espacio V_3 mediante sus coordenadas cartesianas. Hállense dos vectores linealmente independientes ortogonales al vector a.
3. Los vectores linealmente independientes a, b están dados en el espacio V_3 mediante sus coordenadas cartesianas. Hállense el vector no nulo que sea ortogonal respecto de ambos vectores.
4. ¿Qué representa en sí el lugar geométrico de vectores ortogonales a un vector dado?

§ 27. Espacio euclídeo

Los espacios lineales abstractos estudiados anteriormente son, en cierto sentido, más pobres de conceptos y propiedades en comparación con el espacio de segmentos dirigidos. Esta pobreza se debe, en primer lugar, a que en los primeros no se han reflejado los hechos más importantes relacionados con la medición de longitudes, de ángulos, de volúmenes, etc. Los conceptos métricos pueden ser extendidos a los espacios lineales abstractos por medio de varios procedimientos. No obstante, un método más eficaz para posibilitar la realización de mediciones consiste en la introducción *axiomática* de un producto escalar de vectores. Comenzaremos nuestras investigaciones con los espacios lineales reales.

Un espacio lineal real E se llama *euclídeo*, si a todo par de vectores x, y de E se le ha puesto en correspondencia un número real (x, y), llamado *producto escalar*, considerándose cumplidos los siguientes axiomas:

$$\begin{aligned}
&1)\ (x,\ y)=(y,\ x),\\
&2)\ (\lambda x,\ y)=\lambda(x,\ y),\\
&3)\ (x+y,\ z)=(x,\ z)+(y,\ z),\\
&4)\ (x,\ x)>0 \text{ para } x\neq 0;\ (0,\ 0)=0
\end{aligned}\tag{27.1}$$

para los vectores arbitrarios x, y, z de E y el número real arbitrario λ.

Como ya sabemos de estos axiomas se desprende que con el producto escalar se pueden realizar transformaciones algebraicas formales, es decir,

$$\left(\sum_{i=1}^{r}\alpha_i x_i,\ \sum_{j=1}^{s}\beta_j y_j\right)=\sum_{i=1}^{r}\sum_{j=1}^{s}\alpha_i\beta_j(x_i,\ y_j)$$

para cualesquiera vectores x_i, y_j, los números α_i, β_j y para todo número r ó s de sumandos.

Todo subespacio lineal L del espacio euclídeo E se convierte por sí mismo en un espacio euclídeo, si se conserva en L el producto escalar que se ha introducido en E.

Es fácil indicar el método general de introducción del producto escalar en un espacio real arbitrario K. Sea e_1, e_2, . . ., e_n una base de este espacio. Elijamos dos vectores arbitrarios x, y de K y supongamos que

$$x=\xi_1e_1+\xi_2e_2+\ldots+\xi_ne_n,$$
$$y=\eta_1e_1+\eta_2e_2+\ldots+\eta_ne_n.$$

El producto escalar de vectores puede ser introducido ahora del modo siguiente, por ejemplo:

$$(x,\ y)=\xi_1\eta_1+\xi_2\eta_2+\ldots+\xi_n\eta_n.\tag{27.2}$$

El cumplimiento de todos los axiomas se comprueba sin dificultad alguna. Por consiguiente, el espacio lineal K provisto del producto escalar (27.2) es euclídeo.

Observemos que el producto escalar puede ser introducido en el espacio K con ayuda de otros métodos. Por ejemplo, en dicho espacio la siguiente expresión será también un producto escalar:

$$(x,\ y)=\alpha_1\xi_1\eta_1+\alpha_2\xi_2\eta_2+\ldots+\alpha_n\xi_n\eta_n$$

para cualesquiera números positivos fijados α_1, α_2, . . ., α_n. La multiformidad semejante no debe disturbarnos. En efecto, no hay nada asombroso en que las longitudes pueden medirse en metros y en pulgadas, los ángulos en grados y radianes, etc. Precisamente esta multiformidad permite emplear al máximo las propiedades de los espacios concretos, cuando en éstos se introduce el producto escalar.

Al introducir el producto escalar en los espacios de segmentos dirigidos, tuvimos que definirlo por separado, cuando por lo menos uno de los segmentos era igual a cero. En aquel caso se suponía que l producto escalar era nulo. Ahora el hecho dado pasa a ser una eropiedad que proviene de los axiomas (27.1). Si x es un vector arbitrario de E, entonces

$$(0, x) = (0x, x) = 0\,(x, x) = 0.$$

Por supuesto, en virtud del primer axioma de (27.1), $(x, 0) = 0$.

Un vector x de un espacio euclídeo se denomina *normalizado*, si $(x, x) = 1$. Todo vector *no nulo* y puede ser normalizado, si lo multiplicamos por cierto número λ. En efecto, por hipótesis

$$(\lambda y, \lambda y) = \lambda^2 (y, y) = 1,$$

y por eso, a título del factor de normalización, podemos tomar

$$\lambda = (y, y)^{-1/2}$$

Un sistema de vectores se llama *normalizado*, si están normalizados todos sus vectores. Según se deduce de lo dicho, todo sistema de vectores no nulos puede ser normalizado.

Una de las propiedades más importantes del producto escalar la enuncia el siguiente

TEOREMA 27.1 (*desigualdad de Cauchy—Buniakovski*). *Para cualesquiera dos vectores x, y de un espacio euclídeo es válida la desigualdad*

$$(x, y)^2 \leqslant (x, x)\,(y, y).$$

DEMOSTRACION. El teorema tiene lugar, a ciencia cierta, si $y = 0$, por lo cual convengamos en considerar que $y \neq 0$. Examinemos un vector $x - \lambda y$, donde λ es ún número real arbitrario. Tenemos

$$(x - \lambda y, x - \lambda y) = (x, x) - 2\lambda\,(x, y) + \lambda^2\,(y, y).$$

En el primer miembro de la igualdad figura un producto escalar de vectores iguales. Por esta razón el trinomio de segundo grado en el segundo miembro es no negativo, cualquiera que sea λ, en particular, para

$$\lambda = \frac{(x, y)}{(y, y)}. \tag{27.3}$$

De este modo,

$$(x, x) - 2\frac{(x, y)}{(y, y)}(x, y) + \frac{(x, y)^2}{(y, y)^2}(y, y) = (x, x) - \frac{(x, y)^2}{(y, y)} \geqslant 0,$$

de donde se hace obvia la afirmación del teorema.

Por analogía con los espacios de segmentos dirigidos, llamemos *colineales* dos vectores x, y de cualquier espacio lineal, si o bien $x = \lambda y$ o bien $y = \mu x$ para ciertos números λ, μ. En virtud de la igualdad $0 = 0x$ concluimos que los dos vectores son colineales, a ciencia cierta, si por lo menos uno de ellos es nulo. Para la comprobación del carácter colineal de los vectores resulta muy cómoda la desigualdad de Cauchy—Buniakovski. En particular, es válido el

TEOREMA 27.2. *La desigualdad de Cauchy—Buniakovski se convierte en una igualdad si, y sólo si, los vectores x, y son colineales.*

DEMOSTRACION. Supongamos que los vectores x, y son colineales y que, para concretar, $x = \lambda y$. Hallamos

$$(x, y)^2 = (\lambda y, y)^2 = \lambda^2 (y, y)^2,$$
$$(x, x)(y, y) = (\lambda y, \lambda y)(y, y) = \lambda^2 (y, y)^2.$$

La comparación de estas igualdades muestra que la suficiencia de la afirmación del teorema tiene lugar.

Supongamos ahora que para ciertos vectores x, y se verifica la igualdad:

$$(x, y)^2 = (x, x)(y, y). \tag{27.4}$$

Si $y = 0$, los vectores son colineales. En cambio, si $y \neq 0$, entonces, al tomar λ de acuerdo con (27.3) y teniendo presente (27.4), obtenemos

$$(x - \lambda y, x - \lambda y) = 0.$$

En vista del último axioma en (27.1), esto significa que $x - \lambda y = 0$, o bien $x = \lambda y$, es decir, los vectores x, y son colineales. La necesidad de la afirmación del teorema también tiene lugar.

A título de ejemplo consideremos el espacio R_n. Se le puede transformar en espacio euclídeo, si para los vectores

$$x = (\alpha_1, \alpha_2, \ldots, \alpha_n),$$
$$y = (\beta_1, \beta_2, \ldots, \beta_n)$$

el producto escalar se introduce de la manera siguiente:

$$(x, y) = \sum_{i=1}^{n} \alpha_i \beta_i. \tag{27.5}$$

Es evidente que los axiomas (27.1) se cumplen. La desigualdad de Cauchy—Buniakovski significa, en el caso dado, que

$$\left(\sum_{i=1}^{n} \alpha_i \beta_i\right)^2 \leqslant \left(\sum_{i=1}^{n} \alpha_i^2\right)\left(\sum_{i=1}^{n} \beta_i^2\right) \tag{27.6}$$

para cualesquiera números reales α_i, β_i.

Ejercicios.

1. ¿Cómo se debe introducir el producto escalar en un espacio de polinomios de una variable con coeficientes reales?

2. ¿Se hará euclídeo el espacio R_n, si el producto escalar se introduce en él de la manera siguiente:

$$(x, y) = \sum_{i=1}^{n} |\alpha_i| \, |\beta_i| ?$$

3. ¿Cuál es el sentido geométrico de la desigualdad de Cauchy—Buniakovski en el espacio de segmentos dirigidos?

4. Demuéstrese que $x = y$ cuando, y sólo cuando, $(x, d) = (y, d)$ para todo vector d.

§ 28. Ortogonalidad

La relación más importante entre los vectores de un espacio euclídeo es la ortogonalidad.

Por definición, los vectores x, y se denominan *ortogonales* si $(x, y) = 0$. En virtud del primer axioma de (27.1), la relación de ortogonalidad de dos vectores es simétrica. En el espacio de segmentos dirigidos la noción de ortogonalidad coincide, en lo principal, con la de perpendicularidad. Por esta razón, la ortogonalidad puede considerarse como una generalización de la noción de perpendicularidad para los espacios euclídeos abstractos.

Un sistema de vectores de un espacio euclídeo se llama *ortogonal* si o bien consiste en un solo vector o bien sus vectores son ortogonales dos a dos. Si un sistema ortogonal consta de vectores no nulos, se le puede normalizar. Un sistema ortogonal normalizado se denomina *ortonormalizado*.

El interés hacia los sistemas ortogonales y ortonormalizados se debe a las ventajas que ofrecen en la investigación de espacios euclídeos.

Así por ejemplo, cualquier sistema ortogonal de vectores no nulos y, por supuesto, un sistema ortonormalizado, es linealmente independiente. En efecto, supongamos que el sistema $x_1, x_2, \ldots, x_k$ es ortogonal y que $x_i \neq 0$ para todo i. Esto significa que $(x_i, x_j) = 0$ para $i \neq j$, pero $(x_i, x_j) \neq 0$ cuando $i = j$. Escribamos la ecuación

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k = 0.$$

Al multiplicarla escalarmente por cualquiera de los vectores x_i hallamos

$$\alpha_1 (x_i, x_1) + \alpha_2 (x_i, x_2) + \ldots + \alpha_k (x_i, x_k) = 0.$$

Por consiguiente,

$$\alpha_i (x_i, x_i) = 0 \tag{28.1}$$

y, naturalmente, $\alpha_i = 0$. De este modo, el sistema de los vectores $x_1, x_2, \ldots, x_k$ es linealmente independiente.

De la igualdad (28.1) obtenemos, en particular, que si la suma de los vectores ortogonales dos a dos es igual a cero, todos los vectores son nulos.

Una variedad de corolarios útiles se deduce de la suposición que cierto sistema ortonormalizado $e_1, e_2, \ldots, e_s$ puede formar una base del espacio euclídeo E. En este caso todo vector x de E ha de representarse de un modo único, en forma de una combinación lineal

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_s e_s.$$

Pero al multiplicar escalarmente la igualdad dada por e_i obtenemos la expresión explícita para los coeficientes de la descomposición según la base. A saber,

$$\alpha_i = (x, e_i). \tag{28.2}$$

Si para otro vector y tiene lugar la descomposición

$$y = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_s e_s,$$

entonces, al realizar ciertas transformaciones sencillas, hallamos que

$$(x, y) = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_s\beta_s. \tag{28.3}$$

En particular,

$$(x, x) = \alpha_1^2 + \alpha_2^2 + \ldots + \alpha_s^2. \tag{28.4}$$

Antes de continuar unas investigaciones semejantes, aclaremos si existe o no una base que se compone de vectores ortonormalizados.

Una base cuyos vectores forman un sistema ortonormalizado, se denomina *ortonormalizada*. La existencia de tal base en el espacio euclídeo la demuestra el

TEOREMA 28.1. *En todo espacio euclídeo E de dimensión finita existe una base ortonormalizada.*

DEMOSTRACIÓN. Sea $\dim E = n$. Un sistema ortonormalizado es linealmente independiente y, por ende, no puede contener más que n vectores. Supongamos que el sistema $e_1, e_2, \ldots, e_s$ contiene el número máximo de vectores ortonormalizados. Esto significa que en el espacio E no hay ni un solo vector no nulo que sea ortogonal a todos los vectores $e_1, e_2, \ldots, e_s$. Si existe un vector ortogonal a todos estos vectores, el mismo debe ser nulo.

Elijamos un vector arbitrario x de E. Si el sistema ortonormalizado $e_1, e_2, \ldots, e_s$ fuera una base, el vector x tendría que coincidir con el vector y, donde

$$y = (x, e_1)\, e_1 + (x, e_2)\, e_2 + \ldots + (x, e_s)\, e_s.$$

Examinemos, por ello, un vector $x - y$. Tenemos

$$(x-y, e_i) = \left(x - \sum_{p=1}^{s} (x, e_p)\, e_p,\ e_i\right) =$$

$$= (x, e_i) - \sum_{p=1}^{s} (x, e_p)(e_p, e_i) = (x, e_i) - (x, e_i) = 0.$$

El vector $x - y$ resulta ser ortogonal a todos los vectores $e_1, e_2, \ldots$ $\ldots, e_s$. Por consiguiente, $x - y = 0$ ó $x = y$.

Así pues, un sistema linealmente independiente $e_1, e_2, \ldots, e_s$ posee la propiedad que, en términos de sus vectores, se expresa linealmente todo vector del espacio E, es decir el sistema forma una base.

COROLARIO. *Cualquier sistema ortonormalizado de vectores $e_1, e_2, \ldots$ $\ldots e_k$ puede ser complementado hasta obtener una base ortonormalizada.*

Efectivamente, de los sistemas ortonormalizados que contienen un sistema dado tomemos el que cuenta con el máximo número de vectores. Supongamos que este sistema es $e_1, \ldots, e_k, e_{k+1}, \ldots, e_s$. Repitiendo textualmente la demostración del teorema 28.1, establecemos que el sistema nuevo es una base.

Además de los vectores ortogonales en el espacio euclídeo, consideraremos también *conjuntos ortogonales de vectores*. Dos conjuntos F y G de vectores en el espacio euclídeo E se denominan ortogonales, si todo vector de F es ortogonal a todo vector de G. La ortogonalidad de F y G se designa con el símbolo $F \perp G$.

Por supuesto, un conjunto puede componerse también de un solo vector. Si cierto vector de un conjunto es ortogonal a todo el conjunto, es, en particular, ortogonal a sí mismo. Por consiguiente, sólo puede ser nulo.

LEMA 28.1. *Para que el vector x sea ortogonal al subespacio L, es necesario y suficiente que sea ortogonal respecto de todos los vectores de una base cualquiera del subespacio L.*

DEMOSTRACION. Fijemos la base $y_1, y_2, \ldots, y_k$ del subespacio L. Si $x \perp L$, entonces x es ortogonal a todos los vectores de L y, en particular, a los vectores $y_1, y_2, \ldots, y_k$. Sea, ahora, $(x, y_i) = 0$ para todo i. Elijamos un vector arbitrario z de L y descompongámoslo según los vectores de la base. Si

$$z = \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_k y_k$$

para ciertos números $\alpha_1, \alpha_2, \ldots, \alpha_k$, entonces

$$(x, z) = (x, \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_k y_k) =$$
$$= \alpha_1 (x, y_1) + \alpha_2 (x, y_2) + \ldots + \alpha_k (x, y_k) = 0.$$

Esto significa que $x \perp L$.

COROLARIO. *Para que dos subespacios sean ortogonales, es necesario y suficiente que todo vector de cualquier base de un subespacio sea ortogonal a todos los vectores de cualquier base de otro subespacio.*

La suma K de los subespacios lineales $L_1, L_2, \ldots, L_m$ se llama *ortogonal*, si los subespacios son ortogonales dos a dos. Para designar la suma ortogonal se empleará la siguiente notación:

$$K = L_1 \oplus L_2 \oplus \ldots \oplus L_m.$$

LEMA 28.2. *Una suma ortogonal de los subespacios no nulos es siempre una suma directa.*

DEMOSTRACION. Elijamos una base ortonormalizada en cada subespacio y consideremos un sistema de vectores que representa en sí la reunión de bases de todos los subespacios. Está claro que todo vector de la suma ortogonal se expresa linealmente en términos de los vectores del sistema construido. Mas, este sistema es linealmente independiente, ya que consiste en los vectores no nulos ortogonales dos a dos. Ahora la afirmación del lema se deduce del teorema 20.1.

Supongamos que el espacio euclídeo K figura en forma de una suma ortogonal de sus subespacios $L_1, L_2, \ldots, L_m$, caso en el que la totalidad de todos estos subespacios puede considerarse como una base ortogonal generalizada. En particular, si, para cualesquiera dos vectores x, y de K, escribimos sus descomposiciones según los subespacios $L_1, L_2, \ldots, L_m$, es decir, si representamos los vectores citados en la forma

$$x = x_1 + x_2 + \ldots + x_m,$$
$$y = y_1 + y_2 + \ldots + y_m,$$

donde $x_i, y_i \in L_i$, se obtiene con facilidad que

$$(x, y) = (x_1, y_1) + (x_2, y_2) + \ldots + (x_m, y_m). \qquad (28.5)$$

La fórmula obtenida es análoga a la (28.3).

Examinemos un conjunto no vacío arbitrario de vectores F del espacio euclídeo E. La totalidad de todos los vectores, ortogonales al conjunto F, se llama *complemento ortogonal* del conjunto F y se indica con $F^{\perp}$. El complemento ortogonal es un subespacio. En efecto, si los vectores $x, y \in F^{\perp}$, entonces $x, y \perp F$. Pero, en este caso, $\alpha x + \beta y \perp F$ para cualesquiera números α, β, es decir, $\alpha x + \beta y \in F^{\perp}$.

TEOREMA 28.2. *El espacio euclídeo E es la suma ortogonal de cualquier subespacio lineal suyo L y su complemento ortogonal $L^{\perp}$, es decir,*

$$E = L \oplus L^{\perp}.$$

DEMOSTRACIÓN. Sea $\dim L = s$, $\dim L^{\perp} = m$. Elijamos una base ortonormalizada $e_1, \ldots, e_s$ del subespacio L y una base ortonormalizada $r_1, \ldots, r_m$ del subespacio $L^{\perp}$. El sistema de vectores $e_1, \ldots, e_s, r_1, \ldots, r_m$ es ortonormalizado y, por lo tanto, linealmente independiente.

Si este sistema no es la base de E, se le puede complementar hasta que se obtenga la base ortonormalizada de E. Sea e uno de los vectores complementarios. Es ortogonal a los vectores $e_1, \ldots, e_s$, por lo cual $e \perp L$, es decir, $e \in L^{\perp}$. Pero, por otra parte, el vector e es ortogonal a los vectores $r_1, \ldots, r_m$ y, por ende, $e \perp L^{\perp}$. Así pues, el vector e simultáneamente pertenece a $L^{\perp}$ y es ortogonal a $L^{\perp}$. Por consiguiente, $e = 0$, lo que demuestra la afirmación del teorema.

La descomposición de un espacio en una suma ortogonal de sus subespacios permite realizar eficazmente muchas investigaciones. Ilustrémoslo con el ejemplo siguiente.

Elijamos un espacio euclídeo E y examinemos en él un sistema fijado de vectores $x_1, x_2, \ldots, x_k$. Si el rango de este sistema es igual a la dimensión de E, entonces, evidentemente, el único vector de E, ortogonal a todos los vectores del sistema dado, será el vector nulo. Tiene lugar también el lema inverso:

LEMA 28.3. *Si en el espacio euclídeo E se ha dado cierto sistema de vectores $x_1, x_2, \ldots, x_k$ y si el único vector de E, ortogonal a estos vecto-*

res, es nulo, entonces el rango del sistema equivale a la dimensión de E.

DEMOSTRACIÓN. Designemos con L la cápsula lineal del sistema de vectores $x_1, x_2, \ldots, x_k$. Todo vector, ortogonal a los vectores dados, es ortogonal a L, es decir, pertenece al complemento ortogonal de $L^\perp$. De acuerdo con la hipótesis del lema, el subespacio $L^\perp$ sólo consta de un vector nulo. Puesto que $E = L \oplus L^\perp$, de aquí se deduce que la dimensión de L coincide con la de E. Mas, la dimensión de L es igual al rango del sistema de vectores $x_1, x_2, \ldots, x_k$. El lema queda demostrado.

Ejercicios.

1. Demuéstrese que si el producto escalar de cualesquiera dos vectores de un espacio euclídeo se expresa mediante la igualdad (28.3), la base, respecto de la cual se han tomado las coordenadas, es ortonormalizada.
2. Demuéstrese que si el producto escalar de cualquier vector de un espacio euclídeo por sí mismo se expresa mediante la igualdad (28.4), la base, respecto de la cual se han tomado las coordenadas, es ortonormalizada.
3. Demuéstrese que si dos conjuntos compuestos por un número finito de vectores son ortogonales, serán ortogonales también las cápsulas lineales construidas en estos conjuntos.
4. Demuéstrese que la intersección de dos subespacios ortogonales sólo consta de un vector nulo.
5. Demuéstrese que si un espacio euclídeo es la suma directa de sus subespacios y para cualesquiera dos vectores tiene lugar la igualdad (28.5), los subespacios son ortogonales dos a dos.
6. Demuéstrese que para cualesquiera subespacios L, M de un espacio euclídeo E se verifican las correlaciones

$$\dim L + \dim L^\perp = \dim E,$$
$$(L^\perp)^\perp = L,$$
$$(L + M)^\perp = L^\perp \cap M^\perp,$$
$$(L \cap M)^\perp = L^\perp + M^\perp.$$

§ 29. Longitudes, ángulos, distancias

Haremos extender, ahora, las nociones de longitud, ángulo y distancia a los elementos del espacio euclídeo. Realizándolo partiremos de la analogía con los espacios de segmentos dirigidos.

Se llama *longitud* $|x|$ *del vector* x en el espacio euclídeo E la magnitud

$$|x| = +(x, x)^{1/2}.$$

Todo vector tiene su longitud. De acuerdo con el último axioma (27.1), ésta es positiva para los vectores no nulos y es igual a cero, para el vector nulo. Luego, la igualdad

$$|\lambda x| = (\lambda x, \lambda x)^{1/2} = (\lambda^2 (x, x))^{1/2} = |\lambda|\, |x|$$

demuestra la posibilidad de sacar la magnitud absoluta del factor numérico λ fuera del signo de la longitud del vector. Como se ha indicado anteriormente, el vector no nulo puede ser normalizado, es decir, puede ser multiplicado por un número tal que la longitud del vector resultante se haga igual a la unidad.

Se llama *ángulo* $\{x, y\}$ *formado por los vectores no nulos* x, y del espacio euclídeo E aquel que se define por las correlaciones

$$\cos\{x, y\} = \frac{(x, y)}{|x|\,|y|}, \quad 0 \leqslant \{x, y\} \leqslant \pi.$$

Si entre los vectores x, y existe al menos uno no nulo, el ángulo formado por tales vectores se considera *indeterminado*.

La desigualdad de Cauchy—Buniakovski permite afirmar que la expresión, llamada coseno del ángulo entre vectores, no es superior, en modulo, a la unidad. Por ello el ángulo formado por cualesquiera vectores no nulos está siempre definido y, además, unívocamente. Este ángulo no se altera cuando los vectores se multiplican por cualesquiera números positivos y, conforme al teorema 27.2, es igual a 0 ó π si, y sólo si, los vectores no nulos son colineales. Todo esto concuerda plenamente con la noción de ángulo formado por segmentos dirigidos.

Elijamos dos vectores no nulos x, y. Teniendo presente la analogía con los segmentos dirigidos, consideraremos x, y como dos de los lados de cierto triángulo. Como tercer lado del triángulo resulta natural tomar el vector $x - y$. Al hacer uso de la definición de longitud de un vector y de ángulo entre los vectores, hallamos

$$\begin{aligned} |x - y|^2 &= (x - y, x - y) = (x, x) - 2(x, y) + (y, y) = \\ &= |x|^2 + |y|^2 - 2\,|x|\,|y| \cos\{x, y\}. \end{aligned} \qquad (29.1)$$

Hemos probado, pues, que en el espacio euclídeo el cuadrado de la longitud de cualquier lado de un triángulo es igual a la suma de cuadrados de las longitudes de sus otros dos lados menos el producto duplicado de las longitudes de estos lados por el coseno del ángulo entre ellos.

Si el triángulo es rectángulo, es decir, si el ángulo formado por los vectores x, y es recto, entonces, evidentemente,

$$|x - y|^2 = |x|^2 + |y|^2. \qquad (29.2)$$

Esto no es otra cosa que la expresión formal del teorema conocido de Pitágoras.

Consideremos de nuevo un triángulo rectángulo. Puesto que el coseno del ángulo formado por los vectores no sobrepasa en módulo la unidad, entonces, de (29.1) proviene que

$$|x - y|^2 \leqslant (|x| + |y|)^2,$$
$$|x - y|^2 \geqslant (|x| - |y|)^2,$$

o bien

$$|x-y| \leqslant |x|+|y|, \tag{29.3}$$
$$|x-y| \geqslant ||x|-|y||.$$

De esto modo, en el espacio euclídeo la longitud del lado de un triángulo no es superior a la suma de longitudes de otros dos lados, pero no es inferior a la diferencia entre sus longitudes.

Se llama *distancia* $\rho(x, y)$ *entre los vectores*, x, y de un espacio euclídeo la magnitud

$$\rho(x, y)=|x-y|. \tag{29.4}$$

Esta magnitud satisface tres propiedades naturales de la distancia entre los vectores (¡en la interpretación puntual!) en los espacios de segmentos dirigidos. A saber, para cualesquiera vectores x, y, z de un espacio euclídeo

$$\begin{aligned}&1)\ \rho(x, y)=\rho(y, x),\\ &2)\ \rho(x, y)>0,\ \text{si}\ x \neq y;\ \rho(x, y)=0,\ \text{si}\ x=y,\\ &3)\ \rho(x, y) \leqslant \rho(x, z)+\rho(z, y).\end{aligned} \tag{29.5}$$

Las primeras dos propiedades son obvias. La última propiedad no es otra cosa que la generalización de la conocida «desigualdad triangular». La validez de la propiedad 3 se desprende de la primera desigualdad (29.3), si x e y se sustituyen por $x-z$ e $y-z$, respectivamente.

Se llama *distancia* $\rho(A, B)$ *entre los conjuntos* A, B *de vectores* de un mismo espacio la magnitud

$$\rho(A, B)=\inf_{x \in A,\, y \in B} \rho(x, y).$$

En conclusión hemos de notar la siguiente circunstancia. Supongamos que en el espacio euclídeo E se ha fijado una base ortonormalizada $e_1, e_2, \ldots, e_s$. Para cualesquiera dos vectores x, y, dados mediante sus coordenadas

$$x=(\alpha_1, \alpha_2, \ldots, \alpha_s), \qquad y=(\beta_1, \beta_2, \ldots, \beta_s)$$

respecto de esta base, tendremos, de acuerdo con (28.3),

$$|x|=(\alpha_1^2+\alpha_2^2+\ldots+\alpha_s^2)^{1/2}.$$

Por consiguiente,

$$\cos\{x, y\}=\frac{\alpha_1\beta_1+\alpha_2\beta_2+\ldots+\alpha_s\beta_s}{(\alpha_1^2+\ldots+\alpha_s^2)^{1/2}(\beta_1^2+\ldots+\beta_s^2)^{1/2}}.$$

La analogía total con las fórmulas (25.4), (25.5) es evidente.

De este modo, las nociones introducidas de longitud, ángulo y distancia concuerdan enteramente con las nociones análogas en el espacio de segmentos dirigidos.

Ejercicios.

1. Demuéstrese que la longitud de una suma de cualquier número de vectores no es superior a la suma de longitudes de estos vectores.
2. Demuéstrese que el cuadrado de la longitud de una suma de cualquier número de vectores ortogonales es igual a la suma de cuadrados de las longitudes de dichos vectores.
3. En un espacio euclídeo de polinomios, que dependen de una variable t, hállense los ángulos del triángulo formado por los vectores 1, t^2, $1 - t^2$.
4. ¿Cuál es la distancia entre los polinomios $3t^2 + 6$ y $2t^2 + t + 1$?
5. Demuéstrese que un triángulo en el espacio euclídeo es rectángulo cuando, y sólo cuando, la longitud de un lado es igual al producto de la longitud del otro lado por el coseno del ángulo entre ellos.

§ 30. Línea oblicua, perpendicular, proyección

Antes de extender las nociones de oblicua, perpendicular y proyección a los espacios euclídeos abstractos, considerarémoslas en un espacio de segmentos dirigidos.

Sea dado un plano L. Desde un punto M tracemos una perpendicular al plano y designemos con M_L su base (fig. 30.1). Con el fin de comunicar al problema una interpretación vectorial, elijamos en el plano L un punto O y consideremos el espacio V_3 de segmentos dirigidos sujetados al punto O. El plano L forma un subespacio, por lo cual la construcción de la perpendicular bajada desde el punto M al plano L se reduce a la descomposición del vector $\overrightarrow{OM}$ del espacio en la suma

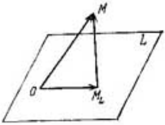


Fig. 30.1.

$$\overrightarrow{OM} = \overrightarrow{OM_L} + \overrightarrow{M_L M}, \tag{30.1}$$

donde $\overrightarrow{OM_L} \in L$ y $\overrightarrow{M_L M} \perp L$. De los razonamientos geométricos se ve que la descomposición (30.1) existe siempre y es única.

