A.G. KUROSCH

CURSO DE ALGEBRA SUPERIOR

EDITORIAL MIR

А. Г. КУРОШ

КУРС
ВЫСШЕЙ АЛГЕБРЫ

ИЗДАТЕЛЬСТВО «НАУКА»

На испанском языке

A. G. KUROSCH

CURSO de ALGEBRA SUPERIOR

Traducido del ruso por

EMILIANO APARICIO BERNARDO,

Candidado a Doctor en Ciencias Físico-Matemáticas,

Catedrático de Matemáticas Superiores.

EDITORIAL MIR • MOSCU 1968

CDU 512.8 (075.8) = 60

INDICE

PALABRAS DE PRESENTACION

La versión castellana de la obra del profesor A. G. Kurosch «Curso de álgebra superior» que ofrecemos al lector, es el primer libro del autor que se traduce al español.

El conocimiento del álgebra superior es indispensable para la formación matemática del estudiante que ha decidido consagrarse al estudio de las matemáticas. El presente libro marca un camino relativamente corto para pasar del álgebra elemental al estudio de los métodos abstractos del álgebra moderna.

En los primeros capítulos se estudian detalladamente los determinantes y sistemas de ecuaciones lineales, se introducen los números complejos y las operaciones sobre las matrices, y se hace una exposición de la teoría de los polinomios y formas cuadráticas. En los capítulos VII y VIII, el autor nos da una idea primordial del álgebra lineal. En el capítulo X vemos que el álgebra lineal, el álgebra de los polinomios y las funciones racionales pueden generalizarse para el caso de un campo fundamental arbitrario. Precisamente en este capítulo, el autor nos enseña los principios del álgebra moderna. Aquí nos encontramos con los conceptos importantes de anillo y campo. Estos conceptos permiten exponer con mayor generalidad la teoría de los polinomios en varias indeterminadas, suponiendo que los coeficientes de estos polinomios pertenecen a un campo fundamental arbitrario. A continuación, las matrices polinomiales también se estudian sobre un campo fundamental arbitrario y se aplican para la elaboración de la teoría de las matrices de Jordan. El último capítulo está dedicado a los grupos; éste es el comienzo de una rama muy importante del álgebra moderna, denominada teoría de los grupos.

El autor de este libro es un gran especialista en teoría de grupos. Su libro «Teoría de los grupos», desempeñó un papel muy importante en el desarrollo de las investigaciones sobre este tema en la Unión Soviética. Hace unos años, el prof. A. G. Kurosch publicó una original obra, titulada «Lecciones de álgebra general», que fue favorablemente acogida por los algebristas soviéticos.

El prof. A. G. Kurosch es jefe de la cátedra de álgebra superior de la Universidad de Moscú desde el año 1949.

El presente libro es un compendio de álgebra superior que comprende los conocimientos de esta ciencia obligatorios para los estudiantes de matemáticas de la Universidad de Moscú. Desde la aparición de su primera edición en ruso, en el año 1946, ya ha sido reeditado ocho veces. En la Unión Soviética éste es uno de los mejores libros sobre el tema considerado. Esperamos que tenga buena acogida en los países de habla hispánica.

Agradeceremos al lector sus observaciones sobre la presente traducción, que trataremos de tener en cuenta en el futuro.

Moscú. Febrero de 1968. *E. Aparicio Bernardo*

CAPITULO I

SISTEMAS DE ECUACIONES LINEALES. DETERMINANTES

§ 1. Método de eliminación consecutiva de las incógnitas

Comenzamos el curso de álgebra superior con el estudio de los sistemas de ecuaciones de primer grado con varias incógnitas o, como suele decirse, *de los sistemas de ecuaciones lineales* *.

La teoría de los sistemas de ecuaciones lineales origina una amplia e importante rama del álgebra, el álgebra lineal, a la que están dedicados una gran parte de los capítulos de este libro y, en particular, los tres primeros. Se supone que son reales los coeficientes de las ecuaciones que se consideran en estos tres capítulos, los valores de las incógnitas y, en general, todos los números que aparecen. En realidad, todo el contenido de estos capítulos se generaliza, palabra por palabra, al caso de números complejos arbitrarios, ya conocidos por el lector en el curso de la escuela media.

A diferencia del álgebra elemental, aquí se estudian los sistemas con un número arbitrario de ecuaciones e incógnitas. Además, suponemos que el número de ecuaciones del sistema *no* coincide con el número de incógnitas.

Sea dado un sistema de s ecuaciones lineales con n incógnitas. Convengamos en emplear las siguientes notaciones: las incógnitas las designaremos con la letra x con subíndices $1, 2, \ldots n$: $x_1, x_2, \ldots, x_n$; supondremos que las ecuaciones están numeradas así: la primera, la segunda, ..., la s-ésima; el coeficiente de la incógnita x_j en la i-ésima ecuación, se señalará mediante a_{ij}**; finalmente, el término independiente de la i-ésima ecuación se designará con b_i.

---

* Esta denominación se debe a que, en la geometría analítica, una ecuación de primer grado con dos incógnitas determina una recta en el plano.

** Por consiguiente, se emplearán dos subíndices, el primero de los cuales indicará el número de la ecuación, y el segundo, el número de la incógnita. Para abreviar, estos índices no se separarán con una coma; claro que, en el caso de a_{11}, no se debe leer «a once», sino «a uno uno», y en el caso de a_{34}, no se debe leer «a treinta y cuatro», sino «a tres cuatro».

Nuestro sistema se escribirá ahora en la forma general siguiente:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s. \end{array}\right\} \qquad (1)$$

Los coeficientes se pueden colocar formando un cuadro

$$\begin{pmatrix} a_{11}a_{12} & \ldots & a_{1n} \\ a_{21}a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots \\ a_{s1}a_{s2} & \ldots & a_{sn} \end{pmatrix}, \qquad (2)$$

denominado *matriz de s filas* y *n columnas*; los números a_{ij} se llaman *elementos* de la matriz*. Si $s = n$ (o sea, que el número de filas es igual al número de columnas), se dice que la matriz es *cuadrada* y *de orden n*. La diagonal de esta matriz que une el ángulo superior izquierdo con el ángulo inferior derecho (o sea, formada por los elementos a_{11}, a_{22}, $\ldots$, a_{nn}), se llama *diagonal principal*. Una matriz cuadrada de orden n se llamará *matriz unidad de orden n*, si todos los elementos de su diagonal principal son iguales a la unidad, y todos los elementos que están fuera de esta diagonal son iguales a cero.

Se llama *solución* de un sistema de ecuaciones lineales (1) a un sistema de n números k_1, k_2, $\ldots$, k_n, en el que cada una de las ecuaciones del sistema (1) se convierte en una identidad, después de haber sustituido en ella las incógnitas x_i por los números correspondientes k_i, $i = 1, 2, \ldots, n$**.

Un sistema de ecuaciones lineales puede no tener solución alguna, y entonces se llama *incompatible*. Tal es, por ejemplo, el sistema

$$x_1+5x_2=1,$$
$$x_1+5x_2=7;$$

los primeros miembros de estas ecuaciones son iguales, mientras que los segundos son distintos. Por lo tanto, ningún sistema de valores de las incógnitas puede satisfacer simultáneamente a las dos ecuaciones.

Si el sistema de ecuaciones lineales tiene solución, se llama *compatible*. Se dice que un sistema compatible es *determinado*, si posee una solución única (en el álgebra elemental solamente se estudian

---

* De este modo, si la matriz (2) se examina sin relación con el sistema (1), el primer subíndice del elemento a_{ij} indica el número de su fila, y el segundo, el número de su columna.

** Hay que subrayar, que los número k_1, k_2, ..., k_n forman *una* solución del sistema y no n soluciones.

tales sistemas), e *indeterminado*, si tiene más de una solución. En este caso, como veremos más adelante, hay una infinidad de soluciones Así, el sistema

$$\left.\begin{aligned} x_1+2x_2&=7, \\ x_1+x_2&=4 \end{aligned}\right\}$$

es determinado y $x_1 = 1$, $x_2 = 3$ es una solución. Por el método de eliminación de la incógnita, se puede comprobar fácilmente que esta solución es única. Por otra parte, el sistema

$$\left.\begin{aligned} 3x_1-x_2&=1, \\ 6x_1-2x_2&=2 \end{aligned}\right\}$$

es indeterminado, puesto que tiene infinitas soluciones de la forma

$$x_1=k, \quad x_2=3k-1, \tag{3}$$

donde el número k es arbitrario. Con las soluciones obtenidas por las fórmulas (3) se agotan todas las soluciones de nuestro sistema.

El problema de la teoría de los sistemas de ecuaciones lineales consiste en la elaboración de métodos que permitan establecer si es compatible o no un sistema dado de ecuaciones, y en caso de compatibilidad, indicar el número de soluciones y señalar un método para hallar todas ellas.

Comenzaremos por el método más cómodo para hallar prácticamente las soluciones de los sistemas con coeficientes numéricos, es decir, *con el método de eliminación consecutiva de las incógnitas o método de Gauss**.

Hagamos primero una observación. A continuación, tendremos que hacer las siguientes transformaciones del sistema de ecuaciones lineales: ambos miembros de una de las ecuaciones del sistema, multiplicados previamente por un mismo número, se van a restar de los miembros correspondientes de otra de las ecuaciones del sistema. Supongamos, por ejemplo, que ambos miembros de la primera ecuación del sistema (1), multiplicados por el número c, se restan de los correspondientes miembros de la segunda ecuación. Obtendremos un nuevo sistema de ecuaciones lineales:

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1, \\ a'_{21}x_1+a'_{22}x_2+\ldots+a'_{2n}x_n&=b'_2, \\ a_{31}x_1+a_{32}x_2+\ldots+a_{3n}x_n&=b_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n&=b_s, \end{aligned}\right\} \tag{4}$$

* También se llama método de reducción. (*Nota del T.*)

donde

$$a'_{2j}=a_{2j}-ca_{1j} \text{ para } j=1,\ 2,\ \ldots,\ n,\ b'_2=b_2-cb_1.$$

Los sistemas de ecuaciones (1) *y* (4) *son equivalentes, es decir, son simultáneamente incompatibles o son simultáneamente compatibles y, en el último caso, poseen las mismas soluciones.* En efecto, sea $k_1,\ k_2,\ \ldots,\ k_n$ una solución arbitraria del sistema (1). Es evidente, que estos números satisfacen a todas las ecuaciones del sistema (4), menos a la segunda. Sin embargo, también satisfacen a la segunda ecuación del sistema (4): es suficiente recordar que esta ecuación se expresa mediante la segunda y la primera de las ecuaciones del sistema (1). Recíprocamente, toda solución del sistema (4) satisface también al sistema (1). En efecto, la segunda ecuación del sistema (1) se obtiene restando de ambos miembros de la segunda ecuación del sistema (4) los miembros correspondientes de la primera ecuación de este sistema, multiplicados por el número $-c$.

Es comprensible que, *si en el sistema* (1) *se efectúan unas cuantas veces las transformaciones del tipo considerado, el sistema obtenido de ecuaciones se mantendrá equivalente al sistema inicial* (1).

Puede ocurrir que después de efectuar tales transformaciones aparezca en nuestro sistema una ecuación, cuyos coeficientes en el primer miembro sean iguales a cero. Si el término independiente de esta ecuación es también igual a cero, la ecuación se satisface con cualesquiera valores de las incógnitas. Por lo tanto, *suprimiendo esta ecuación, llegamos a un sistema de ecuaciones que es equivalente al inicial.* Si el término independiente de la ecuación considerada es diferente de cero, la ecuación no puede ser satisfecha por ninguno de los valores de las incógnitas y, por esto, *el sistema obtenido de ecuaciones, al igual que el sistema inicial equivalente, será incompatible.*

Expongamos ahora el método de Gauss.

Sea dado un sistema arbitrario de ecuaciones lineales (1). Supongamos para precisar que $a_{11} \neq 0$; claro, puede ocurrir que a_{11} sea igual a cero y, entonces, tendríamos que comenzar por cualquier otro coeficiente de la primera ecuación del sistema, diferente de cero.

Transformemos ahora el sistema (1), eliminando la incógnita x_1 de todas las ecuaciones, menos de la primera. Para esto, multipliquemos ambos miembros de la primera ecuación por $\frac{a_{21}}{a_{11}}$ y restémoslos de los miembros correspondientes de la segunda ecuación. Después, multipliquemos ambos miembros de la primera ecuación por $\frac{a_{31}}{a_{11}}$, y restémoslos de los miembros correspondientes de la tercera ecuación, etc., etc.

De este modo, obtendremos un nuevo sistema de s ecuaciones lineales con n incógnitas:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n &= b'_2, \\ a'_{32}x_2 + a'_{33}x_3 + \ldots + a'_{3n}x_n &= b'_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a'_{s2}x_2 + a'_{s3}x_3 + \ldots + a'_{sn}x_n &= b'_s. \end{aligned}\right\} \tag{5}$$

No tenemos necesidad de escribir explícitamente las expresiones de los coeficientes nuevos a'_{ij} y de los términos independientes nuevos b'_i, mediante los coeficientes y los términos independientes del sistema inicial (1).

Como ya sabemos, el sistema de ecuaciones (5) es equivalente al sistema (1). Transformemos ahora el sistema (5). Pero no tocaremos más la primera ecuación, y *las transformaciones solamente las efectuaremos* con la parte del sistema (5) formada por todas las ecuaciones, menos la primera. Se sobrentiende que entre ellas no hay ecuaciones cuyos coeficientes de los primeros miembros sean iguales a cero: tales ecuaciones las habríamos suprimido, si sus términos independientes fuesen iguales a cero, y en caso contrario, quedaría ya demostrada la incompatibilidad de nuestro sistema. Por lo tanto, hay coeficientes a'_{ij} diferentes de cero; supongamos, para precisar, que $a'_{22} \neq 0$. Transformemos ahora el sistema (5), restando de ambos miembros de la tercera ecuación y de cada una de las siguientes ecuaciones, ambos miembros de la segunda ecuación, multiplicados por

$$\frac{a'_{32}}{a'_{22}}, \quad \frac{a'_{42}}{a'_{22}}, \quad \ldots, \quad \frac{a'_{s2}}{a'_{22}},$$

respectivamente. De este modo, quedará eliminada x_2 de todas las ecuaciones, menos de la primera y de la segunda. Obtendremos el siguiente sistema de ecuaciones, que es equivalente al sistema (5), y por consiguiente, también al sistema (1):

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

Nuestro sistema contiene ahora t ecuaciones, $t \leqslant s$, puesto que, posiblemente, algunas de las ecuaciones hayan sido suprimidas. Es evidente que después de eliminar la incógnita x_1, puede disminuir el número de ecuaciones del sistema. A continuación, transformaremos solamente aquella parte del sistema obtenido que contiene todas las ecuaciones, menos las dos primeras.

¿Cuándo se terminará este proceso de eliminación consecutiva de las incógnitas?

Si llegamos a un sistema tal, en el que una de sus ecuaciones tenga un término independiente diferente de cero, mientras que todos los coeficientes del primer miembro sean iguales a cero, entonces, como ya sabemos, nuestro sistema inicial será incompatible.

En caso contrario, obtendremos el siguiente sistema de ecuaciones, equivalente al sistema (1):

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,k-1}x_{k-1} + a_{1k}x_k + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + \ldots + a'_{2,k-1}x_{k-1} + a'_{2k}x_k + \ldots + a'_{2n}x_n &= b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a^{(k-2)}_{k-1,k-1}x_{k-1} + a^{(k-2)}_{k-1,k}x_k + \ldots + a^{(k-2)}_{k-1,n}x_n &= b^{(k-2)}_{k-1}, \\ a^{(k-1)}_{kk}x_k + \ldots + a^{(k-1)}_{kn}x_n &= b^{(k-1)}_k. \end{aligned}\right\} \quad (6)$$

Aquí, $a_{11} \neq 0$, $a'_{22} \neq 0$, $\ldots$, $a^{(k-2)}_{k-1,k-1} \neq 0$, $a^{(k-1)}_{kk} \neq 0$. Señalemos también que $k \leqslant s$ y, evidentemente, $k \leqslant n$.

En este caso, el sistema (1) *es compatible. Es determinado para* $k = n$, *e indeterminado, para* $k < n$.

En efecto, si $k = n$, el sistema (6) tiene la forma

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1, \\ a'_{22}x_2 + \ldots + a'_{2n}x_n &= b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a^{(n-1)}_{nn}x_n &= b^{(n-1)}_n. \end{aligned}\right\} \quad (7)$$

De la última ecuación obtenemos un valor absolutamente determinado de la incógnita x_n. Sustituyéndolo en la penúltima ecuación, hallaremos un valor unívocamente determinado de la incógnita x_{n-1}. Continuando de este modo, hallaremos que el sistema (7), y por tanto el sistema (1), poseen solución única, es decir, son compatibles y determinados.

Si $k < n$, tomamos valores numéricos arbitrarios para las incógnitas «independientes» $x_{k+1}, \ldots, x_n$, después de lo cual, avanzando por el sistema (6) de abajo arriba hallaremos, como anteriormente, unos valores unívocamente determinados para las incógnitas x_k, $x_{k-1}, \ldots, x_2, x_1$. Como los valores para las incógnitas independientes se pueden elegir de infinitos modos, el sistema (6) y, por consiguiente, el sistema (1), serán compatibles, pero indeterminados. Es fácil comprobar, que con el método indicado (eligiendo de todos los modos posibles los valores para las incógnitas independientes), se hallan todas las soluciones del sistema (1).

A primera vista puede parecer que con el método de Gauss el sistema de ecuaciones lineales se puede reducir a otra forma más: la que resulta de agregar al sistema (7) unas cuantas ecuaciones que contengan solamente a la incógnita x_n. Sin embargo, lo que ocurre en realidad es que las transformaciones no se han llevado hasta el fin: como $a_{nn}^{(n-1)} \neq 0$, se puede eliminar la incógnita x_n de todas las ecuaciones, comenzando desde la $(n+1)$-ésima.

Se debe advertir, que la forma «triangular» del sistema de ecuaciones (7), o la forma «trapezoidal» del sistema de ecuaciones (6) (para $k < n$), se obtuvo debido a la suposición de que los coeficientes a_{11}, a'_{22}, etc., etc, eran diferentes de cero. En el caso general, el sistema de ecuaciones a que llegaremos después de realizar hasta el fin el proceso de eliminación de las incógnitas, tomará una forma triangular o trapezoidal sólo después de un cambio debido de la numeración de las incógnitas.

Haciendo un resumen de todo lo expuesto anteriormente, llegamos a la conclusión de que *el método de Gauss se puede aplicar a cualquier sistema de ecuaciones lineales. Además, el sistema será incompatible, si en el proceso de las transformaciones obtenemos una ecuación en la que los coeficientes de las incógnitas son iguales a cero, mientras que el término independiente es diferente de cero; si no nos encontramos con tal ecuación, el sistema será compatible. Un sistema compatible de ecuaciones es determinado, si se reduce a la forma triangular* (7), *e indeterminado, si se reduce a la forma trapezoidal* (6) *siendo* $k < n$.

Apliquemos lo expuesto al caso de un sistema de ecuaciones lineales *homogéneas*, es decir, de ecuaciones, cuyos términos independientes son iguales a cero. Tal sistema siempre es compatible, puesto que posee la *solución nula* (0, 0, . . ., 0). Supongamos que en el sistema considerado, el número de ecuaciones es *menor* que el número de incógnitas. Entonces, este sistema no podrá reducirse a la forma triangular, puesto que en el proceso de transformaciones por el método de Gauss, el número de ecuaciones del sistema sólo puede disminuir pero no aumentar; por consiguiente, éste se reducirá a la forma trapezoidal, es decir, será indeterminado.

En otras palabras: *si en un sistema de ecuaciones lineales homogéneas, el número de ecuaciones es menor que el número de incógnitas, este sistema, además de la solución nula, poseerá también soluciones no nulas*, es decir, soluciones, en las que los valores de ciertas incógnitas (o incluso de todas) serán diferentes de cero; *habrá una infinidad de soluciones de éstas.*

Para la resolución práctica de un sistema de ecuaciones lineales por el método de Gauss, se debe escribir la matriz de los coeficientes del sistema y agregarle una columna formada por los términos independientes, separada para mayor comodidad por una raya vertical. Todas las transformaciones se deben efectuar con las filas de esta matriz «ampliada».

Ejemplos. 1. Resolver el sistema

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

Efectuemos las transformaciones en la matriz ampliada del sistema:

$$\left(\begin{array}{rrr|r} 1 & 2 & 5 & -9\\ 1 & -1 & 3 & 2\\ 3 & -6 & -1 & 25 \end{array}\right) \rightarrow \left(\begin{array}{rrr|r} 1 & 2 & 5 & -9\\ 0 & -3 & -2 & 11\\ 0 & -12 & -16 & 52 \end{array}\right) \rightarrow \left(\begin{array}{rrr|r} 1 & 2 & 5 & -9\\ 0 & -3 & -2 & 11\\ 0 & 0 & -8 & 8 \end{array}\right)$$

Por consiguiente, llegamos al siguiente sistema de ecuaciones:

$$\left.\begin{aligned} x_1+2x_2+5x_3&=-9,\\ -3x_2-2x_3&=11,\\ -8x_3&=8, \end{aligned}\right\}$$

que posee la solución única:

$$x_1=2,\quad x_2=-3,\quad x_3=-1.$$

Por consiguiente, el sistema inicial es determinado.

2. Resolver el sistema

$$\left.\begin{aligned} x_1-5x_2-8x_3+x_4&=3,\\ 3x_1+x_2-3x_3-5x_4&=1,\\ x_1-7x_3+2x_4&=-5,\\ 11x_2+20x_3-9x_4&=2. \end{aligned}\right\}$$

Transformemos la matriz ampliada del sistema:

$$\left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 3 & 1 & -3 & -5 & 1\\ 1 & 0 & -7 & 2 & -5\\ 0 & 11 & 20 & -9 & 2 \end{array}\right) \rightarrow \left(\begin{array}{rrrr|r} 1 & -5 & 8 & 1 & 3\\ 0 & 16 & 21 & -8 & -8\\ 0 & 5 & 1 & 1 & -8\\ 0 & 11 & 20 & -9 & 2 \end{array}\right) \rightarrow$$

$$\rightarrow \left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & -89 & 0 & -29 & 160\\ 0 & 5 & 1 & 1 & -8\\ 0 & -89 & 0 & -29 & 162 \end{array}\right) \rightarrow \left(\begin{array}{rrrr|r} 1 & -5 & -8 & 1 & 3\\ 0 & -89 & 0 & -29 & 160\\ 0 & 5 & 1 & 1 & -8\\ 0 & 0 & 0 & 0 & 2 \end{array}\right)$$

Hemos llegado a un sistema que contiene la ecuación $0=2$. Por consiguiente, el sistema inicial es incompatible.

3. Resolver el sistema

$$\begin{aligned} 4x_1+x_2-3x_3-x_4&=0,\\ 2x_1+3x_2+x_3-5x_4&=0,\\ x_1-2x_2-2x_3+3x_4&=0. \end{aligned}$$

Este es un sistema de ecuaciones homogéneas, donde el número de ecuaciones es menor que el número de incógnitas; por lo tanto, tiene que ser indeterminado. Como todos los términos independientes son iguales a cero, vamos a transformar solamente la matriz de los coeficientes del sistema:

$$\begin{pmatrix} 4 & 1 & -3 & -1 \\ 2 & 3 & 1 & -5 \\ 1 & -2 & -2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 9 & 5 & -13 \\ 0 & 7 & 5 & -11 \\ 1 & -2 & -2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 2 & 0 & -2 \\ 0 & 7 & 5 & -11 \\ 1 & -2 & -2 & 3 \end{pmatrix}.$$

Hemos obtenido el sistema de ecuaciones:

$$\left.\begin{aligned} 2x_2 \qquad - 2x_4 &= 0, \\ 7x_2 + 5x_3 - 11x_4 &= 0, \\ x_1 - 2x_2 - 2x_3 + 3x_4 &= 0. \end{aligned}\right\}$$

Cualquiera de las incógnitas x_2 y x_4, se puede tomar como incógnita independiente. Sea $x_4 = \alpha$; entonces, de la primera ecuación se deduce que $x_2 = \alpha$; de la segunda ecuación obtenemos, $x_3 = \frac{4}{5}\alpha$; y, por fin, de la tercera ecuación, $x_1 = \frac{3}{5}\alpha$. Por lo tanto, la forma general de las soluciones del sistema de ecuaciones dado es:

$$\frac{3}{5}\alpha,\ \alpha,\ \frac{4}{5}\alpha,\ \alpha.$$

§ 2. Determinantes de segundo y tercer orden

El método de resolución de los sistemas de ecuaciones lineales, expuesto en el párrafo anterior, es muy sencillo y requiere la realización de cálculos de un mismo tipo, que fácilmente se efectúan en las máquinas calculadoras. Sin embargo, su defecto esencial consiste en que no da la posibilidad de formular las condiciones de compatibilidad o de determinabilidad de un sistema mediante sus coeficientes y términos independientes. Por otra parte, incluso en el caso de un sistema determinado, con este método no se pueden hallar fórmulas para expresar la solución del sistema mediante sus coeficientes y términos independientes. Sin embargo, estos sistemas encuentran aplicación en diversas cuestiones teóricas y, en particular, en las investigaciones geométricas. De aquí la necesidad de desarrollar la teoría de los sistemas de ecuaciones lineales con otros métodos más profundos. El caso general va a ser estudiado en el capítulo siguiente, mientras que el contenido del presente capítulo está dedicado al estudio de los sistemas determinados que tienen igual número de ecuaciones y de incógnitas. Comenzaremos por los sistemas con dos y tres incógnitas, ya estudiados en el álgebra elemental.

Sea dado un *sistema de dos ecuaciones lineales con dos incógnitas*

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2, \end{aligned}\right\} \qquad (1)$$

cuyos coeficientes forman una matriz cuadrada de segundo orden

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{2}$$

Aplicando al sistema (1), el método de igualación de los coeficientes obtenemos:

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

Supongamos que $a_{11}a_{22}-a_{12}a_{21}\neq 0$. Entonces,

$$x_1=\frac{b_1a_{22}-a_{12}b_2}{a_{11}a_{22}-a_{12}a_{21}}, \quad x_2=\frac{a_{11}b_2-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}. \tag{3}$$

Sustituyendo en las ecuaciones (1) los valores obtenidos de las incógnitas, es fácil comprobar que (3) es solución del sistema (1); el problema de la unicidad de esta solución se estudiará en el § 7.

El común denominador de los valores de las incógnitas (3) está expresado sencillamente por los elementos de la matriz (2), o sea, es precisamente igual al producto de los elementos de la diagonal principal menos el producto de los elementos de la segunda diagonal. Este número se llama *determinante* de la matriz (2). Se suele decir que es un *determinante de segundo orden*, puesto que la matriz (2) es de segundo orden. Para designar el determinante de la matriz (2), se emplea la siguiente notación: se escribe la matriz (2), pero en lugar de los paréntesis se ponen unas barras verticales; de este modo

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}-a_{12}a_{21}. \tag{4}$$

Ejemplos.

1) $$\begin{vmatrix} 3 & 7 \\ 1 & 4 \end{vmatrix} = 3\cdot 4-7\cdot 1=5;$$

2) $$\begin{vmatrix} 1 & -2 \\ 3 & 5 \end{vmatrix} = 1\cdot 5-(-2)\cdot 3=11$$

Es menester subrayar otra vez más que, mientras la matriz representa una tabla de números, el determinante es un número completamente determinado por la matriz cuadrada. Señalemos que los productos $a_{11}a_{22}$ y $a_{12}a_{21}$, se llaman *términos* del determinante de segundo orden.

Los numeradores de las expresiones (3) tienen la misma forma que el denominador, o sea, también son determinantes de segundo orden: el numerador de la expresión para x_1 es el determinante de una matriz, que se obtiene de la matriz (2) sustituyendo su primera columna por la columna de los términos independientes del sistema

(1); el numerador de la expresión para x_2 es el determinante de una matriz, que se obtiene de la matriz (2) por la misma sustitución de su segunda columna. Las fórmulas (3) se pueden escribir ahora en la forma siguiente:

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2=\frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{5}$$

Esta regla de resolución de un sistema de dos ecuaciones lineales con dos incógnitas (denominada *regla de Cramer*) se expresa del modo siguiente:

Si el determinante (4) *de los coeficientes del sistema de ecuaciones* (1) *es diferente de cero, la solución del sistema* (1) *se obtiene tomando por valores de las incógnitas las fracciones, cuyo común denominador es el determinante* (4) *y cuyo numerador, para la incógnita* x_i $(i=1, 2)$, *es el determinante que se obtiene sustituyendo en el determinante* (4) *la columna i-ésima (o sea, la columna de los coeficientes de la incógnita buscada) por la columna de los términos independientes del sistema* (1)*.

Ejemplo. Resolver el sistema

$$\left.\begin{aligned} 2x_1+x_2&=7, \\ x_1-3x_2&=-2. \end{aligned}\right\}$$

El determinante de los coeficientes es

$$d=\begin{vmatrix} 2 & 1 \\ 1 & -3 \end{vmatrix}=-7;$$

es decir, éste es diferente de cero, por lo cual, se puede aplicar la regla de Cramer al sistema.

Los numeradores para las incógnitas son los determinantes:

$$d_1=\begin{vmatrix} 7 & 1 \\ -2 & -3 \end{vmatrix}=-19, \quad d_2=\begin{vmatrix} 2 & 7 \\ 1 & -2 \end{vmatrix}=-11.$$

Por lo tanto, la solución de nuestro sistema es:

$$x_1=\frac{d_1}{d}=\frac{19}{7}, \quad x_2=\frac{d_2}{d}=\frac{11}{7}.$$

La introducción de los determinantes de segundo orden no aporta simplificaciones esenciales en la resolución de un sistema de dos ecuaciones lineales con dos incógnitas. Sin embargo, los métodos análogos para el caso de *sistemas de tres ecuaciones lineales con tres*

* En este enunciado, para abreviar, se habla de la sustitución de las columnas «en el determinante». A continuación, se va a hablar de un modo semejante, si resulta conveniente, de las filas y columnas del determinante, de sus elementos, diagonales, etc.

incógnitas resultan ya prácticamente útiles. Sea dado un sistema

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{aligned}\right\} \qquad (6)$$

con la matriz de los coeficientes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \qquad (7)$$

Fácilmente se comprueba que, si multiplicamos ambos miembros de la primera de las ecuaciones (6) por $a_{22}a_{33} - a_{23}a_{32}$, ambos miembros de la segunda ecuación por $a_{13}a_{32} - a_{12}a_{33}$, ambos miembros

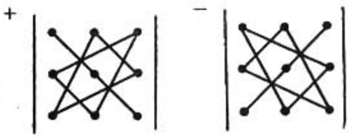

Fig 1.

de la tercera ecuación por $a_{12}a_{23} - a_{13}a_{22}$, y después sumamos estas tres ecuaciones, los coeficientes de x_2 y x_3 resultarán iguales a cero, es decir, que estas incógnitas se eliminarán simultáneamente. De este modo, obtenemos la igualdad

$$(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32})\, x_1 =$$
$$= b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - a_{13}a_{22}b_3 - a_{12}b_2a_{33} - b_1a_{23}a_{32}. \qquad (8)$$

El coeficiente de x_1 en esta igualdad se llama *determinante de tercer orden*, correspondiente a la matriz (7). Para escribirlo, se emplean los mismos símbolos que en el caso de los determinantes de segundo orden; por lo tanto,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \qquad (9)$$

A pesar de que la expresión del determinante de tercer orden es bastante complicada, la ley de su formación con los elementos de la matriz (7) es muy sencilla. En efecto, uno de los tres términos del determinante que figuran en la expresión (9) con el signo más es el producto de los elementos de la diagonal principal; cada uno de los otros dos, es el producto de los elementos situados en la paralela a esta diagonal, por el elemento situado en el ángulo opuesto de la matriz. Los términos que figuran en (9) con signo menos, se forman del mismo modo, pero con respecto a la segunda diagonal. De este modo, obtenemos un método de cálculo de los determinantes de tercer orden, que conduce (teniendo cierta práctica) a un rápido resultado. En la fig. 1, en el esquema de la izquierda, se señala la regla para el cálculo de los términos positivos del determinante de tercer orden, y en el de la derecha, la regla para el cálculo de sus términos negativos.

Ejemplos.

1) $$\begin{vmatrix} 2 & 1 & 2 \\ -4 & 3 & 1 \\ 2 & 3 & 5 \end{vmatrix} = 2\cdot3\cdot5+1\cdot1\cdot2+2\cdot(-4)\cdot3- \\ -2\cdot3\cdot2-1\cdot(-4)\cdot5-2\cdot1\cdot3= \\ =30+2-24-12+20-6=10.$$

2) $$\begin{vmatrix} 1 & 0 & -5 \\ -2 & 3 & 2 \\ 2 & -2 & 0 \end{vmatrix} = 1\cdot3\cdot0+0\cdot2\cdot1+(-5)\cdot(-2)\cdot(-2)- \\ -(-5)\cdot3\cdot1-0\cdot(-2)\cdot0-1\cdot2\cdot(-2)= \\ =-20+15+4=-1.$$

El segundo miembro de la igualdad (8) es también un determinante de tercer orden: es precisamente el determinante de la matriz que se obtiene de la matriz (7), sustituyendo su primera columna por la columna de los términos independientes del sistema (6). Si designamos con la letra d el determinante (9) y con el símbolo d_j ($j = 1, 2, 3$), el determinante que se obtiene de este último, al sustituir su j-ésima columna por la columna de los términos independientes del sistema (6), la igualdad (8) toma la forma $dx_1 = d_1$, de donde, para $d \neq 0$, se deduce que

$$x_1 = \frac{d_1}{d}. \tag{10}$$

Del mismo modo, multiplicando las ecuaciones (6) por los números $a_{23}a_{31} - a_{21}a_{33}$, $a_{11}a_{33} - a_{13}a_{31}$, $a_{13}a_{21} - a_{11}a_{23}$, respectivamente, obtenemos para x_2 la siguiente expresión (siendo $\neq 0$):

$$x_2 = \frac{d_2}{d} \tag{11}$$

Finalmente, multiplicando estas ceuaciones por $a_{21}a_{32} - a_{22}a_{31}$, $a_{12}a_{31} - a_{11}a_{32}$, $a_{11}a_{22} - a_{12}a_{21}$, respectivamente, llegamos a la siguiente expresión para x_3:

$$x_3 = \frac{d_3}{d}. \tag{12}$$

Sustituyendo las expresiones (10), (11) y (12) en la ecuación (6) (se sobrentiende que los determinantes d y todos los d_j están escritos en forma desarrollada), después de unos cálculos bastante complicados, pero asequibles, vemos que se satisfacen todas estas ecuaciones, es decir, que los números (10) (11) y (12) son la solución del sistema. Por lo tanto, *si el determinante de los coeficientes de un sistema de tres ecuaciones lineales con tres incógnitas es diferente de cero, su solución se puede hallar por la regla de Cramer, formulada igualmente que en el caso de un sistema de dos ecuaciones.* En el § 7, el lector hallará, para un caso más general, otra demostración de esta afirmación (que no se basa en los cálculos omitidos), y también la demostración de la unicidad de la solución (10) (11) y (12) del sistema (6).

Ejemplo. Resolver el sistema

$$\left.\begin{aligned} 2x_1-x_2+x_3&=0,\\ 3x_1+2x_2-5x_3&=1,\\ x_1+3x_2-2x_3&=4. \end{aligned}\right\}$$

El determinante de los coeficientes del sistema es diferente de cero:

$$d=\begin{vmatrix} 2 & -1 & 1\\ 3 & 2 & -5\\ 1 & 3 & -2 \end{vmatrix}=28,$$

por eso, se le puede aplicar la regla de Cramer. Los numeradores para las incógnitas son los determinantes

$$d_1=\begin{vmatrix} 0 & -1 & 1\\ 1 & 2 & -5\\ 4 & 3 & 2 \end{vmatrix}=13,\quad d_2=\begin{vmatrix} 2 & 0 & 1\\ 3 & 1 & -5\\ 1 & 4 & -2 \end{vmatrix}=47,$$

$$d_3=\begin{vmatrix} 2 & -1 & 0\\ 3 & 2 & 1\\ 1 & 3 & 4 \end{vmatrix}=21,$$

o sea, que la solución del sistema es el sistema de números

$$x_1=\frac{13}{28},\quad x_2=\frac{47}{28},\quad x_3=\frac{21}{28}=\frac{3}{4}.$$

§ 3. Permutaciones y sustituciones

Para definir y estudiar los determinantes de orden n, necesitamos unos cuantos conceptos y datos referentes a los conjuntos finitos. Sea dado un conjunto finito M, compuesto de n elementos. Estos se pueden numerar, empleando para ello los primeros n números naturales $1, 2, \ldots, n$. Como en las cuestiones que nos interesan, las propiedades individuales de los elementos del conjunto M no van a jugar ningún papel, supondremos simplemente que los mismos números $1, 2, \ldots, n$, representan a los elementos del conjunto M.

Además de la ordenación normal de los números $1, 2, \ldots, n$, éstos pueden ser ordenados de muchos modos. Así, los números $1, 2, 3, 4$ se pueden ordenar también de los modos siguientes: $3, 1, 2, 4$, o bien, $2, 4, 1, 3$, etc. Toda disposición de los números $1, 2, \ldots, n$ en un orden determinado, se llama *permutación* de n números (o de n símbolos).

