

(ÁLGEBRA)

TOMO II

M.Sc. Teresita Noriega Sánchez
Lic. Luis Ramiro Piñeiro Díaz

Edición: Virginia Rojas Lucero
Diseño: María Elena Gil Mac Beath y Sonia Acosta Milián
Diseño de cubierta: Olmer Buchholz
Ilustraciones: Judiht Siklós

ÁLGEBRA

TOMO II

Primera edición: Editorial Pueblo y Educación, 1985
Segunda edición: Editorial Félix Varela, 2007

© Teresita Noriega Sánchez y Luis Ramiro Piñero Díaz, 1985
© Sobre la presente edición:
Editorial Félix Varela, 2007

ISBN 978-959-07-0654-7 Obra completa
ISBN 978-959-07-0642-4 Tomo II

Editorial Félix Varela
San Miguel No. 1111
e/ Mazón y Basarrate, Vedado,

Índice

- 1 Formas cuadráticas 7
 - 1.1 Formas cuadráticas. Representación matricial 7
 - 1.2 Forma canónica de una forma cuadrática 12
 - 1.3 Formas cuadráticas reales de signo constante y formas definidas positivas 22Ejercicios
- 2 Espacios euclidianos y unitarios 30
 - 2.1 Introducción 30
 - 2.2 Espacios vectoriales con producto escalar 32
 - 2.3 Norma o longitud de un vector en un espacio con producto escalar 38
 - 2.4 Ortogonalidad en un espacio con producto escalar 44
 - 2.5 Formas lineales sobre un espacio vectorial 64
 - 2.6 Formas lineales sobre un espacio con producto escalar. Endomorfismo adjunto 85
 - 2.7 Endomorfismos de un espacio con producto escalar de dimensión finita 96
- 3 Grupos 127
 - 3.1 Introducción 127
 - 3.2 Leyes internas 129
 - 3.3 Grupos 136
 - 3.4 Propiedades de los grupos 141
 - 3.5 Grupos finitos expresados mediante tablas 145
 - 3.6 Subgrupos 150
 - 3.7 Grupos de transformaciones 156
 - 3.8 Isomorfismos de grupos y teorema de Cayley 161
 - 3.9 Relaciones de equivalencia 170
 - 3.10 Subgrupos invariantes 185

• 3.11 Grupo cociente	194
• 3.12 Homomorfismos de grupos y teorema fundamental de homomorfismo	198
• 3.13 Grupos cíclicos	206
• 3.14 Nociones sobre la teoría de grupos abelianos	222
4 Anillos y cuerpos	236
• 4.1 Introducción	236
• 4.2 Definiciones, ejemplos y primeras propiedades	237
• 4.3 Distintos tipos de anillos	245
• 4.4 Extensión de un dominio de integridad a un cuerpo	253
• 4.5 Ideales	259
• 4.6 Anillo cociente	266
• 4.7 Homomorfismos de anillos	270
• 4.8 Ideales primos y maximales	276
• 4.9 El anillo de los polinomios en una indeterminada sobre un cuerpo k	282
• 4.10 Dominios de factorización única, dominios euclideos y dominios principales	304
• 4.11 Breve introducción a la extensión de cuerpos	323
Bibliografía	347

Prólogo

El objetivo del presente libro es servir de texto para el curso de Álgebra III que se imparte en el primer semestre del segundo año, en la especialidad de Matemática, en los centros universitarios del país.

Este libro es el resultado de la experiencia en la impartición de esta asignatura durante varios años en el Departamento de Matemática de la Universidad de La Habana y comprende los temas: "Formas cuadráticas", "Espacios euclideos y unitarios", "Grupos" y "Anillos y cuerpos".

La selección de los capítulos en el libro está determinada por el programa de la asignatura para la cual el libro ha sido concebido como texto. De igual forma se hace referencia a la materia tratada en el libro *Álgebra Tomo I* que saldrá a la luz en el próximo año.

Tomando en consideración las dificultades que presentan los estudiantes con el razonamiento abstracto, hemos procurado, siempre que ha sido posible, introducir los conceptos mediante ejemplos conocidos y usando el método constructivo en muchas demostraciones.

También hemos incluido breves reseñas del origen histórico de los conceptos, en los capítulos sobre estructuras algebraicas, ya que en lo que se refiere a la parte de Álgebra lineal, esto se hará en el libro *Álgebra Tomo I*.

El importante aspecto de la relación de los temas tratados con las aplicaciones prácticas no ha sido olvidado, destacándolo dentro de las posibilidades dadas por la ubicación de la asignatura en el *currículum*.

En aras de ayudar a la profundización de los conocimientos y dar una idea más amplia de esta disciplina, hemos tratado algunos temas de forma más detallada a la considerada estrictamente en el programa; específicamente los distintos tipos de anillo y la introducción a la extensión de cuerpos en el capítulo 4. Con el mismo fin incluimos numerosas aclaraciones sobre los resultados correspondientes a los espacios vectoriales de dimensión infinita en el capítulo 2.

En cuanto a los ejercicios, en su mayoría han sido utilizados en la impartición de la asignatura, y se incluyen sugerencias en aquellos que según nuestra experiencia resultan más difíciles a los estudiantes.

Agradecemos a los compañeros Héctor de Arazosa y Mayra Solana del Departamento de Matemática de la Universidad de La Habana por la revisión detallada del manuscrito y las oportunas sugerencias que nos han ofrecido y también al compañero José Madruga por su paciente labor en mecanografiarlo.

Por último, no podemos dejar de agradecer a la Revolución, de la cual somos producto y a la cual nos debemos.

Ciudad de La Habana, julio de 1982

Los autores

1 Formas cuadráticas

1.1 Formas cuadráticas. Representación matricial

En Geometría analítica se estudian las curvas de segundo orden. La ecuación de una curva central de segundo orden, considerando el origen de coordenadas en el centro de la curva, tiene la forma

$$Ax^2 + 2Bxy + Cy^2 = D. \quad (1)$$

Al efectuar una rotación de los ejes de coordenadas en un ángulo α , las coordenadas de los puntos se transforman según las fórmulas siguientes:

$$x = x' \cos \alpha - y' \sin \alpha$$

$$y = x' \sin \alpha + y' \cos \alpha,$$

donde x, y son las coordenadas iniciales de un punto cualquiera y x', y' son sus nuevas coordenadas. Si se selecciona adecuadamente el ángulo de rotación α , la ecuación de la curva puede reducirse a la llamada *forma canónica*

$$A'x'^2 + B'y'^2 = D. \quad (2)$$

En la cual el coeficiente de $x'y'$ es igual a 0. Esto es, mediante una transformación de coordenadas adecuada, hemos transformado el primer miembro de la ecuación (1), que consiste en una expresión polinomial homogénea de segundo grado en las variables x, y , en el primer miembro de la ecuación (2), que también es una ecuación polinomial homogénea de grado dos, en las variables x', y' , y en la cual no aparece el término $x'y'$. Estas expresiones polinomiales homogéneas de segundo grado, llamadas *formas cuadráticas*, aparecen en muchos otros problemas y con un número mayor de variables, por ejemplo, en la fórmula de la energía cinética $\frac{m}{2} (u^2 + v^2 + w^2)$ de un cuerpo moviéndose en el espacio con velocidad de componentes u, v y w , y en Geometría diferencial en la fórmula de la longitud de arco en coordenadas esféricas: $ds^2 = dr^2 + r^2 d\phi^2 + r^2 \sin^2 \phi d\theta^2$.

Un ejemplo con un número mayor de variables es el siguiente:

En una población determinada se desea hacer un análisis sobre el comportamiento de las estaturas de sus elementos. Una primera idea acerca de dicho comportamiento lo puede dar el promedio de las estaturas. Ahora bien, hasta qué punto es dicho valor representativo de la estatura de esta población? Si en la población existen muchos miembros de elevada estatura y muchos de baja estatura, este valor no será realmente representativo.

Para el análisis de esta representatividad, el estadístico analiza cómo se desvían las estaturas de cada uno de los elementos de la población respecto de la media calculada, mediante el análisis de la expresión

$$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

donde x_i representa la estatura del elemento i de la población, $\bar{x}$ la estatura promedio y n el número de elementos de la población.

La expresión anterior posee un carácter cuadrático y depende de más de dos variables. Expresiones de este tipo aparecen, además, en Mecánica, Óptica e Investigación de operaciones, entre otras. Es por ello que generalizaremos la teoría de curvas y superficies, estudiada en Geometría analítica, a un número mayor de variables, llegando así al concepto de forma cuadrática.

DEFINICIÓN 1.1.1

Sea E un K -espacio de dimensión finita n . Se dice que una aplicación $q: E \rightarrow K$ es una *forma cuadrática* si para $x = x_1 a_1 + \dots + x_n a_n$, donde $(a_i)_{1 \leq i \leq n}$ es una base de E se cumple

$$q(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j, \text{ donde } a_{ij} = a_{ji} \text{ y } a_{ij} \in K. \quad (3)$$

Observemos que en la expresión anterior aparecen los productos entre sí, por pares, de las diferentes coordenadas de x , lo que le da un aspecto cuadrático. Como $a_{ij} = a_{ji}$ y $x_i x_j = x_j x_i$, podemos escribir la expresión de $q(x)$ en la forma

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j. \quad (4)$$

Ejemplos

1. $q(x, y, z) = 3x^2 - y^2 + z^2 - 4xz + 2yz$ define una forma cuadrática sobre $\mathbb{R}^3$.
2. $q(x, y, z) = 3x^2 - y^2 + z^2$ define una forma cuadrática sobre $\mathbb{R}^3$ donde no aparecen términos rectángulos.
3. $q(x, y, z) = -4xz + 2yz$ define una forma cuadrática sobre $\mathbb{R}^3$ donde no aparecen términos cuadrados.

De la expresión de la forma cuadrática surgen de forma natural las consideraciones siguientes:

$$\text{Sea } A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

donde los a_{ij} son los coeficientes de la expresión $q(x) = \sum_i \sum_j a_{ij} x_i x_j$, y sea

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ el vector de coordenadas de } x \text{ en la base } \{a_1, \dots, a_n\}.$$

Entonces

$$AX = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{pmatrix}$$

$$\begin{aligned} {}^tX(AX) &= (x_1 \dots x_n) \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \dots \\ a_{n1}x_1 + \dots + a_{nn}x_n \end{pmatrix} \\ &= x_1(a_{11}x_1 + \dots + a_{1n}x_n) + \dots + x_n(a_{n1}x_1 + \dots + a_{nn}x_n) \\ &= a_{11}x_1^2 + \dots + a_{1n}x_1x_n + \dots + a_{n1}x_nx_1 + \dots + a_{nn}x_n^2. \end{aligned}$$

Luego

$${}^tXAX = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j.$$

Por tanto, podemos obtener $q(x)$ mediante la expresión matricial

$${}^tXAX = q(x). \quad (5)$$

DEFINICIÓN 1.1.2

Dada la forma cuadrática q sobre el espacio E

$$q(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j,$$

donde $x_1, \dots, x_n$ son las coordenadas del vector x en la base $\{a_1, \dots, a_n\}$ de E , se llama *matriz de q con respecto a la base (a_i)* a la matriz

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

Como $a_{ij} = a_{ji}$, la matriz A es una matriz simétrica que cumple que $'XAX = q(x)$, donde X es el vector de coordenadas de x en la base (a_i) de E .

Recíprocamente, si tenemos una matriz simétrica A y definimos $q(x) = 'XAX$, se verifica fácilmente que q es una forma cuadrática sobre E , y A es la matriz de q con respecto a la base (a_i) de E .

Concluyendo podemos afirmar que la matriz A de la forma cuadrática q respecto a la base (a_i) de E es la única matriz simétrica tal que $q(x) = 'XAX$, donde X es el vector de coordenadas de x en la base (a_i) .

Ejemplos

1. La matriz respecto a la base canónica de la forma cuadrática definida sobre $\mathbb{R}^3$ por $q(x_1, x_2, x_3) = 3x_1^2 - x_2^2 + x_3^2 - 4x_1x_2 + 2x_2x_3$, es

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 1 \\ -2 & 1 & 1 \end{pmatrix}$$

2. Dada la matriz simétrica $A = \begin{pmatrix} 0 & -2 \\ -2 & 1 \end{pmatrix}$,

$$q(x_1, x_2) = (x_1, x_2) \begin{pmatrix} 0 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -4x_1x_2 + x_2^2$$

es una forma cuadrática sobre $\mathbb{R}^2$.

Observemos que la matriz de la forma cuadrática se obtiene muy fácilmente a partir de su expresión. Los coeficientes a_{ij} son los términos de la diagonal y cada uno de los coeficientes a_{ij} y a_{ji} es la mitad del coeficiente del término $x_i x_j$ en la expresión de la forma cuadrática, pues en esta aparecen reducidos los términos $x_i x_j$ y $x_j x_i$.

Así, en el ejemplo 1 el coeficiente de $x_1 x_2$ es -4 y por esto $a_{12} = a_{21} = -2$.

Recíprocamente, dada una matriz simétrica A , para buscar una forma cuadrática asociada a esta se pueden utilizar dos vías: multiplicar $'XAX$ como en el ejemplo 2 o por inspección, recordando que el coeficiente a_{ii} de la matriz será el del término x_i^2 y que el del término $x_i x_j$ será $2a_{ij} = 2a_{ji}$.

Así, en el ejemplo 2 el coeficiente de $x_1 x_2$ será $2a_{12} = -4$.

Realmente no es necesaria la suposición de que $a_{ij} = a_{ji}$, o sea, que su matriz es simétrica, en la definición de una forma cuadrática.

Si la matriz de una forma cuadrática es una matriz arbitraria A , podemos expresarla por $A = S + K$ donde si $'S = S$ y $'K = -K$ entonces $'XAX = 'XSX + 'XKX$. Y como para $K = (b_{ij})$ se cumple que $b_{ij} = -b_{ji}$, tenemos que $'XKX = 0$.

Luego $'XAX = 'XSX$, lo cual significa que siempre es posible considerar A simétrica.

Para conocer una forma cuadrática basta tener su expresión explícita como aplicación de E en K o una de sus matrices representativas. Ahora bien, ¿qué relación existe entre las matrices que representan a una forma cuadrática en diferentes bases? ¿Son iguales? ¿Son semejantes?

Sean (a_i) y (b_i) bases de E , y sean A y B las matrices de una forma cuadrática q sobre E con respecto a las bases (a_i) y (b_i) , respectivamente, es decir,

$q(x) = 'XAX$, donde X es el vector de coordenadas de x en la base (a_i) ,

$q(x) = 'YBY$, donde Y es el vector de coordenadas de x en la base (b_i) .

Como dada una base del espacio, la matriz de una forma cuadrática en dicha base es única, para determinar B bastará encontrar para $q(x)$ una expresión matricial en la que aparezca el vector de coordenadas de x en (b_i) . Sea P la matriz de cambio de base de la base (a_i) a la base (b_i) ; entonces $PY = X$ donde P es invertible, lo que permite obtener

$$q(x) = 'XAX = '(PY)A(PY) = 'Y(PAP)Y = 'YBY.$$

Luego $B = 'PAP$.

La relación entre las matrices asociadas a una misma forma cuadrática en bases distintas, cumple una relación diferente a la que cumplían las matrices asociadas a un endomorfismo en bases distintas.

En el caso de las formas cuadráticas obtenemos $B = 'PAP$ con P invertible, mientras que en el caso de las aplicaciones lineales obtuvimos $B = P^{-1}AP$ con P invertible.

En el próximo capítulo veremos que estos dos resultados están relacionados entre sí.

DEFINICIÓN 1.1.3

Dos matrices cuadradas A y B con coeficientes en K son *congruentes*, si existe una matriz invertible P , con coeficientes en K , tal que $B = 'PAP$.

De acuerdo con esta definición, queda demostrada la primera parte del teorema siguiente:

TEOREMA 1.1.1

Dos matrices representativas de la misma forma cuadrática son matrices congruentes y dos matrices simétricas congruentes representan la misma forma cuadrática en diferentes bases.

Demostración

Solo demostraremos la segunda parte, ya que la primera ha sido demostrada anteriormente.

Supongamos que A y B son matrices simétricas congruentes, entonces existe una matriz P invertible tal que $B = 'PAP$. Como toda matriz invertible es una matriz de cambio de base, supongamos que P sea la matriz de cambio de base de la base (a_i) a la (b_i) .

Como B es una matriz simétrica, representa a una forma cuadrática q en la base (b) . Sea Y el vector de coordenadas de x en la base (b) ; entonces

$$q(x) = Y^T B Y = Y^T (PAP) Y = (PY)^T A (PY) = X^T A X,$$

donde X es el vector de coordenadas de x en la base (a) y, por tanto, A es la matriz de q en la base (a) .

Es conocido que si se multiplica una matriz A por una matriz inversible, P , el rango del producto es igual al rango de A , de donde podemos concluir que matrices congruentes tienen igual rango. Esto nos permite dar la definición siguiente:

DEFINICIÓN 1.1.4

Se llama *rango de la forma cuadrática q* al rango de cualesquiera de sus representaciones matriciales.

Ejemplos

1. La forma cuadrática sobre $\mathbb{R}^2$ definida por $q(x, y) = x^2 - 2xy + 3y^2$ es de rango 2, porque la matriz de q en la base canónica,

$$\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix},$$

es de rango 2.

2. La forma cuadrática sobre $\mathbb{R}^3$ definida por $q(x, y, z) = x^2 + 2y^2 + 3z^2$ es de rango 3.

1.2 Forma canónica de una forma cuadrática

En geometría es de gran importancia la reducción de una curva de segundo orden a la forma canónica, o sea, a una expresión en que no aparezcan términos en xy . Mediante esta reducción se puede hacer más fácil el estudio de estas curvas.

En el estudio de los máximos o mínimos de una función de dos variables $z = f(x, y)$ en un punto (x_0, y_0) en el cual se anulen $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$, obtenemos mediante el desarrollo de Taylor que

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(x_0, y_0) h^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) h k + \frac{\partial^2 f}{\partial y^2}(x_0, y_0) k^2 \right].$$

Para h y k pequeños, esto es, en un entorno de (x_0, y_0) , está claro que el carácter de posible máximo o mínimo de (x_0, y_0) está determinado por el signo de la expresión dentro del corchete, que no es más que una forma cua-

drática en las variables h y k . Este signo, se podría analizar más fácilmente si no apareciera el término en hk .

En el estudio de los máximos y mínimos de funciones de tres o más variables se obtienen resultados análogos.

Lo anteriormente expuesto conduce al análisis de este problema para una forma cuadrática arbitraria.

DEFINICIÓN 1.2.1

Se llama *forma canónica* de una forma cuadrática, a toda expresión de esta en la cual no aparezcan términos en $x_i y_j$, es decir, a una expresión del tipo

$$q(x) = a_1 x_1^2 + a_2 x_2^2 + \dots + a_n x_n^2.$$

Analicemos ahora en qué casos es posible reducir una forma cuadrática a su forma canónica.

En el caso de un espacio vectorial de dimensión 1, esto es trivial. Veamos qué ocurre para una forma cuadrática en dos variables, esto es, definida sobre un espacio vectorial de dimensión 2.

Sea $q(x) = a_{11} x_1^2 + 2a_{12} x_1 x_2 + a_{22} x_2^2$, con $a_{11} \neq 0$.

Utilizando el conocido recurso de completar el cuadrado, obtenemos

$$\begin{aligned} q(x) &= a_{11} \left(x_1^2 + 2 \frac{a_{12}}{a_{11}} x_1 x_2 + \left(\frac{a_{12}}{a_{11}} \right)^2 x_2^2 \right) - \frac{a_{12}^2}{a_{11}} x_2^2 + a_{22} x_2^2 \\ &= a_{11} \left(x_1 + \frac{a_{12}}{a_{11}} x_2 \right)^2 + \left(\frac{a_{12}^2}{a_{11}} + a_{22} \right) x_2^2. \end{aligned}$$

Haciendo el cambio de coordenadas

$$y_1 = x_1 + \frac{a_{12}}{a_{11}} x_2 \quad \text{y} \quad y_2 = x_2,$$

obtenemos

$$q(x) = a_{11} y_1^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}} \right) y_2^2,$$

que está dada en forma canónica.

Es lógico pensar en la posibilidad de demostrar el resultado anterior en un espacio vectorial de dimensión n cualquiera, lo cual haremos en el siguiente teorema, en el que generalizaremos el método que acabamos de utilizar para el caso $n=2$.

TEOREMA 1.2.1

Sea q una forma cuadrática definida sobre un espacio vectorial E de dimensión n . Entonces existe una base $(a_i)_{1 \leq i \leq n}$ de E , en la que $q(x)$ se escribe en forma canónica.

Demostración

Realizaremos la demostración por inducción en $n = \dim E$.

$$\text{Sea } q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j \quad (2)$$

Para el caso $n=1$, $q(x) = a_{11}x_1^2$ está dada en forma canónica.

Supongamos que para el caso $n-1$ es cierto el teorema y probemos que se cumple para n .

En la demostración debemos diferenciar dos casos:

1ro. Existe algún $a_{ii} \neq 0$.

Supongamos que $a_{11} \neq 0$. (Obsérvese que con esto no se pierde generalidad, pues existe algún $a_{ii} \neq 0$ y podemos lograr mediante un cambio de coordenadas que $a_{11} \neq 0$.)

Entonces

$$q(x) = (a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n) + \sum_{i=2}^n a_{ii} x_i^2 + 2 \sum_{2 \leq i < j \leq n} a_{ij} x_i x_j$$

pero

$$\begin{aligned} a_{11}x_1^2 + 2x_1(a_{12}x_2 + \dots + a_{1n}x_n) &= a_{11} \left\{ x_1^2 + 2x_1 \left[a_{11}^{-1} (a_{12}x_2 + \dots + a_{1n}x_n) \right] + \right. \\ &\quad \left. + \left[a_{11}^{-1} \left(\sum_{i=2}^n a_{1i} x_i \right) \right]^2 \right\} - a_{11}^{-1} \left[\sum_{i=2}^n a_{1i} x_i \right]^2 \\ &= a_{11} (x_1 + a_{11}^{-1} a_{12} x_2 + \dots + a_{11}^{-1} a_{1n} x_n)^2 - a_{11}^{-1} \left[\sum_{i=2}^n a_{1i} x_i \right]^2 \end{aligned} \quad (3)$$

sustituyendo en (3) tenemos

$$q(x) = a_{11} (x_1 + a_{11}^{-1} a_{12} x_2 + \dots + a_{11}^{-1} a_{1n} x_n)^2 + h(x),$$

donde

$$h(x) = \sum_{i=2}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j - a_{11}^{-1} \left[\sum_{i=2}^n a_{1i} x_i \right]^2$$

$$h(x) = \sum_{i=2}^n \beta_{ii} x_i^2 + 2 \sum_{2 \leq i < j \leq n} \beta_{ij} x_i x_j$$

es una forma cuadrática que depende de $n-1$ coordenadas, la cual se puede reducir a la forma canónica según la hipótesis de inducción.

Sea

$$\begin{aligned} y_1 &= x_1 + a_{11}^{-1} a_{12} x_2 + \dots + a_{11}^{-1} a_{1n} x_n \\ y_2 &= x_2 \\ &\vdots \\ y_n &= x_n \end{aligned} \quad (4)$$

La expresión (4) es una verdadera transformación de coordenadas, ya que la matriz de la transformación

$$P = \begin{pmatrix} 1 & a_{11}^{-1} a_{12} & \dots & a_{11}^{-1} a_{1n} & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

es inversible.

$$\text{Entonces } q(x) = a_{11} y_1^2 + h(y_2, \dots, y_n). \quad (5)$$

Hemos obtenido (5) aplicando a (2) la transformación de coordenadas (4), es decir, hemos hallado una base del espacio vectorial E tal que la expresión de la forma cuadrática q en dicha base es (5). Y como por la hipótesis de inducción $h(y_2, \dots, y_n)$ se puede reducir a una suma de cuadrados, hemos demostrado que es posible hallar una base (a) de W tal, que la expresión de q en dicha base es $q(x) = a_1 z_1^2 + \dots + a_n z_n^2$ con $z_1 = y_1$, $z_i = \sum_{j=2}^n y_j a_{ij}$

$$2 \leq i \leq n.$$

2do. Para todo i , $a_{ii} = 0$.

Sea $a_{ii} = 0$, para $i=1, 2, \dots, n$. Entonces

$$q(x) = 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

Supongamos que $a_{12} \neq 0$ (existe algún $a_{ij} \neq 0$ porque si no, $q(x) = 0$ y esta sería su forma canónica).

Al igual que en el caso anterior, podemos suponer que $a_{12} \neq 0$ sin perder generalidad.

Realicemos la siguiente transformación de coordenadas:

$$x_1 = y_1 + y_2$$

$$x_2 = y_1 - y_2$$

$$x_i = y_i, \quad 3 \leq i \leq n$$

De nuevo la matriz Q de la transformación es inversible.

$$Q = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 1 & -1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Entonces

$$q(x) = 2a_{12}(y_1^2 - y_2^2) + 2(y_1 + y_2)(a_{13}y_3 + \dots + a_{1n}y_n) + \\ + 2(y_1 - y_2)(a_{23}y_3 + \dots + a_{2n}y_n) + 2 \sum_{3 \leq i < j \leq n} a_{ij}y_i y_j \\ = 2a_{12}y_1^2 - 2a_{12}y_2^2 + 2 \sum_{1 \leq i < j \leq n} \beta_{ij} y_i y_j$$

con lo que se reduce el problema al primer caso.

La demostración que hemos hecho por inducción se traduce en un proceso iterativo en la práctica, esto es, vamos repitiendo el proceso para las diferentes variables hasta obtener la reducción completa. Veamos el siguiente ejemplo:

Ejemplos

1. Sea $q(x) = x_1^2 + 4x_2^2 + 8x_3^2 - x_4^2 - 4x_1x_2 + 6x_1x_3 - 12x_2x_4 + 2x_3x_4$.

Hacemos

$$q_1(x) = (x_1^2 - 4x_1x_2 + 6x_1x_3) + 4x_2^2 + 8x_3^2 - x_4^2 - 12x_2x_4 + 2x_3x_4 \\ = (x_1^2 - 4x_1x_2 + 6x_1x_3 + 4x_2^2 + 9x_3^2 - 12x_2x_4) - 4x_2^2 - 9x_3^2 + 4x_2^2 + 8x_3^2 - x_4^2 + 12x_2x_4 \\ + 2x_3x_4 - 12x_2x_4 \\ = (x_1 + 2x_2 + 3x_3)^2 - x_3^2 - x_4^2 + 12x_2x_4 + 2x_3x_4 - 12x_2x_4$$

para

$$y_1 = x_1 + 2x_2 + 3x_3$$

$$y_2 = x_2$$

$$y_3 = x_3$$

$$y_4 = x_4$$

obtenemos

$$q(x) = y_1^2 - y_3^2 - y_4^2 + 12y_2y_3 + 2y_2y_4 - 12y_3y_4 = y_1^2 + q_1(y_2, y_3, y_4)$$

donde $q_1(y_2, y_3, y_4) = -y_3^2 - y_4^2 + 12y_2y_3 + 2y_2y_4 + 12y_3y_4$.

Escribimos

$$q(x) = y_1^2 - (y_3^2 - 12y_2y_3 - 2y_2y_4 - y_4^2 - 12y_3y_4) \\ = y_1^2 - (y_3^2 - 12y_2y_3 - 2y_2y_4 + 36y_3^2 + y_4^2 + 12y_2y_4 + y_4^2 + 36y_3^2 - y_4^2 + \\ + 12y_2y_4 - 12y_3y_4) \\ = y_1^2 - (y_3 - 6y_2 + y_4)^2 + 36y_3^2$$

Con la transformación

$$z_1 = y_1$$

$$z_2 = y_2$$

$$z_3 = y_3 - 6y_2 + y_4$$

$$z_4 = y_4$$

obtenemos

$$q(x) = z_1^2 + 36z_2^2 - z_3^2$$

La transformación final es

$$z_1 = x_1 - 2x_2 + 3x_3$$

$$z_2 = x_2$$

$$z_3 = x_3 - 6x_2 + x_4$$

$$z_4 = x_4$$

Si en lugar de la segunda transformación hubiéramos hecho la siguiente:

$$z_1 = y_1$$

$$z_2 = y_1 - 6y_2 + y_4$$

$$z_3 = y_3$$

$$z_4 = y_4$$

habríamos obtenido

$$q(x) = z_1^2 - z_2^2 + 36z_3^2$$

que es una forma canónica diferente a la anterior.

2. Reducir a la forma canónica la forma cuadrática sobre $\mathbb{R}^3$.

$$q(x, y, z) = 2xy - 6yz$$

Consideremos la transformación de coordenadas:

$$x = (x_1 + x_2), \quad y = (x_1 - x_2), \quad z = x_3$$

$$q(x, y, z) = 2(x_1 + x_2)(x_1 - x_2) - 6(x_1 - x_2)x_3 \quad x_3 = 2x_1^2 - 2x_2^2 - 6x_1x_3 + 6x_2x_3 =$$

$$= 2x_1^2 - 6x_1x_3 - 2x_2^2 + 6x_2x_3 = 2\left(x_1^2 - 2\left[\frac{3}{2}x_1\right]x_3\right) - 2x_2^2 + 6x_2x_3$$

$$= 2\left(x_1 - \frac{3x_3}{2}\right)^2 - \frac{9}{2}x_3^2 - 2x_2^2 + 6x_2x_3$$

$$= 2\left(x_1 - \frac{3}{2}x_3\right)^2 - 2\left(x_2^2 - 3x_2x_3 + \frac{9}{4}x_3^2\right)$$

$$= 2\left(x_1 - \frac{3}{2}x_3\right)^2 - 2\left(x_2 - \frac{3}{2}x_3\right)^2$$

Realizando el cambio de coordenadas

$$y_1 = x_1 - \frac{3}{2}x_3$$

$$y_2 = x_2 - \frac{3}{2}x_3$$

$$y_3 = x_3$$

obtenemos

$$q(x) = 2y_1^2 - 2y_2^2$$

que es una expresión en forma canónica de la forma cuadrática q .

La transformación final es

$$x = y_1 + y_2 + 3y_3$$

$$y = y_1 - y_2$$

$$z = y_3$$

En términos matriciales, las transformaciones realizadas son

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

o sea,

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

es la matriz de la primera transformación

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

donde

$$Q = \begin{pmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & 1 & -\frac{3}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

es la matriz de la segunda transformación.

Si llamamos

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad X' = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

tenemos que $X = PX'$ y $Y = QX'$, por lo cual $X = PQ^{-1}Y$ y PQ^{-1} es la matriz de la transformación final que nos permite expresar x, y, z en función de y_1, y_2, y_3 .

Con respecto a estos ejemplos queremos recordar que las matrices de transformación se hallan por simple inspección de la relación que existe entre las coordenadas, y que cada matriz de transformación es a su vez una matriz de cambio de base. Por tanto, si suponemos que x, y, z son las coordenadas de un vector de R^3 en una base $(a_i)_{i=1,2,3}$, x_1, x_2, x_3 las coordenadas de dicho vector en la base $(h_i)_{i=1,2,3}$ y $Y = PX'$, entonces P es la matriz de cambio de base de la base (a_i) a la base (h_i) .

En muchos casos es importante conocer no solo la forma canónica sino la transformación de coordenadas mediante la cual se obtuvo esta. Dicha transformación puede hallarse directamente o por vía matricial, según vimos en el ejemplo precedente.

Haciendo el siguiente cambio de coordenadas en el ejemplo 2:

$z_1 = \sqrt{2} y_1, z_2 = \sqrt{2} y_2, z_3 = y_3$, obtenemos $q(x) = z_1^2 - z_2^2$, que es otra expresión en forma canónica de la forma cuadrática q .

De los ejemplos anteriores podemos concluir que la expresión en forma canónica de una forma cuadrática no es única. Veamos entonces qué tienen de común las diversas formas canónicas de una forma cuadrática dada.

Lo primero que observamos es que en la forma canónica aparecen exactamente r coeficientes no nulos, donde r es el rango de q , es decir, $q(x) = a_1 x_1^2 + \dots + a_r x_r^2$ con $a_i \neq 0$ para $1 \leq i \leq r$, independientemente del cambio de coordenadas que se realice. Esto se infiere de la siguiente interpretación matricial de la reducción de una forma cuadrática a forma canónica.

Suponiendo que A es la matriz de la forma cuadrática q en una cierta base (a_i) , cuando reducimos q a la forma canónica encontramos una expresión diagonal de q en otra base, o sea, encontramos una matriz diagonal D y una inversible P tales que $D = P^t A P$.

Ya señalamos anteriormente que dos matrices congruentes tienen el mismo rango, luego D y A tienen el mismo rango, que es el de q . Al ser D diagonal, se infiere la afirmación anterior sobre el número de elementos no nulos en cualquier forma canónica de una forma cuadrática.

En el ejemplo 2 vimos un tipo de transformación que conduce a una expresión en forma canónica donde en los coeficientes, además de los posibles 0, solo aparecen 1 y -1. Esto simplifica el análisis de los posibles invariantes de la forma cuadrática al ser reducida a la forma canónica. Veamos este tipo de transformación para una forma cuadrática arbitraria.

Analicemos por separado los casos complejos y real.

1ro. $K = \mathbb{C}$ (caso complejo)

Como es posible hallar la raíz cuadrada de todo número complejo, se puede hacer el siguiente cambio de coordenadas:

$$y_1 = \sqrt{a_1} x_1, \dots, y_r = \sqrt{a_r} x_r, y_{r+1} = x_{r+1}, \dots, y_n = x_n$$

de donde

$$q(x) = y_1^2 + \dots + y_r^2$$

Resumiendo podemos decir que en el caso complejo es posible reducir la forma cuadrática a una suma de cuadrados donde todos los coeficientes son iguales a 1 o 0 (r coeficientes iguales a 1 y $n-r$ coeficientes iguales a 0). Esta forma de expresar a q se denomina *forma normal*.

Como conclusión podemos afirmar que la forma normal de una forma cuadrática compleja depende sólo del rango de la forma cuadrática, de ahí que todas las formas cuadráticas complejas del mismo rango puedan ser reducidas a la misma forma normal.

2do. $K = \mathbb{R}$ (caso real)

En el caso real, la forma canónica de q puede ser escrita en la forma

$$q(x) = a_1 x_1^2 + \dots + a_q x_q^2 - a_{q+1} x_{q+1}^2 - \dots - a_r x_r^2$$

donde $a_i > 0$, $i = 1, 2, \dots, r$.

Debido a que es posible hallar la raíz cuadrada de todo número real positivo, podemos hacer el siguiente cambio de coordenadas:

$$y_1 = \sqrt{a_1} x_1, \dots, y_r = \sqrt{a_r} x_r, y_{r+1} = x_{r+1}, \dots, y_n = x_n,$$

de donde

$$q(x) = y_1^2 + \dots + y_q^2 - y_{q+1}^2 - \dots - y_r^2.$$

Esta forma de expresar a q se denomina también *forma normal*.

En los ejemplos 1 y 2 en que redujimos a la forma canónica formas cuadráticas reales, observamos que las distintas formas canónicas obtenidas tenían, además de igual rango, la misma cantidad de términos positivos y negativos. Esto no es casual, como demostraremos en el teorema siguiente:

TEOREMA 1.2.2. (Ley de inercia o de Sylvester para formas cuadráticas reales)

En una forma cuadrática real, el número de cuadrados positivos y el número de cuadrados negativos que aparece en cualquiera de sus formas canónicas es siempre el mismo.

La ley de inercia para formas cuadráticas reales asegura que el número de coeficientes positivos y negativos de cualquier forma canónica no depende de los cambios de coordenadas que se efectúen, sino que están determinados únicamente por la propia forma cuadrática. En particular, esto es cierto para la forma normal, con lo que queda demostrado que esta es única salvo el orden de las variables.

Demostración

Es conocido que el número de coeficientes no nulos es r , donde r es el rango de q , por lo que basta probar que el número de coeficientes positivos (o negativos) de dos expresiones canónicas distintas de $q(x)$ es el mismo.

Sean

$$q(x) = a_1 x_1^2 + \dots + a_k x_k^2 - a_{k+1} x_{k+1}^2 - \dots - a_r x_r^2, \quad a_i > 0$$

(1)

la expresión canónica de $q(x)$ en una base (a_i) y

$$q(x) = \beta_1 y_1^2 + \dots + \beta_k y_k^2 - \beta_{k+1} y_{k+1}^2 - \dots - \beta_r y_r^2, \quad \beta_i > 0$$

(2)

la expresión canónica de $q(x)$ en una base (b_i) .

Supongamos que $k' < k$. Entonces $q(x) > 0$ siempre que

$$x_{k+1} = \dots = x_r = 0.$$

(3)

Las x que satisfacen las $r-k$ ecuaciones dadas por (3), forman un subespacio S_1 de E de dimensión $n-(r-k)$, exactamente el dado por los vectores de forma $(x_1, \dots, x_k, 0, \dots, 0, x_{r+1}, \dots, x_n)$.

Análogamente de (2) resulta que $q(x) < 0$ para los x tales que

$$y_1 = \dots = y_k = y_{k+1} = \dots = y_r = 0$$

(4)

Los x que satisfacen (4) forman un subespacio S_2 de dimensión $r-k'$ de E .

La suma de las dimensiones de S_1 y S_2 es

$$n-(r-k) + (r-k') = n+k-k' > n.$$

Luego $S_1 \cap S_2 \neq \{0\}$ y existe $z \neq 0$ tal que $z \in S_1 \cap S_2$. Para este vector z , se cumple $q(z) > 0$, y $q(z) < 0$, lo que sería una contradicción.

Si suponemos $k < k'$, llegaremos a una contradicción semejante. Luego $k = k'$ según queríamos demostrar y, por consiguiente, el número de términos positivos y negativos que aparecen en toda expresión canónica de una forma cuadrática depende únicamente de dicha forma y no del sistema de coordenadas en que esta aparece expresada.

Se denomina *índice positivo de inercia* al número de cuadrados positivos, *índice negativo de inercia* al número de cuadrados negativos y *signatura* o *signo de q* a la diferencia entre el índice positivo y el índice negativo de inercia. Evidentemente, si se conoce el rango de una forma cuadrática y cualquiera de estos tres números, se pueden determinar los otros dos.

Ejemplo

$$q(x) = 2xy - 6yz \text{ se reduce a la forma normal } q(x) = z_1^2 - z_2^2.$$

El índice positivo de inercia es igual al índice negativo e igual a 1. La signatura de q es igual a 0.

Para concluir observemos que todas las formas cuadráticas reales que tengan el mismo rango y la misma signatura, posean la misma forma normal. Por tanto, podemos asegurar que la forma normal de una forma cuadrática real depende sólo del rango y del signo.

Por ejemplo, en un espacio vectorial de dimensión 2 las únicas formas normales de rango 2 son

$$x^2 + y^2, x^2 - y^2, -x^2 - y^2,$$

que en el caso del plano representarán una circunferencia, una hipérbola y el conjunto vacío, respectivamente, si consideramos las ecuaciones $x^2 + y^2 = 1$, $x^2 - y^2 = 1$ y $-x^2 - y^2 = 1$.

La única forma normal de rango 1 es x^2 , que en el caso del plano da lugar a las rectas $x=1$ y $x=-1$, si consideramos $x^2=1$, y el conjunto vacío si consideramos $-x^2=1$, lo cual significa que cualquier cónica con centro dada por una ecuación del tipo $ax^2+bx+cy^2=1$ puede ser reducida, mediante una transformación de coordenadas conveniente, a uno de los lugares geométricos anteriormente citados.

Se puede demostrar que esta transformación consiste en una combinación de traslaciones, compresiones (dilataciones) y rotaciones, lo que geoméricamente significa que, por ejemplo, una elipse puede transformarse hasta obtener una circunferencia, pero que no es posible mediante una transformación afín reducir una circunferencia a una hipérbola, ya que sus formas normales son distintas.

Anteriormente hicimos referencia a la utilización de la reducción de formas cuadráticas en el estudio de los máximos o mínimos de funciones de dos o más variables. Vimos, entonces, que dada una función $z=f(x, y)$ de $\mathbb{R}^2$, el carácter de máximo o mínimo de un punto (x_0, y_0) está dado por el signo de la expresión

$$q(\xi) = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)hk + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)k^2,$$

que es una forma cuadrática en las variables h y k . Si esta forma tiene rango 2, por ejemplo, sus posibles formas normales son

$$h^2+k^2, h^2-k^2 \text{ y } -h^2-k^2.$$

Como tenemos que

$$f(x_0+h, y_0+k) = f(x_0, y_0) + \frac{q(x)}{2},$$

resulta que si la forma normal es h^2+k^2 , $f(x_0, y_0) < f(x_0+h, y_0+k)$ y (x_0, y_0) es un punto de mínimo; si la forma normal es h^2-k^2 , resulta lo contrario y (x_0, y_0) es un punto de máximo, y por último, si la forma normal es $-h^2-k^2$, (x_0, y_0) no es ni máximo ni mínimo, y este tipo de punto es el conocido por *punto de ensilladura*.

1.3 Formas cuadráticas reales de signo constante y formas definidas positivas

En el epígrafe anterior hemos visto las ventajas de que una forma cuadrática dada definida sobre E tenga el mismo signo para todo elemento de E . Esta situación se repite frecuentemente en la práctica, por lo que estudiaremos las llamadas *formas de signo constante*, en particular las *formas definidas positivas*.

DEFINICIÓN 1.3.1 Una forma cuadrática real se llama *forma definida positiva* si el número de cuadrados positivos de su forma normal es n , es decir, si el rango y el índice positivo de inercia coincide con el número de variables.

Ejemplos

1. En $\mathbb{R}^1$, $q(x, y, z) = x^2 + y^2 + z^2$ es definida positiva.
2. En $\mathbb{R}^1$, $q(x, y, z) = x^2 + z^2$ no es definida positiva.
3. En $\mathbb{R}^2$, $q(x, y) = x^2 + 2xy + y^2$ no es definida positiva porque haciendo $x_1 = x+y$ y $y_1 = y$, se obtiene $q(x_1, y_1) = x_1^2$.
4. En $\mathbb{R}^2$, $q(x, y) = x^2 - y^2$ no es definida positiva.

Para determinar si una forma es definida positiva, de acuerdo con la definición, debemos determinar previamente su forma normal o reducirla a la forma canónica. El siguiente teorema nos da una caracterización de las formas cuadráticas definidas positivas sin necesidad de reducirlas a la forma canónica y, además, justifica que las denominemos de ese modo.

TEOREMA 1.3.1

La forma cuadrática real q es definida positiva, si y solo si $q(x) > 0$ para todo $x \neq 0$.

Demostración

Si q es definida positiva, entonces

$$q(x) = y_1^2 + \dots + y_n^2 \text{ y evidentemente } q(x) > 0 \text{ si } x \neq 0.$$

Recíprocamente, probemos que si $q(x) > 0$ para todo $x \neq 0$, q es definida positiva.

Supongamos que en la forma normal no aparece alguna coordenada al cuadrado, por ejemplo,

$$q(x) = y_1^2 + \dots + y_{n-1}^2 - y_n^2$$

Entonces tomando x tal que $y_1 = \dots = y_{n-1} = 0$ y $y_n = 1$, tendremos $q(x) = 0$ y $x \neq 0$, lo cual contradice nuestra suposición.

Supongamos ahora que alguna coordenada aparece con signo negativo, por ejemplo, $q(x) = y_1^2 + \dots + y_{n-1}^2 - y_n^2$.

En este caso tomando x tal que $y_1 = \dots = y_{n-1} = 0$ y $y_n = 1$, tendremos $q(x) = -1$, lo que también es contrario a nuestra hipótesis.

Por tanto, en la forma normal de q aparecen todas las coordenadas al cuadrado con signo positivo.

Obsérvese que esta demostración es independiente del sistema de coordenadas en el cual se exprese la forma cuadrática.

El teorema 1.3.1 es muy importante y se emplea mucho en el trabajo con formas cuadráticas definidas positivas; sin embargo, en la práctica no siempre resulta fácil utilizarlo.