El ejemplo examinado sugiere cómo, en el caso general, debe plantearse el problema sobre una perpendicular. Supongamos que en el espacio euclídeo E se ha fijado cierto subespacio L. Tomemos un vector arbitrario f de E e investiguemos la posibilidad de su descomposición en la suma

$$f = g + h, \tag{30.2}$$

donde $g \in L$ y $h \perp L$.

Con este problema ya nos encontramos anteriormente. Efectivamente, la condición $h \perp L$ es equivalente a la condición $h \in L^{\perp}$.

De acuerdo con el teorema 28.2, el espacio euclídeo E es la suma directa de los subespacios L y $L^{\perp}$. Por esta razón, la descomposición (30.2) siempre existe y es única.

Al tomar en consideración la analogía con la descomposición (30.1), el vector g en la descomposición (30.2) se llamará *proyección del vector f sobre el subespacio L*; h, *perpendicular trazada desde el vector f al subespacio L*; el propio vector f, *línea oblicua al subespacio L*.

Se sabe que en la geometría elemental la longitud de una perpendicular nunca supera la longitud de una línea oblicua. Una situación análoga tiene lugar también en el espacio euclídeo. Los vectores g, h en la descomposición (30.2) son ortogonales. Por ello, de acuerdo con el teorema de Pitágoras,

$$|f|^2 = |g|^2 + |h|^2,$$

de donde se deduce que

$$|h| \leqslant |f|.$$

Está claro que la longitud de la perpendicular h al subespacio L equivale a la longitud de la línea oblicua f al mismo subespacio cuando, y sólo cuando, $f \perp L$.

Al problema sobre la perpendicular se le puede dar también otra interpretación. Consideraremos otra vez un vector arbitrario f de E. Este vector no pertenece forzosamente al subespacio L. Por consiguiente, se puede plantear el problema sobre la ubicación en L de un vector cuya disposición fuera más próxima a f en el sentido de la distancia introducida anteriormente.

Tomemos un vector arbitrario z de L. Al sustraerlo de ambos miembros de la igualdad (30.2) obtendremos

$$f - z = (g - z) + h.$$

Ya que el vector h es ortogonal al vector $g - z$, de acuerdo con el teorema de Pitágoras, tenemos

$$|f - z|^2 = |g - z|^2 + |h|.$$

Por esto

$$|f - z| \geqslant h,$$

con la particularidad de que la igualdad es verídica cuando, y sólo cuando, $z = g$.

Así pues, de todos los vectores del subespacio L la proyección del vector f sobre L es la más próxima al vector f. Esto es testimonio de que

$$\rho(f, L) = \rho(f, g).$$

Por analogía con los segmentos dirigidos llamemos *ángulo entre el vector f y el subespacio L* el menor de los ángulos formados por el vector f y los vectores z de L. Tomando en consideración la desigualdad de Cauchy—Buniakovski y la descomposición (30.2), hallamos

$$\cos\{f, z\} = \frac{(f, z)}{|f|\,|z|} = \frac{(g+h, z)}{|f|\,|z|} = \frac{(g, z)}{|f|\,|z|} \leqslant \frac{|g|}{|f|}.$$

Es evidente que esta desigualdad se convierte en una igualdad cuando, y sólo cuando, los vectores z y g forman un ángulo nulo.

De este modo, el ángulo entre el vector f y el subespacio L coincide con el ángulo formado por el vector f y la proyección de éste sobre el subespacio L.

Las propiedades citadas que poseen la perpendicular y la proyección reflejan el aspecto geométrico de estas nociones. Ahora considerémoslas desde el punto de vista algebraico. Con el subespacio L fijado, todo vector f del espacio euclídeo E define unívocamente respecto de L sus dos componentes. Por consiguiente, se puede considerar que la descomposición (30.2) prefija dos funciones

$$g = \mathrm{pr}_L f,$$
$$h = \mathrm{ort}_L f.$$

Como «argumento» de la función puede figurar cualquier vector de E; como «valor» de la función $\mathrm{pr}_L f$ sirve un vector de L; como «valor» de la función $\mathrm{ort}_L f$, un vector de $L^{\perp}$.

En vista de la correlación $(L^{\perp})^{\perp} = L$, la perpendicular y la proyección están ligadas mediante las igualdades

$$\begin{aligned} \mathrm{pr}_L f &= \mathrm{ort}_{L^{\perp}} f, \\ \mathrm{ort}_L f &= \mathrm{pr}_{L^{\perp}} f. \end{aligned} \tag{30.3}$$

Por ello, el estudio de estas funciones siempre se reduce, de hecho, al estudio de una de ellas.

Tomemos dos vectores arbitrarios x, y de E. Conforme a la descomposición (30.2), se tiene

$$\begin{aligned} x &= \mathrm{pr}_L x + \mathrm{ort}_L x, \\ y &= \mathrm{pr}_L y + \mathrm{ort}_L y. \end{aligned} \tag{30.4}$$

Al sumar estas igualdades término a término y al multiplicar la primera de ellas por un número real artibrario λ, obtenemos

$$x + y = (\mathrm{pr}_L x + \mathrm{pr}_L y) + (\mathrm{ort}_L x + \mathrm{ort}_L y),$$
$$\lambda x = (\lambda\, \mathrm{pr}_L x) + (\lambda\, \mathrm{ort}_L x).$$

Por comprobación inmediata nos convencemos de que los vectores dentro del primer paréntesis pertenecen a L, mientras que los encerrados dentro del segundo paréntesis son perpendiculares a L. Debido a la unicidad de las descomposiciones del tipo (30.2) esto significa la validez de las siguientes correlaciones

$$\begin{aligned} \mathrm{pr}_L(x+y) &= \mathrm{pr}_L x + \mathrm{pr}_L y, \\ \mathrm{pr}_L \lambda x &= \lambda\, \mathrm{pr}_L x \end{aligned} \tag{30.5}$$

para la función $\mathrm{pr}_L y$, por supuesto, de las correlaciones análogas

$$\begin{aligned} \mathrm{ort}_L(x+y) &= \mathrm{ort}_L x + \mathrm{ort}_L y, \\ \mathrm{ort}_L(\lambda x) &= \lambda\, \mathrm{ort}_L x \end{aligned} \tag{30.6}$$

para la función ort_L. Aquí tiene lugar una completa coincidencia de las fórmulas (25.9) y (30.5).

Observemos que $\mathrm{ort}_L z = 0$ para todo vector z de L. Por eso, de la primera igualdad (30.6) se deduce que

$$\mathrm{ort}_L(x+z) = \mathrm{ort}_L(x).$$

Por consiguiente, el valor de la función ort_L no se altera, si al argumento se adiciona cualquier vector del subespacio L. En particular, si tomamos $z = -\mathrm{pr}_L x$, entonces, teniendo en cuenta (30.4), obtenemos

$$\mathrm{ort}_L(\mathrm{ort}_L x) = \mathrm{ort}_L x. \tag{30.7}$$

Una correlación análoga tiene lugar también para la proyección. A saber,

$$\mathrm{pr}_L(\mathrm{pr}_L x) = \mathrm{pr}_L x. \tag{30.8}$$

Supongamos que el subespacio L es la suma ortogonal de los subespacios L_1 y L_2. Elijamos un vector arbitrario x de E y representémoslo en forma de la suma

$$x = (\mathrm{pr}_{L_1} x + \mathrm{pr}_{L_2} x) + (x - \mathrm{pr}_{L_1} x - \mathrm{pr}_{L_2} x).$$

El vector dentro del primer paréntesis pertenece, evidentemente, al subespacio $L_1 \oplus L_2$. El vector dentro del segundo paréntesis es ortogonal a $L_1 \oplus L_2$, de lo que es fácil convencerse transformándolo con ayuda de las correlaciones (30.4) del modo siguiente

$$x - \mathrm{pr}_{L_1} x - \mathrm{pr}_{L_2} x = \mathrm{ort}_{L_1} x - \mathrm{pr}_{L_2} x = \mathrm{ort}_{L_2} x - \mathrm{pr}_{L_1} x. \tag{30.9}$$

Concluimos, a partir de lo dicho, que

$$\mathrm{pr}_{L_1 \oplus L_2} x = \mathrm{pr}_{L_1} x + \mathrm{pr}_{L_2} x.$$

Una perpendicular bajada desde el vector x sobre el subespacio $L_1 \oplus L_2$ es igual a una de las expresiones (30.9). Si, en particular, $x \perp L_1$, entonces

$$\mathrm{ort}_{L_1 \oplus L_2} x = \mathrm{ort}_{L_2} x. \tag{30.10}$$

Ejercicios.

1. ¿Tendrá lugar en un espacio euclídeo el análogo del teorema sobre los tres perpendiculares?
2. Demuéstrese que la suma de dos ángulos, formados por el vector f y los subespacios L y $L^{\perp}$, es igual a $\pi/2$.
3. Hállense la perpendicular y la proyección del vector f sobre los subespacios triviales.
4. Demuéstrese que si para los espacios fijados L_1, L_2 y para cualquier vector x se verifica la igualdad

$$\mathrm{pr}_{L_1+L_2}\, x = \mathrm{pr}_{L_1}\, x + \mathrm{pr}_{L_2}\, x,$$

la suma $L_1 + L_2$ será ortogonal.

5. Demuéstrese que si los subespacios $L_1, L_2, \ldots, L_m$ son ortogonales dos a dos, entonces para todo vector x de E

$$|x|^2 \geqslant \sum_{i=1}^{m} |\mathrm{pr}_{L_i}\, x|^2.$$

§ 31. Isomorfismo euclídeo

En el transcurso de nuestras investigaciones ya hemos observado, más de una vez, que las propiedades de un espacio euclídeo abstracto coinciden con las de los espacios de segmentos dirigidos. Se podría continuar extendiendo los hechos y teoremas de la geometría elemental al espacio euclídeo. Sin embargo, no hay necesidad en esto.

Introduzcamos la noción de isomorfismo euclídeo. Diremos que los espacios euclídeos E y E' son *isomorfos según Euclides*, si son isomorfos como espacios lineales reales y, además, para todo par de vectores x, y de E y los vectores correspondientes x', y' de E' se verifica la igualdad

$$(x, y) = (x', y').$$

TEOREMA 31.1. *Para que dos espacios euclídeos sean isomorfos según Euclides, es necesario y suficiente que sean iguales sus dimensiones.*

DEMOSTRACION. Si dos espacios euclídeos E y E' son isomorfos según Euclides, serán isomorfos también como espacios reales lineales. Pero, los espacios lineales de tal índole son de igual dimensión.

Consideraremos ahora dos espacios euclídeos E y E' de una misma dimensión n. Sea $e_1, e_2, \ldots, e_n$ la base ortonormalizada en E y sea $e'_1, e'_2, \ldots, e'_n$ la base otronormalizada en E'. A todo vector

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n$$

del espacio E ponemos en correspondencia el vector

$$x' = \alpha_1 e'_1 + \alpha_2 e'_2 + \ldots + \alpha_n e'_n$$

del espacio E'. De acuerdo con lo demostrado anteriormente, esta correspondencia representa un isomorfismo. Tomemos ahora otro par de vectores correspondientes de E y E'

$$y = \beta_1 e_1 + \beta_2 e_2 + \ldots + \beta_n e_n,$$
$$y' = \beta_1 e_1' + \beta_2 e_2' + \ldots + \beta_n e_n'.$$

Según (28.3) tenemos

$$(x, y) = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_n\beta_n = (x', y').$$

El teorema queda demostrado.

Nos interesan siempre sólo aquellas propiedades de los espacios lineales que son consecuencias de las operaciones fundamentales que actúan en los espacios. Desde este punto de vista los espacios isomorfos según Euclides tienen propiedades iguales. Por ello, cualquier teorema geométrico, demostrado para el espacio V_3, será verídico también en cualquier subespacio tridimensional del espacio euclídeo. Por consiguiente, será válido también en todo espacio euclídeo. Por supuesto, como espacio euclídeo tipo puede servir el espacio aritmético $\mathbf{R}_n$ con un producto escalar introducido conforme a (27.2).

Ejercicios.

1. Constrúyase un isomorfismo euclídeo entre los espacios V_2 y $\mathbf{R}_2$.
2. Demuéstrese que en los espacios isomorfos según Euclides un sistema ortonormalizado de vectores pasa a ser de nuevo un sistema ortonormalizado.
3. Demuéstrese que en los espacios isomorfos según Euclides los ángulos entre los pares de correspondientes vectores son iguales.
4. Demuéstrese que en los espacios isomorfos según Euclides una perpendicular y una proyección pasan a ser una perpendicular y una proyección, respectivamente.

§ 32. Espacio unitario

Los conceptos métricos fundamentales han sido extendidos sólo a los espacios lineales reales. Resultados análogos tienen lugar también en un espacio lineal complejo.

El espacio lineal complejo U se denomina *unitario*, si a todo par de vectores x, y de U se le ha puesto en correspondencia un número complejo (x, y), llamado *producto escalar*, con la particularidad de que se cumplen los siguientes axiomas:

1) $(x, y) = \overline{(y, x)}$,
2) $(\lambda x, y) = \lambda (x, y)$,
3) $(x + y, z) = (x, z) + (y, z)$,
4) $(x, x) > 0$ cuando $x \neq 0$; $(0, 0) = 0$

para unos vectores arbitrarios x, y, z de U y un número complejo arbitrario λ.

La raya en el primer axioma es un signo de conjugación compleja. Esta única diferencia de los axiomas del espacio euclídeo no lleva consigo distinciones profundas, sin embargo no se debe olvidarla. Así por ejemplo, si en un espacio euclídeo tiene lugar la igualdad $(x, \lambda y) = \lambda (x, y)$, en el unitario se verificará la igualdad $(x, \lambda y) = \bar{\lambda} (x, y)$.

En el espacio unitario U se pueden introducir algunos conceptos métricos. Al igual que en el caso real, llamemos longitud del vector la magnitud

$$|x| = +(x, x)^{1/2}.$$

Todo vector no nulo tiene una longitud positiva, la longitud de un vector nulo es igual a cero. Para cualquier λ complejo se verifica la correlación

$$|\lambda x| = |\lambda| \cdot |x|.$$

Es válida también la desigualdad de Cauchy—Buniakovski

$$|(x, y)|^2 \leqslant (x, x)(y, y).$$

La demostración se lleva a cabo siguiendo el mismo esquema que en el caso real.

En el espacio unitario el concepto de ángulo entre los vectores, por regla general, no se introduce. Se considera solamente el caso en que los vectores x e y son ortogonales. Al igual que en el caso real, se entiende en tal circunstancia el cumplimiento de la igualdad

$$(x, y) = 0.$$

Es evidente que $(y, x) = \overline{(x, y)} = 0$.

En esencia, toda la teoría del espacio euclídeo, examinada más arriba, se aplica a un espacio unitario sin que cambien las definiciones y los esquemas generales de las demostraciones.

A título del espacio unitario puede servir el espacio aritmético C_n, siempre que para los vectores

$$x = (\alpha_1, \alpha_2, \ldots, \alpha_n),$$
$$y = (\beta_1, \beta_2, \ldots, \beta_n)$$

el producto escalar se introduzca de la manera siguiente:

$$(x, y) = \sum_{i=1}^{n} \alpha_i \bar{\beta}_i. \tag{32.1}$$

Con el ejemplo de este espacio se muestra fácilmente la significacion de la conjugación compleja en el primer axioma. Si en el espacio C_n introdujéramos el producto escalar según la fórmula (27.2), entonces, en el espacio C_3, por ejemplo, para el vector

$$x = (3, 4, 5i)$$

tendríamos

$$(x, x) = 9 + 16 + 25i^2 = 0.$$

El cuarto axioma, que es muy importante, resultaría incumplido.

Ejercicios.

1. Compárense el espacio euclídeo R_1 y el espacio unitario C_1.
2. Escríbase la desigualdad de Cauchy—Buniakovski en el espacio C_n.
3. Si en un espacio complejo el producto escalar se introduce de acuerdo con los axiomas (27.1), ¿podrá cumplirse en tal espacio la desigualdad de Cauchy—Buniakovski?
4. Si en un espacio complejo el producto escalar se introduce de acuerdo con los axiomas (27.1), ¿podrá existir en tal espacio una base ortogonal?

§ 33. Dependencia lineal y sistemas ortonormalizados

Ya hemos visto en el § 22 que la independencia lineal de un sistema de vectores de una base puede perturbarse como resultado de *pequeñas* alteraciones en los mismos vectores. Este fenómeno conduce a complicaciones bastante grandes cuando el concepto de base se emplea en la resolución de los problemas prácticos. No obstante, conviene subrayar que no todas las bases poseen esta desfavorable propiedad. En particular, *no* la *tiene* ninguna base ortonormalizada.

Supongamos que en un espacio euclídeo o en un espacio unitario se ha elegido una base ortonormalizada arbitraria $e_1, e_2, \ldots, e_n$. Si para cierto vector b tiene lugar la descomposición

$$b = \sum_{i=1}^{n} \alpha_i e_i,$$

entonces, según (28.4)

$$|b|^2 = \sum_{i=1}^{n} |\alpha_i|^2. \tag{33.1}$$

Examinaremos ahora el sistema de vectores $e_1 + \varepsilon_1, e_2 + \varepsilon_2, \ldots$ $\ldots, e_n + \varepsilon_n$ y supondremos que el sistema es linealmente dependiente. Esto significa que existen tales números $\beta_1, \beta_2, \ldots, \beta_n$, no nulos a la vez, que

$$\sum_{i=1}^{n} \beta_i (e_i + \varepsilon_i) = 0.$$

De aquí se desprende que

$$\sum_{i=1}^{n} \beta_i e_i = -\sum_{i=1}^{n} \beta_i \varepsilon_i.$$

Al hacer uso de la igualdad (33.1) y desigualdad (27.6), tenemos

$$\sum_{i=1}^{n} |\beta_i|^2 = \Big|\sum_{i=1}^{n} \beta_i e_i\Big|^2 = \Big|\sum_{i=1}^{n} \beta_i \varepsilon_i\Big|^2 \leqslant$$

$$\leqslant \Big(\sum_{i=1}^{n} |\beta_i|\, |\varepsilon_i|\Big)^2 \leqslant \Big(\sum_{i=1}^{n} |\beta_i|^2\Big)\Big(\sum_{i=1}^{n} |\varepsilon_i|^2\Big).$$

Comparando los miembros primero y segundo de las correlaciones obtenidas, llegamos a la conclusión que

$$\sum_{i=1}^{n} |\varepsilon_i|^2 \geqslant 1.$$

De este modo, la desigualdad obtenida significa que al cumplirse la condición

$$\sum_{i=1}^{n} |\varepsilon_i|^2 < 1, \tag{33.2}$$

el sistema de vectores

$$e_1 + \varepsilon_1,\ e_2 + \varepsilon_2,\ \ldots,\ e_n + \varepsilon_n$$

será, a ciencia cierta, linealmente independiente.

La peculiaridad de los sistemas ortonormalizados que acabamos de indicar ha determinado el amplio uso de estos sistemas en la construcción de los más diversos algoritmos de cómputo relacionados con la descomposición según la base.

Ejercicios.

1. Sea $e_1, e_2, \ldots, e_n$ una base ortogonal de un espacio euclídeo. Demuéstrese que el sistema de vectores $x_1, x_2, \ldots, x_n$ es linealmente independiente, si

$$\sum_{i=1}^{n} \cos\{e_i, x_i\} > n - \frac{1}{2}.$$

2. Supongamos que los vectores $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ para $i = 1, 2, \ldots, n$ están definidos por medio de sus coordenadas en una base arbitraria. Demuéstrese que si

$$|x_{ii}|^2 > n \sum_{k \neq i} |x_{ik}|^2$$

para cualquier valor de i, el sistema $x_1, x_2, \ldots, x_n$ es linealmente independiente.

CAPITULO 4 VOLUMEN DEL SISTEMA DE VECTORES EN UN ESPACIO LINEAL

§ 34. Productos vectorial y mixto

De nuevo comencemos nuestras investigaciones por un espacio de segmentos dirigidos. Suponemos, como siempre, que se ha fijado cierto sistema rectangular cartesiano de coordenadas cuyo origen se ubica en O y la base es i, j, k.

Tres vectores se denominan *terna*, si se ha indicado cuál de estos vectores es el primero, cuál es el segundo y cuál el tercero. Al inscribir una terna de vectores, dispondremos los propios vectores de izquierda a derecha en el orden de seguimiento.

Una terna de vectores no coplanares a, b, c se llama *derecha* (*izquierda*), si dichos vectores se disponen igual que los dedos pulgar, índice no doblado y del corazón, respectivamente, de la mano *derecha* (*izquierda*).

De cualesquiera tres vectores no coplanares a, b, c se pueden componer las siguientes seis ternas

$$abc,\ bca,\ cab,\ bac,\ acb,\ cba.$$

Las primeras tres ternas son de la misma denominación que la terna abc, las demás ternas son de denominación contraria. Cabe notar que si en una terna se cambian de lugares cualesquiera dos vectores, dicha terna cambiará su denominación.

Un sistema de coordenadas, afín o cartesiano, se llama *derecho* (*izquierdo*), si los vectores básicos forman una terna derecha (izquierda). Hasta el presente momento nuestras investigaciones no dependían de la denominación que tenía la base del sistema de coordenadas. Ahora en las investigaciones aparecerán ciertas diferencias. Es por eso que para concretar consideraremos en lo sucesivo solamente sistemas de coordenadas derechos.

Sean dados dos vectores no colineales a, b. Pondremos en correspondencia a estos vectores el tercer vector c que satisfaga las siguientes condiciones:

1) el vector c es ortogonal a cada uno de los vectores a, b,
2) la terna abc es derecha,

3) la longitud del vector c es igual numéricamente al área S del paralelogramo construido en los vectores a, b, reducidos a un origen común. Si los vectores a, b son colineales, a tal par de vectores le pondremos en correspondencia el vector nulo.

La correspondencia construida es una operación algebráica en el espacio V_3. Se llama *multiplicación vectorial de los vectores* a, b y se indica con el símbolo

$$c = [a, b].$$

Consideremos los vectores básicos i, j, k. Por definición del producto vectorial tendremos

$$\begin{array}{lll} [i, i] = 0, & [i, j] = k, & [i, k] = -j, \\ [j, i] = -k, & [j, j] = 0, & [j, k] = i, \\ [k, i] = j, & [k, j] = -i, & [k, k] = 0. \end{array} \qquad (34.1)$$

De estas correlaciones se deduce, en particular, que la operación del producto vectorial no es conmutativa.

Toda terna abc de vectores no coplanares aplicados a un punto común O define cierto paralelepípedo. El punto O es uno de los vértices y los vectores a, b, c son las aristas. Designaremos el volumen de este paralelepípedo mediante el símbolo $V(a, b, c)$, subrayando de esta manera, que el volumen depende de los vectores a, b, c. Si la terna a, b, c es coplanar, el volumen se considerará igual a cero. Atribuyamos al volumen el signo más, si la terna no coplanar abc es derecha y el signo menos, si es izquierda. Este nuevo concepto, definido de tal modo, se denominará *volumen orientado* del paralelepípedo y se indicará con el símbolo $V\pm(a, b, c)$.

El volumen y el volumen orientado pueden considerarse como ciertas funciones numéricas de tres argumentos vectoriales que toman determinados valores reales para cada terna de vectores abc. El volumen es siempre no negativo, mientras que el volumen orientado puede tener un signo cualquiera. En lo que sigue se pondrá de manifiesto el sentido bien determinado que radica en la distinción de estas nociones.

Sean dados tres vectores arbitrarios a, b, c. Si a se multiplica vectorialmente a la derecha por b y luego el vector $[a, b]$ se multiplica escalarmente por c, el número que se obtiene $([a, b], c)$ se llama *producto mixto de los vectores* a, b, c.

TEOREMA 34.1. *Un producto mixto* $([a, b], c)$ *es igual al volumen orientado del paralelepípedo construido sobre los vectores* a, b, c *reducidos a un origen común.*

DEMOSTRACIÓN. Sin restringir la generalidad podemos considerar que los vectores a, b no son colineales, puesto que en el caso contrario $[a, b] = 0$ y la afirmación del teorema resulta evidente. Sea, como antes, S el área del paralelogramo construido sobre los vectores a, b.

Conforme a (26.4) se tiene

$$([a, b], c) = |[a, b]| \{\mathrm{pr}_{[a,b]}\, c\} = S\, \{\mathrm{pr}_{[a,b]}\, c\}. \qquad (34.2)$$

Supongamos que los vectores a, b, c son no coplanares. Entonces, $\{\mathrm{pr}_{[a,b]}\, c\}$ es igual, salvo el signo, a la altura h del paralelepípedo construido sobre los vectores a, b, c trasladados al origen común, a condición de que de base sirve el paralelogramo construido sobre los vectores a, b (fig. 34.1). De este modo, el segundo miembro de (34.2) es igual, salvo el signo, al volumen del paralelepípedo construido sobre los vectores a, b, c.

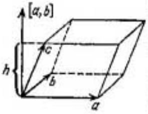


Fig. 34.1.

Evidentemente, $\{\mathrm{pr}_{[a,b]}\, c\} = +h$, si los vectores $[a, b]$ y c se disponen por un lado respecto al plano definido por los vectores a, b. Mas, en este caso, la terna abc también es derecha. En el caso contrario $\{\mathrm{pr}_{[a,b]}\, c\} = -h$. Si los vectores abc son coplanares, entonces c se dispone en el plano definido por los vectores a, b por lo cual $\{\mathrm{pr}_{[a,a]}\, c\} = 0$. El teorema queda demostrado.

COROLARIO. *Para cualesquiera tres vectores a, b, c se verifica la correlación*

$$([a, b], c) = (a, [b, c]). \qquad (34.3)$$

Efectivamente, de la propiedad de simetría del producto escalar se desprende que $(a, [b, c]) = ([b, c], a)$, por lo cual basta señalar que $([a, b], c) = ([b, c], a)$. Mas, la última igualdad es evidente, puesto que las ternas abc y bca son de una misma denominación y a ellas corresponde un mismo paralelepípedo.

La correlación (34.3) permite realizar con gran eficacia las investigaciones algebraicas. Demostremos primeramente que para cualesquiera vectores a, b, c y para todo número real α tienen lugar las siguientes propiedades de la multiplicación vectorial:

1) $[a, b] = -[b, a]$,
2) $[\alpha a, b] = \alpha [a, b]$,
3) $[a + b, c] = [a, c] + [b, c]$,
4) $[a, a] = 0$.

La propiedad 4 proviene de forma evidente de la definición. Con el fin de demostrar las propiedades restantes, haremos uso del hecho que los vectores x e y son iguales entre sí cuando, y sólo cuando,

$$(x, d) = (y, d)$$

para todo vector d.

Sea d un vector arbitrario. Las ternas abd y bad son de diferentes denominaciones. Por consiguiente, en virtud del teorema 34.1 y las propiedades del producto escalar, concluimos que

$$([a, b], d) = -([b, a], d) = (-[b, a], d).$$

Como d es un vector arbitrario, esto es un indicio de que $[a, b] = -[b, a]$ y la primera propiedad queda demostrada.

Para demostrar las propiedades segunda y tercera procedemos de manera análoga, tomando en consideración, además, la correlación (34.3). Tenemos

$$([\alpha a, b], d) = (\alpha a, [b, d]) = \alpha (a, [b, d]) = \\ = \alpha ([ab], d) = (\alpha [a, b], d),$$

lo que significa la validez de la propiedad 2. Luego,

$$([a + b, c], d) = (a + b, [c, d]) = (a, [c, d]) + (b, [c, d]) = \\ = ([a, c], d) + ([b, c], d) = ([a, c] + [b, c], d)$$

y la propiedad 3 resulta también lícita. Respecto del segundo factor tienen lugar las igualdades correspondientes:

$$[a, \alpha b] = -[\alpha b, a] = -\alpha [b, a] = \alpha [a, b],$$
$$[a, b + c] = -[b + c, a] = -[b, a] - [c, a] = [a, b] + [a, c].$$

Ahora podemos investigar las propiedades algebraicas de un volumen orientado considerándolo como una función definida en las ternas de vectores. Sea, por ejemplo, el vector a una combinación lineal de ciertos vectores a', a''. En este caso

$$([\alpha a' + \beta a'', b], c) = (\alpha [a', b] + \beta [a'', b], c) = \\ = \alpha ([a', b], c) + \beta ([a'', b], c).$$

Por consiguiente,

$$V^{\pm} (\alpha a' + \beta a'', b, c) = \alpha V^{\pm} (a', b, c) + \beta V^{\pm} (a'', b, c)$$

para cualesquiera vectores a', a'' y números reales α, β.

Cuando dos argumentos se cambian de lugar, el volumen orientado sólo cambia de signo, a consecuencia de lo cual la propiedad análoga referente a la combinación lineal es verídica para cada argumento. Teniendo presente precisamente esta propiedad, diremos que el volumen orientado representa en sí *una función lineal* respecto de todo argumento.

Si los vectores abc son linealmente dependientes, serán coplanares y, por ende, el volumen orientado en el caso dado es igual a cero. Luego, al tomar en consideración las correlaciones (34.1) hallamos que

$$V^{\pm} (i, j, k) = ([i, j], k) = (k, k) = 1.$$

Así pues, podemos concluir que el volumen orientado posee, en su calidad de función, las siguientes propiedades:

A) el volumen orientado es una función lineal respecto de todo argumento,

B) el volumen orientado es igual a cero en todos los sistemas linealmente dependientes, (34.4)

C) el volumen orientado es igual a la unidad al menos en un sistema ortonormalizado fijado de vectores.