El número de permutaciones diversas de n símbolos es igual al producto $1 \cdot 2 \ldots n$, designado por $n!$ (se lee: «factorial de n»). En efecto, la forma general de una permutación de n símbolos es $i_1, i_2, \ldots, i_n$, donde cada uno de los símbolos i_s representa uno de los números $1, 2, \ldots, n$; además, ninguno de estos números se repite. En calidad de i_1 se puede tomar cualquiera de los números $1, 2, \ldots, n$; esto ofrece n diferentes posibilidades. Sin embargo, si se ha elegido ya i_1, en calidad de i_2 se puede tomar solamente uno de los $n-1$ números restantes, es decir, que el número de modos de elección de los símbolos i_1 y i_2 es igual al producto $n(n-1)$, etc.

Por lo tanto, el número de permutaciones de n símbolos, para $n = 2$, es igual a $2! = 2$ (las permutaciones 12 y 21; en los ejemplos donde $n \leqslant 9$, no separaremos con comas los símbolos que se permutan); para $n = 3$, este número es igual a $3! = 6$, para $n = 4$, es igual a $4! = 24$. A continuación, con el aumento de n, el número de permutaciones crece extraordinariamente; así, para $n = 5$, es igual a $5! = 120$, y para $n = 10$, es ya igual a 3 628 800.

Si en una permutación cambiamos de lugar dos símbolos cualesquiera (no necesariamente situados uno al lado del otro), permaneciendo todos los demás en sus sitios, obtenemos, evidentemente, una nueva permutación. Esta transformación de la permutación se denomina *trasposición*.

Todas las $n!$ permutaciones de n símbolos se pueden colocar en un orden tal, que cada permutación siguiente se obtenga de la anterior mediante una trasposición, pudiendo además comenzar por cualquiera de ellas.

Esta afirmación es justa para $n = 2$: si se pide empezar por la permutación 12, la disposición buscada es 12, 21; si se pide empezar por la permutación 21, la disposición es 21, 12. Supongamos, que nuestra afirmación ya está demostrada para $n - 1$, y que queremos demostrarla para n. Supongamos, además, que tenemos que comenzar con la permutación

$$i_1, i_2, \ldots, i_n. \tag{1}$$

Consideremos todas las permutaciones de n símbolos en las que i_1 ocupa el primer lugar. En total, resultan $(n-1)!$ permutaciones. Según la tesis del teorema, éstas pueden ser ordenadas del modo indicado y, además, comenzando por la permutación (1), puesto que, en realidad, esto se reduce a la ordenación de todas las permutaciones de $n-1$ símbolos y, por la suposición inductiva, se puede comenzar por cualquier permutación y, en particular, por la permutación $i_2, \ldots, i_n$. En la última de las permutaciones de n símbolos obtenidas de este modo, efectuamos una trasposición del símbolo i_1 con cualquier otro símbolo, por ejemplo, con i_2, y comenzando con la permutación nueva obtenida, ordenamos del modo necesario todas las permutaciones en las que i_2 ocupa el primer lugar, etc. Es evidente, que de este modo se pueden obtener todas las permutaciones de n símbolos.

De este teorema se deduce que *de cualquier permutación de n símbolos se puede pasar a cualquier otra permutación de los mismos símbolos, mediante unas cuantas trasposiciones.*

Se dice que, en una permutación dada, los números i y j forman una *inversión*, si $i > j$, pero en esta permutación, i está antes que j. Una permutación se llama *par*, si sus símbolos forman un número par de inversiones, e *impar*, en el caso contrario. Así, la permutación $1, 2, \ldots, n$ es par para cualquier n, puesto que en ella el número de inversiones es igual a cero. La permutación 451362 ($n = 6$) contiene 8 inversiones y, por consiguiente, es par; la permutación 38524671 ($n = 8$) contiene 15 inversiones y, por consiguiente, es impar.

Toda trasposición cambia la paridad de la permutación.

Para demostrar este importante teorema, consideremos primero el caso en que los símbolos i y j que se trasponen estén uno al lado del otro, es decir, que la permutación tiene la forma$\ldots, i, j, \ldots$, donde los puntos sustituyen a los símbolos que no se alteran con la trasposición. La trasposición convierte a nuestra permutación en la permutación$\ldots, j, i, \ldots$; se comprende, además, que en ambas permutaciones, cada uno de los símbolos i, j, forma unas mismas inversiones con los símbolos que se mantienen en el sitio. Si antes los símbolos i y j no formaban inversión, en la nueva permutación aparece una nueva inversión, o sea, el número de inversiones aumenta en una unidad; si antes los símbolos i y j formaban inversión, ésta ahora desaparece, o sea, el número de inversiones disminuye en una unidad. En ambos casos, cambia la paridad de la permutación.

Supongamos ahora, que entre los símbolos i y j que se trasponen hay intercalados s símbolos, $s > 0$, es decir, que la permutación tiene la forma

$$\ldots, i, k_1, k_2, \ldots, k_s, j, \ldots \tag{2}$$

Se puede obtener la trasposición de los símbolos i y j como resultado de la ejecución consecutiva de $2s+1$ trasposiciones de elementos vecinos. Estas son, precisamente, las trasposiciones que permutan los símbolos i y k_1, a continuación, i (que ya ocupa el lugar del símbolo k_1) y k_2, etc, hasta que i llegue a ocupar el lugar del símbolo k_s. Después de estas s trasposiciones viene la trasposición que permuta los símbolos i y j, y luego, s trasposiciones del símbolo j con todos los k, a consecuencia de lo cual j ocupa el lugar del símbolo i, mientras que los símbolos k vuelven a sus lugares antiguos. Por lo tanto, la paridad de la permutación fue cambiada un número impar de veces, y por esto, la permutación (2) y

$$\ldots,\ j,\ k_1,\ k_2,\ \ldots,\ k_s i,\ \ldots \tag{3}$$

tienen diferente paridad.

Para $n \geqslant 2$, *el número de permutaciones pares de n símbolos es igual al número de permutaciones impares, es decir, es igual a* $\frac{1}{2}n!$

En efecto, basándonos en lo demostrado anteriormente, ordenemos todas las permutaciones de n símbolos de tal modo, que cada una de ellas se obtenga de la anterior mediante una trasposición. Las permutaciones vecinas tendrán entonces paridad contraria, es decir, las permutaciones estarán colocadas de tal manera, que las permutaciones pares e impares se alternarán. Nuestra afirmación se deduce ahora de la observación evidente que para $n \geqslant 2$, el número $n!$ es par.

Introduzcamos ahora un nuevo concepto, el *de sustitución de grado n*. Escribamos, una debajo de otra, dos permutaciones de n símbolos, colocando entre paréntesis las dos filas obtenidas; por ejemplo, para $n = 5$:

$$\begin{pmatrix} 3 & 5 & 1 & 4 & 2 \\ 5 & 2 & 3 & 4 & 1 \end{pmatrix}. \tag{4}$$

En este ejemplo*, bajo el número 3 figura el número 5, bajo el número 5, el número 2, etc. Diremos que el número 3 *se sustituye* por el 5 (o también, que la sustitución transporta el 3 sobre el 5); el número 5, por el 2; el número 1, por el 3; número 4, por el 4 (o que *se queda en el sitio*); y, por fin, el número 2, por el 1. Por lo tanto, dos permutaciones, escritas una bajo la otra en la forma (4), determinan una *aplicación biyectiva*** *del conjunto* de los primeros cinco números

* Por su aspecto, se parece a una matriz de dos filas y 5 columnas, pero tiene un significado totalmente distinto.

** A continuación se empleará frecuentemente la siguiente terminología, admitida en la teoría de conjuntos.

Sean dados dos conjuntos M y N (finitos o infinitos) de elementos de cualquier naturaleza.

Si de un modo determinado, a cada elemento x de M (con la notación $x \in M$ se denota que el elemento x pertenece a M) se pone en correspondencia un elemento y de N ($y \in N$), y sólo uno, se dice que se ha definido una aplinaturales sobre sí mismo, es decir, una aplicación que, a cada uno de los números naturales 1, 2, 3, 4, 5, pone en correspondencia uno de estos mismos números naturales, y que, a diversos números pone en correspondencia números diferentes. Como en total hay cinco números de éstos, o sea, es un conjunto finito, a *cada uno* de estos números corresponderá uno de los números 1, 2, 3, 4, 5, es decir, precisamente el número por el que «se sustituye».

Está claro, que la aplicación biyectiva del conjunto de los cinco primeros números naturales que hemos obtenido mediante (4), se podría haber obtenido también escribiendo, una bajo la otra, otros pares de permutaciones de los cinco símbolos. Estas expresiones se obtienen de (4) mediante unas cuantas trasposiciones de las columnas; tales son, por ejemplo,

$$\begin{pmatrix} 2 & 1 & 5 & 3 & 4 \\ 1 & 3 & 2 & 5 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5 & 2 & 4 & 3 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 5 & 1 & 4 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}. \tag{5}$$

En todas estas expresiones, 3 se sustituye por 5, 5 por 2, etc.

De modo análogo, dos permutaciones de n símbolos, escritas una bajo la otra, determinan una aplicación biyectiva del conjunto de los primeros n números naturales sobre sí mismo. Toda aplicación biyectiva A del conjunto de los primeros n números naturales sobre sí mismo, se llama *sustitución de grado n*. Es evidente, que toda sustitución A se puede expresar mediante dos permutaciones, escritas una bajo la otra

$$A = \begin{pmatrix} i_1, & i_2, & \dots & i_n \\ \alpha_{i_1}, & \alpha_{i_2}, & \dots & \alpha_{i_n} \end{pmatrix}; \tag{6}$$

aquí, mediante α_i se denota el número que en la sustitución A sustituye al número i, $i = 1, 2, \dots, n$.

La sustitución A posee una multitud de expresiones de la forma (6). Así, pues, (4) y (5) son diversas expresiones de una misma sustitución de 5° grado.

Se puede pasar de una expresión de la sustitución A a otra, realizando unas cuantas trasposiciones de las columnas. También se puede obtener una expresión de la forma (6), en la que figure, en la fila superior (o inferior), una permutación prefijada de n símbolos. En

---

cación (o una representación) de M en N. El elemento y se llama en este caso imagen o representación de x.

Si todo elemento de N es imagen de al menos un elemento de M, se dice que se tiene una aplicación de M sobre N (pudiendo ser pluriunívoca, cuando varios elementos de M tienen una misma imagen). En este caso, la aplicación se llama exhaustiva (o sobreyectiva). Si distintos elementos de M tienen distintas imágenes, la aplicación es inyectiva. Una aplicación exhaustiva e inyectiva se llama biyectiva (también suele decirse que entre M y N se ha establecido una correspondencia biunívoca). En este caso, cada elemento $y \in N$ es imagen de un sólo elemento $x \in M$. (*Nota del T.*).

particular, toda sustitución A de grado n se puede expresar en la forma

$$A=\begin{pmatrix}1 & 2 & \dots & n\\ \alpha_1 & \alpha_2 & \dots & \alpha_n\end{pmatrix}, \tag{7}$$

o sea, con la ordenación natural de los números en la fila superior. Escribiéndolas de este modo, las diversas sustituciones se diferenciarán unas de otras por las permutaciones que figuran en las filas inferiores, con lo que llegamos a la conclusión de que el *número de sustituciones de grado n es igual al número de permutaciones de n símbolos, es decir, es igual a n!.*

Un ejemplo de sustitución de grado n es la *sustitución unidad (o idéntica)*

$$E=\begin{pmatrix}1 & 2 & \dots & n\\ 1 & 2 & \dots & n\end{pmatrix},$$

en la que todos los símbolos permanecen en su sitio.

Señalemos que, en la expresión (6) de la sustitución A, las filas superior e inferior desempeñan papeles diferentes y que, por lo general, cambiándolas de sitio, obtenemos una sustitución diferente. Así pues, las sustituciones de 4° grado

$$\begin{pmatrix}2 & 1 & 4 & 3\\ 4 & 3 & 1 & 2\end{pmatrix} \text{ y } \begin{pmatrix}4 & 3 & 1 & 2\\ 2 & 1 & 4 & 3\end{pmatrix}$$

son distintas: en la primera, el número 2 se sustituye por el 4, mientras que en la segunda, por el 3.

Tomemos una expresión arbitraria (6) de una sustitución A de grado n. Las permutaciones que forman las filas superior e inferior de esta expresión pueden ser de igual paridad o de paridad contraria. Como ya sabemos, el paso a cualquier otra expresión de la sustitución A se puede realizar mediante la ejecución consecutiva de unas cuantas trasposiciones en la fila superior y las trasposiciones correspondientes en la fila inferior. Por otra parte, al efectuar una trasposición en la fila superior de la expresión (6) y una trasposición de los elementos correspondientes en la fila inferior, las paridades de ambas permutaciones cambian simultáneamente, manteniéndose así la coincidencia o la contrariedad de estas paridades. De aquí se deduce que, *en todas las expresiones de la sustitución, las paridades de las filas superior e inferior coinciden, o bien, en todas estas expresiones, las paridades son contrarias.* En el primer caso, se dice que la sustitución A es *par*, en el segundo, que es *impar*. En particular, la sustitución unidad es par.

Si la sustitución A está escrita en la forma (7), es decir, que en la fila superior figura la permutación par $1, 2, \dots, n$, entonces, la paridad de la sustitución A se determinará por la paridad de la permutación $\alpha_1, \alpha_2, \ldots, \alpha_n$ que figura en la fila inferior. De esto se deduce que el *número de las permutaciones pares de grado n es igual al número de las impares, es decir, es igual a* $\frac{1}{2}n!$.

A la definición de la paridad de las sustituciones se le puede dar otra forma un poco diferente. Si en la expresión (6), las paridades de ambas filas coinciden, el número de inversiones, o es par en ambas filas, o es impar en las dos, es decir, el número total de inversiones en las dos filas de la expresión (6) es par; si las paridades de las filas de la expresión (6) son contrarias, el número total de inversiones en estas dos filas es impar. Por lo tanto, *la sustitución A será par, si el número total de inversiones en las dos filas de cualquiera de sus expresiones es par, e impar, en el caso contrario.*

Ejemplo. Sea dada la sustitución de quinto grado

$$\begin{pmatrix} 3 & 1 & 4 & 5 & 2 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix}$$

En la fila superior hay 4 inversiones y en la inferior 7. El número total de inversiones en las dos filas es igual a 11 y, por consiguiente, la sustitución es impar.

Escribamos esta sustitución en la forma

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 4 & 3 \end{pmatrix}.$$

El número de inversiones en la fila superior es 0 y en la inferior es 5, es decir, el número total de inversiones es de nuevo impar. Vemos, pues, que en diversas expresiones de la sustitución se conserva la paridad, pero no el mismo número total de inversiones.

Queremos señalar ahora otras formas de definición de la paridad de las sustituciones que son equivalentes a las expuestas anteriormente*. Para este fin, definiremos el *producto de sustituciones* (que es también de particular interés). Como ya sabemos, la sustitución de grado n es una aplicación biyectiva del conjunto de los números $1, 2, \ldots, n$, sobre sí mismo. El resultado de la realización consecutiva de dos aplicaciones biyectivas del conjunto $1, 2, \ldots, n$ sobre sí mismo es, evidentemente, una nueva aplicación biyectiva de este conjunto sobre sí mismo, es decir, la realización consecutiva de dos sustituciones de grado n da lugar a otra sustitución tercera de grado n, absolutamente determinada. Esta última se llama *producto* de la primera de las sustituciones dadas por la segunda. Así, sean dadas las sustituciones de cuarto grado

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}.$$

* Estas se necesitarán solamente en el capítulo 14 y por esto, en la primera lectura, este material se puede omitir.

entonces

$$AB=\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 1 & 2 & 3\end{pmatrix}.$$

En efecto, en la sustitución A el símbolo 1 se sustituye por el 3, pero en la sustitución B el símbolo 3 se sustituye por el 4, por lo tanto, en la sustitución AB el símbolo 1 se sustituye por el 4, etc.

Solamente se pueden multiplicar las sustituciones de un mismo grado. *Para $n \geqslant 3$, el producto de las sustituciones de grado n no es conmutativo.* En efecto, para las sustituciones A y B consideradas anteriormente, el producto BA tiene la forma

$$BA=\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 4 & 2 & 1\end{pmatrix}$$

o sea, la sustitución BA es diferente de la sustitución AB. Para todas las n ($n \geqslant 3$) se pueden mostrar ejemplos de este tipo, a pesar de que para algunos pares de sustituciones se pueda cumplir eventualmente la ley conmutativa.

El producto de las sustituciones es asociativo, es decir, que se puede hablar del producto de un número finito cualquiera de sustituciones de grado n, tomados (en vista de que no se cumple la ley conmutativa) en un orden determinado. En efecto, sean dadas las sustituciones A, B y C. Supongamos que en la sustitución A el símbolo i_1, $1 \leqslant i_1 \leqslant n$ se sustituye por el símbolo i_2; en la sustitución B, el símbolo i_2 se sustituye por el símbolo i_3 y en la sustitución C, este último se sustituye por el símbolo i_4. Entonces, en la sustitución AB, el símbolo i_1 se sustituirá por el i_3, en la sustitución BC, el símbolo i_2 se sustituirá por el i_4. Por consiguiente, en la sustitución $(AB)C$, así como en la sustitución $A\,(BC)$, el símbolo i_1 se sustituirá por el símbolo i_4.

Es evidente, que el *producto de cualquier sustitución A por la sustitución unidad E, y también el producto de E por A, son iguales a A*:

$$AE=EA=A.$$

Finalmente, denominaremos *inversa* de la sustitución A a una sustitución del mismo grado A^{-1}, que cumpla las condiciones

$$AA^{-1}=A^{-1}A=E.$$

Fácilmente se observa que la inversa de la sustitución

$$A=\begin{pmatrix}1 & 2 & \dots & n\\ \alpha_1 & \alpha_2 & & \alpha_n\end{pmatrix}$$

es la sustitución

$$A^{-1}=\begin{pmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_n\\ 1 & 2 & \dots & n\end{pmatrix},$$

que se obtiene de la sustitución A, permutando la fila superior con la inferior.

Veamos ahora unas sustituciones de una forma especial, que se obtienen de la sustitución unidad E mediante una trasposición efectuada en su fila inferior. Tales sustituciones son impares, se llaman *trasposiciones* y tienen la forma:

$$\begin{pmatrix}\dots & i & \dots & j & \dots\\ \dots & j & \dots & i & \dots\end{pmatrix}, \tag{8}$$

donde los puntos suspensivos sustituyen a los símbolos que permanecen en su sitio. Convengamos en designar esta trasposición con la notación (i, j). La aplicación de la trasposición de los símbolos i, j a la fila inferior de la expresión (7) de una sustitución arbitraria A, es equivalente a multiplicar la sustitución A a la derecha por la sustitución (8), es decir, por (i, j). Ya sabemos que todas las permutaciones de n símbolos se pueden obtener de una de ellas, por ejemplo, de la permutación $1, 2, \dots, n$, realizando trasposiciones consecutivas; por eso, toda sustitución se puede obtener de la sustitución idéntica mediante la realización sucesiva de unas cuantas trasposiciones en la fila inferior, es decir, mediante una multiplicación sucesiva por sustituciones de la forma (6). Por consiguiente, se puede afirmar (omitiendo el factor E), que *toda sustitución se puede representar en forma de un producto de trasposiciones.*

Toda sustitución se puede descomponer de muchas maneras diversas en un producto de trasposiciones. Por ejemplo, siempre se pueden agregar dos factores iguales de la forma (i, j) (i, j) que, al multiplicarlos, darán las sustitución E, es decir, que se eliminan mutuamente. Señalemos un ejemplo menos trivial:

$$\begin{pmatrix}1 & 2 & 3 & 4 & 5\\ 2 & 5 & 4 & 3 & 1\end{pmatrix}=(12)(15)(34)=(14)(24)(45)(34)(13).$$

El nuevo método de determinación de la paridad de una sustitución se basa en el teorema siguiente:

En todas las descomposiciones de una sustitución en producto de trasposiciones, la paridad del número de estas trasposiciones es la misma, y coincide con la paridad de la sustitución misma.

Así, la sustitución del ejemplo considerado anteriormente es impar, como se puede comprobar calculando el número de inversiones.

El teorema quedará demostrado si se muestra que el *producto de cualesquiera k trasposiciones es una sustitución, cuya paridad* *coincide con la paridad del número k.* Para $k = 1$ esto es cierto, puesto que una trasposición es una sustitución impar. Supongamos que ya está demostrada nuestra afirmación para el caso de $k - 1$ factores. Entonces, su validez para k factores se deducirá de que los números $k - 1$ y k son de paridad contraria, y el producto de una sustitución (en el caso considerado, del producto de los primeros $k - 1$ factores) por una trasposición es equivalente a la realización de esta trasposición en la fila inferior de la sustitución, es decir, cambia su paridad.

Un método muy cómodo de expresión de las sustituciones, que permite hallar fácilmente su paridad, es la *descomposición en ciclos.* Toda sustitución de grado n puede dejar en el sitio algunos de los símbolos $1, 2, \ldots, n$, otros, verdaderamente los puede transportar.

Una sustitución se llama *sustitución circular o ciclo* si al repetirla un número suficiente de veces, cada uno de los símbolos que verdaderamente se transportan puede ser transportado sobre cualquiera otro de estos símbolos. Tal es, por ejemplo, la sustitución de octavo grado

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 8 & 6 & 4 & 5 & 2 & 7 & 3 \end{pmatrix};$$

ésta verdaderamente transporta los símbolos 2, 3, 6 y 8, a saber, el símbolo 2 sobre el 8, el símbolo 8 sobre el 3, el símbolo 3 sobre el 6 y el símbolo 6 de nuevo sobre el 2.

Todas las trasposiciones pertenecen al conjunto de los ciclos. Por analogía con la forma abreviada de expresión de las trasposiciones que se había empleado anteriormente, para los ciclos se usa la siguiente forma de expresión: los símbolos que verdaderamente son transportados se escriben entre paréntesis uno tras otro, en el mismo orden en que se sustituyen unos por otros al repetir la sustitución; la expresión comienza por cualquiera de los símbolos que verdaderamente se transportan y termina con el símbolo que se transporta sobre el primero. Así, para el ejemplo indicado anteriormente, esta expresión tiene la forma:

$$(2\ 8\ 3\ 6).$$

El número de símbolos que verdaderamente son transportados en el ciclo se llama *longitud* del mismo.

Se dice que dos ciclos de grado n son *independientes*, si no tienen símbolos comunes que verdaderamente sean transportados. Se comprende que, al multiplicar ciclos independientes, el orden de los factores no influye en el resultado.

Toda sustitución se puede descomponer de modo único en un producto de ciclos independientes dos a dos. La demostración de esta afirmación no representa dificultad alguna y la omitimos. La descomposición se realiza del modo siguiente: comenzamos por cualquiera de los símbolos que verdaderamente se transportan y escribimos tras él aquellos símbolos sobre los que éste se transporta al repetir la sustitución. Continuamos así, hasta que volvamos a obtener el símbolo inicial. Después de que «se cierre» este ciclo, comenzamos con uno de los símbolos que quedan y que verdaderamente se transportan, obteniendo así el segundo ciclo, etc.

Ejemplos

1) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix} = (13)\,(254).$

2) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 2 & 8 & 7 & 6 & 1 & 4 & 3 \end{pmatrix} = (156)\,(38)\,(47).$

Recíprocamente, para cada sustitución, dada mediante su descomposición en ciclos independientes, se puede hallar una expresión en la forma ordinaria (con la condición de que se conozca el grado de la sustitución). Por ejemplo:

$$3)\ (1372)\,(45)=\begin{pmatrix}1&2&3&4&5&6&7\\3&1&7&5&4&6&2\end{pmatrix},$$

si se sabe que el grado de esta sustitución es igual a 7.

Sea dada una sustitución de grado n, y sea s el número de ciclos independientes en su descomposición, más el número de símbolos que permanecen en su sitio*. La diferencia n-s se llama *decremento* de la sustitución. Es evidente, que el decremento es igual al número de los símbolos que verdaderamente se transportan, menos el número de ciclos independientes que forman parte de la descomposición de la sustitución. Para los ejemplos 1), 2) y 3), considerados anteriormente, el decremento es igual a 3, 4, y 4, respectivamente.

La paridad de una sustitución coincide con la paridad del decremento de ella.

En efecto, todo ciclo de longitud k se puede representar en forma de un producto de $k-1$ trasposiciones del modo siguiente:

$$(i_1,\ i_2,\ \ldots,\ i_k)=(i_1,\ i_2)\,(i_1,\ i_3)\ \ldots\ (i_1,\ i_k).$$

Supongamos dada la descomposición de la sustitución A en ciclos independientes. Si se descompone cada uno de los ciclos en el producto de las trasposiciones que acabamos de indicar, obtendremos la expresión de la sustitución A en forma de un producto de trasposiciones. El número de estas trasposiciones será, evidentemente, menor que el número de los símbolos que verdaderamente son transportados por la sustitución A, en un número igual al número de los ciclos independientes en la descomposición de la sustitución. De aquí se deduce, que la sustitución A se puede descomponer en un producto de trasposiciones, cuyo número es igual al decremento. Por consiguiente, la paridad de la sustitución se determina por la paridad del decremento.

§ 4. Determinantes de n-ésimo orden

Queremos generalizar ahora para el caso de un n arbitrario, los resultados obtenidos en el § 2 para $n = 2$ y 3. Con este fin, es necesario definir los determinantes de n-ésimo orden. Sin embargo, es imposible hacer esto del mismo modo que se introdujeron los determinantes de segundo y tercer orden, es decir, resolviendo en forma general un sistema de ecuaciones lineales, pues, a medida que aumentase n, los cálculos se harían más y más complicados, y siendo n arbitrario, éstos serían prácticamente irrealizables. Procederemos de otro modo. Examinaremos los determinantes de segundo y tercer orden ya conocidos. Procuraremos establecer una ley general, de acuerdo a la cual se expresan estos determinantes mediante los elementos de las matrices correspondientes y tomaremos esta ley por definición para el determinante de orden n. Después demostraremos que con esta definición sigue cumpliéndose la regla de Cramer.

* A todo símbolo que se mantiene en su sitio se podía haber puesto en correspondencia un «ciclo» de longitud 1, es decir, que en el ejemplo 2), indicado anteriormente, se podría escribir: (156) (38) (47) (2). Sin embargo, no procederemos de este modo.

Recordemos las expresiones de los determinantes de segundo y tercer orden:

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

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

Obsérvese que todo término del determinante de segundo orden es un producto de dos elementos, situados en diversas filas y en diversas columnas. Además, todos los productos de este tipo que se pueden formar con los elementos de la matriz de segundo orden (en total son dos), se han utilizado como términos del determinante. De modo semejante, todo término del determinante de tercer orden representa un producto de tres elementos, tomados también uno a uno de cada fila y de cada columna. Todos los productos de estos se utilizan también como términos del determinante.

Sea dada ahora una matriz cuadrada de orden n

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \tag{1}$$

Consideremos todos los productos posibles de n elementos de esta matriz, situados en diferentes filas y en diferentes columnas, o sea, los productos de la forma

$$a_{1\alpha_1}\ a_{2\alpha_2}\ \dots\ a_{n\alpha_n}, \tag{2}$$

donde los subíndices $\alpha_1, \alpha_2, \dots, \alpha_n$ forman una de las permutaciones de los números $1, 2, \dots, n$. El número de estos productos es igual al número de las diversas permutaciones de n símbolos, es decir, es igual a $n!$. Vamos a tomar todos estos productos por términos del futuro determinante de n-ésimo orden, correspondiente a la matriz (1).

Para determinar el signo con que figura el producto (2) en el determinante, observemos que con los subíndices de este producto se puede formar la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}, \tag{3}$$

donde i se sustituye por α_i, si el elemento situado en la i-ésima fila y en la α_i-ésima columna de la matriz (1) forma parte del producto (2). Examinando las expresiones de los determinantes de segundo y tercer orden, observamos que en ellos figuran con signo más los términos cuyos subíndices forman una sustitución par, y con signo menos, los términos cuyos subíndices forman una sustitución impar. Resulta natural conservar también esta ley en la definición del determinante de n-ésimo orden.

Por lo tanto, llegamos a la siguiente definición: se llama determinante de n-ésimo orden, correspondiente a la matriz (1), a la suma algebraica de $n!$ términos, constituida del modo siguiente: son términos de ella todos los productos posibles de n elementos de la matriz, tomados uno de cada fila y de cada columna, tomando el término *con signo más, si sus subíndices forman una sustitución par, y con* signo menos, en el caso contrario.

Para escribir el determinante de n-ésimo orden correspondiente a la matriz (1) se empleará la notación que se usó en el caso de los determinantes de segundo y tercer orden:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \qquad (4)$$

Los determinantes de n-ésimo orden, para $n = 2$ y $n = 3$, se convierten en los determinantes de segundo y tercer orden considerados anteriormente; para $n = 1$, es decir, para las matrices constituidas de un sólo elemento, el determinante es igual al elemento mismo. Sin embargo, por ahora, todavía no sabemos si para $n > 3$ se pueden utilizar los determinantes de n-ésimo orden para la resolución de sistemas de ecuaciones lineales. Esto se mostrará en el § 7; pero previamente tenemos que estudiar detalladamente los determinantes de n-ésimo orden y, en particular, tenemos que hallar un método para su cálculo, puesto que sería muy difícil calcular los determinantes partiendo de su definición, incluso para n no muy grandes.

Ahora estableceremos las propiedades elementales de los determinantes de n-ésimo orden, relativas fundamentalmente a una de las dos cuestiones. Por una parte, nos interesarán las condiciones para que el determinante sea igual a cero; por otra parte, señalaremos unas transformaciones de la matriz que no alteran a su determinante o que proporcionan una alteración de éste, fácilmente calculable.

Llamaremos *transposición* de la matriz A a una transformación de la misma, según la cual sus filas se sustituyen por sus columnas del mismo orden, es decir, el paso de la matriz (1) a la matriz

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}; \qquad (5)$$

se puede decir que transponer la matriz (1) es hacerla girar alrededor de la diagonal principal. Correspondientemente, se dice, que el determinante

$$\begin{vmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} \tag{6}$$

se obtiene transponiendo el determinante (4).

Propiedad 1. *El determinante no varía al transponerlo.*

En efecto, todo término del determinante (4) es de la forma

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n}, \tag{7}$$

donde los segundos subíndices forman una permutación de los símbolos $1, 2, \dots, n$. Pero, todos los factores del producto (7) se mantienen también en el determinante (6) en diferentes filas y en diferentes columnas, es decir, que (7) es también un término del determinante transpuesto. Es evidente que lo recíproco también es justo. Por lo tanto, los determinantes (4) y (6) están constituidos por los mismos términos. El signo del término (7) en el determinante (4) se determina por la paridad de la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}; \tag{8}$$

en el determinante (6), los primeros subíndices de los elementos indican el número de orden de la columna, mientras que los segundos subíndices indican el número de orden de la fila. Por consiguiente, en el determinante (6) al término (7) corresponde la sustitución

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ 1 & 2 & \dots & n \end{pmatrix}. \tag{9}$$

Por lo general, las sustituciones (8) y (9) son diferentes, pero, evidentemente, tienen una misma paridad y, por lo tanto, el término (7) tiene un mismo signo en ambos determinantes. Por consiguiente, los determinantes (4) y (6) representan sumas de términos iguales, tomados con signos iguales, es decir, son iguales entre sí.

De la propiedad 1 se deduce que cualquier afirmación sobre las filas del determinante es válida también para sus columnas y viceversa, es decir, en el *determinante* (a distinción de las matrices), las *filas y las columnas gozan de los mismos derechos*. Partiendo de esto, las siguientes ocho propiedades (2-9) se enunciarán y se demostrarán solamente para las filas del determinante; las propiedades análogas para las columnas no necesitarán una demostración especial.

Propiedad 2. *Si una de las filas del determinante está constituida por ceros, el determinante es igual a cero.*

En efecto, *supongamos que* todos los elementos de la i-ésima fila del determinante son iguales a cero. En cada uno de los términos del determinante tiene que estar incluido uno de los elementos de la i-ésima fila, por lo cual, en nuestro caso, todos los términos del determinante son iguales a cero.

Propiedad 3. *Si un determinante se obtiene de otro permutando dos filas, todos los términos del primer determinante serán términos del segundo, pero con signos contrarios, es decir, al permutar dos filas, el determinante sólo cambia de signo.*

En efecto, supongamos que en el determinante (4) se permutan la i-ésima y la j-ésima filas, $i \neq j$, y que todas las demás filas se mantienen en su sitio. Obtenemos el determinante

$$\begin{array}{l} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \end{array} \begin{array}{l} \\ \\ (i) \\ \\ (j) \\ \\ \\ \end{array} \qquad (10)$$

(al margen están señalados los números de las filas). Si

$$a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} \qquad (11)$$

es un término del determinante (4), evidentemente, todos sus factores se mantienen también en el determinante (10) en diferentes filas y columnas. Por lo tanto, los determinantes (4) y (10) constan de los mismos términos. En el determinante (4) al término (11) le corresponde la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_i & \dots & \alpha_j & \dots & \alpha_n \end{pmatrix}, \qquad (12)$$

mientras que en el determinante (10), la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & j & \dots & i & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_i & \dots & \alpha_j & \dots & \alpha_n \end{pmatrix}, \qquad (13)$$

puesto que el elemento $a_{i\alpha_i}$, por ejemplo, está ahora en la j-ésima fila, pero se mantiene en la α_i*-ésima columna anterior. Sin embargo,* la sustitución (13) se obtiene de la sustitución (12) mediante una trasposición en la fila superior, o sea, tiene paridad contraria. De esto se deduce, que todos los términos del determinante (4) forman parte del determinante (10), pero con signos contrarios, es decir, los determinantes (4) y (10) se diferencian entre sí solamente en el signo.

Propiedad 4. *Un determinante que tiene dos filas iguales es igual a cero.*

En efecto, supongamos que el valor del determinante es igual a d y que son iguales entre sí los elementos correspondientes de su i-ésima y j-ésima filas ($i \neq j$). En virtud de la propiedad 3, después de permutar estas dos filas, el determinante se hace igual a $-d$. Sin embargo, como las filas que se permutan son iguales el determinante, en realidad, no varía, o sea, $d = -d$; de donde $d = 0$.

Propiedad 5. *Si se multiplican todos los elementos de una fila del determinante por un número k, el mismo determinante queda multiplicado por k.*

Supongamos que se han multiplicado por k todos los elementos de la i-ésima fila. Cada término del determinante contiene exactamente un elemento de la i-ésima fila. Por lo tanto, todo términc adquiere el factor k, es decir, el mismo determinante queda multiplicado por k.

Esta propiedad también se puede expresar así: *el factor común de todos los elementos de una fila del determinante se puede sacar fuera del signo de éste.*

Propiedad 6. *Un determinante que tiene dos filas proporcionales es igual a cero.*

Supongamos que los elementos de la j-ésima fila del determinante se diferencian de los elementos correspondientes de la i-ésima fila ($i \neq j$) en un mismo factor k. Sacando este factor común k de la j-ésima fila fuera del signo del determinante, obtenemos un determinante con dos filas iguales. Este será igual a cero, por la propiedad 4.

La propiedad 4, así como la propiedad 2 para $n > 1$, son, evidentemente, casos particulares de la propiedad 6 (para $k = 1$ y $k = 0$).

Propiedad 7. *Si todos los elementos de la i-ésima fila de un determinante de n-ésimo orden representan una suma de dos sumandos:*

$$a_{ij} = b_j + c_j, \quad j = 1, \ldots, n,$$

el determinante es igual a la suma de dos determinantes, en los que todas las filas, menos la i-ésima, coinciden con las del determinante dado, mientras que la i-ésima fila de uno de los sumandos consta de los elementos b_j y la del otro, de los elementos c_j.

Todo término del determinante dado se puede representar de la forma

$$a_{1\alpha_1}a_{2\alpha_2} \ldots a_{i\alpha_i} \ldots a_{n\alpha_n} = a_{1\alpha_1}a_{2\alpha_2} \ldots (b_{\alpha_i} + c_{\alpha_i}) \ldots a_{n\alpha_n} =$$
$$= a_{1\alpha_1}a_{2\alpha_2} \ldots b_{\alpha_i} \ldots a_{n\alpha_n} + a_{1\alpha_1}a_{2\alpha_2} \ldots c_{\alpha_i} \ldots a_{n\alpha_n}.$$

Rouniendo los primeros términos de estas sumas (con los mismos signos que tenían los términos correspondientes en el determinante dado), obtenemos un determinante de orden n, que solamente se diferencia del dado, en que en la i-ésima fila, en lugar de los elementos a_{ij}, figuran los elementos b_j. Correspondientemente, los segundos sumandos forman un determinante en cuya i-ésima fila figuran los elementos c_j. Por lo tanto,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_1+c_1 & b_2+c_2 & \dots & b_n+c_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ b_1 & b_2 & \dots & b_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ c_1 & c_2 & \dots & c_n \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

La propiedad 7 se generaliza sin dificultad al caso en que todo elemento de la i-ésima fila es una suma, no de dos, sino de m sumandos, $m \geqslant 2$.

Se dice que la i-ésima fila de un determinante es *combinación lineal* de las demás filas, si para cada fila del número de orden j, $j = 1, \dots, i-1, i+1, \dots, n$, se puede señalar un número k_j tal, que multiplicando la j-ésima fila por k_j y agregando después todas las filas, menos la i-ésima (la suma de las filas se debe entender como la suma por separado de los elementos de todas estas filas en cada columna), se obtiene la i-ésima fila. Algunos de los coeficientes k_j pueden ser iguales a cero, es decir, en realidad, la i-ésima fila es combinación lineal, no de todas, sino de algunas filas restantes. En particular, si solamente uno de los coeficientes k_j es diferente de cero, obtenemos el caso de proporcionalidad de dos filas. Finalmente, si una fila se compone totalmente de ceros, ésta siempre será combinación lineal de las demás filas: caso en que todos los k_j son iguales a cero.

