Ya.S.Bugrov
S.M.Nikolski

Matemáticas superiores

Elementos de álgebra lineal y geometría analítica

Editorial
Mir
Moscú

ИЗДАТЕЛЬСТВО • МИР •

Я. С. БУГРОВ С. М. НИКОЛЬСКИЙ

ВЫСШАЯ МАТЕМАТИКА
ЭЛЕМЕНТЫ ЛИНЕЙНОЙ АЛГЕБРЫ И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

ИЗДАТЕЛЬСТВО «НАУКА»

Ya. S. Bugrov
S. M. Nikolski

Matematicas superiores

Elementos de álgebra lineal y geometría analítica

Editorial Mir Moscú

Versión española por
ANTONIO APARICIO CORTÉS,
Licenciado en Ciencias Matemáticas

Revisada por
EMILIANO APARICIO BERNARDO
Candidato a Doctor en Ciencias Físico-Matemáticas
por la Universidad Estatal Lomonósov de Moscú,
Doctor en Ciencias Matemáticas
por la Universidad de Bilbao.

На испанском языке

Impreso en la URSS

Indice

Prólogo 7

§ 1. Determinantes de segundo orden 8
§ 2. Determinantes de tercero y n-ésimo orden 9
2.1. Determinantes de tercer orden 9
2.2. Determinantes de n-ésimo orden 11
§ 3. Matrices 18
§ 4. Sistema de ecuaciones lineales. Teoría de Kronecker—Capelli 20
4.1. Sistema de n ecuaciones lineales con n incógnitas 20
4.2. Fórmulas de Cramer 21
4.3. Sistema homogéneo 22
4.4. Reglas para la resolución de un sistema de ecuaciones lineales 23
4.5. Ejemplos de aplicación de las reglas 24
4.6. Fundamentación de las reglas 28
4.7. Método de resolución de un sistema mediante eliminación de incógnitas 30
4.8. Cálculo del rango de una matriz 33
§ 5. Espacio tridimensional. Vectores. Sistema cartesiano de coordenadas 36
5.1. Concepto de vector 36
5.2. Proyección de un vector 38
5.3. Propiedades de las proyecciones de los vectores 39
5.4. Producto escalar de vectores 40
5.5. Sistema rectangular de coordenadas 41
5.6. Problemas 45
§ 6. Espacio euclídeo n-dimensional. Producto escalar 45
6.1. Espacio n-dimensional R_n 45
6.2. Producto escalar en el espacio real R_n . . . 47
6.3. Producto escalar en el espacio complejo R_n 47
6.4. Desigualdad de Buniakovski 48
6.5. Desigualdad de Minkowski 59
§ 7. Segmento. División de un segmento en una razón dada 51
§ 8. La recta 53
§ 9. Ecuación del plano 60
9.1. Ecuación del plano en la forma normal . . . 60
9.2. Ecuación del plano en la forma general . . 61
9.3. Ecuación «segmentaria» del plano 63
9.4. Ecuación del plano que pasa por un punto 63
9.5. Ecuación del plano que pasa por tres puntos 64
9.6. Angulo formado por dos planos 65
9.7. Distancia de un punto a un plano 66
9.8. Problemas 67
§ 10. La recta en el espacio 68
10.1. Ecuación de la recta en la forma canónica 68

10.2. Posición recíproca de dos planos 69
10.3. Problemas 70
§ 11. Orientación de los sistemas rectangulares de coordenadas 71
11.1. Sistema bidimensional de coordenadas . . . 71
11.2. Sistema tridimensional de coordenadas . . 72
§ 12. Producto vectorial 73
12.1. Dos definiciones de producto vectorial . . . 73
12.2. Significado geométrico del determinante de segundo orden 76
12.3. Propiedades del producto vectorial 77
§ 13. Producto vectorial-escalar (mixto) 79
§ 14. Sistema de vectores linealmente independiente . . 80
§ 15. Operadores lineales 87
§ 16. Bases en R_n 94
§ 17. Bases ortogonales en R_n 98
§ 18. Propiedades invariantes de los productos vectorial y escalar 105
§ 19. Transformación de las coordenadas rectangulares en el plano 107
§ 20. Subespacios lineales en R_n 110
§ 21. Teoremas de tipo Fredholm 115
§ 22. Operador autoconjugado. Forma cuadrática . . . 122
§ 23. Forma cuadrática en el espacio bidimensional . . 130
§ 24. Curvas de segundo orden 134
§ 25. Superficie de segundo orden en el espacio tridimensional 146
§ 26. Teoría general de la superficie de segundo orden en el espacio tridimensional 161

Indice alfabético 167

Prólogo

El presente libro es la primera parte de nuestra obra «Matemáticas Superiores». En él se exponen las cuestiones principales de la teoría de los determinantes, los elementos de la teoría de matrices, teoría de sistemas de ecuaciones lineales, álgebra vectorial. También se estudian los temas fundamentales del álgebra lineal: operadores lineales, transformaciones ortogonales, operadores autoconjugados (hermíticos), forma cuadrática y su reducción a la forma canónica.

Se incluyen los elementos de geometría analítica: la recta, el plano, la recta en el espacio y las curvas y superficies de segundo orden.

Por regla general, los razonamientos van acompañados de demostraciones completas. No obstante, la exposición se hace de tal modo que puedan ser omitidas las demostraciones del caso general n-dimensional, conservando no sólo los enunciados de los teoremas sino también la explicación detallada de lo que ocurre en cada caso para dos o tres dimensiones.

Las formas canónicas de las curvas y superficies de segundo orden se exponen muy abreviadamente, ya que se supone que en adelante éstas se estudiarán complementariamente como problemas por los métodos del análisis matemático. La forma cuadrática se estudia por los métodos del análisis matemático o, mejor dicho, por los métodos del análisis funcional.

Aunque a este libro lo llamamos primero de nuestra serie, en realidad, el material del mismo y el del segundo libro (dedicado al cálculo diferencial e integral) tienen una estrecha conexión. Es bien sabido el orden en que se debe presentar el material contenido en ambos.

El libro abarca todas las cuestiones que comprenden los programas de los centros de enseñanza técnica superior (con un volumen de 400—500 horas lectivas).

§ 1. Determinantes de segundo orden

Sean dados los números a_1, a_2, b_1, b_2 (reales o complejos). Estos determinan un número $a_1b_2 - a_2b_1$, que se llama *determinante de segundo orden* y se escribe así:

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1. \tag{1}$$

Los números a_1, a_2, b_1, b_2 se llaman *elementos del determinante*. En el determinante (1) se distinguen *la primera fila* a_1, a_2 y la *segunda fila* b_1, b_2, la *primera columna* a_1, b_1 y la *segunda columna* a_2, b_2.

Fácilmente se comprueban *las siguientes propiedades del determinante*:

El valor del determinante

a) *no varía al sustituir las filas por las columnas correspondientes*:

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix};$$

b) *cambia su signo al permutar las filas* (*o las columnas*):

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = -\begin{vmatrix} b_1 & b_2 \\ a_1 & a_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = -\begin{vmatrix} a_2 & a_1 \\ b_2 & b_1 \end{vmatrix};$$

c) *queda multiplicado por k si se multiplican los elementos de una fila o columna por k (real o complejo)*. Por ejemplo:

$$\begin{vmatrix} ka_1 & b_1 \\ ka_2 & b_2 \end{vmatrix} = k\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},$$

o sea, *el factor común a todos los elementos de una fila o una columna se puede sacar como factor fuera del determinante*;

d) *es igual a cero si los elementos de alguna fila o columna son iguales a cero*. Por ejemplo:

$$\begin{vmatrix} 0 & 0 \\ b_1 & b_2 \end{vmatrix} = 0b_2 - 0b_1 = 0,$$

e) *es igual a cero si los elementos de dos filas o dos columnas son respectivamente iguales*

$$\begin{vmatrix} a_1 & a_1 \\ a_2 & a_2 \end{vmatrix} = a_1a_2 - a_1a_2 = 0.$$

A continuación se introducen los determinantes de tercero y, en general, de n-ésimo orden. Para éstos se conservan las propiedades a), b), c), d) y e).

§ 2. Determinantes de tercero y n-ésimo orden

2.1. Determinantes de tercer orden

El número

$$\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \quad (1)$$

expresado en la forma

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad (2)$$

donde a_{kl} son unos números (reales o complejos), se llama *determinante de tercer orden*.

En el determinante (2) se distinguen la primera, segunda y tercera filas, así como la primera, segunda y tercera columnas. El número a_{kl} se llama *elemento* del determinante; el primer subíndice k denota el número de orden de la fila, mientras que el segundo subíndice l, el número de orden de la columna. También diremos que el elemento a_{kl} está situado en la intersección de la k-ésima fila y l-ésima columna. Los elementos a_{11}, a_{22}, a_{33} forman *la diagonal principal* del determinante, y los elementos a_{13}, a_{22}, a_{31}, la *diagonal secundaria*.

La estructura de la expresión (1) es bastante sencilla. Representa un número que se calcula según los elementos a_{kl} de acuerdo con la clara regla (de Sarrus) siguiente: formemos la tabla (de Sarrus) obtenida por los elementos del determinante (2) añadiendo la primera y segunda columnas del determinante (fig. 1). Vemos que hay que tomar todos los productos posibles de los elementos borrados por las rectas; los tres productos correspondientes a las rectas paralelas a la diagonal principal se toman con el signo más, mientras que los otros tres productos correspondientes a las rectas paralelas a la diagonal secundaria se toman con el signo menos.

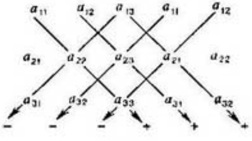

Fig. 1

Cada uno de los productos junto con el signo señalado se llama *término del determinante* (2). Entre los elementos que figuran en un producto hay representantes de cada una de las filas y de cada una de las columnas. Estos elementos pueden colocarse en cada uno de los términos en orden de crecimiento del primer subíndice, o sea, de los números de orden de las filas a las que pertenecen. Precisamente esto se ha hecho en la suma (1). En lo que se refiere a los subíndices de las columnas a las que pertenecen estos elementos, su ordenación viene dada a continuación:

$$\left.\begin{matrix} 1 & 2 & 3, \\ 2 & 3 & 1, \\ 3 & 1 & 2, \end{matrix}\right\} \tag{3}$$

$$\left.\begin{matrix} 3 & 2 & 1, \\ 1 & 3 & 2, \\ 2 & 1 & 3. \end{matrix}\right\} \tag{4}$$

Estas son todas las permutaciones posibles de los números 1, 2, 3. La permutación

$$1,\ 2,\ 3 \tag{5}$$

se llama *principal*.

Se dice que en una permutación se ha efectuado una *transposición de dos elementos determinados*, si estos elementos han sido intercambiados de sitio. Después de una transposición la permutación se convierte en otra permutación. A su vez, en esta última se puede hacer una transposición, obteniéndose una tercera permutación (no se excluye que resulte la primera). Por ejemplo, la permutación

$$3,\ 2,\ 1 \tag{6}$$

se ha obtenido por transposición del primero y tercer elementos de la permutación (5), y la permutación

$$2,\ 3,\ 1 \tag{7}$$

por transposición del primero y segundo elementos de la permutación (6).

Es importante señalar que, si una permutación se ha obtenido de la principal mediante N transposiciones y esta misma permutación se ha obtenido de la principal de otro modo mediante N_1 transposiciones, entonces ambos números N y N_1 son simultáneamente pares o impares. Una permutación de los números 1, 2, 3 se llama *par* (o *impar*), si se obtiene de la permutación principal mediante un número par (impar) de transposiciones.

Sea dada una permutación $j = (j_1, j_2, j_3)$, donde j_1, j_2, j_3 son los números 1, 2, 3, tomados en cierto orden. El *número de transposiciones* mediante los cuales se puede obtener esta permutación a partir de la permutación principal lo denotaremos por $t(j)$. Entonces, la permutación j es par (impar), si $t(j)$ es un número par (impar).

Las permutaciones (3) son pares, mientras que las permutaciones (4) son impares.

Después de todo lo dicho, se puede dar otra definición equivalente del determinante de 3-er orden.

Se llama *determinante de 3-er orden* (2) el número Δ, igual a la suma de todos los productos de la forma $(-1)^{t(j)} a_{1j_1} a_{2j_2} a_{3j_3}$, donde $j = (j_1, j_2, j_3)$ recorre todas las permutaciones posibles de la permutación principal 1, 2, 3:

$$\Delta = \sum_j (-1)^{t(j)} a_{1j_1} a_{2j_2} a_{3j_3}. \tag{8}$$

Esta definición se generaliza a los determinantes de n-ésimo orden.

2.2. Determinantes de n-ésimo orden

Se llama *determinante de n-ésimo orden* y se denota por

$$\Delta = |a_{ik}| = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \tag{9}$$

el número que se calcula a partir de unos números dados a_{ik}, (reales o complejos) denominados elementos del determinante, según la regla siguiente: Δ es la suma

$$\Delta = \sum_j (-1)^{t(j)} a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

extendida a todas las permutaciones distintas posibles $j = (j_1, \ldots, j_n)$ de los números $1, 2, \ldots, n$. El valor de $t(j)$ es igual *al número de transposiciones* que hay que realizar para pasar de la permutación principal $1, 2, \ldots, n$ a la permutación $j = (j_1, \ldots, j_n)$. El producto $(-1)^{t(j)} a_{1j_1} \ldots a_{nj_n}$ se llama *término del determinante.*

Los determinantes de n-ésimo orden satisfacen las propiedades a), b), c), d), e) *enunciadas en el párrafo anterior.*

DEMOSTRACION. a) Después de cambiar en el determinante las filas por las columnas correspondientes, los números de las filas quedarán denotados como segundos subíndices. Por ejemplo, para el determinante de tercer orden (2), tendremos:

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

$$- a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} = \Delta$$

En el caso general, el término general del nuevo determinante se expresa así:

$$(-1)^{t(j)} a_{j_1 1} a_{j_2 2} \ldots a_{j_n n}.$$

Reordenemos los factores del producto $a_{j_1 1} a_{j_2 2} \ldots a_{j_n n}$ según el primer subíndice, para lo que tendremos que pasar de la permutación $j = (j_1, \ldots, j_n)$ a la permutación principal $1, 2, \ldots, n$. Entonces habrá que realizar $t(j)$ transposiciones. Con ello, la permutación principal de los segundos subíndices se transformará en una cierta permutación $i = (i_1, \ldots, i_n)$ y el número $t(i)$ será de la misma paridad que el número $t(j)$. Por lo tanto,

$$(-1)^{t(j)} a_{j_1 1} \ldots a_{j_n n} = (-1)^{t(i)} a_{1 i_1} \ldots a_{n i_n}$$

Fácilmente se observa que a distintas permutaciones $j_1, \ldots, j_n$ les corresponden distintas permutaciones $i_1, \ldots, i_n$. Pero, entonces

$$\sum_j (-1)^{t(j)} a_{j_1 1} \ldots a_{j_n n} = \sum_i (-1)^{t(i)} a_{1 i_1} \ldots a_{n i_n} = \Delta$$

b) Intercambiemos, por ejemplo, la primera y tercera filas del determinante de tercer orden (2). Entonces obtendremos un determinante que denotaremos por Δ'; éste será igual a

$$\Delta' = \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = \sum_j (-1)^{t(j)} a_{3j_1} a_{2j_2} a_{1j_3} =$$

$$= \sum_j (-1)^{t(j)} a_{1j_3} a_{2j_2} a_{3j_1} = - \sum_{j'=(j_3, j_2, j_1)} (-1)^{t(j')} a_{1j_3} a_{2j_2} a_{3j_1} = -\Delta,$$

puesto que la permutación $j = (j_1, j_2, j_3)$ se diferencia de la permutación $j' = (j_3, j_2, j_1)$ en una transposición.

Diremos que una fila (o columna) del determinante se ha multiplicado por un número k, si todos los elementos de esta fila (o columna) se han multiplicado por k.

c) La multiplicación de una fila (o columna) del determinante por un número k se reduce a la multiplicación de todos sus términos por k, puesto que cada uno de los términos contiene un elemento de la fila (o columna) en cuestión. Pero, entonces, el valor de la suma de los términos queda multiplicado por k.

d) Un determinante en el que los elementos de alguna fila o columna son iguales a cero, es también igual a cero, puesto que todos sus términos, evidentemente, son iguales a cero.

e) El determinante es igual a cero si tiene dos filas o dos columnas respectivamente iguales. Esto es consecuencia de la propiedad b) ($\Delta' = -\Delta$, $\Delta' = \Delta$, de donde $\Delta = 0$).

Suprimamos en el determinante de n-ésimo orden (9) la i-ésima fila y la k-ésima columna. La expresión restante engendra un determinante de $(n-1)$-ésimo orden M_{ik}, denominado *menor del elemento* a_{ik}. El valor

$$A_{ik} = (-1)^{i+k} M_{ik}$$

se llama *complemento algebraico o adjunto del elemento* a_{ik}.

PROPIEDAD f). *La suma de los productos de los elementos* a_{ik} *de una fila (o columna) del determinante por los complementos algebraicos de estos elementos es igual al valor del determinante*:

$$\Delta = \sum_{k=1}^{n} a_{ik}A_{ik} \quad (i = 1, \ldots, n), \tag{10}$$

$$\Delta = \sum_{i=1}^{n} a_{ik}A_{ik} \quad (k = 1, \ldots, n). \tag{10'}$$

Demostremos esta propiedad para los determinantes de tercer orden en el caso de la tercera fila. Se tiene

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

La suma (10) se llama *desarrollo del determinante por los elementos de la i-ésima fila*, y la suma (10'), *desarrollo del determinante por los elementos de la k-ésima columna*.

EJEMPLO 1. Si en el determinante Δ (véase (9)), $a_{21} = a_{31} = \ldots \ldots = a_{n1} = 0$, entonces $\Delta = a_{11}A_{11}$, o sea, el cálculo de este determinante se reduce al cálculo de un adjunto del mismo, que es un determinante de $(n-1)$-ésimo orden.

EJEMPLO 2. Si todos los elementos de Δ, situados debajo (encima) de la diagonal principal de Δ, son iguales a cero ($a_{kl} = 0$, si $k > l$ ($k < l$)), entonces $\Delta = a_{11}a_{22} \ldots a_{nn}$. Esto se deduce del ejemplo anterior.

PROPIEDAD g) *La suma de los productos de los elementos* a_{ik} *de una fila (o columna) del determinante por los adjuntos correspondientes de* *los elementos de otra fila (o columna) es igual a cero:*

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \sum_{k=1}^{n} a_{kj}A_{kl} = 0 \tag{11}$$

$(i \neq j;\ i,\ j = 1,\ \ldots,\ n)$.

En efecto, fijemos nuestra atención en la primera suma. Esta no depende de los elementos de la j-ésima fila. Sustituyamos en el determinante los elementos de la j-ésima fila por los elementos correspondientes de la i-ésima fila. Con ello, la suma en cuestión no varía. Sin embargo, ésta puede considerarse como el desarrollo del nuevo determinante por los elementos de la j-ésima fila, y entonces, es igual al valor del nuevo determinante. Pero este último es igual a cero según la propiedad e), ya que tiene dos filas iguales, la i-ésima y la j-ésima.

PROPIEDAD h) *Sean dados dos determinantes de n-ésimo orden Δ_1 y Δ_2, tales que todas sus filas (o columnas) son iguales salvo una determinada. La suma de tales determinantes es igual a un determinante Δ de n-ésimo orden en el que la fila (columna) en cuestión está formada por la suma de los elementos correspondientes de esta fila (columna) de los determinantes Δ_1 y Δ_2.* Por ejemplo:

$$\Delta_1 + \Delta_2 = \begin{vmatrix} a_{11} & \ldots & a_{1,\, n-1} a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,\, n-1} a_{nn} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{11} & \ldots & a_{1,\, n-1} b_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,\, n-1} b_{1n} \end{vmatrix} = \begin{vmatrix} a_{11} & \ldots & a_{1,\, n-1} (a_{1n} + b_{1n}) \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,\, n-1} (a_{nn} + b_{nn}) \end{vmatrix} = \Delta.$$

En efecto, desarrollando los determinantes dados por los elementos de la n-ésima columna, obtenemos:

$$\Delta_1 + \Delta_2 = \sum_{k=1}^{n} a_{kn}A_{kn} + \sum_{k=1}^{n} b_{kn}A_{kn} = \sum_{k=1}^{n} (a_{kn} + b_{kn}) A_{kn} = \Delta.$$

PROPIEDAD i) *El valor de un determinante no varía si a los elementos de una fila (o columna) se les añaden los elementos correspondientes de otra fila (columna), multiplicados por un número k.* Por ejemplo,

$$\begin{vmatrix} a_{11} & \ldots & a_{1,\, n-1} a_{1n} + k a_{11} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,\, n-1} a_{nn} + k a_{n1} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} + k \begin{vmatrix} a_{11} & \ldots & a_{1,\, n-1} a_{11} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{n,\, n-1} a_{n1} \end{vmatrix} = |a_{kl}| + k \cdot 0 = |a_{kl}|,$$

en virtud de las propiedades h), c) y e).

La aplicación adecuada de esta propiedad reduce el cálculo de un determinante al cálculo de otro determinante de orden inferior.

Ejemplo 3.

$$\begin{vmatrix} 1 & 4 & 5 \\ 2 & 3 & 1 \\ 8 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 5 \\ 0 & -5 & -9 \\ 8 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 5 \\ 0 & -5 & -9 \\ 0 & -31 & -39 \end{vmatrix} =$$

$$= \begin{vmatrix} -5 & -9 \\ -31 & -39 \end{vmatrix} = \begin{vmatrix} 5 & 9 \\ 31 & 39 \end{vmatrix} = 5\cdot 39 - 9\cdot 31 = 84.$$

Ejemplo 4.

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ -5 & 3 & -1 \\ -7 & 4 & 2 \end{vmatrix} = -\begin{vmatrix} -5 & -1 \\ -7 & 2 \end{vmatrix} = \begin{vmatrix} 5 & 1 \\ -7 & 2 \end{vmatrix} = 17.$$

Ejemplo 5. El determinante

$$\Delta_n = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix},$$

engendrado por los números $a_1, a_2, \dots, a_n$, se llama *determinante de Vandermonde* [1]).

Este determinante es igual a cero si algún par de números a_j y a_k son iguales entre sí. Si todos los a_j son distintos, entonces,

$$\begin{aligned} \Delta_n = {} & (a_n - a_1)(a_{n-1} - a_1) \dots (a_2 - a_1) \cdot \\ & \cdot (a_n - a_2)(a_{n-1} - a_2) \dots (a_3 - a_2) \cdot \\ & \dots\dots\dots\dots\dots\dots \\ & \cdot (a_n - a_{n-2})(a_{n-1} - a_{n-2}) \cdot \\ & \cdot (a_n - a_{n-1}). \end{aligned} \tag{12}$$

En efecto, para $n = 2$,

$$\Delta_2 = \begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = a_2 - a_1,$$

o sea, es válida la fórmula (12). Supongamos que esta fórmula es válida para $n = k - 1$ y demostremos que también es válida para $n = k$. Aplicaremos las propiedades i) y c) del determinante. Multipliquemos la $(k-1)$-ésima columna del determinante Δ_k por a_1 y restémosla de la k-ésima, multipliquemos la $(k-2)$-ésima columna

[1]) A. T. Vandermonde (1735—1796), matemático francés.

también por a_1 y restémosla de la $(k-1)$-ésima, etc.; entonces obtendremos:

$$\Delta_k = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_2-a_1 & a_2(a_2-a_1) & \dots & a_2^{k-2}(a_2-a_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & a_k-a_1 & a_k(a_k-a_1) & \dots & a_k^{k-2}(a_k-a_1) \end{vmatrix} =$$

$$= (a_k-a_1)(a_{k-1}-a_1)\dots(a_2-a_1)\begin{vmatrix} 1 & a_2 & \dots & a_2^{k-2} \\ 1 & a_3 & \dots & a_3^{k-2} \\ \dots & \dots & \dots & \dots \\ 1 & a_k & \dots & a_k^{k-2} \end{vmatrix}.$$

El último determinante también es un determinante de Vandermonde de orden $(k-1)$, engendrado por los números $a_2, \dots, a_k$, por lo cual, según la hipótesis, se tiene:

$$\begin{aligned} \Delta_k = {} & (a_k-a_1)(a_{k-1}-a_1)\dots(a_2-a_1)\cdot \\ & \cdot(a_k-a_2)(a_{k-1}-a_2)\dots(a_3-a_2)\cdot \\ & \dots\dots\dots\dots\dots\dots\dots \\ & \cdot(a_k-a_{k-2})(a_{k-1}-a_{k-2})\cdot \\ & \cdot(a_k-a_{k-1}). \end{aligned}$$

Así pues, en virtud del método de inducción matemática la fórmula (12) es válida para cualquier $n \geqslant 2$.

PROPIEDAD j) *Sean*

$$\Delta_1 = |b_{kl}|, \quad \Delta_2 = |a_{kl}|.$$

El producto de dos determinantes de orden n con los elementos b_{kl}, a_{kl}, es a su vez un determinante de orden n con los elementos

$$\gamma_{kl} = \sum_{j=1}^{n} b_{kj}a_{jl},$$

o sea,

$$\Delta_1\Delta_2 = \left|\sum_{j=1}^{n} b_{kj}a_{jl}\right| = \Delta.$$

Así pues, el elemento γ_{kl}, perteneciente a la k-ésima fila y la l-ésima columna del determinante Δ, es igual, como suele decirse, *al producto de la k-ésima fila del determinante Δ_1 por la l-ésima columna del determinante Δ_2*. En realidad, esto es *la suma de los productos* de los elementos de la k-ésima fila del determinante Δ_1 por los elementos correspondientes de la l-ésima columna del determinante Δ_2.

Como en los determinantes Δ_1 y Δ_2 se pueden cambiar filas por columnas resulta, evidentemente, que los elementos γ_{kl} del producto Δ también se puede obtener tomando el producto de la k-ésima fila de Δ_1 por la l-ésima fila de Δ_2, o bien, tomando el producto de la k-ésima columna de Δ_1 por la l-ésima columna de Δ_2, o finalmente, tomando el producto de la k-ésima columna de Δ_1 por la l-ésima fila de Δ_2.

DEMOSTRACIÓN. Comprobemos la propiedad en el caso de determinantes de segundo orden:

$$\Delta_1=\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}, \quad \Delta_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta=\begin{vmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{vmatrix},$$

donde

$$\gamma_{11}=b_{11}a_{11}+b_{12}a_{21}, \quad \gamma_{12}=b_{11}a_{12}+b_{12}a_{22},$$
$$\gamma_{21}=b_{21}a_{11}+b_{22}a_{21}, \quad \gamma_{22}=b_{21}a_{12}+b_{22}a_{22}.$$

En virtud de las propiedades h), c) y e), se tiene:

$$\Delta=\begin{vmatrix} b_{11}a_{11} & b_{11}a_{12}+b_{12}a_{22} \\ b_{21}a_{11} & b_{21}a_{12}+b_{22}a_{22} \end{vmatrix}+\begin{vmatrix} b_{12}a_{21} & b_{11}a_{12}+b_{12}a_{22} \\ b_{22}a_{21} & b_{21}a_{12}+b_{22}a_{22} \end{vmatrix}=$$
$$=a_{11}a_{12}\begin{vmatrix} b_{11} & b_{11} \\ b_{21} & b_{21} \end{vmatrix}+a_{11}a_{22}\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}+a_{12}a_{21}\begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{vmatrix}+$$
$$+a_{21}a_{22}\begin{vmatrix} b_{12} & b_{12} \\ b_{22} & b_{22} \end{vmatrix}=a_{11}a_{12}\cdot 0+a_{11}a_{22}\Delta_1+a_{12}a_{21}\begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{vmatrix}+a_{21}a_{22}\cdot 0=$$
$$=a_{11}a_{22}\Delta_1-a_{12}a_{21}\Delta_1=\Delta_1(a_{11}a_{22}-a_{12}a_{21})=\Delta_1\Delta_2.$$

En el caso general de determinantes de n-ésimo orden, podemos escribir:

$$\Delta=\begin{vmatrix} \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_1 1} & \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_1 2} \cdots & \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_1 n} \\ \sum\limits_{s_2=1}^{n} a_{2s_2}b_{s_2 1} & \sum\limits_{s_2=1}^{n} a_{2s_2}b_{s_2 2} \cdots & \sum\limits_{s_2=1}^{n} a_{2s_2}b_{s_2 n} \\ \cdots & \cdots & \cdots \\ \sum\limits_{s_n=1}^{n} a_{ns_n}b_{s_n 1} & \sum\limits_{s_n=1}^{n} a_{ns_n}b_{s_n 2} \cdots & \sum\limits_{s_n=1}^{n} a_{ns_n}b_{s_n n} \end{vmatrix}=$$

$$=\sum_{s_1=1}^{n}\sum_{s_2=1}^{n}\cdots\sum_{s_n=1}^{n} a_{1s_1}a_{2s_2}\cdots a_{ns_n}\begin{vmatrix} b_{s_1 1} & b_{s_1 2} & \cdots & b_{s_1 n} \\ b_{s_2 1} & b_{s_2 2} & \cdots & b_{s_2 n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{s_n 1} & b_{s_n 2} & \cdots & b_{s_n n} \end{vmatrix}=$$

$$=\sum_{s=(s_1,\, s_2,\, \ldots,\, s_n)} a_{1s_1}a_{2s_2}\cdots a_{ns_n}(-1)^{t(s)}\Delta_1=\Delta_2\Delta_1.$$

Al calcular cada uno de los elementos de Δ podemos tomar cualquier subíndice de sumación s $(\gamma_{kl} = \sum_{s=1}^{n} a_{ks}b_{sl})$, pero, para lo que sigue, resulta más cómodo tomar para la primera fila de Δ el subíndice s_1, para la segunda fila, el subíndice s_2, y así sucesivamente. La segunda igualdad se verifica en virtud de las propiedades h) y c); además, la suma múltiple $\sum_{s_1=1}^{n} \sum_{s_2=1}^{n} \dots \sum_{s_n=1}^{n}$ se extiende a todas las permutaciones posibles $(s_1, s_2, \dots, s_n)$, donde $1 \leqslant s_j \leqslant n$. No obstante, si en alguno de los sistemas $(s_1, s_2, \dots, s_n)$ dos componentes s_i y s_j son iguales entre sí $(s_i = s_j,\ i \neq j)$, entonces el determinante $|b_{s_k l}| = 0$. Por esta razón, en la realidad, en la suma múltiple se pueden dejar solamente los términos que corresponden a distintas permutaciones $(s_1, \dots, s_n)$ de los números naturales $(1, \dots, n)$. Además, evidentemente, resultará que el determinante

$$|b_{s_k l}| = (-1)^{I(s)}\Delta_1.$$

§ 3. Matrices

Una tabla de números α_{ij} (reales o complejos) de la forma

$$A = \begin{Vmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \alpha_{m1} & \dots & \alpha_{mn} \end{Vmatrix} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} = \| \alpha_{ij} \| = (\alpha_{ij}), \tag{1}$$

compuesta por m filas y n columnas, se llama *matriz*. Los números α_{ij} se llaman *elementos* de la misma. Esta es una matriz rectangular. Si $m = n$, se llama *matriz cuadrada* de orden n.

Una segunda matriz $B = \| \beta_{ij} \|$, con los elementos β_{ij}, compuesta por m filas y n columnas, se considera *igual* a la matriz A si, y sólo si, los elementos correspondientes de ambas matrices son iguales $(\alpha_{ij} = \beta_{ij})$. En este caso, se escribe $A = B$. Una matriz $\| \alpha_{ij} \|$ no es un número, es una tabla. Sin embargo, para una matriz cuadrada se puede considerar un número $| \alpha_{ij} |$, *el determinante engendrado por esta matriz.*

Sea k un número natural no superior a m y a n $(k \leqslant m, n)$. Suprimamos en la matriz (1) k columnas y k filas cualesquiera. Los elementos α_{js} situados en la intersección de las columnas y filas suprimidas forman una matriz cuadrada, la cual engendra un determinante de k-ésimo orden. El determinante obtenido se llama *determinante de k-ésimo orden engendrado por la matriz A*.

Se llama *rango* de la matriz A al máximo número natural k tal para el cual existe un determinante no nulo de k-ésimo orden engendrado por la matriz A (véase § 4).

Si en la matriz A se cambian las filas por las columnas del mismo orden, resulta la matriz

$$A' = \left\| \begin{matrix} \alpha_{11} & \dots & \alpha_{m1} \\ \cdot & \cdot\cdot\cdot\cdot & \cdot \\ \alpha_{1n} & \dots & \alpha_{mn} \end{matrix} \right\|, \tag{2}$$

denominada *matriz traspuesta de la matriz A*.

Si en la matriz A se sustituyen sus elementos a_{kl} por los complejos conjugados, entonces resulta la matriz

$$\overline{A} = \left\| \begin{matrix} \overline{a}_{11} & \dots & \overline{a}_{1n} \\ \cdot & \cdot\cdot\cdot & \cdot \\ \overline{a}_{n1} & \dots & \overline{a}_{nn} \end{matrix} \right\| = \| \overline{a}_{kl} \|,$$

denominada *matriz compleja conjugada de A*.

La matriz

$$\overline{A'} = (\overline{A})' = \left\| \begin{matrix} \overline{a}_{11} & \dots & \overline{a}_{n1} \\ \cdot & \cdot\cdot\cdot\cdot & \cdot \\ \overline{a}_{1n} & \dots & \overline{a}_{nn} \end{matrix} \right\| = A^*$$

se llama *matriz conjugada de A*.

Si A es una matriz real, o sea, con elementos reales ($\overline{a}_{kl} = a_{kl}$), entonces, evidentemente,

$$A = \overline{A}, \quad A' = A^*.$$

Las matrices de una misma dimensión, o sea, compuestas por el mismo número de filas y columnas, pueden sumarse. Se llama *suma* de dos matrices tales $A = \| \alpha_{ij} \|$ y $B = \| \beta_{ij} \|$ la matriz $C = \| \gamma_{ij} \|$ cuyos elementos son iguales a la suma de los elementos correspondientes de las matrices A y B: $\gamma_{ij} = \alpha_{ij} + \beta_{ij}$. Simbólicamente se escribe así:

$$A + B = C.$$

Fácilmente se observa que

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

Se llama *producto de un número λ por la matriz A* (o producto de la matriz A por el número λ), la matriz cuyos elementos son iguales al producto del número λ por los elementos correspondientes de la matriz A. Por lo tanto, $\lambda A = A\lambda$.

EJEMPLO. Sean

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

Hallar la matriz $\lambda A + \mu B$.

Según la definición de suma de matrices y del producto de una matriz por un número, se tiene:

$$\lambda A=\begin{pmatrix}\lambda & 0 & 2\lambda\\ 2\lambda & 3\lambda & \end{pmatrix},\quad \mu B=\begin{pmatrix}2\mu & 0 & \mu\\ 2\mu & \mu & \mu\end{pmatrix},$$

$$\lambda A+\mu B=\begin{pmatrix}\lambda+2\mu & 0 & 2\lambda+\mu\\ 2\lambda+2\mu & 3\lambda+\mu & \lambda+\mu\end{pmatrix}.$$

§ 4. Sistema de ecuaciones lineales. Teoría de Kronecker—Capelli [1]

4.1. Sistema de n ecuaciones lineales con n incógnicas

Un sistema ordenado arbitrario de n números $(x_1, \ldots, x_n)$ se llama *vector n-dimensional*; lo denotaremos también por una sola letra (negrita) $\boldsymbol{x}$:

$$\boldsymbol{x} = (x_1, \ldots, x_n).$$

Los números x_j (reales o complejos) se llaman *componentes* del vector $\boldsymbol{x}$. El vector

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

se llama *vector nulo*.

Consideremos un sistema de n ecuaciones lineales con n incógnitas

$$\left.\begin{array}{l} a_{11}x_1+\ \ldots +a_{1n}x_n=y_1,\\ \ldots\ldots\ldots\ldots\ldots\\ a_{n1}x_1+\ldots+a_{nn}x_n=y_n.\end{array}\right\} \tag{1}$$

Los números a_{kl} (reales o complejos), denominados *coeficientes del sistema* (1) se consideran dados. También se dice que el sistema (1) se determina por la matriz de sus coeficientes

$$A=\|a_{kl}\|=\left\|\begin{array}{c} a_{11}\ldots a_{1n}\\ \ldots\ldots\\ a_{n1}\ldots a_{nn}\end{array}\right\|. \tag{2}$$

A continuación nos va a interesar el problema de la resolubilidad del sistema (1) para cada vector (o sistema de números)

$$\boldsymbol{y} = (y_1, \ldots, y_n).$$

[1]) L. Kronecker (1823—1891), matemático alemán. A. Capelli (1855—1910), matemático italiano.

Un sistema de números (un vector)

$$x = (x_1, \ldots, x_n)$$

se llama *solución del sistema de ecuaciones* (1), si los números x_j satisfacen estas ecuaciones.

4.2. Fórmulas de Cramer

TEOREMA 1. *Si el determinante del sistema* (1) *es distinto de cero:*

$$\Delta = | a_{kl} | \neq 0,$$

el sistema (1) *admite solución única para cualquier vector* y; *esta solución se calcula por las fórmulas de Cramer* *):

$$x_j = \Delta^j/\Delta \quad (j = 1, \ldots, n), \tag{3}$$

donde Δ^j *es el determinante que se obtiene del determinante* Δ *al sustituir en el mismo los números de la j-ésima columna por los números* $y_1, \ldots, y_n$, *respectivamente*:

$$\Delta^j = \begin{vmatrix} a_{11} \ldots a_{1,j-1} & y_1 & a_{1,j+1} \ldots a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} \ldots a_{n,j-1} & y_n & a_{n,j+1} \ldots a_{nn} \end{vmatrix}. \tag{4}$$

Por lo tanto,

$$x_j = \frac{1}{\Delta} \sum_{s=1}^{n} A_{sj} y_s \quad (j = 1, \ldots, n), \tag{3'}$$

donde A_{sj} *es el adjunto del elemento* a_{sj} *en el determinante* Δ.

DEMOSTRACIÓN. Sea $(x_1, \ldots, x_n)$ una solución del sistema (1). Para hallar la incógnita x_1, multiplicamos la primera ecuación del sistema (1) por el adjunto A_{11}, la segunda por $A_{21}, \ldots$, la n-ésima por A_{n1} y sumamos todas las ecuaciones así obtenidas. Entonces, teniendo en cuenta que

$$\sum_{k=1}^{n} a_{k1} x_1 A_{k1} = x_1 \sum_{k=1}^{n} a_{k1} A_{k1} = x_1 \Delta$$

y que

$$\sum_{k=1}^{n} a_{kj} x_j A_{k1} = x_j \sum_{k=1}^{n} a_{kj} A_{k1} = x_j \cdot 0 = 0 \quad (j \neq 1),$$

obtenemos $x_1 \Delta = \Delta^1$, donde

$$\Delta^1 = \sum_{s=1}^{n} y_s A_{s1} = \begin{vmatrix} y_1 & a_{12} \ldots a_{1n} \\ \cdots & \cdots \\ y_n & a_{n2} \ldots a_{nn} \end{vmatrix}.$$

Por consiguiente, como según la hipótesis $\Delta \neq 0$, resulta $x_1 = \Delta^1/\Delta$.

*) G. Cramer (1704—1752), matemático suizo.

En el caso general, para un j arbitrario, multiplicamos la primera ecuación del sistema (1) por A_{1j}, la segunda por A_{2j}, . . ., la n-ésima por A_{nj}, y sumamos estas ecuaciones, de donde, en virtud de las propiedades f) y g), obtenemos la igualdad

$$x_j \sum_{k=1}^{n} a_{kj}A_{kj} = \sum_{k=1}^{n} y_k A_{kj},$$

o sea,

$$x_j\Delta = \Delta^j,$$

siendo

$$\Delta^j = \sum_{k=1}^{n} y_k A_{kj} = \begin{vmatrix} a_{11} \dots a_{1,\,j-1} & y_1 & a_{1,\,j+1} \dots a_{1n} \\ \dots\dots\dots & \dots & \dots\dots\dots \\ a_{n1} \dots a_{n,\,j-1} & y_n & a_{n,\,j+1} \dots a_{nn} \end{vmatrix}.$$

De aquí, como $\Delta \neq 0$, se deduce la igualdad (3).

Hemos demostrado que si $(x_1, \dots, x_n)$ es una solución del sistema (1), entonces los números x_j se determinan por las fórmulas (3′).

Recíprocamente, el conjunto de los números $x_j = \frac{\Delta^j}{\Delta}$ $(j = 1, \dots, n)$ es una solución del sistema (1). En efecto, sustituyendo x_j $(j = 1, \dots, n)$ en el primer miembro de la k-ésima ecuación $(k = 1, \dots, n)$ del sistema (1), en virtud de las propiedades f), g) de los determinantes, se tiene:

$$\sum_{j=1}^{n} a_{kj}\frac{\Delta^j}{\Delta} = \frac{1}{\Delta}\sum_{j=1}^{n} a_{kj}\sum_{s=1}^{n} y_s A_{sj} = \frac{1}{\Delta}\sum_{s=1}^{n} y_s \sum_{j=1}^{n} a_{kj}A_{sj} = \frac{1}{\Delta}\, y_k\Delta = y_k.$$

Por lo tanto, los números (3′) verdaderamente forman una solución del sistema (1).

4.3. Sistema homogéneo

Un sistema de ecuaciones de la forma

$$\left.\begin{array}{l} a_{11}x_1 + \dots + a_{nn}x_n = 0, \\ \dots\dots\dots\dots\dots \\ a_{n1}x_1 + \dots + a_{nn}x_n = 0 \end{array}\right\} \qquad (5)$$

se llama *homogéneo*. Este es un caso particular del sistema (1) para $y_1 = \dots = y_n = 0$. Está claro que el vector nulo

$$x_1 = 0, \ \dots, \ x_n = 0$$

satisface el sistema homogéneo (5). Pero puede ocurrir que se satisfaga el sistema homogéneo (5) por un vector no nulo $x = (x_1, \dots, x_n)$, o sea, un vector que tenga al menos una componente no nula $x_i \neq 0$.

Este vector se llama entonces *solución no trivial del sistema homogéneo* (5), mientras que el vector nulo se llama *solución trivial del mismo.*

TEOREMA 2. *Si el determinante Δ del sistema homogéneo* (5) *es distinto de cero* ($\Delta \neq 0$), *entonces este sistema sólo admite solución trivial.*

En efecto, en virtud de la propiedad d), todos los determinantes $\Delta^j = 0$ (véase (4)), por lo cual, en virtud de las igualdades (3), $x_j = 0$ ($j = 1, \ldots, n$).

TEOREMA 3. *Si el sistema de ecuaciones* (5) *tiene solución no trivial, entonces su determinante Δ necesariamente es igual a cero* ($\Delta = 0$).

En efecto, si fuese $\Delta \neq 0$, entonces, según el teorema 2, el sistema (5) sólo tendría solución trivial.

Antes habíamos estudiado el sistema lineal (1) en el caso en que su determinante $\Delta \neq 0$. Como se demostró (teorema 1), en este caso el sistema (1) admite solución única, la cual puede calcularse por las fórmulas (3), para cualquier segundo miembro $y = (y_1, \ldots, y_n)$.

4.4. Reglas para la resolución de un sistema de ecuaciones lineales

Estudiaremos el sistema (1) en el caso en que su determinante $\Delta = 0$. Supondremos que al menos un elemento de la matriz A (véase (2)) es distinto de cero y denotaremos el rango de A por k ($k =$ rango A). Por lo tanto, $1 \leqslant k < n$.

Nuestro objetivo es demostrar las siguientes reglas (de una manera explícita fueron enunciadas y demostradas por Kronecker y Capelli).

Si queremos resolver un sistema (1) del que se conoce que el rango de la matriz A es igual a k, tendremos que hallar el rango de la *matriz ampliada*

$$B = \left\| \begin{matrix} a_{11} \ldots a_{1n} & y_1 \\ \ldots\ldots & \ldots \\ a_{n1} \ldots a_{nn} & y_n \end{matrix} \right\|,$$

obtenida de la matriz A por adjunción de la columna

$$\left\| \begin{matrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{matrix} \right\| = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}.$$

1) Si el rango de la matriz B es mayor que el rango de la matriz A (rango $B >$ rango $A = k$), entonces el sistema (1) no admite soluciones. Este sistema es *contradictorio*; no existe ningún vector $x = (x_1, \ldots, x_n)$ que satisfaga simultáneamente todas las ecuaciones (1).

2) Si el rango de la matriz B es igual al rango de la matriz A (rango B = rango $A = k$), entonces el sistema (1) tiene soluciones. Para hallar las soluciones se deben tomar en el sistema (1) k ecuaciones tales que la matriz de sus coeficientes sea de rango k; después se deben resolver estas ecuaciones. Este sistema de k ecuaciones puede tener infinitas soluciones, pero pueden escribirse de una forma expresiva. En este caso, cualquier solución de las k ecuaciones tomadas será también solución de las demás $n - k$ ecuaciones del sistema.

Las reglas 1) y 2) agotan todas las situaciones posibles ya que el rango de B no puede ser menor que k. Hay que tener en cuenta que, por hipótesis, la matriz A engendra un determinante no nulo de k-ésimo orden, el cual es engendrado también por la matriz B.

4.5. Ejemplos de aplicación de las reglas

EJEMPLO 1. El determinante del sistema

$$\left.\begin{array}{l} x+y=1, \\ x-y=2 \end{array}\right\}$$

es

$$\Delta=\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}=-2\neq 0$$

por lo que el sistema tiene solución única, la cual puede calcularse por las fórmulas

$$x=\Delta^1/\Delta, \quad y=\Delta^2/\Delta,$$

donde

$$\Delta^1=\begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix}=-3, \quad \Delta^2=\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}=1,$$

o sea,

$$x=\frac{-3}{-2}=\frac{3}{2}, \quad y=\frac{1}{-2}=-\frac{1}{2}.$$

EJEMPLO 2. El sistema

$$\left.\begin{array}{l} x+y=1, \\ x+y=2 \end{array}\right\} \tag{6}$$

tiene el determinante nulo. La matriz

$$A=\begin{Vmatrix} 1 & 1 \\ 1 & 1 \end{Vmatrix}$$

es de rango 1, mientras que la matriz

$$B=\begin{Vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{Vmatrix}$$

es de rango 2. Como rango $B >$ rango A, el sistema (6) no tiene solución. Por cierto, esto está claro sin necesidad de teoría alguna, pues un mismo número no puede ser igual a 1 y a 2 simultáneamente.

EJEMPLO 3. El sistema

$$\left.\begin{array}{l} 2x+2y=2, \\ 3x+2y=3 \end{array}\right\} \tag{7}$$

tiene el determinante nulo, $\Delta = 0$. La matriz

$$A=\begin{Vmatrix} 2 & 2 \\ 3 & 3 \end{Vmatrix}$$

tiene el rango 1, rango $A = 1$. La matriz ampliada

$$B=\begin{Vmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \end{Vmatrix}$$

también tiene el rango igual a 1. Como rango A = rango $B = 1$, tomamos una ecuación

$$2x + 2y = 2. \tag{8}$$

El coeficiento de y es distinto de cero, por lo cual, en esta ecuación se puede despejar y:

$$y=\frac{2-2x}{2}=1-x. \tag{9}$$

La fórmula (9) proporciona todas las soluciones de la ecuación (8). Podemos asignar a x cualquier valor $(-\infty < x < \infty)$ y calcular el valor y según la fórmula (9). Obtendremos un sistema (un vector) (x, y) que satisface la ecuación (8). El conjunto de todos los sistemas $(x, 1-x)$, donde $x \in (-\infty, \infty)$, es el conjunto de todas las soluciones de la ecuación (8). Estas también son soluciones de la segunda ecuación del sistema (7), pues rango A = rango B. En este caso, este resultado es trivial sin necesidad de aplicar la teoría de los rangos de las matrices. Los coeficientes de las ecuaciones (7) junto con sus segundos miembros son proporcionales, respectivamente, por lo que queda claro que toda solución de una de estas ecuaciones también es solución de la otra.

EJEMPLO 4. El sistema

$$\left.\begin{array}{r} x+y+\ z=1, \\ 2x+y+2z=1, \\ x+y+3z=2 \end{array}\right\}$$

tiene el determinante

$$\Delta=\begin{vmatrix}1 & 1 & 1\\ 2 & 1 & 2\\ 1 & 1 & 3\end{vmatrix}=-2\neq 0$$

por lo cual tiene solución única, que puede calcularse por las fórmulas

$$x=\frac{\Delta^1}{\Delta}=\frac{-1}{2}\begin{vmatrix}1 & 1 & 1\\ 1 & 1 & 2\\ 2 & 1 & 3\end{vmatrix}=-\frac{1}{2},\quad y=\frac{\Delta^2}{\Delta}=\frac{-1}{2}\begin{vmatrix}1 & 1 & 1\\ 2 & 1 & 2\\ 1 & 2 & 3\end{vmatrix}=1,$$

$$z=\frac{\Delta^3}{\Delta}=\frac{-1}{2}\begin{vmatrix}1 & 1 & 1\\ 2 & 1 & 1\\ 1 & 1 & 2\end{vmatrix}=\frac{1}{2}.$$

EJEMPLO 5. El determinante del sistema

$$\left.\begin{aligned}x+y+z&=1,\\ x+y+2z&=1,\\ x+y+3z&=2\end{aligned}\right\}$$

es

$$\Delta=\begin{vmatrix}1 & 1 & 1\\ 1 & 1 & 2\\ 1 & 1 & 3\end{vmatrix}=0.$$

La matriz

$$A=\begin{Vmatrix}1 & 1 & 1\\ 1 & 1 & 2\\ 1 & 1 & 3\end{Vmatrix}$$

es de rango 2, pues $\Delta=0$, pero existe un determinante de segundo orden, engendrado por la matriz A, que es distinto de cero. Por ejemplo,

$$\begin{vmatrix}1 & 1\\ 1 & 2\end{vmatrix}=1\neq 0.$$

La matriz

$$B=\begin{Vmatrix}1 & 1 & 1 & 1\\ 1 & 1 & 2 & 1\\ 1 & 1 & 3 & 2\end{Vmatrix}$$

es de rango 3, pues el determinante engendrado por esta matriz

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{vmatrix} = 1 \neq 0.$$

Como rango $B >$ rango A, el sistema no tiene solución.