Por supuesto, estamos lejos de haber enunciado todas las propiedades del volumen orientado. Las propiedades destacadas en (34.4), como también las otras, pueden establecerse con facilidad, si se conoce la expresión explícita de los productos vectorial y mixto en términos de las coordenadas de los vectores a, b, c.

TEOREMA 34.2. *Si los vectores a, b están dados mediante sus coordenadas rectangulares cartesianas*

$$a = (x_1,\ y_1,\ z_1),$$
$$b = (x_2,\ y_2,\ z_2),$$

el producto vectorial tendrá las coordenadas siguientes:

$$[a,\ b] = (y_1z_2 - y_2z_1,\ z_1x_2 - z_2x_1,\ x_1y_2 - x_2y_1). \qquad (34.5)$$

DEMOSTRACION. Teniendo en cuenta que las coordenadas de los vectores prefijados determinan las descomposiciones

$$a = x_1i + y_1j + z'k,$$
$$b = x_2i + y_2j + z_2k,$$

y apoyándonos en las propiedades algebraicas del producto vectorial, hallamos

$$\begin{aligned}[a,\ b] = x_1x_2\,[i,\ i] + x_1y_2\,[i,\ j] + x_1z_2\,[i,\ k] +& \\ + y_1x_2\,[j,\ i] + y_1y_2\,[j,\ j] + y_1z_2\,[j,\ k] +& \\ + z_1x_2\,[k,\ i] + z_1y_2\,[k,\ j] + z_1z_2\,[k,\ k].&\end{aligned}$$

La validez de la afirmación del teorema se desprende ahora de la correlación (34.1).

COROLARIO. *Si el vector c está dado también mediante las coordenadas x_3, y_3, z_3 en el mismo sistema cartesiano, entonces*

$$\begin{aligned}([a,\ b],\ c) = x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 -& \\ - x_1y_3z_2 - x_2y_1z_3 - x_3y_2z_1.&\end{aligned} \qquad (34.6)$$

La introducción del volumen orientado y la investigación de sus propiedades algebraicas permiten llegar a deducciones de importancia referentes a la longitud, el área y el volumen.

Observemos que las bases derechas e izquierdas determinan la partición del conjunto de todas las bases del espacio en dos clases. La propia denominación «derechas e izquierdas» no lleva ningún sentido profundo, sino que está relacionada con un método cómodo para destinguir la clase a que pertenece tal o cual base. A estas dos clases está ligado, en esencia, el propio concepto de volumen orientado.

Con hechos de tal índole ya nos hemos encontrado. Todas las bases en una recta se pueden dividir también en dos clases, reuniendo en una clase los vectores dirigidos hacia un mismo lado. En este caso resulta que la magnitud de un segmento dirigido es un análogo completo del volumen orientado, siempre que ambos conceptos se consideren como funciones de los sistemas de vectores. La propiedad **A** tiene lugar de acuerdo con las correlaciones (9.8). La propiedad **B** es válida, puesto que la magnitud del segmento nulo es igual a cero. El cumplimiento de la propiedad **C** es obvio.

Una investigación análoga independiente se podría realizar también en el caso de un plano. Sin embargo, resulta más ventajoso hacer uso de los resultados ya obtenidos anteriormente. Fijemos un sistema rectangular cartesiano de coordenadas Oxy. Completémoslo hasta obtener el sistema derecho de coordenadas $Oxyz$ en el espacio. Prestemos atención a que según sea la disposición de los ejes Ox y Oy, el eje Oz pueda tener una de las dos direcciones posibles. Esto también determina la partición del conjunto de las bases de un plano en dos clases. El área orientada $S^{\pm}(a, b)$ del paralelogramo, construido sobre los vectores a, b en el plano Oxy, puede ser definida, por ejemplo, mediante la igualdad $S^{\pm}(a, b) = V^{\pm}(a, b, k)$. Naturalmente, las propiedades **A**, **B**, **C** subsisten en este caso también.

De este modo, atribuyendo a las longitudes, las áreas y los volúmenes ciertos signos y considerando dichos conceptos como funciones definidas en los sistemas de vectores, podemos conseguir que todas estas funciones tengan las mismas propiedades algebraicas **A**, **B**, **C** de (34.4).

Ejercicios.

1. Demuéstrese que los vectores a, b, c son coplanares cuando, y sólo cuando, su producto mixto es igual a cero.
2. Demuéstrese que para cualesquiera tres vectores a, b, c se verifica la correlación

$$[a, [b, c]] = (a, c)\, b - (a, b)\, c.$$

3. Demuéstrese que la multiplicación vectorial no es una operación asociativa.
4. Hállese la expresión del área orientada de un paralelogramo en términos de las coordenadas cartesianas de vectores en un plano.
5. ¿Cambiarán las fórmulas (34.5), (34.6), si el sistema de coordenadas, respecto del cual están dados los vectores, es izquierdo?

§ 35. Volumen y volumen orientado del sistema de vectores

En los espacios lineales de segmentos dirigidos el área y el volumen son las nociones que se derivan de la longitud del segmento. El concepto de longitud ya lo hemos extendido al espacio euclídeo abstracto. Ahora examinemos un problema análogo en relación al área y al volumen.

Supongamos que en un plano están dados dos vectores no colineales x_1, x_2. Construyamos sobre estos vectores un paralelogramo, tomando por base el vector x_1 (fig. 35.1). Empleando el extremo del vector x_2 como punto de partida, tracemos la perpendicular h a la base. El área $S(x_1, x_2)$ del paralelogramo se expresará mediante la fórmula

$$S(x_1, x_2) = |x_1|\,|h|. \tag{35.1}$$

Designaremos con L_0 el subespacio nulo y con L_1, la cápsula lineal construida sobre el vector x_1. Puesto que

$$|x_1| = |\mathrm{ort}_{L_0} x_1|,$$

la fórmula (35.1) puede ser escrita en la forma siguiente:

$$S(x_1, x_2) = |\mathrm{ort}_{L_0} x_1|\,|\mathrm{ort}_{L_1} x_2|. \tag{35.2}$$

Luego, tomemos en el espacio tres vectores no coplanares x_1, x_2, x_3. Construyamos un paralelepípedo sobre estos vectores, tomando por

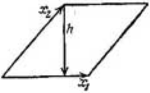


Fig. 35.1.

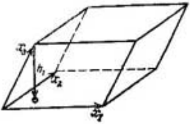


Fig. 35.2.

base el paralelogramo formado por los vectores x_1, x_2 (fig. 35.2). Desde el extremo del vector x_3 tracemos sobre la base la perpendicular h_1. El volumen $V(x_1, x_2, x_3)$ del paralelepípedo se expresará mediante la fórmula

$$V(x_1, x_2, x_3) = S(x_1, x_2)\,|h_1|.$$

Al designar con L_2 la cápsula lineal construida sobre los vectores x_1, x_2, tendremos, conforme a (35.2),

$$V(x_1, x_2, x_3) = |\mathrm{ort}_{L_0} x_1|\,|\mathrm{ort}_{L_1} x_2|\,|\mathrm{ort}_{L_2} x_3|.$$

De este modo, la longitud del vector, el área del paralelogramo y el volumen del paralelepípedo en los espacios lineales V_1, V_2, V_3 se expresan mediante las fórmulas en las cuales no es difícil descubrir cierta regularidad:

$$\begin{aligned} |x_1| &= |\mathrm{ort}_{L_0} x_1|, \\ S(x_1, x_2) &= |\mathrm{ort}_{L_0} x_1|\,|\mathrm{ort}_{L_1} x_2|, \\ V(x_1, x_2, x_3) &= |\mathrm{ort}_{L_0} x_1|\,|\mathrm{ort}_{L_1} x_2|\,|\mathrm{ort}_{L_2} x_3|. \end{aligned} \tag{35.3}$$

En particular, en todo caso el número de factores coincide con la dimensión del espacio.

Estas fórmulas dictan cómo se debe introducir el concepto de volumen en un espacio euclídeo E_n de dimensión n. Supongamos que en E_n se ha dado un sistema arbitrario de vectores $x_1, x_2, \ldots, x_n$. Designemos con L_0 el subespacio nulo y con L_i, la cápsula lineal formada por los vectores $x_1, \ldots, x_i$. Entonces, por analogía con los espacios de segmentos dirigidos diremos que:

Se llama *volumen* $V(x_1, x_2, \ldots, x_n)$ *del sistema de vectores* $x_1, x_2, \ldots, x_n$ del espacio euclídeo E_n el valor que toma en este sistema una función real que depende de n argumentos vectoriales de E_n y está definida del modo siguiente:

$$V(x_1, x_2, \ldots, x_n) = \prod_{i=0}^{n-1} |\mathrm{ort}_{L_i} x_{i+1}|. \qquad (35.4)$$

Naturalmente, por ahora no podemos afirmar que el volumen de un sistema de vectores posee todas las propiedades que son propias precisamente al volumen para cualquier valor de n. Pero, para los espacios euclídeos de dimensiones 1, 2, 3, respectivamente, en virtud del isomorfismo euclídeo y las correlaciones (35.3), el volumen tiene, a ciencia cierta, las mismas propiedades que la longitud de un segmento, el área de un paralelogramo y el volumen de un paralelepípedo, respectivamente.

Veamos ahora un acceso algo más diferente al concepto de volumen de un sistema de vectores del espacio euclídeo E_n. Como ya se ha indicado, la atribución de determinados signos transforma la longitud, el área y el volumen en funciones algebraicas con ciertas propiedades comunes. Por esta razón podemos esperar que la analogía correspondiente tenga lugar también en un espacio euclídeo arbitrario. Teniendo presente precisamente esta analogía, daremos a conocer la siguiente definición:

Se llama *volumen orientado* $V^{\pm}(x_1, x_2, \ldots, x_n)$ *del sistema de vectores* $x_1, x_2, \ldots, x_n$ del espacio euclídeo E_n el valor que toma en este sistema una función real que depende de n argumentos vectoriales de E_n y posee las propiedades (34.4).

En esta definición tampoco todo está claro. Se desconoce si existe o no el volumen orientado para cualquier sistema de vectores en un espacio euclídeo arbitrario cuando $n \geqslant 4$. Si incluso existe, ¿será unívoco el modo en que lo definen las propiedades (34.4)? Por fin, ¿qué relación existe, en el caso general, entre el volumen y el volumen orientado? Por ahora podemos responder solamante a la última pregunta y sólo en el caso en que $n = 1, 2, 3$.

A veces tendremos que considerar el volumen y el volumen orientado en el espacio E_n para los sistemas que contienen menos que n vectores. Esto será el indicio de que en realidad nos enfrentamos no con todo el espacio, sino con un subespacio suyo en el cual se toma el sistema dado. Correspondientemente, las propiedades (34.4) serán consideradas sólo respecto de los vectores del mismo subespacio. Puede surgir la necesidad de considerar el volumen y el volumen orientado para unos sistemas que contienen más que n vectores. De conformidad con la fórmula (35.4) y la propiedad **B** de (34.4), ambas funciones en tales sistemas han de ser iguales a cero.

En conclusión observemos que el empleo de dos diferentes conceptos relacionados con el volumen permitirá simplificar, en grado considerable, su investigación, puesto que un concepto refleja el aspecto geométrico del problema a resolver, mientras que el otro, el aspecto algebraico. Muy pronto se pondrá de manifiesto la relación existente entre ellos. Descubriremos, a continuación, que la importancia que proviene de la introducción de estos conceptos consiste, además, en que ellos generan cierto aparato matemático cuya significación no se limita al problema de un volumen.

Ejercicios.

1. Demuéstrese que en un espacio de segmentos dirigidos la longitud, el área y el volumen orientados se definen por las condiciones (34.4) de un modo unívoco.
2. ¿Será unívoco el modo en que se definen los mismos conceptos, si se excluye una de las condiciones (34.4)?
3. Demuéstrese que en todo espacio euclídeo $V(x_1, x_2) = |x_1|\cdot|x_2|$ cuando, y sólo cuando, los vectores x_1 y x_2 sean ortogonales.
4. Demuéstrese que en cualquier espacio euclídeo $V(x_1, x_2) = V(x_2, x_1)$.

§ 36. Propiedades geométricas y algebraicas del volumen

La investigación del concepto de volumen en el espacio euclídeo E_n la empezamos con el estudio de sus propiedades geométricas y algebraicas que provienen de la definición.

PROPIEDAD 1. *Siempre* $V(x_1, x_2, \ldots, x_n) \geqslant 0$. *La igualdad* $V(x_1, x_2, \ldots, x_n) = 0$ *tiene lugar cuando, y sólo cuando, el sistema de vectores* $x_1, x_2, \ldots, x_n$ *es linealmente dependiente.*

La primera parte de la afirmación se deduce obviamente de (35.4) y la demostración se necesita sólo para la segunda parte. Supongamos que el sistema $x_1, x_2, \ldots, x_n$ es linealmente dependiente. Si es que $x_1 = 0$, entonces, de la definición se desprende que el volumen es también igual a cero. En cambio, si $x_1 \neq 0$, entonces cierto vector x_{k+1} se expresa linealmente en términos de los vectores anteriores $x_1, \ldots, x_k$. Pero, en tal caso, ort $_{L_k}x_{k+1} = 0$ y el volumen de nuevo es igual a cero.

Supongamos ahora que el volumen es igual a cero. Por definición esto significa que es igual a cero uno de los factores en el segundo miembro de (35.4). Sea $i = k$ para este factor. Si $k = 0$, se tiene $x_1 = 0$. En el caso de que $k \neq 0$, la condición $\operatorname{ort}_{L_k} x_{k+1} = 0$ significa que el vector x_{k+1} pertenece a la cápsula lineal formada por los vectores $x_1, \ldots, x_k$, es decir, el sistema $x_1, \ldots, x_{k+1}$ es linealmente dependiente. En ambos casos todo el sistema de vectores $x_1, x_2, \ldots, x_n$ será también linealmente dependiente.

PROPIEDAD 2. *Para cualquier sistema de vectores $x_1, x_2, \ldots, x_n$ es válida la desigualdad de Hadamard*

$$V(x_1, x_2, \ldots, x_n) \leqslant \prod_{i=0}^{n-1} |x_{i+1}|, \tag{36.1}$$

con la particularidad de que la igualdad tiene lugar cuando, y sólo cuando, el sistema $x_1, x_2, \ldots, x_n$ es ortogonal o contiene el vector nulo.

Debido a las propiedades que poseen una perpendicular y una proyección, es válida, a ciencia cierta, la desigualdad

$$|\operatorname{ort}_{L_i} x_{i+1}| \leqslant |x_{i+1}|, \tag{36.2}$$

con la particularidad de que la desigualdad se convierte en una igualdad cuando, y sólo cuando, $x_{i+1} \perp L_i$, o bien, que es lo mismo, cuando el vector x_{i+1} es ortogonal a los vectores $x_1, x_2, \ldots, x_i$. Consideremos los productos de los miembros primero y segundo de las desigualdades del tipo (36.2) para todo i. Tenemos

$$\prod_{i=0}^{n-1} |\operatorname{ort}_{L_i} x_{i+1}| \leqslant \prod_{i=0}^{n-1} |x_{i+1}|.$$

Si todos los vectores del sistema $x_1, x_2, \ldots, x_n$ son no nulos, esta desigualdad se convierte en una igualdad cuando, y sólo cuando, el sistema es ortogonal. La presencia del vector nulo hace el caso trivial.

De la desigualdad de Hadamard se pueden deducir unos cuantos corolarios útiles. Supongamos que el sistema $x_1, x_2, \ldots, x_n$ está normalizado. En este caso, evidentemente,

$$V(x_1, x_2, \ldots, x_n) \leqslant 1.$$

Será cierta también la afirmación siguiente. Si el sistema $x_1, x_2, \ldots, x_n$ está normalizado y su volumen es igual a la unidad, será ortonormalizado. El hecho de que el volumen de cualquier sistema normalizado no es superior a la unidad, es testimonio de que entre todos los sistemas normalizados el sistema ortonormalizado tiene el volumen máximo.

PROPIEDAD 3. *Para cualesquiera dos conjuntos ortogonales de vectores $x_1, x_2, \ldots, x_p$ e $y_1, y_2, \ldots, y_r$ se verifica la igualdad*

$$V(x_1, x_2, \ldots, x_p, y_1, y_2, \ldots, y_r) = \\ = V(x_1, x_2, \ldots, x_p)\, V(y_1, y_2, \ldots, y_r).$$

Designemos mediante L_i una cápsula lineal formada por los primeros i vectores del sistema unido $x_1, \ldots, x_p, y_1, \ldots, y_r$, y mediante K_t, la cápsula lineal formada por los vectores $y_1, \ldots, y_t$. Por hipótesis, cada uno de los vectores del sistema $y_1, \ldots, y_r$ es ortogonal a todos los vectores del sistema $x_1, \ldots, x_p$. Por esta razón

$$L_{p+t} = L_p \oplus K_t$$

para cualquier valor de t desde 0 hasta r. Ahora, al tomar en consideración la igualdad (30.10), tenemos

$$V(x_1, \ldots, x_p, y_1, \ldots, y_r) = \left(\prod_{i=0}^{p-1} |\mathrm{ort}_{L_i} x_{i+1}|\right)\left(\prod_{t=0}^{r-1} |\mathrm{ort}_{L_{p+t}} y_{t+1}|\right) =$$
$$= \left(\prod_{i=0}^{p-1} |\mathrm{ort}_{L_i} x_{i+1}|\right)\left(\prod_{t=0}^{r-1} |\mathrm{ort}_{K_t} y_{t+1}|\right) = V(x_1, \ldots, x_p)\, V(y_1, \ldots, y_r).$$

Antes de pasar a las investigaciones ulteriores haremos una observación. El volumen de un sistema se expresa sólo en términos de las perpendiculares a las cápsulas lineales formadas por los vectores anteriores. Teniendo presente las propiedades de las perpendiculares, se puede deducir que el volumen del sistema no se alterará, si a un vector cualquiera se le añade una combinación lineal de los vectores anteriores. En particular, el volumen quedará intacto, si cualquier vector se sustituye por una perpendicular trazada desde este vector a cualquier cápsula lineal formada por los vectores anteriores.

PROPIEDAD 4. *El volumen del sistema de vectores no se altera cualquiera que sea la permutación de los vectores del sistema.*

Examinemos primero el caso en que dentro del sistema de vectores $x_1, \ldots, x_n$ conmutan dos vectores contiguos x_{p+1}, x_{p+2}. De conformidad con la observación recién mencionada, el volumen no se alterará, si los vectores x_{p+1}, x_{p+2} son sustituidos por los vectores $\mathrm{ort}_{L_p} x_{p+1}, \mathrm{ort}_{L_p} x_{p+2}$ y los vectores $x_{p+3}, \ldots, x_n$ por los vectores $\mathrm{ort}_{L_{p+2}} x_{p+3}, \ldots, \mathrm{ort}_{L_{p+2}} x_n$. En estas condiciones tres conjuntos de vectores

$$x_1, \ldots, x_p,$$
$$\mathrm{ort}_{L_p} x_{p+1},\ \mathrm{ort}_{L_p} x_{p+2}, \qquad \mathrm{ort}_{L_{p+2}} x_{p+3}, \ldots, \mathrm{ort}_{L_{p+2}} x_n$$

son ortogonales dos a dos y, en virtud de la propiedad 3, tendremos

$$V(x_1, \ldots, x_{p+1}, x_{p+2}, \ldots, x_n) = V(x_1, \ldots, x_p) \times$$
$$\times V(\mathrm{ort}_{L_p} x_{p+1}, \mathrm{ort}_{L_p} x_{p+2}) \cdot V(\mathrm{ort}_{L_{p+2}} x_{p+3}, \ldots, \mathrm{ort}_{L_{p+2}} x_n).$$

Está claro que las cápsulas lineales de los vectores $x_1, \ldots, x_p, x_{p+1}, x_{p+2}$ y $x_1, \ldots, x_p, x_{p+2}, x_{p+1}$ coinciden y, por consiguiente,

$$V(x_1, \ldots, x_{p+2}, x_{p+1}, \ldots, x_n) = V(x_1, \ldots, x_p) \times$$
$$\times V(\mathrm{ort}_{L_p} x_{p+2}, \mathrm{ort}_{L_p} x_{p+1}) \cdot V(\mathrm{ort}_{L_{p+2}} x_{p+3}, \ldots, x_{L_{p+2}} x_n).$$

En vista del isomorfismo euclídeo, el volumen de un sistema de dos vectores posee las mismas propiedades que el área de un paralelogramo. En particular, no depende del orden de los vectores del sistema. Comparando los primeros miembros de dos últimas igualdades, concluimos que

$$V(x_1, \ldots, x_{p+1}, x_{p+2}, \ldots, x_n) = V(x_1, \ldots, x_{p+2}, x_{p+1}, \ldots, x_n).$$

Un poco más tarde demostraremos que cualquier permutación $x_{j_1}, x_{j_2}, \ldots, x_{j_n}$ de vectores del sistema $x_1, x_2, \ldots, x_n$ puede ser obtenida por la permutación consecutiva de los vectores contiguos. Por ello, la propiedad 4 para una permutación arbitraria se desprende del caso particular examinado.

PROPIEDAD 5. *El volumen de un sistema de vectores es una función absolutamente homogénea, es decir,*

$$V(x_1, \ldots, \alpha x_p, \ldots, x_n) = |\alpha| V(x_1, \ldots, x_p, \ldots, x_n)$$

para cualquier valor de p.

De acuerdo con la propiedad 4, no disminuiremos la generalidad, si consideramos que $p = n$. Pero, en tal caso, al tomar en consideración (30.6), obtendremos

$$V(x_1, \ldots, x_{n-1}, \alpha x_n) = \left(\prod_{i=0}^{n-2} |\mathrm{ort}_{L_i} x_{i+1}|\right) |\mathrm{ort}_{L_{n-1}} (\alpha x_n)| =$$

$$= |\alpha| \prod_{i=0}^{n-1} |\mathrm{ort}_{L_i} x_{i+1}| = |\alpha| V(x_1, \ldots, x_{n-1}, x_n)$$

PROPIEDAD 6. *El volumen de un sistema de vectores no cambia, si a uno de los vectores del sistema se le añade una combinación lineal de todos los demás vectores.*

Recurriendo de nuevo a la propiedad 4, podemos considerar que al último vector se le agrega una combinación lineal de los vectores anteriores. Pero, según se ha observado, en el caso dado el volumen no cambia.

El volumen de un sistema de vectores es una función real. Esta función posee ciertas propiedades, una parte de las cuales ya la hemos establecido. Se ha confirmado nuestra suposición de que el volumen del sistema de vectores, determinado por nosotros para el espacio euclídeo, posee todas las propiedades propias precisamente del volumen para cualquier valor de n. Pero, lo más importante consiste, quizás, en que las propiedades establecidas definen el volumen de modo unívoco. Más preciso es válido el

TEOREMA 36.1. *Si una función real* $F(x_1, x_2, \ldots, x_n)$, *que depende de* n *argumentos vectoriales de* E_n, *posee las siguientes propiedades:*

A) *no cambia, si a cualquier argumento se le añade cualquiera combinación lineal de los argumentos restantes,*

B) *es absolutamente homogénea,* (36.3)

C) *es igual a la unidad para todos los sistemas ortonormalizados, entonces coincide con el volumen del sistema de vectores.*

DEMOSTRACIÓN. Si entre los argumentos $x_1, \ldots, x_n$ se tiene al menos uno que es nulo, entonces, de conformidad con la propiedad **B**,

$$F(x_1, x_2, \ldots, x_n) = V(x_1, x_2, \ldots, x_n) = 0. \qquad (36.4)$$

Supongamos ahora que el sistema $x_1, x_2, \ldots, x_n$ es arbitrario. Al sustraer de todo vector x_i su proyección sobre un subespacio formado por los vectores $x_1, \ldots, x_{i-1}$ y teniendo presente la propiedad **A**, concluimos que

$$F(x_1, x_2, \ldots, x_n) = F(\mathrm{ort}_{L_0} x_1, \mathrm{ort}_{L_1} x_2, \ldots, \mathrm{ort}_{L_{n-1}} x_n). \qquad (36.5)$$

Si el sistema $x_1, x_2, \ldots, x_n$ es linealmente dependiente, entonces entre los vectores $\mathrm{ort}_{L_{i-1}} x_i$ se tiene por lo menos un vector nulo y la igualdad (36.4) también tiene lugar. Supongamos que el sistema $x_1, x_2, \ldots, x_n$ es linealmente independiente; en este caso todos los vectores del sistema

$$\mathrm{ort}_{L_0} x_1, \mathrm{ort}_{L_1} x_2, \ldots, \mathrm{ort}_{L_{n-1}} x_n$$

serán no nulos. Como que dicho sistema es, además, ortogonal, existe un sistema ortonormalizado $e_1, e_2, \ldots, e_n$ para el cual

$$\mathrm{ort}_{L_{i-1}} x_i = |\mathrm{ort}_{L_{i-1}} x_i| \, e_i.$$

Debido a la propiedad **C**, $F(e_1, e_2, \ldots, e_n) = 1$.

Por eso, de la propiedad **B** se desprende que

$$F(x_1, x_2, \ldots, x_n) = \left(\prod_{i=0}^{n-1} |\mathrm{ort}_{L_i} x_{i+1}|\right) F(e_1, e_2, \ldots, e_n) =$$
$$= V(x_1, x_2, \ldots, x_n).$$

El teorema demostrado permite afirmar que si construimos de tal o cual modo una función que posea las propiedades (36.3), entonces esta función será el volumen del sistema de vectores.

Ejercicios.

1. Establézcase el sentido geométrico de las propiedades 2, 3, 6 en los espacios de segmentos dirigidos.
2. Establézcase el sentido geométrico de la igualdad (36.5) en los espacios de segmentos dirigidos.
3. ¿Podrá una función, que satisface las condiciones (36.3), ser nula en un sistema de vectores linealmente independiente?
4. Supongamos que respecto de una base ortonormalizada $e_1, e_2, \ldots, e_n$ el sistema de vectores $x_1, x_2, \ldots, x_n$ posee la propiedad

$$(x_i, e_j) = 0$$

para $i = 2, 3, \ldots, n$ y $j < i$ ($i = 1, 2, \ldots, n-1$ y $j > i$). Hállese la expresión para el volumen $V(z_1, z_2, \ldots, z_n)$ en términos de las coordenadas de los vectores $z_1, z_2, \ldots, z_n$ en la base $e_1, e_2, \ldots, e_n$.

5. ¿Qué cambiará, si consideramos el concepto de volumen en un espacio complejo?

§ 37. Propiedades algebraicas del volumen orientado

Pasemos, ahora a la investigación de las propiedades algebraicas del volumen orientado, dejando de lado hasta el momento la cuestión de su existencia. Tomemos como base de las investigaciones las condiciones **A**, **B**, **C** de (34.4).

PROPIEDAD 1. *El volumen orientado de un sistema de vectores es nulo, si dos vectores cualesquiera coinciden.*

Esta propiedad es un corolario directo de la condición **B**. Es fácil demostrar que, cumpliéndose la condición **A**, la condición **B** y la propiedad enunciada 1 son equivalentes.

PROPIEDAD 2. *El volumen orientado de un sistema de vectores cambia su signo, al cambiar de lugar dos vectores cualesquiera.*

La demostración es análoga para los vectores de toda clase, por lo que con el fin de simplificar las anotaciones nos limitaremos al examen del caso cuando cambiamos de lugar los vectores primero y segundo. De acuerdo con la propiedad 1,

$$V_{\pm}(x_1 + x_2,\ x_1 + x_2,\ x_3,\ \ldots,\ x_n) = 0.$$

Pero, por otro lado, según la condición **A**

$$\begin{aligned} V_{\pm}(x_1 + x_2, x_1 + x_2, x_3, \ldots, x_n) = V_{\pm}(x_1, x_1, x_3, \ldots, x_n) + \\ + V_{\pm}(x_2, x_2, x_3, \ldots, x_n) + V_{\pm} \\ \pm(x_1, x_2, x_3, \ldots, x_n) + V_{\pm}(x_2, x_1, x_3, \ldots, x_n). \end{aligned}$$

En el segundo miembro de esta igualdad los primeros dos sumandos son nulos, de donde se deduce la validez de la propiedad 2. Esta vez tampoco es difícil demostrar que si se cumple la condición **A**, la condición **B** y la propiedad enunciada 2 son equivalentes.

PROPIEDAD 3. *El volumen orientado de un sistema de vectores queda inalterable, si a cierto vector se le agrega una combinación cualquiera de los vectores restantes.*

Para simplificar, consideremos sólo el primer vector. Según la condición **A** tenemos $V_{\pm}\left(x_1 + \sum_{i=2}^{n} \alpha_i x_i,\ x_2,\ \ldots,\ x_n\right) =$

$$= V_{\pm}(x_1, x_2, \ldots, x_n) + \sum_{i=2}^{n} \alpha_i V_{\pm}(x_i, x_2, \ldots, x_n).$$

En esta igualdad todos los sumandos del segundo miembro, a excepción del primero, son nulos, de acuerdo con la propiedad 1.

PROPIEDAD 4. *El volumen orientado es una función homogénea, es decir, para cualquier valor de p se verifica la correlación*

$$V \pm (x_1, \ldots, \alpha x_p, \ldots, x_n) = \alpha V \pm (x_1, \ldots, x_p, \ldots, x_n).$$

Esta propiedad es un corolario inmediato de la condición A.