Propiedad 8. *Si una de las filas del determinante es combinación lineal de las demás, el determinante es igual a cero.*

Sea, por ejemplo, la i-ésima fila, combinación lineal de las otras s filas, $1 \leqslant s \leqslant n-1$. Entonces, todo elemento de la i-ésima fila será una suma de s términos. Por lo tanto, aplicando la propiedad 7, representamos nuestro determinante en forma de una suma de determinantes, en cada uno de los cuales la i-ésima fila será proporcional a una de las otras filas. Según la propiedad 6, todos estos determinantes son iguales a cero; por consiguiente, también será igual a cero el determinante dado.

Esta propiedad es una generalización de la propiedad 6, y, como se demostrará en el § 10, es el caso más general de igualdad a cero del determinante.

Propiedad 9. *El determinante no varía si a los elementos de una de sus filas se agregan los elementos correspondientes de otra fila, multiplicados por un mismo número.*

Supongamos que a la i-ésima fila del determinante d se le agrega la j-ésima fila, $j \neq i$, multiplicada por el número k, es decir, que en el nuevo determinante todo elemento de la i-ésima fila tiene la forma $a_{is} + ka_{js}$, $s = 1, 2, \ldots, n$. Entonces, de acuerdo a la propiedad 7, este determinante es igual a la suma de dos determinantes, el primero de los cuales es d, mientras que el segundo contiene dos filas proporcionales y, por ello, es igual a cero.

Como el número k puede ser negativo, el *determinante tampoco variará al restar de una de sus filas otra fila, multiplicada por un número.* En general, *el determinante no varía si a una de sus filas se agrega cualquier combinación lineal de las demás.*

Veamos el siguiente ejemplo. Un determinante se llama *antisimétrico*, si sus elementos, situados simétricamente respecto de la diagonal principal, se diferencian entre sí solamente en el signo, es decir, si para todos i y j se tiene $a_{ji} = -a_{ij}$; de esto se deduce que para todo i será $a_{ii} = -a_{ii} = 0$. Por lo tanto, el determinante tiene la forma

$$d = \begin{vmatrix} 0 & a_{12} & a_{13} & \ldots & a_{1n} \\ -a_{12} & 0 & a_{23} & \ldots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \ldots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -a_{1n} & -a_{2n} & -a_{3n} & \ldots & 0 \end{vmatrix}.$$

Multiplicando cada fila de este determinante por -1, obtenemos el determinante transpuesto, que es de nuevo igual a d, de donde, en virtud de la propiedad 5, resulta:

$$(-1)^n d = d.$$

Para n impar, se deduce que: $-d = d$, es decir, $d = 0$. Por lo tanto, *todo determinante antisimétrico (o hemisimétrico) de orden impar es igual a cero.*

§ 5. Los menores y sus complementos algebraicos

Antes se había indicado que sería difícil calcular un determinante de n-ésimo grado aplicando directamente su definición, o sea, escribiendo cada vez todos los $n!$ términos, determinando sus signos, etc. Existen métodos más sencillos para calcular los determinantes, basados en el hecho de que un determinante de orden n se puede expresar mediante determinantes de órdenes inferiores. Introduzcamos, con este fin, el siguiente concepto.

Sea dado un determinante d de orden n. Tomemos un número entero k que satisfaga la condición $1 \leqslant k \leqslant n - 1$, y elijamos arbitrariamente en el determinante d, k filas y k columnas. Los elementos situados en las intersecciones de estas filas y de estas columnas, es decir, pertenecientes a una de las filas y a una de las columnas elegidas, forman, evidentemente, una matriz de orden k. El determinante de esta matriz se llama *menor de orden k* del determinante d. Se puede decir también que el menor de orden k es el determinante que se obtiene después de suprimir $n - k$ filas y $n - k$ columnas en el determinante d. En particular, después de haber suprimido en el determinante una fila y una columna, obtenemos un menor de orden $(n - 1)$; por otra parte, los mismos elementos del determinante d por separado representan menores de primer orden.

Supongamos que en un determinante d de n-ésimo orden se ha tomado un menor M de orden k. Suprimiendo las filas y columnas, en cuyas intersecciones figura este menor, resulta un menor M' de $(n - k)$-ésimo orden, denominado *menor complementario* del menor M. Suprimiendo, por el contrario, las filas y columnas en las que están situados los elementos del menor M', obtendremos el menor M. Por lo tanto, se puede hablar de un *par de menores complementarios entre sí* del determinante. En particular, el elemento a_{ij} y el menor de $(n - 1)$-ésimo orden que se obtiene suprimiendo en el determinante la i-ésima fila y la j-ésima columna, formarán un par de menores complementarios entre sí.

Si un menor M de k-ésimo orden está situado en las filas de orden $i_1, i_2, \ldots, i_k$ y en las columnas de orden $j_1, j_2, \ldots, j_k$, entonces, denominaremos *complemento algebraico* del menor M a su menor complementario M', tomado con el signo más o menos, según que sea par o impar la suma de los números de orden de todas las filas y columnas en las que está situado el menor M, es decir la suma

$$s_M = i_1 + i_2 + \ldots + i_k + j_1 + j_2 + \ldots j_k. \tag{1}$$

En otras palabras, el complemento algebraico del menor M es el número $(-1)^{s_M} M'$.

El producto de cualquier menor M de k-ésimo orden por su complemento algebraico en el determinante d es una suma algebraica, cuyos sumandos, obtenidos al multiplicar los términos del menor M por los términos del menor complementario M' tomados con el signo $(-1)^{s_M}$, son ciertos términos del determinante d, coincidiendo sus signos en esta suma con los signos que tienen en el determinante.

Comenzaremos la demostración de este teorema con el caso en que el menor M está situado en el ángulo superior de la izquierda del determinante:

$$d = \left|\begin{array}{ccc|ccc} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \ldots & M & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ \hline a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & M' & \ldots \\ a_{n1} & \ldots & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{array}\right|,$$

es decir, en las filas cuyos números de orden son $1, 2, \ldots, k$ y en las columnas que tienen los mismos números de orden. Entonces, el menor M' ocupará el ángulo inferior de la derecha del determinante. En este caso, el número s_M es par:

$$s_M = 1+2+\ldots+k+1+2+\ldots+k = 2(1+2+\ldots+k),$$

por eso, el mismo menor M' sirve de complemento algebraico para M.

Tomemos un término arbitrario del menor M

$$a_{1\alpha_1}a_{2\alpha_2}\ldots a_{k\alpha_k}; \tag{2}$$

su signo en M será $(-1)^l$, donde l es el número de inversiones en la sustitución

$$\begin{pmatrix} 1 & 2 & \ldots & k \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k \end{pmatrix}. \tag{3}$$

El término arbitrario del menor M'

$$a_{k+1,\ \beta_{k+1}}\ a_{k+2,\ \beta_{k+2}}\ \ldots\ a_n\beta_n \tag{4}$$

tiene en éste el signo $(-1)^{l'}$, donde l' es el número de inversiones en la sustitución

$$\begin{pmatrix} k+1 & k+2 & \ldots & n \\ \beta_{k+1} & \beta_{k+2} & \ldots & \beta_n \end{pmatrix}. \tag{5}$$

Multiplicando los términos (2) y (4), obtenemos el producto de n elementos

$$a_{1\alpha_1}a_{2\alpha_2}\ldots a_{k\alpha_k}a_{k+1,\beta_{k+1}}a_{k+2,\beta_{k+2}}\ldots a_{n\beta_n}, \tag{6}$$

situados en diferentes filas y columnas del determinante; por consiguiente, éste será un término del determinante d. El signo del término (6) en el producto MM' será igual al producto de los signos de los términos (2) y (4), o sea, $(-1)^l\cdot(-1)^{l'} = (-1)^{l+l'}$. Sin embargo, el término (6) tiene también este mismo signo en el determinante d. En efecto, la fila inferior de la sustitución

$$\begin{pmatrix} 1 & 2 & \ldots & k & k+1 & k+2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k & \beta_{k+1} & \beta_{k+2} & \ldots & \beta_n \end{pmatrix},$$

formada por los índices de este término, contiene solamente $l+l'$ inversiones, puesto que ningún α puede formar inversión con ningún β: todos los α no son mayores que k, mientras que todos los β no son menores que $k+1$.

De este modo, queda demostrado el caso particular considerado del teorema. Pasemos a examinar el caso general. Supongamos que el menor M está situado en las filas que tienen los números de orden

$i_1, i_2, \ldots, i_k$ y en las columnas que tienen los números de orden $j_1, j_2, \ldots, j_k$, siendo

$$i_1 < i_2 < \ldots < i_k, \quad j_1 < j_2 < \ldots < j_k.$$

Trasponiendo las filas y las columnas, procuremos llevar el menor M al ángulo superior de la izquierda, de modo que no se altere el menor complementario. Con este fin, trasponemos la i_1-ésima fila con la (i_1-1)-ésima, después, con la (i_1-2)-ésima, etc., hasta que la i_1-ésima fila ocupe el lugar de la primera; para esto, tendremos que trasponer las filas i_1-1 veces. Después, trasponemos sucesivamente la i_2-ésima fila con todas las filas situadas sobre ella, hasta que se sitúe directamente debajo de la i_1-ésima fila, es decir, en el sitio que ocupaba la segunda fila antes de todas las transformaciones; como es fácil comprobar, para ello tenemos que trasponer las filas i_2-2 veces. De modo análogo, trasladamos la i_3-ésima fila al lugar de la tercera fila, etc., hasta que la i_k-ésima ocupe el lugar de la k-ésima fila. En total, tendremos que efectuar

$$(i_1-1)+(i_2-2)+\ldots+(i_k-k)=$$
$$=(i_1+i_2+\ldots+i_k)-(1+2+\ldots+k)$$

trasposiciones de las filas.

El menor M ya está situado en las primeras k filas del nuevo determinante. Ahora trasponemos sucesivamente las columnas del determinante: la j_1-ésima con todas las precedentes hasta que ocupe el primer lugar, después, la j_2-ésima, hasta que ocupe el segundo lugar, etc. En total, las columnas serán traspuestas

$$(j_1+j_2+\ldots+j_k)-(1+2+\ldots+k)$$

veces.

Después de todas estas transformaciones llegamos a un determinante nuevo d', en el cual, el menor M ocupa el ángulo superior de la izquierda. Como habíamos traspuesto cada vez solamente las filas y columnas contiguas, no sufrirá ninguna alteración la colocación mutua de las filas y columnas que contenían el menor M' en el determinante d. Por lo tanto, el menor M' se mantiene también como menor complementario del menor M en el determinante d', ocupando ya, sin embargo, el ángulo inferior de la derecha. Como hemos demostrado, el producto MM' es una suma de cierto número de términos del determinante d', tomados con los mismos signos que tenían en d'. No obstante, el determinante d' se ha obtenido del determinante d mediante

$$[(i_1+i_2+\ldots+i_k)-(1+2+\ldots+k)]+$$
$$+[(j_1+j_2+\ldots+j_k)-(1+2+\ldots+k)]=s_M-2(1+2+\ldots+k)$$

trasposiciones de las filas y columnas. Por ello, como sabemos por el párrafo anterior, los términos del determinante d' solamente se diferencian de los términos correspondientes del determinante d en el signo $(-1)^{s_M}$ (se comprende que el número par $2(1+2+\ldots+k)$ no influye en el signo). De aquí se deduce que el producto $(-1)^{s_M}MM'$ se compone de una cierta cantidad de términos del determinante d, tomados con los mismos signos que tenían en este determinante. De esta manera, el teorema queda demostrado.

Obsérvese que si los menores M y M' son complementarios entre sí, los números s_M y $s_{M'}$ son de una misma paridad. En efecto, el número de orden de cada fila y de cada columna está incluido como sumando en uno, y sólo en uno, de estos números. Por consiguiente, la suma $s_M + s_{M'}$ es igual a la suma de los números de orden de todas las filas y columnas del determinante, es decir, es igual a la paridad del número $2(1+2+\ldots+n)$.

§ 6. Cálculo de determinantes

Los resultados del párrafo anterior ofrecen la posibilidad de reducir el cálculo de un determinante de n-ésimo orden al cálculo de unos cuantos determinantes de $(n-1)$-ésimo orden. Introduzcamos, primero, las siguientes notaciones: si a_{ij} es un elemento del determinante d, designaremos con M_{ij} el menor complementario, o abreviando, el *menor de este elemento*, es decir, el menor de $(n-1)$-ésimo orden obtenido después de suprimir la i-ésima fila y la j-ésima columna en el determinante. Designaremos con A_{ij} el complemento algebraico del elemento a_{ij},

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Como se ha demostrado anteriormente, el producto $a_{ij}A_{ij}$ representa una suma de unos cuantos términos del determinante d, incluidos en esta suma con los mismos signos que tenían en el determinante d. Es fácil calcular el número de estos términos: es igual al número de términos en el menor M_{ij}, es decir, es igual a $(n-1)!$

Elijamos ahora una fila i-ésima cualquiera del determinante d y tomemos el producto de cada elemento de esta fila por su complemento algebraico:

$$a_{i1}A_{i1},\quad a_{i2}A_{i2},\ \ldots,\ a_{in}A_{in}. \tag{1}$$

Ningún término del determinante d puede estar incluido en dos productos diferentes (1): todos los términos del determinante incluidos en el producto $a_{i1}A_{i1}$ contienen el elemento a_{i1} de la i-ésima fila.

Por ello, se diferencian de los términos que forman parte del producto $a_{i2}A_{i2}$, que contienen el elemento a_{i2} de la i-ésima fila, etc.

Por otra parte, el número total de términos del determinante d, incluidos en todos los productos (1), es igual a

$$(n-1)!\cdot n=n!.$$

Con éstos se agotan por completo todos los términos del determinante d. Por lo tanto, hemos demostrado que se verifica el siguiente *desarrollo del determinante d por los elementos de la i-ésima fila*:

$$d_i=a_{i1}A_{i1}+a_{i2}A_{i2}+\ldots a_{in}A_{in}, \tag{2}$$

lo que significa que el *determinante d es igual a la suma de los productos de todos los elementos de una fila arbitraria de él por sus complementos algebraicos*. Se puede obtener un desarrollo análogo del determinante por los elementos de cualquiera de sus columnas.

Sustituyendo en el desarrollo (2) los complementos algebraicos por los menores correspondientes con los signos más o menos, **reduciremos el cálculo del determinante de n-ésimo orden al cálculo de unos cuantos determinantes de $(n-1)$-ésimo orden.** Obsérvese que si algunos de los elementos de la i-ésima fila son iguales a cero, no habrá que calcular, naturalmente, sus menores correspondientes. En virtud de esto, es conveniente transformar previamente el determinante, aplicando la propiedad 9 (véase el § 4), para que en una de las filas o de las columnas haya un número suficientemente grande de elementos sustituidos por ceros. En realidad, la *propiedad 9 da la posibilidad de sustituir por ceros todos los elementos, menos uno, de cualquier fila o de cualquier columna.* En efecto, si $a_{ik}\neq 0$, cualquier elemento a_{ij}, $j\neq k$, de la i-ésima fila quedará sustituido por cero después de restar de la j-ésima columna la k-ésima columna multiplicada por $\frac{a_{ij}}{a_{ik}}$. De este modo, el cálculo de un determinante de n-ésimo orden se puede reducir al cálculo de **un solo** determinante de $(n-1)$-ésimo orden.

Ejemplos.

1. Calcular el determinante de cuarto orden

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

Desarrollémoslo por los elementos de la tercera fila, aprovechando la existencia de un cero:

$$d=(-1)^{3+1}\cdot 2\cdot\begin{vmatrix} 1 & -1 & 2\\ 1 & 3 & -4\\ -5 & 3 & -3\end{vmatrix}+$$

$$+(-1)^{3+3}\cdot 1\cdot\begin{vmatrix} 3 & 1 & 2\\ -5 & 1 & -4\\ 1 & -5 & -3\end{vmatrix}+(-1)^{3+4}\cdot(-1)\cdot\begin{vmatrix} 3 & 1 & -1\\ -5 & 1 & 3\\ 1 & -5 & 3\end{vmatrix}.$$

Calculando los determinantes obtenidos de tercer orden, obtenemos:

$$d=2\cdot 16-40+48=40.$$

2. Calcular el determinante de quinto orden

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

Agregando a la *segunda fila la quinta, multiplicada por tres*, y restando de la cuarta fila la quinta, multiplicada por cuatro, obtenemos:

$$d=\begin{vmatrix} -2 & 5 & 0 & -1 & 3\\ 1 & -9 & 0 & 13 & 7\\ 3 & -1 & 0 & 5 & -5\\ 2 & 18 & 0 & -7 & -10\\ 0 & -3 & -1 & 2 & 3\end{vmatrix}.$$

Desarrollando este determinante por los elementos de la tercera columna, que contiene solamente un elemento diferente de cero (con la suma de índices $5+3$, *es decir, par*), *obtenemos:*

$$d=(-1)\cdot\begin{vmatrix} -2 & 5 & -1 & 3\\ 1 & -9 & 13 & 7\\ 3 & -1 & 5 & -5\\ 2 & 18 & -7 & -10\end{vmatrix}.$$

Transformamos de nuevo el determinante obtenido, agregando a la primera fila la segunda, multiplicada por dos, restando de la tercera fila la segunda, multiplicada por tres, y de la cuarta, la segunda multiplicada por dos:

$$d=-\begin{vmatrix} 0 & -13 & 25 & 17\\ 1 & -9 & 13 & 7\\ 0 & 26 & -34 & -26\\ 0 & 36 & -33 & -24\end{vmatrix}.$$

Después, desarrollamos éste por los elementos de la primera columna, teniendo además en cuenta, que al único elemento de esta columna, diferente de cero, le corresponde una suma impar de índices. Resulta:

$$d=\begin{vmatrix} -13 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix}.$$

Calculemos este determinante de tercer orden, desarrollándolo previamente por los elementos de su tercera fila:

$$d=36\cdot\begin{vmatrix} 25 & 17 \\ -34 & -26 \end{vmatrix}-(-33)\cdot\begin{vmatrix} -13 & 17 \\ 26 & -26 \end{vmatrix}+(-24)\cdot\begin{vmatrix} -13 & 25 \\ 26 & -34 \end{vmatrix}=$$
$$=36\cdot(-72)-(-33)\cdot(-104)+(-24)\cdot(-208)=-1032.$$

3. *Si todos los elementos de un determinante, situados a un lado de la diagonal principal, son iguales a cero, el determinante es igual al producto de los elementos situados en la diagonal principal.*

Para un determinante de segundo orden, esta afirmación es evidente. Por ello, la vamos a demostrar por el método de inducción; supongamos que está demostrada ya para los determinantes de $(n$-$1)$-ésimo orden. Consideremos el determinante de n-ésimo orden:

$$d=\begin{vmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Desarrollándolo por los elementos de la primera columna, obtenemos

$$d=a_{11}\cdot\begin{vmatrix} a_{22} & a_{23} & \dots & a_{2n} \\ 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}.$$

Al menor que figura en el segundo miembro se le puede aplicar la hipótesis de inducción, es decir, es igual a $a_{22}\, a_{33} \dots a_{nn}$; de donde

$$d=a_{11}a_{22}a_{33}\dots a_{nn}.$$

4. Se llama *determinante de Vandermonde* al siguiente:

$$d=\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ \dots & \dots & \dots & \dots & \dots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \end{vmatrix}.$$

Demostremos que para *cualquier n el determinante de Vandermonde es igual al producto de todas las diferencias posibles $a_i - a_j$*, donde $1 \leqslant j < i \leqslant n$. En efecto, para $n = 2$, se tiene

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

Supongamos que nuestra afirmación está demostrada ya para los determinantes de Vandermonde de $(n-1)$-ésimo orden. Transformemos el determinante d del modo siguiente: de la n-ésima (la última) fila restamos la $(n-1)$-ésima, multiplicada por a_1; después de la $(n-1)$-ésima restamos la $(n-2)$-ésima, multiplicada también por a_1, etc., finalmente, de la segunda fila restamos la primera, multiplicada por a_1. Obtenemos:

$$d = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \dots & a_n - a_1 \\ 0 & a_2^2 - a_1 a_2 & a_3^2 - a_1 a_3 & \dots & a_n^2 - a_1 a_n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & a_2^{n-1} - a_1 a_2^{n-2} & a_3^{n-1} - a_1 a_3^{n-2} & \dots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix}.$$

Desarrollando este determinante por los elementos de la primera columna, llegamos a un determinante de $(n-1)$-ésimo orden; después de sacar fuera del determinante todos los factores comunes de todas las columnas, éste toma la forma:

$$d = (a_2 - a_1)(a_3 - a_1) \dots (a_n - a_1) \cdot \begin{vmatrix} 1 & 1 & \dots & 1 \\ a_2 & a_3 & \dots & a_n \\ a_2^2 & a_3^2 & \dots & a_n^2 \\ \cdot & \cdot & \cdot & \cdot \\ a_2^{n-2} & a_3^{n-2} & \dots & a_n^{n-2} \end{vmatrix}.$$

El último factor es el determinante de Vandermonde de $(n-1)$-ésimo orden que, por la suposición hecha, es igual al producto de todas las diferencias $a_i - a_j$ para $2 \leqslant j < i \leqslant n$. Por consiguiente, empleando el símbolo Π para indicar el producto, se puede escribir:

$$d = (a_2 - a_1)(a_3 - a_1) \dots (a_n - a_1) \prod_{2 \leqslant j < i \leqslant n} (a_i - a_j) = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j).$$

Del mismo modo se puede demostrar que el determinante

$$d' = \begin{vmatrix} a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \dots & a_n^{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_1^2 & a_2^2 & a_3^2 & \dots & a_n^2 \\ a_1 & a_2 & a_3 & \dots & a_n \\ 1 & 1 & 1 & \dots & 1 \end{vmatrix}$$

es igual al producto de todas las diferencias posibles $a_i - a_j$, donde $1 \leqslant i < j \leqslant n$, es decir,

$$d' = \prod_{1 \leqslant i < j \leqslant n} (a_i - a_j).$$

Generalizando los desarrollos del determinante por los elementos de una fila o columna, obtenidos anteriormente, demostraremos el siguiente teorema del *desarrollo del determinante por los menores de unas cuantas filas o columnas.*

Teorema de Laplace. *Supongamos que en un determinante d de orden n se han elegido arbitrariamente k filas (o k columnas), $1 \leqslant k \leqslant n-1$. Entonces, la suma de los productos de todos los menores de k-ésimo orden, contenidos en las filas elegidas, por sus complementos algebraicos es igual al determinante d.*

Demostración. Supongamos que en el determinante d se han elegido las filas, cuyos números de orden son $i_1, i_2, \ldots, i_k$. Sabemos que el producto de cualquier menor M de k-ésimo orden, situado en estas filas, por su complemento algebraico consta de cierta cantidad de términos del determinante d, tomados con los mismos signos que tenían en el determinante. Por consiguiente, el teorema quedará demostrado, si demostramos que haciendo recorrer a M todos los menores de k-ésimo orden, situados en las filas elegidas, obtenemos todos los términos del determinante, no encontrándose ninguno de ellos dos veces.

Sea

$$a_{1\alpha_1}a_{2\alpha_2}\ldots a_{n\alpha_n} \tag{3}$$

un término arbitrario del determinante d. Tomemos aparte el producto de los elementos de este término, pertenecientes a las filas elegidas, y cuyos números de orden son $i_1, i_2, \ldots, i_k$. Esto será el producto

$$a_{i_1\alpha_{i_1}}a_{i_2\alpha_{i_2}}\ldots a_{i_k\alpha_{i_k}}; \tag{4}$$

k factores de este producto están en k columnas diferentes, precisamente en las columnas con los números de orden $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$. Por consiguiente, estos números de orden de las columnas se determinan por el término (3). Si designamos con M el menor de k-ésimo orden, situado en la intersección de las columnas que tienen estos números de orden $\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k}$, y de las filas elegidas anteriormente, con los números de orden $i_1, i_2, \ldots, i_k$, el producto (4) será uno de los términos del menor M. El producto de todos los elementos del término (3), no incluidos en (4), será un término de su menor complementario. Por lo tanto, todo término del determinante forma parte del producto de un menor determinado de k-ésimo orden situado en las filas elegidas por su menor complementario, y además es un producto de unos términos determinados de estos dos menores. Finalmente, para obtener el término tomado del determinante, con el mismo signo que tiene en el determinante, no queda más que sustituir el menor complementario por el complemento algebraico. Con esto termina la demostración del teorema.

Se podía haber demostrado el teorema de otro modo. A saber: el producto de cualquier menor M de k-ésimo orden, situado en las filas elegidas, por su complemento algebraico, consta de $k!\,(n-k)!$ términos. Esto es debido a que el menor M de k-ésimo orden se compone de $k!$ términos, y su complemento algebraico, diferenciándose posiblemente solamente en el signo del menor de orden $n-k$, contiene $(n-k)!$ términos. Por otra parte, el número de menores de k-ésimo orden, contenidos en las filas que hemos elegido, es igual al número de combinaciones de n sobre k, es decir, es igual al número

$$\frac{n!}{k!\,(n-k)!}.$$

Multiplicando, obtenemos que la suma de los productos de todos los menores de k-ésimo orden de las filas elegidas, por sus complementos algebraicos, consta de $n!$ sumandos. Sin embargo, éste es también, el número total de términos del determinante d. Por consiguiente, el teorema quedará demostrado, si demostramos que cualquier término del determinante d está incluido por lo menos una vez (y entonces, será una vez, exactamente) en la suma considerada de productos de menores por sus complementos algebraicos. Para esto no le queda más al lector que repetir (con ciertas simplificaciones) los razonamientos expuestos en la demostración precedente.

El teorema de Laplace permite reducir el cálculo de un determinante de n-ésimo orden al cálculo de unos cuantos determinantes de órdenes k y $n-k$. Resultará que habrá muchos determinantes nuevos de éstos y, por lo tanto, tiene sentido aplicar el teorema de Laplace solamente en el caso en que se puedan elegir en el determinante k filas (o columnas), de modo que muchos de los menores de k-ésimo orden situados en estas filas sean iguales a cero.

Ejemplos.

1. Sea dado un determinante, cuyos elementos situados en las primeras k filas y últimas $n-k$ columnas son iguales a cero:

$$d=\begin{vmatrix} a_{11} & \dots & a_{1k} & & & \\ \cdot & \cdot & \cdot & & 0 & \\ a_{k1} & \dots & a_{kk} & & & \\ a_{k+1,1} & \dots & a_{k+1,k} & a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nk} & a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

Este determinante es igual al producto de dos de sus menores:

$$d=\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \cdot & \cdot & \cdot \\ a_{k1} & \dots & a_{kk} \end{vmatrix}\cdot\begin{vmatrix} a_{k+1,k+1} & \dots & a_{k+1,n} \\ \cdot & \cdot & \cdot \\ a_{n,k+1} & \dots & a_{nn} \end{vmatrix}.$$

Para la demostración es suficiente desarrollar el determinante por los menores de las primeras k filas.

2. Sea dado un determinante d de orden $2n$, en cuyo ángulo superior de la izquierda figura un menor formado totalmente por ceros. Si los menores de n-ésimo orden, situados en los ángulos superior de la derecha, inferior de la izquierda e inferior de la derecha del determinante, se designan con M, M' y M'' respectivamente, es decir, que el determinante d se puede escribir simbólicamente en la forma $d = \begin{vmatrix} O & M \\ M' & M'' \end{vmatrix}$, entonces, $d = (-1)^n MM'$.

Para la demostración, desarrollamos el determinante por las primeras n filas y observamos que

$$s_M = (1+2+\ldots+n) + [(n+1)+(n+2)\ldots+2n] = n + 2n^2,$$

es decir, s_M y n tienen una misma paridad.

3. Calcular el determinante

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

Desarrollándolo por los menores de la primera y tercera columnas, que contienen ceros colocados adecuadamente, obtenemos:

$$d = (-1)^{1+3+1+3}\begin{vmatrix} -4 & 2 \\ 2 & 1 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & -5 \\ -1 & -1 & 0 \\ 4 & 2 & 5 \end{vmatrix} +$$

$$+ (-1)^{1+4+1+3}\begin{vmatrix} -4 & 2 \\ -1 & 3 \end{vmatrix}\cdot\begin{vmatrix} 3 & 1 & -5 \\ -3 & -3 & 1 \\ 4 & 2 & 5 \end{vmatrix} +$$

$$+ (-1)^{3+4+1+3}\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix}\cdot\begin{vmatrix} 1 & -2 & 1 \\ 3 & 1 & -5 \\ 4 & 2 & 5 \end{vmatrix} =$$

$$= (-8)\cdot(-20) - (-10)\cdot(-62) - 7\cdot 87 = -1069.$$

§ 7. Regla de Cramer

La teoría de los determinantes de n-ésimo orden expuesta anteriormente, permite mostrar que estos determinantes, introducidos solamente por analogía con los determinantes de segundo y tercer orden, pueden ser utilizados del mismo modo que estos últimos para la resolución de sistemas de ecuaciones lineales. Sin embargo, primero haremos una observación complementaria, ligada con los desarrollos de los determinantes por los elementos de una fila o columna; en adelante, esta observación va a ser empleada a menudo.

Desarrollemos el determinante

$$d=\begin{vmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2j} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nj} & \ldots & a_{nn} \end{vmatrix}$$

por la j-ésima columna:

$$d=a_{1j}A_{1j}+a_{2j}A_{2j}+\ldots+a_{nj}A_{nj},$$

y sustituyamos después en este desarrollo los elementos de **la** j-ésima columna por el sistema de n números arbitrarios b_1, $b_2, \ldots, b_n$. La expresión

$$b_1A_{1j}+b_2A_{2j}+\ldots+b_nA_{nj},$$

representa el desarrollo por los elementos de la j-ésima columna del determinante

$$d'=\begin{vmatrix} a_{11} & \ldots & b_1 & \ldots & a_{1n} \\ a_{21} & \ldots & b_2 & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & b_n & \ldots & a_{nn} \end{vmatrix},$$

obtenido del determinante d sustituyendo su j-ésima columna por la columna de los números $b_1, b_2, \ldots, b_n$. En efecto, la sustitución de la j-ésima columna del determinante d no afecta a los menores de los elementos de esta columna y, por lo tanto, no afecta a sus complementos algebraicos.

Apliquemos esto al caso en que en lugar de los números $b_1, b_2, \ldots, b_n$ se toman los elementos de la k-ésima columna del determinante d para $k\neq j$. El determinante que se obtiene después de esta sustitución contendrá dos columnas iguales (la j-ésima y la k-ésima) y, por eso, será igual a cero. Por consiguiente, será igual a cero también el desarrollo de este determinante por los elementos de su j-ésima columna, es decir,

$$a_{1k}A_{1j}+a_{2k}A_{2j}+\ldots+a_{nk}A_{nj}=0 \text{ para } j\neq k.$$

Por lo tanto, la *suma de los productos de todos los elementos de una columna del determinante por los complementos algebraicos de los elementos correspondientes de otra columna es igual a cero.* Naturalmente, este resultado es válido también para las *filas* del determinante.

Pasemos a estudiar los sistemas de ecuaciones lineales. Por ahora nos limitaremos al caso de sistemas en los que el **número de**

ecuaciones es igual al número de incógnitas, o sea, a los sistemas de la forma

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

Además, supondremos que el determinante d de los coeficientes de las incógnitas del sistema, denominado abreviadamente *determinante del sistema*, es diferente de cero. En estas condiciones demostraremos que el sistema (1) es compatible e incluso determinado.

En el § 2, al resolver un sistema de tres ecuaciones con tres incógnitas, multiplicábamos cada una de las ecuaciones por cierto factor y después sumábamos estas ecuaciones, resultando iguales a cero los coeficientes de dos de las tres incógnitas. Ahora vemos claramente que los factores que empleábamos eran los complementos algebraicos en el determinante del sistema, del elemento que en la ecuación dada es coeficiente de la incógnita buscada. Este mismo método se va a emplear para la resolución del sistema (1).

Supongamos primero que el sistema (1) es compatible y que $\alpha_1, \alpha_2, \ldots, \alpha_n$ es una de sus soluciones. Por consiguiente, se cumplen las igualdades

$$\left.\begin{array}{l} a_{11}\alpha_1+a_{12}\alpha_2+\ldots+a_{1n}\alpha_n=b_1, \\ a_{21}\alpha_1+a_{22}\alpha_2+\ldots+a_{2n}\alpha_n=b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}\alpha_1+a_{n2}\alpha_2+\ldots+a_{nn}\alpha_n=b_n. \end{array}\right\} \quad (2)$$

Sea j cualquiera de los números $1, 2, \ldots, n$. Multipliquemos **ambos miembros de la primera de las igualdades (2) por A_{1j}, es decir, por el complemento algebraico del elemento a_{1j} en el determinante d del sistema; ambos miembros de la segunda igualdad, por A_{2j}, etc., y finalmente, ambos miembros de la última, por A_{nj}. Sumando después por separado los primeros miembros y los segunpos miembros de todas estas igualdades, llegamos a la siguiente igualdad:**

$$\begin{aligned} (a_{11}A_{1j}+a_{21}A_{2j}+\ldots+a_{n1}A_{nj})\,\alpha_2 &+ \\ +(a_{12}A_{1j}+a_{22}A_{2j}+\ldots+a_{n2}A_{nj})\,\alpha_1 &+ \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ +(a_{1j}A_{1j}+a_{2j}A_{2j}+\ldots+a_{nj}A_{nj})\,\alpha_j &+ \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ +(a_{1n}A_{1j}+a_{2n}A_{2j}+\ldots+a_{nn}A_{nj})\,\alpha_n &= \\ =b_1A_{1j}+b_1A_{2j}+\ldots b_nA_{nj}.& \end{aligned}$$

El coeficiente de α_j en esta igualdad es igual a d, mientras que, en virtud de la observación hecha anteriormente, los coeficientes de los demás α son iguales a cero; el miembro independiente es igual al determinante que se obtiene del determinante d después de sustituir en él la j-ésima columna por la columna de los términos independientes del sistema (1). Si designamos este último determinante, igual que en el § 2, con d_j, nuestra igualdad toma la forma

$$d\alpha_j = d_j,$$

de donde

$$\alpha_j = \frac{d_j}{d},$$

puesto que $d \neq 0$. De este modo, queda demostrado que si el sistema (1) es compatible, éste posee solución única:

$$\alpha_1 = \frac{d_1}{d}, \ \alpha_2 = \frac{d_2}{d}, \ \ldots, \ \alpha_n = \frac{d_n}{d}. \tag{3}$$

Demostremos ahora que el sistema de números (3) satisface realmente al sistema de ecuaciones (1), es decir, que el sistema (1) es compatible. A continuación emplearemos las siguientes notaciones muy usuales.

Toda suma de la forma $a_1 + a_2 + \ldots + a_n$ se indicará abreviadamente mediante $\sum\limits_{i=1}^{n} a_i$. Si se considera una suma, cuyos sumandos a_{ij} *están provistos de dos subíndices, siendo* $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, se pueden tomar primero las sumas de elementos con el primer subíndice fijado, o sea, las sumas $\sum\limits_{j=1}^{m} a_{ij}$, donde $i = 1, 2, \ldots, n$, y después, sumar todas estas sumas. Entonces, para la suma de todos los elementos a_{ij}, obtenemos la expresión

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}.$$

No obstante, se podrían sumar primero los sumandos a_{ij} con el segundo subíndice fijado y sumar después las sumas obtenidas. Por lo tanto,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij},$$

o sea, en la *suma doble se puede cambiar el orden de los sumandos.*

Pongamos ahora en la i-ésima ecuación del sistema (1) los valores (3) de las incógnitas. Como el primer miembro de la i-ésima ecuación se puede escribir de la forma $\sum_{j=1}^{n} a_{ij}x_j$ y como $d_j = \sum_{k=1}^{n} b_k A_{kj}$, obtenemos:

$$\sum_{j=1}^{n} a_{ij}\cdot\frac{d_j}{d} = \frac{1}{d}\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_k A_{kj}\right) = \frac{1}{d}\sum_{k=1}^{n} b_k\left(\sum_{j=1}^{n} a_{ij}A_{kj}\right).$$

Respecto a estas transformaciones, observemos que el número $\frac{1}{d}$ es un factor común de todos los sumandos, por lo cual, se le ha sacado fuera de la suma; además, después de haber cambiado el orden de los sumandos, el factor b_k se ha sacado fuera de la suma interior, ya que no depende del subíndice j de la suma interior.

Ya sabemos que la expresión $\sum_{j=1}^{n} a_{ij}A_{kj} = a_{i1}A_{k1} + a_{i2}A_{k2} + \ldots + a_{in}A_{kn}$ es igual a d para $k = i$, e igual a 0 para los demás k. Por lo tanto, en nuestra suma exterior respecto a k quedará un sumando, precisamente $b_i d$:

$$\sum_{j=1}^{n} a_{ij}\cdot\frac{d_j}{d} = \frac{1}{d}\cdot b_i d = b_i.$$

De este modo, queda demostrado que el sistema de números (3) es, verdaderamente, solución del sistema de ecuaciones (1).

Hemos obtenido el siguiente resultado importante:

Un sistema de n ecuaciones lineales con n incógnitas, cuyo determinante es diferente de cero, tiene solución, la cual, además, es única. Esta solución se obtiene por las fórmulas (3), es decir, por la *regla de Cramer*; la formulación de esta regla es igual que en el caso de un sistema de dos ecuaciones (véase § 2).