EJEMPLO 6. El determinante del sistema

$$\left.\begin{aligned} x + y + z &= 1, \\ x + y + 2z &= 1, \\ 2x + 2y + 4z &= 2 \end{aligned}\right\} \tag{10}$$

es

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{vmatrix} = 0.$$

Es fácil comprobar que las matrices

$$A = \begin{Vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{Vmatrix}, \quad B = \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 2 & 4 & 2 \end{Vmatrix}$$

tidnen el mismo rango, siendo rango $A =$ rango $B = 2$. Tomenos en el sistema (10) dos ecuaciones de tal modo que el rango de la matriz A' de los coeficientes de estas ecuaciones sea igual a 2. En este caso se pueden tomar la primera y segunda ecuaciones o la primera y tercera. Así pues, consideremos el sistema

$$\left.\begin{aligned} x + y + z &= 1, \\ x + y + 2z &= 1. \end{aligned}\right\} \tag{11}$$

Pasemos una de las incógnitas a los segundos miembros de estas ecuaciones de tal modo que los coeficientes de las incógnitas restantes formen una matriz A'' tal que rango $A'' = 2$. En este caso se puede pasar x o y.

En resumen, el sistema no homogéneo

$$\left.\begin{aligned} x + z &= 1 - y, \\ x + 2z &= 1 - y, \end{aligned}\right\} \tag{12}$$

tiene el determinante

$$\Delta'' = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \neq 0.$$

por lo cual, tiene solución única para cualesquiera segundos miembros:

$$x = \frac{1}{\Delta''} \begin{vmatrix} 1-y & 2 \\ 1-y & 2 \end{vmatrix} = 1 - y, \quad z = \frac{1}{\Delta''} \begin{vmatrix} 1 & 1-y \\ 1 & 1-y \end{vmatrix} = 0.$$

Así pues, la terna de números $(1-y, y, 0)$, para $y \in (-\infty, \infty)$, dan todas las soluciones del sistema (12), que también son soluciones de la tercera ecuación del sistema (10) (esta ecuación se obtiene de la segunda multiplicando por 2).

4.6. Fundamentación de las reglas

TEOREMA 4. *Si el sistema* (1) *es compatible, o sea, que admite al menos una solución* x, *entonces, necesariamente rango* B = *rango* A.

La regla 1) se basa en el teorema 4 (véase la pág. 23), puesto que si rango $B >$ rango A, entonces no se cumple la condición necesaria de compatibilidad del sistema (1).

DEMOSTRACIÓN DEL TEOREMA 4. Supongamos que el sistema (1) admite solución y que rango $A = k$. Tenemos que demostrar que rango $B = k$.

Como, por hipótesis, rango $A = k$, existe un determinante no nulo de k-ésimo orden engendrado por la matriz A y, por consiguiente, también por la matriz B. Por lo tanto, rango $B \geqslant k$. No queda más que demostrar que todo determinante de $(k+1)$-ésimo orden engendrado por la matriz B, es igual a cero. Si tal determinante consta sólo de los elementos a_{ij}, entonces, naturalmente, es igual a cero, ya que también es engendrado por la matriz A, la cual, por hipótesis es de rango k. Así pues, hay que demostrar que cualquier determinante de $(k+1)$-ésimo orden engendrado por la matriz B y continente una columna compuesta por los números y_j, es igual a cero. Sin restringir generalidad se puede suponer que éste es el determinante

$$\begin{vmatrix} a_{11} & \dots & a_{1k} & y_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{k+1,1} & \dots & a_{k+1,k} & y_{k+1} \end{vmatrix}.$$

Siempre se puede reducir a este caso cualquier otro, reordenando de un modo adecuado las ecuaciones y las incógnitas x_j.

Por hipótesis, el sistema (1) es compatible, o sea, existe un vector $x = (x_1, \ldots, x_n)$ que satisface a las ecuaciones del sistema. Pero, entonces, en particular, el vector x satisface las $k+1$ ecuaciones del sistema nuevamente reordenado. Por consiguiente,

$$\left.\begin{array}{l} a_{11}x_1 + \ldots + a_{1k}x_k + \lambda_1 = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{k+1,1}x_1 + \ldots + a_{k+1,k}x_k + \lambda_{k+1} = 0, \end{array}\right\} \qquad (13)$$

donde

$$\left.\begin{array}{l} \lambda_1 = a_{1,k+1}x_{k+1} + \ldots + a_{1n}x_n - y_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \lambda_{k+1} = a_{k+1,k+1}x_{k+1} + \ldots + a_{k+1,n}x_n - y_{k+1}. \end{array}\right\} \qquad (14)$$

Formemos el siguiente sistema con las incógnitas $z_1, z_2, \ldots, z_{k+1}$:

$$\left.\begin{array}{l} a_{11}z_1 + \ldots + a_{1k}z_k + \lambda_1 z_{k+1} = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{k+1,1}z_1 + \ldots + a_{k+1,k}z_k + \lambda_{k+1}z_{k+1} = 0. \end{array}\right\} \qquad (15)$$

En virtud de (13) y (14), este sistema se satisface por los números $x_1, \ldots x_k$, 1, entre los cuales, en todo caso, hay uno no nulo. Pero, entonces, el determinante del sistema (15) es igual a cero (véase el teorema 3), o sea,

$$0=\begin{vmatrix} a_{11} & \dots a_{1k} & a_{1,\,k+1}x_{k+1}+\dots+a_{1n}x_n-y_1 \\ \dots & \dots & \dots \\ a_{k+1,\,1} & \dots a_{k+1,\,k} & a_{k+1,\,k+1}x_{k+1}+\dots+a_{k+1,\,n}x_n-y_{k+1} \end{vmatrix}=$$

$$=\sum_{s=k+1}^{n} x_s \begin{vmatrix} a_{11} & \dots a_{1k} & a_{1s} \\ \dots & \dots & \dots \\ a_{k+1,\,1} & \dots a_{k+1,\,k} & a_{k+1,\,s} \end{vmatrix} - \begin{vmatrix} a_{11} & \dots a_{1k} & y_1 \\ \dots & \dots & \dots \\ a_{k+1,\,1} & \dots a_{k+1,\,k} & y_{k+1} \end{vmatrix}=$$

$$=-\begin{vmatrix} a_{11} & \dots a_{1k} & y_1 \\ \dots & \dots & \dots \\ a_{k+1,\,1} & \dots a_{k+1,\,k} & y_{k+1} \end{vmatrix}. \tag{16}$$

Esto último se debe a que los determinantes (¡ de $(k+1)$-ésimo orden !) que figuran en la suma $\sum\limits_s$ son iguales a cero, ya que el rango de la matriz A es igual a k.

Queda demostrado que cualquier determinante de $(k+1)$-ésimo orden engendrado por la matriz B, es igual a cero, como se quería demostrar.

Veamos ahora cómo se puede fundamentar la regla 2) (pág. 24). Como el rango de la matriz A es igual a k (rango $A = k$), el sistema (1) contiene k ecuaciones tales que la matriz de sus coeficientes engendra un determinante de k-ésimo orden no nulo. Reordenando estas ecuaciones y las incógnitas, se puede conseguir que las primeras k ecuaciones del sistema (1)

$$\left.\begin{array}{l} a_{11}x_1+\dots+a_{1n}x_n=y_1, \\ \dots\dots\dots\dots \\ a_{k1}x_1+\dots+a_{kn}x_n=y_k \end{array}\right\} \tag{1'}$$

tengan un determinante no nulo:

$$\sigma=\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}\neq 0.$$

El sistema reordenado (1′) lo escribiremos también así:

$$\left.\begin{array}{l} a_{11}x_1+\dots+a_{1k}x_k=y_1-a_{1,\,k+1}x_{k+1}-\dots-a_{1n}x_n, \\ \dots\dots\dots\dots\dots\dots \\ a_{k1}x_1+\dots+a_{kk}x_k=y_k-a_{k,\,k+1}x_{k+1}-\dots-a_{kn}x_n. \end{array}\right\} \tag{17}$$

Como el determinante $\sigma\neq 0$, a cualquier sistema de números $x_{k+1}, \dots, x_n$ le corresponde un sistema único de números $x_1, \dots, x_k$, los cuales, evidentemente pueden expresarse así:

$$\left.\begin{array}{l} x_1=\dfrac{\sigma^1}{\sigma}=\dfrac{1}{\sigma}\sum\limits_{s=1}^{k}(y_s-a_{s,\,k+1}x_{k+1}-\dots-a_{sn}x_n)\,A_{s1}, \\ \dots\dots\dots\dots\dots\dots \\ x_k=\dfrac{\sigma^k}{\sigma}=\dfrac{1}{\sigma}\sum\limits_{s=1}^{k}(y_s-a_{s,\,k+1}x_{k+1}-\dots-a_{sn}x_n)\,A_{sk}, \end{array}\right\} \tag{18}$$

donde A_{st} son los adjuntos de los elementos a_{st} en el determinante σ. Por consiguiente, todas las soluciones del sistema (17) se expresan según las fórmulas (18). A $x_{k+1}, \dots, x_n$ se les pueden asignar valores arbitrarios y los valores de

$x_1, \ldots, x_k$ se calculan según las fórmulas (18). Vemos pues, que el sistema (17) tiene infinitas soluciones.

Queremos comprobar ahora que si, rango A = rango $B = k$, entonces cualquier solución hallada x de las primeras k ecuaciones del sistema (1) es simultáneamente solución de las demás ecuaciones del sistema. Para precisar, demostremos que es solución de la $(k+1)$-ésima ecuación. Así pues, consideremos las primeras $(k+1)$ ecuaciones del sistema (1), que las escribiremos en la forma (13). Hay que demostrar que toda solución x de las primeras k ecuaciones (13) es simultáneamente solución de la $(k+1)$-ésima ecuación. Sea x un vector que cumpla las primeras k ecuaciones en (13). Consideremos las ecuaciones (15) en las incógnitas $z_1, \ldots, z_{k+1}$, donde $\lambda_1, \ldots, \lambda_{k+1}$ se calculan según las fórmulas (14) mediante las componentes $x_{k+1}, \ldots, x_n$ del vector x. El determinante del sistema (15) es igual a cero. Esto se ve por las ecuaciones (16) que hay que leerlas de derecha a izquierda. Según la hipótesis, el determinante de la derecha es igual a cero. Pero, entonces, el sistema (15) admite solución no trivial $z_1, z_2, \ldots, z_{k+1}$. Aquí $z_{k+1} \neq 0$, ya que si se admite que el sistema (15) tiene una solución de la forma $z_1, \ldots, z_k, 0$, entonces $z_1, \ldots, z_k$ tiene que anularse pues el determinante $\sigma \neq 0$. Pero, entonces, $z_1 = \ldots = z_k = z_{k+1} = 0$ y el sistema $z_1, \ldots, z_{k+1}$ sería trivial. Como el sistema (15) es homogéneo, resulta que no sólo $z_1, \ldots, z_{k+1}$ satisfacen este sistema, sino que también los números $(z_{k+1} \neq 0)$

$$z_1' = z_1/z_{k+1}, \ldots, z_k' = z_k/z_{k+1},$$

poseen esta misma propiedad. Pero, entonces, $z_1', \ldots, z_k'$ satisfacen el sistema de las k primeras ecuaciones (13) cuyo determinante $\sigma \neq 0$. Ya sabemos que este último sistema admite la solución $x_1, \ldots, x_k$ y, en virtud de la unicidad,

$$z_1' = x_1, \ldots, z_k' = x_k.$$

Volviendo a la última ecuación en (15), vemos que la satisfacen los números $(x_1, \ldots, x_k, 1)$, o sea, que los números $(x_1, \ldots, x_k)$ satisfacen la $(k+1)$-ésima ecuación del sistema (13) y, en virtud de (14), el vector considerado $x = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)$ satisface la $(k+1)$-ésima ecuación del sistema (1). Con esto, la proposición 2) queda demostrada.

4.7. Método de resolución de un sistema mediante eliminación de incógnitas

Se puede recomendar el siguiente método de resolución de un sistema de ecuaciones lineales, que es el *método de eliminación de incógnitas* (o método de Gauss [1]). Sea dado el sistema

$$\left.\begin{array}{c} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n. \end{array}\right\} \qquad (1'')$$

Multiplicando una ecuación cualquiera del sistema (1″) por una constante y añadiéndola a otra ecuación del mismo sistema, obtenemos un nuevo sistema equivalente al anterior. El nuevo sistema de ecuaciones tendrá su matriz B', que resulta de la matriz B mediante una transformación correspondiente $(B \Rightarrow B')$. La transformación

[1]) K. F. Gauss (1777—1855), matemático alemán.

consiste en que cierta fila de la matriz B se transforma añadiendo a ella otra fila multiplicada por el número correspondiente.

Del mismo modo, multiplicando una ecuación cualquiera del sistema por un número $k \neq 0$, dejando inalterables las demás ecuaciones, obtenemos, evidentemente, un nuevo sistema equivalente al sistema inicial. El nuevo sistema tendrá una matriz B' que será correspondientemente transformada de la matriz B ($B \Rightarrow B'$). Esta vez, la transformación consiste en que la fila correspondiente a la ecuación en cuestión se multiplica por k.

También surge la necesidad de permutar dos ecuaciones del sistema (1"), obteniendo, por lo tanto, formalmente, un nuevo sistema, pero equivalente al sistema inicial. En este caso, la transformación $B \Rightarrow B'$ se reduce a la permutación de dos filas de la matriz B.

Las tres transformaciones señaladas $B \Rightarrow B'$ se llaman *transformaciones elementales de la matriz*.

En la práctica, en lugar de escribir el nuevo sistema de ecuaciones, se limitan a escribir solamente la matriz correspondiente B'. Aplicando adecuadamente operaciones elementales sobre el sistema de ecuaciones, o lo que es lo mismo, sobre la matriz B, siempre se puede conseguir resolver el sistema dado (1"), o bien, obtener un sistema claramente contradictorio. Como este último sistema es equivalente al sistema (1"), esto demuestra quel el sistema (1") es contradictorio.

A continuación se exponen ejemplos de aplicación de este método.

La operación $B \Rightarrow B'$ denota que B' se obtiene de B mediante una o varias transformaciones elementales.

EJEMPLO 7. Resolver el sistema

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 5, \\ x_2 + 2x_3 + 3x_4 &= 1, \\ x_1 \qquad\;\; + 3x_3 + 4x_4 &= 2, \\ x_1 + \;\; x_2 + 5x_3 + 6x_4 &= 1. \end{aligned}\right\}$$

Naturalmente, según el teorema 1, podríamos calcular todos los cinco determinantes de cuarto orden y hallar x_1, x_2, x_3, x_4. Aquí se repetirían muchos cálculos.

Formemos la matriz B:

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 1 \\ 1 & 0 & 3 & 4 & 2 \\ 1 & 1 & 5 & 6 & 1 \end{pmatrix},$$

donde, como vemos, la última columna consta de los segundos miembros del sistema. Multiplicando la primera fila por (-1) y añadiéndola a la tercera y cuarta filas, obtenemos la matriz

$$B_1'=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -2 & 0 & 0 & -3 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

En la matriz B_1' los elementos de la tercera fila, que representan los coeficientes de las incógnitas, son todos iguales a cero, salvo uno; permutemos esta fila con la segunda. Entonces, el elemento no nulo quedará sobre la diagonal principal:

$$B_1''=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -2 & 0 & 0 & -3 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

También se puede mutliplicar la segunda fila por -1, para que se simplifique su expresión:

$$B_1=\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

Las transformaciones siguientes de las matrices son evidentes:

$$B_1 \Rightarrow B_2=\begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 2 & 2 & -5/2 \end{pmatrix} \Rightarrow B_3=$$

$$=\begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 0 & -1 & -2 \end{pmatrix} \Rightarrow B_4=\begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \Rightarrow$$

$$\Rightarrow B_5=\begin{pmatrix} 1 & 0 & 3 & 0 & -6 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 & -13/2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \Rightarrow B_6=\begin{pmatrix} 1 & 0 & 0 & 0 & 15/4 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 & -13/2 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

De aquí, $x_4=2$, $x_3=-13/4$, $x_2=3/2$, $x_1=15/4$.

Para no cometer errores, se recomienda hacer la prueba, sustituyendo los valores obtenidos en las ecuaciones iniciales del sistema.

Consideremos desde este punto de vista el ejemplo 5:

$$\left.\begin{aligned} x_1+x_2+\ \ x_3&=1,\\ x_1+x_2+2x_3&=1,\\ x_1+x_2+3x_3&=2,\end{aligned}\right\}$$

$$B=\begin{pmatrix}1&1&1&1\\1&1&2&1\\1&1&3&2\end{pmatrix}\Rightarrow B_1=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&2&1\end{pmatrix}\Rightarrow B_2=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

Por lo tanto, el sistema inicial es equivalente al siguiente:

$$\left.\begin{aligned} x_1+\quad x_2+\quad x_3&=1,\\ 0\cdot x_1+0\cdot x_2+\quad x_3&=0,\\ 0\cdot x_1+0\cdot x_2+0\cdot x_3&=1.\end{aligned}\right\}$$

En la última fila el término independiente es igual a la unidad, mientras que los coeficientes de las incógnitas son iguales a cero, por lo cual el sistema es incompatible.

Finalmente, en el ejemplo 6,

$$B=\begin{pmatrix}1&1&1&1\\1&1&2&1\\2&2&4&2\end{pmatrix}\Rightarrow B_1=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&2&0\end{pmatrix}\Rightarrow B_2=$$

$$=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&0&0\end{pmatrix}\Rightarrow B_3=\begin{pmatrix}1&1&1&1\\0&0&1&0\end{pmatrix}\Rightarrow B_4=\begin{pmatrix}1&1&0&1\\0&0&1&0\end{pmatrix}.$$

De aquí que $x_3=0$, $x_1+x_2=1$, o sea, el sistema tiene un conjunto infinito de soluciones: $x_1=x_1$, $x_2=1-x_1$, $x_3=0$, donde x_1 es arbitrario $(-\infty<x_1<\infty)$.

4.8. Cálculo del rango de una matriz

Si sólo nos interesa averiguar el rango de la matriz B, entonces las operaciones elementales indicadas anteriormente $B\Rightarrow B'$ las extendemos no solamente a las filas, sino que también a las columnas de la matriz. Además, si en el proceso de estas transformaciones aparece en la matriz una fila o una columna compuesta totalmente por ceros, entonces éstos hay que suprimirlos de la matriz, o sea, hay que considerar luego una matriz de orden inferior.

Los siguientes ejemplos ilustran este método.

EJEMPLO 8. Hallar el rango de la matriz

$$B=\begin{pmatrix}1&2&3&4&5&6&7\\0&1&0&1&0&1&1\\1&2&1&2&1&2&1\\0&0&0&0&0&4&5\end{pmatrix}.$$

Está claro que el rango de la matriz B no es mayor que 4. En este caso, $a_{11}=1\neq 0$. Multiplicando la primera fila por (-1) y añadiéndola a la tercera, obtenemos:

$$B_1=\begin{pmatrix}1&2&3&4&5&6&7\\0&1&0&1&0&1&1\\0&0&-2&-2&-4&-4&-6\\0&0&0&0&0&4&5\end{pmatrix}.$$

Multiplicando ahora la primera columna por los números correspondientes y añadiéndola a las restantes columnas, obtenemos la matriz

$$B_2=\begin{pmatrix}1&0&0&0&0&0&0\\0&1&0&1&0&1&1\\0&0&-2&-2&-4&-4&-6\\0&0&0&0&0&4&5\end{pmatrix}.$$

La segunda columna ya consta de ceros, salvo el elemento $a_{22}=1\neq$ $\neq 0$. Multiplicando la segunda columna por (-1) y añadiéndola a la 4, 6, 7 columnas, obtenemos:

$$B_3=\begin{pmatrix}1&0&0&0&0&0&0\\0&1&0&0&0&0&0\\0&0&-2&-2&-4&-4&-6\\0&0&0&0&0&4&5\end{pmatrix}\Rightarrow B_4=$$

$$=\begin{pmatrix}1&0&0&0&0&0&0\\0&1&0&0&0&0&0\\0&0&-2&0&0&0&0\\0&0&0&0&0&4&5\end{pmatrix}\Rightarrow B_5=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&-2&0&0\\0&0&0&4&5\end{pmatrix}\Rightarrow B_6=$$

$$=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&-2&0&0\\0&0&0&4&0\end{pmatrix}\Rightarrow B_7=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-2&0\\0&0&0&4\end{pmatrix}.$$

El determinante de cuarto orden de la matriz B_7 es distinto de cero y, por consiguiente,

rango $B =$ rango $B_7 = 4$.

EJEMPLO 9. Hallar el rango de la matriz

$$B=\begin{pmatrix}1&2&1&1\\0&1&1&1\\1&1&0&0\end{pmatrix}\Rightarrow B_1=\begin{pmatrix}1&2&1&1\\0&1&1&1\\0&-1&-1&-1\end{pmatrix}\Rightarrow B_2=$$
$$=\begin{pmatrix}1&0&0&0\\0&1&1&1\\0&-1&-1&-1\end{pmatrix}\Rightarrow B_3=\begin{pmatrix}1&0&0&0\\0&1&1&1\\0&0&0&0\end{pmatrix}\Rightarrow B_4=$$
$$=\begin{pmatrix}1&0&0&0\\0&1&1&1\end{pmatrix}\Rightarrow B_5=\begin{pmatrix}1&0&0&0\\0&1&0&0\end{pmatrix}\Rightarrow B_6=\begin{pmatrix}1&0\\0&1\end{pmatrix},$$

o sea, el rango de la matriz B es igual a dos.

Los razonamientos en los ejemplos 8 y 9 se basan en la siguiente regla general:

En una transformación elemental $B \Rightarrow B'$ se conserva el rango de la matriz, o sea, se cumple la igualdad

rango $B =$ rango B'.

Esta regla es evidente si la transformación elemental se reduce a la permutación de filas o columnas de la matriz o a la eliminación de la matriz de una fila o columna compuesta de ceros.

Queda un caso más, que expresaremos en forma de un teorema.

TEOREMA 5. *Supongamos que la matriz B se ha sometido a una transformación $B \Rightarrow B'$, que consiste en que a una de sus filas (o columnas) se le ha añadido alguna otra fila (o columna), multiplicada por un número c.*

Entonces, los rangos de las matrices B y B' son iguales.

DEMOSTRACIÓN. Vamos a demostrar este teorema para las filas (para las columnas los razonamientos son similares).

Sea k el número de orden de la fila de la matriz $B = \| b_{kl} \|$ que se multiplica por el número c y que se añade a otra fila de B, cuyo número orden supondremos igual a l (por lo tanto, la l-ésima fila de la matriz B' consta de los elementos $cb_{kj} + b_{lj}$, $j = 1, \ldots, n$).

Supongamos que

rango $B = r$, rango $B' = r'$.

Es suficiente demostrar que $r' \leqslant r$, ya que por analogía se demuestra que $r \leqslant r'$, de donde resulta $r = r'$.

Si $r = 0$, entonces todos los elementos de la matriz B son iguales a cero y, por consiguiente, también son iguales a cero todos los elementos de la matriz B', de donde $r = r' = 0$.

Sea ahora $r > 0$. Entonces existe una matriz A de orden r, engendrada por la matriz B, con el determinante distinto de cero ($|A| \neq 0$), mientras que todas las matrices A engendradas por la matriz B y de orden mayor que r tienen el determinante igual a cero. En la transformación $B \Rightarrow B'$ la matriz A se transforma en una matriz A' ($A \Rightarrow A'$). Supongamos que la matriz A es de orden mayor que r.

Si la l-ésima fila de la matriz B no participa en la formación de la matriz A, entonces, evidentemente, $A = A'$ y $0 = |A| = |A'|$. Si en la formación de la matriz A participan la k-ésima y l-ésima filas de la matriz B, entonces $0 = |A| = |A'|$. En efecto, para obtener el determinante $|A'|$ hay que añadir a alguna fila del determinante $|A|$ otra fila determinada del mismo multiplicada por el número c, con lo que el valor del determinante no varía.

Finalmente, supongamos que en la formación de la matriz A participa la l-ésima fila, pero no participa la k-ésima fila. Está claro (véase la propiedad h) de los determinantes) que

$$|A'| = |A| + c\,|\Lambda|, \tag{19}$$

donde Λ es una matriz de orden mayor que r, engendrada de A por sustitución de los elementos de la l-ésima fila por los elementos correspondientes de la k-ésima fila de la matriz B. Evidentemente, si trasladamos al k-ésimo lugar la fila obtenida de este modo en el l-ésimo lugar, resulta una matriz de orden superior a r, engendrada por la matriz B. Pero, entonces, $|\Lambda| = 0$.

De (19) obtenemos $|A'| = 0 + c0 = 0$.

Hemos examinado todos los casos en que el orden de la matriz A' es mayor que r y siempre ha resultado que $|A'| = 0$. Esto muestra que $r' = \text{rango } B' \leqslant r$, como se quería demostrar.

§ 5. Espacio tridimensional. Vectores. Sistema cartesiano de coordenadas [1]

5.1. Concepto de vector

En este párrafo consideraremos el espacio real. El concepto de vector en el espacio real ya lo conoce el lector por la geometría elemental.

Se llama *vector* (en el espacio real) a un segmento orientado $\overrightarrow{AB}$, con el origen en el punto A y con el extremo en el punto B, que se puede trasladar paralelamente a sí mismo. Por lo tanto, se considera

[1]) Señalemos que en este libro primero se expone el producto escalar de vectores, luego la geometría analítica en la recta y en el plano y después de esto, en los §§ 11—13 se dan los conceptos de producto vectorial y producto mixto de vectores. Si se desea, estos párrafos pueden exponerse inmediatamente después del § 6.

que dos segmentos orientados $\overrightarrow{AB}$ y $\overrightarrow{A_1B_1}$, que tienen longitudes iguales ($|AB| = |A_1B_1|$) y una misma dirección, determinan un mismo vector $\boldsymbol{a}$, y en este sentido se escribe: $\boldsymbol{a} = \overrightarrow{AB} = \overrightarrow{A_1B_1}$ (fig. 2).

Se llama *longitud* $|\overrightarrow{AB}| = |\boldsymbol{a}|$ del vector $\overrightarrow{AB} = \boldsymbol{a}$, al número (no negativo) que es igual a la longitud del segmento AB que une los puntos A y B. También escribiremos así: $|\overrightarrow{AB}| = |AB|$.

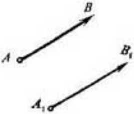

Fig. 2

Los vectores que están situados sobre una misma recta o sobre rectas paralelas se llaman *colineales*.

Si los puntos A y B coinciden, entonces $\overrightarrow{AB} = \overrightarrow{AA} = \boldsymbol{0}$ también se considera como un vector; éste es el *vector nulo*. Su longitud es igual a cero ($|\boldsymbol{0}| = 0$), y la dirección carece de sentido.

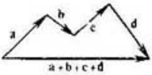

Fig. 3

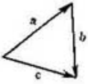

Fig. 4

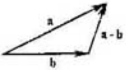

Fig. 5

En geometría se considera la suma y diferencia de vectores, así como el producto de vectores por números reales. Por definición, el producto $\boldsymbol{a}\alpha = \alpha\boldsymbol{a}$ de un vector $\boldsymbol{a}$ por un número α, o bien, el producto del número α por el vector $\boldsymbol{a}$, es un vector cuya longitud es igual a $|\alpha\boldsymbol{a}| = |\alpha|\cdot|\boldsymbol{a}|$ y cuya dirección coincide con la de $\boldsymbol{a}$ si $\alpha > 0$ y es opuesta a la de $\boldsymbol{a}$ si $\alpha < 0$. Si $\alpha = 0$, la longitud $|\alpha\boldsymbol{a}|$ es igual a cero y el vector $\alpha\boldsymbol{a}$ se convierte en el vector nulo (en un punto), el cual carece de dirección.

Un vector $\boldsymbol{e}$ se llama *unitario*, si su longitud es igual a 1, o sea, si $|\boldsymbol{e}| = 1$. Si $\boldsymbol{b} = \alpha\cdot\boldsymbol{e}$ y $\boldsymbol{e}$ es un vector unitario, entonces $|\boldsymbol{b}| = |\alpha|$, ya que $|\boldsymbol{b}| = |\alpha|\cdot|\boldsymbol{e}| = |\alpha|\cdot 1 = |\alpha|$.

Por definición, un número finito de vectores $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots$, se suman según la regla de la clausura de la cadena de estos vectores. Las figs. 3 y 4 nos recuerdan cómo se hace esto. En la fig. 5 se muestra además cómo se restan los vectores.

5.2. Proyección de un vector

Se llama *proyección de un punto A sobre una recta L* (fig. 6) al punto A' en el que se corta la recta L con el plano que pasa por el punto A y es perpendicular a la recta L.

Tomemos una recta orientada L (fig. 7) y un vector $\boldsymbol{a} = \overrightarrow{AB}$.

Se llama *proyección del vector $\boldsymbol{a} = \overrightarrow{AB}$ sobre la recta orientada L* al vector $\overrightarrow{A'B'}$, donde A', B' son las proyecciones de los puntos A y B, respectivamente, sobre L (véase la fig. 7).

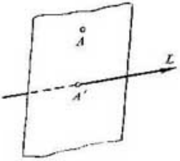

Fig. 6

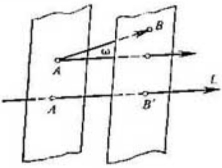

Fig. 7

La proyección del vector $\boldsymbol{a}$ sobre la recta orientada L se denota por $\overrightarrow{\mathrm{pr}_L \boldsymbol{a}}$.

Dada una recta orientada L, las proyecciones $\overrightarrow{A'B'}$ de cualesquiera vectores $\overrightarrow{AB}$ sobre L están situadas en L y llevan la misma dirección que L o la dirección opuesta.

Por cierto, si el vector $\overrightarrow{AB}$ es nulo o es perpendicular a L, entonces, evidentemente, su proyección sobre L es un vector nulo, que no tiene dirección.

Junto con la proyección del vector $\boldsymbol{a}$ sobre la recta orientada L, que representa un vector, introduciremos un nuevo concepto, el de *proyección numérica del vector $\boldsymbol{a}$ sobre la recta orientada L*. Es un número que se denota por $\mathrm{pr}_L\, \boldsymbol{a}$ (sin flecha) y se define del modo siguiente.

Se llama *proyección numérica del vector $\boldsymbol{a} = \overrightarrow{AB}$ sobre la recta orientada L al producto de la longitud del vector $\boldsymbol{a} = \overrightarrow{AB}$ por el coseno del ángulo ω formado por el vector $\boldsymbol{a}$ y la dirección de L:*

$$\mathrm{pr}_L\, \boldsymbol{a} = |\boldsymbol{a}| \cos(\boldsymbol{a}, L) = |\boldsymbol{a}| \cos \omega \quad (0 \leqslant \omega \leqslant \pi).$$

Señalemos los casos siguientes:

Si $\boldsymbol{a} = 0$, o bien, si $\omega = \frac{\pi}{2}$, entonces $\mathrm{pr}_L\, \boldsymbol{a} = 0$; si

$a \neq 0$ y $0 \leqslant \omega < \frac{\pi}{2}$, entonces la proyección numérica es positiva ($\text{pr}_L\, a > 0$) y, evidentemente, es igual a la longitud del vector $\overrightarrow{A'B'}$: $\text{pr}_L\, a = |\overrightarrow{AB}| \times \cos \omega = |\overrightarrow{A'B'}|$; en este caso, el propio vector $\overrightarrow{A'B'}$ lleva la misma dirección que L; ahora bien, si $a \neq 0$ y $\frac{\pi}{2} < \omega \leqslant \pi$, entonces la proyección numérica es negativa ($\text{pr}_L\, a < 0$) y, evidentemente, es igual a la longitud del vector $\overrightarrow{A'B'}$ tomada con signo menos:

$$\text{pr}_L\, a = |\overrightarrow{AB}| \cos \omega = -|\overrightarrow{A'B'}| = -|A'B'|;$$

en este caso, el propio vector $\overrightarrow{A'B'}$ lleva la dirección opuesta a la de L.

Se verifica le igualdad evidente, que expresa la relación entre la proyección del vector a sobre la dirección de L y su proyección numérica sobre L:

$$\overrightarrow{pr_L a} = e\, pr_L a.$$

Aquí e es el vector unitario que lleva la dirección de L.

Si los vectores a y b están situados en la recta orientada L, entonces éstos pueden expresarse en la forma

$$a = \alpha e, \quad b = \beta e,$$

donde e es el vector unitario que lleva la dirección de L, mientras que α y β son unos números. Estos pueden ser positivos, negativos o iguales a cero.

Se verifican las igualdades evidentes:

$$\alpha e \pm \beta e = (\alpha \pm \beta)\, e \tag{1}$$

que muestran que la suma y la diferencia de los vectores en cuestión se reduce a la suma y la diferencia de los números correspondientes α y β.

5.3. Propiedades de las proyecciones de los vectores

Las proyecciones numéricas de los vectores a y b sobre una dirección dada L poseen las propiedades siguientes:

$$\text{pr}_L\, a + \text{pr}_L\, b = \text{pr}_L\, (a + b), \tag{2}$$

$$\text{pr}_L\, (\alpha a) = \alpha \cdot \text{pr}_L\, a. \tag{3}$$

La propiedad (2) se muestra en la fig. 8:

$$\overrightarrow{pr_L a}+\overrightarrow{pr_L b}=\overrightarrow{A'B'}+\overrightarrow{B'C'}=\overrightarrow{A'C'}=\overrightarrow{pr_L c}=\overrightarrow{pr_L (a+b)}.$$

Pero, entonces,

$$\boldsymbol{e}\cdot \mathrm{pr}_L\, \boldsymbol{a} + \boldsymbol{e}\cdot \mathrm{pr}_L\, \boldsymbol{b} = \boldsymbol{e}\cdot \mathrm{pr}_L\, (\boldsymbol{a}+\boldsymbol{b}),$$

de donde se deduce (2) (véase (1)).

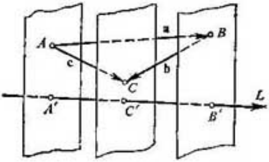

Fig. 8

Como $\boldsymbol{a} = (\boldsymbol{a} - \boldsymbol{b}) + \boldsymbol{b}$ (véase fig. 5), se tiene

$$\mathrm{pr}_L\, (\boldsymbol{a}-\boldsymbol{b}) + \mathrm{pr}_L\, \boldsymbol{b} = \mathrm{pr}_L\, \boldsymbol{a},$$

y, por consiguiente,

$$\mathrm{pr}_L\, \boldsymbol{a} - \mathrm{pr}_L\, \boldsymbol{b} = \mathrm{pr}_L\, (\boldsymbol{a}-\boldsymbol{b}). \qquad (2')$$

Demostremos (3). Considerando que el ángulo formado por el vector $\boldsymbol{a}$ y la dirección de L es igual a ω, se tiene:

si $\alpha > 0$: $\mathrm{pr}_L\, (\alpha\boldsymbol{a}) = |\alpha\boldsymbol{a}| \cos\omega = \alpha\, |\boldsymbol{a}| \cos\omega = \alpha\cdot\mathrm{pr}_L\, \boldsymbol{a}$;

si $\alpha < 0$: $\mathrm{pr}_L\, (\alpha\boldsymbol{a}) = |\alpha\boldsymbol{a}| \cos(\pi-\omega) =$

$$= -\alpha\, |\boldsymbol{a}| \cos(\pi-\omega) = \alpha\, |\boldsymbol{a}| \cos\omega = \alpha\, \mathrm{pr}_L\, \boldsymbol{a}.$$

Hay que tener en cuenta que, si $\alpha < 0$, la dirección del vector $\alpha\boldsymbol{a}$ es opuesta a la del vector $\boldsymbol{a}$, y si $\boldsymbol{a}$ forma con L el ángulo ω, entonces $\alpha\boldsymbol{a}$ forma con L el ángulo $\pi - \omega$.

Si $\alpha = 0$, el primero y segundo miembros de (3) se anulan.

5.4. Producto escalar de vectores

Se llama *producto escalar de dos vectores* $\boldsymbol{a}$ y $\boldsymbol{b}$ al número $(\boldsymbol{a}, \boldsymbol{b})$ que es igual al producto de las longitudes de estos vectores por el coseno del ángulo ω formado por ellos:

$$\boldsymbol{ab} = (\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}|\, |\boldsymbol{b}| \cos(\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}|\, |\boldsymbol{b}| \cos\omega. \qquad (4)$$

Está claro que también se puede decir que *el producto escalar de dos vectores* $\boldsymbol{a}$ *y* $\boldsymbol{b}$ *es el producto de la longitud del vector* $\boldsymbol{b}$ *por la* *proyección numérica del vector* $\boldsymbol{a}$ *sobre la dirección de* $\boldsymbol{b}$, *o bien, es el producto de la longitud del vector* $\boldsymbol{a}$ *por la proyección numérica de* $\boldsymbol{b}$ *sobre la dirección de* $\boldsymbol{a}$:

$$\boldsymbol{ab} = (\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}| \operatorname{pr}_{\boldsymbol{a}} \boldsymbol{b} = |\boldsymbol{b}| \operatorname{pr}_{\boldsymbol{a}} \boldsymbol{a}.$$

El producto escalar posee las propiedades:

$$(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a}), \tag{5}$$

$$(\boldsymbol{a}, \boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{b}) + (\boldsymbol{a}, \boldsymbol{c}), \tag{6}$$

$$(\boldsymbol{a}, \alpha\boldsymbol{b}) = \alpha (\boldsymbol{a}, \boldsymbol{b}). \tag{7}$$

La igualdad (5) es consecuencia inmediata de la definición de producto escalar.

La igualdad (6) se demuestra así:

$$(\boldsymbol{a}, \boldsymbol{b} + \boldsymbol{c}) = |\boldsymbol{a}| \operatorname{pr}_{\boldsymbol{a}} (\boldsymbol{b} + \boldsymbol{c}) = |\boldsymbol{a}| \operatorname{pr}_{\boldsymbol{a}} \boldsymbol{b} + |\boldsymbol{a}| \operatorname{pr}_{\boldsymbol{a}} \boldsymbol{c} = \\ = (\boldsymbol{a}, \boldsymbol{b}) + (\boldsymbol{a}, \boldsymbol{c}).$$

La igualdad (7) se demuestra del modo siguiente:

$$(\boldsymbol{a}, \alpha\boldsymbol{b}) = |\boldsymbol{a}| \operatorname{pr}_{\boldsymbol{a}} (\alpha\boldsymbol{b}) = |\boldsymbol{a}| \alpha \operatorname{pr}_{\boldsymbol{a}} \boldsymbol{b} = \alpha (\boldsymbol{a}, \boldsymbol{b}).$$

De (6) y (7), teniendo en cuenta (5), se deduce que

$$(\boldsymbol{a} + \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{c}) + (\boldsymbol{b}, \boldsymbol{c}), \tag{6'}$$

$$(\alpha\boldsymbol{a}, \boldsymbol{c}) = \alpha (\boldsymbol{a}, \boldsymbol{b}). \tag{7'}$$

EJEMPLO DE FÍSICA. Si un cuerpo, bajo la acción de una fuerza $\boldsymbol{a}$, ha realizado una traslación rectilínea a lo largo de un vector $\boldsymbol{b}$, entonces, como ya se sabe por la física, el trabajo W ejercido por la fuerza $\boldsymbol{a}$ es igual al producto de la magnitud de la fuerza $|\boldsymbol{a}|$ por el camino $|\boldsymbol{b}|$ y por el coseno del ángulo formado por los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$:

$$W = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \cos (\boldsymbol{a}, \boldsymbol{b}).$$

Pero, entonces,

$$W = (\boldsymbol{a}, \boldsymbol{b}),$$

de modo que el trabajo indicado es igual al producto escalar de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$.

5.5. Sistema rectangular de coordenadas

Ahora vamos a pasar a la descripción analítica de los vectores y los puntos del espacio mediante números. Introduzcamos en el espacio un *sistema rectangular de coordenadas* x, y, z, o sea, tres rectas orientadas y perpendiculares entre sí que pasen por un punto O, denominadas *ejes de coordenadas* x, y, z (fig. 9). Se supone que para un sistema dado de coordenadas se ha elegido un segmento unidad mediante el cual se miden los demás segmentos. El punto O se llama *origen de coordenadas*.

Tomemos un punto arbitrario A del espacio tridimensional. El segmento orientado $\overrightarrow{OA}$ se llama *radio vector* (o vector de posición) del punto A. A su vez, el radio vector determina un vector $\boldsymbol{a}$ ($\boldsymbol{a} = \overrightarrow{OA}$) que se puede trasladar en el espacio paralelamente a sí mismo. Las proyecciones numéricas del radio vector $\boldsymbol{a}$ sobre los ejes x, y, z las denotaremos por x, y, z, respectivamente. Estas son las *coordenadas* del punto A; la coordenada x se llama *abscisa*, la coordenada y se llama *ordenada* y la coordenada z se llama *z-coordenada* (cota) del punto A.

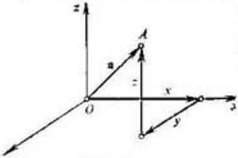

Fig. 9

Entre los puntos A del espacio y sus radios vectores $\overrightarrow{OA}$, o lo que es lo mismo, las ternas de números (x, y, z) que representan las coordenadas del punto A, o bien, las proyecciones de $\overrightarrow{OA}$ sobre los ejes, existe una correspondencia biyectiva. En virtud de esto, no habrá lugar a confusiones si a la terna de números (x, y, z) la llamamos punto A, cuyas coordenadas son los números dados, o vector $\boldsymbol{a}$ cuyas proyecciones son estos mismos números.

Escribiremos $\boldsymbol{a} = (x, y, z)$ y diremos que $\boldsymbol{a}$ ó (x, y, z) es un vector, que es igual al radio vector del punto A y que tiene las coordenadas x, y, z. Claro que se puede considerar que el vector $\boldsymbol{a}$ es igual a algún otro segmento orientado $\overrightarrow{CD}$ igual a $\overrightarrow{OA}$ ($\overrightarrow{CD} = \overrightarrow{OA}$), o sea, que tenga la misma dirección y la misma longitud que $\overrightarrow{OA}$. En este caso, las proyecciones de $\boldsymbol{a}$ sobre los ejes coordenados se denotan frecuentemente mediante a_x, a_y, a_z y se escribe $\boldsymbol{a} = (a_x, a_y, a_z)$.

De la definición de vector como segmento orientado que se puede trasladar en el espacio paralelamente a sí mismo, se deduce que dos vectores $\boldsymbol{a} = (x_1, y_1, z_1)$ y $\boldsymbol{b} = (x_2, y_2, z_2)$ son iguales si, y sólo si, se cumplen simultáneamente las igualdades:

$$x_1 = x_2, \quad y_1 = y_2, \quad z_1 = z_2.$$

Son válidas las igualdades

$$(x, y, z) \pm (x', y', z') = (x \pm x', y \pm y', z \pm z'), \tag{8}$$

$$\alpha (x, y, z) = (\alpha x, \alpha y, \alpha z). \tag{9}$$

La igualdad (8) se deduce de que la proyección de la suma o la diferencia de dos vectores (sobre el eje x, sobre el eje y y sobre el eje z)

es igual a la suma o a la diferencia de las proyecciones de los sumandos.

La igualdad (9) se deduce de que la proyección del vector $\alpha\boldsymbol{a}$ (sobre el eje x, sobre el eje y y sobre el eje z) es igual al producto de α por la proyección de $\boldsymbol{a}$.

Denotemos por $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ los vectores unitarios (de longitudes iguales a uno) cuyas direcciones coinciden con las direcciones de los ejes x, y, z, respectivamente. Estos vectores se llaman *vectores unitarios* (o *versores*) de los ejes x, y, z. Un vector arbitrario (x, y, z) puede expresarse en la forma

$$(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}. \tag{10}$$

En efecto,

$$\boldsymbol{i} = (1, 0, 0), \quad \boldsymbol{j} = (0, 1, 0), \quad \boldsymbol{k} = (0, 0, 1).$$

Por lo tanto, en virtud de (8) y (9), se tiene

$$x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) =$$
$$= (x, 0, 0) + (0, y, 0) + (0, 0, z) = (x, y, z).$$

Señalemos las igualdades que se verifican para los productos escalares de los vectores unitarios de los ejes

$$\boldsymbol{ii} = \boldsymbol{jj} = \boldsymbol{kk} = 1, \quad \boldsymbol{ij} = \boldsymbol{ik} = \boldsymbol{jk} = 0.$$

Sean ahora $\boldsymbol{a} = (x, y, z)$, $\boldsymbol{b} = (x', y', z')$. Entonces,

$$\boldsymbol{ab} = (\boldsymbol{a}, \boldsymbol{b}) = xx' + yy' + zz'. \tag{11}$$

En efecto, en virud de (6), (7), (6'), (7'), se tiene

$$\boldsymbol{ab} = (x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})(x'\boldsymbol{i} + y'\boldsymbol{j} + z'\boldsymbol{k}) = xx'\boldsymbol{ii} + xy'\boldsymbol{ij} +$$
$$+ xz'\boldsymbol{ik} + yx'\boldsymbol{ji} + yy'\boldsymbol{jj} + uz'\boldsymbol{jk} + zx'\boldsymbol{ki} +$$
$$+ zy'\boldsymbol{kj} + zz'\boldsymbol{kk} = xx' + yy' + zz'.$$

En particular, haciendo en esta fórmula $\boldsymbol{b} = \boldsymbol{a}$, obtenemos que

$$|\boldsymbol{a}|^2 = \boldsymbol{aa} = x^2 + y^2 + z^2,$$

de donde la longitud del vector $\boldsymbol{a}$ es igual a

$$|\boldsymbol{a}| = +\sqrt{x^2 + y^2 + z^2}.$$

De aquí resulta que la distancia entre los puntos $\boldsymbol{a} = (x, y, z)$ y $\boldsymbol{b} = (x', y', z')$ es igual a (fig. 10):

$$|AB| = |\boldsymbol{b} - \boldsymbol{a}| = +\sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2}.$$

Para lo que sigue, es conveniente hacer un resumen de lo expuesto. Para ello, introduciremos una notación algo distinta de las coordenadas. Precisamente, consideraremos en el espacio un sistema rectangular de coordenadas x_1, x_2, x_3. En virtud de esto, todo punto del espacio se expresa por una terna de números

$$\boldsymbol{x} = (x_1,\ x_2,\ x_3).$$

Este punto lo hemos denotado con la letra negrita $\boldsymbol{x}$ y lo llamamos también vector $\boldsymbol{x}$ con las componentes x_1, x_2, x_3.

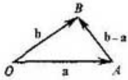

Fig. 10

Hemos demostrado que la suma y la diferencia de vectores, así como el producto de vectores por números se expresan mediante las ternas $(x_1,\ x_2,\ x_3)$ del modo siguiente:

$$\left.\begin{aligned} &(x_1,\ x_2,\ x_3) \pm (x_1',\ x_2',\ x_3') = \\ &\qquad = (x_1 \pm x_1',\ x_2 \pm x_2',\ x_3 \pm x_3'), \\ &\alpha(x_1,\ x_2,\ x_3) = (\alpha x_1,\ \alpha x_2,\ \alpha x_3). \end{aligned}\right\} \tag{12}$$

El producto escalar de los vectores $\boldsymbol{x} = (x_1, x_2, x_3)$ y $\boldsymbol{x}' = (x_1', x_2', x_3')$ se expresa mediante las coordenadas de los vectores $\boldsymbol{x}$ y $\boldsymbol{x}'$ según la fórmula

$$\boldsymbol{x}\boldsymbol{x}' = (\boldsymbol{x},\ \boldsymbol{x}') = x_1 x_1' + x_2 x_2' + x_3 x_3'. \tag{13}$$

La longitud del vector $\boldsymbol{x} = (x_1,\ x_2,\ x_3)$ es un número no negativo, igual a

$$|\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}. \tag{14}$$

Finalmente, la distancia entre los puntos $\boldsymbol{x} = (x_1,\ x_2,\ x_3)$ y $\boldsymbol{x}' = (x_1',\ x_2',\ x_3')$ es igual a

$$|\boldsymbol{x} - \boldsymbol{x}'| = \sqrt{(x_1' - x_1)^2 + (x_2' - x_2)^2 + (x_3' - x_3)^2}. \tag{15}$$

El espacio real cuya geometría hemos estudiado, se llama *espacio tridimensional*, pues sus puntos se expresan de un modo natural mediante ternas de números reales. Lo denotaremos por R_3.

Un plano arbitrario lo denotaremos, naturalmente, por R_2. En R_2 se puede considerar un sistema rectangular de coordenadas x_1, x_2, mediante el cual, un punto arbitrario de R_2, o bien, su radio vector se puede expresar por un par de números $(x_1,\ x_2)$. Las operaciones de adición, sustracción y multiplicación por un número para los vectores pertenecientes al plano cumplen, evidentemente, las condiciones (12) que hemos obtenido anteriormente, donde entre paréntesis hay que suprimir siempre las terceras componentes. El producto escalar de vectores pertenecientes al plano en cuestión también se expresa por la fórmula (13), donde en el segundo miembro hay que suprimir el tercer término. Lo mismo se refiere a las fórmulas (14) y (15).

La generalización de los espacios R_2 y R_3 es el espacio R_n, donde n es un número natural arbitrario.

El espacio R_n para $n > 3$ es una invención matemática. Por cierto, es una invención genial que ayuda a comprender matemáticamente muchos fenómenos complicados.

5.6. Problemas

1. Hallar las longitudes de los vectores (1, 1, 1), (1, —1, 1), (1, 2, 3).
2. Hallar el ángulo formado por los vectores (1, 0, 1), (1, 2, 2).
3. Sea dado un cubo unidad (con la longitud de la arista igual a 1). Hallar el ángulo formado por: a) la diagonal principal y la diagonal de una cara que parten de un mismo vértice; b) dos diagonales de dos caras que parten de un mismo vértice.

