Fundamentos de Algebra Lineal

A. I. Máltsev

EDITORIAL MIR MOSCU

ИЗДАТЕЛЬСТВО • МИР •

ОСНОВЫ ЛИНЕЙНОЙ АЛГЕБРЫ

А. И. Мальцев

ИЗДАТЕЛЬСТВО
«НАУКА»

A. I. Máltsev

Fundamentos de Algebra lineal

Traducido del ruso por
Carlos Vega,
candidato a Doctor en ciencias físico-matemáticas

Tercera edición

EDITORIAL
MIR
MOSCÚ

A NUESTROS LECTORES:

"Mir" edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas extranjeros. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia-ficción.
Dirijan sus opiniones a la Editorial "Mir", 1 Rizhski per., 2, 129 820, Moscú, I- I-110, GSP, URSS

Primera edición, 1972

Segunda edición, 1976

на испанском языке

Impreso en la URSS

INDICE

Capítulo V. Espacios unitarios y euclídeos

INTRODUCCION

El Algebra lineal es una rama de las Matemáticas tan antigua como la propia Matemática. El problema de la solución de la ecuación lineal $ax+b=0$ puede ser considerado como el problema primario del Algebra lineal. Aunque este problema no representa dificultad alguna, el método de su solución, así como las propiedades de la función lineal correspondiente $y=ax+b$, son los modelos de partida para las ideas y los métodos de toda el álgebra lineal. Por ejemplo, la teoría de la solución de un sistema de ecuaciones con varias incógnitas se basa en la idea de la sustitución del sistema dado por una cadena de ecuaciones del tipo indicado y de la forma más sencilla.

La importancia de los sistemas de ecuaciones lineales aumentó particularmente con la creación de la Geometría analítica, que permitió reducir todos los problemas principales sobre la posición de planos y rectas en el espacio al estudio de sistemas de ecuaciones lineales. Ya en el siglo XVIII la búsqueda de fórmulas generales para la solución de un sistema de n ecuaciones con n incógnitas llevó a Leibniz y a Cramer al concepto de determinante. En el siglo XIX, además del Algebra y de la Geometría analítica, los determinantes penetraron también en el Análisis con los trabajos de Ostrogradski, Jacobi (determinantes funcionales), Wronski y otros. Paralelamente en la Geometría analítica, en la teoría de los números y especialmente en la mecánica teórica adquiría cada vez mayor importancia el problema de transformación de formas cuadráticas mediante sustituciones lineales de las variables. Este problema resultó ser también uno de los centrales en el desarrollo de las ideas geométricas de Lobachevski y de Riemann, que llevó a la creación de la teoría de espacios lineales multidimensionales (Grassmann). A mediados del siglo pasado y en relación con el estudio de álgebras no conmutativas (Hamilton) apareció en los trabajos de Cayley y de Sylvester el cálculo de matrices, que en el desarrollo posterior del Algebra lineal pasó a ocupar uno de los puestos principales. A finales del siglo XIX quedaron creados los capítulos principales del cálculo de matrices: forma normal de una matriz de una transformación lineal (Jordan), divisores elementales (Weierstrass), pares de formas cuadráticas (Weierstrass, Kronecker), formas hermitianas (Hermite). El desarrollo de la geometría diferencial de espacios multidimensionales y de la teoría de transformaciones de formas algebraicas de órdenes superiores llevó, a finales del siglo XIX, a la creación del cálculo tensorial.

En el siglo actual los métodos del Algebra lineal han encontrado amplia aplicación y han sido desarrollados en la teoría de anillos y módulos, en la teoría de representaciones de grupos, así como en la teoría de espacios topológicos vectoriales y otros capítulos del Análisis funcional. Ya en las dos últimas décadas la teoría de desigualdades lineales y la teoría de espacios afines multidimensionales, estrechamente ligada a la primera, han ocupado uno de los lugares centrales en una rama tan conocida de la matemática aplicada como es la teoría de operaciones. Gracias a ello los elementos de la teoría de espacios afines multidimensionales constituyen ahora un momento indispensable en la formación matemática de ingenieros y economistas.

En el Algebra lineal se estudian objetos de tres géneros: matrices, espacios y formas algebraicas. Las teorías de estos objetos están estrechamente vinculadas. La mayoría de los problemas del Algebra lineal admite un enunciado natural en términos de cada una de las tres teorías señaladas. El enunciado matricial es generalmente el más cómodo para los fines de cálculo. Por otra parte, en la geometría y en la mecánica la mayoría de los problemas del Algebra lineal aparece como problemas de estudio de formas algebraicas. Sin embargo, la comprensión más clara de las relaciones internas de diferentes problemas del Algebra lineal se alcanza solamente al considerar los espacios lineales correspondientes que son por ello el objeto principal de estudio en el Algebra lineal.

Desde el punto de vista de la teoría de formas el contenido del Algebra lineal se descompone de modo natural en la teoría de formas lineales, cuadráticas y de órdenes superiores. El álgebra lineal propiamente dicha se relaciona, en general, solamente con la teoría de formas lineales y cuadráticas, así como con los elementos iniciales de la teoría de formas polilineales y del álgebra tensorial.

Capítulo I

Matrices y determinantes

§ 1. Operaciones con matrices

1.1. Matrices. Campo principal. Los objetos principales de estudio en lo sucesivo serán las matrices, los espacios lineales y los polinomios de varias variables, llamados también formas algebraicas. En la definición de cada uno de estos objetos participa un conjunto K de números o elementos de otra índole que debe ser previamente elegido. La elección concreta de K depende de los problemas que se resuelven y de la disciplina científica. Por ejemplo, desde el punto de vista algebraico los resultados obtienen la forma más completa, si K es el conjunto de todos los números complejos. Por el contrario, en la geometría y en la mecánica es preciso considerar generalmente los números reales, mientras que en la teoría de los números resulta natural aceptar que K es el conjunto de los números racionales o, incluso, solamente el conjunto de los números racionales enteros. Para que los resultados puedan ser aplicados a un número de problemas lo más amplio posible conviene, por lo tanto, no fijar de antemano qué conjunto concreto se comprende por K. En algunas secciones será suficiente suponer que K es un anillo asociativo. En varios capítulos aceptaremos que K es un cuerpo conmutativo arbitrario o un cuerpo conmutativo arbitrario ordenado, mientras que varios teoremas importantes serán demostrados solamente en la suposición de que K es el conjunto de todos los números reales o el conjunto de todos los números complejos. Para las aplicaciones geométricas y físicas son de mayor importancia precisamente los casos en que K es el cuerpo de los números reales o el cuerpo de los números complejos.

Los elementos del conjunto K se llamarán *números* incluso cuando K sea un anillo arbitrario. Los representaremos por letras griegas minúsculas α, β, ..., τ.

Un sistema arbitrario de elementos del conjunto K dispuestos en forma de una tabla rectangular de m filas y de n columnas se denomina (m, n)-matriz o simplemente *matriz* sobre K. Para representar una matriz, los símbolos que designan sus elementos suelen escribirse en el orden adecuado y la tabla obtenida se incluye entre paréntesis, corchetes o barras verticales dobles. Por consiguiente, la forma general de una (m, n)-matriz será

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{bmatrix}, \quad \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix}, \quad \begin{Vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{Vmatrix},$$

donde α_{ij} representan elementos de K. En vez de esta notación detallada con frecuencia se emplea la notación abreviada: $\|\alpha_{ij}\|$ o $\|\alpha_{ij}\|_{m,n}$.

Si el número de filas coincide con el número de columnas, la matriz se llama *cuadrada* y el número de sus filas, igual al de sus columnas, se llama *orden* de la matriz cuadrada. En particular, una matriz cuadrada de orden 1 es simplemente un elemento de K.

Una matriz compuesta de una sola fila se llama simplemente *fila* y el número de sus elementos se denomina *longitud* de fila. En lo sucesivo las matrices serán representadas por letras mayúsculas latinas.

Dos matrices se llaman *iguales*, si son iguales los números de sus filas y de sus columnas respectivamente y si coinciden los números que ocupan posiciones correspondientes en estas matrices. Por consiguiente, una igualdad entre dos (m, n)-matrices equivale a mn igualdades entre sus elementos.

Las operaciones matriciales principales son: la multiplicación de un número por una matriz o de una matriz por un número, la adición y la multiplicación de dos matrices. Por definición, para multiplicar el número α por la matriz A o la matriz A por el número α hay que multiplicar por α todos los elementos de la matriz A. Por ejemplo,

$$\alpha \cdot \begin{bmatrix} \lambda & \mu \\ \xi & \eta \end{bmatrix} = \begin{bmatrix} \alpha\lambda & \alpha\mu \\ \alpha\xi & \alpha\eta \end{bmatrix}, \qquad \begin{bmatrix} \lambda & \mu \\ \xi & \eta \end{bmatrix} \cdot \alpha = \begin{bmatrix} \lambda\alpha & \mu\alpha \\ \xi\alpha & \eta\alpha \end{bmatrix}.$$

Si las matrices se consideran sobre un anillo conmutativo K, es válida la igualdad $\alpha A = A\alpha$ para cualquier matriz A y para cualquier $\alpha \in K$. En el caso de un anillo no conmutativo K puede resultar que $\alpha A \neq A\alpha$. Siendo K el anillo de todos los números enteros, tenemos, por ejemplo,

$$5 \cdot \begin{bmatrix} 2 & 3 \\ 7 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 7 & -1 \end{bmatrix} \cdot 5 = \begin{bmatrix} 10 & 15 \\ 35 & -5 \end{bmatrix}.$$

La matriz, en la cual todos los elementos son iguales a cero, se llama *matriz nula* y se designa O. Si se quiere indicar de manera explícita el número de filas y de columnas de la matriz nula, se escribe O_{mn}.

Está claro que para toda matriz A sobre K y para cualesquiera α, $\beta \in K$ tienen lugar las relaciones:

1. $1 \cdot A = A \cdot 1 = A$.
2. $0 \cdot A = A \cdot 0 = O$; $\alpha \cdot O = O \cdot \alpha = O$.
3. $\alpha(\beta A) = (\alpha\beta) A$; $(A\alpha)\beta = A(\alpha\beta)$.

Se llama suma de dos matrices A y B de igual cantidad de filas y de columnas respectivamente una matriz con el mismo número de filas y de columnas, cuyos elementos son iguales a las sumas de los elementos correspondientes de las matrices A y B. Por ejemplo,

$$\begin{bmatrix} 2 & 1 & 3 \\ 1 & -3 & 5 \end{bmatrix} + \begin{bmatrix} 0 & 2 & -2 \\ 1 & 2 & -5 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 1 \\ 2 & -1 & 0 \end{bmatrix}.$$

De esta definición se desprenden inmediatamente las relaciones:

4. $A + (B + C) = (A + B) + C$;
5. $A + B = B + A$;
6. $(\alpha + \beta) A = \alpha A + \beta A$; $A(\alpha + \beta) = A\alpha + A\beta$;
7. $\alpha(A + B) = \alpha A + \alpha B$; $(A + B)\alpha = A\alpha + B\alpha$;
8. $A + O = O + A = A$;

la demostración queda a cargo del lector. En particular, empleando las propiedades 1 y 6, obtenemos

$$A + A = 2A, \quad A + A + A = 3A, \ \ldots$$

Introduciendo la notación $(-1)A = -A$, tendremos también

$$A + (-A) = O, \quad (-\alpha) A = -\alpha A, \quad -(A + B) = -A - B,$$
$$-(-A) = A.$$

Para abreviar, en lugar de $A + (-B)$ suele escribirse $A - B$.

1.2. Multiplicación de matrices. A diferencia de las operaciones de adición y de multiplicación por un número, la operación de multiplicación de una matriz por otra se define de forma más compleja. A saber, sean dadas dos matrices A y B, tales que el número de columnas de la primera coincide con el número de filas de la segunda. Si

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \beta_{11} & \beta_{12} & \ldots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \ldots & \beta_{2p} \\ \cdot & \cdot & \cdot & \cdot \\ \beta_{n1} & \beta_{n2} & \ldots & \beta_{np} \end{bmatrix},$$

la matriz

$$C = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \ldots & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & \ldots & \gamma_{2p} \\ \cdot & \cdot & \cdot & \cdot \\ \gamma_{m1} & \gamma_{m2} & \ldots & \gamma_{mp} \end{bmatrix},$$

donde

$$\gamma_{ij}=\alpha_{i1}\beta_{1j}+\alpha_{i2}\beta_{2j}+\ldots+\alpha_{in}\beta_{nj}\ (i=1,\ \ldots,\ m;\ j=1,\ \ldots,\ p),$$

se denomina *producto de A por B* y se designa AB. Por ejemplo,

$$\begin{bmatrix}\alpha & \beta\\ \alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{bmatrix}\begin{bmatrix}\gamma & \delta & \varepsilon\\ \lambda & \mu & \nu\end{bmatrix}=\begin{bmatrix}\alpha\gamma+\beta\lambda & \alpha\delta+\beta\mu & \alpha\varepsilon+\beta\nu\\ \alpha_1\gamma+\beta_1\lambda & \alpha_1\delta+\beta_1\mu & \alpha_1\varepsilon+\beta_1\nu\\ \alpha_2\gamma+\beta_2\lambda & \alpha_2\delta+\beta_2\mu & \alpha_2\varepsilon+\beta_2\nu\end{bmatrix}.$$

La regla de multiplicación de matrices se enuncia, a veces, de la siguiente forma: *para obtener el elemento, que se encuentra en la i-ésima fila y j-ésima columna del producto de dos matrices, hay que multiplicar los elementos de la i-ésima fila de la primera matriz por los elementos correspondientes de la j-ésima columna de la segunda matriz y sumar los productos obtenidos.*

El producto de dos matrices, hablando en términos generales, depende del orden de los factores incluso en el caso en que el anillo K es conmutativo. Por ejemplo,

$$\begin{bmatrix}1 & 2\\ 3 & 2\end{bmatrix}\cdot\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}=\begin{bmatrix}3 & 4\\ 5 & 8\end{bmatrix},$$

$$\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}\cdot\begin{bmatrix}1 & 2\\ 3 & 2\end{bmatrix}=\begin{bmatrix}7 & 6\\ 4 & 4\end{bmatrix}.$$

Si se consideran matrices no cuadradas, puede ocurrir incluso que el producto de dos matrices tomadas en un orden tenga sentido y tomadas en el orden contrario, no lo tenga.

Demostremos ahora las propiedades principales de la multiplicación de matrices.

9. $\alpha(AB)=(\alpha A)B;\quad A(\alpha B)=(A\alpha)B;\quad (AB)\alpha=A(B\alpha).$

Sean $A=\|\alpha_{ij}\|_{mn}$ y $B=\|\beta_{jk}\|_{np}$. Para el elemento que se encuentra en la i-ésima fila y k-ésima columna de la matriz $\alpha(AB)$ $(i=1,\ \ldots,\ m;\ k=1,\ \ldots,\ p)$, obtenemos, empleando la regla de multiplicación de matrices, la expresión siguiente:

$$\alpha(\alpha_{i1}\beta_{1k}+\ldots+\alpha_{in}\beta_{nk}).$$

Análogamente, para el elemento que se encuentra en la misma i-ésima fila y en la misma k-ésima columna de la matriz $(\alpha A)B$, obtenemos la expresión

$$(\alpha\alpha_{i1})\beta_{1k}+\ldots+(\alpha\alpha_{in})\beta_{nk}.$$

Como ambas expresiones coinciden, queda demostrada la primera de las igualdades 9. Con cálculos semejantes se demuestran las otras dos igualdades de 9, así como las propiedades:

10. $(A+B)C=AC+BC.$
11. $C(A+B)=CA+CB.$

De las propiedades 10 y 11 se desprende directamente la siguiente regla general: *para multiplicar una suma de matrices por otra hay que multiplicar cada matriz de la primera suma por cada matriz de la segunda suma y sumar los productos obtenidos.*

Hemos visto que para el producto de matrices no se cumple la ley conmutativa: AB puede ser distinto de BA. Sin embargo, la segunda ley aritmética—la ley asociativa de la multiplicación—se cumple para la multiplicación de matrices[1].

12. $A(BC)=(AB)C$.

Para la demostración tomemos

$$AB=M \quad \text{y} \quad BC=N$$

y representemos mediante μ_{ik} y ν_{jl} los elementos de las matrices M y N. Según la regla de multiplicación de matrices tenemos

$$\mu_{ik}=\alpha_{i1}\beta_{1k}+\alpha_{i2}\beta_{2k}+\ldots+\alpha_{in}\beta_{nk},$$
$$\nu_{jl}=\beta_{j1}\gamma_{1l}+\beta_{j2}\gamma_{2l}+\ldots+\beta_{jp}\gamma_{pl},$$

donde α_{ij}, β_{jk} y γ_{kl} son los elementos de las matrices A, B y C. Efectuando la multiplicación de M por C, obtendremos en la i-ésima fila y l-ésima columna de la matriz $(AB)C$ la suma

$$\mu_{i1}\gamma_{1l}+\mu_{i2}\gamma_{2l}+\ldots+\mu_{ip}\gamma_{pl}=\sum_k\sum_j\alpha_{ij}\beta_{jk}\gamma_{kl}.$$

Análogamente, efectuando la multiplicación de A por N, obtendremos en la i-ésima fila y l-ésima columna del producto $A(BC)$ la suma

$$\alpha_{i1}\nu_{1l}+\alpha_{i2}\nu_{2l}+\ldots+\alpha_{in}\nu_{nl}=\sum_j\sum_k\alpha_{ij}\beta_{jk}\gamma_{kl}.$$

Puesto que estas dos sumas difieren solamente en el orden de los sumandos, la fórmula 12 queda demostrada.

De la fórmula 12 se deduce que el producto de varias matrices dispuestas en un orden determinado no depende de cómo se coloquen los paréntesis. Por esto podemos hablar no sólo sobre el producto de dos matrices, sino también sobre el producto de un número mayor de matrices. Por ejemplo, podemos hablar simplemente del producto $ABCD$ de cuatro matrices, ya que las cinco formas diferentes de calcular este producto

$$((AB)C)D, \quad (A(BC))D, \quad A((BC)D), \quad A(B(CD)), \quad (AB)(CD)$$

llevan al mismo resultado. En efecto, cada producto siguiente se obtiene del anterior aplicando directamente la ley asociativa 12.

[1] Puesto que no se pueden sumar y multiplicar matrices arbitrarias, sino solamente aquellas en las que el número de filas y de columnas está sujeto a condiciones determinadas, las igualdades 10, 11 y 12 deben comprenderse de manera que si las operaciones indicadas en uno de los miembros son posibles, las operaciones indicadas en el otro miembro también son posibles y los resultados obtenidos en ambos miembros coinciden.

Ya hemos señalado que cualesquiera dos matrices no pueden ser sumadas o multiplicadas, ya que para poder realizar estas operaciones es preciso que se cumplan determinadas relaciones entre los números de filas y de columnas. Este inconveniente desaparece si se consideran solamente matrices cuadradas de un orden fijo n. Cualesquiera dos matrices de este tipo pueden ser sumadas o multiplicadas, así como multiplicadas por cualesquiera números de K, y el resultado será otra vez una matriz cuadrada del mismo orden n. Las propiedades 1—12 indican que *el conjunto de todas las matrices cuadradas de orden dado n sobre un anillo asociativo arbitrario K forman, a su vez, un anillo asociativo respecto a las operaciones matriciales de adición y de multiplicación.*

En lo que sigue aceptaremos que el conjunto numérico principal K es un anillo asociativo con el elemento unidad 1. La matriz cuadrada, en la que todos los elementos diagonales son iguales a 1 y los restantes son iguales a cero, se denomina matriz *unidad* y se designa mediante E o E_n, donde n es su orden. Por consiguiente,

$$E=\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & 1\end{bmatrix}.$$

Efectuando el cálculo directo, es fácil obtener para cualquier matriz cuadrada A la igualdad

$$AE=EA=A,$$

que expresa la propiedad fundamental de la matriz E. Las matrices de tipo

$$\begin{bmatrix}\alpha & 0 & \dots & 0\\ 0 & \beta & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \gamma\end{bmatrix},$$

se denominan *diagonales*.

De las reglas para las operaciones se desprende directamente que la suma y el producto de matrices diagonales son también matrices diagonales:

$$\begin{bmatrix}\alpha & 0 & \dots & 0\\ 0 & \beta & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \gamma\end{bmatrix}+\begin{bmatrix}\alpha_1 & 0 & \dots & 0\\ 0 & \beta_1 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \gamma_1\end{bmatrix}=\begin{bmatrix}\alpha+\alpha_1 & 0 & \dots & 0\\ 0 & \beta+\beta_1 & \dots & 0\\ \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & \dots & \gamma+\gamma_1\end{bmatrix},$$

$$\begin{bmatrix} \alpha & 0 & \dots & 0 \\ 0 & \beta & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \gamma \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \beta_1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \gamma_1 \end{bmatrix} = \begin{bmatrix} \alpha\alpha_1 & 0 & \dots & 0 \\ 0 & \beta\beta_1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \gamma\gamma_1 \end{bmatrix}.$$

Consideremos ahora una matriz cuadrada cualquiera X de orden n formada por elementos del anillo K. Por definición, tomamos

$$X^0 = E, \quad X^1 = X, \quad X^2 = XX, \quad X^3 = XXX, \dots$$

Puesto que en los productos de varías matrices los paréntesis pueden ser dispuestos arbitrariamente, tenemos para cualesquiera enteros no negativos p y q y para cualquier matriz cuadrada X sobre el anillo asociativo K

$$X^p X^q = X^{p+q}, \tag{1}$$

$$(X^p)^q = X^{pq}. \tag{2}$$

Las matrices A y B se llaman *conmutables*, si

$$AB = BA. \tag{3}$$

De la relación (1) obtenemos

$$X^p X^q = X^{p+q} = X^q X^p,$$

y, por consiguiente, *todas las potencias naturales de una misma matriz son conmutables entre sí.*

Es válida incluso una afirmación más general: *si las matrices A y B son conmutables, cualesquiera potencias naturales de las mismas también son conmutables y para cualquier p natural se tiene*

$$(AB)^p = A^p B^p. \tag{4}$$

En efecto, para cualesquiera naturales p y q se tiene

$$A^p B^q = AA \dots AB \dots B.$$

Por hipótesis, en este producto pueden ser permutados cualesquiera dos factores contiguos. Pero mediante permutaciones de este tipo los factores pueden ser dispuestos en cualquier orden y, en particular, todos los factores iguales a B pueden ser llevados a las posiciones primeras. Análogamente se demuestra también la fórmula (4).

Consideremos ahora un polinomio

$$\varphi(\lambda) = \alpha_0 + \alpha_1\lambda + \dots + \alpha_n\lambda^n$$

en la letra λ, cuyos coeficientes pertenecen al anillo K. Si A es una matriz cuadrada sobre K, la expresión

$$\alpha_0 E + \alpha_1 A + \dots + \alpha_n A^n$$

se denomina *valor* del polinomio $\varphi(\lambda)$ para $\lambda = A$ o, brevemente, polinomio correspondiente en la matriz A. Suponiendo que el anillo K es conmutativo, llegamos fácilmente a la conclusión de que *el valor de una suma de polinomios en λ para $\lambda = A$ es igual a la suma de los valores de los sumandos y de que el valor de un producto de polinomios es igual al producto de los valores de los factores.*

Sean

$$\varphi(\lambda) = \alpha_0 + \alpha_1\lambda + \ldots + \alpha_n\lambda^n,$$
$$\psi(\lambda) = \beta_0 + \beta_1\lambda + \ldots + \beta_n\lambda^n.$$

Entonces,

$$f(\lambda) = \varphi(\lambda) + \psi(\lambda) = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)\lambda + \ldots + (\alpha_n + \beta_n)\lambda^n,$$
$$g(\lambda) = \varphi(\lambda)\psi(\lambda) = \alpha_0\beta_0 + (\alpha_0\beta_1 + \alpha_1\beta_0)\lambda + \ldots + \alpha_n\beta_n\lambda^{2n}.$$

Nuestras afirmaciones consisten en que

$$f(A) = \varphi(A) + \psi(A),$$
$$g(A) = \varphi(A)\psi(A).$$

Para la demostración es suficiente escribir las expresiones para $\varphi(A)$, $\psi(A)$, $f(A)$ y $g(A)$ y siguiendo las reglas 1—12 del cálculo matricial efectuar la adición y la multiplicación de $\varphi(A)$ y $\psi(A)$.

A título de ejemplo consideremos la igualdad

$$\lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Tomando los valores del primero y segundo miembro para $\lambda = A$, obtenemos la igualdad matricial

$$A^2 - E = (A - E)(A + E).$$

De manera análoga de la igualdad

$$\lambda^3 + 1 = (\lambda + 1)(\lambda^2 - \lambda + 1)$$

obtenemos la relación

$$A^3 + E = (A + E)(A^2 - A + E).$$

En general, de esta forma se puede obtener de toda relación entre polinomios en λ una identidad matricial. En particular, según las reglas de operaciones con polinomios, se tiene

$$\varphi(\lambda)\psi(\lambda) = \psi(\lambda)\varphi(\lambda).$$

Tomando aquí en lugar de λ una matriz cuadrada A, obtenemos

$$\varphi(A)\psi(A) = \psi(A)\varphi(A).$$

Por consiguiente, los polinomios en una misma matriz son conmutables.

1.3. Transposición de matrices. Consideremos una matriz arbitraria

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

La matriz

$$A' = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{m2} \\ \dots & \dots & \dots & \dots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{mn} \end{bmatrix},$$

que se obtiene de A al cambiar las filas por las columnas, se llama *transpuesta* respecto de A. En lo sucesivo la raya siempre indicará el paso a la matriz transpuesta.

Para dos matrices arbitrarias A y B tienen lugar las siguientes reglas de transposición:

$$(\alpha A + \beta B)' = \alpha A' + \beta B',$$
$$(AB)' = B'A',$$

donde α y β son números cualesquiera. Demostremos, por ejemplo, la segunda de estas igualdades. El elemento, que se encuentra en la i-ésima fila y j-ésima columna de la matriz $(AB)'$, es igual al elemento que aparece en la j-ésima fila e i-ésima columna de la matriz AB, es decir, es igual a

$$\alpha_{j1}\beta_{1i} + \alpha_{j2}\beta_{2i} + \dots + \alpha_{jn}\beta_{ni},$$

donde α_{ij} y β_{ij} son los elementos de las matrices A y B. Pero esta expresión es la suma de los productos de los elementos de la i-ésima fila de la matriz B' por los elementos correspondientes de la j-ésima columna de la matriz A'; es decir, $(AB)' = B'A'$.

Si A es una matriz cuadrada cualquiera y

$$A' = A,$$

se dice que A *es simétrica*; en cambio, si

$$A' = -A,$$

se dice que A es *antisimétrica*. Los elementos simétricos respecto de la diagonal principal coinciden en una matriz simétrica y son opuestos en una matriz antisimétrica. En particular, todos los elementos diagonales de una matriz antisimétrica son iguales a cero.

De la regla de transposición de una suma se deduce directamente que la suma de matrices simétricas es una matriz simétrica y que la suma de matrices antisimétricas es una matriz antisimétrica. El producto de matrices simétricas puede no ser una matriz simétrica; por ejemplo,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 7 & 5 \end{bmatrix}.$$

Sin embargo, si dos matrices simétricas A y B son conmutables, el producto también será una matriz simétrica. En efecto, en este caso se tiene

$$(AB)' = B'A' = BA = AB.$$

De aquí se deduce que las potencias de una matriz simétrica son matrices simétricas y que los polinomios en una matriz simétrica también son matrices simétricas.

Una matriz cuadrada A sobre el anillo K se llama *invertible* (sobre K), si existe una matriz cuadrada X sobre K, que satisface las relaciones

$$AX = XA = E. \tag{1}$$

Toda matriz X que verifica las condiciones (1) se denomina matriz *inversa* de A o *inversión* de la matriz A. Para toda matriz invertible A existe solamente una inversión. En efecto, si además de la matriz X hay otra matriz Y que satisface las condiciones (1), multiplicando a la izquierda por X ambos miembros de la igualdad

$$AY = E,$$

obtenemos

$$XA \cdot Y = XE$$

o $Y = X$.

La inversión de la matriz A, si es que existe, se designa mediante A^{-1}. Es decir, por definición,

$$A \cdot A^{-1} = A^{-1} \cdot A = E. \tag{2}$$

En las condiciones (1) las matrices A y X figuran simétricamente y, por ello, si X es la inversión de A, A es la inversión de X; en otras palabras,

$$(A^{-1})^{-1} = A. \tag{3}$$

Si las matrices cuadradas A, B y C son de un mismo orden e invertibles, su producto ABC también es invertible y

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1},$$

es decir, *la inversión de un producto de matrices es igual al producto de las inversiones de los factores tomado en el orden contrario.*

Para la demostración es preciso comprobar las igualdades

$$ABC \cdot C^{-1}B^{-1}A^{-1} = C^{-1}B^{-1}A^{-1} \cdot ABC = E,$$

que son consecuencias evidentes de las relaciones (2) y de las relaciones análogas para las matrices B y C.

Para toda matriz invertible A, además de sus potencias naturales $A^0=E$, $A^1=A$, $A^2=AA$, ..., se consideran también sus potencias negativas enteras, tomando por definición

$$A^{-2}=A^{-1}A^{-1}, \quad A^{-3}=A^{-1}A^{-1}A^{-1}, \ \ldots \tag{4}$$

Las potencias fraccionarias de matrices se consideran raramente, debido a que en muchos casos las definiciones corrientes no ofrecen valores unívocos para estas potencias (véase el p. 16.2 del cap. IV).

De las relaciones (2) y (4) se deduce que para cualquier matriz invertible A y cualesquiera números enteros (no necesariamente positivos) p y q tienen lugar las reglas comunes de operaciones con potencias

$$A^pA^q=A^{p+q},$$
$$(A^p)^q=A^{pq},$$

y además, si las matrices A y B son invertibles y $AB=BA$, se tiene

$$(AB)^p=A^pB^p.$$

Veamos ahora la relación que existe entre las operaciones de transposición e inversión. Aplicando a las relaciones (1) la regla de transposición del producto de matrices, obtenemos

$$X'A'=A'X'=E,$$

es decir, al transponer una matriz invertible A se obtiene de nuevo una matriz invertible y

$$(A')^{-1}=(A^{-1})'. \tag{5}$$

Una matriz cuadrada A se llama *ortogonal*, si

$$AA'=A'A=E, \tag{6}$$

es decir, si la matriz transpuesta es inversa de la inicial. De aquí se deduce, en particular, que *toda matriz ortogonal es invertible.*

Puesto que $(A')'=A$, de (6) se deduce que *la inversión de una matriz ortogonal es una matriz ortogonal.*

Además, si las matrices A y B son ortogonales, se tiene

$$A'=A^{-1}, \quad B'=B^{-1}$$

y, por consiguiente,

$$(AB)'=B'A'=B^{-1}A^{-1}=(AB)^{-1}.$$

En otras palabras, *el producto de matrices ortogonales es una matriz ortogonal.*

Consideremos una operación matricial más. Sea A una matriz arbitraria, cuyos elementos son números complejos. Sustituyamos en A todo elemento por el número complejo conjugado. La matriz obtenida por este procedimiento se llama *conjugada compleja* de A y se designa por $\overline{A}$. La operación consistente en el paso a la matriz conjugada compleja posee las propiedades siguientes:

$$\overline{\alpha A+\beta B}=\overline{\alpha}\,\overline{A}+\overline{\beta}\,\overline{B},$$
$$\overline{AB}=\overline{A}\,\overline{B},$$
$$\overline{A'}=(\overline{A})',$$
$$\overline{A^{-1}}=(\overline{A})^{-1};$$

la demostración es muy sencilla y queda a cargo del lector.

Las matrices A y $\overline{A}'$ se denominan *conjugadas según Hermite* [1]). Si $A=\overline{A}'$, la matriz A se llama *hermitiana* o *simétrica según Hermite.*

Una matriz A que satisface la relación

$$\overline{A}'A=A\overline{A}'=E,$$

se llama *unitaria.*

Se puede demostrar, por el mismo procedimiento que el empleado para las matrices ortogonales, que *la matriz inversa de una matriz unitaria es unitaria* y que *el producto de matrices unitarias es también una matriz unitaria.*

Si todos los elementos de la matriz A son números reales, se tiene $\overline{A}=A$ y, por consiguiente, para las matrices reales los conceptos de simetría y de simetría según Hermite, así como los de matriz unitaria y de matriz ortogonal, coinciden.

1.4. Matrices celulares. Dividamos una matriz A en partes mediante un sistema de rectas verticales y horizontales. Estas partes pueden ser consideradas como matrices de órdenes inferiores que forman, interpretadas como elementos, la propia matriz; se denominan *células, cajas* o *bloques* de la matriz A, mientras que la propia matriz A, dividida de un modo determinado en células, se denomina, respectivamente, celular, de caja o de bloque. Una misma matriz puede ser dividida en células de diferentes maneras; por ejemplo:

$$\left[\begin{array}{cc|cc}1&8&7&6\\3&5&0&2\\ \hline 1&4&9&3\end{array}\right],\qquad \left[\begin{array}{ccc|c}1&8&7&6\\3&5&0&2\\ \hline 1&4&9&3\end{array}\right],\qquad \left[\begin{array}{cc|cc}1&8&7&6\\ \hline 3&5&0&2\\ \hline 1&4&9&3\end{array}\right].$$

La conveniencia de la división en células consiste en que las operaciones principales sobre matrices celulares se realizan formalmente siguiendo las mismas reglas que en el caso de matrices co-

[1]) En el capítulo V estas matrices se llaman también conjugadas transpuestas o anticonjugadas *(N. del Tr.)*

rrientes. En efecto, supongamos una matriz A dividida de algún modo en células:

$$A=\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{bmatrix}.$$

Al multiplicar todas las células por un número α multiplicaremos, al mismo tiempo, todos los elementos de la matriz A por α. Por consiguiente,

$$\alpha A=\begin{bmatrix} \alpha A_{11} & \alpha A_{12} & \dots & \alpha A_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha A_{m1} & \alpha A_{m2} & \dots & \alpha A_{mn} \end{bmatrix}.$$

Sea B una matriz dividida en el mismo número de células que la matriz A:

$$B=\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ B_{m1} & B_{m2} & \dots & B_{mn} \end{bmatrix}.$$

Supongamos, además, que las correspondientes células de las matrices A y B son del mismo número de filas y de columnas respectivamente.