Ejemplo. *Resolver el sistema de ecuaciones lineales*

$$\left.\begin{aligned} 2x_1 + x_2 - 5x_3 + x_4 &= 8,\\ x_1 - 3x_2 \qquad - 6x_4 &= 9,\\ 2x_2 - x_3 + 2x_4 &= -5,\\ x_1 + 4x_2 - 7x_3 + 6x_4 &= 0. \end{aligned}\right\}$$

El determinante de este sistema es diferente de cero:

$$d = \begin{vmatrix} 2 & 1 & -5 & 1\\ 1 & -3 & 0 & -6\\ 0 & 2 & -1 & 2\\ 1 & 4 & -7 & 6 \end{vmatrix} = 27,$$

por lo que se puede aplicar al sistema la regla de Cramer. Los valores de las incógnitas tendrán en los numeradores los determinantes

$$d_1=\begin{vmatrix} 8 & 1 & -5 & 1 \\ 9 & -3 & 0 & -6 \\ -5 & 2 & -1 & 2 \\ 0 & 4 & -7 & 6 \end{vmatrix}=81, \quad d_2=\begin{vmatrix} 2 & 8 & -5 & 1 \\ 1 & 9 & 0 & -6 \\ 0 & -5 & -1 & 2 \\ 1 & 0 & -7 & 6 \end{vmatrix}=-108,$$

$$d_3=\begin{vmatrix} 2 & 1 & 8 & 1 \\ 1 & -3 & 9 & -6 \\ 0 & 2 & -5 & 2 \\ 1 & 4 & 0 & 6 \end{vmatrix}=-27, \quad d_4=\begin{vmatrix} 2 & 1 & -5 & 8 \\ 1 & -3 & 0 & 9 \\ 0 & 2 & -1 & -5 \\ 1 & 4 & -7 & 0 \end{vmatrix}=27.$$

Por lo tanto

$$x_1=3,\ x_2=-4,\ x_3=-1,\ x_4=1$$

será la solución de nuestro sistema y, además, la única.

Hemos excluido el caso en que el determinante del sistema de n ecuaciones lineales con n incógnitas (1) es igual a cero. Este caso lo dejamos para el cap. 2, donde hallará su sitio en la teoría general de los sistemas de cualquier número de ecuaciones con cualquier número de incógnitas.

Referente a los sistemas de n ecuaciones lineales con n incógnitas, haremos otra observación más. Sea dado un sistema de n ecuaciones lineales *homogéneas* con n incógnitas (véase el § 1):

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0. \end{array}\right\} \tag{4}$$

En este caso, todos los determinates d_j, $j=1, 2, \ldots, n$, contienen una columna formada por ceros y, por eso, son iguales a cero. Por lo tanto, si el determinante del sistema (4) es diferente de cero, es decir, si a este sistema se le puede aplicar la regla de Cramer, su única solución será la solución nula

$$x_1=0,\ x_2=0,\ \ldots,\ x_n=0. \tag{5}$$

De aquí se desprende la siguiente conclusión:

Si un sistema de n ecuaciones lineales homogéneas con n incógnitas tiene soluciones diferentes de la nula, entonces el determinante de este sistema es necesariamente igual a cero.

En el § 12 se mostrará que, viceversa, si el determinante de un sistema de éstos es igual a cero, además de la solución nula, cuya existencia es evidente para cualquier sistema de ecuaciones homogéneas, tendrán que existir también otras soluciones.

Ejemplo. ¿Para qué valores de k, el sistema de ecuaciones

$$\left.\begin{aligned} kx_1+x_2=0, \\ x_1+kx_2=0 \end{aligned}\right\}$$

puede tener soluciones no nulas?

El determinante de este sistema

$$\begin{vmatrix} k & 1 \\ 1 & k \end{vmatrix} = k^2-1$$

será igual a cero solamente para $k=\pm 1$. *Es fácil comprobar que para cada* uno de estos dos valores de k, el sistema dado posee verdaderamente soluciones diferentes de la nula.

La importancia de la regla de Cramer consiste fundamentalmente *en que, en los casos en que es aplicable esta regla, ésta da una expresión explícita para la solución del sistema mediante los coeficientes del mismo.* Sin embargo, la aplicación práctica de la regla de Cramer va aparejada con cálculos muy complicados: en el caso de un sistema *de n ecuaciones lineales con n incógnitas, se tienen que calcular* $n+1$ determinantes de n-ésimo orden. El método de eliminación sucesiva de las incógnitas, expuesto en el § 1, es en este sentido mucho más cómodo, puesto que los cálculos que se necesitan para aplicar este método son, en esencia, equivalentes a los que se tienen que realizar al calcular *un solo* determinante de n-ésimo orden.

En algunas aplicaciones aparecen sistemas de ecuaciones lineales cuyos coeficientes y términos independientes son números reales, obtenidos al hacer mediciones de algunas cantidades físicas, es decir, que se conocen sólo aproximadamente, con cierta exactitud. A veces, los métodos expuestos anteriormente para la resolución de tales sistemas son inadecuados, debido a que proporcionan resultados poco exactos. En su lugar, se han elaborado diversos *métodos de iteración*, o sea, métodos que permiten resolver los sistemas indicados de ecuaciones mediante una aproximación sucesiva de las incógnitas. La exposición de estos métodos puede consultarla el lector en las obras sobre la teoría de las aproximaciones.

CAPITULO II

SISTEMAS DE ECUACIONES LINEALES (TEORIA GENERAL)

§ 8. Espacio vectorial de n dimensiones

Para la elaboración de la teoría general de los sistemas de ecuaciones lineales no es suficiente el aparato construido que nos sirvió satisfactoriamente para la resolución de los sistemas en que se puede aplicar la regla de Cramer. Además de los determinantes y las matrices tenemos que utilizar un nuevo concepto que, posiblemente, sea de mayor interés para la matemática en general: el concepto de *espacio vectorial de varias dimensiones.*

Hagamos primero unas cuantas observaciones previas. Por el curso de geometría analítica se sabe que todo punto en el plano se determina (dados los ejes coordenados) por sus dos coordenadas, o sea, por un sistema ordenado de dos números reales; todo vector en el plano se determina por sus dos componentes, o sea, nuevamente, por un sistema ordenado de dos números reales. De modo análogo, todo punto en el espacio de tres dimensiones se determina por sus tres coordenadas, y todo vector en el espacio se determina por sus tres componentes.

En la geometría, y también en la mecánica y en la física, se suelen estudiar frecuentemente algunos objetos, para cuya determinación no son suficientes tres números reales. Veamos, por ejemplo, el conjunto de las esferas en el espacio. Para que la esfera esté determinada por completo, es necesario que estén dadas las coordenadas de su centro y el radio, o sea, hay que señalar un sistema ordenado de cuatro números reales, de los cuales el último (el radio) sólo puede tomar, a su vez, valores positivos. Examinemos, por otra parte, las diferentes posiciones de un cuerpo sólido en el espacio. La posición del cuerpo quedará determinada por completo, si se indican las coordenadas de su centro de gravedad (o sea, tres números reales), la dirección de un eje fijo que pase por el centro de gravedad (dos números: dos, de los tres cosenos directores) y, por fin, el ángulo de rotación alrededor de este eje. Por lo tanto, la posición de un sólido en el espacio se determina por un sistema ordenado de seis números reales.

Estos ejemplos nos sugieren la oportunidad de estudiar el conjunto de todos los sistemas ordenados posibles de n números reales. Precisamente este conjunto, después de haber introducido en él las operaciones de adición y multiplicación (cosa que se hará a continuación por analogía con las operaciones correspondientes sobre los vectores del espacio tridimensional, expresadas mediante las componentes), se denomina espacio vectorial de n dimensiones. Por consiguiente, el espacio de n dimensiones es solamente una formación algebraica que conserva ciertas propiedades elementales del conjunto de los vectores del espacio de tres dimensiones, que parten del origen de coordenadas.

Un sistema ordenado de n números

$$\alpha = (a_1, a_2, \ldots, a_n) \tag{1}$$

se llama *vector de n dimensiones.* Los números a_i, $i = 1, 2, \ldots, n$, se denominarán *componentes* del vector α. Se dirá que los vectores α y

$$\beta = (b_1, b_2, \ldots, b_n) \tag{2}$$

son *iguales*, si coinciden sus componentes situadas en lugares iguales, o sea, si $a_i = b_i$ para $i = 1, 2, \ldots, n$. Para designar los vectores se emplearán en adelante las letras griegas minúsculas, mientras que las letras latinas minúsculas se utilizarán para designar los números.

Como ejemplos de vectores, señalemos los siguientes: 1) Los vectores-segmentos que parten del origen de coordenadas, en el plano o en el espacio de tres dimensiones, estando fijado el sistema de coordenadas, serán vectores de dos y tres dimensiones, respectivamente, en el sentido de la definición dada anteriormente. 2) Los coeficientes de cualquier ecuación lineal con n incógnitas forman un vector de n dimensiones. 3) Toda solución de cualquier sistema de ecuaciones lineales con n incógnitas es un vector de n dimensiones. 4) Dada una matriz de s filas y n columnas, sus filas son vectores de n dimensiones y sus columnas, vectores de s dimensiones. 5) La misma matriz de s filas y n columnas se puede considerar como un vector de sn dimensiones: es suficiente leer seguidamente los elementos de la matriz, fila por fila; en particular, toda matriz cuadrada de orden n se puede considerar como un vector de n^2 dimensiones. Es evidente, además, que cualquier vector de n^2 dimensiones se puede obtener de este modo de una matriz cuadrada de orden n.

Se llama *suma* de los vectores (1) y (2) al vector

$$\alpha + \beta = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n), \tag{3}$$

cuyas componentes son iguales a las sumas de las componentes correspondientes de los vectores que se suman. La adición de vectores está sujeta a las leyes conmutativa y asociativa, puesto que la adición de los números está sujeta a estas leyes.

El *vector nulo* desempeña el papel de cero

$$0 = (0, 0, \ldots, 0).$$

En efecto

$$\alpha + 0 = (a_1 + 0, a_2 + 0, \ldots, a_n + 0) = (a_1, a_2, \ldots, a_n) = \alpha.$$

Para designar el vector nulo emplearemos el mismo símbolo 0 que se emplea para el número cero; nunca encontraremos dificultad alguna para averiguar si en el momento dado se trata del número cero o del vector nulo; sin embargo, al estudiar los próximos párrafos, el lector tiene que recordar que el símbolo 0 se puede emplear en diversos sentidos.

El vector

$$-\alpha = (-a_1, -a_2, \ldots, -a_n). \tag{5}$$

se denominará vector *opuesto* del vector (1). Es evidente, que $\alpha + (-\alpha) = 0$. Ahora, es fácil demostrar que para la adición de vectores existe la operación inversa: la sustracción; la *diferencia* de los vectores (1) y (2) es el vector $\alpha - \beta = \alpha + (-\beta)$, o sea,

$$\alpha - \beta = (a_1 - b_1, a_2 - b_2, \ldots, a_n - b_n). \tag{6}$$

La suma de vectores de n dimensiones, definida por la fórmula (3), fue originada por la suma geométrica de vectores en el plano o en el espacio de tres dimensiones, efectuada de acuerdo a la regla del paralelogramo. En la geometría se define también el producto de un vector por un número real (por un «escalar»): multiplicar el vector α por el número k significa, siendo $k > 0$, que el vector α se alarga k veces (o que se contrae, si $k < 1$), y siendo $k < 0$, que se alarga $|k|$ veces y se cambia su dirección por la opuesta. Expresando esta regla mediante las componentes del vector y pasando al caso general considerado, obtenemos la definición siguiente:

Se llama *producto del vector* (1) *por el número* k, al vector

$$k\alpha = \alpha k = (ka_1, ka_2, \ldots, ka_n), \tag{7}$$

cuyas componentes son iguales al producto de las correspondientes componentes del vector α *por el número* k.

De esta definición se deducen las siguientes importantes propiedades, cuyas demostraciones se dejan al lector:

$$k(\alpha \pm \beta) = k\alpha \pm k\beta; \tag{8}$$

$$(k \pm l)\alpha = k\alpha \pm l\alpha; \tag{9}$$

$$k(l\alpha) = (kl)\alpha; \tag{10}$$

$$1 \cdot \alpha = \alpha. \tag{11}$$

Con la misma facilidad se comprueban, aunque pueden obtenerse también como consecuencia de las propiedades (8) — (11), las propiedades siguientes:

$$0 \cdot \alpha = 0; \tag{12}$$

$$(-1) \cdot \alpha = -\alpha; \tag{13}$$

$$k \cdot \alpha = 0; \tag{14}$$

$$\text{si } k\alpha = 0, \text{ entonces } k = 0, \text{ o bien } \alpha = 0. \tag{15}$$

El conjunto de todos los vectores de n dimensiones con componentes reales, considerado junto con las operaciones de suma de vectores y de multiplicación de un vector por un número, determinadas en el mismo, se llama *espacio vectorial de n dimensiones*.

Subrayemos que en la definición de espacio vectorial de n dimensiones no está incluida ninguna multiplicación de un vector por otro vector. Sería fácil definir el producto de vectores: se podría suponer, por ejemplo, que las componentes del producto de vectores fuesen iguales a los productos de las componentes correspondientes de los factores. Sin embargo, una tal multiplicación no tendría aplicaciones serias. Así, pues, los segmentos-vectores que parten del origen de coordenadas, en el plano o en el espacio de tres dimensiones, (se supone que se ha fijado un sistema de coordenadas), forman un espacio vectorial de dos y de tres dimensiones, respectivamente. Como se ha señalado anteriormente, en este ejemplo, la suma de vectores y el producto de un vector por un número tienen un sentido geométrico importante, mientras que al producto de vectores definido mediante la multiplicación de sus componentes no se le puede dar ninguna significación geométrica racional.

Veamos otro ejemplo más. El primer miembro de una ecuación lineal con n incógnitas, es decir, la expresión de la forma

$$f = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

se llama *forma lineal* en las incógnitas $x_1, x_2, \ldots, x_n$. Es evidente que la forma lineal f queda completamente determinada por el vector $(a_1, a_2, \ldots, a_n)$ de sus coeficientes; recíprocamente, todo vector n-dimensional determina unívocamente una forma lineal. La suma de vectores y el producto de un vector por un número se convierten en las operaciones correspondientes con las formas lineales; estas operaciones fueron empleadas eficazmente por nosotros en el § 1. La multiplicación de los vectores definida mediante el producto de sus componentes, no tiene tampoco en este ejemplo ningún sentido.

§ 9. Dependencia lineal de vectores

Se dice que el vector β, de un espacio vectorial de n dimensiones, es *proporcional* al vector α, si existe un número k tal que $\beta = k\alpha$ (véase la fórmula (7) del párrafo anterior). En particular, el vector nulo es proporcional a cualquier vector α, debido a la igualdad $0 = 0 \cdot \alpha$. Si $\beta = k\alpha$ y $\beta \neq 0$, de donde $k \neq 0$, entonces $\alpha = k^{-1}\beta$; es decir, para los vectores no nulos, la proporcionalidad posee la propiedad de simetría.

Una generalización del concepto de proporcionalidad de vectores es la noción siguiente (con la que ya nos encontramos en el § 4, para el caso de las filas de las matrices): se dice que el vector β es una *combinación lineal* de los vectores $\alpha_1, \alpha_2, \ldots, \alpha_s$, si existen unos números $l_1, l_2, \ldots, l_s$ tales que

$$\beta = l_1\alpha_1 + l_2\alpha_2 + \ldots + l_s\alpha_s.$$

Por lo tanto, la j-ésima componente del vector β, $j = 1, 2, \ldots, n$, en virtud de la definición de la suma de vectores y del producto de un vector por un número, es igual a la suma de los productos de las j-ésimas componentes de los vectores $\alpha_1, \alpha_2, \ldots, \alpha_s$ por los números $l_1, l_2, \ldots, l_s$ correspondientemente.

Se dice que el sistema de vectores

$$\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, \alpha_r \ (r \geqslant 2) \tag{1}$$

es *linealmente dependiente*, si al menos uno de estos vectores puede expresarse como combinación lineal de los demás vectores del sistema (1); en caso contrario, se dice que el sistema (1) es *linealmente independiente*.

Señalemos otra forma de esta importantísima definición: el sistema de vectores (1) es linealmente dependiente, si existen unos números $k_1, k_2, \ldots, k_r$, *entre los cuales al menos uno es diferente de cero*, de modo que se verifica la igualdad

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r = 0. \tag{2}$$

La demostración de la equivalencia de estas dos definiciones no representa dificultad alguna. Sea, por ejemplo, el vector α_r del sistema (1), combinación lineal de los demás vectores:

$$\alpha_r = l_1\alpha_1 + l_2\alpha_2 + \ldots + l_{r-1}\alpha_{r-1}.$$

De aquí se deduce la igualdad

$$l_1\alpha_1 + l_2\alpha_2 + \ldots + l_{r-1}\alpha_{r-1} - \alpha_r = 0,$$

es decir, una igualdad de la forma (2), donde $k_i = l_i$ para $i = 1, 2, \ldots, r-1$ y $k_r = -1$, es decir, $k_r \neq 0$. Recíprocamente, supongamos que los vectores (1) están ligados por la relación (2), en la que, por ejemplo, $k_r \neq 0$. Entonces,

$$\alpha_r = \left(-\frac{k_1}{k_r}\right)\alpha_1 + \left(-\frac{k_2}{k_r}\right)\alpha_2 + \ldots + \left(-\frac{k_{r-1}}{k_r}\right)\alpha_{r-1},$$

es decir, resulta que el vector α_r es combinación lineal de los vectores $\alpha_1, \alpha_2, \ldots, \alpha_{r-1}$.

Ejemplo. El sistema de vectores

$$\alpha_1 = (5, 2, 1), \quad \alpha_2 = (-1, 3, 3), \quad \alpha_3 = (9, 7, 5), \quad \alpha_4 = (3, 8, 7)$$

es linealmente dependiente, puesto que los vectores están ligados por la *relación*

$$4\alpha_1 - \alpha_2 - 3\alpha_3 + 2\alpha_4 = 0.$$

En esta relación todos los coeficientes son diferentes de cero. Por otra parte, entre nuestros vectores existen también otras dependencias lineales, en las que algunos de los coeficientes son iguales a cero, por ejemplo

$$2\alpha_1 + \alpha_2 - \alpha_3 = 0, \quad 3\alpha_2 + \alpha_3 - 2\alpha_4 = 0.$$

La segunda de las definiciones de dependencia lineal dada anteriormente, se puede aplicar cuando $r = 1$, o sea, al caso de un sistema compuesto de un solo vector α: *este sistema será linealmente dependiente cuando, y sólo cuando,* $\alpha = 0$. En efecto, si $\alpha = 0$, entonces, por ejemplo, para $k = 1$, se tiene $k\alpha = 0$. Recíprocamente, si $k\alpha = 0$ y $k \neq 0$, entonces, $\alpha = 0$.

Señalemos la siguiente propiedad del concepto de dependencia lineal.

Si un subsistema del sistema de vectores (1) es linealmente dependiente, lo es también todo el sistema (1).

En efecto, supongamos que los vectores $\alpha_1, \alpha_2, \ldots, \alpha_s$ del sistema (1), donde $s < r$, están ligados por la relación

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_s\alpha_s = 0,$$

en la que no todos los coeficientes son iguales a cero. De aquí se deduce la relación

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_s\alpha_s + 0 \cdot \alpha_{s+1} + \ldots + 0 \cdot \alpha_r = 0,$$

es decir, el sistema (1) es linealmente dependiente.

De esta propiedad se deduce la *dependencia lineal de cualquier sistema de vectores que contenga dos vectores iguales o, en general, dos vectores proporcionales, así como de cualquier sistema que contenga al vector nulo.* Obsérvese que la propiedad que acabamos de demostrar se puede formular de otra manera: *si el sistema de vectores (1) es linealmente independiente, cualquier subsistema del mismo es también linealmente independiente.*

Aquí surgen las preguntas: ¿puede contener muchos vectores un sistema linealmente independiente de vectores de n dimensiones? y en particular, ¿existen tales sistemas con un número arbitrariamente grande de vectores? Para responder a estas preguntas, consideremos en el espacio vectorial de n dimensiones los vectores

$$\left.\begin{array}{l} \varepsilon_1=(1,\ 0,\ 0,\ \ldots,\ 0), \\ \varepsilon_2=(0,\ 1,\ 0,\ \ldots,\ 0), \\ \cdots\cdots\cdots\cdots\cdots \\ \varepsilon_n=(0,\ 0,\ 0,\ \ldots,\ 1), \end{array}\right\} \tag{3}$$

denominados *vectores unitarios* de este espacio.

El sistema de vectores unitarios es linealmente independiente. Sea

$$k_1\varepsilon_1+k_2\varepsilon_2+\ldots+k_n\varepsilon_n=0;$$

como el primer miembro de esta igualdad es igual al vector $(k_1,\ k_2,\ \ldots,\ k_n)$, se tiene

$$(k_1,\ k_2,\ \ldots,\ k_n)=0,$$

o sea, $k_i=0$, $i=1,\ 2,\ \ldots,\ n$, puesto que todas las componentes del vector nulo son iguales a cero y la igualdad de vectores es equivalente a la igualdad de sus componentes correspondientes.

Por lo tanto, en el espacio vectorial de n dimensiones hemos hallado un sistema linealmente independiente, compuesto de n vectores. El lector verá más adelante que en realidad, en este espacio existen infinitos sistemas de éstos. Demostremos, por otra parte, el siguiente teorema:

Cualesquiera s vectores del espacio vectorial de n dimensiones forman, para $s>n$, un sistema linealmente dependiente.

En efecto, supongamos que se han dado los vectores

$$\begin{array}{l} \alpha_1=(a_{11},\ a_{12},\ \ldots,\ a_{1n}), \\ \alpha_2=(a_{21},\ a_{22},\ \ldots,\ a_{2n}), \\ \cdots\cdots\cdots\cdots\cdots \\ \alpha_s=(a_{s1},\ a_{s2},\ \ldots,\ a_{sn}). \end{array}$$

Tenemos que elegir unos números $k_1,\ k_2,\ \ldots,\ k_s$, no todos iguales a cero, de modo que

$$k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s=0. \tag{4}$$

Pasando de la igualdad (4) a las igualdades correspondientes entre las componentes, obtenemos

$$\left.\begin{array}{l} a_{11}k_1+a_{21}k_2+\ldots+a_{s1}k_s=0, \\ a_{12}k_1+a_{22}k_2+\ldots+a_{s2}k_s=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1n}k_1+a_{2n}k_2+\ldots+a_{sn}k_s=0. \end{array}\right\} \tag{5}$$

Las igualdades (5) forman, sin embargo, un sistema de n ecuaciones lineales homogéneas respecto a s incógnitas $k_1, k_2, \ldots, k_s$. El número de ecuaciones en este sistema es menor que el número de incógnitas y, por consiguiente, como se ha demostrado al final del § 1, este sistema tiene soluciones no nulas. Por lo tanto, se pueden elegir unos números $k_1, k_2, \ldots, k_s$, no todos iguales a cero, que satisfaga la condición (4). El teorema queda demostrado.

Un sistema linealmente independiente de vectores de n dimensiones

$$\alpha_1, \alpha_2 \ldots, \alpha_r \tag{6}$$

se llamará sistema linealmente independiente, *maximal*, si al agregarle cualquier vector β de n dimensiones, resulta un sistema linealmente dependiente. Como en cualquier dependencia lineal que liga los vectores $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta$, el coeficiente de β tiene que ser diferente de cero (puesto que, en caso contrario, el sistema (6) sería linealmente dependiente), el vector β se expresará linealmente mediante los vectores (6). Por ello, el sistema de vectores (6) es un sistema linealmente independiente *maximal*, cuando, y sólo cuando, los vectores (6) son linealmente independientes, y cualquier vector β de n dimensiones se expresa como combinación lineal de ellos.

De los resultados que hemos obtenido anteriormente se deduce que en el *espacio de n dimensiones, todo sistema, linealmente independiente, compuesto de n vectores, siempre es maximal*, y también, que *cualquier sistema de vectores linealmente independiente maximal no consta de más de n vectores.*

Todo sistema de vectores de n dimensiones, linealmente independiente, está contenido, al menos, en un sistema linealmente independiente maximal. En efecto, si el sistema dado de vectores no es maximal, se le puede agregar un vector de tal modo que el sistema obtenido se mantenga linealmente independiente. Si este sistema nuevo no es todavía maximal, se le puede agregar otro vector más, etc. Naturalmente, este proceso no se puede continuar indefinidamente, puesto que cualquier sistema de vectores de n dimensiones, compuesto de $n + 1$ vectores, es ya linealmente dependiente.

Como cualquier sistema que consta de un sólo vector no nulo es linealmente independiente, resulta que *cualquier vector no nulo está contenido en un sistema linealmente independiente maximal*. Por consiguiente, en el *espacio vectorial de n dimensiones existe una infinidad de diversos sistemas de vectores linealmente independientes maximales.*

Surge la pregunta: ¿existen en este espacio sistemas linealmente independientes maximales que contengan menos de n vectores, o el número de vectores en cualquier sistema de éstos tiene que ser, indispensablemente, igual a n? La respuesta a esta importante pregunta se dará un poco más adelante, después de hacer algunas observaciones.

Se dice frecuentemente, que el vector β se expresa *linealmente mediante el sistema de vectores*

$$\alpha_1, \alpha_2, \ldots, \alpha_r, \tag{7}$$

si β es una combinación lineal de ellos. Se comprende que, si el vector β se expresa linealmente mediante un subsistema de este sistema, entonces se expresa también linealmente mediante el sistema (7). Para demostrar esto, es suficiente tomar los otros vectores con los coeficientes iguales a cero. Generalizando esta terminología, se dice que el *sistema de vectores*

$$\beta_1, \beta_2, \ldots, \beta_s \tag{8}$$

se expresa *linealmente mediante el sistema* (7), si cada vector β_i, $i = 1, 2, \ldots, s$ es combinación lineal de los vectores del sistema (7).

Demostremos que para este concepto se cumple la ley transitiva: *si el sistema* (8) *se expresa linealmente mediante el sistema* (7), *y el sistema de vectores*

$$\gamma_1, \gamma_2, \ldots, \gamma_t \tag{9}$$

se expresa linealmente mediante el sistema (8), *entonces el sistema* (9) *también se expresa linealmente mediante el sistema* (7).

En efecto

$$\gamma_j = \sum_{i=1}^{s} l_{ji}\beta_i, \quad j = 1, 2, \ldots, t, \tag{10}$$

pero $\beta_i = \sum_{m=1}^{r} k_{im}\alpha_m$, $i = 1, 2, \ldots, s$. Sustituyendo en (10) estas expresiones, obtenemos:

$$\gamma_j = \sum_{i=1}^{s} l_{ji}\Big(\sum_{m=1}^{r} k_{im}\alpha_m\Big) = \sum_{m=1}^{r}\Big(\sum_{i=1}^{s} l_{ji}k_{im}\Big)\alpha_m,$$

o sea, cualquier vector γ_j, $j = 1, 2, \ldots, t$ es combinación lineal de los vectores del sistema (7).

Dos sistemas de vectores se llaman *equivalentes*, si cada uno de ellos se expresa linealmente mediante el otro. De la ley transitiva que acabamos de demostrar, a la que satisface la propiedad de los sistemas de vectores de expresarse linealmente entre sí, se deduce el cumplimiento de la misma ley para el concepto de equivalencia de los sistemas de vectores. De aquí también se deduce la afirmación siguiente: *siendo equivalentes dos sistemas de vectores, si un vector se expresa linealmente mediante uno de estos sistemas, entonces se expresa también linealmente mediante el otro.*

No se puede afirmar que siendo linealmente independiente uno de dos sistemas de vectores, equivalentes entre sí, lo es también el otro. Si ambos sistemas son linealmente independientes, se puede enunciar una proposición importante sobre el número de vectores que forman parte de ellos. Pero, demostremos primero el siguiente teorema, que debido al papel que va a desempeñar en adelante, lo denominaremos **teorema fundamental.**

Si en el espacio vectorial de n dimensiones se han dado dos sistemas de vectores:

$$\text{(I)} \qquad \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_r,$$

$$\text{(II)} \qquad \beta_1,\ \beta_2,\ \ldots,\ \beta_s,$$

el primero de los cuales es linealmente independiente y se expresa linealmente mediante el segundo, entonces el número de vectores del primer sistema no es superior al número de vectores del segundo sistema, es decir, $r \leqslant s$.

En efecto, supongamos que $r > s$. Por la hipótesis, cada vector del sistema (I) se expresa linealmente mediante el sistema (II):

$$\left.\begin{array}{l} \alpha_1 = \alpha_{11}\beta_1 + \alpha_{12}\beta_2 + \ldots + \alpha_{1s}\beta_s, \\ \alpha_2 = \alpha_{21}\beta_1 + \alpha_{22}\beta_2 + \ldots + \alpha_{2s}\beta_s, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_r = \alpha_{r1}\beta_1 + \alpha_{r2}\beta_2 + \ldots + \alpha_{rs}\beta_s. \end{array}\right\} \qquad (11)$$

Los coeficientes de estas expresiones lineales forman un sistema de r vectores de s dimensiones:

$$\begin{array}{l} \gamma_1 = (a_{11},\ a_{12},\ \ldots,\ a_{1s}), \\ \gamma_2 = (a_{21},\ a_{22},\ \ldots,\ a_{2s}), \\ \ldots\ldots\ldots\ldots\ldots \\ \gamma_r = (a_{r1},\ a_{r2},\ \ldots,\ a_{rs}). \end{array}$$

Como $r > s$, estos vectores son linealmente dependientes, o sea,

$$k_1\gamma_1 + k_2\gamma_2 + \ldots + k_r\gamma_r = 0,$$

donde no todos los coeficientes $k_1,\ k_2,\ \ldots,\ k_r$ son iguales a cero. De aquí, llegamos a las siguientes igualdades entre las componentes:

$$\sum_{i=1}^{r} k_i a_{ij} = 0,\ j = 1,\ 2,\ \ldots,\ s. \qquad (12)$$

Consideremos ahora la siguiente combinación lineal de los vectores del sistema (I):

$$k_1\alpha_1 + k_2\alpha_2 + \ldots + k_r\alpha_r$$

o, abreviadamente, $\sum\limits_{i=1}^{r} k_i\alpha_i$. Aplicando (11) y (12), resulta:

$$\sum_{i=1}^{r} k_i\alpha_i = \sum_{i=1}^{r} k_i \left(\sum_{j=1}^{s} a_{ij}\beta_j\right) = \sum_{j=1}^{s} \left(\sum_{i=1}^{r} k_i a_{ij}\right) \beta_j = 0;$$

lo que, sin embargo, contradice a la independencia lineal del sistema (I).

Del teorema fundamental que acabamos de demostrar se deduce el resultado siguiente:

Dos sistemas equivalentes de vectores cualesquiera, linealmente independientes, contiene el mismo número de vectores.

Es evidente que dos sistemas maximales cualesquiera de vectores de n dimensiones linealmente independientes, son equivalentes. Por consiguiente, se componen de un mismo número de vectores, y como existen sistemas de este género compuestos de n vectores, obtenemos por fin la respuesta a la pregunta que se hizo anteriormente: *todo sistema de vectores linealmente independiente maximal del espacio vectorial de n dimensiones consta de n vectores.*

De los resultados obtenidos se pueden deducir también otras consecuencias.

Si en un sistema dado de vectores, linealmente dependiente, se han tomado dos subsistemas linealmente independientes maximales, o sea, dos subsistemas a los cuales no se les puede agregar otro vector del sistema sin violar la independencia lineal, entonces estos subsistemas contienen un número igual de vectores.

En efecto, si en el sistema de vectores

$$\alpha_1, \alpha_2, \ldots, \alpha_r \tag{13}$$

el subsistema

$$\alpha_1, \alpha_2, \ldots, \alpha_s, \ s < r, \tag{14}$$

es linealmente independiente maximal, entonces cualquiera de los vectores $\alpha_{s+1}, \ldots, \alpha_r$ se expresará linealmente mediante el sistema (14). Por otra parte, cualquier vector α_i del sistema (14) se expresa linealmente mediante este sistema: es suficiente tomar el mismo vector α_i con el coeficiente 1, y todos los demás vectores del sistema con el coeficiente 0. Ahora se ve fácilmente que los sistemas (13) y (14) son equivalentes. De aquí se deduce que el sistema (13) es equivalente a cualquiera de sus subsistemas linealmente independiente maximales, por consiguiente, todos estos subsistemas son equivalentes entre sí y, siendo linealmente independientes, contienen un mismo número de vectores.

El número de vectores de cualquier subsistema linealmente independiente maximal de un sistema dado de vectores, se llama *rango* de este sistema. Empleando esta noción, deduzcamos otra consecuencia más del teorema fundamental.

Sean dados dos sistemas de vectores de n dimensiones

$$\alpha_1, \alpha_2, \ldots, \alpha_r \tag{15}$$

y

$$\beta_1, \beta_2, \ldots, \beta_s, \tag{16}$$

no necesariamente linealmente independientes, y sea k, el rango del sistema (15) y l, el rango del sistema (16). Si el primer sistema se expresa linealmente mediante el segundo, entonces $k \leqslant l$. Si estos sistemas son equivalentes, $k = l$.

En efecto, sean

$$\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_k} \tag{17}$$

y

$$\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_l} \tag{18}$$

subsistemas arbitrarios linealmente independientes maximales de los sistemas (15) y (16), respectivamente. Entonces, los sistemas (15) y (17) son equivalentes entre sí; esto mismo se refiere a los sistemas (16) y (18). Como el sistema (15) se expresa linealmente mediante el sistema (16), resulta ahora que el sistema (17) también se expresa linealmente mediante el sistema (16) y, por consiguiente, mediante el sistema (18), equivalente a él, después de lo cual no queda más que aplicar el teorema fundamental, empleando la independencia lineal del sistema (17). La segunda afirmación de la consecuencia que demostramos se deduce inmediatamente de la primera.

§ 10. Rango de una matriz

Dado un sistema de vectores de n dimensiones, surge la pregunta natural. ¿Es linealmente dependiente este sistema o no lo es? No se puede esperar que en cada caso concreto se obtenga sin dificultad la solución de este problema. Con un examen superficial sería difícil observar alguna dependencia lineal del sistema de vectores

$$\alpha = (2, -5, 1, -1), \ \beta = (1, 3, 6, 5), \ \gamma = (-1, 4, 1, 2),$$

a pesar de que, en realidad, estos vectores están ligados por la relación

$$7\alpha - 3\beta + 11\gamma = 0.$$

El § 1 proporciona un método para la resolución de este problema; como son conocidas las componentes de los vectores considerados, llamando incógnitas a los coeficientes de la dependencia lineal buscada, obtenemos un sistema de ecuaciones lineales homogéneas, que se resuelve por el método de Gauss. En el presente párrafo se indicará otro método para abordar el problema considerado; a la vez, nos aproximaremos considerablemente a nuestro objetivo principal, que consiste en resolver sistemas arbitrarios de ecuaciones lineales.

Sea dada la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{s1} & a_{s2} & \dots & a_{sn} \end{pmatrix},$$

con s filas y n columnas, donde los números s y n no están ligados de ningún modo. Las columnas de esta matriz, consideradas como vectores de s dimensiones, pueden ser, en general, linealmente dependientes. El rango del sistema de columnas, o sea, el número máximo de columnas linealmente independientes de la matriz A (con mayor precisión: el número de columnas que abarca cualquier subsistema linealmente independiente maximal del sistema de columnas), se llama *rango* de esta matriz.

Se sobreentiende, que se podrían considerar de modo semejante las filas de la matriz A como vectores de n dimensiones. Resulta que el rango del sistema de filas de la matriz es igual al rango del sistema de sus columnas, es decir, es igual al rango de esta matriz. La demostración de esta inesperada afirmación se obtendrá después de que indiquemos otra forma más de definir el rango de la matriz, lo que proporcionará a la vez un método para su cálculo.

Generalicemos primero el concepto de menor al caso de matrices rectangulares. Elijamos arbitrariamente en la matriz A, k filas y k columnas, $k \leqslant \min(s, n)$. Los elementos situados en las intersecciones de estas filas y columnas forman una matriz cuadrada de k-ésimo orden, cuyo determinante se llama *menor de k-ésimo orden* de la matriz A. A continuación, nos van a interesar los órdenes de los menores de la matriz A, que son diferentes de cero, y, precisamente, el **mayor de estos órdenes**. Para hallarlo es conveniente tener en cuenta la siguiente observación: si todos *los menores de k-ésimo orden de la matriz A son iguales a cero, entonces también son iguales a cero todos los menores de orden superior*. En efecto, desarrollando cualquier menor de orden $k+j$, $k < k+j \leqslant \min(s, n)$, por los menores de cualesquiera k filas, representamos este menor, según al teorema de Laplace, en forma de una suma de menores de orden k, multiplicados por ciertos menores de orden j, con lo que se demuestra que el menor de orden $k+j$ es igual a cero.

Demostremos ahora el siguiente **teorema sobre el rango de una matriz:**

El orden superior de los menores, diferentes de cero, de una matriz A, es igual al rango de esta matriz.

Demostración. Sea r el orden superior de los menores de la matriz A, diferentes de cero. Supongamos —lo que no restringe la generalidad de la demostración—, que el menor D, de r-ésimo orden, situado en el ángulo superior de la izquierda de la matriz A

$$A = \begin{pmatrix} \boxed{\begin{matrix} a_{11} & \ldots & a_{1r} \\ \ldots & D & \ldots \\ a_{r1} & \ldots & a_{rr} \end{matrix}} & \begin{matrix} a_{1,\,r+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{r,\,r+1} & \ldots & a_{rn} \end{matrix} \\ \begin{matrix} a_{r+1,\,1} & \ldots & a_{r+1,\,r} \\ \ldots & \ldots & \ldots \\ a_{s1} & \ldots & a_{sr} \end{matrix} & \begin{matrix} a_{r+1,\,r+1} & \ldots & a_{r+1,\,n} \\ \ldots & \ldots & \ldots \\ a_{s,\,r+1} & \ldots & a_{sn} \end{matrix} \end{pmatrix},$$

es diferente de cero, $D \neq 0$. Entonces, las primeras r columnas de la matriz A serán linealmente independientes entre sí. Si hubiese alguna dependencia lineal entre éstas, entonces, como al sumar los vectores *se suman sus componentes*, entre las columnas del menor D existiría la misma dependencia lineal y, por consiguiente, el menor D sería igual a cero.