PROPIEDAD 5. *La igualdad $V \pm (x_1, x_2, \ldots, x_n) = 0$ tiene lugar cuando, y sólo cuando, el sistema de vectores $x_1, x_2, \ldots, x_n$ es linealmente dependiente.*

Evidentemente, es preciso sólo demostrar que de la igualdad $V \pm (x_1, x_2, \ldots, x_n) = 0$ se desprende la dependencia lineal de los vectores $x_1, x_2, \ldots, x_n$. Supondremos lo contrario. Admitamos que el volumen orientado es igual a cero para cierto sistema linealmente independiente $y_1, y_2, \ldots, y_n$. Este sistema es una base de E_n, razón por la cual para todo vector z de E_n se tiene

$$z = \alpha_1 y_1 + \alpha_2 y_2 + \alpha_n y_n.$$

Sustituyamos en el sistema $y_1, y_2, \ldots, y_n$ un vector cualquiera, por ejemplo, y_1 por el vector z. Haciendo uso sucesivamente de las propiedades 3 y 4, encontramos que

$$V \pm (z, y_2, \ldots, y_n) = V \pm (\alpha_1 y_1 + \sum_{i=2}^{n} \alpha_i y_i, y_2, \ldots, y_n) =$$
$$= V \pm (\alpha_1 y_1, y_2, \ldots, y_n) = \alpha_1 V \pm (y_1, y_2, \ldots, y_n) = 0.$$

Por definición, el volumen orientado no es nulo al menos en un sistema linealmente independiente $z_1, z_2, \ldots, z_n$. Pero, al sustituir sucesivamente los vectores $y_1, y_2, \ldots, y_n$ por los vectores $z_1, z_2, \ldots, z_n$, llegamos a que, de conformidad con el teorema 15.2, en este sistema el volumen orientado también es igual a cero. La contradicción obtenida demuestra la propiedad en consideración.

PROPIEDAD 6. *Si dos volúmenes orientados coinciden al menos en un sistema linealmente independiente de vectores, su coincidencia es idéntica.*

Supongamos que para nosotros es conocido que los volúmenes orientados $V_1^{\pm} (x_1, x_2, \ldots, x_n)$ y $V_2^{\pm} (x_1, x_2, \ldots, x_n)$ coinciden en el sistema linealmente independiente $z_1, z_2, \ldots, z_n$. Veamos la diferencia $F (x_1, x_2, \ldots, x_n) = V_1^{\pm} (x_1; x_2, \ldots, x_n) - V_2^{\pm} \times \times (x_1, x_2, \ldots, x_n)$. Esta función satisface las propiedades 3 y 4 del volumen orientado. Además, es nula en todos los sistemas linealmente dependientes y por lo menos en un sistema linealmente independiente $z_1, z_2, \ldots, z_n$. Repitiendo los razonamientos empleados en la demostración de la propiedad 5, concluimos que $F (x_1, x_2, \ldots$

$\ldots, x_n)$ es igual a cero en todos los sistemas linealmente independientes, es decir, que es nula idénticamente.

De la propiedad 6 se infiere que el volumen orientado se define por las condiciones (34.4) de modo único, si se fija aquel sistema ortonormalizado en que el volumen debe ser igual a la unidad.

PROPIEDAD 7. *El módulo del volumen orientado de un sistema de vectores coincide con el volumen del mismo sistema.*

Supongamos que el volumen orientado es igual a la unidad en el sistema ortonormalizado $z_1, z_2, \ldots, z_n$. Examinemos las funciones $|V \pm (x_1, x_2, \ldots, x_n)|$ y $V(x_1, x_2, \ldots, x_n)$. Ambas satisfacen las condiciones **A** y **B** de (36.3) y coinciden en el sistema linealmente independiente $z_1, z_2, \ldots, z_n$. La función

$$\Phi(x_1, x_2, \ldots, x_n) = \| V \pm (x_1, x_2, \ldots, x_n) | - V(x_1, x_2, \ldots, x_n) |$$

satisface también las condiciones **A**, **B**, de (36.3), es nula en todos los sistemas linealmente dependientes y por lo menos en un sistema linealmente independiente $z_1, z_2, \ldots, z_n$. Al repetir de nuevo los razonamientos realizados en la demostración de la propiedad 5, concluimos que $\Phi(x_1, x_2, \ldots, x_n)$ es nula idénticamente.

La última propiedad es de mucha importancia, puesto que permite afirmar que el módulo del volumen orientado debe poseer las mismas propiedades que tiene el volumen. En particular, debe ser igual en magnitud absoluta a la unidad en todos los sistemas ortonormalizados y no sólo en uno de ellos. Para el módulo es verídica la desigualdad de Hadamard, etc. Esta propiedad nos da la respuesta definitiva a todas las preguntas planteadas respecto del volumen y del volumen orientado. Lo único que nos falta es la demostración de *existencia* del volumen orientado.

Ejercicios.

1. Demuéstrese que si se cumple la condición **A** de (34.4), la condición **B** es equivalente tanto a la propiedad 1, como a la 2.

2. Demuéstrese que, cualquiera que sea el número real α, existe un sistema de vectores, para el cual el volumen orientado es igual a α.

3. Supongamos que la condición C de (34.4) ha sido sustituida por la condición de igualdad a cualquier número fijado en todo sistema fijado linealmente independiente. ¿Cómo cambiará el volumen orientado?

4. ¿Se usaban en la deducción de las propiedades del volumen orientado el hecho de que en el espacio lineal existe el producto escalar y que es real el volumen orientado? ¿Qué cambiará, si consideramos el volumen orientado en un espacio complejo?

§ 38. Permutaciones

Examinaremos un sistema $x_1, x_2, \ldots, x_n$ y otro sistema, $x_{j_1}, x_{j_2}, \ldots, x_{j_n}$, obtenido del primero realizando diversas permutaciones de vectores. Supongamos que de un sistema puede obtenerse el otro realizando permutaciones consecutivas sólo entre los pares de elementos de cada sistema. En tal caso los volúmenes de dichos sistemas serán iguales y los volúmenes orientados, o bien iguales o bien se diferenciarán en el signo, lo cual depende de la cantidad de permutaciones que han de realizarse.

En cuestiones de interés sobre permutaciones, las propiedades individuales de los vectores no desempeñarán ningún papel, pero será importante el orden de éstos. Por eso, en lugar de los propios vectores consideraremos sus números $1, 2, \ldots, n$. La totalidad de números

$$j_1, j_2, \ldots, j_n,$$

entre los cuales no hay números iguales y cada uno de los cuales es uno de los números $1, 2, \ldots, n$ se denomina *permutación* de estos números. La permutación $1, 2, \ldots, n$ se denomina *normal*.

Es fácil señalar que en un conjunto de n números la cantidad total de toda clase de permutaciones es igual a $n!$. En efecto, para $n = 1$ esto es evidente. Supongamos que la afirmación es justa para cualquier conjunto de $n - 1$ números. Todas las permutaciones de n números se pueden dividir en n clases, incluyendo en una clase sólo aquellas permutaciones cuyo primer número es el mismo. La cantidad de permutaciones en toda clase coincide con la cantidad de permutaciones de $n - 1$ números, es decir, es igual a $(n - 1)!$. Por consiguiente, la cantidad de todas las permutaciones de n números es igual a $n!$.

Suele decirse que en una permutación dada los números i, j forman *una inversión*, si $i > j$, pero i precede en la permutación a j. Una permutación se llamará *par*, siempre que sus números constituyan una cantidad par de inversiones y se llamará *impar*, en el caso contrario. Si en una permutación se cambian de lugar, dos números cualesquiera, no necesariamente contiguos, y todos los demás números quedan en sus lugares, obtendremos una permutación nueva. Esta transformación de la permutación se denomina *transposición*.

Demostremos que toda transposición altera la paridad de la permutación. Para los números contiguos esta afirmación es obvia. Su disposición mutua respecto de los demás números quedó la misma, mientras que la permutación de los propios números cambia la cantidad total de inversiones en una unidad.

Supongamos ahora que entre los números a permutar i y j se hallan otros s números $k_1, k_2, \ldots, k_s$, es decir, la permutación tiene por expresión

$$\ldots, i, k_1, k_2, \ldots, k_s, j, \ldots$$

Cambiaremos sucesivamente de lugar el número i con los números contiguos $k_1, k_2, \ldots, k_s, j$. A continuación, el número j, que ya precede a i, lo traslademos a la izquierda con ayuda de s transposiciones con los números k_s, k_{s-1}, . . ., k_1. De este modo, en total llevaremos a cabo $2s + 1$ transposiciones de los números que están junto uno al otro. Por consiguiente, la paridad de la transposición cambiará.

TEOREMA 38.1. *Todas las $n!$ permutaciones de n números se pueden disponer en un orden tal que cada permutación siguiente se obtenga de la anterior con ayuda de una transposición, con la particularidad de que para el inicio de proceso sirve cualquier permutación.*

DEMOSTRACIÓN. Esta afirmación es justa para $n = 2$. Si es necesario empezar por la permutación 1, 2, entonces la disposición buscada será 1, 2, 2, 1; en cambio, si empezamos con la permutación, 2, 1, entonces la disposición buscada será 2, 1, 1, 2.

Supongamos que el teorema ya se ha demostrado para cualesquiera permutaciones que contienen no más de $n - 1$ números. Consideremos las permutaciones de n números. Supongamos que debemos empezar por la permutación i_1, i_2, . . ., i_n. La disposición de las permutaciones se realizará según la siguiente regla. Comencemos por las permutaciones que en el primer lugar tienen el número i_1. Por hipótesis, se puede ordenar todas estas permutaciones en concordancia con los requisitos del teorema, puesto que, de hecho, es preciso disponer en el orden necesario todas las permutaciones de $n - 1$ números.

En la última permutación, obtenida de esta manera, realizamos una transposición, trasladando al primer lugar el número i_2. A continuación, ordenamos, igual que en el caso precedente, todas las permutaciones que en el primer lugar tienen un número dado, etc. Por el procedimiento indicado podemos probar todas las permutaciones de n números.

Con tal sistema de disposición de las permutaciones de n números las permutaciones contiguas tendrán paridades contrarias. Tomando en consideración la paridad del número $n!$ para $n \geqslant 2$, podemos concluir que en este caso la cantidad de permutaciones pares de n números es igual a la cantidad de permutaciones impares y equivale a $\frac{1}{2}\, n!$.

Ejercicios.

1. ¿Cuál es la paridad de la permutación 5, 2, 3, 1, 4?

2. Demuéstrese que ninguna de las permutaciones pares (impares) puede ser reducida a la normal en el transcurso de un número impar (par) de transposiciones.

3. Examinemos un par de permutaciones i_1, i_2, . . ., i_n y 1, 2, . . ., n. Reduciremos a la forma normal la primera permutación con ayuda de transposiciones, realizando en cada una de éstas una transposición de cualesquiera elementos en la segunda permutación. Demuéstrese que terminado el proceso, la segunda permutación tendrá la misma paridad que la permutación i_1, i_2, ..., i_n.

§ 39. Existencia de un volumen orientado

Examinemos ahora la cuestión sobre la existencia de un volumen orientado de un sistema de vectores. Supongamos que en el espacio E_n se ha elegido un sistema ortonormalizado $z_1, z_2, \ldots, z_n$ en el cual el volumen orientado ha de ser igual a uno, de acuerdo con la condición **C** de (34.4). Tomemos un sistema arbitrario $x_1, x_2, \ldots, x_n$ de vectores de E_n. Puesto que el sistema $z_1, z_2, \ldots, z_n$ es la base en E_n, para todo vector x_i existe la siguiente descomposición según la base citada:

$$x_i = a_{i1}z_1 + a_{i2}z_2 + \ldots + a_{in}z_n, \tag{39.1}$$

donde a_{ij} son ciertos números.

Si el volumen orientado existe, de conformidad con la condición **A** de (34.4), podemos transformarlo sucesivamente teniendo presente las descomposiciones (39.1). A saber,

$$\begin{aligned} V^{\pm}(x_1, x_2, \ldots, x_n) &= V^{\pm}\Big(\sum_{j_1=1}^{n} a_{1j_1}z_{j_1}, \sum_{j_2=1}^{n} a_{2j_2}\times \\ &\qquad \times z_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}z_{j_n}\Big) = \\ &= \sum_{j_1=1}^{n} a_{1j_1}V^{\pm}\Big(z_{j_1}, \sum_{j_2=1}^{n} a_{2j_2}z_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}z_{j_n}\Big) = \\ &= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} a_{1j_1}a_{2j_2}V^{\pm}\Big(z_{j_1}, z_{j_2}, \ldots, \sum_{j_n=1}^{n} a_{nj_n}z_{j_n}\Big) = \ldots \\ \ldots &= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\ldots\sum_{j_n=1}^{n} a_{1j_1}a_{2j_2}\ldots a_{nj_n}V^{\pm}(z_{j_1}, z_{j_2}, \ldots, z_{j_n}). \end{aligned} \tag{39.2}$$

En la última suma múltiple de n la mayor parte de los sumandos es igual a cero, puesto que, de acuerdo con la propiedad 1, el volumen orientado del sistema de vectores es nulo, siempre que dos vectores cualesquiera coincidan. Por ello, entre los sistemas $z_{j_1}, z_{j_2}, \ldots, z_{j_n}$ se deben considerar sólo aquellos, para los cuales el surtido de índices $j_1, j_2, \ldots, j_n$ representa una permutación de n números $1, 2, \ldots, n$. Mas, es este caso

$$V^{\pm}(z_{j_1}, z_{j_2}, \ldots, z_{j_n}) = \pm 1$$

según sea par o impar la permutación de los índices.

De este modo, si el volumen orientado existe, debe expresarse en términos de las coordenadas de los vectores $x_1, x_2, \ldots, x_n$

en la base $z_1, z_2, \ldots, z_n$ mediante la siguiente fórmula:

$$V^{\pm}(x_1, x_2, \ldots, x_n) = \sum \pm a_{1j_1}a_{2j_2}, \ldots, a_{nj_n}. \qquad (39.3)$$

Aquí la sumación se realiza según todas las permutaciones de índices $j_1, j_2, \ldots, j_n$ de los números $1, 2, \ldots, n$ y el signo más o el signo menos se toma en dependencia de la paridad o imparidad de la permutación.

Demostremos que una función, prefijada por el segundo miembro de la igualdad (39.3), satisface todas las condiciones que definen un volumen orientado. Sea el vector x_p una combinación lineal de los vectores x_p' y x_p'', es decir,

$$x_p = \alpha x_p' + \beta x_p''$$

para ciertos números α, β. Designemos con a_{pj}' y a_{pj}'' las coordenadas de los vectores x_p' y x_p'', respectivamente, en la base $z_1, z_2, \ldots, z_n$. Entonces, es evidente que

$$a_{pj} = \alpha a_{pj}' + \beta a_{pj}''$$

para todo j desde 1 hasta n. Luego hallamos

$$\sum \pm a_{1j_1} \ldots, a_{pj_p} \ldots, a_{nj_n} = \sum \pm a_{1j_1} \ldots$$
$$\ldots (\alpha a_{pj_p}' + \beta a_{pj_p}'') \ldots a_{nj_n} = \alpha \sum \pm a_{1j_1} \ldots$$
$$\ldots a_{pj_p}' \ldots a_{nj_n} + \beta \sum \pm a_{1j_1} \ldots, a_{pj_p}'' \ldots a_{nj_n},$$

y la condición A de (34.4) queda cumplida.

Supongamos que permutamos dos vectores cualesquiera del sistema $x_1, x_2, \ldots, x_n$. En este caso la función (39.3) cambiará de signo, puesto que variará la paridad de cada permutación. Como ya se ha indicado antes, si es verídica la propiedad de linealidad respecto de cada argumento, la propiedad demostrada equivaldrá al cumplimiento de la condición B de (34.4).

Y, por fin, examinemos el significado de la función construida sobre el sistema de vectores $z_1, z_2, \ldots, z_n$. Para este sistema, las coordenadas a_{ij} tienen la forma

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

Por consiguiente, entre los sumandos de (39.3) será no nulo solamente el sumando $a_{11}a_{22} \ldots a_{nn}$. La permutación $1, 2, \ldots, n$ es par, los elementos $a_{11}, a_{22}, \ldots, a_{nn}$ son iguales a uno, por lo cual el valor de la función en el sistema ortonormalizado $z_1, z_2, \ldots, z_n$ es igual a la unidad.

Así pues, todas las condiciones (34.4) quedan cumplidas y la función (39.3) representa la expresión del volumen orientado del sistema de vectores en los términos de las coordenadas. Esta expresión es *única* en virtud de la unicidad del volumen orientado.

Ejercicios.

1. ¿Se ha usado en esencia el carácter ortonormalizado del sistema $z_1, z_2, \ldots, z_n$ en la deducción de la fórmula (39.3)?
2. ¿Qué cambios tendrán lugar en la fórmula (39.3), si en la condición C de (34.4) el volumen orientado se considera distinto de uno?
3. ¿Cómo esencial fue el uso de la condición B de (34.4) al deducir la fórmula (39.3)?
4. ¿Cambiará la fórmula (39.3) su forma, si el volumen orientado se considera en un espacio complejo?

§ 40. Determinantes

Supongamos que los vectores $x_1, x_2, \ldots, x_n$ del espacio euclídeo $\mathbf{R}_n$ están dados por sus coordenadas

$$x_i = (a_{i1}, a_{i2}, \ldots, a_{in})$$

en la base (21.7). Dispondremos los números a_{ij} en forma de una tabla A del modo siguiente:

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

Esta tabla se denomina *matriz cuadrada de orden n*, los números a_{ij} son *elementos* de la matriz. Si numeramos las filas de una matriz consecutivamente de arriba abajo y las columnas, de izquierda a derecha, entonces el primer índice de un elemento significa el número de la fila en que se halla el elemento, y el segundo índice, número de la columna. Con respecto a los elementos $a_{11}, a_{22}, \ldots, a_{nn}$ se dice que ellos forman *la diagonal principal* de la matriz A.

Cualesquiera n^2 números pueden disponerse en forma de una matriz cuadrada de orden n. Si los elementos de la fila de una matriz se consideran como coordenadas de cierto vector de $\mathbf{R}_n$ en la base (21.7), entonces entre todas las matrices cuadradas de orden n y los sistemas ordenados de n vectores del espacio $\mathbf{R}_n$ se establece una correspondencia biunívoca.

En el espacio $\mathbf{R}_n$, al igual que en cualquier otro espacio, existe un volumen orientado. Será, además, el único, si exigimos que se cumpla la condición C de (34.4) en el sistema de vectores (21.7). Al tomar en consideración la citada correspondencia biunívoca, concluimos que en el conjunto de todas las matrices cuadradas se genera una función bien determinada. Teniendo presente (39.3), llegamos a la siguiente definición de esta función.

Se llama *determinante* de orden n, correspondiente a la matriz A, la suma algebraica de $n!$ términos compuesta de la siguiente manera. Como términos del determinante intervienen toda una serie de productos de n elementos de la matriz, uno en cada fila y en cada columna. El término se toma con signo más, si los índices de las columnas de sus elementos forman una permutación par, a condición de que los propios elementos están dispuestos en orden creciente de los números de las filas; el signo menos se toma en el caso contrario. Para designar un determinante usaremos el símbolo siguiente:

$$\det\begin{pmatrix} a_{11}a_{12}\ldots 1_{1n} \\ a_{21}a_{22}\ldots a_{2n} \\ \cdots\cdots\cdots \\ a_{n1}a_{n2}\ldots a_{nn}\end{pmatrix}$$

si es preciso indicar la forma explícita de los elementos de la matriz. En cambio, si esto no es necesario, se usará un símbolo más sencillo $\det A$, donde se indica sólo la designación de la matriz A. Los elementos de la matriz del determinante se llamarán también *elementos del determinante*.

El determinante coincide con el volumen orientado del sistema de filas de la matriz. Por esta razón, al investigarlo se pueden utilizar todos los datos conocidos referentes al volumen y al volumen orientado. En particular, un determinante es igual a cero cuando, y sólo cuando, las filas de la matriz son linealmente dependientes; el determinante cambia de signo cuando se intercambian de lugar dos filas cualesquiera, etc. Ahora nuestras investigaciones tocarán aquellas propiedades del determinante que son difíciles de demostrar sin emplear la expresión explícita del determinante en términos de los elementos de la matriz.

Llamemos *transposición* de una matriz la transformación de esta última en que las filas se convierten en columnas y las columnas, en filas con los mismos números. Una matriz, transpuesta respecto a la matriz A, se indica con A'. Suele decirse en este caso que el determinante

$$\det\begin{pmatrix} a_{11}a_{21}\ldots a_{n1} \\ a_{12}a_{22}\ldots a_{n2} \\ \cdots\cdots\cdots \\ a_{1n}a_{2n}\ldots a_{nn}\end{pmatrix}$$

o $\det A'$ se ha obtenido por transposición del determinante (40.1). Con relación a la transposición, un determinante posee las siguientes propiedades de importancia.

Un determinante de cualquier matriz no varía durante la operación de transposición.

En efecto, el determinante de la matriz A está compuesto por los términos del tipo

$$a_{1j_1}a_{2j_2}\ldots a_{nj_n}, \tag{40.2}$$

cuyo signo se determina por la paridad de la permutación $j_1, j_2, \ldots, j_n$. Todos los factores del producto (40.2) en la matriz transpuesta A' quedan en distintas filas y distintas columnas, es decir, su producto es un término del determinante transpuesto. Designemos los elementos de la matriz A' mediante a'_{ij}. Está claro que $a'_{ij} = a_{ji}$, por lo cual

$$a_{1j_1}a_{2j_2}\ldots a_{nj_n} = a'_{j_1 1}a'_{j_2 2}\ldots a'_{j_n n}. \tag{40.3}$$

Ordenemos los elementos del segundo miembro de (40.3) según el orden de crecimiento de los números de filas. En este caso la permutación de los índices de columnas tendrá la misma paridad que la permutación $j_1, j_2, \ldots, j_n$. Mas, esto significa que el signo del término (40.2) en el determinante transpuesto será el mismo que en el determinante inicial. Por consiguiente, ambos determinantes se componen de unos mismos términos de signos iguales, es decir, ellos coinciden.

De la propiedad demostrada proviene que en un determinante las filas y columnas se encuentran en las mismas condiciones. Por esto, todas las propiedades demostradas anteriormente respecto a las filas serán justas también, para las columnas.

Consideremos un determinante d de orden n. Elijamos arbitrariamente en su matriz k filas y k columnas. Los elementos que se hallan en la intersección de las filas y columnas elegidas forman una matriz de orden k. El determinante de esta matriz se denomina *menor* de orden k del determinante d. El menor dispuesto en las primeras k columnas y las primeras k filas se llama menor *principal* o menor *angular*.

Supongamos ahora que en el determinante d de orden n se ha elegido un menor M de orden k. Si suprimimos aquellas filas y columnas en cuyas intersecciones se halla el menor M, quedará el menor N de orden $n-k$. Éste se llamará menor *complementario* de M. En cambio, si suprimimos las filas y columnas en las cuales se ubican los elementos del menor N, quedará, evidentemente, el menor M. De este modo, se puede hablar de un par de menores recíprocamente complementarios.

Si el menor M de orden k se halla en las filas de números $i_1, i_2, \ldots, i_k$ y en las columnas de números $j_1, j_2, \ldots, j_k$, entonces el número

$$(-1)^{\sum\limits_{p=1}^{k}(i_p+j_p)} N \tag{40.4}$$

se llamará *complemento algebraico* del menor M.

TEOREMA 40.1: (*teorema de Laplace*). *Supongamos que en el determinante d de orden n se han elegido arbitrariamente k filas (columnas), donde* $1 \leqslant k \leqslant n-1$. *Entonces, la suma de los productos de todos los menores de k-ésimo orden, contenidos en las filas (columnas) elegidas, por sus complementos algebraicos es igual al determinante d.*

DEMOSTRACION. Consideremos las columnas de la matriz del determinante d como vectores $x_1, x_2, \ldots, x_n$ del espacio $\mathbf{R}_n$. La suma de los productos de todos los menores de k-ésimo orden, contenidos en las filas elegidas, por sus complementos algebraicos puede considerarse como cierta función $F(x_1, x_2, \ldots, x_n)$, dependiente de los vectores $x_1, x_2, \ldots, x_n$.

Esta función es, a ciencia cierta, lineal respecto de todo argumento, ya que la propiedad dada la poseen tanto los menores como los complementos algebraicos. Dicha función es igual a uno en el sistema ortonormalizado (21.7), de lo que es fácil convencerse por comprobación inmediata. Si demostramos que $F(x_1, x_2, \ldots, x_n)$ cambia de signo, al cambiar de lugar dos vectores cualesquiera, la función coincidirá con el volumen orientado del sistema de vectores $x_1, x_2, \ldots, x_n$. Pero el volumen orientado coincide con el determinante de la matriz en la que las coordenadas de los vectores están dispuestas en las filas. Puesto que el determinante de la matriz coincide con el de la matriz transpuesta, la demostración del teorema de Laplace se dará por terminado.

Evidentemente, sólo es necesario considerar un caso en que conmutan dos vectores contiguos, pues la permutación de cualesquiera dos vectores siempre se reduce a un número impar de permutaciones de vectores contiguos. La demostración de esta afirmación se ha aducido en el párrafo 38.

Supongamos que conmutan los vectores x_i y x_{i+1}. Establezcamos una correspondencia biunívoca entre los menores en las filas elegidas del determinante inicial y del determinante con las columnas permutadas. Designemos mediante ω la totalidad de los números de las columnas que definen el menor. Son posibles los siguientes casos:

1) $i,\ i+1 \in \omega$,

2) $i,\ i+1 \bar{\in} \omega$,

3) $i \in \omega,\ i+1 \bar{\in} \omega$,

4) $i+1 \in \omega,\ i \bar{\in} \omega$.

En los casos 1, 2 a todo menor le pondremos en correspondencia otro menor, dispuesto en las columnas con la totalidad de números ω; en los casos 3, 4, un menor, dispuesto en las columnas con la totalidad de números obtenida de ω mediante la sustitución de i por $i+1$ e $i+1$ por i, respectivamente.

Hemos de notar que en todos los casos los menores correspondientes se determinan mediante una misma totalidad de elementos. Más aún, en los casos 2—4 los menores coinciden y en el caso 1, sólo difieren en el signo, puesto que en uno de ellos están permutadas, respecto del otro, dos columnas. En virtud de las razones análogas, en el caso 2 los menores complementarios correspondientes se diferencian en el signo, siendo coincidentes en los casos restantes. Los complementos algebraicos se diferencian de los menores complementarios sólo en el signo, determinado por la paridad de la suma de los números de las filas y columnas en las cuales se encuentra el menor. En los casos 1, 2 estos números de los menores correspondientes son idénticos y en los casos 3, 4 difieren en unidad.

Comparando ahora los sumandos correspondientes de la función $F(x_1, x_2, \ldots, x_n)$ y de una función obtenida por permutación de los vectores x_i y x_{i+1}, observamos que difieren sólo en el signo. Por consiguiente, en la permutación de dos vectores contiguos la función $F(x_1, x_2, \ldots, x_n)$ cambia de signo.

El teorema demostrado se emplea con frecuencia cuando se elige sólo una fila o bien una columna. El determinante de una matriz de orden 1 coincide con su único elemento. Por ello, el menor dispuesto en la intersección de la i-ésima fila y la j-ésima columna equivale al elemento a_{ij}. Designemos mediante A_{ij} el complemento algebraico del elemento a_{ij}. De acuerdo con el teorema de Laplace, para todo i se tiene

$$a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} = d. \tag{40.5}$$

Esta fórmula se llama *desarrollo del determinante por la i-ésima fila*. Análogamente, para todo j

$$a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj} = d, \tag{40.6}$$

lo que nos proporciona *el desarrollo del determinante por la j-ésima columna*.

En el desarrollo (40.5) sustituyamos los elementos de la i-ésima fila por una totalidad de los n números arbitrarios $b_1, b_2, \ldots, b_n$. La expresión

$$b_1A_{i1} + b_2A_{i2} + \ldots + b_nA_{in}$$

representa el desarrollo del determinante

$$\det\begin{pmatrix} a_{11}a_{12}\ldots a_{1n} \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ b_1b_2\ldots b_n \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ a_{n1}a_{n2}\ldots a_{nn}\end{pmatrix} \tag{40.7}$$

por la i-ésima fila. Este último se obtiene del determinante d sustituyendo la i-ésima fila por una fila de los números $b_1, b_2, \ldots, b_n$. Tomemos ahora, a título de dichos números, los elementos de la k-ésima fila del determinante d para $k \neq i$. El determinante correspondiente (40.7) es igual a cero, ya que tiene dos filas idénticas. Por consiguiente,

$$a_{k1}A_{i1} + a_{k2}A_{i2} + \ldots + a_{kn}A_{in} = 0, \; k \neq i. \qquad (40.8)$$

Por analogía,

$$a_{1k}A_{1j} + a_{2k}A_{2j} + \ldots + a_{nk}A_{nj} = 0, \; k \neq j. \qquad (40.9)$$

Así pues, la suma de los productos de todos los elementos de cualquier fila (columna) de un determinante por los complementos algebraicos de los elementos correspondientes de otra fila (columna) del mismo determinante es igual a cero.

Observemos, como conlusión, que toda la teoría de los determinantes es aplicable, sin cambio alguno, al caso de las matrices complejas. Lo único que se pierde es la claridad relacionada con el concepto de volumen.

Ejercicios.

1. Escríbanse las expresiones de los determinantes de los órdenes segundo y tercero en términos de los elementos de las matrices. Compárense con la expresión (34.6).
2. Escríbanse la desigualdad de Hadamard para el determinante de las matrices A y A'.
3. Un determinante de n-ésimo orden, todos los elementos del cual son iguales en módulo a la unidad, equivale a $n^{n/2}$. Demuéstrese que sus filas (columnas) forman una base otrogonal.
4. ¿A qué es igual un determinante, si sus elementos satisfacen las condiciones $a_{ij} = 0$ para $i > j$ ($i < j$, $i \geqslant j$, $i \leqslant j$)?
5. Los elementos de un determinante satisfacen las condiciones $a_{ij} = 0$ para $i > k$ y $j \leqslant k$. Demuéstrese que el determinante es igual al producto del menor principal de orden k por su menor complementario.
6. Supongamos que los elementos de una matriz compleja A satisfacen las condiciones $a_{ij} = \overline{a_{ji}}$ para cualesquiera i, j. Demuéstrese que el determinante de esta matriz es un número real.