A continuación estudiaremos una nueva caracterización de las formas definidas positivas que depende solo de los coeficientes de la forma.

Queremos recordar previamente que dada una matriz $A=(a_{ij})$ cuadrada de orden n , se llaman *menores principales* de A a los menores de orden 1, 2, ..., n de esta matriz situados en el ángulo superior izquierdo, es decir,

$$a_{11}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

TEOREMA 1.3.2

Una forma cuadrática real q es definida positiva si y solo si todos los menores principales de la matriz asociada a q son estrictamente positivos.

Demostración

Haremos la demostración por inducción sobre el número de coordenadas, o lo que es lo mismo, sobre la dimensión del espacio.

Si $n=1$, entonces $q(x)=ax^2$, y es evidente que q es definida positiva si y solo si $a>0$.

Supongamos que el teorema se cumple para $n-1$ y probémoslo para n . La forma cuadrática

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

se puede escribir de la forma

$$q(x) = q_1(x_1, \dots, x_{n-1}) + 2 \sum_{i=1}^{n-1} a_{in} x_i x_n + a_{nn} x_n^2,$$

donde q_1 es una forma cuadrática en $n-1$ coordenadas.

Supongamos que q es definida positiva y probemos que también lo es q_1 .

Tomemos $q(x_1, \dots, x_{n-1}) \leq 0$ para algún vector de coordenadas $(x_1, \dots, x_{n-1}) \neq 0$.

Sea $x = (x_1, \dots, x_{n-1}, 0)$, $x_n = 0$. Entonces $q(x) \leq 0$ y $x \neq 0$, lo cual contradice que q sea definida positiva. Por tanto, q_1 es definida positiva.

Ahora bien, q_1 es definida positiva si y solo si todos los menores principales de la matriz asociada a q_1 son estrictamente positivos. Entonces

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n-1} \\ a_{21} & a_{22} & \dots & a_{2n-1} \\ \dots & \dots & \dots & \dots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,n-1} \end{vmatrix} > 0$$

y estos menores de la matriz asociada a q_1 coinciden con los menores de la matriz asociada a q , debido a la forma en que tomamos q_1 .

Solo queda por analizar

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

Como q es definida positiva, haciendo un cambio de coordenadas conveniente tendremos

$$q(x) = z_1^2 + \dots + z_n^2.$$

La matriz de q en la nueva base es la matriz I_n , que es congruente con A , es decir:

$$A = {}^t P I_n P = {}^t P P,$$

donde P representa la matriz de cambio de base correspondiente. $\det A = (\det {}^t P)(\det P)$: pero como $\det {}^t P = \det P \neq 0$, $\det A = (\det P)^2 > 0$

Con esto demostramos que si q es definida positiva, todos sus menores principales son estrictamente positivos.

Hagamos la demostración inversa, es decir, supongamos que todos los menores principales de la matriz de q son estrictamente positivos y probemos que q es definida positiva.

Como

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \begin{vmatrix} a_{11} & \dots & a_{1n-1} \\ \dots & \dots & \dots \\ a_{n-1,1} & \dots & a_{n-1,n-1} \end{vmatrix} > 0$$

y estos menores coinciden con los de la matriz de q_1 , entonces $q_1(x_1, \dots, x_{n-1})$ es definida positiva según la hipótesis de inducción.

Realizando un cambio de coordenadas adecuado, tendremos

$$q_1(x_1, \dots, x_{n-1}) = y_1^2 + \dots + y_{n-1}^2.$$

Haciendo $x_n = y_n$,

$$q(x) = \sum_{i=1}^{n-1} y_i^2 + 2 \sum_{i=1}^{n-1} b_{in} y_i y_n + b_{nn} y_n^2.$$

y como

$$y_i^2 + 2b_{in} y_i y_n = (y_i + b_{in} y_n)^2 - b_{in}^2 y_n^2, \text{ entonces}$$

$$q(x) = \sum_{i=1}^{n-1} (y_i + b_{in} y_n)^2 + (b_{nn} - b_{1n}^2 - \dots - b_{n-1,n}^2) y_n^2.$$

Sea $c = b_{nn} - b_{1n}^2 - \dots - b_{n-1,n}^2$. Haciendo el cambio de coordenadas:

$$z_i = y_i + b_{in} y_n, \quad i = 1, \dots, n-1$$

$$z_n = y_n$$

obtenemos

$$q(x) = z_1^2 + \dots + z_{n-1}^2 + cz_n^2$$

La matriz de q en la nueva base es

$$D = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{pmatrix}$$

que es congruente con A , es decir, $D = 'QAQ$, donde Q es la correspondiente matriz de cambio de base:

$$\det D = (\det 'Q)(\det A)(\det Q) = (\det A) > 0.$$

Por tanto, q es definida positiva ya que su índice positivo de inercia es n .

Una definición análoga a la de forma definida positiva, se tiene cuando en la forma normal aparecen todos los cuadrados con signo negativo. En este caso se dice que la forma es definida negativa y se cumplen resultados similares a los dados en los teoremas 1.3.1 y 1.3.2.

Se acostumbra a decir que una forma real es positiva cuando el índice positivo de inercia es igual al rango de q y menor que n . Análogamente se definen las formas negativas.

Por supuesto, hay formas que no son positivas ni negativas, pues en sus formas normales aparecen cuadrados positivos y negativos, es decir, sus índices positivos y negativos de inercia son mayor o igual a 1.

Ejercicios

1. a) Sea $q(x)$ la forma cuadrática definida sobre $\mathbb{R}^2$ cuya matriz con respecto a la base $(a_1 = (1, 2), a_2 = (3, 5))$ es

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

Encuentre $q(1, 0)$.

- b) Considere $a_1 = (1, -2), a_2 = (1, -1), A = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ y halle $q(x, y)$.

2. Halle la matriz simétrica asociada a cada una de las formas cuadráticas siguientes:

a) $2x^2 + 3xy + 6y^2$

b) $8xy + 4y^2$

c) $4xy$

d) $x^2 + 2xy + 4xz + 3y^2 + yz + 7z^2$

e) $x^2 + 4xy + 4y^2 + 2xz + z^2 + 4yz$

3. Sea q la forma cuadrática definida sobre $\mathbb{R}^3$ por la expresión $q(x, y, z) = xy + y^2 + 4xz + z^2$.

- a) Encuentre la matriz de q con respecto a la base canónica de $\mathbb{R}^3$.

- b) Encuentre la matriz de q con respecto a la base (a) de $\mathbb{R}^3$, donde $a_1 = (1, 1, 1), a_2 = (1, 0, 1), a_3 = (1, 0, 0)$ y tiene dicha forma canónica y $a_4 = (1, 0, 0)$.
- c) Verifique la relación de congruencia existente entre ambas matrices.

4. Dada la matriz simétrica

$$A = \begin{pmatrix} 1 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 \end{pmatrix}$$

halle una matriz diagonal D y una inversible P tales que $A = 'PDP$.

5. Encuentre una base de $\mathbb{R}^4$ con respecto a la cual sea diagonal la forma cuadrática cuya matriz en la base canónica es

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 2 & -1 \end{pmatrix}$$

6. Reduzca a dos formas canónicas diferentes la forma cuadrática $x_1^2 + x_2^2$ considerada sobre un espacio vectorial de dimensión 2 sobre $\mathbb{Q}$.

7. Reduzca a la forma canónica las formas cuadráticas siguientes:

a) $x_1^2 + 2x_1x_2 + x_2^2 + 4x_2x_3 + 5x_3^2$

b) $x_1^2 - 4x_1x_2 + 2x_2x_3 + 4x_3^2 + x_2^2$

c) $x_1x_2 + x_2x_3 + x_1x_3$

d) $x_1^2 - 2x_1x_2 + 2x_2x_3 - 2x_1x_4 + x_2^2 + 2x_2x_4 - 4x_2x_4 - 2x_4^2$

e) $x_1^2 + x_1x_2 + x_2x_4$

f) $x_1^2 + 4x_2^2 + 8x_3^2 - 4x_1x_2 + 6x_2x_3 - 12x_2x_4 + 2x_2x_4 + x_3x_4 + x_4x_4$

g) $3x_1^2 + 4x_1x_2 - x_2^2 + 8x_1x_3 - 6x_2x_3 + x_3^2$

8. Reduzca a la forma canónica las formas cuadráticas siguientes:

a) $\sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j$

b) $\sum_{1 \leq i < j \leq n} x_i x_j$

c) $x_1x_n + x_2x_{n-1} + \dots + x_{n-1}x_2$ ($s = \lfloor \frac{n}{2} \rfloor$)

9. Reduzca a una forma canónica la forma cuadrática definida sobre $\mathbb{R}^3$ mediante la expresión

$$q(x_1, x_2, x_3) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_1)^2$$

e indique la relación existente entre las coordenadas en las cuales se obtiene dicha forma canónica y las coordenadas iniciales (x_i) .

10. Reduzca $x_1x_2 + x_2x_3 + \dots + x_{n-1}x_n$ a la forma canónica.

11. Reduzca las siguientes formas cuadráticas reales a la forma canónica y halle rango, índice positivo de inercia e índice negativo de inercia en cada caso.

a) $9x_1^2 + 12x_1x_2 + 79x_2^2$

b) $2x_1^2 - 12x_1x_2 + 18x_2^2$

c) $-2x_1^2 - 4x_1x_2 + 22x_2^2 + 12x_2x_3 + 6x_3x_1 - x_3^2$

12. Sea q la forma cuadrática definida sobre $\mathbb{R}^3$ mediante la relación

$$q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + tx_3^2 + 2x_1x_2 - 4x_2x_3.$$

a) Determine los valores de t para los cuales q es definida positiva.

b) Para $t=4$ y mediante la reducción de q a la forma normal, encuentre un subespacio V de $\mathbb{R}^3$, de dimensión máxima, tal que la restricción de q a V sea definida positiva.

c) Para $t=9$, encuentre una base (a_i) de $\mathbb{R}^3$ en que la matriz de q resulte diagonal.

d) Apoyándose en c) encuentre una matriz inversible P tal que se cumpla la relación

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & -2 \\ 0 & -2 & 9 \end{pmatrix} = {}^tPP.$$

13. Sea q la forma cuadrática definida sobre $\mathbb{R}^3$ mediante la relación

$$q(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4(x_1x_2 + x_1x_3 + x_2x_3).$$

a) Muestre que q no es definida positiva, reduciendo q a su forma normal.

b) Encuentre un subespacio V de $\mathbb{R}^3$, de dimensión máxima, tal que la restricción de q a V sea definida positiva.

14. Determine, si existen, los valores de los parámetros para los cuales las formas cuadráticas que se indican a continuación resultan definidas positivas.

a) $q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + 3x_3^2 + 2tx_1x_2 + 2x_1x_3$

b) $q(x_1, x_2, x_3, x_4) = x_1^2 - 3x_2^2 - 4x_3^2 + \alpha x_4^2 + 2\beta x_1x_2$

15. Sea A una matriz con coeficientes reales de orden n , simétrica y tal que todos sus menores principales son estrictamente mayores que cero. Demuestre que existe una matriz inversible P tal que $A = {}^tPP$.

16. Reduzca a la forma canónica:

a) $2x_1^2 + x_2^2 - 4x_1x_2 - 4x_2x_3$

b) $x_1^2 + 2x_2^2 + 3x_3^2 - 4x_1x_2 - 4x_2x_3$

c) $3x_1^2 + 4x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_2x_3$

d) $2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$

17. Diga cuáles de las formas cuadráticas del ejercicio anterior son definidas positivas y halle una base tal, que la expresión de q respecto a dicha base sea la forma normal.

18. Describa los lugares geométricos que corresponden a los posibles casos de forma canónica de formas cuadráticas reales en tres dimensiones.

19. Pruebe que $ax^2 + bxy + cy^2$ es definida positiva si y solo si $4ac - b^2 > 0$.

20. Una forma cuadrática se llama positiva cuando el índice positivo de inercia es igual al rango y ambos son menores que la dimensión del espacio. Enuncie y demuestre para esta clase de forma, teoremas análogos a los 1.3.1 y 1.3.2.

2. Espacios euclidianos y unitarios

2.1 Introducción

Al estudiar los espacios vectoriales partimos de conceptos geométricos bien conocidos y, en particular, de los espacios $\mathbb{R}^2$ y $\mathbb{R}^3$. Sin embargo, no generalizamos ciertos conceptos muy usados en Geometría analítica, como son: *ángulo entre dos vectores*, *longitud de un vector*, y otros.

Para tener una idea de cómo con la estructura vectorial solamente no podemos caracterizar los conceptos geométricos anteriormente citados, observemos que intuitivamente la longitud del vector $3x$ debe ser tres veces la longitud de un vector x dado; sin embargo, el hecho de que la aplicación en un espacio vectorial E dada por $x \rightarrow 3x$ sea un isomorfismo de espacios vectoriales, indicaría que los vectores x y $3x$ pueden identificarse desde un punto puramente algebraico.

Es necesario, pues, dotar a la estructura de espacio vectorial de nuevos elementos que nos permita considerar los conceptos geométricos que mencionamos al inicio.

Recordemos brevemente cómo estudiamos estas nociones en Geometría analítica. Por ejemplo, en $\mathbb{R}^2$ la longitud o norma de un vector $x = (x_1, x_2)$ estaba dada por $\|x\| = \sqrt{x_1^2 + x_2^2}$, y el ángulo θ entre los vectores $x = (x_1, x_2)$ y $y = (y_1, y_2)$ por

$$\cos \theta = \frac{x_1 y_1 + x_2 y_2}{\|x\| \|y\|}$$

donde a la expresión $x_1 y_1 + x_2 y_2$ la llamamos *producto escalar* de x y y , y la denotamos por $x \cdot y$.

En realidad, a partir del producto escalar en $\mathbb{R}^2$ pueden definirse los otros conceptos, pues conocido $x \cdot y \forall x, y \in \mathbb{R}^2$, la norma es

$$\|x\| = \sqrt{x \cdot x}$$

y el ángulo entre dos vectores se expresa a partir del producto escalar y la norma.

Es lógico pensar que si queremos generalizar a espacios vectoriales cualesquiera los conceptos *ángulo entre vectores*, *norma* o *longitud*, etc., baste generalizar el concepto *producto escalar*, por lo que precisaremos cuáles son las propiedades definitorias del producto escalar en $\mathbb{R}^2$, y a partir de estas propiedades generalizaremos el concepto a espacios vectoriales cualesquiera.

En primer lugar, el producto escalar puede considerarse como una aplicación de $\mathbb{R}^2 \times \mathbb{R}^2$ en $\mathbb{R}$, ya que a cada par de vectores de $\mathbb{R}^2$ le podemos asignar su producto escalar que es un número real. Además, sabemos que dados $x, y \in \mathbb{R}^2$, $\alpha \in \mathbb{R}$, se cumple:

- (1) $(x+y) \cdot z = x \cdot z + y \cdot z$
- (2) $(\alpha x) \cdot y = \alpha(x \cdot y)$
- (3) $x \cdot y = y \cdot x$
- (4) $x \cdot x \geq 0$
- (5) $x \cdot x = 0 \iff x = 0$

Hay otras propiedades, como $x \cdot (y+z) = x \cdot y + x \cdot z$ y $x \cdot (\alpha y) = \alpha(x \cdot y)$, pero estas pueden deducirse de las que hemos dado.

Como un primer intento de generalización, investiguemos si en $\mathbb{C}^2$ como espacio vectorial sobre $\mathbb{C}$ es válida la fórmula conocida para el producto escalar y tiene las mismas propiedades.

En $\mathbb{C}^2$, dados $x = (x_1, x_2)$, $y = (y_1, y_2)$, definamos $x \cdot y = x_1 y_1 + x_2 y_2$, expresión análoga a la conocida en $\mathbb{R}^2$. Entonces no habrá problemas con las propiedades 1, 2 y 3, pero en la 4 tendremos que

$$ix \cdot ix = i^2 x \cdot x = -x \cdot x,$$

lo que entraría en contradicción con nuestra interpretación intuitiva de la norma o longitud de un vector, la cual nos indica que debe ser una cantidad real positiva, pues si $\|x\| = \sqrt{x \cdot x}$ es positiva se tiene, sin embargo, que $\|ix\| = \sqrt{ix \cdot ix}$ es negativa. Esto se resolvería definiendo

$$x \cdot y = x_1 \bar{y}_1 + x_2 \bar{y}_2$$

De aquí

$$x \cdot y = y \cdot x,$$

lo que conduce a

$$x \cdot \alpha y = \alpha y \cdot x = \alpha x \cdot y.$$

De esta forma se obtiene

$$ix \cdot ix = i \bar{i} x \cdot x = |i|^2 x \cdot x = x \cdot x.$$

Las consideraciones anteriores nos sirven para dar una definición axiomática de producto escalar en un espacio vectorial real o complejo cualquiera. Nos limitaremos a espacios vectoriales reales o complejos, ya que el considerar espacios vectoriales sobre $\mathbb{Q}$ puede ocasionar problemas al tratar de

generalizar el concepto de norma, pues aparecería una raíz cuadrada que no siempre está definida para todo número racional.

2.2 Espacios vectoriales con producto escalar

En primer lugar vamos a definir lo que llamaremos *producto escalar* en un espacio vectorial E sobre $\mathbb{R}$ o sobre $\mathbb{C}$. Por comodidad denotaremos $\mathbb{R}$ y $\mathbb{C}$ por K , y aclararemos de cuál de los dos se trata cuando sea necesario.

DEFINICIÓN 2.2.1

Dado un espacio vectorial E sobre K , se dice que está definido un producto escalar en E si a cada par x, y de elementos de E se le puede asignar un elemento de K , que se denota por $\langle x, y \rangle$, de forma tal que para todo $x, y, z \in E$ y $\alpha \in K$ se cumple:

- (1) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- (2) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- (3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- (4) $\langle x, x \rangle \geq 0$
- (5) $\langle x, x \rangle = 0 \iff x = 0$

En un lenguaje más formal, el producto escalar no es más que una aplicación de $E \times E$ en K que cumple las propiedades (1) a (5).

La denominación *producto escalar* se utiliza por analogía con la terminología de la Física, pero también es muy usual el nombre de *producto interior* o *interno*.

Si $K = \mathbb{R}$, la propiedad (3) se convierte simplemente en $\langle x, y \rangle = \langle y, x \rangle$, que es la conocida propiedad de simetría del producto escalar en $\mathbb{R}^2$ estudiada anteriormente.

Utilizamos la notación $\langle x, y \rangle$ para denotar el producto escalar de x y y porque otras notaciones usuales como las de $x \cdot y$ o (x, y) pueden prestarse a confusión en determinados casos.

Si $K = \mathbb{C}$, $\langle x, y \rangle$ es un número complejo; sin embargo, $\langle x, x \rangle$ siempre es real pues por la propiedad (3) se tiene $\langle x, x \rangle = \overline{\langle x, x \rangle}$, de ahí que tenga sentido exigir al producto escalar la propiedad (4).

De la definición de producto escalar se deducen:

- a) $\langle 0, x \rangle = 0 \quad \forall x \in E$
- b) $\langle x, 0 \rangle = 0 \quad \forall x \in E$
- c) $\langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad \forall x, y, z \in E$
- d) $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle \quad \forall x, y \in E \quad \forall \alpha \in K$
- e) $\langle x, y \rangle = 0 \quad \forall y \in E \implies x = 0$
- f) $\langle x, y \rangle = 0 \quad \forall x \in E \implies y = 0$

Demostración

Demostraremos a, d y e, y dejamos al estudiante que demuestre de manera similar b, c y f.

- a) $\langle 0, x \rangle = \langle y-y, x \rangle = \langle y, x \rangle - \langle y, x \rangle = 0 \quad \forall x \in E$
- d) $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle = \overline{\alpha} \langle y, x \rangle = \overline{\alpha} \langle x, y \rangle \quad \forall x, y \in E \quad \forall \alpha \in K$
- e) $\langle x, y \rangle = 0 \quad \forall y \in E \implies \langle x, x \rangle = 0 \implies x = 0$ por la propiedad (5) de la definición de producto escalar.

Ejemplos

1. En $\mathbb{R}^2$ como espacio vectorial sobre $\mathbb{R}$, $\langle x, y \rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2$ donde $x = (x_1, x_2)$ y $y = (y_1, y_2)$ es un producto escalar.
En efecto, veamos que se cumplen las propiedades (1) a (5), ya que la fórmula dada asigna a cada x, y un número real.
Sean $z = (z_1, z_2)$ y $\alpha \in \mathbb{R}$.

$$\begin{aligned} (1) \quad \langle x+z, y \rangle &= (x_1+z_1)y_1 - (x_1+z_1)y_2 - (x_2+z_2)y_1 + 3(x_2+z_2)y_2 \\ &= (x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2) + (z_1 y_1 - z_1 y_2 - z_2 y_1 + 3z_2 y_2) \\ &= \langle x, y \rangle + \langle z, y \rangle \end{aligned}$$

$$(2) \quad \langle \alpha x, y \rangle = (\alpha x_1)y_1 - (\alpha x_1)y_2 - (\alpha x_2)y_1 + 3(\alpha x_2)y_2 = \alpha(x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2) = \alpha \langle x, y \rangle.$$

$$(3) \quad \langle x, y \rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2 = y_1 x_1 - y_1 x_2 - y_2 x_1 + 3y_2 x_2 = \langle y, x \rangle$$

$$(4) \quad \langle x, x \rangle = x_1^2 - 2x_1 x_2 + 3x_2^2 = (x_1 - x_2)^2 + 2x_2^2 \geq 0$$

$$(5) \quad \langle x, x \rangle = 0 \iff x_1 - x_2 = 0, \quad x_2 = 0 \iff x_1 = x_2 = 0 \iff x = 0$$

Con este ejemplo vemos que la definición dada de producto escalar no solo generaliza los conceptos conocidos en $\mathbb{R}^2$ y $\mathbb{R}^3$ a espacios más generales, sino que inclusive amplía el concepto *producto escalar* en el propio $\mathbb{R}^2$ (análogamente en $\mathbb{R}^3$), y que el producto escalar en $\mathbb{R}^2$ conocido: $x \cdot y = x_1 y_1 + x_2 y_2$, es un caso particular, aunque importante según veremos, de producto escalar, al cual nos referiremos como *producto escalar usual*.

2. En $\mathbb{C}^2$ como $\mathbb{C}$ espacio, dados $z = (z_1, z_2)$ y $w = (w_1, w_2)$, $\langle z, w \rangle = z_1 \overline{w_1} + (1+i)z_1 \overline{w_2} + (1-i)z_2 \overline{w_1} + 3z_2 \overline{w_2}$ es un producto escalar, lo cual puede verificar el estudiante siguiendo el esquema dado en el ejemplo 1 y observando que se trata de un espacio vectorial sobre $\mathbb{C}$.
3. En $\mathbb{R}^2$ como $\mathbb{R}$ espacio, dados $x = (x_1, x_2)$ y $y = (y_1, y_2)$, $\langle x, y \rangle = 3x_1 y_1 + x_1 y_2 + x_2 y_1 - 2x_2 y_2$ no es un producto escalar, ya que para $x = (0, 1)$ se cumple $\langle x, x \rangle = -2 < 0$.
4. En $\mathbb{C}^2$ como $\mathbb{C}$ espacio, dados $z = (z_1, z_2)$ y $w = (w_1, w_2)$, $\langle z, w \rangle = z_1 \overline{w_1} + iz_1 \overline{w_2} + iz_2 \overline{w_1} + z_2 \overline{w_2}$

no es un producto escalar, ya que no se cumple la propiedad (3):

$$\begin{aligned}\overline{(w, z)} &= \overline{w_1 z_1 + i w_1 z_2 + i w_2 z_1 + w_2 z_2} \\ &= \overline{z_1 w_1 - i z_2 w_1 - i z_1 w_2 + z_2 w_2} \\ &\neq (z, w).\end{aligned}$$

En $\mathbb{R}^n$ como espacio vectorial sobre $\mathbb{R}$, dados $x = (x_1, \dots, x_n)$ y $y = (y_1, \dots, y_n)$,

$$(x, y) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

es un producto escalar que llamaremos producto escalar usual en $\mathbb{R}^n$.

6. En $\mathbb{C}^n$ como espacio vectorial sobre $\mathbb{C}$, dados $z = (z_1, \dots, z_n)$ y $w = (w_1, \dots, w_n)$,

$$(z, w) = z_1 \overline{w_1} + \dots + z_n \overline{w_n}$$

es un producto escalar que llamaremos producto escalar usual en $\mathbb{C}^n$.

7. En el espacio $M_n(\mathbb{C})$ de las matrices cuadradas de orden n con coeficientes en $\mathbb{C}$ como espacio vectorial sobre $\mathbb{C}$, dadas $A, B \in M_n(\mathbb{C})$,

$$(A, B) = \text{tr } {}^t \overline{B} \cdot A,$$

donde ${}^t \overline{A}$ es la matriz transpuesta de la matriz cuyos coeficientes son los conjugados de los coeficientes de A , es un producto escalar ya que $\text{tr } {}^t \overline{B} A$ es un número complejo, y se cumple:

$$\begin{aligned}(1) \quad (A + C, B) &= \text{tr } {}^t \overline{B} (A + C) = \text{tr } ({}^t \overline{B} \cdot A + {}^t \overline{B} \cdot C) = \\ &= \text{tr } {}^t \overline{B} \cdot A + \text{tr } {}^t \overline{B} \cdot C = (A, B) + (C, B)\end{aligned}$$

$$(2) \quad (\alpha A, B) = \text{tr } {}^t \overline{B} (\alpha A) = \alpha \text{tr } ({}^t \overline{B} \cdot A) = \alpha \text{tr } ({}^t \overline{B} \cdot A) = \alpha (A, B)$$

$$(3) \quad (A, B) = \text{tr } {}^t \overline{B} \cdot A = \text{tr } ({}^t \overline{A} \cdot \overline{B}) = \text{tr } ({}^t \overline{A} \cdot \overline{B}) = \text{tr } {}^t \overline{A} \cdot B = \text{tr } {}^t \overline{A} \cdot B = (B, A)$$

$$(4) \quad (A, A) = \text{tr } {}^t \overline{A} \cdot A$$

Si $A = (a_{ij})$, entonces

$$\text{tr } {}^t \overline{A} \cdot A = \sum_{i=1}^n \overline{a_{ii}} a_{ii} = \sum_{i=1}^n |a_{ii}|^2 \geq 0.$$

(5) Sea de nuevo $A = (a_{ij})$.

$$\text{Si } \text{tr } {}^t \overline{A} \cdot A = 0 \Leftrightarrow \sum_{i=1}^n \overline{a_{ii}} a_{ii} = \sum_{i=1}^n |a_{ii}|^2 = 0, \text{ entonces } a_{ii} = 0 \forall i, j.$$

De aquí que $(A, A) = 0 \Rightarrow A = 0$.

El recíproco es evidente.

8. Sea $E = C[-\pi, \pi] = \{f: [-\pi, \pi] \rightarrow \mathbb{R} \text{ continuas}\}$, considerado como espacio vectorial real para la suma y el producto por un escalar de funciones. Si $f, g \in E$, entonces $(f, g) = \int_{-\pi}^{\pi} f(t) g(t) dt$ es un producto escalar.

9. Sea el espacio l_2 de las sucesiones $(a_n)_{n \in \mathbb{N}}$ de números reales tales que $\sum_{n=1}^{\infty} a_n^2$ es convergente, con la suma y el producto por un número real de sucesiones.

Entonces, dados $a = (a_n)_{n \in \mathbb{N}}$, $b = (b_n)_{n \in \mathbb{N}}$ en l_2 , $(a, b) = \sum_{n=1}^{\infty} a_n \cdot b_n$ es un producto escalar en l_2 .

Los ejemplos 8 y 9 deben ser verificados por el estudiante, para lo cual utilizará sus conocimientos de Análisis matemático.

Hemos definido el producto escalar para un espacio vectorial de dimensión finita o infinita, pero realmente en muchos resultados tendremos que limitarnos al caso de espacios vectoriales de dimensión finita, ya que para los espacios de dimensión no finita se requieren otras consideraciones de carácter no algebraico. No obstante, siempre que sea posible haremos referencia a los resultados en el caso general.

Los espacios vectoriales reales con un producto escalar se llaman *espacios euclidianos* y los espacios complejos con un producto escalar, *espacios unitarios*. Como en muchas ocasiones es posible obtener resultados para ambos tipos de espacios simultáneamente, nos referiremos a ellos como *espacios con producto escalar* y aclararemos si son euclidianos o unitarios solamente cuando sea necesario.

Representación matricial del producto escalar cuando E es de dimensión finita

En los ejemplos 1 y 2 de productos escalares se pone de manifiesto que en la definición de estos son fundamentales las coordenadas (x_i) y (y_i) de los vectores y ciertos números reales o complejos que acompañan a las combinaciones $x_i y_j$.

En el ejemplo sobre $\mathbb{R}^2$,

$$(x, y) = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3 x_2 y_2,$$

aparece de forma natural la matriz $\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}$ y no es difícil verificar que si llamamos

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, A = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix},$$

entonces

$$(x, y) = {}^t Y A X.$$

En el ejemplo 2, sobre $\mathbb{C}^2$,

$$(z, w) = z_1 \overline{w_1} + (1+i) z_1 \overline{w_2} + (1-i) z_2 \overline{w_1} + 3 z_2 \overline{w_2},$$

Llamando

$$Z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, W = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, A = \begin{pmatrix} 1 & 1+i \\ i-i & 3 \end{pmatrix}$$

se verifica también fácilmente que

$$\langle z, w \rangle = {}^t \bar{W} A Z$$

Ahora veremos que en el caso de un espacio con producto escalar de dimensión finita, si fijamos una base del espacio, siempre es posible encontrar una representación matricial para el producto escalar. Lo haremos en el caso general, esto es para $K = \mathbb{C}$, y los resultados serán aplicables para $K = \mathbb{R}$ con los ajustes correspondientes.

Supongamos que $(a_i)_{1 \leq i \leq n}$ es una base de E y sean $x = \sum_{i=1}^n x_i a_i$, $y = \sum_{j=1}^n y_j a_j$ elementos de E . Entonces

$$\langle x, y \rangle = \left\langle \sum_{i=1}^n x_i a_i, \sum_{j=1}^n y_j a_j \right\rangle = \sum_{i=1}^n x_i \langle a_i, y \rangle$$

$$= \sum_{i=1}^n x_i \langle a_i, \sum_{j=1}^n y_j a_j \rangle = \sum_{i,j=1}^n x_i y_j \langle a_i, a_j \rangle$$

Si llamamos $a_{ij} = \langle a_i, a_j \rangle$, podemos escribir

$$\langle x, y \rangle = \sum_{i,j=1}^n y_j \bar{a}_{ji} x_i$$

llamando

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, A = (a_{ij}),$$

entonces

$$\langle x, y \rangle = {}^t \bar{Y} A X.$$

A se llama *matriz del producto escalar en la base (a_i)* .

En el caso real (euclídeano) se obtiene $\langle x, y \rangle = {}^t Y A X$.

De las propiedades del producto escalar, específicamente de (3), (4) y (5) resulta que $A = {}^t \bar{A}$ y ${}^t \bar{X} A X > 0$ para $X \neq 0$. (Recordemos que una matriz tal que $A = {}^t \bar{A}$ se llama *hermítica*.)

En el caso real podemos precisar más, pues $A = {}^t A$ y ${}^t X A X > 0$ para todo $X \neq 0$, que según los resultados obtenidos en el capítulo 1 significa que la forma cuadrática asociada a la matriz simétrica A es definida positiva.

Recíprocamente, si $A \in M_n(\mathbb{C})$ tal que ${}^t \bar{A} = A$ y ${}^t \bar{X} A X > 0$ para $X \neq 0$ ($A \in M_n(\mathbb{R})$ tal que $A = {}^t A$ y ${}^t X A X > 0$ para $X \neq 0$), puede comprobarse que si fijamos una base (a_i) de E y X, Y son los vectores de coordenadas de x, y de E en la base (a_i) , entonces $\langle x, y \rangle = {}^t \bar{Y} A X$ define un producto escalar en E tal que su matriz en la base (a_i) es A (en el caso real $\langle x, y \rangle = {}^t Y A X$).

Esta interpretación matricial es muy adecuada en el tratamiento de algunos problemas. En particular, para comprobar si en un espacio vectorial de dimensión finita una expresión dada es un producto escalar o no, basta comprobar si es de la forma $\langle x, y \rangle = {}^t \bar{Y} A X$ para cierta A y si A tiene las propiedades analizadas anteriormente.

En los ejemplos 1 y 2 hubiera bastado comprobar que

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix} \text{ es simétrica definida positiva en el caso real,}$$

$$B = \begin{pmatrix} 1 & 1+i \\ 1-i & 3 \end{pmatrix} \text{ es hermítica y } {}^t \bar{X} B X > 0 \text{ para } X \neq 0, \text{ siendo } {}^t \bar{X} B X = 0 \text{ si y solo si } X = 0, \text{ respectivamente.}$$

Ejercicios

1. ¿Para qué valores de los parámetros correspondientes, son productos escalares en $\mathbb{R}^2$ o $\mathbb{C}^2$, dados $x = (x_1, x_2)$ y $y = (y_1, y_2)$?
 - a) $\langle x, y \rangle = x_1 y_1 - 3x_1 y_2 - 3x_2 y_1 + kx_2 y_2$, $k \in \mathbb{R}$
 - b) $\langle x, y \rangle = ax_1 y_1 + bx_1 y_2 + cx_2 y_1 + dx_2 y_2$, $a, b, c, d \in \mathbb{R}$
 - c) $\langle x, y \rangle = ax_1 y_1 + bx_1 y_2 + cx_2 y_1 + dx_2 y_2$, $a, b, c, d \in \mathbb{C}$
2. Sea E un espacio con producto escalar. Pruebe que si $\langle x, y \rangle = 0$ para todo $x \in E$, entonces $y = 0$.
3. Sea E un espacio vectorial sobre K . Pruebe que la suma de dos productos escalares definidos en E es un producto escalar en E . (Considere el producto escalar como una aplicación de $E \times E$ en K , y que la suma y otras operaciones se refieren a las correspondientes operaciones entre aplicaciones.)

¿Qué ocurrirá con la diferencia y el producto por un elemento $p \in K$?
4. Describa explícitamente todos los productos escalares en los espacios vectoriales $\mathbb{R}$ y $\mathbb{C}$.
5. a) Considere $\mathbb{R}^2$ con el producto escalar usual y $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $T(x_1, x_2) = (-x_2, x_1)$. Verifique que $\langle x, T(x) \rangle = 0$ para todo $x \in \mathbb{R}^2$ y encuentre todos los productos escalares en $\mathbb{R}^2$ tales que $\langle x, T(x) \rangle = 0$ para todo $x \in \mathbb{R}^2$.

b) Si en lugar de $\mathbb{R}^2$ consideramos $\mathbb{C}^2$ con el producto escalar usual, ¿existirá una aplicación lineal de $\mathbb{C}^2$ en $\mathbb{C}^2$ tal que $\langle x, T(x) \rangle = 0$ para todo $x \in \mathbb{C}^2$?

6. Sean E y F espacios vectoriales sobre F y suponga que en F hay definido un producto escalar. Si $T: E \rightarrow F$ es una aplicación lineal inversible, pruebe que $\langle x, y \rangle = \langle T(x), T(y) \rangle$ define un producto escalar en E . Aplique el resultado para obtener un producto escalar en $C[-\pi, \pi]$ a partir del producto escalar $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$ y la aplicación $T: C[-\pi, \pi] \rightarrow C[-\pi, \pi]$ dada por

$$T(f(t)) = if(t).$$

7. Pruebe que en un espacio unitario E (espacio vectorial complejo con producto escalar), $\langle x, y \rangle = \operatorname{Re} \langle x, y \rangle + i \operatorname{Re} \langle x, iy \rangle$ para todo $x, y \in E$.

8. Sean $A \in M_n(\mathbb{R})$; $X, Y \in M_{(2,1)}(\mathbb{R})$; $A = (a_{ij})$ y $\langle X, Y \rangle = YAX$. Pruebe que $\langle X, Y \rangle$ es un producto escalar en $M_{(2,1)}(\mathbb{R})$ si y solo si $A = {}^t A$, $a_{11} > 0$, $a_{22} > 0$ y $\det A > 0$.

9. a) Pruebe que en el espacio vectorial de los polinomios con coeficientes reales $\mathbb{R}[X]$, la fórmula

$$\left\langle \sum_{j=1}^n a_j x^j, \sum_{k=1}^m b_k x^k \right\rangle = \sum_{j,k} \frac{a_j b_k}{j+k+1}$$

define un producto escalar.

b) Considere $W = \{f(x) \in \mathbb{R}[X] : \operatorname{gr} f(x) \leq n\}$ con el producto escalar definido en a) y halle la matriz de dicho producto escalar con respecto a la base $1, x, x^2, \dots, x^n$ de W .

10. Sea E un espacio vectorial sobre K con producto escalar n -dimensional y sea $(a_i)_{1 \leq i \leq n}$ una base de E . Si $k_1, k_2, \dots, k_n$ son n elementos de K pruebe que existe exactamente un vector x en E tal que $\langle x, a_i \rangle = k_i$, $i=1, \dots, n$.

2.3 Norma o longitud de un vector en un espacio con producto escalar

Sobre la base de nuestros conocimientos geométricos en $\mathbb{R}^2$ y $\mathbb{R}^3$, podemos definir a partir del producto escalar el concepto *norma* o *longitud* de un vector.

DEFINICIÓN 2.3.1

Dado un espacio con producto escalar E , se llama *norma* o *longitud* de $x \in E$ y se denota por $\|x\|$ a

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Esta norma está perfectamente definida para todo x de E , ya que como sabemos $\langle x, x \rangle \geq 0$ para todo x .

La norma está asociada a un producto escalar. Por ejemplo, si en $\mathbb{R}^2$ consideramos el producto escalar usual $x \cdot y = x_1 y_1 + x_2 y_2$ y el producto $\langle x, y \rangle = x_1 y_1 - x_2 y_2 - x_3 y_1 + 3x_2 y_2$, tendremos que $\|(1, 2)\| = \sqrt{1+4} = \sqrt{5}$ con respecto al primer producto escalar, mientras que $\|(1, 2)\| = \sqrt{1-2+2+12} = 3$ con respecto al segundo.

Las propiedades de la norma coinciden con nuestra idea intuitiva de *longitud*. Veamos las más simples.

$$(1) \|x\| = 0 \iff x = 0,$$

significa que el único vector de longitud nula es el nulo.

(2) $\|ax\| = \sqrt{\langle ax, ax \rangle} = \sqrt{|a|^2 \langle x, x \rangle} = |a| \|x\|$, es decir, la longitud del vector x aparece multiplicada por $|a|$ al multiplicar el vector x por el escalar a .

Otras propiedades de la norma son las siguientes:

(3) *Desigualdad de Cauchy-Buniakovski-Schwarz*. Esta propiedad es importante, ya que relaciona la norma con el producto escalar y se aplica frecuentemente. Se enuncia así:

PROPOSICIÓN 2.3.1

En un espacio E con producto escalar,

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \text{ para todo } x, y \in E,$$

y la igualdad se cumple si y solo si x y y son linealmente dependientes.

Demostración

Sea $\lambda \in K$, entonces

$\langle x - \lambda y, x - \lambda y \rangle \geq 0$ por la propiedad (4) del producto escalar.

$$\begin{aligned} \langle x - \lambda y, x - \lambda y \rangle &= \langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda \bar{\lambda} \langle y, y \rangle \\ &= \langle x, x \rangle - \bar{\lambda} \langle x, y \rangle - \lambda \langle x, y \rangle + \lambda \bar{\lambda} \langle y, y \rangle \geq 0. \end{aligned}$$

Si $y=0$, la desigualdad que queremos demostrar se cumple de forma trivial; así que supongamos $y \neq 0$.

Si tomamos $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle}$, tenemos

$$\langle x, x \rangle - \frac{\bar{\langle x, y \rangle}}{\langle y, y \rangle} \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle x, y \rangle + \frac{\langle x, y \rangle \bar{\langle x, y \rangle}}{\langle y, y \rangle \langle y, y \rangle} \langle y, y \rangle \geq 0,$$

de donde

$$\langle x, x \rangle \langle y, y \rangle - \langle x, y \rangle \langle x, y \rangle - \langle x, y \rangle \langle x, y \rangle + \langle x, y \rangle \langle x, y \rangle \geq 0 \text{ y}$$

$$\|x\|^2 \cdot \|y\|^2 \geq |\langle x, y \rangle|^2,$$

o sea,

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

según queríamos demostrar.

En cuanto a la segunda parte de la proposición, podemos demostrarla como sigue.

Si x, y son linealmente dependientes, entonces $y = kx$.

$$\langle x, y \rangle = \langle x, kx \rangle = \bar{k} \langle x, x \rangle$$

$$|\langle x, y \rangle| = |\langle x, kx \rangle| = |\bar{k} \langle x, x \rangle| = |k| \cdot \|x\|^2 = |k| \cdot \|x\| \cdot \|x\|$$

$$= \|x\| \cdot \|y\|$$

Recíprocamente, probemos que si se cumple la igualdad, x y y son linealmente dependientes.

Supongamos lo contrario, o sea, que x, y son linealmente independientes. Entonces $x - \lambda y \neq 0$ para todo λ y tendremos $\langle x - \lambda y, x - \lambda y \rangle > 0$, lo que conduce a $|\langle x, y \rangle| < \|x\| \cdot \|y\|$, en contradicción con el hecho de que se cumple la igualdad.

Existen diversas demostraciones de la desigualdad anterior, pero todas son más o menos artificiosas. En la que hemos escogido se usa una técnica muy usual en los espacios euclidianos: el considerar el producto $\langle x - \lambda y, x - \lambda y \rangle$ y la propiedad (4), que volverá a encontrar el estudiante en propiedades y ejercicios posteriores.

Si el estudiante aplica la desigualdad de Cauchy-Buniakovski-Schwarz a los ejemplos de producto escalar estudiados, encontrará algunas desigualdades conocidas.

(4) $\|x + y\| \leq \|x\| + \|y\|$ para todo x, y del espacio con producto escalar E .

Esta propiedad se conoce como *desigualdad triangular*, ya que expresa la siguiente propiedad geométrica:

La longitud de la suma de dos vectores es menor que la suma de las longitudes de estos (figura 2.1).

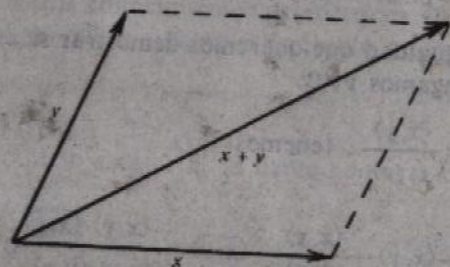


Fig. 2.1

Para demostrar esta propiedad aplicaremos precisamente la desigualdad de Cauchy-Buniakovski-Schwarz y una técnica similar a la usada en su demostración.

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle \bar{x}, y \rangle + \langle y, y \rangle$$

$$= \langle x, x \rangle + 2\operatorname{Re}\langle x, y \rangle + \langle y, y \rangle$$

considerando $\langle x, y \rangle = \operatorname{Re}\langle x, y \rangle + i \operatorname{Im}\langle x, y \rangle$.

Como

$$\operatorname{Re}\langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\| \cdot \|y\|,$$

tenemos

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

De aquí

$$\|x + y\| \leq \|x\| + \|y\|.$$

Aunque esta demostración es válida para $\mathbb{R}$ y $\mathbb{C}$, recomendamos al estudiante que realice la demostración en el caso real.

Una consecuencia importante de la desigualdad de Cauchy-Buniakovski-Schwarz es la posibilidad de definir en un espacio real con producto escalar el ángulo entre dos vectores no nulos. En efecto

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \Leftrightarrow \frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \leq 1, \text{ de donde } -1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1,$$

y podemos interpretar $\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$ como el coseno de un ángulo del intervalo $[0, \pi]$.