Demostremos ahora que *cualquier l-ésima columna de la matriz A*, $r < l \leqslant n$, es combinación lineal de las primeras r columnas. Tomemos cualquier i, $1 \leqslant i \leqslant s$, y formemos el determinante auxiliar de $(r+1)$-ésimo orden

$$\Delta_i = \begin{vmatrix} a_{11} \ldots a_{1r} a_{1l} \\ \ldots\ldots\ldots \\ a_{r1} \ldots a_{rr} a_{rl} \\ a_{i1} \ldots a_{ir} a_{il} \end{vmatrix},$$

que se obtiene «orlando» el menor D con los elementos correspondientes de la l-ésima columna y de la i-ésima fila. Para cualquier i, el determinante Δ_i es igual a cero. En efecto, si $i > r$, entonces Δ_i será un menor de $(r+1)$-*ésimo orden de nuestra matriz A*, y, por lo tanto, es igual a cero, en virtud de la elección del número r. Si $i \leqslant r$, entonces Δ_i no será ya un menor de la matriz A, puesto que no puede ser obtenido de esta matriz suprimiendo algunas de sus filas y columnas; sin embargo, el determinante Δ_i contendrá ahora dos filas iguales y, *por consiguiente*, será de nuevo igual a cero.

Consideremos los complementos algebraicos de los elementos de la última fila del determinante Δ_i. Es evidente que el menor D sirve de complemento algebraico para *el elemento* a_{il}. Si $1 \leqslant j \leqslant r$, el complemento algebraico del elemento a_{ij} en Δ_i será el número

$$A_j = (-1)^{(r+1)+j} \begin{vmatrix} a_{11} \ldots a_{1,\,j-1} a_{1,\,j+1} \ldots a_{1r} a_{1l} \\ \ldots\ldots\ldots\ldots\ldots \\ a_{r1} \ldots a_{r,\,j-1} a_{r,\,j+1} \ldots a_{rr} a_{rl} \end{vmatrix};$$

éste no depende de i y por eso se ha designado con A_j. Por lo tanto, desarrollando el determinante Δ_i por los elementos de su última fila e igualando a cero este desarrollo, puesto que $\Delta_i = 0$, obtenemos:

$$a_{i1}A_1 + a_{i2}A_2 + \ldots + a_{ir}A_r + a_{il}D = 0,$$

de donde, en virtud de que $D \neq 0$,

$$a_{il} = -\frac{A_1}{D} a_{i1} - \frac{A_2}{D} a_{i2} - \ldots - \frac{A_r}{D} a_{ir}.$$

Esta igualdad se verifica para todos los i, $i = 1, 2, \ldots, s$ y como sus coeficientes no dependen de i, resulta que toda la l-ésima columna de la matriz A es una suma de sus primeras r columnas, tomadas respectivamente con los coeficientes $-\frac{A_1}{D}, -\frac{A_2}{D}, \ldots, \frac{A_r}{D}$.

Por lo tanto, en el sistema de las columnas de la matriz A hemos hallado un subsistema linealmente independiente maximal compuesto de r columnas. Con esto queda demostrado que el rango de la matriz A es igual a r, es decir, queda demostrado el teorema sobre el rango.

Este teorema proporciona un método para el cálculo práctico del rango de la matriz, y también para la solución del problema sobre la existencia de dependencia lineal en un sistema dado de vectores; formando una matriz para la que los vectores dados sirvan de columnas, y calculando el rango de esta matriz, obtenemos el número mayor de vectores de nuestro sistema, linealmente independientes.

El método para el cálculo del rango de una matriz, basado en el teorema sobre el rango, requiere el cálculo de un número de menores de esta matriz que, aunque es finito, puede ser muy grande. Sin embargo, la siguiente observación da la posibilidad de introducir en este método simplificaciones considerables. Si el lector examina otra vez más la demostración del teorema sobre el rango de la matriz, observará que al efectuarla no se aplicó la igualdad a cero de *todos* los menores de $(r+1)$-ésimo orden de la matriz A, sino que se usaron solamente los menores de $(r+1)$-ésimo orden que orlaban al menor dado D de r-ésimo orden, diferente de cero (o sea, que lo contienen totalmente dentro de sí). Por lo tanto, de la igualdad a cero solamente de estos menores, se deduce que r es el máximo número de columnas linealmente independientes de la matriz A. Esto último trae consigo la igualdad a cero de todos los menores de $(r+1)$-ésimo orden de esta matriz. Llegamos a la siguiente **regla para el cálculo del rango de una matriz:**

Al calcular el rango de una matriz se debe pasar de los menores de menor orden a los de orden mayor. Habiendo hallado un menor D de k-ésimo orden diferente de cero, se deben calcular solamente *los menores de $(k+1)$-ésimo orden que orlan al menor D: si todos éstos son iguales a cero, el rango de esta matriz es igual a k.*

Ejemplos.

1. Hallar el rango de la matriz

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

El menor de segundo orden, situado en el ángulo superior de la izquierda de esta matriz, es igual a cero. Sin embargo, en esta matriz hay también menores de segundo orden, diferentes de cero, por ejemplo,

$$d=\begin{vmatrix} -4 & 3 \\ -2 & 1 \end{vmatrix}\neq 0.$$

El menor de tercer orden

$$d'=\begin{vmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{vmatrix},$$

es un orlado del menor d, diferente de cero, $d'=1$, no obstante, los orlados de cuarto orden del menor d' son iguales a cero:

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

Por lo tanto, el rango de la matriz A es igual a tres.

2. Hallar un subsistema, linealmente independiente, maximal en el sistema de vectores

$$\alpha_1=(2,\ -2,\ -4),\ \alpha_2=(1,\ 9,\ 3),\ \alpha_3=(-2,\ -4,\ 1),\ \alpha_4=(3,\ 7,\ -1).$$

Formamos la matriz

$$\begin{pmatrix} 2 & 1 & -2 & 3 \\ -2 & 9 & -4 & 7 \\ -4 & 3 & 1 & -1 \end{pmatrix},$$

en la que los vectores dados sirven de columnas. El rango de esta matriz es igual a dos: el menor de segundo orden situado en el ánguo superior de la izquierda es diferente de cero, pero los dos menores orlados de él, de tercer orden, son iguales a cero. De aquí se deduce que los vectores α_1, α_2 forman en el sistema dado uno de los subsistemas linealmente independientes maximales.

Como consecuencia del teorema sobre el rango de una matriz, demostremos la afirmación ya enunciada anteriormente:

El máximo número de filas linealmente independientes de cualquier matriz es igual al máximo número de sus columnas linealmente independientes, es decir, es igual al rango de la matriz.

Para la demostración, trasponemos la matriz, o sea, sus filas las hacemos columnas, conservando su numeración. En la transposición, el máximo orden de los menores de la matriz diferentes de cero no puede alterarse, puesto que la trasponsición no altera al determinante y para cualquier menor de la matriz inicial, el menor obtenido de él por transposición está contenido en la nueva matriz y viceversa. De aquí se deduce que el rango de la nueva matriz es igual al rango de la matriz inicial; éste a su vez es igual al máximo número de columnas linealmente independientes de la nueva matriz, es decir, igual al máximo número de filas linealmente independientes de la matriz inicial.

Ejemplo. En el § 8 se introdujo el concepto de la forma lineal en n incógnitas y se dio la definición de suma de formas lineales y de su producto por un número. Esta definición permite generalizar el concepto de dependencia lineal, con todas sus propiedades, para el caso de formas lineales.

Sea dado el sistema de formas lineales

$$\begin{aligned} f_1 &= x_1+2x_2+x_3+3x_4, \\ f_2 &= 4x_1-x_2+5x_3-6x_4, \\ f_3 &= x_1-3x_2-4x_3-7x_4, \\ f_4 &= 2x_1-x_2-x_3. \end{aligned}$$

Se necesita elegir en él un subsistema linealmente independiente maximal.

Formemos la matriz de los coeficientes de estas formas:

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

y hallemos su rango. El menor de segundo orden, situado en el ángulo superior de la izquierda, es diferente de cero. Pero, como fácilmente se comprueba, sus cuatro determinantes orlados de tercer orden son iguales a cero. De aquí se deduce que las primeras dos filas de nuestra matriz son linealmente independientes, mientras que la tercera y la cuarta son combinaciones lineales de ellas. Por consiguiente, el sistema f_1, f_2 es el subsistema buscado del sistema dado de formas lineales.

Señalemos otra consecuencia importante del teorema sobre el rango de una matriz.

Un determinante de n-ésimo orden es igual a cero cuando, y sólo cuando, entre sus filas existe una dependencia lineal.

En una dirección, esta afirmación ya está demostrada en el § 4 (propiedad 8). Supongamos ahora que se ha dado un determinante de n-ésimo orden igual a cero o, en otras palabras, una matriz cuadrada de n-ésimo orden, cuyo único menor de máximo orden es igual a cero. De aquí se deduce que el máximo orden de los menores de la matriz que son diferentes de cero es menor que n, o sea, que el rango es menor que n, y, por lo demostrado anteriormente, las filas de esta matriz son linealmente dependientes.

Se sobreentiende que en el enunciado de la consecuencia que hemos demostrado se puede hablar de las columnas del determinante, en lugar de las filas.

Existe también otro método para calcular el rango de una matriz que no está ligado con el teorema sobre el rango y que no requiere el cálculo de determinantes. Pero se puede aplicar solamente cuando se quiera determinar el mismo rango y no interese saber qué columnas (o filas) son las que precisamente forman un sistema linealmente independiente maximal. Veamos este método.

Se llaman *transformaciones elementales* de una matriz A a las siguientes:

(a) la permutación (trasposición) de dos filas o de dos columnas;

(b) la multiplicación de una fila (o de una columna) por un número arbitrario diferente de cero;

(c) la suma a una fila (o a una columna) de otra fila (columna) multiplicada por un número.

Fácilmente se observa que las *transformaciones elementales no alteran el rango de la matriz*. En efecto, si, por ejemplo, se aplican estas transformaciones a las columnas de la matriz, entonces el sistema de columnas, consideradas como vectores, se sustituye por otro equivalente. Demostremos esto solamente para la transformación (c), puesto que para las (a) y (b), es evidente. Supongamos que a la i-ésima columna se agrega la j-ésima columna, multiplicada por el número k. Si antes de la transformación, los vectores

$$\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_n, \tag{1}$$

servían de columnas de la matriz, después de la transformación servirán de columnas los vectores

$$\alpha_1, \ldots, \alpha_i' = \alpha_i + k\alpha_j, \ldots, \alpha_j, \ldots, \alpha_n. \tag{2}$$

El sistema (2) se expresa linealmente mediante el sistema (1). La igualdad

$$\alpha_i = \alpha_i' - k\alpha_j$$

muestra a su vez, que el sistema (1) se expresa linealmente mediante el (2). Por consiguiente, estos sistemas son equivalentes, y sus subsistemas, linealmente independientes maximales están compuestos de un mismo número de vectores.

Por lo tanto, para calcular el rango de una matriz, se puede simplificar previamente mediante una combinación de transformaciones elementales.

Se dice que una matriz que consta de s filas y n columnas es de *forma diagonal*, si todos sus elementos son iguales a cero, a excepción de los elementos $a_{11}, a_{22}, \ldots, a_{rr}$ (donde $0 \leqslant r \leqslant \min(s, n)$), que son iguales a la unidad. Es evidente que el rango de esta matriz es igual a r.

Toda matriz se puede reducir a la forma diagonal mediante transformaciones elementales.

En efecto, sea dada la matriz

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{s1} & \ldots & a_{sn} \end{pmatrix}.$$

Si todos sus elementos son iguales a cero, esta tiene la forma diagonal. Si en ella hay elementos diferentes de cero, trasponiendo filas y columnas se puede conseguir que el elemento a_{11} sea diferente de cero. Multiplicando después la primera fila por a_{11}^{-1}, convertimos el elemento a_{11} en la unidad. Restando ahora de la j-ésima columna, $j > 1$, la primera columna, multiplicada por a_{1j}, se sustituye por cero el elemento a_{1j}. Efectuando esta transformación con todas las columnas, comenzando con la segunda, y también con todas las filas, llegaremos a la matriz de la forma:

$$A' = \begin{pmatrix} 1 & 0 & & 0 \\ 0 & a'_{22} & \ldots & a'_{2n} \\ . & . & . & . \\ 0 & a'_{s2} & \ldots & a'_{sn} \end{pmatrix}.$$

Luego efectuamos estas mismas transformaciones con la matriz que queda en el ángulo inferior de la derecha, etc., etc., Después de reiterar este proceso una cantidad finita de veces, llegaremos a la matriz diagonal que tiene el mismo rango que la matriz inicial A.

Por lo tanto, *para hallar el rango de una matriz hay que reducir esta matriz mediante transformaciones elementales a la forma diagonal y calcular el número de unidades que hay en su diagonal principal.*

Ejemplo. Hallar el rango de la matriz

$$A = \begin{pmatrix} 0 & 2 & -4 \\ -1 & -4 & 5 \\ 3 & 1 & 7 \\ 0 & 5 & -10 \\ 2 & 3 & 0 \end{pmatrix}.$$

Transponiendo en esta matriz la primera y segunda columna, y multiplicando la primera fila por el número $\frac{1}{2}$, llegamos a la matriz

$$\begin{pmatrix} 1 & 0 & -2 \\ -4 & -1 & 5 \\ 1 & 3 & 7 \\ 5 & 0 & -10 \\ 3 & 2 & 0 \end{pmatrix}.$$

Agregando a su tercera columna la primera duplicada y agregando después a cada una de las demás filas un múltiplo de la nueva primera fila, obtenemos la matriz

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -3 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \\ 0 & 2 & 6 \end{pmatrix}.$$

Finalmente, multiplicando la segunda fila por -1, restando de la tercera columna la segunda, multiplicada por tres, y restando de la tercera y quinta filas unos múltiplos de la segunda fila nueva, llegaremos a la forma diagonal buscada

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Por lo tanto, el rango de la matriz A es igual a dos.

En el cap. 13 nos encontraremos otra vez con las transformaciones elementales y con la forma diagonal de la matriz; pero serán matrices cuyos elementos no serán números sino polinomios.

§ 11. Sistemas de ecuaciones lineales

Estudiaremos ahora sistemas arbitrarios de ecuaciones lineales, sin suponer que el número de ecuaciones sea igual al número de incógnitas. Nuestros resultados se podrán aplicar también al caso (que quedó sin examinar en el § 7) en el que el número de ecuaciones sea igual al número de incógnitas, siendo el determinante igual a cero.

Sea dado un sistema de ecuaciones lineales

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n = b_s. \end{array}\right\} \qquad (1)$$

Como sabemos por el § 1, ante todo se debe resolver el problema sobre la compatibilidad de este sistema. Con este fin, tomemos la matriz A de los coeficientes del sistema y la matriz «ampliada» $\overline{A}$, obtenida al agregar a la matriz A la columna de los términos independientes,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{s1} & a_{s2} & \ldots & a_{sn} \end{pmatrix}, \quad \overline{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{s1} & a_{s2} & \ldots & a_{sn} & b_s \end{pmatrix},$$

y calculemos los rangos de estas matrices. Es fácil ver que *el rango de la matriz* $\overline{A}$, *o es igual al rango de la matriz* A, *o es mayor en una unidad.* En efecto, tomemos un sistema máximal de columnas de la matriz A, linealmente independiente. Este también será linealmente independiente en la matriz $\overline{A}$. Si conserva también la propiedad de ser maximal, o sea, que la columna de los términos independientes se expresa linealmente mediante el mismo, entonces los rangos de las matrices A y $\overline{A}$ son iguales; en caso contrario, agregando a este sistema la columna de los términos independientes, obtenemos el sistema linealmente independiente de columnas de la matriz $\overline{A}$, que en ésta será maximal.

El problema de la compatibilidad de un sistema de ecuaciones lineales se resuelve difinitivamente con el siguiente teorema.

Teorema de Kronecker-Capelli. *El sistema de ecuaciones* (1) *es compatible cuando, y sólo cuando, el rango de la matriz ampliada $\overline{A}$ es igual al rango de la matriz A.*

Demostración. 1. Supongamos que el sistema (1) es compatible y que $k_1, k_2, \ldots, k_n$ es una de sus soluciones. Sustituyendo estos números en lugar de las incógnitas del sistema (1), obtenemos s identidades, que muestran que la última columna de la matriz $\overline{A}$ es una suma de todas las demás columnas, tomadas con los coeficientes $k_1, k_2, \ldots, k_n$, respectivamente. Cualquiera otra columna de la matriz $\overline{A}$ forma parte también de la matriz A y, por eso, se expresa linealmente mediante todas las columnas de esta matriz. Recíprocamente, toda columna de la matriz A es también columna de la matriz $\overline{A}$, o sea, se expresa linealmente mediante las columnas de esta matriz. De aquí se deduce que los sistemas de columnas de las matrices A y $\overline{A}$ son equivalentes entre sí. Por consiguiente, como se demostró al final del § 9, estos dos sistemas de vectores de s dimensiones tienen un mismo rango; en otras palabras, los rangos de las matrices A y $\overline{A}$ son iguales entre sí.

2. Supongamos ahora que las matrices A y $\overline{A}$ tienen un mismo rango. De esto se deduce, que cualquier sistema linealmente independiente maximal de columnas de la matriz A se mantiene también en la matriz $\overline{A}$ como sistema linealmente independiente maximal. Por lo tanto, la última columna de la matriz $\overline{A}$ se expresa linealmente mediante este sistema y, por consiguiente, mediante el sistema de columnas de la matriz A. Así que existe un sistema de coeficientes $k_1, k_2, \ldots, k_n$ tal que la suma de las columnas de la matriz A, tomadas con estos coeficientes, es igual a la columna de los términos independientes. De aquí que los números $k_1, k_2, \ldots, k_n$ formen una solución del sistema (1). Por ello, la coincidencia de los rangos de las matrices A y $\overline{A}$ trae consigo la compatibilidad del sistema (1).

El teorema queda demostrado.

Al aplicar este teorema en los ejercicios prácticos, es necesario calcular primero el rango de la matriz A. Para esto hay que hallar uno de los menores de la matriz que sea diferente de cero y cuyos orlados sean iguales a cero; sea éste el menor M. Después se deben calcular todos los menores de la matriz $\overline{A}$ que son orlados de M, pero que no estén contenidos en A (los llamados *determinantes característicos* del sistema (1)). Si todos éstos son iguales a cero, el rango de la matriz $\overline{A}$ es igual al rango de la matriz A y, por consiguiente, el sistema (1) es compatible; en caso contrario, es incompatible. De aquí que el teorema de Kronecker-Capelli se pueda enunciar del modo siguiente: el *sistema de ecuaciones lineales (1) es compatible cuando, y sólo cuando, todos sus determinantes característicos son iguales a cero.*

Supongamos ahora que el **sistema (1) es compatible.** El teorema de Kronecker—Capelli, mediante el que establecemos la compatibilidad de este sistema, afirma la existencia de una solución; mas éste no proporciona ningún método para la **averiguación práctica de todas las soluciones** del sistema. Pasemos ahora a resolver este problema.

Supongamos que la matriz A es de rango r. Como se ha demostrado en el párrafo anterior, r es el máximo número de filas linealmente independientes de la matriz A. Para precisar, supongamos que las primeras r filas de la matriz A son linealmente independientes, y que cada una de las demás es combinación lineal de ellas. Entonces, las primeras r filas de la matriz $\overline{A}$ serán también linealmente independientes: toda dependencia lineal entre ellas sería también una dependencia lineal entre las primeras r filas de la matriz A (¡véase la definición de la suma de vectores!). De la coincidencia de los rangos de las matrices A y $\overline{A}$ se deduce que las primeras r filas de la matriz $\overline{A}$ forman en ésta un sistema maximal de filas linealmente independiente, o sea, que cualquier otra fila de esta matriz es combinación lineal de ellas.

De aquí se deduce que cualquier ecuación del sistema (1), se puede representar como una suma de *las primeras r ecuaciones, tomadas* con ciertos coeficientes. En consecuencia, cualquier solución simultánea de las primeras r ecuaciones satisface también a todas las ecuaciones del sistema (1). Por consiguiente, es suficiente hallar todas las soluciones del sistema

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n = b_r. \end{array}\right\} \qquad (2)$$

Como las filas formadas por los coeficientes de las incógnitas en las ecuaciones (2) son linealmente independientes, o sea, la matriz de los coeficientes es de rango r, se tiene que $r \leqslant n$. Además, al menos uno de los menores de r-ésimo orden de esta matriz es diferente de cero. Si $r = n$, entonces (2) será un sistema con igual número de ecuaciones que de incógnitas y con un determinante diferente de cero, por lo que éste y también el sistema (1), tendrán solución única, que es precisamente la que se calcula por la regla de Cramer.

Supongamos ahora que $r < n$, y para precisar, supongamos que es diferente de cero el menor de r-ésimo orden, formado por los coeficientes de las primeras r incógnitas. Traslademos al segundo miembro, en cada una de las ecuaciones (2), todos los términos que contienen las incógnitas $x_{r+1}, \ldots, x_n$ y elijamos para estas incógnitas algunos valores $c_{r+1}, \ldots, c_n$. Obtenemos un sistema de r ecuaciones

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1 - a_{1,\,r+1}c_{r+1} - \ldots - a_{1n}c_n, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r = b_2 - a_{2,\,r+1}c_{r+1} - \ldots - a_{2n}c_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = b_r - a_{r,\,r+1}c_{r+1} - \ldots - a_{rn}c_n \end{array}\right\} \quad (3)$$

con respecto a r incógnitas, $x_1, x_2, \ldots, x_r$. A este sistema se le puede aplicar la regla de Cramer, poseyendo por lo tanto, una solución única, $c_1, c_2, \ldots, c_r$; es evidente que el sistema de números $c_1, c_2, \ldots, c_r, c_{r+1}, \ldots, c_n$ representa una solución del sistema (2). Como los valores $c_{r+1}, \ldots, c_n$, para las incógnitas $x_{r+1}, \ldots, x_n$, llamadas *incógnitas independientes*, podían ser elegidos arbitrariamente, se pueden obtener de este modo infinitas soluciones distintas del sistema (2).

Por otra parte, toda solución del sistema (2) se puede obtener por el método indicado: si se ha obtenido alguna solución $c_1, c_2, \ldots, c_n$ del sistema (2), como valores para las incógnitas independientes tomamos los números $c_{r+1}, \ldots, c_n$. Entonces, los números $c_1, c_2, \ldots, c_r$ serán solución del sistema (3) y, por consiguiente, formarán la única solución de este sistema, que se calcula por la regla de Cramer.

Todo lo expuesto anteriormente se resume en la siguiente **regla para la solución de un sistema arbitrario de ecuaciones lineales:**

Sea dado un sistema compatible de ecuaciones lineales (1) *y sea r el rango de la matriz A de los coeficientes del sistema. Elijamos en A, r filas linealmente independientes y dejemos en el sistema* (1) *solamente aquellas ecuaciones, cuyos coeficientes forman parte de las filas elegidas. Dejemos en los primeros miembros de estas ecuaciones r incógnitas, de modo que el determinante formado por los coeficientes de ellas sea diferente de cero, mientras que las otras incógnitas las consideramos independientes, trasladándolas a los segundos miembros de las ecuaciones. Dando valores numéricos arbitrarios a las incógnitas* *independientes y calculando los valores de las demás incógnitas por la regla de Cramer, obtenemos todas las soluciones del sistema* (1).

He aquí de nuevo el enunciado del resultado obtenido:

Un sistema compatible (1) *tiene solución única cuando, y sólo cuando, el rango de la matriz A es igual al número de las incógnitas.*

Ejemplos. 1. Resolver el sistema

$$\left.\begin{aligned} 5x_1 - x_2 + 2x_3 + x_4 &= 7, \\ 2x_1 + x_2 + 4x_3 - 2x_4 &= 1, \\ x_1 - 3x_2 - 6x_3 + 5x_4 &= 0. \end{aligned}\right\}$$

El rango de la matriz de los coeficientes es igual a dos: el menor de segundo orden, situado en el ángulo superior de la izquierda de esta matriz, es diferente de cero, pero ambos menores orlados de tercer orden son iguales a cero. El rango de la matriz ampliada es igual a tres, puesto que

$$\begin{vmatrix} 5 & -1 & 7 \\ 2 & 1 & 1 \\ 1 & -3 & 0 \end{vmatrix} = -35 \neq 0.$$

De aquí se deduce que el sistema es incompatible.

2. Resolver el sistema

$$\left.\begin{aligned} 7x_1 + 3x_2 &= 2, \\ x_1 - 2x_2 &= -3, \\ 4x_1 + 9x_2 &= 11. \end{aligned}\right\}$$

El rango de la matriz de los coeficientes es igual a dos, o sea, es igual al número de incógnitas; el rango de la matriz ampliada también es igual a dos. Por lo tanto, el sistema es compatible y tiene solución única. Los primeros miembros de las primeras dos ecuaciones son linealmente independientes; resolviendo el sistema de estas dos ecuaciones, obtenemos los siguientes valores para las incógnitas:

$$x_1 = -\frac{5}{17}, \quad x_2 = \frac{23}{17}.$$

Vemos fácilmente que esta solución satisface también a la tercera ecuación.

3. Resolver el sistema

$$\left.\begin{aligned} x_1 + x_2 - 2x_3 - x_4 + x_5 &= 1, \\ 3x_1 - x_2 + x_3 + 4x_4 + 3x_5 &= 4, \\ x_1 + 5x_2 - 9x_3 - 8x_4 + x_5 &= 0. \end{aligned}\right\}$$

El sistema es compatible, puesto que el rango de la matriz ampliada al igual que el de la matriz de los coeficientes es igual a dos. Los primeros miembros de la primera y tercera ecuaciones son linealmente independientes, puesto que los coeficientes de las incógnitas x_1 y x_2 forman un menor de segundo orden diferente de cero. El sistema de estas dos ecuaciones lo resolvemos suponiendo que las incógnitas x_3, x_4, x_5 son independientes; para ello, trasladamos éstas a los segundos miembros de las ecuaciones y suponemos que ya se les han atribuido valores numéricos. Aplicando la regla de Cramer, obtenemos:

$$x_1 = \frac{5}{4} + \frac{1}{4}x_3 - \frac{3}{4}x_4 - x_5,$$

$$x_2 = -\frac{1}{4} + \frac{7}{4}x_3 + \frac{7}{4}x_4.$$

Estas igualdades determinan la *solución general* del sistema dado: asignando a las incógnitas independientes valores numéricos arbitrarios, obtenemos todas las soluciones de nuestro sistema. Así pues, son soluciones de nuestro sistema, por ejemplo, los vectores $(2, 5, 3, 0, 0)$, $(3, 5, 2, 1, -2)$, $(0, -\frac{1}{4}, -1, 1, \frac{1}{4})$, etc. Por otra parte, sustituyendo las expresiones para x_1 y x_2 de la solución general en cualquiera de las ecuaciones del sistema, por ejemplo, en la segunda, que fue anteriormente excluida, obtenemos una identidad.

4. Resolver el sistema

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

A pesar de que el número de ecuaciones es igual al número de incógnitas, no se puede aplicar la regla de Cramer, pues el determinante del sistema es igual a cero. El rango de la matriz de los coeficientes es igual a tres: en el ángulo superior de la derecha de esta matriz está situado un menor de tercer orden, diferente de cero. El rango de la matriz ampliada también es igual a tres, es decir, el sistema es compatible. Examinando solamente las primeras tres ecuaciones y tomando la incógnita x_1 como independiente, obtenemos la solución general en la forma:

$$x_2 = -\frac{1}{5} - \frac{2}{5}x_1, \quad x_3 = -\frac{8}{5} + \frac{9}{5}x_1, \quad x_4 = 0.$$

5. Sea dado un sistema compuesto de $n+1$ ecuaciones respecto a n incógnitas. La matriz ampliada $\overline{A}$ de este sistema es cuadrada, de orden $n+1$. Si nuestro sistema es compatible, entonces, según el teorema de Kronecker-Capelli, el determinante de la matriz $\overline{A}$ tiene que ser igual a cero.

Así, pues, sea dado el sistema

$$\left.\begin{aligned} x_1-8x_2 &= 3,\\ 2x_1+x_2 &= 1,\\ 4x_1+7x_2 &= -4. \end{aligned}\right\}$$

El determinante de los coeficientes y de los términos independientes de estas ecuaciones es diferente de cero:

$$\begin{vmatrix} 1 & -8 & 3\\ 2 & 1 & 1\\ 4 & 7 & -4 \end{vmatrix} = -77,$$

por lo tanto, el sistema es incompatible.

En general, la afirmación recíproca no es justa: de la igualdad a cero del determinante de la matriz $\overline{A}$ no se deduce la coincidencia de los rangos de las matrices A y $\overline{A}$.

§ 12. Sistemas de ecuaciones lineales homogéneas

Apliquemos los resultados del párrafo anterior al caso de un *sistema de ecuaciones lineales homogéneas*:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n &= 0. \end{aligned}\right\} \quad (1)$$

Del teorema de Kronecker-Capelli se deduce que este sistema siempre es compatible, puesto que, agregando una columna de ceros, no se puede elevar el rango de la matriz. Por cierto, esto se observa inmediatamente, ya que el sistema (1) siempre posee la *solución nula* $(0, 0, \ldots, 0)$.

Supongamos que la matriz A de los coeficientes del sistema (1) es de rango r. *Si $r=n$, la solución nula es la única solución del sistema* (1); *si $r<n$, el sistema posee también soluciones diferentes de la nula*; para hallar todas estas soluciones se emplea el mismo método que anteriormente se usó en el caso de un sistema arbitrario de ecuaciones. En particular, un *sistema de n ecuaciones lineales homogéneas con n incógnitas tiene soluciones diferentes de la nula cuando, y sólo cuando, el determinante de este sistema es igual a cero**. En efecto, la igualdad a cero de este determinante es equivalente a la afirmación de que el rango de la matriz A es menor que n. Por otra parte, *si en el sistema de ecuaciones lineales homogéneas el número de ecuaciones es menor que el número de incógnitas, el sistema posee indispensablemente soluciones diferentes de la nula*, puesto que en este caso el rango no puede ser igual al número de las incógnitas; este resultado fue obtenido en el § 1 por medio de otros razonamientos.

Veamos en particular el caso de un sistema, compuesto de $n-1$ ecuaciones homogéneas con respecto a n incógnitas, en el que se supone que los primeros miembros de las ecuaciones son linealmente independientes entre sí. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1,1} & a_{n-1,2} & \ldots & a_{n-1,n} \end{pmatrix}$$

la matriz de los coeficientes de este sistema; designemos con M_i el menor de $(n-1)$-ésimo orden, obtenido después de suprimir en la matriz A la i-ésima columna, $i=1, 2, \ldots, n$. *Entonces, una de las soluciones de nuestro sistema será el sistema de números*

$$M_1, \ -M_2, \ M_3, \ -M_4, \ \ldots, \ (-1)^{n-1} M_n, \qquad (2)$$

* Una mitad de esta afirmación se demostró ya en el § 7.

y cualquier otra solución será proporcional a ésta.

Demostración. Como, por hipótesis, el rango de la matriz A es igual a $n-1$, uno de los menores M_i tiene que ser diferente de cero; supongamos que sea M_n. Tomemos la incógnita x_n como independiente y la traslademos al segundo miembro en cada una de las ecuaciones, obteniendo:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1,\,n-1}x_{n-1} = -a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2,\,n-1}x_{n-1} = -a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n-1,\,1}x_1 + a_{n-1,\,2}x_2 + \ldots + a_{n-1,\,n-1}x_{n-1} = -a_{n-1,\,n}x_n. \end{array}$$

Aplicando ahora la regla de Cramer, obtenemos la solución general del sistema dado de ecuaciones, la cual, después de sencillas transformaciones, puede ser expresada de la forma siguiente:

$$x_i = (-1)^{n-1}\frac{M_i}{M_n}x_n, \quad i = 1, 2, \ldots, n-1. \tag{3}$$

Haciendo $x_n = (-1)^{n-1} M_n$, obtenemos: $x_i = (-1)^{2n-i-1} M_i$, $i = 1, 2, \ldots, n-1$, o bien, como la diferencia $(2n-i-1)-(i-1) = 2n-2i$ es un número par, $x_i = (-1)^{i-1} M_i$, es decir, el sistema de números (2) es verdaderamente una solución de nuestro sistema de ecuaciones. Cualquier otra solución de este sistema se obtiene de las fórmulas (3) con otro valor númerico de la incógnita x_n, por lo que será proporcional a la solución (2). Se comprende que la afirmación considerada es justa también cuando $M_n = 0$, siendo diferente de cero uno de los menores M_i, $1 \leqslant i \leqslant n-1$.

Las soluciones de un sistema de ecuaciones lineales homogéneas poseen las siguientes propiedades. Si el vector $\beta = (b_1, b_2, \ldots, b_n)$ es una solución del sistema (1), entonces, para cualquier número k, el vector $k\beta = (kb_1, kb_2, \ldots, kb_n)$ también será solución de este sistema. Esto se comprueba inmediatamente por sustitución en cualquiera de las ecuaciones (1). Si, además, el vector $\gamma = (c_1, c_2, \ldots, c_n)$ es otra solución del sistema (1), el vector $\beta + \gamma = (b_1 + c_1, b_2 + c_2, \ldots, b_n + c_n)$ también será solución de este sistema:

$$\sum_{j=1}^{n} a_{ij}(b_j + c_j) = \sum_{j=1}^{n} a_{ij}b_j + \sum_{j=1}^{n} a_{ij}c_j = 0, \quad i = 1, 2, \ldots, s.$$

Por eso, en general, *cualquier combinación lineal de soluciones del sistema homogéneo* (1) *será también solución de este sistema.* Obsérvese que en el caso de un sistema *no homogéneo*, o sea, de un sistema de ecuaciones lineales, cuyos términos independientes no son todos iguales a cero, la afirmación correspondiente no se cumple: ni la suma de dos soluciones de un sistema de ecuaciones no homogéneas, ni el producto de una solución de este sistema por un número, será ya solución de este sistema.

Por el § 9 se sabe que cualquier sistema de vectores de n dimensiones, compuesto de más de n vectores, es linealmente dependiente. De aquí se deduce que entre las soluciones del sistema homogéneo (1) (como es sabido, éstas son vectores de n dimensiones) se puede elegir un sistema *finito* linealmente independiente maximal de modo que sea maximal en el sentido de que cualquier otra solución del sistema (1) sea combinación lineal de las soluciones que forman parte de este sistema elegido. Todo sistema maximal de soluciones linealmente independientes del sistema de ecuaciones homogéneas (1), se llama *sistema fundamental de soluciones*.

Subrayemos otra vez más que *un vector de n dimensiones es solución del sistema* (1) *si, y sólo si, éste es combinación lineal de los vectores que forman el sistema fundamental dado*.

Se entiende que existirá un sistema fundamental solamente en el caso en que el sistema (1) tenga soluciones no nulas, o sea, *cuando el rango de su matriz de los coeficientes sea menor que* el número de las incógnitas. En tal caso, el sistema (1) puede tener muchos sistemas de soluciones fundamentales diversas. Sin embargo, todos estos sistemas serán equivalentes entre sí, puesto que cada vector de cada uno de estos sistemas se expresa linealmente mediante cualquier otro sistema. Por ello, los sistemas *constan de un mismo número de soluciones*.

Subsiste el siguiente teorema:

Si el rango r de la matriz de los coeficientes de un sistema de ecuaciones lineales homogéneas (1) *es menor que el número de las incógnitas n, entonces cualquier sistema fundamental de soluciones del sistema* (1) *consta de $n - r$ soluciones.*

Para la demostración, observemos que $n - r$ es el número de incógnitas independientes en el sistema (1); supongamos que las incógnitas independientes son: $x_{r+1}, x_{r+2}, \ldots, x_n$. Consideremos un determinante cualquiera d, de orden $n - r$ diferente de cero, que lo escribiremos de la forma siguiente:

$$d = \begin{vmatrix} c_{1,\,r+1}, & c_{1,\,r+2}, & \ldots, & c_{1n} \\ c_{2,\,r+1}, & c_{2,\,r+2}, & \ldots, & c_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n-r,\,r+1}, & c_{n-r,\,r+2}, & \ldots, & c_{n-r,\,n} \end{vmatrix}.$$

Tomando los elementos de la i-ésima fila de este determinante, $1 \leqslant i \leqslant n - r$, como valores para las incógnitas independientes, obtenemos como es sabido, unos valores unívocamente determinados para las incógnitas $x_1, x_2, \ldots, x_r$, o sea, llegaremos a una solución completamente determinada del sistema de ecuaciones (1); escribamos esta solución en forma de vector

$$\alpha_i = (c_{i1}, c_{i2}, \ldots, c_{ir}, c_{i,\,r+1}, c_{i,\,r+2}, \ldots, c_{in}).$$

El sistema obtenido de vectores $\alpha_1, \alpha_2, \ldots, \alpha_{n-r}$ representa un sistema fundamental de soluciones del sistema de ecuaciones (1). En efecto, este sistema de vectores es linealmente independiente, puesto que la matriz formada por estos vectores como filas contiene un menor d, de orden $n-r$, diferente de cero. Por otra parte, sea

$$\beta=(b_1,\ b_2,\ \ldots,\ b_r,\ b_{r+1},\ b_{r+2},\ \ldots,\ b_n)$$

una solución arbitraria del sistema de ecuaciones (1). Demostremos que el vector β se expresa linealmente mediante los vectores $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{n-r}$.