§ 41. Dependencia lineal y determinantes

Las más difundidas aplicaciones de los determinantes las encontramos en los problemas relacionados con la dependencia lineal. Supongamos que en un espacio K_n de dimensión n vienen dados m vectores $x_1, x_2, \ldots, x_m$ y es necesario determinar la base de estos vectores. Elijamos una base cualquiera en K_n y consideremos una tabla rectangular

$$A = \begin{pmatrix} a_{11} a_{12} \ldots a_{1n} \\ a_{21} a_{22} \ldots a_{2n} \\ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \\ a_{m1} a_{m2} \ldots a_{mn} \end{pmatrix}, \qquad (41.1)$$

cuyas filas representan las coordenadas de los vectores dados en la base elegida.

Esta tabla se denomina *matriz rectangular*. El primer índice del elemento a_{ij} significa, como hasta ahora, el número de la fila de la matriz en la que se halla dicho elemento, el segundo índice significa el número de la columna. Si se desea subrayar la cantidad de filas y columnas en la matriz A, se escribirá A $(m \times n)$ y de la matriz A se dirá que es una matriz $m \times n$. La matriz A $(n \times n)$ se llamará, como antes, matriz cuadrada de orden n. Conjuntamente con la matriz A consideraremos la matriz transpuesta A'. Si las dimensiones de A son $m \times n$, las dimensiones de A' serán $n \times m$.

En una matriz rectangular A $(m \times n)$ pueden también indicarse diferentes menores cuyo orden no es superior, naturalmente, al menor de los números m, n. Si una matriz tiene no sólo elementos nulos, el orden superior r de los menores, distintos de cero, se denominará *rango* de la matriz A. Cualquier menor de rango r distinto de cero se llama menor *básico*; *básicas* serán también las filas y las columnas en las cuales se encuentra el menor básico. Está claro que pueden haber varios menores básicos. El rango de la matriz nula es igual a cero lo que se desprende de la definición.

Consideraremos que las filas de la matriz A son vectores. Es evidente que si determinamos la base de estos vectores filas, los vectores correspondientes del espacio K_n formarán la base de los vectores $x_1, x_2, \ldots, x_m$.

TEOREMA 41.1. *Cualesquiera filas básicas forman la base de los vectores filas de la matriz.*

DEMOSTRACIÓN. Para convencerse de la validez del teorema hace falta mostrar que las filas básicas son linealmente independientes y toda fila de la matriz se expresa linealmente en términos de estas últimas.

Si las filas básicas fueran linealmente dependientes, entonces una de estas filas se expresaría linealmente en términos de las filas básicas restantes. Pero en este caso el menor debe ser igual a cero lo que contradice la hipótesis.

Ahora agreguemos a las filas básicas otra fila cualquiera de la matriz A. Entonces, por definición del menor básico, todos los menores de orden $r + 1$, dispuestos en dichas filas, serán iguales a cero. Supongamos que las filas citadas son linealmente independientes. Al completarlas hasta obtener una base, habremos construid cierta matriz cuadrada cuyo determinante no debe ser nulo. Pero, por otro lado, desarrolando este determinante por $r+1$ filas iniciales llegamos a la conclusión de que es igual a cero. La contradicción obtenida es testimonio de que toda fila de la matriz A se expresa linealmente en términos de las filas básicas.

El teorema demostrado permite reducir el problema de búsqueda de la base de un sistema de vectores a la búsqueda del menor básico de la matriz. Puesto que el determinante de la matriz transpuesta coincide con el de la matriz inicial, está claro que el teorema 41.1 es válido no sólo para las filas, sino también para las columnas. Esto significa que para cualquier matriz rectangular el rango de su sistema de sus vectores filas es igual al rango de su sistema de vectores columnas. *Lo expuesto no es obvio, si se tiene en cuenta sólo el concepto de rango de un sistema de vectores.*

En un espacio dotado de un producto escalar la dependencia o independencia lineal del sistema de vectores $x_1, x_2, \ldots, x_m$ puede establecerse sin recurrir al desarrollo por la base. Examinemos un determinante

$$G(x_1, x_2, \ldots, x_m) = \det \begin{pmatrix} (x_1, x_1) & (x_1, x_2) & \ldots & (x_1, x_m) \\ (x_2, x_1) & (x_2, x_2) & \ldots & (x_2, x_m) \\ \cdot & \cdot & \cdot & \cdot \\ (x_m, x_1) & (x_m, x_2) & \ldots & (x_m, x_m) \end{pmatrix},$$

que se llama *determinante de Gram* del sistema de vectores $x_1, x_2, \ldots, x_m$.

TEOREMA 41.2. *Un sistema de vectores es linealmente dependiente cuando, y sólo cuando, su determinante de Gram es igual a cero.*

DEMOSTRACIÓN. Supongamos que el sistema de vectores $x_1, x_2, \ldots, x_m$ es linealmente dependiente. En este caso existen tales números $\alpha_1, \alpha_2, \ldots, \alpha_m$, no todos iguales a cero, que

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m = 0.$$

Al multiplicar esta igualdad escalarmente por x_i para todo valor de i, concluimos que las columnas del determinante de Gram son también linealmente dependientes, en otras palabras, el propio determinante es igual a cero.

Supongamos ahora que el determinante de Gram es nulo. Entonces sus columnas son linealmente dependientes, es decir, existen tales números $\alpha_1, \alpha_2, \ldots, \alpha_m$, no todos iguales a cero, que

$$\alpha_1 (x_i, x_1) + \alpha_2 (x_i, x_2) + \ldots + \alpha_m (x_i, x_m) = 0$$

para todo i. Escribamos estas igualdades de la forma siguiente:

$$(x_i, \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m) = 0.$$

Multiplicándolas término a término por $\alpha_1, \alpha_2, \ldots, \alpha_m$ y sumándolas, obtenemos

$$|\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m|^2 = 0.$$

Esto significa que

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_m x_m = 0,$$

es decir, que los vectores $x_1, x_2, \ldots, x_m$ son linealmente dependientes.

Ejercicios.

1. ¿Qué representa en sí una matriz, todos los menores de la cual son nulos?
2. ¿Serán las filas básicas y columnas básicas sistemas equivalentes de vectores para una matriz cuadrada?
3. ¿Cambiarán el rango de las matrices las transformaciones elementales examinadas en el § 15?
4. Demuéstrese la desigualdad

$$0 \leqslant G(x_1, x_2, \ldots, x_n) \leqslant \prod_{i=1}^{n} (x_i, x_i).$$

¿Cuáles son los casos en que se consiguen igualdades?

5. Es evidente que

$$\det \begin{pmatrix} \sqrt{2} & 1 \\ 2 & \sqrt{2} \end{pmatrix} = 0.$$

Demuéstrese que con cualquier aproximación del número $\sqrt{2}$ por medio de una fracción decimal p

$$\det \begin{pmatrix} p & 1 \\ 2 & p \end{pmatrix} \neq 0.$$

§ 42. Cálculo de los determinantes

El cálculo directo de un determinante, durante el cual se usa la expresión explícita de éste en términos de los elementos de la matriz, es de rara aplicación en la práctica por ser muy laborioso. Un determinante de orden n se compone de $n!$ términos, mientras que para calcular cualquier término y sumarlo con los otros es preciso efectuar n operaciones aritméticas. Si todos estos cálculos se realizaran con ayuda de una computadora moderna, capaz de ejecutar 10^6 operaciones aritméticas por segundo, incluso en este caso, para calcular un determinante de orden 100, por ejemplo, tendríamos que esperar el resultado durante varios millones de años.

Uno de los métodos más efectivos para calcular determinantes está basado en la siguiente idea. Supongamos que en la matriz A existe un elemento a_{kp} distinto de cero. Llamémoslo *elemento rector*. Si a toda i-ésima fila, $i \neq k$, agregamos la k-ésima fila multiplicada por un número arbitrario α_i, el determinante, como se sabe, no variará. Tomemos

$$\alpha_i = -\frac{a_{ip}}{a_{kp}}$$

y llevemos a cabo el procedimiento citado para todo $i \neq k$. Entonces, en la matriz nueva todos los elementos de la p-ésima columna, a excepción del rector, serán iguales a cero. Desarrollando el determinante nuevo por la p-ésima columna, reduzcamos el cálculo del determinante de n-ésimo orden al cálculo de un solo determinante de orden $(n - 1)$. Con este último procedamos de un modo análogo, etc.

El algoritmo descrito se llama *método de Gauss*. Para calcular un determinante de n-ésimo orden, rigiéndose por este método, se requiere cumplir en total alrededor de $\frac{2}{3}n^3$ operaciones aritméticas. De esta manera un determinante de orden 100 puede ser calculado en menos de un segundo en una máquina computadora que realiza 10^6 operaciones aritméticas por segundo.

En conclusión observemos que bajo las condiciones en que las operaciones aritméticas se calculan aproximadamente y los datos se prefijan también de modo aproximado, los resultados de cálculos de los determinantes deben tratarse con cierta cautela. Si las deducciones sobre la dependencia o independencia lineal de un sistema de vectores se basan sólo en el hecho de si es o no igual a cero el determinante, entonces en presencia de la inestabilidad de la que se ha tratado en el § 22, las deducciones pueden resultar erróneas. Al operar con los determinantes, esto siempre debe tenerse en cuenta.

Ejercicios.

1. ¿A qué se debe que los cálculos de un determinante por el método de Gauss son más rápidos que los cálculos directos?

2. Supongamos que todos los elementos de un determinante no sobrepasan, en módulo, la unidad y que al calcular cada término del mismo se comete un error del orden ε. ¿Con qué n el cálculo directo del determinante tendrá sentido desde el punto de vista de su precisión?

3. Constrúyase el algoritmo, basándose en el método de Gauss, para calcular el rango de una matriz rectangular. ¿Qué significa la aplicación de este algoritmo a las condiciones de cálculos aproximados?

CAPITULO 5 LÍNEA RECTA Y PLANO EN EL ESPACIO LINEAL

§ 43. Ecuaciones de la línea recta y del plano

El objeto principal de nuestras investigaciones inmediatas serán una recta y un plano en los espacios de segmentos dirigidos. Si fijamos cierto sistema de coordenadas, las coordenadas de los puntos ubicados en una línea recta o en un plano ya no pueden ser arbitrarias, sino que han de satisfacer correlaciones determinadas. Pasamos ahora a la deducción de dichas correlaciones.

Fijemos en un plano el sistema rectangular cartesiano de coordenadas Oxy y una recta L. Consideraremos un vector no nulo

$$n = (A, B), \tag{43.1}$$

que es perpendicular a L. Evidentemente, todos los demás vectores perpendiculares a dicha recta serán colineales a n.

Elijamos un punto arbitrario $M_0(x_0, y_0)$ en la línea recta. Todos los puntos $M(x, y)$ de la recta L, y sólo ellos, poseen la propiedad de perpendicularidad de los vectores $\overrightarrow{M_0M}$ y n, es decir,

$$(\overrightarrow{M_0M}, n) = 0. \tag{43.2}$$

Puesto que

$$\overrightarrow{M_0M} = (x - x_0, y - y_0),$$

entonces, de (43.1), (43.2) se deduce que

$$A(x - x_0) + B(y - y_0) = 0.$$

Al introducir la designación

$$-Ax_0 - By_0 = C,$$

concluimos que en el sistema dado Oxy las coordenadas de los puntos de la recta L, y sólo ellas, satisfacen la ecuación

$$Ax + By + C = 0. \tag{43.3}$$

Entre los números A, B hay uno que no es igual a cero. Por esto, llamaremos la ecuación (43.3) ecuación de *primer grado* respecto a las variables x, y.

Demostremos ahora que toda ecuación de primer grado (43.3) define una línea recta con relación al sistema fijado de coordenadas Oxy. Puesto que la ecuación (43.3) es de primer grado, entonces de las constantes A, B aunque una es distinta de cero. Por consiguiente, la ecuación (43.3) tiene al menos una solución x_0, y_0, por ejemplo,

$$x_0 = -\frac{AC}{A^2+B^2}, \qquad y_0 = -\frac{BC}{A^2+B^2},$$

siendo en este caso

$$Ax_0 + By_0 + C = 0.$$

Al sustraer de la ecuación (43.3) la identidad dada, obtendremos la ecuación

$$A(x - x_0) + B(y - y_0) = 0,$$

equivalente a la ecuación (43.3). Mas, esto significa que cualquier punto $M(x, y)$, cuyas coordenadas satisfacen la ecuación (43.3), se halla en una recta que pasa por el punto $M_0(x_0, y_0)$ y es perpendicular al vector (43.1).

Así pues, si tenemos un sistema fijado de coordenadas en un plano, cualquier ecuación de primer grado define una recta y las coordenadas de los puntos de cada recta satisfacen una ecuación de primer grado. La ecuación (43.3) se denomina *ecuación general de la línea recta en un plano*; el vector n de (43.1) se llama vector *normal* de la recta.

Las investigaciones de un plano en un espacio se realizan sin introducir los cambios importantes. Fijemos un sistema rectangular cartesiano de coordenadas $Oxyz$ y consideremos el plano π. Tomemos de nuevo un vector no nulo

$$n = (A, B, C), \tag{43.4}$$

perpendicular a π. Repitiendo los razonamientos anteriores llegamos a la conclusión que todos los puntos $M(x, y, z)$ del plano π, y sólo ellos, satisfacen la ecuación

$$Ax + By + Cz + D = 0, \tag{43.5}$$

que se llamará también ecuación de *primer grado* respecto de las variables x, y, z.

Al examinar nuevamente la ecuación arbitraria de primer grado (43.5), descubriremos que ésta tiene también por lo menos una solución x_0, y_0, z_0, por ejemplo,

$$x_0 = -\frac{AD}{A^2+B^2+C^2}, \qquad y_0 = -\frac{BD}{A^2+B^2+C^2},$$

$$z_0 = -\frac{CD}{A^2+B^2+C^2}.$$

A continuación establecemos que cualquier punto $M(x, y, z)$, cuyas coordenadas satisfacen la ecuación dada (43.5), se dispone en un plano que pasa por el punto $M_0(x_0, y_0, z_0)$ y es perpedicular al vector (43.4).

De este modo, si se fija un sistema de coorenadas en un espacio, toda ecuación de primer grado define un plano y las coordenadas de los puntos de cualquier plano satisfacen la ecuación de primer grado. La ecuación (43.5) se llama *ecuación general del plano en el espacio*; el vector n de (43.4) se denomina vector *normal* del plano.

Veamos ahora cómo se relacionan dos ecuaciones generales que definen una misma línea recta o un plano. Para concretar, sean dadas dos ecuaciones del plano π

$$A_1x + B_1y + C_1z + D_1 = 0,$$
$$A_2x + B_2y + C_2z + D_2 = 0. \tag{43.6}$$

Los vectores

$$n_1 = (A_1, B_1, C_1), \qquad n_2 = (A_2, B_2, C_2)$$

son perpendiculares a un mismo plano π y, por ende, son colineales. Como que son, además, no nulos, existe un número t tal que, por ejemplo,

$$n_1 = tn_2$$

o bien

$$A_1 = tA_2, \qquad B_1 = tB_2, \qquad C_1 = tC_2. \tag{43.7}$$

Multiplicando la segunda ecuación de (43.6) por t y restando de ella la primera, en virtud de las correlaciones (43.7), obtendremos

$$D_1 = tD_2.$$

Por consiguiente, *los coeficientes de las ecuaciones generales que definen una misma recta o un plano son proporcionales.*

Una ecuación general se llama *completa*, si todos sus coeficientes son distintos de cero. Una ecuación no completa se llama *incompleta*. Examinemos la ecuación completa de la recta (43.3). Como todos los coeficientes son distintos de cero, la ecuación puede escribirse en la siguiente forma:

$$\frac{x}{-\frac{C}{A}} + \frac{y}{-\frac{C}{B}} = 1.$$

Al designar

$$a = -\frac{C}{A}, \quad b = -\frac{C}{B},$$

obtendremos una ecuación nueva de la recta

$$\frac{x}{a} + \frac{y}{b} = 1.$$

Esta ecuación lleva el nombre de *ecuación de la recta "en segmentos"*. Los números a, b tienen un sentido geométrico muy simple. Ellos expresan las magnitudes de los segmentos cortados por la recta en los semiejes de coordenadas (fig. 43.1). Por supuesto, la ecuación completa de un plano puede reducirse a una forma análoga

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1.$$

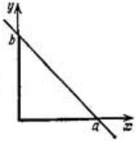


Fig. 43.1.

Las diferentes ecuaciones incompletas definen ciertos casos particulares de la disposición de la recta o del plano. Es útil recordarlas, pues, se encuentran con frecuencia. Por ejemplo, cuando $C = 0$, la ecuación (43.3) define la recta que pasa por el origen de coordenadas; cuando $B = C = 0$, la recta coincide con el eje Oy, etc. Cuando $A = 0$, la ecuación (43.5) define un plano paralelo al eje Ox; cuando $A = B = D = 0$, el plano coincide con el plano coordenado Oxy, etc.

Todo vector no nulo, paralelo a una línea recta, se llamará *vector director*. Consideremos, por ejemplo, el caso de un espacio y hallemos la ecuación de la recta que pasa por un punto dado M_0 (x_0, y_0, z_0) y tiene un vector director dado

$$q = (l, m, n).$$

Evidentemente, el punto M (x, y, z) se ubica en la recta citada cuando, y sólo cuando, los vectores $\overrightarrow{M_0M}$ y q son colineales, es decir, cuando, y sólo cuando, las coordenadas de estos vectores son proporcionales, es decir,

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}. \tag{43.8}$$

Estas son las ecuaciones buscadas de la recta. Se denominan comúnmente *ecuaciones canónicas de la recta*. Está claro que en el caso de la recta en un plano la ecuación tendrá la forma:

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}, \tag{43.9}$$

si la recta pasa por el punto M_0 (x_0, y_0) y tiene un vector director $q = (l, m)$.

De las ecuaciones canónicas se obtienen con facilidad las ecuaciones de la recta que pasa por dos puntos prefijados M_0, M_1. Con este fin es suficiente tomar, como vector director, el vector $\overrightarrow{M_0M_1}$, expresando sus coordenadas en términos de las coordenadas de los puntos M_0, M_1 y sustituyéndolas en las ecuaciones (43.8), (43.9).

Por ejemplo, en el caso de una recta en el plano tendremos la ecuación siguiente:

$$\frac{x-x_0}{x_1-x_0}=\frac{y-y_0}{y_1-y_0},$$

y en el caso de un espacio:

$$\frac{x-x_0}{x_1-x_0}=\frac{y-y_0}{y_1-y_0}=\frac{z-z_0}{z_1-z_0}.$$

Observemos que en las ecuaciones canónicas de una recta los denominadores pueden resultar ser nulos. Por esto, toda proporción $a/b = c/d$ se entenderá en lo sucesivo como la igualdad $ad = bc$. Por consiguiente, la anulación de una de las coordenadas del vector director significa la reducción a cero del numerador correspondiente en las ecuaciones canónicas.

Con el objeto de representar una recta analíticamente, las coordenadas de sus puntos se escriben, a menudo, como funciones de cierto parámetro auxiliar t. Tomemos por t cada una de las razones iguales en (43.8) y (43.9). Entonces, en el caso de un espacio tendremos las siguientes ecuaciones para una recta:

$$\begin{aligned} x &= x_0 + lt, \\ y &= y_0 + mt, \\ z &= z_0 + nt \end{aligned} \tag{43.10}$$

y las ecuaciones análogas en el caso de una recta en el plano

$$\begin{aligned} x &= x_0 + lt, \\ y &= y_0 + mt. \end{aligned} \tag{43.11}$$

Las últimas se denominan *ecuaciones paramétricas de una recta.* Asignando al parámetro t diferentes valores, obtendremos diferentes puntos de la recta.

Grandes posibilidades y comodidades en la inscripción de diferentes ecuaciones de una recta y de un plano ofrece el uso del concepto de determinante. Deduzcamos, por ejemplo, la ecuación de un plano que pasa por tres diferentes puntos no dispuestos en una misma recta. Así pues, sean dados tres puntos $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ $M_3(x_3, y_3, z_3)$. Como estos puntos no se hallan en una misma recta, entonces los vectores

$$\overrightarrow{M_1M_2}=(x_2-x_1,\ y_2-y_1,\ z_2-z_1),\ \overrightarrow{M_1M_3}=(x_3-x_1,\ y_3-y_1,\ z_3-z_1)$$

no son colineales. Por esta razón el punto $M(x, y, z)$ se halla en un mismo plano con los puntos M_1, M_2, M_3, cuando, y sólo cuando, los vectores $\overrightarrow{M_1M_2}$, $\overrightarrow{M_1M_3}$ y

$$\overrightarrow{M_1M}=(x-x_1,\ y-y_1,\ z-z_1)$$

son coplanares, es decir, cuando, y sólo cuando, el determinante compuesto de sus coordenadas es igual a cero. Por lo tanto, la ecuación

$$\det \begin{pmatrix} x-x_1 & y-y_1 & z-z_1 \\ x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \end{pmatrix} = 0 \qquad (43.12)$$

es la ecuación del plano buscado que pasa por los tres puntos dados.

Consideremos, por fin, la ecuación de la recta en un espacio la cual es perpendicular a dos rectas no paralelas y pasa por un punto dado. Supongamos que ambas rectas están dadas por sus ecuaciones canónicas

$$\frac{x-x_1}{l_1} = \frac{y-y_1}{m_1} = \frac{z-z_1}{n_1},$$

$$\frac{x-x_2}{l_2} = \frac{y-y_2}{m_2} = \frac{z-z_2}{n_2}.$$

El vector director q de la recta buscada debe ser perpendicular a dos vectores

$$q_1 = (l_1, m_1, n_1), \qquad q_2 = (l_2, m_2, n_2).$$

Estos vectores no son colineales y, por tanto, a título de q puede tomarse, por ejemplo, el producto vectorial $[q_1, q_2]$. Recordando la expresión de las coordenadas de un producto vectorial en términos de las coordenadas de los factores y, usando para su inscripción los determinantes de segundo orden, obtenemos

$$q = \left(\det \begin{pmatrix} m_1 n_1 \\ m_2 n_2 \end{pmatrix}, \quad \det \begin{pmatrix} n_1 l_1 \\ n_2 l_2 \end{pmatrix}, \quad \det \begin{pmatrix} l_1 m_1 \\ l_2 m_2 \end{pmatrix}\right).$$

Si la recta buscada pasa por el punto $M_0 (x_0, y_0, z_0)$, las ecuaciones canónicas de la recta serán:

$$\frac{x-x_0}{\det \begin{pmatrix} m_1 n_1 \\ m_2 n_2 \end{pmatrix}} = \frac{y-y_0}{\det \begin{pmatrix} n_1 l_1 \\ n_2 l_2 \end{pmatrix}} = \frac{z-z_0}{\det \begin{pmatrix} l_1 m_1 \\ l_2 m_2 \end{pmatrix}}.$$

Naturalmente, desde el punto de vista de principio, muchas deducciones respecto a la ecuación de una recta o de un plano quedan en vigor, cuando se emplea cualquier sistema afín de coordenadas. Nuestro deseo de utilizar los sistemas rectangulares cartesianos de coordenadas se debe, en lo principal, a que los razonamientos, en este último caso, son más simples.

Ejercicios.

1. Escríbase la ecuación de la recta en un plano, que pasa por dos puntos dados, utilizando el determinante de segundo orden. Compárese con (43.12).
2. ¿Será justa la afirmación de que la ecuación (43.12) representa siempre la ecuación de un plano?

3. Escríbase, por analogía con las ecuaciones (43.10), las ecuaciones paramétricas de un plano en un espacio. ¿Cuántos parámetros deben contener estas ecuaciones?

4. Hállense las coordenadas del vector normal de un plano que pasa por tres puntos dados no dispuestos en una misma recta.

5. ¿Qué representa en sí el lugar geométrico de los puntos en un espacio cuyas coordenadas son las soluciones de un sistema de dos ecuaciones lineales algebraicas con tres incógnitas?

§ 44. Disposición conjunta

Cuando se consideran simultáneamente varias rectas y varios planos surgen diferentes problemas relacionados, en primer lugar, con la necesidad de determinar su disposición mutua.

Supongamos que en un espacio dos planos que se cortan están dados por sus ecuaciones generales

$$A_1x + B_1y + C_1z + D_1 = 0,$$
$$A_2x + B_2y + C_2z + D_2 = 0.$$

Estos planos forman dos ángulos adyacentes que en suma dan dos rectas. Hallemos uno de ellos. Los vectores $n_1 = (A_1, B_1, C_1)$ y $n_2 = (A_2, B_2, C_2)$ son normales y, por ende, la determinación del ángulo entre los planos se reduce a la determinación del ángulo φ entre los vectores n_1, n_2. Conforme a (25.5) tenemos

$$\cos\varphi = \frac{A_1A_2 + B_1B_2 + C_1C_2}{(A_1^2 + B_1^2 + B_1^2)^{1/2}\,(A_2^2 + B_2^2 + C_2^2)^{1/2}}.$$

Por analogía completa se deduce la fórmula para hallar el ángulo entre dos rectas en un plano, dadas por sus ecuaciones generales

$$A_1x + B_1y + C_1 = 0,$$
$$A_2x + B_2y + C_2 = 0.$$

Uno de los ángulos φ, formados por dichas rectas, se calcula según la fórmula

$$\cos\varphi = \frac{A_1A_2 + B_1B_2}{(A_1^2 + B_1^2)^{1/2}\,(A_2^2 + B_2^2)^{1/2}}.$$

La condición de paralelismo de unas rectas, dadas por sus ecuaciones generales, es la condición de carácter colineal de los vectores normales, es decir, la condición de proporcionalidad de sus coordenadas

$$\frac{A_1}{A_2} = \frac{B_1}{B_2}.$$

La condición de perpendicularidad de las rectas coincide con la condición $\cos\varphi = 0$, o bien, lo que es igual, con la condición

$$A_1A_2 + B_1B_2 = 0.$$

Por supuesto, una forma análoga la tiene la condición

$$\frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}$$

de paralelismo de los planos y la condición

$$A_1A_2+B_1B_2+C_1C_2=0$$

de perpendicularidad de los planos dados también mediante las ecuaciones generales.

Supongamos ahora que dos rectas en un espacio, por ejemplo, están dadas mediante las ecuaciones canónicas

$$\frac{x-x_1}{l_1}=\frac{y-y_1}{m_1}=\frac{z-z_1}{n_1},$$

$$\frac{x-x_2}{l_2}=\frac{y-y_2}{m_2}=\frac{z-z_2}{n_2}.$$

Puesto que los vectores $q_1=(l_1, m_1, n_1)$, $q_2=(l_2, m_2, n_2)$ son los vectores directores para estas rectas, concluimos de nuevo que uno de los ángulos φ entre las rectas coincidirá con el ángulo entre los vectores q_1, q_2. Por consiguiente,

$$\cos\varphi=\frac{l_1l_2+m_1m_2+n_1n_2}{(l_1^2+m_1^2+n_1^2)^{1/2}\,(l_2^2+m_2^2+n_2^2)^{1/2}}.$$

Correspondientemente, la proporcionalidad de las coordenadas

$$\frac{l_1}{l_2}=\frac{m_1}{m_2}=\frac{n_1}{n_2}$$

es la condición de paralelismo de las rectas y la igualdad

$$l_1l_2+m_1m_2+n_1n_2=0$$

es la condición de perpendicularidad de las mismas.

Está claro que si las rectas y los planos vienen dados mediante un método en el que se indica de modo explícito el vector director o el normal, entonces la determinación del ángulo entre las rectas y los planos se reduce siempre a la del ángulo entre dichos vectores. Supongamos, por ejemplo, que en un espacio se han dado el plano π, mediante su ecuación general

$$Ax+By+Cz+D=0,$$

y la recta L, mediante la ecuación canónica

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}.$$

Puesto que el ángulo φ entre la recta y el plano es complementario al ángulo ψ entre el vector director de la recta y el vector normal del plano (fig. 44.1), entonces

$$\operatorname{sen}\varphi=\frac{|Al+Bm+Cn|}{(A^2+B^2+C^2)^{1/2}\,(l^2+m^2+n^2)^{1/2}}.$$

Es evidente la condición

$$Al+Bm+Cn=0$$

de paralelismo entre la recta y el plano y la condición

$$\frac{A}{l}=\frac{B}{m}=\frac{C}{n}$$

de perpendicularidad de la recta respecto del plano.

La forma de ecuaciones generales, mediante la cual se dan una recta y un plano, permite resolver con toda eficacia un importante problema, a saber, el cálculo de la distancia desde un punto hasta la recta, o bien desde un punto hasta el plano. La deducción de las fórmulas es, en ambos casos, enteramente análoga y nos limitamos de nuevo a la consideración detallada de una sola de ellas.

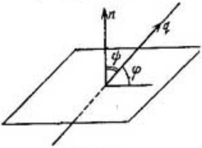


Fig. 44.1.

Supongamos que el plano π en el espacio está dado por su ecuación general (43.5). Tomemos un punto arbitrario $M_0(x_0, y_0, z_0)$. Del punto M_0 tracemos una perpendicular al plano y designemos con $M_1(x_1, y_1, z_1)$ la base de ésta. Está claro que la distancia $\rho(M_0, \pi)$ del punto M_0 al plano es igual a la longitud del vector $\overrightarrow{M_0M_1}$.