Esta fórmula coincide con la conocida en el caso de $\mathbb{R}^2$ y $\mathbb{R}^3$ en la Geometría analítica y puede demostrarse que si definimos el ángulo θ entre x y y por el coseno de dicho ángulo dado por $\cos \theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$, se satis-

facen una serie de propiedades conocidas de los ángulos. No trataremos esto en detalle, pues solamente nos interesa el caso particular en que el ángulo es $\frac{\pi}{2}$, que estudiaremos en el próximo epígrafe.

En el caso del espacio con producto escalar complejo, no tiene sentido la consideración anterior, ya que $\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$ es un número complejo.

A partir de la norma puede definirse el concepto *distancia* entre dos vectores, el cual generaliza los conceptos de distancia en el plano y en el espacio.

DEFINICIÓN 2.3.2

Dado un espacio E con producto escalar y $x, y \in E$, se llama *distancia* entre x y y y se denota por $d(x, y)$ a

$$d(x, y) = \|x - y\|.$$

Es fácil verificar que la distancia así definida cumple las siguientes propiedades, que corresponden a la idea intuitiva de distancia que conocemos.

- (1) $d(x, x) = 0$, $d(x, y) \neq 0$, si $x \neq y$
- (2) $d(x, y) = d(y, x) \forall x, y \in E$
- (3) $d(x, y) + d(y, z) \geq d(x, z) \forall x, y, z \in E$

Por último, queremos recordar que en todo espacio con producto escalar es posible definir la norma o longitud de un vector, pero pudiéramos plantearnos el problema inverso, esto es:

Si en un espacio vectorial E sobre K tenemos una función de E en K que a $x \in E$ le asocie un elemento de K , que designamos por $\|x\|$ y llamamos *norma* de x , tal que

- $$\left. \begin{array}{l} (n1) \|x\| \geq 0 \text{ y } \|x\| = 0 \Leftrightarrow x = 0 \\ (n2) \|ax\| = |a| \|x\| \\ (n3) \|x+y\| \leq \|x\| + \|y\| \end{array} \right\} \forall x, y \in E, \forall a \in K,$$

que son las propiedades básicas que caracterizan el carácter geométrico del concepto *longitud de un vector*, entonces ¿podrá definirse en E un producto interior tal que $\|x\| = \sqrt{\langle x, x \rangle}$? La respuesta no es en general afirmativa, como veremos en el ejemplo siguiente:

Ejemplo

En $\mathbb{R}^n$ definamos $\|x\| = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$.

Es fácil comprobar que cualquiera sea p , se cumplen los axiomas (n1), (n2) y (n3) mencionados anteriormente; sin embargo, dicha norma no cumple la identidad

$$\|x-y\|^2 + \|x+y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

que cumplen las normas definidas a partir del producto escalar para $p \neq 2$. Por tanto, podemos afirmar que no existe producto escalar en $\mathbb{R}^n$ tal que

$$\|x\| = \sqrt{\langle x, x \rangle} \text{ para } p \neq 2.$$

Ejercicios

1. a) Encuentre la norma de $(6, -7) \in \mathbb{R}^2$ con respecto al producto escalar usual y al producto $\langle x, y \rangle = x_1 y_1 - x_2 y_2 - x_3 y_3 + 3x_4 y_4$.
- b) Con respecto a los productos escalares dados en a), halle vectores unitarios (de norma 1) a partir de $(1, -1)$, $(2, 3)$ y $(-3, 4)$.

- c) Encuentre la norma de $(1 - 2i, 2 + 3i) \in \mathbb{C}^2$ con respecto al producto escalar usual y al producto $\langle x, y \rangle = x_1 y_1 + (1+i)x_1 y_2 + (1-i)x_2 y_1 + 3x_2 y_2$.
- d) Con respecto a los productos dados en c), halle vectores unitarios a partir de $(1, 1+i)$, $(1+i, i-i)$ y $(7i, 2)$.

2. Encuentre el coseno del ángulo comprendido entre x, y , si

- a) $x = (1, -3, 2)$, $y = (2, 1, 5)$ en $\mathbb{R}^3$ con el producto escalar usual.

$$\text{b) } x = 2t - 1, y = t^2 \text{ en } C[-\pi, \pi] \text{ con } \langle f, g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$$

$$\text{c) } X = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix} Y = \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix} \text{ en } M_2(\mathbb{R}) \text{ con } \langle X, Y \rangle = \text{tr } Y^T X$$

3. Si E es un espacio con producto escalar, pruebe:

- a) $\|x \pm y\|^2 = \|x\|^2 \pm 2\text{Re} \langle x, y \rangle + \|y\|^2 \forall x, y \in E$
- b) Si $K = \mathbb{R}$

$$\langle x, y \rangle = \frac{1}{4} \|x+y\|^2 - \frac{1}{4} \|x-y\|^2 \forall x, y \in E$$

- c) Si $K = \mathbb{C}$

$$\langle x, y \rangle = \frac{1}{4} \|x+y\|^2 - \frac{i}{4} \|x-y\|^2 + \frac{1}{4} \|x+iy\|^2 - \frac{i}{4} \|x-iy\|^2$$

$$\text{d) } \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \forall x, y \in E$$

4. Sea E un espacio vectorial sobre K y considere la definición de norma como una aplicación de E en K tal que a todo $x \in E$ le asigna $\|x\| \in K$ y que cumple las propiedades (n1), (n2) y (n3).

- a) Dado $p \in \mathbb{N}$, $p > 0$, diga bajo qué condiciones de p , siendo $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\|x\| = (|x_1|^p + |x_2|^p)^{1/p}$$

define una norma en $\mathbb{R}^2$.

- b) Pruebe que $\|x\| = \max\{|x_1|, |x_2|\}$ define una norma en $\mathbb{R}^2$. ¿A cuál de las normas definidas en a) y b) corresponde un producto escalar en $\mathbb{R}^2$ tal que $\|x\|^2 = \langle x, x \rangle$ para todo x de $\mathbb{R}^2$?

5. a) Pruebe que una condición necesaria y suficiente para que en un espacio vectorial real E con una norma, según definimos en el ejercicio 2, exista un producto escalar que satisfaga $\|x\|^2 = \langle x, x \rangle$ para todo x de E , es

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \forall x, y \in E.$$

- b) Analice la afirmación hecha en a) para un espacio complejo provisto de una norma.

- c) Pruebe que una condición necesaria y suficiente impuesta a una norma en $\mathbb{R}^2$ para que exista un producto interior que satisfaga $\|x\|^2 = \langle x, x \rangle$ para todo $x \in \mathbb{R}^2$, es que el lugar geométrico de la ecuación $\|x\|=1$ sea una elipse.

2.4 Ortogonalidad en un espacio con producto escalar

Un concepto de gran importancia y aplicabilidad es el de ortogonalidad entre vectores en un espacio con producto escalar, el cual aparece de forma natural en la fórmula del ángulo θ entre dos vectores:

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$$

Daremos, pues, la definición siguiente:

DEFINICIÓN 2.4.1

Dado un espacio E con producto escalar, se dice que los vectores x, y de E son *ortogonales* si y solo si $\langle x, y \rangle = 0$.

Según la definición anterior, se puede afirmar que x y y son ortogonales sin tener que especificar un orden, esto es: x es ortogonal a y o y es ortogonal a x , ya que $\langle x, y \rangle = 0 \iff \langle y, x \rangle = 0$ por la propiedad (3) del producto escalar.

Una confirmación de cómo la definición de ortogonalidad entre vectores coincide con las ideas geométricas intuitivas, nos la da el hecho de que el 0 es el único vector que es ortogonal a todo vector del espacio, ya que $\langle 0, x \rangle = 0 \quad \forall x \in E$ y si $\langle x, y \rangle = 0 \quad \forall x \in E$, entonces $y = 0$.

Aunque no introdujimos el concepto *ángulo entre vectores* en el caso de los espacios con producto escalar complejo, es evidente que no hay dificultad en considerar en este caso la ortogonalidad entre vectores.

En los espacios con producto escalar desempeñan una función importante los sistemas de vectores ortogonales dos a dos. Precisemos esto en la siguiente definición:

DEFINICIÓN 2.4.2

Se dice que un sistema de vectores S del espacio con producto escalar E es *ortogonal* si y solo si $\langle x, y \rangle = 0$ para todo $x \neq y$ de S .

Asumiremos que un sistema formado por un solo vector es ortogonal. En particular, analizaremos los sistemas de vectores ortogonales en los cuales todos los vectores tengan norma 1.

DEFINICIÓN 2.4.3

Se dice que un sistema de vectores S del espacio con producto escalar E es *ortonormal*, si y solo si S es un sistema ortogonal y $\|x\|=1$ para todo x de S .

Los vectores de norma 1 se llaman *vectores unitarios*. A partir de todo vector no nulo x , podemos encontrar un múltiplo unitario de x : el vector

$$\frac{x}{\|x\|}$$

Aunque en este texto nos limitaremos a trabajar, en general, con sistemas finitos de vectores ortogonales y ortonormales, ya que el tratamiento del caso infinito, se aparta del objetivo puramente algebraico, hemos dado las definiciones de sistemas ortogonales y ortonormales en el caso de un sistema de vectores no necesariamente finito, y siempre que sea posible estudiaremos los resultados en el caso general.

Ejemplos

1. En $\mathbb{R}^2$ con el producto escalar usual la base canónica $e_1 = (1, 0)$, $e_2 = (0, 1)$ es un sistema ortonormal; sin embargo, no es ni siquiera ortogonal en el espacio formado por $\mathbb{R}^2$ con el producto escalar $\langle x, y \rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + 3x_2 y_2$, ya que $\langle e_1, e_2 \rangle = -1 \neq 0$.
2. La base canónica $(e_i)_{1 \leq i \leq n}$ donde $e_i = (0, \dots, \underset{i}{1}, \dots, 0)$ es un sistema ortonormal en $\mathbb{R}^n$ y $\mathbb{C}^n$ con sus respectivos productos escalares usuales.
3. En $M_n(\mathbb{C})$ con el producto escalar dado por $\langle A, B \rangle = \text{tr } \bar{B} \cdot A$. El sistema de matrices $(E_{ij})_{1 \leq i, j \leq n}$, donde E_{ij} es la matriz tal que todos sus coeficientes son nulos excepto el i, j que es 1, es un sistema ortonormal.
4. En $C[-\pi, \pi]$ con el producto escalar $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t) \overline{g(t)} dt$, el sistema de vectores $\{1, \cos t, \cos 2t, \dots, \sin t, \sin 2t, \dots\}$ es un sistema ortogonal. La ortogonalidad del sistema del ejemplo 4 desempeña un importante papel en Análisis. Observe el estudiante que para comprobar dicha ortogonalidad debe probar que las integrales del tipo $\int_{-\pi}^{\pi} \cos nt \cos mt dt$, $\int_{-\pi}^{\pi} \cos nt \sin mt dt$ y $\int_{-\pi}^{\pi} \sin nt \sin mt dt$ son nulas, lo cual le dejemos como ejercicio.
5. En el espacio l_2 de las sucesiones $(a_n)_{n \in \mathbb{N}}$ de números reales tales que $\sum_{n=1}^{\infty} a_n^2$ es convergente, con el producto escalar $\langle a, b \rangle = \sum_{n=1}^{\infty} a_n \cdot b_n$, el sistema $\{e_k\}_{k \in \mathbb{N}}$ donde $e_k = (0, 0, \dots, \underset{k}{1}, 0, \dots)$ es un sistema ortonormal de vectores.

Seguramente al estudiante le habrá llamado la atención el hecho de que los ejemplos dados de sistemas ortogonales u ortonormales, eran conocidos por él como ejemplos de sistemas linealmente independientes de los correspondientes espacios. Realmente existe relación entre los conceptos de orto-

gonalidad y de independencia lineal en un espacio con producto escalar, exactamente:

PROPOSICIÓN 2.4.1

En un espacio con producto escalar E , todo sistema ortogonal de vectores no nulos S es linealmente independiente.

Demostración

Sea S un sistema ortogonal de vectores no nulos de E . Probemos que si $x_1, x_2, \dots, x_n$ son vectores cualesquiera de S , entonces

$$\sum_{i=1}^n \alpha_i x_i = 0 \Rightarrow \alpha_i = 0, \quad i=1, \dots, n.$$

(Recordemos que la independencia lineal de un sistema no necesariamente finito de vectores de E se demuestra como acabamos de plantear.) Pero

$$\sum_{i=1}^n \alpha_i x_i = 0 \Rightarrow \left\langle \sum_{i=1}^n \alpha_i x_i, x_k \right\rangle = 0 \Leftrightarrow \sum_{i=1}^n \alpha_i \langle x_i, x_k \rangle = \alpha_k \langle x_k, x_k \rangle = 0,$$

y como los x_i son no nulos, $\langle x_k, x_k \rangle \neq 0$, de donde $\alpha_k = 0$.

Lo anterior podemos repetirlo para todo α_k , $k=1, \dots, n$, con lo que queda probado que S es linealmente independiente.

El recurso de multiplicar escalarmente por un vector ortogonal utilizado en esta demostración, es típico en el trabajo con los sistemas ortogonales, como comprobará el estudiante al encontrarlo frecuentemente.

De la proposición 2.4.1 se infiere inmediatamente que en un espacio vectorial con producto escalar, el número máximo de vectores mutuamente ortogonales no puede superar a la dimensión del espacio, lo cual coincide con nuestra intuición geométrica, ya que si interpretamos que los vectores mutuamente ortogonales dan las posibles direcciones perpendiculares en el plano y en el espacio, sabemos que estas no exceden a dos y tres respectivamente, esto es, las dimensiones de $\mathbb{R}^2$ y $\mathbb{R}^3$.

En la demostración de la proposición 2.4.1 también aparece otro hecho interesante: al suponer una combinación lineal de vectores de un sistema

ortogonal de vectores no nulos S , por ejemplo $\sum_{i=1}^n \alpha_i x_i$ con $x_i \in S$, $\alpha_i \in K$, y calcular convenientemente, encontramos que si $x = \sum_{i=1}^n \alpha_i x_i$, entonces

$$x = \sum_{i=1}^n \alpha_i x_i \Rightarrow \langle x, x_k \rangle = \sum_{i=1}^n \alpha_i \langle x_i, x_k \rangle = \alpha_k \langle x_k, x_k \rangle,$$

o sea,

$$\alpha_k = \frac{\langle x, x_k \rangle}{\|x_k\|^2},$$

y esto puede hacerse para todo k .

Si en particular $\|x_k\|=1$, entonces $\alpha_k = \langle x, x_k \rangle$, o sea, que si S es ortonormal, $\alpha_k = \langle x, x_k \rangle$ para todo k . Este es un resultado importante, pues indica que en un espacio vectorial con producto escalar se tiene una fórmula general para calcular los coeficientes de la expresión de un vector como combinación lineal de un sistema ortogonal y ortonormal; de aquí que pasemos a estudiar dentro de los espacios con producto escalar, las bases que sean a la vez sistemas ortogonales y ortonormales, precisando si es que estas existen y estudiando qué resultados podemos obtener a partir de su existencia.

Estudiaremos la existencia de base ortonormal solamente para el caso de espacios vectoriales con producto escalar de dimensión finita, ya que el estudio en dimensión no finita requiere de consideraciones no solo de tipo algebraico. No obstante, plantearemos algunos resultados relacionados con este problema que el estudiante puede encontrar en forma más general en cualquier texto de Análisis funcional.

DEFINICIÓN 2.4.4

Se llama *base ortogonal* (ortonormal) en un espacio E con producto escalar de dimensión finita, a una base de E que sea a la vez un sistema ortogonal (ortonormal) de dicho espacio.

En los ejemplos de sistemas ortogonales y ortonormales hemos encontrado algunos que son bases, pero el problema consiste en establecer si en cualquier espacio vectorial con producto escalar de dimensión finita existe una base ortogonal, ya que una vez encontrada una base ortogonal, multiplicando cada vector de esta por el inverso de su norma obtenemos una base ortonormal. La respuesta a este problema es afirmativa y se demuestra en el siguiente teorema:

TEOREMA 2.4.1

En todo espacio con producto escalar de dimensión finita E existe una base ortonormal.

Demostración

Como siempre es posible afirmar la existencia de un sistema ortogonal de vectores no nulos, ya que el sistema formado por un vector no nulo es ortogonal, probaremos que partiendo de un sistema ortogonal del espacio E puede encontrarse una base ortogonal de E . (Ya aclaramos anteriormente que basta demostrar la existencia de una base ortogonal, pues a partir de esta se encuentra una ortonormal.)

Sea, pues, $\dim E = n$ y $\{x_1, x_2, \dots, x_m\}$ un sistema de vectores ortogonales no nulos de E . Sabemos también que $m \leq n$.

Si $m = n$, entonces como $\{x_1, x_2, \dots, x_n\}$ es linealmente independiente, sería una base de E y quedaría probada nuestra afirmación.

Sea pues $m < n$. Intuitivamente nos damos cuenta que el problema consiste en demostrar la posibilidad de añadir elementos a $\{x_1, x_2, \dots, x_m\}$ de manera que los nuevos sistemas de vectores continúen siendo ortogonales y de esta forma encontrar una base de E .

Si tomamos un vector y cualquiera de E , entonces es fácil verificar que $x = y - \sum_{k=1}^n \frac{\langle y, x_k \rangle}{\langle x_k, x_k \rangle} x_k$ es ortogonal a todos los x_i , $i=1, \dots, n$. En efecto,

$$\langle x, x_i \rangle = \langle y, x_i \rangle - \sum_{k=1}^n \frac{\langle y, x_k \rangle \langle x_k, x_i \rangle}{\langle x_k, x_k \rangle} = \langle y, x_i \rangle - \langle y, x_i \rangle = 0.$$

Como el número de vectores de un sistema ortogonal está acotado por la dimensión del espacio, podemos afirmar que en un número finito de pasos llegaremos a un sistema ortogonal que no puede estar contenido en forma estricta en otro sistema ortogonal. Si demostramos que este sistema genera todo el espacio, será una base de E y, por tanto, tendrá n elementos.

(Observemos que en nuestro razonamiento aún no hemos probado que completando el sistema ortogonal llegamos a una base; de ahí que debemos probar que al agotar la posibilidad de seguir aumentando el sistema ortogonal, obtenemos un sistema generador.)

Tenemos, pues, un sistema ortogonal de vectores $\{x_1, x_2, \dots, x_r\}$ de E tal que no está incluido estrictamente en ningún otro sistema ortogonal de vectores de E , y debemos probar que $\{x_1, x_2, \dots, x_r\}$ genera el espacio E .

Sea $z \in E$. Según vimos,

$$u = z - \sum_{i=1}^r \frac{\langle z, x_i \rangle}{\langle x_i, x_i \rangle} x_i$$

es ortogonal a $\{x_1, x_2, \dots, x_r\}$. Si $u \neq 0$ entraríamos en contradicción con el hecho de que $\{x_1, x_2, \dots, x_r\}$ no puede ser ampliado a un conjunto ortogonal que lo contenga estrictamente; luego $u=0$ y $z = \sum_{i=1}^r \frac{\langle z, x_i \rangle}{\langle x_i, x_i \rangle} x_i$, o sea, z pertenece al espacio generado por $\{x_1, x_2, \dots, x_r\}$.

Proceso de ortogonalización (ortonormalización) de Gram-Schmidt

En la demostración anterior aparece de forma implícita otra vía para resolver el problema de la existencia de base ortonormal, la cual tiene gran interés práctico ya que ofrece un método de construcción de una base ortonormal en un espacio con producto escalar. Este método con el cual se construye un sistema ortogonal a partir de un sistema linealmente independiente, se conoce con el nombre de *proceso de ortogonalización de Gram-Schmidt* y se basa en un hecho geométrico bien conocido. Trataremos de dar una idea intuitiva de este proceso.

Supongamos que en $\mathbb{R}^2$ tenemos dos vectores cualesquiera b_1, b_2 linealmente independientes. A partir de $\{b_1, b_2\}$ podemos hallar geoméricamente un sistema ortogonal mediante los vectores a_1, a_2 , donde $a_1 = b_1$ y $a_2 = b_2 - b'_1$ (figura 2.2).

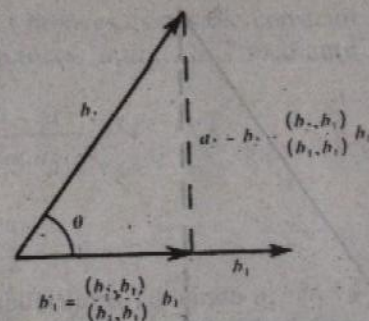


Fig. 2.2

Ahora bien, como b'_1 es un vector en la dirección de b_1 , solamente tenemos que hallar su longitud.

Sabemos que

$$\cos \theta = \frac{\langle b_2, b_1 \rangle}{\|b_2\| \cdot \|b_1\|} \text{ y } \|b'_1\| = |\cos \theta| \|b_2\|.$$

Luego

$$b'_1 = \left[\frac{\langle b_2, b_1 \rangle}{\|b_1\|^2} \right] b_1$$

pues debemos tomar el vector unitario en la dirección de b_1 . De aquí que

$$b'_1 = \frac{\langle b_2, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1$$

y el sistema ortogonal construido a partir de $\{b_1, b_2\}$ sea

$$a_1 = b_1$$

$$a_2 = b_2 - \frac{\langle b_2, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1.$$

Hemos trabajado de manera muy intuitiva, pero una vez encontrado el sistema $\{a_1, a_2\}$ es muy fácil verificar que es ortogonal lo cual dejamos al estudiante.

Observamos que el vector b_1 es lo que conocíamos en Geometría analítica por *proyección ortogonal* de b_2 sobre la recta determinada por b_1 .

Supongamos ahora que en $\mathbb{R}^3$ tenemos tres vectores linealmente independientes, b_1, b_2, b_3 , y queremos construir a partir de ellos un sistema ortogonal. Si tomamos $a_1 = b_1$ y $a_2 = b_2 - \frac{\langle b_2, a_1 \rangle}{\langle a_1, a_1 \rangle} a_1$ obtenemos un sistema ortogonal $\{a_1, a_2\}$, luego faltaría solamente construir el tercer vector ortogonal. Veamos gráficamente cuál es la situación (figura 2.3).

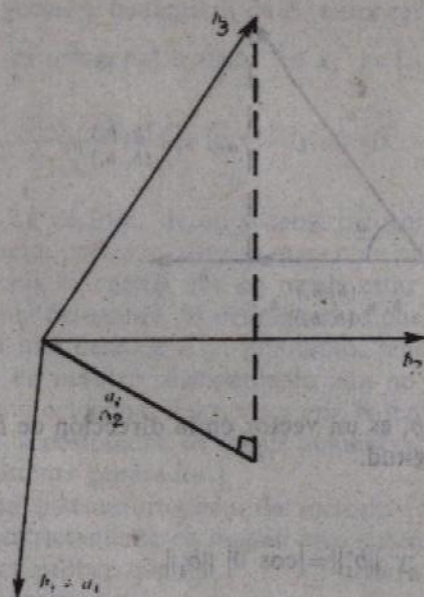


Fig. 2.3

De nuevo la dirección ortogonal a a_1 y a_2 está dada por el vector $h_1 - b'_1$; pero b'_1 es un vector del plano determinado por a_1 y a_2 que constituyen una base ortogonal de dicho plano y tenemos una fórmula para hallarlo, exactamente:

$$b'_1 = \frac{\langle b'_1, a_1 \rangle}{\|a_1\|^2} a_1 + \frac{\langle b'_1, a_2 \rangle}{\|a_2\|^2} a_2$$

$$= \frac{\langle b_1, a_1 \rangle}{\langle a_1, a_1 \rangle} a_1 + \frac{\langle b_1, a_2 \rangle}{\langle a_2, a_2 \rangle} a_2,$$

de donde resulta que

$$a_1 = b_1 - \frac{\langle b_1, a_1 \rangle}{\langle a_1, a_1 \rangle} a_1 - \frac{\langle b_1, a_2 \rangle}{\langle a_2, a_2 \rangle} a_2.$$

El estudiante puede verificar que a_1 es ortogonal a a_1 y a a_2 , y de aquí que $\{a_1, a_2, a_3\}$ sea un sistema ortogonal de vectores.

Observemos de nuevo que para obtener b_1 hemos hallado lo que geométricamente se llama *proyección* de b_1 sobre el plano determinado por a_1 y a_2 .

Sobre la base de las consideraciones anteriores no es difícil plantear el teorema siguiente:

TEOREMA 2.4.2 (Proceso de ortogonalización de Gram-Schmidt)

Sean E un espacio con producto escalar y $b_1, b_2, \dots, b_k$ vectores linealmente

independientes de E . Entonces es posible construir a partir de los b_i un sistema ortogonal de vectores $\{a_1, a_2, \dots, a_k\}$ mediante la fórmula

$$a_{i+1} = b_{i+1} - \sum_{j=1}^i \frac{\langle b_{i+1}, a_j \rangle}{\langle a_j, a_j \rangle} a_j, \quad 1 \leq i < k$$

$$\text{y } a_1 = b_1.$$

Demostración

La haremos por inducción. Tomando $a_1 = b_1$ ya vimos anteriormente que para $n=2$ y $n=3$ se cumple el teorema. Supongamos que el sistema $\{a_1, a_2, \dots, a_n\}$ construido según la fórmula de la hipótesis es ortogonal para $n \leq m$, y probemos que mediante esta fórmula podemos encontrar a_{m+1} tal que $\{a_1, a_2, \dots, a_{m+1}\}$ sea ortogonal, esto es.

$$a_{m+1} = b_{m+1} - \sum_{j=1}^m \frac{\langle b_{m+1}, a_j \rangle}{\langle a_j, a_j \rangle} a_j,$$

donde a_{m+1} es no nulo, pues si no sería una combinación lineal de los a_j y luego de los b_j , lo que está en contradicción con que $b_1, \dots, b_k$ sean linealmente independientes. Por otra parte,

$$\langle a_{m+1}, a_i \rangle = \langle b_{m+1}, a_i \rangle - \sum_{j=1}^m \frac{\langle b_{m+1}, a_j \rangle \langle a_j, a_i \rangle}{\langle a_j, a_j \rangle}$$

$$= \langle b_{m+1}, a_i \rangle - \langle b_{m+1}, a_i \rangle = 0 \quad \text{para } i=1, \dots, m.$$

y el sistema $\{a_1, a_2, \dots, a_{m+1}\}$ será un sistema ortogonal de vectores no nulos.

Una vez encontrado $\{a_1, a_2, \dots, a_k\}$ ortogonal, es evidente que

$$\left\{ \frac{a_1}{\|a_1\|}, \frac{a_2}{\|a_2\|}, \dots, \frac{a_k}{\|a_k\|} \right\}$$

es un sistema ortonormal obtenido a partir de $\{a_1, a_2, \dots, a_k\}$.

El proceso de Gram-Schmidt da una vía para obtener un sistema ortonormal de vectores a partir de un sistema de vectores linealmente independiente.

Si partimos de una base del espacio en un espacio E con producto escalar de dimensión finita, aplicando el proceso de Gram-Schmidt y la normalización (multiplicación por el inverso de la norma) de los vectores obtenidos en dicho proceso obtenemos una base ortonormal del espacio, por lo que resulta la demostración por otra vía, en este caso constructiva, de la existencia de base ortonormal en un espacio E con producto escalar de dimensión finita.

Veamos algunos ejemplos prácticos de aplicación del proceso de Gram-Schmidt.

Ejemplos

1. En $\mathbb{R}^3$ con el producto escalar usual consideremos los vectores $b_1=(3,0,4)$, $b_2=(-1,0,7)$, $b_3=(2,9,11)$ y construyamos mediante el proceso de Gram-Schmidt un sistema ortonormal de $\mathbb{R}^3$.

El sistema $\{b_1, b_2, b_3\}$ es linealmente independiente, lo que puede verificarse con facilidad.

Aplicemos el proceso de Gram-Schmidt.

$$a_1 = b_1 = (3, 0, 4)$$

$$a_2 = b_2 - \frac{\langle b_2, a_1 \rangle}{\langle a_1, a_1 \rangle} a_1 = (-1, 0, 7) - \frac{\langle (-1, 0, 7), (3, 0, 4) \rangle}{25} (3, 0, 4) = (-4, 0, 3)$$

$$\begin{aligned} a_3 &= b_3 - \frac{\langle b_3, a_1 \rangle}{\langle a_1, a_1 \rangle} a_1 - \frac{\langle b_3, a_2 \rangle}{\langle a_2, a_2 \rangle} a_2 \\ &= (2, 9, 11) - \frac{\langle (2, 9, 11), (3, 0, 4) \rangle}{25} (3, 0, 4) - \frac{\langle (2, 9, 11), (-4, 0, 3) \rangle}{25} (-4, 0, 3) = (0, 9, 0) \end{aligned}$$

El sistema ortonormal $\{c_1, c_2, c_3\}$ estará dado por

$$c_1 = \frac{a_1}{\|a_1\|}, \quad c_2 = \frac{a_2}{\|a_2\|}, \quad c_3 = \frac{a_3}{\|a_3\|},$$

esto es,

$$c_1 = \left(\frac{3}{5}, 0, \frac{4}{5}\right), \quad c_2 = \left(-\frac{4}{5}, 0, \frac{3}{5}\right), \quad c_3 = (0, 1, 0).$$

Como $\{b_1, b_2, b_3\}$ es una base de $\mathbb{R}^3$, hemos obtenido mediante este proceso una base ortonormal de $\mathbb{R}^3$.

2. En $\mathbb{C}^2$ con el producto escalar

$$\langle z_1, z_2 \rangle = z_1 \bar{w}_1 + i z_1 \bar{w}_2 - i z_2 \bar{w}_1 + 2 z_2 \bar{w}_2,$$

hallemos un sistema ortogonal a partir de $a_1=(1,0)$, $A_2=(1,i)$.

El sistema $\{(1,0), (1,i)\}$ es un sistema linealmente independiente. El sistema ortogonal estará dado por

$$b_1 = a_1 = (1, 0)$$

$$b_2 = a_2 - \frac{\langle a_2, b_1 \rangle}{\langle b_1, b_1 \rangle} b_1 = (1, i) - \frac{\langle (1, i), (1, 0) \rangle \langle 1, 0 \rangle}{\langle (1, 0), (1, 0) \rangle} = (-1, i)$$

El estudiante puede comprobar que el sistema $\{(1,0), (-1,i)\}$ es ortogonal en este espacio calculando $\langle (1,0), (-1,i) \rangle$.

3. El proceso de ortogonalización de Gram-Schmidt puede aplicarse a un sistema de vectores de un espacio de dimensión infinita. Por ejemplo, sea $C[-1,1]$ el espacio de las funciones continuas en $[-1,1]$ con el producto

escalar $\langle f, g \rangle = \int_{-1}^1 f(t) g(t) dt$. Podemos construir un sistema ortogonal a partir del sistema linealmente independiente de dicho espacio, formado por las funciones

$$f_1(x) = 1, \quad f_2(x) = x, \dots, \quad f_n(x) = x^n \dots \quad (n \in \mathbb{N}).$$

Recomendamos al estudiante que halle algunos elementos de dicho sistema ortogonal.

En el caso de un espacio con producto escalar de dimensión finita, cuando aplicamos el proceso de Gram-Schmidt a una base $(b_i)_{i=1, \dots, n}$ para obtener una base ortogonal (a_i) , la matriz P de cambio de base de (b_i) a (a_i) es una matriz triangular. Verifique el estudiante esta afirmación buscando dicha matriz.

Al motivar geoméricamente, para los casos de dos y tres vectores, el proceso de ortogonalización de Gram-Schmidt, encontramos que la solución estaba dada mediante la utilización de un concepto geométrico conocido: la proyección perpendicular de un vector sobre una recta o un plano. En estos casos sabemos que la proyección está relacionada con la búsqueda de la mejor aproximación a un vector por vectores de una recta o un plano, ya que, por ejemplo, en el caso del plano π y $\beta \notin \pi$ al proyectar β perpendicularmente sobre dicho plano encontramos que $\|\beta - \alpha\| \leq \|\beta - \gamma\| \quad \forall \gamma \in \pi$ (figura 2.4).

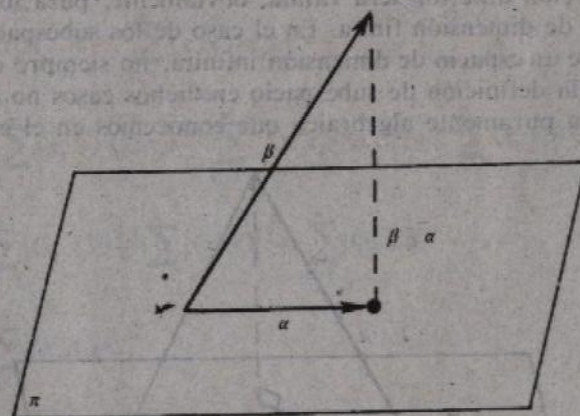


Fig. 2.4

En las aplicaciones aparece frecuentemente una generalización de este problema; esto es, dados un subespacio W de un espacio E con producto escalar y un vector β de E tal que $\beta \notin W$, se plantea la búsqueda de un vector α de W tal que $\|\beta - \alpha\| \leq \|\beta - \gamma\|$ para todo $\gamma \in W$. Si seguimos la intuición, es lógico pensar que si este vector α existe, debe ser un vector que sea ortogonal a todo vector de W .

El proceso de Gram-Schmidt da la respuesta a este problema en el caso de un subespacio de dimensión finita de un espacio con producto escalar

PROPOSICIÓN 2.4.2

Si E es un espacio con producto escalar y W un subespacio de dimensión finita de E , y $\{a_1, a_2, \dots, a_n\}$ una base ortonormal de W , entonces si $\beta \in E$, el vector $\alpha \in W$ dado por

$$\alpha = \sum_{i=1}^n \langle \beta, a_i \rangle a_i \text{ es tal que } \|\beta - \alpha\| \leq \|\beta - \gamma\| \quad \forall \gamma \in W.$$

Demostración

Sabemos que $\beta - \alpha$ es ortogonal a todo a_i por la demostración del teorema 2.4.1. luego $\beta - \alpha$ es ortogonal a todo vector de W , pues si $x \in W$

$$x = \sum_{i=1}^n x_i a_i, \text{ y } \langle \beta - \alpha, x \rangle = \sum_{i=1}^n x_i \langle \beta - \alpha, a_i \rangle = 0.$$

Probemos que entonces $\|\beta - \alpha\| \leq \|\beta - \gamma\| \quad \forall \gamma \in W$.
Podemos escribir

$$\beta - \gamma = \beta - \alpha + \alpha - \gamma \quad \forall \gamma \in W.$$

Como $\alpha - \gamma \in W$ y $\beta - \alpha$ es ortogonal a todo vector de W , tenemos que $\|\beta - \gamma\|^2 = \|\beta - \alpha\|^2 + \|\alpha - \gamma\|^2$.

De aquí, $\|\beta - \gamma\| \geq \|\beta - \alpha\| \quad \forall \gamma \in W$.

La proposición anterior será válida, obviamente, para todo subespacio de un espacio de dimensión finita. En el caso de los subespacios de dimensión infinita de un espacio de dimensión infinita, no siempre es válida. (ya que de hecho la definición de subespacio en dichos casos no coincide exactamente con la puramente algebraica que conocemos en el caso finito).

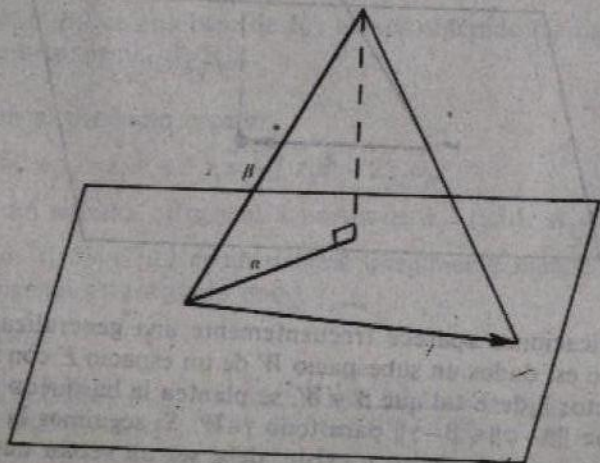


Fig. 2.5

Observemos que $\|\beta - \alpha\| \leq \|\beta - \gamma\|$ significa que α es el vector de W que permite obtener un vector $\beta - \gamma$ de longitud mínima. En el caso de un plano puede interpretarse geométricamente como el punto del plano que permite obtener la mínima distancia entre este y un punto exterior (figura 2.5).

Otras propiedades de los sistemas y bases ortogonales y ortonormales

Veremos ahora algunos resultados de utilidad acerca de los sistemas y bases ortogonales y ortonormales, fundamentalmente para espacios con producto escalar de dimensión finita. El estudiante comprobará posteriormente que en espacios de dimensión infinita, algunos de estos resultados sufren cambios, pero conocerlos en el caso finito le servirá de base para el estudio posterior del caso infinito.

(1) *Desigualdad de Bessel.* Sean E un espacio con producto escalar y $\{x_1, \dots, x_n\}$ un sistema ortonormal de vectores de E . Entonces

$$\sum_{i=1}^n |\langle x, x_i \rangle|^2 \leq \|x\|^2 \quad \forall x \in E.$$

Demostración

Tomemos $x = \sum_{i=1}^n \langle x, x_i \rangle x_i$ y calculemos $\langle x, x \rangle$:

$$\langle x - \sum_{i=1}^n \langle x, x_i \rangle x_i, x - \sum_{i=1}^n \langle x, x_i \rangle x_i \rangle =$$

$$= \langle x, x \rangle - \sum_{i=1}^n \langle \overline{\langle x, x_i \rangle} \langle x, x_i \rangle \rangle - \sum_{i=1}^n \langle x, x_i \rangle \langle x, x_i \rangle + \sum_{i=1}^n \langle x, x_i \rangle \langle \overline{\langle x, x_i \rangle} \langle x, x_i \rangle \rangle$$

$$= \|x\|^2 - \sum_{i=1}^n |\langle x, x_i \rangle|^2 - \sum_{i=1}^n |\langle x, x_i \rangle|^2 + \sum_{i=1}^n |\langle x, x_i \rangle|^2$$

$$= \|x\|^2 - \sum_{i=1}^n |\langle x, x_i \rangle|^2.$$

Como $\langle x, x \rangle \geq 0$ para todo $x \in E$, tenemos

$$\|x\|^2 - \sum_{i=1}^n |\langle x, x_i \rangle|^2 \geq 0.$$

y de aquí resulta

$$\sum_{i=1}^n |\langle x, x_i \rangle|^2 \leq \|x\|^2.$$

La técnica que usamos en esta demostración combina las propiedades del producto escalar con la definición de sistema ortogonal, y se utiliza en la demostración de muchas propiedades. El estudiante debe tenerla en cuenta para utilizarla en la obtención de resultados similares. La desigualdad de Bessel puede ser enunciada también para un sistema ortogonal, como sigue:

Si $\{y_1, y_2, \dots, y_n\}$ es un sistema ortogonal de un espacio con producto escalar E y $y \in E$, entonces

$$\sum_{k=1}^n \frac{|\langle y, y_k \rangle|^2}{\|y_k\|^2} \leq \|y\|^2.$$

La demostración no difiere de la anterior, por lo que la dejamos para que sea realizada por el estudiante.

La desigualdad de Bessel tiene una interpretación geométrica que veremos más adelante, pero el estudiante puede pensar desde este momento en qué consiste.

(2) En un espacio con producto escalar de dimensión finita E , dada una base ortonormal $\{a_1, a_2, \dots, a_n\}$ de E , se tiene que si $\bar{x} = \sum_{i=1}^n x_i a_i$, $y = \sum_{j=1}^n y_j a_j$, entonces

$$\langle x, y \rangle = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n.$$

Demostración

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_{i=1}^n x_i a_i, \sum_{j=1}^n y_j a_j \right\rangle = \sum_i \sum_j x_i \bar{y}_j \langle a_i, a_j \rangle = \\ &= x_1 \bar{y}_1 + \dots + x_n \bar{y}_n \end{aligned}$$

Este resultado reafirma la importancia de las bases ortonormales y de que en todo espacio con producto escalar de dimensión finita sea posible encontrar una base ortonormal, pues acabamos de probar que si expresamos los vectores del espacio en función de una base ortonormal de este, la expresión del producto escalar no difiere de la del producto escalar usual en K^n . Esto lleva implícito un hecho que formalizaremos posteriormente: que todo espacio vectorial con producto escalar de dimensión n sobre K es isomorfo a K^n , no solo en el sentido de la estructura de espacio vectorial, sino también desde el punto de vista de la estructura de espacio con producto escalar.

Vimos anteriormente que si tenemos un sistema ortonormal $\{a_1, \dots, a_k\}$ y $x \in E$ es tal que $x = \sum_{i=1}^k x_i a_i$, entonces $x_i = \langle x, a_i \rangle$ lo cual es aplicable a una base ortonormal del espacio.

(3) Si $\{a_1, a_2, \dots, a_n\}$ es una base ortonormal del espacio con producto escalar de dimensión finita E , entonces el coeficiente i -ésimo x_i de $x \in E$ en la base

(a) $\{a_i\}_{i=1}^n$ se calcula mediante $x_i = \langle x, a_i \rangle$. En otras palabras, $x = \sum_{i=1}^n \langle x, a_i \rangle a_i$.

(4) *Identidad de Parseval*. Aplicando (2) y (3) obtenemos la siguiente igualdad, conocida como identidad de Parseval:

Si $\{a_1, a_2, \dots, a_n\}$ es una base ortonormal de un espacio con producto escalar de dimensión finita E , entonces para todo $x \in E$,

$$\|x\|^2 = \sum_{i=1}^n |\langle x, a_i \rangle|^2.$$

Demostración

$$\text{Sea } x = \sum_{i=1}^n x_i a_i.$$

$$\|x\|^2 = \langle x, x \rangle = \sum_{i=1}^n x_i \bar{x}_i = \sum_{i=1}^n \langle x, a_i \rangle \langle x, a_i \rangle = \sum_{i=1}^n |\langle x, a_i \rangle|^2.$$

En un espacio vectorial con producto escalar, la identidad de Parseval es una condición suficiente para que un sistema ortonormal sea una base.

(5) Si E es un espacio vectorial con producto escalar de dimensión finita y si $\{a_1, a_2, \dots, a_k\}$ es un sistema ortonormal de E tal que para todo $x \in E$,

$$\|x\|^2 = \sum_{i=1}^k |\langle x, a_i \rangle|^2, \text{ entonces}$$

$\dim E = k$ y $\{a_1, a_2, \dots, a_k\}$ es una base de E .

Demostración

Supongamos que se cumple lo contrario. Si $\{a_1, a_2, \dots, a_k\}$ no es una base de E , entonces es posible encontrar un vector x_0 no nulo que sea ortogonal a todos los (a_i) , pero según la identidad de Parseval

$$\|x_0\|^2 = \sum_{i=1}^k |\langle x_0, a_i \rangle|^2 = 0 \Rightarrow x_0 = 0$$

y esto contradice lo supuesto.

La propiedad (5) tiene un valor fundamentalmente teórico, pues en la práctica en un espacio con producto escalar de dimensión finita, para comprobar que un sistema ortonormal es una base basta comprobar si el número de sus elementos coincide con la dimensión del espacio. Sin embargo, en el caso de los espacios con producto escalar de dimensión infinita desempeñan una función importante las diferentes caracterizaciones de base ortonormal que hemos ido desarrollando de forma implícita en las demostraciones.

(6) Las siguientes propiedades equivalentes que caracterizan a una base ortonormal en un espacio de dimensión finita con producto escalar, las hemos ido demostrando implícitamente en el desarrollo del epígrafe, pero invitamos al estudiante a probar de nuevo, como resumen, la equivalencia entre estas propiedades.