Designemos con α_i', $i=1,\ 2,\ \ldots,\ n-r$, la i-ésima fila del determinante d, considerada como un vector de $(n-r)$ dimensiones y sea

$$\beta'=(b_{r+1},\ b_{r+2},\ \ldots,\ b_n).$$

Los vectores α_i', $i=1,\ 2,\ \ldots,\ n-r$ son linealmente independientes, ya que $d\neq 0$. Sin embargo, el sistema de vectores de $(n-r)$ dimensiones

$$\alpha_1',\ \alpha_2',\ \ldots,\ \alpha_{n-r}',\ \beta'$$

es linealmente dependiente, debido a que en éste el número de vectores es mayor que las dimensiones de ellos. Por consiguiente, existen unos números $k_1,\ k_2,\ \ldots,\ k_{n-r}$ tales que

$$\beta'=k_1\alpha_1'+k_2\alpha_2'+\ldots+k_{n-r}\alpha_{n-r}'. \tag{4}$$

Examinemos ahora el vector de n dimensiones

$$\delta=k_1\alpha_1+k_1\alpha_2+\ldots+k_{n-r}\alpha_{n-r}-\beta.$$

El vector δ, siendo combinación lineal de las soluciones del sistema de ecuaciones homogéneas (1), representa también una solución del mismo. De la igualdad (4) se deduce que en la solución δ los valores para todas las incógnitas independientes son iguales a cero. No obstante, la única solución del sistema de ecuaciones (1) que resulta con los valores iguales a cero para las incógnitas independientes, es la solución nula. Por lo tanto, $\delta=0$, de donde

$$\beta=k_1\alpha_1+k_2\alpha_2+\ldots+k_{n-r}\alpha_{n-r}.$$

El teorema queda demostrado.

Obsérvese que la demostración expuesta nos permite afirmar que tomando **por d todos** los determinantes posibles de orden $n-r$, diferentes de cero, **obtenemos todos** los sistemas fundamentales de soluciones del sistema de ecuaciones homogéneas (1).

Ejemplo. Sea dado el sistema de ecuaciones lineales homogéneas

$$\left.\begin{aligned}3x_1+\ x_2-\ 8x_3+\ 2x_4+\ x_5&=0,\\ 2x_1-\ 2x_2-\ 3x_3-\ 7x_4+2x_5&=0,\\ x_1+11x_2-12x_3+34x_4-5x_5&=0,\\ x_1-\ 5x_2+\ 2x_3-16x_4+3x_5&=0.\end{aligned}\right\}$$

El *rango de la matriz de los coeficientes es igual a dos; el número de* incógnitas es igual a cinco. Por consiguiente, cualquier sistema fundamental de soluciones de este sistema de ecuaciones consta de tres soluciones. Resolvamos el sistema limitándonos a las dos primeras ecuaciones linealmente independientes y tomando como independientes las incógnitas x_3, x_4, x_5. Obtenemos la solución general de la forma:

$$x_1=\frac{19}{8}x_3+\frac{3}{8}x_4-\frac{1}{2}x_5,$$

$$x_2=\frac{7}{8}x_3-\frac{25}{8}x_4+\frac{1}{2}x_5.$$

Tomemos luego los siguientes tres vectores de tres dimensiones, linealmente independientes: (1, 0, 0), (0, 1, 0), (0, 0, 1). Sustituyendo las componentes de cada uno de ellos en la solución general, en calidad de valores para las incógnitas independientes, y, calculando los valores para x_1 y x_2, obtenemos el siguiente sistema fundamental de soluciones del sistema dado de ecuaciones:

$$\alpha_1=\left(\frac{19}{8},\ \frac{7}{8},\ 1,\ 0,\ 0\right),\quad \alpha_2=\left(\frac{3}{8},\ -\frac{25}{8},\ 0,\ 1,\ 0\right),$$

$$\alpha_3=\left(-\frac{1}{2},\ \frac{1}{2},\ 0,\ 0,\ 1\right).$$

Por último, consideremos la relación que existe entre las soluciones *de los sistemas homogéneos* y *no homogéneos*. Sea dado un sistema de ecuaciones lineales no homogéneas:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots\quad a_{2n}x_n=b_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=b_s. \end{array}\right\} \qquad (5)$$

El sistema de ecuaciones lineales homogéneas

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{s1}x_1+a_{s2}x_2+\ldots+a_{sn}x_n=0, \end{array}\right\} \qquad (6)$$

obtenido del sistema (5) al sustituir por ceros los términos independientes, se llama *sistema reducido*. Entre las soluciones de los sistemas (5) y (6) existe una notable relación, como lo muestran perfectamente los dos teoremas siguientes:

I. *La suma de cualquier solución del sistema* (5) *con cualquier solución del sistema reducido* (6) *será nuevamente solución del sistema* (5).

En efecto, sea $c_1, c_2, \ldots, c_n$ una solución del sistema (5) y $d_1, d_2, \ldots, d_n$, una solución del sistema (6). Tomemos cualquiera de las ecuaciones del sistema (5), por ejemplo la k-ésima, y sustituyamos en ella, en lugar de las incógnitas, los números $c_1 + d_1$, $c_2 + d_2$, . . ., $c_n + d_n$. Resulta:

$$\sum_{j=1}^{n} a_{kj}(c_j + d_j) = \sum_{j=1}^{n} a_{kj}c_j + \sum_{j=1}^{n} a_{kj}d_j = b_k + 0 = b_k.$$

II. La diferencia de dos soluciones cualesquiera del sistema es solución del sistema reducido (6).

En efecto, sean $c_1, c_2, \ldots, c_n$ y $c'_1, c'_2, \ldots, c'_n$ dos soluciones del sistema (5). Tomemos cualquiera de las ecuaciones del sistema (6), por ejemplo la k-ésima, y sustituyamos en ella, en lugar de las incógnitas, los números

$$c_1 - c'_1,\ c_2 - c'_2,\ \ldots,\ c_n - c'_n.$$

Resulta:

$$\sum_{j=1}^{n} a_{kj}(c_j - c'_j) = \sum_{j=1}^{n} a_{kj}c_j - \sum_{j=1}^{n} a_{kj}c'_j = b_k - b_k = 0.$$

De estos teoremas se deduce que, *hallando una solución del sistema de ecuaciones lineales no homogéneas* (5) *y sumándola con cada una de las soluciones del sistema reducido* (6), *obtenemos todas las soluciones del sistema* (5).

CAPITULO III

ALGEBRA DE LAS MATRICES

§ 13. Multiplicación de matrices

En los capítulos anteriores el concepto de matriz se había empleado como un instrumento auxiliar, esencial para el estudio de los sistemas de ecuaciones lineales. Las numerosas y diversas aplicaciones de este concepto contribuyeron a convertirlo en el objetivo de una amplia teoría particular que, en gran parte, sale fuera de los márgenes de nuestro curso. Ahora nos ocuparemos de los fundamentos de esta teoría que comienza definiendo de un modo original, pero bien fundamentado, dos operaciones algebraicas: la suma y la multiplicación, aplicables al conjunto de todas las matrices cuadradas de un orden dado. Examinemos primero la definición del producto de matrices; la suma de matrices la veremos en el § 15.

Por el curso de geometría analítica se sabe que al girar los ejes de un sistema rectangular de coordenadas en el plano, en un ángulo α, las coordenadas de los puntos se transforman según las fórmulas siguientes:

$$x = x' \cos\alpha - y' \operatorname{sen}\alpha,$$
$$y = x' \operatorname{sen}\alpha + y' \cos\alpha,$$

donde x, y son las coordenadas primitivas del punto, mientras que x', y' son sus coordenadas nuevas; por lo tanto, x e y se expresan linealmente mediante x' e y', con ciertos coeficientes numéricos. En diversas ocasiones también nos encontramos con la necesidad de efectuar una transformación de las indeterminadas (o de las variables) tal, que las indeterminadas primitivas queden expresadas linealmente mediante las nuevas; ordinariamente, esta sustitución de las indeterminadas se llama transformación lineal (o sustitución lineal). Por consiguiente, llegamos a la siguiente definición:

Se llama *transformación lineal de las indeterminadas* al paso del sistema de n indeterminadas $x_1, x_2, \ldots, x_n$ al sistema de n indeterminadas $y_1, y_2, \ldots, y_n$, de manera que las indeterminadas primitivas queden expresadas linealmente mediante las nuevas con ciertos coeficientes numéricos:

$$\left.\begin{array}{l} x_1 = a_{11}y_1 + a_{12}y_2 + \ldots + a_{1n}y_n, \\ x_2 = a_{21}y_1 + a_{22}y_2 + \ldots + a_{2n}y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = a_{n1}y_1 + a_{n2}y_2 + \ldots + a_{nn}y_n. \end{array}\right\} \qquad (1)$$

La transformación lineal (1) se determina completamente por la matriz de los coeficientes

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

puesto que dos transformaciones lineales con una misma matriz pueden diferenciarse entre sí solamente en las letras que designan las indeterminadas; sin embargo, nosotros supondremos que la elección de estas notaciones corre a nuestro cargo. Recíprocamente, dada una matriz arbitraria de n-ésimo orden, podemos escribir inmediatamente una transformación lineal, para la que esta matriz sirva de matriz de sus coeficientes. Por lo tanto, entre las transformaciones lineales de n indeterminadas y las matrices cuadradas de n-ésimo orden existe una correspondencia biunívoca, y por ello, a cualquier noción ligada con las transformaciones lineales y a cualquier propiedad de estas transformaciones tiene que corresponder una noción o una propiedad análoga, referente a las matrices.

Examinemos la cuestión sobre la realización consecutiva de dos transformaciones lineales. Supongamos que después de la transformación lineal (1), se ha realizado la transformación lineal

$$\left.\begin{array}{l} y_1 = b_{11}z_1 + b_{12}z_2 + \ldots + b_{1n}z_n, \\ y_2 = b_{21}z_1 + b_{22}z_2 + \ldots + b_{2n}z_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n = b_{n1}z_1 + b_{n2}z_2 + \ldots + b_{nn}z_n, \end{array}\right\} \qquad (2)$$

que sustituye al sistema de indeterminadas $y_1, y_2, \ldots, y_n$ por el sistema $z_1, z_2, \ldots, z_n$; designemos con B la matriz de esta transformación. Sustituyendo en (1) las expresiones para $y_1, y_2, \ldots, y_n$, dadas en (2), llegaremos a unas expresiones lineales para las indeterminadas $x_1, x_2, \ldots, x_n$ mediante las indeterminadas $z_1, z_2, \ldots, z_n$. Por lo tanto, el *resultado de la realización consecutiva de dos transformaciones lineales de las indeterminadas, es de nuevo una transformación lineal.*

Ejemplo. El resultado de la realización consesutiva de las transformaciones lineales

$$x_1 = 3y_1 - y_2, \qquad y_1 = z_1 + z_2,$$
$$x_2 = y_1 + 5y_2, \qquad y_2 = 4z_1 + 2z_2$$

es la transformación lineal

$$x_1 = 3(z_1 + z_2) - (4z_1 + 2z_2) = -z_1 + z_2,$$
$$x_2 = (z_1 + z_2) + 5(4z_1 + 2z_2) = 21z_1 + 11z_2.$$

Designemos con C la matriz de la transformación lineal que representa el resultado de la realización consecutiva de las transformaciones (1) y (2), y hallemos la ley por la que se expresan los elementos c_{ik}, $i, k = 1, 2, \ldots, n$ mediante los elementos de las matrices A y B. Escribiendo abreviadamente las transformaciones (1) y (2) en la forma

$$x_i = \sum_{j=1}^{n} a_{ij} y_j \quad i = 1, 2, \ldots, n; \quad y_j = \sum_{k=1}^{n} b_{jk} z_k, \quad j = 1, 2, \ldots, n,$$

obtenemos

$$x_i = \sum_{j=1}^{n} a_{ij} \Big(\sum_{k=1}^{n} b_{jk} z_k \Big) = \sum_{k=1}^{n} \Big(\sum_{j=1}^{n} a_{ij} b_{jk} \Big) z_k, \quad i = 1, 2, \ldots, n.$$

En consecuencia, el coeficiente de z_k en la expresión para x_i, es decir, el elemento c_{ik} de la matriz C, tiene la forma

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} = a_{i1} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk}; \tag{3}$$

el elemento de la matriz C, situado en la i-ésima fila y en la k-ésima columna, es igual a la suma de los productos de los correspondientes elementos de la i-ésima fila de la matriz A y de la k-ésima columna de la matriz B.

La fórmula (3), que da la expresión de los elementos de la matriz C mediante los elementos de las matrices A y B, permite escribir inmediatamente la matriz C, siendo dadas las matrices A y B, sin recurrir a las transformaciones lineales correspondientes a estas matrices. De este modo, a cualquier par de matrices cuadradas de n-ésimo orden se pone en correspondencia una tercera matriz unívocamente determinada. Se puede decir que hemos definido una operación algebraica en el conjunto de todas las matrices cuadradas de n-ésimo orden; ésta se llama *multiplicación de las matrices*, y la matriz C, *producto* de la matriz A por la matriz B:

$$C = AB.$$

Enunciemos una vez más la relación entre las transformaciones lineales y el producto de las matrices:

La transformación lineal de las indeterminadas obtenida como resultado de la realización consecutiva de dos transformaciones lineales con las matrices A y B, tiene a la matriz AB por matriz de sus coeficientes.

Ejemplos

1) $$\begin{pmatrix} 4 & 9 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 4\cdot 1+9\cdot(-2) & 4\cdot(-3)+9\cdot 1 \\ (-1)\cdot 1+3\cdot(-2) & (-1)\cdot(-3)+3\cdot 1 \end{pmatrix} = \begin{pmatrix} -14 & -3 \\ -7 & 6 \end{pmatrix}.$$

2) $$\begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} -3 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} = \begin{pmatrix} -6 & 1 & 3 \\ 6 & 2 & 9 \\ -12 & -3 & 14 \end{pmatrix}.$$

3) $$\begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix}^2 = \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 51 & 16 \\ 8 & 3 \end{pmatrix}.$$

4) Hallar el resultado de la realización consecutiva de las transformaciones lineales

$$\begin{aligned} x_1 &= 5y_1 - y_2 + 3y_3, \\ x_2 &= y_1 - 2y_2, \\ x_3 &= 7y_2 - y_3 \end{aligned}$$

y

$$\begin{aligned} y_1 &= 2z_1 + z_3, \\ y_2 &= z_2 - 5z_3, \\ y_3 &= 2z_2. \end{aligned}$$

Multiplicando las matrices, obtenemos:

$$\begin{pmatrix} 5 & -1 & 3 \\ 1 & -2 & 0 \\ 0 & 7 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -5 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 5 & 10 \\ 2 & -2 & 11 \\ 0 & 5 & -35 \end{pmatrix},$$

de donde, la transformación lineal buscada tiene la forma:

$$\begin{aligned} x_1 &= 10z_1 + 5z_2 + 10z_3, \\ x_2 &= 2z_1 - 2z_2 + 11z_3, \\ x_3 &= 5z_2 - 35z_3. \end{aligned}$$

Tomemos uno de los ejemplos que acabamos de estudiar de multiplicación de las matrices, por ejemplo el 2), y hallemos el producto de las mismas matrices, pero tomadas en orden inverso:

$$\begin{pmatrix} -3 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 1 \\ -2 & 3 & 2 \\ 4 & -1 & 5 \end{pmatrix} = \begin{pmatrix} -8 & 3 & -1 \\ 0 & 5 & 9 \\ 14 & -6 & 13 \end{pmatrix}.$$

Vemos, pues, que el producto de las matrices depende del orden de los factores, es decir, que la *multiplicación de las matrices no es conmutativa*. Por cierto, esto era de esperar, aunque sólo sea por el hecho de que, en la definición de la matriz C, dada anteriormente mediante la fórmula (3), las matrices A y B no figuran de un modo equivalente: en A se toman las filas, mientras que en B, las columnas.

Se pueden señalar, para todos los n, comenzando desde $n = 2$, ejemplos de matrices de n-ésimo orden no conmutables, o sea, de matrices cuyo producto se altera al permutar los factores (las matrices de segundo orden en el ejemplo 1) no son conmutables). Por otra parte, *dos matrices pueden ser ocasionalmente conmutables*, como muestra el siguiente ejemplo:

$$\begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix} \cdot \begin{pmatrix} 26 & 45 \\ 15 & 26 \end{pmatrix} = \begin{pmatrix} 26 & 45 \\ 15 & 26 \end{pmatrix} \cdot \begin{pmatrix} 7 & -12 \\ -4 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}.$$

El producto de las matrices es asociativo; por consiguiente, se puede hablar del producto, unívocamente determinado, de cualquier número finito de matrices de n-ésimo orden, tomadas (en virtud de que el producto no es conmutativo) en un orden determinado.

Demostración. Sean dadas tres matrices arbitrarias de n-ésimo orden, A, B, y C. Escribámoslas del modo abreviado siguiente, donde se indica la forma general de sus elementos: $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. Introduzcamos luego las siguientes notaciones:

$$AB = U = (u_{ij}), \qquad BC = V = (v_{ij}),$$
$$(AB)C = S = (s_{ij}), \qquad A(BC) = T = (t_{ij}).$$

Tenemos que demostrar que se cumple la igualdad $(AB)C = A(BC)$, es decir, $S = T$. Sin embargo,

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

de donde, en virtud de las igualdades $S = UC$ y $T = AV$,

$$s_{ij} = \sum_{l=1}^{n} u_{il} c_{lj} = \sum_{l=1}^{n} \sum_{k=1}^{n} a_{ik} b_{kl} c_{lj},$$

$$t_{ij} = \sum_{k=1}^{n} a_{ik} v_{kj} = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} b_{kl} c_{lj},$$

o sea, $s_{ij} = t_{ij}$ para $i, j = 1, 2, \ldots, n$.

Para el estudio ulterior de las propiedades del producto de las matrices se necesita el empleo de los determinantes. Además, para abreviar, convendremos en designar con $|A|$ el determinante de la matriz A. Si en cada uno de los ejemplos considerados anteriormente el lector calcula los determinantes de las matrices que se multiplican y compara el producto de estos determinantes con el determinante del producto de las matrices dadas, puede observar una ley bastante curiosa que se expresa con el siguiente importante **teorema sobre el producto de los determinantes:**

El determinante del producto de varias matrices de n-ésimo orden es igual al producto de los determinantes de estas matrices.

Es suficiente demostrar este teorema para el caso de dos matrices. Sean dadas las matrices de n-ésimo orden $A = (a_{ij})$ y $B = (b_{ij})$, y sea $AB = C = (c_{ij})$. Formemos el siguiente determinante auxiliar Δ de orden $2n$: en su ángulo superior de la izquierda colocamos la matriz A, en el ángulo inferior de la derecha, la matriz B, todo el ángulo superior de la derecha lo ocupamos con ceros. Finalmente, formamos la diagonal principal del ángulo inferior de la izquierda con el número -1, ocupando todos los demás lugares también con ceros. Por consiguiente, el determinante Δ tiene la forma siguiente:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix}.$$

La aplicación del teorema de Laplace al determinante — su desarrollo por los menores de las primeras n filas — nos lleva a la siguiente igualdad:

$$\Delta = |A| \cdot |B|.$$

Procuremos, a su vez, transformar el determinante Δ de tal modo que, sin cambiar su valor, todos los elementos b_{ij}, $i, j = 1, 2, \dots, n$, queden sustituidos por ceros. Con este fin, agreguemos a la $(n+1)$-ésima columna del determinante Δ su primera columna, multiplicada por b_{11}, su segunda columna, multiplicada por b_{21}, etc., y finalmente, su n-ésima columna, multiplicada por b_{n1}. Después, agreguemos a la $(n+2)$-ésima columna del determinante Δ la primera columna, multiplicada por b_{12}, la segunda columna, multiplicada por b_{22}, etc. En general, agreguemos a la $(n+j)$-

ésima columna del determinante Δ, donde $j = 1, 2, \ldots, n$, la suma de las primeras n columnas, tomadas con los coeficientes $b_{1j}, b_{2j}, \ldots$ $\ldots, b_{nj}$, respectivamente.

Fácilmente se ve que estas transformaciones, no alterando el determinante, dan lugar a la sustitución de todos los elementos b_{ij} por ceros. A la vez, en lugar de los ceros que figuraban en el ángulo superior de la derecha del determinante, aparecerán los números siguientes: en la intersección de la i-ésima fila y $(n + j)$-ésima columna del determinante, $i, j = 1, 2, \ldots, n$, estará ahora el número $a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$, que en virtud de (3) es igual al elemento c_{ij} de la matriz $C = AB$. Por consiguiente, el ángulo superior de la derecha del determinante lo ocupa ahora la matriz C:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} & c_{11} & c_{12} & \ldots & c_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} & c_{21} & c_{22} & \ldots & c_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} & c_{n1} & c_{n2} & \ldots & c_{nn} \\ -1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 \end{vmatrix}$$

Apliquemos otra vez más el teorema de Laplace, desarrollando el determinante por los menores de las últimas n columnas. Como el menor complementario para el menor $|C|$ es igual a $(-1)^n$, y el menor $|C|$ está situado en las filas cuyos números de orden son $1, 2, \ldots, n$ y en las columnas cuyos números de orden son $n + 1, n + 2, \ldots, 2n$, aplicando la igualdad

$$1 + 2 + \ldots + n + (n + 1) + (n + 2) + \ldots + 2n = 2n^2 + n,$$

se tiene

$$\Delta = (-1)^{2n^2+n}(-1)^n|C| = (-1)^{2(n^2+n)}|C|$$

o bien, como el número $2(n^2 + n)$ es par,

$$\Delta = |C|. \tag{5}$$

Finalmente, de (4) y (5) se deduce la igualdad que queríamos demostrar

$$|C| = |A| \cdot |B|.$$

El teorema sobre el producto de los determinantes podría haber sido demostrado también sin la utilización del teorema de Laplace. El lector hallará una de estas demostraciones al final del § 16.

§ 14. Matriz inversa

Una matriz cuadrada se llama *degenerada* (o *singular*), si su determinante es igual a cero, y *no degenerada* (o *no singular*)*, en el caso contrario. Correspondientemente, una transformación lineal de las indeterminadas se llama *degenerada* o *no degenerada*, según que el determinante de los coeficientes de esta transformación sea igual a cero o no. Del teorema demostrado al final del párrafo anterior, se deduce la afirmación siguiente:

El producto de matrices, al menos una de las cuales es degenerada, es también una matriz degenerada.

El producto de cualesquiera matrices no degeneradas también es una matriz no degenerada. De aquí se deduce, en virtud de la relación existente entre el producto de las matrices y la realización consecutiva de las transformaciones lineales, la proposición siguiente: *el resultado de la realización consecutiva de unas cuantas transformaciones lineales será una transformación no degenerada cuando, y sólo cuando, todas las transformaciones dadas sean no degeneradas.*

En el producto de las matrices, el papel de la unidad lo desempeña la matriz

$$E=\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & & 1 \end{pmatrix},$$

que es además conmutable con cualquier matriz A del orden dado,

$$AE=EA=A. \tag{1}$$

Estas igualdades se demuestran aplicando directamente la regla de multiplicación de las matrices, o basándose en la observación de que la matriz unidad corresponde a la transformación lineal *idéntica* de las indeterminadas

$$\begin{aligned} x_1 &= y_1, \\ x_2 &= \quad y_2, \\ &\cdots\cdots \\ x_n &= \qquad y_n, \end{aligned}$$

cuya realización, antes o después de cualquier otra transformación lineal, no cambia, evidentemente, esta última.

Obsérvese que la *matriz E es la única que satisface a la condición* (1) *para cualquier matriz A*. En efecto, si existiese otra matriz E' con esta propiedad, tendríamos que

$$E'E=E', \quad E'E=E,$$

de donde, $E'=E$.

* También se llama matriz regular. (*Nota del T.*)

El problema de la existencia de la *matriz inversa* para una matriz dada A es más complicado. Como el producto de las matrices no es conmutativo, hablaremos ahora de la matriz inversa a la *derecha*, o sea, de una matriz A^{-1} tal, que el producto de la matriz A por esta matriz a la derecha es igual a la matriz unidad,

$$AA^{-1} = E. \tag{2}$$

Si la matriz A es degenerada y existiese la matriz A^{-1}, el producto que figura en el primer miembro de la igualdad (2) sería, como ya sabemos, una matriz degenerada. En realidad, la matriz E que figura en el segundo miembro de esta igualdad no es degenerada, puesto que su determinante es igual a la unidad. Por lo tanto, una matriz degenerada no puede tener matriz inversa a la derecha. Estos mismos razonamientos muestran que ésta no puede tener tampoco matriz inversa a la izquierda, *no existiendo, por lo tanto, matriz inversa para una matriz degenerada.*

Refiriéndonos al caso de una matriz no degenerada, introduzcamos primero el siguiente concepto auxiliar. Sea dada una matriz de n-ésimo orden

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

La matriz

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

formada por los complementos algebraicos de los elementos de la matriz A, donde el complemento algebraico del elemento a_{ij} está situado en la intersección de la j-ésima fila y de la i-ésima columna, se llama matriz *adjunta* de la matriz A.

Hallemos los poductos AA^* y A^*A. Aplicando la fórmula estudiada en el § 6, sobre el desarrollo de un determinante por los elementos de una fila o columna, y también el teorema del § 7, sobre la suma de los productos de los elementos de cualquier fila (o columna) de un determinante por los complementos algebraicos de los elementos correspondientes de otra fila (columna), y designando con d el determinante de la matriz A,

$$d = |A|,$$

obtenemos las siguientes igualdades:

$$AA^* = A^*A = \begin{pmatrix} d & 0 \dots 0 \\ 0 & d \dots 0 \\ \dots & \dots \\ 0 & 0 \dots d \end{pmatrix}. \qquad (3)$$

De aquí se deduce que, si la *matriz A no es degenerada, su matriz adjunta A^* tampoco lo será, siendo además, el determinante d^* de la matriz A^* igual a la $(n-1)$-ésima potencia del determinante d de la matriz A.*

En efecto, pasando de las igualdades (3) a la igualdad entre los determinantes, obtenemos:

$$dd^* = d^n,$$

de donde, como $d \neq 0$,

$$d^* = d^{n-1}\text{*}).$$

Ahora es fácil demostrar la existencia de una matriz inversa para cualquier matriz A **no degenerada,** y hallar su forma. Obsérvese primero que si se considera el producto de dos matrices AB y se dividen por un mismo número d todos los elementos de uno de los factores, por ejemplo B, entonces, todos los elementos del producto AB también se dividirán por este mismo número: para la demostración solamente hay que recordar la definición del producto de las matrices. Por lo tanto, si

$$d = |A| \neq 0,$$

de las igualdades (3) se deduce, que la *inversa de la matriz A es la matriz que resulta de la matriz adjunta A^* al dividir todos sus elementos por el número d*:

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

En efecto, de (3) se deducen las igualdades

$$AA^{-1} = A^{-1}A = E. \qquad (4)$$

Subrayemos una vez más que en la i-ésima **fila** de la matriz A^{-1} figuran los complementos algebraicos de los elementos de la i-ésima **columna** del determinante $|A|$, divididos por $d = |A|$.

* Se podría demostrar que si la matriz A es degenerada, su matriz adjunta A^* también lo es, teniendo además un rango no superior al número 1.

Es fácil demostrar que la *matriz* A^{-1} *es la única que satisface a la condición* (4) *para una matriz dada* A, *no degenerada*. En efecto, si la matriz C es tal que

$$AC = CA = E,$$

entonces,

$$CAA^{-1} = C(AA^{-1}) = CE = C,$$
$$CAA^{-1} = (CA)A^{-1} = EA^{-1} = A^{-1},$$

de donde $C = A^{-1}$.

De (4) y del teorema sobre el producto de los determinantes, se deduce que el *determinante de la matriz* A^{-1} *es igual a* $\frac{1}{|A|}$. *Así, pues, esta matriz tampoco es degenerada; la inversa para ella es la misma matriz* A.

Si se dan ahora las matrices cuadradas A y B de n-ésimo orden, de las cuales A no es degenerada, mientras que B es arbitraria, podemos efectuar la división **por la derecha y por la izquierda** de B por A, es decir, resolver las ecuaciones matriciales

$$AX = B, \quad YA = B. \tag{5}$$

Para esto, en virtud de la asociatividad del producto de las matrices, es suficiente hacer

$$X = A^{-1}B, \quad Y = BA^{-1};$$

como el producto de las matrices no es conmutativo, por lo general, estas soluciones de las ecuaciones (5) serán diferentes.

Ejemplos. 1) Se da la matriz

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

Su determinante $|A| = 5$, por consiguiente, la matriz inversa A^{-1} existe:

$$A^{-1} = \begin{pmatrix} 1 & \frac{4}{5} & -\frac{1}{5} \\ 2 & \frac{12}{5} & -\frac{3}{5} \\ 0 & \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

2) Se dan las matrices

$$A = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 7 \\ 3 & 5 \end{pmatrix}.$$

La matriz A no es degenerada, y además,

$$A^{-1}=\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix},$$

por lo tanto, las soluciones de las ecuaciones $AX=B$, $YA=B$ serán las matrices

$$X=\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}\cdot\begin{pmatrix} -1 & 7 \\ 3 & 5 \end{pmatrix}=\begin{pmatrix} -9 & 11 \\ 13 & -13 \end{pmatrix},$$

$$Y=\begin{pmatrix} -1 & 7 \\ 3 & 5 \end{pmatrix}\cdot\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}=\begin{pmatrix} -31 & 23 \\ -11 & 9 \end{pmatrix}.$$

Multiplicación de matrices rectangulares. El producto de las matrices definido en el párrafo anterior solamente para las matrices cuadradas de igual orden, puede generalizarse también para el caso de matrices rectangulares A y B, siempre que sea posible aplicar la fórmula (3) del párrafo anterior, o sea, cuando cada fila de la matriz A contenga tantos elementos como haya en cada columna de la matriz B. En otras palabras, *se puede hablar del producto de las matrices rectangulares A y B cuando el número de columnas de la matriz A es igual al número de filas de la matriz B. En este caso, el número de filas de la matriz AB es igual al número de filas de la matriz A, y el número de columnas de la matriz AB es igual al número de columnas de la matriz B.*

Ejemplos.

1) $$\begin{pmatrix} 5 & -1 & 3 & 1 \\ 2 & 0 & -1 & 4 \end{pmatrix}\cdot\begin{pmatrix} -1 & 3 & 0 \\ -2 & 1 & 1 \\ 3 & 0 & -2 \\ 4 & 1 & 2 \end{pmatrix}=\begin{pmatrix} 10 & 15 & -5 \\ 11 & 10 & 10 \end{pmatrix}.$$

2) $$\begin{pmatrix} 0 & -3 & 1 \\ 2 & 1 & 5 \\ -4 & 0 & -2 \end{pmatrix}\cdot\begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}=\begin{pmatrix} 8 \\ 14 \\ -16 \end{pmatrix}.$$

3) $$(5 \quad 1 \quad 0 \quad -3)\cdot\begin{pmatrix} 2 & 0 \\ 1 & -4 \\ 3 & 1 \\ 0 & -1 \end{pmatrix}=(11 \quad -1).$$

El producto de las matrices rectangulares se puede ligar con la ejecución consecutiva de las transformaciones lineales de las indeterminadas solamente si en la definición de estas últimas no se insiste en que se conserve el número de indeterminadas en la transformación lineal.

Repitiendo palabra por palabra la demostración dada anteriormente para el caso de las matrices cuadradas, se comprueba fácilmente que la *ley asociativa se cumple también para el producto de matrices rectangulares.*

Ahora utilizaremos el producto de las matrices rectangulares y las propiedades de la matriz inversa para **deducir de nuevo la regla de Cramer,** evitando los complicados cálculos que se realizaron en el § 7. *Sea dado un sistema de n ecuaciones lineales con n incógnitas:*

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n, \end{array}\right\} \tag{6}$$

donde el determinante del sistema es diferente de cero. Designemos con A la matriz de los coeficientes del sistema (6); esta matriz no es degenerada, ya que, por la suposición hecha, $d = |A| \neq 0$. Designemos con X *la columna de las incógnitas; con* B, *la columna* de los términos independientes del sistema (6), es decir,

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

El producto AX tiene sentido, puesto que el número de columnas de la matriz A es igual al número de filas de la matriz X, y además, este producto será una columna, formada por los primeros miembros de las ecuaciones del sistema (6). Por lo tanto, el sistema (6) puede escribirse en forma de una ecuación matricial

$$AX = B. \tag{7}$$

Multiplicando a la izquierda ambos miembros de la ecuación (7) por la matriz A^{-1}, cuya existencia es consecuencia de que la matriz cuadrada A no es degenerada, obtenemos:

$$X = A^{-1}B. \tag{8}$$

El producto que figura en el segundo miembro de esta igualdad es una matriz de una sola columna; su j-ésimo elemento es igual a la suma de los productos de los elementos de la j-ésima fila de la matriz A^{-1} por los elementos correspondientes de la matriz B, es decir, es igual al número

$$\frac{A_{1j}}{d} b_1 + \frac{A_{2j}}{d} b_2 + \ldots + \frac{A_{nj}}{d} b_n = \frac{1}{d}(A_{1j}b_1 + A_{2j}b_2 + \ldots + A_{nj}b_n).$$

Sin embargo, la expresión que figura entre paréntesis en el segundo miembro de la igualdad es el desarrollo por los elementos de la j-ésima columna del determinante d_j, que se obtiene sustituyendo la j-ésima columna del determinante d por la columna B. Por lo tanto, las fórmulas (8) son equivalentes a las fórmulas (3) del § 7, que expresan la solución del sistema (6) obtenida por la regla de Cramer.

Queda por demostrar que los valores obtenidos de las incógnitas forman verdaderamente una solución del sistema (6). Con este fin, es suficiente poner la expresión (8) en la ecuación matricial (7), lo que da lugar, evidentemente, a la identidad $B = B$.

El rango del producto de las matrices. En el caso de matrices degeneradas, el teorema del producto de los determinantes no nos lleva a ningún enunciado más de que tal producto también es degenerado, a pesar de que las matrices cuadradas degeneradas se pueden diferenciar también por su rango. Obsérvese que no existe una dependencia absolutamente determinada entre los rangos de los factores y el rango del producto, como se muestra en los ejemplos siguientes:

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

en ambos casos se multiplican matrices de rango 1. Sin embargo, en un caso el producto tiene el rango 1, mientras que en el otro, el rango es 0. Subsiste solamente el siguiente teorema, que no sólo es justo para las matrices cuadradas, sino también para las matrices rectangulares:

El rango del producto de varias matrices no es superior al rango de cada uno de los factores.

Es suficiente demostrar este teorema para el caso de dos factores. Sean dadas las matrices A y B, para las cuales tiene sentido el producto AB; emplearemos la notación $AB = C$. Veamos la fórmula (3) del § 13, que da la expresión de los elementos de la matriz C. Tomando esta fórmula para un k dado y todos los i posibles, $(i = 1, 2, \ldots)$ obtenemos que la k-ésima columna de la matriz C representa una suma de todas las columnas de la matriz A, tomadas con ciertos coeficientes (precisamente con los coeficientes b_{1k}, b_{2k}, ...). De este modo, queda demostrado que el sistema de columnas de la matriz C se expresa linealmente mediante el sistema de columnas de la matriz A, y, por consiguiente, como se ha demostrado en el § 9, el rango del primer sistema es menor o igual al rango del segundo sistema; en otras palabras, el rango de la matriz C no es mayor que el rango de la matriz A. Por otra parte, como de la misma fórmula (3) del

§ 13, para un i dado y todos los k, se deduce que toda i-ésima fila de la matriz C es combinación lineal de las filas de la matriz B, con razonamientos análogos obtenemos que el rango de C no es mayor que el rango de B.

Cuando uno de los factores representa una matriz cuadrada no degenerada, se obtiene un resultado más exacto.

El rango del producto a la derecha y a la izquierda de una matriz A por una matriz cuadrada no degenerada Q, es igual al rango de la matriz A.

Sea, por ejemplo,

$$AQ = C. \tag{9}$$

Del teorema precedente se deduce que el rango de la matriz C no es mayor que el rango de la matriz A. Por otra parte, multiplicando a la derecha la igualdad (9) por Q^{-1}, llegamos a la igualdad

$$A = CQ^{-1},$$

y, por consiguiente, otra vez por el teorema precedente, el rango de A no es mayor que el rango de C. Comparando estos dos resultados obtenemos la coincidencia de los rangos de las matrices A y C.

§ 15. Suma de matrices y multiplicación de una matriz por un número

Para las matrices cuadradas de orden n, la **suma** se define del modo siguiente:

Se llama *suma* $A + B$ *de dos matrices cuadradas* $A = (a_{ij})$ *y* $B = (b_{ij})$ *de orden* n, a una matriz $C = (c_{ij})$ tal, que cualquier elemento de ella es igual a la suma de los elementos correspondientes de las matrices A y B;

$$c_{ij} = a_{ij} + b_{ij} \text{*}$$

Es evidente, que la suma de matrices definida es conmutativa y asociativa. Para ella existe la operación inversa: la resta, llamándose diferencia de las matrices A y B a la matriz formada por las diferencias de los elementos correspondientes de las matrices dadas. En este caso, el papel de cero lo desempeña la *matriz nula*, compuesta totalmente de ceros; a continuación, esta matriz se designará con el símbolo 0: no hay peligro de confundir la matriz nula con el número cero.

La suma de las matrices cuadradas y el producto de éstas, definido en el § 13, están ligados con las leyes distributivas.

* Por supuesto, se podría definir también el producto de matrices multiplicando los elementos correspondientes. Sin embargo, esta multiplicación, a diferencia de la que se definió en el § 13, carecería de aplicaciones serias.