Los vectores $n=(A, B, C)$ y $\overrightarrow{M_0M_1}=(x_1-x_0,\ y_1-y_0,\ z_1-z_0)$ son perpendiculares al mismo plano, razón por la cual son colineales. Por eso, existe un número t tal que $\overrightarrow{M_0M_1}=tn$, es decir,

$$x_1-x_0=tA,$$
$$y_1-y_0=tB,$$
$$z_1-z_0=tC.$$

El punto $M_1(x_1, y_1, z_1)$ se halla en el plano π. Al expresar sus coordenadas a base de las correlaciones obtenidas y sustituirlas en la ecuación del plano, tenemos

$$t=-\frac{Ax_0+By_0+Cz_0+D}{A^2+B^2+C^2}.$$

Pero, la longitud del vector n es igual a $(A^2 + B^2 + C^2)^{1/2}$ y, por ende, $|\overrightarrow{M_0M_1}| = |t|(A^2 + B^2 + C^2)^{1/2}$. Por consiguiente

$$\rho(M_0, \pi) = \frac{|Ax_0 + By_0 + Cz_0 + D|}{(A^2+B^2+C^2)^{1/2}}.$$

En particular,

$$\rho(0, \pi) = \frac{|D|}{(A^2+B^2+C^2)^{1/2}}.$$

A la par con la ecuación general (43.5) del plano consideraremos, además, las siguientes ecuaciones suyas

$$\pm(A^2+B^2+C^2)^{-1/2}(Ax+By+Cz+D)=0.$$

De los dos signos posibles en el primer miembro elijamos el opuesto al signo de D. Si es que $D = 0$, elijamos un signo cualquiera. Entonces, el término independiente de esta ecuación será un número no positivo $-p$, mientras que los coeficientes de x, y, z serán cosenos de los ángulos entre el vector normal y los ejes de coordenadas. La ecuación

$$x \cos\alpha + y \cos\beta + z \cos\gamma - p = 0 \tag{44.1}$$

se llama *ecuación normalizada del plano*. Es evidente que

$$\rho(M_0, \pi) = |x_0 \cos\alpha + y_0 \cos\beta + z_0 \cos\gamma - p|,$$
$$\rho(0, \pi) = p.$$

La distancia $\rho(M_0, L)$ desde el punto $M_0(x_0, y_0)$ hasta la recta L en el plano, dada mediante su ecuación general (43.3), se determina por la fórmula análoga

$$\rho(M_0, L) = \frac{|Ax_0+By_0+C|}{(A^2+B^2)^{1/2}}.$$

La ecuación normalizada de una línea recta tiene por expresión

$$x \cos\alpha + y \operatorname{sen}\alpha - p = 0. \tag{44.2}$$

Aquí, α es el ángulo formado por el vector normal con el eje Ox.

Ejercicios.

1. ¿En qué condiciones, impuestas en las coordenadas de los vectores normales, dos rectas en un plano (tres planos en un espacio) se intersecan en un punto?
2. ¿Bajo qué condición la recta (43.8) pertenece al plano (43.5)?
3. ¿Bajo qué condición dos rectas en un espacio, dadas mediante ecuaciones canónicas, pertenecen a un plano?
4. Calcúlense los ángulos entre la diagonal de un cubo y las caras de éste.
5. Dedúzcase la fórmula para calcular la distancia entre un punto y una recta en un espacio, dada esta última mediante sus ecuaciones canónicas.

§ 45. Plano en el espacio lineal

Ya hemos subrayado más de una vez que una recta y un plano que pasan por el origen de coordenadas se identifican en los espacios de los segmentos dirigidos con la imagen geométrica de un subespacio. Pero según sus propiedades, difieren poco de cualesquiera otras rectas y otros planos que se obtienen por traslación paralela o por desplazamiento de estos subespacios. Deseando generalizar este hecho para cualesquiera espacios lineales, llegamos al concepto de plano en el espacio lineal.

Sea L cierto subespacio del espacio lineal K. Fijemos en K un vector arbitrario x_0. En particular, este vector puede pertenecer a L. El conjunto H de vectores z que se obtienen según la fórmula

$$z = x_0 + y, \tag{45.1}$$

donde y es cualquier vector de L, se denomina *plano* en el espacio lineal K. El vector x_0 se llama *vector de desplazamiento* y el subespacio L, *director*. En cuanto al plano H, diremos que está formado por el desplazamiento del subespacio L al vector x_0.

El concepto formal de un plano incluye en sí las nociones de rectas y planos (¡en la interpretación vectorial!) en los espacios de segmentos dirigidos. Pero, por ahora no sabemos si posee las propiedades análogas.

Todo vector del plano H se representa del modo *único* en forma de la suma (45.1). Si $z = x_0 + y$ y $z = x_0 + y'$, donde $y, y' \in L$, resulta que $y = y'$. De la fórmula (45.1) se deduce, además, que la diferencia entre dos vectores cualesquiera del plano H pertenece al subespacio L.

Elijamos en el plano H un vector arbitrario z_0. Sea $z_0 = x_0 + y_0$. Representemos la igualdad (45.1) en la forma:

$$z = z_0 + (y - y_0).$$

Los conjuntos de vectores y e $y - y_0$ describen el mismo subespacio L. La última igualdad significa, por ello, que el plano H puede obtenerse por desplazamiento del subespacio L a *todo* vector fijado del mismo plano.

El plano H es un conjunto de vectores de K generado por el subespacio L y el vector de desplazamiento x_0, de conformidad con (45.1). Una circunstancia de mucha importancia consiste en que cualquier plano puede ser generado por un solo subespacio. Supongamos que no es así, es decir, que existen un subespacio director más L', y un vector de desplazamiento x_0', que forman el mismo plano H. En tal caso para todo $z \in H$ tenemos $z = x_0 + y$, donde $y \in L$, y al mismo tiempo $z = x_0' + y'$, donde $y' \in L'$. De aquí se desprende que el subespacio L' es un conjunto de vectores de K, definidos mediante la fórmula

$$y' = (x_0 - x_0') + y.$$

Como el vector nulo pertenece a L', de la última fórmula se infiere que el vector $(x_0 - x_0')$ pertenece a L. Mas, esto significa que el subespacio L' se compone de los mismos vectores que el subespacio L.

Ya hemos indicado que el vector de desplazamiento se define mediante un plano y, además de un modo *no unívoco*. No obstante, en este caso también la cuestión acerca de la univocidad puede resolverse de una manera bien natural.

Convengamos en considerar que en el espacio lineal K se ha introducido un producto escalar. Si en lugar del vector x_0 tomamos $\mathrm{ort}_L x_0$, está claro que obtendremos el mismo plano. Por ello, sin restringir la generalidad, se puede considerar que $x_0 \perp L$. El vector x_0 en este caso se llamará vector *ortogonal* de desplazamiento. Ahora podemos demostrar que todo plano es generado sólo por *un* vector de desplazamiento.

Efectivamente, supongamos que existen dos vectores de desplazamiento x_0', x_0'', que son ortogonales al subespacio L y generan, sin embargo, un mismo plano H. Entonces, para cualquier vector $y' \in L$ debe existir un vector $y'' \in L$ tal que $x_0' + y' = x_0'' + y''$. De aquí se deduce que $x_0' - x_0'' \in L$. Pero, según la hipótesis, $x_0' - x_0'' \perp L$. Por consiguiente, $x_0' - x_0'' = 0$, es decir, $x_0' = x_0''$. Esto significa, en particular, que en un espacio dotado de un producto escalar cualquier plano cuenta sólo con un vector ortogonal al subespacio director.

Dos planos se denominan *paralelos*, si el subespacio director de uno de ellos forma parte del subespacio director del otro.

Esta afirmación se justifica con facilidad. Cualesquiera dos planos paralelos H_1, H_2 *o bien no contienen ningún vector común* o bien uno de ellos forma parte del otro. Supongamos que H_1, H_2 tienen un vector común z_0. Puesto que todo plano puede ser obtenido por desplazamiento del subespacio director a cualquier vector suyo, entonces tanto H_1 como H_2 se obtienen por desplazamiento de los subespacios correspondientes al vector z_0. Pero uno de los subespacios forma parte del otro, por lo cual uno de los planos forma parte del otro.

Un subespacio es un caso particular del plano. Es evidente que el subespacio L es paralelo a todo plano H, obtenido por desplazamiento de L a cierto vector x_0. De la propiedad demostrada para los planos paralelos se infiere que H coincide con L cuando, y sólo cuando, $x_0 \in L$.

Consideraremos ahora dos planos no paralelos H_1, H_2. Estos o bien no tienen ningún vector común o bien tienen el vector común.

En el primer caso los planos H_1, H_2 se llamarán planos *que se cruzan* en el segundo caso llamémoslos planos *que se intersecan.*

Al igual que en el caso de subespacios, un conjunto de vectores pertenecientes simultáneamente a H_1 y H_2 se denominará *intersección* de estos planos y se indicará $H_1 \cap H_2$. Supongamos que el plano H_1 se ha formado por desplazamiento del subespacio L_1 y el plano H_2, por desplazamiento del subespacio L_2. Designaremos

$$H = H_1 \cap H_2, \qquad L = L_1 \cap L_2.$$

TEOREMA 45.1. *Si la intersección H contiene el vector z_0, representa en sí un plano formado por desplazamiento de la intersección L a dicho vector.*

DEMOSTRACION. Según la hipótesis del teorema, existe un vector z_0 perteneciente a la intersección H. Supongamos que existe un vector más $z_1 \in H$. Representémoslo en la forma:

$$z_1 = z_0 + (z_1 - z_0).$$

Ahora, de la sucesión de correlaciones

$$z_1, \quad z_0 \in H \rightarrow z_1, \quad z_0 \in H_1; \quad z_1 z_0 \in H_2 \rightarrow z_1 - z_0 \in L_1;$$
$$z_1 - z_0 \in L_2 \rightarrow z_1 - z_0 \in L$$

conluimos que todo vector de la intersección H puede ser representado como la suma del vector z_0 y cierto vector de la intersección L.

Tomemos, luego, un vector arbitrario f de L. Se tiene

$$f \in L \rightarrow f \in L_1; \quad f \in L_2 \rightarrow z_0 + f \in H_1;$$
$$z_0 + f \in H_2 \rightarrow z_0 + f \in H,$$

es decir, todo vector del subespacio L desplazado al vector z_0, pertenece a la intersección H. El teorema queda demostrado.

Un plano no ha de ser forzosamente un subespacio. No obstante, se le puede atribuir *una dimensión* igual a la del subespacio director. Un plano de dimensión nula contiene sólo un vector: el vector de desplazamiento. Al determinar la dimensión de la intersección de los planos será útil el teorema 19.1. De los teoremas 19.1, 45.1 se desprende que la dimensión de la intersección H no es superior a la mínima de las dimensiones H_1, H_2.

Si en un espacio de segmentos dirigidos vienen dados dos (tres) vectores, entonces, al imponer ciertas condiciones adicionales, podemos construir sólo un plano de dimensión 1 (2) que contenga los vectores dados. Dichas condiciones adicionales pueden enunciarse del modo siguiente. Los dos vectores dados no deben ser coincidentes, es decir, no deben pertenecer a un plano de dimensión nula. Los tres vectores dados no deben pertenecer a un plano de dimensión uno.

Hechos análogos tienen lugar en un espacio lineal arbitrario.

Supongamos que en un espacio lineal están dados los vectores $x_0, x_1, \ldots, x_k$. Diremos que estos vectores se encuentran en *la posición general*, si no pertenecen a un plano de dimensión $k-1$.

TEOREMA 45.2. *Si los vectores $x_0, x_1, \ldots, x_k$ se encuentran en la posición general, existe un único plano H de dimensión k que los contiene.*

DEMOSTRACIÓN. Examinemos los vectores $x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0$. Si fueran linealmente dependientes, pertenecerían a cierto subespacio de dimensión no superior a $k-1$. Por consiguiente, los propios vectores $x_0, x_1, \ldots, x_k$ pertenecerían a un plano obtenido por desplazamiento de este subespacio al vector x_0, lo que contradice la hipótesis del teorema.

Así pues, los vectores $x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0$ son linealmente independientes. Indiquemos con L su cápsula lineal. El subespacio L tiene una dimensión k. Al desplazarlo al vector x_0, obtendremos cierto plano H de la misma dimensión a la que pertenecen todos los vectores dados $x_0, x_1, \ldots, x_k$.

El plano construido H es único. En efecto, supongamos que los vectores $x_0, x_1, \ldots, x_k$ pertenecen a dos planos H_1, H_2, ambos de dimensión k. El plano no cambia, si el vector de desplazamiento se sustituye por cualquier otro vector del plano. Por ello, sin restringir la generalidad, se puede considerar que H_1, H_2 se han obtenido por desplazamiento respectivo de los subespacios L_1, L_2 a un mismo vector x_0. Pero en este caso resulta que ambos subespacios coinciden, puesto que tienen la misma dimensión k y contienen un mismo sistema linealmente independiente $x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0$.

Ejercicios.

1. Sean H_1, H_2 dos planos cualesquiera. Llamemos *suma* $H_1 + H_2$ *de los planos* H_1, H_2 el conjunto de todos los vectores del tipo $z_1 + z_2$, donde $z_1 \in H_1$, $z_2 \in H_2$. Demuéstrese que la suma de los planos es un plano.
2. Sea H un plano, y sea λ un número. Llamemos *producto* λH *del plano* H *por el número* λ el conjunto de todos los vectores del tipo λz, donde $z \in H$. Demuéstrese que este producto es un plano.
3. ¿Será espacio lineal el conjunto de todos los planos de un mismo espacio con las operaciones sobre ellos introducidas arriba?
4. Demuéstrese que los vectores $x_0, x_1, \ldots, x_k$ se encuentran en la posición general cuando, y sólo cuando, los vectores $x_1 - x_0, x_2 - x_0, \ldots, x_k - x_0$ son linealmente independientes.
5. Demuéstrese que un plano de dimensión k que contiene los vectores de posición general $x_0, x_1, \ldots, x_k$ es un subespacio cuando, y sólo cuando, estos vectores son linealmente dependientes.

§ 46. Recta e hiperplano

En el espacio lineal K de dimensión m existen dos clases de planos que ocupan una posición especial. Se trata de los planos de dimensión 1 y de los planos de dimensión

$m-1$. De acuerdo con la imagen geométrica en los espacios de segmentos dirigidos, todo plano de dimensión 1 se llama *línea recta*. Un plano de dimensión $m-1$ se llama *hiperplano*.

Consideremos una recta arbitraria H en el espacio lineal K. Designemos mediante x_0 el vector de desplazamiento y mediante q, el vector básico del subespacio director unidimensional. Supongamos que estos vectores están dados por sus coordenadas

$$x_0 = (x_1, x_2, \ldots, x_m),$$
$$q = (q_1, q_2, \ldots, q_m)$$

respecto de cierta base del espacio K. Evidentemente, todo vector z de la recta H puede ser dado en la forma:

$$z = x_0 + tq, \tag{46.1}$$

donde t es un número. Por esto la correlación (46.1) se puede considerar ecuación vectorial de la recta H en el espacio K. Si el vector z tiene en la misma base las coordenadas

$$z = (z_1, z_2, \ldots, z_m),$$

entonces, escribiendo la igualdad (46.1) según las coordenadas, obtendremos

$$\begin{aligned} z_1 &= x_1 + q_1 t, \\ z_2 &= x_2 + q_2 t, \\ &\cdots\cdots\cdots \\ z_m &= x_m + q_m t. \end{aligned} \tag{46.2}$$

Al comparar ahora estas ecuaciones con (43.10), (43.11), resulta natural denominarlas ecuaciones *paramétricas* de la recta H. Diremos que la recta H pasa por el vector x_0 y tiene el vector *director* q.

De acuerdo con el teorema 45.2, por cualesquiera dos vectores no coincidentes x_0, y_0 siempre se puede trazar una recta y, además, una sola. Supongamos que en cierta base del espacio K los vectores x_0, y_0 vienen dados por sus coordenadas

$$x_0 = (x_1, x_2, \ldots, x_m),$$
$$y_0 = (y_1, y_2, \ldots, y_m).$$

Como que a título de vector director puede tomarse, por ejemplo, el vector $y_0 - x_0$, entonces, de la ecuación (46.2) obtenemos las ecuaciones paramétricas

$$\begin{aligned} z_1 &= x_1 + (y_1 - x_1)\, t, \\ z_2 &= x_2 + (y_2 - x_2)\, t, \\ &\cdots\cdots\cdots\cdots \\ z_m &= x_m + (y_m - x_m)\, t \end{aligned} \tag{46.3}$$

de una recta que pasa por los dos vectores dados.

Cuando $t = 0$, estas ecuaciones definen el vector x_0; cuando $t = 1$, el vector y_0. Si el espacio K es real, el conjunto de vectores dados mediante las ecuaciones (46.3) para $0 \leqslant t \leqslant 1$ se llamará *segmento* que une los vectores x_0, y_0. Por supuesto, esta denominación está ligada con la imagen geométrica del conjunto dado en los espacios de segmentos dirigidos.

Supongamos que la recta H se corta con cierto plano. En tal caso, de acuerdo con el corolario del teorema 45.1, la intersección será o bien una recta o bien un vector. Si la intersección resulta ser una recta, ésta coincidirá, por supuesto, con la recta H. Pero esto significa que al cortarse una línea recta con un plano, la recta o bien se mantiene íntegramente en el plano o bien tiene con éste sólo un vector común.

La noción de hiperplano tiene sentido en todo espacio lineal, sin embargo la emplearemos sólo en los espacios dotados de producto escalar.

Examinemos un hiperplano arbitrario H. Supongamos que se ha formado por desplazamiento del subespacio $(m-1)$-dimensional L al vector x_0. El complemento ortogonal $L^{\perp}$ será, en este caso, un subespacio unidimensional. Indiquemos con n cualquiera de sus vectores básicos. El vector z pertenece al hiperplano H cuando, y sólo cuando, el vector $z - x_0$ pertenezca al subespacio L. A su vez, esta condición se cumple cuando, y sólo cuando, el vector $z - x_0$ sea ortogonal al vector n, es decir,

$$(n,\ z - x_0) = 0. \tag{46.4}$$

De este modo, hemos obtenido una ecuación la cual se satisface por todos los vectores del hiperplano H. Con el fin de definir un hiperplano en forma de tal ecuación, basta indicar cualquier vector n, ortogonal al subespacio director, y el vector de desplazamiento x_0.

El hecho de que la ecuación tenga una forma explícita permite simplificar, en grado considerable, diversas investigaciones. Sean dados los vectores $n_1, n_2, \ldots, n_k$ y $x_1, x_2, \ldots, x_k$. Investiguemos un plano R que es la intersección de los hiperplanos

$$\begin{aligned} (n_1,\ z - x_1) &= 0, \\ (n_2,\ z - x_2) &= 0, \\ &\cdots\cdots \\ (n_k,\ z - x_k) &= 0. \end{aligned} \tag{46.5}$$

Este problema puede considerarse como resolución del sistema de ecuaciones (46.5) respecto de los vectores z. Supongamos que la intersección de los hiperplanos no es vacía, es decir, que el sistema (46.5) tiene por lo menos una solución z_0. Entonces, como se sabe, el plano buscado se determina también por el siguiente sistema:

$$\begin{aligned} &(n_1, z-z_0)=0, \\ &(n_2, z-z_0)=0, \\ &\cdots\cdots\cdots \\ &(n_k, z-z_0)=0 \end{aligned} \tag{46.6}$$

en virtud de que ningún plano varía, si el vector de desplazamiento se sustituye por cualquier otro vector del plano.

El vector $y = z - z_0$ es un vector arbitrario de la intersección L de los subespacios directores de todos los k hiperplanos. Es evidente que los vectores y del subespacio L satisfacen el siguiente sistema:

$$\begin{aligned} &(n_1, y)=0, \\ &(n_2, y)=0, \\ &\cdots\cdots \\ &(n_k, y)=0. \end{aligned} \tag{46,7}$$

La intersección L, dada en forma del sistema (46.7), permite resolver con facilidad la cuestión sobre la dimensión de L. Según se desprende del mismo sistema, el subespacio L es un complemento ortogonal de la cápsula lineal del sistema de vectores $n_1, n_2, \ldots$ $\ldots, n_k$. Si r es el rango de dicho sistema, la dimensión de L y, por lo tanto, de R será igual a $m - r$, donde m es la dimensión del espacio. En particular, si los vectores $n_1, n_2, \ldots, n_k$ son linealmente independientes, la dimensión del plano (46.5) es $m - k$ En este caso se presupone, desde luego, que el plano existe, es decir, el sistema (46.5) tiene al menos una solución. Con el fin de prefijar el subespacio L, definido por el sistema (46.7), basta indicar su base, es decir, cualquier sistema de $m - r$ vectores linealmente independientes, ortogonales a los vectores $n_1, n_2, \ldots, n_k$.

La ecuación (46.4) del hiperplano puede escribirse también de una forma algo diferente. Designemos $(n, x_0) = b$, entonces la ecuación

$$(n, z) = b \tag{46.8}$$

definirá el mismo hiperplano que la ecuación (46.4). Hemos de notar que las ecuaciones generales (43.3), (43.5) de una recta y de un plano son, en esencia, las mismas ecuaciones. Resulta importante subrayar que *toda* ecuación del tipo (46.8) puede ser reducida al tipo (46.4), eligiendo de una manera adecuada el vector x_0. Para ello es suficiente, por ejemplo, tomar x_0 en la forma:

$$x_0 = \alpha n.$$

Al sustituir esta expresión en (46.4) y comparando con (46.8), concluimos que debe verificarse

$$\alpha = \frac{b}{(n, n)}$$

Ahora podemos hacer una deducción de que si un sistema del tipo (46.5) define cierto plano, el mismo plano puede ser definido también por el sistema del tipo siguiente:

$$\begin{aligned}(n_1, z) &= b_1,\\ (n_2, z) &= b_2,\\ &\dots\dots\\ (n_k, z) &= b_k\end{aligned} \qquad (46.9)$$

con los correspondientes números $b_1, b_2, \ldots, b_k$. Evidentemente, será lícita también la afirmación inversa. El sistema (46.9) define el mismo plano que el sistema (46.5) con los vectores correspondientes $x_1, x_2, \ldots, x_k$.

La recta y el plano son hiperplanos en los espacios V_2 y V_3, respectivamente. Anteriormente se ha establecido la relación que existe entre la distancia desde un punto hasta dichos hiperplanos y el resultado de la sustitución de las coordenadas del punto en las ecuaciones generales. Una relación análoga tiene lugar también en el caso de un hiperplano arbitrario.

Sea H un hiperplano dado mediante la ecuación (46.4). Como hasta ahora, designemos con $\rho(v, H)$ la distancia entre el vector v y H. Tomando en consideración la ecuación (46.4), obtenemos que

$$(n, v - x_0) = (n, v - x_0) - (n, z - x_0) = (n, v - z)$$

para cualquier vector z de H. De conformidad con la desigualdad de Cauchy—Buniakovski

$$|(n, v - x_0)| \leqslant |n|\,|v - z|. \qquad (46.10)$$

Por esto

$$|v - z| \geqslant \frac{|(n, v - x_0)|}{|n|}$$

Si mostramos que en el hiperplano H existe un vector z^* y, además, sólo uno, para el cual la desigualdad (46.10) se convierte en una igualdad, esto será indicio de que, en primer lugar,

$$\rho(v, H) = \frac{|(n, v - x_0)|}{|n|} \qquad (46.11)$$

y, en segundo lugar, el valor $\rho(v, H)$ se logra en un solo vector z^*.

Designemos con L el subespacio director del hiperplano H. Está claro que todo vector, ortogonal a L, será colineal con n, y viceversa. La desigualdad (46.10) se convierte en igualdad cuando, y sólo cuando, los vectores n y $v - z$ son colineales, es decir, $v - z = \alpha n$ para cierto número α. Supongamos que la igualdad se verifica para dos vectores z_1, z_2 de H, es decir,

$$v - z_1 = \alpha_1 n,$$
$$v - z_2 = \alpha_2 n.$$

De aquí se desprende que

$$z_1 - z_2 = (\alpha_2 - \alpha_1)\, n.$$

Por consiguiente, $z_1 - z_2 \perp L$. Pero $z_1 - z_2 \in L$ como la diferencia entre dos vectores del hiperplano. Por ello $z_1 - z_2 = 0$ ó $z_1 = z_2$.

Designemos mediante z_0 un vector de H ortogonal a L. Como se sabe, este vector existe y es único. Escribamos el vector z en forma de la suma

$$z = z_0 + y,$$

donde $y \in L$. Representemos el vector v en la forma

$$v = f + s,$$

donde $f \in L$, $s \perp L$. Ahora

$$v - z = (s - z_0) + (f - y).$$

Si hacemos $z = z_0 + f$,

$$v - z = s - z_0.$$

El vector $h = s - z_0$ es ortogonal a L y la fórmula (46.11) queda establecida.

Al mismo tiempo hemos demostrado que todo vector v del espacio puede ser representado de modo *único* en forma de la suma

$$v = z + h,$$

donde el vector z pertenece al hiperplano H y el vector h es ortogona al subespacio director L. Por analogía con los espacios de segmentos dirigidos, el vector z en esta descomposición se denomina *proyección* del vector v sobre el hiperplano H; h es *la perpendicular* trazada del vector v sobre H. El proceso en el que el vector z se obtiene de v se llama *proyección* de v sobre H. Si el hiperplano está dado por la ecuación (46.4), el vector n se llama vector *normal* del hiperplano. Para los vectores dados x_0 y n existe *un único* hiperplano que contiene el vector x_0 y es ortogonal al vector n.

Ejercicios.

1. Demuéstrese que cualquier plano distinto de todo el espacio puede ser dado como una intersección de los hiperplanos (46.9).
2. Demuéstrese que la suma de hiperplanos será un hiperplano cuando, y sólo cuando, los hiperplanos sumados son paralelos.
3. Demuéstrese que el producto de un hiperplano por un número no nulo es un hiperplano.
4. ¿Cuáles son las condiciones de paralelismo entre una recta y un hiperplano?
5. Dedúzcase la formula de la distancia entre un vector y una recta dada mediante la ecuación (46.1).

§ 47. Semiespacio

Con las nociones de recta e hiperplano de un cuerpo está relacionada la noción de los así llamados conjuntos convexos. Como que estos conjuntos son de amplio uso en las más diversas ramas de la matemática, fijemos nuestra atención en la investigación de una parte de ellos.

Un conjunto de vectores de un espacio lineal real se denomina *convexo*, si junto con cada dos vectores contiene también todo el segmento que los une.

Como conjuntos convexos pueden intervenir, por ejemplo, un vector, un segmento, una recta, un subespacio, un plano, un hiperplano y varios más.

Supongamos que el hiperplano en un espacio real viene dado mediante la ecuación

$$(n, z) - b = 0.$$

Un conjunto de vectores z que satisfacen la desigualdad

$$(n, z) - b < 0 \tag{47.1}$$

ó

$$(n, z) - b > 0. \tag{47.2}$$

se llama *semiespacio abierto*. El semiespacio (47.1) se denomina *negativo* y el (47.2), *positivo*.

TEOREMA 47.1. *Un semiespacio es un conjunto convexo.*

DEMOSTRACION. Tomemos dos vectores x_0, y_0 y designemos

$$\Phi_1 = (n, x_0) - b, \qquad \Phi_2 = (n, y_0) - b.$$

Si z es un vector cualquiera de una recta que pasa por x_0, y_0, entonces

$$z = x_0 + t\,(y_0 - x_0).$$

Para $0 \leqslant t \leqslant 1$ obtenemos los vectores del segmento que une x_0, y_0. Tenemos

$$(n, z) - b = \Phi_1\,(1 - t) + \Phi_2 t. \tag{47.3}$$

Si Φ_1 y Φ_2 tienen signos idénticos, es decir, si los vectores x_0, y_0 pertenecen a un mismo semiespacio, el segundo miembro de la correlación (47.3) será del mismo signo que Φ_1 y Φ_2 para todos los valores de t que satisfacen las desigualdades $0 \leqslant t \leqslant 1$.

De este modo, todo hiperplano divide el espacio lineal en tres conjuntos convexos disjuntos, a saber, el propio hiperplano y dos semiespacios abiertos.

Supongamos que los vectores x_0, y_0 pertenecen a diferentes semiespacios, es decir, Φ_1 y Φ_2 tienen signos opuestos. La transformación formal de la correlación (47.3) conduce a la desigualdad siguiente:

$$(n,\ z)-b=(\Phi_2-\Phi_1)\left(t-\frac{1}{1-\Phi_2/\Phi_1}\right).$$

De aquí se infiere que una recta que pasa por los vectores x_0, y_0 corta el hiperplano. La intersección se determina por el valor

$$t=\frac{1}{1-\Phi_2/\Phi_1},$$

que satisface las desigualdades $0\leqslant t\leqslant 1$. Así pues,

Si dos vectores pertenecen a diferentes semiespacios el segmento que une dichos vectores corta el hiperplano que define los semiespacios.

Teniendo presente la propiedad enunciada, es fácil comprender qué representan en sí los semiespacios en los espacios de segmentos dirigidos. En un plano los extremos de los vectores del semiespacio se disponen de un lado de la línea recta; en un espacio, de un lado del plano.

A la par con los semiespacios abiertos se consideran, con frecuencia, *los cerrados*. Se definen como conjuntos de los vectores z que satisfacen la desigualdad

$$(n,\ z)-b\leqslant 0 \qquad (47.4)$$

ó

$$(n,\ z)-b\geqslant 0. \qquad (47.5)$$

El semiespacio (47.4) se llama *no positivo*, el (47.5) *no negativo*. Por supuesto, los semiespacios cerrados son también conjuntos convexos.

TEOREMA 47.2. *Una intersección de conjuntos convexos es un conjunto convexo.*

DEMOSTRACIÓN. Es suficiente, evidentemente, examinar el caso de dos conjuntos U_1, U_2. Sea $U=U_1\cap U_2$ su intersección. Tomemos cualesquiera dos vectores x_0, y_0 de U y designemos mediante S el segmento que los une. Los vectores x_0, y_0 pertenecen tanto a U_1 como a U_2. Por ello, por ser los conjuntos U_1, U_2 convexos, el segmento S pertenece íntegramente tanto a U_1 como a U_2, es decir, el segmento S pertenece a la intersección U.

El teorema demostrado es de gran importancia, cuando se estudian los conjuntos convexos. En particular, permite afirmar que un conjunto no vacío de vectores z que satisfacen simultáneamente el sistema de desigualdades

$$\begin{gathered}(n_1,\ z)-f_1\gtreqless 0,\\ (n_2,\ z)-f_2\gtreqless 0,\\ \dots\dots\dots\\ (n_k,\ z)-f_k\gtreqless 0\end{gathered}$$

es convexo. Semejantes sistemas de desigualdades sirven de elemento principal en la descripción de muchos problemas referentes a la planificación de la producción, al mando, etc.