§ 6. Espacio euclídeo n-dimensional. Producto escalar

6.1. Espacio n-dimensional R_n

El conjunto de todos los sistemas posibles de números reales (complejos) $(x_1, \ldots, x_n)$ se llama *espacio real* (complejo) *n-dimensional* y se denota por R_n. Cada uno de los sistemas lo denotaremos por una letra (negrita) sin subíndice:

$$\boldsymbol{x} = (x_1, \ldots, x_n)$$

y lo llamaremos *punto* o *vector* de R_n. Los números $x_1, \ldots, x_n$ se llaman *coordenadas* del punto (vector) $\boldsymbol{x}$ o también, *componentes* del vector $\boldsymbol{x}$.

Dos puntos

$$\boldsymbol{x} = (x_1, \ldots, x_n), \quad \boldsymbol{x}' = (x'_1, \ldots, x'_n)$$

se consideran iguales si sus coordenadas correspondientes son iguales:

$$x_j = x'_j \quad (j = 1, \ldots, n).$$

En los demás casos, $\boldsymbol{x}$ y $\boldsymbol{x}'$ son distintos ($\boldsymbol{x} \neq \boldsymbol{x}'$).

Los sistemas (vectores) $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ se pueden sumar y restar, así como multiplicar por números reales $\alpha, \beta, \ldots$, si R_n es el espacio real, y por números complejos si R_n es el espacio complejo.

Por definición, se llama *suma de los vectores* x e y al vector

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \tag{1}$$

y diferencia, al vector

$$x - y = (x_1 - y_1, \ldots, x_n - y_n). \tag{2}$$

Se llama *producto del número* α *por el vector* x, o bien, producto del vector x por el número α, al vector

$$\alpha x = x\alpha = (\alpha x_1, \ldots, \alpha x_n).$$

Finalmente, el vector $-x$ se define por la igualdad

$$-x = (-1)\,x = (-x_1, \ldots, -x_n).$$

También se introduce el concepto de vector nulo, cuyas componentes son iguales a cero:

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

Evidentemente, se cumplen las propiedades:

1) $x + y = y + x$,
2) $(x + y) + z = x + (y + z)$,
3) $x - y = x + (-1)\,y$,
4) $\alpha x + \alpha y = \alpha\,(x + y)$,
5) $\alpha x + \beta x = (\alpha + \beta)\,x$,
6) $\alpha\,(\beta x) = (\alpha\beta)\,x$,
7) $1 \cdot x = x$,
8) $x + (-x) = \mathbf{0}$,

donde α, β son números arbitrarios, mientras que $x, y \in R_n$.

El espacio R_n se llama *lineal* puesto que en el mismo se cumplen las propiedades 1) — 8) enunciadas anteriormente (véase más adelante la nota 1).

El número (no negativo)

$$|x| = \sqrt{\sum_{k=1}^{n} x_k^2} \tag{3}$$

se llama *longitud* o *norma* del vector $x = (x_1, \ldots, x_n)$ en el espacio real R_n.

La distancia entre los puntos $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ del espacio real R_n se define según la fórmula

$$|x - y| = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}. \tag{4}$$

6.2. Producto escalar en el espacio real R_n

Se llama *producto escalar de dos vectores* $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ en el espacio real R_n (los números x_i, y_i son reales) al número

$$(x, y) \sum_{i=1}^{n} x_i y_i. \tag{5}$$

El producto escalar en el espacio real R_n cumple las propiedades:

a) $(x, x) \geqslant 0$; además, la igualdad a cero se verifica si, y sólo si, $x = 0 = (0, \ldots, 0)$,

b) $(x, y) = (y, x)$,

c) $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z)$.

En efecto,

$$(x, x) = \sum_{j=1}^{n} x_j^2 \geqslant 0$$

y la igualdad sólo es posible si $x_1 = \ldots = x_n = 0$. Por otra parte,

$$(x, y) = \sum_{j=1}^{n} x_j y_j = \sum_{j=1}^{n} y_j x_j = (y, x)$$

y

$$(\alpha x + \beta y, z) = \sum_{j=1}^{n} (\alpha x_j + \beta y_j) z_j = \alpha \sum_{j=1}^{n} x_j z_j + \beta \sum_{j=1}^{n} y_j z_j = \alpha (x, z) + \beta (y, z).$$

6.3. Producto escalar en el espacio complejo R_n

Se llama *producto escalar de dos vectores* $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ en el espacio complejo R_n al número

$$(x, y) = \sum_{j=1}^{n} x_j \bar{y}_j, \tag{5'}$$

donde $\bar{y}_j$ es el número complejo conjugado de y_j (por definición, si $z = \alpha + \beta i$, donde α y β son reales, entonces, $\bar{z} = \alpha - \beta i$).

El producto escalar en el espacio complejo R_n cumple las propiedades:

a') $(x, x) \geqslant 0$; además, la igualdad a cero se verifica si, y sólo si, $x = \mathbf{0} = (0, \ldots, 0)$,

b') $\overline{(x, y)} = (y, x)$,

c') $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z)$.

En efecto,

$$(x, x) = \sum_{j=1}^{n} x_j \bar{x}_j = \sum_{j=1}^{n} |x_j|^2 \geqslant 0,$$

y la igualdad sólo es posible si $x_1 = \ldots = x_n = 0$. Por otra parte,

$$\overline{(x, y)} = \overline{\left(\sum_{j=1}^{n} x_j \bar{y}_j\right)} = \sum_{j=1}^{n} \bar{x}_j y_j = \sum_{j=1}^{n} y_j \bar{x}_j = (y, x).$$

Aquí se han aplicado las propiedades de la operación de conjugación $\overline{z + z_1} = \bar{z} + \bar{z}_1$, $\overline{z z_1} = \bar{z} \cdot \bar{z}_1$. Finalmente,

$$(\alpha x + \beta y, z) = \sum_{j=1}^{n} (\alpha x_j + \beta y_j) \bar{z}_j =$$

$$= \alpha \sum_{j=1}^{n} x_j \bar{z}_j + \beta \sum_{j=1}^{n} y_j \bar{z}_j = \alpha (x, z) + \beta (y, z).$$

En el espacio complejo R_n la longitud del vector x se define mediante la igualdad

$$|x| = \sqrt{\sum_{j=1}^{n} |x_j|^2} = \sqrt{\sum_{j=1}^{n} x_j \bar{x}_j}, \tag{3'}$$

y la distancia entre los puntos $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ así:

$$|x - y| = \sqrt{\sum_{j=1}^{n} |x_j - y_j|^2}. \tag{4'}$$

Fácilmente se observa que para valores reales de x_j, y_j, las expresiones (3) y (4) son casos particulares de las expresiones (3'), (4').

El espacio R_n (real o complejo) en el que se ha introducido el producto escalar según la fórmula (5) (o (5'), respectivamente) se llama *espacio euclídeo n-dimensional.*

6.4. Desigualdad de Buniakovski

En adelante nos limitaremos al estudio del espacio real n-dimensional. En este espacio los vectores se multiplican por números reales. Nos permitiremos omitir a veces el término «real».

De las propiedades a), b), c) del producto escalar se deduce la importante desigualdad de Buniakovski [1]):

$$|(x, y)| \leqslant \sqrt{(x, x)}\sqrt{(y, y)}. \tag{6}$$

En efecto, para cualquier número real λ,

$$0 \leqslant (x + \lambda y, x + \lambda y) = (x, x) + \lambda (y, x) + \lambda (x, y) +$$
$$+ \lambda^2 (y, y) = (x, x) + 2 (x, y) \lambda + (y, y) \lambda^2 = a + 2b\lambda + c\lambda^2,$$

donde $a = (x, x)$, $b = (x, y)$, $c = (y, y)$. Vemos que el polinomio cuadrático

$$a + 2b\lambda + c\lambda^2 \geqslant 0 \quad (-\infty < \lambda < \infty)$$

es no negativo para cualquier λ real. Por esta razón, toda su gráfica está situada por encima del eje λ, y esto sólo es posible si el discriminante del polinomio es negativo o es igual a cero, o sea, si $b^2 - ac \leqslant 0$, de modo que $b^2 \leqslant ac$, con lo que obtenemos la desigualdad de Buniakovski (6).

En términos de componentes de los vectores x e y, la desigualdad (6) puede escribirse así:

$$\left| \sum_{j=1}^{n} x_j y_j \right| \leqslant \sqrt{\sum_{j=1}^{n} x_j^2} \sqrt{\sum_{j=1}^{n} y_j^2}. \tag{7}$$

Por lo tanto, para cualesquiera números reales x_j, y_j se verifica la desigualdad (7).

En virtud de (3), la desigualdad de Buniakovski puede escribirse así

$$|(x, y)| \leqslant |x|\,|y|.$$

Pero, entonces, existe un λ único que satisface las desigualdades $-1 \leqslant \lambda \leqslant 1$ y tal que se verifica la igualdad exacta

$$(x, y) = \lambda\, |x|\,|y|.$$

Señalemos que en $[0, \pi]$ la función $\cos t$ admite una función inversa uniforme y estrictamente decreciente con el campo de valores sobre $[-1, 1]$. Por esto, para todo λ $(-1 \leqslant \lambda \leqslant 1)$ existe un ángulo único ω $(0 \leqslant \omega \leqslant \pi)$ tal que $\lambda = \cos \omega$. Por lo tanto, hemos demostrado la igualdad

$$(x, y) = |x|\,|y| \cos \omega. \tag{8}$$

El número ω se llama *ángulo formado por los vectores n-dimensionales* x e y, aunque en realidad, para $n > 3$ los vectores x e y no son segmentos reales sino abstracciones matemáticas.

---

[1]) V. Ya. Buniakovski (1804—1889), matemático ruso.

Los vectores x e y se llaman *ortogonales* si su producto escalar es igual a cero:

$$(x, y) = \sum_{j=1}^{n} x_j y_j = 0.$$

De (8) se deduce que, *para que dos vectores no nulos x e y sean ortogonales es necesario y suficiente que el ángulo formado por ellos sea igual a* $\omega = \frac{\pi}{2}$.

6.5. Desigualdad de Minkowski [1)]

Señalemos una *desigualdad importante de Minkowski*

$$|x + y| \leqslant |x| + |y|, \tag{9}$$

o bien, en términos de componentes

$$\sqrt{\sum_{j=1}^{n} (x_j + y_j)^2} \leqslant \sqrt{\sum_{j=1}^{n} x_j^2} + \sqrt{\sum_{j=1}^{n} y_j^2}. \tag{10}$$

Esta puede demostrarse así:

$$|x + y|^2 = (x + y, x + y) = (x, x) + 2(x, y) + (y, y).$$

Aplicando la desigualdad de Buniakovski (6), se tiene:

$$|x+y|^2 \leqslant (x, x) + 2\sqrt{(x, x)}\sqrt{(y, y)} + (y, y) = \left(\sqrt{(x, x)} + \sqrt{(y, y)}\right)^2.$$

de donde se deduce (9). De (9) se deduce la desigualdad

$$||x| - |y|| \leqslant |x - y|, \tag{11}$$

pues

$$|x| = |x - y + y| \leqslant |x - y| + |y|,$$
$$|y| = |y - x + x| \leqslant |x - y| + |x|.$$

PROBLEMA. Hallar el ángulo formado por los vectores (1, 0, 1, 0) y (1, 1, 1, 1).

Nota 1. Un conjunto arbitrario E, compuesto por elementos de cualquier naturaleza $x, y, z, \ldots$, se llama *espacio lineal*, si existe una ley según la cual, para cualesquiera dos elementos $x, y \in E$ están definidos unos elementos $x + y \in E$ y $x - y \in E$, denominados suma y diferencia, respectivamente, de los elementos x e y, y para cualquier número real o complejo α y cualquier elemento $x \in E$ está definido un elemento $\alpha x = x\alpha \in E$, denominado producto de α por x, o bien, producto de x por α, de modo que se cumplen las propiedades 1) — 8) enunciadas anteriormente en el ap. 6.1, donde $-x = (-1)x$ y $\mathbf{0} = 0x$, $\forall x \in E$.

---

1) H. Minkowski (1864—1909), matemático alemán.

R_n se puede considerar como un ejemplo de espacio lineal, pero también existen otros ejemplos interesantes. Por ejemplo, el conjunto C de las funciones $f, \varphi, \psi, \ldots$ continuas en un segmento $[a, b]$, considerando que $f + \varphi$, $f - \varphi$, αf se han definido como $f(x) + \varphi(x)$, $f(x) - \varphi(x)$, $\alpha f(x)$, es un espacio lineal.

Un espacio lineal con el producto de sus elementos por números reales (o complejos) se llama *espacio lineal real* (*complejo*).

Nota 2. Si en algún espacio lineal E, a cada par de elementos x, y se les ha puesto en correspondencia un número (x, y) tal que se cumplen las propiedades enunciadas anteriormente a), b), c), en el caso real, y a'), b'), c'), en el caso complejo, entonces se dice que en E se ha introducido un producto escalar.

Anteriormente se dio la definición de espacio euclídeo n-dimensional; éste es el espacio R_n en el que se ha introducido el producto escalar según las fórmulas (5) ó (5'), respectivamente.

§ 7. Segmento. División de un segmento en una razón dada

Tomemos unos puntos arbitrarios $x, y \in R_n$ y consideremos el conjunto de puntos (vectores)

$$z = \lambda x + \mu y, \quad (\lambda, \mu \geqslant 0, \quad \lambda + \mu = 1), \tag{1}$$

determinados por los números no negativos λ, μ cuya suma es igual a 1. Se tiene:

$$z = (1 - \mu) x + \mu y = x + \mu (y - x) \quad (0 \leqslant \mu \leqslant 1) \tag{2}$$

o bien,

$$z = y + \lambda (x - y) \quad (0 \leqslant \lambda \leqslant 1). \tag{2'}$$

De la igualdad (2) vemos que en el espacio tridimensional los puntos z completan el segmento que une x e y. Esto se debe a que el radio

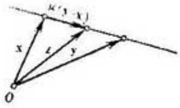

Fig. 11

vector z es la suma del vector x y del vector $\mu(y - x)$, el cual es colineal con $y - x$ (fig. 11). Por lo tanto, el conjunto de puntos (1) representa un segmento $[x, y]$ en R_3 que une los puntos x e y. Para

$\mu = 0$ $z = x$, para $\lambda = 0$ $z = y$, y para cualquier $\lambda > 0$ ($\mu = 1 - \lambda > 0$) z es un punto arbitrario de $[x, y]$.

Por definición, se llama *segmento* $[x, y]$ *que une los puntos* x, $y \in R_n$ al conjunto de todos los puntos z de la forma (1). Se verifica el

TEOREMA 1. *El punto*

$$z = \lambda x + \mu y \quad |(\lambda,\ \mu \geqslant 0,\quad \lambda + \mu = 1)$$

divide al segmento $[x, y]$ *que une los puntos* x, $y \in R_n$ *en segmentos cuya razón de longitudes es igual a* $\mu : \lambda$.

DEMOSTRACIÓN. De (2) se deduce que $z - x = \mu (y - x)$, por lo que la distancia entre los puntos x y z es igual a

$$|z - x| = \mu\, |y - x|. \tag{3}$$

Por otra parte, de (2') resulta $z - y = \lambda (x - y)$, y por lo tanto, la distancia entre los puntos z e y es igual a

$$|z - y| = \lambda\, |x - y|. \tag{4}$$

De (3) y (4) se deduce que

$$|z - x| : |z - y| = \mu : \lambda,$$

como se quería demostrar.

PROBLEMA. Se pide hallar en el segmento $[x, y]$, que une los puntos x, $y \in R_n$, un punto z que divida este segmento en la razón $p : q$ ($p > 0$, $q > 0$).

SOLUCIÓN. Tomemos los números

$$\lambda = \frac{q}{p+q}, \quad \mu = \frac{p}{p+q}.$$

Estos cumplen las propiedades λ, $\mu \geqslant 0$, $\lambda + \mu = 1$, $\mu/\lambda = p/q$. Por esto, en virtud del teorema 1, el punto pedido es

$$z = \lambda x + \mu y = \frac{qx + py}{p+q}. \tag{5}$$

Sus coordenadas $z = (z_1, \ldots, z_n)$ se expresan mediante las coordenadas $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ por las fórmulas

$$z_j = \frac{qx_j + py_j}{p+q} \quad (j = 1, \ldots, n). \tag{5'}$$

En particular, el punto medio del segmento se obtiene para $p = q = 1$, o sea, para $\lambda = \mu = 1/2$.

Señalemos que, como se demuestra en la mecánica, el punto z determinado por las fórmulas (5) o (5') es el centro de gravedad del sistema de puntos x e y en los que están concentradas las masas q y p, respectivamente.

Obsérvese que en R_3, el conjunto de puntos

$$z = \lambda x + \mu y, \quad \lambda + \mu = 1,$$

donde λ y μ son de signo arbitrario, representa la recta que pasa por los puntos x e y. Esto se ve en la igualdad (2').

En lo que se refiere al espacio R_n ($n > 3$), este conjunto se llama recta por definición.

EJEMPLO 1. Hállense las coordenadas del centro de gravedad de un sistema de puntos materiales $x^k = (x_1^k, x_2^k, x_3^k)$ de masas p_k ($k = 1, \ldots, M$), respectivamente.

Aplicando las fórmulas (5') a los puntos x^1, x^2, hallamos el centro de gravedad z^1 de estos dos puntos. Luego hallamos el centro de gravedad z^2 de los puntos z^1 y x^3 de masas $p_1 + p_2$ y p_3, respectivamente. Continuando este proceso, después de $(M - 1)$ pasos, obtenemos:

$$z_j^{M-1} = \frac{1}{p_1 + \ldots + p_M} \sum_{k=1}^{M} p_k x_j^k \quad (j = 1, 2, 3).$$

§ 8. La recta

El concepto de recta es primario en la geometría. De los axiomas de geometría sabemos que por dos puntos pasa una sola recta y que por un punto situado en una recta dada se puede trazar una recta única perpendicular a la dada.

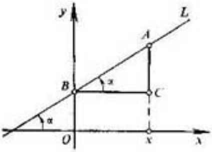

Fig. 12

En el plano tomamos un sistema rectangular de coordenadas (x, y) y una recta L no paralela al eje y (fig. 12). Por el curso escolar sabemos que la ecuación de la recta L tiene la forma

$$y = kx + l, \tag{1}$$

donde $k = \operatorname{tg} \alpha$ y α es el ángulo formado por la recta L con la dirección positiva del eje x, y l es la ordenada del punto de intersección de L con el eje y ($l = OB$).

Cuando se dice que (1) es la ecuación de la recta L, con esto quieren expresar que L es el lugar geométrico de los puntos cuyas coordenadas (x, y) satisfacen la ecuación (1), lo cual se observa fácilmente en la fig. 12. El punto A es un punto arbitrario (variable) de la recta L cuyas coordenadas son (x, y), $BC = x$, $AC = y - l$, y

$$k = \operatorname{tg} \alpha = \frac{y - l}{x}, \tag{1'}$$

de donde se deduce (1). Recíprocamente, la igualdad (1) es equivalente a la igualdad (1'), y esta última expresa el hecho evidente de que el punto (x, y) está situado sobre la recta L. En la fig. 12 el ángulo α es agudo. En el caso de un ángulo obtuso α se pueden hacer unos razonamientos semejantes.

Consideremos la ecuación

$$Ax + By + C = 0, \tag{2}$$

donde A, B, C son ciertos números y además, A y B no son simultáneamente iguales a cero.

Si $B \neq 0$, entonces la ecuación (2) puede expresarse en la forma siguiente:

$$y = -\frac{A}{B}x - \frac{C}{B} \tag{2'}$$

o bien, haciendo

$$k = -\frac{A}{B}, \quad l = -\frac{C}{B}$$

en la forma (1). Como las ecuaciones (2) y (2′) son equivalentes (ya que cualquier punto (x, y) que satisface una de estas ecuaciones satisface también la otra), la ecuación (2) para $B \neq 0$ es la ecuación de la recta que forma con la dirección positiva del eje x un ángulo α cuya tangente es igual a $-\frac{A}{B}$ ($\operatorname{tg} \alpha = -A/B$), y que se corta con el eje y en el punto de ordenada $-C/B$ ($l = -C/B$). Para $B = 0$ la ecuación (2) toma la forma

$$Ax + C = 0 \quad (A \neq 0),$$

o bien,

$$x = a \quad (a = -C/A).$$

Esta también es la ecuación de una recta, pero paralela al eje y. Es, precisamente, el lugar geométrico de los puntos (x, y) cuyas abscisas x son iguales a un mismo número a. En la fig. 13 viene representada una tal recta para $a > 0$.

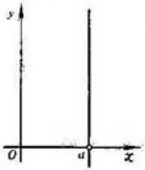
Fig. 13

De lo expuesto se deduce que la ecuación (2), donde A, B, C son unos números dados y A y B no son simultáneamente nulos, es la ecuación de una recta. Si $B \neq 0$, esta recta no es paralela al eje y. En particular, si $A = 0$ es paralela al eje x ($y = -C/B$). Si $B = 0$ es paralela al eje y. Señalemos que la ecuación del eje x es, evidentemente, $y = 0$, mientras que la del eje y es, $x = 0$.

La ecuación (2) se llama *ecuación general de la recta*. Cualquier recta, situada arbitrariamente con respecto al sistema de coordenadas, puede expresarse por una ecuación de la forma (2) para valores adecuados de las constantes A, B, C. Subrayemos que los *números A* *y B en la ecuación* (2) *de la recta no son simultáneamente iguales a cero.* El número k en la ecuación (1) se llama *coeficiente angular de la recta.*

Resolvamos varios problemas importantes.

PROBLEMA 1. Escribir la ecuación de la recta con el coeficiente angular k y que pasa por un punto dado (x_0, y_0).

SOLUCIÓN. La recta con el coeficiente angular k tiene la forma

$$y = kx + l, \tag{3}$$

donde l puede ser cualquier número. Como el punto (x_0, y_0) tiene que estar situado sobre la recta dada, tiene que cumplirse la igualdad

$$y_0 = kx_0 + l. \tag{4}$$

Restando (4) de (3), obtenemos la ecuación pedida de la recta

$$y - y_0 = k(x - x_0) \tag{5}$$

que pasa por el punto (x_0, y_0) y tiene el coeficiente angular k.

PROBLEMA 2. Escribir la ecuación de la recta que pasa por dos puntos dados (x_1, y_1), (x_2, y_2). Se supone que los puntos son distintos.

SOLUCIÓN. Supongamos que $x_1 \neq x_2$. Entonces, evidentemente, la recta buscada no es paralela al eje y, por lo cual, puede expresarse en la forma

$$y - y_1 = k(x - x_1), \tag{6}$$

donde k es un número. La ecuación (6) ya expresa que la recta pasa por el punto (x_1, y_1). Para que pase también por el punto (x_2, y_2), es necesario que se cumpla la igualdad

$$y_2 - y_1 = k(x_2 - x_1). \tag{7}$$

Dividiendo (6) por (7), obtenemos

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \tag{8}$$

Esta es la ecuación de la recta que pasa por los puntos (x_1, y_1) y (x_2, y_2).

Nota. Podría suceder que $y_2 - y_1 = 0$; entonces, formalmente obtendríamos la igualdad

$$\frac{y - y_1}{0} = \frac{x - x_1}{x_2 - x_1}.$$

A pesar de lo absurdo de esta igualdad, sin embargo, la escriben así, pues se considera que es muy cómoda. Despejando, obtenemos una igualdad verdadera

$$y - y_1 = 0 \cdot \frac{x - x_1}{x_2 - x_1} = 0$$

o bien,

$$y = y_1. \tag{9}$$

El caso $x_1 = x_2 = c$ da lugar a la solución $x = c$.

PROBLEMA 3. Hallar el ángulo ω formado por las rectas

$$y = k_1 x + l_1, \quad y = k_2 x + l_2.$$

SOLUCIÓN. Se tiene, $k_1 = \operatorname{tg} \alpha_1$, $k_2 = \operatorname{tg} \alpha_2$, donde α_1, α_2 son los ángulos respectivos, formados por las rectas dadas con la dirección

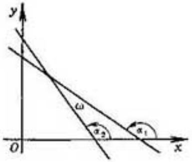

Fig. 14

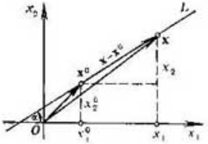

Fig. 15

positiva del eje x. Entonces (fig. 14),

$$\operatorname{tg} \omega = \operatorname{tg}(\alpha_1 - \alpha_2) = \frac{\operatorname{tg} \alpha_1 - \operatorname{tg} \alpha_2}{1 + \operatorname{tg} \alpha_1 \operatorname{tg} \alpha_2} = \frac{k_1 - k_2}{1 + k_1 k_2}. \tag{10}$$

y hemos obtenido la *fórmula para el ángulo formado por las rectas.*

El caso $1 + k_1 k_2 = 0$, o bien,

$$k_1 k_2 = -1 \tag{11}$$

expresa la *condición de perpendicularidad de las rectas. La condición de paralelismo de las rectas* ($\operatorname{tg} \omega = 0$), se expresa así:

$$k_1 = k_2. \tag{12}$$

Escribamos la ecuación (2) en la forma siguiente:

$$A_1 x_1 + A_2 x_2 + C = 0, \tag{2'}$$

donde haremos $x = x_1$, $y = x_2$, $A = A_1$, $B = A_2$. Para $A_1 \neq 0$, $A_2 \neq 0$, $C \neq 0$, la ecuación (2') puede escribirse en la forma

$$\frac{x_1}{a} + \frac{x_2}{b} = 1, \tag{13}$$

donde $a = -C/A_1$, $b = -C/A_2$. Esta ecuación se llama *ecuación de la recta «segmentaria»*. Esta recta se corta con el eje x_1 (la recta $x_2 = 0$) en el punto $(a, 0)$ y con el eje x_2 (la recta $x_1 = 0$) en el punto $(0, b)$.

Si la recta que satisface la ecuación (2′) pasa por el punto (x_1^0, x_2^0), entonces

$$A_1x_1^0 + A_2x_2^0 + C = 0. \tag{14}$$

Restando (14) de (2′), obtenemos

$$A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) = 0. \tag{15}$$

Esta se llama *ecuación de la recta que pasa por el punto* (x_1^0, x_2^0).

Si se introducen los vectores $A = (A_1, A_2)$, $x = (x_1, x_2)$, $x^0 = (x_1^0, x_2^0)$, entonces el primer miembro de (15) se puede considerar como el producto escalar del vector A por el vector $x - x^0$. Por lo tanto, la ecuación (15) en forma vectorial se expresa así:

$$A(x - x^0) = 0. \tag{15′}$$

El vector $x - x^0$ pertenece a la recta L (fig. 15). Así pues, en (15′) se observa que el vector $A = (A_1, A_2)$ es ortogonal (perpendicular) a la recta dada, con lo que queda establecido el significado geométrico de los coeficientes A_1 y A_2.

Consideremos dos rectas

$$A_1x_1 + A_2x_2 + C_1 = 0 \quad (L_1), \tag{16}$$

$$B_1x_1 + B_2x_2 + C_2 = 0 \quad (L_2). \tag{17}$$

Como los vectores $A = (A_1, A_2)$ y $B = (B_1, B_2)$ son perpendiculares a las rectas (16) y (17), respectivamente, el ángulo φ formado por las rectas (16) y (17) es igual al ángulo formado por los vectores A y B (fig. 16). El ángulo φ se puede calcular por la fórmula

$$\cos\varphi = \frac{(A,\ B)}{|A|\,|B|} = \frac{A_1B_1 + A_2B_2}{\sqrt{A_1^2 + A_2^2}\,\sqrt{B_1^2 + B_2^2}}. \tag{18}$$

Nota. Si φ es el ángulo formado por las rectas, entonces $\pi - \varphi$ también es un ángulo formado por estas rectas. El número (18) puede ser positivo o negativo. Uno de ellos corresponde al ángulo φ y el otro, al ángulo $\pi - \varphi$.

De (18) obtenemos la *condición de perpendicularidad de* L_1 y L_2 ($\varphi = \pi/2$, $\cos\varphi = 0$):

$$A_1B_1 + A_2B_2 = 0. \tag{19}$$

Si las rectas L_1 y L_2 *son paralelas*, entonces los vectores A y B son colineales y $A = \lambda B$, donde λ es un número real. De aquí que la condición de paralelismo de rectas se expresa por la igualdad

$$\frac{A_1}{B_1} = \frac{A_2}{B_2}. \tag{20}$$

Sea dada una recta arbitraria L en un sistema rectangular de coordenadas (fig. 17), que no pase por el origen de coordenadas, y sea $\boldsymbol{a}$ una vector con el comienzo en el origen de coordenadas, perpendicular a la recta L y tal que su extremo esté situado sobre la recta. El vector $\boldsymbol{a}$ determina totalmente la recta L (por el extremo del vector $\boldsymbol{a}$ pasa una recta única perpendicular al mismo). Sea p la longitud del vector $\boldsymbol{a}$ ($p = |\boldsymbol{a}|$) y $\mathbf{v} = (\cos\alpha_1, \cos\alpha_2)$, el vector unitario que lleva la misma dirección que el vector $\boldsymbol{a}$. Aquí α_i es el ángulo formado

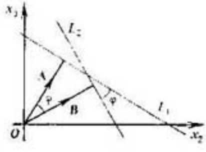

Fig. 16

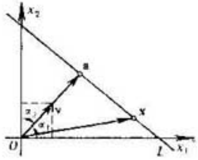

Fig. 17

por el vector $\boldsymbol{a}$ (o el vector $\mathbf{v}$) y la dirección positiva del eje x_i ($i = 1, 2$); $\cos^2\alpha_1 + \cos^2\alpha_2 = 1$ ($\cos\alpha_2 = \operatorname{sen}\alpha_1$). Denotemos por $\boldsymbol{x} = (x_1, x_2)$ el radio vector (el vector de posición) de un punto arbitrario (un punto variable) de la recta L. La proyección del vector $\boldsymbol{x}$ sobre el vector unitario $\mathbf{v}$, evidentemente, es igual a p, o bien, el producto escalar del radio vector de un punto arbitrario $\boldsymbol{x}$ de la recta L por el vector $\mathbf{v}$ es igual a p:

$$(\boldsymbol{x}, \mathbf{v}) = |\mathbf{v}|\,\mathrm{pr}_{\mathbf{v}}\ \boldsymbol{x} = p. \tag{21}$$

En resumen, hemos obtenido la ecuación vectorial de la recta L, ya que, recíprocamente, si el radio vector de un punto satisface la ecuación (21), entonces éste está situado sobre L (un punto no situado sobre L tiene una proyección sobre $\mathbf{v}$ distinta de p).

Si la recta L pasa por el origen de coordenadas, entonces su ecuación también puede expresarse en la forma (21), donde $\mathbf{v}$ es un vector unitario perpendicular a la misma.

En coordenadas, la ecuación (21) tiene la forma

$$x_1\cos\alpha_1 + x_2\cos\alpha_2 = p \quad (p \geqslant 0) \tag{21'}$$

o bien,

$$x_1\cos\alpha_1 + x_2\operatorname{sen}\alpha_1 = p \quad (p \geqslant 0). \tag{21''}$$

La ecuación (21′) (o la (21″) se llama *ecuación de la recta en la forma normal.*

Si la recta L viene dada por la ecuación general (2′)

$$A_1x_1 + A_2x_2 + C = 0,$$

entonces, ésta puede reducirse a la forma normal multiplicándola por el número

$$M = \pm 1/\sqrt{A_1^2 + A_2^2}, \tag{22}$$

donde se debe tomar el signo contrario al de C ($p = -MC \geqslant 0$). El número M se llama *factor normalizador*. Como

$$(MA_1)^2 + (MA_2)^2 = 1,$$

existe un ángulo único α_1 que cumple las desigualdades $0 \leqslant \alpha_1 < 2\pi$ y tal que

$$MA_1 = \cos \alpha_1, \quad MA_2 = \operatorname{sen} \alpha_1. \tag{23}$$

De este modo, obtenemos la ecuación (21), donde $p = -MC \geqslant 0$. Señalemos una vez más que el número p es igual a la distancia del origen de coordenadas a la recta.

PROBLEMA. Hallar la distancia d de un punto $\boldsymbol{x}^0 = (x_1^0, x_2^0)$ a la recta L dada por la ecuación

$$A_1x_1 + A_2x_2 + C = 0. \tag{24}$$

SOLUCIÓN. Sea

$$(\boldsymbol{x}, \boldsymbol{\nu}) - p = 0 \tag{25}$$

la ecuación normal de la recta (24). Entonces, si $C \neq 0$ se tiene que p ($p > 0$) es la longitud del vector $\overrightarrow{OA}$ que une el origen de coordenadas con la recta L, mientras que $\boldsymbol{\nu}$ es el vector unitario que lleva la dirección de $\overrightarrow{OA}$ ($p = |\overrightarrow{OA}|$, $\boldsymbol{\nu} = \overrightarrow{OA}/p$, véase la fig. 18). Sea $\boldsymbol{x}$ el radio vector de un punto arbitrario de L. Entonces, evidentemente para hallar la distancia del punto cuyo radio vector es $\boldsymbol{x}^0$ a L, hay que hallar la proyección del vector $\boldsymbol{x} - \boldsymbol{x}^0$ sobre la dirección del vector $\boldsymbol{\nu}$ y tomar el valor absoluto de la proyección:

$$\begin{aligned} d &= |\operatorname{pr}_{\boldsymbol{\nu}} (\boldsymbol{x} - \boldsymbol{x}^0)| = |(\boldsymbol{x} - \boldsymbol{x}^0, \boldsymbol{\nu}| = \\ &= |(\boldsymbol{x}, \boldsymbol{\nu}) - (\boldsymbol{x}^0, \boldsymbol{\nu})| = |p - (\boldsymbol{x}^0, \boldsymbol{\nu})| = |(\boldsymbol{x}^0, \boldsymbol{\nu}) - p|. \end{aligned}$$

Queda demostrada la fórmula

$$d = |(\boldsymbol{x}^0, \boldsymbol{\nu}) - p|. \tag{26}$$

Por lo tanto, para obtener la distancia d, hay que reducir la ecuación (24) a la forma normal, trasladar p al primer miembro, sustituir en el primer miembro x_1 y x_2 por las coordenadas correspondientes x_1^0, x_2^0 del punto $\boldsymbol{x}^0$ y tomar el valor absoluto de la expresión obtenida.

Mediante los coeficientes A_1, A_2, C, la fórmula (26) se expresa así:

$$d = \frac{|A_1x_1^0 + A_2x_2^0 + C|}{\sqrt{A_1^2 + A_2^2}}. \tag{26'}$$

Para $C = 0$ sigue siendo válida la fórmula (26) y, por consiguiente, también la fórmula (26'). En este caso, $p = 0$, $\mathbf{v}$ es uno de los

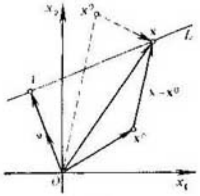

Fig. 18

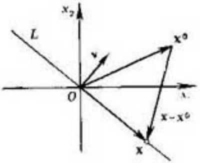

Fig. 19

dos vectores unitarios que son perpendiculares a L (fig. 19). Ahora

$$d = | \operatorname{pr}_{\mathbf{v}} (x - x^0) | = |(\mathbf{v}, x - x^0) | = \\ = | (\mathbf{v}, x) - (\mathbf{v}, x^0) | = | (\mathbf{v}, x^0) |$$

o bien,

$$d = \frac{| A_1 x_1^0 + A_2 x_2^0 |}{\sqrt{A_1^2 + A_2^2}},$$

o sea, resulta la fórmula (26') para $C = 0$.

Nota. En la fig. 18 se observa que, si el origen de coordenadas O y el punto x^0 están situados:

a) hacia un mismo lado de L, entonces el ángulo formado por $\mathbf{v}$ y $x - x^0$ es agudo y $d = p - (x^0, \mathbf{v})$,

b) a distintos lados de L, entonces el ángulo formado por $\mathbf{v}$ y $x - x^0$ es obtuso y $d = (x^0, \mathbf{v}) - p$.

PROBLEMA. Hallar la distancia del punto (1, 1) a la recta $2x + \sqrt{5}y - \sqrt{5} = 0$.

§ 9. Ecuación del plano

9.1. Ecuación del plano en la forma normal

Tomemos en el espacio R_3, en el que se ha introducido un sistema rectangular de coordenadas (x_1, x_2, x_3), un vector a con el origen en O. Tracemos un plano L por el extremo del vector a, que sea perpendicular a este vector (fig. 20). Denotaremos por $x = (x_1, x_2, x_3)$ un punto arbitrario (un punto variable) del plano L. La letra x denota también el radio vector del punto x.

Sea $p = |\boldsymbol{a}|$ la longitud del vector $\boldsymbol{a}$ y

$$\mathbf{v} = (\cos\alpha_1,\ \cos\alpha_2,\ \cos\alpha_3)$$

el vector unitario que lleva la misma dirección que $\boldsymbol{a}$. Aquí $\alpha_1, \alpha_2, \alpha_3$ son los ángulos que forma el vector $\mathbf{v}$ con las direcciones positivas de los ejes x_1, x_2, x_3, respectivamente. Evidentemente, la proyección de cualquier punto $x \in L$ sobre el vector $\mathbf{v}$ es una cantidad constante e igual a p:

$$(x, \mathbf{v}) = p \quad (p \geqslant 0). \tag{1}$$

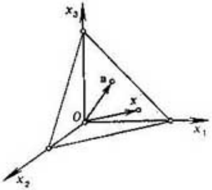

Fig. 20

La ecuación (1) también tiene sentido para $p = 0$. En este caso, el plano pasa por el origen de coordenadas O $(\boldsymbol{a} = 0)$ y el vector unitario $\mathbf{v}$ que parte del origen O perpendicularmente a L en una cualquiera de las dos direcciones posibles. La ecuación (1) es la *ecuación del plano L en la forma vectorial.*

En coordenadas se escribe así:

$$x_1\cos\alpha_1 + x_2\cos\alpha_2 + x_3\cos\alpha_3 = p \quad (p \geqslant 0) \tag{1'}$$

y se llama *ecuación del plano en la forma normal.*

9.2. Ecuación del plano en la forma general

Multiplicando la ecuación (1′) por algún número distinto de cero, obtenemos una ecuación equivalente de la forma

$$A_1x_1 + A_2x_2 + A_3x_3 + B = 0, \tag{2}$$

que determina el mismo plano. Aquí los números A_1, A_2, A_3 no son simultáneamente iguales a cero. La ecuación (2), donde los números A_1, A_2, A_3 no son todos iguales a cero, se llama *ecuación del plano en la forma general.*

Cualquier ecuación de la forma (2), donde los números A_1, A_2, A_3 no son simultáneamente iguales a cero, puede reducirse a la forma normal multiplicándola por el número

$$M = \pm 1/\sqrt{A_1^2 + A_2^2 + A_3^2},$$

donde el signo se toma opuesto al signo de B. Entonces, el número $p = -MB$ resultará negativo y la ecuación (2) se transformará en la siguiente ecuación equivalente

$$MA_1x_1 + MA_2x_2 + MA_3x_3 = p. \tag{3}$$

Aquí

$$(MA_1)^2 + (MA_2)^2 + (MA_3)^2 = 1.$$

Esto muestra que el vector

$$\mathbf{v} = (MA_1, MA_2, MA_3)$$

es unitario ($|\mathbf{v}| = 1$). Sus proyecciones sobre los ejes coordenados son iguales a

$$MA_1 = \cos\alpha_1,$$
$$MA_2 = \cos\alpha_2,$$
$$MA_3 = \cos\alpha_3,$$

donde α_1, α_2, α_3 son los ángulos formados por el vector $\mathbf{v}$ con las direcciones positivas de los ejes x_1, x_2, x_3, respectivamente. En virtud de las notaciones introducidas, la ecuación (3) tiene la forma

$$x_1\cos\alpha_1 + x_2\cos\alpha_2 + x_3\cos\alpha_3 = p \quad (p \geqslant 0), \tag{3'}$$

o sea, hemos obtenido la ecuación del plano (2) en la forma normal.

Si, dada la ecuación de un plano en la forma general (2), se necesita conocer su posición respecto del sistema de coordenadas, resulta suficiente reducir la ecuación (2) a la forma normal, multiplicándola por su factor normalizador M.

De la propia ecuación (2), sin hacer cálculo alguno sólo se pueden establecer dos hechos: 1) si $B = 0$, entonces el plano pasa por el origen de coordenadas, mientras que si $B \neq 0$, el plano no pasa por el origen de coordenadas; 2) el vector $A = (A_1, A_2, A_3)$ es perpendicular al plano, pues es colineal al vector unitario $\mathbf{v} = (MA_1, MA_2, MA_3)$, el cual es perpendicular al plano dado.

La ecuación

$$A_1x_1 + A_2x_2 + B = 0 \tag{4}$$

es un caso particular de la ecuación (2). En el plano (x_1, x_2) la ecuación (4) determina una recta, y en el espacio (x_1, x_2, x_3) es la ecuación del plano L que es perpendicular al plano coordenado (x_1, x_2) y que pasa por esta recta. Cualquiera que sea el punto (x_1, x_2, x_3), perteneciente al plano L, sus coordenadas x_1, x_2 satisfacen la ecuación (4) independientemente de la tercera coordenada x_3 que tenga. La ecuación

$$A_1x_1 + B = 0 \quad (A_1 \neq 0) \tag{5}$$

es un caso particular de la ecuación (4). Esta puede escribirse también en la forma

$$x_1 = C \quad (C = -B/A_1). \tag{5'}$$

La ecuación (5') en el espacio (x_1, x_2, x_3) es el lugar geométrico de los puntos (x_1, x_2, x_3) que tienen la primera coordenada igual al número C. Las coordenadas x_2, x_3 pueden ser en este caso cualesquiera. Está claro que (5') determina un plano paralelo al plano coordenado x_2, x_3 (o es perpendicular al eje x_1).

9.3. Ecuación "segmentaria" del plano

Si $A_i \neq 0$ $(i = 1, 2, 3)$ y $B \neq 0$, entonces la ecuación (2) se puede escribir así:

$$\frac{x_1}{a_1}+\frac{x_2}{a_2}+\frac{x_3}{a_3}=1, \tag{6}$$

donde $a_i = -B/A_i$ $(i = 1, 2, 3)$. La ecuación (6) se llama *ecuación «segmentaria» del plano.*

Este plano (fig. 21) corta el eje Ox_1 en el punto $(a_1, 0, 0)$, el eje Ox_2 en el punto $(0, a_2, 0)$ y el eje Ox_3 en el punto $(0, 0, a_3)$.

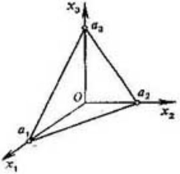

Fig. 21

Según sea la ecuación (6) puede uno figurarse la posición del plano respecto del sistema de coordenadas.

9.4. Ecuación del plano que pasa por un punto

Si el punto $x^0 = (x_1^0, x_2^0, x_3^0)$ está situado en el plano (2), entonces sus coordenadas satisfacen la ecuación (2):

$$\sum_{i=1}^{3} A_i x_i^0 + B = 0. \tag{7}$$

Restando (7) de (2), obtenemos:

$$\sum_{i=1}^{3} A_i (x_i - x_i^0) = 0. \tag{8}$$

La ecuación (8) se llama *ecuación del plano que pasa por el punto* $x^0 = (x_1^0, x_2^0, x_3^0)$. En forma vectorial la ecuación (8) se escribe así:

$$A\,(x - x^0) = 0. \tag{8'}$$

Como el vector $x - x^0$, aplicado al punto x^0, pertenece al plano (2), la igualdad (8') denota que el vector A es ortogonal al plano (2), que es lo que se estableció anteriormente mediante otros razonamientos.

9.5. Ecuación del plano que pasa por tres puntos

Sean dados tres puntos no situados en una recta

$$x^1 = (x_1^1, x_2^1, x_3^1),$$
$$x^2 = (x_1^2, x_2^2, x_3^2),$$
$$x^3 = (x_1^3, x_2^3, x_3^3).$$

Se pide hallar la ecuación del plano que pasa por estos tres puntos. Por la geometría se sabe que tal plano existe y es único. Como pasa por el punto x^1, su ecuación tiene la forma

$$A_1 (x_1 - x_1^1) + A_2 (x_2 - x_2^1) + A_3 (x_3 - x_3^1) = 0, \tag{9}$$

donde A_1, A_2, A_3 no son simultáneamente iguales a cero. Ahora bien, como también pasa por los puntos x^2, x^3, tienen que cumplirse las condiciones

$$\left.\begin{aligned} A_1 (x_1^2 - x_1^1) + A_2 (x_2^2 - x_2^1) + A_3 (x_3^2 - x_3^1) = 0, \\ A_1 (x_1^3 - x_1^1) + A_2 (x_2^3 - x_2^1) + A_3 (x_3^3 - x_3^1) = 0. \end{aligned}\right\} \tag{10}$$

Formemos el sistema de ecuaciones lineales homogéneas respecto de las incógnitas (z_1, z_2, z_3):

$$\left.\begin{aligned} (x_1 - x_1^1) z_1 + (x_2 - x_2^1) z_2 + (x_3 - x_3^1) z_3 = 0, \\ (x_1^2 - x_1^1) z_1 + (x_2^2 - x_2^1) z_2 + (x_3^2 - x_3^1) z_3 = 0, \\ (x_1^3 - x_1^1) z_1 + (x_2^3 - x_2^1) z_2 + (x_3^3 - x_3^1) z_3 = 0. \end{aligned}\right\} \tag{11}$$

Aquí $x = (x_1, x_2, x_3)$ es un punto arbitrario, que satisface la ecuación del plano (9). En virtud de (9) y (10), el vector no trivial $A = (A_1, A_2, A_3)$ satisface el sistema (11), por lo que el determinante de este sistema es igual a cero

$$\begin{vmatrix} x_1 - x_1^1 & x_2 - x_2^1 & x_3 - x_3^1 \\ x_1^2 - x_1^1 & x_2^2 - x_2^1 & x_3^2 - x_3^1 \\ x_1^3 - x_1^1 & x_2^3 - x_2^1 & x_3^3 - x_3^1 \end{vmatrix} = 0. \tag{12}$$

Hemos obtenido una ecuación de la forma (9), o sea, la ecuación de un plano; esto se comprueba fácilmente desarrollando el determinante obtenido por los elementos de la primera fila. Además, este plano pasa por los puntos x^1, x^2, x^3, como se deduce de las propiedades de los determinantes. El problema planteado queda resuelto.

La ecuación (12) también puede escribirse en la forma siguiente:

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ x_1^1 & x_2^1 & x_3^1 & 1 \\ x_1^2 & x_2^2 & x_3^2 & 1 \\ x_1^3 & x_2^3 & x_3^3 & 1 \end{vmatrix} = 0. \tag{13}$$

El valor del determinante (13) no varía si de la primera, tercera y cuarta filas restamos la segunda. Desarrollando el resultado después por los elementos de la cuarta columna, obtenemos la ecuación (12).

9.6. Angulo formado por dos planos

Consideremos dos planos

$$\left.\begin{array}{l} A_1x_1+A_2x_2+A_3x_3+B=0, \\ A_1'x_1+A_2'x_2+A_3'x_3+B'=0. \end{array}\right\} \qquad (14)$$

Es sabido que los vectores $A=(A_1, A_2, A_3)$ y $A'=(A_1', A_2', A_3')$ son perpendiculares a los planos dados, respectivamente, por lo cual, el ángulo φ formado por A y A' es igual al ángulo (diedro) formado por estos planos. Pero el producto escalar es igual a

$$AA'=|A|\,|A'|\cos\varphi,$$

por lo que

$$\cos\varphi=\frac{AA'}{|A|\,|A'|}=\frac{A_1A_1'+A_2A_2'+A_3A_3'}{\sqrt{A_1^2+A_2^2+A_3^2}\sqrt{A_1'^2+A_2'^2+A_3'^2}}. \qquad (15)$$

Es suficiente considerar que $0\leqslant\varphi\leqslant\pi$.

Señalemos que, en realidad, dos planos que se cortan forman dos ángulos diedros φ_1 y φ_2. Su suma es igual a π ($\varphi_1+\varphi_2=\pi$), y sus cosenos son iguales en valor absoluto pero se diferencian en el signo ($\cos\varphi_1=-\cos\varphi_2$). Si en la primera ecuación (14) se sustituyen los números A_1, A_2, A_3 por los números $-A_1, -A_2, -A_3$, respectivamente, entonces la ecuación obtenida determinará el mismo plano, pero en (15) el ángulo φ quedará sustituido por $\pi-\varphi$.

Evidentemente, la *condición de perpendicularidad de dos planos* (14) es $\cos\varphi=0$, o sea,

$$AA'=A_1A_1'+A_2A_2'+A_3A_3'=0. \qquad (16)$$

Dos planos (14) son *paralelos* si, y sólo si, los vectores A y A' (perpendiculares a ellos) son colineales, o sea, que se verifican las condiciones de proporcionalidad

$$\frac{A_1}{A_1'}=\frac{A_2}{A_2'}=\frac{A_3}{A_3'}. \qquad (17)$$

Si, además, se verifica la condición ampliada de proporcionalidad

$$\frac{A_1}{A_1'}=\frac{A_2}{A_2'}=\frac{A_3}{A_3'}=\frac{B}{B'}, \qquad (18)$$

entonces, esto significa que los planos (14) *coinciden*, o sea, que ambas ecuaciones (14) determinan un mismo plano.

A pesar de que no se puede dividir por cero, sin embargo, resulta cómodo escribir las proporciones simbólicas (17) o (18) con ceros en los denominadores. Pero, entonces, si, por ejemplo, $A_2' = 0$, hay que considerar que también $A_2 = 0$. O bien, si $B' = 0$, entonces $B = 0$.

EJEMPLOS. Las ecuaciones

$$2x + 4y + 1 = 0,$$
$$x + 2y + 5 = 0$$

determinan un par de planos paralelos, y las ecuaciones

$$2x + 4y + 2 = 0,$$
$$x + 2y + 1 = 0,$$

un par de planos que coinciden.