Para sumar las matrices A y B hay que sumar, según la definición, sus elementos correspondientes. Pero lo mismo ocurrirá, si sumamos las células correspondientes de estas matrices. Por esto

$$A+B=\begin{bmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \dots & A_{1n}+B_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & \dots & A_{mn}+B_{mn} \end{bmatrix}.$$

La situación es menos evidente en el caso de la multiplicación. Consideremos las matrices

$$U=\begin{bmatrix} U_{11} & \dots & U_{1n} \\ \cdot & \cdot & \cdot \\ U_{m1} & & U_{mn} \end{bmatrix}, \qquad V=\begin{bmatrix} V_{11} & \dots & V_{1p} \\ \cdot & \cdot & \cdot \\ V_{n1} & & V_{np} \end{bmatrix},$$

divididas en células U_{ij} y V_{jk} de manera que el número de columnas de la célula U_{ij} sea igual al número de filas de la célula V_{jk} ($i=1, \dots, m$; $j=1, \dots, n$; $k=1, \dots, p$). En estas condiciones las expresiones

$$W_{ik}=U_{i1}V_{1k}+U_{i2}V_{2k}+\dots+U_{in}V_{nk}$$

tienen sentido. Es fácil demostrar que

$$UV=\begin{bmatrix} W_{11} & \dots & W_{1p} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ W_{m1} & \dots & W_{mp} \end{bmatrix}, \tag{1}$$

es decir, las matrices divididas de manera adecuada en células pueden ser multiplicadas de la forma corriente: *las células del producto son iguales a las sumas de los productos de las células de las filas de U por las células correspondientes de las columnas de V.*

Demostremos primero esta regla en el siguiente caso particular:

$$[AB]\begin{bmatrix} C \\ D \end{bmatrix} = AC + BD. \tag{2}$$

Sean α_{ij}, β_{ik}, γ_{jl} y δ_{kl} los elementos de las matrices A, B, C y D respectivamente ($i=1, \dots, m$; $j=1, \dots, n$; $k=1, \dots, s$; $l=1, \dots, t$). Efectuando las operaciones indicadas en el primer miembro de la igualdad (2), obtendremos que el elemento que se halla en la i-ésima fila y l-ésima columna es igual a

$$\alpha_{i1}\gamma_{1l}+\dots+\alpha_{in}\gamma_{nl}+\beta_{i1}\delta_{1l}+\dots+\beta_{is}\delta_{sl}.$$

Por otra parte, calculando el elemento correspondiente del segundo miembro, obtendremos la misma expresión y, por consiguiente, la igualdad (2) queda demostrada.

Empleando la fórmula (2), es fácil demostrar ahora una fórmula más general

$$[A_1A_2 \dots A_n]\begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = A_1B_1 + A_2B_2 + \dots + A_nB_n, \tag{3}$$

donde A_i y B_i son células. Para $n=2$ esta fórmula coincide con (2). Apliquemos la inducción. Supongamos que para los valores de n menores que un valor dado la fórmula (3) ha sido ya demostrada y sea

$$C=[A_2 \dots A_n], \quad D=\begin{bmatrix} B_2 \\ \vdots \\ B_n \end{bmatrix}.$$

En este caso de (2) obtenemos

$$[A_1A_2 \dots A_n]\begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix} = [A_1C]\begin{bmatrix} B_1 \\ D \end{bmatrix} = A_1B_1 + CD = A_1B_1 + A_2B_2 + \dots + A_nB_n.$$

De manera análoga se obtienen las fórmulas

$$A[B_1\ B_2\ \dots\ B_n] = [AB_1\ AB_2\ \dots\ AB_n], \tag{4}$$

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} B = \begin{bmatrix} A_1B \\ A_2B \\ \vdots \\ A_nB \end{bmatrix}, \tag{5}$$

Para deducir ahora de las fórmulas particulares (3), (4) y (5) la fórmula general (1), designemos mediante $U_1, \dots, U_m$ las filas de células de la matriz U y mediante

$V_1, \ldots, V_p$ las columnas de células de la matriz V. En base a la fórmula (5) tenemos

$$UV = \begin{bmatrix} U_1 \\ \vdots \\ U_m \end{bmatrix} V = \begin{bmatrix} U_1V \\ \vdots \\ U_mV \end{bmatrix}.$$

Tomando aquí en lugar de la matriz V su división en células $[V_1 \ldots V_p]$ y empleando la fórmula (4), obtenemos

$$UV = \begin{bmatrix} U_1V_1 & \ldots & U_1V_p \\ \cdot & \cdot\cdot\cdot & \cdot \\ U_mV_1 & \ldots & U_mV_p \end{bmatrix}. \tag{6}$$

Por otra parte, en virtud de (3) tenemos

$$U_iV_k = [U_{i1} \ \ldots \ U_{in}] \begin{bmatrix} V_{1k} \\ \vdots \\ V_{nk} \end{bmatrix} = U_{i1}V_{1k} + \ldots + U_{in}V_{nk} = W_{ik}.$$

Colocando estas expresiones en (6), obtendremos la fórmula (1).

En el caso de matrices cuadradas resulta necesario, como regla general, dividirlas de manera que las células diagonales también sean cuadradas. Es fácil ver que, divididas dos matrices cuadradas en células de manera que las células diagonales sean cuadradas y que los ordenes de las células diagonales correspondientes coincidan, esta división satisface tanto las condiciones en las que es posible la adición célula por célula, como las condiciones que son necesarias para poder multiplicarlas como matrices celulares.

Toda matriz celular de tipo

$$A = \begin{bmatrix} A_1 & O & \ldots & O \\ O & A_2 & \ldots & O \\ \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ O & O & \ldots & A_s \end{bmatrix},$$

donde $A_1, \ldots, A_s$ son células cuadradas y O son matrices nulas de dimensiones adecuadas, se llama matriz *celular diagonal*. En el mismo sentido se dice también que A se descompone en partes $A_1, \ldots, A_s$ o que A es *la suma directa* de las matrices $A_1, \ldots, A_s$; simbólicamente

$$A = A_1 \dotplus A_2 \dotplus \ldots \dotplus A_s.$$

Las operaciones con matrices descompuestas se reducen a las operaciones con sus células diagonales. De aquí, a su vez, se desprende que siendo $f(\lambda)$ un polinomio y A una matriz celular diagonal de células diagonales $A_1, \ldots, A_s$, se tiene

$$f(A)=\begin{bmatrix} f(A_1) & & & & \\ & f(A_2) & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \quad f(A_s) \end{bmatrix}. \tag{7}$$

1.5. Cuaternios. Las matrices constituyen un instrumento cómodo mediante el cual se pueden construir a partir de un anillo dado, por ejemplo, del anillo de los números reales, anillos de estructura más compleja. De forma sistemática este problema se estudia en la teoría de anillos y nos vamos a limitar a considerar solamente dos casos particulares.

Consideremos el anillo R_2 de todas las matrices cuadradas de segundo orden sobre el cuerpo R de los números reales. Tomemos

$$E=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{e} \quad I=\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Sea C el conjunto de todas las matrices de R_2 de tipo

$$\alpha E+\beta I=\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \quad (\alpha, \beta \in R). \tag{1}$$

Elevando al cuadrado la matriz I, obtenemos $I^2=-E$ y por ello las operaciones con las matrices (1) se pueden efectuar siguiendo las fórmulas

$$(\alpha E+\beta I) \pm (\gamma E+\delta I)=(\alpha \pm \gamma) E+(\beta \pm \delta) I,$$
$$(\alpha E+\beta I)(\gamma E+\delta I)=(\alpha\gamma-\beta\delta) E+(\alpha\delta+\beta\gamma) I,$$

es decir, las mismas fórmulas que para los números complejos $\alpha+\beta i$ y $\gamma+\delta i$. De las fórmulas señaladas se deduce también que el conjunto de matrices C es un anillo y que la correspondencia

$$\alpha E+\beta I \rightarrow \alpha+\beta i$$

es una aplicación isomorfa del anillo C sobre el anillo de todos los números complejos.

A partir de las matrices de segundo orden E e I construimos ahora cuatro matrices cuadradas de orden 4:

$$e=\begin{bmatrix} E & O \\ O & E \end{bmatrix}, \quad i=\begin{bmatrix} O & -E \\ E & O \end{bmatrix}, \quad j=\begin{bmatrix} I & O \\ O & -I \end{bmatrix}, \quad k=\begin{bmatrix} O & I \\ I & O \end{bmatrix}. \tag{2}$$

La matriz e es la matriz unidad corriente de orden 4 y, por ello,

$$ei=ie=i, \quad ej=je=j, \quad ek=ke=k. \tag{3}$$

Multiplicando las matrices (2) según las reglas de multiplicación de matrices celulares y teniendo en cuenta que $I^2 = -E$, obtenemos

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik \tag{4}$$

y, además,

$$i^2 = j^2 = k^2 = -e. \tag{5}$$

Es fácil memorizar las fórmulas (4): significan que en la secuencia i, j, k, i, j, k el producto de dos elementos consecutivos es igual al elemento que les sigue.

Las matrices de tipo

$$\alpha e + \beta i + \gamma j + \delta k = \begin{bmatrix} \alpha & -\gamma & -\beta & -\delta \\ \gamma & \alpha & \delta & -\beta \\ \beta & -\delta & \alpha & \gamma \\ \delta & \beta & -\gamma & \alpha \end{bmatrix}, \tag{6}$$

donde α, β, γ, δ son números reales arbitrarios, se llaman *cuaternios* (o, a veces, matrices cuaternios). De (6) se deduce que la representación de los cuaternios en la forma $\alpha e + \beta i + \gamma j + \delta k$ es unívoca. En otras palabras, la relación

$$\alpha e + \beta i + \gamma j + \delta k = \alpha_1 e + \beta_1 i + \gamma_1 j + \delta_1 k$$

equivale a cuatro igualdades:

$$\alpha = \alpha_1, \quad \beta = \beta_1, \quad \gamma = \gamma_1, \quad \delta = \delta_1.$$

La adición y la sustracción de los cuaternios, representados en la forma algebraica normal $\alpha e + \beta i + \gamma j + \delta k$, se realiza por la regla corriente:

$$\begin{aligned} &(\alpha e + \beta i + \gamma j + \delta k) \pm (\alpha_1 e + \beta_1 i + \gamma_1 j + \delta_1 k) = \\ &= (\alpha \pm \alpha_1) e + (\beta \pm \beta_1) i + (\gamma \pm \gamma_1) j + (\delta \pm \delta_1) k. \end{aligned} \tag{7}$$

Para multiplicar dos cuaternios, representados en la forma algebraica normal, es suficiente recurrir a la ley distributiva y a las tablas de multiplicación (3), (4) y (5). Como resultado llegamos a la fórmula

$$\begin{aligned} &(\alpha e + \beta i + \gamma j + \delta k)(\alpha_1 e + \beta_1 i + \gamma_1 j + \delta_1 k) = \\ &= (\alpha\alpha_1 - \beta\beta_1 - \gamma\gamma_1 - \delta\delta_1) e + (\alpha\beta_1 + \beta\alpha_1 + \gamma\delta_1 - \delta\gamma_1) i + \\ &+ (\alpha\gamma_1 + \gamma\alpha_1 + \delta\beta_1 - \beta\delta_1) j + (\alpha\delta_1 + \delta\alpha_1 + \beta\gamma_1 - \gamma\beta_1) k. \end{aligned} \tag{8}$$

Las fórmulas (7) y (8) muestran que el conjunto Q de los cuaternios matriciales es un anillo con unidad e, que constituye un subanillo del anillo de todas las matrices reales de orden 4. Los cuaternios e, i, j y k suelen llamarse cuaternios unidades. De las relaciones (4) se deduce que el anillo de los cuaternios es no conmutativo. El hecho más notable consiste en que el anillo de los cuaternios es un cuerpo, es decir, que en el anillo de los cuaternios pueden ser resueltas todas las ecuaciones de tipo

$$ax = b, \quad ya = b, \tag{9}$$

donde a y b son cualesquiera dos cuaternios dados, de los cuales a es diferente de cero. Más abajo se dan unas fórmulas cómodas para la solución de estas ecuaciones.

Consideremos un cuaternio cualquiera

$$q = \alpha e + \beta i + \gamma j + \delta k.$$

El número real

$$N(q) = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$$

se llama *norma* del cuaternio q. Puesto que α, β, γ y δ son números reales, la norma de un cuaternio es un número real no negativo e igual a cero sólo para el cuaternio nulo.

El cuaternio

$$q^* = \alpha e - \beta i - \gamma j - \delta k$$

se denomina *conjugado* de q. Es claro que $q^{**} = q$. Mediante la multiplicación directa de los cuaternios q y q^* (según la fórmula (8)) se obtienen las igualdades principales

$$qq^* = q^*q = N(q) \cdot e,$$

de donde

$$q \cdot \frac{1}{N(q)} \cdot q^* = e, \tag{10}$$

o

$$q^{-1} = \frac{1}{N(q)} q^* \tag{11}$$

(siempre que $q \neq 0$).

Volvamos ahora a las ecuaciones (9). Multiplicándolas por a^{-1}, la primera por la izquierda y la segunda por la derecha, obtenemos

$$x = a^{-1}b, \quad y = ba^{-1}. \tag{12}$$

Con la sustitución de estos valores en las ecuaciones (9) se demuestra que las fórmulas (12) ofrecen efectivamente las soluciones buscadas de estas ecuaciones.

En el caso general las soluciones $a^{-1}b$ y ba^{-1} son distintas. Por esto suelen llamarse cocientes por la izquierda y por la derecha de b por a, designándose mediante $a \diagdown b$ y $b \diagup a$, respectivamente. Efectuando el cálculo directo, es fácil demostrar las fórmulas

$$(\alpha a + \beta b)^* = \alpha a^* + \beta b^*,$$
$$(ab)^* = b^* a^*,$$

de donde es fácil deducir la importante relación

$$N(ab) = N(a)\,N(b) \quad (a,\ b \in Q).$$

En efecto,

$$N(ab)\cdot e = abb^*a^* = aa^*N(b) = N(a)\,N(b)\cdot e.$$

La aplicación $\alpha \to \alpha e$ es un isomorfismo del cuerpo (conmutativo) de los números reales en el cuerpo de los cuaternios. Esto permite identificar un cuaternio de tipo αe con el número α y en lugar de $\alpha e + \beta i + \gamma j + \delta k$ escribir simplemente $\alpha + \beta i + \gamma j + \delta k$. Así, por ejemplo, se tiene

$$(j+k)(1+i-j)^{-1} = (j+k)(1-i+j)\frac{1}{3} = -\frac{1}{3} - \frac{1}{3}i + \frac{2}{3}k.$$

La mayor parte de los resultados de las secciones siguientes del libro estará relacionada con objetos cuya definición depende de un cuerpo K dado de antemano. Aunque tendrán interés principal los casos en que K es un cuerpo conmutativo, los razonamientos se realizarán de manera que no quede excluido el caso de cuerpos no conmutativos. Al analizar estos razonamientos conviene tener en cuenta que el cuerpo K puede ser precisamente el cuerpo de los cuaternios Q.

Ejemplos y problemas

1. Sean

$$\varphi(\lambda) = -2 - 5\lambda + 3\lambda^2 \quad \text{y} \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix};$$

entonces

$$\varphi(A) = \begin{bmatrix} 14 & 2 \\ 3 & 14 \end{bmatrix}.$$

2. Demuéstrese que la matriz

$$U = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

es unitaria. ¿Bajo qué condiciones una matriz diagonal resulta ser ortogonal? ¿unitaria?

3. Hállese la inversa de la matriz

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 3 & 1 & 5 \\ 3 & 2 & 3 \end{bmatrix}.$$

4. Si

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

se tiene

$$A^n = \begin{bmatrix} 1 & n & \frac{1}{2}n(n-1) \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}.$$

5. Demuéstrese que todas las matrices conmutables con

$$A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

son de tipo

$$B=\begin{bmatrix} \alpha & \beta & \gamma & \delta \\ 0 & \alpha & \beta & \gamma \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & 0 & \alpha \end{bmatrix}.$$

6. Si una matriz posee dos de las tres propiedades siguientes: es *real*, es *ortogonal*, es *unitaria*, posee también la tercera.

7. Toda matriz cuadrada puede ser representada como la suma de una matriz simétrica y otra antisimétrica.

8. Una matriz I se llama *involutiva*, si $I^2=E$. Demuéstrese que si una matriz posee dos de las propiedades: es *simétrica*, es *ortogonal*, es *involutiva*, posee también la tercera.

9. Una matriz P se llama *idempotente*, si $P^2=P$. Demuéstrese que las matrices

$$\begin{bmatrix} -26 & -18 & -27 \\ 21 & 15 & 21 \\ 12 & 8 & 13 \end{bmatrix} \text{ y } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

son idempotentes.

10. Si P es idempotente, la matriz

$$I=2P-E$$

es involutiva y viceversa, si I es involutiva, la matriz

$$P=\frac{1}{2}(I+E)$$

es idempotente.

11. Consideremos las matrices cuadradas de orden n. Sea E_{ij} $(i, j=1, \ldots, \ldots, n)$ la matriz en la que el elemento de la i-ésima fila y j-ésima columna es igual a 1, mientras que los demás elementos son iguales a 0. En estas condiciones, para $A=\|\alpha_{ij}\|$ se tiene

$$A\cdot E_{ij}=\alpha_{1i}E_{1j}+\ldots+\alpha_{ni}E_{nj},$$
$$E_{ij}\cdot A=\alpha_{j1}E_{i1}+\ldots+\alpha_{jn}E_{in}.$$

Dedúzcase de aquí que la matriz A es conmutable con cada una de las matrices E_{ij} si, y sólo si, A es de la forma αE.

Empleando este resultado, demuéstrese que la matriz A es conmutable con una matriz cualquiera cuadrada de orden n si, y sólo si, $A=\alpha E$, donde α es un elemento del anillo principal K conmutable con cualquier elemento de K.

12. A veces, además de matrices de orden finito, se consideran también matrices de orden infinito que tienen la forma de una tabla infinita de dos entradas:

$$A=\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots \\ \alpha_{21} & \alpha_{22} & \cdots \\ \cdot & \cdot & \cdot \end{bmatrix}.$$

Una matriz de este tipo se llama *finita en fila*, si cada una de sus filas contiene solamente un número finito de elementos diferentes de cero. Las operaciones con las matrices finitas en fila (así como con las matrices finitas en columna, que se definen análogamente) se realizan siguiendo las mismas reglas que tienen lugar en el caso de matrices cuadradas de orden finito. Es fácil ver que el resultado es de nuevo una matriz finita en fila.

Sea $A=\|\alpha_{ij}\|$ una matriz de orden infinito tal que $1=\alpha_{12}=\alpha_{23}=\ldots$ y $\alpha_{st}=0$ para $t-s\neq 1$. Demuéstrese que $AA'=E$ y $A'A\neq E$.

§ 2. Determinantes

2.1. Definición. El concepto de determinante surgió en relación con el problema de solución de sistemas de ecuaciones lineales. Tomemos un cuerpo conmutativo K y consideremos sistemas elementales de ecuaciones de primer grado de dos y tres incógnitas y con coeficientes de K. Un sistema de dos ecuaciones lineales con dos incógnitas ξ_1 y ξ_2 se representa en la forma siguiente:

$$\begin{aligned}\alpha_{11}\xi_1+\alpha_{12}\xi_2=\beta_1,\\ \alpha_{21}\xi_1+\alpha_{22}\xi_2=\beta_2,\end{aligned} \tag{1}$$

donde α_{ij} y β_i son números de K dados. Las matrices

$$A=\begin{bmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{bmatrix} \quad \text{y} \quad B=\begin{bmatrix}\alpha_{11} & \alpha_{12} & \beta_1\\ \alpha_{21} & \alpha_{22} & \beta_2\end{bmatrix}$$

se llaman matriz *principal* y matriz *ampliada*, respectivamente, del sistema (1). Con el fin de eliminar la incógnita ξ_2, multipliquemos la primera ecuación por α_{22} y la segunda por $-\alpha_{12}$ y sumemos ambas. Obtendremos la ecuación

$$(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})\,\xi_1=\beta_1\alpha_{22}-\beta_2\alpha_{12}.$$

Si $\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21}\neq 0$, obtenemos de esta ecuación y de una ecuación análoga que se obtiene eliminando ξ_1

$$\xi_1=\frac{\alpha_{22}\beta_1-\alpha_{12}\beta_2}{\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21}}, \qquad \xi_2=\frac{\alpha_{11}\beta_2-\alpha_{21}\beta_1}{\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21}}. \tag{2}$$

Los denominadores de las expresiones de las incógnitas ξ_1 y ξ_2 coinciden y representan un polinomio en los elementos de la matriz principal A. El valor de este polinomio se llama *determinante* de la matriz A y se designa $\det A$ o $|A|$. Si la matriz viene dada por su tabla, el determinante se designa escribiendo la tabla entre barras verticales.

Es decir, por definición para cualquier matriz cuadrada de orden dos se tiene

$$\begin{vmatrix}\alpha & \beta\\ \gamma & \delta\end{vmatrix}=\alpha\delta-\beta\gamma. \tag{3}$$

Empleando los determinantes, la fórmula (2) puede ser escrita en la forma

$$\xi_1 = \frac{\begin{vmatrix} \beta_1 & \alpha_{12} \\ \beta_2 & \alpha_{22} \end{vmatrix}}{\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}}, \qquad \xi_2 = \frac{\begin{vmatrix} \alpha_{11} & \beta_1 \\ \alpha_{21} & \beta_2 \end{vmatrix}}{\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}}. \tag{4}$$

Resolviendo de forma análoga el sistema de tres ecuaciones

$$\begin{aligned} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \alpha_{13}\xi_3 &= \beta_1, \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \alpha_{23}\xi_3 &= \beta_2, \\ \alpha_{31}\xi_1 + \alpha_{32}\xi_2 + \alpha_{33}\xi_3 &= \beta_3 \end{aligned} \tag{5}$$

con tres incógnitas ξ_1, ξ_2 y ξ_3, obtenemos

$$\xi_1 = \frac{\beta_1\alpha_{22}\alpha_{33} - \beta_1\alpha_{23}\alpha_{32} + \beta_2\alpha_{32}\alpha_{13} - \beta_2\alpha_{12}\alpha_{33} + \beta_3\alpha_{12}\alpha_{23} - \beta_3\alpha_{22}\alpha_{13}}{\alpha_{11}\alpha_{22}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32} + \alpha_{12}\alpha_{23}\alpha_{31} - \alpha_{12}\alpha_{21}\alpha_{33} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31}} \tag{6}$$

y unas expresiones análogas para ξ_2 y ξ_3. Por supuesto estas expresiones tienen sentido sólo en el caso en que su denominador sea diferente de cero. Las matrices

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \beta_1 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \beta_2 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \beta_3 \end{bmatrix}$$

también se llaman matriz *principal* y matriz *ampliada* del sistema de ecuaciones (5). El denominador de la fórmula (6) se llama *determinante* de la matriz cuadrada A de orden tres. Luego, por definición,

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \alpha_{11}\alpha_{22}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32} + \alpha_{12}\alpha_{23}\alpha_{31} - \alpha_{12}\alpha_{21}\alpha_{33} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31}. \tag{7}$$

Uniendo en el segundo miembro los términos que contienen α_{11}, α_{12} y α_{13} y recordando la fórmula (3), obtenemos

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \alpha_{11}\begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12}\begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13}\begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix}. \tag{8}$$

Es fácil memorizar la fórmula (8). Para abreviar, en lugar de *determinante de una matriz* de orden dos o tres suele decirse *determinante* de segundo o tercer orden. Los tres determinantes de segundo orden de la fórmula (8) se obtienen suprimiendo del determinante del orden tres, que figura en la misma fórmula, la primera fila y, respectivamente, la primera, la segunda y la tercera columna. A continuación, el determinante de segundo orden que se obtiene suprimiendo la primera fila y la j-ésima columna debe ser multiplicado por el elemento que se halla en la primera fila y en la j-ésima columna y los productos obtenidos deben ser tomados con signos alternos y sumados. Como resultado obtendremos el determinante de orden tres.

Esta regla nos sugiere la idea de cómo debe ser definido el concepto de determinante de una matriz cuadrada de orden cuatro, cinco y de órdenes superiores. Por lo tanto, introducimos la siguiente definición principal:

DEFINICIÓN. *Se llama determinante de una matriz de primer orden, formada por el número α, el propio número α. Supongamos ahora que para un número natural cualquiera $n \geqslant 1$ conocemos ya el elemento del anillo K que representa el determinante de una matriz arbitraria cuadrada de orden n sobre K. Entonces, para una matriz cuadrada arbitraria $A = \| \alpha_{ij} \|$ de orden $n+1$ sobre K tomamos, por definición,*

$$\det \| \alpha_{ij} \| = \alpha_{11} | A_1^1 | - \alpha_{12} | A_1^2 | + \alpha_{13} | A_1^3 | - \ldots + (-1)^n \alpha_{1,\, n+1} | A_1^{n+1} |, \tag{9}$$

donde $| A_1^j |$ es el valor del determinante de la matriz de orden n que se obtiene de la matriz inicial A suprimiendo la primera fila y la j-ésima columna ($j = 1, \ldots, n+1$).

Aplicando esta definición en el caso en que $n = 1$, obtenemos la fórmula (3) para determinantes de orden dos. Conociendo la expresión para los determinantes de orden dos, podemos emplear la definición principal para obtener la expresión (8) para los determinantes de orden tres. De la expresión (8) obtenemos mediante (3) la fórmula definitiva (7).

Veamos ahora cuál es la fórmula definitiva para los determinantes de orden cuatro. De acuerdo con la definición principal, el determinante de una matriz cuadrada arbitraria $A = \| \alpha_{ij} \|$ de orden cuatro coincide con la expresión

$$\alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{42} & \alpha_{43} & \alpha_{44} \end{vmatrix} - \alpha_{12} \begin{vmatrix} \alpha_{21} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{43} & \alpha_{44} \end{vmatrix} + \alpha_{13} \begin{vmatrix} \alpha_{21} & \alpha_{22} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{44} \end{vmatrix} - \\ - \alpha_{14} \begin{vmatrix} \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} \end{vmatrix} \tag{10}$$

Introduciendo aquí las expresiones de los determinantes de orden tres según la fórmula (7) y suprimiendo los paréntesis, obtendremos la fórmula definitiva que buscabamos para un determinante de orden cuatro. No la escribiremos puesto que no tiene sentido memorizarla. Según (7) todo determinante de tercer orden es igual a la suma de seis términos tomados con signos alternados. Por eso, si tomamos en (10) en lugar de los determinantes de orden tres sus expresiones y suprimimos los paréntesis, obtendremos en total $4 \cdot 6 = 24$ términos. La mitad de ellos aparecerá con el signo más y la otra mitad con el signo menos. Una afirmación análoga es válida también para los determinantes de orden cualquiera.

TEOREMA 1. *Para una matriz cuadrada arbitraria se tiene*

$$\begin{vmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \dots & \alpha_{nn} \end{vmatrix} = \sum \pm \alpha_{1i_1}\alpha_{2i_2}\dots\alpha_{ni_n}, \tag{11}$$

donde la suma se extiende a las permutaciones arbitrarias $(i_1, i_2, \dots, i_n)$ *de los números* $1, 2, \dots, n$. *El signo más o menos se toma según sea par o impar la permutación* $(i_1, i_2, \dots, i_n)$, *es decir, en la mitad de los casos se toma el signo más y en la otra mitad se toma el signo menos; el número total de términos en la suma* (11) *es igual a* $1 \cdot 2 \cdot \ldots \cdot n = n!$.

Para $n = 1, 2, 3, 4$ la veracidad del teorema ya ha sido comprobada. Ahora aplicamos la inducción. Supongamos que el teorema es cierto para los determinantes de un orden cualquiera n y sea $A = \|\alpha_{ij}\|$ una matriz cuadrada de orden $n+1$. Por hipótesis, cada determinante $|A_1^j|$ $(j = 1, \dots, n+1)$ de la fórmula (9) es de la forma:

$$|A_1^j| = \sum \pm \alpha_{2m_1}\alpha_{3m_2}\dots\alpha_{n+1,\, m_n}, \tag{12}$$

donde la suma se extiende a todas las permutaciones $(m_1, \dots, m_n)$ de los números $1, \dots, j-1, j+1, \dots, n+1$. El número de términos en la suma (12) es igual a $n!$. Sustituyendo en (9) $|A_1^j|$ por sus expresiones y suprimiendo los paréntesis, obtendremos un total de $n!\,(n+1) = (n+1)!$ términos. La mitad de ellos tendrá el signo más y la otra mitad el signo menos. No habrá términos semejantes, puesto que los términos que se obtienen al suprimir distintos paréntesis difieren en el primer factor. Es evidente que el término arbitrario de tipo $\alpha_{1i_1}\alpha_{2i_2}\dots\alpha_{n+1,\, i_{n+1}}$ se obtiene al suprimir los paréntesis en el producto $\alpha_{1i_1}|A_1^{i_1}|$ y, por consiguiente, la fórmula (11) es verídica.

De la fórmula (11) se puede deducir fácilmente el siguiente corolario importante. Supongamos que los elementos de la matriz $A = \|\alpha_{ij}\|$ son números complejos. En virtud de la fórmula (11), tenemos

$$|\overline{A}| = \sum \pm \bar{\alpha}_{1i_1}\dots\bar{\alpha}_{ni_n} = \overline{\sum \pm \alpha_{1i_1}\dots\alpha_{ni_n}} = \overline{|A|},$$

es decir,

$$|\overline{A}| = \overline{|A|}.$$

Según la fórmula (11) el valor del determinante de una matriz es igual a una suma algebraica de términos, cada uno de los cuales es el producto de elementos, tomados de manera que haya uno de cada fila y uno de cada columna. Por ello, si todos los elementos de una fila o de una columna de la matriz son iguales a cero, también serán iguales a cero todos los términos del determinante. Es decir, hemos obtenido el siguiente corolario:

COROLARIO. *Si una fila o una columna de una matriz cuadrada está compuesta íntegramente por ceros, su determinante es igual a cero.*

Una matriz cuadrada se llama *semidescompuesta*, si sus elementos pueden ser divididos mediante una línea vertical y otra horizontal en cuatro matrices de manera que a lo largo de la diagonal figuren matrices cuadradas y una de las otras dos matrices esté compuesta íntegramente por ceros. En otras palabras, la matriz A es semidescompuesta si tiene una de las dos formas siguientes

$$\begin{bmatrix} \alpha_{11} & \dots & \alpha_{1r} & \alpha_{1,\,r+1} & \dots & \alpha_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{r1} & \dots & \alpha_{rr} & \alpha_{r,\,r+1} & \dots & \alpha_{rn} \\ 0 & \dots & 0 & \alpha_{r+1,\,r+1} & \dots & \alpha_{r+1,\,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \dots & 0 & \alpha_{n,\,r+1} & \dots & \alpha_{nn} \end{bmatrix},$$

$$\begin{bmatrix} \alpha_{11} & \dots & \alpha_{1r} & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{r1} & \dots & \alpha_{rr} & 0 & \dots & 0 \\ \alpha_{r+1,\,1} & \dots & \alpha_{r+1,\,r} & \alpha_{r+1,\,r+1} & \dots & \alpha_{r+1,\,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{n1} & \dots & \alpha_{nr} & \alpha_{n,\,r+1} & \dots & \alpha_{nn} \end{bmatrix}.$$

A veces, la primera de estas matrices suele llamarse matriz semidescompuesta *superior* y la segunda, matriz semidescompuesta *inferior*.

TEOREMA 2. *El determinante de una matriz semidescompuesta es igual al producto de los determinantes de sus células diagonales.*

Para una matriz de orden dos esta proposición es evidente, ya que

$$\begin{vmatrix} \alpha & \beta \\ 0 & \delta \end{vmatrix} = \begin{vmatrix} \alpha & 0 \\ \gamma & \delta \end{vmatrix} = \alpha\delta.$$

Apliquemos ahora la inducción aceptando que el teorema 2 es cierto para las matrices de un orden cualquiera $n-1$. Consideremos una matriz arbitraria semidescompuesta A de orden n. Supongamos que es de la forma

$$A = \|\alpha_{ij}\| = \begin{bmatrix} B & D \\ O_{sr} & C \end{bmatrix},$$

donde B y C son matrices cuadradas de orden r y s respectivamente $(r+s=n)$ mientras que O_{sr} y D son matrices rectangulares y la matriz O_{sr} es nula. Aplicando la fórmula (9), obtenemos

$$|A| = \alpha_{11}|A_1^1| - \alpha_{12}|A_1^2| + \dots + (-1)^{n-1}\alpha_{1n}|A_1^n|, \qquad (13)$$

donde $|A_1^i|$ es la matriz de orden $n-1$ que se obtiene suprimiendo en A la primera fila y la i-ésima columna. Es fácil ver que todas las matrices A_1^i son semidescompuestas y, por ello, según la suposición inductiva, el determinante de cada una de ellas es igual al producto de los determinantes de las células diagonales. Pero las matrices $A_1^1, \ldots, A_1^r$ y las matrices $A_1^{r+1}, \ldots, A_1^n$ se descomponen de modos distintos. Para las primeras tenemos

$$|A_1^i| = |B_1^i| \cdot |C|, \quad i = 1, \ldots, r, \tag{14}$$

mientras que las matrices $A_1^{r+1}, \ldots, A_1^n$ pueden ser divididas mediante una línea vertical y otra horizontal de manera que el cuadrado superior de la izquierda sea de orden r. Puesto que la última fila de este cuadrado está compuesta de ceros, su determinante es igual a cero. El determinante de la matriz A_1^{r+i} es igual al producto del determinante del cuadrado indicado por el determinante del cuadrado complementario. Luego, $|A_1^{r+i}| = 0$ y, por consiguiente,

$$\begin{aligned} |A| &= \alpha_{11}|B_1^1| \cdot |C| - \ldots + (-1)^{r-1}\alpha_{1r}|B_1^r| \cdot |C| = \\ &= (\alpha_{11}|B_1^1| - \ldots + (-1)^{r-1}\alpha_{1r}|B_1^r|) \cdot |C| = |B| \cdot |C|, \end{aligned} \tag{15}$$

que es lo que queríamos demostrar. El caso en que la matriz A es de la forma

$$A = \|\alpha_{ij}\| = \begin{bmatrix} B & O_{rs} \\ D & C \end{bmatrix},$$

es aun más sencillo, ya que ahora se tiene $\alpha_{1,r+1} = \ldots = \alpha_{1n} = 0$ y por esto de (13) y (14) obtenemos inmediatamente (15).

TEOREMA 3 *Si en las matrices cuadradas $A = \|\alpha_{ij}\|$ y $B = \|\beta_{ij}\|$ de un mismo orden n coinciden todos los elementos correspondientes menos los elementos de una fila i-ésima cualquiera, se tiene*

$$\begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{i1} & \ldots & \alpha_{in} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix} + \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \beta_{i1} & \ldots & \beta_{in} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{i1}+\beta_{i1} & \ldots & \alpha_{in}+\beta_{in} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix}. \tag{16}$$

Este teorema suele llamarse a veces teorema de adición de determinantes. Pasando a su demostración, designemos mediante C la matriz del segundo miembro de la igualdad (16). Para matrices de primer orden la igualdad (16) es evidente. Aplicando la inducción, supongamos que para matrices de orden $n-1$ el teorema 3 es cierto. Si en la fórmula (16) $i=1$, desarrollando el determinante de la matriz C según la fórmula (9), obtenemos

$$|C| = (\alpha_{11}+\beta_{11})|C_1^1| - \ldots + (-1)^{n-1}(\alpha_{1n}+\beta_{1n})|C_1^n|.$$

Es evidente que $C_1^i = A_1^i = B_1^i$ y, por consiguiente, se tiene

$$|C| = (\alpha_{11}|A_1^1| - \ldots + (-1)^{n-1}\alpha_{1n}|A_1^n|) +$$
$$+ (\beta_{11}|B_1^1| - \ldots + (-1)^{n-1}\beta_{1n}|B_1^n|) = |A| + |B|.$$

Supongamos ahora que en la fórmula (16) $i > 1$. Entonces, se tiene

$$|C| = \alpha_{11}|C_1^1| - \ldots + (-1)^{n-1}\alpha_{1n}|C_1^n|. \tag{17}$$

Las matrices $C_1^1, \ldots, C_1^n$ son de orden $n-1$ y por esto para ellas es válido el desarrollo de tipo (16). Como resultado, obtenemos

$$|C_1^i| = |A_1^i| + |B_1^i| \quad (i = 1, \ldots, n)$$

y, en virtud de la relación (17), tenemos $|C| = |A| + |B|$.