En un espacio con producto escalar de dimensión finita, las siguientes propiedades son equivalentes a que el subconjunto ortonormal $B = \{a_1, a_2, \dots, a_n\}$ de E sea una base.

a) B es un conjunto ortonormal completo, esto es, no existe otro conjunto ortonormal de vectores de E que lo contenga estrictamente.

b) $\langle x, a_i \rangle = 0$ para $i = 1, \dots, n \Rightarrow x = 0$

c) $\langle B \rangle = E$

d) $\forall x \in E, x = \sum_{i=1}^n \langle x, a_i \rangle a_i$

e) $\forall x, y \in E, \langle x, y \rangle = \sum_{i=1}^n \langle x, a_i \rangle \langle y, a_i \rangle$

f) $\forall x \in E, \|x\|^2 = \sum_{i=1}^n |\langle x, a_i \rangle|^2$

Complemento ortogonal. Proyecciones

En $\mathbb{R}^2$ y $\mathbb{R}^3$ vimos los siguientes resultados que corresponden totalmente con la intuición geométrica:

En $\mathbb{R}^2$ el conjunto de vectores ortogonales a una recta dada es una recta vectorial de dirección ortogonal a la dirección de la recta inicial (figura 2.6a).

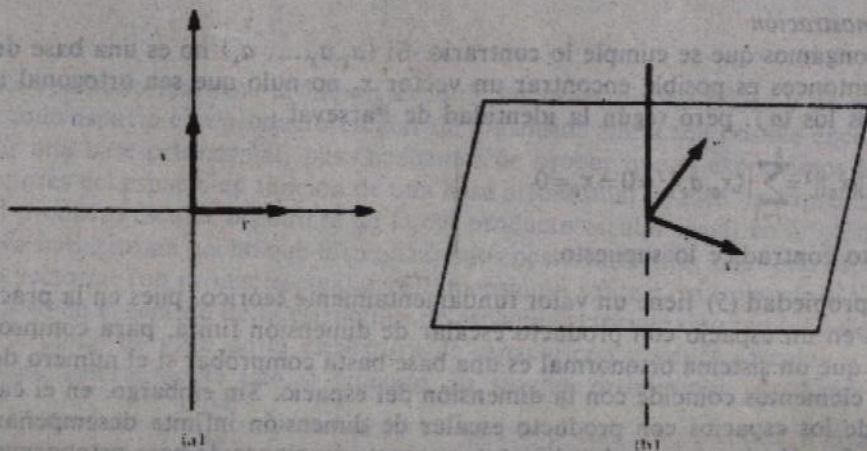


Fig. 2.6

En $\mathbb{R}^3$ el conjunto de vectores ortogonales a un plano vectorial es una recta vectorial de dirección ortogonal a todo vector del plano (figura 2.6b).

Una generalización natural de la idea anterior y que es de gran aplicación en diferentes ramas como el Análisis y la Estadística, consiste en lo siguiente:

DEFINICIÓN 2.4.5

Dados un espacio vectorial con producto escalar E y $W \subseteq E$, se llama *ortogonal* de W y se denota por $W^\perp$ (se lee: W ortogonal) al conjunto

$$W^\perp = \{y \in E : \langle y, x \rangle = 0 \forall x \in W\}.$$

Ejemplos

1. Si en $\mathbb{R}^4$ con el producto escalar usual consideramos

$W = \{w_1 = (1, 2, 3, -1, 2), w_2 = (2, 4, 7, 2, -1)\}$ entonces

$W^\perp = \{v = (x, y, z, t, u) : \langle v, w_i \rangle = 0 \forall w_i \in W\}.$

Esto es v debe cumplir

$$\langle v, w_1 \rangle = 0$$

$$\langle v, w_2 \rangle = 0$$

De aquí resulta el sistema

$$\begin{cases} x + 2y + 3z - t + 2u = 0 \\ 2x + 4y + 7z + 2t - u = 0 \end{cases}$$

y

$$W^\perp = \{(2\alpha + 13\beta - 17\gamma, -\alpha, -4\beta + 5\gamma, \beta, \gamma) : \alpha, \beta, \gamma \in \mathbb{R}\}.$$

Analice el estudiante cuál será $W^\perp$ si $W = \langle w_1, w_2 \rangle$.

2. Sea $E = C[-1, 1]$ con el producto escalar $\langle f, g \rangle = \int_{-1}^1 f(t) g(t) dt$ y sea $W = \{f \in C[-1, 1] \text{ tales que } f(-t) = -f(t)\}$, esto es, el subespacio de E de las funciones impares.

Hallemos $W^\perp$. Para ello razonaremos mediante algunas consideraciones de tipo analítico.

$$W^\perp = \{g \in C[-1, 1] : \langle f, g \rangle = \int_{-1}^1 f(t) g(t) dt = 0 \forall f \in W\}.$$

Si f es impar, su gráfico en $[-1, 1]$ es aproximadamente de la forma representada en la figura 2.7.

$$\text{Por tanto, } \int_{-1}^1 f(t) dt = 0.$$

Por otra parte, es fácil verificar, y lo dejamos al estudiante, que el producto de dos funciones impares es una función par, esto es, una función g tal que $g(-t) = g(t)$, y el producto de una función par por una impar es una función impar.

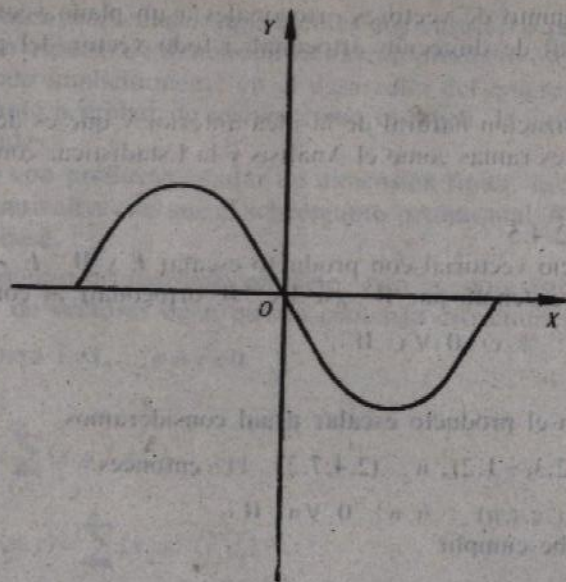


Fig. 2.7

Es evidente también que el gráfico de una función par es simétrico con respecto al eje de las y ; por tanto, si g es par, $\int_{-1}^1 g(t) dt \neq 0$ para g no nula.

Además, es conocido que toda función (en particular las de E) se puede expresar en forma única como la suma de una función par y una impar, exactamente.

Si $f \in E$, $f = f_1 + f_2$ con f_1 par y f_2 impar, donde

$$f_1(t) = \frac{f(t) + f(-t)}{2}, \quad f_2(t) = \frac{f(t) - f(-t)}{2}.$$

Ahora podemos buscar $W^\perp$.

Si $g \in W^\perp$, entonces

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt = 0 \quad \forall f \in W.$$

Como f es impar, tomemos $g = g_1 + g_2$ con g_1 par y g_2 impar; tendremos

$$\langle f, g \rangle = \int_{-1}^1 f(t)g_1(t) dt + \int_{-1}^1 f(t)g_2(t) dt.$$

$\int_{-1}^1 f(t)g_1(t) dt = 0$ ya que $f \cdot g_1$ es impar. Entonces, es necesario y suficiente que

$$\int_{-1}^1 f(t)g_2(t) dt = 0 \quad \forall f \in E.$$

Pero g_2 es impar, luego

$$\int_{-1}^1 g_2(t)g_2(t) dt = 0, \text{ de donde } g_2 = 0.$$

$W^\perp$ será, pues, el conjunto de las funciones pares.

Aunque el ejemplo anterior resulta un poco laborioso, hemos querido que el estudiante pueda darse cuenta de cómo hallar el ortogonal de un subconjunto infinito en el caso de un espacio de dimensión infinita. En el caso de los espacios de dimensión finita, la búsqueda del ortogonal no tiene ninguna dificultad, pues como veremos se puede reducir a la solución de un sistema de ecuaciones lineales homogéneo.

Algunas propiedades del ortogonal de un subconjunto son las siguientes:

Sea E un espacio con producto escalar y $W \subset E$. Entonces:

(1) $W^\perp$ es un subespacio vectorial de E .

(2) $\langle W \rangle^\perp = W^\perp$

(3) $\{0\}^\perp = E$ y $E^\perp = \{0\}$

(4) TEOREMA 2.4.3

Si E es de dimensión finita, entonces para todo F subespacio de E se tiene que $E = F \oplus F^\perp$ y $(F^\perp)^\perp = F$

Probaremos algunas de estas propiedades para mostrar las técnicas de trabajo, y dejamos las restantes como ejercicio para el estudiante.

Demostración de (2)

Debemos probar que $\langle W \rangle^\perp \subset W^\perp$ y reciprocamente.

Recordemos que $\langle W \rangle = \{ \alpha_1 x_1 + \dots + \alpha_n x_n : \alpha_i \in K, x_i \in W, n \in \mathbb{N} \}$. Si $y \in \langle W \rangle^\perp$, entonces $\langle y, \alpha_1 x_1 + \dots + \alpha_n x_n \rangle = 0$ para todo elemento de esta forma, en particular $\langle y, x \rangle = 0$ para todo $x \in W$; luego $\langle W \rangle^\perp \subset W^\perp$.

Sea ahora $y \in W^\perp$. Debemos probar que $\langle y, \alpha_1 x_1 + \dots + \alpha_n x_n \rangle = 0$ para todo $\alpha_i \in K, x_i \in W$. Entonces

$$\langle y, \alpha_1 x_1 + \dots + \alpha_n x_n \rangle = \bar{\alpha}_1 \langle y, x_1 \rangle + \dots + \bar{\alpha}_n \langle y, x_n \rangle = 0,$$

ya que $\langle y, x_i \rangle = 0$ para todo i por ser $y \in W^\perp$.

Esta propiedad significa que para hallar el ortogonal de un subespacio es suficiente hallar el ortogonal de un generador, lo cual en dimensión finita es equivalente a plantear un sistema de ecuaciones lineales homogéneo dado por $\langle x, a_i \rangle = 0, i = 1, \dots, n$, ya que $W = \langle a_1, a_2, \dots, a_n \rangle$.

Demostración de (3)

$E^\perp = \{0\}$. Sabemos que $0 \in E^\perp$ y que $E^\perp = \{ y \in E : \langle x, y \rangle = 0 \quad \forall x \in E \}$. Luego $y \in E^\perp \Rightarrow \langle y, y \rangle = 0 \Rightarrow y = 0$.

Demostración del teorema 2.4.3

Debemos probar que $\forall z \in E$ puede escribirse como $z = x + y$ con $x \in F, y \in F^\perp$ y que $F \cap F^\perp = \{0\}$.

Como F es un subespacio de E , será un espacio con producto escalar de dimensión finita. Luego podemos afirmar que en él existe una base ortonormal.

Sea $\{h_1, h_2, \dots, h_n\}$ una base ortonormal de F . Tomemos $x = \sum_{i=1}^n \langle z, h_i \rangle h_i$, $y = z - x$, entonces $x \in F$ y $y = z - x \in F^\perp$, según vimos en la demostración del teorema 2.4.1 (existencia de base ortonormal).

Problemos ahora que $F \cap F^\perp = \{0\}$. Sea $z \in F \cap F^\perp$. Entonces $\langle z, z \rangle = 0 \Rightarrow z = 0$. En cuanto a que $(F^\perp)^\perp = F$, es evidente que $F \subset (F^\perp)^\perp$ por la definición de $(F^\perp)^\perp$.

Problemos lo contrario, sea $z \in (F^\perp)^\perp$. Por la descomposición $E = F \oplus F^\perp$, se cumple $z = x + y$ con $x \in F$, $y \in F^\perp$. Luego $\langle z, y \rangle = 0 = \langle x + y, y \rangle = \langle x, y \rangle + \langle y, y \rangle = 0 + \|y\|^2$, $\langle y, y \rangle = 0 \Rightarrow y = 0$.

De aquí que $z = x$, sea $z \in F$, con lo que concluimos que $(F^\perp)^\perp = F$. A partir de la descomposición $E = F \oplus F^\perp$, para todo subespacio F de E se definen las proyecciones asociadas a dicha descomposición, esto es, las aplicaciones:

$p_F : E \rightarrow E$ tal que dada $z \in E$ con $z = x + y$, $x \in F$, $y \in F^\perp$, se cumple $p_F(z) = x$.

$p_{F^\perp} : E \rightarrow E$ dada por $p_{F^\perp}(z) = y$.

Como $y = z - x$, tenemos que

$p_{F^\perp}(z) = \text{id}_E(z) - p_F(z) = (\text{id}_E - p_F)(z)$.

p_F y $\text{id}_E - p_F$ se llaman *proyecciones ortogonales sobre F y $F^\perp$ respectivamente*.

Tanto el teorema 2.4.3 como la definición de las proyecciones ortogonales, son de gran utilidad en la práctica. Si volvemos a la proposición 2.4.2, en la cual estudiamos la "mejor aproximación" de un vector por vectores de un subespacio, y unimos el resultado obtenido allí con la demostración del teorema 2.4.3, podemos obtener un interesante resultado geométrico.

En la demostración anterior, dados F subespacio de E y $z \in F$,

$p_F(z) = \sum_{i=1}^n \langle z, h_i \rangle h_i$, donde $F = \langle h_1, \dots, h_n \rangle$ y $\{h_i\}_{i=1}^n$ es una base ortonormal de F , y en la proposición 2.3.2 obtuvimos que bajo estas mismas premisas el vector $p_F(z)$ es tal que

$\|z - p(z)\| \leq \|z - y\|$ para todo $y \in F$.

Luego, se confirma un resultado general que coincide con lo conocido para el plano y la recta, y que puede interpretarse como en el caso del plano, en el sentido de que la *distancia mínima* al subespacio se obtiene mediante la proyección ortogonal. Esperamos que después de estas últimas consideraciones, el estudiante pueda dar la interpretación geométrica de la desigualdad de Bessel que mencionamos anteriormente.

Ejercicios

- En $\mathbb{R}^4$ con el producto escalar usual, halle el ortogonal de $W = \langle (1, 2, 3, 4), (3, -5, 7, 8) \rangle$.
- Halle en $\mathbb{R}^3$ con respecto al producto escalar usual, una base ortonormal de $W = \langle (1, i, 1), (1 + i, 0, 2) \rangle$.
- Encuentre una base ortonormal de $\mathbb{C}^2$ para el producto escalar $\langle x, y \rangle = x_1 y_1 + i x_1 y_2 - i x_2 y_1 + 2 x_2 y_2$.
- Sea E el espacio de los polinomios con coeficientes reales de grado ≤ 2 con el producto escalar del ejercicio 8 del epígrafe 1.
 - Halle una base del ortogonal a $\{2t + 1\}$.
 - Halle una base ortonormal de E a partir de la base $\{1, t, t^2\}$.
- Sea $M_n(\mathbb{R})$ con el producto escalar $\langle A, B \rangle = \text{tr } BA$. Encuentre el ortogonal de
 - el subespacio de las matrices diagonales.
 - el subespacio de las matrices simétricas.
- Sean E un espacio con producto escalar de dimensión finita n , W un subespacio de E y $\{a_1, a_2, \dots, a_r\}$ una base de W . Si $\{v_1, \dots, v_s\}$ es un sistema de vectores linealmente independientes tales que $\langle a_i, v_j \rangle = 0$ para todo i, j , pruebe que $\{v_1, \dots, v_s\}$ es una base de $W^\perp$.
- Sea E un espacio con producto escalar real (euclideo). Pruebe que
 - $\|x\| = \|y\| \Leftrightarrow \langle x + y, x - y \rangle = 0$
 - $\|x + y\|^2 = \|x\|^2 + \|y\|^2 \Leftrightarrow \langle x, y \rangle = 0$
 ¿Se cumplirán a) y b) en un espacio unitario (sobre $\mathbb{C}$)?
- Sea E un espacio con producto escalar de dimensión finita. Pruebe que si V y W son subespacios de E , entonces
 - $(V + W)^\perp = V^\perp \cap W^\perp$
 - $(V \cap W)^\perp = V^\perp + W^\perp$
 Analice si a) y b) se cumplen para un espacio E con producto escalar de dimensión infinita.
- Pruebe que si E es un espacio con producto escalar complejo (unitario), x y y de E son ortogonales si y solo si $\|\alpha x + \beta y\|^2 = \|\alpha x\|^2 + \|\beta y\|^2$ para todo α, β de números complejos.
- Si E es un espacio con producto escalar y $\{x, y\}$ es un sistema ortonormal en E , calcule $d(x, y)$.
- Si E es un espacio con producto escalar y W un subespacio de E tal que $\langle v, w \rangle + \langle w, v \rangle \leq \langle w, w \rangle$ para todo $w \in W$, pruebe que $\langle v, w \rangle = 0$ para todo $w \in W$.

12. Sean E un espacio con producto escalar y $\alpha, \beta \in E$. Pruebe que $\alpha = \beta \iff \langle \alpha, r \rangle = \langle \beta, r \rangle \forall r \in E$.

13. Sea en $\mathbb{R}^2$ con el producto escalar usual, el subespacio $W = \langle (3, 4) \rangle$. Halle

a) $p_W(x_1, x_2)$.

b) la matriz de p_W en la base canónica.

c) $W^\perp$.

d) una base ortonormal en la cual p_W esté representada por la matriz

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

e) a, b, c, d para $\mathbb{R}^2$ con el producto escalar

$$\langle x, y \rangle = x_1 y_1 - x_1 y_2 - x_2 y_1 + 4 x_2 y_2.$$

14. Sea $E = M_n(\mathbb{C})$ con el producto escalar $\langle A, B \rangle = \text{tr } B^* A$. Halle el ortogonal del subespacio de las matrices diagonales.

15. Sea W un subespacio de dimensión finita de un espacio con producto escalar E y sea p_W la proyección ortogonal sobre W . Pruebe que $\langle p_W(x), y \rangle = \langle x, p_W(y) \rangle$ para todo $x, y \in E$.

16. Sea E un espacio con producto escalar n dimensional y $\{a_1, a_2, \dots, a_n\}$ una base ortonormal de E . Si f es una aplicación lineal de E en E y $A = M(f(a_i)) = (a_j)$, pruebe que $a_j = \langle f(a_i), a_j \rangle$.

17. Sea E un espacio con producto escalar y W un subespacio de E de dimensión finita. Hay muchas proyecciones que tienen a W como imagen. En particular, la proyección ortogonal cumple que $\|p_W(x)\| \leq \|x\|$ para todo $x \in E$. Pruebe que si p es una proyección con imagen W tal que $\|p(x)\| \leq \|x\| \forall x \in E$, entonces $p = p_W$.

2.5 Formas lineales sobre un espacio vectorial

Hemos visto en los epígrafes anteriores que los espacios con producto escalar poseen características que los diferencian de los espacios vectoriales simples. Estas características se acentúan en el caso de las aplicaciones lineales y, en particular, en los endomorfismos de dichos espacios.

Antes de comenzar el estudio de las aplicaciones lineales en un espacio con producto escalar, profundizaremos en un tipo especial de aplicaciones lineales sobre un espacio vectorial cualquiera: las llamadas *formas lineales* o *funcionales lineales*.

El concepto *forma lineal* es un poderoso instrumento de trabajo en el estudio de diversos problemas en Álgebra, Análisis funcional, Investigación

operacional y otras ramas, y en particular, en los espacios con producto escalar sirve de base al estudio de las propiedades de importantes clases de aplicaciones lineales.

Formas o funcionales lineales

DEFINICIÓN 2.5.1

Dado un espacio vectorial E sobre K , se llama *forma lineal sobre E* a una aplicación lineal de E en K .

Recordemos que esto significa que si f es una forma lineal sobre E , entonces es una aplicación $f: E \rightarrow K$ tal que

$$f(x+y) = f(x) + f(y) \text{ para todo } x, y \in E$$

$$f(ax) = a f(x) \text{ para todo } a \in K$$

o de forma equivalente

$$f(ax + by) = af(x) + bf(y) \forall a, b \in K \forall x, y \in E.$$

A las formas lineales sobre un espacio E suele llamárseles también *funcionales lineales*, especialmente cuando el espacio E se trata de un espacio de funciones.

El hecho de que una forma lineal sea un caso particular de aplicación lineal entre espacios vectoriales, permite aplicarle las propiedades generales estudiadas para las aplicaciones lineales; pero como la forma lineal está definida del espacio E en K hace que posea propiedades específicas.

Veamos primero algunos ejemplos. La verificación de la linealidad la dejaremos al estudiante, ya que no es más que un caso particular de lo realizado al tratar aplicaciones lineales en general.

Ejemplos

1. En $\mathbb{R}^n$, dados $a_1, a_2, \dots, a_n$ números reales fijos, la aplicación $f(x) = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ donde $x = (x_1, \dots, x_n)$, es una forma lineal sobre $\mathbb{R}^n$.

Toda forma lineal sobre $\mathbb{R}^n$ es de la forma anterior, ya que si recordamos que toda aplicación lineal de $\mathbb{R}^n$ en K puede obtenerse a partir de una matriz A de tamaño $(1, n)$ mediante la expresión $Y = AX$, tendremos que, si $A = (a_1, \dots, a_n)$,

$$f(x) = (a_1, \dots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + \dots + a_n x_n.$$

2. Sea $\mathbb{R}[X]$ el espacio vectorial real de los polinomios con coeficientes reales. Entonces la aplicación $T_a: \mathbb{R}[X] \rightarrow \mathbb{R}$ donde $a \in \mathbb{R}$ fijo, dada por $T_a(p(x)) = p(a)$, es una forma lineal sobre $\mathbb{R}[X]$.

Con mas generalidad, si $a_1, a_2, \dots, a_n$ y $a_1, a_2, \dots, a_n$ son números reales fijos, la aplicación $T: \mathbb{R}[X] \rightarrow \mathbb{R}$ dada por $T(p(x)) = a_1 p(a_1) + \dots + a_n p(a_n)$ es una forma lineal sobre $\mathbb{R}[X]$.

3. En $M_n(K)$ con $K = \mathbb{R}$ o $K = \mathbb{C}$ como espacio vectorial sobre K , la aplicación $\text{tr}: M_n(K) \rightarrow K$ definida por:

Si $A = (a_{ij})$, entonces $\text{tr}(A) = a_{11} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}$ es una forma lineal que aparece frecuentemente en diferentes resultados.

4. En $C[a, b] = \{f: [a, b] \rightarrow \mathbb{R} \text{ con } f \text{ continua}\}$ como espacio vectorial sobre $\mathbb{R}$ para las operaciones usuales de funciones, es una forma lineal la aplicación $I: C[a, b] \rightarrow \mathbb{R}$ tal que $I(f) = \int_a^b f(t) dt$.

También es una forma lineal en $C[a, b]$, dada $f_0 \in C[a, b]$ fija, la aplicación $I_{f_0}: C[a, b] \rightarrow \mathbb{R}$ definida por $I_{f_0}(f) = \int_a^b f(t) f_0(t) dt$.

En la mecánica cuántica es de mucho interés la forma lineal definida sobre $C[a, b]$ por $x_{t_0}(f) = f(t_0)$ donde $t_0 \in (a, b)$ fijo.

5. En $l_2 = \{(x_n)_{n \in \mathbb{N}} : x_n \in \mathbb{R} \text{ y } \sum_{n=1}^{\infty} x_n^2 \text{ es convergente}\}$ como espacio vectorial sobre $\mathbb{R}$, para las operaciones usuales de sucesiones se define una forma lineal para cada $k \in \mathbb{N}$ fijo mediante $f_k(x_n) = x_k$.

Algunas propiedades específicas de las formas lineales

Veamos algunas propiedades de las formas lineales que las diferencian de las aplicaciones lineales en general.

- (1) Toda forma lineal no nula sobre E es sobreyectiva. En efecto, sea $y^* \in E^*$, entonces $\dim \text{Im } y^* = \dim K$. De aquí que $\dim \text{Im } y^* = 1$ y de esto se deduce que $\text{Im } y^* = K$. De lo anterior resulta que la propiedad (1) también puede enunciarse así: Toda forma lineal no nula tiene rango 1.

- (2) Si E es un espacio de dimensión finita n , entonces el núcleo de toda forma lineal no nula es un subespacio de dimensión $n-1$ de E .

Esto es una consecuencia inmediata de la propiedad (1) y del teorema del rango, y da lugar a una interpretación geométrica de las formas lineales.

Por analogía con el espacio usual, en el cual podemos interpretar los subespacios de dimensión 2 como planos, llamamos *hiperplano* a un subespacio

de dimensión $n-1$ de E . De esta forma la propiedad (2) puede ser enunciada como sigue:

El núcleo de toda forma lineal no nula sobre un espacio E de dimensión finita es un hiperplano de dicho espacio.

Recíprocamente, dado un hiperplano H de E y $x_0 \in E: x_0 \notin H$, existe una única forma lineal y^* sobre E tal que $\text{Ker } y^* = H$ y $y^*(x_0) = 1$. Para probarlo basta tomar en E una base $\{a_1, a_2, \dots, a_{n-1}, x_0\}$, donde $\{a_1, \dots, a_{n-1}\}$ es una base de H , y tomar y^* como la única forma lineal tal que $y^*(a_i) = 0$, $i = 1, \dots, n-1$ y $y^*(x_0) = 1$.

Dado el hiperplano H de E , ¿será esta forma lineal la única sobre E tal que su núcleo sea H ? De la misma demostración de la existencia de y^* resulta que no es así. Veamos si existe alguna relación entre las formas lineales que tengan por núcleo al hiperplano H y la forma y^* encontrada.

Tenemos que $E = H \oplus Kx_0$ donde $Kx_0 = \{ax_0: a \in K\}$; luego $\forall x \in E$, x se expresa en forma única como $x = ax_0 + y$ con $y \in H$. De aquí que

$$y^*(x) = ay^*(x_0) + y^*(y) = ay^*(x_0) \text{ y } x = y^*(x)x_0 + y, \text{ con } y \in H.$$

Sea z^* una forma lineal tal que su núcleo sea H . Entonces

$$z^*(x) = y^*(x) z^*(x_0) + z^*(y) = z^*(x_0) y^*(x) \quad \forall x \in E.$$

De aquí que $z^* = z^*(x_0) y^*$, y z^* y y^* sean proporcionales.

Es evidente que, recíprocamente, si z^* y y^* no nulas, son dos formas proporcionales, o sea, $z^* = ay^*$, entonces tienen el mismo núcleo, que es un hiperplano.

Podemos interpretar geométricamente lo anterior considerando que un hiperplano H de E (figura 2.8a) determina en E^* una recta vectorial (figura 2.8b).

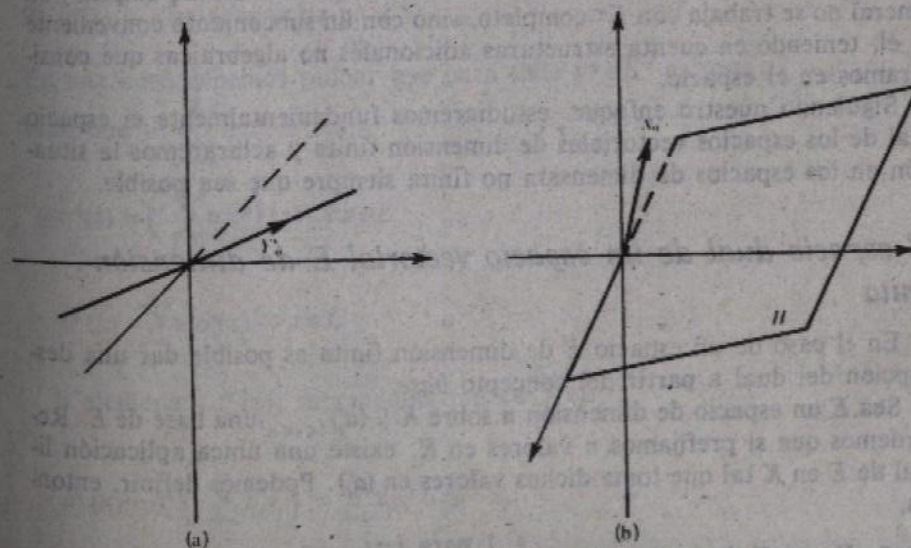


Fig. 2.8

Realmente la relación entre formas lineales e hiperplanos es un poco más compleja, pues en todo momento hemos trabajado bajo la premisa de que x_0 es un punto fijo fuera del hiperplano. Quedaría por analizar qué ocurre si tomamos otro punto diferente de x_0 , pero creemos que con lo estudiado se tiene la idea básica de la relación entre hiperplano y formas lineales.

Todo lo que acabamos de analizar para espacios de dimensión finita puede extenderse a espacios de dimensión infinita, pero al no contar en este caso con el teorema del rango, es necesario usar conceptos que no hemos estudiado anteriormente y que introduciremos en próximos capítulos; estos son: *relación de equivalencia y cociente*.

Sabemos que la suma de dos aplicaciones lineales entre espacios vectoriales es una aplicación lineal, y lo mismo ocurre con el producto de una aplicación lineal por un número. También sabemos que el conjunto $L_K(E, F)$ de las aplicaciones lineales entre los espacios vectoriales E y F con las operaciones

$$(f+g)(x) = f(x) + g(x) \quad \forall f, g \in L_K(E, F)$$

$$(af)(x) = af(x) \quad \forall a \in K,$$

es un espacio vectorial sobre K .

Luego el conjunto de las formas lineales sobre E , que no es más que $L_K(E, K)$, es un espacio vectorial sobre K .

DEFINICIÓN 2.5.2

Se llama *espacio dual de E* y se denota por E^* al espacio vectorial de las formas lineales sobre E con las operaciones de suma de aplicaciones y producto de un escalar por una aplicación.

Los elementos de E^* se denotan por x^* , y^* , z^* , etcétera.

Si E no es de dimensión finita, E^* es usualmente demasiado amplio y en general no se trabaja con E^* completo, sino con un subconjunto conveniente de él, teniendo en cuenta estructuras adicionales no algebraicas que consideramos en el espacio.

Siguiendo nuestro enfoque, estudiaremos fundamentalmente el espacio dual de los espacios vectoriales de dimensión finita y aclararemos la situación en los espacios de dimensión no finita siempre que sea posible.

El espacio dual de un espacio vectorial E de dimensión finita

En el caso de un espacio E de dimensión finita es posible dar una descripción del dual a partir del concepto *base*.

Sea E un espacio de dimensión n sobre K y $(a_i)_{1 \leq i \leq n}$ una base de E . Recordemos que si prefijamos n valores en K , existe una única aplicación lineal de E en K tal que toma dichos valores en (a_i) . Podemos definir, entonces,

$$a_i^* : E \longrightarrow K \text{ dada por } a_i^*(a_j) = \begin{cases} 1 & \text{para } i=j \\ 0 & \text{para } i \neq j \end{cases}$$

Por esta vía tenemos n formas lineales definidas sobre E perfectamente determinadas.

De antemano sabemos que $\dim E^* = n$, ya que $E^* = L_K(E, K)$ y $\dim L_K(E, K) = (\dim E) \cdot (\dim K) = n$. Podemos probar que la familia de n formas lineales $(a_i^*)_{1 \leq i \leq n}$ es una base de E^* . Para ello basta probar, ya que sabemos previamente la dimensión de E^* , que $(a_i^*)_{1 \leq i \leq n}$ es linealmente independiente o que genera a E^* . Probaremos, sin embargo, ambas afirmaciones, ya que nos interesa que el estudiante observe el método de trabajo seguido.

1) $(a_i^*)_{1 \leq i \leq n}$ es linealmente independiente.

Debemos probar que $\sum_{i=1}^n \alpha_i a_i^* = 0 \Rightarrow \alpha_i = 0, i = 1, \dots, n$.

$$\sum_{i=1}^n \alpha_i a_i^* = 0 \Leftrightarrow$$

$$\left(\sum_{i=1}^n \alpha_i a_i^* \right)(x) = 0 \quad \forall x \in E \Leftrightarrow \sum_{i=1}^n \alpha_i a_i^*(x) = 0 \quad \forall x \in E.$$

En particular, para $x = a_j, 1 \leq j \leq n$, tenemos

$$\sum_{i=1}^n \alpha_i a_i^*(a_j) = \alpha_j = 0.$$

Hemos demostrado que la familia $(a_i^*)_{1 \leq i \leq n}$ es linealmente independiente.

El recurso usado anteriormente es típico del trabajo en el espacio dual y debe ser observado atentamente por el estudiante.

2) $(a_i^*)_{1 \leq i \leq n}$ es generatriz en E^* .

En este caso debemos probar que para todo $y^* \in E^*$ existen $\alpha_1, \dots, \alpha_n$ en K tales que $y^* = \sum_{i=1}^n \alpha_i a_i^*$, o sea,

$$y^*(x) = \left(\sum_{i=1}^n \alpha_i a_i^* \right)(x) \quad \forall x \in E$$

$$y^*(x) = \sum_{i=1}^n \alpha_i a_i^*(x) \quad \forall x \in E.$$

Calculemos $a_i^*(x)$, suponiendo que $x = \sum_{j=1}^n x_j a_j$:

$$a_i^*(x) = a_i^* \left(\sum_{j=1}^n x_j a_j \right) = \sum_{j=1}^n x_j a_i^*(a_j) = x_i,$$

es decir, que a_i^* le asigna a un vector x de E su coordenada i -ésima en la

base $(a_i)_{1 \leq i \leq n}$. Es por esto que a las formas lineales a_i^* , $i=1, \dots, n$, se les llama *formas coordenadas asociadas a la base* $(a_i)_{1 \leq i \leq n}$.

Teníamos, pues, que encontrar las α_i tales que $y^*(x) = \sum_{i=1}^n \alpha_i a_i^*(x)$. Pero como $x = \sum_{i=1}^n x_i a_i$, entonces

$$y^*(x) = y^*\left(\sum_{i=1}^n x_i a_i\right) = \sum_{i=1}^n x_i y^*(a_i) = \sum_{i=1}^n y^*(a_i) a_i^*(x).$$

Evidentemente, tomamos $\alpha_i = y^*(a_i)$.

Así queda demostrado que $(a_i^*)_{1 \leq i \leq n}$ es generatriz en E^* y, por tanto, una base de E^* .

DEFINICIÓN 2.5.3

Dada una base $(a_i)_{1 \leq i \leq n}$ de un espacio vectorial E sobre K , se llama *base dual de E a la base* $(a_i^*)_{1 \leq i \leq n}$ de E^* formada por las formas coordenadas asociadas a $(a_i)_{1 \leq i \leq n}$.

Observemos la siguiente relación entre una base y su dual, la cual reafirma su nombre pues intercambia las funciones de los elementos del espacio y del dual.

PROPOSICIÓN 2.5.1

Si $(a_i)_{1 \leq i \leq n}$ es una base de E y $(a_i^*)_{1 \leq i \leq n}$ es la base dual de $(a_i)_{1 \leq i \leq n}$ en E^* , entonces

$$\text{para todo } y^* \in E^*, y^* = \sum_{i=1}^n y^*(a_i) a_i^*$$

$$\text{para todo } x \in E, x = \sum_{i=1}^n a_i^*(x) a_i$$

Demostración

En la demostración de que $(a_i^*)_{1 \leq i \leq n}$ es generatriz en E^* se encuentran incluidas ambas afirmaciones.

La proposición anterior nos permite describir el dual de un espacio E de dimensión finita en términos de una base de este, pues si, por ejemplo, $x \in E$ se expresa como $x = \sum_{i=1}^n x_i a_i$ y $y^* \in E^*$, entonces, llamando $y^*(a_i) = \alpha_i$, tenemos

$$y^*(x) = \sum_{i=1}^n y^*(a_i) a_i^*(x) = \sum_{i=1}^n \alpha_i x_i = \alpha_1 x_1 + \dots + \alpha_n x_n$$

lo cual generaliza el resultado que encontramos para K^n en el ejemplo 1 del presente epígrafe.

Es conveniente señalar que hemos hallado que todo espacio de dimensión finita es isomorfo a su dual, e incluso una forma de hallar isomorfismo entre ellos: asignando a cada base su base dual. Pero estos isomorfismos dependen de las bases halladas y, por tanto, no podemos independizar los resultados obtenidos al usar dichos isomorfismos, de las bases escogidas en el espacio.

En caso de que el espacio E no sea de dimensión finita, de todas formas es posible definir para una base $(a_i)_{i \in I}$ de E dada, la familia de elementos del dual $(a_i^*)_{i \in I}$ mediante la misma fórmula usada en el caso finito, esto es:

$$\text{Para todo } i, a_i^*(a_j) = \begin{cases} 1, & \text{si } i=j \\ 0, & \text{si } i \neq j \end{cases}$$

A esta familia se le llama *familia de formas coordenadas asociada a la base* $(a_i)_{i \in I}$, ya que al igual que en el caso finito, dado $x \in E$ tal que $x = \sum_{i \in I} \alpha_i a_i$, con $\alpha_i = 0$, salvo para un número finito, $a_i^*(x) = \alpha_i$.

La demostración de que la familia $(a_i^*)_{i \in I}$ es linealmente independiente no difiere (salvo por las consideraciones particulares de la independencia lineal en el caso infinito) de la realizada en el caso finito.

Sin embargo, la familia $(a_i^*)_{i \in I}$ no es generatriz en el caso infinito, ya que si consideramos, por ejemplo, la forma lineal $y_0^* \in E^*$ dada por $y_0^*(a_i) = 1$ para todo i , obtendremos $y_0^* = \sum_{i \in I} y_0^*(a_i) a_i^*$, donde todo $y_0^*(a_i) \neq 0$, lo cual es imposible. Esto es, no es posible expresar y_0^* como combinación lineal de los (a_i^*) . Esto nos da un indicio de que, según señalamos anteriormente, el dual de un espacio de dimensión infinita es complejo, e intuitivamente corroboramos la idea expresada de que resulta, en general, "demasiado grande".

Ejemplos

1. En $\mathbb{R}^3$ la base dual de la base de $\mathbb{R}^3$:

$a_1 = (1, -1, 3)$, $a_2 = (0, 1, -1)$, $a_3 = (0, 3, -2)$ se halla como sigue:

$(a_i^*)_{1 \leq i \leq 3}$ está dada por

$$a_i^*(a_j) = \begin{cases} 1, & \text{si } i=j \\ 0, & \text{si } i \neq j \end{cases}$$

Además, sabemos que dada la forma general de las formas lineales sobre $\mathbb{R}^3$, podemos escribir:

$$a_1^*(x_1, x_2, x_3) = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$$

$$a_2^*(x_1, x_2, x_3) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$a_3^*(x_1, x_2, x_3) = \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3$$

para todo $(x_1, x_2, x_3) \in \mathbb{R}^3$, en particular para a_1, a_2, a_3 .

La búsqueda de a_1^*, a_2^*, a_3^* puede reducirse, por tanto, a la solución de los sistemas de ecuaciones lineales siguientes:

$$\begin{cases} a_1^*(a_1) = a_1^*(1, -1, 3) = \alpha_1 - \alpha_2 + 3\alpha_3 = 1 \\ a_1^*(a_2) = a_1^*(0, 1, -1) = \alpha_2 - \alpha_1 = 0 \\ a_1^*(a_3) = a_1^*(0, 3, -2) = 3\alpha_2 - 2\alpha_1 = 0 \end{cases}$$

$$\begin{cases} a_2^*(a_1) = a_2^*(1, -1, 3) = \beta_1 - \beta_2 + 3\beta_3 = 0 \\ a_2^*(a_2) = a_2^*(0, 1, -1) = \beta_2 - \beta_1 = 1 \\ a_2^*(a_3) = a_2^*(0, 3, -2) = 3\beta_2 - 2\beta_1 = 0 \end{cases}$$

$$\begin{cases} a_3^*(a_1) = a_3^*(1, -1, 3) = \gamma_1 - \gamma_2 + 3\gamma_3 = 0 \\ a_3^*(a_2) = a_3^*(0, 1, -1) = \gamma_2 - \gamma_1 = 0 \\ a_3^*(a_3) = a_3^*(0, 3, -2) = 3\gamma_2 - 2\gamma_1 = 1 \end{cases}$$

Resolviendo estos sistemas obtenemos:

$$a_1^*(x_1, x_2, x_3) = x_1$$

$$a_2^*(x_1, x_2, x_3) = 7x_1 - 2x_2 - 3x_3$$

$$a_3^*(x_1, x_2, x_3) = -2x_1 + x_2 + x_3$$

2. Un ejemplo importante es hallar la base dual de la base canónica de $\mathbb{R}^3$. Recordemos que el vector e_i de dicha base es $e_i = (0, \dots, 1, \dots, 0)$.

La base dual está dada por

$$e_i^*(e_j) = \begin{cases} 1, & \text{si } i=j \\ 0, & \text{si } i \neq j \end{cases}$$

donde $e_i^*(x_1, x_2, \dots, x_n) = a_{i1}x_1 + \dots + a_{in}x_n$, lo que dará lugar a los sistemas de ecuaciones siguientes:

$$(1) \begin{cases} a_{11} = 1 \\ \vdots \\ a_{j1} = 0 \\ \vdots \\ a_{n1} = 0 \end{cases} \quad (i) \begin{cases} a_{1i} = 0 \\ \vdots \\ a_{ii} = 1 \\ \vdots \\ a_{ni} = 0 \end{cases} \quad (n) \begin{cases} a_{1n} = 0 \\ \vdots \\ a_{jn} = 0 \\ \vdots \\ a_{nn} = 1 \end{cases}$$

De aquí obtenemos la respuesta:

$$e_i^*(x_1, x_2, \dots, x_n) = x_i$$

Hemos resuelto el caso general. Si el estudiante al leerlo se encuentra un poco "perdido" por causa de los subíndices, le recomendamos que resuelva los casos particulares de las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$, siguiendo la vía dada en el ejemplo 1.

Por supuesto, hemos trabajado más de lo necesario con toda intención, pues si recordamos que la forma coordenada i -ésima le asigna a cada vector su i -ésima coordenada, hubiéramos podido dar el resultado anterior sin realizar cálculos.

3. Planteemos el problema inverso: dada una base buscar su base dual, y observemos atentamente la solución, pues la interpretaremos posteriormente.

Sea $E = \{a_0 + a_1x + a_2x^2 : a_i \in \mathbb{R}\}$, esto es, el espacio de los polinomios de grado ≤ 2 , que sabemos tiene dimensión 3, y sean p_0^*, p_1^*, p_2^* las formas lineales definidas sobre E por

$$p_0^*(p) = p(0)$$

$$p_1^*(p) = p(1)$$

$$p_2^*(p) = p(2)$$

p_0^*, p_1^* y p_2^* son linealmente independientes en E^* , pues si suponemos $a_0p_0^* + a_1p_1^* + a_2p_2^* = 0$, es decir, $(a_0p_0^* + a_1p_1^* + a_2p_2^*)(p) = 0 \forall p \in E$, y tomamos para p los valores 1, X , X^2 , obtenemos el sistema:

$$\begin{cases} a_0 + a_1 + a_2 = 0 \\ a_1 + 2a_2 = 0 \\ a_1 + 4a_2 = 0 \end{cases}$$

El determinante del sistema es

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

Luego el sistema tiene solo la solución trivial $a_0 = a_1 = a_2 = 0$. Como sabemos que $\dim E^* = 3$, podemos afirmar que p_0^*, p_1^*, p_2^* forman una base de E^* .

¿Existirá una base $\{p_0, p_1, p_2\}$ de E tal que $\{p_0^*, p_1^*, p_2^*\}$ sea su dual? Tratemos de resolver este problema.