Ejercicios.

1. Demuéstrese que el conjunto de vectores z que satisfacen la condición $(z, z) \leqslant \alpha$, es convexo.
2. Demuéstrese que si un vector z pertenece al hiperplano (46.8), el vector $z + n$ se ubica en un semiespacio positivo.
3. Demuéstrese que si los hiperplanos están dados mediante las ecuaciones normadas (44.1), (44.2), el origen de coordenadas se dispone siempre en un semiespacio no positivo.

§ 48. Sistemas de ecuaciones lineales

Volvamos nuevamente a la investigación de los sistemas de ecuaciones algebraicas lineales, esta vez en relación con el problema de la intersección de hiperplanos.

Examinemos un espacio K, real o complejo, de dimensión m. Supongamos que en el espacio se ha introducido un producto escalar. Elijamos una base ortonormalizada y admitamos que los vectores normales $n_1, n_2, \ldots, n_k$ de los hiperplanos $\Pi_1, \Pi_2, \ldots, \Pi_k$ de (46.9) están dados mediante sus coordenadas

$$\begin{aligned} n_1 &= (a_{11}, a_{12}, \ldots, a_{1m}), \\ n_2 &= (a_{21}, a_{22}, \ldots, a_{2m}), \\ &\cdots\cdots\cdots\cdots \\ n_k &= (a_{k1}, a_{k2}, \ldots, a_{km}) \end{aligned} \tag{48.1}$$

respecto de dicha base. Convengamos en considerar que los vectores

$$z = (z_1, z_2, \ldots, z_m),$$

que pertenecen a la intersección de los hiperplanos, se determinan también por sus coordenadas en la misma base.

En el caso de un espacio real, el producto escalar de vectores en la base ortonormalizada es igual a la suma de los productos de coordenadas tomados dos a dos. Por ello, el sistema (46.9), expresado en coordenadas, tendrá la forma siguiente:

$$\begin{aligned} a_{11}z_1 + a_{12}z_2 + \ldots + a_{1m}z_m &= b_1, \\ a_{21}z_1 + a_{22}z_2 + \ldots + a_{2m}z_m &= b_2, \\ &\cdots\cdots\cdots\cdots \\ a_{k1}z_1 + a_{k2}z_2 + \ldots + a_{km}z_m &= b_k. \end{aligned} \tag{48.2}$$

En el caso de un espacio complejo obtendremos de nuevo un sistema análogo, a excepción de que los coeficientes ante las incógnitas y los segundos miembros se sustituirán por unos números conjugados complejos.

Así pues, el problema de la intersección de hiperplanos se reduce a otro problema en el que se resuelve un sistema de ecuaciones algebraicas lineales y que ya estudiamos en el párrafo 22. Evidentemente, cualquier sistema de ecuaciones de coeficientes complejos o reales puede también investigarse desde el punto de vista de la intersección de unos hiperplanos en el espacio complejo o real P_m.

Una cuestión de importancia es la investigación de un sistema de ecuaciones algebraicas lineales relacionada con *la compatibilidad* de éste. Precisamente con esta cuestión está ligada la respuesta a la pregunta sobre si es conjunto vacío o no la intersección de hiperplanos. Desde luego, para efectuar tal investigación se puede aprovechar el método de Gauss. No obstante, este método no es siempre cómodo.

Al estudiar los sistemas de ecuaciones algebraicas lineales nos vemos obligados a tratar dos matrices. Una de ellas se compone de los coeficientes de las incógnitas y se llama *matriz del sistema*. Esta matriz tiene la forma siguiente

$$\begin{pmatrix} a_{11}a_{12}\ldots a_{1m} \\ a_{21}a_{22}\ldots a_{2m} \\ \cdots\cdots\cdots \\ a_{k1}a_{k2}\ldots a_{km} \end{pmatrix}.$$

La otra se obtiene de la primera por adición de una columna de los segundos miembros

$$\begin{pmatrix} a_{11}a_{12}\ldots a_{1m}b_1 \\ a_{21}a_{22}\ldots a_{2m}b_2 \\ \cdots\cdots\cdots \\ a_{k1}a_{k2}\ldots a_{km}b_k \end{pmatrix}$$

y se denomina *matriz ampliada del sistema*. Observemos, en particular, que el rango de la matriz del sistema coincide con el del sistema de vectores (48.1).

TEOREMA 48.1. (*de Kronecker—Capelli*). *Para que un sistema de ecuaciones algebraicas lineales sea compatible, es necesario y suficiente que el rango de la matriz ampliada del sistema sea igual al rango de la matriz del sistema.*

DEMOSTRACION. Haremos uso de las designaciones aceptadas en el párrafo 22. Los vectores $a_1, a_2, \ldots, a_m, b$ representan en sí, salvo la disposición, las columnas de las matrices en consideración. Puesto que el rango de la matriz coincide con el del sistema de sus vectores columna, entonces para demostrar el teorema basta probar que el sistema será compatible cuando, y sólo cuando, el rango del sistema $a_1, a_2, \ldots, a_m$ coincide con el del sistema $a_1, a_2, \ldots, a_m, b$.

Supongamos que el sistema (48.2) es compatible. Esto significa que la igualdad (22.1) se verifica para cierto surtido de números $z_1, z_2, \ldots, z_m$, es decir, el vector b es una combinación lineal de vectores $a_1, a_2, \ldots, a_m$. Mas, de aquí se deduce que cualquier base del sistema $a_1, a_2, \ldots, a_m$ será, a la vez, la base del sistema $a_1, a_2, \ldots, a_m, b$, es decir, los rangos de ambos sistemas coinciden.

Supongamos ahora que los rangos de estos sistemas coinciden. Elijamos una base cualquiera de $a_1, a_2, \ldots, a_m$. Esta será también la base del sistema $a_1, a_2, \ldots, a_m, b$. Por consiguiente, el vector b se expresa linealmente en términos de una parte de los vectores del sistema $a_1, a_2, \ldots, a_m$. Ya que puede también representarse en forma de una combinación lineal de todos los vectores $a_1, a_2, \ldots, a_m$, esto significa la compatibilidad del sistema (48.2). El teorema queda demostrado.

El sistema de ecuaciones algebraicas lineales (48.2) se llama *no homogéneo*, si no todos los miembros segundos son nulos. En el caso contrario, se llama *homogéneo*. Todo sistema homogéneo es siempre compatible, puesto que una de sus soluciones es $z_1 = z_2 = \ldots \ldots = z_m = 0$. Un sistema, obtenido del sistema (48.2) como resultado de sustituir todos los segundos miembros por ceros, se denomina sistema homogéneo *reducido*. Si el sistema (48.2) es compatible, cada solución de éste se llamará solución *particular*. El conjunto de todas las soluciones particulares se denominará *solución general del sistema*.

Sirviéndonos de la información obtenida anteriormente, acerca de los planos y sistemas (46.6), (46.7), (46.9) que describen la intersección de los hiperplanos, podemos hacer toda una serie de deducciones referentes a la solución general del sistema de ecuaciones algebraicas lineales. A saber,

La solución general de un sistema homogéneo reducido forma en el espacio P_m un subespacio de dimensión $m - r$, donde r es el rango de la matriz del sistema. Cualquier base de este subespacio lleva el nombre de *sistema fundamental de soluciones*.

La solución general de un sistema no homogéneo es un plano en el espacio P_m, obtenido por desplazamiento de la solución general del sistema homogéneo reducido a cualquier solución particular del sistema no homogéneo.

La diferencia entre cualesquiera dos soluciones particulares de un sistema no homogéneo es una solución particular del sistema homogéneo reducido.

Entre las soluciones particulares de un sistema no homogéneo hay una única solución que es ortogonal a todas las soluciones del sistema homogéneo reducido. Esta solución se llama *normal*.

Para que un sistema compatible tenga una única solución, es necesario y suficiente que el rango de la matriz del sistema sea igual al número de incógnitas.

Para que un sistema homogéneo tenga una solución no nula, es necesario y suficiente que el rango de la matriz del sistema sea inferior al número de incógnitas.

En la investigación realizada de los sistemas de ecuaciones algebraicas lineales el concepto de determinante se ha usado sólo de manera indirecta, principalmente, por intermedio del concepto de rango de la matriz. Sin embargo, en la teoría de los sistemas de ecuaciones el determinante desempeña un papel considerablemente mayor.

Supongamos que la matriz de un sistema es cuadrada. Para que el rango de la matriz del sistema sea inferior al número de incógnitas, es necesario y suficiente que el determinante del sistema sea igual a cero. Por ello,

Un sistema homogéneo tiene una solución no nula cuando, y sólo cuando, el determinante del sistema es igual a cero.

Supongamos ahora que el determinante del sistema es distinto de cero. Esto es indicio de que el rango de la matriz del sistema es m. El rango de la matriz ampliada no puede ser inferior a m. Pero tampoco puede ser superior a m, puesto que no hay menores de orden $m+1$. Por consiguiente, los rangos de ambas matrices son idénticos, es decir, el sistema en este caso es obligatoriamente compatible. Más aún, tiene una solución única. De este modo,

Si el determinante de un sistema es distinto de cero, el sistema siempre tiene solución y esta solución es única.

Desde el punto de vista de la investigación de la intersección de hiperplanos, a este hecho se le puede atribuir la siguiente forma:

Si los vectores normales de unos hiperplanos forman la base de un espacio, la intersección de dichos hiperplanos no es vacía y contiene sólo un vector.

Indiquemos con d el determinante de un sistema y con d_j, un determinante que se distingue de d sólo en que la j-ésima columna en este último se ha sustituido por la columna de los segundos miembros $b_1, b_2, \ldots, b_m$. En este caso la única solución del sistema puede calcularse según las fórmulas

$$z_j = \frac{d_j}{d}, \quad j = 1, 2, \ldots, n. \tag{48.3}$$

En efecto, designemos mediante A_{ij} el complemento algebraico del elemento a_{ij} del determinante del sistema. Desarrollando d_j por los elementos de la j-ésima columna, obtenemos

$$d_j = \sum_{i=1}^{n} b_i A_{ij}$$

Sustituyamos ahora las expresiones (48.3) en la k-ésima ecuación arbitraria del sistema

$$\sum_{j=1}^{n} a_{kj} \frac{d_j}{d} = \frac{1}{d} \sum_{j=1}^{n} a_{kj} \sum_{i=1}^{n} b_i A_{ij} = \frac{1}{d} \sum_{j=1}^{n} \sum_{i=1}^{n} a_{kj} b_i A_{ij} =$$
$$= \frac{1}{d} \sum_{i=1}^{n} \sum_{j=1}^{n} b_i a_{kj} A_{ij} = \frac{1}{d} \sum_{i=1}^{n} b_i \sum_{j=1}^{n} a_{kj} A_{ij} = b_k.$$

La suma interior en la última igualdad es igual, de acuerdo con (40.5), (40.8), o bien a d o bien a cero, lo que depende de si $i = k$ ó $i \neq k$.

De este modo, las fórmulas (48.3) proporcionan una expresión explícita para la solución del sistema en términos de sus elementos. En virtud de la unicidad, no existe ninguna otra solución. Estas *fórmulas* llevan el nombre *de Cramer.*

Desde el punto de vista formal, basándose en los cálculos de los determinantes, se pueden resolver cualesquiera sistemas de ecuaciones del tipo (48.2). Al principio, calculando varios menores de la matriz del sistema y de la matriz ampliada, comprobamos la compatibilidad del sistema. Supongamos que el sistema resulta ser compatible y el rango de ambas matrices es igual a r. Sin restringir la generalidad, se puede considerar que el menor principal de orden r de la matriz del sistema es distinto de cero. De acuerdo con el teorema 41.1, las últimas $k - r$ filas de la matriz ampliada son combinaciones lineales de sus r primeras filas. Por consiguiente, el sistema

$$\begin{array}{|l|l} \hline a_{11}z_1 + \ldots + a_{1r}z_r & + a_{1,r+1}z_{r+1} + \ldots + a_{1m}z_m = b_1, \\ a_{21}z_1 + \ldots + a_{2r}z_r & + a_{2,r+1}z_{r+1} + \ldots + a_{2m}z_m = b_2, \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}z_1 + \ldots + a_{rr}z_r & + a_{r,r+1}z_{r+1} + \ldots + a_{rm}z_m = b_r \\ \hline \end{array} \qquad (48.4)$$

es equivalente al sistema (48.2).

Las incógnitas $z_{r+1}, \ldots, z_m$ se llamarán, como antes, *independientes.* Al atribuirles cualesquiera valores, se pueden determinar todas las demás incógnitas, resolviendo el sistema con la matriz del menor principal según, por ejemplo, las fórmulas de Cramer.

Este método de resolver el sistema posee cierta valía sólo en las investigaciones *teóricas.* En cuanto al aspecto *práctico,* resulta mucho más ventajoso servirse del método de Gauss, descrito en el § 22.

Ejercicios.

1. Demuéstrese que la solución general es un conjunto convexo.
2. Demuéstrese que entre todas las soluciones particulares de un sistema no homogéneo la solución normal es de longitud mínima.
3. Demuéstrese que el sistema fundamental es una totalidad de cualesquiera $m - r$ soluciones del sistema homogéneo reducido, para las cuales el determinante compuesto de los valores de las incógnitas independientes es distinto de cero.
4. En el § 22 se ha observado que las pequeñas variaciones en las coordenadas pueden conducir a la violación de la dependencia o independencia lineal de los vectores. ¿Qué deducciones pueden hacerse de este hecho con relación a los problemas referentes a la intersección de hiperplanos?

CAPÍTULO 6 LÍMITE EN EL ESPACIO LINEAL

§ 49. Espacio métrico

Uno de los conceptos fundamentales del análisis matemático lo constituye el concepto de límite. Está basado en que para los puntos de un eje numérico se ha definido la noción de "cercanía" o, con más precisión, de distancia entre los puntos.

La comparación de las "cercanías" se puede introducir también en los conjuntos cuya naturaleza es totalmente diferente. En el § 29 ya hemos definido la distancia entre los vectores de los espacios lineales dotados de producto escalar, descubriéndose que la distancia citada posee las mismas propiedades (29.5) que tiene la distancia entre los puntos de un eje numérico. La distancia entre los vectores se definía mediante el producto escalar que, a su vez, se introducía axiomáticamente.

Parece natural tratar de introducir *axiomáticamente* la propia distancia, al exigir que se cumplan sin falta las propiedades (29.5).

Ha de ser notado que muchos hechos fundamentales de la teoría del límite en el análisis matemático no están relacionados con el hecho que para los números están definidas unas operaciones algebraicas. Por eso empezaremos por la extensión del concepto de distancia a unos conjuntos arbitrarios de elementos que no son forzosamente vectores de un espacio lineal.

Un conjunto se denomina *espacio métrico*, si a todo par de sus elementos se le ha puesto en correspondencia un número real no negativo, llamado *distancia*, con la particularidad de que se cumplen los axiomas siguientes:

1) $\rho(x, y) = \rho(y, x)$,
2) $\rho(x, y) > 0$, si $x \neq y$; $\quad \rho(x, y) = 0$, si $x = y$,
3) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$

para cualesquiera elementos x, y, z. Estos axiomas se llaman axiomas *de la métrica*; el primero de ellos se denomina *axioma de simetría* y el tercero, *axioma triangular*.

Todo conjunto de elementos en el que se ha definido una relación de igualdad puede ser convertido formalmente en un espacio métrico al poner

$$\rho(x, y) = \begin{cases} 0, & \text{si } x = y, \\ 1, & \text{si } x \neq y. \end{cases} \tag{49.1}$$

Es fácil comprobar que todos los axiomas de la métrica están cumplidos.

El vector x_0 del espacio métrico X se llama *límite* de la sucesión $\{x_n\}$ de elementos $x_1, x_2, \ldots, x_n, \ldots$ de X, si la sucesión de distancias $\rho(x_0, x_1), \rho(x_0, x_2), \ldots, \rho(x_0, x_n) \ldots$ converge a cero En este caso suele escribirse

$$x_n \to x_0$$

o bien

$$\lim_{n \to \infty} x_n = x_0$$

y decirse que la sucesión $\{x_n\}$ se llama *convergente en X* o simplemente *convergente*.

Observemos que una misma sucesión de elementos de un mismo conjunto X puede ser convergente o divergente, según sea la métrica introducida en X. Supongamos, por ejemplo, que en un espacio métrico X se ha elegido una sucesión convergente $\{x_n\}$, compuesta por unos elementos distintos dos a dos. Cambiemos la métrica en X, al introducirla de acuerdo con (49.1). En este caso la sucesión $\{x_n\}$ ya no será convergente. Efectivamente, supongamos que $x_n \to x_0'$, es decir, $\rho(x_n, x_0') \to 0$. Con la métrica nueva esto es posible sólo en el caso en que todos los elementos de $\{x_n\}$, a excepción de su número finito, coinciden con x_0'. La contradicción obtenida confirma la afirmación enunciada.

Las dos propiedades que siguen son comunes para cualesquiera sucesiones convergentes.

Si una sucesión $\{x_n\}$ converge, será convergente y tendrá el mismo límite cualquiera de sus subsucesiones. La sucesión puede tener no más que un límite.

La primera propiedad es obvia. Supongamos que la sucesión $\{x_n\}$ tiene dos límites, x_0 e y_0. En este caso, para cualquier número $\varepsilon > 0$, tan pequeño como se quiera, se puede elegir un número N tal que

$$\rho(x_0, x_n) < \frac{\varepsilon}{2}, \qquad \rho(y_0, x_n) < \frac{\varepsilon}{2}$$

para todo $n > N$. De aquí, haciendo uso del axioma triangular, hallamos

$$\rho(x_0, y_0) \leqslant \rho(x_0, x_n) + \rho(x_n, y_0) < \varepsilon.$$

Por ser ε arbitrario, esto significa que $\rho(x_0, y_0) = 0$, es decir, $x_0 = y_0$.

Se denomina *bola* $S(a, r)$ en el espacio métrico X el conjunto de todos los elementos $x \in X$ que satisfacen la condición

$$\rho(a, x) < r. \tag{49.2}$$

El elemento a se llama *centro* de la bola; el número r, *radio* de la bola. Toda bola con el centro en a llamemos *entorno* *) del elemento a. Un conjunto de elementos se denomina *limitado*, si pertenece íntegramente a cierta bola.

Es fácil ver que el elemento x_0 es el límite de la sucesión $\{x_n\}$ cuando, y sólo cuando, cualquier entorno del elemento x_0 contiene todos los elementos de la sucesión que se examina, a partir de cierto número.

En el espacio métrico pueden ser introducidos también muchos otros conceptos de importancia, con los cuales nos encontramos en los conjuntos numéricos. Así por ejemplo, dado el conjunto $M \subset X$, el elemento $x \in X$ se llamará *punto límite* de este conjunto, si cualquier entorno del elemento x contiene al menos un solo elemento del conjunto M que no coincide con x. Un conjunto obtenido por agregación a M de todos sus puntos límite se denomina *adherencia* del conjunto M y se indica con $\overline{M}$. El conjunto M se llama *cerrado*, si $M = \overline{M}$.

Examinemos los puntos límite de la bola (49.2). Probemos que todos ellos satisfacen la condición

$$\rho(a, x) \leqslant r. \tag{49.3}$$

Efectivamente, supongamos que existe por lo menos un punto límite x' para la bola (49.2), para el cual $\rho(a, x') > r$. Por definición de punto límite, todo entorno de elemento x' debe contener al menos un elemento de la bola (49.2) que no sea coincidente con x'. Pero el entorno cuyo radio es $0{,}5\,(\rho(a, x') - r)$ no contiene, a ciencia cierta, ninguno de los elementos de tal índole. De conformidad con lo dicho:

El conjunto $\overline{S}(a, r)$ de todos los elementos x que satisfacen la condición (49.3) *se llama bola cerrada.*

Ejercicios.

1. Demuéstrese que si $x_n \to x$, entonces $\rho(x_n, z) \to \rho(x, z)$ para todo elemento z.

2. ¿Será el conjunto de todos los números reales un espacio métrico, si para los números x, y ponemos

$$\rho(x, y) = \operatorname{arc\,tg} |x - y|?$$

3. ¿Podrá un conjunto compuesto por un número finito de elementos tener puntos límite?

*) Se usa también el término «vecindad». (*N. del Tr.*)

§ 50. Espacio completo

La sucesión $\{x_n\}$ de elementos de un espacio métrico se llama *fundamental* o *convergente en sí*, si para cualquier número $\varepsilon > 0$ existe tal número N que $\rho(x_n, x_m) < \varepsilon$ para $n, m > N$.

Toda sucesión fundamental es limitada. En efecto, siendo ε dado, elijamos un número N, de acuerdo con la definición, y tomemos un número arbitrario $n_0 > N$. Todos los elementos de la sucesión, a partir de x_{n_0}, pertenecen, a ciencia cierta, a una bola de radio ε y centro x_{n_0}. Mientras tanto todos los elementos pertenecen a una bola cuyos centro y radio, respectivamente, son x_{n_0} y el máximo de los números

$$\varepsilon, \ \rho(x_1, x_{n_0}), \ \ldots, \ \rho(x_{n_0-1}, x_{n_0}).$$

Si la sucesión es convergente, será fundamental. Supongamos que la sucesión $\{x_n\}$ converge a x_0. Entonces para todo $\varepsilon > 0$ existe N tal que

$$\rho(x_n, x_0) < \frac{\varepsilon}{2}$$

para $n > N$. De acuerdo con el axioma triangular,

$$\rho(x_n, x_m) \leqslant \rho(x_n, x_0) + \rho(x_0, x_m) < \varepsilon$$

para $n, m > N$, lo que significa precisamente el carácter fundamental de la sucesión $\{x_n\}$.

Para el conjunto de todos los números reales es cierta también la afirmación recíproca. A saber, toda sucesión fundamental es convergente. No obstante, en el caso general esto ya no es cierto, lo que se confirma con el ejemplo de un espacio métrico del cual se ha eliminado por lo menos un solo punto límite.

Un espacio métrico se denomina *completo*, si toda sucesión fundamental en él es convergente.

En los espacios métricos completos tiene lugar un teorema que es análogo al teorema sobre segmentos encajados para los números reales. Sea dada una sucesión de bolas. Llamaremos estas bolas *encajadas* una dentro de la otra, si toda bola consecutiva está contenida dentro de la anterior.

TEOREMA 50.1 *Supongamos que en el espacio métrico completo X sea dada la sucesión $\{\overline{S}(a_n, \varepsilon_n)\}$ de bolas cerradas encajadas una dentro de la otra. Si la sucesión de radios tiende hacia cero, existe un único elemento de X perteneciente a todas estas bolas.*

DEMOSTRACIÓN. Examinemos la sucesión $\{a_n\}$. Puesto que $\overline{S}(a_{n+p}, \varepsilon_{n+p}) \subset \overline{S}(a_n, \varepsilon_n)$ para cualquier $p \geqslant 0$, entonces $a_{n+p} \in$

$\in \overline{S}(a_n, \varepsilon_n)$. Por consiguiente,

$$\rho(a_{n+p}, a_n) \leqslant \varepsilon_n,$$

de donde se desprende que la sucesión $\{a_n\}$ es fundamental.

El espacio X es completo y, por ende, la sucesión $\{a_n\}$ converge a cierto límite a de X. Tomemos una bola cualquiera $\overline{S}(a_k, \varepsilon_k)$. A esta bola le pertenecen todos los términos de la sucesión $\{a_n\}$, a partir de a_k. En vista de que las bolas son cerradas, el límite de esta sucesión también pertenece a $\overline{S}(a_k, \varepsilon_k)$. De este modo, a pertenece a todas las bolas.

Admitamos luego que existe otro elemento, b, perteneciente también a todas las bolas. De acuerdo con el axioma triangular

$$\rho(a, b) \leqslant \rho(a, a_n) + \rho(a_n, b) \leqslant 2\varepsilon_n.$$

Puesto que ε_n puede ser tan pequeño como se quiera, esto significa que $\rho(a, b) = 0$, es decir, $a = b$.

Como ejemplos más importantes de espacios completos sirven los conjuntos de números reales y complejos. En este caso se supone que la distancia entre los números coincide con el módulo de la diferencia entre ellos. La completitud del conjunto de números reales se demuestra en el curso del análisis matemático. Mostremos la completitud del conjunto de números complejos.

Convengamos en considerar que los números complejos vienen dados en forma algebraica. La distancia entre los números

$$z = a + ib, \qquad v = c + id$$

se introducirá de conformidad con la regla

$$\rho(z, v) = |z - v|, \tag{50.1}$$

donde

$$|z - v|^2 = (a - c)^2 + (b - d)^2. \tag{50.2}$$

Es evidente que los axiomas de la métrica están cumplidos.

Consideremos una sucesión $\{z_k = a_k + ib_k\}$ de números complejos. Supongamos que es fundamental. Siendo $\varepsilon > 0$, hallemos tal N que para cualesquiera $n, m > N$ sea

$$|z_n - z_m| < \varepsilon.$$

De (50.2) se infiere que en este caso

$$|a_n - a_m| < \varepsilon, \qquad |b_n - b_m| < \varepsilon, \tag{50.3}$$

es decir, las sucesiones $\{a_k\}$ y $\{b_k\}$ son también fundamentales. Dado que el conjunto de números reales es completo, existen unos números a, b para los cuales

$$a_k \to a, \qquad b_k \to b.$$

Pasando al límite en las desigualdades (50.3), obtendremos

$$|a_n - a| \leqslant \varepsilon, \qquad |b_n - b| \leqslant \varepsilon.$$

Al designar

$$z = a + ib,$$

encontramos que

$$\rho(z_n, z) \leqslant \sqrt{2}\,\varepsilon$$

para todo $n > N$. Mas, esto significa que la sucesión fundamental $\{z_k\}$ es convergente.

Observemos, a título de corolario, que la sucesión $\{z_k = a_k + ib_k\}$ converge al número $z = a + ib$ cuando, y sólo cuando, las sucesiones $\{a_k\}$ y $\{b_k\}$ convergen a los números a y b, respectivamente.

El espacio completo de números complejos tiene mucho en común con el espacio de números reales. En particular, toda sucesión acotada de números complejos tiene una subsucesión convergente. En efecto, esta afirmación es verídica para toda sucesión acotada de números reales. Luego, es evidente que si es acotada la sucesión $\{z_k = a_k + ib_k\}$, lo serán también las sucesiones $\{a_k\}$ y $\{b_k\}$. Puesto que la sucesión $\{a_k\}$ es acotada, tiene subsucesión convergente $\{a_{\nu_k}\}$. Consideremos la sucesión $\{b_{\nu_k}\}$. Es acotada y por esta razón también tiene subsucesión convergente $\{b_{\nu_{k_n}}\}$. Está claro que $\{a_{\nu_{k_n}}\}$ será convergente. Por consiguiente, la subsucesión $\{z_{\nu_{k_n}}\}$ será también convergente.

Por analogía con el espacio real, en el espacio complejo se introduce la noción de sucesión infinita creciente. A saber, la sucesión $\{z_k\}$ se denomina *infinita creciente*, si para un número A, tan grande como se quiera, se puede indicar un número N tal que para todo $k > N$ se cumple la desigualdad $|z_k| > A$. Es evidente que en toda sucesión no acotada siempre puede elegirse una subsucesión infinita creciente.

Ejercicios.

1. ¿Será un espacio completo el conjunto de todos los números reales, si para los números x, y hacemos

$$\rho(x, y) = \operatorname{arctg} |x - y|?$$

2. Demuéstrese que todo conjunto cerrado de un espacio completo es por sí mismo un espacio completo.

3. ¿Será necesariamente un espacio completo todo conjunto cerrado de un espacio métrico arbitrario?

4. Constrúyase una métrica, para la cual el conjunto de todos los números complejos no sea un espacio completo.

§ 51. Desigualdades auxiliares

Obtengamos algunas desigualdades que se emplearán en las investigaciones inmediatas. Tomemos un número positivo arbitrario α y examinemos una función exponencial $y = \alpha^x$ (fig. 51.1). Sean x_1, x_2 dos números reales distintos. Tracemos una recta por los puntos cuyas coordenadas son (x_1, α^{x_1}), (x_2, α^{x_2}). Teniendo presentes las propiedades de una función exponencial, concluimos que con el cambio de argumento en el segmento $[x_1, x_2]$ todos los puntos de la misma no serán superiores a los puntos de la recta construida.

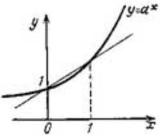


Fig. 51.1.

Ahora, sea $x_1 = 0$ y sea $x_2 = 1$. En este caso la ecuación de la recta a examinar será $y = \alpha x + (1 - x)$. Por consiguiente,

$$\alpha^x \leqslant \alpha x + (1 - x) \tag{51.1}$$

para $0 \leqslant x \leqslant 1$.

Llamemos *conjugados* los números positivos p, q, si satisfacen la correlación

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{51.2}$$

Está claro que p, $q > 1$.

Para cualesquiera números positivos a, b el número $a^p b^{-q}$ será también positivo y se le puede tomar en calidad de α de (51.1). Si se considera que $x = p^{-1}$, entonces $1 - x = q^{-1}$. Ahora, de (51.1) se deduce la validez de la desigualdad siguiente:

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q} \tag{51.3}$$

para cualesquiera a, b positivos y p, q conjugados. Evidentemente, esta desigualdad tiene lugar para todos los números no negativos a, b.