Desde el punto de vista formal, la definición principal de determinante sirve también para matrices formadas por elementos de un anillo asociativo K cualquiera (no necesariamente conmutativo). En la demostración de los teoremas 1, 2 y 3 tampoco se ha empleado la conmutatividad de K. Sin embargo, una serie de propiedades de los determinantes, de importancia para las aplicaciones, dependen de la conmutatividad del anillo principal K. Una de estas propiedades se indica en el siguiente teorema.

TEOREMA 4. *Sea A una matriz cuadrada formada por elementos de un anillo conmutativo K. Si se multiplican todos los elementos de una fila de la matriz A por un elemento $\lambda \in K$, el determinante de la matriz también se multiplicará por λ. En otras palabras, se tiene*

$$\begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \lambda\alpha_{i1} & \ldots & \lambda\alpha_{in} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix} = \lambda \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{i1} & \ldots & \alpha_{in} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix}.$$

La demostración se realiza igual que en el teorema anterior y no la vamos a repetir. Señalemos solamente el siguiente corolario importante:

COROLARIO. *Para cualquier matriz cuadrada A de orden n formada por elementos de un anillo conmutativo K y para cualquier $\lambda \in K$ se tiene*

$$|\lambda A| = \lambda^n |A|.$$

En efecto, al multiplicar la matriz A por λ, cada una de sus n filas se multiplicará por λ. Por consiguiente, el determinante de la matriz A se multiplicará por λ^n.

2.2. Propiedades principales de los determinantes. Consideremos una matriz cuadrada arbitraria $A = \|\alpha_{ij}\|$ de orden n formada por elementos de un anillo cualquiera K. Según la definición principal, se tiene

$$|A| = \alpha_{11}|A_1^1| - \alpha_{12}|A_1^2| + \ldots + (-1)^{n+1}\alpha_{1n}|A_1^n|, \qquad (1)$$

donde A_1^j es la matriz que se obtiene de A suprimiendo la primera fila y la j-ésima columna. Aplicando ahora al determinante de la matriz A_1^j la misma fórmula, obtenemos

$$\begin{aligned}|A_1^j| = \alpha_{21}|A_{12}^{j1}| - \ldots + (-1)^j \alpha_{2,\,j-1}|A_{12}^{j,\,j-1}| + \\ + (-1)^{j+1}\alpha_{2,\,j+1}|A_{12}^{j,\,j+1}| + \ldots + (-1)^n \alpha_{2n}|A_{12}^{jn}|,\end{aligned} \qquad (2)$$

donde A_{12}^{jk} es la matriz que se obtiene de A suprimiendo la primera y la segunda filas y la j-ésima y la k-ésima columnas ($j \neq k$).

Introduzcamos en la fórmula (1) en lugar de los valores $|A_1^j|$ sus valores (2). Como resultado, llegamos a la relación

$$|A| = \sum_{j \neq k} \pm \alpha_{1j}\alpha_{2k}|A_{12}^{jk}|. \qquad (3)$$

Aquí la suma se extiende a todos los pares posibles j, k de diferentes números, pertenecientes al conjunto $1, \ldots, n$ y el signo más o menos se toma de acuerdo con las fórmulas (1) y (2). Las matrices A_{12}^{jk} y A_{12}^{kj} coinciden evidentemente y, por lo tanto, la fórmula (3) puede ser representada en la forma siguiente:

$$|A| = \sum_{j<k} (\pm \alpha_{1j}\alpha_{2k} \pm \alpha_{1k}\alpha_{2j})|A_{12}^{jk}|. \qquad (4)$$

Calculemos ahora con mayor exactitud cuáles son los signos que deben tomarse en la última fórmula. En virtud de (2), para el j-ésimo término de (1) tenemos

$$(-1)^{j+1}\alpha_{1j}|A_1^j| = \ldots + (-1)^{j+1}\alpha_{1j}(-1)^k \alpha_{2k}|A_{12}^{jk}| + \ldots$$

Análogamente y teniendo en cuenta que $j < k$, obtenemos para el k-ésimo término de (1)

$$(-1)^{k+1}\alpha_{1k}|A_1^k| = \ldots + (-1)^{k+1}\alpha_{1k}(-1)^{j+1}\alpha_{2j}|A_{12}^{jk}| + \ldots$$

Por consiguiente, el coeficiente del término $|A_{12}^{jk}|$ de la fórmula (4) es igual a

$$(-1)^{j+k+1}(\alpha_{1j}\alpha_{2k} - \alpha_{1k}\alpha_{2j}) = (-1)^{j+k+1+2}\begin{vmatrix}\alpha_{1j} & \alpha_{1k}\\ \alpha_{2j} & \alpha_{2k}\end{vmatrix}$$

y, en consecuencia,

$$|A| = \sum_{j<k} (-1)^{j+k+1+2}\begin{vmatrix}\alpha_{1j} & \alpha_{1k}\\ \alpha_{2j} & \alpha_{2k}\end{vmatrix} \cdot |A_{12}^{jk}|. \qquad (5)$$

La fórmula (5) se denomina desarrollo del determinante según los elementos de la primera y segunda filas. Es fácil memorizarla:

se toman todos los determinantes posibles de orden dos formados por los elementos que se hallan en la primera y segunda filas y en las columnas j-ésima y k-ésima ($j < k$) y se multiplican por los determinantes de las matrices correspondientes que se obtienen al suprimir en la matriz A las filas y columnas indicadas. Los productos se multiplican además por $(-1)^{j+k+1+2}$, donde el exponente es igual a la suma de los números que corresponden a las filas y a las columnas suprimidas, y después se suman. La suma algebraica así obtenida es igual al determinante de la matriz dada.

Reglas semejantes tienen lugar también para los desarrollos según las tres primeras, las cuatro primeras, etc. filas. Pero en lo sucesivo no las necesitaremos.

Emplearemos ahora la fórmula (5) para deducir una serie de propiedades principales de los determinantes. En lo que sigue se supone que los elementos de las matrices consideradas se toman de un anillo conmutativo K.

TEOREMA 1. *Si en una matriz cuadrada se cambian entre sí dos filas cualesquiera, el determinante de la matriz nueva será igual al determinante de la matriz inicial tomado con el signo menos.*

Para matrices de segundo orden esta proposición se comprueba directamente:

$$\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \alpha\delta - \beta\gamma, \qquad \begin{vmatrix} \gamma & \delta \\ \alpha & \beta \end{vmatrix} = \gamma\beta - \delta\alpha = -(\alpha\delta - \beta\gamma).$$

A continuación, suponemos por inducción que el teorema es justo para las matrices de orden $n-1$ y que la matriz dada $A = \|\alpha_{ij}\|$ es de orden n. Supongamos que en A se cambian entre sí las dos primeras filas. De la fórmula (5) vemos que todos los determinantes de orden dos cambian de signo, mientras que los factores adicionales no varían. Luego, toda la suma adquiere el factor -1 que es lo que queríamos demostrar.

Consideremos el caso en el que se cambian entre sí la j-ésima y la k-ésima filas, donde $1 < j < k$. Entonces, la primera fila permanece invariable y del desarrollo (1) deducimos que cada factor $|A_1^j|$ obtendrá después del intercambio el valor opuesto; por ello, toda la suma obtendrá después del intercambio de las filas el valor opuesto.

Finalmente, supongamos que se cambian entre sí la primera y la i-ésima fila, donde $i > 2$. Este mismo resultado obtendremos si cambiamos entre sí primero la primera y la segunda filas, luego la segunda y la i-ésima y, finalmente, la segunda y la primera. Según hemos demostrado el determinante cambiará cada vez de signo y después de los tres cambios el determinante se multiplicará por -1.

COROLARIO 1. *El determinante de una matriz cuadrada en la que coinciden dos filas es igual a cero.*

Cambiando las filas entre sí se puede conseguir que las fillas coincidentes sean la primera y la segunda. El determinante de la matriz con las filas cambiadas o bien coincidirá con el determinante de la matriz inicial o bien diferirá de él en el factor -1. De la fórmula (5) se ve directamente que el determinante de una matriz en la que coinciden las dos primeras filas es igual a cero. Por esto será también igual a cero el determinante de la matriz inicial.

COROLARIO 2. *Si a los elementos de la r-ésima fila cualquiera de la matriz cuadrada $A=\|\alpha_{ij}\|$ se agregan los elementos correspondientes de otra fila s-ésima cualquiera $(s\neq r)$ multiplicados por un factor arbitrario λ, el determinante de la matriz nueva será igual al determinante de la inicial.*

En efecto, según los teoremas 3 y 4 del punto anterior, se tiene

$$\begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{r1} & \ldots & \alpha_{rn} \\ \cdot & \cdot & \cdot \\ \alpha_{s1}+\lambda\alpha_{r1} & \ldots & \alpha_{sn}+\lambda\alpha_{rn} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{r1} & \ldots & \alpha_{rn} \\ \cdot & \cdot & \cdot \\ \alpha_{s1} & \ldots & \alpha_{sn} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix} + \lambda \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{r1} & \ldots & \alpha_{rn} \\ \cdot & \cdot & \cdot \\ \alpha_{r1} & \ldots & \alpha_{rn} \\ \cdot & \cdot & \cdot \\ \alpha_{n1} & \ldots & \alpha_{nn} \end{vmatrix}.$$

Aquí el último determinante tiene dos filas coincidentes y, por consiguiente, es igual a cero.

Hasta el momento todos nuestros resultados estaban relacionados con las filas de los determinantes. Hagamos el primer paso para introducir en el juego las columnas.

TEOREMA 2. *Para toda matriz cuadrada $A=\|\alpha_{ij}\|$ de orden n es válido el siguiente desarrollo según los elementos de la primera columna:*

$$|A|=\alpha_{11}|A_1^1|-\alpha_{21}|A_2^1|+\ldots+(-1)^{n-1}\alpha_{n1}|A_n^1|, \tag{6}$$

donde A_r^1 es la matriz que se obtiene de A suprimiendo la primera columna y la r-ésima fila.

Consideremos que los elementos α_{ij} de la matriz dada son letras. Entonces, el determinante de la matriz A será un polinomio en estas letras cuya forma general se ha establecido en el teorema 1 del punto anterior. En particular, hemos señalado allí que cada término del polinomio $|A|$ contiene un factor, y sólo uno, perteneciente al conjunto $\alpha_{11}, \alpha_{21}, \ldots, \alpha_{n1}$. Agrupemos en $|A|$ todos los términos que contienen el factor α_{r1}, saquemos fuera de los paréntesis este factor común y designemos mediante A_{r1} la expresión comprendida entre los paréntesis. De esta forma obtendremos

$$|A|=\alpha_{11}A_{11}+\alpha_{21}A_{21}+\ldots+\alpha_{n1}A_{n1}. \tag{7}$$

Comparando las fórmulas (7) y (1), llegamos a la conclusión de que

$$A_{11} = |A_1^1|.$$

Para hallar la expresión análoga para A_{r1}, es suficiente recurrir ahora al teorema 1. En efecto, cambiemos sucesivamente la r-ésima fila de la matriz A con cada una de las anteriores, elevándola más y más. Después de $r-1$ intercambios de esta índole obtendremos una matriz B que difiere de la matriz A sólo en el orden de filas y, por consiguiente, tendremos

$$|A| = (-1)^{r-1}|B|.$$

Desarrollando el determinante B según los elementos de la primera fila, obtenemos

$$|B| = \alpha_{r1}|B_1^1| - \alpha_{r2}|B_1^2| + \ldots + (-1)^{n-1}\alpha_{rn}|B_1^n|,$$

y por lo tanto

$$|A| = \alpha_{r1}(-1)^{r-1}|B_1^1| - \ldots + (-1)^{r+n-2}\alpha_{rn}|B_1^n|.$$

Comparando este desarrollo con el de la fórmula (7), llegamos a la relación

$$A_{r1} = (-1)^{r-1}|B_1^1|.$$

La matriz B_1^1 se obtiene suprimiendo en la matriz B la primera fila y la primera columna. Está claro que la misma matriz se obtendrá al suprimir en A la r-ésima fila y la primera columna, es decir, que $B_1^1 = A_r^1$. Sustituyendo en la fórmula (7) el valor A_{r1} por el valor $(-1)^{r-1}|A_r^1|$, obtenemos el desarrollo deseado (6).

Anteriormente hemos indicado que la matriz, en la que la primera, la segunda, etc. filas coinciden respectivamente con la primera, la segunda, etc. columnas de la matriz A, se llama *transpuesta* respecto de A y se designa por A'. Es evidente, que siendo A una matriz cuadrada de orden n, la matriz A' es también una matriz cuadrada de orden n.

TEOREMA 3. *El determinante de una matriz cuadrada no varía en la transposición, es decir,*

$$|A'| = |A|. \tag{8}$$

Para matrices de orden uno la proposición es evidente. Recurriendo, al igual que antes, a la inducción, aceptemos que la matriz dada $A = \|\alpha_{ij}\|$ es de orden $n > 1$ y que el teorema es cierto para las matrices de orden $n-1$. Desarrollando el determinante de la matriz A según los elementos de la primera fila y el determinante de la matriz transpuesta A' según los elementos de la primera columna, obtenemos

$$|A| = \alpha_{11}|A_1^1| - \alpha_{12}|A_1^2| + \ldots + (-1)^{n-1}\alpha_{1n}|A_1^n|, \tag{9}$$
$$|A'| = \alpha_{11}|(A')_1^1| - \alpha_{12}|(A')_2^1| + \ldots + (-1)^{n-1}\alpha_{1n}|(A')_n^1|.$$

Sin embargo, es fácil ver que $(A')_i^1=(A_1^i)'$. Las matrices A_1^i son de orden $n-1$ y, por la hipótesis de inducción, se tiene

$$|(A')_i^1|=|(A_1^i)'|=|A_1^i| \qquad (i=1,\ \ldots,\ n).$$

Comparando estas relaciones con los desarrollos (9) obtenemos (8).

Por consiguiente, al calcular el determinante de una matriz las columnas y las filas pueden ser sustituidas unas por otras. Esto significa que de todo teorema referente a las propiedades del determinante de una matriz, enunciado en términos de filas o de columnas, se puede obtener un nuevo teorema sustituyendo las filas y las columnas unas por otras. En particular, de las propiedades de los determinantes indicadas anteriormente y relacionadas con las filas, obtenemos el resultado siguiente.

COROLARIO. *Al cambiar entre sí dos columnas de una matriz su determinante adquiere el valor opuesto. El determinante de una matriz cuadrada que tiene dos columnas idénticas es igual a cero. Si se multiplican todos los elementos de una columna de la matriz por λ, el determinante de la matriz también quedará multiplicado por λ. Si se agregan a todos los elementos de una columna de la matriz los elementos correspondientes de otra columna, multiplicados por un número fijo, el determinante de la matriz nueva será igual al determinante de la inicial.*

Análogamente, sustituyendo las filas por las columnas obtenemos del desarrollo (5) el desarrollo según las dos primeras columnas:

$$|A|=\sum_{i<j}(-1)^{i+j+1+2}\begin{vmatrix}\alpha_{i1} & \alpha_{i2}\\ \alpha_{j1} & \alpha_{j2}\end{vmatrix}\cdot|A_{ij}^{12}|.$$

Finalmente, del teorema de adición de determinantes, enunciado en términos de filas, obtenemos la fórmula correspondiente

$$\begin{vmatrix}\alpha_{11} & \ldots & \alpha_{1i} & \ldots & \alpha_{1n}\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \alpha_{n1} & \ldots & \alpha_{ni} & \ldots & \alpha_{nn}\end{vmatrix}+\begin{vmatrix}\alpha_{11} & \ldots & \beta_{1i} & \ldots & \alpha_{1n}\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \alpha_{n1} & \ldots & \beta_{ni} & \ldots & \alpha_{nn}\end{vmatrix}=$$
$$=\begin{vmatrix}\alpha_{11} & \ldots & \alpha_{1i}+\beta_{1i} & \ldots & \alpha_{1n}\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \alpha_{n1} & \ldots & \alpha_{ni}+\beta_{ni} & \ldots & \alpha_{nn}\end{vmatrix}$$

referente a las columnas.

Hasta el momento nos hemos valido de los desarrollos de un determinante según los elementos de su *primera* fila o su *primera* columna. Al mismo tiempo, conocemos la ley de variación del determinante al cambiar en él entre sí filas o columnas. Esto ofrece la posibilidad de obtener inmediatamente de los desarrollos según los elementos de la primera fila o la primera columna los desarrollos análogos según los elementos de cualquier fila o columna.

TEOREMA 4. *Para una matriz cuadrada arbitraria $A=\|\alpha_{ij}\|$ de orden n son válidos los siguientes desarrollos según los elementos de la r-ésima fila y s-ésima columna:*

$$|A|=(-1)^{r+1}\alpha_{r1}|A_r^1|+\ldots+(-1)^{r+n}\alpha_{rn}|A_r^n|, \tag{10}$$
$$|A|=(-1)^{s+1}\alpha_{1s}|A_1^s|+\ldots+(-1)^{s+n}\alpha_{ns}|A_n^s|, \tag{11}$$

donde A_i^j es la matriz que se obtiene de A suprimiendo la i-ésima fila y la j-ésima columna.

Puesto que las filas y las columnas se encuentran en las mismas condiciones basta demostrar sólo una de las fórmulas (10) y (11), por ejemplo, la fórmula (10). Cambiando sucesivamente la r-ésima fila de la matriz A con cada una de las anteriores, después de $r-1$ intercambios obtendremos la matriz

$$B=\begin{vmatrix} \alpha_{r1} & \alpha_{r2} & \ldots & \alpha_{rn} \\ \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{r-1,1} & \alpha_{r-1,2} & \ldots & \alpha_{r-1,n} \\ \alpha_{r+1,1} & \alpha_{r+1,2} & \ldots & \alpha_{r+1,n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix}.$$

Según el teorema sobre el intercambio de filas tenemos $|A|=$ $=(-1)^{r-1}|B|$. Desarrollando aquí el determinante B según los elementos de la primera fila y empleando las relaciones evidentes $B_1^i=A_r^i$, llegamos al desarrollo (10).

El determinante de la matriz A_i^j se llama *menor* del determinante de la matriz A correspondiente al elemento α_{ij}. La expresión $(-1)^{i+j}|A_i^j|$ se llama *adjunto* del elemento α_{ij} en $|A|$ y se designa con frecuencia mediante $|A|_{ij}$. Empleando el concepto de adjunto, las fórmulas (10) y (11) pueden ser representadas en forma más breve:

$$\alpha_{i1}|A|_{i1}+\alpha_{i2}|A|_{i2}+\ldots+\alpha_{in}|A|_{in}=|A|, \tag{12}$$
$$\alpha_{1i}|A|_{1i}+\alpha_{2i}|A|_{2i}+\ldots+\alpha_{ni}|A|_{ni}=|A|. \tag{13}$$

En estas igualdades cada elemento α_{ij} se multiplica por su adjunto $|A|_{ij}$. ¿Qué sucederá si se toma la suma de los productos de los elementos de la i-ésima fila por los adjuntos de los elementos correspondientes de otra fila cualquiera? Con el fin de obtener la respuesta, sustituyamos la j-ésima fila de la matriz dada A por su i-ésima fila sin alterar todas las filas restantes, incluyendo también la i-ésima. Obtendremos una matriz B en la que son idénticas las filas i-ésima y j-ésima. El determinante de esta matriz es igual a cero. Al mismo tiempo, es evidente que los menores de los elementos de la j-ésima fila en los determinantes de las matrices

A y B coinciden. Desarrollando B según los elementos de la i-ésima fila, obtenemos

$$\alpha_{i1}|A|_{j1}+\ldots+\alpha_{in}|A|_{jn}=0 \qquad (i\neq j). \tag{14}$$

Sustituyendo en los razonamientos las filas por las columnas obtenemos la segunda serie de relaciones:

$$\alpha_{1i}|A|_{1j}+\ldots+\alpha_{ni}|A|_{nj}=0 \qquad (i\neq j). \tag{15}$$

Las relaciones (12), (13), (14) y (15) suelen enunciarse de la forma siguiente.

TEOREMA 5. *La suma de los productos de los elementos de una fila (de una columna) de un determinante por sus adjuntos es igual al valor del determinante. La suma de los productos de los elementos de una fila (de una columna) de un determinante por los adjuntos de los elementos correspondientes de otra fila (columna) cualquiera es igual a cero.*

Hagamos ahora algunas observaciones acerca de los métodos de cálculo de determinantes. Los determinantes de orden dos y tres se calculan generalmente mediante las fórmulas iniciales. Por ejemplo,

$$\begin{vmatrix} 7 & -2 & 5 \\ 2 & 4 & 1 \\ 3 & -1 & 8 \end{vmatrix} = 7\begin{vmatrix} 4 & 1 \\ -1 & 8 \end{vmatrix} + 2\begin{vmatrix} 2 & 1 \\ 3 & 8 \end{vmatrix} + 5\begin{vmatrix} 2 & 4 \\ 3 & -1 \end{vmatrix} = 187.$$

Aquí hemos desarrollado el determinante de orden tres según los elementos de su primera fila. Si uno de los elementos del determinante dado fuese igual a cero, resultaría más conveniente recurrir al desarrollo según aquella fila (o columna) que contenga este elemento nulo. Este mismo método se puede aplicar al calcular determinantes de orden elevado que contienen muchos elementos iguales a cero. En particular, tenemos

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ 0 & \alpha_{22} & \ldots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \alpha_{nn} \end{vmatrix} = \alpha_{11}\begin{vmatrix} \alpha_{22} & \ldots & \alpha_{2n} \\ \cdot & \cdot & \cdot \\ 0 & \ldots & \alpha_{nn} \end{vmatrix} = \alpha_{11}\alpha_{22}\ldots\alpha_{nn}$$

y análogamente

$$\begin{vmatrix} 0 & 0 & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \alpha_{n-1,2} & \ldots & \alpha_{n-1,n} \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{vmatrix} = (-1)^{n+1}\alpha_{n1}\begin{vmatrix} 0 & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{n-1,2} & \ldots & \alpha_{n-1,n} \end{vmatrix} =$$

$$= (-1)^{\frac{n(n-1)}{2}}\alpha_{n1}\ldots\alpha_{1n}.$$

La situación es más compleja si el orden del determinante dado es relativamente elevado (por ejemplo, 7 y más) y entre los elementos del determinante hay pocos ceros o no los hay. En este caso, al determinante se aplican primero las transformaciones indicadas en el corolario 2 del teorema 1, tratando de escoger el coeficiente λ de modo que el determinante nuevo contenga en algunos lugares ceros o tenga en cierto sentido una estructura más sencilla. Estos métodos se ven con claridad en el ejemplo siguiente. Supongamos que debemos calcular el determinante

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

Restando en este determinante la quinta fila de la sexta, restando en el determinante obtenido la cuarta fila de la quinta, restando después la tercera fila de la cuarta, etc., obtenemos

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

Restando ahora la sexta fila de la primera, de la segunda, ..., de la quinta, obtenemos

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

En el determinante obtenido de orden cinco agregamos a la primera columna las restantes y después desarrollamos el determinante según los elementos de la primera columna. Tendremos

$$d=-6^4\cdot\begin{vmatrix} 21 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{vmatrix}=-6^4\cdot 21.$$

Mediante cálculos análogos se obtiene también la fórmula general

$$\begin{vmatrix} 1 & 2 & 3 & \dots & n-1 & n \\ 2 & 3 & 4 & \dots & n & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ n & 1 & 2 & \dots & n-2 & n-1 \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} \frac{n(n+1)}{2} n^{n-2}. \qquad (16)$$

Este determinante lleva el nombre de *circulador* de orden n, debido a que sus filas se obtienen por permutaciones cíclicas de los elementos de la primera fila. Es esta regularidad en la disposición de los elementos del determinante en la que se basa la deducción de la fórmula (16) que, a propósito, no será empleada en lo sucesivo. Actualmente para hallar el valor numérico de los determinantes de orden elevado se recurre a los servicios de los centros de cómputo en los que existen programas standard para el cálculo de determinantes adaptados a aquellos tipos de máquinas con las que está equipado el centro. Hallemos la estimación aproximada del número de operaciones aritméticas suficiente indudablemente para el cálculo del determinante de una matriz arbitraria $\|\alpha_{ij}\|$ de orden n.

Supongamos que $\alpha_{11} \neq 0$. Calculamos α_{11}^{-1} y representamos el determinante dado en la forma

$$\alpha_{11} \cdot \begin{vmatrix} 1 & \alpha_{11}^{-1}\alpha_{12} & \dots & \alpha_{11}^{-1}\alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix}.$$

Durante esta operación realizamos una inversión y $n-1$ multiplicaciones. A continuación, agregamos a los elementos de la segunda, de la tercera, ..., de la n-ésima filas los elementos de la primera fila multiplicados respectivamente por $-\alpha_{21}$, $-\alpha_{31}$, ..., $-\alpha_{n1}$ (un total de $(n-1)^2$ multiplicaciones e igual número de adiciones). Después de esto el problema se reduce al cálculo de un determinante de orden $n-1$:

$$\alpha_{11} \cdot \begin{vmatrix} \alpha'_{11} & \dots & \alpha'_{1,\, n-1} \\ \cdot & \cdot & \cdot \\ \alpha'_{n-1,1} & \dots & \alpha'_{n-1,\, n-1} \end{vmatrix}.$$

Continuando este proceso, tendremos que calcular al final del mismo solamente el producto $\alpha_{11}\alpha'_{11} \dots \alpha_{11}^{(n-1)}$ (es decir, $n-1$ multiplicaciones). Al aplicar el algoritmo expuesto necesitaremos realizar un total de $n-1$ inversiones de números, de

$$n(n-1) + (n-1)(n-2) + \dots + 2 \cdot 1 = \frac{(n+1)\, n\, (n-1)}{3}$$

multiplicaciones, de $\frac{(2n-1)\, n\, (n-1)}{6}$ adiciones y de $n-1$ multiplicaciones finales. Por lo tanto, para calcular un determinante de orden n es necesario efectuar aproximadamente cerca de $\frac{1}{3}n^3$ multiplicaciones y de un número igual de adiciones. No nos detendremos aquí en el estudio más detallado de esta cuestión. Varias definiciones y resultados exactos se indican en el artículo de *В. Я. Пан*, „О способах вычисления значений многочленов", Успехи математических наук **21**, № 1 (1966), 103—134 (V. Ya. Pan, Acerca de los métodos de cálculo de valores de polinomios), así como en la bibliografía señalada en este artículo.

2.3. Determinante de un producto. Matrices inversas. Un papel importante lo desempeña en la teoría de matrices el siguiente teorema.

TEOREMA 1 (sobre la multiplicación de determinantes). *El determinante del producto de matrices cuadradas (formadas por elementos pertenecientes a un anillo conmutativo K) es igual al producto de los determinantes de las matrices.*

Sean dadas dos matrices cuadradas $A=\|\alpha_{ij}\|$ y $B=\|\beta_{ij}\|$ de orden n. En virtud del teorema sobre las matrices semidescompuestas (p. 2.1), se tiene

$$|A|\cdot|B|=\begin{vmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & 0 & 0 & \dots & 0 \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ 0 & -1 & \dots & 0 & \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & -1 & \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{vmatrix}.$$

En el segundo miembro aparece el determinante de una matriz de tipo especial de orden $2n$. Sin alterar el valor del determinante, con esta matriz se pueden realizar sucesivamente las siguientes transformaciones: agregar a su primera fila los elementos de la $(n+1)$-ésima fila multiplicados por α_{11}, los elementos de la $(n+2)$-ésima fila multiplicados por α_{12}, etc., los elementos de la $2n$-ésima fila multiplicados por α_{1n}. Así surgirá una matriz de orden $2n$, en la que las primeras n posiciones de la primera fila estarán ocupadas por ceros, mientras que las otras n posiciones estarán ocupadas por los productos de la primera fila de la matriz A por las columnas de la matriz B. En la nueva matriz de orden $2n$ agregamos ahora a los elementos de su segunda fila los elementos de la $(n+1)$-ésima fila, de la $(n+2)$-ésima fila, etc., multiplicados respectivamente por α_{21}, α_{22}, ..., α_{2n}. Después realizamos transformaciones análogas con la tercera, ..., con la n-ésima filas.

Después de esto obtenemos la igualdad siguiente:

$$|A|\cdot|B|=\begin{vmatrix} 0 & \ldots & 0 & \sum\alpha_{1i}\beta_{i1} & \ldots & \sum\alpha_{1i}\beta_{in} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & 0 & \sum\alpha_{ni}\beta_{i1} & \ldots & \sum\alpha_{ni}\beta_{in} \\ -1 & \ldots & 0 & \beta_{11} & \ldots & \beta_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \ldots & -1 & \beta_{n1} & \ldots & \beta_{nn} \end{vmatrix}.$$

Para reducir el determinante de la derecha a la forma semidescompuesta, cambiamos en él la primera columna con la $(n+1)$-ésima, la segunda con la $(n+2)$-ésima, ..., la n-ésima con la $2n$-ésima. Así obtendremos la igualdad

$$|A|\cdot|B|=(-1)^n\begin{vmatrix} \sum\alpha_{1i}\beta_{i1} & \ldots & \sum\alpha_{1i}\beta_{in} & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sum\alpha_{ni}\beta_{i1} & \ldots & \sum\alpha_{ni}\beta_{in} & 0 & \ldots & 0 \\ \beta_{11} & \ldots & \beta_{1n} & -1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \beta_{n1} & \ldots & \beta_{nn} & 0 & \ldots & -1 \end{vmatrix}. \qquad (1)$$

Puesto que el determinante de una matriz semidescompuesta es igual al producto de los determinantes de las células diagonales, de la relación (1) se deduce que

$$|A|\cdot|B|=(-1)^n(-1)^n\begin{vmatrix} \sum\alpha_{1i}\beta_{i1} & \ldots & \sum\alpha_{1i}\beta_{in} \\ \cdot & \cdot & \cdot \\ \sum\alpha_{ni}\beta_{i1} & \ldots & \sum\alpha_{ni}\beta_{in} \end{vmatrix},$$

es decir,

$$|A|\cdot|B|=|AB|, \qquad (2)$$

que es lo que queríamos demostrar.

Hemos demostrado el teorema 1 para el producto de dos matrices. Está claro que de aquí se desprende su validez también en el caso del producto de un número finito cualquiera de matrices. Por ejemplo,

$$|ABC|=|(AB)C|=|AB|\cdot|C|=|A|\cdot|B|\cdot|C|.$$

En particular, para cualquier matriz cuadrada A se tiene

$$|A^k|=|A|^k \quad (k=0, 1, 2, \ldots). \qquad (3)$$

Sabemos que la transposición de una matriz cuadrada no altera su determinante. Por ello, para dos matrices cuadradas cualesquiera de un mismo orden, se tiene

$$|A|\cdot|B|=|AB|=|A'B|=|AB'|=|A'B'|.$$

Sea, por ejemplo,

$$A = \begin{vmatrix} x & y \\ u & v \end{vmatrix}.$$

En este caso, se tiene

$$|A|^2 = |AA'| = \begin{vmatrix} x^2+y^2 & xu+yv \\ xu+yv & u^2+v^2 \end{vmatrix} = (x^2+y^2)(u^2+v^2)-(xu+yv)^2,$$

es decir,

$$(xv-yu)^2 = (x^2+y^2)(u^2+v^2)-(xu+yv)^2.$$

Esta relación suele escribirse en forma de la siguiente identidad

$$(x^2+y^2)(u^2+v^2) = (xv-yu)^2+(xu+yv)^2$$

que se denomina *identidad de Lagrange*.

Empleemos ahora el teorema 1 para un estudio más detallado de las propiedades de la matriz inversa.

Consideremos una matriz cuadrada cualquiera $A = \|\alpha_{ij}\|$ de orden n. Formemos una matriz con los adjuntos de los elementos del determinante A y tomemos su transpuesta. La matriz así obtenida se llama *adjunta* de A y se designa mediante A^*. En otras palabras,

$$A^* = \begin{bmatrix} |A|_{11} & |A|_{21} & \dots & |A|_{n1} \\ |A|_{12} & |A|_{22} & \dots & |A|_{n2} \\ . & . & . & . \\ |A|_{1n} & |A|_{2n} & \dots & |A|_{nn} \end{bmatrix}.$$

Calculando los productos $A \cdot A^*$ y $A^* \cdot A$ según la regla de multiplicación de matrices y teniendo en cuenta las fórmulas de (12) a (15) del punto anterior, obtenemos directamente las relaciones importantes

$$A \cdot A^* = A^* \cdot A = |A| \cdot E, \tag{4}$$

donde E es la matriz unidad.

El determinante de la matriz A es un elemento del anillo principal K, del cual se toman los elementos de todas las matrices consideradas. Supongamos que para el elemento $|A|$ existe en K su inverso $|A|^{-1}$. Entonces, multiplicando las relaciones (4) por $|A|^{-1}$, llegamos a las igualdades

$$A \cdot |A|^{-1} A^* = |A|^{-1} A^* \cdot A = E,$$

de donde se deduce que

$$A^{-1} = |A|^{-1} \cdot A^*. \tag{5}$$

Por otra parte, si la ecuación matricial

$$AX = E$$

tiene solución, obtenemos pasando a los determinantes

$$|A|\cdot|X|=1,$$

es decir, el elemento $|A|$ es invertible en el anillo K.

Uniendo los resultados obtenidos, llegamos a la proposición siguiente.

TEOREMA 2. *Una matriz cuadrada A, formada por elementos de un anillo conmutativo K con el elemento unidad, posee la matriz inversa, también formada por elementos de K, si, y sólo si, el determinante de A es invertible en el anillo K. Si la matriz inversa existe, su determinante es igual al inverso del valor del determinante de la matriz dada.*

Por ejemplo, en el anillo de todos los números enteros $K=\{0, \pm 1, \ldots\}$ poseen inverso solamente los elementos 1 y -1. Por lo tanto, *una matriz cuadrada formada por números enteros posee la matriz inversa, también formada por números enteros, si, y sólo si, el determinante de la matriz dada es igual a ± 1.*

Supongamos que se consideran matrices formadas por elementos de un cuerpo conmutativo. En un cuerpo conmutativo, todo elemento no nulo posee su inverso. Por ello, sobre un cuerpo conmutativo son invertibles aquellas matrices, y sólo aquéllas, cuyos determinantes son diferentes de cero.

Una matriz cuadrada cuyo determinante es igual a cero se llama *singular*. En el caso contrario, la matriz se llama *regular*. Por esto, la condición indicada anteriormente para que una matriz sea invertible puede ser enunciada de la forma siguiente: *una matriz cuadrada sobre un cuerpo conmutativo posee inversa si, y sólo si, es regular.*

Hemos visto que para matrices A regulares sobre el cuerpo conmutativo K, se tiene

$$|A^{-1}|=|A|^{-1}.$$

De aquí obtenemos

$$|A^{-m}|=|A^{-1}\ \ldots\ A^{-1}|=|A|^{-m} \qquad (m=1, 2, \ldots),$$

de modo que la fórmula (3) es válida no sólo para los valores naturales de k, sino también para cualesquiera valores enteros de k (siempre que la matriz A sea invertible).