Si dicha base existe, será de la forma

$$p_0 = a_0 + a_1X + a_2X^2$$

$$p_1 = b_0 + b_1X + b_2X^2$$

$$p_2 = c_0 + c_1X + c_2X^2$$

y deberá cumplir

$$p_i^*(p_j) = \begin{cases} 1, & \text{si } i=j \\ 0, & \text{si } i \neq j \end{cases}$$

Luego podemos plantear los sistemas de ecuaciones siguientes:

$$\begin{cases} p_0^*(p_0) = p_0(0) = a_0 & = 1 \\ p_1^*(p_0) = p_0(1) = a_0 + a_1 + a_2 & = 0 \\ p_2^*(p_0) = p_0(2) = a_0 + 2a_1 + 4a_2 & = 0 \end{cases}$$

$$\begin{cases} p_0^*(p_1) = p_1(0) = b_0 & = 0 \\ p_1^*(p_1) = p_1(1) = b_0 + b_1 + b_2 & = 1 \\ p_2^*(p_1) = p_1(2) = b_0 + 2b_1 + 4b_2 & = 0 \end{cases}$$

$$\begin{cases} p_0^*(p_2) = p_2(0) = c_0 & = 0 \\ p_1^*(p_2) = p_2(1) = c_0 + c_1 + c_2 & = 0 \\ p_2^*(p_2) = p_2(2) = c_0 + 2c_1 + 4c_2 & = 1 \end{cases}$$

Luego de resolver los sistemas, obtenemos:

$$p_0 = \frac{X^2}{2} - \frac{3X}{2} + 1$$

$$p_1 = -X^2 + 2X$$

$$p_2 = \frac{X^2}{2} - \frac{X}{2}$$

Queremos hacer notar al estudiante que aunque los ejemplos 1 y 3 responden a la solución de problemas inversos entre sí, el método de solución es en esencia el mismo, ya que solo aparecen intercambiados los papeles de la base del espacio y la del dual.

En los ejemplos precedentes hemos resuelto para un caso concreto el problema de hallar, a partir de una base del espacio dual de un espacio de dimensión finita, una base del espacio de la cual la dada fuese dual. Ahora podemos plantearnos un problema más amplio, esto es: dada una base cualquiera del dual de un espacio, ¿existirá una base de dicho espacio de la cual la dada sea dual?

En la solución de este problema en los ejemplos dados surgió una analogía en los métodos que pudiera interpretarse como explicaremos a continuación.

Si en el ejemplo 3 consideramos E^* como espacio inicial y observamos los sistemas de ecuaciones planteados, podemos decir que en la práctica hemos hallado una base dual de la dada, pero esto hay que explicarlo teóricamente, pues si bien E^* es un espacio vectorial y como tal tiene sentido considerar su dual, nada nos permite afirmar que este sea E .

Sin embargo, las consideraciones que daremos aquí, no del todo formales, nos permitirán comprender la importancia de estudiar el espacio dual

de un espacio dual E^* , al que llamaremos espacio *bidual* o *doble dual*, y sus relaciones con el propio E^* y con E .

DEFINICIÓN 2.5.4

Dado un espacio vectorial E , se llama espacio *bidual* o *doble dual* de E , y se denota por E^{**} , el espacio dual de E^* .

Es inmediato de los resultados obtenidos para el espacio dual de un espacio vectorial de dimensión finita, que $\dim E^{**} = \dim E^* = \dim E$, y de aquí E , E^* y E^{**} son isomorfos. Pero según vimos, el tipo de isomorfismo que encontramos entre un espacio y su dual dependía de las bases seleccionadas en dichos espacios; por tanto, será mucho mejor obtener información de la correspondencia existente entre E y E^{**} sin el "lastre" de las bases, lo cual nos permitirá "identificar" los elementos de E y E^{**} y formalizar nuestras consideraciones sobre el ejemplo 3.

Buscaremos, pues, un isomorfismo entre E y E^{**} que dependa solamente de los propios espacios. Es decir, debemos encontrar un modo de hacer corresponder a cada elemento x de E un elemento de E^{**} , o sea, una aplicación de E en E^{**} .

Si para cada $x \in E$ fijo consideramos $y^*(x)$ con $y^* \in E^*$, entonces $y^*(x)$ será un elemento de K que dependerá de y^* . Esto nos da una idea de cómo encontrar un elemento del bidual y que esté ligado a x .

Sea $x \in E$ y consideremos $\varphi_x : E^* \rightarrow K$ dada por $\varphi_x(y^*) = y^*(x)$. Probemos que $\varphi_x \in E^{**}$, para lo que bastará probar que φ_x es lineal.

$$\varphi_x(y^* + z^*) = (y^* + z^*)(x) = y^*(x) + z^*(x) = \varphi_x(y^*) + \varphi_x(z^*) \quad \forall y^*, z^* \in E^*$$

$$\varphi_x(\alpha y^*) = (\alpha y^*)(x) = \alpha y^*(x) = \alpha \varphi_x(y^*) \quad \forall \alpha \in K, y^* \in E^*$$

Ya estamos en condiciones de definir una correspondencia entre E y E^{**} . Definamos

$$\chi : E \rightarrow E^{**} \text{ por } \chi(x) = \varphi_x.$$

$\chi(x) \in E^{**}$ está caracterizado por

$$\chi(x)(y^*) = \varphi_x(y^*) = y^*(x) \quad \forall x \in E \quad \forall y^* \in E^*$$

Evidentemente, la definición de χ no depende de la selección de ninguna base en el espacio E , sino de la propia estructura de espacio vectorial de E , por lo que a esta aplicación se le suele llamar *aplicación canónica de E sobre E^{**}* .

Comprobemos que χ es un isomorfismo de E sobre E^{**} en caso de que E sea un espacio de dimensión finita.

PROPOSICIÓN 2.5.2

Sea E un espacio de dimensión finita. La aplicación $\chi : E \rightarrow E^{**}$ caracterizada por $\chi(x)(y^*) = y^*(x) \quad \forall x \in E, \forall y^* \in E^*$, es un isomorfismo de E sobre E^{**} .

Demostración

Debemos probar que χ es lineal, inyectiva y sobreyectiva.

Para probar que χ es lineal tenemos que verificar que

$$\chi(ax + \beta y) = a\chi(x) + \beta\chi(y) \quad \forall a, \beta \in K \quad \forall x, y \in E,$$

o sea,

$$\chi(ax + \beta y)(y^*) = a\chi(x)(y^*) + \beta\chi(y)(y^*) \quad \forall y^* \in E^*$$

$$\chi(ax + \beta y)(y^*) = y^*(ax + \beta y) = ay^*(x) + \beta y^*(y) = a\chi(x)(y^*) + \beta\chi(y)(y^*).$$

con esto queda probado que χ es lineal.

Probemos que χ es inyectiva, demostrando que $\text{Ker } \chi = \{0\}$.

Sea $x \in \text{Ker } \chi \Leftrightarrow \chi(x) = 0 \Leftrightarrow \chi(x)(y^*) = y^*(x) = 0 \quad \forall y^* \in E^*$.

Si $x = \sum_{i=1}^n x_i a_i$ donde (a_i) es una base de E , entonces $y^*(x) = 0 \quad \forall y^* \in E^*$.

En particular,

$$a_i^*(x) = x_i = 0, \quad i = 1, \dots, n. \quad \text{luego } x = 0.$$

Como sabemos que $\dim E = \dim E^*$, del hecho de que χ sea inyectiva se infiere que χ es sobreyectiva, y así queda demostrada la proposición.

El isomorfismo χ en su carácter de canónico antes señalado, permite "identificar" E y E^* y escribir $E = E^*$. Y como todo $x^{**} \in E^{**}$ puede escribirse como $x^{**} = \chi(x)$ podemos sustituir $x^{**}(y^*)$ por $y^*(x)$ ya que $\chi(x)(y^*) = y^*(x)$.

Esta identificación permite formalizar lo que señalamos en el ejemplo 3, esto es, que al identificar E con E^* el problema de buscar una base de E dual de una conocida de E^* puede considerarse equivalente al de buscar, dada una base de E^* , su base dual en E^{**} .

El problema anterior puede enunciarse en términos más formales mediante el lenguaje de los isomorfismos. Esto lo haremos en los ejercicios.

¿Qué ocurrirá con la correspondencia entre E y E^{**} cuando E es de dimensión infinita? El estudiante puede analizar que en este caso la linealidad y la inyectividad de χ pueden obtenerse por la misma vía que en el caso finito. Sin embargo, no ocurre así con la sobreyectividad, lo cual podemos intuir del hecho de que la familia de formas coordenadas asociada a una base $(a_i)_{i \in I}$ no genera a E^* , ya que si tomamos el subespacio W de E^* dado por $W = \langle (a_i^*)_{i \in I} \rangle$ y la forma lineal y_0^* definida por $y_0^*(a_i) = 1$ para todo i , se cumple que $y_0^* \notin W$. Es posible demostrar, entonces, que tomando una forma lineal y_0^* sobre E^* que se anule en W y tal que $y_0^*(y_0^*) = 1$, dicha forma no pertenece a $\text{Im } \chi$; luego χ no es sobreyectiva.

Ortogonalidad y dualidad

Ahora estudiaremos un problema que coincidirá con otro ya conocido en el caso de las formas lineales en un espacio con producto escalar. Exactamente, dado un subconjunto W de un espacio vectorial E , analizaremos el

subconjunto de E^* de las formas lineales que se anulan en todo elemento de W . Mediante este análisis hallaremos relaciones interesantes y útiles entre dicho subconjunto de E^* y W . En particular, aplicaremos los resultados obtenidos al estudio de los sistemas de ecuaciones lineales homogéneas.

DEFINICIÓN 2.5.5

Sean E un espacio vectorial sobre K y W un subconjunto de E . Entonces se llama *anulador* u *ortogonal* de W en E^* y se denota por W^0 , el conjunto

$$W^0 = \{y^* \in E^* : y^*(x) = 0 \quad \forall x \in W\}.$$

No tenemos elementos todavía para justificar la denominación de *ortogonal* dado a W^0 , pero veremos que las propiedades de W^0 son similares a las del ortogonal de un subconjunto en un espacio con producto escalar. W^0 posee las propiedades siguientes:

- (1) W^0 es un subespacio vectorial de E^* .
- (2) $(W^0)^0 = W$.
- (3) $\{0\}^0 = E^*$.
- (4) $E^0 = \{0\} \subset E^*$.

Estas propiedades se infieren inmediatamente de la definición de W^0 , lo que dejamos como ejercicio al estudiante.

Ejemplos

1. Sea en $\mathbb{R}^4$, $W = \{(1, 2, -3, 4), (0, 1, 4, -1)\}$. Hallemos W^0 .

$$W^0 = \{y^* \in \mathbb{R}^4 : y^*(x) = 0 \quad \forall x \in W\}$$

De la propiedad (2) resulta que los elementos de W^0 están dados por las formas lineales y^* de $\mathbb{R}^4$ tales que *

$$y^*(1, 2, -3, 4) = 0$$

$$y^*(0, 1, 4, -1) = 0.$$

Pero sabemos que si $y^* \in \mathbb{R}^4$, y^* tiene la forma

$$y^*(x_1, x_2, x_3, x_4) = a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4.$$

Nuestro problema se reduce, pues, a resolver el siguiente sistema homogéneo de ecuaciones lineales.

$$\begin{cases} a_1 + 2a_2 - 3a_3 + 4a_4 = 0 \\ a_2 + 4a_3 - a_4 = 0 \end{cases}$$

y $W^0 = \langle \{y_1^*, y_2^*\} \rangle$ donde

$$y_1^*(x_1, x_2, x_3, x_4) = 11x_1 - 4x_2 + x_3$$

$$y_2^*(x_1, x_2, x_3, x_4) = -6x_1 + x_2 + x_4.$$

Este método es posible aplicarlo a la búsqueda del ortogonal o anulador de cualquier subespacio W de K^n , ya que si $W = \langle \{a_1, \dots, a_m\} \rangle$ donde

$a_i = (a_{i1}, \dots, a_{in})$, $1 \leq i \leq m$, como la expresión general de una forma lineal sobre K^n es

$$y^*(x_1, x_2, \dots, x_n) = a_1 x_1 + \dots + a_n x_n,$$

entonces para que y^* pertenezca a W^0 será necesario y suficiente que $a_1, a_2, \dots, a_n$ satisfagan

$$a_1 a_{1i} + \dots + a_n a_{ni} = 0, \quad i = 1, \dots, m,$$

y $a_1, a_2, \dots, a_n$ estarán dados por las soluciones del sistema de ecuaciones lineales homogéneo.

En la práctica para buscar el anulador u ortogonal de un subespacio W de un espacio vectorial E cualquiera, existe la dificultad de que no se conoce la expresión general de las formas lineales sobre E . Si E es de dimensión finita n sobre K , vimos que mediante la proposición 2.5.1 es posible describir el dual usando una base del espacio, lo que nos permite trabajar en la práctica como si se tratara de K^n .

2. Hallemos el ortogonal del subespacio de $M_2(\mathbb{R})$ dado por

$$W = \left\langle \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle$$

Bastará tomar la base usual en dicho espacio y así podremos identificar $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ y $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ con sus vectores de coordenadas $(1, 1, 1, 0)$ y $(1, 0, 0, 1)$.

Explicaremos ahora una nueva vía de solución del ejemplo 1, que si bien en un caso práctico tiene un mayor volumen de cálculo, nos indicará el camino para obtener un resultado más general.

Teníamos que $W = \langle b_1 = (1, 2, -3, 4), b_2 = (0, 1, 4, -1) \rangle$ y queremos hallar W^0 . Dada la definición de base dual, una vía lógica de solución es la siguiente.

Completemos b_1, b_2 hasta formar una base de $\mathbb{R}^4$ de la forma más sencilla posible. Así, sea

$$b_1 = (1, 2, -3, 4)$$

$$b_2 = (0, 1, 4, -1)$$

$$b_3 = (0, 0, 1, 0)$$

$$b_4 = (0, 0, 0, 1).$$

Hallemos en $\mathbb{R}^4$ la base dual de $(b_i)_{1 \leq i \leq 4}$ que está dada, como sabemos, por

$$h_i(b_j) = \begin{cases} 1, & \text{para } i=j \\ 0, & \text{para } i \neq j \end{cases}$$

Suponiendo que

$$b_1^*(x_1, x_2, x_3, x_4) = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4$$

$$b_2^*(x_1, x_2, x_3, x_4) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

$$b_3^*(x_1, x_2, x_3, x_4) = \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3 + \gamma_4 x_4$$

$$b_4^*(x_1, x_2, x_3, x_4) = \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 + \delta_4 x_4$$

obtenemos los siguientes sistemas de ecuaciones lineales:

$$(1) \quad \begin{aligned} \alpha_1 + 2\alpha_2 - 3\alpha_3 + 4\alpha_4 &= 1 \\ \alpha_2 + 4\alpha_3 - \alpha_4 &= 0 \\ \alpha_3 &= 0 \\ \alpha_4 &= 0 \end{aligned}$$

$$(3) \quad \begin{aligned} \gamma_1 + 2\gamma_2 - 3\gamma_3 + 4\gamma_4 &= 0 \\ \gamma_2 + 4\gamma_3 - \gamma_4 &= 0 \\ \gamma_3 &= 1 \\ \gamma_4 &= 0 \end{aligned}$$

$$(2) \quad \begin{aligned} \beta_1 + 2\beta_2 - 3\beta_3 + 4\beta_4 &= 0 \\ \beta_2 + 4\beta_3 - \beta_4 &= 1 \\ \beta_3 &= 0 \\ \beta_4 &= 0 \end{aligned}$$

$$(4) \quad \begin{aligned} \delta_1 + 2\delta_2 - 3\delta_3 + 4\delta_4 &= 0 \\ \delta_2 + 4\delta_3 - \delta_4 &= 0 \\ \delta_3 &= 0 \\ \delta_4 &= 1 \end{aligned}$$

La solución es:

$$b_1^*(x_1, x_2, x_3, x_4) = x_1$$

$$b_2^*(x_1, x_2, x_3, x_4) = -2x_1 + x_2$$

$$b_3^*(x_1, x_2, x_3, x_4) = 11x_1 - 4x_2 + x_3$$

$$b_4^*(x_1, x_2, x_3, x_4) = -6x_1 + x_2 + x_4$$

No es difícil verificar que $W^0 = \langle b_3^*, b_4^* \rangle$, con lo que encontramos de nuevo la respuesta hallada por la primera vía.

Por cualquiera de las dos vías de solución del ejemplo 1, hemos obtenido que la dimensión de W^0 es 2, lo que nos permite, observando que la dimensión de W también es 2, escribir en este caso:

$$\dim W + \dim W^0 = \dim \mathbb{R}^4.$$

Por otra parte, si recordamos las propiedades del ortogonal de un subespacio en un espacio con producto escalar y las analogías que hemos encontrado entre el concepto anulador ortogonal en el dual y el concepto ortogonal de un subespacio de un espacio con producto escalar, no es difícil suponer que se cumplirá la siguiente proposición para los espacios de dimensión finita, en cuya demostración nos servirá de guía la segunda vía de solución del ejemplo 1.

PROPOSICIÓN 2.5.3

Sean E un espacio vectorial de dimensión finita n sobre K y W un subespacio de E . Entonces se cumple: $\dim W + \dim W^0 = \dim E$.

Demostración

Sea $\{a_1, a_2, \dots, a_n\}$ una base de E tal que $\{a_1, a_2, \dots, a_m\}$ es una base de W .

Tomemos la base $\{a_1^*, a_2^*, \dots, a_n^*\}$ dual de la base $\{a_i\}_{1 \leq i \leq n}$ y probemos que $\{a_{m+1}^*, \dots, a_n^*\}$ constituye una base de W^0 , con lo que quedaría probada la proposición.

$a_{m+1}^*, \dots, a_n^*$ pertenecen a W^0 ya que para todo $x \in W^0$, $x = a_1 a_1 + \dots + a_m a_m$, y evidentemente $a_i^*(x) = 0$ para $i = m+1, \dots, n$.

Como $a_{m+1}^*, \dots, a_n^*$ son linealmente independientes, queda solo por probar que generan a W^0 .

$y^* \in W^0$. Debemos probar que existen $\alpha_{m-1}, \dots, \alpha_n \in K$ tales que

$$y^* = \sum_{k=m+1}^n a_k a_k^*. \text{ Pero}$$

$$y^* = \sum_{i=1}^n y^*(a_i) a_i^* \quad y^* \in W^n \Rightarrow y^*(a_1) = \dots = y^*(a_m) = 0.$$

Luego

$$y^* \in W^0 \Rightarrow y^* = \sum_{k=m+1}^n y^*(a_k) a_k^*,$$

con lo que la proposición queda demostrada, ya que $\dim W^0 = n - m$.

Tomemos ahora en $\mathbb{R}^4$ el subespacio $\langle \{y_1^*, y_2^*\} \rangle$, donde

$$y_1^*(x_1, x_2, x_3, x_4) = 11x_1 - 4x_2 + x_3$$

$$y_2^*(x_1, x_2, x_3, x_4) = -6x_1 + x_2 + x_4,$$

o sea, $\langle \{y_1^*, y_2^*\} \rangle = W^n$, donde $W = \langle (1, 2, -3, 4), (0, 1, 4, -1) \rangle$ es el subespacio de $\mathbb{R}^4$ del ejemplo que presentamos antes de la proposición 2.4.3.

Como R^4 es un espacio vectorial, tiene sentido plantear la búsqueda de $(W^0)^0$ que será un subespacio de R^{4**} . Pero vimos que $(W^0)^0$ se puede identificar, mediante la aplicación χ de E en E^{**} , con un subespacio de R^4 , y plantear en lugar del conjunto

$$\{x^{**} \in \mathbb{R}^{4**} : x^{**}(y^*) = 0 \quad \forall y^* \in W^0\}.$$

el conjunto

$$\{x \in \mathbb{R}^4 : y^*(x) = 0 \quad \forall y^* \in W^0\}.$$

sustituyendo x^{**} por su preimagen en $\mathbb{R}^1$ mediante la aplicación χ .

Asi.

$$(W^0)^0 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : y_1^*(x_1, x_2, x_3, x_4) = 0, y_2^*(x_1, x_2, x_3, x_4) = 0\}$$

lo que conduce al sistema de ecuaciones lineales homogéneo siguiente:

$$\begin{cases} 11x_1 - 4x_2 + x_3 = 0 \\ -6x_1 + x_2 + x_4 = 0 \end{cases}$$

Al resolver este sistema encontramos que el espacio solución, esto es, $(W^n)^0$ está dado por

$$(W^{\eta})^0 = \langle \{(0, 1, 4, -1), (1, 2, -3, 4)\} \rangle = W.$$

La respuesta es muy interesante pero no inesperada, dados los otros resultados que hemos encontrado en el espacio dual. Observemos que también el método de trabajo *es el mismo* en la búsqueda de W^0 y $(W^0)^0$, ya que solamente se han intercambiado los papeles de vectores y formas lineales. Este resultado no es casual, como se demuestra en la proposición siguiente:

PROPOSICIÓN 2.5.4

En un espacio vectorial E de dimensión finita n sobre K , para todo subespacio W de E se cumple que $(W^0)^0 = W$.

Demostración:

Tomemos como guía los razonamientos del ejemplo anterior, esto es:

$$(W^0)^0 = \{x^{**} \in E^{**} : x^{**}(Y^*) = 0 \quad \forall y^* \in W^0\},$$

y como $x^{**} = \chi(x)$ con $x \in W$, entonces

$$(W^0)^0 = \{\chi(x) : \chi(x)(y^*) = y^*(x) = 0 \ \forall y^* \in W^0\} \\ = \{x \in W : y^*(x) = 0 \ \forall y^* \in W^0\}.$$

Ahora bien, de la definición de $(W^0)^0$ resulta que

$$W \in (W^0)^0 \text{ y}$$

$$\dim (W^0)^0 = \dim E^{**} - \dim W^0 = \dim E - \dim W^0 = \dim W.$$

Luego $W = (W^0)^0$, como queríamos demostrar.

Quando el espacio vectorial es de dimensión infinita, no es posible identificar $(W^0)^0$ con un subconjunto de E . Se define, sin embargo, el subconjunto $(W^0)' = \{x \in E : y^*(x) = 0 \ \forall y^* \in W^0\}$ y se puede probar que $(W^0)' = W$ para todo subespacio de E .

Dualidad y sistemas de ecuaciones lineales homogéneos

La búsqueda del ortogonal o anulador de un subespacio W de E y la del ortogonal de dicho ortogonal, que de hecho nos mostró cómo buscar el ortogonal de un subespacio de E^* , puso en evidencia la relación entre formas lineales y sistemas de ecuaciones lineales homogéneos, lo que permite encontrar el resultado sobre la dimensión del subespacio solución de dichos sistemas mediante las formas lineales.

En efecto, sea el sistema

$$S \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + \dots + a_{2n}x_n = 0 \\ \dots\dots\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$$

Cada miembro izquierdo de las ecuaciones puede interpretarse como la imagen por una forma lineal de un vector $(x_1, x_2, \dots, x_n)$ de K^n . Así,

$$\alpha_{i1}x_1 + \dots + \alpha_{in}x_n = f_i(x) \quad 1 \leq i \leq m,$$

y hallar la solución de S sería equivalente a hallar el subconjunto de K^n dado por

$$\{(x_1, \dots, x_n) \in K^n : f_i(x_1, \dots, x_n) = 0, i = 1, \dots, m\},$$

o sea, W° donde W es el subespacio de $(K^n)^*$ generado por $\{f_1, \dots, f_m\}$.

La dimensión de este subespacio, llamémoslo S_0 , está dada por el rango del sistema de vectores de E^* , $\{f_1, f_2, \dots, f_m\}$, y como para cada f_i , $\alpha_{i1}, \dots, \alpha_{in}$ representan las coordenadas de f_i en la base e_i^* , dual de la canónica de K^n , el rango de dicho sistema no es más que el rango de la matriz

$$\begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

que es la transpuesta de lo que conocemos como matriz del sistema. En este caso,

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix}$$

Pero para toda matriz, $\text{rg } A = \text{rg } A^t$ y $\text{rg } \{f_1, \dots, f_m\} = \text{rg } A = \text{rg } A^t$. Luego $\dim S_0 = \dim E - \text{rg } A = n - \text{rg } A$.

y encontramos de nuevo por esta vía el conocido resultado sobre la dimensión del espacio solución de un sistema de ecuaciones lineales homogéneo.

Ejercicios

1. Encuentre la base dual de las siguientes bases de $\mathbb{R}^3$:

a) $a_1 = (1, -2, 3)$, $a_2 = (1, -1, 1)$, $a_3 = (2, -4, 7)$

b) $b_1 = (1, 0, -1)$, $b_2 = (1, 1, 1)$, $b_3 = (2, 2, 0)$

2. Sean E el espacio de los polinomios de grado ≤ 2 y p_1^* , p_2^* , p_3^* formas lineales sobre E . Pruebe en cada caso que p_1^* , p_2^* , p_3^* forman una base de E^* y halle una base de E de la cual la dada sea dual.

a) $p_1^*(p) = \int_0^1 p(t) dt$, $p_2^*(p) = p'(1)$, $p_3^*(p) = p(0)$

b) $p_1^*(p) = \int_0^1 p(t) dt$, $p_2^*(p) = \int_0^2 p(t) dt$, $p_3^*(p) = \int_0^{-1} p(t) dt$

3. Sean E un espacio vectorial sobre K y y_1^* , $y_2^* \in E^*$, y suponga que $y_1^*(x) = 0 \Rightarrow y_2^*(x) = 0$ para todo $x \in E$. Pruebe que $y_2^* = k y_1^*$ para algún $k \in K$.

4. Sea E un espacio vectorial sobre K y W un subespacio de E .

a) Pruebe que para cada forma lineal y^* en W° existe una forma lineal z^* en E^* tal que $y^*(x) = z^*(x) \forall x \in W$.

b) Sea $x_0 \in E : x_0 \notin W$. Pruebe que existe una forma lineal y^* sobre E , no nula, tal que $\text{Ker } y^* \supset W$ y $y^*(x_0) = 1$.

Sugerencia: Considere Kx_0 y pruebe que existe un hiperplano H de E tal que $H \supset W$.

5. Sean E un espacio vectorial sobre $\mathbb{R}$ y y_1^* , $y_2^* \in E^*$, y suponga que $\phi : E \rightarrow \mathbb{R}$ dada por $\phi(x) = y_1^*(x) \cdot y_2^*(x)$ es una forma lineal sobre E . Pruebe que $y_1^* = 0$ o $y_2^* = 0$.

6. Sean m y n números enteros positivos y $y_1^*, \dots, y_m^*$ formas lineales sobre K^n . Definamos para $x \in K^n$

$$T(x) = (y_1^*(x), \dots, y_m^*(x)).$$

Pruebe que T es una aplicación lineal de K^n en K^m y que cada aplicación lineal de K^n en K^m es de la forma anterior para algún $y_1^*, \dots, y_m^*$.

7. Si E es un espacio vectorial de dimensión finita sobre K y x_1, x_2 son vectores de E tales que $x_1 \neq x_2$, pruebe que existe $y^* \in E^*$ tal que $y^*(x_1) \neq y^*(x_2)$.

8. Sea E un espacio vectorial de dimensión finita sobre K . Pruebe que cada base $(a_i)^*$ de E^* es la base dual de alguna base de E .

Sugerencia: Utilice el isomorfismo χ para buscar las preimágenes en E de la base dual de (a_i^*) en E^{**} y pruebe que dicha base es la buscada.

9. Sea n un entero positivo y W el conjunto de los vectores $(x_1, \dots, x_n)$ en K^n tales que $x_1 + \dots + x_n = 0$.

a) Pruebe que W° está formado por las formas lineales sobre K^n de la

$$\text{forma } y^*(x_1, \dots, x_n) = C \sum_{j=1}^n x_j, \quad C \in K.$$

b) Pruebe que el espacio dual W^* de W puede ser naturalmente identificado con las funcionales lineales $y^*(x_1, \dots, x_n) = C_1 x_1 + \dots + C_n x_n$ sobre K^n que satisfacen $C_1 + C_2 + \dots + C_n = 0$.

10. Sean E un espacio vectorial de dimensión finita sobre E , $(a_i)_{1 \leq i \leq n}$ y $(b_j)_{1 \leq j \leq n}$ bases de E , y $(a_i^*)_{1 \leq i \leq n}$, $(b_j^*)_{1 \leq j \leq n}$ sus bases duales respectivas. Suponga que P es la matriz de cambio de base de $(a_i)_{1 \leq i \leq n}$ a $(b_j)_{1 \leq j \leq n}$. Pruebe que P^{-1} es la matriz de cambio de base de $(a_i^*)_{1 \leq i \leq n}$ a $(b_j^*)_{1 \leq j \leq n}$.

11. a) Pruebe que si S_1 y S_2 son subconjuntos de un espacio vectorial E sobre K tales que $S_1 \subset S_2$, entonces $S_2^\circ \subset S_1^\circ$.

b) Pruebe que $(S_1^\circ)^\circ = \langle S_1 \rangle$ para todo $S_1 \subset E$.

12. Sean W_1 y W_2 subespacios de un espacio vectorial E sobre K . Pruebe que $(W_1 + W_2)^0 = W_1^0 \cap W_2^0$.

13. Sea W el subespacio de $\mathbb{R}^4$ dado por $W = \langle (1, 2, -3, 4), (1, 3, -2, 6), (1, 4, -1, 8) \rangle$. Halle W^0 .

14. Sean W_1 y W_2 subespacios de un espacio vectorial de dimensión finita E sobre K . Pruebe que $(W_1 \cap W_2)^0 = W_1^0 + W_2^0$.
Sugerencia: Utilice el ejercicio 12.

15. Sea E un espacio vectorial sobre K y suponga que W_1 y W_2 son subespacios de E tales que $E = W_1 \oplus W_2$. Pruebe que $E^* = W_1^0 \oplus W_2^0$.

16. Sea E un espacio vectorial sobre $\mathbb{R}$.

Se llama *segmento* que une a los puntos x_1, x_2 de E a

$$\overline{x_1 x_2} = \{kx_1 + (1-k)x_2, 0 \leq k \leq 1\}.$$

Un subconjunto S de E se llama *convexo* si $x_1, x_2 \in S \Rightarrow \overline{x_1 x_2} \subset S$.

Sea $y^* \in E^*$. Pruebe que son convexos los subconjuntos:

$$S_1 = \{x \in E : y^*(x) > 0\}$$

$$S_2 = \{x \in E : y^*(x) = 0\}$$

$$S_3 = \{x \in E : y^*(x) < 0\}$$

Pruebe que si E es un espacio vectorial de dimensión finita sobre K , todo subespacio W de E es la intersección de un número finito de hiperplanos de E .

18. Sean $E = M_n(\mathbb{R})$ y $B = \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix}$. Sean

$W = \{A \in M_n(\mathbb{R}) : A \cdot B = 0\}$ y $y^* \in W^0$. Suponga que $y^*(I_2) = 0$ y

$$y^* \left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] = 3, \text{ y halle } y^*(B).$$

Sean E, F espacios vectoriales sobre K , y $f : E \rightarrow F$ una aplicación lineal.

a) Pruebe que la expresión $f(y^*) = y^* \circ f$ para $y^* \in F^*$ define una aplicación lineal de F^* en E^* , a la cual se le llama *transpuesta de f* .

b) $\text{Ker } f' = [\text{Im } f]^0$

c) Si E y F son de dimensión finita, entonces

$$\text{i) } \text{rg } f = \text{rg } f'$$

$$\text{ii) } \text{Im } f' = [\text{Ker } f]^0$$

iii) Si $A = M(f, B, B')$, entonces $M(f', B^*, B^*) = A$ (donde B^* y B'^* representan las bases duales de B y B' respectivamente).

20. Aplique el ejercicio 19 (inciso c, iii) para hallar las aplicaciones transpuestas de las siguientes aplicaciones de K^2 :

$$\text{a) } f(x_1, x_2) = (x_1, 0)$$

$$\text{b) } f(x_1, x_2) = (-x_2, x_1)$$

$$\text{c) } f(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$$

21. Sean E un espacio vectorial de dimensión finita sobre K , f un endomorfismo de E y $\alpha \in K$. Suponga que existe un vector no nulo x de E tal que $f(x) = \alpha x$. Pruebe que existe una forma lineal no nula y^* en E^* tal que $f(y^*) = \alpha y^*$.

22. Sea A una matriz de tamaño $m \times n$ con coeficientes reales. Pruebe que $A = 0$ si y solo si $AA^0 = 0$.

23. Sea E el espacio de los polinomios de grado $\leq n$ con coeficientes reales, y suponga $T : E \rightarrow E$ dado por $T(f) = f'$ (polinomio derivado). Halle $\text{Ker } T$.

24. Pruebe que si $f : E \rightarrow F$ y $g : F \rightarrow G$ son aplicaciones lineales entre los espacios vectoriales sobre K , E y F , F y G respectivamente, entonces, $(g \circ f)' = f' \circ g'$.

25. Sea $f : E \rightarrow F$ una aplicación lineal, $x \in F$. Pruebe que $[x \in \text{Im } f]$ o [existe $y^* \in E^*$ tal que $f(y^*) = 0$ y $y^*(x) = 1$].

26. Sea E un espacio de dimensión n sobre K . Pruebe que la aplicación $f \mapsto f'$ es un isomorfismo del espacio de los endomorfismos de E sobre el espacio de los endomorfismos de E^* .

27. Sea $E = M_n(\mathbb{R})$.

a) Sea $B \in M_n(\mathbb{R})$ fija. Pruebe que $f_B : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ dada por $f_B(A) = \text{tr } B \cdot A$ para $A \in M_n(\mathbb{R})$ es una forma lineal sobre $M_n(\mathbb{R})$.

b) Pruebe que todo $y^* \in (M_n(\mathbb{R}))^*$ es de la forma anterior.

c) Pruebe que la correspondencia $B \mapsto f_B$ es un isomorfismo de $M_n(\mathbb{R})$ sobre $(M_n(\mathbb{R}))^*$.

2.6 Formas lineales sobre un espacio con producto escalar. Endomorfismo adjunto

En el epígrafe anterior vimos que, aun en el caso de los espacios de dimensión finita, para los cuales se cumple que el espacio y su dual son isomorfos, no podemos encontrar un isomorfismo natural entre estos. En otras palabras, no podemos dar una descripción del espacio dual sin tomar una base del espacio a la cual referirnos.

Esta situación cambia en el caso de los espacios con producto escalar de dimensión finita, ya que en estos toda forma lineal puede expresarse en tér-

minos del producto escalar, lo cual nos permitirá encontrar una correspondencia natural entre el espacio y su dual.

En el ejercicio 27 del epígrafe 2.5 planteamos que en el espacio $M_n(\mathbb{R})$ toda forma lineal y^* sobre dicho espacio se puede expresar como $y^*(A) = \text{tr } B \cdot A$ donde B es una matriz fija. Recordemos que en el espacio $M_n(\mathbb{R})$ hemos estudiado la expresión $\text{tr } B \cdot A$ como un ejemplo de producto escalar, de modo que $y^*(A) = (A, B)$ donde B es una matriz fija.

Un ejemplo más trivial sería el de $\mathbb{R}^n$, en el cual toda forma lineal se expresa como $y^*(x_1, x_2, \dots, x_n) = a_1 x_1 + \dots + a_n x_n$ o lo que es lo mismo, $y^*(x_1, x_2, \dots, x_n) = (x, a)$ donde $x = (x_1, \dots, x_n)$ y $a = (a_1, \dots, a_n)$ para el producto escalar usual en $\mathbb{R}^n$.

Esta situación que encontramos en los ejemplos anteriores, puede demostrarse de forma más general mediante el teorema siguiente:

TEOREMA 2.6.1

En un espacio con producto escalar E de dimensión finita, a toda forma lineal y^* sobre E le corresponde un vector x_0 de E único, tal que $y^*(x) = (x, x_0)$ para todo $x \in E$.

Demostración

Hemos visto que un elemento fundamental en los espacios con producto escalar son las bases ortonormales.

Sea, pues, $(a_i)_{1 \leq i \leq n}$ una base ortonormal de E y tratemos de encontrar x_0 . Supongamos que $x_0 = \sum_{i=1}^n a_i$, y que $y^* \in E^*$; entonces debe ocurrir que $y^*(x) = (x, x_0)$ para todo x de E .

Ante todo señalemos que esta igualdad tiene sentido, pues a partir de las propiedades del producto escalar, se verifica fácilmente que dado $x_0 \in E$ fijo, la aplicación de E en $\mathbb{R}$ dada por $x \rightarrow (x, x_0)$ es una forma lineal sobre E . (Dejamos al estudiante que realice estos cálculos.) Llamemos a dicha aplicación f_{x_0} .

Una vez aclarado esto, estamos ante la situación de encontrar $x_0 \in E$ de forma tal que las formas lineales y^* y f_{x_0} coincidan, y para ello basta que coincidan en una base de E . De aquí que

$$y^*(a_j) = (a_j, x_0) = (a_j, \sum_{i=1}^n a_i) = \bar{a}_j$$

de donde

$$x_0 = \sum_{i=1}^n y^*(a_i) a_i$$

Hemos encontrado x_0 por una vía constructiva. No obstante, el estudiante puede comprobar que efectivamente la expresión de x_0 que acabamos de dar es la requerida.

Probemos ahora que un tal x_0 es único.

Supongamos que existe otro $y_0 \in E$ tal que $y^*(x) = (x, y_0) \forall x \in E$. Entonces

$$(x, x_0) = (x, y_0) \quad \forall x \in E$$

$$(x, x_0 - y_0) = 0 \quad \forall x \in E.$$

de lo que se deduce que $x_0 - y_0 = 0$, o sea, $x_0 = y_0$.

Así queda demostrado el teorema.

El teorema 2.6.1 no es válido en general cuando el espacio es de dimensión infinita; por ejemplo, en el espacio $C[X]$ de los polinomios complejos con el producto escalar

$$(p, q) = \int_0^1 p(t) \bar{q}(t) dt, \quad \text{para la forma lineal } y^*(p) = p(z_0) \text{ donde } z_0 \text{ es un número complejo}$$

fijo, no existe ningún elemento p_0 de $C[X]$ tal que $y^*(p) = (p, p_0) \forall p \in C[X]$.

Sin embargo, en ciertos espacios de dimensión infinita a los cuales se dota de elementos adicionales, como son los llamados *espacios de Hilbert*, se encuentra un resultado completamente análogo al que acabamos de demostrar.

Del teorema 2.6.1 resulta una correspondencia biyectiva entre un espacio vectorial con producto escalar de dimensión finita y su dual, correspondencia que depende solamente de la estructura intrínseca de dichos espacios, pero que en el caso de los espacios con producto escalar sobre $\mathbb{C}$ (unitarios) no es lineal. Esta correspondencia entre E y E^* , llamémosla ϕ , estará dada por $y \rightarrow f_y$ donde $f_y(x) = (x, y) \forall x \in E$, y cumple $\phi(y+y') = f_{y+y'}$. Y como

$$f_{y+y'}(x) = (x, y+y') = (x, y) + (x, y') = f_y(x) + f_{y'}(x)$$

tenemos que

$$\phi(y+y') = f_{y+y'} = f_y + f_{y'} = \phi(y) + \phi(y').$$

$$\text{Pero } \phi(ay) = f_{ay} \text{ y } f_{ay}(x) = (x, ay) = \bar{a}(x, y).$$

Por tanto,

$$\phi(ay) = f_{ay} = \bar{a} f_y = \bar{a} \phi(y).$$

La aparición de $\bar{a}$ al analizar la linealidad de la aplicación, realmente no resta validez a la correspondencia natural entre E y E^* . En el caso de un espacio con producto escalar sobre $\mathbb{R}$ (euclideo) ϕ es un verdadero isomorfismo, y en el caso de un espacio con producto escalar sobre $\mathbb{C}$ (unitario) se puede demostrar que las aplicaciones que cumplen $f(x+y) = f(x) + f(y)$ y $f(ax) = \bar{a}f(x)$, que son llamadas *semilineales*, tienen muchas de las propiedades fundamentales de las lineales, lo cual nos permite considerar aún en este caso a E y E^* como naturalmente isomorfos. Es por esto que algunas veces se suele usar para las imágenes por una forma lineal la siguiente notación: dado $x \in E$ y $y^* \in E^*$ se denota $y^*(x)$ por (x, y^*) , generalizando así los resultados de los espacios con producto escalar en los cuales sabemos identificar naturalmente y^* con un único y de E .

Aunque no hemos entrado en todos los detalles de la correspondencia natural entre espacios con producto escalar de dimensión finita y sus duales, creemos que el estudiante ha recibido una explicación satisfactoria de algunas semejanzas encontradas entre las propiedades de la ortogonalidad en un espacio con producto escalar y las de la ortogonalidad que definimos con respecto al dual.

Por ejemplo, del teorema de existencia de base dual para cualquier base, aplicado a un espacio con producto escalar de dimensión finita, resulta la existencia de una base ortonormal en un espacio con producto escalar. Recomendamos al estudiante que piense en esta afirmación y confirme su veracidad, y lo mismo podemos decir para la obtención de las propiedades de $W^\perp$, dado W un subconjunto de E , a partir de las propiedades de W^0 .

Hay todavía otro resultado interesante con respecto a esta correspondencia natural entre E y E^* para los espacios con producto escalar de dimensión finita: la relación que tiene este tipo de correspondencia con la correspondencia natural entre E y E^{**} , que encontramos en el epígrafe 2.5.

Ante todo, para que tenga sentido comparar ambas correspondencias habría que pasar de E^* a E^{**} , pero siendo E^* un espacio con producto escalar. No es difícil verificar, y lo plantearemos detalladamente en un ejercicio al final de este epígrafe, que E^* puede ser convertido en un espacio con producto escalar, dado el producto escalar de E^* mediante la definición siguiente:

$$\langle y_1^*, y_2^* \rangle = \langle y_1, y_2 \rangle = \langle y_2, y_1 \rangle,$$

donde y_1, y_2 son los elementos de E que le corresponden a y_1^*, y_2^* según la correspondencia natural entre E y E^* .

De la misma definición de producto escalar en E^* y de que $\bar{\alpha} = \alpha$ para todo $\alpha \in \mathbb{C}$, es posible intuir que yendo de E a E^{**} por la vía de la correspondencia definida en el epígrafe 2.4 o por la vía de la correspondencia entre elementos de un espacio con producto escalar y su dual, obtendremos el mismo resultado.

En efecto, sea $x_0 \in E$. A x_0 le corresponde $y_0^* \in E^*$ dado por $y_0^*(x) = \langle x, x_0 \rangle \forall x \in E$ y a y_0^* le corresponde $y_0^{**} \in E^{**}$ dado por $y_0^{**}(y^*) = \langle y^*, y_0^* \rangle \forall y^* \in E^*$. Por otra parte, a x_0 le corresponde $x_0^{**} \in E^{**}$ dado por $x_0^{**}(y^*) = y^*(x_0) \forall y^* \in E^*$.

Probemos que y_0^* satisface esta última igualdad, ya que el punto x_0^{**} que corresponde a x_0 y satisface dicha igualdad es único.

$$y_0^{**}(y^*) = \langle y^*, y_0^* \rangle = \langle x_0, y \rangle = y^*(x_0) \quad \forall y^* \in E^*.$$

De aquí que ambas correspondencias coincidan.

Endomorfismo adjunto o conjugado de un endomorfismo de un espacio con producto escalar

Al iniciar el estudio de las formas lineales, señalamos que estas constituyen un poderoso instrumento para el estudio de los endomorfismos de un

espacio con producto escalar. El nexo entre estos dos problemas aparentemente diferentes, comenzó a evidenciarse al caracterizar las formas lineales de un espacio con producto escalar de dimensión finita mediante el producto escalar por un vector fijo del espacio. Ahora bien, realmente el instrumento que nos permitirá estudiar importantes clases de endomorfismos de un espacio con producto escalar, es el llamado *endomorfismo adjunto*, que asociaremos a cada endomorfismo y que nos posibilitará caracterizar las clases de endomorfismos en función de la relación de estos con su adjunto.

Según veremos, el endomorfismo adjunto o conjugado desempeña, con respecto a los endomorfismos de un espacio con producto escalar, una función similar al conjugado de un número complejo.