En efecto, sean dadas tres matrices de orden n, $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$. Entonces, para cualesquiera i y j, se cumple la igualdad

$$\sum_{s=1}^{n} (a_{is} + b_{is})\, c_{sj} = \sum_{s=1}^{n} a_{is}c_{sj} + \sum_{s=1}^{n} b_{is}c_{sj}.$$

El primer miembro de esta igualdad es el elemento situado en la i-ésima fila y j-ésima columna de la matriz $(A + B)\,C$, el segundo miembro es el elemento situado en el mismo lugar, pero en la matriz $AC + BC$. Con esto, queda demostrada la igualdad

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

La igualdad $C\,(A + B) = CA + CB$ se demuestra del mismo modo. Como el producto de matrices no es conmutativo, se deben demostrar estas dos leyes distributivas.

Introduzcamos la siguiente definición de producto de una matriz por un número.

Se llama *producto kA de una matriz cuadrada $A = (a_{ij})$ por el número k*, a la matriz $A' = (a'_{ij})$ que se obtiene multiplicando por k todos los elementos de la matriz A:

$$a'_{ij} = ka_{ij}.$$

En el párrafo anterior ya tratamos un ejemplo de multiplicación de una matriz por un número: si la matriz A no es degenerada, siendo $|A| = d$, su matriz inversa A^{-1} y su matriz adjunta A^* están ligadas por la igualdad

$$A^{-1} = d^{-1}A^*.$$

Como ya sabemos, toda matriz cuadrada de orden n se puede considerar como un vector de n^2 dimensiones, siendo biunívoca esta correspondencia entre las matrices y los vectores. En este caso, las operaciones definidas de suma de matrices y de producto de una matriz por un número, se convierten en la suma de vectores y en el producto de un vector por un número. Por lo tanto, el *conjunto de las matrices cuadradas de orden n se puede considerar como un espacio vectorial de n^2 dimensiones.*

De aquí se deduce el cumplimiento de las igualdades siguientes (aquí, A, B son matrices de orden n; k, l son unos números; 1 es el número uno):

$$k\,(A + B) = kA + kB, \tag{1}$$

$$(k + l)\,A = kA + lA, \tag{2}$$

$$k\,(lA) = (kl)\,A, \tag{3}$$

$$1 \cdot A = A. \tag{4}$$

Las propiedades (1) y (2) ligan el producto de una matriz por un número con la suma de matrices. Al mismo tiempo, existe una ligazón importante entre el producto de una matriz por un número y el producto de las matrices mismas, que se expresa así:

$$(kA)\,B = A\,(kB) = k\,(AB), \tag{5}$$

o sea, *si en el producto de las matrices, uno de los factores se multiplica por el número k, todo el producto queda multiplicado por k.*

En efecto, sean dadas las matrices $A = (a_{ij})$ y $B = (b_{ij})$ y el número k. Entonces, para cualesquiera i y j, se tiene:

$$\sum_{s=1}^{n} (ka_{is})\,b_{sj} = k \sum_{s=1}^{n} a_{is}b_{sj}.$$

Pero, el primer miembro de esta igualdad es el elemento situado en la i-ésima fila y j-ésima columna de la matriz $(kA)\,B$, y el segundo miembro es el elemento situado en el mismo sitio en la matriz $k\,(AB)$. Con esto queda demostrada la igualdad

$$(kA)\,B = k\,(AB).$$

La igualdad $A(kB) = k\,(AB)$ se demuestra del mismo modo.

La *multiplicación de una matriz por un número* permite introducir un nuevo método de expresión de las matrices. Designemos con E_{ij} *la matriz en la que, en la intersección de la* i-ésima fila y j-ésima columna figura la unidad, mientras que todos los demás elementos *son iguales a cero.* Haciendo $i = 1, 2, \ldots, n$ y $j = 1, 2, \ldots, n$, obtenemos n^2 matrices de éstas, E_{ij}, que, como fácilmente se comprueba, están ligadas por la siguiente tabla de multiplicar:

$$E_{is}E_{sj} = E_{ij}, \quad E_{is}E_{tj} = 0 \text{ para } s \neq t.$$

La matriz kE_{ij} solamente se diferencia de la matriz E_{ij} en que en ella figura el número k en la intersección de la i-ésima fila y j-ésima columna. Teniendo esto en cuenta y aplicando la definición de suma de matrices, obtenemos la siguiente expresión para una matriz cuadrada arbitraria A:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}E_{ij}, \tag{6}$$

poseyendo, evidentemente, la matriz A una sola expresión de la forma (6).

La matriz kE, donde E es la matriz unidad, según la definición del producto de una matriz por un número, tiene la forma siguiente:

$$kE = \begin{pmatrix} k & & & 0 \\ & k & & \\ & & \ddots & \\ 0 & & & k \end{pmatrix},$$

o sea, en la diagonal principal figura un mismo número k, mientras que todos los elementos situados fuera de esta diagonal son iguales a cero. Tales matrices se llaman *escalares.*

La definición de la suma de matrices conduce a la siguiente igualdad

$$kE + lE = (k + l)\,E. \tag{7}$$

Por otra parte, aplicando la definición del producto de matrices o basándose en la igualdad (5), obtenemos:

$$kE \cdot lE = (kl)\,E. \tag{8}$$

El producto de una matriz A por un número k se puede interpretar como el producto obtenido al multiplicar A por la matriz escalar kE, en el sentido de la multiplicación de matrices. En efecto, según (5)

$$(kE)\,A = A\,(kE) = kA.$$

De aquí se deduce también, que *toda matriz escalar es conmutable con cualquier matriz A*. Es de gran importancia tener en cuenta que las matrices escalares son las únicas que poseen esta propiedad:

Si una matriz $C = (c_{ij})$ de n-ésimo orden es conmutable con cualquier matriz del mismo orden, la matriz C es escalar.

En efecto, supongamos que $i \neq j$ y consideremos los productos CE_{ij} y $E_{ij}C$ (véase más arriba la definición de la matriz E_{ij}) que, por hipótesis, son iguales entre sí. Fácilmente se observa que todas las columnas de la matriz CE_{ij}, menos la j-ésima, se componen de ceros, y que la j-ésima columna coincide con la i-ésima columna de la matriz C; en particular, en la intersección de la i-ésima fila y j-ésima columna de la matriz CE_{ij} está situado el elemento c_{ii}. Análogamente, todas las filas de la matriz $E_{ij}C$, menos la i-ésima, se componen de ceros, y la i-ésima fila coincide con la j-ésima fila de la matriz C; en la intersección de la i-ésima fila y j-ésima columna de la matriz $E_{ij}C$ está situado el elemento c_{jj}. Aplicando la igualdad $CE_{ij} = E_{ij}C$, obtenemos que: $c_{ii} = c_{jj}$ (como elementos situados en lugares iguales de matrices que son iguales entre sí); o sea, todos los elementos de la diagonal principal de la matriz C son iguales entre sí. Por otra parte, en la intersección de la j-ésima fila y j-ésima columna de la matriz CE_{ij} está el elemento c_{ji}; pero, en la matriz $E_{ij}C$, en este sitio está el cero (en virtud de que $i \neq j$), de donde: $c_{ji} = 0$, es decir, cualquier elemento de la matriz C, situado fuera de la diagonal principal, es igual a cero. El teorema está demostrado.

§ 16. Construcción axiomática de la teoría de los determinantes

Un determinante de n-ésimo orden representa un número, unívocamente determinado por una matriz cuadrada dada de n-ésimo orden. La definición de este concepto, expuesta en el § 4, da la regla, según la cual el determinante se expresa mediante los elementos de la matriz dada. Sin embargo, esta definición constructiva se puede sustituir por una axiomática; mejor dicho, entre las propiedades del determinante establecidas en los §§ 4 y 6, se pueden indicar algunas, de modo que la única función de la matriz con valores reales que posea estas propiedades sea su determinante.

La definición más simple de este género consiste en la utilización de los desarrollos del determinante por los elementos de una fila. Consideremos las matrices cuadradas de cualesquiera órdenes y supongamos que a cada matriz M de éstas se le ha puesto en correspondencia un número d_M, cumpliéndose las condiciones siguientes:

1) Si la matriz M es de primer orden, o sea, que consta de un elemento a, entonces $d_M = a$.

2) Si los elementos $a_{11}, a_{12}, \ldots, a_{1n}$ forman la primera fila de la matriz M de n-ésimo orden y si se ha designado con M_i, $i = 1, 2, \ldots, n$, la matriz de $(n-1)$-ésimo orden que queda después de suprimir en M la primera fila y la i-ésima columna, entonces

$$d_M = a_{11}d_{M_1} - a_{12}d_{M_2} + a_{13}d_{M_3} - \ldots + (-1)^{n-1} a_{1n}d_{M_n}.$$

Por lo tanto, para cualquier matriz M, el número d_M es igual al determinante de esta matriz. La demostración de esta afirmación la dejamos a cuenta del lector. Esta se efectúa por el método de inducción sobre n y se basa en los resultados del § 6.

Mucho más interesantes son otras formas de definición axiomática de los determinantes, que se refieren al caso de un orden dado n y se basan en algunas de las propiedades simples de los determinantes, establecidas en el § 4. Examinaremos ahora una de estas definiciones.

Supongamos que a cada matriz cuadrada M de n-ésimo orden se pone en correspondencia un número d_M, cumpliéndose las condiciones siguientes:

I. *Si una de las filas de la matriz M se multiplica por un número k, el número d_M queda multiplicado por k.*

II. *El número* d_M *no varía si a una de las filas de la matriz* M *se le agrega otra fila de esta matriz.*

III. *Si* E *es la matriz unidad, entonces* $d_E = 1$.

Demostremos que *para cualquier matriz* M, *el número* d_M *es igual al determinante de esta matriz.*

Establezcamos primero, partiendo de las condiciones I—III, unas propiedades del número d_M, que son análogas a las propiedades correspondientes de los determinantes.

(1) *Si una de las filas de la matriz* M *consta de ceros, entonces* $d_M = 0$.

En efecto, multiplicando la fila compuesta de ceros por el número 0, la matriz no varía. Sin embargo, en virtud de la condición I, el número d_M adquiere el factor 0, de donde,

$$d_M = 0 \cdot d_M = 0.$$

(2) *El número* d_M *no varía si a la i-ésima fila de la matriz* M *se le agrega la j-ésima fila,* $j \neq i$, *multiplicada por el número* k.

Si $k = 0$, todo está demostrado. Si $k \neq 0$, multiplicamos la j-ésima fila por k y obtenemos la matriz M', para la que $d_{M'} = kd_M$, en virtud de la condición I. Después agregamos a la i-ésima fila de la matriz M' su j-ésima fila y obtenemos la matriz M'', y en virtud de la condición II, se tiene $d_{M''} = d_{M'}$. Finalmente, multiplicamos la j-ésima fila de la matriz M'' por el número k^{-1}, obteniendo la matriz M''', que en realidad resulta de M mediante la transformación indicada en el enunciado de la propiedad que estamos demostrando; además.

$$d_{M'''} = k^{-1}d_{M''} = k^{-1}d_{M'} = k^{-1} \cdot kd_M = d_M.$$

(3) *Si las filas de la matriz* M *son linealmente dependientes, entonces* $d_M = 0$.

En efecto, si una de las filas, por ejemplo la i-ésima, es combinación lineal de las otras filas, entonces aplicando unas cuantas veces la transformación (2), se puede sustituir la i-ésima fila por ceros. La transformación (2) no altera el número d_M, por lo cual, en virtud de la propiedad (1), se tiene $d_M = 0$.

(4) *Si la i-ésima fila de la matriz* M *es la suma de dos vectores* β *y* γ, *y si las matrices* M' *y* M'' *se obtienen de la matriz* M *sustituyendo su i-ésima fila por los vectores* β *y* γ, *respectivamente, entonces*

$$d_M = d_{M'} + d_{M''}.$$

En efecto, sea S el sistema de todas las filas de la matriz M, excluyendo la i-ésima. Si existe en S una dependencia lineal, las filas de cada una de las matrices M, M' y M'' son linealmente dependientes, de donde, por la propiedad (3), $d_M = d_{M'} = d_{M''} = 0$.

De aquí se deduce la validez de la propiedad en cuestión. Si el sistema S constituido de $n-1$ vectores es linealmente independiente, entonces, como muestran los resultados del § 9, éste se puede completar con un vector α de modo que resulte un sistema linealmente independiente maximal de vectores del espacio de n dimensiones. Los vectores β y γ se pueden expresar linealmente mediante este sistema. Supongamos que el vector α figura en esta expresión con los coeficientes k y l, respectivamente; por consiguiente, en la expresión del vector $\beta+\gamma$, o sea, en la i-ésima fila de la matriz M, el vector α figurará con el coeficiente $k+l$. Ahora se pueden transformar las matrices M, M' y M'', restando de sus i-ésimas filas ciertas combinaciones lineales de las otras filas, de modo que en las i-ésimas filas resulten los vectores $(k+l)\alpha$, $k\alpha$, y $l\alpha$, respectivamente. En consecuencia, designando con M^0 la matriz que resulta de la matriz M sustituyendo su i-ésima fila por el vector α, y, teniendo en cuenta las propiedades (2) y I, llegamos a las igualdades:

$$d_M=(k+l)\,d_{M^0},\quad d_{M'}=kd_{M^0},\quad d_{M''}=ld_{M^0}.$$

Con esto, la propiedad (4) queda demostrada.

(5) *Si la matriz* $\overline{M}$ *se ha obtenido de la matriz* M *trasponiendo dos filas, entonces* $d_{\overline{M}}=-d_M$.

En efecto, supongamos que en la matriz M hay que trasponer las filas que tienen los números de orden i y j. Esto se puede conseguir mediante una cadena de transformaciones siguientes: primero agregamos a la i-ésima fila de la matriz M su j-ésima y obtenemos la matriz M', resultando, por la condición II, $d_{M'}=d_M$. Después restamos de la j-ésima fila de la matriz M' su i-ésima fila y obtenemos la matriz M'' para la que, en virtud de la propiedad (2), se tiene $d_{M''}=d_{M'}$; la j-ésima fila de la matriz M'' se diferenciará de la i-ésima fila de la matriz M en el signo. Agreguemos ahora a la i-ésima fila de la matriz M'' su j-ésima fila. Para la matriz M''' que se obtiene con esta transformación, por la condición II, se tiene $d_{M'''}=$ $=d_{M''}$, además, la i-ésima fila de esta matriz coincide con la j-ésima fila de la matriz M. Finalmente, multiplicando la j-ésima fila de la matriz M''', por el número -1, obtendremos la matriz buscada $\overline{M}$. Por consiguiente, en virtud de la condición I, se tiene

$$d_{\overline{M}}=-d_{M'''}=-d_M.$$

(6) *Si la matriz* M' *se ha obtenido de la matriz* M *trasponiendo las filas, y la* i*-ésima fila de la matriz* M', $i=1, 2, \ldots, n$, *es la* α_i*-ésima fila de la matriz* M, *entonces*,

$$d_{M'}=\pm d_M;$$

donde el signo más corresponde al caso en que la sustitución

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \end{pmatrix}$$

es par, y el signo menos, al caso en que ésta es impar.

En efecto, la matriz M' se puede obtener de la matriz M realizando cierto número de trasposiciones de dos filas, pudiéndose, por consiguiente, aplicar la propiedad (5). Como se sabe por el § 3, la paridad del número de estas trasposiciones determina la paridad de la sustitución indicada anteriormente.

Veamos ahora las matrices $M = (a_{ij})$, $N = (b_{ij})$ y su producto $Q = MN$, en el sentido del § 13. Hallemos el número d_Q. Sabemos que cualquier i-ésima fila de la matriz Q representa una suma de todas las filas de la matriz N, tomadas con los coeficientes a_{i1}, a_{i2}, ..., a_{in}, respectivamente (véase, por ejemplo, el § 14). Sustituyamos todas las filas de la matriz Q por sus expresiones lineales indicadas mediante las filas de la matriz N y apliquemos unas cuantas veces la propiedad (4). Obtendremos que el número d_Q será igual a la suma de los números d_T para todas las matrices posibles T de la forma siguiente: la i-ésima fila de la matriz T, $i = 1, 2, \dots, n$, es igual a la α_i-ésima fila de la matriz N, multiplicada por el número $a_{i\alpha_i}$. Además, en virtud de la propiedad (3), se pueden excluir todas las matrices T para las que existen unos índices i y j, $i \neq j$ tales que $\alpha_i = \alpha_j$; en otras palabras, quedan solamente las matrices T para las que los índices $\alpha_1, \alpha_2, \dots, \alpha_n$ forman una permutación de los números $1, 2, \dots, n$. En virtud de las propiedades I y (6), el número d_T para esta matriz tiene la forma

$$d_T = \pm a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n} d_N,$$

donde el signo se determina por la paridad de la sustitución de los índices. De aquí llegamos a la expresión para el número d_Q; después de sacar el factor común d_N de todos los sumandos de la forma d_T, entre paréntesis queda, evidentemente, el determinante $|M|$ de la matriz M en el sentido de la definición constructiva dada en el § 4, es decir,

$$d_Q = |M| \cdot d_N. \qquad (*)$$

Si ahora tomamos por matriz N la matriz unidad E, se tendrá, $Q = M$, de donde, por la propiedad III, $d_N = d_E = 1$, o sea *para cualquier matriz M se cumple la igualdad*

$$d_M = |M|,$$

que es lo que se quería demostrar. **Simultáneamente, sin haber utilizado el teorema de Laplace, queda demostrado de nuevo el teorema del producto de los determinantes:** para esto es suficiente sustituir los números d_Q y d_N en la igualdad (*) por los determinantes de las matrices correspondientes.

Terminemos estas consideraciones axiomáticas con la demostración de la **independencia** de las condiciones I—III, o sea, con la demostración de que ninguna de estas condiciones es consecuencia de las otras dos.

Para la demostración de la independencia de la condición III, hagamos $d_M = 0$ para cualquier matriz M de n-ésimo orden. Es evidente que las condiciones I y II se cumplen, mientras que la condición III no.

Para la demostración de la independencia de la condición II, supongamos que para cualquier matriz M el número d_M es igual al producto de los elementos situados en la diagonal principal de esta matriz. *Las condiciones I y III se cumplen, mientras que la condición* II ya no tiene lugar. Finalmente, para la demostración de la independencia de la condición I, hagamos $d_M = 1$ para cualquier matriz M. En este caso, las condiciones II y III se cumplirán, mientras que la condición I no.

CAPITULO IV

NUMEROS COMPLEJOS

§ 17. El sistema de los números complejos

En el curso del álgebra elemental varias veces se efectúa un enriquecimiento de las reservas de los números. El alumno que comienza el estudio del álgebra ya conoce por la aritmética los números enteros y quebrados positivos. En esencia, el álgebra comienza con la introducción de los números negativos, o sea, con la formación del primero de los sistemas numéricos fundamentales: del sistema de los *números enteros* que consta de todos los números enteros, positivos y negativos, incluyendo el cero, y del sistema más amplio de los *números racionales*, que consta de todos los números enteros y quebrados, tanto positivos como negativos.

Posteriormente, se efectúa una ampliación del conjunto de los números introduciendo los números irracionales. El sistema, compuesto de todos los números racionales e irracionales, se llama sistema de números *reales*. Ordinariamente, el curso universitario de análisis matemático contiene una construcción rigurosa del sistema de números reales; sin embargo, para nuestra exposición bastan los conocimientos de los números reales que tiene el lector que comienza a estudiar el álgebra superior.

Finalmente, al terminar el curso del álgebra elemental, se amplía el sistema de números reales obteniendo el sistema de *números complejos*. Naturalmente, este sistema de números sigue siendo menos habitual para el lector que el sistema de números reales, a pesar de que posee unas propiedades muy útiles. En el presente capítulo se expondrá de nuevo la teoría de los números complejos con la extensión y plenitud debida.

La introducción de los números complejos es debida al problema siguiente. Es sabido que los números reales no son suficientes para resolver cualquier ecuación cuadrática con coeficientes reales. La ecuación cuadrática más simple, que carece de raíces en el conjunto de los números reales, es

$$x^2 + 1 = 0; \tag{1}$$

ahora, nos va a interesar solamente esta ecuación. El problema que se nos plantea es: *hay que ampliar el sistema de números reales hasta obtener un sistema tal de números, en el que la ecuación (1) tenga ya raíz.*

Los puntos del plano se tomarán como material de construcción de este nuevo sistema de números. Recordemos que la representación de los números reales por puntos de una línea (basada en que se obtiene una correspondencia biunívoca entre el conjunto de todos los puntos de la recta y el conjunto de todos los números reales, al poner en correspondencia a cada punto de la recta su abscisa; se suponen dados el origen de coordenadas y la unidad de medida) se utiliza sistemáticamente en todas las ramas de las matemáticas y es tan habitual que, ordinariamente, no hacemos distinción alguna entre un número real y el punto que le corresponde.

Por lo tanto, *queremos definir un sistema de números que se representen por todos los puntos del plano.* Hasta ahora, no hemos tenido que sumar o multiplicar los puntos del plano, lo que nos da derecho de elegir la definición de las operaciones con los puntos, preocupándose solamente de que el nuevo sistema de números posea las propiedades que son el motivo de su creación. Al principio, estas definiciones nos parecerán artificiales, sobre todo la del producto. Sin embargo, en el capítulo X se demostrará que **ningunas otras definiciones de las operaciones**, incluso las que a primera vista parecen más naturales, **nos conducirían al objetivo**, que consiste en la construcción de una ampliación del sistema de números reales, para que la ecuación (1) tenga raíz. Allí mismo se demostrará que, *en esta construcción, la sustitución de los puntos del plano por otro material, no nos conduciría a un sistema de números diferente, por sus propiedades algebraicas, del sistema de números complejos que vamos a construir a continuación.*

Supongamos que en el plano se ha elegido un sistema rectangular de coordenadas. Convengamos en designar los puntos del plano con las letras α, β, γ, ... y en representar con la notación (a, b) el punto α de abscisa a y ordenada b, es decir, que apartándonos un poco de lo convenido en la geometría analítica, escribiremos $\alpha = (a, b)$. Dados los puntos $\alpha = (a, b)$ y $\beta = (c, d)$, **llamaremos** *suma* de estos puntos al punto que tiene la abscisa $a + c$ y la ordenada $b + d$, o sea,

$$(a, b) + (c, d) = (a + c,\ b + d); \tag{2}$$

llamaremos *producto* de los puntos $\alpha = (a, b)$ y $\beta = (c, d)$ al punto de abscisa $ac - bd$ y ordenada $ad + bc$, o sea,

$$(a, b)(c, d) = (ac - bd,\ ad + bc). \tag{3}$$

De este modo, hemos definido dos operaciones algebraicas en el conjunto de todos los puntos del plano. Demostremos que *estas operaciones poseen todas las propiedades principales que ellas mismas tienen en el sistema de números reales o en el sistema de números racionales; ambas son conmutativas y asociativas, están ligadas por la ley distributiva y para ellas existen las operaciones inversas: la resta y la división (excluyendo la división por cero).*

Las leyes conmutativa y asociativa de la suma son evidentes (o más exatamente, se deducen de las propiedades correspondientes de la suma de los números reales), puesto que al sumar los puntos en el plano, se suman por separado sus abscisas y sus ordenadas. La conmutatividad del producto se basa en que en la definición del producto los puntos α y β gozan de simetría. Las siguientes igualdades:

$$[(a, b)(c, d)](e, f) = (ac - bd,\ ad + bc)(e, f) =$$
$$= (ace - bde - adf - bcf,\ acf - bdf + ade + bce),$$
$$(a, b)[(c, d)(e, f)] = (a, b)(ce - df,\ cf + de) =$$
$$= (ace - adf - bcf - bde,\ acf + ade + bce - bdf),$$

demuestran que para el producto se cumple la ley asociativa. La ley distributiva se deduce de las igualdades:

$$[(a, b) + (c, d)](e, f) = (a + c,\ b + d)(e, f) =$$
$$= (ae + ce - bf - df,\ af + cf + be + de),$$
$$(a, b)(e, f) + (c, d)(e, f) = (ae - bf,\ af + be) + (ce - df,\ cf + de) =$$
$$= (ae - bf + ce - df,\ af + be + cf + de).$$

Veamos la cuestión de las operaciones inversas. Si se han dado los puntos $\alpha = (a, b)$ y $\beta = (c, d)$, su diferencia será un punto (x, y) tal que

$$(c, d) + (x, y) = (a, b).$$

De aquí, en virtud de (2), se deduce que

$$c + x = a,\ d + y = b.$$

Por lo tanto, la *diferencia* de los puntos $\alpha = (a, b)$ y $\beta = (c, d)$ es el punto

$$\alpha - \beta = (a - c,\ b - d) \tag{4}$$

y esta diferencia queda definida unívocamente. En particular, el origen de coordenadas (0,0) sirve de *cero*, y el punto *opuesto* al punto $\alpha = (a, b)$ será el punto

$$-\alpha = (-a,\ -b). \tag{5}$$

Supongamos ahora que se dan los puntos $\alpha = (a, b)$ y $\beta = (c, d)$, y que el punto β es diferente de cero, o sea, que al menos una de las coordenadas c, d no es igual a cero y, por consiguiente, $c^2 + d^2 \neq 0$. El cociente de la división de α por β tiene que ser un punto (x, y) tal que $(c, d)(x, y) = (a, b)$. De aquí, en virtud de (3), se tiene que,

$$cx - dy = a,$$
$$dx + cy = b.$$

Resolviendo este sistema de ecuaciones, obtenemos:

$$x = \frac{ac + bd}{c^2 + d^2}, \quad y = \frac{bc - ad}{c^2 + d^2}.$$

Por lo tanto, para $\beta \neq 0$, el *cociente* $\frac{\alpha}{\beta}$ existe y se determina unívocamente:

$$\frac{\alpha}{\beta} = \left(\frac{ac + bd}{c^2 + d^2}, \ \frac{bc - ad}{c^2 + d^2} \right). \tag{6}$$

Poniendo aquí $\beta = \alpha$, obtenemos que en nuestra multiplicación de los puntos la *unidad* es el punto $(1, 0)$, situado en el eje de abscisas a la distancia 1 del origen de coordenadas a la derecha. Poniendo luego en (6) $\alpha = 1 = (1, 0)$, obtenemos que, para $\beta \neq 0$, el punto *opuesto* a β es:

$$\beta^{-1} = \left(\frac{c}{c^2 + d^2}, \ \frac{-d}{c^2 + d^2} \right). \tag{7}$$

Por lo tanto, hemos construido un sistema de números representados por puntos del plano, donde las operaciones con ellos quedan definidas por las fórmulas (2) *y* (3); *éste se denomina sistema de números complejos.*

Demostremos que *este sistema representa una ampliación del sistema de números reales.* Con este fin, veamos los puntos situados en el eje de abscisas, o sea, los puntos de la forma $(a, 0)$; poniendo en correspondencia al *punto* $(a, 0)$, *el número real* a, *obtenemos* evidentemente una correspondencia biunívoca entre el conjunto considerado de puntos y el conjunto de todos los números reales. (Véase la nota del T. en la pág. 25) La aplicación de las fórmulas (2) y (3) a estos puntos proporciona las igualdades

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

o sea, los puntos $(a, 0)$ se suman y se multiplican entre sí, igual que los números reales correspondientes. Por lo tanto, el *conjunto de puntos situados en el eje de abscisas, considerado como una parte del sistema de números complejos, no se diferencia en nada por sus* *propiedades algebraicas del sistema de números reales, representado ordinariamente por puntos de una recta.* Esto nos permite no hacer a continuación ninguna distinción entre el punto $(a, 0)$ y el número real a, o sea, que pondremos $(a, 0) = a$. En particular, el cero $(0, 0)$ y la unidad $(1, 0)$ del sistema de números complejos resultan ser los números reales ordinarios 0 y 1.

Tenemos que mostrar ahora que *entre los números complejos está contenida una raíz de la ecuación* (1), es decir, un número cuyo cuadrado sea igual al número real -1. Este será, por ejemplo, el punto $(0, 1)$, o sea, el punto situado en el eje de ordenadas a la distancia 1 del origen de coordenadas, hacia arriba. En efecto, aplicando (3), obtenemos:

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

Designemos este punto con la letra i, de modo que $i^2 = -1$.

Finalmente, demostremos que *para los números complejos introducidos se puede obtener su expresión ordinaria.* Para esto, hallemos primero el producto del número real b por el punto i.

$$bi = (b, 0)\cdot(0, 1) = (0, b);$$

por consiguiente, éste es el punto que tiene la ordenada b y está situado en el eje de ordenadas; además todos los puntos del eje de ordenadas se representan en forma de productos de éstos. Si ahora (a, b) es un punto arbitrario, en virtud de la igualdad

$$(a, b) = (a, 0) + (0, b),$$

se tiene:

$$(a, b) = a + bi,$$

o sea, que verdaderamente llegamos a la expresión ordinaria de los números complejos; por supuesto, en la expresión $a + bi$, la suma y el producto se deben entender en el sentido de las operaciones definidas en el sistema de números complejos construido.

Una vez introducidos los números complejos, el lector comprobará fácilmente que *todo el contenido de los capítulos precedentes del libro* (la teoría de los determinantes, la teoría de los sistemas de ecuaciones lineales, la teoría de la dependencia lineal de los vectores y la teoría de las operaciones con las matrices) *se generaliza sin restricciones al caso en que se permite el uso de cualesquiera números complejos, y no sólo de los números reales.*

Por último, obsérvese que la construcción expuesta del sistema de números complejos nos lleva a la siguiente pregunta: ¿Se puede definir la suma y el producto de los puntos del espacio de tres dimensiones, de modo que el conjunto de éstos forme un sistema de números que contenga al sistema de números complejos o, al menos, al sistema de números reales? Esta cuestión sale fuera de los márgenes de nuestro curso y solamente señalaremos que la respuesta es negativa.

Por otra parte, observando que la suma de los números complejos definida anteriormente coincide en su esencia con la suma de vectores en el plano que parten del origen de coordenadas (véase el siguiente párrafo), resulta natural la siguiente pregunta: ¿Es posible definir para ciertos valores de n el producto de vectores del espacio vectorial real de n dimensiones, de modo que éste sea, con respecto a esta multiplicación y a la adición ordinaria de los vectores, un sistema numérico que contenga al sistema de números reales? Se puede demostrar que esto no se puede hacer si se quiere que *se cumplan todas las propiedades de las operaciones que tienen* lugar en los sistemas de números racionales, reales y complejos. *En el espacio de cuatro dimensiones* esta construcción *es posible* si se prescinde de la conmutatividad de la multiplicación; el sistema de números obtenido se denomina *sistema de cuaterniones*. También es posible una construcción análoga en el espacio de ocho dimensiones, resultando el llamado sistema de *números de Cayley*. Desde luego, en este caso no hay que prescindir solamente de la conmutatividad del producto, sino también de su asociatividad, sustituyendo esta última por otra menos rigurosa.

§ 18. Estudio posterior de los números complejos

De acuerdo a la tradición histórica, al número complejo i lo llamaremos *unidad imaginaria*, y a los números de la forma bi, *números imaginarios puros*, a pesar de que no dudamos de la existencia de ellos, pudiendo señalar los puntos del plano (que están en el eje de ordenadas) que los representan. En la expresión del número complejo α en la forma $\alpha = a + bi$, el número a se denomina *parte real* del número α, y el número bi, *parte imaginaria*. El plano, cuyos puntos se han identificado con los números complejos según el método expuesto en el § 17, se llamará *plano complejo*. El eje de abscisas de este plano se llama *eje real*, puesto que sus puntos representan a los números reales; respectivamente, el eje de ordenadas del plano complejo se *llama eje imaginario*.

La suma, resta, multiplicación y división de los números complejos expresados en la forma $a + bi$, como se deduce de las fórmulas (2), (4), (3), y (6) del párrafo anterior, se efectúan del modo siguiente:

$$(a+bi)+(c+di)=(a+c)+(b+d)\,i;$$
$$(a+bi)+(c-di)=(a-c)+(b-d)\,i;$$
$$(a+bi)\,(c+di)=(ac-bd)+(ad+bc)\,i;$$
$$\frac{a+bi}{c+di}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}\,i.$$

Se puede decir que *al sumar los números complejos, se suman por separado sus partes reales y sus partes imaginarias*; para la resta se cumple una regla análoga. Las expresiones verbales de las fórmulas para multiplicar y dividir serían muy complicadas y las omitimos. No hay necesidad de recordar la última de estas fórmulas:

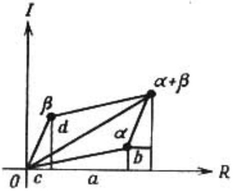

Fig. 2.

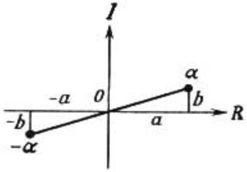

Fig. 3.

solamente hay que tener en cuenta que ésta se puede deducir multiplicando el numerador y denominador del quebrado dado por un número, que se diferencia del denominador solamente por el signo de la parte imaginaria.

En efecto,

$$\frac{a+bi}{c+di}=\frac{(a+bi)(c-di)}{(c+di)(c-di)}=\frac{(ac+bd)+(bc-ad)i}{c^2+d^2}=\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.$$

Ejemplos.

1) $(2+5i)+(1-7i)=(2+1)+(5-7)i=3-2i;$

2) $(3-9i)-(7+i)=(3-7)+(-9-1)i=-4-10i;$

3) $(1+2i)(3-i)=[1\cdot3-2\cdot(-1)]+[1\cdot(-1)+2\cdot3]i=5+5i;$

4) $\frac{23+i}{3+i}=\frac{(23+i)(3-i)}{(3+i)(3-i)}=\frac{70-20i}{10}=7-2i.$

La representación de los números complejos por puntos del plano conduce al deseo natural de obtener una interpretación geométrica de las operaciones definidas para los números complejos. Esta es fácil de lograr para la suma. Sean dados los números $\alpha = a+bi$ y $\beta = c+di$. Unamos con segmentos el origen de coordenadas con los puntos (a, b) y (c, d) correspondientes a dichos números, y sobre estos segmentos, como lados, trazemos un paralelogramo (fig. 2). Es evidente que el cuarto vértice de este paralelogramo será el punto $(a+c, b+d)$. Por lo tanto, la *suma de números complejos se efectúa geométricamente por la regla del paralelogramo, o sea por la regla de la suma de vectores que parten del origen de coordenadas.*

El número opuesto al número $\alpha = a + bi$ es el punto del plano complejo que es simétrico al punto α con respecto del origen de coordenadas (fig. 3). De aquí se puede obtener sin dificultad alguna la interpretación geométrica de la resta.

La interpretación geométrica de la multiplicación y división de los números complejos quedará clara solamente después de que introduzcamos una nueva *expresión para los números complejos*. Para la expresión del número α en la forma $\alpha = a + bi$ utilizamos las coordenadas cartesianas del punto correspondiente a este número. Sin embargo, la posición del punto en el plano queda también determinada, si se conocen sus *coordenadas polares: la distancia* r del origen de coordenadas al punto y el ángulo φ que forma la dirección positiva del eje de abscisas con la dirección que va desde el origen de coordenadas hacia este punto (fig. 4).

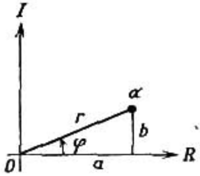

Fig. 4.

El número r es real y no negativo, siendo además igual a cero solamente para el punto 0. Para un número α situado en el eje real, o sea, para un número real, el número r es el valor absoluto de α; por esto, a veces, para *cualquier número complejo α, a r también se le llaman valor absoluto* o *módulo* del número α, representándose por la notación $|\alpha|$.

El ángulo φ se llamará *argumento* del número α y se designará con la notación: $\arg \alpha$*. El ángulo φ puede tomar cualesquiera valores reales, tanto positivos como negativos, teniendo que medir*se los ángulos positivos en dirección contraria a la del movimien*to de las agujas del reloj; sin embargo, si los ángulos se diferencian entre sí en 2π o en un número múltiplo de 2π, sus puntos correspondientes del plano coinciden.

De este modo, el argumento de un número complejo α tiene *infinitos valores, que se diferencian entre sí en números enteros* múltiplos de 2π; por consiguiente, de la igualdad de dos números complejos, representados por sus módulos y sus argumentos, solamente se puede hacer la conclusión de que sus argumentos se diferencian en un número entero múltiplo de 2π, mientras que sus módu*los son iguales. Solamente para el número 0 el argumento es indefi*nido; sin embargo, este número queda completamente determinado por la igualdad: $|0| = 0$.

El argumento del número complejo es una generalización natural del signo del número real. En efecto, el argumento de un número *real positivo es igual a cero, el argumento de un número real nega-*

---

* No recurrimos a las denominaciones corrientes de las coordenadas polares: radio polar y ángulo polar.

tivo es igual a π; en el eje real, del origen de coordenadas parten solamente dos direcciones, las cuales se pueden distinguir por los símbolos: + y —. En el plano complejo hay infinitas direcciones que parten del punto 0, diferenciándose por el ángulo que forman con la dirección positiva del eje real.

Las coordenadas cartesianas y polares de un punto están ligadas por las relaciones siguientes:

$$a = r\cos\varphi, \quad b = r\operatorname{sen}\varphi, \tag{1}$$

que se cumplen independientemente de la posición del punto en el plano.

De aquí que

$$r = +\sqrt{a^2+b^2}. \tag{2}$$

Apliquemos las fórmulas (1) a un número complejo arbitrario $\alpha = a+bi$:

$$\alpha = a+bi = r\cos\varphi + (r\operatorname{sen}\varphi)\,i,$$

o sea,

$$\alpha = r(\cos\varphi + i\operatorname{sen}\varphi). \tag{3}$$

Recíprocamente, supongamos que el número $\alpha = a+bi$ se expresa en la forma $\alpha = r_0(\cos\varphi_0 + i\operatorname{sen}\varphi_0)$, donde r_0 y φ_0 son unos números reales, siendo $r_0 \geqslant 0$. Entonces, $r_0\cos\varphi_0 = a$, $r_0\operatorname{sen}\varphi_0 = b$, de donde $r_0 = +\sqrt{a^2+b^2}$, y, en virtud de (2), $r_0 = |\alpha|$. De aquí, aplicando (1), obtenemos: $\cos\varphi_0 = \cos\varphi$, $\operatorname{sen}\varphi_0 = \operatorname{sen}\varphi$, o sea, $\varphi_0 = \arg\alpha$. Por lo tanto, *todo número complejo α se expresa unívocamente en la forma* (3), *donde $r = |\alpha|$,* $\varphi = \arg\alpha$ (por supuesto, el argumento φ está definido salvo un sumando, múltiplo de 2π). Esta expresión del número α se llama *forma trigonométrica* y se empleará frecuentemente a continuación.