9.7. Distancia de un punto a un plano

Se pide calcular la distancia de un punto $x^0 = (x_1^0, x_2^0, x_3^0)$ a un plano L determinado por la ecuación

$$A_1x_1 + A_2x_2 + A_3x_3 + B = 0.$$

Para ello, reducimos la ecuación de L a la forma normal:

$$(x, \mathbf{v}) = p \quad (p \geqslant 0).$$

En la fig. 22 se observa que la diferencia $x - x^0$ entre el radio vector de un punto arbitrario x del plano L y el radio vector del punto x^0

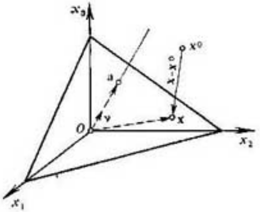

Fig. 22

es un vector tal, que el valor absoluto de su proyección sobre $\mathbf{v}$ es igual a la distancia buscada d de x^0 a L:

$$d = |(x - x^0, \mathbf{v})|.$$

Pero
$$(x - x^0, \nu) = (x, \nu) - (x^0, \nu) = p - (x^0, \nu),$$
y, por consiguiente,
$$d = |(x^0, \nu) - p|.$$
Vemos que, para calcular la distancia d del punto x^0 al plano L hay que escribir la ecuación del plano L en la forma normal, trasladar p al primer miembro y sustituir en la última ecuación x por x^0. El valor absoluto de la expresión obtenida es precisamente el número buscado d.

Evidentemente, en términos de parámetros del plano
$$d = \frac{|A_1x_1^0 + A_2x_2^0 + A_3x_3^0 + B|}{\sqrt{A_1^2 + A_2^2 + A_3^2}}.$$

Fácilmente se observa que si el punto x^0 y el origen de coordenadas están situados a distintos lados del plano L (como en la fig. 22), entonces el vector $x - x^0$ forma con ν un ángulo obtuso y por ello,
$$d = -(x - x^0, \nu) = (x^0, \nu) - p > 0.$$
Ahora bien, si el punto x^0 y el origen de coordenadas están situados a un mismo lado de L, entonces el ángulo indicado es agudo y
$$d = (x - x^0, \nu) = p - (x^0, \nu) > 0.$$
Por consiguiente, en el primer caso $(x^0, \nu) > p$, y en el segundo $(x^0, \nu) < p$.

EJEMPLO 1. La distancia d del punto (1, 1, 1) al plano L
$$x + 2y + z = 3$$
es igual a
$$d = \left|\frac{x + 2y + z - 3}{\sqrt{1 + 4 + 1}}\right|_{x=1,\ y=1,\ z=1} = \frac{1}{\sqrt{6}}.$$
En este caso, el punto (1, 1, 1) y el origen de coordenadas están situados a distintos lados del plano L, ya que $M > 0$, y
$$[x + 2y + z - 3]_{x=1,\ y=1,\ z=1} = 1 > 0.$$

9.8. Problemas

1. Reducir a la forma normal las ecuaciones de los planos
$$x - y + z = 2, \quad 2x - y + \sqrt{20}z = 10.$$
2. Hallar el ángulo formado por los planos
$$x + y + z = 1 \quad \text{y} \quad y = 0;$$
$$x - y + z = 2 \quad \text{y} \quad x + y + z = 3.$$
3. Escribir la ecuación del plano que pasa por los puntos (0, 0, 1), (1, 0, 1), (1, 1, 0).
4. Escribir la ecuación de la esfera con el centro en el origen de coordenadas y que es tangente al plano
$$2x - 6y + 7z = 2.$$

§ 10. La recta en el espacio

10.1. Ecuación de la recta en la forma canónica

Consideremos en el espacio una recta arbitraria L. Señalemos en ella un punto x^0 y un vector situado en la misma $a = (a_1, a_2, a_3) \neq 0$, aplicado al punto x^0 (fig. 23). Denotaremos por x un punto arbitrario (un punto variable) de la recta L. El vector $x - x^0$ se puede expresar en la forma $x - x^0 = ta$, donde t es un número (un escalar). Si la variable real t recorre el intervalo $(-\infty, \infty)$, entonces $x = x^0 + ta$ recorre toda la recta L. Por eso, se dice que la igualdad

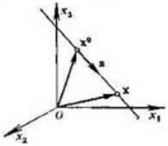

Fig. 23

$$x - x^0 = ta \quad (-\infty < t < \infty) \tag{1}$$

es la *ecuación vectorial de la recta que pasa por el punto x^0 y que lleva la dirección del vector a.*

En términos de coordenadas la ecuación (1) se descompone en tres ecuaciones:

$$\left.\begin{aligned} x_1 - x_1^0 &= ta_1, \\ x_2 - x_2^0 &= ta_2, \\ x_3 - x_3^0 &= ta_3. \end{aligned}\right\} \tag{1'}$$

Eliminando entre ellas el parámetro t, obtenemos las ecuaciones de la recta (un sistema de dos ecuaciones)

$$\frac{x_1 - x_1^0}{a_1} = \frac{x_2 - x_2^0}{a_2} = \frac{x_3 - x_3^0}{a_3}, \tag{1''}$$

donde los números a_1, a_2, a_3 no son simultáneamente iguales a cero. Las ecuaciones (1″) se llaman *ecuaciones de la recta en la forma canónica.*

Nota. Puede ocurrir que uno o dos de los números a_1, a_2, a_3 sean iguales a cero. Entonces, a pesar de todo, está convenido escribir las igualdades (1″) con el cero o con los dos ceros en los denominadores. En tal caso, esta expresión resulta simbólica, pero muy práctica.

EJEMPLO 1. Las ecuaciones

$$\frac{x-1}{1} = \frac{y-2}{0} = \frac{z-3}{2} \tag{2}$$

determinan una recta en el espacio, que pasa por el punto $(1, 2, 3)$ y lleva la dirección del vector $(1, 0, 2)$.

Estas ecuaciones se pueden sustituir por las equivalentes siguientes:

$$y - 2 = 0 \cdot (x - 1), \quad 2(x - 1) = z - 3,$$

o sea,

$$y = 2, \quad z = 2x + 1. \tag{2'}$$

Por lo tanto, la recta considerada es la intersección de los dos planos que se determinan por las ecuaciones (2').

EJEMPLO 2. Las ecuaciones de la recta

$$\frac{x-1}{1} = \frac{y-2}{0} = \frac{z-3}{0}$$

son equivalentes a las siguientes:

$$y - 2 = 0, \quad z - 3 = 0.$$

10.2. Posición recíproca de dos planos

Sean dadas las ecuaciones de dos planos

$$A_1x + A_2y + A_3z + B = 0, \tag{3}$$

$$A_1'x + A_2'y + A_3'z + B' = 0. \tag{4}$$

Si los coeficientes A_j de la primera ecuación son respectivamente proporcionales a los coeficientes A_j' de la segunda, ($A_1 : A_2 : A_3 = A_1' : A_2' : A_3'$), entonces los planos (3) y (4) son paralelos o incluso coinciden (si se cumple la condición $A_1 : A_2 : A_3 : B = A_1' : A_2' : A_3' : B'$) (véase § 9, (17) y (18)). En caso contrario, los planos (3) y (4) se cortan por una recta. En este caso uno de los determinantes

$$\begin{vmatrix} A_1 & A_2 \\ A_1' & A_2' \end{vmatrix}, \quad \begin{vmatrix} A_1 & A_3 \\ A_1' & A_3' \end{vmatrix}, \quad \begin{vmatrix} A_2 & A_3 \\ A_2' & A_3' \end{vmatrix}$$

es distinto de cero. Para precisar, supongamos que el primero:

$$\begin{vmatrix} A_1 & A_2 \\ A_1' & A_2' \end{vmatrix} \neq 0. \tag{5}$$

Entonces las ecuaciones (3), (4) se pueden resolver respecto de x, y, obteniendo

$$\left.\begin{aligned} x &= \alpha z + \mu, \\ y &= \beta z + \nu, \end{aligned}\right\} \tag{6}$$

donde α, β, μ, ν son ciertos números. Las ecuaciones (6) son equivalentes a las siguientes:

$$\frac{x-\mu}{\alpha} = \frac{y-\nu}{\beta} = \frac{z}{1}. \tag{7}$$

Vemos que con la condición (5), las ecuaciones de los dos planos (3), (4) determinan una recta (7), o sea, el lugar geométrico de los puntos cuyas coordenadas satisfacen las ecuaciones (3), (4). Esta recta pasa por el punto $(\mu, \nu, 0)$ y lleva la dirección del vector $(\alpha, \beta, 1)$. Los números α, β o uno de ellos pueden ser iguales a cero, pero entonces las ecuaciones (7) tienen carácter simbólico.

EJEMPLO 3. Evidentemente, la recta determinada por las ecuaciones $x = 0$, $y = 0$ es el eje coordenado z. Se podía haber obtenido este resultado operando formalmente. Se tiene,

$$x = 0\cdot z, \quad y = 0\cdot z,$$

de donde

$$\frac{x}{0} = \frac{y}{0} = \frac{z}{1},$$

o sea, hemos obtenido las ecuaciones de la recta que pasa por el origen de coordenadas $(0, 0, 0)$ en dirección del vector $(0, 0, 1)$. Está claro que esta recta es el eje z.

EJEMPLO 4. Hallar el ángulo formado por las rectas

$$\frac{x_1 - x_1^1}{a_1} = \frac{x_2 - x_2^1}{a_2} = \frac{x_3 - x_3^1}{a_3}, \tag{8}$$

$$\frac{x_1 - x_1^2}{b_1} = \frac{x_2 - x_2^2}{b_2} = \frac{x_3 - x_3^2}{b_3}. \tag{9}$$

Los vectores $\boldsymbol{a} = (a_1, a_2, a_3)$, $\boldsymbol{b} = (b_1, b_2, b_3)$ están situados sobre las rectas en cuestión y, como convenimos, están aplicados a los puntos $\boldsymbol{x}^1 = (x_1^1, x_2^1, x_3^1)$, $\boldsymbol{x}^2 = (x_1^2, x_2^2, x_3^2)$, respectivamente. Por ello, el ángulo φ formado por estos vectores es precisamente el ángulo formado por las rectas (8) y (9):

$$\cos\varphi = \frac{a_1b_1 + a_2b_2 + a_3b_3}{|\boldsymbol{a}|\,|\boldsymbol{b}|}.$$

10.3. Problemas

1. Escribir las ecuaciones de la recta que pasa por el punto $(2, -1, 0)$ y es perpendicular al plano $2x + z - 4y = 7$.

2. Escribir la ecuación del plano que pasa por el punto $(1, -1, 2)$ y es perpendicular a la recta determinada por las ecuaciones $2x + 3y = 1$, $3x - z = 1$.

§ 11. Orientación de los sistemas rectangulares de coordenadas

11.1. Sistema bidimensional de coordenadas

En las figs. 24 y 25 están representados sendos sistemas de coordenadas (x_1, x_2). Estos sistemas son distintos; se dice que son de *orientación contraria*. En el caso de la fig. 24, mediante una rotación del eje x_1 en torno del punto O un ángulo $\frac{\pi}{2}$, se puede hacer coincidir las direcciones de los ejes x_1 y x_2 solamente si la rotación se efectúa en sentido contrario a la del movimiento de las agujas del reloj.

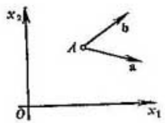

Fig. 24

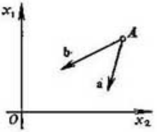

Fig. 25

Sin embargo, en el caso de la fig. 25 se puede conseguir esto sólo si se gira el eje x_1 en el sentido del movimiento de las agujas del reloj. Trasladando como un sólido, sobre el plano considerado, el sistema de coordenadas representado en la fig. 24, resulta imposible (!) hacerlo coincidir con el sistema representado en la fig. 25 de tal modo que coincidan las direcciones de los ejes correspondientes.

En la fig. 24, así como en la fig. 25, está representado un par de vectores no colineales $\boldsymbol{a}$ y $\boldsymbol{b}$, aplicados a un punto A. Trasladando este par de vectores como un sólido sobre el plano, podemos conseguir que el punto A coincida con el origen de coordenadas O. Propongámonos conseguir mediante una rotación de cada uno de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ en torno del punto O, que el vector $\boldsymbol{a}$ tome la dirección del eje x_1 y que el vector $\boldsymbol{b}$ resulte situado en el eje x_2. Se exige que durante este proceso los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ permanezcan situados constantemente en el plano considerado y que el ángulo formado por ellos no sea igual a 0 o a π. Evidentemente, siempre se puede conseguir esto. Primero giramos el sistema de vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ como un sólido en torno del punto O hasta que coincida el vector $\boldsymbol{a}$ con la dirección positiva del eje x_1. Como los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ no son colineales, el vector $\boldsymbol{b}$ resultará situado en el semiplano superior o inferior. Después giramos el vector $\boldsymbol{b}$ un ángulo necesario para que quede situado en el eje x_2; en este caso, no se permite que el vector $\boldsymbol{b}$ se sitúe sobre el eje x_1. Por esta razón, puede ocurrir que la dirección del vector $\boldsymbol{b}$

coincida con la dirección del eje x_2 (esto es posible si el vector b estaba situado en el semiplano superior), o bien, que el vector b resulte dirigido hacia la dirección negativa del eje x_2 (véase la fig. 26,

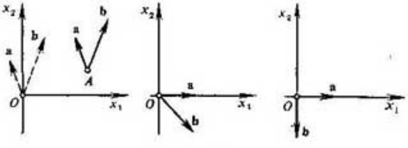

Fig. 26

donde se muestra la dinámica del proceso). En el primer caso se dice que el *par de vectores* (a, b) *está orientado igual que el sistema de coordenadas* (x_1, x_2), y en el segundo caso, que la *orientación del par* (a, b) *es contraria a la de* (x_1, x_2).

11.2. Sistema tridimensional de coordenadas

Los sistemas rectangulares de coordenadas en el espacio x_1, x_2, x_3 representados en las figs. 27 y 28, también son distintos. Considerando el sistema de coordenadas de la fig. 27 como un sólido, después de una traslación adecuada del mismo, se puede conseguir que

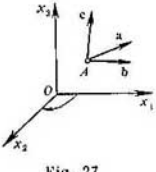

Fig. 27

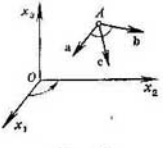

Fig. 28

coincidan los ejes x_1 y x_2 de ambos sistemas. Pero la dirección positiva del eje x_3 del primer sistema no coincidirá con la dirección positiva del eje x_3 del segundo sistema. Se dice que los sistemas de las figs. 27 y 28 son *de orientación contraria*. El sistema de la fig. 27 se llama *sistema inverso de coordenadas* (sinistrórsum), mientras que el de la fig. 28, *sistema directo de coordenadas* (dextrórsum). Si un tornillo con rosca de a la derecha (a la izquierda) se atornilla en dirección del eje x_3, haciéndole girar según indica la flecha de la fig. 28 (fig. 27), entonces realizará un desplazamiento en esta dirección.

También se puede distinguir el sistema de coordenadas según la siguiente regla. Si desde algún punto del semieje positivo Ox_3 se observa la dirección positiva del semieje Ox_2, entonces, el semieje positivo Ox_1 puede quedar hacia la izquierda o hacia la derecha. En el primer caso, el sistema de coordenadas se llama *sinistrórsum* (fig. 27), y en el segundo, *dextrórsum* (fig. 28).

Los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ se llaman *coplanares* si están situados en un plano o en planos paralelos.

Tomemos un sistema de vectores no coplanares $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, aplicados a un punto A. Giremos el vector $\boldsymbol{b}$, en el plano de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$, en torno del punto A hasta que $\boldsymbol{b}$ resulte perpendicular a $\boldsymbol{a}$. Nos preocuparemos que durante el movimiento el ángulo formado por $\boldsymbol{a}$ y $\boldsymbol{b}$ nunca sea igual a cero o a π. Después de esto, giramos el vector $\boldsymbol{c}$ en torno de A con el fin de que obtenga una dirección perpendicular a los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$. También tendremos cuidado que el vector $\boldsymbol{c}$ nunca se sitúe en el plano de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$. Como resultado, los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ quedarán perpendiculares entre sí. Traslademos ahora esta terna como un sólido al punto O y hagámosla girar en torno del punto O con el objeto de que los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ obtengan las direcciones de los ejes x_1 y x_2, respectivamente. Pueden resultar dos casos:

1) que el vector $\boldsymbol{c}$ tenga la dirección del semieje positivo x_3;

2) que tenga la dirección opuesta.

En el primer caso, se dice que el sistema inicial de vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ tiene *la misma orientación que el sistema de coordenadas* x_1, x_2, x_3, y en el segundo caso, que la *orientación es contraria* (véanse las figs. 27 y 28, respectivamente).

§ 12. Producto vectorial

12.1. Dos definiciones de producto vectorial

Tomemos dos vectores en un sistema rectangular de coordenadas del espacio real tridimensional

$$\boldsymbol{a} = (a_1,\ a_2,\ a_3), \quad \boldsymbol{b} = (b_1,\ b_2,\ b_3)$$

y llamemos *producto vectorial* de $\boldsymbol{a}$ y $\boldsymbol{b}$ al vector

$$\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = [\boldsymbol{a} \times \boldsymbol{b}] = (a_2b_3 - a_3b_2)\,\boldsymbol{i} + (a_3b_1 - a_1b_3)\,\boldsymbol{j} + (a_1b_2 - a_2b_1)\,\boldsymbol{k} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \qquad (1)$$

El último término está escrito en forma de un «determinante generalizado» cuya primera fila consta de los vectores $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ (los vectores unitarios del sistema de coordenadas). En el segundo término se muestra cómo se debe entender este determinante generalizado (el determinante de tercer orden se desarrolla por los elementos de la primera fila, como si $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ fuesen números).

Evidentemente,

$$[\boldsymbol{a} \times (-\boldsymbol{b})] = -[\boldsymbol{a} \times \boldsymbol{b}].$$

El producto vectorial de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ también puede definirse del modo siguiente:

1) si $\boldsymbol{a}$ y $\boldsymbol{b}$ son colineales, entonces su producto vectorial es igual a cero;

2) si $\boldsymbol{a}$ y $\boldsymbol{b}$ no son colineales, entonces el vector $\boldsymbol{c}$ es perpendicular a estos vectores y, además, de tal modo que la orientación del siste-

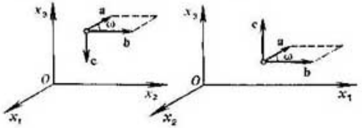

Fig. 29

ma $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ sea igual que la del sistema de coordenadas dado. La longitud del vector $\boldsymbol{c}$ es igual a

$$|\boldsymbol{c}| = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \operatorname{sen} \omega \quad (0 \leqslant \omega \leqslant \pi), \tag{2}$$

donde ω es el ángulo formado por $\boldsymbol{a}$ y $\boldsymbol{b}$, o sea, la longitud de $\boldsymbol{c}$ es igual al área del paralelogramo construido sobre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ (fig. 29).

Demostremos la equivalencia de las dos definiciones enunciadas.

Si el vector $\boldsymbol{c} = \boldsymbol{0}$, entonces de (1) se deduce que las componentes de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ son proporcionales:

$$a_1 : a_2 : a_3 = b_1 : b_2 : b_3,$$

o sea, los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ son colineales; pero, entonces, el producto vectorial también es igual a cero según la segunda definición. Recíprocamente, si los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ son colineales, entonces, según la segunda definición $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{0}$. Como, en este caso, las componentes de los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ son proporcionales, resulta según la primera definición que $\boldsymbol{c} = \boldsymbol{0}$.

Supongamos ahora que $\boldsymbol{a}$ y $\boldsymbol{b}$ son vectores no colineales y sea $\boldsymbol{c}$ su producto vectorial según (1). Está claro que

$$(\boldsymbol{c}, \boldsymbol{a}) = (a_2b_3 - a_3b_2)\, a_1 + (a_3b_1 - a_1b_3)\, a_2 + \\ + (a_1b_2 - a_2b_1)\, a_3 = 0$$

y de un modo similar,

$$(c,\ b) = 0.$$

Así pues, el vector c es perpendicular a a y a b.

Demostremos que el sistema de vectores $(a,\ b,\ c)$ tiene la misma orientación que el sistema de coordenadas $(x_1,\ x_2,\ x_3)$. Hagamos girar continuamente los vectores a y b en torno del punto O, calculando siempre según ellos el vector c, pero de tal modo que a y b sigan siendo todo el tiempo no colineales. Entonces, el vector c siempre será distinto de cero ($c \neq 0$) y el sistema $(a,\ b,\ c)$ siempre será no coplanario. Realicemos unas rotaciones tales que los vectores a y b obtengan las direcciones de los ejes x_1 y x_2, respectivamente, o sea, que tengan la forma $a = (|a|,\ 0,\ 0)$, $b = (0,\ |b|,\ 0)$. Esto siempre puede conseguirse, ya que en el caso dado el plano de los vectores a, b puede irar en el espacio. Pero, entonces, el vector c, calculado según la fórmula (1) tiene la forma $c = (0,\ 0,\ |a|\,|b|)$. Vemos que finalmente los vectores $(a,\ b,\ c)$ han quedado orientados igual que los ejes $(x_1,\ x_2,\ x_3)$. Según la definición de orientación (véase § 11), entonces, el sistema inicial a, b, c tiene la misma orientación que el sistema de coordenadas $(x_1,\ x_2,\ x_3)$.

En resumen, el producto vectorial $c = a \times b$, definido según la fórmula (1), es un vector perpendicular a los vectores a y b y el sistema de vectores $(a,\ b,\ c)$ tiene la misma orientación que el sistema considerado de coordenadas x_1, x_2, x_3.

Todavía tenemos que demostrar la fórmula (2). Supongamos que los vectores a y b forman con los ejes coordenados los ángulos α_1, α_2, α_3 y β_1, β_2, β_3, respectivamente. Como

$$a_j = |a| \cos \alpha_j, \quad b_j = |b| \cos \beta_j \quad (j = 1,\ 2,\ 3),$$

resulta

$$\begin{aligned}
|c|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 = \\
= |a|^2\,|b|^2\,[(\cos \alpha_2 \cos \beta_3 - \cos\alpha_3 \cos \beta_2)^2 + \\
+ (\cos \alpha_3 \cos \beta_1 - \cos \alpha_1 \cos \beta_3)^2 + (\cos \alpha_1 \cos \beta_2 - \cos \alpha_2 \cos \beta_1)^2] = \\
= |a|^2\,|b|^2\,[(\cos^2 \alpha_1 + \cos^2 \alpha_2 + \cos^2 \alpha_3) \times \\
\times (\cos^2 \beta_1 + \cos^2 \beta_2 + \cos^2 \beta_3) - \\
- (\cos \alpha_1 \cos \beta_1 + \cos \alpha_2 \cos \beta_2 + \cos \alpha_3 \cos \beta_3)^2] = \\
= |a|^2\,|b|^2\,[1 \cdot 1 - \cos^2 \omega] = |a|^2\,|b|^2 \sin^2 \omega
\end{aligned}$$

donde ω es el ángulo formado por los vectores a y b

$$\left(\cos \omega = \sum_{j=1}^{3} \cos\alpha_j \cos\beta_j\right).$$

Así pues, hemos demostrado (2).

Por lo tanto, hemos demostrado completamente que de la definición de producto vectorial según la fórmula (1) se deduce su segunda definición.

Recíprocamente, si un vector se determina según la segunda definición, entonces, éste es único, ya que sólo puede haber un vector que sea perpendicular a $\boldsymbol{a}$ y a $\boldsymbol{b}$ y cuya longitud sea igual al área del paralelogramo construido sobre $\boldsymbol{a}$ y $\boldsymbol{b}$ de tal modo que el sistema $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ tenga la misma orientación que el sistema (x_1, x_2, x_3). Pero, entonces, éste es el vector $\boldsymbol{c}$, determinado según la fórmula (1), puesto que, como ya hemos comprobado, este último posee las propiedades indicadas.

Señalemos una vez más que la condición $\boldsymbol{a} \times \boldsymbol{b} = 0$ es *condición necesaria y suficiente para que los vectores* $\boldsymbol{a}$ *y* $\boldsymbol{b}$ *sean colineales.*

12.2. Significado geométrico del determinante de segundo orden

Consideremos ahora especialmente dos vectores no nulos

$$\boldsymbol{a} = (a_1, a_2), \quad \boldsymbol{b} = (b_1, b_2) \tag{3}$$

en un sistema rectangular de coordenadas (x_1, x_2) (figs. 30, 31). Vamos a considerar que el plano en cuestión está situado en el espacio y agreguemos a los ejes x_1, x_2 un eje x_3 perpendicular a ellos. Los

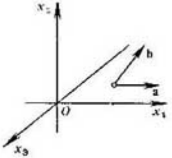

Fig. 30

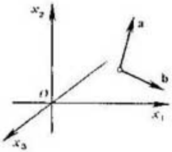

Fig. 31

vectores $\boldsymbol{a}$, $\boldsymbol{b}$ tendrán ahora las coordenadas

$$\boldsymbol{a} = (a_1, a_2, 0), \quad \boldsymbol{b} = (b_1, b_2, 0)$$

y su producto vectorial será igual a

$$\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \boldsymbol{k}. \tag{4}$$

donde $\boldsymbol{k}$ es el vector unitario del eje x_3. Por definición de producto vectorial, el sistema $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ está orientado igual que el sistema de coordenadas (x_1, x_2, x_3). Por esto, si el determinante

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} > 0,$$

el sistema de vectores $(\boldsymbol{a}, \boldsymbol{b})$ tiene que tener la misma orientación que el sistema de coordenadas (x_1, x_2). Pero si el determinante

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} < 0,$$

la orientación del sistema $(\boldsymbol{a}, \boldsymbol{b})$ es contraria a la de (x_1, x_2). En la fig. 30 está representada la posición de los vectores $(\boldsymbol{a}, \boldsymbol{b})$ en el primer caso, y en la fig. 31, en el segundo caso. También sabemos que el área del paralelogramo construido sobre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es igual a (véase (4))

$$S = |\boldsymbol{c}| = \left\| \begin{matrix} a_1 & a_2 \\ b_1 & b_2 \end{matrix} \right\|,$$

o sea, es igual al valor absoluto del determinante

$$\Delta = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}. \tag{5}$$

Así pues, hemos demostrado que:

1) el valor absoluto del determinante (5) es igual al área del paralelogramo construido sobre los vectores $\boldsymbol{a} = (a_1, a_2)$ y $\boldsymbol{b} = (b_1, b_2)$;

2) el signo del determinante (5) depende de posición de estos vectores respecto del sistema rectangular de coordenadas (x_1, x_2). Al signo + le corresponde un sistema $(\boldsymbol{a}, \boldsymbol{b})$ de la misma orientación que (x_1, x_2), y al signo —, de orientación contraria.

12.3. Propiedades del producto vectorial

Se verifican las propiedades siguientes:

$$\boldsymbol{a} \times \boldsymbol{b} = -[\boldsymbol{b} \times \boldsymbol{a}], \tag{6}$$

$$\boldsymbol{a} \times (\alpha \boldsymbol{b}) = \alpha\, [\boldsymbol{a} \times \boldsymbol{b}], \tag{7}$$

$$\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = [\boldsymbol{a} \times \boldsymbol{b}] + [\boldsymbol{a} \times \boldsymbol{c}], \tag{8}$$

donde $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ son vectores arbitrarios y α es un escalar.

Estas igualdades se obtienen fácilmente expresando los productos vectoriales que figuran en (6), (7) y (8) mediante las componentes de los vectores

$$\boldsymbol{a} = (a_1, a_2, a_3), \quad \boldsymbol{b} = (b_1, b_2, b_3), \quad \boldsymbol{c} = (c_1, c_2, c_3),$$

según la fórmula (1)

Las fórmulas (6) y (7) también se deducen fácilmente mediante razonamientos geométricos. Sean $\boldsymbol{a}$ y $\boldsymbol{b}$ unos vectores no colineales. Si en el producto vectorial se permutan $\boldsymbol{a}$ y $\boldsymbol{b}$, entonces el área del paralelogramo construido sobre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$, así como la perpendicular a $\boldsymbol{a}$ y a $\boldsymbol{b}$, no varían. Solamente cambiará la dirección de $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$ por la opuesta, lo que implicará el cambio de orientación de la terna $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b})$.

La multiplicación del vector $\boldsymbol{b}$ por un número positivo α solamente implica el aumento del área del paralelogramo, construido sobre $\boldsymbol{a}$ y $\boldsymbol{b}$, α veces, mientras que la dirección del producto vectorial permanece invariable. Si $\alpha < 0$, se tiene

$$\boldsymbol{a} \times (\alpha \boldsymbol{b}) = \boldsymbol{a} \times (-\mid \alpha \mid \boldsymbol{b}) = \mid \alpha \mid [\boldsymbol{a} \times (-\boldsymbol{b})] =$$
$$= -\mid \alpha \mid [\boldsymbol{a} \times \boldsymbol{b}] = \alpha [\boldsymbol{a} \times \boldsymbol{b}].$$

Obsérvese también que de (6) y (7) se deduce que

$$(\alpha \boldsymbol{a}) \times \boldsymbol{b} = -[\boldsymbol{b} \times (\alpha \boldsymbol{a})] = -\alpha [\boldsymbol{b} \times \boldsymbol{a}] = \alpha [\boldsymbol{a} \times \boldsymbol{b}].$$

EJEMPLO 1. Hallar el ángulo A del triángulo ABC con los vértices $A = (1, 2, 3)$, $B = (2, 2, 2)$, $C = (1, 2, 4)$.

Denotemos por ω el ángulo buscado. Por lo tanto, ω es el ángulo formado por los vectores $\overrightarrow{AB}$ y $\overrightarrow{AC}$. De la segunda definición de producto vectorial, se tiene

$$\operatorname{sen} \omega = \frac{\mid \overrightarrow{AB} \times \overrightarrow{AC} \mid}{\mid \overrightarrow{AB} \mid \mid \overrightarrow{AC} \mid},$$

donde

$$\overrightarrow{AB} = (1, 0, -1), \quad \overrightarrow{AC} = (0, 0, 1); \quad \mid \overrightarrow{AB} \mid = \sqrt{2}, \quad \mid \overrightarrow{AC} \mid = 1,$$
$$\overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{vmatrix} = -\boldsymbol{j}.$$

De aquí

$$\operatorname{sen} \omega = \frac{1}{\sqrt{2}}, \quad \omega = \frac{\pi}{4}, \quad \omega = \frac{3\pi}{4}.$$

Como

$$\overrightarrow{BC} = (-1, 0, 2) \quad \text{y} \quad \mid \overrightarrow{BC} \mid^2 = 5 > \mid \overrightarrow{AC} \mid^2 + \mid \overrightarrow{AB} \mid^2 = 1 + 2 = 3,$$

hay que tomar $\omega = 3\pi/4$.

Observación. Si en el triángulo ABC el ángulo A es recto, entonces $\mid \overrightarrow{BC} \mid^2 = BC^2 = AC^2 + AB^2$; si A es un ángulo obtuso, entonces

$BC^2 > AC^2 + AB^2$; si A es un ángulo agudo, entonces $BC^2 < AC^2 + AB^2$.

EJEMPLO 2. (de mecánica). Sean dados dos puntos A y B. Al punto B ha sido aplicada una fuerza, determinada por el vector $\boldsymbol{F}$.

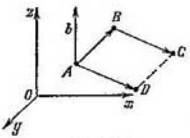

Fig. 31a

Hagamos $\boldsymbol{a} = \overrightarrow{AB}$. Se llama *momento de la fuerza* $\boldsymbol{F}$ *respecto del punto* O al producto vectorial del vector $\boldsymbol{a}$ por el vector $\boldsymbol{F}$

$$\boldsymbol{b} = \boldsymbol{a} \times \boldsymbol{F}$$

(véase la fig. 31a, $\overrightarrow{AD} = \overrightarrow{BC}$).

El vector $\boldsymbol{b}$ (el momento de la fuerza $\boldsymbol{F}$) es perpendicular a los vectores $\boldsymbol{a}$ y $\boldsymbol{F}$ y su longitud es igual al área del paralelogramo construído sobre los vectores $\boldsymbol{a}$ y $\boldsymbol{F}$. La dirección del vector $\boldsymbol{b}$ depende del sistema rectangular de coordenadas que se tome en este problema.

En la fig. 31a se ha tomado un sistema de coordenadas sinistrórsum. La dirección del vector $\boldsymbol{b}$ se ha tomado de tal modo que los vectores $\boldsymbol{a}$, $\boldsymbol{F}$, $\boldsymbol{b}$ también formen un sistema de sinistrórsum.

§ 13. Producto vectorial-escalar (mixto)

Se llama *producto vectorial-escalar (producto mixto) de los vectores* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ en el espacio real tridimensional, a un escalar que es igual al producto escalar del vector $\boldsymbol{a} \times \boldsymbol{b}$ por el vector $\boldsymbol{c}$:

$$(\boldsymbol{a} \times \boldsymbol{b})\,\boldsymbol{c} = (a_2b_3 - a_3b_2)\,c_1 + (a_3b_1 - a_1b_3)\,c_2 + (a_1b_2 - a_2b_1)\,c_3 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \qquad (1)$$

En virtud de la definición del producto escalar,

$$(\boldsymbol{a} \times \boldsymbol{b})\,\boldsymbol{c} = |\boldsymbol{a} \times \boldsymbol{b}|\,\mathrm{pr}_{\boldsymbol{a}\times\boldsymbol{b}}\boldsymbol{c} = (|\boldsymbol{a}|\cdot|\boldsymbol{b}|\,\mathrm{sen}\,\omega)\,\mathrm{pr}_{\boldsymbol{a}\times\boldsymbol{b}}\boldsymbol{c}.$$

Por esto, evidentemente, también se puede decir que el producto mixto $(\boldsymbol{a} \times \boldsymbol{b})\,\boldsymbol{c}$ es igual al volumen del paralelepípedo construido sobre los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, pero tomado con el signo + ó — según que el sistema de vectores a, b, c tenga la misma orientación que el sistema de coordenadas x_1, x_2, x_3 o la orientación contraria. Señalemos que $|\operatorname{pr}_{a\times b} c|$ es igual a la altura del paralelepípedo.

Se verifican las igualdades

$$(a \times b)\,c = (c \times a)\,b = (b \times c)\,a, \tag{2}$$

que se deducen fácilmente de las propiedades del determinante (1).

Si los vectores a, b, c están situados en un plano, entonces

$$(a \times b)\,c = 0,$$

puesto que $a \times b$ es perpendicular al vector c. Recíprocamente, si $(a \times b)\,c = 0$, entonces el vector c es perpendicular al vector $a \times b$ y, por consiguiente, está situado en el plano de los vectores a y b, o bien, en un plano paralelo a este plano.

Por lo tanto, la condición

$$(a \times b)\,c = 0$$

es *necesaria y suficiente para que los tres vectores* a, b, c *sean coplanares.*

EJEMPLO 1. Hallar la condición para que cuatro puntos pertenezcan a un plano.

Sean dados cuatro puntos $A_j = (x_j, y_j, z_j)$ $(j = 1, 2, 3, 4)$. Si estos puntos están situados en un plano, entonces los vectores $\overrightarrow{A_1A_2}$, $\overrightarrow{A_1A_3}$, $\overrightarrow{A_1A_4}$ también están situados en este plano y, por consiguiente, su producto mixto es igual a cero:

$$(\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3})\,\overrightarrow{A_1A_4} = \begin{vmatrix} x_2-x_1 & y_2-y_1 & z_2-z_1 \\ x_3-x_1 & y_3-y_1 & z_3-z_1 \\ x_4-x_1 & y_4-y_1 & z_4-z_1 \end{vmatrix} = 0.$$

Esta es la condición para que cuatro puntos pertenezcan a un plano (compárese con el § 9, (12)).

§ 14. Sistema de vectores linealmente independiente

Consideremos en R_n (real o complejo) un sistema de k vectores

$$a^s = (a_{s1}, a_{s2}, \ldots a_{sn}) \quad (s = 1, \ldots, k). \tag{1}$$

Por definición, *este sistema es linealmente independiente*, si la igualdad vectorial

$$\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_k a^k = 0. \tag{2}$$

donde $\lambda_1, \lambda_2, \ldots, \lambda_k$ son unos números (reales o complejos, respectivamente), implica que

$$\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

El sistema de vectores (1) se llama *linealmente dependiente*, si existen unos números $\lambda_1, \lambda_2, \ldots, \lambda_k$, no simultáneamente iguales a cero, tales que se verifica la igualdad (2). Si, para precisar, consideramos que $\lambda_k \neq 0$, entonces, de (2) se deduce que

$$a^k = \mu_1 a^1 + \ldots + \mu_{k-1} a^{k-1}, \qquad (3)$$

donde

$$\mu_1 = -\frac{\lambda_1}{\lambda_k}, \ \ldots, \ \mu_{k-1} = -\frac{\lambda_{k-1}}{\lambda_k}.$$

Por lo tanto, *si un sistema de k vectores es linealmente dependiente, entonces, uno de ellos es una combinación lineal de los demás*, o como suele decirse, *depende de los demás*.

Como siempre vamos a tratar de la dependencia *lineal*, a veces, omitiremos el vocablo *lineal*. También diremos que los *vectores son dependientes* o *independientes*, en lugar de decir que el *sistema de vectores es dependiente* o *independiente*.

Un vector a^1 también forma un sistema, que es linealmente independiente si $a^1 \neq 0$, y dependiente si $a^1 = 0$.

Si un sistema de vectores $a^1, \ldots, a^k$ es linealmente independiente, entonces cualquier parte de este sistema $a^1, \ldots, a^s$ $(s < k)$ también es linealmente independiente. En caso contrario, existiría un sistema no trivial de números $\lambda_1, \ldots, \lambda_s$ tales que se cumpliría la igualdad

$$\lambda_1 a^1 + \ldots + \lambda_s a^s = 0,$$

pero, entonces, para el sistema $\lambda_1, \ldots, \lambda_s, \overbrace{0, \ldots, 0}^{k\text{-}s \text{ veces}}$, que tampoco es trivial, se verificaría la igualdad

$$\lambda_1 a^1 + \ldots + \lambda_s a^s + 0 \cdot a^{s+1} + \ldots + 0 \cdot a^k = 0.$$

De todo esto se deduce que, si un sistema de vectores $a^1, \ldots, a^s$ es linealmente dependiente, entonces, cualquier sistema completado

$$a^1, \ldots, a^s, a^{s+1}, \ldots, a^k$$

posee la misma propiedad. En particular, el sistema de vectores que contiene el vector nulo siempre es linealmente dependiente.

Formemos la matriz determinada por el sistema de vectores (1)

$$A = \left\| \begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{matrix} \right\|.$$

TEOREMA 1. *Si* rango $A = k$, *o sea, que el rango de A es igual al número de vectores, entonces el sistema* (1) *es linealmente independiente.*

Si rango $A < k$, *entonces el sistema* (1) *es linealmente dependiente.*

EJEMPLO 1. Dos vectores $a^1 = (a_{11}, a_{12})$, $a^2 = (a_{21}, a_{22})$ en el espacio real R_2 forman un sistema linealmente independiente si el determinante

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0,$$

puesto que la ecuación vectorial

$$\lambda_1 a^1 + \lambda_2 a^2 = 0 \tag{4}$$

es equivalente a dos ecuaciones para las componentes correspondientes

$$\left.\begin{aligned} a_{11}\lambda_1 + a_{21}\lambda_2 = 0, \\ a_{12}\lambda_1 + a_{22}\lambda_2 = 0. \end{aligned}\right\} \tag{5}$$

Pero si $\Delta \neq 0$, entonces el sistema (5) admite únicamente la solución trivial

$$\lambda_1 = \lambda_2 = 0. \tag{6}$$

Si $\Delta = 0$, entonces las ecuaciones (5) tiene solución no trivial (λ_1, λ_2), de modo que si $\Delta = 0$, el sistema de vectores a^1, a^2 es linealmente dependiente.

Evidentemente, decir que en R_2 los vectores a^1 y a^2 son colineales o linealmente dependientes, es lo mismo. Pero, entonces también es lo mismo decir que los vectores a^1 y a^2 no son colineales o que son linealmente independientes.

EJEMPLO 2. Un sistema de vectores $a^1, a^2, \ldots, a^k$ ($k \geqslant 3$) en el espacio real R_2 siempre es linealmente dependiente. Geométricamente esto está claro de la fig. 32: si c es un vector arbitrario y a, b

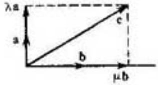

Fig. 32

son vectores no colineales, entonces siempre se pueden señalar unos números λ y μ, tales que

$$c = \lambda a + \mu b.$$

Esto muestra que el sistema a, b, c es linealmente dependiente. Ahora bien, si a y b son vectores colineales, entonces son linealmente dependientes, y por lo tanto, también son linealmente dependientes los vectores a, b, c.

Según el teorema 1, para estudiar un par de vectores a^1, a^2 hay que escribir la matriz de sus coordenadas

$$A=\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}.$$

En el caso considerado, $k = 2$.

a) Si rango $A = 1 < 2 = k$, entonces el teorema afirma que los vectores a^1, a^2 son linealmente dependientes.

b) Si rango $A = 2 = k$, entonces los vectores a^1, a^2 son linealmente independientes.

Esto coincide con las deducciones hechas anteriormente, ya que en este caso, a) $\Delta = 0$, b) $\Delta \neq 0$.

Del teorema también se deduce que tres vectores arbitrarios a^1, a^2, a^3 en R_2 son linealmente dependientes, pues

$$\text{rango } A \leqslant 2 < 3 = k.$$

EJEMPLO 3. En el espacio real tridimensional R_3, dos vectores

$$a^1 = (a_{11}, a_{12}, a_{13}), \quad a^2 = (a_{21}, a_{22}, a_{23})$$

son linealmente dependientes si, y sólo si, son colineales.

En efecto, supongamos que a^1, a^2 son colineales. Si uno de los vectores dados es nulo, entonces estos vectores son linealmente dependientes. Si a^1 y a^2 son colineales y no nulos, entonces

$$a^1 = \lambda a^2,$$

donde λ es un número. Esto último significa que a^1 y a^2 son linealmente dependientes.

Recíprocamente, si a^1, a^2 son linealmente dependientes, entonces uno de ellos depende del otro, por ejemplo,

$$a^2 = \mu a^1,$$

o sea, que los vectores son colineales.

Si consideramos la matriz

$$A=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{Vmatrix},$$

entonces los elementos de las filas son proporcionales, por lo cual,

$$\text{rango } A = 1 < 2 = k,$$

o sea, la afirmación que hemos hecho concuerda con el teorema **1**.

EJEMPLO 4. Consideremos ahora tres vectores en R_3:

$$a^1 = (a_{11}, a_{12}, a_{13}),$$
$$a^2 = (a_{21}, a_{22}, a_{23}),$$
$$a^3 = (a_{31}, a_{32}, a_{33}).$$

Sea

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

La ecuación vectorial

$$\lambda_1 a^1 + \lambda_2 a^2 + \lambda_3 a^3 = 0 \tag{7}$$

es equivalente al sistema de tres ecuaciones

$$\left.\begin{aligned} a_{11}\lambda_1 + a_{21}\lambda_2 + a_{31}\lambda_3 = 0, \\ a_{12}\lambda_1 + a_{22}\lambda_2 + a_{32}\lambda_3 = 0, \\ a_{13}\lambda_1 + a_{23}\lambda_2 + a_{33}\lambda_3 = 0. \end{aligned}\right\} \tag{7'}$$

Si $\Delta \neq 0$, entonces el sistema (7′) tiene únicamente la solución trivial $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Pero, entonces, la ecuación (7) tiene también únicamente la solución trivial, y el sistema de vectores a^1, a^2, a^3 es linealmente independiente.

Si $\Delta = 0$, entonces el sistema (7′) y, por consiguiente, también la ecuación (7) admiten solución no trivial $(\lambda_1, \lambda_2, \lambda_3)$. Pero, entonces, el sistema de vectores (a^1, a^2, a^3) es linealmente dependiente. Pero aquí se puede detallar más:

1) Sea rango $A = 1$, donde

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

Entonces al menos una de las filas de A, supongamos para precisar que la primera, tiene siquiera un elemento no nulo. Consideremos la matriz

$$A'=\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{Vmatrix} \tag{8}$$

Esta es de rango 1, ya que todos los determinantes de segundo orden engendrados por ella son iguales a cero:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = 0.$$

Pero, entonces, evidentemente, las componentes de los vectores a^1 y a^2 son proporcionales

$$\frac{a_{21}}{a_{11}} = \frac{a_{22}}{a_{12}} = \frac{a_{23}}{a_{13}} = \lambda,$$

o sea,

$$a_{21} = a_{11}\lambda, \quad a_{22} = a_{12}\lambda, \quad a_{23} = a_{13}\lambda$$

o bien,

$$a^2 = \lambda a^1.$$

Similarmente, teniendo en cuenta que en la matriz

$$A'' = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

todos los determinantes de segundo orden también son iguales a cero, obtenemos que

$$a^3 = \mu a^1,$$

donde μ es un número. Así pues, en este caso los vectores a^1, a^2, a^3 son colineales.

2) Supongamos ahora que rango $A = 2$. Entonces una de las matrices compuestas por dos filas de la matriz A es de rango 2. Supongamos para precisar que es la matriz A' (véase (8)). En virtud del ejemplo 3, los vectores a^1 y a^2 son linealmente independientes. Pero el sistema a^1, a^2, a^3 es dependiente, o sea, que para cierta terna no trivial del números $(\lambda_1, \lambda_2, \lambda_3)$

$$\lambda_1 a^1 + \lambda_2 a^2 + \lambda_3 a^3 = 0. \tag{9}$$

Aquí $\lambda_3 \neq 0$, pues en caso contrario tendríamos $\lambda_1 a^1 + \lambda_2 a^2 = 0$, y en virtud de la independencia del sistema a^1, a^2, resultaría $\lambda_1 = \lambda_2 = 0$. Pero, entonces, en la igualdad (9) se puede despejar a^3:

$$a^3 = \mu_1 a^1 + \mu_2 a^2, \quad \mu_1 = -\frac{\lambda_1}{\lambda_3}, \qquad \mu_2 = -\frac{\lambda_2}{\lambda_3}.$$

Así pues, si $\triangle = 0$ y rango $A' = 2$ (véase (8)), entonces los vectores a^1 y a^2 son no colineales y el vector a^3 está situado en el plano de estos vectores.

Los razonamientos expuestos no contradicen al teorema 1. En efecto, si $\triangle \neq 0$, entonces rango $A = 3 = k$ y según el teorema 1, el sistema de vectores a^1, a^2, a^3 es linealmente independiente. Ahora bien, si $\triangle = 0$, entonces rango $A < 3$ y el sistema de vectores a^1, a^2, a^3 es linealmente dependiente.

DEMOSTRACION DEL TEOREMA 1. La igualdad vectorial (2) es equivalente a las siguientes n ecuaciones para las componentes de los vectores:

$$\left.\begin{aligned} &\lambda_1 a_{11} + \lambda_2 a_{21} + \ldots + \lambda_k a_{k1} = 0, \\ &\lambda_1 a_{12} + \lambda_2 a_{22} + \ldots + \lambda_k a_{k2} = 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\lambda_1 a_{1n} + \lambda_2 a_{2n} + \ldots + \lambda_k a_{kn} = 0. \end{aligned}\right\} \tag{2'}$$

Sea rango $A = k$. Entonces, evidentemente, $k \leqslant n$. Existe un determinante no nulo de k-ésimo orden engendrado por la matriz A. Sin restringir generalidad, se puede suponer que éste es el determinante de los coeficientes de las primeras k ecuaciones del sistema (2'):

$$\left.\begin{array}{l} \lambda_1 a_{11} + \ldots + \lambda_k a_{k1} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1 a_{1k} + \ldots + \lambda_k a_{kk} = 0, \end{array}\right\} \tag{10}$$

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

Como el determinante del sistema homogéneo (10) es distinto de cero, éste sólo admite solución trivial

$$\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0,$$

y el sistema de vectores (1) es linealmente independiente. En este caso, el teorema queda demostrado.

Supongamos ahora que rango $A = s < k$. Reordenemos la numeración de las variables λ_j y de los coeficientes a_{ij} del sistema (2') de tal modo que sea

$$\begin{vmatrix} a_{11} & \ldots & a_{s1} \\ \ldots & \ldots & \ldots \\ a_{1s} & \ldots & a_{ss} \end{vmatrix} \neq 0.$$

Entonces, el sistema de las primeras s ecuaciones puede escribirse así:

$$\left.\begin{array}{l} a_{11}\lambda_1 + \ldots + a_{s1}\lambda_s = -a_{s+1,\,1}\lambda_{s+1} - \ldots - a_{k1}\lambda_k, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1s}\lambda_1 + \ldots + a_{ss}\lambda_s = -a_{s+1,\,s}\lambda_{s+1} - \ldots - a_{ks}\lambda_k. \end{array}\right\}$$

Haciendo $\lambda_{s+1} = \ldots = \lambda_k = 1$, en el sistema dado se pueden despejar $\lambda_1, \ldots, \lambda_s$. Así obtenemos una solución no trivial de las primeras s ecuaciones (2'):

$$\lambda_1, \ldots, \lambda_k \tag{11}$$

Ahora podemos añadir a estas s ecuaciones cualesquiera otras $k - s$ ecuaciones de (2') y el sistema obtenido también tendrá como solución el sistema no trivial hallado. Para explicar esta afirmación, escribamos formalmente el sistema (2') del modo siguiente:

$$\left.\begin{array}{l} \lambda_1 a_{11} + \ldots + \lambda_k a_{k1} + \lambda_{k+1}\cdot 0 + \ldots + \lambda_n \cdot 0 = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1 a_{1n} + \ldots + \lambda_k a_{kn} + \lambda_{k+1}\cdot 0 + \ldots + \lambda_n \cdot 0 = 0. \end{array}\right\} \tag{2'}$$

La matriz de los coeficientes de las ecuaciones obtenidas también es de rango s, así como la matriz ampliada (¡con ceros a la derecha!). Las primeras s ecuaciones del sistema (2') se satisfacen por los números hallados $\lambda_1, \ldots, \lambda_k$ (véase (11)) y por cualesquiera números $\lambda_{k+1}, \ldots, \lambda_n$. En virtud de la afirmación 2) § 4 (regla para la resolución de un sistema), los números $\lambda_1, \ldots, \lambda_n$ también satisfacen las demás ecuaciones del sistema (2′), o sea, los números $\lambda_1, \ldots, \lambda_k$ (no todos iguales a cero) satisfacen las demás ecuaciones del sistema (2).

Así pues, los vectores $a^1, \ldots, a^k$ son linealmente dependientes y el teorema queda demostrado también en este caso.