Sea E un espacio con producto escalar de dimensión finita y f un endomorfismo de E . Consideremos el producto $\langle f(x), y \rangle$. De las propiedades del producto escalar resulta que para cada $y \in E$ fijo, la aplicación ϕ de E en K dada por $x \mapsto \langle f(x), y \rangle$ es una forma lineal sobre E , pues

$$\begin{aligned} \langle f(\alpha x + \beta x'), y \rangle &= \langle \alpha f(x) + \beta f(x'), y \rangle \\ &= \alpha \langle f(x), y \rangle + \beta \langle f(x'), y \rangle. \end{aligned}$$

o sea,

$$\phi_y(\alpha x + \beta x') = \alpha \phi_y(x) + \beta \phi_y(x') \quad \forall \alpha, \beta \in K, \quad \forall x, x' \in E$$

Como toda forma lineal sobre E puede expresarse en términos del producto escalar por un vector fijo, podemos plantear, dada la caracterización de las formas lineales en un espacio con producto escalar, el resultado siguiente:

TEOREMA 2.6.2

Para todo endomorfismo f de un espacio con producto escalar de dimensión finita E , existe un único endomorfismo f^* de E tal que

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle \quad \forall x, y \in E.$$

Demostración

Ya vimos que dado f y $y \in E$, la aplicación $\phi_y: E \rightarrow K$ dada por $x \mapsto \langle f(x), y \rangle$ es una forma lineal sobre E .

Luego existe para cada ϕ_y y por consiguiente para cada $y \in E$, un único vector y' tal que

$$\langle f(x), y \rangle = \phi_y(x) = \langle x, y' \rangle \quad \forall x \in E.$$

Esto nos permite definir una aplicación de E en E , asignándole a cada y el vector y' . Definimos, pues, $f^*: E \rightarrow E$ por $f^*(y) = y'$.

De la definición de f^* resulta que $\langle f(x), y \rangle = \langle x, f^*(y) \rangle \forall x, y \in E$. Por otra parte, existe una única aplicación de E en E que satisface $\langle f(x), y \rangle = \langle x, g(y) \rangle \forall x, y \in E$; esta es f^* .

En otras palabras, la identidad anterior caracteriza a f^* , pues si suponemos que exista tal g tendremos

$$\langle x, f^*(y) \rangle = \langle x, g(y) \rangle \quad \forall x \in E, \quad \forall y \in E$$

$$\langle x, f^*(y) - g(y) \rangle = 0 \quad \forall x, y \in E.$$

De aquí que $f^*(y) - g(y) = 0 \quad \forall y \in E$, ya que el único vector ortogonal a todo el espacio es el nulo y se deduce que $f^*(y) = g(y) \quad \forall y \in E$, o sea, $g = f^*$.

En el trabajo con f^* se utiliza generalmente la identidad planteada en el teorema, la cual usaremos primeramente para probar que f^* es lineal y mostrar así el método de trabajo con f^* . Esta identidad es llamada frecuentemente *identidad fundamental de f^** .

Tenemos que

$$\langle x, f^*(\alpha y + \beta y') \rangle = \langle f(x), \alpha y + \beta y' \rangle = \alpha \langle f(x), y \rangle + \beta \langle f(x), y' \rangle = \alpha \langle x, f^*(y) \rangle + \beta \langle x, f^*(y') \rangle = \langle x, \alpha f^*(y) + \beta f^*(y') \rangle \quad \forall x \in E.$$

Como esta igualdad se cumple para todo $x \in E$, podemos afirmar que $f^*(\alpha y + \beta y') = \alpha f^*(y) + \beta f^*(y')$. Pero esto, a su vez, es válido $\forall \alpha, \beta \in K, \forall y, y' \in E$.

Luego, f^* es lineal.

El teorema queda así demostrado.

El teorema anterior da lugar a la siguiente definición:

DEFINICIÓN 2.6.1

Sean E un espacio con producto escalar y f un endomorfismo de E . Se dice que el endomorfismo f posee adjunto o conjugado en E si existe un endomorfismo f^* de E tal que

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle \quad \forall x, y \in E.$$

A f^* , si es que existe, se le llama *endomorfismo adjunto o conjugado de f* .

El estudiante debe observar que no hemos restringido la definición de endomorfismo adjunto al caso de los endomorfismos de espacios con producto escalar de dimensión finita.

El teorema 2.5.2 plantea que en un espacio con producto escalar de dimensión finita todo endomorfismo posee adjunto o conjugado en E y que este es único.

En el caso de los espacios vectoriales de dimensión infinita no siempre existe el endomorfismo adjunto, aunque en tipos especiales de espacios, por ejemplo los de Hilbert, puede demostrarse la existencia del adjunto para todo endomorfismo u *operador*, que es el nombre que se suele dar a los endomorfismos en los espacios de funciones, los cuales son los ejemplos más frecuentes de espacios de dimensión infinita.

Si pensamos en la demostración de existencia del adjunto hecha en el caso de los espacios con producto escalar de dimensión finita, y en el señalamiento que hicimos después del teorema 2.6.1 sobre la existencia en los espacios de Hilbert de un resultado completamente análogo a dicho teorema, no nos debe sorprender la afirmación sobre la existencia del adjunto para todo endomorfismo de dichos espacios.

Por otra parte, en cuanto a la unicidad del adjunto, en caso de que este existiera para un endomorfismo de un espacio con producto escalar de dimensión infinita, el estudiante puede analizar la demostración de unicidad

en el caso finito, hecha en el teorema 2.5.2 y comprobar que es completamente válida en el caso infinito.

Ejemplos

1. Sea $E = M_n(\mathbb{C})$ con el producto escalar $\langle A, B \rangle = \text{tr } {}^t \bar{B} \cdot A$. Dada una matriz T fija de $M_n(\mathbb{C})$, la aplicación $I_T: M_n(\mathbb{C}) \rightarrow M_n(\mathbb{C})$ dada por $I_T(A) = TA$ es un endomorfismo de $M_n(\mathbb{C})$, lo cual puede ser verificado fácilmente por el estudiante.

Como E es de dimensión finita, sabemos que existe I_T^* , y podemos buscarlo mediante la identidad fundamental. Si I_T^* existe debe satisfacer

$$\langle I_T(A), B \rangle = \langle A, I_T^*(B) \rangle \quad \forall A, B \in M_n(\mathbb{C}).$$

El método práctico de trabajo consiste en desarrollar el primer miembro de dicha identidad de forma tal que obtengamos una expresión del tipo $\langle A, f(B) \rangle$, pues por la unicidad del adjunto podemos encontrar la expresión de este.

$$\begin{aligned} \langle I_T(A), B \rangle &= \text{tr } {}^t \bar{B} \cdot TA = \text{tr } (TA {}^t \bar{B}) = \text{tr } (A {}^t \bar{B} T) = \text{tr } (A {}^t \bar{T} B) = \\ &= \text{tr } {}^t \bar{T} B \cdot A = \langle A, {}^t \bar{T} B \rangle. \end{aligned}$$

$$\text{Luego, } I_T^*(B) = {}^t \bar{T} \cdot B = I_{{}^t \bar{T}}.$$

2. Sea E el espacio de los polinomios con coeficientes reales con el producto escalar $\langle p, q \rangle = \int_0^1 p(t) q(t) dt$, y sea D la aplicación de E en E que a un polinomio le asocia su derivada. Esta aplicación no tiene adjunto, pues si suponemos que existe D^* tal que

$$\langle D(p), q \rangle = \langle p, D^*(q) \rangle \quad \forall p, q \in E. \quad (1)$$

entonces

$$\langle D(p), q \rangle = \int_0^1 D(p(t)) q(t) dt = p(1)q(1) - p(0)q(0) - \langle p, Dq \rangle. \quad (2)$$

De (1) y (2) obtenemos

$$\langle p, (D^* + D)q \rangle = p(1)q(1) - p(0)q(0). \quad (3)$$

Para cada q , la aplicación que a p le asigna $p(1)q(1) - p(0)q(0)$ es una forma lineal sobre E , y puede demostrarse que para ella no existe $p_0 \in E$ tal que $y^*(p) = p(1)q(1) - p(0)q(0) = \langle p, p_0 \rangle$, a menos que $y^* = 0$; pero de (3) se infiere que si existe D^* , existe tal vector.

Luego $y^*(p) = 0 \quad \forall p$, de donde $q(1) = q(0) = 0$.

Como la identidad fundamental del adjunto (1) tiene que cumplirse para todo p y q en E , basta tomar q tal que $q(1) \neq 0$ o $q(0) \neq 0$, para que no exista D^* .

Los dos ejemplos anteriores muestran las técnicas de trabajo fundamentales en la búsqueda del endomorfismo adjunto. En los espacios con producto escalar de dimensión finita, en los cuales sabemos que existe el endomorfismo adjunto de cualquier endomorfismo, la búsqueda del adjunto no presenta dificultades. En los espacios de dimensión infinita debe de tenerse más cuidado, pues en general no podemos afirmar si existe o no el adjunto; por ejemplo, en el espacio con producto escalar del ejemplo 2, el operador dado por $I_q(p) = qp$ donde q es un polinomio fijo, posee adjunto pues

$$\langle I_q(p), r \rangle = \int_0^1 q(t)p(t)r(t) dt = \int_0^1 p(t)q(t)r(t) dt = \langle p, q(t)r(t) \rangle.$$

Luego $I_q^*(r) = qr$, es decir, $I_q^* = I_q$.

La forma del adjunto en todos los casos y su existencia en el caso de espacios con producto escalar de dimensión infinita, dependen del producto escalar considerado en el espacio y del propio endomorfismo. Esto se hizo evidente en el caso de los ejemplos del espacio de los polinomios estudiados.

En el caso finito dimensional existe, además, otra vía para hallar el adjunto: la vía matricial, la cual es importante pues como sabemos la representación matricial es muy conveniente para el estudio de las aplicaciones lineales. El problema consiste en conocer, dada la matriz de un endomorfismo en una base, cuál sería la matriz del adjunto en esa misma base, lo que nos permitiría comparar el endomorfismo y su adjunto mediante sus matrices.

Utilizaremos las ventajas de los espacios con producto escalar que encontramos en el epígrafe 2.4. Esto es, si tomamos una base ortonormal $(a_i)_{1 \leq i \leq n}$ de un espacio n -dimensional E con producto escalar, entonces el coeficiente α_{ij} de la matriz de un endomorfismo f de E en base (a_i) puede calcularse mediante $\alpha_{ij} = \langle f(a_j), a_i \rangle$.

Entonces, conocida la matriz $A = (\alpha_{ij})$ de f en la base ortonormal $(a_i)_{1 \leq i \leq n}$, calculemos la matriz de f^* , que denotaremos por $A^* = (\alpha_{ji}^*)$.

$$\alpha_{ji}^* = \langle f^*(a_i), a_j \rangle = \langle a_i, f(a_j) \rangle = \langle f(a_j), a_i \rangle = \alpha_{ji} = \bar{\alpha}_{ij},$$

de donde $\alpha_{ji}^* = \bar{\alpha}_{ij}$ y $A^* = {}^t\bar{A}$.

Hemos demostrado pues, la siguiente proposición:

PROPOSICIÓN 2.6.1

Sean E un espacio con producto escalar de dimensión n sobre K , $(a_i)_{1 \leq i \leq n}$ una base ortonormal de E , f un endomorfismo de E y $A = M(f, (a_i))$. Entonces

$$M(f^*(a_i)) = {}^t\bar{A}.$$

Esta proposición facilita la búsqueda del adjunto en dimensión finita.

Ejemplo

En $\mathbb{C}^2$ con el producto escalar usual, dada $f(x_1, x_2) = (x_1 + ix_2, -2x_1 - x_2)$, si queremos hallar f^* basta tomar la matriz de f en la base canónica, que es ortonormal, esto es:

$$A = M(f, (e_i)) = \begin{pmatrix} 1 & i \\ -2 & -1 \end{pmatrix}$$

$$\text{Entonces } {}^t\bar{A} = \begin{pmatrix} 1 & -2 \\ -i & -1 \end{pmatrix} \text{ será la matriz de } f^* \text{ en } (e_i).$$

$$\text{De aquí que } f^*(y_1, y_2) = (y_1 - 2y_2, -iy_1 - y_2).$$

Queremos destacar que la relación entre las matrices de f y f^* se halló a partir del hecho de que la base (a_i) en la cual se toman dichas matrices es ortonormal, lo que se pone en evidencia en el ejemplo siguiente:

Ejemplo

Sea $\mathbb{R}^2$ con el producto escalar usual y $p: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por $p(x_1, x_2) = (x_1, 0)$, o sea, la proyección ortogonal con respecto al eje OX .

Hallemos p^* :

$$\langle p(x_1, x_2), (y_1, y_2) \rangle = \langle (x_1, 0), (y_1, y_2) \rangle = x_1 \cdot y_1 = \langle (x_1, y_1), (y_1, 0) \rangle.$$

$$\text{Luego } p^*(y_1, y_2) = (y_1, 0) = p(y_1, y_2), \text{ es decir, } p = p^*.$$

Tomemos una base no ortonormal de $\mathbb{R}^2$, por ejemplo $b_1 = (1, 1)$, $b_2 = (1, 0)$.

$$M(p, B) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = A \text{ y}$$

$$M(p^*, B) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = A.$$

Sin embargo, $A \neq {}^t\bar{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Luego no se mantiene la relación entre la matriz de p y p^* encontrada en la proposición 2.6.1.

Propiedades del endomorfismo adjunto

Se pueden demostrar las siguientes propiedades del endomorfismo adjunto:

PROPOSICIÓN 2.6.2

Sean E un espacio con producto escalar sobre K , f, g endomorfismos de E que posean adjunto y $\alpha \in K$. Entonces

- (1) $f+g$ tiene adjunto y $(f+g)^* = f^* + g^*$
- (2) αf tiene adjunto y $(\alpha f)^* = \bar{\alpha} f^*$
- (3) $f \circ g$ tiene adjunto y $(f \circ g)^* = g^* \circ f^*$
- (4) f^* tiene adjunto y $(f^*)^* = f$

Probaremos (1) y (4) para mostrar la técnica de trabajo, y dejaremos (2) y (3) como ejercicio para el estudiante. Usaremos la identidad fundamental.

Demostración

$$(1) \langle (f+g)(x), y \rangle = \langle f(x) + g(x), y \rangle = \langle f(x), y \rangle + \langle g(x), y \rangle = \\ = \langle x, f^*(y) \rangle + \langle x, g^*(y) \rangle = \langle x, (f+g)^*(y) \rangle \quad \forall x, y \in E.$$

Luego, dada la unicidad del adjunto, tenemos $(f+g)^*(y) = (f^*+g^*)(y)$ $\forall y \in E \Rightarrow (f+g)^* = f^*+g^*$.

$$(4) \langle f^*(x), y \rangle = \langle y, f^*(x) \rangle = \overline{\langle f(y), x \rangle} = \langle x, f(y) \rangle \quad \forall x, y \in E$$

Como f^* es único y debe satisfacer $\langle f^*(x), y \rangle = \langle x, f^*(y) \rangle \quad \forall x, y \in E$,

se cumple que

$$f^{**}(y) = f(y) \quad \forall y \in E, \text{ de donde } f^{**} = f.$$

Hemos enunciado las propiedades anteriores de forma tal, que sean válidas en un espacio de dimensión finita o infinita, siempre que los endomorfismos considerados posean adjunto. Está claro que dichas propiedades son válidas en un espacio con producto escalar de dimensión finita, cualesquiera sean los endomorfismos de dicho espacio.

De estas propiedades resulta una interesante analogía entre los números complejos y el conjunto de los endomorfismos de un espacio con producto escalar, la cual nos ayudará en la búsqueda de las clases de endomorfismos de un espacio con producto escalar que tengan propiedades especiales.

En el conjunto $\mathbb{C}$ de los números complejos, para cada $\alpha \in \mathbb{C}$ existe el número $\bar{\alpha}$ complejo conjugado de α .

Recordemos que en $\mathbb{C}$ se cumplen las propiedades siguientes:

$$(1) \overline{(\alpha + \beta)} = \bar{\alpha} + \bar{\beta}$$

$$(2) \overline{(\alpha \cdot \beta)} = \bar{\alpha} \cdot \bar{\beta}$$

$$(3) \bar{\bar{\alpha}} = \alpha$$

las cuales hemos encontrado en las propiedades del adjunto, salvo algunas diferencias (el orden inverso en la composición de aplicaciones para $(f \circ g)^* = (g^* \circ f^*)$, haciendo un paralelo entre la suma de números complejos y la de endomorfismos y entre el producto de números complejos y la composición de endomorfismos. De ahí el nombre de *conjugado* que también se da al adjunto.

Esta analogía básica la usaremos como guía en el estudio de los endomorfismos de un espacio con producto escalar de dimensión finita.

Ejercicios

1. Halle el adjunto de los endomorfismos siguientes:

a) En $\mathbb{C}^2$ con el producto escalar usual:

$$f(z_1, z_2) = ((1+i)z_1 + iz_2, 2z_1 + iz_2).$$

b) En $\mathbb{C}^3$ con el producto escalar usual:

$$f(z_1, z_2, z_3) = (2z_1 + (1-i)z_2, (3+2i)z_1 - 4iz_2, (4-3i)z_2 - 3z_3).$$

2. Sea E el espacio de los polinomios con coeficientes reales de grado ≤ 2 con el producto escalar

$$\langle p, q \rangle = \int_0^1 p(t)q(t)dt.$$

Para $y^*: E \rightarrow \mathbb{R}$ dada por $y^*(p) = p(1)$, halle un vector p_0 de E tal que $y^*(p) = \langle p, p_0 \rangle \quad \forall p \in E$.

3. Sea $\mathbb{C}^3$ con el producto escalar usual y sea f el endomorfismo de $\mathbb{C}^3$ cuya matriz en la base canónica de $\mathbb{C}^3$ es

$$A = (a_{ij}) \text{ con } a_{ij} = i^{i+j} \text{ con } i^2 = -1.$$

Halle una base de $\text{Ker } f^*$.

4. Sean E un espacio con producto escalar y $\alpha, \beta \in E$.

a) Pruebe que $f: E \rightarrow E$ dada por $f(x) = \langle x, \alpha \rangle \beta$, define una aplicación lineal en E .

b) Pruebe que f tiene adjunto y halle la forma explícita de este.

c) Suponga que $E = \mathbb{C}^n$ con el producto escalar usual y considere $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, $\beta = (\beta_1, \beta_2, \dots, \beta_n)$. Halle la matriz de f con respecto a la base canónica de $\mathbb{C}^n$.

5. Sea E el espacio de los polinomios con coeficientes reales de grado ≤ 3 con el producto escalar

$$\langle p, q \rangle = \int_0^1 p(t)q(t)dt.$$

Halle el adjunto del endomorfismo de E que a cada polinomio le asigna su polinomio derivado.

6. Sean $E = M_n(\mathbb{C})$ con el producto escalar $\langle A, B \rangle = \text{tr } \bar{B}A$, P una matriz inversible fija en E y $L_P: M_n(\mathbb{C}) \rightarrow M_n(\mathbb{C})$ dado por $L_P(A) = P^{-1}AP$. Encuentre L_P^* .

7. Pruebe que en un espacio con producto escalar E de dimensión finita se cumple:

a) $(\text{id}_E)^* = \text{id}_E$

b) $(0)^* = 0$

c) Si f es un endomorfismo inversible, $(f^*)^{-1} = (f^{-1})^*$.

8. Sea E un espacio con producto escalar de dimensión finita y W un subespacio de E invariante por un endomorfismo f de E ($f(W) \subset W$). Pruebe que $W^\perp$ es invariante por f^* .

9. Sea E un espacio con producto escalar de dimensión finita y f un endomorfismo de E . Pruebe que $\text{Im } f^* = [\text{Ker } f]^\perp$.
10. Pruebe que si f es un endomorfismo de un espacio con producto escalar de dimensión finita, entonces $f^* \circ f = 0 \Rightarrow f = 0$.
11. Compare la definición y las propiedades del adjunto de un endomorfismo en un espacio E con producto escalar de dimensión finita, con la definición y las propiedades de la transpuesta de una aplicación lineal (en este caso, un endomorfismo) enunciadas en el ejercicio 19 del epígrafe 2.5.

2.7 Endomorfismo de un espacio con producto escalar de dimensión finita

Mediante el endomorfismo adjunto o conjugado, hemos hecho notar ciertas analogías entre la estructura de los números complejos y la estructura de los endomorfismos de un espacio con producto escalar, en la cual el endomorfismo adjunto desempeña la función del conjugado de un número complejo en $\mathbb{C}$.

Vamos ahora a estudiar las clases más importantes de endomorfismos de un espacio con producto escalar. Para ello nos guiaremos por determinados subconjuntos de $\mathbb{C}$, buscando su equivalencia en clases del conjunto de los endomorfismos de un espacio con producto escalar, pues es lógico pensar que de esta forma encontraremos importantes propiedades. La práctica confirma esta suposición, ya que muchos problemas pueden ser expresados mediante modelos matemáticos que se basan en dichas clases de endomorfismos.

En $\mathbb{C}$ podemos considerar, entre otros, los siguientes subconjuntos notables: el eje real, el eje imaginario y el círculo unitario (números complejos de módulo 1).

Como las analogías que buscamos se basan en los conceptos *conjugado de un número complejo* y *adjunto de un endomorfismo*, caractericemos mediante dichos conceptos los conjuntos anteriores.

El eje real es $\{z \in \mathbb{C}; z = \bar{z}\}$.

El eje imaginario es $\{z \in \mathbb{C}; z = -\bar{z}\}$.

El círculo unitario es $\{z \in \mathbb{C}; |z| = z \cdot \bar{z} = 1\}$.

Pasando a los endomorfismos de un espacio con producto escalar, obtendremos las siguientes clases de endomorfismos:

$\{f; f = f^*\}$ para el eje real.

$\{f; f = -f^*\}$ para el eje imaginario.

$\{f; f \circ f^* = \text{id}_E\}$ para el círculo unitario.

Estudiaremos ahora estas clases de endomorfismos.

Podíamos haber considerado otros subconjuntos en $\mathbb{C}$, por ejemplo el semieje real positivo, cuyos elementos se caracterizan por ser de la forma $x = z \cdot z$ con $z \neq 0$, y obtener así la correspondiente clase de endomorfismos dada por los f tales que $f = g \circ g^*$ con g inversible. Pero hemos tenido en cuenta solo los subconjuntos fundamentales, y en los ejercicios estudiaremos la clase de endomorfismos que acabamos de mencionar.

Endomorfismos autoadjuntos

Comenzaremos por estudiar los endomorfismos que pueden considerarse análogos a los números reales. El nombre que les daremos resulta de forma natural de la definición siguiente:

DEFINICIÓN 2.7.1

En un espacio E con producto escalar se llama *endomorfismo autoadjunto* o *autoconjugado* a un endomorfismo de E tal que $f = f^*$.

Ejemplo

Sea E el espacio de los polinomios con coeficientes reales y con el producto escalar dado por $\langle p, q \rangle = \int_0^1 p(t) q(t) dt$.

Entonces el endomorfismo T de E dado por $T(p(t)) = tp(t)$ es autoadjunto, pues

$$\langle T(p(t)), q(t) \rangle = \int_0^1 tp(t) q(t) dt = \int_0^1 p(t) tq(t) dt = \langle p(t), Tq(t) \rangle,$$

de donde

$$T^*(q(t)) = tq(t) \text{ y } T = T^*.$$

De la definición de endomorfismo autoconjugado o autoadjunto, está claro que un endomorfismo autoconjugado o autoadjunto f está caracterizado por $\langle f(x), y \rangle = \langle x, f(y) \rangle \forall x, y \in E$.

A los operadores autoconjugados o autoadjuntos se les suele llamar *simétricos*, si están definidos en un espacio con producto escalar real, y *hermíticos*, si están definidos en un espacio con producto escalar complejo.

Como estamos estudiando los endomorfismos autoadjuntos de un espacio E con producto escalar de dimensión finita, nos interesa su representación matricial, y como ya vimos anteriormente, para buscar la matriz del endomorfismo adjunto lo más conveniente es considerar una base ortonormal de E .

Sea, pues, $(a_i)_{1 \leq i \leq n}$ una base ortonormal de E y sea f un endomorfismo autoadjunto de E . Supongamos que $M(f, (a_i)) = A$; entonces $M(f^*, (a_i)) = {}^t\bar{A}$, y de aquí, como $f = f^*$ se deduce que $A = {}^t\bar{A}$.

Recíprocamente, supongamos que A es una matriz tal que $A = {}^t\bar{A}$. Podemos encontrar un endomorfismo f de K^n tal que A lo represente en la base canónica $(e_i)_{1 \leq i \leq n}$ de dicho espacio, esto es, $A = M(f, (e_i))$.

Entonces $\bar{A} = m(f^*(e))$, y de $A = \bar{A}$ se deduce que $M(f(e)) = M(f^*(e))$, por lo que podemos afirmar que $f = f^*$.

Lo expuesto anteriormente se resume en la proposición siguiente:

PROPOSICIÓN 2.7.1

En un espacio E con producto escalar, un endomorfismo f es autoadjunto o autoconjugado si y solo si su matriz A en cualquier base ortonormal de E cumple $A = \bar{A}$.

Esta proposición sugiere la definición siguiente:

DEFINICIÓN 2.7.2

Una matriz con coeficientes en K ($K = \mathbb{R}, \mathbb{C}$) se llama *matriz autoadjunta* o *autoconjugada* si y solo si $A = \bar{A}$.

Con el objetivo de unificar los resultados válidos indistintamente para $\mathbb{R}$ y $\mathbb{C}$, hemos estado trabajando en el caso más general y tomando la matriz $\bar{A}$. Está claro que en el caso real, $\bar{A}$ se convierte en A .

Existen terminologías específicas para las matrices autoconjugadas en los casos real y complejo; una matriz autoconjugada real, $A = \bar{A}$, se denomina *simétrica*, y una matriz autoconjugada compleja, $A = \bar{A}$, *hermítica*. Más adelante estudiaremos propiedades exclusivas de los casos real y complejo, y entonces usaremos la terminología específica del caso.

Es evidente que la representación matricial simplifica en un caso práctico: cuando sea posible obtener fácilmente una representación matricial del endomorfismo en una base ortonormal, la determinación de si el endomorfismo es autoadjunto o no, ya que basta analizar si su matriz en dicha base es autoconjugada.

Estudiemos ahora algunas propiedades de los endomorfismos autoadjuntos que reafirman nuestra idea intuitiva de la semejanza de estos con los números reales.

Sea E un espacio con producto escalar. Entonces se cumple:

- (1) La suma de dos endomorfismos autoadjuntos de E es un endomorfismo autoadjunto de E .
- (2) El producto de un endomorfismo autoadjunto por un elemento α de K ($K = \mathbb{R}$ o $\mathbb{C}$) es autoadjunto si y solo si α es real.
- (3) La composición de dos endomorfismos autoadjuntos es un endomorfismo autoadjunto si y solo si dichos endomorfismos conmutan. (En esta propiedad hay cierta variación con respecto a su homóloga para el producto de números.)

Las propiedades anteriores se deducen de las propiedades del endomorfismo adjunto, estudiadas en la proposición 2.5.2 y se expresan como sigue:

- (1) $(f + g)^* = f^* + g^*$

$$(2) (\alpha f)^* = \bar{\alpha} f^*$$

$$(3) (f \circ g)^* = g^* \circ f^*$$

y se cumplen para todo par de endomorfismos de E y todo $\alpha \in K$.

Las propiedades (1) y (2) de los endomorfismos autoadjuntos se deducen de las dos primeras propiedades anteriores. En cuanto a la propiedad (3):

Si $f \circ g$ es autoadjunto, es decir, $(f \circ g)^* = f \circ g$, tenemos que $f \circ g = (f \circ g)^* = g^* \circ f^* = g \circ f$. Luego f y g conmutan. Recíprocamente, supongamos que f y g conmutan, es decir, $f \circ g = g \circ f$; entonces $(f \circ g)^* = (g \circ f)^* = f^* \circ g^* = f \circ g$ y $f \circ g$ es autoadjunto. (Lo mismo ocurre si tomamos $g \circ f$.)

Quizás la analogía fundamental entre los números reales y los endomorfismos autoadjuntos esté dada por el hecho de que estos últimos permiten expresar todo endomorfismo de un espacio con producto escalar de una manera similar a como expresamos un número complejo mediante su parte real y su parte imaginaria.

Como ya tenemos un análogo para endomorfismos de los números reales, es lógico que antes de tratar de expresar un endomorfismo de manera similar a la forma $a + bi$ de los números complejos, busquemos un análogo de los números imaginarios puros.

Un número imaginario puro es de la forma ib con b real; así que es lógico suponer que el análogo para endomorfismos sea un endomorfismo de la forma ig con g autoadjunto. Tratemos de caracterizar dichos endomorfismos mediante su relación con el adjunto:

$$(ig)^* = \bar{i}g^* = -ig^* = -(ig).$$

Esta relación tiene sentido solo en un espacio con producto escalar sobre $\mathbb{C}$ (unitario), es decir, lo anterior nos lleva a considerar endomorfismos g que satisfagan la condición $f^* = -f$. Esta condición caracteriza a los endomorfismos que cumplen $g = if$ con f autoadjunto, ya que si $f^* = -f$, entonces $f = i(if^*)$ y $(if^*)^* = -if = if^*$. Luego if^* es autoadjunto.

Vamos a considerar endomorfismos tales que $f^* = -f$ y a investigar si resuelven el problema que planteamos de encontrar un análogo de la forma $a + bi$ para los endomorfismos de un espacio con producto escalar sobre $\mathbb{C}$.

DEFINICIÓN 2.7.3

Se llama *endomorfismo antiautoadjunto* a un endomorfismo f de un espacio con producto escalar E sobre K tal que $f^* = -f$.

En el caso particular de que $K = \mathbb{R}$, se le llama *endomorfismo antisimétrico*, y en el caso $K = \mathbb{C}$, *endomorfismo antihermítico*. En este último caso vimos que una definición equivalente sería $f = ig$ con g autoadjunto.

Los endomorfismos antiautoadjuntos tienen propiedades similares a los autoadjuntos, que dejamos que el estudiante enuncie y demuestre, teniendo cuidado con las posibles modificaciones. En particular, estos endomorfismos están caracterizados porque su matriz A en cualquier base ortonormal sa-

tisface $A = -{}^t\bar{A}$; dichas matrices se llaman *antiautoadjuntas* (antisimétricas si $K = \mathbb{R}$ y *antihermíticas* si $K = \mathbb{C}$).

Tratemos ahora de expresar un endomorfismo f arbitrario de un espacio con producto escalar sobre $\mathbb{C}$, como la suma de uno autoadjunto y otro antiautoadjunto. Supongamos que $f = f_1 + f_2$ (1) con f_1 autoadjunto y f_2 antiautoadjunto, de donde

$$f^* = f_1^* + f_2^* = f_1 - f_2 \quad (2).$$

Y trabajando formalmente con el sistema formado por (1) y (2), obtenemos

$$f_1 = \frac{f + f^*}{2}, \quad f_2 = \frac{f - f^*}{2}.$$

De nuevo hemos trabajado en forma constructiva y no sería necesario verificar nada, pero de todas formas el estudiante puede verificar que f_1 es autoadjunto y f_2 antiautoadjunto. Es evidente que $f = f_1 + f_2$, y de la forma en que los hemos obtenido se infiere que f_1 y f_2 son únicas. Entonces:

$$f = f_1 + i(f_2^*) = f_1 + i f_2' \text{ con } f_1 \text{ y } f_2' \text{ autoadjuntas,}$$

que es el análogo para endomorfismos de un espacio con producto escalar complejo, de la expresión de un número complejo en la forma $a + ib$.

Aunque más adelante volveremos a tratar los endomorfismos autoadjuntos al estudiar el problema de la reducción de endomorfismos en un espacio con producto escalar, como una última analogía entre los endomorfismos autoadjuntos y los números reales y los antiautoadjuntos y los números imaginarios puros, estudiemos los valores propios de ambas clases de endomorfismos. Es lógico esperar que los valores propios de los autoadjuntos sean reales y los de los antiautoadjuntos, imaginarios puros. Veamos la siguiente proposición:

PROPOSICIÓN 2.7.2

En un espacio E con producto escalar, los valores propios de un endomorfismo autoadjunto son reales y los de un endomorfismo antiautoadjunto son imaginarios puros.

Demostración

Probemos primero la afirmación referente a los valores propios de un endomorfismo autoadjunto.

Sea f autoadjunto, es decir, $f = f^*$ y supongamos que λ es un valor propio de f , o sea, existe x no nulo de E tal que $f(x) = \lambda x$. Entonces

$$\langle f(x), x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle \text{ y}$$

$$\langle f(x), x \rangle = \langle x, f(x) \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle.$$

De aquí

$$\lambda \langle x, x \rangle = \bar{\lambda} \langle x, x \rangle \text{ y como } x \neq 0, \text{ será } \langle x, x \rangle \neq 0,$$

de donde $\lambda = \bar{\lambda}$; luego λ es real.

Pasemos ahora a los valores propios de los endomorfismos antiautoadjuntos.

Sean $f = f^*$ y λ un valor propio de f y $x \neq 0$ tal que $f(x) = \lambda x$. Entonces

$$\langle f(x), x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle.$$

Por otra parte,

$$\langle f(x), x \rangle = \langle x, f^*(x) \rangle = \langle x, -f(x) \rangle = \langle x, -\lambda x \rangle = -\bar{\lambda} \langle x, x \rangle.$$

De manera similar al caso conjugado se cumple que $\lambda = -\bar{\lambda}$, esto es, λ es imaginario puro.

Observe el estudiante que en ambos casos la técnica de trabajo usada es la misma, y es la que en general se usa en la caracterización de los valores propios de los endomorfismos de un espacio con producto escalar.

Endomorfismos unitarios

Investiguemos ahora la clase de los endomorfismos de un espacio con producto escalar que sería análoga a la de los números complejos con módulo uno, y que al inicio de este epígrafe vimos que se trata de los endomorfismos caracterizados por $f \circ f^* = \text{id}_E$.

En un espacio con producto escalar de dimensión finita, la condición $f \circ f^* = \text{id}_E$ se puede expresar en varias formas equivalentes, por lo cual las daremos previamente a la definición de los endomorfismos que llamaremos *endomorfismos unitarios*.

PROPOSICIÓN 2.7.3

En un espacio con producto escalar de dimensión finita E , son equivalentes las siguientes condiciones para un endomorfismo U de E :

- (1) $U \circ U^* = \text{id}_E$
- (2) U es inversible y $U^* = U^{-1}$.
- (3) $U^* \circ U = \text{id}_E$
- (4) $\langle U(x), U(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E$
- (5) $\|U(x)\| = \|x\| \quad \forall x \in E$
- (6) U transforma toda base ortonormal de E en una base ortonormal de E .
- (7) U transforma alguna base ortonormal de E en una base ortonormal de E .

Demostración

$$(1) \Rightarrow (2) \Rightarrow (3)$$

$U \circ U^* = \text{id}_E \iff U(U^*(x)) = x \quad \forall x \in E$, lo que evidentemente significa que es sobreyectiva, ya que para todo $y \in E$ existe $z = U^*(y)$ tal que $U(z) = y$.

Como el espacio es de dimensión finita, U sobreyectiva es equivalente a U biyectiva, luego U es inversible. Además, $U^{-1} = U^*$, y como la inversa es única, se tiene que $U^* \circ U = \text{id}_E$.

(3) $\Rightarrow$ (4)

$$\langle U(x), U(y) \rangle = \langle x, U^*(U(y)) \rangle = \langle x, y \rangle \quad \forall x, y \in E$$

(4) $\Rightarrow$ (5)

Si $\langle U(x), U(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E$, para $y = x$ tenemos

$$\langle U(x), U(x) \rangle = \langle x, x \rangle \quad \forall x \in E,$$

o sea,

$$\|U(x)\| = \|x\| \quad \forall x \in E.$$

(5) $\Rightarrow$ (6)

Si $\|U(x)\| = \|x\| \quad \forall x \in E$, U es inyectiva, pues,

$$U(x) = 0 \Rightarrow \|U(x)\| = 0 = \|x\| \Rightarrow x = 0.$$

Luego, según lo que señalamos en (1) $\Rightarrow$ (2) $\Rightarrow$ (3), U es inversible y es un isomorfismo.

Sea $(a_i)_{1 \leq i \leq n}$ una base ortonormal de E . Como U es un isomorfismo, $(U(a_i))_{1 \leq i \leq n}$ es también una base de E . Queda probar que $(U(a_i))_{1 \leq i \leq n}$ es ortonormal.

Primero veamos que

$$\|U(x)\| = \|x\| \quad \forall x \in E \Rightarrow U^* \circ U = \text{id}_E.$$

$$\langle U^* \circ U - \text{id}_E \rangle(x), x \rangle = \langle U^* \circ U \rangle(x), x \rangle - \langle x, x \rangle$$

$$= \langle U(x), U(x) \rangle - \langle x, x \rangle = 0 \quad \forall x \in E$$

De aquí, como $U^* \circ U - \text{id}_E$ es autoadjunto,

$$(U^* \circ U - \text{id}_E)(x) = 0 \quad \forall x \in E \Rightarrow U^* \circ U = \text{id}_E$$

Como $U^* \circ U = \text{id}_E \Rightarrow \langle U(x), U(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E$.

Calculemos

$$\langle U(a_i), U(a_j) \rangle = \langle a_i, a_j \rangle = \begin{cases} 1, & \text{si } i=j \\ 0, & \text{si } i \neq j \end{cases}$$

y queda probado que $(U(a_i))_{1 \leq i \leq n}$ es ortonormal y (5) $\Rightarrow$ (6).

(6) $\Rightarrow$ (7) es evidente.

(7) $\Rightarrow$ (1)

Supongamos que $(a_i)_{1 \leq i \leq n}$ es una base ortonormal de E tal que $(U(a_i))_{1 \leq i \leq n}$ es también una base ortonormal. Debemos probar que $U \circ U^* = \text{id}_E$.

En este caso la vía es un poco más larga. Probaremos exactamente (7) $\Rightarrow$ (4) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (1).

Sea pues (7). Entonces consideremos $x, y \in E$ y supongamos que

$$x = \sum_{i=1}^n x_i a_i, \quad y = \sum_{j=1}^n y_j a_j$$

$$\langle U(x), U(y) \rangle = \langle U\left(\sum_{i=1}^n x_i a_i\right), U\left(\sum_{j=1}^n y_j a_j\right) \rangle$$

$$= \sum_{i,j=1}^n x_i \bar{y}_j \langle U(a_i), U(a_j) \rangle = \sum_{i=1}^n x_i \bar{y}_i$$

Como $(a_i)_{1 \leq i \leq n}$ es ortonormal, se cumple

$$\sum_{i=1}^n x_i \bar{y}_i = \langle x, y \rangle.$$

De aquí que

$$\langle U(x), U(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E.$$

Por tanto, para cada $y \in E$ se cumple

$$\langle U(x), U(y) \rangle = \langle x, U^*(U(y)) \rangle = \langle x, y \rangle \quad \forall x \in E.$$

Luego

$$U^*(U(y)) = y \quad \forall y \in E.$$

y tenemos

$$U^* \circ U = \text{id}_E.$$

De aquí que U sea inyectiva, pues $U(x) = 0 \Rightarrow U^*(U(x)) = x = 0$.

Como el espacio es de dimensión finita, U inyectiva equivale a U sobreyectiva.

De lo anterior se deduce que U es inversible y su inversa es U^* , y de la unicidad de la inversa, $U \circ U^* = \text{id}_E$.

Hemos escogido una vía de demostración que, por supuesto, no es la única y quizás tampoco la más corta. Dejamos como ejercicio al estudiante que trate de encontrar otra vía de demostración. Es posible que obtenga una vía más rápida y de todas formas constituye un ejercicio que le permitirá profundizar en las relaciones entre las diferentes condiciones dadas.

DEFINICIÓN 2.7.4

Dado un espacio E con producto escalar de dimensión finita, se llama *endomorfismo unitario de E* a un endomorfismo U que satisface cualquiera de las caracterizaciones equivalentes dadas en la proposición anterior.

Cuando $K = \mathbb{R}$, a los endomorfismos unitarios se les llama también *ortogonales*.

Ejemplos

1. En $\mathbb{R}^2$ con el producto escalar usual, $U_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$U_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$$

y que geoméricamente representa una rotación en el plano, es un endomorfismo unitario, ya que

$$\begin{aligned} \|U_\theta(x, y)\| &= \sqrt{\langle U_\theta(x, y), U_\theta(x, y) \rangle} \\ &= \sqrt{(x \cos \theta - y \sin \theta)^2 + (x \sin \theta + y \cos \theta)^2} \\ &= \sqrt{x^2 + y^2} \\ &= \|(x, y)\| \quad \forall (x, y) \in \mathbb{R}^2. \end{aligned}$$

2. En $E = M_n(\mathbb{C})$ con el producto escalar $\langle A, B \rangle = \text{tr } {}^t \bar{B} \cdot A$, el endomorfismo T_M de E dado por $T_M(A) = M(A)$ con $M = M^{-1}$, es un endomorfismo unitario, pues

$$(T_M)^* = T_{M^*} = T_{M^{-1}} = (T_M)^{-1}.$$

Los endomorfismos unitarios también son llamados *isometrías*. Este nombre proviene de que al definir una distancia a partir de la norma, según se hace en el espacio usual, los endomorfismos unitarios conservan las distancias.

En el caso de los espacios con producto escalar, nos interesan no solamente los isomorfismos del espacio en el sentido de la estructura vectorial, sino los que conservan los productos escalares. Cuando el espacio E es de dimensión finita, todo endomorfismo de E que conserva los productos escalares es un isomorfismo. Cuando el espacio es de dimensión infinita, esto no ocurre así, como lo muestra el ejemplo siguiente:

Ejemplo

En el espacio $C[0, 1]$ de las funciones con valores reales definidas sobre $[0, 1]$ consideremos el producto escalar

$$\langle f, g \rangle = \int_0^1 f(t)g(t)t^2 dt,$$

el endomorfismo T de $C[0, 1]$ con este producto, en $C[0, 1]$ con el producto usual, dado por

$$T(f(t)) = tf(t),$$

conserva los productos escalares, pues

$$\langle T(f(t)), T(g(t)) \rangle = \int_0^1 tf(t)tg(t)dt = \int_0^1 f(t)g(t)t^2 dt = \langle f, g \rangle.$$

Pero T no es un isomorfismo ya que no es sobreyectiva, pues la función 1, por ejemplo, no es imagen de ninguna función de $C[0, 1]$.

En general, dados dos espacios con producto escalar E y F sobre el mismo cuerpo K , se llama *isomorfismo de espacios con producto escalar entre E y F* a un isomorfismo f de E sobre F tal que $\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E$.

Si el estudiante estudia detenidamente la demostración de la proposición 2.7.3, podrá demostrar que si $\dim E = \dim F$ entonces las propiedades (1) a (4) son equivalentes.

$$(1) \langle f(x), f(y) \rangle = \langle x, y \rangle \quad \forall x, y \in E$$

$$(2) \|f(x)\| = \|x\| \quad \forall x \in E$$

(3) f transforma cada base ortonormal de E en una base ortonormal de F .

(4) f transforma alguna base ortonormal de E en una base ortonormal de F .

Lo anterior nos permite obtener un resultado lógico:

PROPOSICIÓN 2.7.4

Dos espacios E y F con producto escalar de dimensión finita sobre un mismo cuerpo K son isomorfos, en el sentido de los isomorfismos de espacios con producto escalar, si y solo si tienen la misma dimensión.

Demostración

Si E y F son isomorfos en el sentido de los espacios con producto escalar, entonces es evidente que tienen la misma dimensión.

Recíprocamente, si tienen la misma dimensión, tenemos:

Sean $(a_i)_{1 \leq i \leq n}$ una base ortonormal de E y $(b_j)_{1 \leq j \leq n}$ una base ortonormal de F . Entonces la aplicación lineal f de E en F tal que $f(a_i) = b_i, i = 1, \dots, n$, es un isomorfismo de espacios con producto escalar.

Pasemos a estudiar la representación matricial de los endomorfismos unitarios de un espacio con producto escalar de dimensión finita. Para ello tomaremos, como es usual, una base ortonormal.

Sean E un espacio con producto escalar de dimensión finita, U un endomorfismo unitario de E y $(a_i)_{1 \leq i \leq n}$ una base ortonormal de E . Busquemos la matriz de U^* en $(a_i)_{1 \leq i \leq n}$.