Según el p.1.3, una matriz cuadrada A se llama ortogonal, si

$$AA'=A'A=E.$$

Pasando a los determinantes, obtenemos $|A|^2=1$ y, por ello, si las matrices se consideran sobre un cuerpo conmutativo K, se tiene $|A|=\pm 1$. Gracias a esto todas las matrices ortogonales se descomponen en dos conjuntos de matrices: las matrices *ortogonales propias* de determinante $+1$ y las matrices *ortogonales impropias* de determinante -1. Todas las matrices ortogonales propias forman, respecto a la operación de multiplicación, un grupo que constituye un subgrupo del grupo de todas las matrices ortogonales.

Recordemos que una matriz cuadrada A definida sobre el cuerpo de los números complejos se llama unitaria si $A\bar{A}'=E$. Tomando los determinantes, obtenemos $|A|\cdot\overline{|A|}=1$, es decir, *el módulo del determinante de una matriz unitaria es igual a la unidad.*

Para concluir hallemos el determinante de la matriz adjunta. Aceptemos que el anillo K es el anillo de los polinomios en letras x_{ij} $(i, j=1, \ldots, n)$ con coeficientes enteros y consideremos la matriz $X=\|x_{ij}\|$. Según la fórmula (4) se tiene

$$X\cdot X^*=|X|\cdot E$$

y por lo tanto

$$|X|\cdot|X^*|=|X|^n.$$

Puesto que el polinomio $|X|$ es diferente de cero, la igualdad puede ser dividida por $|X|$ y así obtenemos la fórmula deseada

$$|X^*|=|X|^{n-1}, \tag{6}$$

ambos miembros de la cual son polinomios en las variables x_{ij} con coeficientes enteros. Pero, si coinciden dos polinomios de este tipo, también coinciden sus valores para cualesquiera valores de las variables pertenecientes a un anillo conmutativo arbitrario. Por esto, la fórmula (6) es válida para matrices cuadradas arbitrarias sobre cualquier anillo conmutativo.

2.4. Sistemas cramerianos de ecuaciones lineales. Consideremos el conjunto de condiciones

$$\left.\begin{array}{c} \alpha_{11}\xi_1+\ldots+\alpha_{1n}\xi_n=\beta_1, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \alpha_{m1}\xi_1+\ldots+\alpha_{mn}\xi_n=\beta_m, \end{array}\right\} \tag{1}$$

donde $\xi_1, \ldots, \xi_n$ son letras, mientras que α_{ij} y β_i son elementos de un anillo K. El conjunto de estas condiciones se llama *sistema de ecuaciones lineales* con las incógnitas $\xi_1, \ldots, \xi_n$ sobre el anillo K. Una sucesión $\xi_1^0, \ldots, \xi_n^0$ de elementos del anillo K se llama *solución* del sistema (1) si al tomar en las condiciones (1) en lugar de las variables $\xi_1, \ldots, \xi_n$ los elementos correspondientes $\xi_1^0, \ldots, \xi_n^0$ todas las condiciones resultan verídicas. Introduciendo las matrices

$$A=\begin{bmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot & \cdot \\ \alpha_{m1} & \ldots & \alpha_{mn} \end{bmatrix}, \quad x=\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \quad \text{y} \quad b=\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix},$$

el sistema (1) puede ser representado en la siguiente forma matricial

$$Ax=b. \tag{2}$$

La matriz A se llama matriz del sistema (1) y la matriz

$$B=\begin{bmatrix}\alpha_{11} & \dots & \alpha_{1n} & \beta_1\\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{m1} & \dots & \alpha_{mn} & \beta_m\end{bmatrix}$$

se llama matriz *ampliada* del sistema (1). Las propiedades de la matriz B nos harán falta en el p. 5.3 al estudiar de forma más detallada las propiedades del sistema (1); por ahora nos limitaremos a considerar el caso en que $m=n$, es decir, en que la matriz A es cuadrada. Supongamos que la matriz A es invertible. En este caso, multiplicando ambos miembros de la igualdad (2) a la izquierda por A^{-1}, obtenemos

$$x=A^{-1}b. \tag{3}$$

Viceversa, multiplicando la relación (3) a la izquierda por A, obtenemos (2). Por consiguiente, las condiciones (2) y (3) son equivalentes y por ello la fórmula (3) puede ser considerada como la fórmula que ofrece la solución del sistema de ecuaciones (1) siempre que la matriz del sistema sea invertible.

Hasta el momento no hemos supuesto siquiera que el anillo K sea conmutativo. Supongamos ahora que K es un cuerpo conmutativo. En este caso, el hecho de que la matriz A sea invertible equivale a que sea regular; la matriz inversa se expresará en términos de la matriz adjunta en la forma anteriormente señalada y la fórmula (3) podrá ser representada de la forma siguiente

$$x=|A|^{-1}A^*b.$$

Tomando en lugar de la matriz A^* su expresión en términos de los adjuntos del determinante $|A|$ y realizando la multiplicación en el segundo miembro, obtenemos

$$\xi_i=\frac{|A|_{1i}\beta_1+\ldots+|A|_{ni}\beta_n}{|A|}$$

o bien, en forma definitiva,

$$\xi_i=\frac{1}{|A|}\begin{vmatrix}\alpha_{11} & \dots & \alpha_{1,i-1} & \beta_1 & \alpha_{1,i+1} & \dots & \alpha_{1n}\\ \alpha_{21} & \dots & \alpha_{2,i-1} & \beta_2 & \alpha_{2,i+1} & \dots & \alpha_{2n}\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \alpha_{n1} & \dots & \alpha_{n,i-1} & \beta_n & \alpha_{n,i+1} & \dots & \alpha_{nn}\end{vmatrix} (i=1, \ldots, n). \tag{4}$$

Para comprobar bastará desarrollar el determinante que figura en el numerador según los elementos de la i-ésima columna.

Un sistema de ecuaciones lineales sobre un cuerpo conmutativo se llama *crameriano* si el número de ecuaciones coincide con el número de incógnitas y si el determinante de la matriz del sistema es diferente de cero. Vemos, por consiguiente, que *todo sistema crameriano de ecuaciones lineales tiene una solución, y sólo una,* $(\xi_1, \ldots, \xi_n)$ *y que esta solución viene dada por las fórmulas* (4). Las fórmulas (4) se llaman *fórmulas de Cramer*, en memoria de Cramer, matemático de mediados del siglo pasado, ya que aparecieron, por lo visto, por primera vez en uno de sus trabajos. Para $n=2$ y $n=3$ las fórmulas de Cramer han sido dadas de forma detallada en el p. 2.1.

Las fórmulas de Cramer dan una expresión «cerrada» de la solución de un sistema crameriano de ecuaciones en términos de los coeficientes de estas ecuaciones. Sin embargo, no es conveniente hallar la solución de un sistema lineal empleando las fórmulas de Cramer. En efecto, recurriendo a las fórmulas de Cramer, habrá que calcular los valores de $n+1$ determinantes de orden n. Según las estimaciones aproximadas, realizadas en el p. 2.2, habrá que efectuar para ello cerca de n^4 multiplicaciones y adiciones. Al mismo tiempo, el método corriente de eliminación sucesiva de incógnitas requiere, como ahora veremos, un número de operaciones considerablemente menor. Este método, ordenado de modo adecuado, se llama en la actualidad *algoritmo de Gauss*.

Consideremos, pues, de nuevo el sistema de ecuaciones lineales (1) cuyos coeficientes han sido tomados de un cuerpo conmutativo K. Si todos los coeficientes de una de las ecuaciones son iguales a cero podemos suprimirla del sistema. Si los coeficientes de las incógnitas son iguales a cero, mientras que $\beta \neq 0$, el sistema no tiene solución. Por esto se puede aceptar que algunos de los coeficientes de las incógnitas en la primera ecuación son diferentes de cero. En este caso, cambiando si es necesario la numeración de las incógnitas, podemos llevar el sistema a la forma (1) en la que $\alpha_{11} \neq 0$. Calculemos α_{11}^{-1} y multipliquemos todos los miembros de la primera ecuación por α_{11}^{-1}. A continuación, de la segunda, de la tercera, ..., de la m-ésima ecuaciones restamos término por término la primera ecuación multiplicada por $\alpha_{21}, \alpha_{31}, \ldots, \alpha_{m1}$, respectivamente. Después de esta operación nuestro sistema inicial resultará ser equivalente al sistema de tipo

$$\left.\begin{aligned} \xi_1 + \alpha'_{12}\xi_2 + \ldots + \alpha'_{1n}\xi_n &= \beta'_1, \\ \alpha'_{22}\xi_2 + \ldots + \alpha'_{2n}\xi_n &= \beta'_2, \\ \cdots\cdots\cdots\cdots\cdots & \\ \alpha'_{m2}\xi_2 + \ldots + \alpha'_{mn}\xi_n &= \beta'_m. \end{aligned}\right\} \tag{5}$$

Al pasar del sistema (1) al sistema (5) hemos realizado una inversión, nm multiplicaciones y $n(m-1)$ adiciones. Procediendo de la misma forma con las últimas $m-1$ ecuaciones del sistema (5), obtendremos — después de una inversión, de $(n-1)(m-1)$ multiplicaciones y de $(n-1)(m-2)$ adiciones — el sistema de tipo

$$\left.\begin{array}{r}\xi_1+\alpha'_{12}\xi_2+\alpha'_{13}\xi_3+\ldots+\alpha'_{1n}\xi_n=\beta'_1,\\ \xi_2+\alpha''_{23}\xi_3+\ldots+\alpha''_{2n}\xi_n=\beta''_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ \alpha''_{m3}\xi_3+\ldots+\alpha''_{mn}\xi_n=\beta''_m.\end{array}\right\}$$

Continuando el proceso indicado, podremos encontrarnos con tres casos: a) en uno de los pasos llegamos a la conclusión de que el sistema no tiene solución; b) obtendremos un sistema de tipo

$$\left.\begin{array}{r}\xi_1+\gamma_{12}\xi_2+\ldots+\gamma_{1r}\xi_r+\gamma_{1,\,r+1}\xi_{r+1}+\ldots+\gamma_{1n}\xi_n=\delta_1,\\ \xi_2+\ldots+\gamma_{2r}\xi_r+\gamma_{2,\,r+1}\xi_{r+1}+\ldots+\gamma_{2n}\xi_n=\delta_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \xi_r+\gamma_{r,\,r+1}\xi_{r+1}+\ldots+\gamma_{rn}\xi_n=\delta_n;\end{array}\right\}$$

c) obtendremos un sistema de tipo

$$\left.\begin{array}{r}\xi_1+\gamma_{12}\xi_2+\gamma_{13}\xi_3+\ldots+\gamma_{1n}\xi_n=\delta_1,\\ \xi_2+\gamma_{23}\xi_3+\ldots+\gamma_{2n}\xi_n=\delta_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ \xi_{n-1}+\gamma_{n-1,\,n}\xi_n=\delta_{n-1},\\ \xi_n=\delta_n.\end{array}\right\}\qquad(6)$$

En el caso b), expresamos ξ_r, empleando la última ecuación, en términos de las incógnitas «libres» $\xi_{r+1}, \ldots, \xi_n$. Introduciendo el valor obtenido de ξ_r en la ecuación anterior, hallamos de ella ξ_{r-1}, etc. De esta forma encontraremos unas expresiones lineales de las incógnitas $\xi_1, \ldots, \xi_r$ en términos de las incógnitas «libres». Dando valores arbitrarios a las incógnitas «libres», podremos, a partir de las fórmulas señaladas, calcular los valores correspondientes de las incógnitas $\xi_1, \ldots, \xi_r$ y obtener de esta forma una solución $\xi_1, \ldots, \xi_r$, $\xi_{r+1}, \ldots, \xi_n$ del sistema dado. Lo mismo ocurrirá en el caso c); la diferencia consistirá solamente en que en este caso obtendremos unos valores determinados para todas las incógnitas $\xi_1, \ldots, \xi_n$.

Calculemos el número de operaciones que hay que efectuar en el caso c) para obtener la solución. Para reducir el sistema a la forma (6) (aceptando que $m=n$) tendremos que realizar n inversiones.

$$n^2+(n-1)^2+\ldots+1^2=\frac{1}{6}(2n+1)(n+1)n$$

multiplicaciones y

$$n(n-1)+(n-1)(n-2)+\ldots+2\cdot 1=\frac{1}{3}(n+1)n(n-1)$$

adiciones. Además, para hallar del sistema triangular (6) las incógnitas habrá que realizar adicionalmente

$$1+2+\ldots+(n-1)=\frac{1}{2}n(n-1)$$

multiplicaciones y un número igual de adiciones. Por consiguiente, para resolver el sistema (1) empleando el método de Gauss, es necesario realizar, en el caso general, n inversiones,

$$\frac{1}{3}n(n^2+3n-1)\approx\frac{1}{3}n^3$$

multiplicaciones y

$$\frac{1}{6}n(2n^2+3n-5)\approx\frac{1}{3}n^3$$

adiciones.

Observemos ahora que empleando el método de Gauss puede ser calculada también la matriz inversa A^{-1}. Esto se logra de la forma siguiente. Se escribe el sistema de ecuaciones lineales $Ax=b$, en el que las columnas x y b se consideran formadas por letras. Después, a partir del sistema dado y empleando el método de Gauss, las variables $\xi_1, \ldots, \xi_n$ se expresan en términos de las variables $\beta_1, \ldots, \beta_n$. Así se obtienen fórmulas de tipo

$$\xi_i=\gamma_{i1}\beta_1+\ldots+\gamma_{in}\beta_n \quad (i=1, \ldots, n).$$

Puesto que de $Ax=b$ se deduce que $x=A^{-1}b$, la matriz $\|\gamma_{ij}\|$ será precisamente la inversa de A. El cálculo del número de operaciones muestra que en este caso es suficiente realizar n inversiones, $\frac{1}{3}n(4n^2-1)\approx\frac{4}{3}n^3$ multiplicaciones y $\frac{2}{3}(n^2-n)(2n-1)\approx\frac{4}{3}n^3$ adiciones.

En los cálculos que hemos realizado acerca del número de operaciones necesarias para hallar el valor del determinante de una matriz, la solución de un sistema de ecuaciones lineales o la inversión de una matriz, no se han tenido en cuenta algunas operaciones secundarias, así como el aumento del número de cifras decimales en los números que se multiplican o se suman. Sin embargo, los resultados obtenidos pueden resultar útiles al decidir, si un sistema de ecuaciones lineales de interés práctico debe ser resuelto a mano o conviene pasar el encargo a un centro de cálculo.

Complementos y ejercicios

1. Empleando las fórmulas de Cramer, resuélvase el sistema

$$\left.\begin{aligned} x+y+z&=-2,\\ x+3y-2t&=-4,\\ 2x+z-t&=-1,\\ 2y-z-3t&=-3.\end{aligned}\right\}$$

2. Empleando el método de Gauss hállense las inversas de las matrices

$$\begin{bmatrix}1 & 2 & -3\\ 2 & -3 & 5\\ 3 & -1 & 1\end{bmatrix} \quad \text{y} \quad \begin{bmatrix} a+1 & a & a & a\\ a & a+1 & a & a\\ a & a & a+1 & a\\ a & a & a & a+1\end{bmatrix}.$$

3. Demuéstrese la fórmula (determinante de Vandermonde)

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

4. Demuéstrese que

$$\begin{vmatrix} a+x_1 & a & \dots & a \\ a & a+x_2 & \dots & a \\ \dots & \dots & \dots & \dots \\ a & a & \dots & a+x_n \end{vmatrix} = x_1 x_2 \dots x_n \left(1+\frac{a}{x_1}+\dots+\frac{a}{x_n}\right).$$

5. Si en el anillo principal K para cualquier α de $2\alpha=0$ se deduce que $\alpha=0$, el determinante de cualquier matriz antisimétrica sobre K de orden impar es igual a cero.

6. Consideremos una matriz cuadrada arbitraria A (sobre un anillo conmutativo K) de orden n. Suprimiendo en A las filas con los números $i_1, \dots, i_r$ $(1 \leqslant i_1 < \dots < i_r \leqslant n)$ y las columnas con los números $j_1, \dots, j_r$ $(1 \leqslant j_1 < < \dots < j_r \leqslant n)$, obtendremos una matriz nueva $A_{i_1 \dots i_r}^{j_1 \dots j_r}$ de orden $n-r$. Los elementos de la matriz A que se hallan en los cruces de las filas y de las columnas suprimidas también forman una matriz cuadrada. El orden de esta última matriz es igual a r. Su determinante se denomina *menor* de orden r del determinante de la matriz A perteneciente a las filas con los números $i_1, \dots, i_r$ y a las columnas con los números $j_1, \dots, j_r$. Se llama *adjunto* de este menor el determinante de la matriz $A_{i_1 \dots i_r}^{j_1 \dots j_r}$ multiplicado por $(-1)^{s+t}$, donde $s=i_1+ + \dots + i_r$ y $t=j_1+\dots+j_r$. Tiene lugar la siguiente generalización del desarrollo de un determinante según los elementos de la primera y de la segunda filas: *si en un determinante de orden n se fijan r filas cualesquiera, la suma de los productos de los menores de orden r, pertenecientes a las filas escogidas, por sus adjuntos es igual al valor del determinante dado.*

7. Si la matriz cuadrada dada X es simétrica o antisimétrica, la matriz adjunta X^* también es simétrica o antisimétrica.

8. Sea A una matriz cuadrada de elementos enteros y sean x y b unas columnas, cuyos elementos son letras: $\xi_1, \dots, \xi_n$ y $\beta_1, \dots, \beta_n$, respectivamente. El sistema de ecuaciones lineales $Ax=b$ tiene soluciones enteras $\xi_1, \dots, \xi_n$ para cualesquiera valores enteros de $\beta_1, \dots, \beta_n$ si, y sólo si, $|A| = \pm 1$.

§ 3. Polinomios característico y mínimo

Todas las matrices que se consideran en este parrágrafo se suponen cuadradas y de un mismo orden n. Los elementos de estas matrices se toman de un cuerpo conmutativo K cualquiera, pero fijo.

3.1. Semejanza de matrices. La matriz A se llama *semejante* a la matriz B, si existe una matriz no degenerada X, tal, que

$$A = X^{-1}BX. \tag{1}$$

En este caso también suele decirse que la matriz A es *la transformada de B por X*. Multiplicando la igualdad (1) a la izquierda por X y a la derecha por X^{-1}, obtenemos

$$B = XAX^{-1} = (X^{-1})^{-1} AX^{-1}.$$

Por consiguiente, *la semejanza de A con B implica la semejanza de B con A*. Además, si

$$A = X^{-1}BX \quad \text{y} \quad B = Y^{-1}CY,$$

realizando la sustitución obtenemos

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

Por consiguiente, *dos matrices, semejantes cada una a otra tercera, son también semejantes*. Es obvio, finalmente, que toda matriz es semejante a sí misma.

Estas propiedades demuestran que el conjunto de todas las matrices cuadradas de un orden dado n, formadas por elementos de un cuerpo conmutativo K, se descompone de un modo natural en clases de matrices semejantes. Uno de los problemas centrales de la teoría de matrices es el de hallar las condiciones necesarias y suficientes de semejanza de matrices. La solución de este problema será dada en el capítulo IV. Aquí estableceremos solamente algunas propiedades previas de matrices semejantes.

Para transformar una suma de matrices por X es suficiente transformar por X cada sumando.

En efecto,

$$X^{-1}(A_1 + A_2 + \ldots + A_k)X = X^{-1}A_1X + X^{-1}A_2X + \ldots + X^{-1}A_kX.$$

Para transformar un producto de matrices por X es suficiente transformar cada factor

Efectivamente,

$$X^{-1}A_1X \cdot X^{-1}A_2X \ \ldots \ X^{-1}A_kX = X^{-1}A_1A_2 \ \ldots \ A_kX,$$

ya que los productos XX^{-1}, que aparecen en el primer miembro, son iguales a E y pueden ser omitidos.

Para transformar una potencia de una matriz es suficiente transformar la base de la potencia, es decir, se tiene

$$X^{-1}A^mX = (X^{-1}AX)^m.$$

Si $m \geqslant 0$, esta fórmula es un caso particular de la anterior. Si $m < 0$, tomemos $k = -m$. Entonces tendremos

$$X^{-1}A^mX = X^{-1}(A^{-1})^k X = (X^{-1}A^{-1}X)^k = (X^{-1}AX)^{-k} = (X^{-1}AX)^m.$$

El valor transformado de un polinomio en una matriz es igual al valor del polinomio en la matriz transformada, en otras palabras,

$$X^{-1}f(A)X = f(X^{-1}AX).$$

Esta proposición es resultado inmediato de las anteriores, ya que el valor de un polinomio en A se obtiene a partir de A realizando operaciones de elevación a potencias, de multiplicación por número y de adición.

Las reglas de la transformación de expresiones permiten en muchas ocasiones simplificar considerablemente los cálculos. Sean, por ejemplo,

$$A=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} \quad \text{y} \quad X=\begin{bmatrix}2 & 1\\ 3 & 2\end{bmatrix}.$$

Por inducción es fácil demostrar que

$$A^n=\begin{bmatrix}1 & n\\ 0 & 1\end{bmatrix}.$$

Puesto que

$$X^{-1}AX=\begin{bmatrix}7 & 4\\ -9 & -5\end{bmatrix},$$

tendremos, aplicando la regla de transformación de una potencia,

$$\begin{bmatrix}7 & 4\\ -9 & -5\end{bmatrix}^n=X^{-1}A^nX=\begin{bmatrix}1+6n & 4n\\ -9n & 1-6n\end{bmatrix}.$$

3.2. Polinomio característico. Sea A una matriz cuadrada y sean α_{ij} $(i,\ j=1,\ 2,\ \dots,\ n)$ sus elementos. La matriz

$$\lambda E-A=\begin{bmatrix}\lambda-\alpha_{11} & -\alpha_{12} & \dots & -\alpha_{1n}\\ -\alpha_{21} & \lambda-\alpha_{22} & \dots & -\alpha_{2n}\\ \dots & \dots & \dots & \dots\\ -\alpha_{n1} & -\alpha_{n2} & \dots & \lambda-\alpha_{nn}\end{bmatrix},$$

donde λ es una variable independiente, se llama *matriz característica* de A. Su determinante

$$\varphi(\lambda)=|\lambda E-A| \tag{1}$$

es evidentemente un polinomio en λ y se llama *polinomio característico* de la matriz A.

Para hallar los primeros términos de este polinomio recurramos a que el valor de un determinante es igual a la suma de los productos de sus elementos tomados de modo que haya uno de cada fila y uno de cada columna y provistos de signo adecuado. Para hallar el término de mayor grado respecto a λ es necesario tomar los productos de los elementos de mayor grado. Es nuestro caso este producto será uno solamente, a saber, el producto $(\lambda-\alpha_{11})(\lambda-\alpha_{22})\dots$ $\dots(\lambda-\alpha_{nn})$ de los elementos diagonales. Todos los demás productos que componen el determinante serán de grado no mayor que $n-2$, ya que siendo $-\alpha_{ij}$ $(i\neq j)$ uno de los factores de alguno de estos productos, el último no contendrá los factores $\lambda-\alpha_{ii}$ y $\lambda-\alpha_{jj}$ y, por consiguiente, será de grado no mayor que $n-2$. De este modo, $\varphi(\lambda)=(\lambda-\alpha_{11})\ \dots\ (\lambda-\alpha_{nn})+$ términos de grado no mayor que $n-2$, o

$$\varphi(\lambda)=\lambda^n-(\alpha_{11}+\dots+\alpha_{nn})\lambda^{n-1}+\dots \tag{2}$$

La suma de los elementos diagonales de una matriz se llama *traza* de la matriz. Según la fórmula (2), *el grado del polinomio característico de una matriz es igual al orden de esta matriz, el coeficiente principal del polinomio característico es igual a 1, mientras que el coeficiente de λ^{n-1} es igual a la traza de la matriz tomada con el signo contrario.* Si en la fórmula (2) tomamos $\lambda = 0$, tendremos

$$\varphi(0) = |-A| = (-1)^n |A|.$$

Pero $\varphi(0)$ es el término independiente del polinomio característico. Es decir, el término independiente del polinomio característico de una matriz A es igual al determinante de esta matriz multiplicado por $(-1)^n$, donde n es el orden de la matriz A.

El teorema que sigue describe una de las propiedades más importantes del polinomio característico.

TEOREMA 1. *Los polinomios característicos de matrices semejantes coinciden.*

En efecto, sea A una matriz semejante a la matriz B:

$$A = X^{-1}BX.$$

Entonces para el polinomio característico de A tenemos

$$|\lambda E - A| = |\lambda E - X^{-1}BX| = |X^{-1}(\lambda E - B)X| = |X^{-1}|\cdot|\lambda E - B|\cdot|X|.$$

Los determinantes $|X^{-1}|$ y $|X|$ son números recíprocos cuyo producto es igual a la 1; por esto, se tiene

$$|\lambda E - B| = |\lambda E - A|$$

que es lo que se quería demostrar.

De este teorema se deduce, en particular, que *las matrices semejantes tienen trazas y determinantes iguales*, ya que la traza y el determinante de una matriz, provistos de signo adecuado, son coeficientes de su polinomio característico.

La igualdad de los polinomios característicos es una condición necesaria pero, como regla general, no suficiente de la semejanza de matrices. Por ejemplo, los polinomios característicos de las matrices

$$E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

coinciden. Sin embargo, A no puede ser semejante a E ya que para cualquier matriz X se tiene

$$X^{-1}EX = X^{-1}X = E.$$

Las raíces del polinomio característico de una matriz se llaman *números característicos* o *valores propios* de la misma. Las raíces múltiples del polinomio característico se llaman valores propios múltiples de la matriz. Es conocido que la suma de todas las raíces reales y complejas de un polinomio de grado n, cuyo coeficiente principal es la 1, es igual al coeficiente de la $(n-1)$-ésima potencia de la variable tomado con el signo contrario. De la fórmula (2) resulta, por ello, que *en el cuerpo de los números complejos la suma de todos los valores propios de una matriz es igual a su traza.*

Observemos que la traza de una suma de matrices es igual a la suma de las trazas de los sumandos y que la traza del producto de un número por una matriz es igual al producto de este número por la traza de la matriz. Ambas afirmaciones pueden ser resumidas en una fórmula:

$$\text{traza}\,(\alpha A+\beta B)=\alpha\cdot\text{traza}\,A+\beta\cdot\text{traza}\,B$$

para cuya demostración basta considerar las matrices correspondientes y calcular sus trazas.

En el p. 1.2 a todo polinomio $\varphi(\lambda)$ ha sido puesta en correspondencia una matriz $\varphi(A)$ llamada valor del polinomio $\varphi(\lambda)$ para $\lambda=A$. Si $\varphi(A)=O$, se dice que A es *raíz* de $\varphi(\lambda)$.

TEOREMA DE HAMILTON-CAYLEY. *Toda matriz es raíz de su polinomio característico.*

Sea A una matriz. Sea B la matriz adjunta de la matriz característica $\lambda E-A$ (véase el p. 2.3). Sean β_{ij} $(i,\ j=1,\ 2,\ \ldots,\ n)$ los elementos de la matriz B. Estos elementos son los adjuntos del determinante $|\lambda E-A|$ y, por lo tanto, representan polinomios en λ de grado no superior a $n-1$. Sea

$$\beta_{ij}=\beta_{ij}^{(0)}+\beta_{ij}^{(1)}\lambda+\ldots+\beta_{ij}^{(n-1)}\lambda^{(n-1)}.$$

Consideremos las matrices numéricas auxiliares

$$B^{(k)}=\begin{bmatrix}\beta_{11}^{(k)} & \beta_{12}^{(k)} & \ldots & \beta_{1n}^{(k)}\\ \beta_{21}^{(k)} & \beta_{22}^{(k)} & \ldots & \beta_{2n}^{(k)}\\ \cdot & \cdot & \cdot & \cdot\\ \beta_{n1}^{(k)} & \beta_{n2}^{(k)} & \ldots & \beta_{nn}^{(k)}\end{bmatrix}\quad (k=0,\ 1,\ \ldots,\ n-1).$$

Es evidente que la matriz B puede ser representada en este caso en la forma siguiente:

$$B=B^{(0)}+\lambda B^{(1)}+\ldots+\lambda^{n-1}B^{(n-1)}.$$

En virtud de la propiedad principal de las matrices adjuntas, se tiene

$$B(\lambda E-A)=|\lambda E-A|\cdot E. \tag{3}$$

Aquí $|\lambda E-A|$ es el polinomio característico de la matriz A que representaremos mediante $\varphi(\lambda)$. Sea

$$\varphi(\lambda)=\alpha_0+\alpha_1\lambda+\ldots+\alpha_{n-1}\lambda^{n-1}+\lambda^n.$$

Ahora podemos escribir la igualdad (3) en forma más detallada:

$$(B^{(0)}+\lambda B^{(1)}+\ldots+\lambda^{n-1}B^{(n-1)})(\lambda E-A)=$$
$$=(\alpha_0+\alpha_1\lambda+\ldots+\alpha_{n-1}\lambda^{n-1}+\lambda^n)\cdot E.$$

Suprimiendo los paréntesis y comparando los coeficientes de potencias iguales de λ, obtenemos

$$\begin{aligned}
-B^{(0)}A \qquad\qquad &= \alpha_0 E,\\
-B^{(1)}A+B^{(0)} \quad &= \alpha_1 E,\\
-B^{(2)}A+B^{(1)} \quad &= \alpha_2 E,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
-B^{(n-1)}A+B^{(n-2)} &= \alpha_{n-1}E,\\
B^{(n-1)} &= E.
\end{aligned}$$

Multipliquemos a la derecha estas igualdades por E, A, ..., A^n, respectivamente, y sumémoslas. Todos los términos del primer miembro se anularán y obtendremos

$$O=\alpha_0E+\alpha_1A+\alpha_2A^2+\ldots+A^n,$$

es decir, $\varphi(A)=O$ que es lo que se quería.

3.3. Polinomio mínimo. Consideremos todos los polinomios no nulos $f(\lambda)$ para los cuales una matriz A dada es una raíz. Existen polinomios de este tipo; entre ellos figura, por ejemplo, el polinomio característico de la matriz A. El polinomio no nulo de menor grado y de coeficiente principal igual a la 1, para el cual la matriz A es una raíz, se llama *polinomio mínimo* de esta matriz.

Toda matriz A tiene sólo un polinomio mínimo. En efecto, si hubiese dos, digamos $\psi_1(\lambda)$ y $\psi_2(\lambda)$, la diferencia $\psi_1(\lambda)-\psi_2(\lambda)$ sería un polinomio no nulo de grado menor, para el cual la matriz A también sería una raíz. Dividiendo esta diferencia por su coeficiente principal, tendríamos un polinomio de coeficiente principal igual a la 1, la matriz A sería una de sus raíces y su grado sería inferior al de los polinomios mínimos $\psi_1(\lambda)$ y $\psi_2(\lambda)$, lo cual estaría en contradicción con la definición de los polinomios mínimos.

Todo polinomio $f(\lambda)$ para el cual la matriz A es una raíz es divisible por el polinomio mínimo $\psi(\lambda)$ de esta matriz.

Efectivamente, supongamos, al contrario, que $f(\lambda)$ no es divisible por $\psi(\lambda)$. Designando mediante $q(\lambda)$ el cociente y mediante $r(\lambda)$ el resto de la división de $f(\lambda)$ por $\psi(\lambda)$, tendremos

$$f(\lambda)=\psi(\lambda)\,q(\lambda)+r(\lambda).$$

Tomando aquí $\lambda=A$ y valiéndonos de que $\psi(A)=f(A)=O$, obtendremos $r(A)=O$. Pero el grado del resto $r(\lambda)$ es menor que el grado del divisor $\psi(\lambda)$. Por consiguiente, $r(\lambda)$ es un polinomio no nulo, la matriz A es su raíz y su grado es inferior al grado del polinomio mínimo $\psi(\lambda)$, lo cual es contradictorio. Hemos demostrado nuestra afirmación. En particular, *el polinomio mínimo de una matriz es un divisor de su polinomio característico.*

Sabemos que las matrices semejantes tienen un mismo polinomio característico. Esta misma propiedad la tiene el polinomio mínimo: las matrices semejantes tienen polinomios mínimos idénticos. En efecto, sea A semejente a B: $A=X^{-1}BX$. Si $f(\lambda)$ es un polinomio que tiene a B como raíz, tendremos en virtud del p. 3.1

$$f(A)=f(X^{-1}BX)=X^{-1}f(B)X=O.$$

Por consiguiente, el conjunto de polinomios que tienen como raíz una de las matrices semejantes coincide con el conjunto de polinomios que tienen como raiz otra de las matrices semejantes. Por esto, el polinomio de grado menor y de coeficiente principal igual a la 1, perteneciente a este conjunto, será el polinomio mínimo de ambas matrices.

La igualdad de los polinomios mínimos es una condición necesaria más de la semejanza de matrices. Sin embargo, esta condición tampoco es suficiente. Consideremos por ejemplo, las matrices

$$A=\begin{bmatrix}2&0&0\\0&3&0\\0&0&3\end{bmatrix}\quad \text{y}\quad B=\begin{bmatrix}2&0&0\\0&2&0\\0&0&3\end{bmatrix}.$$

Sus polinomios característicos son iguales, respectivamente, a

$$(\lambda-2)(\lambda-3)^2 \quad \text{y} \quad (\lambda-2)^2(\lambda-3).$$

Puesto que estos polinomios son diferentes, las matrices A y B no son semejantes. El polinomio mínimo de la matriz A debe ser divisor de su polinomio característico, es decir, debe coincidir con uno de los polinomios siguientes: $\lambda-2$, $\lambda-3$, $(\lambda-3)^2$, $(\lambda-2)(\lambda-3)$ o $(\lambda-2)(\lambda-3)^2$. Tomando aquí en lugar de λ la matriz A, encontraremos que el cero se obtiene por primera vez en el caso del polinomio $(\lambda-2)(\lambda-3)$. Por consiguiente, este polinomio será precisamente el polinomio mínimo de la matriz A. De la misma forma determinaremos que el polinomio mínimo de la matriz B es también el polinomio $(\lambda-2)(\lambda-3)$. Es decir, los polinomios mínimos de las matrices A y B coinciden, mientras que las matrices A y B no son semejantes.

En el p. 1.4 hemos dicho que las matrices de tipo

$$A=\begin{bmatrix}A_1 & & & & \\ & A_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & A_s\end{bmatrix},$$

donde $A_1, \ldots, A_s$ son matrices cuadradas, se llaman *descompuestas en células* $A_1, \ldots, A_s$ o *descompuestas en la suma directa de estas células.*

El polinomio característico de una matriz descompuesta es igual al producto de los polinomios característicos de sus células diagonales.

En efecto, si A se descompone en células $A_1, \ldots, A_s$, es fácil ver que su matriz característica es de la forma

$$\lambda E - A = \begin{bmatrix} \lambda E_1 - A_1 & & & & \\ & \lambda E_2 - A_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda E_s - A_s \end{bmatrix},$$

donde $E_1, \ldots, E_s$ son las matrices unidades de ordenes correspondientes. Se sabe del p. 2.1 que el determinante de una matriz descompuesta es igual al producto de los determinantes de sus células diagonales. Por consiguiente,

$$|\lambda E - A| = |\lambda E_1 - A_1| \cdot |\lambda E_2 - A_2| \ldots |\lambda E_s - A_s|$$

que es lo que se quería demostrar.