§ 15. Operadores lineales

Tomemos una matriz cuadrada arbitraria

$$A = \| a_{kl} \| = \begin{Vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{Vmatrix}. \tag{1}$$

La matriz A se puede considerar como un *operador* que pone en correspondencia a cada vector $x = (x_1, \ldots, x_n) \in R_n$ un vector $y = (y_1, \ldots, y_n) \in R_n$ cuyas componentes se calculan por las fórmulas

$$\left.\begin{array}{l} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = a_{n1}x_1 + \ldots + a_{nn}x_n, \end{array}\right\} \tag{2}$$

o abreviadamente,

$$y_i = \sum_{j=1}^{n} a_{ij}x_j \quad (i = 1, \ldots, n), \tag{2′}$$

o sea, la i-ésima coordenada del vector y se expresa mediante la i-ésima fila de la matriz A. Este operador lo escribiremos abreviadamente así:

$$y = Ax \quad (x, y \in R_n). \tag{2″}$$

En particular, en el caso del espacio unidimensional R_1, los vectores x e y son números, y el operador (2″) se convierte en la función

$$y = a_{11}x.$$

Nota 1. Si a_{kl} y, más adelante, b_{kl}, son números complejos, entonces hay que considerar que R_n es el espacio complejo. Pero si a_{kl}, b_{kl} son números reales, entonces R_n puede ser un espacio real o complejo.

El *operador* (2″) *es lineal.* Esto significa que cumple la propiedad

$$A\ (\alpha x + \beta x') = \alpha Ax + \beta Ax'$$

para cualesquiera vectores x, $x' \in R_n$ y cualesquiera números α, β. En efecto,

$$\sum_{j=1} a_{ij}(\alpha x_j+\beta x_j') = \alpha \sum_{j=1}^{n} a_{ij}x_j + \beta \sum_{j=1}^{n} a_{ij}x_j' \quad (i=1, \ldots, n).$$

Si las matrices $A = \| a_{kl} \|$ y $B = \| b_{kl} \|$ son iguales, o sea, que sus elementos correspondientes son iguales, $a_{kl} = b_{kl}$, entonces éstas determinan operadores idénticamente iguales:

$$Ax = Bx, \ \forall x \in R_n. \tag{3}$$

Recíprocamente, la igualdad (3) implica que

$$a_{kl} = b_{kl} \quad (k, \ l = 1, \ldots, n),$$

o sea, la igualdad de las matrices A y B ($A = B$). Esto se comprueba fácilmente haciendo en (3)

$$x = x^l = (0, \ldots, 0, 1, 0, \ldots, 0),$$

donde en el l-ésimo lugar está la unidad ($l = 1, \ldots, n$).

Por lo tanto, a distintas matrices A les corresponden distintos operadores; si dos matrices A_1 y A_2 se diferencian al menos en un elemento, entonces, necesariamente existe un vector x tal que

$$A_1x \neq A_2x.$$

Supongamos que, además de A, se ha dado otro operador B, determinado por una matriz cuadrada de n-ésimo orden

$$B = \| b_{kl} \|.$$

A cada vector $x \in R_n$ le corresponde mediante el operador A un vector $y \in R_n$, al cual, mediante el operador B, le corresponde un vector z cuyas componentes se calculan por las fórmulas

$$z_k = \sum_{l=1}^{n} b_{kl}y_l \quad (k = 1, \ldots, n).$$

Como resultado, obtenemos un operador lineal compuesto

$$z = BAx \quad (x \in R_n), \tag{4}$$

donde

$$z_k = \sum_{l=1}^{n} b_{kl}y_l = \sum_{l=1}^{n} b_{kl} \sum_{j=1}^{n} a_{lj}x_j = \sum_{j=1}^{n} \Big(\sum_{l=1}^{n} b_{kl}a_{lj} \Big) x_j = \sum_{j=1}^{n} \gamma_{kj}x_j$$

con la matriz $\| \gamma_{kj} \|$, denominada *producto de las matrices* B y A y que se denota por

$$BA = \| \gamma_{kj} \|, \tag{5}$$

siendo

$$\gamma_{kj}=\sum_{i=1}^{n} b_{ki}a_{ij} \quad (k,\ j=1,\ \ldots,\ n). \tag{6}$$

Por lo tanto, para obtener el elemento γ_{kj} de la matriz BA (perteneciente a su k-ésima fila y j-ésima columna), hay que multiplicar los elementos de la k-ésima fila de la matriz B por los elementos correspondientes de la j-ésima columna de la matriz A y sumar los resultados.

El determinante de la matriz BA es igual al producto de los determinantes de las matrices B y A:

$$|BA| = |B|\cdot|A|. \tag{7}$$

Esta propiedad se deduce de la fórmula para el producto de determinantes (véase § 2, propiedad j)).

Supongamos que el determinante de la matriz del operador A (véase (1)) no es igual a cero:

$$\triangle = |a_{kl}| \neq 0.$$

En este caso (véase § 4, teorema 1), el sistema de ecuaciones (2) o, lo que es lo mismo, la ecuación operacional $y = Ax$ admite solución única $x \in R_n$ para cualquier $y \in R_n$. Además, las fórmulas por las que se halla x para un $y \in R_n$ dado, tienen la forma

$$x_j=\sum_{s=1}^{n} b_{js}y_s \quad (j=1,\ \ldots,\ n). \tag{8}$$

Aquí

$$b_{js} = A_{sj}/\triangle \quad (s,\ j = 1,\ \ldots,\ n) \tag{9}$$

(véase § 4, (3′)), donde A_{sj} es el complemento algebraico del elemento a_{sj} en el determinante $\triangle$.

Por cierto, ahora sólo es importante señalar que los números b_{js} son los elementos de una matriz

$$B = \| b_{js} \|,$$

que posee las siguientes propiedades interesantes:

$$BAx=x, \quad \forall x \in R_n, \tag{10}$$

$$ABy=y, \quad \forall y \in R_n. \tag{11}$$

En efecto, un vector arbitrario $x \in R_n$ se transforma mediante el operador A en un vector y, que a su vez se transforma mediante el operador B nuevamente en x. Por otra parte, a cada vector $y \in R_n$ le corresponde mediante el operador B (véase (8)) un vector x tal que $Ax = y$.

Evidentemente, en la igualdad (11) se puede sustituir y por cualquier otra letra, por lo que hemos obtenido la identidad

$$BAx = ABx = x, \ \forall x.$$

El operador $x = Ex$, $\forall\, x \in R_n$, se llama *operador unidad*. La matriz correspondiente tiene la forma

$$E=\begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{vmatrix}$$

y se llama *matriz unidad*. Hemos demostrado que

$$AB = BA = E.$$

El operador B que posee esta propiedad se llama *inverso* del operador A y se denota por A^{-1}. Respectivamente, su matriz se llama *matriz inversa* de la matriz A y también se denota por A^{-1}. Los elementos de la matriz A^{-1} se hallan mediante los elementos de la matriz A por las fórmulas (9).

Hemos demostrado que, *si el determinante* $|A|$ *de una matriz cuadrada* A *es distinto de cero, entonces existe la matriz inversa* A^{-1}. Por lo tanto, para A^{-1} se cumple la propiedad

$$A^{-1}A = AA^{-1} = E.$$

Si el determinante de la matriz A es igual a cero ($|A| = 0$), entonces ésta no tiene matriz inversa. Es suficiente señalar que, en este caso la ecuación $y = Ax$ no tiene solución para cualquier y. Sin embargo, la propiedad $AA^{-1}y = y$, si se cumple, afirma que *a cada* $y \in R_n$ le corresponde (mediante el operador A^{-1}) un x tal que es la solución de la ecuación $y = Ax$.

Nota. La operación de multiplicación de matrices se puede generalizar también a las matrices no cuadradas B y A, siempre que el número de columnas de la matriz B sea igual al número de filas de la matriz A. Entonces la multiplicación de las matrices se efectúa según unas fórmulas semejantes a (6). Por ejemplo, si

$$B=\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A=\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

entonces

$$BA=\begin{pmatrix} 4 & 3 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

En este caso, el producto AB carece de sentido, ya que la matriz A tiene dos columnas y la matriz B tiene tres filas.

Para las matrices cuadradas A y B tienen sentido los productos AB y BA, pero no siempre AB es igual a BA. Por ejemplo, si

$$A=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix},\quad B=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},$$

entonces

$$AB=\begin{pmatrix}0 & 1\\ 0 & 1\end{pmatrix},\quad BA=\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}.$$

o sea, $AB \neq BA$.

Fácilmente se comprueba que

$$(AB)\,C = A\,(BC).$$

Si A es un operador lineal, entonces, la expresión Ax se puede considerar como el producto de la matriz A por la matriz de una columna

$$x=\begin{pmatrix}x_1\\ \cdots\\ x_n\end{pmatrix}.$$

Sean dados los operadores lineales A y B. Se llama *suma* de estos operadores al operador $A+B$ determinado por la igualdad

$$(A+B)\,x = Ax + Bx,\quad \forall\, x\in R_n.$$

Está claro que la matriz del operador $A+B$ coincide con la matriz que es igual a la suma de las matrices de los operadores A y B.

Fácilmente se comprueba que

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

EJEMPLO 1. Hallar la matriz inversa de la matriz

$$A=\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 1\\ 1 & 1 & 0\end{pmatrix}.$$

La matriz A determina un operador lineal $y = Ax$, que pone en correspondencia a cada vector $x = (x_1, x_2, x_3)$ un vector $y = (y_1, y_2, y_3)$ mediante las igualdades

$$\begin{aligned} y_1 &= x_1 + 2x_2,\\ y_2 &= x_2 + x_3,\\ y_3 &= x_1 + x_2. \end{aligned}$$

Estas igualdades se pueden considerar también como un sistema de tres ecuaciones lineales en las incógnitas x_1, x_2, x_3. El determinante de este sistema es distinto de cero. Pero, entonces, este sistema admite solución para cualesquiera valores de y_1, y_2, y_3. Como resultado, obtenemos las igualdades:

$$\begin{aligned} x_1 &= -y_1 + 2y_3, \\ x_2 &= y_1 - y_3, \\ x_3 &= -y_1 + y_2 + y_3, \end{aligned}$$

las cuales determinan un operador $x = A^{-1}y$, inverso al operador A. La matriz de este operador es

$$A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Esta es la matriz inversa de la matriz A.

Los elementos de la matriz A^{-1} se pueden obtener mediante cálculos aplicando las fórmulas (9). Hagamos estos cálculos.

Denotemos los elementos de la matriz inversa A^{-1} mediante b_{js}. Se tiene $\triangle = 1$, $b_{js} = A_{sj}$,

$$A_{11} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1, \quad A_{12} = -\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1,$$

$$A_{13} = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} = -1,$$

$$A_{21} = -\begin{vmatrix} 2 & 0 \\ 1 & 0 \end{vmatrix} = 0, \quad A_{22} = \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} = 0,$$

$$A_{23} = -\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = 1,$$

$$A_{31} = \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 2, \quad A_{32} = -\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1,$$

$$A_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1.$$

Por lo tanto,

$$\begin{aligned} &b_{11} = A_{11} = -1, \quad && b_{12} = A_{21} = 0, \quad && b_{13} = A_{31} = 2, \\ &b_{21} = A_{12} = 1, && b_{22} = A_{22} = 0, && b_{23} = A_{32} = -1, \\ &b_{31} = A_{13} = -1, && b_{32} = A_{23} = 1, && b_{33} = A_{33} = 1. \end{aligned}$$

En resumen,

$$A^{-1}=\begin{pmatrix}-1 & 0 & 2\\ 1 & 0 & -1\\ -1 & 1 & 1\end{pmatrix}.$$

EJEMPLO 2. Calcular el producto de las matrices BA, donde

$$A=\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 2\\ 1 & 1 & 1\end{pmatrix},\quad B=\begin{pmatrix}2 & 0 & 0\\ 0 & 3 & 2\\ 2 & 1 & -1\end{pmatrix}.$$

El cálculo se puede realizar según las fórmulas (5), (6), pero se puede razonar también del modo siguiente.

La matriz A determina un operador $y = Ax$, que pone en correspondencia a los vectores $x = (x_1, x_2, x_3)$ los vectores $y = (y_1, y_2, y_3)$ mediante las igualdades

$$\begin{aligned} y_1 &= x_1 + x_3,\\ y_2 &= x_2 + 2x_3,\\ y_3 &= x_1 + x_2 + x_3. \end{aligned}$$

Por otra parte, la matriz B determina un operador $z = By$, que pone en correspondencia a los vectores $y = (y_1, y_2, y_3)$ los vectores $z = (z_1, z_2, z_3)$ mediante las igualdades

$$\begin{aligned} z_1 &= 2y_1,\\ z_2 &= 2y_2 + 2y_3,\\ z_3 &= 2y_1 + y_2 - y_3. \end{aligned}$$

Pero, entonces, el operador

$$z = BAx$$

se determina por las igualdades

$$\begin{aligned} z_1 &= 2\,(x_1 + x_3) &&= 2x_1 \qquad +2x_3,\\ z_2 &= 3\,(x_2 + 2x_3) + 2\,(x_1 + x_2 + x_3) &&= 2x_1 + 5x_2 + 8x_3,\\ z_3 &= 2\,(x_1 + x_3) + (x_2 + 2x_3) - (x_1 + x_2 + x_3) &&= \ x_1 \qquad +3x_3. \end{aligned}$$

Por consiguiente, el producto BA de las matrices B y A es la matriz

$$BA=\begin{pmatrix}2 & 0 & 2\\ 2 & 5 & 8\\ 1 & 0 & 3\end{pmatrix}.$$

§ 16. Bases en R_n

Introduzcamos en el espacio R_n (real o complejo) n vectores

$$\left.\begin{array}{l} i^1=(1,\ 0,\ \ldots,\ 0), \\ i^2=(0,\ 1,\ 0,\ \ldots,\ 0), \\ \cdots\cdots\cdots\cdots \\ i^n=(0,\ 0,\ \ldots,\ 0,\ 1), \end{array}\right\} \qquad (1)$$

denominados *vectores unitarios coordenados (versores) del espacio* R_n.

Se llama *eje* x_k del espacio R_n al conjunto de puntos de la forma $(0, \ldots, 0, x_k, 0, \ldots, 0)$, donde x_k ocupa el k-ésimo lugar y recorre todos los valores reales; el vector i^k se llama *vector unitario* (o *versor*) *del eje* x_k.

Si $a = (x_1, \ldots, x_n)$ es un vector arbitrario (real en el espacio real R_n o complejo en el espacio complejo R_n), entonces, evidentemente, éste puede expresarse en forma de una combinación lineal de los vectores (1) de la forma siguiente:

$$a = x_1 i^1 + x_2 i^2 + \ldots + x_n i^n. \qquad (2)$$

Como la igualdad $a = (x_1, \ldots, x_n) = 0$ implica que $x_1 = \ldots \ldots = x_n = 0$, el sistema $(i^1, \ldots, i^n)$ es linealmente independiente.

Tomemos un sistema arbitrario formado por n vectores linealmente independientes

$$\left.\begin{array}{l} a^1=(a_{11},\ \ldots,\ a_{1n}), \\ \cdots\cdots\cdots\cdots \\ a^n=(a_{n1},\ \ldots,\ a_{nn}). \end{array}\right\} \qquad (3)$$

Como ya sabemos (véase § 14, teorema 1), el sistema (3) es linealmente independiente si el determinante

$$\triangle = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \neq 0. \qquad (4)$$

Si $\triangle = 0$, el sistema (3) es linealmente dependiente.

Según el teorema 1 del § 14, cualesquiera $n + 1$ vectores en el espacio R_n son linealmente dependientes, puesto que el rango de la matriz formada por las componentes de estos vectores no es superior a n. Por esto, si $a = (x_1, \ldots, x_n)$ es un vector arbitrario y el sistema de vectores (3) es linealmente independiente ($\triangle \neq 0$), entonces el sistema de vectores $a^1, \ldots, a^n, a$ es linealmente dependiente, o sea, existen unos números $\lambda_1, \ldots, \lambda_n, \lambda_{n+1}$, no simultáneamente iguales a cero, tales que

$$\lambda_1 a^1 + \ldots + \lambda_n a^n + \lambda_{n+1} a = 0,$$

donde $\lambda_{n+1} \neq 0$ (en caso contrario, el sistema (3) sería linealmente dependiente). De aquí

$$a = \sum_{k=1}^{n} x'_k a^k, \tag{5}$$

donde $x'_k = \frac{-\lambda_k}{\lambda_{n+1}}$ $(k=1, \ldots, n)$ son unos números. Expresemos la suma (5) mediante los vectores unitarios coordenados i^k (véase (2)):

$$a = \sum_{k=1}^{n} x'_k \left(\sum_{l=1}^{n} a_{kl} i^l\right) = \sum_{l=1}^{n} \left(\sum_{k=1}^{n} a_{kl} x'_k\right) i^l.$$

Por otra parte, según (2),

$$a = \sum_{l=1}^{n} x_l i^l.$$

En virtud de la independencia lineal del sistema $i^1, \ldots, i^n$, los coeficientes de vectores iguales i^l tienen que ser iguales:

$$x_l = \sum_{k=1}^{n} a_{kl} x'_k \quad (l = 1, \ldots, n). \tag{6}$$

Por lo tanto, si se conocen las componentes x_l del vector a en el sistema $i^1, \ldots, i^n$, entonces las componentes x'_k de este vector en el sistema $a^1, \ldots, a^n$ se hallan por las fórmulas (6) y, además, unívocamente, ya que el determinante del sistema (6) es distinto de cero, $\Delta \neq 0$.

Hemos demostrado que *cualquiera que sea el sistema linealmente independiente de vectores $a^1, \ldots, a^n$, cualquier vector $a \in R_n$ se puede desarrollar por este sistema, o sea, se puede expresar en forma de una suma* (5), *donde $x'_1, \ldots, x'_n$ son unos números, que se determinan en* (6) *y, además, unívocamente.*

En este sentido el sistema de vectores $a^1, \ldots, a^n$ se llama *base* en R_n. Esto significa que cualquier vector $a \in R_n$ se puede expresar en forma de una combinación lineal (5) de estos vectores y, además, unívocamente. Hemos demostrado que *cualquier sistema linealmente independiente de n vectores de R_n es una base en R_n.*

Un sistema linealmente independiente formado por m vectores

$$a^1 = (a_{11}, \ldots, a_{1m}, a_{1,m+1}, \ldots, a_{1n}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a^m = (a_{m1}, \ldots, a_{mm}, a_{m,m+1}, \ldots, a_{mn}),$$

donde $m < n$, no es una base en R_n. En efecto, el rango de la matriz de las componentes de estos vectores es igual a m. Supongamos que las primeras m columnas de esta matriz forman un determinante distinto de cero. Ampliemos esta matriz, añadiendo abajo la fila

$$a^{m+1} = (0, \ldots, 0, 1, 0, \ldots, 0),$$

donde el 1 ocupa el $(m+1)$-ésimo lugar. La matriz ampliada es de rango $m+1$ y, por consiguiente, el sistema de vectores $a^1, \ldots, a^m, a^{m+1}$ es linealmente independiente. Pero, entonces, el vector a^{m+1} no puede ser una combinación lineal de los vectores del sistema $a^1, \ldots, a^m$, y este sistema no es una base en R_n.

Denotemos por

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix}$$

la matriz de los vectores $(a^1, \ldots, a^n)$[1]).

El cambio de la base $(i^1, \ldots, i^n)$ a la base $(a^1, \ldots, a^n)$ se realiza mediante la matriz A:

$$a^k = \sum_{s=1}^{n} a_{ks} i^s \quad (k = 1, \ldots, n), \tag{7}$$

o sea, que el vector a^k se expresa mediante los vectores i^s aplicando la k-ésima fila de la matriz A. El cambio inverso de $(a^1, \ldots, a^n)$ a $(i^1, \ldots, i^n)$ se realiza mediante la matriz inversa A^{-1} (véase § 15, (9))

$$i^s = \sum_{l=1}^{n} b_{sl} a^l \quad (s = 1, \ldots, n), \tag{8}$$

cuyos elementos se calculan según las fórmulas $b_{sl} = \frac{A_{ls}}{\triangle}$, donde A_{ls} es el adjunto del elemento a_{ls} en el determinante $\triangle$ (señalemos que el elemento b_{sl}, perteneciente a la s-ésima fila y a la l-ésima columna, se expresa mediante el adjunto A_{ls} del elemento a_{ls}, perteneciente a la l-ésima fila y s-ésima columna). Señalemos también que

$$a = \sum_{l=1}^{n} x'_l a^l = \sum_{s=1}^{n} x_s i^s = \sum_{s=1}^{n} x_s \sum_{l=1}^{n} b_{sl} a^l = \sum_{l=1}^{n} \left(\sum_{s=1}^{n} b_{sl} x_s \right) a^l,$$

de donde

$$x'_l = \sum_{s=1}^{n} b_{sl} x_s, \tag{9}$$

o sea, el cambio de las coordenadas $(x_1, \ldots, x_n)$ a $(x'_1, \ldots, x'_n)$ se realiza mediante la matriz (véase § 3, (2))

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

---

[1]) Aunque el texto que se expone a continuación se presenta con letra pequeña, recomendamos leerlo. En el § 17 se cita (9).

En (9) se observa que x_l' se expresa mediante $x_1, \ldots, x_n$ aplicando la l-ésima columna de la matriz A^{-1} o la l-ésima fila de la matriz $(A^{-1})'$, que es la traspuesta a la matriz A^{-1}.

Por otra parte, en la fórmula (6)

$$x_s = \sum_{l=1}^{n} a_{ls} x_l' \quad (s = 1, \ldots, n)$$

se observa que el cambio de $(x_1', \ldots, x_n')$ a $(x_1, \ldots, x_n)$ se realiza mediante la matriz A', que es la traspuesta de A, de modo que x_s se expresa mediante $x_1', \ldots, x_n'$ aplicando la s-ésima fila de la matriz A' o la s-ésima columna de la matriz A.

Nota. En el § 14 se había establecido que una matriz cuadrada arbitraria

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \tag{10}$$

determina un operador lineal $y = Ax$ $(x \in R_n, y \in R_n)$, definido según las fórmulas

$$y_k = \sum_{j=1}^{n} a_{kj} x_j \quad (k = 1, \ldots, n). \tag{11}$$

Pero también se verifica lo recíproco: cualquiera que sea el operador lineal $y = Ax$ $(x \in R_n, y \in R_n)$, éste se determina por una matriz (10) de tal modo que el vector $y = Ax$ se calcula según el vector $y = Ax$ por las fórmulas (11).

En efecto, sea dado un operador lineal arbitrario $y = Ax$ $(x \in R_n, y \in R_n)$. Denotemos las imágenes de los versores i^s mediante el operador A del modo siguiente:

$$A(i^s) = (a_{1s}, a_{2s}, \ldots, a_{ns}) = \sum_{k=1}^{n} a_{ks} i^k$$

$$(s = 1, \ldots, n).$$

Entonces, como A es lineal, un vector arbitrario

$$x = x_1 i^1 + \ldots + x_n i^n = \sum_{s=1}^{n} x_s i^s$$

se transforma mediante A en un vector y, determinado por las igualdades

$$y = Ax = A\Big(\sum_{s=1}^{n} x_s i^s\Big) = \sum_{s=1}^{n} x_s A(i^s) = \sum_{s=1}^{n} x_s \sum_{k=1}^{n} a_{ks} i^k = \sum_{k=1}^{n} \Big(\sum_{s=1}^{n} a_{ks} x_s\Big) i^k,$$

de donde se deduce que la k-ésima componente de y se determina por la fórmula (11). Por lo tanto, el operador A engendra una matriz (10) tal que las columnas constan de las coordenadas de las imágenes de los vectores de la base (versores) mediante el operador A.

EJEMPLO 1. Hallar la matriz del operador lineal (de la transformación) A que consiste en una rotación de los vectores del plano R_2, que parten del origen, un ángulo α $(0 < \alpha < \pi/2)$ en sentido contrario al del movimiento de las agujas del reloj.

Consideremos la base formada por los vectores $i^1 = (1, 0)$, $i^2 = (0, 1)$. Entonces, evidentemente, (fig. 33),

$$A\ (i^1) = (\cos\alpha,\ \operatorname{sen}\alpha),$$
$$A\ (i^2) = (-\operatorname{sen}\alpha,\ \cos\alpha).$$

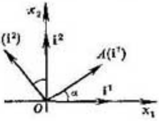

Fig. 33

Por lo tanto, la matriz del operador en cuestión tiene la forma

$$A = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}.$$

§ 17. Bases ortogonales en R_n

Se dice que dos vectores no nulos $x,\ y \in R_n$, tienen *la misma dirección*, si existe un número *positivo* λ tal que $x = \lambda y$.

Un vector arbitrario no nulo $x \in R_n$ se puede *normalizar* sustituyéndolo por el vector unitario

$$y = \frac{1}{|x|}\, x \quad (|y| = 1),$$

que tiene la misma dirección que el vector x.

Un vector unitario (cuya norma (o longitud) es igual a 1) se llama *normal*.

Dos vectores x e y en el espacio R_n se llaman *ortogonales*, si su producto escalar es igual a cero: $(x,\ y) = 0$. Aquí R_n puede ser real o complejo. En el caso de R_n complejo, el producto escalar se define como en el § 6, (5′).

Un sistema de vectores

$$x^1,\ \ldots,\ x^\nu \in R_n \tag{1}$$

se llama *ortogonal*, si cualesquiera dos de sus vectores son ortogonales. Un sistema de vectores (1) se llama *ortogonal* y *normal* o también *ortonormal*, si

$$(x^k,\ x^l) = \delta_{kl} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases} \quad (k,\ l = 1,\ \ldots,\ \nu),$$

o sea, si todos los vectores del sistema son normales y ortogonales dos a dos. Si un sistema de vectores (1) es ortogonal y ninguno de ellos es nulo, entonces normalizándolos, obtenemos, evidentemente, un sistema ortonormal. *Un sistema ortonormal* (1) *es linealmente independiente.* En efecto, supongamos que

$$\lambda_1 x^1 + \lambda_2 x^2 + \ldots + \lambda_\nu x^\nu = 0,$$

donde $\lambda_1, \ldots, \lambda_\nu$ son unos números. Multiplicando escalarmente esta igualdad por x^s, obtenemos

$$\lambda_s (x^s, x^s) = \lambda_s = 0 \quad (s = 1, \ldots, \nu),$$

lo que muestra que el sistema (1) es linealmente independiente. Entonces, el sistema ortonormal formado por n vectores de R_n es una base y, por consiguiente, todo vector $a \in R_n$ se puede expresar como una combinación lineal

$$a = \sum_{k=1}^{n} \lambda_k x^k. \tag{2}$$

Multiplicando escalarmente esta igualdad por x^s, obtenemos

$$(a, x^s) = \lambda_s \quad (s = 1, \ldots, n)$$

y, por consiguiente,

$$a = \sum_{k=1}^{n} (a, x^k)\, x^k, \quad \forall a \in R_n.$$

El número (a, x^k) $(|x|^k = 1)$ se llama *proyección del vector* a sobre la dirección del vector x^k. En el espacio real R_3, la magnitud (a, x^k) es la proyección numérica ordinaria del vector a sobre la dirección del vector x^k.

TEOREMA 1. *Un sistema ortonormal de vectores*

$$x^1, x^2, \ldots, x^\nu \quad (\nu < n)$$

se puede completar hasta una base ortonormal en R_n. *En otras palabras, siempre se pueden señalar unos vectores* $x^{\nu+1}, \ldots, x^n$ *tales que el sistema*

$$x^1, \ldots, x^\nu, x^{\nu+1}, \ldots, x^n \tag{3}$$

sea ortonormal y, por consiguiente, represente una base en R_n.

DEMOSTRACIÓN. Como $\nu < n$, existe en R_n un vector a que no depende linealmente de $x^1, \ldots, x^\nu$. Pero, entonces,

$$a = \sum_{k=1}^{\nu} (a, x^k)\, x^k + y,$$

donde $y \neq 0$. El vector y es ortogonal a todos los vectores $x^1, \ldots, x^\nu$. En efecto,

$$(y, x^s) = \left(a - \sum_{k=1}^{\nu} (a, x^k) x^k, x^s\right) =$$
$$= (a, x^s) - (a, x^s) = 0 \quad (s = 1, \ldots, \nu). \tag{4}$$

Normalizando el vector y, obtenemos el vector

$$x^{\nu+1} = \frac{1}{|y|} y \quad (|x^{\nu+1}| = 1),$$

y el sistema

$$x^1, \ldots, x^\nu, x^{\nu+1}$$

es ortonormal. Si $\nu + 1 = n$, obtenemos una base en R_n. En caso contrario, continuamos este proceso. En la $(n - \nu)$-ésima etapa obtendremos una base (3) en R_n.

El sistema de versores en R_n

$$\begin{aligned} i^1 &= (1, 0, 0, \ldots, 0, 0), \\ i^2 &= (0, 1, 0, \ldots, 0, 0), \\ &\ldots\ldots\ldots\ldots\ldots \\ i^n &= (0, 0, 0, \ldots, 0, 1) \end{aligned}$$

puede servir de ejemplo de base ortonormal en R_n.

Un vector arbitrario $a = (x_1, \ldots, x_n) \in R_n$ admite un desarrollo por los vectores unitarios coordenados de la forma

$$a = x_1 i^1 + \ldots + x_n i^n = \sum_{k=1}^{n} x_k i^k, \tag{5}$$

donde $x_k = (a, i^k)$ $(k = 1, \ldots, n)$ es la proyección del vector a sobre la dirección del vector coordenado unitario i^k.

Sea dado un sistema ortonormal determinado de n vectores

$$\left.\begin{aligned} a^1 &= (a_{11}, \ldots, a_{1n}), \\ &\ldots\ldots\ldots\ldots \\ a^n &= (a_{n1}, \ldots, a_{nn}), \end{aligned}\right\} \tag{6}$$

o bien,

$$a^k = \sum_{s=1}^{n} a_{ks} i^s \quad (k = 1, \ldots, n) \tag{7}$$

El cambio de los vectores $(i^1, \ldots, i^n)$ a los vectores $(a^1, \ldots, a^n)$ se realiza en este caso mediante la matriz

$$\Lambda = \left\| \begin{array}{ccc} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot \cdot \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{array} \right\|, \tag{8}$$

o sea, el vector a^k se expresa mediante $i^1, \ldots, i^n$ aplicando la k-ésima fila de la matriz Λ.

A continuación, consideraremos que el espacio R_n y la matriz Λ son reales (véase más adelante la nota 2).

Esta matriz es ortogonal, o sea, posee la propiedad siguiente:

$$\sum_{s=1}^{n} a_{ks} a_{ls} = \delta_{kl} \quad (k, l = 1, \ldots, n). \tag{9}$$

En efecto, como en el caso dado el sistema $a^1, \ldots, a^n$ es ortonormal, se tiene:

$$\delta_{kl} = (a^k, a^l) = \left(\sum_{s=1}^{n} a_{ks} i^s, \sum_{r=1}^{n} a_{lr} i^r\right) =$$
$$= \sum_{s=1}^{n} \sum_{r=1}^{n} a_{ks} a_{lr} (i^s, i^r) = \sum_{s=1}^{n} \sum_{r=1}^{n} a_{ks} a_{lr} \delta_{sr} = \sum_{s=1}^{n} a_{ks} a_{ls}. \tag{10}$$

Vemos que, recíprocamente, la ortogonalidad de la matriz (8) implica la ortonormalidad del sistema de vectores $a^1, \ldots, a^n$, determinados por las fórmulas (7).

Esto muestra que las *fórmulas* (7), *donde* $\| a_{ks} \|$ son *matrices ortogonales arbitrarias, determinan todas las bases ortonormales posibles en* R_n.

Multipliquemos escalarmente el vector i^l por el vector a^k:

$$(i^l, a^k) = a_{kl}.$$

De aquí, obtenemos

$$i^l = \sum_{k=1}^{n} (i^l, a^k)\, a^k = \sum_{k=1}^{n} a_{kl} a^k. \tag{11}$$

Por lo tanto, el cambio de la base $(a^1, \ldots, a^n)$ a la base $(i^1, \ldots, i^n)$ se realiza mediante la matriz Λ', que es traspuesta a la matriz Λ. Como las transformaciones (11) son inversas a las transformaciones (7) (véase § 15), queda demostrado también que una matriz ortogonal Λ tiene la interesantísima propiedad siguiente:

$$\Lambda^{-1} = \Lambda'.$$

De (11) se deduce que

$$\delta_{kl}=(i^k,\ i^l)=\left(\sum_{s=1}^{n} a_{sk}a^s,\ \sum_{r=1}^{n} a_{rl}a^r\right)=$$
$$=\sum_{s=1}^{n}\sum_{r=1}^{n} a_{sk}a_{rl}\,(a^s,\ a^r)=\sum_{s=1}^{n} a_{sk}a_{sl}. \qquad (12)$$

Habíamos definido la matriz ortogonal como una matriz tal, que sus filas (los vectores que éstas representan) son normales y las distintas filas son ortogonales. Como se observa en (12), de esta definición inmediatamente se deduce que en una matriz ortogonal las columnas también son normales y que las distintas columnas son ortogonales.

El cambio de $(x_1, \ldots, x_n)$ a $(x_1', \ldots, x_n')$ se realiza mediante la matriz (véase (9) § 16, $x' = \Lambda x$)

$$(\Lambda')^{-1} = \Lambda'' = \Lambda,$$

o sea, (considerando que $a=\sum_{l=1}^{n} x_l'a^l$)

$$x_l'=\sum_{s=1}^{n} a_{ls}x_s, \qquad (13)$$

mientras que el cambio de $(x_1', \ldots, x_n')$ a $(x_1, \ldots, x_n)$ se realiza (véase (6) § 16) mediante la matriz Λ' que es traspuesta a Λ, o sea,

$$x_s=\sum_{l=1}^{n} a_{ls}x_l'.$$

Señalemos que el determinante de una matriz ortogonal arbitraria Λ (véase (6)) es en valor absoluto igual a 1: $|\,|\Lambda|\,| = |\,|a_{kl}|\,| = 1$.

Esto se debe a que

$$|\Lambda|^2=|a_{kl}|\cdot|a_{ij}|=\left|\sum_{s=1}^{n} a_{ks}a_{ls}\right|=\begin{vmatrix} 1 & 0 & 0 & \ldots & 0\\ 0 & 1 & 0 & \ldots & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \ldots & 1\end{vmatrix}=1.$$

Aquí se ha tenido en cuenta que el elemento γ_{kl} del producto de dos determinantes es igual a la suma de los productos de los elementos de la k-ésima fila por los elementos correspondientes de la l-ésima fila (véase § 2, propiedad j)).

Señalemos también que el determinante formado por las componentes de los vectores de la base $i^1, \ldots, i^n$ es igual a 1:

$$\begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} = 1.$$

Si el determinante de la base ortogonal $a^1, \ldots, a^n$ es igual a 1, $|\Lambda| = 1$ (véase (6)), entonces se dice que esta *base tiene la misma orientación que la base* $i^1, \ldots, i^n$. Si $|\Lambda| = -1$, se dice que la *orientación es contraria*. Estas definiciones concuerdan con las definiciones respectivas en los casos bidimensional y tridimensional expuestos en el § 11.

Nota 1. Si en la matriz Λ (véase (8)) las coordenadas del vector a^k en la base $(i^1, \ldots, i^n)$ las hubiésemos colocado en la k-ésima columna ($k = 1, \ldots, n$), entonces el cambio de las coordenadas del vector x' a las del vector x se realizaría mediante las filas de la matriz Λ. El cambio de x a x' en la fórmula (13) se realizaría mediante la matriz Λ'.

Nota 2. En el espacio complejo R_n la matriz (8), siendo a_{kl} complejos, se llama ortogonal si

$$\sum_{s=1}^{n} a_{ks}\overline{a}_{ls} = \delta_{kl} \quad (k,\ l = 1, \ldots, n). \tag{9'}$$

Demostremos que un sistema ortonormal de vectores (6) (ó (7)) en el espacio complejo R_n engendra una matriz ortogonal Λ (véase (8)). En efecto, en el espacio complejo R_n el producto escalar de los vectores x, y cumple las propiedades (véase § 6, b'), c')):

$$(\overline{x,\ y}) = (y,\ x), \quad (\alpha x + \beta y,\ z) = \alpha\,(x,\ z) + \beta\,(y,\ z),$$

siendo α, β números complejos. Entonces,

$$(x,\ \alpha y + \beta z) = (\overline{\alpha y + \beta z,\ x}) = (\overline{\alpha y,\ x}) + (\overline{\beta z,\ x}) =$$
$$= \overline{\alpha\,(y,\ x)} + \overline{\beta\,(z,\ x)} = \overline{\alpha}\,(x,\ y) + \overline{\beta}\,(x,\ z).$$

Pero, entonces, para un sistema ortogonal de vectores $a^1, \ldots, a^n$ se verifica

$$\delta_{kl} = (a^k,\ a^l) = \Big(\sum_{s=1}^{n} a_{ks} i^s,\ \sum_{r=1}^{n} a_{lr} i^r\Big) = \sum_{s=1}^{n}\sum_{r=1}^{n} a_{ks}\overline{a}_{lr}\,(i^s,\ i^r) =$$
$$= \sum_{s=1}^{n}\sum_{r=1}^{n} a_{ks}\overline{a}_{lr}\delta_{sr} = \sum_{s=1}^{n} a_{ks}\overline{a}_{ls} \tag{10'}$$

de modo que la matriz Λ es ortogonal. Vemos que, recíprocamente, la ortogonalidad de Λ implica la ortonormalidad de los vectores $a^1, \ldots, a^n$ determinados por las fórmulas (7).

Multipliquemos escalarmente el vector i^l por el vector a^k (véase (7)):

$$(i^l, a^k) = (i^l, \sum_{s=1}^{n} a_{ks} i^s) = \sum_{s=1}^{n} \overline{a}_{ks} (i^l, i^s) = \overline{a_{kl}}.$$

De aquí

$$i^l = \sum_{k=1}^{n} (i^l, a^k) a^k = \sum_{k=1}^{n} \overline{a_{kl}} a^k. \qquad (11')$$

Por lo tanto, el cambio de la base $(a^1, \ldots, a^n)$ a la base $(i^1, \ldots, i^n)$ se realiza mediante la matriz

$$\Lambda^* = \left\| \begin{matrix} \overline{a_{11}} & \ldots & \overline{a_{n1}} \\ \cdot & \cdot \cdot \cdot \cdot & \cdot \\ \overline{a_{n1}} & \ldots & \overline{a_{nn}} \end{matrix} \right\|.$$

Como las transformaciones (11′) son inversas a las transformaciones (7), queda también demostrado que la matriz ortogonal Λ cumple la siguiente propiedad notable:

$$\Lambda^{-1} = \Lambda^*. \qquad (14)$$

De (11′) se deduce que

$$\delta_{kl} = (i^k, i^l) = (\sum_{s=1}^{n} \overline{a}_{sk} a^s, \sum_{r=1}^{n} \overline{a}_{rl} a^r) = \sum_{s=1}^{n} \sum_{r=1}^{n} \overline{a_{sk}} a_{rl} \delta_{sr} = \sum_{s=1}^{n} \overline{a_{sk}} a_{sl}.$$

Por consiguiente, las igualdades

$$\delta_{kl} = \sum_{s=1}^{n} \overline{a}_{sk} a_{sl} \quad (k, l = 1, \ldots, n)$$

(así como (9′)) pueden servir como definición de matriz ortogonal Λ.

En virtud de (14) y de los resultados generales obtenidos en el § 16, señalemos las matrices que realizan las transformaciones ortogonales que se dan a continuación:

$$\begin{aligned} \Lambda&: (i^1, \ldots, i^n) \to (a^1, \ldots, a^n); \\ \Lambda^*&: (a^1, \ldots, a^n) \to (i^1, \ldots, i^n); \\ \overline{\Lambda}&: (x_1, \ldots, x_n) \to (x'_1, \ldots, x'_n); \\ \Lambda'&: (x'_1, \ldots, x'_n) \to (x_1, \ldots, x_n), \end{aligned}$$

donde $(x_1, \ldots, x_n)$ y $(x'_1, \ldots, x'_n)$ son las coordenadas de un vector arbitrario en el espacio complejo R_n respecto de la base $(i^1, \ldots, i^n)$ y de la base ortonormal $(a^1, \ldots, a^n)$, respectivamente.

Finalmente, la igualdad $||\Lambda||=1$ se demuestra en el caso complejo así:

$$||\Lambda||^2=|a_{kl}|\cdot|\bar{a}_{kj}|=\left|\sum_{s=1}^{n} a_{ks}\bar{a}_{ts}\right|=\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}=1.$$

Las matrices ortogonales también se llaman *unitarias.*

§ 18. Propiedades invariantes de los productos vectorial y escalar

Nota. Consideremos dos sistemas rectangulares de coordenadas en el espacio real tridimensional, (x_1, x_2, x_3) e (y_1, y_2, y_3), con los sistemas de versores (i^1, i^2, i^3) y (j^1, j^2, j^3), respectivamente. Supongamos que

$$j^k=\sum_{s=1}^{3} \alpha_{ks}i^s \quad (k=1, 2, 3) \tag{1}$$

(compárese con (7) y (13) § 17). Entonces, la matriz (real)

$$A=||\alpha_{ks}||$$

es ortogonal y (véase (13) § 17)

$$y_l=\sum_{s=1}^{3} \alpha_{ls}x_s \quad (l=1, 2, 3). \tag{2}$$

Un mismo punto (un vector) del espacio tiene en el primer sistema las coordenadas (las componentes) (x_1, x_2, x_3) y en el segundo sistema, las coordenadas (y_1, y_2, y_3). Además, las fórmulas de transformación de las coordenadas (el cambio de (x_1, x_2, x_3) a (y_1, y_2, y_3)) son exactamente iguales que las fórmulas de transformación del sistema de versores (i^1, i^2, i^3) al sistema de versores (j^1, j^2, j^3). En ambos casos se aplica una misma matriz ortogonal (véase (13) y (7) del § 17)

$$A=||\alpha_{kl}||.$$

En el primer caso, la matriz A se aplica al sistema de números (x_1, x_2, x_3) para obtener el sistema de números (y_1, y_2, y_3), y en el segundo caso, la misma matriz A se aplica a los versores (i^1, i^2, i^3) para obtener los versores (j^1, j^2, j^3).

Enunciemos la definición general de vector admitida en el cálculo tensorial.

Se llama vector en el espacio tridimensional a un objeto que se exprese en todo sistema rectangular de coordenadas mediante una terna de números, que se transforman del mismo modo (mediante la misma matriz) que las ternas de los versores de los respectivos sistemas de coordenadas.

Una terminología semejante se aplica también en el caso de los espacios n-dimensionales.

Sean $\boldsymbol{a}$ y $\boldsymbol{b}$ dos vectores del espacio (real) tridimensional, definidos en los sistemas de coordenadas con los versores (i^1, i^2, i^3) y (j^1, j^2, j^3), del modo siguiente:

$$\boldsymbol{a}=a_{x_1}i^1+a_{x_2}i^2+a_{x_3}i^3=a_{y_1}j^1+a_{y_2}j^2+a_{y_3}j^3,$$
$$\boldsymbol{b}=b_{x_1}i^1+b_{x_2}i^2+b_{x_3}i^3=b_{y_1}j^1+b_{y_2}j^2+b_{y_3}j^3.$$

Se verifica la igualdad

$$\boldsymbol{ab}=\sum_{s=1}^{3} a_{x_s}b_{x_s}=\sum_{r=1}^{3} a_{y_r}b_{y_r},$$

la cual muestra que *el producto escalar es invariante respecto de las transformaciones de los sistemas rectangulares de coordenadas.*

En efecto, como los sistemas de vectores (i^1, i^2, i^3) y (j^1, j^2, j^3) son ortonormales, éstos se transforman según las fórmulas (1), donde $\|\alpha_{kl}\|$ es una matriz ortogonal. Las componentes del vector $\boldsymbol{a}=(a_{x_1}, a_{x_2}, a_{x_3})$ se transforman en las componentes $(a_{y_1}, a_{y_2}, a_{y_3})$ mediante la misma matriz (véase (2)). Por lo tanto,

$$\sum_{r=1}^{3} a_{y_r}b_{y_r}=\sum_{r=1}^{3}\Big(\sum_{s=1}^{3}\alpha_{rs}a_{x_s}\cdot\sum_{\nu=1}^{3}\alpha_{r\nu}b_{x_\nu}\Big)=$$
$$=\sum_{s=1}^{3}\sum_{\nu=1}^{3}a_{x_s}b_{x_\nu}\Big(\sum_{r=1}^{3}\alpha_{rs}\alpha_{r\nu}\Big)=\sum_{s=1}^{3}a_{x_s}b_{x_s}=\boldsymbol{ab}. \qquad (3)$$

Hemos demostrado mediante cálculos que el producto escalar $\boldsymbol{ab}$ es invariante. Por cierto, de la otra definición de producto escalar de los vectores (de los segmentos orientados), de la definición geométrica, en virtud de la cual $\boldsymbol{ab}=|\boldsymbol{b}|\,\mathrm{pr}_{\boldsymbol{b}}\,\boldsymbol{a}$, inmediatamente se observa que este número es invariante, ya que esta definición no está relacionada con ningún sistema de coordenadas.

En lo que se refiere al producto vectorial $\boldsymbol{a}\times\boldsymbol{b}$, éste es invariante respecto de los sistemas rectangulares de coordenadas de la misma orientación. En los sistemas con los versores (i^1, i^2, i^3) y (j^1, j^2, j^3), el producto escalar se expresa del modo siguiente:

$$[\boldsymbol{a}\times\boldsymbol{b}]_{i^1,i^2,i^3}=\begin{vmatrix} i^1 & i^2 & i^3\\ a_{x_1} & a_{x_2} & a_{x_3}\\ b_{x_1} & b_{x_2} & b_{x_3}\end{vmatrix},$$

$$[\boldsymbol{a}\times\boldsymbol{b}]_{j^1,j^2,j^3}=\begin{vmatrix} j^1 & j^2 & j^3\\ a_{y_1} & a_{y_2} & a_{y_3}\\ b_{y_1} & b_{y_2} & b_{y_3}\end{vmatrix}.$$

En virtud de las fórmulas (1) y (2), se tiene,

$$[a\times b]_{j^1,j^2,j^3}=\begin{vmatrix}\sum_{s=1}^{3}\alpha_{1s}i^s & \sum_{s=1}^{3}\alpha_{2s}i^s & \sum_{s=1}^{3}\alpha_{3s}i^s\\ \sum_{s=1}^{3}\alpha_{1s}a_{x_s} & \sum_{s=1}^{3}\alpha_{2s}a_{x_s} & \sum_{s=1}^{3}\alpha_{3s}a_{x_s}\\ \sum_{s=1}^{3}\alpha_{1s}b_{x_s} & \sum_{s=1}^{3}\alpha_{2s}b_{x_s} & \sum_{s=1}^{3}\alpha_{3s}b_{x_s}\end{vmatrix}=$$

$$=\begin{vmatrix} i^1 & i^2 & i^3\\ a_{x_1} & a_{x_2} & a_{x_3}\\ b_{x_1} & b_{x_2} & b_{x_3}\end{vmatrix}\begin{vmatrix}\alpha_{11} & \alpha_{12} & \alpha_{13}\\ \alpha_{21} & \alpha_{22} & \alpha_{23}\\ \alpha_{31} & \alpha_{32} & \alpha_{33}\end{vmatrix}=\pm\begin{vmatrix} i^1 & i^2 & i^3\\ a_{x_1} & a_{x_2} & a_{x_3}\\ b_{x_1} & b_{x_2} & b_{x_3}\end{vmatrix}=\pm[a\times b]_{i^1,i^2,i^3}, \qquad (4)$$

donde hay que poner el signo + ó — según que el determinante $|\alpha_{ik}|$ sea igual a $+1$ o a -1, o lo que es lo mismo, según que la transformación considerada de las coordenadas cambie la orientación o no.

En el caso dado, al multiplicar los determinantes se ha utilizado la regla siguiente:

El elemento γ_{ik} de la matriz del producto $\|\gamma_{ik}\|$ se determina como el producto de la i-ésima fila del primer determinante por la k-ésima fila del segundo determinante (véase § 2, propiedad j)).

En resumen, queda demostrado mediante cálculos, que *el producto vectorial de dos vectores es invariante respecto de las transformaciones de los sistemas rectangulares de coordenadas que no varían la orientación.*

Las transformaciones (3), (4) son interesantes porque se generalizan al caso importantísimo del análisis matemático del vector simbólico nabla $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)$.

§ 19. Transformación de las coordenadas rectangulares en el plano

Consideremos el plano R_2 en el que se ha introducido un sistema rectangular de coordenadas (x_1, x_2). Sean

$$i^1 = (1, 0), \quad i^2 = (0, 1)$$

los versores de los ejes x_1, x_2. Los versores i^1, i^2 forman una base ortonormal en R_2.

Un vector unitario (normal) arbitrario b^1 puede expresarse del modo siguiente:

$$b = (\cos\alpha, \operatorname{sen}\alpha) \quad (0 \leqslant \alpha < 2\pi).$$

Un vector unitario ortogonal (perpendicular) a b^1, que denotaremos por b^2, sólo puede corresponder al ángulo $\alpha + \frac{\pi}{2}$ ó $\alpha - \frac{\pi}{2}$.