Tenemos que $U \circ U^* = U^* \circ U = \text{id}_E$. Luego

$$M(U \circ U^*, (a_i)) = M(U^* \circ U, (a_i)) = I_n.$$

De aquí obtenemos que

$$M(U, (a_i)) \cdot M(U^*, (a_i)) = M(U^*, (a_i)) \cdot M(U, (a_i)) = I_n.$$

Si llamamos $A = M(U, (a_i))$, tenemos

$$A \cdot {}^t \bar{A} = {}^t \bar{A} \cdot A = I_n \iff A^{-1} = {}^t \bar{A}.$$

Recíprocamente, si A es una matriz con coeficientes en K tal que $A^{-1} = {}^t \bar{A}$, podemos encontrar un endomorfismo U de K^n considerado con el

producto escalar usual, tal que $M(U, (e_i)) = A$, donde $(e_i)_{1 \leq i \leq n}$ es la base canónica de K^n . De aquí que ${}^tA = M(U^*, (e_i))$ y obtengamos $U^* = U^{-1}$, es decir, U es unitario.

Podemos enunciar el siguiente resultado que acabamos de demostrar:

PROPOSICIÓN 2.7.5

Los endomorfismos unitarios de un espacio con producto escalar, se caracterizan porque su matriz A en una base ortonormal de dicho espacio satisfacen $A^{-1} = {}^tA$.

Al igual que en el caso de los endomorfismos autoadjuntos, hemos considerado el caso más general: $A^{-1} = {}^tA$; cuando $K = \mathbb{R}$ sería $A^{-1} = {}^tA$. Esto da lugar a la definición siguiente:

DEFINICIÓN 2.7.5

Una matriz A de $M_n(K)$, $K = \mathbb{R}$ o $K = \mathbb{C}$, se llama *matriz unitaria* si y solo si $A^{-1} = {}^tA$.

En el caso real sería $A^{-1} = {}^tA$ y dichas matrices suelen llamarse *ortogonales*.

Los isomorfismos unitarios están, pues, representados por matrices unitarias. La representación matricial facilita el trabajo con los endomorfismos unitarios.

Ejemplo

Busquemos las matrices unitarias de orden dos, lo cual nos dará los posibles endomorfismos unitarios de un espacio con producto escalar de dimensión dos.

Diferenciaremos los casos real y complejo.

1ro. $K = \mathbb{R}$

Sea $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$.

Para que A sea ortogonal es necesario y suficiente que

$${}^tA = A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

$\det A = 1$, ya que si A es ortogonal $A \cdot {}^tA = I_2$, entonces

$$\det A \cdot \det {}^tA = \det I_2 = 1.$$

De aquí

$$(\det A)^2 = 1 \text{ y } \det A = \pm 1.$$

(Esto es válido para una matriz ortogonal de orden n cualquiera, como podrá inferir el estudiante del hecho de que lo único que utilizamos es la definición de matriz ortogonal.)

De lo anterior se deduce que A es ortogonal si y solo si

$$A_1 = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ o } A_2 = \begin{pmatrix} -a & b \\ b & a \end{pmatrix} \text{ con } a^2 + b^2 = 1.$$

En particular, el ejemplo que vimos de la rotación en $\mathbb{R}^2$ cuya matriz en la base canónica de $\mathbb{R}^2$ es

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

responde al primero de estos casos. En cuanto al segundo caso, podemos obtenerlo del primero multiplicando por $R = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, esto es, $A_2 = AR$.

La matriz R puede interpretarse geoméricamente como la matriz de una reflexión especular con respecto a OY , ya que

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}.$$

Luego A_2 es el producto de una rotación y una reflexión.

2do. $K = \mathbb{C}$

En este caso, trabajando de forma similar para $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{C})$, tendremos que para que A sea unitaria es necesario y suficiente que

$$\begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

pero $|\det A|^2 = 1$, esto es, $\det A = e^{i\theta}$ con $\theta \in \mathbb{R}$, y calculando se llega a

$$A = \begin{pmatrix} a & b \\ -e^{i\theta}\bar{b} & e^{i\theta}\bar{a} \end{pmatrix},$$

donde $\theta \in \mathbb{R}$ y $|a|^2 + |b|^2 = 1$.

Volviendo a las analogías entre los endomorfismos unitarios y los números complejos de módulo 1, veamos algunas propiedades de los endomorfismos unitarios en los espacios con producto escalar de dimensión finita.

Propiedades

- (1) El producto de dos endomorfismos unitarios es unitario.
- (2) El inverso de un endomorfismo unitario es unitario.
- (3) El endomorfismo identidad es unitario.

Estas resultan de la definición de endomorfismo unitario y son las análogas a:

El producto de dos números complejos de módulo 1 es un número complejo de módulo 1.

El inverso de un número complejo de módulo 1 es de módulo 1.

1 es un número complejo de módulo 1.

(4) Los valores propios de un endomorfismo unitario son números complejos de módulo 1.

Demostración

Sean U un endomorfismo unitario de un espacio E con producto escalar de dimensión finita, λ un valor propio de E y $x \neq 0$ de E tal que $U(x) = \lambda x$. Entonces

$$\langle U(x), U(x) \rangle = \langle \lambda x, \lambda x \rangle = \lambda \bar{\lambda} \langle x, x \rangle = \langle x, x \rangle,$$

y como $x \neq 0$, tenemos

$$\lambda \bar{\lambda} = 1 \iff |\lambda| = 1,$$

según queríamos demostrar.

Hemos visto la importancia de los cambios de base de una base ortonormal a otra ortonormal en un espacio con producto escalar de dimensión finita. Ahora estudiaremos la reducción a formas simples de la matriz de un endomorfismo de un tal espacio y a partir de todos los resultados encontrados hasta el momento, veremos que esta reducción se logra mediante bases ortonormales. Es por esto que nos interesa caracterizar las matrices de cambio de base de una base ortonormal a otra base ortonormal.

Sean, pues, E un espacio con producto escalar de dimensión finita y $(a_i)_{1 \leq i \leq n}, (b_j)_{1 \leq j \leq n}$ dos bases ortonormales de E .

Si P es la matriz de cambio de $(a_i)_{1 \leq i \leq n}$ a $(b_j)_{1 \leq j \leq n}$ tenemos que

$$b_j = a_{1j} a_1 + \dots + a_{nj} a_n, \quad j = 1, \dots, n.$$

Entonces:

$$\langle b_i, b_j \rangle = \sum_{k=1}^n a_{ki} \bar{a}_{kj} = \begin{cases} 0, & \text{si } i \neq j \\ 1, & \text{si } i = j \end{cases}$$

Pero $\sum_{k=1}^n a_{ki} \bar{a}_{kj}$ es el coeficiente j, i de $\bar{P} \cdot P$, de donde $\bar{P} \cdot P = I_n$ y P es unitaria.

Recíprocamente, si P es unitaria y $P = (a_{ij})$ se cumple que

$$\sum_{k=1}^n \bar{a}_{kj} a_{ki} = \begin{cases} 0, & \text{si } i \neq j \\ 1, & \text{si } i = j \end{cases}$$

Si interpretamos las columnas de P como vectores de K^n , resulta de lo anterior que dichas columnas forman una base B ortonormal con respecto al producto escalar usual en K^n ; de aquí que P sea la matriz de cambio de la base canónica de K^n a la base B , esto es, la matriz de cambio de una base ortonormal a otra base ortonormal.

Hemos demostrado el resultado siguiente:

PROPOSICIÓN 2.7.6

Las matrices de cambio de base de una base ortonormal a otra base ortonormal en un espacio unitario con producto escalar de dimensión finita, son exactamente las matrices unitarias (cuando $K = \mathbb{R}$, ortogonales).

El hecho de que las columnas de una matriz unitaria formen una base ortonormal de K^n con respecto al producto escalar usual de dicho espacio, caracteriza a dichas matrices y se usa frecuentemente en la práctica.

Nos interesa el resultado obtenido mediante un cambio de base ortonormal a ortonormal en la matriz de un endomorfismo de un espacio con producto escalar de dimensión finita, hemos visto que estos cambios de base están representados por matrices P tales que $P = P^{-1}$. Si recordamos, además, que dos matrices A y B que representen a un endomorfismo en bases distintas son semejantes, o sea, que existe P inversible tal que $B = P^{-1}AP$, es lógico que consideremos la definición siguiente:

DEFINICIÓN 2.7.6

Dos matrices A y B de $M_n(K)$, $K = \mathbb{R}$ o $K = \mathbb{C}$, se llaman *matrices unitariamente equivalentes* si y solo si existe una matriz P unitaria tal que $B = P^{-1}AP$.

En el caso real P es ortogonal y se dice que A y B son ortogonalmente equivalentes.

Reducción de endomorfismos de un espacio con producto escalar de dimensión finita

Plantearemos el siguiente problema: ¿Bajo qué condiciones, dado un endomorfismo de un espacio con producto escalar de dimensión finita, existe una base ortonormal del espacio, formada por vectores propios de dicho endomorfismo?

En otras palabras:

¿Es posible representar el endomorfismo en una base ortonormal mediante una matriz diagonal? (Sabemos la importancia de la representación matricial, pues nos permite obtener fácilmente información sobre el endomorfismo.)

Busquemos las condiciones que debe cumplir un endomorfismo para que pueda ser representado en una base ortonormal mediante una matriz diagonal.

Sean E un espacio con producto escalar de dimensión finita y f un endomorfismo de E . Supongamos que existe una base ortonormal $(a_i)_{1 \leq i \leq n}$ de

E formada por vectores propios de dicho espacio, esto es, $f(a_i) = \lambda_i a_i$, $i=1, \dots, n$. Entonces

$$M(f, (a_i)) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Como $(a_i)_{1 \leq i \leq n}$ es ortonormal, tenemos

$$M(f^*, (a_i)) = \begin{pmatrix} \bar{\lambda}_1 & & \\ & \ddots & \\ & & \bar{\lambda}_n \end{pmatrix}$$

de donde $f^*(a_i) = \bar{\lambda}_i a_i$, $i=1, \dots, n$.

Si $K = \mathbb{R}$, entonces $\bar{\lambda}_i = \lambda_i$. Luego $M(f, (a_i)) = M(f^*, (a_i))$ y obtenemos la condición necesaria $f = f^*$, o lo que es lo mismo, que f sea autoadjunto.

Si $K = \mathbb{C}$, la obtención de la condición es un poco más compleja, ya que no obtenemos ninguna de las clases de endomorfismos que hemos estudiado. Sin embargo, se reconoce rápidamente que las matrices de f y f^* conmutan; por tanto, $f \circ f^* = f^* \circ f$ es una condición necesaria en este caso.

Note el estudiante que esta condición se satisface por los endomorfismos autoadjuntos y unitarios, y da lugar a una nueva clase de endomorfismos, los llamados *endomorfismos normales*.

DEFINICIÓN 2.7.7

Un endomorfismo f de un espacio E con producto escalar, se llama *endomorfismo normal* si y solo si $f \circ f^* = f^* \circ f$.

De manera similar a lo que hicimos en las clases de endomorfismos anteriores, puede demostrarse que la matriz A en base ortonormal de un endomorfismo de un espacio con producto escalar de dimensión finita satisface $\bar{A} \cdot A = A \cdot \bar{A}$ (en el caso real, $A \cdot A = A \cdot A$), y recíprocamente, toda matriz que satisfaga dicha condición representa un endomorfismo normal. A dichas matrices se les llama *matrices normales*.

Ya tenemos condiciones necesarias en los casos real y complejo. Veamos si son suficientes.

Podríamos estudiar por separado el caso real y el complejo, pero vamos a hacer otro ordenamiento. Estudiaremos primero los endomorfismos autoadjuntos, aunque estos son un caso particular de los endomorfismos normales, y veremos que los resultados obtenidos para los endomorfismos normales en el caso complejo son aplicables a los autoadjuntos. La razón de este ordenamiento es que las características de los endomorfismos autoadjuntos hacen más naturales las vías de razonamiento a seguir y sirven de guía en el caso de los normales.

En primer lugar hay que probar que si f es un endomorfismo autoadjunto del espacio con producto escalar de dimensión finita E , entonces podemos encontrar una base ortonormal formada por vectores propios de f . Recordemos que para buscar una base de vectores propios basta demostrar, luego de verificar que las raíces de p_f están en K , que la dimensión de cada subespacio propio $V(\lambda)$ es igual a su multiplicidad como raíz de p_f (donde p_f es el polinomio característico de f).

En segundo lugar hay que probar que en cada subespacio propio se puede tomar una base ortonormal tal, que la unión de las bases de todos los subespacios propios sea una base ortonormal de E .

Si f es autoadjunto, ya hemos probado que los valores propios de f son reales, por lo cual, aun en el caso en que $K = \mathbb{R}$, se cumple la condición de que las raíces del polinomio característico pertenezcan a K , previa a cualquier otra consideración.

Aquí hemos usado el hecho de que en un espacio vectorial de dimensión finita los valores propios son exactamente las raíces del polinomio característico, y a lo sumo lo que puede pasar es que si $K = \mathbb{R}$ dichas raíces no sean reales. La situación es distinta en un espacio de dimensión infinita, donde un endomorfismo puede no tener valores propios.

Ejemplo

Sea $E = C[0,1]$, esto es, el espacio de las funciones, con valores reales, continuas en $[0,1]$ con el producto escalar $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$, y sea T el endomorfismo dado por $T(f(t)) = tf(t)$, que es autoadjunto. Supongamos que T posee valor propio, es decir, que para $\lambda \in \mathbb{R}$ existe $f \in E$, f no nula tal que $T(f(t)) = \lambda f(t)$.

Entonces tendremos que

$$tf(t) - \lambda f(t) = 0 \quad \forall t \in [0,1],$$

o sea,

$$(t - \lambda) f(t) = 0 \quad \forall t \in [0,1],$$

y de aquí $f(t) = 0$ para $t \neq \lambda$.

Al ser f continua se cumple que $f = 0$, lo cual es una contradicción, de donde se infiere que T no posee valores propios.

Es conocido que para cualquier endomorfismo los vectores propios asociados a valores propios distintos son linealmente independientes. En el caso de un endomorfismo autoadjunto de un espacio E con producto escalar, es válida la proposición siguiente:

PROPOSICIÓN 2.7.7

Sea f un endomorfismo autoadjunto de un espacio E con producto escalar de dimensión finita. Entonces los vectores propios asociados a valores propios distintos son ortogonales.

Demostración:

Supongamos que λ_1 y λ_2 son dos valores propios de f con $\lambda_1 \neq \lambda_2$. Sean x_1 y x_2 vectores propios asociados a λ_1 y λ_2 respectivamente. Probemos que $\langle x_1, x_2 \rangle = 0$.

$$\langle f(x_1), x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = \lambda_1 \langle x_1, x_2 \rangle$$

$$\langle x_1, f(x_2) \rangle = \langle x_1, \lambda_2 x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle$$

Como $\langle f(x_1), x_2 \rangle = \langle x_1, f(x_2) \rangle$ por ser f autoadjunto, resulta

$$\lambda_1 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle$$

Luego $(\lambda_1 - \lambda_2) \langle x_1, x_2 \rangle = 0$, y como $\lambda_1 \neq \lambda_2$, tenemos que $\langle x_1, x_2 \rangle = 0$ según queríamos demostrar.

La proposición 2.7.7 nos asegura que los vectores propios asociados a valores propios distintos son ortogonales entre sí; luego nos queda por probar que en cada subespacio propio $V(\lambda)$ se puede tomar una base ortonormal de dicho subespacio y que la unión de dichas bases es una base de E . Para ello es necesario demostrar la proposición siguiente:

PROPOSICIÓN 2.7.8

Si f es un endomorfismo autoadjunto de un espacio E con producto escalar de dimensión finita y W es un subespacio de E invariante por f , entonces $W^\perp$ es invariante por f .

Demostración

Tenemos que W es invariante por f , es decir, que $f(W) \subset W$ y debemos probar que $f(W^\perp) \subset W^\perp$, esto es, que dado $f(y)$ con $y \in W^\perp$. Entonces $\langle f(y), x \rangle = 0$. En otras palabras:

$$\langle x, f(y) \rangle = 0 \quad \forall x \in W.$$

Pero

$$\langle x, f(y) \rangle = \langle f(x), y \rangle = 0 \quad \forall x \in W.$$

ya que por ser W invariante por f , $f(x) \in W$, y teníamos que $y \in W^\perp$.

Luego hemos obtenido lo que queríamos demostrar.

Es conocida la importancia que tienen los subespacios invariantes en la obtención de formas sencillas para la matriz de un endomorfismo. En particular, recordemos que los subespacios propios de un endomorfismo f de E son subespacios invariantes de E .

Ya estamos en condiciones de demostrar el teorema siguiente:

TEOREMA 2.7.1

Sean E un espacio con producto escalar de dimensión finita y f un endomorfismo autoadjunto de E . Entonces existe una base ortonormal de E formada por vectores propios de f .

Demostración

Según las consideraciones anteriores, basta probar que para cada subespacio propio $V(\lambda)$ del endomorfismo f la dimensión de dicho subespacio coincide con la multiplicidad de λ como raíz del polinomio característico de f , ya que entonces f será diagonalizable y podemos encontrar una base de E formada por la unión de bases de los subespacios propios.

Probemos esto primero y después completaremos los detalles, pues no basta encontrar una base de vectores propios, sino que esta base sea ortonormal.

Supongamos que λ_i es un valor propio de f de multiplicidad k_i en el polinomio característico de f , y probemos que $\dim V(\lambda_i) = k_i$. Llamemos $p_i = \dim V(\lambda_i)$.

Sea $(a_j)_{1 \leq j \leq n}$ una base ortonormal de E tal que $a_1, a_2, \dots, a_{p_i}$ sea una base de $V(\lambda_i)$, y por consiguiente, $a_{p_i+1}, \dots, a_n$ una base de $[V(\lambda_i)]^\perp$.

Como $E = V(\lambda_i) \oplus [V(\lambda_i)]^\perp$ utilizando la proposición 2.6.8 tenemos que

$$M(f, (a_j)) = \begin{pmatrix} \lambda_i & & 0 \\ & \lambda_i & \\ 0 & & B \end{pmatrix}$$

pues, como sabemos, de la invarianza por f de $V(\lambda_i)$ y $[V(\lambda_i)]^\perp$ se infiere la aparición de bloques en $M(f, (a_j))$ y

$$B = M(f|_{[V(\lambda_i)]^\perp}, (a_{p_i+1}, \dots, a_n)).$$

Luego

$$p_f(\lambda) = p_A(\lambda) = (\lambda - \lambda_i)^{p_i} \cdot p_B(\lambda).$$

Como la multiplicidad de λ_i en p_f es k_i , si $p_i < k_i$, entonces λ_i debe ser una raíz de $p_B(\lambda)$. Por tanto, obtendremos un vector propio no nulo de f asociado a λ_i y perteneciente a $[V(\lambda_i)]^\perp$, lo cual es imposible, ya que entonces $\dim V(\lambda_i) \neq p_i$. De aquí que $p_i \geq k_i$; luego $p_i = k_i$ pues siempre $\dim V(\lambda_i) \leq k_i$.

Podemos afirmar que si $\lambda_1, \lambda_2, \dots, \lambda_r$ son los valores propios distintos de f , entonces

$$E = V(\lambda_1) \oplus V(\lambda_2) \oplus \dots \oplus V(\lambda_r),$$

y así obtenemos una base de vectores propios de E mediante la unión de las bases de los $V(\lambda_i)$, $i=1, \dots, r$.

Veamos cómo hallar una base ortonormal de vectores propios.

En cada subespacio $V(\lambda_i)$ podemos tomar una base ortonormal, y como los vectores propios asociados a valores propios distintos son mutuamente ortogonales, la unión de las bases ortonormales de los subespacios

$V(\lambda_i), i=1, \dots, r$, será la base ortonormal de vectores propios de E que queramos hallar.

Este teorema tiene un enunciado equivalente en forma matricial.

PROPOSICIÓN 2.7.9

Para toda matriz autoadjunta A (simétrica si $K = \mathbb{R}$) existe una matriz unitaria P (ortogonal si $K = \mathbb{R}$) tal que $P^{-1}AP$ es diagonal.

Demostración

Toda matriz autoadjunta A puede interpretarse como la matriz de un endomorfismo f autoadjunto de K^n en la base canónica $(e_i)_{1 \leq i \leq n}$ de dicho espacio con el producto escalar usual. De aquí que exista, según el teorema 2.7.1., una base ortonormal $(a_i)_{1 \leq i \leq n}$ de K^n tal que $D = M(f(a_i))$ sea diagonal.

La relación entre A y D será $D = P^{-1}AP$, donde P es la matriz de cambio de base de $(e_i)_{1 \leq i \leq n}$ a $(a_i)_{1 \leq i \leq n}$, que como matriz de cambio de una base ortonormal a otra ortonormal será unitaria. Así queda demostrada nuestra proposición.

Observemos que D será una matriz real, pues sus elementos son los valores propios de f .

Otra forma de enunciar la proposición anterior es:

Toda matriz autoadjunta es unitariamente equivalente a una matriz diagonal real.

En el caso real, esto es, para A simétrica o, lo que es lo mismo, $A = A'$, veremos que los resultados anteriores se aplican a las formas cuadráticas.

Veamos un ejemplo práctico de reducción de un endomorfismo autoadjunto de $\mathbb{R}^3$.

Ejemplo

Sea f de $\mathbb{R}^3$ en $\mathbb{R}^3$ dado por

$$f(x_1, x_2, x_3) = (2x_1 + x_2 + x_3, x_1 + 2x_2 + x_3, x_1 + x_2 + 2x_3).$$

Hallemos $A = M(f, (e_i))$, donde (e_i) es la base canónica de $\mathbb{R}^3$, que es ortonormal con respecto al producto escalar usual.

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Busquemos los valores y vectores propios de f .

$$P_f(\lambda) = P_A(\lambda) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = -(\lambda-1)^2(\lambda-4)$$

$$V(1) = \langle (1, 1, 0), (-1, 0, 1) \rangle$$

$$V(4) = \langle (1, 1, 1) \rangle$$

Queda hallar bases ortonormales en $V(1)$ y $V(4)$, para lo cual usaremos el proceso de Gram-Schmidt en $V(1)$ y normalizaremos el vector $(1, 1, 1)$ en $V(4)$. Así obtenemos la base de $\mathbb{R}^3$

$$a_1 = \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0 \right), a_2 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}} \right),$$

$$a_3 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right),$$

en la cual $V(1) = \langle a_1, a_2 \rangle$, $V(4) = \langle a_3 \rangle$.

La base $\{a_1, a_2, a_3\}$ es ortonormal con respecto al producto escalar usual de $\mathbb{R}^3$.

Ahora volveremos al caso de los endomorfismos normales, que estudiaremos solo en el caso complejo, ya que vimos que, en el caso real, una condición necesaria y suficiente para encontrar una base ortonormal del espacio formada por vectores propios del endomorfismo, es que el endomorfismo sea autoadjunto.

Además, un endomorfismo normal de un espacio con producto escalar real puede no tener valores propios reales. Por ejemplo, puede verificarse fácilmente que es normal el endomorfismo de $\mathbb{R}^2$ representado en la base canónica de dicho espacio por la matriz

$$A = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

ya que $A \cdot A' = A' \cdot A$ y sus valores propios son complejos: $\frac{\sqrt{3}}{2} + i, \frac{\sqrt{3}}{2} - i$.

Recordemos que llamamos *normal* a un endomorfismo f de un espacio con producto escalar E tal que $f \circ f^* = f^* \circ f$. Nuestro objetivo es demostrar que en caso de que E sea un espacio con producto escalar sobre $\mathbb{C}$, para todo endomorfismo normal de E existe una base ortonormal del espacio formada por vectores propios de f .

Como ya hemos demostrado un resultado homólogo al que queremos alcanzar, trataremos de seguir aquí la vía tomada para un endomorfismo autoadjunto, con las modificaciones necesarias.

Veamos primero qué ocurre con los valores propios en el caso autoadjunto. Por coincidir el endomorfismo y su adjunto, dichos valores eran los mismos para ambos endomorfismos, pero no es esta la situación para los endomorfismos normales. En este caso se obtiene el siguiente resultado:

PROPOSICIÓN 2.7.10

Sean E un espacio con producto escalar de dimensión finita sobre $\mathbb{C}$ y f un endomorfismo normal de E . Entonces $\lambda \in \mathbb{C}$ es un valor propio de f si

y solo si $\bar{\lambda}$ es un valor propio de f^* . Más aún, cada vector propio de x asociado a un valor propio λ es un vector propio de f^* asociado a $\bar{\lambda}$.

Demostración

Para todo x de E tenemos, por ser f normal, que

$$\langle f(x), f(x) \rangle = \langle x, f^*(f(x)) \rangle = \langle x, f(f^*(x)) \rangle = \langle f^*(x), f^*(x) \rangle.$$

De aquí *

$$\|f(x)\| = \|f^*(x)\| \quad \forall x \in E. \quad (*)$$

Si $\lambda \in \mathbb{C}$, entonces el endomorfismo $f - \lambda \text{id}_E$ es normal, y aplicando las propiedades del endomorfismo adjunto y (*), tenemos

$$\|(f - \lambda \text{id}_E)(x)\| = \|(f^* - \bar{\lambda} \text{id}_E)(x)\| \quad \forall x \in E.$$

Por lo tanto,

$$(f - \lambda \text{id}_E)(x) = 0 \text{ si y solo si } (f^* - \bar{\lambda} \text{id}_E)(x) = 0.$$

Por tanto, si existe x no nulo tal que $(f - \lambda \text{id}_E)(x) = 0$, se cumplirá que $(f^* - \bar{\lambda} \text{id}_E)(x) = 0$ y recíprocamente, con lo que queda probada la proposición.

Lo segundo que demostramos en el caso de los operadores autoadjuntos es el resultado relativo a la ortogonalidad de los vectores propios asociados a valores propios distintos. Si seguimos la demostración hecha en dicho caso, podemos observar que, con pequeñas modificaciones, se obtiene un resultado idéntico para los endomorfismos normales.

PROPOSICIÓN 2.7.11

Sean E un espacio con producto escalar de dimensión finita sobre $\mathbb{C}$ y f un endomorfismo normal de E . Entonces los vectores propios de f asociados a valores propios distintos son ortogonales.

Demostración

Sean λ_1 y λ_2 valores propios de f tales que $\lambda_1 \neq \lambda_2$ y x_1 y x_2 vectores propios asociados a λ_1 y λ_2 respectivamente. Entonces

$$\langle f(x_1), x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = \lambda_1 \langle x_1, x_2 \rangle \text{ y}$$

$$\langle f(x_1), x_2 \rangle = \langle x_1, f^*(x_2) \rangle = \langle x_1, \bar{\lambda}_2 x_2 \rangle = \bar{\lambda}_2 \langle x_1, x_2 \rangle.$$

Por lo tanto,

$$(\lambda_1 - \bar{\lambda}_2) \langle x_1, x_2 \rangle = 0 \text{ y } \lambda_1 \neq \bar{\lambda}_2 \Rightarrow \langle x_1, x_2 \rangle = 0.$$

Por tanto x_1, x_2 son ortogonales como queríamos demostrar.

El resultado que tendríamos que demostrar ahora se refiere al ortogonal de un subespacio invariante por un endomorfismo normal f . En este caso el resultado que obtenemos es más restringido, pero basta para nuestro objeto.

PROPOSICIÓN 2.7.12

Si f es un endomorfismo normal de un espacio E con producto escalar de dimensión finita sobre $\mathbb{C}$, y $V(\lambda)$ es el subespacio propio asociado al valor propio λ de f , entonces $[V(\lambda)]^\perp$ es invariante por f .

Demostración

Sea $x \in [V(\lambda)]^\perp$. Debemos probar que $f(x) \in [V(\lambda)]^\perp$, es decir, que

$$\langle x, y \rangle = 0 \quad \forall y \in V(\lambda) \Rightarrow \langle f(x), y \rangle = 0 \quad \forall y \in V(\lambda)$$

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle = \langle x, \lambda y \rangle = \lambda \langle x, y \rangle = 0 \quad \forall y \in V(\lambda).$$

De aquí que $[V(\lambda)]^\perp$ sea invariante por f .

Por último, si observamos la demostración del teorema 2.7.1, que prueba que para todo endomorfismo autoadjunto de un espacio con producto escalar de dimensión finita E existe una base ortonormal de E formada por vectores propios del endomorfismo, veremos que en esta demostración solo fue necesario utilizar las proposiciones 2.7.7 y 2.7.8, que hemos encontrado prácticamente idénticas en el caso de un endomorfismo normal. Podemos, pues, enunciar el teorema siguiente:

TEOREMA 2.7.2

Para todo endomorfismo normal f de un espacio E de dimensión finita sobre $\mathbb{C}$, existe una base ortonormal de E formada por vectores propios de f .

Demostración

La demostración no difiere de la realizada en el caso de los endomorfismos autoadjuntos. Señalaremos los puntos fundamentales y dejaremos al estudiante que complete la demostración.

- 1) Demostrar que para cada valor propio, λ_i de f de multiplicidad k_i , en el polinomio característico de f se tiene $\dim V(\lambda_i) = k_i$.
- 2) Establecer que, entonces, si $\lambda_1, \lambda_2, \dots, \lambda_r$ son los valores propios distintos de f , se cumple:

$$E = V(\lambda_1) \oplus \dots \oplus V(\lambda_r).$$

- 3) Utilizar 2) y la proposición 2.7.12 para establecer la existencia de una base ortonormal de vectores propios.

El teorema vuelve a tener una interpretación matricial para las matrices normales, que el estudiante puede enunciar.

Es evidente que el resultado del teorema 2.7.2 es aplicable a los endomorfismos unitarios en el caso complejo. En el caso real se encuentra otro tipo de reducción que el estudiante puede buscar en la bibliografía recomendada.

Al estudiar las diferentes clases de endomorfismos, vimos que sus valores propios tienen las siguientes características específicas:

- (1) Son números reales si el endomorfismo es autoadjunto.
- (2) Son números complejos de módulo 1 si el endomorfismo es unitario.

El teorema 2.7.2 permite caracterizar las diferentes clases de endomorfismos por sus valores propios.

COROLARIO

Si f es un endomorfismo normal de un espacio con producto escalar sobre $\mathbb{C}$ de dimensión finita, entonces:

- (1) Si f tiene todos sus valores reales, f es autoadjunto.
- (2) Si todos los valores propios de f tienen módulo 1, f es unitario.

Demostración

- (1) Debemos probar que si todo valor propio λ de f es real, entonces f es autoadjunto.

Si f es normal existe una base ortonormal $(a_i)_{1 \leq i \leq n}$ de vectores propios de f , es decir, $f(a_i) = \lambda_i a_i$.

Según demostramos, $f^*(a_i) = \bar{\lambda}_i a_i = \lambda_i a_i$, de donde $f(a_i) = f^*(a_i)$, y al coincidir en una base de E , se deduce que $f = f^*$.

- (2) Seguimos el mismo razonamiento que en (1), esto es, tomamos una base ortonormal $(a_i)_{1 \leq i \leq n}$ de vectores propios de f . Entonces

$$\langle f^* \circ f(a_i), a_j \rangle = \langle f(a_i), f(a_j) \rangle = \langle \lambda_i a_i, \lambda_j a_j \rangle = \lambda_i \bar{\lambda}_j \langle a_i, a_j \rangle = \langle a_i, a_j \rangle$$

De aquí que

$$\langle f^* \circ f(a_i), y \rangle = \langle a_i, y \rangle \text{ para todo } y \in E$$

$$(f^* \circ f)(a_i) = a_i = id_E(a_i) \text{ para todo } i.$$

Al coincidir en una base ortonormal, $f^* \circ f = id_E$ y f es unitario.

La demostración de este corolario muestra una vía para caracterizar los endomorfismos normales mediante sus valores propios. Note el estudiante que lo que hemos usado realmente es la existencia de una base en el espacio formado por vectores propios del endomorfismo; luego para toda clase de endomorfismos que posea una propiedad de este tipo, podremos caracterizar los diferentes endomorfismos de la clase mediante sus valores propios.

Reducción de una forma cuadrática a la forma canónica mediante transformaciones ortogonales

El teorema 2.6.1, que nos daba la posibilidad de encontrar una representación diagonal de un endomorfismo simétrico en una base ortonormal, se aplica en la reducción de una forma cuadrática real a la forma canónica.

Recordemos que en esta reducción, estudiada en el capítulo 1, se presentaba el problema de que la forma canónica hallada dependía de la transformación realizada.

Por otra parte, en la aplicación de la reducción a la forma canónica, a veces es necesario que el cambio de base realizado sea de una base ortonormal a otra ortonormal. En este caso la matriz de cambio de base es una matriz ortogonal y las transformaciones realizadas se conocen como ortogonales.

Sea, pues, una forma cuadrática q definida sobre un espacio E con producto escalar, y consideremos su matriz A en una base ortonormal $(a_i)_{1 \leq i \leq n}$ de E . Dicha matriz es simétrica y podemos interpretarla como la matriz en una base $(a_i)_{1 \leq i \leq n}$ de un endomorfismo simétrico f de E .

Por la proposición 2.6.9 tenemos que existe una matriz U unitaria tal que $U^{-1}AU$ es diagonal. Como $U^{-1} = {}^tU$, entonces $U^{-1}AU = {}^tUAU$, y esta matriz diagonal no solo representará al endomorfismo simétrico f , sino también a la forma cuadrática q en la base ortonormal $(b_i)_{1 \leq i \leq n}$.

En la diagonal de la matriz $D = U^{-1}AU$ aparecerán los valores propios de f , por lo cual, cualquiera sea la base ortonormal $(b_i)_{1 \leq i \leq n}$ de vectores propios de f en la cual obtengamos la forma diagonal de A , lo más que podrá ocurrir es que cambie el orden en que aparezcan los valores propios de f en dicha diagonal. Por tanto, podemos afirmar que la forma canónica que obtenemos para q es única.

Este proceso de reducción de una forma cuadrática a la forma canónica mediante transformaciones ortogonales, es conocido como la *reducción de una forma cuadrática real a los ejes principales*. Este nombre tiene su origen en el problema geométrico de la reducción de una cónica o una cuádrica a sus ejes principales, que como es sabido consiste en realizar una o varias transformaciones de coordenadas tal que los ejes de simetría y el centro de la curva o la superficie (en los casos de curvas y superficies con centro) coincidan, respectivamente, con los ejes y el centro del sistema de coordenadas hallado.

El resultado hallado se generaliza a una forma cuadrática cualquiera y es muy usado en las aplicaciones, por ejemplo en Estadística. Veamos cómo se aplica en la práctica.

Ejemplo

Reducir a los ejes principales la forma cuadrática de $\mathbb{R}^3$:

$$q(x_1, x_2, x_3) = 7x_1^2 - 4x_1x_2 + 6x_2^2 - 4x_2x_3 + 5x_3^2.$$

Podemos considerar a $\mathbb{R}^3$ con el producto escalar usual y que la forma está dada en la base canónica. Entonces

$$M(q, (e_i)) = \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}$$

y existirá un endomorfismo f de $\mathbb{R}^3$ tal que

$$A = M(f, (e_i)) = \begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}$$

esto es,

$$f(x_1, x_2, x_3) = (7x_1 - 2x_2, -2x_1 + 6x_2 - 2x_3, -2x_2 + 5x_3).$$

Hallemos una base ortonormal de vectores propios de f .

$$-p_f(\lambda) = \lambda^3 - 18\lambda^2 + 99\lambda - 162$$

y los valores propios son 3, 6 y 9.

$$V(3) = \langle (1, 2, 2) \rangle, \quad V(6) = \langle (2, 1, -2) \rangle, \quad V(9) = \langle (-2, 2, -1) \rangle$$

Una base ortonormal de vectores propios es:

$$b_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right), \quad b_2 = \left(\frac{2}{3}, \frac{1}{3}, \frac{-2}{3} \right), \quad b_3 = \left(\frac{-2}{3}, \frac{2}{3}, \frac{-1}{3} \right)$$

$$y \quad U = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{-2}{3} \\ \frac{-2}{3} & \frac{2}{3} & \frac{-1}{3} \end{pmatrix}$$

es la matriz de cambio de base de la base (e_i) a la (b_i) .

$$M(f, (b_i)) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

Las ecuaciones de transformación son:

$$x_1 = \frac{1}{3}y_1 + \frac{2}{3}y_2 - \frac{2}{3}y_3$$

$$x_2 = \frac{2}{3}y_1 + \frac{1}{3}y_2 + \frac{2}{3}y_3$$

$$x_3 = \frac{2}{3}y_1 - \frac{2}{3}y_2 - \frac{1}{3}y_3$$

y la forma canónica obtenida es

$$q(y) = 3y_1^2 + 6y_2^2 + 9y_3^2.$$

El problema anterior resuelve completamente el problema geométrico siguiente:

Reducir a su forma canónica la ecuación de la superficie

$$7x^2 + 6y^2 + 5z^2 - 4xy - 4yz = 18$$

cuyo primer miembro está dado por la forma cuadrática que acabamos de reducir.

De aquí podemos decir que en un sistema de coordenadas $x'y'z'$ cuyas ecuaciones de transformación con respecto a x, y, z son:

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

$$y = \frac{2}{3}x' + \frac{1}{3}y' + \frac{2}{3}z'$$

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

la superficie tiene por ecuación

$$3x'^2 + 6y'^2 + 9z'^2 = 18$$

$$o \quad \frac{x'^2}{6} + \frac{y'^2}{3} + \frac{z'^2}{2} = 1.$$

Esto es, la superficie es un elipsoide con centro en el origen de coordenadas y ejes coincidentes con los ejes de coordenadas. Las direcciones de los vectores directores de los nuevos ejes están dadas por los vectores propios de la matriz de la forma cuadrática que aparece en el primer miembro de la ecuación.

Reducción simultánea de dos formas cuadráticas

Otro problema relativo a las formas cuadráticas reales, en el cual se aplica el teorema 2.7.1, es el de la *reducción simultánea* de un par de formas cuadráticas. Este es un problema que fue resuelto por Weierstrass en el siglo pasado y que puede plantearse para un par de formas cuadráticas cualesquiera; pero en este texto estudiaremos solo el caso en que una de las formas es definida positiva. En el caso de dos formas cualesquiera no siempre tiene solución.

Este problema tiene gran interés en Estadística y aparece como un resultado algebraico preliminar en numerosos textos de esa disciplina.

El problema consiste en lo siguiente:
 Dado un par de formas cuadráticas q_1 y q_2 definidas sobre un espacio E con producto escalar de dimensión finita y q_2 definida positiva, investigar si es posible encontrar una transformación de coordenadas que reduzca simultáneamente q_1 a los ejes principales y q_2 a la forma normal.
 La respuesta es afirmativa y nos la da el teorema siguiente:

TEOREMA 2.7.3

Si q_1 y q_2 son formas cuadráticas definidas sobre el espacio E con producto escalar de dimensión finita sobre $\mathbb{R}$, y q_2 es definida positiva, entonces existe una transformación de coordenadas que reduce simultáneamente q_1 a los ejes principales y q_2 a la forma normal.

Demostración

Esta demostración se basa en el hecho de que una transformación ortogonal no modifica la forma normal de una forma cuadrática definida positiva, ya que en forma normal la matriz de una forma cuadrática definida positiva es I_n , y si U es una matriz ortogonal,

$$U \cdot I_n U = U^{-1} I_n U = I_n$$

A partir de esto podemos demostrar muy fácilmente el teorema. Esta demostración nos da el método a seguir en la práctica.

Supongamos que $q_1(x_1, x_2, \dots, x_n)$ y $q_2(x_1, x_2, \dots, x_n)$. Realicemos una transformación $X = PY$ tal que lleve q_2 a la forma normal, es decir,

$$q_2(y_1, y_2, \dots, y_n) = y_1^2 + \dots + y_n^2$$

Ahora realicemos una transformación ortogonal $Y = UZ$ tal que reduzca $q_1(y_1, y_2, \dots, y_n)$ a los ejes principales, o sea,

$$q_1(z_1, z_2, \dots, z_n) = \lambda_1 z_1^2 + \dots + \lambda_n z_n^2$$

Pero según lo que señalamos al inicio de la demostración, tenemos

$$q_1(z_1, z_2, \dots, z_n) = z_1^2 + \dots + z_n^2$$

Por tanto, la transformación buscada es

$$X = (PU)Z$$

Ejercicios

- Sean E un espacio con producto escalar de dimensión finita y f un endomorfismo de E tal que $f^2 = f$. Pruebe que f es autoadjunto si y solo si $f \circ f^* = f^* \circ f$.
- Dé un ejemplo de endomorfismos autoadjuntos cuya composición no sea un endomorfismo autoadjunto.

- Sea E el espacio de los polinomios de grado menor o igual que n con coeficientes reales y con el producto escalar $\langle p, q \rangle = \int_0^1 p(t)q(t)dt$ y sea f el endomorfismo de E que a cada polinomio le asigna su derivada. ¿Es f autoadjunto?
- Si f es un endomorfismo antiautoadjunto de un espacio E con producto escalar, ¿lo serán f^2 y f^n ?
- Sean E un espacio con producto escalar y f y g endomorfismos de E .
 - Pruebe que si f y g son antiautoadjuntos, entonces $f \circ g - g \circ f$ es antiautoadjunto.
 - ¿Qué sucederá si uno de ellos es autoadjunto y el otro antiautoadjunto?
- Pruebe que si f es un endomorfismo antisimétrico de un espacio con producto escalar real E , entonces $\langle f(x), x \rangle = 0$ para todo $x \in E$. ¿Se cumplirá el recíproco?
- Pruebe que si f es un endomorfismo autoadjunto de un espacio E con producto real y $f^2(x) = 0$, entonces $f(x) = 0$.
- Pruebe que si f es un endomorfismo antisimétrico de un espacio con producto escalar real de dimensión impar E , entonces $\det f = 0$. (Recuerde que $\det f$ es el determinante de la matriz de f en cualquier base de E .)
- Pruebe que en un espacio E con producto escalar, las siguientes condiciones son equivalentes para un endomorfismo f .
 - $f = g^2$ para algún endomorfismo autoadjunto g .
 - $f = g^* \circ g$ para algún endomorfismo g .
 - f es autoadjunto y $\langle f(x), x \rangle \geq 0$ para todo x de E .
- Los endomorfismos de un espacio E con producto escalar de dimensión finita que satisfagan cualquiera de las condiciones equivalentes dadas en el ejercicio 9, se llaman *endomorfismos positivos*. Si en la condición (3) sustituimos $\langle f(x), x \rangle \geq 0$ por $\langle f(x), x \rangle > 0$ para $x \neq 0$, se dice que el endomorfismo es *definido positivo*.
 - Pruebe que los valores propios de un endomorfismo positivo son reales y positivos y los de un endomorfismo definido positivo son reales y estrictamente positivos.
 - Pruebe que la suma de dos endomorfismos positivos (definidos positivos) es un endomorfismo positivo (definido positivo).
 - Pruebe que un endomorfismo positivo es definido positivo si y solo si es inversible.
- Sean U un endomorfismo unitario de un espacio E con producto escalar y W un subespacio de E invariante por U . Pruebe que $W^\perp$ es invariante por U .

12. Sea A una matriz de orden n y llamemos f_i , $i=1, \dots, n$, a las filas de A y c_j , $j=1, \dots, n$, a las columnas de A . Pruebe que el coeficiente i, j de AA^* es $\langle f_i, f_j \rangle$ y el coeficiente i, j de A^*A es $\langle c_i, c_j \rangle$.

13. Encuentre una matriz ortogonal A cuya primera columna sea

$$c_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right).$$

Sugerencia: Recuerde que las matrices ortogonales caracterizan los cambios de base ortonormal a base ortonormal.