El polinomio mínimo de una matriz descompuesta es igual al mínimo común múltiplo de los polinomios mínimos de sus células diagonales.

Supongamos que la matriz A se descompone en las células $A_1, \ldots, A_s$. Sean $\psi_1(\lambda), \ldots, \psi_s(\lambda)$ sus polinomios mínimos respectivos. Consideremos un polinomio arbitrario $f(\lambda)$. Si $f(A) = O$, de la fórmula (7) del p. 1.4 resulta que $f(A_1) = \ldots = f(A_s) = O$. Pero todo polinomio que tiene como raíz a la matriz A_i es divisible por el polinomio mínimo $\psi_i(\lambda)$ de esta matriz. Por consiguiente, $f(\lambda)$ es un común múltiplo de los polinomios $\psi_1(\lambda), \ldots, \psi_s(\lambda)$. Viceversa, si un polinomio cualquiera $f(\lambda)$ es un común múltiplo de $\psi_1(\lambda), \ldots, \psi_s(\lambda)$, se tiene evidentemente $f(A) = O$. Por consiguiente, para obtener el polinomio mínimo de la matriz A es preciso tomar el común múltiplo de menor grado de los polinomios $\psi_1(\lambda), \ldots, \psi_s(\lambda)$, es decir, el mínimo común múltiplo, que es lo que se quería demostrar.

Una matriz celular A se llama *semidescompuesta* o *celular triangular*, si todas sus células diagonales son cuadradas, mientras que las células que figuran a un lado cualquiera de la diagonal principal están formadas por ceros. En lo sucesivo siempre aceptaremos que las células nulas de la matriz semidescompuesta se hallan sobre la diagonal principal. Por consiguiente, las matrices celulares semidescompuestas tienen la estructura siguiente

$$A=\begin{bmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1} & A_{s2} & \ldots & A_{ss} \end{bmatrix}, \tag{1}$$

donde $A_{11}, \ldots, A_{ss}$ son células cuadradas y todas las células que se hallan encima de éstas están formadas por ceros.

Sea B otra matriz semidescompuesta cualquiera, cuyas células diagonales son del mismo orden que las células correspondientes de la matriz A. En virtud de las reglas de operación con matrices celulares, tenemos

$$\begin{bmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1} & A_{s2} & \ldots & A_{ss} \end{bmatrix}+\begin{bmatrix} B_{11} & & & \\ B_{21} & B_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ B_{s1} & B_{s2} & \ldots & B_{ss} \end{bmatrix}=$$

$$=\begin{bmatrix} A_{11}+B_{11} & & & \\ A_{21}+B_{21} & A_{22}+B_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1}+B_{s1} & A_{s2}+B_{s2} & \ldots & A_{ss}+B_{ss} \end{bmatrix},$$

$$\begin{bmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ A_{s1} & A_{s2} & \ldots & A_{ss} \end{bmatrix}\cdot\begin{bmatrix} B_{11} & & & \\ B_{21} & B_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ B_{s1} & B_{s2} & \ldots & B_{ss} \end{bmatrix}=\begin{bmatrix} A_{11}B_{11} & & & \\ C_{21} & A_{22}B_{22} & & \\ \cdot & \cdot & \cdot & \cdot \\ C_{s1} & C_{s2} & \ldots & A_{ss}B_{ss} \end{bmatrix},$$

donde $C_{ij}=A_{ij}B_{jj}+A_{i,\,j+1}B_{j+1,\,j}+\ldots+A_{ii}B_{ij}$. Por consiguiente, *la suma y el producto de matrices semidescompuestas son matrices semidescompuestas cuyas células diagonales son iguales a las sumas y a los productos de las células correspondientes de las matrices dadas.* En particualr, si $f(\lambda)$ es un polinomio en λ y A es una matriz semidescompuesta de tipo (1), se tiene

$$f(A)=\begin{bmatrix} f(A_{11}) & & & \\ D_{21} & f(A_{22}) & & \\ \cdot & \cdot & \cdot & \cdot \\ D_{s1} & D_{s2} & \ldots & f(A_{ss}) \end{bmatrix} \tag{2}$$

(las células D_{ij} tienen una estructura más compleja). Del § 2 conocemos que el determinante de una matriz semidescompuesta es igual al producto de los determinantes de sus células diagonales.

De aquí se deduce directamente que *el polinomio característico de una matriz semidescompuesta es igual al producto de los polinomios característicos de las células diagonales de esta matriz.*

Si las células de una matriz semidescompuesta A son de orden 1, se dice que A tiene *forma triangular* o, simplemente, que es *una matriz triangular*. Los valores propios de una matriz triangular son iguales a sus elementos diagonales. De las fórmulas (2) resulta también que siendo $\zeta_1, \zeta_2, \ldots, \zeta_n$ los valores propios de una matriz triangular, los valores propios de la matriz $f(A)$ serán $f(\zeta_1)$, $f(\zeta_2), \ldots, f(\zeta_n)$.

Ejemplos y problemas

1. Una matriz se llama *nilpotente* si una de sus potencias es igual a la matriz nula. Demuéstrese que una matriz semidescompuesta es nilpotente si, y sólo si, son nilpotentes sus células diagonales.

2. Empleando el teorema de Hamilton-Cayley demuéstrese que si en el cuerpo de los números complejos todos los valores propios de una matriz son iguales a cero, la matriz es nilpotente.

3. Demuéstrese que si en la suma *directa* de matrices se cambian entre sí los sumandos, la suma nueva será una matriz *semejante* a la suma inicial.

4. Calcúlense los polinomios característicos y los valores propios de las matrices

$$A=\begin{bmatrix}2 & 3\\ 1 & 4\end{bmatrix},\quad B=\begin{bmatrix}\cos\alpha & \operatorname{sen}\alpha\\ -\operatorname{sen}\alpha & \cos\alpha\end{bmatrix}\quad \text{y}\quad C=\begin{bmatrix}1 & 2 & 3\\ 2 & 1 & 3\\ 3 & 3 & 6\end{bmatrix}.$$

5. Hállense los polinomios mínimos de las matrices

$$\begin{bmatrix}3 & 1\\ 0 & 3\end{bmatrix},\quad \begin{bmatrix}3 & 1 & 0\\ 0 & 3 & 0\\ 0 & 0 & 3\end{bmatrix},\quad \begin{bmatrix}2 & 0 & 0\\ 0 & 3 & 1\\ 0 & 0 & 3\end{bmatrix}\quad \text{y}\quad \begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 3 & 0\\ 0 & 0 & 0 & 3\end{bmatrix}.$$

6. Demuéstrese que los valores propios de una matriz diagonal son sus elementos diagonales.

7 Demuéstrese la fórmula

$$\text{traza}\,(AB)=\text{traza}\,(BA).$$

8. Si en una matriz cuadrada de orden n se suprimen m filas con los números $i_1, i_2, \ldots, i_m$ y m columnas con los mismos números, quedará una matriz de orden $n-m$ y su determinante se llama *menor principal* de orden $n-m$. Demuéstrese que el coeficiente de λ^m en el polinomio característico de la matriz A es igual a la suma de sus menores principales de orden $n-m$ multiplicada por $(-1)^{n-m}$.

9. Demuéstrese que el polinomio característico de la matriz AB coincide con el polinomio característico de la matriz BA.

10. Si $\zeta_1, \zeta_2, \ldots, \zeta_n$ son los valores propios de una matriz A, los valores propios de la matriz $f(A)$, donde $f(\lambda)$ es un polinomio, serán iguales a $f(\zeta_1)$, $f(\zeta_2), \ldots, f(\zeta_n)$.

Capítulo II

Espacios lineales

§ 4. Dimensión

4.1. Módulos y espacios vectoriales. Un conjunto arbitrario no vacío de elementos $\mathfrak{L}$ se llama *módulo sobre un anillo K*, si se cumplen las condiciones siguientes:

a) existe una regla que a partir de cualquier par de elementos a y b de $\mathfrak{L}$ permite hallar un elemento de $\mathfrak{L}$ que se llama *suma* de los dos primeros y se designa mediante $a+b$;

b) existe una regla que a partir de cualquier número α de K y de cualquier elemento a de $\mathfrak{L}$ permite hallar en $\mathfrak{L}$ un elemento nuevo que se llama *producto de α por a* y que se designa mediante αa;

c) las operaciones de adición y de multiplicación por número satisfacen los axiomas siguientes:

1° la adición es conmutativa:

$$a+b=b+a;$$

2° la adición es asociativa:

$$a+(b+c)=(a+b)+c;$$

3° es posible realizar la sustracción, es decir, para todo par de elementos a y b de $\mathfrak{L}$ existe en $\mathfrak{L}$ un elemento x tal que

$$a+x=b;$$

4° la multiplicación es asociativa:

$$\alpha(\beta a)=(\alpha\beta)a;$$

5° la multiplicación es distributiva respecto a la adición en $\mathfrak{L}$:

$$\alpha(a+b)=\alpha a+\alpha b;$$

6° la multiplicación es distributiva respecto a la adición de números:

$$(\alpha+\beta)a=\alpha a+\beta a.$$

Si el anillo K está provisto del elemento unidad 1, suele exigirse que se cumpla además la condición

$$7^\circ \qquad 1 \cdot a = a \quad (a \in \mathfrak{L}).$$

En este caso el módulo se llama *unitario*. Un módulo unitario sobre un cuerpo K se llama *espacio lineal* o *vectorial* sobre K. Los elementos de los módulos y de los espacios vectoriales se llaman *vectores* y en lo que sigue se representan con letras latinas minúsculas a, b, x, y, ... Los espacios lineales y los conjuntos de vectores se representarán con letras góticas mayúsculas $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{L}$, $\mathfrak{M}$, ...

En la definición dada más arriba los vectores se multiplican a la izquierda por los elementos del anillo K. Por esta razón, los módulos y espacios definidos más arriba se llaman también módulos a la izquierda y espacios a la izquierda sobre K. Si se exige que estén definidos los productos a la derecha de los elementos de $\mathfrak{L}$ por los elementos de K y si se modifican de modo correspondiente los axiomas de 4° a 7°, se obtendrá una estructura que se llama módulo a la derecha y, respectivamente, espacio lineal a la derecha. Está claro que las propiedades de los espacios a la izquierda y a la derecha son las mismas; tiene importancia distinguir la multiplicación a la izquierda y a la derecha sólo cuando estén definidas simultáneamente. En lo que sigue los espacios a la izquierda sobre K se llaman simplemente espacios sobre K.

Un espacio lineal sobre el cuerpo de los números complejos se llama *espacio lineal complejo* y un espacio lineal sobre el cuerpo de los números reales se llama *espacio lineal real*. Los ejemplos principales de módulos y de espacios lineales se darán más tarde, mientras que ahora consideraremos los corolarios más simples que se desprenden inmediatamente de los axiomas de 1° a 7°.

Ante todo, la condición a) y los axiomas 2° y 3° muestran que todo módulo es un grupo respecto a la operación de adición de vectores y, además, según el axioma 1° este grupo debe ser conmutativo. Por consiguiente, igual que en cualquier grupo conmutativo, la suma de un número finito de vectores no depende ni del orden de los sumandos de esta suma ni de la forma en la que están dispuestos los paréntesis. Por ejemplo,

$$(a+b)+(c+d) = c+(a+(d+b)).$$

Además, entre los vectores de un módulo arbitrario existe, al igual que en todo grupo aditivo, un único vector—designémoslo mediante o—que posee la propiedad de que

$$x+o = o+x = x$$

cualquiera que sea el vector x del módulo considerado. El vector o se llama *vector nulo* o simplemente *cero* del módulo. Además, para todo vector a del módulo $\mathfrak{L}$ existe en el último un vector y, y sólo uno, que satisface la ecuación

$$a+y=o.$$

Este vector y se indica por $-a$ y se llama *opuesto* de a.

De los axiomas 1°, 2° y 3° se desprende que $-o=o$ y que para cualesquiera a y b

$$-(-a)=a \quad \text{y} \quad -(a+b)=(-a)+(-b).$$

En las expresiones de tipo $(-a)+b+(-c)$ suelen omitirse los paréntesis escribiendo $-a+b-c$. La operación binaria $-$, definida mediante la fórmula

$$x-y=x+(-y),$$

se llama operación de *sustracción* de vectores. El hecho de que un mismo símbolo indica dos operaciones diferentes (la del paso al vector opuesto y la operación binaria de sustracción) no originará en lo sucesivo ningún inconveniente.

Hasta el momento hemos considerado sólo aquellas propiedades de los módulos que los caracterizan por ser ellos grupos conmutativos respecto a la adición, es decir, aquellas propiedades que se desprenden de los axiomas 1°, 2° y 3°. Teniendo en cuenta los axiomas de 4° a 7°, obtenemos fácilmente las identidades

$$0\cdot a=o,$$
$$ma=a+a+\ldots+a$$

(m sumandos; m es un número entero positivo),

$$(-\alpha)\cdot a=-(\alpha a),$$
$$\alpha\cdot o=o.$$

En efecto, la primera es consecuencia de que

$$a=(1+0)\,a=1\cdot a+0\cdot a=a+0\cdot a;$$

la segunda se demuestra de modo siguiente:

$$ma=(1+\ldots+1)\,a=1\cdot a+\ldots+1\cdot a=a+\ldots+a.$$

Luego, puesto que

$$\alpha a+(-\alpha)\,a=(\alpha-\alpha)\,a=o,$$

se tiene $(-\alpha)\,a=-(\alpha a)$. Finalmente, la cuarta igualdad es consecuencia de que

$$\alpha\cdot o+\alpha a=\alpha\,(o+a)=\alpha a.$$

Observemos también que, siendo K un cuerpo, de $\alpha a=o$ se deduce que o bien $\alpha=0$ o bien $a=o$. En efecto, si $\alpha\neq 0$, multiplicando la relación $\alpha a=o$ por α^{-1} obtenemos $a=o$.

Las expresiones de tipo

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s$$

se llaman *combinaciones lineales* de los vectores $a_1, a_2, \ldots, a_s$. De nuestras consideraciones se desprende que las combinaciones lineales pueden ser sumadas, restadas y multiplicadas por números y que en ellas se pueden reducir los términos semejantes siguiendo las reglas habituales.

El ejemplo más conocido de un espacio lineal es el conjunto de los segmentos orientados que parten de un punto fijo O de nuestro espacio habitual. Multiplicar un segmento por un número real positivo α significa aumentar su longitud α veces sin alterar su dirección. Si α es negativo, la multiplicación de un segmento por α significa que su longitud aumenta $|\alpha|$ veces y que su dirección cambia por la contraria. Análogamente, sumar dos segmentos significa tomar la diagonal del paralelogramo construido a partir de estos segmentos. El vector nulo será el segmento, cuyo origen y extremo coinciden con el punto O. Puesto que la operación de multiplicación de un segmento por un número está definida sólo para los números reales, el campo principal K es en este caso el cuerpo de todos los números reales y el espacio de segmentos orientados es un espacio lineal *real*. Lo llamaremos siempre espacio de vectores-segmentos corrientes y lo indicaremos por $\mathfrak{R}$.

Un ejemplo más general de espacios lineales, y además básico para toda la teoría de los mismos, es *el espacio de filas*. Consideremos el conjunto de todas las sucesiones de tipo $[\alpha_1, \alpha_2, \ldots, \alpha_n]$, donde $\alpha_1, \ldots, \alpha_n$ son números de un cuerpo K y n es un número entero dado. También llamaremos estas sucesiones *filas*, considerándolas como matrices compuestas de una fila. Dos filas se llaman *iguales* si son iguales sus elementos respectivos. Las operaciones de adición de filas y de multiplicación de filas por un número se definen mediante las fórmulas matriciales correspondientes:

$$\beta[\alpha_1, \ldots, \alpha_n] = [\beta\alpha_1, \ldots, \beta\alpha_n],$$
$$[\alpha_1, \alpha_2, \ldots, \alpha_n] + [\beta_1, \beta_2, \ldots, \beta_n] =$$
$$= [\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n].$$

Está claro que en este caso los axiomas de 1° a 7° se cumplen y que el conjunto de todas las filas de longitud n formadas por elementos de un cuerpo K es *un espacio lineal sobre* K. Siendo K un anillo cualquiera, obtenemos un ejemplo de *un módulo sobre* K.

En lugar de filas se pueden considerar matrices de un número cualquiera, pero fijo, de filas y de columnas con elementos de un anillo K. De acuerdo con las reglas del cálculo de matrices, cualquier matriz de este tipo puede ser multiplicada por un número de K y cualesquiera dos pueden ser sumadas, obteniéndose ambas veces una matriz del mismo tipo. Es obvio que los axiomas de 1° a

7° también se cumplen aquí y, por consiguiente, las matrices de m filas y de n columnas constituyen, respecto a las operaciones de adición y de multiplicación por número, un módulo sobre K o un espacio lineal, si K es un cuerpo.

Consideremos un ejemplo más. Sea $\mathfrak{M}$ un conjunto arbitrario de cualesquiera elementos y sea K un anillo. Supongamos que tenemos a nuestra disposición una ley que pone en correspondencia a cada elemento m de $\mathfrak{M}$ un número determinado de K. Toda ley de esta índole se llama *función* definida sobre el conjunto $\mathfrak{M}$ y con valores en K. Si la función se indica con una letra cualquiera f, mediante $f(m)$ suele indicarse el número que corresponde al elemento m. El número $f(m)$ se llama valor de la función f sobre el elemento m. Dos funciones f y g que satisfacen la igualdad $f(m) = g(m)$ cualquiera que sea m de $\mathfrak{M}$, se llaman *iguales*. Para las funciones definidas sobre $\mathfrak{M}$ se pueden introducir de modo habitual las operaciones de multiplicación por número, de adición y de multiplicación de funciones. Por ejemplo, dados un número α y una función f y poniendo en correspondencia a cada elemento m de $\mathfrak{M}$ el número $\alpha f(m)$, obtenemos una función nueva que se llama producto del número α por la función f. Análogamente se definen la suma y el producto de dos funciones. Si consideramos sólo las dos primeras de estas operaciones — la multiplicación por número y la adición — es fácil ver que se cumplen los axiomas de un módulo. Por consiguiente, el conjunto de todas las funciones, definidas sobre un conjunto dado $\mathfrak{M}$ y con valores en un anillo K, constituye *un módulo sobre K*. Siendo el anillo inicial un cuerpo, obtenemos, al igual que anteriormente, un ejemplo de *espacio vectorial*..

En el caso en que $\mathfrak{M}$ consta de un número finito de elementos, se puede introducir una notación cómoda para las funciones definidas sobre $\mathfrak{M}$. Sean $m_1, m_2, \ldots, m_s$ los elementos del conjunto $\mathfrak{M}$. Indiquemos mediante $[m_i]$ la función que es igual a 1, sobre m_i e igual a 0 sobre todos los demás elementos de $\mathfrak{M}$ $(i = 1, \ldots, \ldots, s)$. En estas condiciones el producto $\alpha[m_i]$ será la función igual a α sobre m_i e igual a 0 sobre todos los demás elementos de $\mathfrak{M}$ y la expresión

$$\alpha_1 [m_1] + \alpha_2 [m_2] + \ldots + \alpha_s [m_s] \tag{1}$$

será, obviamente, la función que es igual a α_1 sobre m_1, a α_2 sobre $m_2, \ldots$, a α_s sobre m_s. Por consiguiente, toda función definida sobre $\mathfrak{M}$ puede ser representada en la forma (1) y esta representación, como es fácil ver, es única. Si no existe el peligro de confusión, los paréntesis en (1) suelen omitirse y en lugar de (1) suele escribirse brevemente:

$$\alpha_1 m_1 + \alpha_2 m_2 + \ldots + \alpha_s m_s$$

con la particularidad de que los términos de coeficientes nulos no se escriben. Por ejemplo, si $\mathfrak{M}$ se compone de las letras a, b y c, la expresión $2a-c$ significa la función que es igual a 2 sobre a, a cero sobre b y a -1 sobre c.

4.2. Dependencia lineal. Sea $\mathfrak{L}$ un espacio vectorial sobre un cuerpo de coeficientes K y sean $a_1, a_2, \ldots, a_m$ unos vectores de este espacio. La relación de tipo

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m = o,$$

donde $\alpha_1, \ldots, \alpha_m$ son números de K, se llama *relación de dependencia lineal* entre los vectores $a_1, \ldots, a_m$. Si todos los coeficientes $\alpha_1, \ldots, \alpha_m$ son iguales a cero, la relación se llama *trivial*. En el caso contrario, es decir, si al menos uno de los coeficientes es diferente de cero, la relación se llama *no trivial*. Está claro que una relación trivial existe entre cualesquiera vectores. La respuesta a la pregunta sobre si existe o no una relación no trivial, depende de los vectores que se consideran. Por ejemplo, en el espacio de filas de longitud tres entre los vectores $a=[1, 4, 6]$, $b=[1, -1, 1]$ y $c=[1, 1, 3]$ existe la siguiente relación de dependencia lineal

$$2a+3b-5c=o.$$

Por otra parte, en este mismo espacio no existe ninguna relación no trivial entre los vectores $e_1=[1, 0, 0]$, $e_2=[0, 1, 0]$ y $e_3=[0, 0, 1]$ ya que la relación

$$\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = o$$

significa que

$$[\alpha_1, \alpha_2, \alpha_3] = [0, 0, 0],$$

de donde se tiene

$$\alpha_1 = \alpha_2 = \alpha_3 = 0.$$

Un sistema finito de vectores $a_1, a_2, \ldots, a_m$ de un espacio lineal se llama linealmente dependiente, si entre ellos existe una relación no trivial.

Si no existe tal relación, es decir, si de toda relación de tipo

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m = o$$

se deduce que $\alpha_1 = \alpha_2 = \ldots = \alpha_m = 0$, el sistema $a_1, a_2, \ldots, a_m$ se llama *linealmente independiente*.

De esta definición se deduce inmediatamente que *si a un sistema linealmente dependiente de vectores $a_1, a_2, \ldots, a_m$ se agregan otros vectores $b_1, \ldots, b_k$, el sistema ampliado seguirá siendo linealmente dependiente*. En efecto, si

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m = o$$

es una relación no trivial entre $a_1, \ldots, a_m$, tendremos que

$$\alpha_1 a_1 + \ldots + \alpha_m a_m + 0 \cdot b_1 + \ldots + 0 \cdot b_k = o$$

será una relación no trivial entre $a_1, \ldots, a_m, b_1, \ldots, b_k$.

En los problemas relacionados con la dependencia lineal el vector nulo o ocupa una posición especial debido a que todo sistema que contenga el vector nulo es linealmente dependiente. Para demostrar esta afirmación basta observar que la relación

$$1 \cdot o + 0 \cdot a_1 + \ldots + 0 \cdot a_m = o$$

es una relación no trivial cualesquiera que sean $a_1, \ldots, a_m$.

La definición que hemos dado de dependencia lineal de un sistema presupone que dicho sistema contiene un número finito de vectores. Sin embargo, resulta necesario, con frecuencia, considerar también sistemas infinitos. Diremos que un sistema infinito de vectores es *linealmente dependiente*, si resulta linealmente dependiente alguna parte suya finita. El siguiente ejemplo muestra que pueden existir sistemas infinitos linealmente independientes de vectores. Consideremos el conjunto de todos los polinomios en letra λ con coeficientes de un cuerpo K. Respecto a las operaciones de adición y de multiplicación por números de K, estos polinomios forman, evidentemente, un espacio lineal. Los polinomios

$$1, \lambda, \lambda^2, \ldots, \lambda^m, \ldots$$

constituyen un sistema linealmente independiente en este espacio. Efectivamente, cualquier parte finita suya

$$\lambda^{m_1}, \lambda^{m_2}, \ldots, \lambda^{m_k}$$

$$(0 \leqslant m_1 < m_2 < \ldots < m_k)$$

es linealmente independiente, ya que de

$$\alpha_1 \lambda^{m_1} + \alpha_2 \lambda^{m_2} + \ldots + \alpha_k \lambda^{m_k} = o$$

se deduce que $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$.

Consideremos un espacio lineal arbitrario $\mathfrak{L}$. Si un vector a de este espacio puede ser representado en la forma

$$a = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m$$

se dice que *a se expresa linealmente en términos de $a_1, a_2, \ldots, a_m$* o que *a depende linealmente de $a_1, \ldots, a_m$*.

Si los vectores $a_1, \ldots, a_m$ se expresan linealmente en términos de $b_1, \ldots, b_n$ y $b_1, \ldots, b_n$ se expresan linealmente en téminos de $c_1, \ldots, c_p$, los vectores $a_1, \ldots a_m$ se expresan linealmente en términos de $c_1, \ldots, c_p$.

Para la demostración basta en las expresiones lineales de los vectores $a_1, \ldots, a_m$ en términos de $b_1, \ldots, b_n$ sustituir $b_1, \ldots b_n$ por sus expresiones en términos de $c_1, \ldots, c_p$ y reducir los términos semejantes.

TEOREMA 1. *Si un sistema de vectores no nulos $a_1, a_2, \ldots, a_m$ considerados en un orden determinado es linealmente dependiente, al menos uno de estos vectores puede ser expresado linealmente en términos de los anteriores. Recíprocamente, si uno de los vectores de esta sucesión se expresa linealmente en términos de los anteriores, el sistema es linealmente dependiente.*

Supongamos que entre los vectores $a_1, a_2, \ldots, a_m$ existe una relación no trivial

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m = o. \tag{1}$$

Sea α_k el último coeficiente diferente de cero. Si $k=1$, la relación (1) se convierte en

$$\alpha_1 a_1 = o \quad (\alpha_1 \neq 0),$$

de donde $a_1 = o$ a pesar de haber aceptado que el sistema no contiene vectores nulos. Por consiguiente, $1 < k \leqslant m$ y la relación (1) puede ser representada en la forma

$$\alpha_1 a_1 + \ldots + \alpha_k a_k = o \quad (\alpha_k \neq 0),$$

de donde se tiene

$$a_k = -\alpha_k^{-1}\alpha_1 a_1 - \ldots - \alpha_k^{-1}\alpha_{k-1}a_{k-1}.$$

Con esto queda demostrada la primera parte del teorema. La afirmación recíproca es evidente.

Sea $\mathfrak{M}$ un conjunto de vectores de un espacio lineal $\mathfrak{L}$. Un sisteme de vectores $a_1, a_2, \ldots$ de este conjunto se llama *sistema de generadores* de $\mathfrak{M}$, si todo vector de $\mathfrak{M}$ puede ser expresado linealmente en términos de un número finito de los vectores $a_1, a_2, \ldots$ Un sistema linealmente independiente de generadores de $\mathfrak{M}$ se llama *base* del conjunto $\mathfrak{M}$. Por ejemplo, en el espacio de todos los polinomios en λ los polinomios

$$1, \lambda, \lambda^2, \ldots, \lambda^m, \ldots \tag{2}$$

constituyen una base ya que estos polinomios, como hemos visto, son linealmente independientes y, por otro lado, todo polinomio es de la forma

$$\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_m\lambda^m,$$

es decir, se expresa linealmente en términos de $1, \lambda, \ldots, \lambda^m$.

LEMA *Si un sistema de generadores $a_1, a_2, \ldots$ de un conjunto $\mathfrak{M}$ contiene un elemento a_i que puede ser expresado linealmente en términos de los demás generadores, entonces suprimiendo a_i en el sistema de generadores se obtiene de nuevo un sistema de generadores de $\mathfrak{M}$.*

En efecto, por hipótesis todo vector de $\mathfrak{M}$ es una combinación lineal de un número finito de generadores. Sustituyendo en estas combinaciones el vector a_i por su expresión en términos de los demás generadores obtendremos para cualquier elemento de $\mathfrak{M}$ una expresión en términos de los generadores diferentes de a_i que es lo que se quería demostrar.

TEOREMA 2. *De todo sistema de generadores de un espacio $\mathfrak{L}$ se puede extraer una base de este espacio.*

Supongamos que $\mathfrak{L}$ tiene un sistema finito de generadores $a_1, a_2, \ldots, a_s$. Omitiendo en este sistema todos los vectores que se expresan linealmente en términos de los anteriores, obtendremos un sistema de vectores $a_{i_1}, a_{i_2}, \ldots a_{i_n}$ que, de acuerdo con el lema, será aun un sistema de generadores del espacio $\mathfrak{L}$. Puesto que ningún vector de este último sistema puede ser expresado linealmente en términos de los anteriores, este sistema es, en virtud del teorema 1, linealmente independiente y, por lo tanto, es una base del espacio.

Hemos demostrado el teorema 2 aceptando que el sistema de generadores contiene un número finito de elementos. Sin embargo, este teorema es válido también en el caso de un sistema infinito de generadores. Para la demostración es suficiente disponer los generadores en una sucesión transfinita $a_1, a_2, \ldots, a_\omega, a_{\omega+1}, \ldots$ y omitir en ella todos los elementos que se expresan linealmente en términos de los anteriores.

Sea $a_1, a_2, \ldots$ una base del espacio $\mathfrak{L}$. Según la definición de una base, todo vector de $\mathfrak{L}$ se expresa linealmente en términos de un número finito de vectores básicos. Demostremos que esta representación es única. En efecto, sea

$$a=\alpha_1a_1+\alpha_2a_2+\ldots+\alpha_sa_s$$

y sea

$$a=\beta_1a_1+\beta_2a_2+\ldots+\beta_sa_s.$$

Restando obtenemos

$$o=(\alpha_1-\beta_1)\,a_1+(\alpha_2-\beta_2)\,a_2+\ldots+(\alpha_s-\beta_s)\,a_s.$$

Como $a_1, a_2, \ldots$ son linealmente independientes, se tiene $\alpha_1-\beta_1=0$, $\alpha_2-\beta_2=0, \ldots, \alpha_s-\beta_s=0$, es decir, $\alpha_1=\beta_1, \ldots, \alpha_s=\beta_s$, que es lo que se quería demostrar.

Si un espacio lineal $\mathfrak{L}$ tiene al menos una base compuesta por un número finito de elementos, se dice que $\mathfrak{L}$ es *de dimensión finita*. Con más detalle, $\mathfrak{L}$ se llama de dimensión finita si se puede escoger en $\mathfrak{L}$ un sistema linealmente independiente finito de vectores, tal, que todos los vectores de $\mathfrak{L}$ se expresan linealmente en términos de este sistema.

El concepto de base no se puede aplicar al espacio nulo. Sin embargo, aceptaremos que el espacio nulo es también de dimensión finita.

TEOREMA 3. *Todas las bases de un espacio lineal no nulo $\mathfrak{L}$ de dimensión finita constan de un mismo número finito de vectores. Este* *número se llama dimensión del espacio $\mathfrak{L}$. La dimensión del espacio nulo es, por definición, el número cero.*

Según la definición, en $\mathfrak{L}$ existe una base que consta de un número finito de vectores. Sea

$$a_1, a_2, \ldots, a_n \tag{3}$$

esta base. Demostremos que el número de vectores de otra base cualquiera

$$x_1, x_2, \ldots, x_m, \ldots \tag{4}$$

no puede ser mayor que n. En efecto, consideremos el sistema

$$x_1, a_1, a_2, \ldots, a_n. \tag{5}$$

Puesto que el sistema (3) es un sistema de generadores de $\mathfrak{L}$, es decir, todo vector de $\mathfrak{L}$ se expresa linealmente en términos de los vectores (3), también el sistema (5) tendrá esta propiedad. Sin embargo, el sistema (5) es linealmente dependiente, ya que su primer vector x_1 puede ser expresado linealmente en términos de los restantes. Aplicando a la sucesión (5) el teorema 3, vemos que uno de los vectores de esta sucesión, digamos a_i, debe expresarse linealmente en términos de los anteriores. Omitiendo en (5) el vector a_i, obtendremos una sucesión nueva

$$x_1, a_1', \ldots, a_{n-1}'. \tag{6}$$

donde $a_1', \ldots, a_{n-1}'$ representan aquellos de los vectores $a_1, \ldots, a_n$ que hemos conservado. En virtud del lema, el sistema (6) será aun un sistema de generadores de $\mathfrak{L}$. Consideremos ahora la sucesión

$$x_2, x_1, a_1', \ldots, a_{n-1}'. \tag{7}$$

Esta sucesión es linealmente dependiente debido a que el vector x_2 se expresa linealmente en términos de sus elementos restantes. Por consiguiente, de acuerdo con el teorema 1 uno de los vectores de este sistema debe expresarse linealmente en términos de *los anteriores*. Este vector debe ser sólo uno de los vectores $a_1', \ldots, a_{n-1}'$, ya que x_2 y x_1 son, por hipótesis, linealmente independientes. Omitiéndolo de la sucesión (7), obtenemos la sucesión

$$x_2, x_1, a_1'', \ldots, a_{n-2}'', \tag{8}$$

donde $a_1'', \ldots, a_{n-2}''$ son aquellos de los vectores $a_1', \ldots, a_{n-1}'$ que hemos conservado. Puesto que la sucesión (7) era un sistema de generadores de $\mathfrak{L}$, también la sucesión (8) será, en virtud del lema, un sistema de generadores de $\mathfrak{L}$. Agregando ahora a la sucesión (8) el vector x_3, omitiremos de la nueva sucesión el vector que se expresa linealmente en términos de los anteriores, etc. Si el número de vectores x_i fuese mayor que n, al cabo de n pasos obtendríamos una sucesión

$$x_n, x_{n-1}, \ldots, x_2, x_1 \tag{9}$$

que no contendría los vectores $a_1, \ldots, a_n$ y que sería un sistema de generadores de $\mathfrak{L}$. Esto significaría que todos los vectores del espacio $\mathfrak{L}$ podrían ser expresados linealmente en términos de los vectores (9). En particular, también el vector x_{n+1} podría ser expresado linealmente en términos de los vectores (9), lo que estaría en contradicción con la independencia lineal de los vectores $x_1, x_2, \ldots$ Por consiguiente la base (4) no puede contener más vectores que la base (3), es decir, todas las bases de $\mathfrak{L}$ constan de un número *finito* de vectores. Por otro lado, la base (3) ha sido escogida arbitrariamente. Por lo tanto, con nuestro razonamiento queda demostrado que el número de vectores de una base no puede ser inferior al número de vectores de otra base, es decir, todas las bases de $\mathfrak{L}$ tienen un mismo número de vectores.

TEOREMA 4. *Cualquiera que sea un sistema linealmente independiente $a_1, \ldots, a_m$ de vectores de un espacio $\mathfrak{L}$ de dimensión finita, se puede encontrar en $\mathfrak{L}$ unos vectores $a_{m+1}, \ldots, a_n$ tales, que el sistema $a_1, \ldots, a_m, a_{m+1}, \ldots, a_n$ sea una base de $\mathfrak{L}$.*

DEMOSTRACIÓN. Escojamos en $\mathfrak{L}$ una base cualquiera $x_1, x_2, \ldots, x_n$ y consideremos la sucesión

$$a_1, a_2, \ldots, a_m, x_1, x_2, \ldots, x_n. \tag{10}$$

Omitamos ahora en esta sucesión todos aquellos vectores que se expresan linealmente en términos de los anteriores. Puesto que $a_1, a_2, \ldots a_m$ son linealmente independientes, no será omitido ninguno de ellos y el sistema que resulte será de la forma

$$a_1, \ldots, a_m, x_{i_1}, \ldots, x_{i_k}. \tag{11}$$

Debido al teorema 1 este sistema es linealmente independiente. Por otra parte, todos los vectores del espacio $\mathfrak{L}$ podían ser expresados linealmente en términos del sistema (10). En virtud del lema, la misma propiedad la debe tener también el sistema (11). Por consiguiente, el sistema (11) es una base del espacio $\mathfrak{L}$ y $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$ son los vectores que queríamos encontrar.