Como

$$\cos\left(\alpha+\frac{\pi}{2}\right)=-\operatorname{sen}\alpha,\quad \operatorname{sen}\left(\alpha+\frac{\pi}{2}\right)=\cos\alpha,$$
$$\cos\left(\alpha-\frac{\pi}{2}\right)=\operatorname{sen}\alpha,\quad \operatorname{sen}\left(\alpha-\frac{\pi}{2}\right)=-\cos\alpha,$$

todos los sistemas ortonormales posibles $\boldsymbol{b}^1$, $\boldsymbol{b}^2$ en R_2 se determinan por las igualdades (fig. 34)

$$\left.\begin{aligned}\boldsymbol{b}^1&=\boldsymbol{i}^1\cos\alpha+\boldsymbol{i}^2\operatorname{sen}\alpha,\\ \boldsymbol{b}^2&=-\boldsymbol{i}^1\operatorname{sen}\alpha+\boldsymbol{i}^2\cos\alpha\end{aligned}\right\}\ (0\leqslant\alpha<2\pi), \tag{1'}$$

correspondientes a la rotación de los ejes en torno del origen un

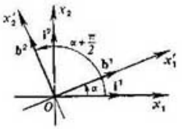

Fig. 34

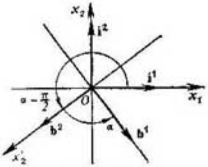

Fig. 35

ángulo α conservando la orientación, o por las igualdades (fig. 35)

$$\left.\begin{aligned}\boldsymbol{b}^1&=\boldsymbol{i}^1\cos\alpha+\boldsymbol{i}^2\operatorname{sen}\alpha,\\ \boldsymbol{b}^2&=\boldsymbol{i}^1\operatorname{sen}\alpha+\boldsymbol{i}^2\cos\alpha,\end{aligned}\right\} \tag{1''}$$

correspondientes a la rotación de los ejes en torno del origen un ángulo α cambiando la orientación.

Ambas transformaciones se reúnen en la siguiente fórmula común:

$$\begin{aligned}\boldsymbol{b}^1&=\alpha_{11}\boldsymbol{i}^1+\alpha_{12}\boldsymbol{i}^2,\\ \boldsymbol{b}^2&=\alpha_{21}\boldsymbol{i}^1+\alpha_{22}\boldsymbol{i}^2,\end{aligned} \tag{1}$$

donde la matriz de la transformación

$$\Lambda=\begin{Vmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{Vmatrix} \tag{2}$$

es ortogonal (la suma de los cuadrados de los elementos de cada una de sus filas o columnas es igual a 1, mientras que el producto escalar de dos filas o columnas distintas es igual a 0).

Cualquier transformación ortogonal determinada (1) es en realidad una de las transformaciones (1') o (1") para cierto α.

Como la matriz (2) es ortogonal, de (1) se deduce que

$$\begin{aligned} i^1 &= \alpha_{11} b^1 + \alpha_{21} b^2, \\ i^2 &= \alpha_{12} b^1 + \alpha_{22} b^2, \end{aligned} \tag{3}$$

con lo que obtenemos la transformación inversa a (1), cuya matriz

$$\begin{Vmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{Vmatrix} = \Lambda'$$

es conjugada a Λ.

Tomemos en el plano un vector (un punto) arbitrario a. Supongamos que en el sistema antiguo tiene las coordenadas (x_1, x_2) y en el sistema nuevo, las coordenadas (x_1', x_2'). Entonces,

$$a = x_1 i^1 + x_2 i^2 = x_1' b^1 + x_2' b^2. \tag{4}$$

En virtud de las fórmulas (3) y (4), se tiene,

$$\begin{aligned} x_1' b^1 + x_2' b^2 &= x_1 (\alpha_{11} b^1 + \alpha_{21} b^2) + x_2 (\alpha_{12} b^1 + \alpha_{22} b^2) = \\ &= (\alpha_{11} x_1 + \alpha_{12} x_2)\, b^1 + (\alpha_{21} x_1 + \alpha_{22} x_2)\, b^2 \end{aligned}$$

e igualando las componentes para los versores b^1, b^2 iguales, respectivamente, obtenemos:

$$\left.\begin{aligned} x_1' &= \alpha_{11} x_1 + \alpha_{12} x_2, \\ x_2' &= \alpha_{21} x_1 + \alpha_{22} x_2. \end{aligned}\right\} \tag{5}$$

Ahora bien, en virtud de las fórmulas (1) y (4),

$$\begin{aligned} x_1 i^1 + x_2 i^2 &= x_1' (\alpha_{11} i^1 + \alpha_{12} i^2) + x_2' (\alpha_{21} i^1 + \alpha_{22} i^2) = \\ &= (\alpha_{11} x_1' + \alpha_{21} x_2')\, i^1 + (\alpha_{12} x_1' + \alpha_{22} x_2')\, i^2, \end{aligned}$$

de donde, igualando las componentes de i^1 e i^2, obtenemos las fórmulas inversas a (5):

$$\left.\begin{aligned} x_1 &= \alpha_{11} x_1' + \alpha_{21} x_2', \\ x_2 &= \alpha_{12} x_1' + \alpha_{22} x_2'. \end{aligned}\right\} \tag{6}$$

Si, además de la transformación (6), hacemos una traslación del origen de los ejes x_1', x_2' al punto O' con las coordenadas $x_1 = x_1^0$, $x_2 = x_2^0$, entonces las fórmulas (6) se complican, evidentemente, del modo siguiente:

$$\left.\begin{aligned} x_1 &= x_1^0 + \alpha_{11} x_1' + \alpha_{21} x_2', \\ x_2 &= x_2^0 + \alpha_{12} x_1' + \alpha_{22} x_2'. \end{aligned}\right\} \tag{7}$$

En resumen, una transformación arbitraria de las coordenadas rectangulares (x_1, x_2) en las coordenadas rectangulares (x_1', x_2') acompañada de una traslación del origen del sistema (x_1, x_2) al punto $O' = (x_1^0, x_2^0)$, se expresa por las fórmulas (7), donde la matriz

$$\begin{Vmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{Vmatrix}$$

es ortogonal.

La transformación correspondiente que conserva la orientación del sistema de coordenadas tiene la forma

$$\left.\begin{aligned} x_1 &= x_1^0 + x_1' \cos\alpha - x_2' \operatorname{sen}\alpha, \\ x_2 &= x_2^0 + x_1' \operatorname{sen}\alpha + x_2' \cos\alpha \end{aligned}\right\} \qquad (7')$$

y la transformación que cambia la orientación, la forma

$$\left.\begin{aligned} x_1 &= x_1^0 + x_1' \cos\alpha + x_2' \operatorname{sen}\alpha, \\ x_2 &= x_2^0 + x_1' \operatorname{sen}\alpha - x_2' \cos\alpha \end{aligned}\right\} \qquad (7'')$$

(las matrices de los coeficientes de x_1' y x_2' en (7') y (7''), son las traspuestas de (1') y (1''), respectivamente).

§ 20. Subespacios lineales en R_n

Un conjunto L en R_n ($L \subset R_n$) se llama *subespacio lineal del espacio* R_n o, abreviadamente, subespacio en R_n, si la pertenencia a L de dos vectores arbitrarios $\boldsymbol{x}$ e $\boldsymbol{y}$ ($\boldsymbol{x}, \boldsymbol{y} \in L$) implica que el vector $\alpha\boldsymbol{x} + \beta\boldsymbol{y}$ también pertenece a L ($\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in L$) para cualesquiera números α y β *). *Un subespacio L* se llama *m-dimensional*, si existe en el mismo un sistema linealmente independiente $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m$ compuesto por m vectores, pero no existe ningún sistema compuesto por $m + 1$ vectores linealmente independientes.

En este caso, si $\boldsymbol{a}$ es un vector arbitrario de L ($\boldsymbol{a} \in L$), entonces el sistema $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m, \boldsymbol{a}$ es linealmente dependiente, o sea, existe un sistema no trivial de números $\lambda_1, \ldots, \lambda_m, \lambda_{m+1}$ tales que

$$\lambda_1\boldsymbol{a}^1 + \ldots + \lambda_m\boldsymbol{a}^m + \lambda_{m+1}\boldsymbol{a} = \boldsymbol{0}. \qquad (1)$$

Aquí $\lambda_{m+1} \neq 0$, pues en caso contrario sería

$$\lambda_1\boldsymbol{a}^1 + \ldots + \lambda_m\boldsymbol{a}^m = 0$$

y en virtud de la independencia lineal del sistema $\boldsymbol{a}', \ldots, \boldsymbol{a}^m$ resultaría $\lambda_1 = \ldots = \lambda_m = 0$, y todo el sistema $\lambda_1, \ldots, \lambda_m,$

*) Fácilmente se comprueba que un subespacio lineal también es un espacio lineal (véase la nota 1 en el ap. 6.5) ya que se cumplen las propiedades 1)—8) enunciadas en el ap. 6.1. (*Nota del Tr.*)

λ_{m+1} sería trivial. Entonces, en la ecuación (1) se puede despejar $\boldsymbol{a}$:

$$\boldsymbol{a}=\mu_1\boldsymbol{a}^1+\ldots+\mu_m\boldsymbol{a}^m \quad (\mu_s=-\lambda_s/\lambda_{m+1}), \tag{2}$$

o sea, que este vector puede expresarse como una combinación lineal de los vectores $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m$. Por otra parte, una combinación lineal de la forma (2) pertenece a L, puesto que L es un subespacio. Por lo tanto, se dice que el sistema $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m$ es una base en L. Está claro que cualquier otro sistema linealmente independiente de vectores $\boldsymbol{b}^1, \ldots, \boldsymbol{b}^m$ pertenecientes a L es una base en L.

Desarrollando los vectores $\boldsymbol{b}^k$ por los vectores $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m$, obtenemos

$$\boldsymbol{b}^k=\sum_{s=1}^{m} b_{ks}\boldsymbol{a}^s \quad (k=1, \ldots, m).$$

Razonando de un modo similar a como se hizo en el § 16 para R_n (donde ahora hay que sustituir $\boldsymbol{i}^s$ y $\boldsymbol{a}^k$ por $\boldsymbol{a}^s$ y $\boldsymbol{b}^k$, respectivamente), se puede demostrar que el sistema $\boldsymbol{b}^1, \ldots, \boldsymbol{b}^m$ es linealmente independiente si, y sólo si, el determinante $|b_{ks}|$ es distinto de cero, $|b_{ks}|\neq 0$, y que cualquier sistema independiente compuesto por $l<m$ vectores ya no puede ser una base en L.

El espacio R_n se puede considerar como un subespacio de R_n que tiene dimensión n.

El conjunto formado por el único vector nulo $\boldsymbol{0}$, es un subespacio lineal $(\alpha\boldsymbol{0}+\beta\boldsymbol{0}=\boldsymbol{0})$; suele decirse que éste tiene dimensión 0. El vector $\boldsymbol{0}$ no forma un sistema linealmente independiente, ya que la igualdad $\lambda\boldsymbol{0}=\boldsymbol{0}$, donde λ es un número, no implica necesariamente que λ sea igual a cero.

Si $\boldsymbol{x}^0$ es un vector no nulo, $\boldsymbol{x}^0\neq\boldsymbol{0}$, entonces el conjunto de los vectores de la forma $\lambda\boldsymbol{x}^0$, donde λ es un número arbitrario, es un *subespacio unidimensional*. En éste se puede tomar por base el vector $\boldsymbol{x}^0$.

Sea L un subespacio lineal en R_n. Se dice que un vector $\boldsymbol{v}\in R_n$ es *ortogonal a* L, si es ortogonal a cualquier vector $\boldsymbol{u}\in L$. Denotemos por L' el conjunto de todos los vectores que son ortogonales a L. Entonces, L' es un subespacio en R_n. En efecto, sean $\boldsymbol{v}, \boldsymbol{v}'\in L'$, o sea, que

$$(\boldsymbol{v}, \boldsymbol{u})=0, \quad \forall\boldsymbol{u}\in L;$$

$$(\boldsymbol{v}', \boldsymbol{u})=0, \quad \forall\boldsymbol{u}\in L.$$

Entonces, para cualesquiera números α y β

$$(\alpha\boldsymbol{v}+\beta\boldsymbol{v}', \boldsymbol{u})=\alpha(\boldsymbol{v}, \boldsymbol{u})+\beta(\boldsymbol{v}', \boldsymbol{u})=0, \quad \forall\boldsymbol{u}\in L,$$

o sea, $\alpha\boldsymbol{v}+\beta\boldsymbol{v}'\in L'$.

Por definición, *un subespacio $L'\subset R_n$ se llama ortogonal de un subespacio dado $L\subset R_n$*, si L' es un conjunto de *todos* los vectores de R_n cada uno de los cuales es ortogonal a L.

A continuación se demuestra un teorema que clarifica la estructura del subespacio arbitrario $L \subset R_n$ y del espacio ortogonal de él $L' \subset R_n$. En particular, de aquí se deduce que *si L' es el subespacio ortogonal de L, entonces viceversa L es el subespacio ortogonal de L'.*

TEOREMA 1. *Sea L un subespacio lineal distinto de R_n y del subespacio nulo. Entonces:*

a) *existen un número entero m que satisface las desigualdades*

$$1 \leqslant m < n \tag{3}$$

y una base ortonormal

$$a^1, \ldots, a^m \tag{4}$$

en L; si esta base se prolonga de cualquier modo hasta la base ortonormal en R_n

$$a^1, \ldots, a^m, a^{m+1}, \ldots, a^n \tag{5}$$

entonces el subespacio lineal L' con base

$$a^{m+1}, \ldots, a^n \tag{6}$$

posee las siguientes propiedades;

b) *L' es un subespacio ortogonal de L;*

c) *L es un subespacio ortogonal de L';*

d) *todo vector $a \in R_n$ puede representarse en forma de la suma*

$$a = u + v$$

donde $u \in L$, $v \in L'$ de un modo único.

DEMOSTRACIÓN. Según la hipótesis L se distingue del subespacio nulo, por consiguiente, en L existe un vector x distinto de 0. Normalizando x obtenemos el vector normal

$$a^1 = \frac{1}{|x|} x.$$

Designemos con a^2 cualquier vector normal perteneciente a L y ortogonal a a^1 ($|a^2| = 1$, $(a^2, a^1) = 0$) si existe tal vector. Luego, designemos con a^3 un vector normal perteneciente a L y ortogonal a a^1 y a^2 ($|a^3| = 1$, $(a^3, a^1) = (a^3, a^2) = 0$) si existe tal vector. Este proceso terminará en una m-ésima etapa donde m satisface las desigualdades (3), es decir, se hallará el sistema ortonormal de vectores (4) pertenecientes a L, pero ya no existirá en L un vector unitario ortogonal a los vectores $a^1, \ldots, a^m$. En rigor, $m \geqslant 1$, puesto que es evidente que $a^1 \in L$. Por otro lado, m no puede ser igual a n. En el caso contrario los vectores $a^1, \ldots, a^n$ pertenecerían a L, perteneciendo conjuntamente a este mismo subespacio L todas las combinaciones lineales $\sum_{k=1}^{n} \lambda_k a^k$ y entonces resultaría que L coincide con R_n, pero L es distinto de R_n. El sistema ortonormal obtenido $a^1, \ldots, a^m$ es base en L. De hecho, junto con los vectores $a^1, \ldots, a^m$ también pertenecen a L todas las combinaciones lineales de ellos $\sum_{k=1}^{m} \lambda_k a^k$. Pero en L no hay otros vectores porque si admitimos que cierto vector $a \in L$ no es una combinación lineal, a puede ser anotado como la suma

$$a = \sum_{k=1}^{m} (a, a^k) a^k + y \tag{7}$$

donde $y \neq 0$. Puesto que los vectores a y a^k pertenecen al subespacio L, concluiríamos que el vector

$$y = a - \sum_{k=1}^{m} (a, a^k) a^k$$

también pertenece a L. Mas, el vector y es ortogonal a todos los a^s $(s = 1, \ldots, m)$ (véase el § 17, (4)). El vector normalizado

$$b = \frac{1}{|y|} y \tag{8}$$

también pertenecería a L y sería ortogonal a todos los a^k $(k = 1, \ldots, m)$. Pero esto es imposible en virtud de la propiedad maximal del número m. Así queda demostrada la afirmación a) del teorema.

La completación del sistema ortonormal (4) hasta una base ortonormal (5) en R_n se realiza de acuerdo al teorema 1 del § 17. Denotemos por L' el subespacio de todas las combinaciones lineales $v = \sum_{k=m+1}^{n} \mu_k a^k$ de los vectores del sistema (6). Está claro que cada uno de estos vectores es ortogonal a cualquier vector $u \in L$, que se expresa en forma de una suma $u = \sum_{k=1}^{m} \lambda_k a^k$. Por otra parte, si $a \in R_n$ es un vector arbitrario, ortogonal a todos los vectores $u \in L$ y, en particular, a los vectores $a^1, \ldots, a^m$, entonces, su desarrollo en la base (5) tiene la forma

$$a = \sum_{k=1}^{n} (a, a^k) a^k = \sum_{k=m+1}^{n} (a, a^k) a^k,$$

o sea, $a \in L'$. Queda demostrada la proposición b) del teorema.

Ahora bien, cualquier vector $u = \sum_{k=1}^{m} \lambda_k a^k \in L$ es ortogonal a todos los vectores $v = \sum_{k=m+1}^{n} \mu_k a^k \in L'$ y, si se sabe que algún vector

$a=\sum_{k=1}^{n}(a,\ a^k)\cdot a^k$ es ortogonal a todos los vectores de L' y, en particular, a los vectores $a^{m+1},\ \ldots,\ a^n$, entonces $a=\sum_{k=1}^{m}(a,\ a^k)\,a^k$, o sea, $a\in L$. Queda demostrada la proposición c).

Finalmente, si $a\in R_n$ es un vector arbitrario, entonces éste puede expresarse unívocamente en forma de una suma

$$a=\sum_{k=1}^{n}(a,\ a^k)\,a^k=u+v,$$

donde

$$u=\sum_{k=1}^{m}(a,\ a^k)\,a^k\in L,\quad v=\sum_{k=m+1}^{n}(a,\ a^k)\,a^k\in L'.$$

Con esto, el teorema queda completamente demostrado.

TEOREMA 2. *Sea L un subespacio en R_n, de dimensión m. Entonces, el complemento ortogonal $L'\subset R_n$ de L es de dimensión $n-m$ y, a su vez, L es ortogonal a L'.*

DEMOSTRACIÓN. Si L es distinto de R_n y del subespacio nulo, entonces, evidentemente, este teorema está contenido en el teorema 1.

Supongamos que L es el subespacio nulo. Como cualquier vector $a\in R_n$ es ortogonal al vector $\mathbf{0}$, resulta, $L'=R_n$ y la dimensión de R_n es igual a $n-0=n$. Recíprocamente, el vector $\mathbf{0}$ es ortogonal a todos los vectores $a\in R_n=L'$. Pero no hay otros vectores que sean ortogonales a todos los vectores de R_n, ya que todo vector distinto del vector $\mathbf{0}$ no es ortogonal a sí mismo. Queda demostrado que L es ortogonal a L'.

Si $L=R_n$, el razonamiento es similar.

COROLARIO 1. *Sea dado un sistema de vectores*

$$x^1,\ \ldots,\ x^m, \tag{9}$$

y sea L' el subespacio formado por todos los vectores $v\in R_n$ que son ortogonales a los vectores de este sistema:

$$(v,\ x^k)=0\quad (k=1,\ \ldots,\ m).$$

Supongamos ahora que a es un vector ortogonal a todos los vectores v, o sea, que es ortogonal al subespacio L'. Entonces, a se expresa como una combinación lineal de los vectores del sistema dado (9):

$$a=\sum_{k=1}^{m}\lambda_k x^k.$$

DEMOSTRACION. Consideremos el subespacio L formado por todas las combinaciones lineales posibles de los vectores del sistema (9), de modo que todo vector $u \in L$ es una combinación lineal

$$u = \sum_{h=1}^{m} \lambda_k x^k.$$

En este caso, diremos también que el subespacio L está engendrado por los vectores del sistema (9).

Como todo vector $v \in L'$ es ortogonal a los vectores del sistema (9), éste, evidentemente, es ortogonal a cualquier vector $u \in L$. Esto muestra que el subespacio L' es ortogonal al subespacio L. Pero, entonces, según el teorema 2, L es también ortogonal a L', o sea, L consta de *todos* los vectores u que son ortogonales a L'. Por hipótesis, $\boldsymbol{a}$ es uno de tales vectores u y, por consiguiente, el vector $\boldsymbol{a}$ es una combinación lineal de los vectores del sistema (9).

§ 21. Teoremas de tipo Fredholm

En este párrafo se expone una teoría de las ecuaciones lineales que es paralela a la teoría expuesta en el § 4.

Esta es una teoría sin determinantes. En sus enunciados, el determinante del sistema de ecuaciones no figura explícitamente. La ventaja de esta teoría consiste en que ésta sirvió de base y analogía para numerosas generalizaciones que se hicieron en el análisis matemático. Las primeras generalizaciones importantes de éstas pertenecen a Fredholm[1]).

Consideremos de nuevo un operador lineal (véase § 15) A:

$$y = Ax \quad (x \in R_n), \tag{1}$$

que pone en correspondencia a cada vector $x \in R_n$ un vector $y \in R_n$ mediante las ecuaciones

$$y_i = \sum_{s=1}^{n} a_{is} x_s \quad (i = 1, \ldots, n). \tag{2}$$

Aquí

$$A = \| a_{is} \| = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix} \tag{3}$$

es una matriz cuadrada dada. Al operador A le corresponde el operador conjugado

$$y = A^* x \quad (x \in R_n), \tag{1*}$$

[1]) E. I. Fredholm (1866—1927), matemático sueco.

determinado por la matriz conjugada a (3):

$$A^* = \left\| \begin{matrix} \bar{a}_{11} & \bar{a}_{21} & \dots & \bar{a}_{n1} \\ \cdot & \cdot & \cdot & \cdot \\ \bar{a}_{1n} & \bar{a}_{2n} & \dots & \bar{a}_{nn} \end{matrix} \right\|. \tag{3*}$$

Mediante las componentes de los vectores x, y, el operador conjugado se expresa en la forma

$$y_j = \sum_{i=1}^{n} \overline{a_{ij}} x_i \quad (j = 1, \dots, n), \tag{2*}$$

o sea, que la componente y_j se expresa mediante las coordenadas del vector x aplicando la j-ésima fila de la matriz A^* o la j-ésima columna de la matriz $\bar{A}$.

Se verifica la igualdad

$$(Ax, z) = (x, A^*z), \quad \forall x, z \in R_n, \tag{4}$$

que es válida para todos los x, $z \in R_n$. En efecto, para R_n y a_{is} reales,

$$(Ax, z) = \sum_{i=1}^{n} y_i z_i = \sum_{i=1}^{n} \Big(\sum_{s=1}^{n} a_{is} x_s\Big) z_i = \sum_{s=1}^{n} \Big(\sum_{i=1}^{n} a_{is} z_i\Big) x_s = (x, A^*z).$$

En el caso complejo,

$$(Ax, z) = \sum_{j=1}^{n} y_j \bar{z}_j = \sum_{j=1}^{n} \Big(\sum_{s=1}^{n} a_{js} x_s\Big) \bar{z}_j =$$

$$= \sum_{s=1}^{n} x_s \sum_{j=1}^{n} a_{js} \bar{z}_j = \sum_{s=1}^{n} x_s \overline{\sum_{j=1}^{n} \bar{a}_{js} z_j} = (x, A^*z).$$

La igualdad (4) es característica para un operador conjugado, ya que, si para un operador B se verifica la igualdad

$$(Ax, z) = (x, Bz), \quad \forall x, z \in R_n, \tag{5}$$

entonces necesariamente $B = A^*$. En efecto, ($B = \| b_{ik} \|$), para a_{is}, b_{is}, R_n reales,

$$(Ax, z) = \sum_{i=1}^{n} \sum_{s=1}^{n} a_{is} x_s z_i,$$

$$(x, Bz) = \sum_{s=1}^{n} \sum_{i=1}^{n} b_{si} z_i x_s = \sum_{i=1}^{n} \sum_{s=1}^{n} b_{si} x_s z_i.$$

De (5) se deduce que

$$\sum_{i=1}^{n} \sum_{s=1}^{n} a_{is} x_s z_i = \sum_{i=1}^{n} \sum_{s=1}^{n} b_{si} x_s z_i, \quad \forall x, z \in R_n, \tag{6}$$

de donde $a_{is} = b_{si}$ $(i, s = 1, \ldots, n)$, lo que se comprueba haciendo en (6) $x = (0, \ldots, 0, 1, 0, \ldots, 0)$, $z = (0, \ldots, 0, 1, 0, \ldots, 0)$, donde en x la unidad ocupa el s-ésimo lugar y en z, el i-ésimo.

En el caso complejo,

$$(Ax, z) = \sum_{j=1}^{n} \sum_{s=1}^{n} a_{js} x_s \bar{z}_j, \quad (x, Bz) = \sum_{s=1}^{n} x_s \overline{\sum_{j=1}^{n} b_{sj} z_j} =$$

$$= \sum_{s=1}^{n} x_s \sum_{j=1}^{n} \bar{b}_{sj} \bar{z}_j = \sum_{j=1}^{n} \sum_{s=1}^{n} \bar{b}_{sj} x_s \bar{z}_j,$$

de donde $a_{js} = \bar{b}_{sj}$, o bien, $b_{sj} = \bar{a}_{js}$.

Por lo tanto, *el operador conjugado A^* de un operador lineal A se puede definir también como un operador lineal tal, que se cumple la igualdad* (4).

Las igualdades (1) y (1*) se pueden considerar como ecuaciones: dado el vector $y \in R_n$, se pide hallar $x \in R_n$ de tal modo que se cumpla la igualdad (1) o (1*).

Las ecuaciones homogéneas correspondientes tienen la forma

$$Ax = 0 \tag{1_0}$$

o bien,

$$\sum_{s=1}^{n} a_{is} x_s = 0 \quad (i = 1, \ldots, n) \tag{2_0}$$

y

$$A^* z = 0 \tag{1_0^*}$$

o bien,

$$\sum_{l=1}^{n} a_{lj} z_l = 0 \quad (j = 1, \ldots, n). \tag{2_0^*}$$

Denotemos por L la imagen del espacio R_n *mediante el operador* A:

$$L = A(R_n),$$

y denotemos por L' *el subespacio de todos los vectores* z *que satisfacen a la ecuación conjugada homogénea* (1_0^*).

Se ha llamado a L' subespacio puesto que junto con z y z' también pertenecen al mismo $\alpha z + \beta z'$ siendo α y β números arbitrarios:

$$A^*(\alpha z + \beta z') = \alpha A^* z + \beta A^* z' = 0.$$

L también es un subespacio, puesto que, si $y, y' \in L$, entonces existen unos vectores $x, x' \in R_n$ tales que $y = Ax$, $y' = Ax'$, y por consiguiente,

$$\alpha y + \beta y' = \alpha A x + \beta A x' = A(\alpha x + \beta x'),$$

de modo que $\alpha y + \beta y' \in L$.

Se verifica el siguiente lema (véase § 20, teorema 2).

LEMA 1. *Los subespacios L y L' son ortogonales uno al otro, es decir, L' es un conjunto de todos los vectores $z \in R_n$ cada uno de los cuales es ortogonal a L y L, a su vez, es un conjunto de todos los vectores $y \in R_n$ cada uno de los cuales es ortogonal a L'. Si L es de dimensión k, entonces, L' es de dimensión $n-k$.*

DEMOSTRACIÓN. Consideremos la igualdad

$$(Ax, z) = (x, A^*z), \tag{7}$$

que es válida para todos los vectores x, $z \in R_n$. Supongamos que z es un vector ortogonal a L; entonces, para éste, el primer miembro en (7) es igual a cero para todo $x \in R_n$. Pero, entonces, el segundo miembro también es igual a cero para todo $x \in R_n$ y, en particular, para $x = A^*z$:

$$(A^*z, A^*z) = 0.$$

Por lo tanto, $A^*z = 0$. Queda demostrado que si un vector z es ortogonal a L, entonces este vector satisface a la ecuación $A^*z = 0$ (o sea, $z \in L'$).

Recíprocamente, supongamos que el vector z satisface a la ecuación $A^*z = 0$. Para tal z, el segundo miembro en (7) es igual a cero para cualesquiera $x \in R_n$; pero, entonces, el primer miembro es igual a cero, de modo que z es ortogonal a todos los vectores de la forma Ax, o sea, a todos los vectores $y \in L$. En otras palabras, z es ortogonal a L.

Queda demostrado que L' es un conjunto de todos los vectores z ortogonales al subespacio L. Pero, entonces, en virtud del teorema 2 § 20, L es a su vez el conjunto de todos los vectores y ortogonales a L' y la suma de las dimensiones de L y L' es igual a n. El lema queda demostrado.

Se verifica el

TEOREMA 1. *Para que la ecuación*

$$y = Ax \tag{1'}$$

admita solución para un vector dado $y \in R_n$, es necesario y suficiente que el vector y sea ortogonal a todos los vectores z que satisfacen a la ecuación conjugada homogénea

$$A^*z = 0. \tag{1_0^*}$$

La solución x de la ecuación (1), *si ésta existe, puede expresarse como la suma*

$$x = x^0 + u,$$

donde x^0 es una solución particular cualquiera de la ecuación no homogénea (1) *y u es una solución arbitraria de la ecuación homogénea*

$$Au = 0. \tag{1_0}$$

Cualquier suma indicada es solución de (1').

DEMOSTRACIÓN. En virtud del lema 1, si $L = A\,(R_n)$ y L' es el conjunto de todos los vectores z que satisfacen a la ecuación $A^*z = 0$, entonces L y L' son subespacios ortogonales uno al otro. Pero, entonces, si para y existe solución de la ecuación (1), se tiene que $y \in L$ y necesariamente todos los vectores $z \in L'$ son ortogonales a y. Por otra parte, si el vector y es ortogonal a todos los vectores $z \in L'$, entonces $y \in L$, o sea, que existe un x tal que $y = Ax$.

Supongamos ahora que para un vector y existe solución de la ecuación (1'). Sea ésta x^0:

$$y = Ax^0.$$

Entonces, evidentemente, la suma $x^0 + u$, donde $Au = 0$, también es solución de la ecuación (1'):

$$A\,(x^0 + u) = Ax^0 + Au = y + 0 = y.$$

Recíprocamente, si x es una solución arbitraria de la ecuación (1') y x^0 es una solución particular determinada, entonces

$$y = Ax, \quad y = Ax^0,$$

y, por consiguiente,

$$0 = Ax - Ax^0 = A\,(x - x^0) = Au,$$

donde $u = x - x^0$, o sea, $x = x^0 + u$, donde u satisface la ecuación $Au = 0$.

Nota. Expliquemos con el ejemplo del espacio real R_2 la relación del teorema 1 con la teoría de Kronecker—Capelli.

Supongamos que el vector $y = (y_1, y_2)$ es ortogonal a todas las soluciones del sistema

$$\left.\begin{aligned} a_{11}z_1 + a_{21}z_2 = 0, \\ a_{12}z_1 + a_{22}z_2 = 0. \end{aligned}\right\} \tag{8}$$

Demostremos que entonces los rangos de la matriz A y de la matriz ampliada

$$B = \begin{Vmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \end{Vmatrix}$$

son iguales entre sí. Si rango $A = 2$, entonces, evidentemente, rango $B = 2$. Supongamos que rango $A = 1$. Siempre rango $B \geqslant$ rango $A = 1$. Por lo tanto, tenemos que demostrar que

$$\Delta_1 = \begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix} = 0, \quad \Delta_2 = \begin{vmatrix} a_{12} & y_1 \\ a_{22} & y_2 \end{vmatrix} = 0.$$

En efecto, como y es ortogonal a las soluciones (no triviales) del sistema (8) se tiene, $y_1z_1 + y_2z_2 = 0$. Por esto, suponiendo que

$z_1 \neq 0$, resulta:

$$\Delta_1 = a_{11}y_2 - a_{21}y_1 = a_{11}y_2 + a_{21}\frac{y_2 z_2}{z_1} = \frac{y_2}{z_1}(a_{11}z_1 + a_{21}z_2) = 0,$$

$$\Delta_2 = a_{12}y_2 - a_{22}y_1 = a_{12}y_2 + a_{22}\frac{y_2 z_2}{z_1} = \frac{y_2}{z_1}(a_{12}z_1 + a_{22}z_2) = 0.$$

De aquí se deduce que rango B = rango $A = 1$.

Recíprocamente, supongamos que el vector $y = (y_1, y_2)$ es tal que rango B = rango A, entonces (1) admite solución (x_1, x_2). Demostremos que y es ortogonal a las soluciones $z = (z_1, z_2)$ del sistema (8). En efecto,

$$y_1 z_1 + y_2 z_2 = (a_{11}x_1 + a_{12}x_2)\, z_1 + (a_{21}x_1 + a_{22}x_2))\, z_2 =$$
$$= (a_{11}z_1 + a_{21}z_2)\, x_1 + (a_{12}z_1 + a_{22}z_2)\, x_2 = 0 \cdot x_1 + 0 \cdot x_2 = 0.$$

TEOREMA 2. *Las ecuaciones homogéneas*

$$Ax = 0 \tag{1_0}$$

y

$$A^* x = 0 \tag{1_0^*}$$

tienen el mismo número de soluciones linealmente independientes.

En particular, si una de estas ecuaciones sólo tiene la solución trivial **0**, *o sea, que no tiene soluciones linealmente independientes, entonces, lo mismo ocurre con la otra.*

Observación. En el último caso la ecuación (1) admite solución única.

DEMOSTRACION. Las matrices A y A^* tienen el mismo rango, que denotaremos por k. Estas también tienen un mismo determinante Δ.

Si $k = n$, entonces $\Delta \neq 0$ y las ecuaciones (1_0) y (1_0^*) sólo admiten la solución trivial **0**. En este caso, según el teorema 1, la ecuación (1) tiene solución única para cualesquiera $y \in R_n$.

Supongamos ahora que $1 \leqslant k < n$. Después de una reordenación adecuada de las ecuaciones y de las componentes, para el determinante se tiene:

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

Escribamos ahora las primeras k ecuaciones (1_0) en la forma

$$\left.\begin{array}{l} a_{11}x_1 + \dots + a_{1k}x_k = -a_{1,\,k+1}x_{k+1} - \dots - a_{1n}x_n, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{k1}x_1 + \dots + a_{kk}x_k = -a_{k,\,k+1}x_{k+1} - \dots - a_{kn}x_n. \end{array}\right\} \tag{9}$$

A continuación exponemos una tabla de $n-k$ vectores:

$$\left.\begin{array}{l} x^1=(x_1^1, \ldots, x_k^1, 1, 0, \ldots, 0), \\ x^2=(x_1^2, \ldots, x_k^2, 0, 1, 0, \ldots, 0), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^{n-k}=(x_1^{n-k}, \ldots, x_k^{n-k}, 0, \ldots, 0, 1). \end{array}\right\} \quad (10)$$

Para obtener el primer vector, hacemos en el sistema (9) la sustitución

$$x_{k+1}=1, \ x_{k+2}=0, \ \ldots, \ x_n=0$$

y lo resolvemos respecto de $x_1, \ldots, x_k$. Las únicas soluciones que se obtienen aquí las denotamos por $x_1', \ldots, x_k'$. Para obtener el segundo vector, sustituimos en (9)

$$x_{k+1}=0, \quad x_{k+2}=1, \quad x_{k+3}=0, \ \ldots, \ x_n=0$$

y hallamos los números $x_1^2, \ldots, x_k^2$, etc. Los vectores (10) cumplen las siguientes propiedades.

1) El sistema de vectores (10) es linealmente independiente, ya que el rango de la matriz de estos vectores es igual al número de ellos, $\mu = n-k$.

2) Cada uno de los vectores del sistema (10) es solución de (todas) las ecuaciones (1_0), o sea, $Ax=\mathbf{0}$.

3) Todas las soluciones posibles de la ecuación $Ax=\mathbf{0}$ tienen la forma

$$\lambda_1 x^1 + \ldots + \lambda_{n-k} x^{n-k},$$

donde $\lambda_1, \ldots, \lambda_{n-k}$ son números arbitrarios.

Ordinariamente, estas tres conclusiones se sustituyen por la siguiente:

La ecuación (1_0) tiene $n-k$ soluciones linealmente independientes.

Mediante unos razonamientos semejantes, teniendo en cuenta que rango A = rango A^*, se demuestra que la ecuación $A^*x=\mathbf{0}$ también tiene $n-k$ soluciones linealmente independientes.

El teorema queda demostrado.

TEOREMA 3. *Si una de las ecuaciones homogéneas* (1_0) *ó* (1_0^*) *tiene* k *soluciones linealmente independientes, entonces la otra también tiene* k *soluciones linealmente independientes; las imágenes* $L = A(R_n)$ *y* $L^* = A^*(R_n)$ *del espacio* R_n, *obtenidas mediante los operadores* A *y* A^*, *son subespacios de dimensión* $n-k$.

DEMOSTRACIÓN. La primera tesis del teorema sobre la igualdad del número de soluciones linealmente independientes de las ecuaciones homogéneas (1_0) y (1_0^*) representa el teorema 2, y la segunda tesis, el lema 1, en virtud del cual la dimensión del subespacio L es igual a $n-k$, donde k es la dimensión del subespacio L' de los vectores

z que satisfacen la ecuación $A^*z = 0$. De un modo similar, la dimensión de L^* es igual a $n - k$, donde k es la dimensión del subespacio formado por los vectores u que satisfacen la ecuación $Au = 0$.

§ 22. Operador autoconjugado. Forma cuadrática

Un operador lineal

$$y_k = \sum_{l=1}^{n} a_{kl} x_l \quad (k = 1, \ldots, n) \tag{1}$$

o, abreviadamente,

$$y = Ax \quad (x \in R_n, \quad y \in R_n) \tag{2}$$

se llama *autoconjugado*[1]), si es igual a su conjugado ($A = A^*$), o sea, si

$$Ax = A^*x, \quad \forall x \in R_n, \tag{3}$$

en otras palabras, si la matriz A es simétrica:

$$a_{kl} = a_{lk} \quad (k, l = 1, \ldots, n) \tag{4}$$

(véase (3) y (3*) § 21). Aquí se ha considerado que a_{kl} y R_n son reales (véase más adelante la nota 1).

Para un operador autoconjugado se verifica la igualdad característica

$$(x, Az) = (Ax, z), \quad \forall x, z \in R_n$$

(véase § 21, (4)). Está claro que

$$(x, Ax) = \sum_{k=1}^{n} x_k \sum_{l=1}^{n} a_{kl} x_l = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} x_k x_l \quad (a_{kl} = a_{lk}). \tag{4'}$$

La expresión del último miembro en (4') se llama *forma cuadrática de n-ésimo orden*. Esta es una función continua del vector x, o lo que es lo mismo, de las variables $x_1, \ldots, x_n$.

Vamos a considerar esta función sobre el conjunto S de vectores x con la norma unidad ($|x| = 1$). Este conjunto representa una esfera en R_n de radio 1 con el centro en el punto 0, y por lo tanto, S es un conjunto acotado. Además, es cerrado [2]) puesto que si los puntos de una sucesión $\{x^\nu\}$ ($\nu = 1, 2, \ldots$) pertenecen a S (o sea,

---

[1]) O hermítico (*Nota del Tr.*).

[2]) Véase el § 8.12 del libro de Yu. S. Bugrov, S. M. Nikolski «Cálculo diferencial e integral», Editorial «Mir», 1984, o S. M. Nikolski «Curso de análisis matemático», t. I, Ed. URMO, Bilbao, 1972.

$|x^\nu| = 1$, $\nu = 1, 2, \ldots$) y esta sucesión tiende hacia un punto $x^0 \in R_n$ ($x^\nu \to x^0$, $\nu \to \infty$), entonces necesariamente $x^0 \in S$, o sea, $|x^0| = 1$, ya que $|1 - |x^0|| = ||x^\nu| - |x^0|| \leqslant |x^\nu - x^0| \to 0$, de donde $|x^0| = 1$.

Hallemos el valor máximo de la forma cuadrática (4') sobre la esfera S. Como la forma (4') es una función continua sobre un conjunto cerrado y acotado, su máximo sobre S se alcanza para un cierto vector x^1 ($|y^1| = 1$). Denotemos este máximo por λ_1:

$$\lambda_1 = (Ax^1, x^1) \geqslant (Ax, x), \quad \forall x : |x| = 1. \tag{5}$$

Consideremos el subespacio L' ortogonal al vector x^1 es decir, el conjunto de todos los vectores v que son ortogonales a x^1. Tomemos en L' un vector unitario cualquiera v^0 ($|v^0| = 1$). El vector

$$\cos\alpha \cdot x^1 + \operatorname{sen}\alpha \cdot v^0$$

depende de α y su norma es igual a la unidad

$$|\cos\alpha \cdot x^1 + \operatorname{sen}\alpha \cdot v^0| = (\cos\alpha \cdot x^1 + \operatorname{sen}\alpha \cdot v^0, \cos\alpha \cdot x^1 + \\ + \operatorname{sen}\alpha \cdot v^0)^{1/2} = (\cos^2\alpha + \operatorname{sen}^2\alpha)^{1/2} = 1.$$

Si $\alpha = 0$ este vector se convierte en x^1. Pero, entonces, la función

$$\psi(\alpha) = (A(\cos\alpha \cdot x^1 + \operatorname{sen}\alpha \cdot v^0), \cos\alpha \cdot x^1 + \operatorname{sen}\alpha \cdot v^0)$$

alcanza su máximo en el punto $\alpha = 0$ ($\psi(0) = (Ax^1, x^1)$) y, en virtud de la condición necesaria de extremo,

$$\psi'(0) = 0.$$

Calculemos esta derivada. Se tiene,

$$\psi(\alpha) = \cos^2\alpha\,(Ax^1, x^1) + \operatorname{sen} 2\alpha\,(Ax^1, v^0) + \operatorname{sen}^2\alpha\,(Av^0, v^0).$$

Por consiguiente,

$$\psi'(\alpha) = -\operatorname{sen} 2\alpha\,(Ax^1, x^1) + 2\cos 2\alpha\,(Ax^1, v^0) + \operatorname{sen} 2\alpha\,(Av^0, v^0)$$

y

$$\psi'(0) = 2\,(Ax^1, v^0) = 0.$$

Hemos demostrado que el vector Ax^1 es ortogonal a todos los vectores unitarios $v^0 \in L'$ y, por consiguiente, también a todos los vectores $v \in L'$. Pero, entonces, Ax^1 difiere de x^1 solamente en un factor (véase el corolario 1 al final del § 20), o sea,

$$Ax^1 = \lambda x^1,$$

donde λ es un número.

De la primera igualdad en (5), teniendo en cuenta que $|x^1| = 1$, se deduce que

$$\lambda_1 = (\lambda x^1, x^1) = \lambda.$$

Por lo tanto, hemos demostrado que el máximo de la forma cuadrática (4') sobre la esfera unidad $|x| = 1$ se alcanza en un punto x^1 tal que

$$\max_{|x|=1} (Ax, x) = (Ax^1, x^1) = \lambda_1.$$

Además,

$$Ax^1 = \lambda_1 x^1, \quad |x^1| = 1.$$

Vemos que el vector no trivial (no nulo) x^1 se transforma mediante el operador A en un vector $\lambda_1 x^1$, que es colineal al mismo.

Tal vector se llama *vector propio del operador* A, y el número λ_1, *valor propio perteneciente a este vector*.

Ahora vamos a considerar el operador A sobre el subespacio R^1, definido como el conjunto de los vectores $x \in R_n$ que son ortogonales al vector x^1 (anteriormente lo habíamos denotado por L'). R^1 es un subespacio $(n-1)$-dimensional; en él hay bases ortonormales compuestas por $n-1$ vectores. Nuestro objetivo consiste en hallar una tal base que, como veremos, está naturalmente relacionada con el operador A.

Es importante subrayar que la imagen $A(R^1)$ del subespacio R^1 mediante el operador A pertenece a R^1, ya que, si $(x, x^1) = 0$, entonces,

$$(Ax, x^1) = (x, Ax^1) = (x, \lambda_1 x^1) = \lambda_1 (x, x^1) = 0,$$

o sea, $Ax \in R^1$.

Está claro que el operador A, en forma trivial, sigue siendo autoconjugado sobre R^1, ya que siendo válida la igualdad

$$(Ax, y) = (x, Ay)$$

para todos los vectores $x, y \in R_n$, también es válida para todos los vectores $x, y \in R^1$.

Así pues, ahora consideramos el operador lineal autoconjugado A sobre el subespacio lineal R^1 de dimensión $n-1$. Podemos aplicar a éste todos los razonamientos expuestos anteriormente y demostrar que existe en R^1 un vector unitario x^2 tal que

$$\max_{\substack{|x|=1\\(x, x^1)=0}} (Ax, x) = \underset{\substack{|x^2|=1\\(x^2, x^1)=0}}{(Ax^2, x^2)} = \lambda_2 \leqslant \lambda_1.$$

Hay que tener en cuenta que la esfera unidad S^1 en R^1 se define, evidentemente, como el conjunto de los vectores unitarios x que son ortogonales a x^1. Además,

$$Ax^2 = \lambda_2 x^2.$$

Ya hemos hallado un segundo vector propio del operador A, el vector x^2, y su valor propio correspondiente λ_2, que, evidentemente, no es superior a λ_1 (al reducirse el campo de aplicación del operador, el máximo sólo puede disminuir). Además, $(x^1, x^2) = 0$.

De un modo semejante se puede introducir el subespacio R^2 de dimensión $n - 2$, ortogonal a x^1 y x^2, y demostrar que el operador A transforma R^2 en R^2; también se puede hallar un tercer vector unitario x^3 que sea ortogonal a x^1 y x^2 y tal que se verifique la igualdad

$$\max_{\substack{|x|=1 \\ (x,\, x^1)=0,\ (x,\, x^2)=0}} (Ax,\ x) = (Ax^3,\ x^3) = \lambda_3$$

y la igualdad

$$Ax^3 = \lambda_3 x^3 \quad (\lambda_3 \leqslant \lambda_2 \leqslant \lambda_1).$$

Continuando este proceso por inducción hasta obtener el n-ésimo vector x^n, obtendremos un sistema ortonormal de vectores

$$x^1,\ x^2,\ \ldots,\ x^n \tag{6}$$

y un sistema de números reales

$$\lambda_1,\ \lambda_2,\ \ldots,\ \lambda_n, \tag{7}$$

que cumplen las propiedades

$$\left.\begin{aligned} Ax^k = \lambda_k x^k \quad (k = 1,\ \ldots,\ n), \\ \lambda_1 \geqslant \lambda_2 \ldots \geqslant \lambda_n. \end{aligned}\right\} \tag{8}$$

Hemos obtenido *un sistema completo de vectores propios del operador* A y de valores propios pertenecientes a estos vectores. Como el sistema ortonormal (6) pertenece a R_n y consta de n vectores, éste es una base en R_n (véase § 17). Por ello, un vector arbitrario $x \in R_n$ admite un desarrollo por este sistema

$$x = \sum_{k=1}^{n} (x,\ x^k)\, x^k. \tag{9}$$

Entonces, el operador autoconjugado en cuestión A puede expresarse del modo siguiente:

$$Ax = A\Big(\sum_{k=1}^{n} (x,\ x^k)\, x^k\Big) = \sum_{k=1}^{n} (x,\ x^k)\, Ax^k = \sum_{k=1}^{n} \lambda_k (x,\ x^k)\, x^k. \tag{10}$$

Queda demostrado el teorema:

TEOREMA 1. *A todo operador autoconjugado A en el espacio R_n le corresponde un sistema ortogonal de vectores $x^1, \ldots, x^n$ (una base en R_n) y un sistema de números reales $\lambda_1, \ldots, \lambda_n$ tales que Ax, para cualquier $x \in R_n$, se expresa como la suma* (10).

La forma cuadrática (4′) se expresa, respectivamente, del modo siguiente:

$$(x,\ Ax)=\left(\sum_{h=1}^{n}(x,\ x^h)\,x^h,\ \ \sum_{s=1}^{n}\lambda_s(x,\ x^s)\,x^s=\sum_{s=1}^{n}\lambda_s(x,\ x^s)^2.\qquad (4'')$$

En la práctica, frecuentemente, se parte de una forma cuadrática

$$\sum_{k=1}^{n}\sum_{l=1}^{n}a_{kl}x_kx_l\quad (a_{kl}=a_{lk}).\qquad (4')$$

Para aplicar a ésta los resultados obtenidos, se puede definir en relación con ella el operador lineal

$$y=Ax,$$

determinado por las igualdades

$$y_i=\sum_{j=1}^{n}a_{ij}x_j\quad (i=1,\ \ldots,\ n).$$

En virtud de la condición $a_{kl}=a_{lk}$, éste es un operador autoconjugado y se le puede aplicar entonces el teorema 1. En términos de forma cuadrática, el teorema 1 puede enunciarse del modo siguiente:

TEOREMA 2. *Sea dada una forma cuadrática (4′) en un sistema de coordenadas n-dimensional* $(x_1,\ \ldots,\ x_n)$ *del espacio* R_n *con los versores* $i^1,\ \ldots,\ i^n$ $\left(x=\sum_{k=1}^{n}x_ki^k\right)$. *Entonces, existen un sistema rectangular de coordenadas* $(\xi_1,\ \ldots,\ \xi_n)$ *con los versores* $x^1,\ \ldots,\ x^n$, *que forman una base ortogonal* $\left(x=\sum_{h=1}^{n}\xi_hx^h\right)$, *y un sistema de números reales* $\lambda_1\ \ldots,\ \lambda_n$ *tales que la forma cuadrática (4′) en este sistema es la suma de los cuadrados de las coordenadas* ξ_s *del vector* x, *multiplicados por los números* λ_s, *respectivamente*:

$$\sum_{k=1}^{n}\sum_{l=1}^{n}a_{kl}x_kx_l=\sum_{s=1}^{n}\lambda_s\xi_s^2.\qquad (4'')$$

El cambio del primer miembro en (4′) al segundo se realiza una vez conocidos los desarrollos de los vectores $x^1,\ \ldots,\ x^h$ por los versores $i^1,\ \ldots,\ i^n$. Supongamos que

$$x^j=\sum_{s=1}^{n}\beta_{js}i^s$$

(véase § 17, (7), donde hay que sustituir a_{js} y a^k por β_{js} y x^j, respectivamente). Como $i^1,\ \ldots,\ i^n$ y $x^1,\ \ldots,\ x^n$ son bases ortonormales en R_n, la matriz

$$\Lambda = \| \beta_{js} \|$$

es ortogonal. Suponemos que ésta es conocida. Un mismo vector x puede desarrollarse en las dos bases:

$$x = \sum_{s=1}^{n} x_s i^s = \sum_{j=1}^{n} \xi_j x^j.$$

Pero, entonces,

$$\sum_{j=1}^{n} \xi_j x^j = \sum_{j=1}^{n} \xi_j \sum_{s=1}^{n} \beta_{js} i^s = \sum_{s=1}^{n} \left(\sum_{j=1}^{n} \beta_{js} \xi_j \right) i^s$$

y, en virtud de la independencia lineal del sistema $i^1, \ldots, i^n$, obtenemos:

$$x_s = \sum_{j=1}^{n} \beta_{js} \xi_j \quad (s = 1, \ldots, n). \tag{11}$$

Por lo tanto, el cambio de las coordenadas $\xi_1, \ldots, \xi_n$ a las coordenadas $x_1, \ldots, x_n$ se efectúa mediante la matriz Λ' traspuesta a Λ (o sea, mediante las filas de la matriz Λ' o las columnas de la matriz Λ).