Los números

$$\alpha = 3\left(\cos\frac{\pi}{4} + i\operatorname{sen}\frac{\pi}{4}\right), \quad \beta = \cos\frac{19}{3}\pi + i\operatorname{sen}\frac{19}{3}\pi$$

y

$$\gamma = \sqrt{3}\left[\cos\left(-\frac{\pi}{7}\right) + i\operatorname{sen}\left(-\frac{\pi}{7}\right)\right]$$

están dados en forma trigonométrica; aquí, $|\alpha| = 3$, $|\beta| = 1$, $|\gamma| = \sqrt{3}$; $\arg\alpha = \frac{\pi}{4}$, $\arg\beta = \frac{19}{3}\pi$, $\arg\gamma = -\frac{\pi}{7}$ $\left(\text{o bien, } \arg\beta = \frac{\pi}{3},\ \arg\gamma = \frac{13}{7}\pi\right)$.

Por otra parte, los números complejos

$$\alpha' = (-2)\left(\cos\frac{\pi}{5} + i\operatorname{sen}\frac{\pi}{5}\right), \quad \beta' = 3\left(\cos\frac{2}{3}\pi - i\operatorname{sen}\frac{2}{3}\pi\right),$$

$$\gamma' = 2\left(\cos\frac{\pi}{3} + i\operatorname{sen}\frac{3}{4}\pi\right), \delta' = \operatorname{sen}\frac{3}{4}\pi + i\cos\frac{3}{4}\pi$$

ya no están dados en forma trigonométrica, a pesar de que estas expresiones se parezcan a la expresión (3). Estos números se expresan en forma trigonométrica del modo siguiente:

$$\alpha' = 2\left(\cos\frac{6}{5}\pi + i\ \mathrm{sen}\ \frac{6}{5}\pi\right), \quad \beta' = 3\left(\cos\frac{4}{3}\pi + i\ \mathrm{sen}\ \frac{4}{3}\pi\right),$$

$$\delta' = \cos\frac{7}{4}\pi + i\ \mathrm{sen}\ \frac{7}{4}\pi.$$

La determinación de la forma trigonométrica del número γ' es engorrosa, como casi siempre ocurre al pasar de la expresión ordinaria del número complejo a la trigonométrica y viceversa: salvo unos pocos casos, dados los valores numéricos del seno y del coseno, resulta imposible hallar el valor exacto del ángulo, dado el ángulo, resulta imposible hallar los valores *exactos* de su seno y coseno.

Supongamos que los números complejos α y β se dan en su forma trigonométrica: $\alpha = r\ (\cos\varphi + i\ \mathrm{sen}\ \varphi)$, $\beta = r'\ (\cos\varphi' + i\ \mathrm{sen}\ \varphi')$. Multipliquemos estos números:

$$\alpha\beta = [r(\cos\varphi + i\,\mathrm{sen}\,\varphi)]\cdot[r'(\cos\varphi' + i\,\mathrm{sen}\,\varphi')] =$$

$$= rr'(\cos\varphi\cos\varphi' + i\cos\varphi\,\mathrm{sen}\,\varphi' + i\,\mathrm{sen}\,\varphi\cos\varphi' - \mathrm{sen}\,\varphi\,\mathrm{sen}\,\varphi'),$$

o sea,

$$\alpha\beta = rr'[\cos(\varphi+\varphi') + i\,\mathrm{sen}\,(\varphi+\varphi')]. \tag{4}$$

Hemos obtenido la expresión del producto en la forma trigonométrica, de donde $|\alpha\beta| = rr'$, o sea,

$$|\alpha\beta| = |\alpha|\,|\beta|, \tag{5}$$

es decir, el *módulo del producto de números complejos es igual al producto de los módulos de los factores*; por otra parte, $\arg(\alpha\beta) = \varphi + \varphi'$, o sea,

$$\arg(\alpha\beta) = \arg\alpha + \arg\beta, \tag{6}$$

es decir, el *argumento del producto de números complejos es igual a la suma de los argumentos de los factores**. Claro que estas reglas *se generalizan para cualquier número finito de factores*. *En el caso* de números reales, la fórmula (5) proporciona la conocida propiedad de los valores absolutos de estos números, mientras que la fórmula (6), como fácilmente se comprueba, se convierte en la regla de los signos de la multiplicación de los números reales.

La división también goza de propiedades análogas. En efecto, sea $\alpha = r\ (\cos\varphi + i\ \mathrm{sen}\ \varphi)$, $\beta = r'\ (\cos\varphi' + i\ \mathrm{sen}\ \varphi')$, siendo $\beta \neq 0$, o sea, $r' \neq 0$. Entonces,

$$\frac{\alpha}{\beta} = \frac{r(\cos\varphi + i\,\mathrm{sen}\,\varphi)}{r'(\cos\varphi' + i\,\mathrm{sen}\,\varphi')} = \frac{r(\cos\varphi + i\,\mathrm{sen}\,\varphi)(\cos\varphi' - i\,\mathrm{sen}\,\varphi')}{r'(\cos^2\varphi' + \mathrm{sen}^2\varphi')} =$$

$$= \frac{r}{r'}(\cos\varphi\cos\varphi' + i\,\mathrm{sen}\,\varphi\cos\varphi' - i\cos\varphi\,\mathrm{sen}\,\varphi' + \mathrm{sen}\,\varphi\,\mathrm{sen}\,\varphi'),$$

* Subrayemos que aquí la igualdad se entiende salvo un sumando, múltiplo de 2π.

o sea,

$$\frac{\alpha}{\beta}=\frac{r}{r'}\left[\cos(\varphi-\varphi')+i\operatorname{sen}(\varphi-\varphi')\right]. \tag{7}$$

De aquí se deduce que $\left|\frac{\alpha}{\beta}\right|=\frac{r}{r'}$, o bien,

$$\left|\frac{\alpha}{\beta}\right|=\frac{|\alpha|}{|\beta|}, \tag{8}$$

es decir, el *módulo del cociente de dos números complejos es igual al módulo del dividendo dividido por el módulo del divisor*; por otra parte, $\arg\left(\frac{\alpha}{\beta}\right)=\varphi-\varphi'$, o sea,

$$\arg\left(\frac{\alpha}{\beta}\right)=\arg\alpha-\arg\beta \tag{9}$$

es decir, el *argumento del cociente de dos números complejos se obtiene restando el argumento del divisor del argumento del dividendo.*

El significado geométrico del producto y del cociente se aclara ahora sin dificultad. En efecto, en virtud de las fórmulas (5)

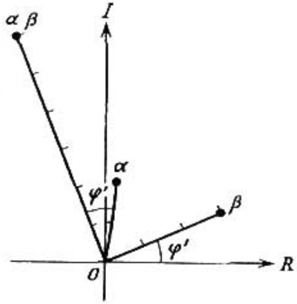

Fig. 5.

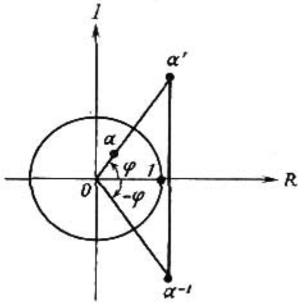

Fig. 6.

y (6), para obtener el punto que representa el producto del número α por el número $\beta=r'(\cos\varphi'+i\operatorname{sen}\varphi')$, hay que hacer girar al vector que va de 0 a α (fig. 5) un ángulo $\varphi'=\arg\beta$ en dirección contraria a la del movimiento de las agujas del reloj, y después hay que alargar este vector $r'=|\beta|$ veces (si $0\leqslant r'<1$, esto no será un alargamiento, sino una contracción). Por otra parte, de (7) se deduce que para $\alpha=r(\cos\varphi+i\operatorname{sen}\varphi)\neq 0$, se tiene,

$$\alpha^{-1}=r^{-1}\left[\cos(-\varphi)+i\operatorname{sen}(-\varphi)\right], \tag{10}$$

o sea, $|\alpha|^{-1}=|\alpha^{-1}|$, $\arg(\alpha^{-1})=-\arg\alpha$. Por lo tanto, para obtener el punto α^{-1} hay que pasar del punto α al punto α', situado en la misma semirrecta que parte del cero y que pasa por el punto α, a la distancia r^{-1} del cero (fig 6)*, y después hay que pasar al punto simétrico a α' con respecto al eje real.

La suma y la diferencia de números complejos, dados en forma trigonométrica, no se pueden expresar por fórmulas semejantes a las fórmulas (4) y (7). Sin embargo, para el módulo de la suma se *cumplen las importantes desigualdades:*

$$|\alpha|-|\beta|\leqslant|\alpha+\beta|\leqslant|\alpha|+|\beta|. \tag{11}$$

es decir, el *módulo de la suma de dos números complejos es menor o igual a la suma de los módulos de los sumandos, pero es mayor o igual a la diferencia de estos módulos.* Las desigualdades (11) se deducen del conocido teorema de geometría elemental sobre los lados del triángulo, puesto que como se sabe, $|\alpha+\beta|$ es igual a la diagonal del paralelogramo de lados $|\alpha|$ y $|\beta|$. Si los puntos α, β y 0 están situados en una recta, se necesita un estudio especial; esto lo dejamos a cuenta del lector. Solamente en este caso se cumple el signo de igualdad en las fórmulas (11).

Como $\alpha-\beta=\alpha+(-\beta)$ y

$$|-\beta|=|\beta| \tag{12}$$

(esta igualdad es consecuencia de la interpretación geométrica del número $-\beta$), de (11) se deducen también las desigualdades

$$|\alpha|-|\beta|\leqslant|\alpha-\beta|\leqslant|\alpha|+|\beta|,^{**} \tag{13}$$

es decir, para el módulo de la diferencia se cumplen también las mismas desigualdades que para el módulo de la uma.

Las desigualdades (11) se podrían obtener también del modo siguiente: Sea $\alpha=r(\cos\varphi+i\,\mathrm{sen}\,\varphi)$, $\beta=r'(\cos\varphi'+i\,\mathrm{sen}\,\varphi')$; supongamos que la forma trigonométrica del número $\alpha+\beta$ es: $\alpha+\beta=R(\cos\psi+i\,\mathrm{sen}\,\psi)$. Sumando por separado las partes reales y las partes imaginarias, obtenemos:

$$r\cos\varphi+r'\cos\varphi'=R\cos\psi,$$

$$r\,\mathrm{sen}\,\varphi+r'\,\mathrm{sen}\,\varphi'=R\,\mathrm{sen}\,\psi;$$

multiplicando ambos miembros de la primera igualdad por $\cos\psi$, ambos miembros de la segunda por $\mathrm{sen}\,\psi$ y sumando, obtenemos: $r(\cos\varphi\cos\psi+$

* La igualdad $|\alpha'|=|\alpha|$ se cumple cuando, y sólo cuando $|\alpha|=1$, o sea, si el punto α está situado en la circunferencia del *círculo unidad.* Si α está situado dentro del círculo unidad, α' estará situado fuera de él, y viceversa, obteniendo de este modo una correspondencia biunívoca entre todos los puntos del plano complejo, situados fuera del círculo unidad, y todos los puntos, situados dentro de este círculo y *diferentes de cero.*

** Por consiguiente, se cumple también la desigualdad

$$\big||\alpha|-|\beta|\big|\leqslant|\alpha-\beta|,$$

que se aplicará en el § 23. (*Nota de T.*)

$+\operatorname{sen}\varphi\operatorname{sen}\psi)+r'(\cos\varphi'\cos\psi+\operatorname{sen}\varphi'\operatorname{sen}\psi)=R(\cos^2\psi+\operatorname{sen}^2\psi),$
o sea,

$$r\cos(\varphi-\psi)+r'\cos(\varphi'-\psi)=R.$$

Como el coseno nunca es mayor que la unidad, de aquí se deduce la desigualdad $r+r'\geqslant R$, o sea, $|\alpha|+|\beta|\geqslant|\alpha+\beta|$. Por otra parte, $\alpha=(\alpha+\beta)-\beta=(\alpha+\beta)+(-\beta)$. De aquí, por lo demostrado y en virtud de (12),

$$|\alpha|\leqslant|\alpha+\beta|+|-\beta|=|\alpha+\beta|+|\beta|,$$

de donde $|\alpha|-|\beta|\leqslant|\alpha+\beta|$.

Es menester observar que los conceptos «mayor» y «menor» no se pueden definir racionalmente para los números complejos, puesto que éstos, a diferencia de los números reales, no se sitúan en una recta, cuyos puntos están ordenados de un modo natural, sino en un plano. Por esto, los números complejos (no nos referimos a sus módulos) no se pueden unir nunca con el signo de desigualdad.

Números conjugados. Sea dado un número complejo $\alpha=a+bi$. El número $a-bi$, que se diferencia de α solamente en el signo de la parte imaginaria, se llama número *conjugado* de α y se designa por $\bar{\alpha}$.

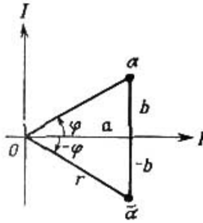

Fig. 7.

Recordemos, que al estudiar la división de los números complejos recurríamos a los números conjugados, a pesar de que no habíamos introducido esta denominación.

Es evidente que el número conjugado de $\bar{\alpha}$ es α, es decir, se puede hablar de pares de números conjugados. Los números reales, y solamente éstos, son conjugados consigo mismos.

Geométricamente, los números conjugados son puntos simétricos entre sí con respecto al eje real (fig. 7). De aquí se deducen las igualdades

$$|\bar{\alpha}|=|\alpha|,\ \arg\bar{\alpha}=-\arg\alpha. \tag{14}$$

La suma y el producto de números complejos conjugados son números reales. En efecto,

$$\left.\begin{aligned}\alpha+\bar{\alpha}&=2a,\\ \alpha\bar{\alpha}&=a^2+b^2=|\alpha|^2.\end{aligned}\right\} \tag{15}$$

La última igualdad muestra que el número $\alpha\bar{\alpha}$ es, incluso, positivo para $\alpha\neq 0$. En el § 24 se verá un teorema que muestra que la propiedad que acabamos de demostrar de los números conjugados es característica para éstos.

La igualdad

$$(a-bi)+(c-di)=(a+c)-(b+d)\,i$$

muestra que el *número conjugado de la suma de dos números es igual a la suma de los números conjugados con los sumandos*:

$$\overline{\alpha+\beta}=\overline{\alpha}+\overline{\beta}. \tag{16}$$

Análogamente, de la igualdad

$$(a-bi)\,(c-di)=(ac-bd)-(ad+bc)\,i$$

resulta que el *número conjugado con producto es igual al producto de los números conjugados con los factores*:

$$\overline{\alpha\beta}=\overline{\alpha}\cdot\overline{\beta}. \tag{17}$$

Una comprobación directa muestra que se verifican también las igualdades

$$\overline{\alpha-\beta}=\overline{\alpha}-\overline{\beta}, \tag{18}$$

$$\overline{\left(\frac{\alpha}{\beta}\right)}=\frac{\overline{\alpha}}{\overline{\beta}}. \tag{19}$$

Demostremos también la siguiente proposición: *si el número α se expresa de cierto modo por los números complejos $\beta_1, \beta_2, \ldots, \beta_n$ mediante la suma, el producto, la resta y la división, entonces, al sustituir en esta expresión todos los números β_k por sus conjugados, se obtiene el número conjugado de α*; en particular, si el número α es real, éste no se altera al sustituir todos los números complejos β_k por sus conjugados.

Esta proposición la demostraremos por inducción sobre n, puesto que para $n=2$ ésta se deduce de las fórmulas (16)—(19).

Supongamos que el número α se expresa por los números $\beta_1, \beta_2, \ldots, \beta_n$, que no son necesariamente diferentes. En esta expresión hay un orden determinado de aplicación de las operaciones de sumar, multiplicar, restar y dividir. El último acto consistirá en la aplicación de una de estas operaciones a un número γ_1, expresado mediante los números $\beta_1, \beta_2, \ldots, \beta_k$, donde $1 \leqslant k \leqslant n-1$, y a un número γ_2, expresado mediante los números $\beta_{k+1}, \ldots, \beta_n$. Por la hipótesis de inducción, la sustitución de los números $\beta_1, \beta_2, \ldots \beta_k$ por los conjugados implica el cambio del número γ_1 por $\overline{\gamma_1}$, y la sustitución de los números $\beta_{k+1}, \beta_{k+2}, \ldots, \beta_n$ por los conjugados, el cambio del número γ_2 por $\overline{\gamma_2}$. Pero, según una de las fórmulas (16)—(19), el cambio de γ_1 y γ_2 por $\overline{\gamma_1}$ y $\overline{\gamma_2}$ convierte al número α en el número $\overline{\alpha}$.

§ 19. Extracción de la raíz de los números complejos

Estudiemos el problema de la elevación de los números complejos a una potencia y de la extracción de una raíz. Para elevar el número $\alpha = a + bi$ a una potencia entera y positiva n, es suficiente aplicar la fórmula del binomio de Newton a la expresión $(a + bi)^n$ (esta fórmula subsiste también para los números complejos, puesto que su demostración se basa solamente en la ley distributiva), y después, las igualdades: $i^2 = -1$, $i^3 = -i$, $i^4 = 1$; en general

$$i^{4k} = 1, \quad i^{4k+1} = i, \quad i^{4k+2} = -1, \quad i^{4k+3} = -i.$$

Si el número α está dado en forma trigonométrica, entonces, siendo n entero y positivo, de la fórmula (4) del párrafo anterior resulta la fórmula siguiente, llamada *fórmula de Moivre*:

$$[r(\cos\varphi + i \operatorname{sen}\varphi)]^n = r^n(\cos n\varphi + i \operatorname{sen} n\varphi), \tag{1}$$

o sea, que *al elevar un número complejo a una potencia, se eleva el módulo a esta potencia y se multiplica el argumento por el exponente de la potencia.* La fórmula (1) es válida también para los exponentes enteros negativos. En efecto, en virtud de la igualdad $\alpha^{-n} = (\alpha^{-1})^n$, es suficiente aplicar la fórmula de Moivre al número α^{-1}, cuya forma trigonométrica viene dada por la fórmula (10) del párrafo anterior.

Ejemplos.

1) $i^{37} = i, \; i^{122} = -1$;

2) $(2 + 5i)^3 = 2^3 + 3\cdot 2^2\cdot 5i + 3\cdot 2\cdot 5^2 i^2 + 5^3 i^3 =$
$= 8 + 60i - 150 - 125i = -142 - 65i$;

3) $\left[\sqrt{2}\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right)\right]^4 = (\sqrt{2})^4(\cos\pi + i\operatorname{sen}\pi) = -4$;

4) $\left[3\left(\cos\frac{\pi}{5} + i\operatorname{sen}\frac{\pi}{5}\right)\right]^{-3} =$

$$= 3^{-3}\left[\cos\left(-\frac{3}{5}\pi\right) + i\operatorname{sen}\left(-\frac{3}{5}\pi\right)\right] = \frac{1}{27}\left(\cos\frac{7}{5}\pi + i\operatorname{sen}\frac{7}{5}\pi\right).$$

De la igualdad

$$(\cos\varphi + i\operatorname{sen}\varphi)^n = \cos n\varphi + i\operatorname{sen} n\varphi,$$

que representa un caso particular de la fórmula de Moivre, fácilmente se obtienen las fórmulas para el seno y el coseno de un ángulo múltiplo. En efecto, aplicando la fórmula del binomio de Newton al primer miembro de esta igualdad e igualando por separado las partes reales e imaginarias de ambos miembros, se tiene:

$$\cos n\varphi = \cos^n \varphi - \binom{n}{2} \cos^{n-2}\varphi \cdot \operatorname{sen}^2 \varphi + \binom{n}{4} \cos^{n-4}\varphi \cdot \operatorname{sen}^4 \varphi - \ldots,$$

$$\operatorname{sen} n\varphi = \binom{n}{1} \cos^{n-1} \varphi \cdot \operatorname{sen} \varphi - \binom{n}{3} \cos^{n-3} \varphi \cdot \operatorname{sen}^3 \varphi + $$
$$+ \binom{n}{5} \cos^{n-5}\varphi \cdot \operatorname{sen}^5\varphi - \ldots;$$

aquí $\binom{n}{k}$ es la notación ordinaria del coeficiente binomial

$$\binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{1\cdot 2\cdot 3\ldots k}.$$

Para $n=2$, se tienen las conocidas fórmulas

$$\cos 2\varphi = \cos^2 \varphi - \operatorname{sen}^2 \varphi,$$
$$\operatorname{sen} 2\varphi = 2 \cos \varphi \operatorname{sen} \varphi;$$

para $n=3$,

$$\cos 3\varphi = \cos^3 \varphi - 3 \cos \varphi \operatorname{sen}^2 \varphi,$$
$$\operatorname{sen} 3\varphi = 3 \cos^2 \varphi \operatorname{sen} \varphi - \operatorname{sen}^3 \varphi.$$

La extracción de la raíz de los números complejos es mucho más complicada. Comencemos por la extracción de la raíz cuadrada del número $\alpha = a + bi$. Todavía no sabemos si existe un número complejo cuyo cuadrado sea igual a α. Suponiendo que tal número existe, por ejemplo, $u + vi$ y empleando la notación ordinaria, se puede escribir:

$$\sqrt{a+bi} = u + vi.$$

De la igualdad

$$(u+vi)^2 = a + bi$$

resulta,

$$\left.\begin{aligned} u^2 - v^2 &= a, \\ 2uv &= b. \end{aligned}\right\} \tag{2}$$

Elevando al cuadrado ambos miembros de las igualdades (2) y sumándolas después, se tiene

$$(u^2-v^2)^2 + 4u^2v^2 = (u^2+v^2)^2 = a^2 + b^2,$$

de donde

$$u^2 + v^2 = +\sqrt{a^2+b^2};$$

se toma el signo más, porque los números u y v son reales, debido a lo cual, el primer miembro de la igualdad es positivo. De esta igualdad y de la primera de las igualdades (2), resulta:

$$u^2 = \frac{1}{2}(a + \sqrt{a^2 + b^2}),$$

$$v^2 = \frac{1}{2}(-a + \sqrt{a^2 + b^2}).$$

Extrayendo las raíces cuadradas se obtienen dos valores para u, que se diferencian en el signo, y también dos valores para v. Todos estos valores son reales, puesto que para cualesquiera a y b, las raíces cuadradas se extraen de números positivos. Los valores obtenidos de u y v no se pueden combinar entre sí de modo arbitrario, puesto que, en virtud de la segunda de las igualdades (2), el signo del producto uv tiene que coincidir con el signo de b. Resultan, pues, dos combinaciones posibles de los valores de u y v, o sea, dos números de la forma $u + vi$, que pueden servir de valores de la raíz cuadrada del número α; estos números se diferencian entre sí en el signo. Una prueba elemental, aunque complicada (elevando al cuadrado los números obtenidos, una vez cuando $b > 0$, y otra vez cuando $b < 0$), muestra que los números obtenidos son, verdaderamente, valores de la raíz cuadrada del número α. Por lo tanto, *siempre es posible la extracción de la raíz cuadrada de un número complejo, proporcionando ésta dos valores que se diferencian entre sí en el signo.*

En particular, ahora resulta posible la extracción de la raíz cuadrada de un número real negativo, siendo los valores de esta raíz números imaginarios puros. En efecto, si $a < 0$ y $b = 0$, entonces $\sqrt{a^2 + b^2} = -a$, puesto que esta raíz tiene que ser positiva, por lo cual, $u^2 = \frac{1}{2}(a - a) = 0$, o sea, $u = 0$, así que $\sqrt{a} = \pm vi$.

Ejemplo. Sea $\alpha = 21 - 20i$. Entonces $\sqrt{a^2 + b^2} = \sqrt{441 + 400} = 29$. Por consiguiente, $u^2 = \frac{1}{2}(21 + 29) = 25$, $v^2 = \frac{1}{2}(-21 + 29) = 4$, de donde $u = \pm 5$, $v = \pm 2$. Los signos de u y v tienen que ser diferentes, puesto que b es negativo; en consecuencia,

$$\sqrt{21 - 20i} = \pm(5 - 2i).$$

Las pruebas de extracción de raíces de grado más elevado de los números complejos, dados en la forma $a + bi$, chocan con dificultades insuperables. Así, pues, si quisiéramos hallar con el mismo método la raíz cúbica del número $a + bi$, tendríamos que resolver una ecuación cúbica auxiliar, cosa que por el momento no sabemos hacer y que, como ya veremos en el § 38, requiere a su vez la extracción de la raíz cúbica de números complejos. Por otra parte, la forma trigonométrica se adapta perfectamente para la extracción de las raíces de cualquier grado, con cuya aplicación se resuelve totalmente este problem.

Supongamos que se necesita extraer la raíz n-ésima del número $\alpha = r(\cos\varphi + i \operatorname{sen}\varphi)$. Supongamos también que ésta se puede hallar, *resultando* $\rho(\cos\theta + i \operatorname{sen}\theta)$, *de modo que*,

$$[\rho(\cos\theta + i \operatorname{sen}\theta)]^n = r(\cos\varphi + i \operatorname{sen}\varphi). \tag{3}$$

Según la fórmula de Moivre, $\rho^n = r$, *o sea*, $\rho = \sqrt[n]{r}$, *donde en* el segundo miembro figura el valor positivo, unívocamente determinado, de la raíz n-ésima del número real positivo r. Por otra parte, el argumento del primer miembro de la igualdad (3) es $n\theta$. Sin embargo, no se puede afirmar que $n\theta$ es igual a φ, porque, en realidad, éstos pueden diferir en un sumando que es múltiplo entero del número 2π. En consecuencia, $n\theta = \varphi + 2k\pi$, donde k es entero y

$$\theta = \frac{\varphi + 2k\pi}{n}.$$

Recíprocamente, tomando el número $\sqrt[n]{r}\left(\cos\frac{\varphi+2k\pi}{n} + i \operatorname{sen}\frac{\varphi+2k\pi}{n}\right)$, se tiene que, **para cualquier** k **entero, positivo o negativo,** la n-ésima potencia de este número es igual a α. Por lo tanto

$$\sqrt[n]{r(\cos\varphi + i \operatorname{sen}\varphi)} = \sqrt[n]{r}\left(\cos\frac{\varphi + 2k\pi}{n} + i \operatorname{sen}\frac{\varphi + 2k\pi}{n}\right). \tag{4}$$

Dando a k diversos valores, *no siempre se obtienen diversos* valores de la raíz buscada. En efecto, para

$$k = 0, 1, 2, \ldots, n-1 \tag{5}$$

se obtienen n valores distintos de la raíz, puesto que el aumento de k en una unidad ocasiona un aumento del argumentos en $\frac{2\pi}{n}$. Supongamos ahora que k es arbitrario. Si $k = nq + r$, $0 \leqslant +r \leqslant n-1$, se tiene,

$$\frac{\varphi + 2k\pi}{n} = \frac{\varphi + 2(nq + r)\pi}{n} = \frac{\varphi + 2r\pi}{n} + 2q\pi,$$

es decir, el valor del argumento para nuestro k difiere del valor del argumento para $k = r$ en un número múltiplo de 2π; por consiguiente, se obtiene el mismo valor de la raíz que resulta para el valor de k igual a r, incluido en el sistema (5).

Por lo tanto, *siempre es posible la extracción de la raíz n-ésima de un número complejo* α, *resultando n valores distintos. Todos los valores de la raíz n-ésima están situados en una circunferencia de radio* $\sqrt[n]{|\alpha|}$ *con el centro en el cero, dividiendo a ésta en n partes iguales.*

En particular, la raíz n-ésima de un número real a tiene también n valores distintos; entre éstos puede haber dos reales, uno *o ninguno, dependiendo del signo de a y de la paridad de n.*

Ejemplos.

1) $$\beta=\sqrt[3]{2\left(\cos\frac{3}{4}\pi+i\operatorname{sen}\frac{3}{4}\pi\right)}=$$

$$=\sqrt[3]{2}\left(\cos\frac{\frac{3}{4}\pi+2k\pi}{3}+i\operatorname{sen}\frac{\frac{3}{4}\pi+2k\pi}{3}\right);$$

$$k=0:\ \beta_0=\sqrt[3]{2}\left(\cos\frac{\pi}{4}+i\operatorname{sen}\frac{\pi}{4}\right);$$

$$k=1:\ \beta_1=\sqrt[3]{2}\left(\cos\frac{11}{12}\pi+i\operatorname{sen}\frac{11}{12}\pi\right);$$

$$k=2:\ \beta_2=\sqrt[3]{2}\left(\cos\frac{19}{12}\pi+i\operatorname{sen}\frac{19}{12}\pi\right).$$

2) $$\beta=\sqrt{i}=\sqrt{\cos\frac{\pi}{2}+i\operatorname{sen}\frac{\pi}{2}}=$$

$$=\cos\frac{\frac{\pi}{2}+2k\pi}{2}+i\operatorname{sen}\frac{\frac{\pi}{2}+2k\pi}{2};$$

$$\beta_0=\cos\frac{\pi}{4}+i\operatorname{sen}\frac{\pi}{4}=\frac{\sqrt{2}}{2}+i\frac{\sqrt{2}}{2};$$

$$\beta_1=\cos\frac{5}{4}\pi+i\operatorname{sen}\frac{5}{4}\pi=-\beta_0.$$

3) $$\beta=\sqrt[3]{-8}=\sqrt[3]{8(\cos\pi+i\operatorname{sen}\pi)}=$$

$$=2\left(\cos\frac{\pi+2k\pi}{3}+i\operatorname{sen}\frac{\pi+2k\pi}{3}\right);$$

$$\beta_0=2\left(\cos\frac{\pi}{3}+i\operatorname{sen}\frac{\pi}{3}\right)=1+i\sqrt{3};$$

$$\beta_1=2(\cos\pi+i\operatorname{sen}\pi)=-2;$$

$$\beta_2=2\left(\cos\frac{5\pi}{3}+i\operatorname{sen}\frac{5\pi}{3}\right)=1-i\sqrt{3}.$$

Raíces de la unidad. El caso de la extracción de la raíz n-ésima del número 1 es de particular importancia. Esta raíz tiene n valores y, en virtud de la igualdad $1 = \cos 0 + i \operatorname{sen} 0$ y de la fórmula (4), todos ellos, o, como nos expresaremos, todas las *raíces n-ésimas de la unidad*, vienen dadas por la fórmula

$$\sqrt[n]{1}=\cos\frac{2k\pi}{n}+i\operatorname{sen}\frac{2k\pi}{n};\ k=0,1,\ldots,n-1. \tag{6}$$

Los valores reales de la raíz n-ésima de la unidad se obtienen de la fórmula (6) para los valores $k = 0$ y $\frac{n}{2}$, si n es par, y para $k = 0$, si n es impar. En el plano complejo las raíces n-ésimas de la unidad están situadas en la circunferencia del círculo unidad y la dividen en n arcos iguales; el número 1 es uno de los puntos de división. De esto se deduce que las raíces n-ésimas de la unidad que no son reales están situadas simétricamente con respecto al eje real, es decir, son conjugadas entre sí.

La raíz cuadrada de la unidad tiene dos valores: 1 y -1. La raíz cuártica de la unidad tiene cuatro valores: 1, -1, i y $-i$. Para lo futuro es conveniente recordar los valores de la *raíz cúbica de la unidad*. En virtud de (6), éstos son los números $\cos\frac{2k\pi}{3}+$ $+i\operatorname{sen}\frac{2k\pi}{3}$, donde $k=0, 1, 2$, o sea, además de la unidad, se tienen los números conjugados entre sí:

$$\left.\begin{aligned} \varepsilon_1 &= \cos\frac{2\pi}{3}+i\operatorname{sen}\frac{2\pi}{3}=-\frac{1}{2}+i\frac{\sqrt{3}}{2}, \\ \varepsilon_2 &= \cos\frac{4\pi}{3}+i\operatorname{sen}\frac{4\pi}{3}=-\frac{1}{2}-i\frac{\sqrt{3}}{2}. \end{aligned}\right\} \tag{7}$$

Todos los valores de la raíz n-ésima del número complejo α se pueden obtener multiplicando uno de estos valores por todas las raíces n-ésimas de la unidad. En efecto, sea β uno de los valores de la raíz n-ésima del número α, o sea, que $\beta^n=\alpha$, y sea ε un valor arbitrario de la raíz n-ésima de la unidad, o sea, que $\varepsilon^n=1$. Entonces, $(\beta\varepsilon)^n=\beta^n\varepsilon^n=\alpha$, es decir, $\beta\varepsilon$ también es uno de los valores de $\sqrt[n]{\alpha}$. Multiplicando β por cada una de las raíces n-ésimas de la unidad, obtenemos n valores diferentes de la raíz n-ésima del número α, o sea, todos los valores de esta raíz.

Ejemplos. 1) Uno de los valores de la raíz cúbica de -8 es -2. En virtud de (7), los otros dos serán los números $-2\varepsilon_1=1-i\sqrt{3}$ y $-2\varepsilon_2=1+i\sqrt{3}$ (véase el ejemplo anterior 3). 2) $\sqrt[4]{81}$ tiene cuatro valores: 3, -3, $3i$, $-3i$.

El producto de dos raíces n-ésimas de la unidad también es una raíz n-ésima de la unidad. En efecto, si $\varepsilon^n=1$ y $\eta^n=1$, se tiene, $(\varepsilon\eta)^n=$ $=\varepsilon^n\eta^n=1$. Por otra parte, el *número recíproco de la raíz n-ésima de la unidad también es una raíz n-ésima de la unidad.* En efecto, sea $\varepsilon^n=1$. Entonces, de la igualdad $\varepsilon^n.\ \varepsilon^{-1}=1$ resulta, $\varepsilon^n(\varepsilon^{-1})^n=$ $=1$, o sea, $(\varepsilon^{-1})^n=1$. En general, *toda potencia de la raíz n-ésima de la unidad también es una raíz n-ésima de la unidad.*

Toda raíz k-ésima de la unidad también es raíz l-ésima de la unidad para cualquier l, múltiplo de k. De esto se deduce que, considerando todo el conjunto de las raíces n-ésimas de la unidad, algunas de estas raíces también son raíces n'-ésimas de la unidad para ciertos n', divisores del número n. Sin embargo, para todo n existen raíces n-ésimas de la unidad que no son raíces de la unidad de orden menor. Estas se llaman *raíces primitivas* n-ésimas de la unidad. Su existencia se deduce de la fórmula (6): si se designa con ε_k el valor de la raíz que corresponde al valor dado de k (de modo que $\varepsilon_0 = 1$), en virtud de la fórmula de Moivre (1), se tiene:

$$\varepsilon_1^k = \varepsilon_k.$$

Por consiguiente, ninguna potencia del número ε_1, menor que la n-ésima, será igual a 1, o sea, el número $\varepsilon_1 = \cos \frac{2\pi}{n} + i \operatorname{sen} \frac{2\pi}{n}$ es una raíz primitiva.

La raíz n-ésima de la unidad ε *es primitiva cuando, y sólo cuando, sus potencias* ε^k, $k = 0, 1, \ldots, n-1$, *son diferentes, es decir, si con ellas se agotan todas las raíces n-ésimas de la unidad.*

En efecto, si todas las potencias indicadas del número ε son diferentes, es evidente que éste es raíz primitiva n-ésima de la unidad. Si, por el contrario, $\varepsilon^k = \varepsilon^l$ para $0 \leqslant k < l \leqslant n-1$, entonces, $\varepsilon^{l-k} = 1$, y en virtud de las desigualdades $1 \leqslant l - k \leqslant n - 1$, la raíz ε no será primitiva.

En el caso general, el número ε_1 hallado anteriormente no será la única raíz primitiva n-ésima de la unidad. Para hallar todas estas raíces se aplica el teorema siguiente:

Si ε *es una raíz primitiva n-ésima de la unidad, el número* ε^k *es una raíz primitiva n-ésima de la unidad cuando, y sólo cuando,* k *es primo con* n.

En efecto, sea d el máximo común divisor de los números k y n. Si $d > 1$ y $k = dk'$, $n = dn'$, entonces,

$$(\varepsilon^k)^{n'} = \varepsilon^{kn'} = \varepsilon^{k'n} = (\varepsilon^n)^{k'} = 1,$$

o sea, la raíz ε^k resulta raíz n'-ésima de la unidad.

Por otra parte, supongamos que $d = 1$ y que el número ε^k es una raíz m-ésima de la unidad, $1 \leqslant m < n$. Por lo tanto,

$$(\varepsilon^k)^m = \varepsilon^{km} = 1.$$

Como el número ε es una raíz **primitiva** n-ésima de la unidad, lo que implica que pueden ser iguales a la unidad solamente las potencias del mismo cuyos exponentes sean múltiplos de n, el número km es múltiplo de n. Sin embargo, como $1 \leqslant m < n$, resulta que los números k y n no pueden ser primos entre sí, lo que contradice a la hipótesis. Por lo tanto, el número de raíces primitivas n-ésimas de la unidad es igual al número de enteros positivos k, menores de n y primos con k. La expresión de este número que, generalmente, se designa mediante $\varphi(n)$, se puede hallar en cualquier tratado sobre la teoría de los números. Si p es un número primo, todas las raíces p-ésimas de la unidad son primitivas, a excepción de la unidad misma. Por otra parte, entre las raíces cuárticas de la unidad son primitivas i y $-i$, pero no 1 y -1.