Consideraremos dos vectores arbitrarios x, y, pertenecientes al espacio $\mathbf{R}_n$ o $\mathbf{C}_n$. Supongamos que estos vectores vienen dados mediante sus coordenadas

$$x = (x_1, x_2, \ldots, x_n),$$
$$y = (y_1, y_2, \ldots, y_n).$$

Demos a conocer la así llamada *desigualdad de Hölder*

$$\sum_{k=1}^{n} |x_k y_k| \leqslant \Big(\sum_{k=1}^{n} |x_k|^p\Big)^{1/p} \Big(\sum_{k=1}^{n} |x_k|^p\Big)^{1/q}. \tag{51.4}$$

Hemos de notar que si entre los vectores x, y hay por lo menos un vector nulo, la desigualdad de Hölder es, evidentemente, lícita. Por esto podemos considerar que $x \neq 0$, $y \neq 0$. Supongamos que la desigualdad se cumple para cualesquiera vectores no nulos x, y. En este caso se cumple también para los vectores λx, μy con cualesquiera λ, μ. Por eso es suficiente demostrarla sólo para el caso en que

$$\sum_{k=1}^{n} |x_k|^p = \sum_{k=1}^{n} |y_k|^q = 1. \qquad (51.5)$$

Haciendo ahora $a = |x_k|$, $b = |y_k|$ en la desigualdad (51.3) y sumando según k desde 1 hasta n, obtendremos, tomando en consideración (51.2), (51.5):

$$\sum_{k=1}^{n} |x_k y_k| \leqslant 1.$$

Esto es precisamente la desigualdad de Hölder para el caso (51.5).

Pasemos ahora a la demostración de *la desigualdad de Minkowski* en la que para cualesquiera vectores x, y de $\mathbf{R}_n$ o $\mathbf{C}_n$ se verifica

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p\right)^{1/p} \leqslant \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{1/p} \qquad (51.6)$$

para todo $p \geqslant 1$.

La desigualdad de Minkowski es obvia para $p = 1$, puesto que el módulo de una suma de dos números no es superior a la suma de sus módulos. Además, se cumple a ciencia cierta, si por lo menos uno de los vectores x, y es nulo. Por ello podemos limitarnos al examen del caso en que $p > 1$ y $x \neq 0$. Escribamos la identidad

$$(|a| + |b|)^p = (|a| + |b|)^{p-1} |a| + (|a| + |b|)^{p-1} |b|.$$

Al poner en ésta $a = x_k$, $b = y_k$ y al sumar según k desde 1 hasta n, obtenemos

$$\sum_{k=1}^{n} (|x_k| + |y_k|)^p = \sum_{k=1}^{n} (|x_k| + |y_k|)^{p-1} |x_k| + \sum_{k=1}^{n} (|x_k| + |y_k|)^{p-1} |y_k|.$$

Apliquemos la desigualdad de Hölder a cada una de las dos sumas que figuran en el segundo miembro de esta correlación. Teniendo presente que $(p-1)q = p$, obtenemos

$$\sum_{k=1}^{n} (|x_k| + |y_k|)^p \leqslant \left(\sum_{k=1}^{n} (|x_k| + |y_k|)^p\right)^{1/q} \times$$

$$\times \left(\left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p\right)^{1/p}\right)$$

Al dividir ambos miembros de la desigualdad por

$$\left(\sum_{k=1}^{n}(|x_k|+|y_k|)^p\right)^{1/q},$$

encontramos que

$$\left(\sum_{k=1}^{n}|x_k|+|y_k|)^p\right)^{1/p}\leqslant\left(\sum_{k=1}^{n}|x_k|^p\right)^{1/p}+\left(\sum_{k=1}^{n}|y_k|^p\right)^{1/p},$$

de donde proviene inmediatamente la desigualdad (51.6).

Ejercicios.

1. Dedúzcase la desigualdad de Cauchy—Buniakovski a partir de la desigualdad de Hölder.
2. Estúdiese la desigualdad de Hölder para $p \to \infty$.
3. Estúdiese la desigualdad de Minkowski para $p \to \infty$.

§ 52. Espacio normalizado

Al concepto de espacio métrico hemos llegado concentrando nuestra atención sólo en una propiedad del conjunto, a saber, en la presencia de una distancia en dicho conjunto. De modo análogo, al concentrar nuestra atención en las operaciones en un conjunto, llegamos al concepto de espacio lineal. Ahora consideraremos los espacios lineales provistos de una métrica.

Evidentemente, si el concepto de distancia no está relacionado de tal o cual modo con las operaciones sobre elementos, resulta imposible construir una teoría enjundiosa cuyos hechos unan juntos los conceptos algebraicos y métricos. Por esta razón impondremos sobre la métrica, introducida en el espacio lineal, unas condiciones complementarias.

Ya nos hemos encontrado en realidad con los espacios lineales métricos. Son, por ejemplo, los espacios euclídeo y unitario con una métrica (29.4). No obstante, la necesidad en tal métrica no surge siempre. La introducción de un producto escalar significa, de hecho, que se introduce no sólo la distancia entre los elementos, sino también los ángulos entre los mismos. Con mayor frecuencia se requiere que en el espacio lineal sea dado sólo el concepto aceptable de distancia. Los más importantes espacios lineales de este tipo son los así llamados *espacios normalizados*.

Un espacio lineal X, sea real o complejo, se llama normalizado, si a todo vector $x \in X$ se le ha puesto en correspondencia un número real $\| x \|$, llamado *norma* del vector x, con la particularidad de que se consideren cumplidos los siguientes axiomas:

1) $\| x \| > 0$, si $x \neq 0$, $\| 0 \| = 0$,
2) $\| \lambda x \| = | \lambda | \, \| x \|$,
3) $\| x + y \| \leqslant \| x \| + \| y \|$

para cualesquiera vectores x, y y todo número λ. El segundo axioma lleva el nombre de *axioma de homogeneidad absoluta de la norma*, el tercer axioma se llama *axioma de la desigualdad triangular*.

Del axioma de homogeneidad absoluta de la norma proviene que para todo vector x no nulo se puede encontrar un número λ tal que la norma del vector λx sea igual a uno. Para ello será suficiente tomar $\lambda = \| x \|^{-1}$. Un vector cuya norma es igual a la unidad se llamará *normalizado*.

De la desigualdad triangular para las normas se desprende una correlación muy útil. Tenemos $\| x \| \leqslant \| y \| + \| x - y \|$ para cualesquiera x, y. Por consiguiente, $\| x \| - \| y \| \leqslant \| x - y \|$. Intercambiando x, y de lugares, obtenemos $\| y \| - \| x \| \leqslant \leqslant \| x - y \|$. Por esto

$$| \| x \| - \| y \| | \leqslant \| x - y \|. \tag{52.1}$$

Un espacio normalizado se convierte con facilidad en un espacio métrico, si ponemos

$$\rho (x, y) = \| x - y \|. \tag{52.2}$$

En efecto, $\rho (x, y) = 0$ quiere decir que $\| x - y \| = 0$, lo que de acuerdo con el axioma 1, significa $x = y$. La simetría de la distancia introducida es obvia. Por fin, la desigualdad triangular para la distancia es simplemente un corolario de la desigualdad triangular para la norma. A saber,

$$\rho (x, y) = \| x - y \| = \| (x - z) + (z - y) \| \leqslant$$
$$\leqslant \| x - z \| + \| z - y \| = \rho (x, z) + \rho (z, y).$$

Observemos que

$$\| x \| = \rho (x, 0). \tag{52.3}$$

La métrica (52.2) definida en el espacio lineal posee, además, las siguientes propiedades:

$$\rho (x + z, y + z) = \rho (x, y)$$

para cualesquiera $x, y, z \in X$, es decir, la distancia no varía con el desplazamiento de los vectores, y

$$\rho (\lambda x, \lambda y) = | \lambda | \rho (x, y)$$

para cualesquiera vectores $x, y \in X$ y todo número λ, es decir, la distancia es una función absolutamente homogénea.

Si en el espacio lineal metrico X una métrica, cualquiera que sea, satisface estas dos exigencias complementarias, entonces X puede considerarse como espacio normalizado, siempre que la norma se defina mediante la igualdad (52.3) para todo $x \in X$.

Tomando en consideración las correlaciones del § 29, es fácil establecer que *un espacio lineal provisto de un producto escalar es un* *espacio normalizado*. En este caso por norma del vector conviene tomar su longitud.

Se pueden aducir también otros ejemplos de introducción de la norma. Supongamos que en un espacio lineal los vectores vienen dados por medio de sus coordenadas respecto de cierta base. Si $x = (x_1, x_2, \ldots, x_n)$, ponemos

$$\| x \|_p = \Big(\sum_{k=1}^{n} |x_k|^p \Big)^{1/p} \tag{52.4}$$

donde $p \geqslant 1$. El cumplimiento de los primeros dos axiomas para la norma es obvio, mientras que el cumplimiento del tercer axioma se deduce de la desigualdad de Minkowski (51.6). Son de mayor uso las siguientes normas:

$$\| x \|_1 = \sum_{k=1}^{n} |x_k|, \quad \| x \|_2 = \Big(\sum_{k=1}^{n} |x_k|^2 \Big)^{1/2},$$

$$\| x \|_\infty = \max_{1 \leqslant k \leqslant n} |x_k|. \tag{52.5}$$

La segunda de dichas normas se llama, a menudo, norma *euclídea* y se indica con $\| x \|_E$.

En adelante se considerarán sólo espacios normalizados los que contengan la métrica (52.2). La convergencia de una sucesión de vectores en tal métrica la llamaremos *convergencia en norma* y la acotación de un conjunto de vectores, *acotación en norma*, etc.

Ejercicios.

1. Demuéstrese que existe una sucesión de vectores cuyas normas forman una sucesión infinita creciente.

2. Demuéstrese que para cualesquiera números λ_i y vectores e_i se verifica la desigualdad

$$\Big\| \sum_{i=1}^{n} \lambda_i e_i \Big\| \leqslant \sum_{i=1}^{n} |\lambda_i| \, \| e_i \|.$$

3. Demuéstrese que si $x_n \to x$, $y_n \to y$, $\lambda_n \to \lambda$, entonces

$$\| x_n \| \to \| x \|, \; x_n + y_n \to x + y, \; \lambda_n x_n \to \lambda x.$$

§ 53. Convergencia en norma y convergencia coordenada

Un espacio lineal de dimensión finita, real o complejo, además de la convergencia en norma admite también la introducción de otro concepto de convergencia. Consideremos el espacio X y sea $e_1, e_2, \ldots, e_n$ su base. Para cualquier sucesión $\{x_m\}$ de los vectores de X existen los desarrollos

$$x_m = \sum_{k=1}^{n} \xi_k^{(m)} e_k. \tag{53.1}$$

Si para el vector

$$x_0 = \sum_{k=1}^{n} \xi_k^{(0)} e_k \tag{53.2}$$

tienen lugar las correlaciones límites

$$\lim_{m\to\infty} \xi_k^{(m)} = \xi_k^{(0)} \tag{53.3}$$

para cualesquiera $k = 1, 2, \ldots, n$, entonces diremos que tiene lugar la convergencia *coordenada* de la sucesión $\{x_m\}$ al vector x_0.

La convergencia coordenada es bien natural. Si dos vectores son "cercanos", se puede suponer que han de ser "cercanas" también las coordenadas correspondientes en el desarrollo según una misma base. Los espacios normalizados de dimensión finita son remarcables por el hecho de que en estos espacios los conceptos de convergencia en norma y de convergencia coordenada son *equivalentes*.

Es fácil probar que de la convergencia coordenada proviene la convergencia en norma. En efecto, supongamos que tienen lugar las correlaciones límites (53.3). Haciendo uso de los axiomas de homogeneidad absoluta de la norma y de la desigualdad triangular, concluimos a base de (53.3), (53.2) que

$$\| x_m - x_0 \| = \Big\| \sum_{k=1}^{n} (\xi_k^{(m)} - \xi_k^{(0)}) e_k \Big\| \leqslant \sum_{k=1}^{n} |\xi_k^{(m)} - \xi_k^{(0)}| \, \| e_k \| \to 0.$$

La demostración de la afirmación inversa es considerablemente más compleja.

LEMA 53.1. *Si en un espacio normalizado de dimensión finita la sucesión de vectores está acotada en norma, serán acotadas también las sucesiones numéricas de todas las coordenadas en el desarrollo de los vectores según cualquier base.*

DEMOSTRACION. Supongamos que todo vector de la sucesión $\{x_m\}$ esté representado en la forma (53.1). Introduzcamos la designación

$$\sigma_m = \sum_{k=1}^{n} |\xi_k^{(m)}|$$

y demostremos que la sucesión $\{\sigma_m\}$ está acotada.

Supóngase que no es así. Entonces, de dicha sucesión se puede elegir una subsucesión infinita creciente $\{\sigma_{m_p}\}$. Pongamos

$$y_p = \frac{1}{\sigma_{m_p}} x_{m_p}.$$

Si

$$y_p = \sum_{k=1}^{n} \eta_k^{(p)} e_k,$$

entonces, desde luego

$$\eta_k^{(p)} = \frac{\xi_k^{(m_p)}}{\sigma_{m_p}}$$

para cualesquiera $k = 1, 2, \ldots, n$ y m_p, y concluimos que

$$\sum_{k=1}^{n} |\eta_k^{(p)}| = 1. \tag{53.4}$$

Las sucesiones $\{\eta_k^{(p)}\}$ son acotadas, puesto que, de conformidad con (53.4), $|\eta_k^{(p)}| \leqslant 1$. Por consiguiente, se puede eligir tal subsucesión de vectores $\{y_{p_1}\}$ que la subsucesión $\{\eta_1^{(p_1)}\}$ será convergente, es decir

$$\lim_{p_1 \to \infty} \eta_1^{(p_1)} = \eta_1$$

para cierto número η_1. En la subsucesión $\{y_{p_1}\}$ puede elegirse, a su vez, una subsucesión $\{y_{p_2}\}$, para la cual

$$\lim_{p_2 \to \infty} \eta_2^{(p_2)} = \eta_2$$

para cierto número η_2. En este caso, como hasta ahora,

$$\lim_{p_2 \to \infty} \eta_1^{(p_2)} = \eta_1.$$

Continuando este proceso, eligimos en la sucesión $\{y_p\}$ tal subsucesión $\{y_{p_n}\}$ que existirán los límites

$$\lim_{p_n \to \infty} \eta_k^{(p_n)} = \eta_k \tag{53.5}$$

para cualesquiera $k = 1, 2, \ldots, n$. De acuerdo con (53.4),

$$\sum_{k=1}^{n} |\eta_k| = 1 \tag{53.6}$$

De la convergencia coordenada se desprende la convergencia en norma, por lo cual, las correlaciones límite (53.5) significan que

$$\lim_{p_n \to \infty} \| y_{p_n} - y \| = 0, \tag{53.7}$$

donde

$$y = \sum_{k=1}^{n} \eta_k e_k.$$

El vector y no debe ser nulo en virtud de (53.6). Por otro lado, se tiene

$$\| y_{p_n} \| = \frac{\| x_{m_{p_n}} \|}{\sigma_{m_{p_n}}} \to 0.$$

puesto que la sucesión $\{x_{m_{p_n}}\}$ está acotada en norma, y la subsucesión $\{\sigma_{m_{p_n}}\}$ es infinita creciente. Por consiguiente de (53.7) se desprende que $\| y \| = 0$, es decir, y es un vector nulo. La contradicción obtenida demuestra la validez de la afirmación del lema.

TEOREMA 53.1. *En el espacio normalizado de dimensión finita, de la convergencia en norma se infiere la convergencia coordenada.*

DEMOSTRACION. Sea dada la sucesión de vectores $\{x_m\}$ que converge en norma al vector x_0. Es evidente que basta examinar el caso en que $x_0 = 0$ y en la sucesión $\{x_m\}$ no haya vectores nulos. Representemos los vectores x_m en forma de los desarrollos (53.1). La sucesión de vectores

$$y_m = \frac{1}{\| x_m \|} x_m$$

será acotada en norma y, de acuerdo con el lema 53.1, deben ser acotadas las sucesiones de números

$$\eta_k^{(m)} = \frac{\xi_k^{(m)}}{\| x_m \|}$$

para cualesquiera $k = 1; 2, \ldots, n$. Puesto que $\| x_m \| \to 0$, esto es posible sólo en el caso en que $\xi_k^{(m)} \to 0$ para todo k. Pero esto es precisamente testimonio de que tiene lugar la convergencia coordenada de la sucesión $\{x_m\}$ al vector x_0.

La convergencia coordenada se emplea con toda la eficacia en las investigaciones *teóricas*, mientras que la convergencia en norma es mucho más cómoda para utilizarla en las aplicaciones *prácticas*. Esto se debe, principalmente, a que en la investigación de los espacios lineales de gran dimensión resulta difícil operar con un elevado número de sucesiones coordenadas. Además, no siempre se conoce aunque sea una sola base. Pero, si incluso sabemos la base, el empleo de la misma nos conduce, en la mayoría de los casos, a cálculos voluminosos y no justificados.

Ejercicios.

1. Cuando se demuestra la equivalencia de dos tipos de convergencia ¿será importante el requisito de que un espacio sea de dimensión finita?
2. Demuéstrese que si un conjunto de vectores de un espacio de dimensión finita está acotado en una norma, estará acotado también en cualquier otra norma.
3. Demuéstrese que si en un espacio de dimensión finita $x_n \to x$ en una norma, entonces $x_n \to x$ en cualquier otra norma.

§ 54. Completitud de los espacios normalizados

Los espacios normalizados de dimensión finita representan espacios en los cuales tienen lugar varias analogías de las afirmaciones relacionadas con el concepto de límite en los conjuntos de números. He aquí algunas de estas analogías.

LEMA 54.1. *De toda sucesión acotada de vectores de un espacio normalizado de dimensión finita se puede elegir una subsucesión convergente en dicho espacio.*

DEMOSTRACION. Sea $\{x_m\}$ una sucesión arbitraria acotada en norma. Representemos los vectores x_m en forma de los desarrollos (53.1). De acuerdo con el lema 53.1, las sucesiones $\{\xi_k^{(m)}\}$ serán acotadas para cualesquiera $k = 1, 2, \ldots, n$. De un modo semejante al usado en la demostración del lema 53.1, elijamos en la sucesión $\{x_m\}$ una subsucesión $\{x_{m_n}\}$, para la cual existen correlaciones límites $\xi_k^{(m_n)} \to \xi_k^{(0)}$ para todo k. De aquí se deduce que la subsucesión $\{x_{m_n}\}$ converge al vector (53.2).

El lema demostrado es una analogía del lema de Bolzano—Weierstrass que es conocido en el curso del análisis matemático. El lema es de mucha importancia en las investigaciones de cualesquiera espacios normalizados de dimensión finita. Ilustrémoslo demostrando algunas afirmaciones.

TEOREMA 54.1. *Todo espacio normalizado de dimensión finita es completo.*

DEMOSTRACION. Sea $\{x_m\}$ una sucesión fundamental. Está acotada. Elijamos en ella una subsucesión convergente $\{x_{m_n}\}$ e indiquemos con x_0 su límite. Entonces

$$\| x_m - x_0 \| \leqslant \| x_m - x_{m_n} \| + \| x_{m_n} - x_0 \|.$$

Tomemos un número arbitrario $\varepsilon > 0$. Ya que la sucesión $\{x_m\}$ es fundamental, existe un número N_1 tal que $\| x_m - x_{m_n} \| < \varepsilon/2$ cuando $m, m_n > N_1$. En vista de que la sucesión $\{x_{m_n}\}$ converge a x_0, existe un número N_2 tal que $\| x_{m_n} - x_0 \| < \varepsilon/2$ cuando $m_n > N_2$. Si N es el máximo de los números N_1, N_2, entonces, siendo $m > N$,

$$\| x_m - x_0 \| < \varepsilon.$$

El número ε es arbitrario. Por lo tanto, la sucesión fundamental $\{x_m\}$ converge en norma al vector x_0.

LEMA 54.2. *Todo subespacio de dimensión finita X_0 de un espacio normalizado X es un conjunto cerrado.*

DEMOSTRACION. Consideremos en el espacio normalizado X un subespacio de dimensión finita X_0. Supongamos que el vector $x \in X$ es un punto límite para X_0. Esto quiere decir que existe una sucesión de vectores $\{x_m\}$ de X_0, no coincidentes con x, tal que

$\| x_m - x \| \to 0$. La sucesión $\{x_m\}$ es acotada y, por ende, podemos elegir de la misma una subsucesión $\{x_{m_p}\}$ que, por ser X_0 completo, converge a cierto vector $x_0 \in X_0$. Ahora tenemos

$$\| x - x_0 \| \leqslant \| x - x_{m_p} \| + \| x_{m_p} - x_0 \| \to 0,$$

es decir, $x = x_0$.

LEMA 54.3. *Sea X un espacio normalizado y sea X_0 su subespacio de dimensión finita no coincidente con X. Existe un vector normalizado $x \notin X_0$ tal que $\| x - x_0 \| \geqslant 1$ para cualquier vector $x_0 \in X_0$.*

DEMOSTRACION. Como que X_0 no coincide con X, existe un vector $x' \notin X_0$. Puesto que X_0 es cerrado, tenemos

$$\inf_{x_0 \in X_0} \| x' - x_0 \| = d > 0. \tag{54.1}$$

Por definición de la cota exacta inferior, en X_0 habrá un vector $x_0^{(k)}$, para el cual

$$d \leqslant \| x' - x_0^{(k)} \| \leqslant \frac{d}{1-2^{-k}}.$$

La sucesión $\{x_0^{(k)}\}$ es acotada. Elijamos en la misma una subsucesión $\{x_0^{k_p}\}$ que converge, en virtud de que X_0 es completo, a cierto vector $x_0' \in X_0$. Para este vector, evidentemente,

$$\| x' - x_0' \| = d. \tag{54.2}$$

Hagamos

$$x = \frac{1}{d} (x' - x_0').$$

Está claro que $\| x \| = 1$. Además, si $x_0 \in X_0$, entonces, de acuerdo con (54.1), tendremos

$$\| x - x_0 \| = \left\| \frac{1}{d} x' - \frac{1}{d} x_0' - x_0 \right\| = \frac{1}{d} \| x' - (x_0' + dx_0) \| \geqslant \frac{1}{d} d = 1,$$

puesto que el vector $x_0' + dx_0$ pertenece a X_0.

Hemos demostrado al mismo tiempo que la cota inferior (54.1) se logra por lo menos en un vector $x_0' \in X_0$. En la correlación $\| x - x_0 \| \geqslant 1$ la igualdad se consigue a ciencia cierta para $x_0 = 0$.

Observemos, como conclusión, que el lema 54.1 que desempeña un papel tan importante en los espacios de dimensión finita, no es válido *en ningún* espacio de dimensión infinita. A saber, es válido

LEMA 54.4. *Si en cada sucesión acotada de vectores del espacio normalizado X se puede elegir una subsucesión convergente, entonces el espacio X es de dimensión finita.*

DEMOSTRACION. Supongamos lo contrario. Sea X un espacio de dimensión infinita. Elijamos un vector normalizado arbitrario x_1 y designemos con L_1 su cápsula lineal. De acuerdo con el lema 54.3, existe un vector normalizado x_2 tal que $\| x_2 - x_1 \| \geqslant 1$. Designemos mediante L_2 la cápsula lineal de los vectores x_1, x_2. Continuando los razonamientos, hallemos la sucesión $\{x_n\}$ de los vectores normalizados que satisfacen las desigualdades $\| x_n - x_k \| \geqslant 1$ para todo $k < n$. Por consiguiente, de esta sucesión no se puede elegir ninguna subsucesión convergente. Esto contradice la hipótesis del lema, por lo cual, la suposición acerca de que el espacio X es de dimensión infinita era errónea.

Ejercicios.

1. Demuéstrese que un plano en el espacio normalizado de dimensión finita es un conjunto cerrado.
2. Demuéstrese que el conjunto de vectores x de un espacio de dimensión finita que satisfacen la condición $\| x \| \leqslant \alpha$, es un conjunto cerrado.
3. Demuéstrese que en un conjunto acotado cerrado de vectores de un espacio de dimensión finita existen unos vectores en los cuales se logran tanto la cota inferior como la superior de los valores de cualquier norma.
4. Demuéstrese que para cualesquiera dos normas $\| x \|_{\mathrm{I}}$, $\| x \|_{\mathrm{II}}$ en un espacio de dimensión finita existen tales números positivos α, β que

$$\alpha \| x \|_{\mathrm{I}} \leqslant \| x \|_{\mathrm{II}} \leqslant \beta \| x \|_{\mathrm{I}}$$

para todos los vectores x. Los números α, β no dependen de x.

§ 55. Límite y procesos de cálculo

En un espacio métrico completo el concepto de límite es de amplio uso al construir y argumentar los más diversos procesos de cálculo. Consideremos, a título de ejemplo, uno de los métodos de resolución de los sistemas de ecuaciones algebraicas lineales.

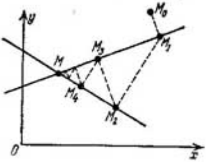


Fig. 96.1.

Sea dado un sistema de dos ecuaciones con dos incógnitas. Convengamos en considerar que el sistema es compatible y tiene una sola solución. Con el fin de simplificar la exposición, supongamos que todos los coeficientes son reales. Cada una de las ecuaciones

$$a_{11}x + a_{12}y = f_1,$$
$$a_{21}x + a_{22}y = f_2$$

del sistema define en el plano una recta. El punto M, donde se cortan estas rectas, determina la solución del sistema (fig. 55.1).

Tomemos un punto arbitrario M_0 que se halla en el plano fuera de las rectas mencionadas. Tracemos desde este punto una perpendicular a cualquiera de las rectas. El pie M_1 de la perpendicular queda más próximo al punto M que al punto M_0, puesto que una proyección es siempre menor que una oblicua. Tracemos a continuación una perpendicular desde el punto M_1 a la otra recta. El pie M_2

de esta última perpendicular se encontrará aún más próximo a la solución. Realizando sucesivamente la proyección, una vez a una recta y otra vez, a la otra, obtendremos una sucesión $\{M_k\}$ de puntos en el plano la cual converge al punto M. Merece subrayar que la convergencia de la sucesión construida tiene lugar para cualquier ubicación inicial del punto M_0.

Este ejemplo sugiere cómo se puede construir el proceso de cálculo para resolver un sistema de ecuaciones algebraicas lineales en la forma general (48.2). Sustituyamos el problema dado por uno equivalente que consiste en la búsqueda de los vectores de la intersección del sistema de hiperplanos (46.9). Supongamos que los hiperplanos contienen por lo menos un vector común y convengamos en considerar, para simplificar, que el espacio lineal es real.

Elijamos un vector arbitrario v_0 y proyectémoslo sobre el primer hiperplano. El vector obtenido v_1 lo proyectemos sobre el segundo hiperplano, etc. Este proceso determina cierta sucesión $\{v_p\}$. Investiguémosla.

El elemento principal de un proceso de cálculo consiste en proyectar cierto vector v_p sobre un hiperplano definido por la ecuación (46.8). Está claro que el vector v_{p+1} satisface esta ecuación y está ligado con el vector v_p mediante la igualdad

$$v_{p+1} = v_p + tn$$

para cierto número t. Al sustituir v_{p+1} en la ecuación (46.8), hallaremos t. De aquí obtenemos que

$$v_{p+1} = v_p + \left(\frac{b_p - (n,\ v_p)}{(n,\ n)}\right) n.$$

De esta fórmula se desprende que todos los vectores de la sucesión $\{v_p\}$ se disponen en un plano obtenido mediante el desplazamiento de una cápsula lineal $L\ (n_1,\ n_2,\ \ldots,\ n_k)$ al vector v_0. Pero todos los vectores pertenecientes a la intersección de los hiperplanos (46.9) se ubican en otro plano, obtenido por desplazamiento del complemento ortogonal $L^{\perp}\ (n_1,\ n_2,\ \ldots,\ n_k)$. Existe un único vector z_0 que pertenece a ambos planos.

Si demostramos que una subsucesión de $\{v_p\}$, cualquiera que sea, converge a cierto vector perteneciente a los hiperplanos (46.9), entonces, por ser el plano cerrado, dicha sucesión convergerá precisamente a z_0. En este caso al vector z_0 convergerá también toda la sucesión $\{v_p\}$.

Para cualquier r los vectores $z_0 - v_{r+1}$, $v_{r+1} - v_r$ son ortogonales, por lo cual, de acuerdo con el teorema de Pitágoras,

$$\rho^2\ (z_0,\ v_r) = \rho^2\ (z_0,\ v_{r+1}) + \rho^2\ (v_r,\ v_{r+1}).$$

Sumando las igualdades obtenidas respecto de r, desde 0 hasta $p-1$, hallamos

$$\rho^2(z_0, v_0) = \rho^2(z_0, v_p) + \sum_{r=0}^{p-1} \rho^2(v_r, v_{r+1}).$$

Por consiguiente,

$$\sum_{r=0}^{p-1} \rho^2(v_r, v_{r+1}) \leqslant \rho^2(z_0, v_0),$$

de donde concluimos que

$$\rho(v_p, v_{p+1}) \to 0. \tag{55.1}$$

Denotemos mediante H_r el hiperplano en la r-ésima fila (46.9). Está claro que la distancia del vector v_p hasta H_r no es superior a la distancia entre v_p y cualquier vector de H_r. De acuerdo con la construcción de $\{v_p\}$, entre cualesquiera k vectores sucesivos suyos hay sin falta un vector perteneciente a cualquiera de los hiperplanos. Haciendo uso de la desigualdad triangular y correlación límite (55.1), obtenemos

$$\rho(v_p, H_r) \leqslant \rho(v_p, v_{p+1}) + \rho(v_{p+1}, v_{p+2}) + \ldots$$
$$\ldots + \rho(v_{p+k-1}, v_{p+k}) \to 0 \tag{55.2}$$

para todo $r = 1, 2, \ldots, k$.

La sucesión $\{v_p\}$ es, evidentemente, acotada. Elijamos de ella una subsucesión convergente. Supongamos que la subsucesión converge al vector z_0'. Pasando al límite en (55.2), encontramos que

$$\rho(z_0', H_r) = 0$$

para todo $r = 1, 2, \ldots, k$. Pero, como ya se ha observado más arriba, el vector z_0' debe coincidir con z_0. Por consiguiente, la sucesión $\{v_p\}$ converge a z_0.

Ejercicios.

1. ¿Se han usado en su esencia los conceptos de completitud y de carácter cerrado en la investigación realizada?
2. ¿Cómo se pueden hallar otras soluciones del sistema, si existen?
3. ¿Cuál es el comportamiento del proceso, si el sistema es incompatible?