Sustituyendo las expresiones (11) para x_s en el primer miembro de (4'''), obtenemos el segundo miembro. Escribamos esta igualdad

$$\sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} \sum_{j=1}^{n} \beta_{jk} \xi_j \sum_{i=1}^{n} \beta_{il} \xi_i =$$
$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{k=1}^{n} \beta_{jk} \left(\sum_{l=1}^{n} a_{kl} \beta_{il} \right) \xi_j \xi_i = \sum_{s=1}^{n} \lambda_s \xi_s^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} \delta_{ji} \lambda_i \xi_j \xi_i,$$

donde $\delta_{ji} = \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases}$ es el símbolo de Kronecker.

Igualando los coeficientes de $\xi_j \xi_i$ en ambos miembros de esta igualdad, obtenemos las igualdades:

$$\sum_{k=1}^{n} \beta_{jk} \left(\sum_{l=1}^{n} a_{kl} \beta_{il} \right) = \delta_{ji} \lambda_i \quad (i, j = 1, \ldots, n),$$

que se pueden interpretar del modo siguiente (véase § 15, (6)). Para la matriz

$$A = \| a_{kl} \| \quad (a_{kl} = a_{lk})$$

de un operador autoconjugado A existe una matriz ortogonal

$$\Lambda = \| \beta_{js} \|$$

tal que

$$\Lambda \cdot A \cdot \Lambda^{-1} = \mathfrak{U}, \tag{12}$$

donde $\mathfrak{U}$ es una matriz diagonal

$$\mathfrak{U} = \begin{Vmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_n \end{Vmatrix} \tag{13}$$

(λ_i son números reales), denominada *canónica*.

Señalemos que para una matriz ortogonal Λ, se tiene

$$\Lambda^{-1} = \Lambda'.$$

Como los determinantes de las matrices ortogonales cumplen la igualdad $|\Lambda| = |\Lambda^{-1}| = \pm 1$, de (12) se deduce que

$$\prod_{j=1}^{n} \lambda_j = |\mathfrak{U}| = |\Lambda| \cdot |A| \cdot |\Lambda^{-1}| = |A|. \tag{14}$$

En particular, queda demostrado el siguiente teorema:

TEOREMA 3. *Si el determinante $|A|$ de una matriz autoconjugada A es distinto de cero ($|A| \neq 0$), entonces, todos sus valores propios $\lambda_1, \ldots, \lambda_n$ son distintos de cero ($\lambda_j \neq 0$, $j = 1, \ldots, n$).*

Del teorema 2 se deduce que

1) Si $\lambda_1 \geqslant \ldots \geqslant \lambda_n > 0$, entonces la forma cuadrática es positiva para cualesquiera vectores $\xi \neq 0$, y por consiguiente, también para cualesquiera vectores $x \neq 0$. En este caso, ésta se llama *estrictamente positiva* (definida positiva).

2) Si $0 > \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, entonces la forma es negativa para cualesquiera $\xi \neq 0$, y por consiguiente, también para cualesquiera $x \neq 0$. En este caso, ésta se llama *estrictamente negativa* (definida negativa).

3) Si $\lambda_1 \geqslant \ldots \geqslant \lambda_n$ y $\lambda_n = 0$, entonces la forma es no negativa. Existe una dirección (el eje ξ_n), a lo largo del cual la forma es igual a cero. Esta es *una forma positiva, pero no estrictamente.*

4) Si $\lambda_1 = 0 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, entonces *es una forma negativa, pero no estrictamente.*

5) Si $\lambda_1 > 0$, mientras que $\lambda_n < 0$, entonces *la forma es indefinida.* Excluyendo el punto nulo, la forma es positiva a lo largo del eje ξ_1 y negativa a lo largo del eje ξ_n.

Resulta que según sea la forma de la matriz $\|A\|$ y según sean los signos de algunos de los determinantes engendrados por ella, se puede averiguar si sus valores propios son todos positivos, todos negativos o entre ellos hay algunos positivos y otros negativos. En esto consiste *el teorema de Sylvester* [1]).

[1]) J. J. Sylvester (1814—1897), matemático inglés.

Formemos la sucesión de los menores principales de la forma cuadrática (Ax, x):

$$\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \ldots,\ \Delta_n = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Según el teorema de Sylvester, que no demostraremos aquí, se verifican las siguientes proposiciones:

1. Si $\Delta_1 > 0$, $\Delta_2 > 0$, $\ldots$, $\Delta_n > 0$, entonces la forma es estrictamente positiva (definida positiva) (caso 1)).

2. Si $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, $\ldots$, $(-1)^n \Delta_n > 0$, entonces, la forma es estrictamente negativa (definida negativa) (caso 2)).

3. Si $\Delta_1 \geqslant 0$, $\Delta_2 \geqslant 0$, $\ldots$, $\Delta_n \geqslant 0$, o bien,

$$\Delta_1 \leqslant 0,\ \Delta_2 \geqslant 0,\ \ldots,\ (-1)^n \Delta_n \geqslant 0,$$

pero existe un j tal que $\Delta_j = 0$, entonces, la forma no es estrictamente definida.

4. En todos los demás casos la forma cuadrática es indefinida.

Nota 1. Si R_n es el espacio complejo y $a_{kl} = a_{lk}$ son de nuevo números reales, entonces los razonamientos expuestos anteriormente varían muy poco. La fórmula (4') se escribirá ahora así:

$$(x, Ax) = \sum_{k=1}^{n} x_k \sum_{l=1}^{n} a_{kl}\bar{x}_l = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl}x_k\bar{x}_l.$$

El número (x, Ax) sigue siendo real, pues

$$\overline{(x, Ax)} = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl}\bar{x}_k x_l = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{lk}x_k\bar{x}_l = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl}x_k\bar{x}_l = (x, Ax).$$

Esto muestra que todo lo expuesto anteriormente (las fórmulas (4') — (10)) sigue siendo válido y, en particular, los números $\lambda_1, \ldots, \lambda_n$ son reales también en el caso de R_n complejo. El teorema 1 se conserva también para R_n complejo. La fórmula (4") tiene ahora la forma

$$(x, Ax) = \sum_{s=1}^{n} \lambda_s |(x, x^s)|^2,$$

de modo que ahora ya hay que sustituir los números (x, x^s) por los cuadrados de sus módulos. La fórmula (4‴) tiene ahora la forma

$$\sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl}x_k\bar{x}_l = \sum_{s=1}^{n} \lambda_s |\xi_s|^2,$$

y en lo demás, sigue siendo válido el teorema 2.

Nota 2. Señalemos que la demostración de que los valores propios de un operador lineal autoconjugado (hermítico) A en R_n (real o complejo) son reales se puede hacer del modo siguiente: Sea λ un valor propio del operador A y sea x^0 ($|x^0|=1$) el vector propio correspondiente. Como $Ax^0=\lambda x^0$, se tiene,

$$\lambda=\lambda(x^0, x^0)=(\lambda x^0, x^0)=(Ax^0, x^0)=(x^0, Ax^0)=$$
$$=(x^0, \lambda x^0)=\overline{\lambda}(x^0, x^0)=\overline{\lambda}.$$

Inmediatamente se demuestra también que los vectores propios del operador, pertenecientes a distintos valores propios, son ortogonales.

En efecto,

$$Ax^1=\lambda_1 x^1,\ Ax^2=\lambda_2 x^2\ (|x^1|=1,\ |x^2|=1,\ \lambda_1\neq\lambda_2),$$

y, por lo tanto,

$$\lambda_1(x^1, x^2)=(\lambda_1 x^1, x^2)=(Ax^1, x^2)=(x^1, Ax^2)=$$
$$=(x^1, \lambda_2 x^2)=\overline{\lambda}_2(x^1, x^2)=\lambda_2(x^1, x^2).$$

Ahora bien, como $\lambda_1\neq\lambda_2$, se tiene,

$$(x^1, x^2)=0.$$

§ 23. Forma cuadrática en el espacio bidimensional

Para $n=2$ la forma cuadrática se escribe así:

$$a_{11}x_1^2+a_{12}x_1x_2+a_{21}x_2x_1+a_{22}x_2^2=a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2, \qquad (1)$$

ya que $a_{12}=a_{21}$. Suponemos que los números a_{kl} son reales.

Para reducir la forma (1) a la suma de los cuadrados de las coordenadas del vector (ξ_1, ξ_2) en una base (x^1, x^2), hay que hallar (véase el § 22) los vectores unitarios básicos x^1, x^2, que son los vectores propios del operador autoconjugado A, engendrado por la matriz simétrica

$$A=\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}.$$

Señalemos un método para hallar los valores propios y los vectores propios del operador A, distinto del método del § 22.

Así pues, si λ_0 es un valor propio del operador A y $x^0=(x_1^{(0)}, x_2^{(0)})\neq$ $\neq 0$ es el vector propio correspondiente, entonces

$$Ax^0=\lambda_0 x^0.$$

Escribamos esta ecuación en coordenadas:

$$\left.\begin{aligned}(a_{11}-\lambda_0)\,x_1^{(0)}+\qquad a_{12}x_2^{(0)}&=0,\\ a_{21}x_1^{(0)}+(a_{22}-\lambda_0)\,x_2^{(0)}&=0,\end{aligned}\right\}\tag{2}$$

o en forma de operadores

$$(A-\lambda_0E)\,\boldsymbol{x}^0=\mathbf{0},\tag{2'}$$

donde E es el operador idéntico.

Por lo tanto, el sistema homogéneo (2) tiene solución no nula $\boldsymbol{x}^0$, lo cual es posible si el determinante del sistema (2) ó (2′) es igual a cero:

$$\begin{vmatrix}a_{11}-\lambda_0 & a_{12}\\ a_{12} & a_{22}-\lambda_0\end{vmatrix}=|A-\lambda_0E|=0.$$

En resumen, el valor propio λ_0 es una raíz de la ecuación

$$|A-\lambda_0E|=0,\tag{3}$$

denominada *ecuación característica* del operador A (o de la forma cuadrática $(A\boldsymbol{x},\ \boldsymbol{x})$).

También es válido lo recíproco. Si λ_0 es una raíz de la ecuación (3), entonces toda solución no trivial del sistema

$$(A-\lambda_0E)\,\boldsymbol{x}=\mathbf{0}\tag{4}$$

es un vector propio del operador autoconjugado A.

Por consiguiente, en el caso dado los valores propios del operador A son las raíces de la ecuación cuadrática (3):

$$(a_{11}-\lambda)(a_{22}-\lambda)-a_{12}^2=0,$$

$$\lambda^2-(a_{11}+a_{22})\,\lambda+a_{11}a_{22}-a_{12}^2=0.$$

Resolviendo esta ecuación, obtenemos:

$$\left.\begin{aligned}\lambda_1&=\frac{1}{2}\left[a_{11}+a_{22}+\sqrt{4a_{12}^2+(a_{11}-a_{22})^2}\right],\\ \lambda_2&=\frac{1}{2}\left[a_{11}+a_{22}-\sqrt{4a_{12}^2+(a_{11}-a_{22})^2}\right].\end{aligned}\right\}\tag{5}$$

De aquí vemos que $\lambda_1\geqslant\lambda_2$, y además, $\lambda_1=\lambda_2$ si $a_{12}=0$, $a_{11}=a_{22}$. Vamos a suponer, para precisar, que $a_{11}\geqslant a_{22}$ (en caso contrario, cambiamos x_1 por x_2 y x_2 por x_1). Entonces,

$$\lambda_1\geqslant a_{11}\ \left(\lambda_1-a_{11}=\frac{1}{2}\left[\sqrt{4a_{12}^2+(a_{11}-a_{22})^2}-(a_{11}-a_{22})\right]\geqslant0.\right.$$

De (5) se deduce que los *valores propios del operador autoconjugado A son números reales.*

Ahora, una vez conocidos los valores propios λ_1 y λ_2, hallamos los vectores unitarios propios como solución del sistema (4). Como $|A - \lambda_1 E| = 0$, se tiene,

$$\text{rango }(A - \lambda_1 E) \leqslant 1.$$

Si $\lambda_1 = \lambda_2$, la matriz $A - \lambda_1 E$ consta sólo de ceros ($\lambda_1 = \lambda_2 = a_{11} = a_{22}$, $a_{12} = 0$), o sea, su rango es igual a cero. En este caso la forma cuadrática ya está reducida a la suma de cuadrados ($a_{12} = a_{21} = 0$). Cualquier vector $x = (x_1, x_2)$ satisface el sistema (4) Por ello, se pueden tomar por vectores propios los vectores unitarios del sistema de coordenadas $x^1 = i = (1, 0)$, $x^2 = j = (0, 1)$. Cualquier otro sistema de vectores ortonormales (x^1, x^2) cumple la propiedad de que en este sistema la forma cuadrática consta solamente de cuadrados.

Ahora bien, si $\lambda_1 > \lambda_2$, entonces, o $a_{12} \neq 0$, o bien, $a_{12} = 0$, $a_{11} \neq a_{22}$. El segundo caso se puede no estudiar, pues la forma (1) ya está reducida a la suma de cuadrados.

Así pues, supongamos que $a_{12} \neq 0$. Entonces,

$$\text{rango }(A - \lambda_1 E) = 1.$$

Por lo tanto, es suficiente considerar una ecuación del sistema (4):

$$(a_{11} - \lambda_1)\, x_1 + a_{12} x_2 = 0.$$

De aquí resulta ($a_{12} \neq 0$),

$$x_2 = [(-a_{11} + \lambda_1)/a_{12}]\, x_1$$

El vector

$$y^1 = \left(x_1, \frac{\lambda_1 - a_{11}}{a_{12}}\, x_1\right)$$

es solución del sistema (4). Normalizando este vector, obtenemos un vector propio

$$x^1 = \left(x_1^{(1)}, x_2^{(1)}\right) = \frac{y^1}{|y^1|} = \\ = \left(\frac{\pm 1}{\sqrt{1 + \left(\frac{\lambda_1 - a_{11}}{a_{12}}\right)^2}}, \frac{\pm(\lambda_1 - a_{11})}{a_{12}\sqrt{1 + \left(\frac{\lambda_1 - a_{11}}{a_{12}}\right)^2}}\right).$$

Haciendo transformaciones elementales se pueden obtener las igualdades

$$\left.\begin{aligned} x_1^{(1)} &= \pm \frac{1}{\sqrt{2}} \sqrt{1 + (a_{11} - a_{22})/\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}}, \\ x_2^{(1)} &= \pm \frac{\operatorname{sign} a_{12}}{\sqrt{2}} \sqrt{1 - (a_{11} - a_{22})/\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}}, \end{aligned}\right\} \quad (6)$$

A continuación es suficiente tomar en las fórmulas (6) el signo $+$.

Similarmente, dado el valor propio λ_2, podemos hallar el vector propio x^2. Resulta que

$$x^2=(x_1^{(2)},\ x_2^{(2)})=(-x_2^{(1)},\ x_1^{(1)}).$$

Formemos ahora la matriz del operador (de la transformación ortogonal) Λ que transforma los versores (i, j) en los versores (x^1, x^2):

$$\Lambda=\begin{vmatrix} x_1^{(1)} & x_2^{(1)} \\ -x_2^{(1)} & x_1^{(1)} \end{vmatrix}$$

(en las filas figuran las coordenadas de las imágenes de los vectores unitarios básicos i, j mediante Λ, o sea, $x^1=x_1^{(1)}i+x_2^{(1)}j$, $x^2=$ $=-x_2^{(1)}i+x_1^{(1)}j$. Entonces, las coordenadas del vector (x_1, x_2) en el sistema (i, j) están relacionadas con las coordenadas (ξ_1, ξ_2) de este vector en el sistema (x^1, x^2) mediante las columnas de la matriz Λ:

$$\left.\begin{aligned} x_1&=x_1^{(1)}\xi_1-x_2^{(1)}\xi_2, \\ x_2&=x_2^{(1)}\xi_1+x_1^{(1)}\xi_2. \end{aligned}\right\} \tag{7}$$

Sustituyendo estos valores en la forma cuadrática (1) y teniendo en cuenta las fórmulas (5) y (6), obtenemos:

$$a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2=\lambda_1\xi_1^2+\lambda_2\xi_2^2. \tag{8}$$

El segundo miembro de esta igualdad se llama *forma canónica de la forma cuadrática*.

Si los números λ_1 y λ_2 son de un mismo signo, entonces se dice que la forma cuadrática es de *tipo elíptico*; si λ_1 y λ_2 son de distinto signo se dice que es de *tipo hiperbólico*, y si uno de los números λ_1 o λ_2 es igual a cero, se dice que es de *tipo parabólico*.

En (5) se observa que $\lambda_1\lambda_2=a_{11}a_{22}-a_{12}^2$. Por esto, el tipo de la forma (1) se puede determinar según el signo de la expresión $a_{11}a_{22}-a_{12}^2$ (del discriminante de la forma cuadrática).

La forma cuadrática será de tipo *elíptico*, *hiperbólico* o *parabólico* según que la expresión $a_{11}a_{22}-a_{12}^2$ (el discriminante) sea mayor, menor o igual a cero.

EJEMPLO 1. Reducir la forma cuadrática

$$x_1^2-\sqrt{3}\,x_1x_2+2x_2^2.$$

a la forma canónica.

En este caso, $a_{11}=1$, $a_{12}=-\frac{\sqrt{3}}{2}$, $a_{22}=2$. Como $a_{11}a_{22}-a_{12}^2=$ $=2-\frac{3}{4}=\frac{5}{4}>0$, la forma es elíptica. Busquemos los vectores propios y sus valores propios según las fórmulas (5) y (6):

$$x_1^{(1)}=\frac{1}{2},\ x_2^{(1)}=-\frac{\sqrt{3}}{2},\ x^1=\left(\frac{1}{2},\ -\frac{\sqrt{3}}{2}\right),\ \lambda_1=\frac{5}{2}.$$

Por otra parte,

$$x^2=\left(\frac{\sqrt{3}}{2},\frac{1}{2}\right),\quad \lambda_2=\frac{1}{2}.$$

En el sistema (x^1, x^2) la forma cuadrática en cuestión toma la forma

$$\frac{5}{2}\xi_1^2+\frac{1}{2}\xi_2^2.$$

Ahora bien, como $x_1^{(1)}=\frac{1}{2}=\cos\left(-\frac{\pi}{3}\right)$, $x_2^{(1)}=-\frac{\sqrt{3}}{2}=$ $=\operatorname{sen}\left(-\frac{\pi}{3}\right)$, la transformación mediante la matriz

$$\Lambda=\begin{Vmatrix} x_1^{(1)} & -x_2^{(1)} \\ x_2^{(1)} & x_1^{(1)} \end{Vmatrix} (x^1=(x_1^{(1)},\ x_2^{(1)}),\ x^2=(-x_2^{(1)},\ x_1^{(1)}))$$

representa una rotación del sistema (x_1, x_2) un ángulo $\alpha=\pi/3$ en torno del origen de coordenadas en sentido del movimiento de las agujas del reloj (véase el ejemplo 1 al final del § 16).

§ 24. Curva de segundo orden

Supongamos que en el plano, respecto de un sistema rectangular de coordenadas (x, y), se ha dado una curva determinada implícitamente por una ecuación de segundo grado

$$Ax^2+2Bxy+Cy^2+2Dx+2Ey+F=0, \tag{1}$$

donde A, B, C, D, E, F son unos números reales dados. Se supone que los números A, B, C no son simultáneamente iguales a cero. Esta curva se llama *curva de segundo orden*. En realidad, puede ocurrir que no haya puntos (x, y) con coordenadas reales que satisfagan la ecuación (1). En este caso, se dice que la ecuación (1) determina una *curva imaginaria de segundo orden*. Aquí no vamos a estudiar este caso. La ecuación

$$x^2+y^2=-1$$

puede servir como ejemplo de ecuación de segundo grado que determina una curva imaginaria.

Enunciemos los seis casos más importantes de la ecuación general (1):

1) Ecuación de la *elipse*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad (a\geqslant b>0)$$

con los semiejes de longitudes a y b. En particular, si $a = b$, resulta la ecuación de la *circunferencia*

$$x^2 + y^2 = a^2$$

de radio a y con el centro en el origen de coordenadas.

2) Ecuación de la *hipérbola*

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (a\geqslant b>0)$$

con los semiejes a y b.

3) Ecuación de la *parábola*

$$y^2 = 2px \quad (p > 0).$$

4) Ecuación de *un par de rectas que se cortan*

$$a^2x^2 - b^2y^2 = 0 \quad (0 < a,\ b).$$

5) Ecuación de *un par de rectas paralelas o coincidentes*

$$x^2 - a^2 = 0 \quad (a \geqslant 0).$$

6) Ecuación que determina un *punto*

$$x^2 + y^2 = 0.$$

Detengámonos brevemente en las curvas mencionadas.

LA ELIPSE

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad (a\geqslant b>0). \tag{2}$$

Si $a = b$ la elipse (2) se convierte en una circunferencia de radio a con el centro en el origen de coordenadas, o sea, en el lugar geométrico de puntos cuyas distancias al origen son iguales a a.

Sea $a > b$. Hagamos $c = \sqrt{a^2 - b^2}$. Señalemos en el eje x los puntos F_1, F_2 de abscisas $x = -c$ y $x = c$. Estos son los *focos de la elipse*. La elipse (2) se puede definir como el lugar geométrico de puntos cuyas sumas de distancias a los focos F_1 y F_2 es una cantidad constante e igual a $2a$.

En efecto (fig. 36),

$$MF_1=\sqrt{(x+c)^2+y^2}, \quad MF_2=\sqrt{(x-c)^2+y^2},$$
$$2a=\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2},$$

de donde

$$2a - \sqrt{(x+c)^2+y^2} = \sqrt{(x-c)^2+y^2}$$

y

$$4a^2 + (x+c)^2 + y^2 - 4a\sqrt{(x+c)^2+y^2} = (x-c)^2+y^2,$$

$$4a^2 + 4cx = 4a\sqrt{(x+c)^2+y^2},$$

$$a^4 + 2a^2cx + c^2x^2 = a^2[x^2 + 2cx + c^2 + y^2],$$

$$-b^2x^2 = -a^2b^2 + a^2y^2,$$

$$a^2b^2 = b^2x^2 + a^2y^2,$$

de donde se deduce la ecuación (2). Efectuando estos cálculos en orden inverso, obtendremos que, si el punto (x, y) satisface la ecuación (2), entonces la suma de sus distancias hasta F_1 y F_2 es igual a $2a$.

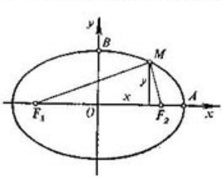

Fig. 36

Sustituyendo en la ecuación (2) x por $-x$, ésta no varía. Esto muestra que la elipse (2) es una curva simétrica respecto del eje y. Similarmente, la elipse (2) es simétrica respecto del eje x, ya que su ecuación no varía al sustituir y por $-y$. Pero, entonces, es suficiente estudiar su ecuación en el primer cuadrante del sistema de coordenadas, o sea, para x, $y \geqslant 0$. La parte de la elipse que está situada en el primer cuadrante se determina por la ecuación

$$y = \frac{b}{a}\sqrt{a^2-x^2}, \quad 0 \leqslant x \leqslant a.$$

De esta ecuación vemos que la elipse considerada pasa por los puntos $(0, b)$ y $(a, 0)$. Además, su ordenada y decrece continuamente cuando x crece continuamente sobre el segmento $[0, a]$.

La elipse es una curva limitada. Está situada en el interior de un circulo de radio a con el centro en el origen de coordenadas, pues, para las coordenadas (x, y) de cualquier punto de la elipse se verifica la desigualdad

$$x^2 + y^2 \leqslant a^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right) = a^2.$$

En la fig. 36 se observa que la elipse es una curva cerrada continua. En el primer cuadrante es una curva convexa hacia arriba. En cualquiera de sus puntos se puede trazar la tangente[1]. Todas

[1] Véase el § 4.2 del libro de Yu. S. Bugrov y S. M. Nikolski «Cálculo diferencial e integral», Editorial «Mir», 1984 o S. M. Nikolski «Curso de análisis matemático», t. I, Ed. URMO, S.A., Bilbao 1979.

estas propiedades y otras muchas pueden estudiarse con éxito por los métodos del análisis matemático que, a su vez, proporciona los medios para la definición exacta de los conceptos enunciados anteriormente como continuidad, convexidad, etc.

La ecuación de la elipse también puede escribirse en forma paramétrica

$$\left.\begin{array}{l} x=a\cos\theta, \\ y=b\operatorname{sen}\theta, \end{array}\right\} \quad (-\infty<\theta<\infty). \tag{3}$$

En efecto,

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\cos^2\theta+\operatorname{sen}^2\theta=1,$$

o sea, el punto (x, y) determinado por las igualdades (3) pertenece a la elipse (2) para cualquier θ. Si θ recorre continuamente el semiintervalo $[0, 2\pi)$, entonces el punto (x, y) recorre la elipse completa. Si θ sigue creciendo, el movimiento se repetirá continuamente.

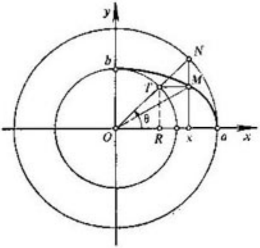

Fig. 37

Averigüemos el significado del parámetro θ y señalemos de paso un método de construcción de la elipse (fig. 37). Tracemos dos circunferencias concéntricas de radios b y a ($b<a$) con centros en el punto O. Tracemos luego el radio vector que forme un ángulo θ con el eje x, y denotemos por T y N sus puntos de intersección con las circunferencias de radios b y a, respectivamente. Tracemos una recta desde el punto N, paralela al eje y, y otra recta desde el punto T, paralela al eje x. El punto de intersección M de estas rectas pertenece a la elipse. En efecto, sea x la abscisa del punto M e y la ordenada. Entonces (véase la fig. 37),

$$x=ON\cdot\cos\theta=a\cos\theta,$$
$$y=TR=OT\cdot\operatorname{sen}\theta=b\operatorname{sen}\theta,$$

o sea, el punto M verdaderamente está situado en la elipse (3) y el parámetro θ es el ángulo formado por el eje x y el rayo ON. Señalemos que θ no es el ángulo polar φ que forma el radio vector OM con el eje x ($\operatorname{tg}\varphi=\frac{b}{a}\operatorname{tg}\theta$). Por ejemplo, si $\varphi=\pi/4$, $a=\sqrt{3}$, $b=1$, entonces $\theta=\pi/3$; si $\varphi=0$, entonces $\theta=0$; si $\varphi=\pi/2$, entonces $\theta=\pi/2$.

LA HIPÉRBOLA

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (0<a,\ b). \tag{4}$$

Hagamos $c=\sqrt{a^2+b^2}$ y señalemos en el eje x los puntos F_1 y F_2, los *focos de la hipérbola* (4), que tienen las abscisas $x=-c$, $x=c$ (fig. 38).

La hipérbola (4) puede definirse también como el lugar geométrico de puntos $A = (x, y)$ cuyas diferencias de distancias a los focos F_1 y F_2 es una cantidad constante e igual a $2a$.

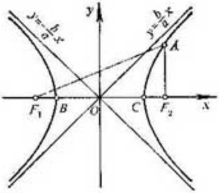

Fig. 38

Se tiene (véase la fig. 38),

$$AF_1 - AF_2 = \sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2} = 2a,$$
$$\sqrt{(x+c)^2+y^2} = \sqrt{(x-c)^2+y^2} + 2a,$$
$$(x+c)^2+y^2 = (x-c)^2+y^2+4a\sqrt{(x-c)^2+y^2}+4a^2,$$
$$4cx - 4a^2 = 4a\sqrt{(x-c)^2+y^2},$$
$$c^2x^2 - 2a^2cx + a^4 = a^2(x^2 - 2cx + c^2) + a^2y^2,$$
$$(a^2+b^2)x^2 = a^2x^2 + a^2b^2 + a^2y^2,$$
$$b^2x^2 - a^2y^2 = a^2b^2,$$

de donde se deduce la ecuación (4).

Hemos obtenido la rama derecha de la hipérbola (véase la fig. 38). Para obtener la rama izquierda hay que comenzar con la igualdad

$$AF_2 - AF_1 = 2a.$$

Mediante unos razonamientos efectuados en orden inverso, partiendo de la ecuación (4), se puede llegar a la conclusión de que los puntos (x, y) que satisfacen esta ecuación pertenecen al lugar geométrico anteriormente indicado.

Por la fórma de la ecuación (4) sacamos la conclusión que la hipérbola (4) es simétrica con respecto al eje x y al eje y. La parte de la hipérbola que está situada en el primer cuadrante tiene la ecuación

$$y = \frac{b}{a}\sqrt{x^2-a^2} \quad (a \leqslant x < \infty). \tag{5}$$

Vemos que la hipérbola en cuestión pasa por el punto $(a, 0)$ y la ordenada y crece y tiende al infinito cuando x crece en el semiintervalo $[a, \infty)$. Los puntos $B = (-a, 0)$ y $C = (a, 0)$, en los que la hipérbola se corta con el eje x, se llaman *vértices de la hipérbola*.

En la fig. 38 están representadas dos rectas:

$$y = \pm \frac{b}{a} x.$$

Estas son las *asíntotas* de la hipérbola en cuestión.

Supongamos dada una curva $y = f(x)$ en el semiintervalo $[a, \infty)$ (o en $(-\infty, a]$). Se dice que la recta $y = mx + n$ es la *asíntota* de esta curva cuando $x \to +\infty$ $(x \to -\infty)$, si

$$\lim_{x \to +\infty} [f(x) - mx - n] = 0$$

$\left(\lim\limits_{x \to -\infty} [f(x) - mx - n] = 0, \right.$ respectivamente).

Consideremos la parte de la hipérbola determinada por la igualdad (5), y comparémosla con la recta $y = \frac{b}{a} x$. Se tiene.

$$\lim_{x \to +\infty} \left[\frac{bx}{a} - \frac{b}{a} \sqrt{x^2 - a^2} \right] = \lim_{x \to +\infty} \frac{b}{a} \frac{a^2}{x + \sqrt{x^2 - a^2}} = 0.$$

Esto muestra que la recta $y = \frac{b}{a} x$ es la asíntota de la parte considerada de la hipérbola cuando $x \to +\infty$. Pero, entonces, se dice que esta recta es la asíntota (de toda) la hipérbola cuando $x \to +\infty$. En virtud de la simetría de la hipérbola respecto de los ejes, así como la simetría del par de rectas $y = \pm \frac{b}{a} x$ respecto de los ejes, se puede decir que ambas rectas son asíntotas de la hipérbola tanto cuando $x \to +\infty$ como cuando $x \to -\infty$.

La rama derecha de la hipérbola (4) puede expresarse en la forma paramétrica

$$\left. \begin{aligned} x &= a \operatorname{ch} u = \frac{a}{2} (e^u + e^{-u}), \\ y &= b \operatorname{sh} u = \frac{b}{2} (e^u - e^{-u}), \end{aligned} \right\} \quad (-\infty < u < \infty). \tag{6}$$

En efecto, como

$$\operatorname{ch}^2 u - \operatorname{sh}^2 u = 1, \tag{7}$$

de las ecuaciones (6), obtenemos

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \operatorname{ch}^2 u - \operatorname{sh}^2 u = 1.$$

La mitad superior de la rama derecha de la hipérbola corresponde a la variación de $u \in [0, \infty)$, y la inferior, a la variación de $u \in \in (-\infty, 0]$.

Veamos cómo está relacionado el parámetro u con el parámetro θ en la ecuación paramétrica de la elipse, y señalemos de paso un método de construcción de la hipérbola mediante la regla y el compás. Como este método de construcción de la hipérbola va a estar basado en el método de construcción de la elipse, expondremos simultáneamente ambos métodos (fig. 39). Nos limitaremos a la construcción de las partes de la elipse (2) y de la hipérbola (6) que están situadas en el primer cuadrante. Tracemos dos circunferencias concéntricas de radios a y b con centro en el origen de coordenadas. Tracemos un rayo desde el origen de coordenadas que forme un ángulo θ_0 con el eje x. Sean T_1 y N_1 los puntos de intersección de este rayo con las circunferencias señaladas ($OT_1 = b$, $ON_1 = a$). Trazando desde los puntos T_1 y N_1 dos rectas paralelas a los ejes x o y, respectivamente, obtendremos el punto de su intersección $M_e = (x_0, y_0)$ perteneciente a la elipse (2). Tracemos ahora el rayo OM_e. Sea N_2 el punto de intersección de este rayo con la circunferencia de radio a, y sea P el punto de intersección de este rayo con la recta paralela al eje y que pasa por el punto de la elipse $A = (a, 0)$. La ecuación del rayo OP se puede escribir así:

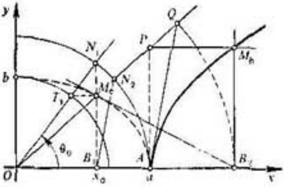

Fig. 39

$$y = \frac{y_0}{x_0}\, x.$$

De aquí se deduce que la ordenada del punto P es igual a $Y_0 = \frac{ay_0}{x_0}$. Unamos ahora el punto $B_1 = (x_0, 0)$ con el punto N_2 y tracemos por el punto A una recta paralela a B_1N_2, que se cortará con el rayo OP en el punto Q. Por la semejanza de los triángulos OAQ y OB_1N_2, obtenemos que $OQ = a^2/x_0$. Con el radio OQ señalamos sobre el eje x el punto $B_2 = (a^2/x_0, 0)$.

Tracemos ahora desde los puntos B_2 y P dos rectas paralelas a los ejes y y x, respectivamente. El punto de intersección de estas rectas $M_h = (X_0, Y_0)$, donde $X_0 = a^2/x_0$, pertenece a la hipérbola (4).

En efecto, como el punto (x_0, y_0) está situado en la elipse (2), se tiene,

$$\frac{X_0^2}{a^2} - \frac{Y_0^2}{b^2} = \frac{a^4}{a^2x_0^2} - \frac{a^2y_0^2}{x_0^2b^2} = \frac{a^2}{x_0^2}\left(1 - \frac{y_0^2}{b^2}\right) = \frac{a^2}{x_0^2}\cdot\frac{x_0^2}{a^2} = 1,$$

o sea, el punto M_h pertenece a la hipérbola (4).

Señalemos que el punto $B_2 = (a^2/x_0, 0)$ es el punto de intersección de la tangente a la elipse en el punto M_e con el eje x (véase la llamada en la pág. 136).

Así pues, a cada punto (x, y) de la elipse (2) le corresponde un punto completamente determinado (X, Y) de la hipérbola (4) y, recíprocamente.

Ahora bien, si la elipse (2) viene dada en forma paramétrica, entonces

$$x = a\cos\theta, \quad y = b \operatorname{sen}\theta.$$

Por lo tanto,

$$X = \frac{a^2}{x} = \frac{a}{\cos\theta}, \quad Y = \frac{ay}{x} = b\operatorname{tg}\theta.$$

De aquí, teniendo en cuenta (6), obtenemos

$$\operatorname{ch} u = \frac{1}{\cos\theta}, \quad \operatorname{sh} u = \operatorname{tg}\theta.$$

También son válidas las siguientes fórmulas:

$$\operatorname{tg}\frac{\theta}{2} = \sqrt{\frac{1-\cos\theta}{1+\cos\theta}} = \sqrt{\frac{\operatorname{ch} u - 1}{\operatorname{ch} u + 1}} = \operatorname{th}\frac{u}{2};$$

$$e^u = \operatorname{ch} u + \operatorname{sh} u = \frac{1}{\cos\theta} + \operatorname{tg}\theta = \frac{1+\operatorname{sen}\theta}{\cos\theta} =$$

$$= \frac{\left(\cos\frac{\theta}{1} + \operatorname{sen}\frac{\theta}{2}\right)^2}{\cos^2\frac{\theta}{2} - \operatorname{sen}^2\frac{\theta}{2}} = \frac{\cos\frac{\theta}{2} + \operatorname{sen}\frac{\theta}{2}}{\cos\frac{\theta}{2} - \operatorname{sen}\frac{\theta}{2}} = \frac{\operatorname{sen}\left(\frac{\pi}{4} + \frac{\theta}{2}\right)}{\cos\left(\frac{\pi}{4} + \frac{\theta}{2}\right)} = \operatorname{tg}\left(\frac{\pi}{4} + \frac{\theta}{2}\right).$$

o sea,

$$u = \ln \operatorname{tg}\left(\frac{\pi}{4} + \frac{\theta}{2}\right).$$

LA PARÁBOLA

$$y^2 = 2px \quad (p > 0). \tag{8}$$

Señalemos en el eje x el punto F de abscisa $x = p/2$, denominado *foco de la parábola* (8), y tracemos la recta $x = -p/2$, denominada *directriz de la parábola* (8) (fig. 40).

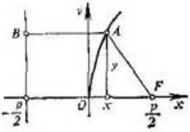

Fig. 40

La parábola puede definirse también como el lugar geométrico de puntos $A = (x, y)$ equidistantes del foco y de la directriz. En efecto (véase la fig. 40),

$$AF^2 = \left(x - \frac{p}{2}\right)^2 + y^2,$$

$$AB^2 = \left(x + \frac{p}{2}\right)^2,$$

y, por consiguiente,

$$\left(x - \frac{p}{2}\right)^2 + y^2 = \left(x + \frac{p}{2}\right)^2$$

$$-px + y^2 = px,$$

o sea,

$$y^2 = 2px.$$

Recíprocamente, de esta ecuación se deduce que los puntos que la satisfacen pertenecen al lugar geométrico de puntos indicado.

En la ecuación (8) se observa que la parábola es simétrica respecto del eje x. Su mitad superior tiene la ecuación

$$y = \sqrt{2px} \quad (0 \leqslant x < \infty), \tag{9}$$

la cual muestra que, cuando x recorre el semiintervalo $[0, \infty)$ creciendo, la ordenada y crece de 0 a ∞.

Señalemos un método sencillo para la construcción de la parábola (9) mediante la regla y el cartabón o mediante la regla y el compás. Tracemos la recta $x = -2p$ (fig. 41). Tomemos sobre esta recta un punto arbitrario $K = (-2p, y)$, $y > 0$. Unamos este punto con el origen de coordenadas mediante una recta y tracemos una recta que pase por el origen de coordenadas y sea perpendicular a la recta OK. Tracemos ahora una recta por el punto K, paralela al eje x. Las dos últimas rectas se cortan en un punto $M = (x, y)$ que pertenece a la parábola (9), ya que $OA = y$ es la media geométrica de los números $2p$ y x ($y = \sqrt{2px}$).

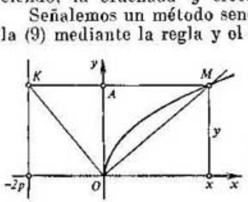

Fig. 41

La parábola no tiene asíntotas[1].

El par de rectas que se cortan

$$a^2x^2 - b^2y^2 = (ax - by)(ax + by) = 0 \quad (0 < a, b). \tag{10}$$

Si algún punto (x, y) satisface la ecuación (10), entonces satisface una de las ecuaciones

$$\left.\begin{aligned} ax - by = 0, \\ ax + by = 0 \end{aligned}\right\} \tag{10'}$$

o ambas. Recíprocamente, si un punto (x, y) satisface una de las ecuaciones (10'), entonces también satisface la ecuación (10). En este sentido se dice que (10) es la *ecuación de un par de rectas.*

Más adelante se demostrará que existe un sistema rectangular de coordenadas tal que en el mismo, la curva (1), si no es imaginaria, tiene una de las ecuaciones mencionadas anteriormente 1) — 6).

Más detalladamente:

[1]) Véase el § 4.20 del libro de Yu. S. Bugrov y S. M. Nikolski «Cálculo diferencial e integral», Editorial «Mir», 1984, o S. M. Nikolski «Curso de análisis matemático», t. I, Ed. URMO, S.A., Bilbao, 1979.

si $AC - B^2 > 0$ la curva (1) es una elipse, un punto (casos 1), 6)) o una curva imaginaría;

si $AC - B^2 < 0$ la curva (1) es una hipérbola o un par de rectas (distintas) que se cortan (casos 2), 4));

si $AC - B^2 = 0$, la curva (1) es una parábola, un par de rectas paralelas o coincidentes o una curva imaginaria (casos 3), 5)).

Nos permitimos hablar de «curvas» incluso en los casos 4), 5), 6), cuando se trata de un par de rectas o de un conjunto compuesto por un punto.

Así pues, sea dada la ecuación

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \tag{1}$$

donde los coeficientes A, B, C no son simultáneamente iguales a cero.

Sin restringir generalidad se puede suponer que $A \geqslant 0$, $A \geqslant C$, $B \geqslant 0$. Siempre se puede conseguir esto mediante las transformaciones ortogonales

$$\left.\begin{array}{l} x = \eta, \\ y = \xi, \end{array}\right\} \quad \left.\begin{array}{l} x = -\xi, \\ y = \eta \end{array}\right\}$$

y multiplicando ambos miembros de (1) por -1.

Si $B = 0$, $A \geqslant C > 0$, entonces (1) se puede escribir en la forma

$$A\left(x + \frac{D}{A}\right)^2 + C\left(y + \frac{E}{C}\right)^2 + F - \frac{E^2}{C} - \frac{D^2}{A} = 0. \tag{11}$$

La traslación paralela

$$\xi = x + \frac{D}{A}, \quad \eta = y + \frac{E}{C}$$

trasforma la ecuación (11) del modo siguiente:

$$A\xi^2 + C\eta^2 = \frac{E^2}{C} + \frac{D^2}{A} - F. \tag{11'}$$

Si $\frac{E^2}{C} + \frac{D^2}{A} - F > 0$, entonces, la ecuación (11') representa la ecuación de una elipse (caso 1)) con los semiejes a, b, siendo

$$a^2 = \left(\frac{E^2}{C} + \frac{D^2}{A} - F\right) \Big/ A, \quad b^2 = \left(\frac{E^2}{C} + \frac{D^2}{A} - F\right) \Big/ C.$$

Señalemos que en el caso dado $AC - B^2 = AC > 0$.

Si el segundo miembro de la ecuación (11') es igual a cero, entonces obtenemos un punto (caso 6)).

Si el segundo miembro de la ecuación (11') es negativo, entonces resulta una curva imaginaria.

Si $C = 0$, $A > 0$, entonces (1) puede escribirse en la forma

$$A\left(x+\frac{D}{A}\right)^2+2Ey+F-\frac{D^2}{A}=0. \tag{12}$$

Sea $E \neq 0$, entonces la traslación paralela

$$\xi=x+\frac{D}{A}, \quad \eta=y+\frac{F}{2E}-\frac{D^2}{2AE}$$

transforma la ecuación (12) en la ecuación

$$A\xi^2+2E\eta=0, \tag{12'}$$

que (después de la sustitución, si fuese necesario, de η por $-\eta$) representa la ecuación de una parábola (caso 3)).

Si $E = 0$, entonces, según cual sea el signo de $\frac{D^2}{A}-F$, obtenemos un par de rectas paralelas o una curva imaginaria. Señalemos que aquí $AC - B^2 = 0$.

Por otra parte, si $C < 0$, $A > 0$, entonces la ecuación (1) puede escribirse así:

$$A\left(x+\frac{D}{A}\right)^2-|C|\left(y-\frac{E}{|C|}\right)^2+F+\frac{E^2}{|C|}-\frac{D^2}{A}=0, \tag{13}$$

cuyo análisis se realiza igual que en el caso de la ecuación (11). La ecuación (13) representa una hipérbola o un par de rectas que se cortan (casos 2) y 4)). Señalemos que, en este caso, $AC - B^2 = AC < 0$.

El caso $A = 0$, $C < 0$ se reduce a una ecuación del tipo (12).

En resumen, si $B = 0$ la ecuación (1) representa siempre uno de los casos particulares 1) — 6)).

Supongamos ahora que $B > 0$, $A \geqslant C$. Entonces, como ya sabemos, (véase el § 23), existe una transformación ortogonal

$$\left.\begin{aligned} x&=x_1\xi-y_1\eta,\\ y&=y_1\xi+x_1\eta, \end{aligned}\right\} \tag{14}$$

donde

$$x_1=\sqrt{\frac{1}{2}+\frac{A-C}{2\sqrt{4B^2+(A-C)^2}}},$$

$$y_1=\sqrt{\frac{1}{2}-\frac{A-C}{2\sqrt{4B^2+(A-C)^2}}},$$

que reduce la forma cuadrática

$$Ax^2+2Bxy+Cy^2$$

a la forma canónica.

Transformemos la ecuación (1) mediante (14):

$$\lambda_1\xi^2 + \lambda_2\eta^2 + D\,(x_1\xi - y_1\eta) + E\,(y_1\xi + x_1\eta) + F = 0, \qquad (15)$$

donde

$$\lambda_1 = \frac{1}{2}\left[A + C + \sqrt{4B^2 + (A-C)^2}\right],$$

$$\lambda_2 = \frac{1}{2}\left[A + C - \sqrt{4B^2 + (A-C)^2}\right]$$

($\lambda_1 > \lambda_2$, $\lambda_1\lambda_2 = AC - B^2$). Escribamos la ecuación (15) en la forma

$$\lambda_1\xi^2 + \lambda_2\eta^2 + (x_1D + y_1E)\,\xi + (x_1E - y_1D)\,\eta + F = 0. \qquad (15')$$

La ecuación (15′) es un caso particular de la ecuación (1) para $B = 0$, que ya hemos estudiado.

Por lo tanto, podemos decir que, si

1) $AC - B^2 = \lambda_1\lambda_2 > 0$, entonces la ecuación (1) representa una elipse, un punto o una curva imaginaria. En este caso se dice que la ecuación (1) pertenece al *tipo elíptico*;

2) $AC - B^2 = \lambda_1\lambda_2 < 0$, entonces la ecuación (1) representa una hipérbola o un par de rectas que se cortan. En este caso se dice que la ecuación (1) pertenece al *tipo hiperbólico*;

3) $AC - B^2 = \lambda_1\lambda_2 = 0$, entonces la ecuación (1) representa una parábola, un par de rectas paralelas o una curva imaginaria. En este caso se dice que la ecuación (1) pertenece al *tipo parabólico*.

Ejemplo 1. Estudiar el carácter de la curva

$$2x^2 + \sqrt{3}\,xy + y^2 + 2x + 2\sqrt{3}y + F = 0,$$

donde F es un número real arbitrario.

En el caso dado $A = 2 > C = 1$, $B = \frac{\sqrt{3}}{2} > 0$, $AC - B^2 = \frac{5}{4} > 0$, o sea, la ecuación pertenece al tipo elíptico. Es fácil calcular (véase el ejemplo del § 23) que

$$x_1 = \frac{\sqrt{3}}{2}, \quad y_1 = \frac{1}{2}, \quad \lambda_1 = \frac{5}{2}, \quad \lambda_2 = \frac{1}{2}.$$

Por lo tanto, mediante la transformación ortogonal

$$x = \frac{1}{2}(\sqrt{3}\xi - \eta), \quad y = \frac{1}{2}(\xi + \sqrt{3}\,\eta)$$

la ecuación considerada se escribe en la forma

$$\frac{5}{2}\xi^2 + \frac{1}{2}\eta^2 + (\sqrt{3}\,\xi - \eta) + \sqrt{3}(\xi + \sqrt{3}\,\eta) + F = 0$$

o bien,

$$\frac{5}{2}\left(\xi + \frac{2}{5}\sqrt{3}\right)^2 + \frac{1}{2}(\eta + 2)^2 = \frac{16}{5} - F.$$

Hagamos también una traslación paralela

$$u=\xi+\frac{2}{5}\sqrt{3},$$
$$v=\eta+2,$$

entonces, tendremos:

$$\frac{5}{2}u^2+\frac{1}{2}v^2=\frac{16}{5}-F. \qquad (16)$$

Si $\frac{16}{5}-F>0$, entonces (16) será la ecuación de una elipse con los semiejes a, b, donde

$$a^2=2\,(16-5F)/25, \quad b^2=2\,(16-5F)/5.$$

Si $\frac{16}{5}-F=0$, entonces la ecuación (16) representa un punto. Si $\frac{16}{5}-F<0$, entonces la ecuación (16) representa una curva imaginaria.

§ 25. Superficie de segundo orden en el espacio tridimensional

La ecuación

$$\sum_{k=1}^{3}\sum_{l=1}^{3}a_{kl}x_kx_l+2\sum_{l=1}^{3}A_lx_l+B=0, \qquad (1)$$

donde $a_{kl}=a_{lk}$, A_l, B son unas constantes dadas y $x=(x_1, x_2, x_3)$ es un punto variable en R_3, determina por lo general un conjunto de puntos en R_3, denominado *superficie de segundo orden*. Si la ecuación (1) no se satisface para ningún punto real $x=(x_1, x_2, x_3)$, entonces se dice que determina una *superficie imaginaria*. Estos casos no nos van a interesar. En algunos casos la ecuación (1) puede determinar un par de planos distintos o coincidentes o un punto único. A tales conjuntos también los llamaremos superficies.

He aquí los casos particulares más importantes de la ecuación (1):

1) *El elipsoide*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \quad (a,\ b,\ c>0).$$

2) *El hiperboloide de una hoja*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a,\ b,\ c>0).$$

3) *El hiperboloide de dos hojas*

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a,\ b,\ c>0).$$

4) *El paraboloide elíptico*

$$\frac{x^2}{p}+\frac{y^2}{q}=2z \quad (p,\ q>0).$$

5) *El paraboloide hiperbólico*

$$\frac{x^2}{p}-\frac{y^2}{q}=2z \quad (p,\ q>0).$$

6) *El cono de segundo orden*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \quad (a,\ b,\ c>0)$$

7) *Un punto*

$$x^2+y^2+z^2=0.$$

8) *Los cilindros de segundo orden*:
el cilindro elíptico

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad (a,\ b>0),$$

el cilindro hiperbólico

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (a,\ b>0),$$

el cilindro parabólico

$$y^2=2px \quad (p>0),$$

un par de planos que se cortan

$$a^2x^2-b^2y^2=0 \quad (a,\ b>0),$$

un par de planos paralelos o coincidentes

$$x^2-a^2=0 \quad (a>0),$$

$$z^2=0,$$

una recta

$$x^2+y^2=0.$$

Al considerar los casos particulares de la ecuación (1) se ha hecho $x=x_1,\ y=x_2,\ z=x_3$.