El teorema 4 puede ser enunciado también en esta forma: *todo sistema linealmente independiente de vectores del espacio $\mathfrak{L}$ o bien es una base o bien forma parte de una base de $\mathfrak{L}$.*

Supongamos que $\mathfrak{L}$ es de dimensión n. Entonces, toda base del espacio $\mathfrak{L}$ contiene n vectores y, por consiguiente, el número de vectores en cualquier sistema linealmente independiente de $\mathfrak{L}$ es o menor de n o igual a n. En el último caso el sistema debe ser una base de $\mathfrak{L}$. En particular, el número máximo de vectores linealmente independientes de $\mathfrak{L}$ es igual a n, es decir, es igual a la dimensión de $\mathfrak{L}$.

Hemos obtenido el teorema siguiente.

TEOREMA 5. *Todo sistema de $n+1$ vectores de un espacio lineal $\mathfrak{L}$ de n dimensiones es linealmente dependiente. Cualesquiera n vectores linealmente independientes de este espacio constituyen una base del mismo. El número máximo de vectores linealmente independientes del espacio $\mathfrak{L}$ es igual a la dimensión de este espacio.*

Para concluir, observemos que los teoremas 3 y 4, que hemos enunciado para el caso de espacios de dimensión finita, tienen lugar también en los espacios de dimensión infinita. Solamente en lugar del número de vectores de una base habrá que considerar la potencia del conjunto de los vectores que la componen y, respectivamente, la dimensión de un espacio deberá entenderse como la potencia del conjunto de vectores que forman una base cualquiera de este espacio. En cuanto al teorema 5, solamente la última de sus afirmaciones puede ser extendida directamente al caso de dimensión infinita. En este libro estudiaremos las propiedades de los espacios de dimensión finita. Por esto, *en lo sucesivo un espacio lineal se comprenderá, siempre que no se diga lo contrario, como un espacio lineal de dimensión finita.*

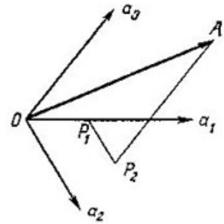

Fig. 1.

Determinemos la dimensión de los espacios considerados en el p.4.1. Sea $\mathfrak{R}$ el espacio habitual de segmentos orientados que parten de un punto O. Como de costumbre, diremos que estos segmentos son vectores. Demostremos que *cualesquiera tres vectores a_1, a_2 y a_3 del espacio $\mathfrak{R}$, que parten del punto O y que no pertenecen a un mismo plano, constituyen una base de $\mathfrak{R}$.* En efecto, los vectores a_1, a_2 y a_3 son linealmente independientes, ya que de lo contrario uno de ellos, digamos a_3, debería expresarse linealmente en términos de los otros dos. Sin embargo, la relación $a_3=\alpha_1 a_1+\alpha_2 a_2$ significa que a_3 es la diagonal del paralelogramo construido a partir de los vectores $\alpha_1 a_1$ y $\alpha_2 a_2$. Puesto que $\alpha_1 a_1$ y $\alpha_2 a_2$ se hallan en el plano a_1Oa_2, también a_3 tendría que pertenecer al mismo plano, lo que estaría en contradicción con la hipótesis. Por otra parte, todo vector $\overrightarrow{OA}$ del espacio $\mathfrak{R}$ puede ser expresado linealmente en términos de a_1, a_2 y a_3 (fig. 1):

$$\overrightarrow{OA}=\overrightarrow{OP_1}+\overrightarrow{P_1P_2}+\overrightarrow{P_2A}=\alpha_1 a_1+\alpha_2 a_2+\alpha_3 a_3,$$

donde α_i es la relación entre la longitud del segmento $P_{i-1}P_i$ ($P_0=O$, $P_3=A$) y la longitud de a_i tomado con el signo adecuado. Por consiguiente, a_1, a_2 y a_3 es una base del espacio $\mathfrak{R}$ y $\mathfrak{R}$ es de dimensión tres.

Análogamente se demuestra que el conjunto de vectores, que parten de un punto O y que pertenecen a un plano que pasa por O, es un espacio lineal de dos dimensiones y que el conjunto de vectores, que parten de O y que pertenecen a una recta que pasa por O, es un espacio de una dimensión.

Consideremos el espacio de filas de longitud n formadas por elementos de un cuerpo K. Sea

$$\begin{gathered} e_1 = [1, 0, \ldots, 0], \\ e_2 = [0, 1, \ldots, 0], \\ \cdots\cdots\cdots \\ e_n = [0, 0, \ldots, 1]. \end{gathered}$$

Multiplicando sucesivamente estas filas por números arbitrarios $\alpha_1, \alpha_2, \ldots, \alpha_n$ y sumando, obtenemos

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = [\alpha_1, \alpha_2, \ldots, \alpha_n].$$

Es decir, una fila arbitraria $[\alpha_1, \ldots, \alpha_n]$ se expresa linealmente en términos de $e_1, \ldots, e_n$. Sin embargo, el sistema $e_1, \ldots, e_n$ es linealmente independiente, debido a que la relación

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots + \alpha_n e_n = o$$

implica

$$[\alpha_1, \alpha_2, \ldots, \alpha_n] = [0, 0, \ldots, 0],$$

de donde se tiene $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Por ello, $e_1, e_2, \ldots, e_n$ es una base del espacio considerado y, por consiguiente, su dimensión es n.

Finalmente, sea $\mathfrak{L}$ el espacio de funciones definidas sobre un conjunto finito $\mathfrak{M}$ y con valores en un cuerpo K. Sean $m_1, m_2, \ldots, m_s$ los elementos de $\mathfrak{M}$. Según el p. 4.1, toda función f de $\mathfrak{L}$ puede ser representada en la forma

$$f = \alpha_1 [m_1] + \alpha_2 [m_2] + \ldots + \alpha_s [m_s],$$

donde $[m_i]$ es la función igual a 1 sobre m_i e igual a 0 sobre los demás elementos del conjunto $\mathfrak{M}$. Es decir, f se expresa linealmente en términos de $[m_1], \ldots, [m_s]$. Por otro lado, la igualdad

$$\alpha_1 [m_1] + \alpha_2 [m_2] + \ldots + \alpha_s [m_s] = o$$

significa que todos los valores de la función que figura en el primer miembro son iguales a cero y, por consiguiente, $\alpha_1 = \ldots = \alpha_s = 0$. Vemos de esta forma que $[m_1], \ldots, [m_s]$ es una base del espacio $\mathfrak{L}$ y que su dimensión es igual a s.

Hemos considerado espacios lineales es decir, módulos (unitarios) sobre un cuerpo K. Ejemplos muy sencillos permiten ver que para los módulos sobre *anillos arbitrarios*, e incluso sobre anillos conmutativos arbitrarios, los teoremas 1 y 2 dejan de tener lugar. Sea, en particular, K el anillo de los números enteros corrientes y sea $\mathfrak{L}$

el conjunto de estos mismos números enteros en el que la operación de adición está definida del modo corriente y la operación de multiplicación de «números» de K por «vectores» de $\mathfrak{L}$ está definida como la multiplicación habitual de números enteros. Convendremos en indicar los números enteros que representan los vectores de $\mathfrak{L}$ con cifras gruesas, mientras que los números de K se indicarán con cifras corrientes. Puesto que para cualquier vector $\boldsymbol{m}$ es válida la fórmula

$$\boldsymbol{m} = m \cdot \mathbf{1},$$

nuestro módulo numérico $\mathfrak{L}$ resulta ser generado por el vector $\mathbf{1}$. Los vectores $\mathbf{2}$ y $\mathbf{3}$ son linealmente dependientes, ya que

$$2 \cdot \mathbf{3} + (-3) \cdot \mathbf{2} = \mathbf{0}.$$

Sin embargo, para cualquier $\alpha \in K$ se tiene $\alpha \cdot \mathbf{3} \neq \mathbf{2}$ y $\alpha \cdot \mathbf{2} \neq \mathbf{3}$, es decir, en $\mathfrak{L}$ ninguno de los vectores $\mathbf{2}$ y $\mathbf{3}$ se expresa linealmente en términos del otro. Las propiedades de los módulos sobre anillos arbitrarios son estudiadas sistemáticamente en la teoría de los módulos, estrechamente ligada a la teoría de los números. Para nosotros, en cambio, tendrán interés solamente las propiedades de los espacios vectoriales y tocaremos la teoría de los módulos sobre anillos sólo cuando esto sea necesario para la teoría de los espacios vectoriales.

4.3. Isomorfismo. El concepto de espacio lineal tiene dos facetas esencialmente diferentes. En primer lugar, un espacio lineal es un conjunto de ciertas entes que se denominan vectores y, en segundo lugar, en un espacio lineal actúan las operaciones de adición y de multiplicación por número. Por esto, o bien podemos limitarnos a estudiar qué es lo que representan los vectores y cuáles son la naturaleza y las propiedades de los mismos, o bien podemos tomar otro punto de vista y estudiar las propiedades de las operaciones indicadas independientemente de la naturaleza de los elementos con los cuales se efectúan estas operaciones. En lo sucesivo nos interesarán solamente las propiedades del segundo género. Por ello, dos espacios de la misma estructura respecto a las operaciones de adición y de multiplicación por número se considerará que tienen las mismas propiedades o que son *isomorfos*. Con más precisión el concepto de *isomorfismo* puede enunciarse del modo siguiente:

Dos espacios lineales sobre un mismo cuerpo de coeficientes se llaman isomorfos, si se puede establecer una correspondencia biyectiva entre sus elementos, tal que a la suma de vectores del primer espacio corresponda la suma de los vectores correspondientes del segundo espacio y al producto de un número por un vector del primer espacio corresponda el producto de este mismo número por el vector correspondiente del segundo espacio.

Toda correspondencia biyectiva que posee las propiedades indicadas se llama *isomorfa* o *isomorfismo*. Consideremos las propiedades elementales de los isomorfismos.

En una correspondencia isomorfa el vector nulo corresponde al vector nulo. En efecto, supongamos que en una aplicación isomorfa de un espacio lineal $\mathfrak{L}$ sobre otro espacio lineal $\mathfrak{L}_1$ el vector a de $\mathfrak{L}$ corresponde al vector a_1 de $\mathfrak{L}_1$. Entonces, según la definición de un isomorfismo, el producto $0 \cdot a$ debe corresponder al producto $0 \cdot a_1$, es decir, el vector nulo del primer espacio debe transformarse en el vector nulo del segundo espacio.

En una aplicación isomorfa un sistema de generadores del primer espacio se transforma en un sistema de generadores del segundo espacio.

Efectivamente, sean $a_1, a_2, \ldots, a_s$ unos generadores del primer espacio y sean $b_1, b_2, \ldots, b_s$ los vectores que les corresponden en el segundo espacio. Tomemos en el segundo espacio un vector arbitrario b y consideremos el vector a del primer espacio que le corresponde. Por hipótesis, el vector a puede ser representado en la forma

$$a = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_s a_s.$$

Según la definición de una aplicación isomorfa, la suma $\alpha_1 a_1 + \ldots + \alpha_s a_s$ debe transformarse en la suma $\alpha_1 b_1 + \ldots + \alpha_s b_s$ y, por consiguiente, el vector b debe coincidir con la suma $\alpha_1 b_1 + \ldots + \alpha_s b_s$, es decir, los vectores $b_1, \ldots, b_s$ constituyen un sistema de generadores del segundo espacio.

En un isomorfismo los vectores linealmente independientes se transforman en vectores linealmente independientes.

En efecto, supongamos que los vectores linealmente independientes $a_1, a_2, \ldots, a_m$ del primer espacio se transforman en los vectores $b_1, b_2, \ldots, b_m$ del segundo espacio. Supongamos que entre los últimos existe una relación de tipo

$$\beta_1 b_1 + \beta_2 b_2 + \ldots + \beta_m b_m = o_1.$$

Según la definición de un isomorfismo, al primer miembro de esta igualdad corresponde en el primer espacio el vector $\beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_m a_m$ y al vector nulo o_1 corresponde en el primer espacio el vector nulo o. Por consiguiente,

$$\beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_m a_m = o.$$

Puesto que los vectores $a_1, \ldots, a_m$ son linealmente independientes se tiene

$$\beta_1 = \beta_2 = \ldots = \beta_m = 0,$$

es decir, los vectores $b_1, \ldots, b_m$ son linealmente independientes.

De las dos propiedades últimas se deduce directamente que *en un isomorfismo una base de un espacio lineal se transforma de nuevo en una base de un espacio lineal* y, por consiguiente, *los espacios lineales isomorfos tienen la misma dimensión.*

La afirmación recíproca es también válida: *si dos espacios lineales sobre un mismo cuerpo de coeficientes tienen la misma dimensión, son isomorfos.*

Para la demostración tomemos una base cualquiera en cada uno de los espacios dados, por ejemplo, $a_1, a_2, \ldots, a_n$ y $b_1, b_2, \ldots, b_n$. Diremos que los vectores

$$a=\alpha_1 a_1+\alpha_2 a_2+\ldots+\alpha_n a_n$$

y

$$b=\beta_1 b_1+\beta_2 b_2+\ldots+\beta_n b_n$$

son *correspondientes*, si $\alpha_1=\beta_1, \ldots, \alpha_n=\beta_n$. Puesto que todo vector de un espacio se expresa linealmente en términos de una base de un modo único, nuestra correspondencia es biyectiva. Sean ahora

$$a=\alpha_1 a_1+\alpha_2 a_2+\ldots+\alpha_n a_n$$

y

$$b=\alpha_1 b_1+\alpha_2 b_2+\ldots+\alpha_n b_n$$

dos vectores correspondientes. Entonces se tiene

$$\alpha a=\alpha\alpha_1 a_1+\alpha\alpha_2 a_2+\ldots+\alpha\alpha_n a_n$$

y

$$\alpha b=\alpha\alpha_1 b_1+\alpha\alpha_2 b_2+\ldots+\alpha\alpha_n b_n.$$

Puesto que en estas descomposiciones coinciden los coeficientes respectivos, los vectores αa y αb serán correspondientes, es decir, en nuestra correspondencia el producto de un número por un vector se transforma en el producto del mismo número por el vector correspondiente. Análogamente se demuestra que la suma de vectores se transforma en la suma de vectores correspondientes. Por esto, la correspondencia construida es un isomorfismo, que es lo que se quería demostrar.

Las propiedades de las correspondencias isomorfas que hemos enunciado muestran que fijado el cuerpo principal K todo espacio lineal queda determinado, salvo un isomorfismo, por su dimensión. Por esta razón, los espacios de filas de longitud n con elementos de un cuerpo K, donde $n=1, 2, \ldots,$ agotan, salvo isomorfismos, todos los espacios de dimensión finita sobre K. En particular, el espacio habitual de segmentos orientados es isomorfo al espacio de filas de longitud tres sobre el cuerpo de los números reales, el espacio de funciones definidas sobre un conjunto $\mathfrak{M}$, compuesto por s elementos, y con valores en un cuerpo K es isomorfo al espacio de filas de longitud s con elementos de K, etc.

Para concluir hagamos una observación más. En el Algebra general desempeñan un papel principal el concepto de álgebra de signatura dada y el concepto de isomorfismo de álgebras de signatura determinada. En la definición que hemos dado del concepto de módulo no se ha indicado, desde el punto de vista formal, la signatura del módulo, es decir, no se han señalado las operaciones que se consideran como principales y respecto a las cuales se define el concepto de isomorfismo. Definiendo (al igual que para los espacios lineales) el concepto de isomorfismo de módulos, fijamos con ello el conjunto de las operaciones principales, aunque de un modo implícito. Indicando estas operaciones explícitamente, obtenemos la siguiente definición de módulo unitario que sólo en la forma difiere de la definición del p. 4.1.

Se llama *módulo unitario* sobre un anillo K con el elemento unidad 1 un álgebra cuya signatura se compone del símbolo $+$ de una operación binaria y de los símbolos — y F_α $(\alpha \in K)$ de operaciones de una posición siempre que en este álgebra se cumplan las identidades

$$M_1\colon\ x+y=y+x;$$
$$M_2\colon\ x+(y+z)=(x+y)+z;$$
$$M_3\colon\ x+(y+(-y))=x;$$
$$M_4\colon\ F_\alpha(F_\beta(a))=F_{\alpha\beta}(a);$$
$$M_5\colon\ F_\alpha(a+b)=F_\alpha(a)+F_\alpha(b);$$
$$M_6\colon\ F_{\alpha+\beta}(a)=F_\alpha(a)+F_\beta(a);$$
$$M_7\colon\ F_1(a)=a.$$

Por consiguiente, la multiplicación de los elementos del conjunto principal (de los vectores) por cualquier elemento fijo $\alpha \in K$ se considera aquí como una operación principal independiente. Si el anillo principal K es infinito, la signatura del módulo es también infinita. Cambiando el anillo K, cambiamos también la signatura de la clase de módulos.

Existe, en general, otra forma de incluir la teoría de módulos en la teoría general de álgebras. Para ello los módulos se consideran como álgebras compuestas de dos conjuntos principales: el conjunto de los números y el conjunto de los vectores. La signatura consta ahora de las operaciones de adición de números, de paso al número opuesto, de adición de vectores, de paso al vector opuesto, de multiplicación de un número por un vector y de la operación 0-aria que despeja la unidad 1 (en total seis operaciones). Las identidades principales son en este caso las identidades que definen un anillo, las identidades que definen un grupo conmutativo y las identidades de 4° a 7° del p. 4.1.

Este nuevo concepto de módulo es diferente del anterior. Todos los módulos tienen, en el sentido nuevo, una misma signatura (las seis operaciones indicadas) y por esto resulta posible preguntar si son o no isomorfos unos módulos definidos sobre distintos anillos. Estos isomorfismos nuevos suelen llamarse, en diferencia de los definidos anteriormente, *antiisomorfismos*. En este orden de ideas se definen también los *antiautomorfismos* y los *antiendomorfismos* y otros conceptos análogos.

Ejemplos y problemas

1. Demuéstrese que la dimensión del espacio de todos los polinomios en una variable de grado no mayor que n es igual a $n+1$.

2. Los polinomios homogéneos en dos variables de grado n constituyen un espacio lineal de dimensión $n+1$.

3. ¿Cuál es la dimensión del espacio de los polinomios homogéneos en k variables de grado n?

4. Las matrices de m filas y de n columnas formadas por elementos de un cuerpo dado K constituyen un espacio lineal respecto a las operaciones usuales matriciales de adición y de multiplicación por número. Demuéstrese que las matrices en las que un elemento es igual a la unidad y todos los demás ele-

mentos son iguales a cero, constituyen una base de este espacio y que, por consiguiente, la dimensión de este espacio es igual a mn.

5. Las matrices simétricas, así como las matrices antisimétricas, de orden n formadas por elementos de un cuerpo K constituyen unos espacios lineales sobre K. Demuéstrese que las dimensiones de estos espacios son iguales a $\frac{n(n+1)}{2}$ y a $\frac{n(n-1)}{2}$, respectivamente.

§ 5. Coordenadas

5.1. Coordenadas de un vector. En el parágrafo anterior hemos considerado las propiedades generales elementales de los espacios lineales. Sin embargo, en las aplicaciones, además de conocer las propiedades generales, es importante saber definir los vectores en términos de números y poder reducir las operaciones vectoriales a operaciones con números. Este problema se resuelve introduciendo las coordenadas de un espacio vectorial.

Toda base de un espacio lineal $\mathfrak{L}$, cuyos vectores se toman en un orden determinado, se llamará *base de coordenados* o *sistema de coordenadas* de $\mathfrak{L}$. Por consiguiente, si

$$a_1, a_2, \ldots, a_n \tag{1}$$

es un sistema de coordenadas de $\mathfrak{L}$, estos mismos vectores, pero tomados en otro orden, representarán otro sistema de coordenadas de $\mathfrak{L}$. Hemos visto que todo vector a de $\mathfrak{L}$ puede ser representado unívocamente en la forma siguiente:

$$a = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n. \tag{2}$$

Los números $\alpha_1, \ldots, \alpha_n$ se llaman *coordenadas del vector* a en el sistema de coordenadas (1). La fila $[\alpha_1, \alpha_2, \ldots, \alpha_n]$ compuesta por las coordenadas del vector a, tomadas en un orden adecuado, se llama *fila de coordenadas* y se indica por $[a]$. Por consiguiente, una vez escogido en el espacio un sistema de coordenadas determinado, a todo vector corresponde una fila de coordenadas y, viceversa, para toda fila de longitud n se obtiene con la fórmula (2) un vector determinado a, para el cual esta fila es su fila de coordenadas.

Supongamos que $[\alpha_1, \ldots, \alpha_n]$ y $[\beta_1, \ldots, \beta_n]$ son las filas de coordenadas de los vectores a y b, es decir,

$$a = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n$$

y

$$b = \beta_1 a_1 + \beta_2 a_2 + \ldots + \beta_n a_n.$$

Es evidente que

$$\alpha a = (\alpha\alpha_1) a_1 + (\alpha\alpha_2) a_2 + \ldots + (\alpha\alpha_n) a_n$$

y

$$a+b=(\alpha_1+\beta_1)\,a_1+(\alpha_2+\beta_2)\,a_2+\ldots+(\alpha_n+\beta_n)\,a_n.$$

Empleando las reglas de operaciones con filas, estas igualdades pueden ser representadas en la forma

$$[\alpha a]=\alpha\,[a] \ \text{ y } \ [a+b]=[a]+[b].$$

Por consiguiente, *la fila de coordenadas de una suma de vectores es igual a la suma de las filas de coordenadas de los sumandos y la fila de coordenadas del producto de un número por un vector es igual al producto de este número por la fila de coordenadas del vector.*

Este resultado puede ser interpretado de la forma siguiente. Sea $\mathfrak{L}$ un espacio lineal de dimensión n sobre un cuerpo K. Sea $\mathfrak{L}_n$ el espacio de filas de longitud n formadas por elementos de K. Tomemos en $\mathfrak{L}$ un sistema de coordenadas determinado y pongamos en correspondencia a todo vector de $\mathfrak{L}$ su fila de coordenadas. Nuestro resultado significa que esta correspondencia es un isomorfismo entre $\mathfrak{L}$ y $\mathfrak{L}_n$. En particular, de aquí se desprende que los vectores linealmente independientes tienen filas de coordenadas linealmente independientes y que toda relación de dependencia lineal entre los vectores dados tiene lugar también para las filas de coordenadas de los mismos.

En un mismo espacio $\mathfrak{L}$ existen diferentes sistemas de coordenadas. Por esto surge la pregunta: ¿cómo varían las coordenadas de un vector al cambiar un sistema de coordenadas por otro? Para resolver este problema, tomemos en $\mathfrak{L}$ dos sistemas de coordenadas $a_1, a_2, \ldots, a_n$ y $a_1', a_2', \ldots, a_n'$ cualesquiera. Puesto que los vectores $a_1, \ldots, a_n$ constituyen una base de $\mathfrak{L}$, los vectores $a_1', \ldots, a_n'$ deben expresarse linealmente en términos de $a_1, \ldots, a_n$. Sean

$$\begin{aligned}
a_1' &= \tau_{11}a_1+\tau_{12}a_2+\ldots+\tau_{1n}a_n;\\
a_2' &= \tau_{21}a_1+\tau_{22}a_2+\ldots+\tau_{2n}a_n,\\
&\cdots\cdots\cdots\\
a_n' &= \tau_{n1}a_1+\tau_{n2}a_2+\ldots+\tau_{nn}a_n.
\end{aligned}$$

estas expresiones. La matriz

$$T=\begin{bmatrix}\tau_{11} & \tau_{12} & \ldots & \tau_{1n}\\ \tau_{21} & \tau_{22} & \ldots & \tau_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ \tau_{n1} & \tau_{n2} & \ldots & \tau_{nn}\end{bmatrix}$$

se llama *matriz del cambio* del sistema de coordenadas $a_1, \ldots, a_n$ por el sistema de coordenadas $a_1', \ldots, a_n'$. Tomemos un vector a cualquiera y sean $[\alpha_1, \ldots, \alpha_n]$ y $[\alpha_1', \ldots, \alpha_n']$ sus filas de coordenadas en los sistemas antiguo y nuevo de coordenadas. Es decir,

tenemos
$$a=\alpha_1 a_1+\alpha_2 a_2+\ldots+\alpha_n a_n$$
y
$$a=\alpha'_1 a'_1+\alpha'_2 a'_2+\ldots+\alpha'_n a'_n.$$

Tomando en la segunda de estas igualdades en lugar de los vectores $a'_1, \ldots, a'_n$ sus expresiones en términos de $a_1, \ldots, a_n$, obtenemos
$$a=(\alpha'_1\tau_{11}+\alpha'_2\tau_{21}+\ldots+\alpha'_n\tau_{n1})\,a_1+(\alpha'_1\tau_{12}+\ldots+\alpha'_n\tau_{n2})\,a_2+\ldots,$$
es decir
$$\begin{aligned}
\alpha_1&=\alpha'_1\tau_{11}+\alpha'_2\tau_{21}+\ldots+\alpha'_n\tau_{n1},\\
\alpha_2&=\alpha'_1\tau_{12}+\alpha'_2\tau_{22}+\ldots+\alpha'_n\tau_{n2},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
\alpha_n&=\alpha'_1\tau_{1n}+\alpha'_2\tau_{2n}+\ldots+\alpha'_n\tau_{nn}.
\end{aligned}$$

Estas igualdades son precisamente las fórmulas de *transformación de coordenadas* que buscábamos. Observando que la expresión para α_i:
$$\alpha_i=\alpha'_1\tau_{1i}+\alpha'_2\tau_{2i}+\ldots+\alpha'_n\tau_{ni}$$
representa el producto de la fila de coordenadas $[\alpha'_1, \ldots, \alpha'_n]$ por la i-ésima columna de la matriz T, vemos que todo el sistema de las fórmulas de transformación de coordenadas puede ser representado brevemente en la forma matricial:
$$[\alpha_1, \ldots, \alpha_n]=[\alpha'_1, \ldots, \alpha'_n]\,T.$$

Hemos obtenido la siguiente regla:

REGLA DE TRANSFORMACIÓN DE COORDENADAS. *La fila de coordenadas antigua de un vector es igual a la nueva multiplicada por la matriz del cambio.*

Demostremos el siguiente lema que en muchas ocasiones resulta útil.

LEMA. *Sean A y B dos matrices cuadradas de orden n formadas por elementos de un cuerpo K. Si para cualquier fila $[\xi_1, \ldots, \xi_n]$ de elementos de K resulta*
$$[\xi_1, \ldots, \xi_n]\,A=[\xi_1, \ldots, \xi_n]\,B, \tag{3}$$
se tiene $A=B$.

En efecto, sean α_{ij} los elementos de la matriz A y sean β_{ij} los elementos de la matriz $B\,(i, j=1, \ldots, n)$; entonces cualquiera que sea i para $\xi_i=1$ y $\xi_j=0$ $(i\neq j)$ de la igualdad (3) resulta que $\alpha_{ik}=\beta_{ik}$ $(i, k=1, \ldots, n)$, que es lo que se quería demostrar.

Consideremos ahora en un espacio lineal $\mathfrak{L}$ de n dimensiones dos sistemas de coordenadas $a_1, \ldots, a_n$ y $a'_1, \ldots, a'_n$. Podemos expresar o bien $a'_1, \ldots, a'_n$ en términos de $a_1, \ldots, a_n$:
$$\begin{aligned}
a'_1&=\tau_{11}a_1+\ldots+\tau_{1n}a_n,\\
&\ldots\ldots\ldots\ldots\\
a'_n&=\tau_{n1}a_1+\ldots+\tau_{nn}a_n,
\end{aligned} \tag{4}$$

o bien, al contrario, $a_1, \ldots, a_n$ en términos de $a'_1, \ldots, a'_n$:

$$a_1 = \sigma_{11}a'_1 + \ldots + \sigma_{1n}a'_n,$$
$$\cdots\cdots\cdots\cdots$$
$$a_n = \sigma_{n1}a'_1 + \ldots + \sigma_{nn}a'_n.$$

La matriz T, formada por los coeficientes τ_{ij}, es la matriz del cambio del sistema $a_1, \ldots, a_n$ por el sistema $a'_1, \ldots, a'_n$ y la matriz S, formada por los coeficientes σ_{ij}, es la matriz del *cambio recíproco* del sistema $a'_1, \ldots, a'_n$ por el sistema $a_1, \ldots, a_n$. Sea $[x]$ la fila de coordenadas de un vector cualquiera x en el sistema de coordenadas $a_1, \ldots, a_n$ y sea $[x]_1$ la fila de coordenadas del vector x en el sistema $a'_1, \ldots, a'_n$. Considerando que $a_1, \ldots, a_n$ es el sistema antiguo y aplicando la regla de transformación de coordenadas, obtenemos

$$[x] = [x]_1 T. \tag{5}$$

Al contrario, considerando que el sistema antiguo es $a'_1, \ldots, a'_n$, tendremos

$$[x]_1 = [x]\, S.$$

Tomando aquí en lugar de $[x]$ su expresión (5), obtenemos

$$[x]_1 = [x]_1\, TS$$

y de forma análoga

$$[x] = [x]\, ST.$$

Aquí el vector x y, por consiguiente, sus filas de coordenadas $[x]_1$ y $[x]$ son arbitrarias. En virtud del lema, obtenemos entonces

$$E = TS = ST \quad \text{y} \quad S = T^{-1}.$$

Es decir, *la matriz de la transformación de coordenadas siempre posee la inversa* que es la matriz de la transformación recíproca de coordenadas.

Nuestros razonamientos permiten aclarar un detalle relacionado con las matrices inversas que hemos dejado sin consideración en el p. 2.3. Según el p. 2.3, una matriz cuadrada S se llama inversa de la matriz T, si se cumplen dos igualdades:

$$ST = E \quad \text{y} \quad TS = E.$$

Si S satisface sólo la primera o sólo la segunda de estas igualdades, se dice que es *inversa a la izquierda* y, respectivamente, *inversa a la derecha*.

TEOREMA *Si una matriz cuadrada T formada por elementos de un cuerpo K posee la matriz inversa a la izquierda o a la derecha S, la matriz S es simplemente la inversa de T.*

Tomemos un espacio $\mathfrak{L}$ de n dimensiones con una base cualquiera $a_1, \ldots, a_n$. Calculemos según las fórmulas (4) los vectores $a'_1, \ldots, a'_n$.

Demostremos que son linealmente independientes. Sea

$$\lambda_1 a_1' + \ldots + \lambda_n a_n' = o.$$

Es obvio que esta relación equivale a la condición matricial

$$[\lambda_1, \ \ldots, \ \lambda_n] \cdot T = o. \tag{6}$$

Si T posee la matriz inversa a la derecha S, multiplicando a la derecha ambos miembros de la igualdad (6) por S y observando que $TS = E$, obtenemos $[\lambda_1, \ \ldots, \ \lambda_n] = o$. Por consiguiente, los vectores $a_1', \ \ldots, \ a_n'$ son linealmente independientes y el sistema $a_1', \ \ldots, \ a_n'$ puede ser considerado como un nuevo sistema de coordenadas de $\mathfrak{L}$, con la particularidad de que la matriz T será la matriz del cambio. La matriz del cambio posee la matriz inversa T^{-1}. Multiplicando la relación $TS = E$ a la izquierda por T^{-1}, obtenemos $S = T^{-1}$. Hemos demostrado que una matriz inversa a la derecha es simplemente la matriz inversa. Análogamente se demuestra que una matriz inversa a la izquierda es también simplemente la inversa.

De los razonamientos realizados se desprende, en particular, que toda matriz invertible de orden n es una matriz del cambio de determinados sistemas de coordenadas.

5.2. Rangos de matrices. Consideremos un espacio lineal $\mathfrak{L}$ de dimensión finita n sobre un cuerpo K. Tomemos en $\mathfrak{L}$ una base cualquiera $a_1, \ \ldots, \ a_n$ a título del sistema de coordenadas y supongamos que los vectores $x_1, \ \ldots, \ x_m$ tienen respectivamente las siguientes filas de coordenadas

$$\begin{gathered} [x_1] = [\alpha_{11}, \ \ldots, \ \alpha_{1n}], \\ \cdots\cdots\cdots\cdots \\ [x_m] = [\alpha_{m1}, \ \ldots, \ \alpha_{mn}]. \end{gathered}$$

Formemos a partir de estas filas la matriz

$$A = \begin{bmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \alpha_{m1} & \ldots & \alpha_{mn} \end{bmatrix}.$$

Sabemos ya que el número máximo de vectores linealmente independientes en el sistema $x_1, \ \ldots, \ x_m$ es igual al número máximo de filas linealmente independientes en el sistema $[x_1], \ \ldots, \ [x_m]$, es decir, coincide con el número máximo de filas linealmente independientes de la matriz A.

El número máximo de filas linealmente independientes de una matriz arbitraria A, formada por elementos de un cuerpo dado K, se llama *rango de la matriz* A, o, con más precisión, *rango según las filas*.

De este modo, el número máximo de vectores linealmente independientes entre los vectores $x_1, \ \ldots, \ x_m$ es igual al rango de la matriz formada por las filas de coordenadas de estos vectores.

Volvamos a considerar una matriz arbitraria A de m filas y de n columnas. Escojamos en A cualesquiera k filas y k columnas. Los elementos de la matriz A, que se hallan en los cruces de estas filas y columnas, tomadas en su orden natural, forman una matriz cuadrada de orden k, que se llama menor de orden k de la matriz A. Considerando todos los menores de primero, segundo, etc. órdenes, de la matriz A, podremos ver que una parte de los mismos serán matrices invertibles y otra parte, matrices no invertibles. El orden máximo de los menores invertibles de la matriz A se llama *rango* de la misma *según los menores* o *rango de menores*.

Los menores de primer orden son los elementos de la matriz A. Puesto que la matriz A se toma sobre un cuerpo, el hecho de que un menor de primer orden no sea invertible significa simplemente que este menor está formado por el número cero. Luego, una matriz A no posee menores invertibles si, y sólo si, está compuesta de ceros. En este caso se dice que el rango de menores de la matriz A es igual a cero. Observemos que el rango según las filas de la matriz nula es también igual a cero, ya que un sistema de vectores nulos no posee subsistemas linealmente independientes. Pretendemos demostrar ahora que el rango según las filas de una matriz arbitraria coincide con su rango de menores. Antes de obtener este resultado, consideremos las matrices cuadradas.

TEOREMA 1 *Para que una matriz cuadrada $A=\|\alpha_{ij}\|_{nn}$, formada por elementos de un cuerpo K, sea invertible es necesario y suficiente que sus filas sean linealmente independientes.*

Sea a_i $(i=1, \ldots, n)$ la i-ésima fila de la matriz. Entonces toda relación de dependencia lineal

$$\lambda_1 a_1 + \ldots + \lambda_n a_n = o \tag{1}$$

entre las filas de la matriz A puede ser representada en la forma

$$[\lambda_1, \ldots, \lambda_n] \cdot A = o. \tag{2}$$

Siendo la matriz A invertible y multiplicando ambos miembros de la igualdad (2) por A^{-1}, obtenemos $[\lambda_1, \ldots, \lambda_n] = o$ y con esto queda demostrado que las condiciones del teorema 1 son necesarias.