Se puede demostrar que para cada caso particular de la ecuación (1), si ésta no determina una superficie imaginaria, se puede hallar un sistema rectangular de coordenadas en el que esta ecuación tiene una de las ocho formas enunciadas anteriormente. Esto se deduce de la teoría general del § 22. La propia transformación de la ecuación (1) se realiza del mismo modo que en el § 24. La búsqueda de los valores propios λ_1, λ_2, λ_3 se reduce a la resolución de una ecuación cúbica.

Señalemos otro método para la búsqueda de los valores y vectores propios, que realmente ya lo habíamos visto en el § 23 en el caso bidimensional. Los valores propios λ_1, λ_2, λ_3 del operador autoconjugado A y los vectores normalizados propios correspondientes x^1, x^2, x^3 ($|x^j| = 1$, $Ax^j = \lambda_j x^j$, $j = 1, 2, 3$) se pueden hallar del modo siguiente (véase más adelante la fundamentación). Se considera el determinante ($a_{kl} = a_{lk}$)

$$D(\lambda) = |A - \lambda E| = \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix},$$

donde E es la matriz unidad. Hallamos las raíces λ_1, λ_2, λ_3 de la ecuación

$$|A - \lambda E| = 0, \tag{2}$$

denominada *ecuación característica* del operador A ($D(\lambda_j) = 0$, $j = 1, 2, 3$). Estos son los valores propios del operador A. Estos son reales, y pueden ser distintos, pero también pueden coincidir siendo entonces múltiples. Por lo tanto,

$$D(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda).$$

Después, para la raíz λ_1 se busca una solución no trivial $x = (x_1, x_2, x_3)$ del sistema homogéneo de ecuaciones:

$$\left.\begin{aligned} (a_{11}-\lambda_1)x_1 + a_{12}x_2 + a_{13}x_3 &= 0, \\ a_{12}x_1 + (a_{22}-\lambda_1)x_2 + a_{23}x_3 &= 0, \\ a_{31}x_1 + a_{32}x_2 + (a_{33}-\lambda_1)x_3 &= 0 \end{aligned}\right\} \tag{3}$$

o sea, de la ecuación correspondiente para el operador $A - \lambda_1 E$:

$$(A - \lambda_1 E)\, x = 0, \tag{3'}$$

donde E es la matriz unidad y $x = (x_1, x_2, x_3)$, $0 = (0, 0, 0)$ son vectores.

Si λ_1 es una raíz simple (o sea, que en este caso λ_1 es distinto de λ_2 y λ_3), entonces el rango de la matriz del sistema (3) necesariamente es igual a dos (rango $(A - \lambda_1 E) = 2$), y se obtiene un vector $x^1 = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$, que es único salvo el signo, y que satisface el sistema (3), o sea,

$$(A - \lambda_1 E)\, x^1 = 0$$

o bien

$$Ax^1 = \lambda_1 x^1.$$

Si λ_1 es una raíz de segundo orden ($\lambda_1 = \lambda_2 \neq \lambda_3$), entonces, la matriz del sistema (3) necesariamente es de rango uno (rango $(A - \lambda_1 E) = 1$), y el sistema tiene entonces dos soluciones ortonormalizadas x^1 y x^2 ($|x^1| = |x^2| = 1$, $(x^1, x^2) = 0$), que son dos vectores propios pertenecientes al valor propio λ_1:

$$Ax^j = \lambda_j x^j \quad (j = 1,\ 2,\quad \lambda_1 = \lambda_2).$$

Finalmente, si λ_1 es una raíz de tercer orden ($\lambda_1 = \lambda_2 = \lambda_3$), entonces la matriz del sistema (3) necesariamente es de rango cero (rango $(A - \lambda_1 E) = 0$), y el sistema tiene tres soluciones ortonormalizadas x^1, x^2, x^3:

$$Ax^j = \lambda_j x^j \quad (j = 1,\ 2,\ 3,\quad \lambda_1 = \lambda_2 = \lambda_3).$$

Tres vectores ortonormales cualesquiera en R_3 pueden tomarse entonces como vectores propios x^1, x^2, x^3 pertenecientes a los valores propios $\lambda_1 = \lambda_2 = \lambda_3$.

Veamos la fundamentación de lo expuesto. Ya sabemos que en R_3 existe un sistema de vectores ortonormales x^1, x^2, x^3 y tres números reales λ_1, λ_2, λ_3 tales que

$$Ax^j = \lambda_j x^j \quad (j = 1,\ 2,\ 3).$$

Además, siempre se puede señalar una matriz ortogonal Λ tal (véase § 22) que

$$\Lambda A \Lambda^{-1} = \begin{Vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{Vmatrix}.$$

De aquí que, para la variable λ se verifica la identidad

$$\Lambda (A - \lambda E) \Lambda^{-1} = \begin{Vmatrix} \lambda_1 - \lambda & 0 & 0 \\ 0 & \lambda_2 - \lambda & 0 \\ 0 & 0 & \lambda_3 - \lambda \end{Vmatrix} \tag{4}$$

$$(\Lambda (A - \lambda E) \Lambda^{-1} = \Lambda A \Lambda^{-1} - \Lambda \lambda E \Lambda^{-1} = \Lambda A \Lambda^{-1} - \lambda \Lambda \Lambda^{-1} = \\ = \Lambda A \Lambda^{-1} - \lambda E).$$

El determinante de la matriz $A - \lambda E$ lo hemos denotado anteriormente por $D(\lambda)$. Este, como se observa en (4), es igual al determinante de la matriz que figura en el último miembro en (4), pues

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) = |\Lambda (A - \lambda E) \Lambda^{-1}| = \\ = |\Lambda|\,|A - \lambda E|\,|\Lambda^{-1}| = |A - \lambda E| = D(\lambda).$$

Por lo tanto,

$$D(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda).$$

Resulta, pues, que las raíces del polinomio $D(\lambda)$ coinciden con los valores propios λ_1, λ_2, λ_3 del operador A. Por consiguiente, estas raíces son *reales*.

Sea λ_1 una raíz simple y, por consiguiente, $\lambda_1 \neq \lambda_2$, $\lambda_1 \neq \lambda_3$. En este caso, la matriz del segundo miembro en (4) para $\lambda = \lambda_1$ es de rango 2 (rango $\Lambda(A - \lambda_1 E)\Lambda^{-1} = 2$); pero, entonces, rango $(A - \lambda_1 E) = 2$. Hay que tener en cuenta que las soluciones del sistema homogéneo (3) y las del sistema

$$\left.\begin{aligned} 0\cdot z_1 + 0\cdot z_2 + 0\cdot z_3 &= 0,\\ 0\cdot z_1 + (\lambda_2 - \lambda_1)z_2 + 0\cdot z_3 &= 0,\\ 0\cdot z_1 + 0\cdot z_2 + (\lambda_3 - \lambda_1)z_3 &= 0, \end{aligned}\right\} \qquad (5)$$

correspondiente a la matriz (4), se transforman unas en otras mediante una matriz ortogonal (los sistemas (3) y (5) son equivalentes). Ahora bien, el sistema (5) sólo tiene una solución normalizada, salvo el signo $(\pm 1, 0, 0)$. Pero esto sólo es posible si rango $(A - \lambda_1 E) = 2$.

Se puede dar una explicación a esto. Si se supone que todos los determinantes de segundo orden engendrados por la matriz $A - \lambda_1 E$ son iguales a cero, entonces, todos los determinantes de segundo orden engendrados por la matriz $\Lambda(A - \lambda_1 E)\Lambda^{-1}$ también serán iguales a cero, ya que estos determinantes son combinaciones lineales de determinantes de segundo orden de la matriz $A - \lambda_1 E$. Pero esto es imposible, ya que para un determinante engendrado por la matriz $\Lambda(A - \lambda_1 E)\Lambda^{-1}$, se tiene

$$\begin{vmatrix} \lambda_2 - \lambda_1 & 0 \\ 0 & \lambda_3 - \lambda_1 \end{vmatrix} \neq 0.$$

Por otra parte, si $\lambda_1 = \lambda_2 \neq \lambda_3$, entonces, razonando de un modo similar, obtenemos que rango $(A - \lambda_1 E) = 1$, y, entonces existen exactamente dos soluciones ortonormales x^1, x^2 del sistema (3) correspondientes a $\lambda_1 = \lambda_2$.

Finalmente, si $\lambda_1 = \lambda_2 = \lambda_3$ se tiene, rango $(A - \lambda_1 E) = 0$, o sea, que todos los elementos de la matriz $A - \lambda_1 E$ son iguales a cero. En este caso, cualquier vector $x = (x_1, x_2, x_3)$ es solución del sistema (3). Esto da lugar a que tres vectores arbitrarios x^1, x^2, x^3 que formen un sistema ortonormal serán vectores propios pertenecientes al valor propio $\lambda_1 = \lambda_2 = \lambda_3$.

Obsérvese que, en este caso, la forma cuadrática ya está reducida a la suma de cuadrados ($a_{12} = a_{13} = a_{23} = 0$, $a_{11} = a_{22} = a_{33} = \lambda_1$).

EJEMPLO 1. Reducir la forma cuadrática.

$$x_1^2 + x_2^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

a la forma canónica.

Aquí $a_{11} = a_{22} = a_{12} = a_{13} = a_{23} = 1$, $a_{33} = 0$.

Formemos la ecuación característica:

$$\begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = 0$$

o bien,

$$-\lambda (1-\lambda)^2 + 2 - 2(1-\lambda) + \lambda = 0, \quad -\lambda (1-\lambda)^2 + 3\lambda = 0.$$

Fácilmente se observa que $\lambda = 0$ es una raíz de esta ecuación. Hallemos las otras dos raíces:

$$3-(1-\lambda)^2=0, \quad (1-\lambda)^2=3, \quad 1-\lambda = \pm\sqrt{3}, \quad \lambda = 1\pm\sqrt{3}.$$

Así pues,

$$\lambda_1 = 1+\sqrt{3}, \quad \lambda_2 = 0, \quad \lambda_3 = 1-\sqrt{3},$$

o sea, ha resultado que $\lambda_1 > \lambda_2 > \lambda_3$.

Busquemos el vector propio x^1. Para ello, formamos el sistema (3):

$$\begin{aligned} -\sqrt{3}\,x_1 + x_2 + x_3 &= 0 \\ x_1 - \sqrt{3}\,x_2 + x_3 &= 0 \\ x_1 + x_2 - (1+\sqrt{3})\,x_3 &= 0 \end{aligned}$$

Dos ecuaciones cualesquiera de este sistema son linealmente independientes. Resolviendo el sistema formado por las dos primeras ecuaciones, obtenemos

$$x_1 = \frac{1+\sqrt{3}}{2}\,x_3, \quad x_2 = \frac{1+\sqrt{3}}{2}\,x_3.$$

Por lo tanto, el vector

$$y^1 = \left(\frac{1+\sqrt{3}}{2}\,x_3, \quad \frac{1+\sqrt{3}}{2}\,x_3, \; x_3\right)$$

es una solución del sistema y, normalizándolo, obtenemos el vector propio

$$x^1 = \frac{y^1}{|y^1|} = \left(\frac{1+\sqrt{3}}{2\sqrt{3+\sqrt{3}}}, \; \frac{1+\sqrt{3}}{2\sqrt{3+\sqrt{3}}}, \; \frac{1}{\sqrt{3+\sqrt{3}}}\right).$$

Busquemos ahora x^2 ($\lambda_2 = 0$):

$$\begin{aligned} x_1 + x_2 + \quad x_3 &= 0 \\ x_1 + x_2 + \quad x_3 &= 0 \\ x_1 + x_2 + 0 \cdot x_3 &= 0. \end{aligned}$$

Resolviendo el sistema formado por las dos ecuaciones últimas (pues el determinante de los coeficientes de x_2 y x_3 es distinto de cero), obtenemos, $x_2 = -x_1$, $x_3 = 0$.

El vector $y^2 = (x_1, -x_1, 0)$ es solución del sistema, y el vector

$$x^2 = \frac{y^2}{|y^2|} = \left(\frac{1}{\sqrt{2}},\ -\frac{1}{\sqrt{2}},\ 0\right)$$

es un vector propio unitario. Fácilmente se comprueba que éste es ortogonal a x^1 (el producto escalar de estos vectores es igual a cero).

Finalmente, para $\lambda_3 = 1 - \sqrt{3}$ hallamos el tercer vector propio

$$x^3 = \left(\frac{1}{\sqrt{6+2\sqrt{3}}},\ \frac{1}{\sqrt{6+2\sqrt{3}}},\ -\frac{1+\sqrt{3}}{\sqrt{6+2\sqrt{3}}}\right).$$

La matriz ortogonal de cambio de las coordenadas del vector $x = (x_1, x_2, x_3)$ en el sistema (i, j, k) a las coordenadas del vector $x = (\xi_1, \xi_2, \xi_3)$ en el sistema (x^1, x^2, x^3), tiene la forma

$$\Lambda = \begin{Vmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} \\ x_1^{(3)} & x_2^{(3)} & x_3^{(3)} \end{Vmatrix}$$

$$(|\Lambda| = 1,\ \Lambda^{-1}(i) = x^1,\ \Lambda^{-1}(j) = x^2,\ \Lambda^{-1}(k) = x^3),$$

o sea,

$$\left.\begin{aligned} x_1 &= x_1^{(1)}\xi_1 + x_1^{(2)}\xi_2 + x_1^{(3)}\xi_3, \\ x_2 &= x_2^{(1)}\xi_1 + x_2^{(2)}\xi_2 + x_2^{(3)}\xi_3, \\ x_3 &= x_3^{(1)}\xi_1 + x_3^{(2)}\xi_2 + x_3^{(3)}\xi_3. \end{aligned}\right\} \tag{6}$$

La transformación dada conserva la orientación (ya que $|\Lambda| = 1 > 0$), o sea, el sistema (x^1, x^2, x^3) tiene la misma orientación que el sistema inicial (i, j, k).

Sustituyendo x_ν, en la forma cuadrática en cuestión, por sus valores según las fórmulas (6), obtenemos su forma canónica

$$(1+\sqrt{3})\xi_1^2 + (1-\sqrt{3})\xi_3^2.$$

Detengámonos ahora solamente en el estudio más detallado de las ecuaciones y de las superficies que éstas representan para los ocho tipos indicados anteriormente.

El elipsoide

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \quad (a,\ b,\ c>0). \tag{7}$$

Si $a=b=c$, el elipsoide (7) se convierte en una esfera de radio a con el centro en el origen de coordenadas, o sea, en el lugar geométrico de puntos cuyas distancias al origen de coordenadas son iguales a a

Las magnitudes a, b, c se llaman *semiejes del elipsoide.*

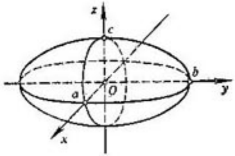

Fig. 42

Si en la ecuación (7) sustituímos (simultáneamente o por separado) x por $-x$, y por $-y$, z por $-z$, entonces la ecuación no varía, lo que muestra que el elipsoide (7) es una superficie simétrica respecto de los planos coordenados $x=0$, $y=0$, $z=0$ y respecto del origen de coordenadas. Por esto, es suficiente estudiar la ecuación del elipsoide (7) en el primer octante del sistema de coordenadas, o sea, para $x\geqslant 0$, $y\geqslant 0$, $z\geqslant 0$. La parte del elipsoide que está situada en el primer octante se determina por la ecuación explícita, por ejemplo,

$$z=c\sqrt{1-\frac{x^2}{a^2}-\frac{y^2}{b^2}}, \quad x\geqslant 0, \quad y\geqslant 0, \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}\leqslant 1.$$

Para precisar, vamos a suponer que $a\geqslant b\geqslant c$.

El elipsoide es una superficie limitada. Está situado en el interior de una bola de radio a con el centro en el origen de coordenadas, ya que para las coordenadas de cualquier punto del elipsoide $(x,\ y,\ z)$ se verifica la desigualdad

$$x^2+y^2+z^2\leqslant a^2\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}\right)=a^2\cdot 1=a^2.$$

Para tener una idea más exacta del elipsoide, consideremos las secciones de éste por planos paralelos a los planos coordenados. Por ejemplo, cortando el elipsoide por los planos $z=h$ $(-c\leqslant h\leqslant c)$ obtenemos en las secciones las elipses

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1-\frac{h^2}{c^2}$$

de semiejes

$$a\sqrt{1-\frac{h^2}{c^2}}, \quad b\sqrt{1-\frac{h^2}{c^2}}.$$

Vemos que la elipse más grande se obtiene en la sección del elipsoide por el plano $z=0$. Lo mismo ocurre en la sección por los planos $x=h$ $(-a\leqslant h\leqslant a)$, $y=h$ $(-b\leqslant h\leqslant b)$.

El elipsoide (7) tiene la forma representada en la fig. 42.

Los puntos $(\pm a,\ 0,\ 0)$, $(0\ \pm b,\ 0)$, $(0,\ 0,\ \pm c)$ están situados en el elipsoide (7) y se llaman *vértices*.

Si algún par de semiejes son iguales entre sí, entonces (7) es un *elipsoide de revolución*, o sea, se obtiene por la rotación de la elipse en torno del eje coordenado correspondiente.

El hiperboloide de una hoja

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a,\ b,\ c>0). \tag{8}$$

Observando la ecuación (8) sacamos la conclusión de que el hiperboloide de una hoja es una superficie simétrica respecto de los planos coordenados y respecto del origen de coordenadas. Los números a, b, c se llaman *semiejes del hiperboloide de una hoja*. Los puntos $(\pm a,\ 0,\ 0)$, $(0,\ \pm b,\ 0)$, situados en la superficie (8), se llaman *vértices* del hiperboloide de una hoja.

Hagamos una sección de la superficie (8) por el plano $z=h$, entonces, en la sección obtendremos una elipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1+\frac{h^2}{c^2}$$

con los semiejes

$$a\sqrt{1+\frac{h^2}{c^2}},\quad b\sqrt{1+\frac{h^2}{c^2}}.$$

Al variar h desde $-\infty$ hasta $+\infty$ esta elipse forma la superficie (8).

Si ahora hacemos una sección de la superficie (8) por el plano $x=h$ (o $y=h$), obtenemos en la sección la hipérbola

$$\frac{y^2}{b^2}-\frac{z^2}{c^2}=1-\frac{h^2}{a^2}\quad\left(\frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{h^2}{b^2}\right).$$

Si $h=\pm a$, la primera hipérbola se descompone en dos rectas $y=\pm\frac{b}{c}z$.

Si $|h|\leqslant a$, entonces, el eje real de simetría de la hipérbola correspondiente es una recta paralela al eje Oy, y si $|h|>a$, es una recta paralela al eje Oz.

Se llama *eje real de simetría* de la hipérbola al eje de simetría que se corta por la hipérbola.

Si $a=b$, entonces, en la sección de la superficie (8) por el plano $z=h$ resulta una circunferencia de radio $a\sqrt{1+(h^2/c^2)}$. En este caso, la superficie (8) se forma por la rotación de la hipérbola $\frac{x^2}{a^2}-\frac{z^2}{c^2}=1$ en torno del eje Oz. La forma general del hiperboloide de una hoja está representada en la fig. 43.

El hiperboloide de dos hojas

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a,\ b,\ c>0). \tag{9}$$

Como la ecuación (9) sólo contiene cuadrados de las variables, la superficie en cuestión es simétrica respecto de los planos $x=0$, $y=0$, $z=0$ y respecto del origen de coordenadas.

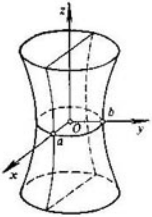

Fig. 43

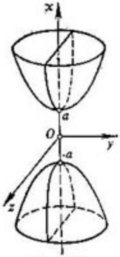

Fig. 44

Escribamos también la ecuación (9) en la forma

$$\frac{y^2}{b^2}+\frac{z^2}{c^2}=-1+\frac{x^2}{a^2}. \tag{9'}$$

De aquí queda claro que cortando la superficie (9′) por el plano $x=$ $=h$ ($|h|\geqslant a$), resulta en la intersección una elipse

$$\frac{y^2}{b^2}+\frac{z^2}{c^2}=-1+\frac{h^2}{a^2}$$

con los semiejes

$$b\sqrt{(h^2/a^2)-1}, \quad c\sqrt{(h^2/a^2)-1}.$$

Si $|h|<a$, entonces $(h^2/a^2)-1<0$, por lo cual, no hay puntos de intersección de la superficie (9′) con el plano $x=h$.

En la sección de la superficie (9) por el plano $z=h$ (o $y=h$), obtenemos la hipérbola

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1+\frac{h^2}{c^2} \quad \left(\frac{x^2}{a^2}-\frac{z^2}{c^2}=1+\frac{h^2}{b^2}\right).$$

Los puntos $(\pm a,\ 0,\ 0)$ están situados en la superficie (9) y se llaman *vértices del hiperboloide de dos hojas.* La superficie (9) viene representada en la fig. 44.

El paraboloide elíptico

$$\frac{x^2}{p}+\frac{y^2}{q}=2z \quad (p,\ q>0). \tag{10}$$

Como en (10) figuran los cuadrados de las variables x e y, la superficie dada es simétrica respecto de los planos coordenados $x=0$, $y=0$. Por otra parte, como se considera que $p,\ q>0$, la superficie (10) está situada en el semiespacio $z\geqslant 0$.

Cortando la superficie (10) por los planos $z=h$ $(h\geqslant 0)$, obtendremos en las secciones elipses

$$\frac{x^2}{p}+\frac{y^2}{q}=2h$$

de semiejes

$$\sqrt{2ph},\ \sqrt{2qh}.$$

Al variar h desde cero hasta ∞ las elipses dadas forman la superficie (10).

Cortando la superficie (10) por los planos $x=h$ (o $y=h$), obtenemos en las secciones parábolas

$$y^2=2q\left(z-\frac{h^2}{2p}\right) \quad \left(x^2=2p\left(z-\frac{h^2}{2q}\right)\right)$$

con los vértices desplazados en los puntos con las z-coordenadas $z=\frac{h^2}{2p}$ $\left(z=\frac{h^2}{2q}\right)$.

Si $p=q$, la superficie (10) resulta una superficie de revolución, que se obtiene por la rotación de la parábola $x^2=2pz$, en torno del eje Oz. En este caso, la superficie (10) se llama *paraboloide de revolución*.

El punto $(0,\ 0,\ 0)$ está situado en la superficie (10) y se llama *vértice del paraboloide elíptico*. El paraboloide elíptico está representado en la fig. 45.

El paraboloide hiperbólico

$$\frac{x^2}{p}-\frac{y^2}{q}=2z \quad (p,\ q>0). \tag{11}$$

A base de la ecuación (11) sacamos la conclusión de que la superficie dada es simétrica respecto de los planos coordenados $x=0$, $y=0$. Cortando la superficie (11) por los planos $z=h$, obtendremos en las secciones las hipérbolas

$$\frac{x^2}{p}-\frac{y^2}{q}=2h,$$

además, para $h>0$ el eje real de simetría de cada hipérbola es paralelo al eje Ox y para $h<0$ es paralelo al eje Oy. Para $h=0$, en la sección resultan dos rectas que se cortan.

En las secciones de la superficie (11) por los planos $x = h$ o $y = h$ obtenemos parábolas, cuyas ramas están dirigidas hacia abajo o hacia arriba, respectivamente:

$$-\frac{y^2}{q} = 2z - \frac{h^2}{p}, \quad \frac{x^2}{p} = 2z + \frac{h^2}{q}.$$

El paraboloide hiperbólico está representado en la fig. 46.

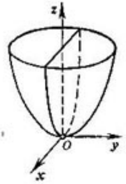

Fig. 45

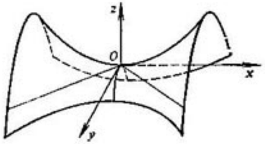

Fig. 46

EL CONO DE SEGUNDO ORDEN

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \quad (a,\ b,\ c > 0). \tag{12}$$

Está claro que esta superficie es simétrica respecto de los planos coordenados $x = 0$, $y = 0$, $z = 0$ y respecto del origen de coordenadas.

En las secciones de la superficie (12) por los planos $z = h$, se obtienen elipses

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{h^2}{c^2}$$

de semiejes $a \mid h \mid / c$, $b \mid h \mid / c$.

En las secciones de la superficie (12) por los planos $x = h$ o $y = h$, resultan hipérbolas

$$\frac{z^2}{c^2} - \frac{y^2}{b^2} = \frac{h^2}{a^2} \quad \left(\frac{z^2}{c^2} - \frac{x^2}{a^2} = \frac{h^2}{b^2}\right).$$

Si ahora hacemos cortes de la superficie (12) por los planos $y = hx$, entonces, en las secciones obtenemos pares de rectas que se cortan

$$z = \pm cx\sqrt{(1/a^2) + (h^2/b^2)}.$$

La forma del cono viene representada en la fig. 47.

El punto

$$x^2 + y^2 + z^2 = 0. \tag{13}$$

A la ecuación (13) sólo la satisface el punto $x = y = z = 0$.

Los cilindros de segundo orden

a) El cilindro elíptico

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a,\ b > 0). \tag{14}$$

La ecuación (14) no contiene la variable z. En el plano xOy la ecuación (14) determina una elipse con los semiejes a y b. Si el punto $(x,\ y)$ está situado en esta elipse, entonces, para cualquier z

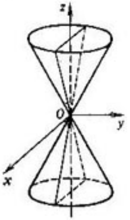

Fig. 47

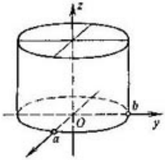

Fig. 48

el punto $(x,\ y,\ z)$ está situado en la superficie (14). El conjunto de tales puntos es la superficie engendrada por una recta paralela al eje Oz y que corta a la elipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

en el plano xOy.

La elipse (14) se llama *directriz de la superficie* dada y todas las rectas en movimiento que engendran la superficie, se llaman *generatrices*.

En general, una superficie engendrada por una recta que permanece paralela a una dirección dada y que corta a una curva dada L, se llama superficie *cilíndrica*. La superficie (14) está representada en la fig. 48.

b) Los cilindros hiperbólico y parabólico

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (a,\ b > 0), \tag{15}$$

$$y^2 = 2px \quad (p > 0). \tag{16}$$

En estos casos, las directrices de las superficies son una hipérbola y una parábola, y las generatrices son rectas paralelas al eje Oz y que pasan por la hipérbola o por la parábola en el plano xOy. Las superficies (15) y (16) están representadas en las figs. 49 y 50.

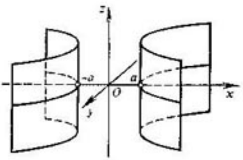

Fig. 49

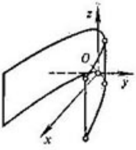

Fig. 50

c) Planos paralelos y planos que se cortan. Recta

$$a^2x^2 - b^2y^2 = 0 \quad (a,\ b > 0), \tag{17}$$

$$x^2 - a^2 = 0 \quad (a > 0), \tag{18}$$

$$z^2 = 0, \tag{19}$$

$$x^2 + y^2 = 0. \tag{20}$$

Para la superficie (17) las directrices son las rectas

$$y = \pm \frac{a}{b} x.$$

Por tanto, la superficie (17) es un par de planos que se cortan.

En ninguna de las ecuaciones de las superficies (18) y (19) figuran dos coordenadas. La ecuación (18) representa en el plano xOy un par de rectas $x = \pm a$.

Tomando $x = \pm a$ y cualesquiera y, z, los puntos $(\pm a,\ y,\ z)$ satisfacen la ecuación (18); por lo tanto, (18) representa un par de rectas paralelas.

La ecuación (19) representa el plano xOy, ya que satisfacen esta ecuación todos los puntos de la forma $(x,\ y,\ 0)$ y todo el conjunto de tales puntos forma el plano xOy.

También se puede considerar $z = 0$ como la directriz de cualquiera de los planos xOz ó yOz, y las generatrices son las rectas paralelas al eje Oy o al eje Ox y que pasan por la recta $z = 0$ en el plano xOz ó yOz.

A la ecuación (20) la satisface cualquier punto con $x = y = 0$ y z arbitrario. Por lo tanto, (20) representa una recta; precisamente el eje Oz.

Superficies regladas

Algunas superficies de segundo orden están formadas por una recta en movimiento. Tales son todas las superficies cilíndricas y el cono de segundo orden. Pero también hay otras superficies que se forman por el movimiento de una recta.

La superficie que se forma por una recta en movimiento, se llama superficie *reglada*, y las rectas que están situadas totalmente sobre la superficie, se llaman *generatrices rectilíneas.*

El hiperboloide de una hoja y el paraboloide hiperbólico son superficies regladas.

La ecuación (8) del hiperboloide de una hoja puede escribirse en la forma

$$\frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{y^2}{b^2}$$

o bien, descomponiendo ambos miembros en factores, así:

$$\left(\frac{x}{a}+\frac{z}{c}\right)\left(\frac{x}{a}-\frac{z}{c}\right)=\left(1+\frac{y}{b}\right)\left(1-\frac{y}{b}\right). \qquad (21)$$

Consideremos el sistema de ecuaciones de primer grado

$$\left.\begin{aligned} \frac{x}{a}+\frac{z}{c}&=k\left(1+\frac{y}{b}\right),\\ \frac{x}{a}-\frac{z}{c}&=\frac{1}{k}\left(1-\frac{y}{b}\right), \end{aligned}\right\} \qquad (22)$$

donde k es un parámetro arbitrario.

Para un valor determinado del parámetro k resulta una recta, y para k variable se obtiene una *familia de rectas.* Multiplicando miembro a miembro las ecuaciones (22), se obtiene la ecuación (21) de la superficie en cuestión. Por lo tanto, cualquier punto (x, y, z) que satisfaga el sistema (22) estará situado en la superficie (21). Por consiguiente, cada una de las rectas de la familia (22) está situada completamente en la superficie del hiperboloide de una hoja.

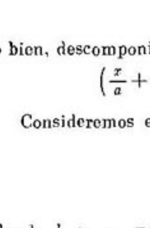

Fig. 51

De un modo similar, el sistema

$$\left.\begin{aligned} \frac{x}{a}+\frac{z}{c}&=l\left(1-\frac{y}{b}\right),\\ \frac{x}{a}-\frac{z}{c}&=\frac{1}{l}\left(1+\frac{y}{b}\right), \end{aligned}\right\} \qquad (23)$$

donde l es un parámetro, también determina una familia de rectas distinta de la familia (22) y que pertenece a la superficie (21).

Por cada punto del hiperboloide (21) pasa una sola recta de cada una de las familias y, por lo general, correspondientes a distintos valores de los parámetros k y l (fig. 51). Por ejemplo, por el punto $\left(\sqrt{\frac{3}{2}}\,a,\ \frac{\sqrt{2}}{2}\,b,\ c\right)$ de la superficie (21) pasa una recta de la familia (22) para $k = (2+\sqrt{6})/(2+\sqrt{2})$ y una recta de la familia (23) para $l = (2+\sqrt{6})/(2-\sqrt{2})$.

Señalemos que los hiperboloides de una hoja encontraron aplicación en las técnicas de la construcción. La construcción de distintas torres altas mediante la aplicación de las generatrices rectilíneas del hiperboloide de una hoja permite conjugar la gran resistencia de la construcción con la sencillez de su edificación. La idea de emplear el hiperboloide de una hoja en la construcción pertenece a nuestro compatriota, el ingeniero V. G. Shújov (1853-1939). La torre de televisión situada en la calle Shabolovka de la ciudad de Moscú ha sido construida según el proyecto de Shújov; ésta consta de varias secciones de hiperboloides de una hoja de revolución.

Fácilmente se comprueba que las dos familias de rectas

$$\left.\begin{aligned} \frac{x}{\sqrt{p}}+\frac{y}{\sqrt{q}}&=2kz,\\ \frac{x}{\sqrt{p}}-\frac{y}{\sqrt{q}}&=\frac{1}{k},\end{aligned}\right\}\tag{24}$$

$$\left.\begin{aligned} \frac{x}{\sqrt{p}}+\frac{y}{\sqrt{q}}&=\frac{1}{l}\\ \frac{x}{\sqrt{p}}-\frac{y}{\sqrt{q}}&=2lz\end{aligned}\right\}\tag{25}$$

forman la superficie del paraboide hiperbólico (11).

Las rectas correspondientes a las familias (24) y (25) están situadas en esta superficie y, recíprocamente, cualquier punto de esta superficie es la intersección de alguna recta de la familia (24) con alguna recta de la familia (25).

§ 26. Teoría general de la superficie de segundo orden en el espacio tridimensional

Sea dada una superficie de segundo orden

$$\sum_{k=1}^{3}\sum_{l=1}^{3} a_{kl}x_kx_l + 2\sum_{l=1}^{3} A_lx_l + B = 0,\tag{1}$$

donde $a_{kl} = a_{lk}$, A_l, B son unas constantes, los coeficientes de la ecuación.

En el § 25 se mencionaron ocho tipos 1) — 8) (casos particulares) de la ecuación (1) y se señaló la posibilidad de demostrar que, para cada una de las ecuaciones (1), siempre que no determine una superficie imaginaria, se puede hallar un sistema rectangular de coordenadas en el que la ecuación tenga la forma de uno de los tipos indicados.

A continuación hacemos la demostración de esta proposición.

Empezamos por considerar la forma cuadrática que figura en el primer miembro de la ecuación (1).

En virtud del teorema 2 § 22, mediante una transformación ortogonal adecuada

$$x_i = \sum_{s=1}^{3} \beta_{is} x'_s \quad (i = 1,\ 2,\ 3) \tag{2}$$

esta forma se puede reducir a la forma:

$$\sum_{k=1}^{3} \sum_{l=1}^{3} a_{kl} x_k x_l = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2,$$

donde λ_1, λ_2, λ_3 son unos números reales determinados.

Subrayemos que las fórmulas (2) determinan una transformación del sistema rectangular inicial de coordenadas $(x_1,\ x_2,\ x_3)$ en otro sistema rectangular $(x'_1\ x'_2,\ x'_3)$. Un punto que en el sistema inicial tenía las coordenadas $(x_1,\ x_2,\ x_3)$, tendrá en el nuevo sistema las coordenadas $(x'_1,\ x'_2,\ x'_3)$, que se obtienen invirtiendo las fórmulas (2).

Evidentemente, en el nuevo sistema de coordenadas la superficie en cuestión tiene la ecuación

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2 + 2 \sum_{k=1}^{3} A'_k x'_k + B = 0, \tag{3}$$

donde A'_k son unas constantes.

Consideremos primero el caso en que los tres números λ_1, λ_2, λ_3 son distintos de cero ($\lambda_i \neq 0$, $i = 1,\ 2,\ 3$).

En este caso, hacemos una traslación del sistema de coordenadas x'_1, x'_2, x'_3 de tal modo que el origen se traslade al punto $(a_1,\ a_2,\ a_3)$; entonces obtendremos un segundo sistema rectangular de coordenadas $(\xi_1,\ \xi_2,\ \xi_3)$, donde

$$x'_i = a_i + \xi_i \quad (i = 1,\ 2,\ 3).$$

En este sistema la ecuación de la superficie en cuestión tiene la forma

$$\lambda_1 (a_1 + \xi_1)^2 + \lambda_2 (a_2 + \xi_2)^2 + \lambda_3 (a_3 + \xi_3)^2 + 2 \sum_{k=1}^{3} A'_k (a_k + \xi_k) + B = 0$$

o bien

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 + 2 \sum_{l=1}^{3} (a_l \lambda_l + A'_l)\, \xi_l + B_1 = 0,$$

donde B_1 es una constante. Haciendo

$$a_l = -\frac{A'_l}{\lambda_l} \quad (l = 1, 2, 3),$$

esta ecuación se simplifica:

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 = -B_1. \tag{4}$$

Supongamos que los números λ_1, λ_2, λ_3 son del mismo signo.

Si, en este caso, $B_1 = 0$, entonces, a la ecuación (4) la satisface un *punto único*, el punto nulo (0, 0, 0) (véase 7) § 25).

Si $B_1 \neq 0$ y tiene el mismo signo que los números λ_1, λ_2, λ_3, entonces, evidentemente, no hay puntos con coordenadas reales que satisfagan la ecuación (4). En este caso, la *superficie* (1) es *imaginaria*.

Si $B_1 \neq 0$, pero tiene el signo contrario al de los números λ_1, λ_2, λ_3, entonces la ecuación (4) puede escribirse en la forma

$$\frac{\xi_1^2}{-\frac{B_1}{\lambda_1}} + \frac{\xi_2^2}{-\frac{B_1}{\lambda_2}} + \frac{\xi_3^2}{-\frac{B_1}{\lambda_3}} = 1 \tag{5}$$

o bien, haciendo

$$a^2 = -\frac{B_1}{\lambda_1}, \quad b^2 = -\frac{B_2}{\lambda_2}, \quad c^2 = -\frac{B_3}{\lambda_3},$$

en la forma

$$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} + \frac{\xi_3^2}{c^2} = 1 \quad (a, b, c > 0).$$

Por lo tanto, la superficie (1) es un *elipsoide* (véase 1) § 25).

Supongamos ahora que dos de los números λ_1, λ_2, λ_3 tienen un signo, mientras que el tercero tiene el signo contrario.

Si, en este caso, $B_1 = 0$, entonces se puede considerar que en la ecuación (4), $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 < 0$, pues, si fuese necesario multiplicaríamos (4) por -1 y permutaríamos los ξ_j. Entonces, haciendo

$$\lambda_1 = \frac{1}{a^2}, \quad \lambda_2 = \frac{1}{b^2}, \quad \lambda_3 = -\frac{1}{c^2} \quad (a, b, c > 0),$$

obtenemos la ecuación del *cono* (véase 6) § 25)

$$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} = 0.$$

Si $B_1 \neq 0$, aplicamos de nuevo la fórmula (5). Seleccionemos dos casos esencialmente distintos:

$\frac{\xi_1^2}{a^2} + \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} = 1$ (*hiperboloide de una hoja* 2) del § 25), y

$\frac{\xi_1^2}{a^2} - \frac{\xi_2^2}{b^2} - \frac{\xi_3^2}{c^2} = 1$ (*hiperboloide de dos hojas* 3) del § 25).

Los demás casos se reducen a éstos mediante un cambio respectivo de las coordenadas ξ_i.

Supongamos ahora que $\Delta = \lambda_1\lambda_2\lambda_3 = 0$. Entonces, al menos uno de los números λ_1, λ_2, λ_3 es igual a cero. Consideraremos que $\lambda_3 = 0$, permutando si fuese necesario los ξ_i.

Así pues, supongamos que $\lambda_3 = 0$. Examinemos primero el caso en que $A_3' = 0$. Entonces, la ecuación (3) tiene la forma

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + 2\,(A_1' x_1' + A_2' x_2') + B = 0,$$

donde los números λ_1, λ_2, A_1', A_2', B son arbitrarios.

En el plano (x_1', x_2') ésta es la ecuación general de una curva de segundo orden. En el espacio (x_1, x_2, x_3) es la *ecuación de una superficie cilíndrica* que pasa por la *curva plana de segundo orden* con generatriz paralela al eje x_3' (véase 8) § 25).

A continuación siempre vamos a suponer que $A_3' \neq 0$ y $\lambda_3 = 0$, y entonces la ecuación (3) tiene la forma

$$(\lambda_1 x_1'^2 + 2A_1' x_1') + (\lambda_2 x_2'^2 + 2A_2' x_2') + 2A_3' x_3' + B = 0. \qquad (3')$$

Son esencialmente distintos los casos siguientes: a) λ_1, $\lambda_2 > 0$; b) $\lambda_1 > 0$, $\lambda_2 < 0$; c) $\lambda_1 > 0$, $\lambda_2 = 0$. Los otros casos reducen a éstos mediante un cambio de coordenadas o multiplicando por -1.

Consideremos el caso a) λ_1, $\lambda_2 > 0$. Saquemos fuera de paréntesis los factores λ_1 y λ_2 y completemos cuadrados en las expresiones entre paréntesis. Entonces, teniendo en cuenta que λ_1, λ_2, A_3' son distintos de cero, obtenemos:

$$\lambda_1\,(x_1' + \alpha)^2 + \lambda_2\,(x_2' + \beta)^2 + 2A_3'\,(x_3' + \gamma) = 0,$$

donde α, β, γ son ciertos números. Haciendo

$$\xi = x_1' + \alpha, \quad \eta = x_2' + \beta, \quad \zeta = x_3' + \gamma,$$

obtenemos

$$\lambda_1 \xi^2 + \lambda_2 \eta^2 = -2A_3' \zeta$$

o bien

$$\frac{\xi^2}{-\frac{A_3'}{\lambda_1}} + \frac{\eta^2}{-\frac{A_3'}{\lambda_2}} = 2\zeta. \qquad (6)$$

Si $-A_3' > 0$, entonces la ecuación (6) tiene la forma

$$\frac{\xi^2}{p} + \frac{\eta^2}{q} = 2\zeta \quad (p,\ q > 0), \qquad (7)$$

(*paraboloide elíptico* 4) del § 25).

Si $-A_3' > 0$, entonces, sustituyendo ζ por $-\zeta$, de nuevo obtenemos una ecuación de la forma (7), o sea, un *paraboloide elíptico*.

Examinemos ahora el caso b) $\lambda_1 > 0$, $\lambda_2 < 0$ $(A_3' \neq 0)$. Consideremos la ecuación (6). Si $A_3' < 0$, entonces esta ecuación se escribe en la forma

$$\frac{\xi^2}{p} - \frac{\eta^2}{q} = 2\zeta \quad (p,\ q > 0), \tag{8}$$

(*paraboloide hiperbólico* 5) del § 25).

Si $A_3' > 0$, entonces, después de permutar ξ y η, de nuevo obtenemos una ecuación de la forma (8), o sea, un *paraboloide hiperbólico*.

Consideremos ahora el caso c) $\lambda_1 > 0$, $\lambda_2 = 0$ ($A_3' \neq 0$). Entonces, la ecuación (3′) se reduce a la siguiente:

$$(\lambda_1 x_1'^2 + 2A_1'x_1') + 2(A_2'x_2' + A_3'x_3') + B = 0.$$

Sacando fuera del primer paréntesis λ_1 y completando el cuadrado en el mismo (teniendo en cuenta que λ_1, $A_3' \neq 0$), obtenemos:

$$\lambda_1 (x_1' + \alpha)^2 + 2A_2'x_2' + 2A_3'(x_3' + \beta) = 0,$$

donde α, β son ciertos números. Después de hacer la sustitución

$$\xi = x_1' + \alpha, \quad \eta = x_2', \quad \zeta = x_3' + \beta,$$

esta ecuación se reduce a la siguiente:

$$\lambda_1 \xi^2 + 2(A_2'\eta + A_3'\zeta) = 0 \tag{9}$$

Consideremos en el plano (η, ζ) el vector $\omega = (A_2', A_3')$. Escribámoslo en la forma

$$(A_2', A_3') = \rho\,(\cos\alpha,\ \operatorname{sen}\alpha),$$

donde $\rho > 0$ es la longitud de ω, y $(\cos\alpha, \operatorname{sen}\alpha)$ es el vector unitario que lleva la dirección de ω. El vector $(\operatorname{sen}\alpha, -\cos\alpha)$ también es unitario y es perpendicular al primero.

Consideremos en el plano (η, ζ) la transformación ortogonal

$$u = \eta\cos\alpha + \zeta\operatorname{sen}\alpha, \quad v = \eta\operatorname{sen}\alpha - \xi\cos\alpha.$$

Esta transformación cambia el sistema rectangular de coordenadas (η, ζ) por el sistema rectangular de coordenadas (u, v), y el sistema rectangular de coordenadas (ξ, η, ζ) se sustituye por el sistema rectangular de coordenadas (ξ, u, v). Como resultado, la ecuación (9), que puede escribirse así

$$\lambda_1\xi^2 + 2\rho\,(\eta\cos\alpha + \zeta\operatorname{sen}\alpha) = 0,$$

toma la forma siguiente:

$$\lambda_1\xi^2 + 2\rho u = 0,$$

o bien, cambiando u por $-u$,

$$\xi^2 = 2pu \quad (p = \rho/\lambda_1 > 0),$$

o bien, finalmente, permutando ξ y u,

$$u^2 = 2p\xi \quad (p > 0).$$

o sea, hemos obtenido la *ecuación de un cilindro parabólico* (en las coordenadas rectangulares (ξ, u, v)).

Quedan estudiados todos los casos que pueden presentarse para la ecuación (1), y en cada uno de ellos hemos hallado el sistema rectangular de coordenadas en el que la ecuación (1) toma de las formas 1) — 8) § 25.

La proposición queda demostrada

INDICE ALFABÉTICO

Abscisa 42
Adjunto o complemento algebraico 13
Angulo formado por dos planos 65
— — — las rectas 56, 70
— — — los vectores 49
Asintota de la hipérbola 139

Base 94, 95
— en el espacio 95
— — — subespacio 110-111
— ortogonal 98

Cilíndros de segundo orden 147, 158
Coeficientes del sistema de ecuaciones 20
Condiciones de paralelismo de las rectas 56-57
— — — — los planos 65
— — perpendicularidad de las rectas 56, 57
— — — — los planos 65
Coordenadas del vector 20, 42, 45
Curvas de segundo orden 134

Desigualdad de Buniakovski 48
— — Minkowski 50
Determinante 8, 9, 11
— de Vandermonde 15
— engendrado por la matriz 18
Diagonal del determinante 9
Diferencia de vectores 39, 44
Directriz de la parábola 141
Distancia de un punto a una recta 59
— — — — — un plano 66, 67
— entre los puntos 43, 44, 46
División de un segmento en una razón dada 51

Ecuación característica 131, 148
— de la recta 53
— — — — con el coeficiente angular 55
— — — — en la forma canónica 68
— — — — — — — normal 58
— — — — — — — vectorial 57
— — — — que pasa por dos puntos 55
— — — — — — — — un punto 57
— — — — «segmentaria» 56
— del plano en la forma normal 61
— — — que pasa por tres puntos 64
Ecuaciones de los cilíndros de segundo orden 146, 147, 155-161

Ecuación general del plano 61
— segmentaria del plano 63
Ejes de coordenadas 42, 94
Elemento del determinante 11
Elementos de la matriz 18
Elipse 135
Elipsoide 146, 153
Espacio euclídeo n-dimensional 45

Factor normalizador 59
Foco de la elipse 135
— — — hipérbola 137
— — — parábola 141
Forma canónica 128
— cuadrática 122, 126, 128, 133
Fórmulas de Cramer 21

Generatrices rectilíneas 160

Imagen del espacio mediante el operador 117
Invariación de los productos vectorial y escalar 105

Longitud o norma del vector 37, 46

Matriz 18
— ampliada 23
— cuadrada 18
— inversa 90
— traspuesta 19
— unidad 90
Menor del elemento 13
Método de eliminación de incognitas 30
— — Gauss 30

Operador autoconjugado 122
— inverso 90
— lineal 87
— unidad 90
Orientación del sistema de vectores 72, 79
— de los sistemas de coordenadas 71, 72, 73
Origen de coordenadas 42

Permutación 10
Producto de la matriz por un número 19

— — las matrices 88
— — los determinantes 16, 89
— escalar 40, 45
— vectorial 73
— vectorial-escalar 79
Proyección del vector sobre la recta 38, 99
Punto del espacio n-dimensional 45

Radio vector 42
Rango de la matriz 18
Recta 53, 68
Regla de Sarrus 9

Segmento 51
Sistema de ecuaciones compatible 28
— — — homogéneo 22
— — vectores ortonormal 98
— — — propios del operador 125
Sistemas de coordenadas rectangulares 41, 42
— — vectores 80, 81
Solución del sistema de ecuaciones 20, 21, 23
Subespacios 110
— lineales 110
— ortogonales 111
Suma de determinantes 14
— — matrices 19
— — vectores 37, 46
Superficie cilíndrica 158
— de segundo orden 146
— reglada 160
Superficies directrices 158
— generatrices 158

Teorema de Sylvester 128
— — tipo Fredholm 115
Teoría de Kronecker-Capelli 20, 23
Término del determinante 11
Transposición de elementos 10
Transformación de las coordenadas rectangulares 107
Transformaciones de la orientación 110
Transformación ortogonal 100, 108

Valor propio del operador 124
Vector 20, 36, 45, 105
Vectores colineares 37, 76
— coplanares 73, 80
— ortogonales 50, 98
Vector normal 98
— nulo 20, 22, 37
— propio del operador 124
Versores de los ejes de coordenadas 43, 94

Z-coordenada 42

A NUESTROS LECTORES:

Mir edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas extranjeros. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia ficción. Dirijan sus opiniones a la Editorial Mir, 1 Rizhski per., 2, 129820, Moscú, I-110, GSP, URSS.

Este es el primer libro de la serie "Matemáticas superiores", compuesta de tres volúmenes: "Elementos de álgebra lineal y de geometría analítica", "Cálculo diferencial e integral" y "Ecuaciones diferenciales. Integrales múltiples. Series. Funciones de variable compleja".

En el presente volumen se desarrollan las cuestiones fundamentales de la teoría de los determinantes, de la teoría de los sistemas de ecuaciones lineales, del álgebra vectorial. También se examinan las partes más importantes del álgebra lineal: operadores lineales; transformaciones ortogonales; operadores autoconjugados; la forma cuadrática y su reducción a la forma canónica.

Se incluyen de la geometría analítica: la linea recta; el plano; la recta en el espacio; curvas y superficies de segundo orden.

Por lo general, los razonamientos van acompañados de demostraciones exhaustivas.

Las formas canónicas de curvas y superficies de segundo orden se desarrollan en este libro en forma muy breve, debido a que se supone que las mismas se estudiarán complementariamente en forma de ejercicios como métodos del análisis matemático.

La forma cuadrática se estudia como métodos del análisis matemático o del análisis funcional.

En este volumen se exponen todos los temas que conforman el programa respectivo para los institutos de enseñanza técnica superior.