Recíprocamente, supongamos que las filas $a_1, \ldots, a_n$ de la matriz A son linealmente independientes. Sea $\mathfrak{L}$ el espacio vectorial de todas las filas de longitud n formadas por elementos del cuerpo K. Las filas

$$\begin{aligned} e_1 &= [1, 0, \ldots, 0], \\ e_2 &= [0, 1, \ldots, 0], \\ &\ldots\ldots\ldots\ldots \\ e_n &= [0, 0, \ldots, 1] \end{aligned}$$

constituyen una base de $\mathfrak{L}$. Por otra parte, el espacio $\mathfrak{L}$ es de n dimensiones y, por hipótesis, los vectores $a_1, \ldots, a_n$ son linealmente independientes. Luego, los vectores $a_1, \ldots, a_n$ forman también una base del espacio $\mathfrak{L}$. Considerando la base $e_1, \ldots, e_n$ como el sistema de coordenadas inicial en $\mathfrak{L}$ y la base $a_1, \ldots, a_n$ como un nuevo sistema de coordenadas, vemos que A es la matriz del cambio del primer sistema de coordenadas por el segundo y que, por consiguiente (véase el p. 5.1), la matriz A es invertible.

COROLARIO. *El determinante de una matriz cuadrada, formada por elementos de un cuerpo conmutativo, es igual a cero si, y sólo si, las filas de esta matriz son linealmente dependientes.*

Efectivamente, para matrices cuadradas, formadas por elementos de un cuerpo conmutativo, la invertibilidad equivale a la regularidad, es decir, a que el determinante de la matriz sea diferente de cero.

Tomando ahora en consideración la observación hecha al principio de este punto, vemos que un sistema de n vectores de un espacio vectorial de n dimensiones sobre un cuerpo conmutativo forma una base de este espacio si, y sólo si, es diferente de cero el determinante de la matriz formada por las filas de coordenadas de los vectores indicados.

El teorema 1 y su corolario tratan de las filas de una matriz. Sin embargo, estas proposiciones son válidas también para las columnas, siempre que la expresión «relación de dependencia lineal de columnas» sea interpretada como una relación de dependencia lineal respecto a la multiplicación a la derecha de las columnas por los elementos de K. Por ello, siempre que no se diga lo contrario, la multiplicación de filas por elementos de K significará la multiplicación a la izquierda de las filas por los elementos, mientras que la multiplicación de columnas por elementos de K significará la multiplicación a la derecha. Está claro que este convenio sobra si K es un cuerpo conmutativo. Para el caso de cuerpos K no conmutativos el convenio aceptado es de importancia.

Es fácil comprobar que todos los razonamientos que hemos empleado para demostrar el teorema 1 permanecen válidos si la palabra «fila» es sustituida en los mismos por la palabra «columna». Así obtenemos el resultado siguiente.

TEOREMA 1a *Para que una matriz cuadrada, formada por elementos de un cuerpo, sea invertible es necesario y suficiente que sus columnas sean linealmente independientes. El determinante de una matriz cuadrada, formada por elementos de un cuerpo conmutativo, es igual a cero si, y sólo si, las columnas de la matriz son linealmente dependientes.*

Es evidente que si se consideran matrices sobre un cuerpo conmutativo, el teorema 1a se obtiene del teorema 1 pasando simplemente a la matriz transpuesta.

TEOREMA 2. *Sea $A = \|\alpha_{ij}\|_{mn}$ una matriz cualquiera formada por elementos de un cuerpo K. Los rangos de la matriz A según las filas, las columnas y los menores no varían si A se somete a una de las transformaciones siguientes:*

a) *se cambian entre sí cualesquiera filas de la matriz A;*

b) *se cambian entre sí cualesquiera columnas de la matriz A;*

c) *se suman a los elementos de una de las filas de la matriz A los elementos correspondientes de otra fila cualquiera multiplicados a la izquierda por un factor arbitrario fijo $\lambda \in K$;*

d) *se suman a los elementos de una de las columnas de la matriz A los elementos correspondientes de otra columna cualquiera multiplicados a la derecha por un factor arbitrario $\lambda \in K$.*

El rango de la matriz A según las filas es igual al número máximo de elementos linealmente independientes en el conjunto de sus filas $a_1, \ldots, a_m$. Al cambiar el orden de las filas de A sólo alteramos la numeración de las últimas; pero los conceptos de dependencia o independencia lineal de vectores no están relacionados con la numeración. Así mismo, está claro que el rango de la matriz A según las filas no varía al cambiar sus columnas. En efecto, sean $b_1, \ldots, b_m$ las filas de la matriz nueva obtenida después del cambio de columnas. Supongamos que entre las filas de la matriz A existía una relación de dependencia lineal $\lambda_1 a_1 + \ldots + \lambda_m a_m = o$; es evidente que las filas nuevas verificarán la relación análoga $\lambda_1 b_1 + \ldots + \lambda_m b_m = o$ y que, viceversa, la última relación implicará la anterior. Lo mismo ocurrirá si a una columna cualquiera de la matriz A se agrega otra columna multiplicada a la derecha por λ. Finalmente, el número máximo de filas linealmente independientes no cambiará si la i-ésima fila de la matriz A es sustituida por la fila $a_i + \lambda a_j$, ya que los vectores del sistema $a_1, \ldots, a_i + \lambda a_j, \ldots, a_j, \ldots, a_m$ se expresan linealmente en términos de los vectores $a_1, \ldots, a_m$ y los vectores del último sistema se expresan linealmente en términos de los vectores del primer sistema.

Hemos visto que el rango de la matriz A *según las filas* no varía en las transformaciones indicadas en el teorema. Razonamientos análogos demuestran que el rango de la matriz A *según las columnas* no varía en las transformaciones a), b), c) y d). Pasemos ahora a demostrar que tampoco varía el rango de la matriz A *según los menores*, lo cual requerirá una mayor atención.

Supongamos que en la matriz A se cambian entre sí cualesquiera filas o columnas. En este caso los menores de la matriz nueva se obtienen de los menores de la matriz antigua mediante cambios de filas o columnas y sólo debemos comprobar que al cambiar filas o columnas en una matriz invertible de nuevo obtenemos una matriz invertible.

Supongamos, pues, que en una matriz cuadrada $B = \| \beta_{ij} \|_{ss}$ se han cambiado entre sí la primera y la i-ésima filas. El cálculo directo deja constancia de que la matriz nueva puede ser representada en la forma DB, donde

$$D = E_{11} + E_{22} + \ldots + E_{i-1,\, i-1} + E_{i1} + E_{i+1,\, i+1} + \ldots + E_{ss}.$$

Aquí E_{ij} es la matriz cuyo elemento de la i-ésima fila y j-ésima columna es la unidad, mientras que todos los demás elementos suyos son ceros. Mediante cálculo directo podemos comprobar que

$$DD=E \quad \text{y} \quad D^{-1}=D.$$

Por esto, si la matriz B tiene inversa, la matriz $B^{-1}D$ será la inversa de DB. Análogamente, la matriz BD se obtiene de B cambiando entre sí la primera y la i-ésima columnas. Si B es invertible, la matriz BD también tiene la inversa, a saber, DB^{-1}.

Hemos demostrado que el rango de menores no varía al cambiar entre sí filas y columnas. Supongamos que el rango de menores de B es igual a r y que a la i-ésima fila de B se ha agregado la j-ésima fila multiplicada por λ. Por hipótesis, la matriz B posee un menor invertible de orden r perteneciente a r filas y a r columnas determinadas. Para no complicar sin necesidad la notación, aceptemos que este menor pertenece a las r filas primeras y a las r columnas primeras de la matriz B. Si la i-ésima fila que se altera no pertenece a las r primeras filas, el menor indicado de orden r continuará siendo un menor de la matriz nueva. Por esto el rango de la matriz nueva es $\geqslant r$. Supongamos que la i-ésima fila es una de las r primeras filas. Consideremos los elementos de las matrices nueva y antigua que se encuentran en las r primeras columnas, en las r primeras filas y en la j-ésima fila ($j>r$). Obtendremos las matrices

$$P=\begin{bmatrix} \beta_{11} & \dots & \beta_{1r} \\ \cdot & \cdot & \cdot \\ \beta_{i1} & \dots & \beta_{ir} \\ \cdot & \cdot & \cdot \\ \beta_{r1} & \dots & \beta_{rr} \\ \beta_{j1} & \dots & \beta_{jr} \end{bmatrix} \quad \text{y} \quad Q=\begin{bmatrix} \beta_{11} & \dots & \beta_{1r} \\ \cdot & \cdot & \cdot \\ \beta_{i1}+\lambda\beta_{j1} & \dots & \beta_{ir}+\lambda\beta_{jr} \\ \cdot & \cdot & \cdot \\ \beta_{r1} & \dots & \beta_{rr} \\ \beta_{j1} & \dots & \beta_{jr} \end{bmatrix}.$$

En virtud del teorema 1, las primeras r filas de la matriz P son linealmente independientes y, por ello, el rango de P según las filas es igual a r. De acuerdo con lo demostrado, de aquí se deduce que el rango de la matriz Q según las filas es también igual a r. Las primera, ..., $(i-1)$-ésima, $(i+1)$-ésima, ..., r-ésima filas de la matriz Q son desde luego linealmente independientes, ya que forman parte del sistema linealmente independiente de las r primeras filas de la matriz Q. Por esto, o bien las r primeras filas de la matriz Q son linealmente independientes o bien la i-ésima fila se expresa linealmente en términos de las restantes y, por consiguiente, son linealmente independientes las restantes r filas de la matriz Q. En otras palabras, uno de los menores

$$\begin{bmatrix} \beta_{11} & \dots & \beta_{1r} \\ \cdot & \cdot & \cdot \\ \beta_{i1}+\lambda\beta_{j1} & \dots & \beta_{ir}+\lambda\beta_{jr} \\ \cdot & \cdot & \cdot \\ \beta_{r1} & \dots & \beta_{rr} \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} \beta_{11} & \dots & \beta_{1r} \\ \cdot & \cdot & \cdot \\ \beta_{i-1,\,1} & \dots & \beta_{i-1,\,r} \\ \beta_{i+1,\,1} & \dots & \beta_{i+1,\,r} \\ \cdot & \cdot & \cdot \\ \beta_{r1} & \dots & \beta_{rr} \\ \beta_{j1} & \dots & \beta_{jr} \end{bmatrix}$$

de la matriz Q está formado por filas linealmente independientes. En virtud del teorema 1, este menor es invertible y el rango de menores de la matriz Q es no menor de r.

Hemos demostrado que la transformación de tipo c) no disminuye el rango de menores de una matriz. Pero, si la matriz C se obtiene de la matriz A agregando a su i-ésima fila la j-ésima fila multiplicada por λ, la matriz A se obtiene de C sumando a la i-ésima fila de la matriz C su j-ésima fila multiplicada por $-\lambda$. En una y otra transformación los rangos de menores no disminuyen y, por consiguiente, no varían. Resta considerar el caso de la transformación d); pero para ella la invariabilidad del rango de la matriz A se demuestra de la misma forma que la invariabilidad en la transformación c).

Del teorema 2 mediante razonamientos sencillos se deduce el siguiente teorema principal.

TEOREMA 3 (SOBRE EL RANGO DE UNA MATRIZ). *Para una matriz arbitraria $A=\|\alpha_{ij}\|_{mn}$, formada por elementos de un cuerpo, el rango según las filas, el rango según las columnas y el rango según los menores coinciden.*

La proposición es evidente si todos los elementos de la matriz A son iguales a cero. Por esto, supondremos que A contiene un elemento α_{ij} diferente de cero. Cambiando entre sí la primera y la i-ésima filas y la primera y la j-ésima columnas de la matriz A, obtendremos una matriz $B=\|\beta_{ij}\|$ en la que $\beta_{11}=\alpha_{ij}\neq 0$ y, además, en virtud del teorema 2, los tres rangos de la matriz B serán iguales a los rangos correspondientes de la matriz A. Sumando ahora a la segunda, ..., m-ésima filas de la matriz B su primera fila multiplicada a la izquierda por $-\beta_{21}\beta_{11}^{-1}$, ..., $-\beta_{m1}$, β_{11}^{-1}, respectivamente, obtendremos una matriz de tipo

$$C_1=\begin{vmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ 0 & \gamma_{22} & \dots & \gamma_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & \gamma_{m2} & \dots & \gamma_{mn} \end{vmatrix},$$

con la particularidad de que los tres rangos de la matriz C_1 serán los mismos que los de la matriz A. Si todos los γ_{ij} son iguales a cero, no realizamos ninguna transformación más. En cambio, si $\gamma_{ij} \neq 0$, $(i \geqslant 2,\ j \geqslant 2)$, cambiamos entre sí la segunda y la i-ésima filas y la segunda y la j-ésima columnas de la matriz C_1, obteniendo así una matriz en la que el elemento γ_{22} sea diferente de cero. Sumando entonces a la tercera, ..., m-ésima filas la segunda fila multiplicada por $-\gamma_{32}\gamma_{22}^{-1}$, ..., $-\gamma_{m2}\gamma_{22}^{-1}$, respectivamente, obtendremos una matriz de tipo

$$C_2 = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} & \dots & \delta_{1n} \\ 0 & \delta_{22} & \delta_{23} & \dots & \delta_{2n} \\ 0 & 0 & \delta_{33} & \dots & \delta_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \delta_{m3} & \dots & \delta_{mn} \end{bmatrix} \quad (\delta_{11}\delta_{22} \neq 0).$$

Los rangos de la matriz C_2 también coincidirán con los rangos correspondientes de la matriz A. Continuando este proceso, al cabo de un número de pasos no mayor que m, obtendremos una matriz de tipo

$$C_r = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1r} & \mu_{1,\,r+1} & \dots & \mu_{1n} \\ 0 & \mu_{22} & \dots & \mu_{2r} & \mu_{2,\,r+1} & \dots & \mu_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \mu_{rr} & \mu_{r,\,r+1} & \dots & \mu_{rn} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \quad (\mu_{11} \dots \mu_{rr} \neq 0).$$

cuyos rangos coincidirán también con los rangos correspondientes de la matriz A. Sin embargo, de la forma de la matriz C_r se deduce directamente que los tres rangos de esta matriz son iguales a un mismo número r. Por esto, los tres rangos de la matriz A también tienen un mismo valor r.

En el p. 2.3 se ha demostrado que una matriz cuadrada con elementos de un cuerpo conmutativo es invertible si, y sólo si, tiene el determinante diferente de cero. Por esto, para matrices sobre un cuerpo conmutativo el teorema 3 se puede enunciar del modo siguiente:

TEOREMA 3a. *En toda matriz con elementos de un cuerpo conmutativo el número máximo de filas linealmente independientes es igual al número máximo de columnas linealmente independientes e igual al orden máximo de sus menores de determinante diferente de cero.*

La determinación práctica del rango de una matriz se realiza generalmente aplicando el método indicado en la demostración del teorema 3 o alguna de sus modificaciones convenientes.

Con frecuencia resultan útiles las dos observaciones siguientes que se deducen directamente del teorema 3.

TEOREMA 4. *Cualesquiera que sean dos matrices* $A=\|\alpha_{ij}\|_{mn}$, *y* $B=\|\beta_{jk}\|_{np}$ *con elementos de un cuerpo arbitrario el rango del producto* AB *es no mayor que el rango de la matriz* A *y no mayor que el rango de la matriz* B.

Efectivamente, de la definición del producto se ve que las filas de la matriz AB son combinaciones lineales de las filas de la matriz B y, por consiguiente, el número máximo de filas linealmente independientes de la matriz AB no puede pasar del número máximo de filas linealmente independientes de la matriz B. Análogamente, las columnas de la matriz AB son combinaciones lineales de las columnas de la matriz A y, por consiguiente, el rango de la matriz AB según las columnas es no mayor que el rango según las columnas de la matriz A.

TEOREMA 5 *Si los elementos de la matriz* A *pertenecen a un cuerpo conmutativo, su rango coincide con el rango de la matriz transpuesta* A'.

Todos los menores de la matriz A' se obtienen por transposición de los menores de la matriz A. Puesto que los elementos de los menores se toman en un cuerpo conmutativo, de la transposición de un menor invertible resulta un menor invertible y, por consiguiente, el rango (según los menores) de la matriz A' coincide con el rango (según los menores) de la matriz A.

En los ejercicios de la pág. 101 se da un ejemplo de una matriz invertible, formada por elementos de un cuerpo no conmutativo, cuya matriz transpuesta no es invertible.

5.3. Sistemas generales de ecuaciones lineales. Consideremos un cuerpo K y un sistema arbitrario de ecuaciones de tipo

$$\left.\begin{array}{l}\alpha_{11}\xi_1+\alpha_{12}\xi_2+\ldots+\alpha_{1n}\xi_n=\beta_1,\\ \alpha_{21}\xi_1+\alpha_{22}\xi_2+\ldots+\alpha_{2n}\xi_n=\beta_2,\\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\\ \alpha_{m1}\xi_1+\alpha_{m2}\xi_2+\ldots+\alpha_{mn}\xi_n=\beta_m,\end{array}\right\}\qquad(1)$$

donde α_{ij} y β_i son elementos dados de K y $\xi_1, \xi_2, \ldots, \xi_n$ son incógnitas, cuyo número n puede ser mayor, menor o igual al número m de ecuaciones. En el p. 2.4 ha sido expuesto uno de los algoritmos de la resolución de estos sistemas, el método de eliminación de incógnitas. El teorema sobre el rango de una matriz, demostrado en el punto anterior, ofrece un acceso diferente al estudio de los sistemas de ecuaciones lineales.

Los coeficientes α_{ij} y los términos independientes β_i de las ecuaciones (1) se pueden disponer de un modo natural formando dos matrices:

$$A=\begin{bmatrix}\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn}\end{bmatrix} \quad \text{y} \quad B=\begin{bmatrix}\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} & \beta_1\\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} & \beta_2\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} & \beta_m\end{bmatrix},$$

llamadas *matriz principal* y *matriz ampliada* del sistema (1). A toda ecuación del sistema (1) corresponde una fila determinada en las matrices A y B. El intercambio de ecuaciones en el sistema (1) lleva al intercambio correspondiente de filas en las matrices A y B y la modificación de la numeración de las incógnitas lleva al intercambio de columnas en las matrices indicadas.

Se dice que la i-ésima ecuación del sistema (1) depende linealmente de las i_1-ésima, ..., i_s-ésima ecuaciones de este sistema, si la i-ésima fila de la matriz B es una combinación lineal de sus i_1-ésima, ..., i_s-ésima filas. Puesto que el cuerpo K no se supone conmutativo, las filas de las matrices se multiplican por elementos de K siempre a la izquierda, mientras que las columnas, a la derecha (véase el p. 5.2).

Se llama *solución* del sistema de ecuaciones (1) una sucesión $\xi_1^0, \dots, \xi_n^0$ de elementos del cuerpo K que al ser introducidos en las ecuaciones (1) en lugar de las letras $\xi_1, \dots, \xi_n$ hacen válidas todas las igualdades. Si disponemos los elementos $\xi_1, \dots, \xi_n$ en una columna $[\xi_1, \dots, \xi_n]' = x$, el sistema (1) puede ser representado en la forma matricial

$$A\cdot x = b, \tag{2}$$

donde b es la columna $[\beta_1, \dots, \beta_n]'$ de los términos independientes.

A) *Si la i-ésima fila de la matriz B es una combinación lineal de sus restantes filas, entonces suprimiendo en el sistema* (1) *la i-ésima ecuación obtendremos un sistema reducido de ecuaciones que tiene el mismo conjunto de soluciones que el sistema* (1).

Está claro que toda solución del sistema (1) es también una solución del sistema reducido. Recíprocamente, supongamos que la i-ésima fila de la matriz B es igual a la suma de las i_1-ésima, ..., i_s-ésima filas multiplicadas a la izquierda, respectivamente, por los elementos $\lambda_1, \dots, \lambda_s$ de $K (i \neq i_1, \dots, i_s)$ y supongamos que los valores $\xi_1^0, \dots, \xi_n^0$ satisfacen las i_1-ésima, ..., i_s-ésima ecuaciones de (1). Entonces, multiplicando estas ecuaciones a la izquierda por $\lambda_1, \dots \lambda_s$, respectivamente, y sumándolas término por término, obtendremos la i-ésima ecuación.

B) *Si el rango (según las filas) de la matriz B es igual a r, el sistema de ecuaciones* (1) *contiene un subsistema de r ecuaciones linealmente independientes que posee el mismo conjunto de soluciones que el sistema* (1).

Puesto que el rango de la matriz B es igual a r, la matriz B posee r filas linealmente independientes y todas sus filas restantes se expresan linealmente en términos de éstas. Tomando las ecuaciones que corresponden a las r filas linealmente independientes señaladas, obtendremos el subsistema deseado.

C) *Si el sistema de ecuaciones* (1) *tiene una solución* $\xi_1^0, \ldots, \xi_n^0$, *el rango de la matriz ampliada B coincide con el rango de la matriz principal A.*

Si las igualdades (1) son válidas para $\xi_1^0, \ldots, \xi_n^0$, esto demuestra que la última columna de la matriz B es igual a la suma de sus primera, ..., n-ésima columnas multiplicadas a la derecha por $\xi_1^0, \ldots, \xi_n^0$, respectivamente. Por esto, el número máximo de columnas linealmente independientes de la matriz B es igual al número máximo de columnas linealmente independientes de la matriz A, es decir, el rango según las columnas de la matriz B es igual al rango según las columnas de la matriz A. Las palabras «según las columnas» pueden ser omitidas, ya que en el p. 5.2 se ha demostrado que el rango según las columnas coincide con el rango según las filas.

De las proposiciones B) y C) resulta fácilmente el siguiente teorema principal.

TEOREMA DE KRONECKER–CAPELLI. [1]) *Para que el sistema de ecuaciones* (1) *tenga solución, es necesario y suficiente que el rango de la matriz ampliada sea igual al rango de la matriz principal de este sistema. Si los rangos de las matrices principal y ampliada coinciden con el número de incógnitas, el sistema tiene una solución única. Si el rango r de las matrices principal y ampliada es inferior al número n de incógnitas, el sistema* (1) *tiene más de una solución y es equivalente a un sistema de tipo*

$$\left.\begin{array}{l} \xi_{i_1} = \gamma_{11}\xi_{j_1} + \cdots + \gamma_{1,\,n-r}\xi_{j_{n-r}} + \gamma_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \xi_{i_r} = \gamma_{r1}\xi_{j_1} + \cdots + \gamma_{r,\,n-r}\xi_{j_{n-r}} + \gamma_r, \end{array}\right\} \qquad (3)$$

donde $(i_1, \ldots, i_r, j_1, \ldots, j_{n-r})$ *es una permutación adecuada de los números* $1, 2, \ldots, n$.

En otras palabras, en el último caso entre las incógnitas $\xi_1, \ldots, \xi_n$ se pueden escoger $n-r$ incógnitas $\xi_{j_1}, \ldots, \xi_{j_{n-r}}$, llamadas libres, y dándoles valores arbitrarios de K se pueden encontrar para las restantes incógnitas unos valores únicos que satisfagan el sistema (1).

La necesidad de las condiciones se deduce de la proposición C). Supongamos, por esto, que los rangos de las matrices A y B coinciden y son iguales a r. La matriz A contiene un total de n columnas, de modo que $r \leqslant n$. Por hipótesis, la matriz A posee un

[1]) En algunos libros españoles este teorema se conoce como el teorema de Rouché—Frobenius. *(N. del Tr.)*

menor invertible M de orden r (véase el p. 5.2) que pertenece a las k_1-ésima, ..., k_r-ésima filas y a las i_1-ésima, ..., i_r-ésima columnas de la matriz A. De aquí se deduce que la k_1-ésima, ..., la k_r-ésima filas de la matriz B son linealmente independientes, mientras que sus filas restantes son combinaciones lineales de las filas señaladas. Por esto, dejando solamente la k_1-ésima, ..., la k_r-ésima ecuaciones del sistema (1) y suprimiendo todas las demás ecuaciones, obtendremos un sistema reducido que tendrá las mismas soluciones que el sistema (1). Dejando ahora en los primeros miembros de cada una de las ecuaciones del sistema reducido los términos que contienen las incógnitas $\xi_{i_1}, \ldots, \xi_{i_r}$ y pasando al otro miembro todos los demás términos, llevaremos el sistema reducido a la forma

$$M \cdot \begin{bmatrix} \xi_{i_1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{i_r} \end{bmatrix} = \begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_r \end{bmatrix}, \tag{4}$$

donde

$$c_s = -\alpha_{k_s, j_1}\xi_{j_1} - \ldots - \alpha_{k_s, j_{n-r}}\xi_{j_{n-r}} + \beta_{k_s} \quad (s = 1, \ldots, r).$$

Por hipótesis, el menor M es una matriz invertible. Multiplicando ambos miembros de la igualdad (4) a la izquierda por M^{-1}, obtenemos

$$\begin{bmatrix} \xi_{i_1} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{i_r} \end{bmatrix} = M^{-1} \begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_r \end{bmatrix}.$$

Realizando la multiplicación en el segundo miembro, llegamos al sistema de tipo (3). El teorema queda demostrado.

El sistema de ecuaciones (1) suele llamarse *determinado*, *indeterminado* y *contradictorio* según tenga, respectivamente, solución única, más de una solución y no tenga solución. Notemos que de acuerdo a esta terminología un sistema, que no es determinado, es o bien indeterminado o bien contradictorio.

COROLARIO 1. *Si el número de ecuaciones del sistema* (1) *es inferior al número de incógnitas, el sistema* (1) *es o bien indeterminado o bien contradictorio.*

COROLARIO 2. *Un sistema de n ecuaciones lineales con n incógnitas es determinado si, y sólo si, la matriz principal de este sistema es invertible.*

Ambos corolarios se desprenden directamente del teorema de Kronecker—Capelli.

Hemos considerado el sistema de ecuaciones lineales (1) en el que los coeficientes aparecen a la izquierda de las incógnitas. ¿Qué puede decirse acerca del sistema de ecuaciones

$$\left.\begin{array}{l}\xi_1\alpha_{11}+\xi_2\alpha_{21}+\ldots+\xi_n\alpha_{n1}=\beta_1,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \xi_1\alpha_{1m}+\xi_2\alpha_{2m}+\ldots+\xi_n\alpha_{nm}=\beta_m\end{array}\right\} \qquad (5)$$

en el que los coeficientes figuran a la derecha de las incógnitas? Este sistema puede ser representado en la forma matricial

$$[\xi_1, \ldots, \xi_n]\cdot A=[\beta_1, \ldots, \beta_m],$$

donde $A=\|\alpha_{ij}\|$ es la matriz de los coeficientes de las incógnitas en el sistema (5). Prestemos atención a que en la matriz A los coeficientes de la i-ésima ecuación forman la i-ésima columna y no la i-ésima fila como sucedía anteriormente. Todos los razonamientos, que hemos realizado al demostrar el teorema de Kronecker — Capelli, siguen siendo válidos también para el sistema (5), siempre que se cambien entre sí en ellos las palabras «columna» «y «fila» y siempre que por la matriz ampliada del sistema (5) se comprenda la matriz

$$B=\begin{bmatrix}\alpha_{11} & \ldots & \alpha_{1m}\\ \cdot & \cdot\ \cdot\ \cdot & \cdot\\ \alpha_{n1} & \ldots & \alpha_{nm}\\ \beta_1 & \ldots & \beta_m\end{bmatrix}.$$

Por consiguiente, para que el sistema (5) sea compatible es necesario y suficiente que el rango de la matriz B coincida con el rango de la matriz A. Si los rangos de las matrices A y B coinciden y son iguales al número de incógnitas, el sistema (5) tiene solución única. Si los rangos de las matrices A y B coinciden y son iguales a un número r menor que el número n de las incógnitas, entre las incógnitas podrán encontrarse $n-r$ incógnitas libres $\xi_{j_1}, \ldots, \xi_{j_{n-r}}$ y todas las demás incógnitas $\xi_{i_1}, \ldots, \xi_{i_r}$ se expresarán en términos de éstas por fórmulas de tipo

$$\xi_{i_s}=\xi_{j_1}\gamma_{1s}+\ldots+\xi_{j_{n-r}}\gamma_{n-r,\,s}+\gamma_s \quad (s=1, \ldots, r).$$

Está claro que toda diferencia entre los sistemas (1) y (5) desaparece, si se consideran ecuaciones lineales sobre un cuerpo conmutativo. En el caso de cuerpos no conmutativos K (por ejemplo, para los cuaternios (véase el p. 1.5)), además de los sistemas standard de tipo (1) y (5) se consideran también ecuaciones de tipo «mixto» como es, por ejemplo, la ecuación $\alpha\xi-\xi\alpha=0$. Sin embargo, los problemas relacionados con la resolución de estas ecuaciones tienen un carácter específico y se salen de los márgenes del Algebra lineal propiamente dicha.

Complementos y ejercicios

1. En el espacio de filas de longitud tres sobre el cuerpo de los números racionales se toma un sistema de coordenadas formado por las filas [1, 3, 5], [6, 3, 2] y [3, 1, 0]. ¿Qué filas de coordenadas tienen en este sistema los vectores [3, 7, 1], [0, 0, 1], [2, 3, 5] y [1, 1, 1]?

2. El espacio $\mathfrak{L}$ está formado por los polinomios en λ de grado no mayor que n. Demuéstrese que los polinomios $1, \lambda-1, (\lambda-1)^2, \ldots, (\lambda-1)^n$ constituyen una base de $\mathfrak{L}$. Hállense en esta base las filas de coordenadas de los polinomios $2-3\lambda+\lambda^2$ y λ^n.

3. En el plano se toma un sistema de coordenadas, compuesto por dos vectores mutuamente perpendiculares de longitud 1, y después se realiza una transformación de coordenadas de matriz $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$. ¿Qué condiciones deben verificar los números α, β, γ y δ para que los nuevos vectores coordenados sean mutuamente perpendiculares y de longitud 1?

4. Hállense los rangos de las matrices

$$\begin{bmatrix} 1 & 0 & 2 & 5 & 7 \\ -2 & 1 & 1 & 2 & 3 \\ 2 & 1 & 7 & 17 & 24 \\ 1 & 3 & 15 & 41 & 55 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ x & y & z & u \\ x^3 & y^3 & z^3 & u^3 \\ x^5 & y^5 & z^5 & u^5 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1+\lambda & 1 & 1 & 1 & 1 \\ 1 & 1+\lambda & 1 & 1 & 1 \\ 1 & 1 & 1 & 1+\lambda & 1 \\ 1 & 1 & 1 & 1 & 1+\lambda \end{bmatrix}.$$

5. Demuéstrese que en el cuerpo de los cuaternios K (p.1.4) el sistema de ecuaciones lineales «a la derecha»

$$\left.\begin{aligned} i\xi_1 + j\xi_2 &= 1, \\ k\xi_1 - \xi_2 &= i \end{aligned}\right\}$$

tiene solución única, mientras que el sistema de ecuaciones «a la izquierda»

$$\left.\begin{aligned} \xi_1 i + \xi_2 j &= 1, \\ \xi_1 k - \xi_2 &= i \end{aligned}\right\}$$

no tiene soluciones. Tenemos así un ejemplo de una matriz invertible sobre un cuerpo, cuya matriz transpuesta no es invertible.

6. Hasta el momento entendíamos la independencia lineal de filas como la independencia lineal respecto a la multiplicación de filas a la izquierda por los elementos del cuerpo K y la independencia lineal de las columnas como la independencia respecto a la multiplicación de columnas a la derecha por los elementos del cuerpo. Esto está relacionado con la regla inicial de multiplicación de las matrices: las filas de la primera se multiplican por las columnas de la segunda. Por supuesto, la teoría no cambiará si las matrices se multiplican por la regla «izquierda»: las columnas de la primera se multiplican por las filas de la segunda. Pero en este caso, habrá que considerar la dependencia lineal de las filas respecto a la multiplicación a la derecha por los elementos de K y la dependencia lineal de las columnas respecto a la multiplicación a la izquierda por los elementos de K. Como resultado, obtendremos los rangos de una matriz según las filas a la derecha y según las columnas a la izquierda. Ambos rangos coincidirán, pero en el caso general serán diferentes de los rangos corrientes según las filas (a la izquierda) y según las columnas (a la derecha). Como ejemplo puede servir la matriz de cuaternios

$$\begin{bmatrix} i & j \\ k & -1 \end{bmatrix}$$

del ejercicio 5, en la que el rango a la izquierda según las filas es igual a 2, mientras que el rango a la derecha según las filas es igual a 1.

7. Sea A una matriz cuadrada invertible de orden n y sean B y C matrices arbitrarias compuestas, respectivamente, por n filas y n columnas. En estas con-

diciones, se tiene

$$\text{rango}\,(AB) = \text{rango}\,B,$$
$$\text{rango}\,(CA) = \text{rango}\,C.$$

8. Dése un ejemplo de matrices cuadradas A y B de un mismo orden 2, para las cuales se tenga rango $(AB) \neq$ rango (BA).

9. Demuéstrese que para cualesquiera matrices cuadradas A y B de orden n se tiene

$$\text{rango}\,(AB) \geqslant \text{rango}\,A + \text{rango}\,B - n.$$

§ 6. Subespacios lineales

6.1. Intersección y suma de subespacios. Un conjunto no vacío $\mathfrak{A}$ de vectores de un espacio lineal $\mathfrak{L}$ se llama *subespacio lineal* de este espacio, si se cumplen las dos condiciones siguientes:

1° *si* $\mathfrak{A}$ *contiene un vector* a, *también contiene todos los múltiplos* λa, *donde* λ *es un número del cuerpo de coeficientes*;

2° *si* $\mathfrak{A}$ *contiene unos vectores* a *y* b, *también contiene su suma* $a + b$.

Está claro que ambas condiciones equivalen a la siguiente: *si* $\mathfrak{A}$ *contiene unos vectores* a *y* b, *también contiene cualquier combinación lineal* $\lambda a + \mu b$ *de los mismos.*

De estas definiciones se desprende que todo subespacio lineal $\mathfrak{A}$ contiene el vector nulo y todas las combinaciones lineales de cualesquiera vectores suyos.

En el p. 4.3 se ha señalado que todo espacio lineal sobre un cuerpo K es un álgebra en la que las operaciones principales son: la adición de vectores, la inversión de vectores y la multiplicación de los vectores (a la izquierda) por elementos de K. Puesto que la inversión de un vector equivale a su multiplicación por el elemento -1 de K, las condiciones 1° y 2° significan simplemente que *los subespacios lineales del espacio* $\mathfrak{L}$ *son subálgebras del álgebra* $\mathfrak{L}$.

El conjunto compuesto solamente del vector nulo posee las propiedades 1° y 2° y, por consiguiente, es un subespacio lineal de $\mathfrak{L}$. Este subespacio se llama subespacio *nulo*. Por otro lado, el propio espacio $\mathfrak{L}$ puede ser considerado como un subespacio lineal de sí mismo. El subespacio nulo y $\mathfrak{L}$ suelen llamarse subespacios *triviales* del espacio $\mathfrak{L}$. Todos los demás subespacios se denominan *no triviales* o *propios*.

El método más sencillo de obtener subespacios lineales consiste en lo siguiente. En el espacio lineal dado $\mathfrak{L}$ se toman unos vectores arbitrarios $a_1, a_2, \ldots, a_m$ y se consideran todas las combinaciones lineales de los mismos

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m.$$

Sea $\mathfrak{A}$ el conjunto de estas combinaciones. Puesto que la suma de combinaciones lineales de $a_1, \ldots, a_m$ y el producto de una combinación lineal por un número son combinaciones lineales de $a_1, \ldots, a_m$, tenemos que $\mathfrak{A}$ es un subespacio lineal. Los vectores $a_1, \ldots, a_m$ son los generadores de este subespacio. Suprimiendo aquellos que dependen linealmente de los anteriores, obtendremos un sistema linealmente independiente de generadores del subespacio $\mathfrak{A}$, es decir, una base de $\mathfrak{A}$. Pero el número de vectores de una base coincide con la dimensión del espacio y, por ello, *la dimensión del subespacio $\mathfrak{A}$ es igual al número máximo de vectores linealmente independientes que contiene el sistema $a_1, \ldots, a_m$*. A veces se dice que el subespacio $\mathfrak{A}$ es el subespacio *tendido* sobre los vectores $a_1, \ldots, a_m$. Por consiguiente, la dimensión del subespacio tendido sobre el sistema de vectores $a_1, \ldots, a_m$ es igual al número máximo de vectores linealmente independientes que contiene este sistema.

El método de tender los subespacios sobre un sistema de vectores dado es general: *todo subespacio lineal $\mathfrak{A}$ de un espacio lineal $\mathfrak{L}$ es el subespacio tendido sobre su base.*

Como que el número de vectores linealmente independientes de $\mathfrak{L}$ no puede superar la dimensión de $\mathfrak{L}$, de aquí se ve que la dimensión de un subespacio lineal no puede superar la dimensión del espacio que lo envuelve. Es más, si la dimensión del espacio lineal $\mathfrak{L}$ es igual a la dimensión del subespacio $\mathfrak{A}$, toda base de $\mathfrak{A}$ será también una base de $\mathfrak{L}$. Por esto, todo vector de $\mathfrak{L}$ se expresará linealmente en términos de una base del subespacio $\mathfrak{A}$, es decir, $\mathfrak{A}$ coincidirá con $\mathfrak{L}$. Por consiguiente, la dimensión de todo subespacio lineal propio es inferior a la dimensión del espacio que lo envuelve.

A título de ejemplo consideremos el espacio corriente $\mathfrak{R}$, formado por los segmentos orientados que parten de un punto O. La dimensión del espacio $\mathfrak{R}$ es igual a tres y, por esto, los subespacios propios pueden ser de dimensión uno o dos. Los subespacios de dimensión uno deben ser tendidos sobre un vector no nulo a, es decir, deben ser el conjunto de los multiplos αa del segmento a. Pero todos los segmentos de tipo αa se hallan sobre la recta que contiene al vector a. Por consiguiente, *los subespacios de $\mathfrak{R}$ de dimensión una son las rectas que pasan por el punto O.*

Los subespacios de dos dimensiones deben ser tendidos sobre dos vectores linealmente independientes, es decir, sobre dos vectores a y b que no pertenecen a una misma recta. Sea $\mathfrak{A}$ el plano que pasa por los vectores a y b. Entonces, todas las combinaciones lineales $\alpha a + \beta b$ pertenecerán al plano $\mathfrak{A}$ y, por otro lado, todo vector perteneciente a $\mathfrak{A}$ será una combinación lineal de los vectores a y b. Por consiguiente, el subespacio tendido sobre los vectores a y b será el conjunto de vectores pertenecientes al plano $\mathfrak{A}$. Luego, *los subespacios de dos dimensiones del espacio $\mathfrak{R}$ son los planos que pasan por el punto O.*

Los espacios de dimensión mayor que tres no admiten una interpretación geométrica tan clara. Sin embargo, a ellos también se aplica la terminología geométrica llamando *rectas* a los subespacios lineales de una dimensión, *planos* a los subespacios de dos dimensiones y *planos de k dimensiones* a los subespacios de dimensión k, para $k \geqslant 3$. Los subespacios lineales de dimensión menor en 1 que la dimensión del espacio llevan el nombre especial de *hiperplanos*.

Con los subespacios de un espacio lineal dado se pueden efectuar determinadas operaciones; las más importantes de éstas son la adición y la intersección. Se llama *intersección* de los subespacios $\mathfrak{A}$, $\mathfrak{B}$, ... de un espacio $\mathfrak{L}$ el conjunto $\mathfrak{P}$ de los vectores que pertenecen simultáneamente a todos estos espacios. La operación de intersección se indica por el símbolo $\cap$, de modo que

$$\mathfrak{P} = \mathfrak{A} \cap \mathfrak{B} \cap \ldots$$

La intersección de cualquier número de subespacios lineales de un espacio $\mathfrak{L}$ es un subespacio lineal de este espacio.

En efecto, el vector nulo pertenece a cada uno de los subespacios dados $\mathfrak{A}$, $\mathfrak{B}$, ... Por esto, pertenece también a la intersección $\mathfrak{P}$ de los mismos que es, por consiguiente, un conjunto no vacío. Por otro lado, si unos vectores a y b pertenecen a la intersección $\mathfrak{P}$, estos vectores, y con ellos cualquier combinación lineal $\alpha a + \beta b$ de los mismos, pertenecerán a cada uno de los subespacios $\mathfrak{A}$, $\mathfrak{B}$, ... Por consiguiente, $\alpha a + \beta b$ está contenido en $\mathfrak{P}$, es decir, $\mathfrak{P}$ es un subespacio lineal.

Se llama *suma* de un número finito de subespacios lineales $\mathfrak{A}_1$, $\mathfrak{A}_2$, ..., $\mathfrak{A}_s$ del espacio $\mathfrak{L}$ el conjunto de vectores que pueden ser representados en la forma

$$a = a_1 + a_2 + \ldots + a_s, \tag{1}$$

donde a_i es un vector de $\mathfrak{A}_i$ ($i = 1, \ldots, s$). La operación de adición de subespacios se indica por el símbolo $+$.

La suma de un número finito de subespacios lineales de un espacio $\mathfrak{L}$ es también un subespacio lineal; contiene todos los vectores de los subespacios dados, así como todas sus combinaciones lineales.

En efecto, sean $\mathfrak{A}_1, \ldots, \mathfrak{A}_s$ los subespacios lineales dados y sea $\mathfrak{A} = \mathfrak{A}_1 + \ldots + \mathfrak{A}_s$. Si a y b son unos vectores de $\mathfrak{A}$, ello significa que pueden ser representados en la forma

$$a = a_1 + a_2 + \ldots + a_s \quad \text{y} \quad b = b_1 + b_2 + \ldots + b_s,$$

donde a_i y b_i pertenecen a $\mathfrak{A}_i$ ($i = 1, \ldots, s$). Pero, en este caso, para cualesquiera α y β de K la expresión

$$\alpha a + \beta b = (\alpha a_1 + \beta b_1) + (\alpha a_2 + \beta b_2) + \ldots + (\alpha a_s + \beta b_s)$$

es la descomposición en la forma (1) del vector $\alpha a + \beta b$, ya que la suma $\alpha a_i + \beta b_i$ pertenece a $\mathfrak{A}_i$. Por consiguiente, $\alpha a + \beta b$ pertenece a $\mathfrak{A}$ y $\mathfrak{A}$ es un subespacio lineal. Tomando ahora en (1) $a_j = o$ para $j \neq i$, vemos que a_i estará contenido en $\mathfrak{A}$, es decir, $\mathfrak{A}$ contiene todos los vectores de $\mathfrak{A}_i$.

Indiquemos, finalmente, sin demostración las propiedades siguientes de la suma de subespacios, que se desprenden directamente de su definición:

1° $\mathfrak{A} + \mathfrak{B} = \mathfrak{B} + \mathfrak{A}$;

2° $\mathfrak{A} + (\mathfrak{B} + \mathfrak{C}) = (\mathfrak{A} + \mathfrak{B}) + \mathfrak{C}$;

3° *si* $\mathfrak{A}$ *está contenido en un subespacio* $\mathfrak{B}$, *se tiene* $\mathfrak{A} + \mathfrak{B} = \mathfrak{B}$.

Considerando la suma de dos subespacios lineales arbitrarios $\mathfrak{A}$ y $\mathfrak{B}$, podremos ver fácilmente que su dimensión depende no sólo de la dimensión de los subespacios $\mathfrak{A}$ y $\mathfrak{B}$, sino también de cuán grande es la parte común de los mismos. El valor exacto de la dimensión de la suma se determina por el teorema siguiente:

TEOREMA. *La dimensión de la suma de dos subespacios lineales de un espacio* $\mathfrak{L}$ *es igual a la suma de las dimensiones de estos subespacios menos la dimensión de su intersección.*

Sean $\mathfrak{A}$ y $\mathfrak{B}$ los subespacios dados y sean r_1 y r_2 sus dimensiones respectivas. Sea m la dimensión de la intersección $\mathfrak{C}$ de estos subespacios. Tomemos en $\mathfrak{C}$ una base cualquiera $c_1, c_2, \ldots, c_m$. Los vectores $c_1, c_2, \ldots, c_m$ son linealmente independientes y pertenecen a $\mathfrak{A}$. Por esto, en $\mathfrak{A}$ se pueden encontrar unos vectores a_1, $a_2, \ldots, a_k$, tales, que el sistema $a_1, a_2, \ldots, a_k, c_1, \ldots, c_m$ sea una base de $\mathfrak{A}$ (véase el p. 4.2). Por esta misma razón, en el subespacio $\mathfrak{B}$ existen vectores $b_1, b_2, \ldots, b_p$ tales, que junto a los vectores $c_1, c_2, \ldots, c_m$ constituyen una base de $\mathfrak{B}$. Puesto que el número de vectores de una base coincide con la dimensión del espacio, entre los números k y p y las dimensiones de los subespacios $\mathfrak{A}$ y $\mathfrak{B}$ existen las relaciones

$$r_1 = k + m \quad \text{y} \quad r_2 = p + m.$$

Si demostramos que el sistema

$$a_1, \ldots, a_k, \quad c_1, \ldots, c_m, \quad b_1, \ldots, b_p \tag{2}$$

es una base del espacio $\mathfrak{A} + \mathfrak{B}$, el teorema 1 quedará demostrado, ya que la dimensión del espacio $\mathfrak{A} + \mathfrak{B}$ será igual a

$$k + m + p = r_1 + r_2 - m.$$

Todo vector a de $\mathfrak{A}$ se expresa linealmente en términos del sistema $a_1, \ldots, a_k, c_1, \ldots, c_m$, que constituye una base de $\mathfrak{A}$ y, por esto, se expresa también en términos del sistema (2). Análogamente, todo vector b de $\mathfrak{B}$ también se expresa linealmente en términos de (2). Pero en este caso la suma $a + b$, es decir, cualquier vector de $\mathfrak{A} + \mathfrak{B}$, se expresará linealmente en términos de (2). Resta demostrar que el sistema (2) es linealmente independiente.

Sea

$$\alpha_1 a_1 + \ldots + \alpha_k a_k + \gamma_1 c_1 + \ldots + \gamma_m c_m + \beta_1 b_1 + \ldots + \beta_p b_p = o \quad (3)$$

una relación de dependencia lineal entre estos vectores. Pongamos

$$b = \beta_1 b_1 + \ldots + \beta_p b_p.$$

El vector b se expresa linealmente en términos de los vectores $b_1, \ldots, b_p$ contenidos en el espacio $\mathfrak{B}$; luego, b pertenece a $\mathfrak{B}$. Por otro lado de (3) se desprende que

$$b = -\alpha_1 a_1 - \ldots - \alpha_k a_k - \gamma_1 c_1 - \ldots - \gamma_m c_m. \quad (4)$$

Puesto que $a_1, \ldots, a_k, c_1, \ldots, c_m$ están contenidos en $\mathfrak{A}$, de aquí se deduce que b también pertenece a $\mathfrak{A}$. Por consiguiente, b figura en la intersección de los espacios $\mathfrak{A}$ y $\mathfrak{B}$, es decir, la expresión (4) del vector b en términos de la base de $\mathfrak{A}$ no contiene términos con $a_1, \ldots, a_k$, en otras palabras,

$$\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0.$$

Tomando estos valores en (3), obtenemos

$$\gamma_1 c_1 + \ldots + \gamma_m c_m + \beta_1 b_1 + \ldots + \beta_p b_p = o.$$

Pero el sistema $c_1, \ldots, c_m, b_1, \ldots, b_p$ es una base de $\mathfrak{B}$ y, por consiguiente, es linealmente independiente, de modo que

$$\gamma_1 = \ldots = \gamma_m = \beta_1 = \ldots = \beta_p = 0.$$

Hemos demostrado que el sistema (2) es linealmente independiente.

Del teorema demostrado se puede deducir una desigualdad que ofrece el valor mínimo de la dimensión de la intersección de unos subespacios. Consideremos unos subespacios lineales $\mathfrak{A}$ y $\mathfrak{B}$ de $\mathfrak{L}$ y sean r_1 y r_2 las dimensiones de estos subespacios, n la dimensión de $\mathfrak{L}$ y m la dimensión de la intersección $\mathfrak{A} \cap \mathfrak{B}$. En virtud del teorema, la dimensión de la suma $\mathfrak{A} + \mathfrak{B}$ es igual a $r_1 + r_2 - m$. Pero la dimensión de la suma $\mathfrak{A} + \mathfrak{B}$ es no mayor que la dimensión del espacio $\mathfrak{L}$. Por consiguiente, $r_1 + r_2 - m \leqslant n$, de donde se tiene $m \geqslant r_1 + r_2 - n$. Es decir, *la dimensión de la intersección de dos subespacios lineales del espacio $\mathfrak{L}$ no puede ser menor que el exceso de la suma de las dimensiones de estos subespacios respecto a la dimensión del espacio $\mathfrak{L}$.*

Por ejemplo, la intersección de dos planos del espacio de tres dimensiones contiene siempre una recta, la intersección de un subespacio de dos dimensiones con un subespacio de tres dimensiones en un espacio de cuatro dimensiones contiene una recta, la intersección de dos subespacios de tres dimensiones de un espacio de cuatro dimensiones contiene un plano, etc.

6.2. Sumas directas. Según la definición, todo vector a que pertenece a la suma de subespacios lineales

$$\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2 + \ldots + \mathfrak{A}_s,$$

puede ser representado en la forma

$$a = a_1 + a_2 + \ldots + a_s \quad (a_i \in \mathfrak{A}_i,\ i = 1,\ \ldots,\ s). \tag{1}$$

Sin embargo, esta representación no será, en general, unívoca. Si todo vector a de $\mathfrak{A}$ admite sólo una representación de tipo (1), es decir, si de las igualdades (1) y de

$$a = a_1' + a_2' + \ldots + a_s' \quad (a_i' \in \mathfrak{A}_i,\ i = 1,\ \ldots,\ s) \tag{2}$$

se deduce que $a_1 = a_1',\ \ldots,\ a_s = a_s'$, la suma se llama *suma directa* y se indica

$$\mathfrak{A} = \mathfrak{A}_1 \dotplus \mathfrak{A}_2 \dotplus \ldots \dotplus \mathfrak{A}_s.$$

Las sumas directas poseen numerosas propiedades específicas, algunas de las cuales serán ahora consideradas. Ante todo observemos que en lugar de exigir que cualquier vector de la suma posee un desarrollo unívoco se puede exigir que el desarrollo unívoco exista sólo para el vector nulo: *si el vector nulo de la suma* $\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2 + \ldots + \mathfrak{A}_s$ *admite un desarrollo único de tipo* (1), *es decir, si*

$$a_1 + a_2 + \ldots + a_s = o \quad (a_i \in \mathfrak{A}_i,\ i = 1,\ \ldots,\ s)$$

implica que $a_1 = a_2 = \ldots = a_s = o$, *la suma es directa.*

Para la demostración tomemos un vector cualquiera a de la suma y supongamos que admite dos desarrollos: (1) y (2). Restando un desarrollo del otro, obtenemos

$$(a_1 - a_1') + (a_2 - a_2') + \ldots + (a_s - a_s') = o.$$

Puesto que aquí $(a_i - a_i') \in \mathfrak{A}_i$, $i = 1,\ 2,\ \ldots,\ s$, tenemos según la hipótesis

$$a_1 - a_1' = a_2 - a_2' = \ldots = a_s - a_s' = o,$$

es decir,

$$a_1 = a_1',\ a_2 = a_2',\ \ldots,\ a_s = a_s'.$$

Hemos demostrado la proposición.

El teorema que sigue está relacionado con el empleo de los paréntesis en las sumas directas.

TEOREMA 1. *Si las descomposiciones de los subespacios lineales*

$$\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2 + \ldots + \mathfrak{A}_s, \tag{3}$$

$$\left.\begin{aligned} \mathfrak{A}_1 &= \mathfrak{A}_{11} + \mathfrak{A}_{12} + \ldots + \mathfrak{A}_{1m_1}, \\ &\ldots\ldots\ldots\ldots\ldots \\ \mathfrak{A}_s &= \mathfrak{A}_{s1} + \mathfrak{A}_{s2} + \ldots + \mathfrak{A}_{sm_s}, \end{aligned}\right\} \tag{4}$$

son directas, la descomposición

$$\mathfrak{A}=\mathfrak{A}_{11}+\mathfrak{A}_{12}+\ldots+\mathfrak{A}_{1m_1}+\ldots+\mathfrak{A}_{s1}+\mathfrak{A}_{s2}+\ldots+\mathfrak{A}_{sm_s} \quad (5)$$

también es directa; en otras palabras, si en una suma directa todo sumando es sustituido por su descomposición directa, se obtendrá de nuevo una descomposición directa. Recíprocamente, si la descomposición (5) es directa, las descomposiciones (4) *y* (3) *también son directas.*

Demostremos primero la proposición directa. Sea

$$a_{11}+a_{12}+\ldots+a_{1m_1}+\ldots+a_{s1}+a_{s2}+\ldots+a_{sm_s}=o.$$

Escribamos esta igualdad en la forma

$$a_1+a_2+\ldots+a_s=o, \quad (6)$$

donde

$$a_i=a_{i1}+a_{i2}+\ldots+a_{im_i} \quad (i=1,\ 2,\ \ldots,\ s). \quad (7)$$

Como que $a_i\in\mathfrak{A}_i$ y la suma (3) es directa, de (6) se deduce que $a_1=a_2=\ldots=a_s=o$ y la igualdad (7) se transforma en

$$a_{i1}+a_{i2}+\ldots+a_{im}=o \quad (i=1,\ 2,\ \ldots,\ s). \quad (8)$$

Pero la descomposición

$$\mathfrak{A}_i=\mathfrak{A}_{i1}+\mathfrak{A}_{i2}+\ldots+\mathfrak{A}_{im_i}$$

es, por hipótesis, directa y por ello de (8) se deduce que $a_{i1}=a_{i2}=\ldots=a_{im_i}=o$. Por consiguiente, la suma (5) es directa que es lo que se quería demostrar. Repitiendo los razonamientos en el orden contrario, obtendremos la demostración de la afirmación recíproca.

El teorema 1 permite considerar la suma directa de varios subespacios como el resultado de sucesivas sumas directas de dos sumandos. Las condiciones que garantizan que la suma de dos subespacios sea directa pueden ser enunciadas de la siguiente forma conveniente.

TEOREMA 2. *Para que la suma de dos subespacios lineales de un espacio* $\mathfrak{L}$ *sea directa es necesario y suficiente que la intersección de estos subespacios sea nula.*

En efecto, si la intersección de dos subespacios $\mathfrak{A}$ y $\mathfrak{B}$ contiene un vector no nulo a, para el vector o se puede escribir la descomposición

$$a+(-a)=o,$$

donde $a\neq o$, $a\in\mathfrak{A}$ y $-a\in\mathfrak{B}$, y, por consiguiente, la suma $\mathfrak{A}+\mathfrak{B}$ no será directa. Viceversa, si la suma $\mathfrak{A}+\mathfrak{B}$ no es directa, para el vector nulo deberá existir la descomposición

$$a+b=o \quad (a\neq o,\ a\in\mathfrak{A} \text{ y } b\in\mathfrak{B}).$$

Puesto que $b = -a$, el vector b figura también en el subespacio $\mathfrak{A}$. Por consiguiente, la intersección $\mathfrak{A} \cap \mathfrak{B}$ contiene, en este caso, un vector no nulo b, que es lo que se quería demostrar.

TEOREMA 3. *La dimensión de una suma directa de subespacios es igual a la suma de las dimensiones de estos subespacios.*

Si tenemos dos sumandos, entonces la dimensión de la suma es igual, de acuerdo con el p. 6.1, a la suma de las dimensiones de los sumandos menos la dimensión de la intersección. Pero, según el teorema 2, la intersección de subespacios en el caso de una suma directa es nula y su dimensión es igual a cero. Por esto, *la dimensión de una suma directa de dos subespacios es igual a la suma de sus dimensiones.* Si se tiene más de dos sumandos, la demostración se realiza fácilmente por inducción.

Del teorema 3 se desprende el siguiente corolario: *si el subespacio $\mathfrak{A}$ es la suma directa de los subespacios $\mathfrak{A}_1, \ldots, \mathfrak{A}_s$, entonces, escogiendo en cada subespacio $\mathfrak{A}_i$ una base $a_{i1}, a_{i2}, \ldots, a_{im_i}$ $(i = 1, 2, \ldots, s)$ y uniendo estas bases en un sistema*

$$a_{11}, a_{12}, \ldots, a_{1m_1}, \ldots, a_{s1}, a_{s2}, \ldots, a_{sm_s}, \tag{9}$$

obtendremos una base del subespacio $\mathfrak{A}$.

Efectivamente, todo vector del subespacio $\mathfrak{A}_i$ se expresa linealmente en términos de los vectores (9) y por esto cualquier vector de $\mathfrak{A}$ también se expresará linealmente en términos de los vectores (9). En virtud del teorema 3, el número de vectores del sistema (9) es igual a la dimensión del subespacio $\mathfrak{A}$. Luego, el sistema (9) es una base de $\mathfrak{A}$, que es lo que se quería demostrar.

El teorema 3 admite la siguiente inversión: *si la dimensión de una suma de subespacios lineales es igual a la suma de las dimensiones de los mismos, la suma es directa.*

Sea primero $\mathfrak{A} = \mathfrak{A}_1 + \mathfrak{A}_2$ y sea

$$\text{dim. } \mathfrak{A} = \text{dim. } \mathfrak{A}_1 + \text{dim. } \mathfrak{A}_2.$$

En virtud del teorema del p. 6.1, de aquí se deduce que la dimensión de la intersección $\mathfrak{A}_1 \cap \mathfrak{A}_2$ es igual a cero, es decir, que esta intersección es nula. A su vez, de aquí se deduce, de acuerdo con el teorema 2, que la suma $\mathfrak{A}_1 + \mathfrak{A}_2$ es directa. Si el número de sumandos es mayor que dos, es suficiente aplicar para la demostración la inducción.

6.3. Sistemas de ecuaciones lineales homogéneas. Con el fin de ilustrar la teoría de subespacios lineales consideremos el problema de resolución de los sistemas de ecuaciones lineales homogéneas, es decir, de los sistemas de tipo

$$\left.\begin{array}{l} \xi_1\alpha_{11} + \xi_2\alpha_{21} + \ldots + \xi_n\alpha_{n1} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \xi_1\alpha_{1m} + \xi_2\alpha_{2m} + \ldots + \xi_n\alpha_{nm} = 0, \end{array}\right\} \tag{1}$$

donde α_{ij} son elementos de un cuerpo K, mientras que $\xi_1, \ldots, \xi_n$ son las incógnitas; los valores de las últimas se buscan en el cuerpo K. El sistema (1) puede ser representado en la siguiente forma matricial

$$x \cdot A = O, \tag{2}$$

donde $x = [\xi_1 \ldots, \xi_n]$ es la fila de incógnitas y $A = \|\alpha_{ij}\|$ es la matriz en la que la j-ésima columna está formada por los coeficientes de la j-ésima ecuación del sistema (1). La fila nula $[0, \ldots, 0]$ ofrece una solución del sistema (1) cualesquiera que sean los coeficientes α_{ij}. Esta fila se llama solución *nula* o *trivial* del sistema (1). La existencia de la solución trivial indica que un sistema de ecuaciones lineales homogéneas nunca es contradictorio.

Para poder estudiar con más detalle las propiedades de las soluciones del sistema (1) indiquemos por $\mathfrak{L}$ el espacio de filas de longitud n sobre el cuerpo K y consideremos toda solución $[\xi_1^0, \ldots, \xi_n^0]$ del sistema (1) como un vector x del espacio $\mathfrak{L}$ que satisface la ecuación (2). *El conjunto de todas las soluciones de la ecuación* (2) *es un subespacio lineal del espacio* $\mathfrak{L}$. Efectivamente, de (2) y de $y \cdot A = O$ se deduce que

$$(\lambda x + \mu y) A = \lambda (xA) + \mu (yA) = O$$

cualesquiera que sean $\lambda, \mu \in K$. En otras palabras, una combinación lineal arbitraria de soluciones del sistema (1) es también una solución del sistema (1). Puesto que el espacio $\mathfrak{L}$ es de dimensión n, el espacio de soluciones del sistema (1) es de dimensión no mayor que n.

Se dice que unas soluciones $x_1, \ldots, x_s$ del sistema (2) forman *un sistema fundamental de soluciones* de (2), si cualquier solución del sistema (2) puede ser representada en forma de una combinación lineal $\lambda_1 x_1 + \ldots + \lambda_s x_s$ de las soluciones indicadas y, al mismo tiempo, ninguna de las soluciones $x_1, \ldots, x_s$ puede ser representada como una combinación lineal de las restantes. En términos de la teoría de espacios lineales esto significa que el conjunto $x_1, \ldots, x_s$ es simplemente *una base del subespacio de soluciones* y que el número de soluciones fundamentales es la dimensión del subespacio de soluciones.

El rango de la matriz A se llama *rango del sistema homogéneo* (1). Los sistemas homogéneos de tipo (1) representan un caso particular de los sistemas lineales generales estudiados ya en el p. 5.3. Empleando el teorema de Kronecker—Capelli demostrado en aquella ocasión se obtiene fácilmente el siguiente teorema principal:

TEOREMA SOBRE LAS ECUACIONES LINEALES HOMOGÉNEAS. *La dimensión del espacio de soluciones del sistema* (1) *de ecuaciones lineales homogéneas con n incógnitas es igual a la diferencia $n-r$, donde r es el rango del sistema* (1).

Si $r=n$, el sistema (1) debe tener, según el teorema de Kronecker—Capelli, una solución única que es, por consiguiente, la solución nula. El subespacio de soluciones consta solamente del vector nulo y su dimensión es igual a 0, es decir, coincide con la diferencia $n-r$.

Sea $r<n$. En virtud del teorema de Kronecker—Capelli, el sistema (1) será en este caso (después de realizar de nuevo una enumeración adecuada de las incógnitas y de las ecuaciones) equivalente a un sistema de tipo

$$\left.\begin{array}{l} \xi_1=\xi_{r+1}\gamma_{r+1,1}+\ldots+\xi_n\gamma_{n1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \xi_r=\xi_{r+1}\gamma_{r+1,r}+\ldots+\xi_n\gamma_{nr}, \end{array}\right\} \qquad (3)$$

donde γ_{st} son determinados elementos de K. A las incógnitas $\xi_{r+1}, \ldots, \xi_n$ se les puede asignar aquí cualesquiera valores de K. Tomando sucesivamente

$$\begin{array}{l} [\xi_{r+1}, \xi_{r+2}, \ldots, \xi_n]=[1, 0, \ldots, 0], \\ [\xi_{r+1}, \xi_{r+2}, \ldots, \xi_n]=[0, 1, \ldots, 0], \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ [\xi_{r+1}, \xi_{r+2}, \ldots, \xi_n]=[0, 0, \ldots, 1] \end{array}$$

y calculando toda vez los valores de las incógnitas $\xi_1, \ldots, \xi_n$ según las fórmulas (3), encontramos el siguiente sistema de soluciones de (1):

$$\begin{array}{l} c_1=[\gamma_{r+1,1}, \ldots, \gamma_{r+1,r}, 1, 0, \ldots, 0], \\ c_2=[\gamma_{r+2,1}, \ldots, \gamma_{r+2,r}, 0, 1, \ldots, 0], \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_{n-r}=[\gamma_{n1}, \ldots, \gamma_{nr}, \quad 0, 0, \ldots, 1]. \end{array} \qquad (4)$$

Sabemos que el número de filas linealmente independientes en el sistema $c_1, \ldots, c_{n-r}$ coincide con el rango de la matriz formada por estas filas. Pero de (4) se ve que la matriz indicada contiene un menor, formado por las últimas $n-r$ columnas, que es la matriz unidad. Por esto, el rango de la matriz es igual a $n-r$ y, por consiguiente, todos los vectores $c_1, \ldots, c_{n-r}$ son linealmente independientes. Por otro lado, las fórmulas (3) muestran que toda solución $x=[\xi_1^0, \ldots, \xi_n^0]$ satisface la relación

$$x=\xi_{r+1}^0c_1+\ldots+\xi_n^0c_{n-r},$$

es decir, es una combinación lineal de las soluciones $c_1, \ldots, c_{n-r}$. Hemos verificado de esta forma que el sistema $c_1, \ldots, c_{n-r}$ es una base del espacio de soluciones del sistema (1) y, por esto, la dimensión de dicho espacio es igual a $n-r$.

Indiquemos dos corolarios directos del teorema demostrado.

COROLARIO 1. *Para que un sistema de n ecuaciones lineales homogéneas de n incógnitas*

$$\left.\begin{array}{l}\xi_1\alpha_{11}+\ldots+\xi_n\alpha_{n1}=0,\\ \ldots\ldots\ldots\ldots\\ \xi_1\alpha_{n1}+\ldots+\xi_n\alpha_{nn}=0\end{array}\right\}$$

con coeficientes de un cuerpo cualquiera K no tenga soluciones no nulas, es necesario y suficiente que la matriz de esta sistema

$$A=\left\|\begin{array}{ccc}\alpha_{11} & \ldots & \alpha_{n1}\\ \ldots & \ldots & \ldots\\ \alpha_{1n} & \ldots & \alpha_{nn}\end{array}\right\|$$

sea invertible.

Efectivamente, la condición $n=r$ significa que el rango de la matriz A debe coincidir con su orden, es decir, la matriz A debe ser invertible.

COROLARIO 2. *Para que un sistema de n ecuaciones lineales homogéneas de n incógnitas con coeficientes de un cuerpo conmutativo tenga solución no nula es necesario y suficiente que el determinante de la matriz de este sistema sea igual a cero.*

En efecto, una matriz cuadrada formada por elementos de un cuerpo conmutativo no es invertible si, y sólo si, su determinante es igual a cero.

Hemos visto que el teorema sobre las ecuaciones lineales homogéneas es un corolario directo del teorema de Kronecker—Capelli. El último teorema, además de ser cierto para los sistemas lineales con coeficientes a la derecha, tiene lugar también para los sistemas lineales con coeficientes a la izquierda. Por esto, junto al teorema sobre las ecuaciones lineales homogéneas y al corolario 1, son válidas las proposiciones análogas relacionadas con sistemas de ecuaciones lineales homogéneas con coeficientes a la izquierda, a saber: el conjunto de soluciones columnas $x=[\xi_1^0, \ldots, \xi_n^0]$ de un sistema de ecuaciones lineales homogéneas

$$\left.\begin{array}{l}\alpha_{11}\xi_1+\ldots+\alpha_{1n}\xi_n=0,\\ \ldots\ldots\ldots\ldots\\ \alpha_{m1}\xi_1+\ldots+\alpha_{mn}\xi_n=0\end{array}\right\} \tag{5}$$

con coeficientes de un cuerpo K es un subespacio lineal del espacio formado por todas las columnas de longitud n sobre K. La dimensión de este espacio de soluciones es igual a $n-r$, donde r es el rango de la matriz $A=\|\alpha_{ij}\|$, en la que la i-ésima fila está formada por los coeficientes de la i-ésima ecuación del sistema (5).

Ejemplos y problemas

1. Hállese la dimensión del subespacio lineal tendido sobre los vectores $a=[1, 3, 2, 1]$, $b=[4, 9, 5, 4]$ y $c=[3, 7, 4, 3]$.

2. Si $e_1, e_2, \ldots, e_n$ es una base de un espacio lineal $\mathfrak{L}$ y $\mathfrak{A}_i$ es el subespacio lineal tendido sobre e_i, se tiene

$$\mathfrak{L}=\mathfrak{A}_1\dot{+}\mathfrak{A}_2\dot{+}\ldots\dot{+}\mathfrak{A}_n.$$

3. Para todo subespacio propio $\mathfrak{A}$ de un espacio $\mathfrak{L}$ existe un subespacio lineal $\mathfrak{B}$, tal, que $\mathfrak{L}=\mathfrak{A}\dot{+}\mathfrak{B}$.

4. Demuéstrese la siguiente generalización del teorema 2 del p. 6.2: para que la suma de varios subespacios lineales dados sea directa es necesario y suficiente que cada uno de los subespacios dados tenga intersección nula con la suma de los restantes.

5. Demuéstrese que la intersección de todos los subespacios lineales de un espacio $\mathfrak{L}$, que contienen a los subespacios lineales dados, es igual a la suma de los últimos.

6. Demuéstrese que para cualesquiera subespacios $\mathfrak{A}$, $\mathfrak{B}$ y $\mathfrak{C}$ de un espacio lineal tienen lugar las igualdades

$$\mathfrak{A}(\mathfrak{A}\mathfrak{B}+\mathfrak{C})=\mathfrak{A}\mathfrak{B}+\mathfrak{A}\mathfrak{C},$$
$$(\mathfrak{A}+\mathfrak{B})(\mathfrak{A}+\mathfrak{C})=\mathfrak{A}+(\mathfrak{A}+\mathfrak{B})\mathfrak{C},$$

donde, para abreviar, la intersección de subespacios se indica igual que el producto.

7. Todo espacio lineal $\mathfrak{L}$ de dimensión infinita contiene un subespacio lineal propio cuya dimensión coincide con la dimensión de todo el espacio $\mathfrak{L}$.

8. ¿Para qué valores de λ (del cuerpo de los números complejos) el sistema de ecuaciones

$$\left.\begin{aligned} x+2\lambda y-\lambda z&=y,\\ 2x+y-z&=\lambda x,\\ x+\lambda y-z&=0 \end{aligned}\right\}$$

tiene solución no nula? ¿Para qué valores de λ el espacio de soluciones de este sistema es de mayor dimensión?

9. ¿Para qué valores del parámetro λ el sistema de ecuaciones

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

es compatible?

10. ¿Qué dimensión tienen los espacios de soluciones de los sistemas

$$\left.\begin{aligned} x-iy+iz&=0\\ jx+ky-z&=0\\ x-iy+(2i-j)z&=0 \end{aligned}\right\} \quad \text{y} \quad \left.\begin{aligned} x-yi+zi&=0\\ xj+yk-z&=0\\ x-yi+z(2i-j)&=0 \end{aligned}\right\}$$

considerados sobre el cuerpo de los cuaternios (p.1.5)?

11. ¿Existe un polinomio en variables cuaternias a, b, c y d y con coeficientes cuaternios, tal, que su anulación sea la condición necesaria y suficiente para que el sistema de ecuaciones lineales

$$ax+by=0 \quad \text{y} \quad cx+dy=0$$

tenga solución no nula?

12. ¿Existe un polinomio en $a, \bar{a}, b, \bar{b}, c, \bar{c}, d$ y $\bar{d}$ que satisfaga las exigencias del ejercicio anterior?