14. Sea f un endomorfismo cualquiera de un espacio con producto escalar de dimensión finita; entonces existe un endomorfismo unitario U y un único endomorfismo P positivo tal que $f = UP$. Pruebe que si, además, f es inversible, entonces U es también único. (Observe la analogía con la expresión de un número complejo cualquiera como $z = pe^{i\theta}$ donde p es el módulo de z . Esta expresión se conoce como la *forma polar* de z , y por eso a la descomposición $f = UP$ se le suele llamar *forma polar* de f .)

Sugerencia:

a) Pruebe que el endomorfismo $f^* \circ f$ es positivo y que existe un único endomorfismo positivo g tal que $g^2 = f^* \circ f$.

b) Si f es inversible, considere $\bar{U} = gf^{-1}$ y pruebe que $\bar{U}$ es unitario. Tome $U = \bar{U}^{-1}$.

c) Si f no es inversible, considere $W = \text{Im } g$ y defina $U_1 : W \rightarrow W$ por $U_1(x) = f(y)$ donde $g(y) = x$. Pruebe que U_1 está bien definido, que $W^\perp$ y $[\text{Im } f]^\perp$ tienen la misma dimensión y considere a U_2 como un isomorfismo cualquiera entre $W^\perp$ y $[\text{Im } f]^\perp$.

d) Considere $E = W \oplus W^\perp$ y defina $U(x) = U_1(x_1) + U_2(x_2)$ donde $x = x_1 + x_2$ con $x_1 \in W$, $x_2 \in W^\perp$ y pruebe que cumple las condiciones requeridas.

15. Encuentre una matriz simétrica y ortogonal tal que su primera columna sea $c_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3} \right)$.

16. Pruebe que si una matriz unitaria (ortogonal) es triangular, entonces es diagonal.

17. Sean E un espacio con producto escalar de dimensión finita y W un subespacio de E . Considere $E = W \oplus W^\perp$ y el endomorfismo $U : E \rightarrow E$ tal que a $x = x_1 + x_2$ con $x_1 \in W$, $x_2 \in W^\perp$ le hace corresponder $U(x) = x_1 - x_2$. Pruebe que U es autoadjunto y unitario.

18. Sean E un espacio con producto escalar de dimensión finita y $U : E \rightarrow E$ una aplicación tal que $\langle U(x), U(y) \rangle = \langle x, y \rangle \forall x, y \in E$ y que sea sobreyectiva. Pruebe que U es lineal y, por tanto, un endomorfismo unitario de E .

19. Sean E un espacio con producto escalar complejo y f un endomorfismo autoadjunto de E . Pruebe que

a) $\|x + if(x)\| = \|x - if(x)\| \forall x \in E$

b) $x + if(x) = y + if(y)$ si y solo si $x = y$.

c) $id_E + if$, $id_E - if$ son inversibles.

d) Si E es de dimensión finita, entonces $U = (id_E - if) \circ (id_E + if)^{-1}$ es un endomorfismo unitario (llamado *transformación de Cayley*).

20. Pruebe que si θ es un número real, entonces las matrices

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \text{ y } \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

son unitariamente equivalentes.

21. Sean E un espacio con producto escalar de dimensión finita y P un endomorfismo positivo de E (ver ejercicio 10 de este epígrafe). Definimos un producto escalar en E que denotamos por $\langle \cdot, \cdot \rangle$ mediante $\langle x, y \rangle = \langle P(x), y \rangle$. Sean U un endomorfismo de E y U^* su adjunto con respecto al producto escalar $\langle \cdot, \cdot \rangle$. Pruebe que U es unitario con respecto a $\langle \cdot, \cdot \rangle$ si y solo si $P = U^* \circ P \circ U$.

22. Sea $\mathbb{C}$ como espacio vectorial sobre $\mathbb{R}$.

a) Pruebe que $\langle x, y \rangle = \text{Re } x\bar{y}$ es un producto escalar en $\mathbb{C}$.

b) Para cada $z \in \mathbb{C}$ considere T_z el endomorfismo de E dado por $T_z(x) = z \cdot x$. Pruebe que $(T_z)^* = T_{\bar{z}}$.

c) Diga para qué valores de z es T_z :

i) autoadjunto, ii) unitario, iii) positivo.

d) Halle la matriz de T_z en la base $\{1, i\}$ de $\mathbb{C}$.

e) Si f es un endomorfismo de E , encuentre una condición necesaria y suficiente para que f sea un T_z .

f) Halle un endomorfismo unitario de E que no sea un T_z .

23. Sea E un espacio con producto escalar de dimensión finita. Para cada $x, y \in E$, sea $T_{x,y}$ el endomorfismo de E definido por $T_{x,y}(z) = \langle z, y \rangle x$. Pruebe que

a) $T_{x,y}^* = T_{y,x}$

b) $T_{x,y} \circ T_{z,u} = T_x \cdot \langle y, z \rangle u$

c) Bajo qué condiciones es $T_{x,y}$ autoadjunto?

24. Para cada una de las siguientes matrices simétricas A , halle una matriz ortogonal P tal que PAP sea diagonal.

$$\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

25. Considere que la matriz $\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ representa un endomorfismo

f de $\mathbb{C}^2$ con producto escalar usual y con respecto a la base canónica de dicho espacio. ¿Existirá una base ortonormal de $\mathbb{C}^2$ formada por vectores propios de f ? Si es así, hállela.

Con el mismo enunciado considere las matrices:

$$\begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix} \text{ y } \begin{pmatrix} 1 & i \\ 1 & 2+i \end{pmatrix}.$$

26. Sea f un endomorfismo cualquiera de un espacio con producto escalar de dimensión finita E sobre $\mathbb{C}$. Pruebe que existe una base ortonormal de E con respecto a la cual f puede ser representado por una matriz triangular.

Sugerencia: Proceda por inducción en $\dim E$. Utilice el hecho de que f tiene al menos un valor propio λ y que, entonces, si consideramos un vector propio x asociado a λ , podemos escribir $E = \langle x \rangle \oplus \langle x \rangle^\perp$.

27. Pruebe que una matriz triangular es normal si y solo si es diagonal.

28. Analice si a partir de los ejercicios 26 y 27, puede encontrar otra demostración de la existencia de una base ortonormal de vectores propios para todo endomorfismo normal de un espacio con producto escalar de dimensión finita.

29. Pruebe que si f es un endomorfismo normal de un espacio con producto escalar E , entonces $\|f(x)\| = \|f^*(x)\| \quad \forall x \in E$. ¿Se cumplirá el recíproco si $K = \mathbb{R}$?

30. Pruebe que si f es un endomorfismo normal, entonces f y f^* tienen el mismo núcleo y la misma imagen.

31. Pruebe que si f y g son endomorfismos normales de un espacio E con producto escalar de dimensión finita, entonces existe una base ortonormal de E formada por vectores propios de f y g .

32. Sea f un endomorfismo de un espacio con producto escalar de dimensión finita E tal que es a la vez positivo y unitario. Pruebe que $f = id_E$.

33. Pruebe que un endomorfismo f normal de un espacio con producto escalar tal que existe $n \in \mathbb{N}$ y $f^n = 0$, tiene que ser nulo.

34. Reduzca a los ejes principales las siguientes formas cuadráticas de $\mathbb{R}^n$:

a) En $\mathbb{R}^3$, $q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 - 4x_1x_2 - 4x_1x_3$

b) En $\mathbb{R}^4$, $q(x_1, x_2, x_3, x_4) = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 + 2x_2x_3 + 2x_2x_4 + 2x_3x_4$

c) En $\mathbb{R}^3$, $q(x_1, x_2, x_3) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3$

d) En $\mathbb{R}^4$, $q(x_1, x_2, x_3, x_4) = 2x_1x_2 + 2x_3x_4$

e) En $\mathbb{R}^n$, $q(x_1, x_2, \dots, x_n) = x_1x_2 + x_2x_3 + \dots + x_{n-1}x_n$

3 Grupos

3.1 Introducción

Los antecedentes históricos de la *teoría de grupos* pueden hallarse, como todos los del Álgebra, en el estudio de la solubilidad de ecuaciones.

Alrededor de 1770, Lagrange (1736-1813) consideró la solubilidad de ecuaciones en términos de las permutaciones de sus raíces, y acerca de esto conjeturó que las ecuaciones de grado superior al 4to. no podían resolverse por los métodos usados hasta entonces. Recordemos que estos métodos consistían en reducir la solución a la de una ecuación de grado inferior, llamada *resolvente*, cuya solución era conocida.

Lagrange intuyó que las permutaciones de las raíces de una ecuación algebraica se comportaban de forma especial con respecto a las operaciones algebraicas, esto es, que poseían una *estructura algebraica*, y que precisamente las propiedades de dicha estructura algebraica daban condiciones para la solubilidad o no de la ecuación algebraica dada.

Galois y Abel continuaron los trabajos de Lagrange, y Abel (1802-1829) llegó a probar, como es conocido, la insolubilidad por radicales de la ecuación de grado superior al 4to. (En este sentido y de forma independiente había trabajado Ruffini, llegando a resultados similares a los de Abel.)

En cuanto a Galois (1811-1832), podemos decir que fue el que le dio el nombre de grupo a la estructura algebraica mencionada, ya intuida por Lagrange. Inspirado en la prueba de Abel sobre la insolubilidad de la ecuación de 5to. orden, Galois descubrió que una ecuación algebraica irreducible es soluble por radicales si y solo si un subgrupo del grupo simétrico (grupo de permutaciones) de sus raíces es soluble. Por ejemplo, aplicando el criterio de Galois a la ecuación $x^5 - ax - 1 = 0$ con $a \in \mathbb{Z}$, obtenemos que $x^5 - 1 = 0$ se resuelve por radicales, mientras que $x^5 - x - 1 = 0$ no tiene solución por radicales. A dichas ecuaciones le corresponden respectivamente los grupos $\mathbb{Z}_5$ y S_5 , que estudiaremos más adelante, y la solubilidad o no solubilidad de ellas depende de las características de los citados grupos.

Posteriormente, con el desarrollo de las geometrías no euclidianas en la segunda mitad del siglo XIX, la teoría de grupos se desarrolló aún más me-

diente los trabajos de Félix Klein (1849-1925), el cual la utilizó para estudiar con un método general los diferentes resultados encontrados en Geometría.

Klein conocía los trabajos de Lagrange y se impresionó por las posibilidades de unificación que brindaba el concepto de grupo, hasta el punto que puede decirse que pasó el resto de su vida desarrollando, aplicando y dando a conocer dicho concepto. En colaboración con el matemático noruego Sophus Lie (1842-1899) trabajó en la teoría de los grupos de transformaciones. En 1872 Klein mostró cómo podía caracterizar mediante esta teoría las diferentes geometrías que aparecieron durante el siglo XIX. Estas ideas las publicó en un trabajo famoso, conocido como "Erlanger Programm", donde describía la Geometría como el estudio de aquellas propiedades de las figuras que quedan invariantes bajo un cierto grupo de transformaciones. De ahí que una clasificación de los grupos de transformaciones se convertía en una clasificación de las geometrías.

Por ejemplo, según este enfoque, la geometría euclídeana plana, es el estudio de las propiedades de las figuras que quedan invariantes bajo el grupo de traslaciones y rotaciones del plano. Estas transformaciones se escriben

$$\left. \begin{aligned} x' &= ax + by + c \\ y' &= dx + ey + f \end{aligned} \right\} \text{ con } ae - bd = 1$$

y forman un grupo para el producto de transformaciones.

Si $ae - bd \neq 0$, pero no necesariamente $ae - bd = 1$, de todas formas dichas transformaciones forman un grupo, pero que no deja invariantes áreas y longitudes; sin embargo, una cónica se transforma en una cónica del mismo tipo. Estas transformaciones se llaman *afines* y caracterizan a la geometría afín, llamada así porque un punto finito se transforma en otro punto finito.

Algunas de estas ideas geométricas habían sido ya intuitivas por otros matemáticos, pero no existía el desarrollo de la matemática que permitió a Klein expresarlas de forma general.

Múltiples son las aplicaciones que se derivan de la introducción del concepto de grupo, que van desde las aplicaciones más teóricas a la Geometría, hasta la solución de ecuaciones y otras aplicaciones más directas a otras ramas de la ciencia como la física, la química, la biología, la estadística, etc. Por ejemplo, existen técnicas estadísticas de cómo organizar un experimento de forma tal, que con un mínimo de recursos se obtengan resultados de cuyo análisis se llegue a conclusiones válidas sobre el fenómeno en estudio. Dichas técnicas utilizan la teoría de grupos finitos en la construcción de los modelos a emplear.

Para el estudio de la teoría molecular es importantísimo el estudio de la simetría interna de las moléculas. Por ejemplo, las moléculas de agua responden a la estructura del grupo llamado *grupo de Klein*, que es un grupo finito de cuatro elementos y que estudiaremos más adelante. Actualmente el

desarrollo de la teoría molecular es muy difícil sin la ayuda de la teoría de grupos.

Otra rama en la que desde fines del siglo pasado se encontró aplicación a la teoría de grupos es la Cristalografía, donde se ha encontrado que la simetría interna de los cristales responden a estructuras de grupo, los llamados *grupos cristalográficos*, que son de utilidad en el estudio de las propiedades físicas de los cristales.

Los ejemplos anteriores son sólo una mínima muestra de las posibles aplicaciones de la teoría de grupos, la cual ocupa un lugar importante en el álgebra contemporánea.

El concepto *grupo*, como todas las estructuras algebraicas, consiste en un conjunto con cuyos elementos podemos operar algebraicamente, esto es, combinar dos elementos del conjunto para obtener un elemento también del conjunto, suponiendo además que estas operaciones algebraicas están sujetas a ciertas reglas que se llaman *axiomas definitorios de la estructura*.

Estos axiomas surgen de la experiencia que resulta de observar muchos ejemplos, de notar analogías entre objetos matemáticos aparentemente dispares y de la investigación de las raíces de estas analogías.

Los grupos como estructuras con una sola operación facilitan el estudio de los conceptos fundamentales, por lo que empezaremos estudiándolos. No obstante, como ya hemos señalado, para hablar de estructura es necesario primero establecer a qué llamamos *operación algebraica* o *ley interna*.

3.2 Leyes internas

La idea de ley interna nos es familiar desde nuestros primeros contactos con la Matemática, al sumar y multiplicar números.

¿Qué es lo que hacemos realmente cuando sumamos o multiplicamos? Tomamos un par de números y le asignamos otro número al cual llamamos *suma* o *producto*, según el caso, de los números dados. Posteriormente nos acostumbramos a operar con otros entes matemáticos; así, sumamos vectores, sumamos y multiplicamos matrices, etc. En cada caso tomamos dos elementos de un conjunto dado y le asignamos un elemento de este mismo conjunto, lo que nos lleva a la definición de ley interna u operación interna en un conjunto, que generaliza las ideas antes expuestas.

DEFINICIÓN 3.2.1

Se llama *ley interna* en un conjunto E , una regla que a cada par ordenado de elementos de E le asigna un único elemento de E .

De la definición se deduce que no puede quedar ningún par de elementos del conjunto sin tener un elemento asignado por la ley. Por otra parte, en la definición se toma un par ordenado para destacar que no necesariamente el resultado de componer dos elementos mediante la ley, es independiente del orden en que esto se haga, de lo cual tenemos ejemplo en la diferencia

de números enteros y el producto de matrices. Por último, decimos que la regla asigna a cada par de elementos de E un único elemento de E , ya que el resultado debe estar perfectamente determinado.

No se excluye la posibilidad de que el resultado de componer dos elementos dé por resultado uno de ellos; por ejemplo, $5+0=5$ en la suma de números naturales.

Las leyes internas suelen denotarse por diferentes símbolos: $*$, $\perp$, $\circ$, pero los más comunes son $+$ y $\cdot$, por analogía con las operaciones en los conjuntos numéricos. De este modo, si $a, b \in E$ y $*$ denota una ley interna en E , el resultado de componer a y b por dicha ley se denota por $a * b$ y se lee: " a compuesto con b por $*$ " (o simplemente " a compuesto con b ", si no hay duda de cuál es la ley).

En las notaciones de $+$ y $\cdot$ se conservan los nombres usuales de *suma* y *producto* para denominar el resultado de la operación.

Ejemplos

Son leyes internas:

1. $+$ en los conjuntos numéricos, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
2. La diferencia en $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
3. El cociente en $\mathbb{Q}^*$, $\mathbb{R}^*$, $\mathbb{C}^*$.
4. La suma y el producto de matrices cuadradas de orden n .
5. La unión y la intersección de conjuntos en $\mathcal{P}(E)$, donde E es un conjunto cualquiera y $\mathcal{P}(E)$ el conjunto de los subconjuntos de E .
6. En $\mathbb{N}$, la ley definida por $a * b = a$.
7. En $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \{f: \mathbb{R} \rightarrow \mathbb{R}\}$, la composición de aplicaciones.

No son leyes internas:

1. En $\mathbb{N}$, la correspondencia que a $n, m \in \mathbb{N}$ le asigna $n-m$, ya que por ejemplo el par $(2, 3)$ no tiene elemento que le corresponda.
2. En $\mathbb{Z}_+$, la correspondencia que a $p, q \in \mathbb{Z}_+$ le asigna r , donde r es mayor que $p+q+5$, ya que aunque todo par de enteros positivos tiene un entero positivo asignado, este no es único, por lo que el resultado de la operación no estaría bien determinado.
3. En $\mathbb{Z}_+$, la correspondencia que a $p, q \in \mathbb{Z}_+$ le asigna el mayor entero menor que el producto $p \cdot q$, ya que por ejemplo al par $(1, 1)$ le correspondería 0 y $0 \notin \mathbb{Z}_+$.

Por último, siempre es posible en un conjunto finito E definir una ley interna usando una "tabla" que nos indique cuál es la correspondencia. Por

ejemplo, si $E = \{a, b, c\}$ definimos una ley $*$ mediante una tabla de doble entrada como la siguiente:

		entrada 3			
	*	a	b	c	
entrada 2	a	a	b	c	
	b	c	a	b	
	c	b	b	a	
					fila 2 columna 3

La regla para hallar el resultado es la siguiente:

$$\left(\begin{array}{c} \text{elemento de} \\ \text{la entrada } i \\ \text{de la izquierda} \end{array} \right) * \left(\begin{array}{c} \text{elemento de la} \\ \text{entrada } j \text{ de} \\ \text{arriba} \end{array} \right) = \left(\begin{array}{c} \text{elemento de la} \\ \text{fila } i \text{ y la} \\ \text{columna } j \text{ de} \\ \text{los resultados} \end{array} \right)$$

De este modo, si queremos hallar $b * c$, que está dado por b en la entrada 2 de la izquierda y c de la entrada 3 de arriba, obtenemos b , que está en la fila 2 y la columna 3 de los resultados.

La definición de ley interna se puede dar de manera más formal mediante el concepto *aplicación*, ya que una ley interna sobre E no es más que una aplicación de $E \times E$ en E .

Tomando la operación diferencia en $\mathbb{Z}$, notamos que $3-2 \neq 2-3$, mientras que en la suma en $\mathbb{Z}$, $a+b=b+a$ para todo $a, b \in \mathbb{Z}$. Esto nos lleva a dar la definición siguiente:

DEFINICIÓN 3.2.2

Se dice que una ley interna $*$ definida sobre un conjunto E es *conmutativa* si y solo si $a * b = b * a$ para todo $a, b \in E$.

En los ejemplos dados anteriormente, son conmutativas las leyes en el 1 la suma en el 4 y la intersección y la unión en el 5, y son no conmutativas las leyes en el 2 y en el 3, el producto en el 4 y las leyes en el 6 y en el 7. Tampoco es conmutativa la ley definida por la tabla.

De igual manera podemos observar que en $\mathbb{Z}$, $2-(3-1) \neq (2-3)-1$ y, sin embargo, $(a+b)+c=a+(a+c)$ para todo $a, b, c \in \mathbb{Z}$, lo que lleva a definir la llamada propiedad asociativa.

DEFINICIÓN 3.2.3

Se dice que una ley interna $*$ definida sobre un conjunto E es *asociativa* si y solo si $(a * b) * c = a * (b * c)$ para todo $a, b, c \in E$.

En los ejemplos dados, son asociativas las leyes de los ejemplos 1, 4, 5, 6 y 7; no son asociativas las del 2 y el 3.

En las operaciones en los conjuntos numéricos, hemos observado la función que desempeña el 0 en la suma ($a+0=0+a=a$ para todo a) y el 1 en el producto ($1 \cdot a=a \cdot 1=a$ para todo a). Esto se repite con otras operaciones, por ejemplo, el $\emptyset$ para la unión en $P(E)$ y la id_R para la composición de aplicaciones en $F(R, R)$. Esta propiedad puede ser generalizada mediante la siguiente definición:

DEFINICIÓN 3.2.4

Se dice que una ley interna $*$ definida sobre E posee elemento neutro, si existe $e \in E$ tal que $a * e = e * a = a$ para todo $a \in E$. El elemento e se llama *neutro de E para la ley interna $*$* .

A veces ocurre que existe e tal que cumple solo $a * e = a$ o $e * a = a$ para todo $a \in E$. Decimos entonces que $*$ posee neutro a la derecha o a la izquierda, respectivamente.

En los ejemplos de leyes internas dados, tenemos que:

En el 4, la matriz nula es neutro para la suma y la matriz identidad I_n neutro para el producto.

En el 5, el conjunto E es neutro para la intersección; en los ejemplos 2, 3 y 6 no hay elemento neutro.

En el 2, el 0 es elemento neutro a la derecha pues $a - 0 = a$ para todo a , pero no lo es a la izquierda pues $0 - a = -a$. Una situación similar se presenta en el ejemplo 3, la cual puede ser analizada por el estudiante.

Por último, si observamos la suma en Z y la composición de aplicaciones en $F(R, R)$, entre otras, vemos que dado un elemento n de Z , existe $-n$ de Z tal que $n + (-n) = 0$. En $F(R, R)$, para f biyectiva existe f^{-1} tal que $f \circ f^{-1} = f^{-1} \circ f = id_R$, lo que lleva a la siguiente definición:

DEFINICIÓN 3.2.5

Dada una ley interna $*$ en E , que posea elemento neutro e , se dice que $a \in E$ posee *simétrico* para $*$ si existe $a' \in E$ tal que $a * a' = a' * a = e$.

De forma similar a lo señalado con respecto al neutro, si existe $a' \in E$ tal que solamente $a * a' = e$ o $a' * a = e$, se dice que a posee simétrico a la derecha o a la izquierda, respectivamente.

En los ejemplos:

En el 1, el simétrico existe para el producto para todo elemento no nulo.

En el 4, el simétrico para la suma de matrices, dada una matriz A , es la matriz $-A$, y en el producto de matrices tienen simétrico A^{-1} las matrices de determinante no nulo.

En el 5, para la unión y la intersección en $P(E)$ solamente tienen simétrico $\emptyset$ y E , respectivamente.

Hemos visto que las leyes internas poseen diversas propiedades y que hay leyes que poseen algunas propiedades y otras no. Frecuentemente se presentan ejemplos de leyes que poseen un mismo sistema de propiedades, lo cual tiene como consecuencia que en el conjunto donde están definidas se pueden resolver una serie de problemas que dependen sola-

mente de la ley y sus propiedades y no da lugar al estudio del concepto *estructura algebraica*, y como ya dijimos anteriormente, una de las estructuras algebraicas más sencillas es la de grupo, que trataremos en el próximo epígrafe.

Ejercicios

1. Determine cuáles de las leyes siguientes son internas en el conjunto dado:

- a) En Z_+ : $a * b = a - b$
- b) En Z_+ : $a * b = a^b$
- c) En R : $a * b = a - b$
- d) En Z_+ : $a * b = a + b + 5$

2. Se definen las aplicaciones $(x, y) \rightarrow mcm(x, y)$ y $(x, y) \rightarrow med(x, y)$ en $N \times N$. Demuestre que ambas definen leyes internas asociativas y conmutativas en $N \times N$. ¿Tienen estas leyes elemento neutro?

3. Demuestre que las aplicaciones de $P(E) \times P(E) \rightarrow P(E)$ definidas por $X, Y \in P(E)$, $(X, Y) \rightarrow X \cap Y$ y $(X, Y) \rightarrow X \cup Y$, son asociativas, conmutativas y admiten elemento neutro.

4. Sea $*$ la operación definida en $S = \{a, b, c, d, e\}$ por la tabla:

*	a	b	c	d	e
a	a	b	c	d	e
b	b	c	a	e	c
c	c	a	b	b	a
d	b	e	b	e	d
e	d	b	a	d	c

- a) Calcule $b * d$, $c * c$ y $[(a * c) * e] * a$
- b) Calcule $(a * b) * c$ y $a * (b * c)$. ¿Se puede asegurar que $*$ es asociativa?
- c) Calcule $(b * d) * c$ y $b * (d * c)$. ¿Se puede asegurar que $*$ es asociativa?
- d) ¿Es $*$ conmutativa? ¿Por qué?

5. Complete la tabla de modo que la operación $*$ definida en $\{a, b, c, d\}$ sea conmutativa.

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

6. Complete la tabla para que la operación $*$ definida en $\{a, b, c, d\}$ sea asociativa.

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

7. En $\mathbb{Q}$ se define la ley interna $a * b = \frac{a}{2} + \frac{b}{2}$, $a, b \in \mathbb{Q}$.

¿Qué propiedades tiene la ley $*$ en $\mathbb{Q}$?

8. En el conjunto $S = \{e, a, b, c, d\}$ se define la ley:

*	e	a	b	c	d
e	e	a	b	c	d
a	e	d	c	b	e
b	b	c	a	e	b
c	c	b	e	b	c
d	d	e	b	c	d

¿Qué propiedades tiene esta ley en S ?

9. En el conjunto $S = \{e, a, b, c, d\}$ se define la ley:

*	e	a	b	c	d
e	e	a	b	c	d
a	a	e	c	b	d
b	b	d	e	a	c
c	c	d	a	e	b
d	d	a	d	b	e

¿Qué propiedades tiene esta ley en S ?

10. Sea S un conjunto con una ley interna $*$, asociativa y conmutativa. Pruebe que para todo $a, b, c, d \in S$ se cumple:

$$(a * b) * (c * d) = [(d * c) * a] * b.$$

11. Para todo par $a, b \in \mathbb{R}^*$ se define la siguiente ley interna $\perp$ en $\mathbb{R}$:

$$x \perp y = ax + by.$$

Encuentre para qué valores de a y b es esta ley asociativa.

12. Sea $\perp$ una ley interna asociativa definida en un conjunto S y $a \in S$ fijo. Demuestre que la ley $*$ definida por

$$x * y = x \perp a \perp y$$

es asociativa.

13. Sea $E = \mathbb{R} \setminus \{0, 1\}$ y $f_i : E \rightarrow \mathbb{R}$ ($i=1, 2, \dots, 6$) definidas por

$$f_1(x) = x, \quad f_2(x) = \frac{1}{x-1}, \quad f_3(x) = \frac{x-1}{x},$$

$$f_4(x) = \frac{1}{x}, \quad f_5(x) = \frac{x}{x-1}, \quad f_6(x) = 1-x.$$

Demuestre que la composición de aplicaciones es una ley interna en $S = \{f_1, f_2, \dots, f_6\}$ y construya la tabla.

14. En $\mathbb{R}^2$ se define la relación $\circ$ por.

$$(x, y) \circ (x', y') = \begin{cases} \text{punto medio del segmento que une } (x, y) \text{ con } (x', y') & \text{si } (x, y) \neq (x', y') \\ (x, y) & \text{si } (x, y) = (x', y') \end{cases}$$

Diga si $\circ$ es una ley interna en $\mathbb{R}^2$.

15. En $\mathbb{R}$ se define $a * b = a + b + ab$. Diga las propiedades de esta ley.

16. En $\mathbb{Q}$ se definen las leyes

$$a * b = a - b + ab, \quad a \perp b = \frac{a+b+ab}{2}, \quad a \vdash b = \frac{a+b}{3}.$$

Analice la conmutatividad y la asociatividad de estas leyes.

17. En $S = \{1, 2, 3\}$ se define la siguiente ley interna:

$$\begin{array}{llll} 1 * 1 = 2 & 1 * 2 = 3 & 1 * 3 = 1 & 2 * 1 = 3 \\ 2 * 2 = 1 & 2 * 3 = 2 & 3 * 1 = 1 & 3 * 2 = 2 \\ & 3 * 3 = 3 & & \end{array}$$

Construya la tabla para esta ley.

18. a) ¿Cuántas leyes internas diferentes se pueden definir en un conjunto de dos elementos? ¿Cuántas leyes conmutativas?

b) Conteste las mismas preguntas para un conjunto de tres elementos.

19. Pruebe que el elemento neutro para una ley interna $*$ definida en un conjunto E es único.

Sugerencia: Suponga la existencia de dos elementos neutros para $*$ y, aplicando la definición de neutro, pruebe que tienen que ser iguales.

20. Pruebe que si una ley interna $*$ en un conjunto E tiene un elemento neutro a la izquierda e' y uno a la derecha e'' , entonces $e' = e''$ y $*$ posee neutro.

21. Sea $*$ una ley interna asociativa definida en E que posee elemento neutro e .

a) Pruebe que si $x \in E$ posee simétrico x' para $*$, este es único.

b) Pruebe que si $x, y \in E$ poseen simétricos x', y' para $*$, respectivamente, entonces $x * y$ posee simétrico $(x * y)'$ para $*$ y $(x * y)' = y' * x'$.

3.3 Grupos

En la introducción señalamos que el origen histórico del concepto grupo estaba precisamente en un conjunto de permutaciones.

Una permutación de orden n no es más que una biyección del conjunto

$\{1, 2, \dots, n\}$ en sí mismo. Estudiemos el problema para $n=3$. En este caso tenemos 6 permutaciones que denotaremos como sigue:

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \mu_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \mu_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

En la notación anterior, la primera fila de cada permutación indica los elementos del conjunto $\{1, 2, 3\}$ y la segunda fila las imágenes respectivas de dichos elementos. Por ejemplo, en ρ_1 la imagen de 1 es 2, la de 2 es 3 y la de 3 es 1.

Sea $S_3 = \{\rho_0, \rho_1, \rho_2, \mu_1, \mu_2, \mu_3\}$.

Obsérvese que mediante el concepto *biyección* formalizamos la idea de "arreglar" los elementos $1, 2, \dots, n$ de todas las formas posibles.

En S_3 se define una ley interna: el producto de permutaciones, que no es más que la composición de aplicaciones. Así, por ejemplo, $\mu_1 \circ \mu_2$ consiste en aplicar μ_2 y luego μ_1 y, entonces, $\mu_1 \circ \mu_2 = \rho_2$.

Resulta, pues, la siguiente tabla de multiplicar:

$\circ$	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_1	ρ_1	ρ_2	ρ_0	μ_3	μ_1	μ_2
ρ_2	ρ_2	ρ_0	ρ_1	μ_2	μ_3	μ_1
μ_1	μ_1	μ_2	μ_3	ρ_0	ρ_1	ρ_2
μ_2	μ_2	μ_3	μ_1	ρ_2	ρ_0	ρ_1
μ_3	μ_3	μ_1	μ_2	ρ_1	ρ_2	ρ_0

S_3 con la ley $\circ$ posee las propiedades siguientes:

(1) $\circ$ es asociativa.

(2) Existe un elemento neutro para $\circ$, éste es ρ_0 .

(3) Todo elemento de S_j posee simétrico para $\circ$.

Notemos que, sin embargo, $\circ$ no es conmutativa.

Por otra parte conocemos otros ejemplos de conjuntos dotados de leyes internas y que poseen las mismas propiedades que $(S_j, \circ)$. En $\mathbb{Z}$ la ley de suma es asociativa, existe elemento neutro para $+$: el 0, y todo $n \in \mathbb{Z}$ posee simétrico para $+$: $-n$. En este caso, además, $+$ es conmutativa.

Del Álgebra lineal, si consideramos las matrices de orden n y $\det \neq 0$ con el producto de matrices, vemos de nuevo que $\cdot$ es asociativa, existe neutro para $\cdot$, y para toda matriz A de $\det \neq 0$ existe A^{-1} . Denotamos por $GL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A \neq 0\}$.

Los tres ejemplos anteriores pueden ser resumidos y generalizados mediante el concepto *grupo*, el cual tiene similitud con operaciones familiares de la Aritmética y posee un gran campo de aplicabilidad.

DEFINICIÓN 3.3.1

Se llama *grupo* a un conjunto G dotado de una ley interna, que denotaremos aditiva o multiplicativamente, y que cumple las propiedades siguientes:

(g1) La ley interna es asociativa.

(g2) Existe un elemento neutro para la ley interna.

(g3) Todo elemento de G posee simétrico para la ley interna.

En símbolos se expresa:

(g1) $a(b \cdot c) = (a \cdot b)c$ para todo a, b, c de G .

(g2) Existe $1 \in G$ tal que $a \cdot 1 = 1 \cdot a = a$ para todo $a \in G$.

(g3) Para todo $a \in G$ existe $a^{-1} \in G$ tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Si la ley interna es conmutativa, se dice que el grupo es *conmutativo* o *abeliano*. En este caso usaremos la notación aditiva.

Recomendamos al estudiante que enuncie las propiedades (g1), (g2) y (g3) con notación aditiva. Además, llamamos su atención sobre lo siguiente: realmente debemos hablar de un grupo como un par: $(G, \cdot)$ o $(G, +)$ o $(G, \perp)$, dado cada uno por un conjunto y una ley interna, pues la estructura de grupo la determina no solo el conjunto, sino también la ley que consideramos sobre dicho conjunto. Sin embargo, para facilitar la notación, siempre que no haya confusión posible denotaremos $(G, \cdot)$ simplemente por G .

Si el grupo posee un número finito de elementos, se dice que el grupo es *finito* y se da la definición siguiente:

DEFINICIÓN 3.3.2

Se llama *orden* de un grupo finito G la cantidad de elementos de G . Se denota por $0(G)$ y se lee: "orden de G ".

Históricamente el concepto *grupo* se trató en sus inicios solo en el caso de grupos finitos, y es posteriormente que, por las necesidades de otras ramas de la Matemática, se desarrolla la teoría para grupos con cualquier cantidad de elementos, no necesariamente finita. Cuando un grupo no tiene una cantidad finita de elementos, diremos simplemente que es un *grupo infinito*.

Ejemplos

Además de $(S_j, \circ)$ y $(GL(n, \mathbb{R}), \cdot)$, otros ejemplos de grupo son los siguientes:

1. Los conjuntos numéricos $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ son grupos abelianos para la ley de suma usual en dichos conjuntos, y $\mathbb{Q}^*, \mathbb{R}^*, \mathbb{C}^*$ son grupos para el producto en dichos conjuntos.

2. $\mathbb{R}$ no es un grupo para la ley de suma, ya que aunque $+$ es conmutativa y posee elemento neutro 0, ningún elemento no nulo posee simétrico para $+$.

3. En $\mathbb{Q}^*$, si definimos $a * b = \frac{ab}{2}$, esta es una ley interna que es una ley de grupo para $\mathbb{Q}^*$, ya que

(g1) $*$ es asociativa pues

$$a * (b * c) = a * \frac{bc}{2} = \frac{abc}{4} \quad \text{para todo } a, b, c \text{ en } \mathbb{Q}^*.$$

$$(a * b) * c = \frac{ab * c}{2} = \frac{abc}{4}.$$

(g2) $*$ posee elemento neutro, el 2.

En efecto, $a * 2 = 2 * a = a$ para $a \in \mathbb{Q}^*$.

(g3) Para todo $a \in \mathbb{Q}^*$ existe $a' \in \mathbb{Q}^*$ tal que

$$a * a' = a' * a = 2, \text{ esto es, } a' = \frac{4}{a}.$$

$$a * \frac{4}{a} = \frac{\frac{4}{a} * a}{2} = 2.$$

4. Todos los ejemplos de espacios vectoriales que estudiamos en los cursos de Álgebra lineal son grupos para la ley de suma del espacio vectorial.

5. El conjunto $G = \{a + \sqrt{2}b \mid a, b \in \mathbb{R}\}$ es un grupo para la suma de números reales.

6. Un ejemplo geométrico de grupo es el de las rotaciones del plano (o del espacio) con centro en un punto fijo O . Se llama *producto* de las rotaciones ρ_1 y ρ_2 a la rotación que resulta de efectuar ρ_1 y ρ_2 sucesivamente. Este producto es una ley interna, el neutro es la rotación que deja cada punto en su posición original y el inverso de una rotación dada ρ es la rotación realizada en sentido opuesto a ρ . Es posible verificar fácilmente la asociatividad y comprobar que la conmutatividad no tiene que cumplirse.

7. El conjunto

$$SL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A = 1\}$$

dotado del producto de matrices es un grupo, ya que

$$\det A = 1, \det B = 1 \Rightarrow \det A \cdot B = 1$$

(g1) El producto de matrices es asociativo.

(g2) $\det I_n = 1$ y $A \cdot I_n = I_n \cdot A = A$ para toda $A \in SL(n, \mathbb{R})$

(g3) Si $\det A = 1$, entonces $\det A^{-1} = \frac{1}{\det A} = 1$, y sabemos que

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

8. Sea E un conjunto y $\sigma_E = \{f : E \rightarrow E \text{ biyección}\}$.

σ_E es un grupo para la composición de aplicaciones, el cual según veremos desempeñará un importante papel, y es la generalización natural de los grupos de permutaciones.

Comprobemos que σ_E es un grupo.

En efecto, la compuesta de dos biyecciones es una biyección, id_E dada por $\text{id}_E(x) = x$ es una biyección de E y $f \circ \text{id}_E = \text{id}_E \circ f = f$ para todo $f \in \sigma_E$. Si $f \in \sigma_E$, entonces f^{-1} es una biyección y $f \circ f^{-1} = f^{-1} \circ f = \text{id}_E$ y la composición de aplicaciones es asociativa.

Ejercicios

1. Diga si $(S, \circ)$ es un grupo

- $S = \{q \in \mathbb{Q}, q > 0\}$ y $\circ$ es la multiplicación usual en $\mathbb{Q}$.
- $S = \{z \in \mathbb{Z} : z = 2\}$ y $\circ$ es la multiplicación usual en $\mathbb{Z}$.
- $S = \mathbb{Z}$ y $a \circ b = 0$ para todo $a, b \in \mathbb{Z}$.
- $S = \mathbb{Z}$ y $a \circ b = b - a$.
- $S = \{1, -1, i, -i\}$ con el producto usual en $\mathbb{C}$.

2. Sea $S = \mathbb{R} \setminus \{-1\}$. Se define en S la ley $a * b = a + b + ab$.

- Demuestre que $(S, *)$ es un grupo.
- Halle la solución de la ecuación $2 * x * 3 = 7$ en S .

3. Demuestre que $(\mathbb{R}^*, \cdot)$ es un grupo abeliano si $a * b = |a| \cdot b$.
¿Es $(\mathbb{R}^*, \cdot)$ un grupo?

4. Sea $S = \{a + b\sqrt{2} : a, b \in \mathbb{Q}, a \text{ y } b \text{ no simultáneamente nulos}\}$. Demuestre que S con el producto usual es un grupo.

5. Sea $S = \{a + b\sqrt{-5} : a, b \in \mathbb{Q}, a \text{ y } b \text{ no simultáneamente nulos}\}$. Pruebe que S es un grupo para la multiplicación usual de números complejos.

6. Sea $S = \{z \in \mathbb{C} : z^m = 1\}$ con $m \in \mathbb{N}$, fijo. Pruebe que S con el producto usual de los números complejos es un grupo de orden m .

7. Sea $\mathbb{R}$ y la operación $x * y = \sqrt[3]{x^3 + y^3}$. Pruebe que $(\mathbb{R}, *)$ es un grupo abeliano.

8. Sea $A_1(\mathbb{R}) = \{\varphi_{a,b} : \mathbb{R} \rightarrow \mathbb{R} : \varphi_{a,b}(x) = ax + b; a, b \in \mathbb{R}, a \neq 0\}$. Pruebe que $A_1(\mathbb{R})$ es un grupo para la composición de aplicaciones.

9. Sea $G = \{(a, b) \in \mathbb{R}^2 : a \neq 0\}$. Se define en G la operación
 $(a, b) \cdot (c, d) = (ac, ad + b)$. Verifique que $(G, \cdot)$ es un grupo.

3.4 Propiedades de los grupos

Veamos algunas propiedades de la estructura de grupo que resultan de los axiomas.

- (1) Unicidad del neutro

El elemento neutro para una ley de grupo es único. En efecto, supongamos que existan e, e' tales que

$$a) \quad ex = xe = x \quad \forall x \in G$$

$$b) \quad e'x = xe' = x \quad \forall x \in G.$$

Entonces $ee' = e'e = e$ por b) y $ee' = e'e = e'$ por a), de donde $e = e'$.

Nótese que lo único que hemos usado es la existencia en G de una ley interna, por lo que puede pensarse en una propiedad más general.

(2) Unicidad del simétrico para cada $a \in G$.

Si $a \in G$, su simétrico es único.

Supongamos que existen dos simétricos a' y a'' para a , o sea,

$$\left. \begin{aligned} a \cdot a' &= a' \cdot a = e \\ a \cdot a'' &= a'' \cdot a = e \end{aligned} \right\} \text{ donde } e \text{ denota el elemento neutro de } G.$$

Entonces

$$a' \cdot a \cdot a'' = a' (a \cdot a'') = a' \cdot e = a'$$

$$a' \cdot a \cdot a'' = (a' \cdot a) \cdot a'' = e \cdot a'' = a''.$$

Luego $a' = a''$.

De forma similar a lo que hicimos en la unicidad del neutro, para la unicidad del simétrico hemos usado solo la existencia de una ley asociativa y que posee neutro.

(3) Propiedad cancelativa

Se cumple la propiedad cancelativa a la izquierda y a la derecha. En otras palabras,

$$\forall a \in G, ax = ay \Rightarrow x = y \text{ (a la izquierda)}$$

$$\forall a \in G, xa = ya \Rightarrow x = y \text{ (a la derecha).}$$

En efecto,

$$ax = ay \Rightarrow a^{-1}ax = a^{-1}ay \Rightarrow ex = ey \Rightarrow x = y.$$

(Análogamente se demuestra a la derecha.)

(4) Las ecuaciones $ax = b$ y $xa = b$ tienen solución única en G para todo $a, b \in G$.

En efecto, tomemos por ejemplo $ax = b$. Hallar una solución en G significa encontrar $x \in G$ tal que satisfaga la ecuación.

Utilizando los axiomas de grupo, tenemos:

$$\text{Multiplicando por } a^{-1}, a^{-1}(ax) = a^{-1}b.$$

$$\text{Asociando convenientemente, } (a^{-1}a)x = a^{-1}b.$$

$$\text{Utilizando } a \cdot a^{-1} = 1 \text{ y } 1 \cdot x = x, x = a^{-1}b.$$

Similarmente obtenemos la solución para la ecuación $xa = b$.

La propiedad (4) es una de las propiedades más importantes de la estructura de grupo; incluso podemos decir que una de las razones por las cuales tomamos la definición de grupo que dimos anteriormente, es que se cumple esta propiedad, o sea, que tenemos la posibilidad de resolver el tipo de ecuaciones $ax = b$ y $xa = b$ para cualquier par de elementos.

El ejemplo más conocido lo constituye $(\mathbb{Z}, +)$, donde podemos resolver las ecuaciones $a + x = b$ y $x + a = b$ para cualquier par de enteros. En $\mathbb{N}$, a pesar

de tener definida una suma con algunas de las propiedades de grupo, no es posible resolver las ecuaciones $a + x = b$ y $x + a = b$ para cualquier par de naturales, ya que la suma en $\mathbb{N}$ carece de la propiedad (g3), como señalamos anteriormente.

La solución $x = a^{-1}b$ obtenida para la ecuación $ax = b$ es única, ya que el inverso a^{-1} es único para cada $a \in G$.

(5) Si G es un conjunto con una ley interna asociativa y tal que para todo $a, b \in G$ las ecuaciones $ax = b$, $xa = b$ tienen solución única en G , entonces $(G, \cdot)$ es un grupo.

(Esta propiedad es, en cierto sentido, recíproca de la anterior.)

En efecto, podemos probar que existe el elemento neutro en G .

Tomemos la ecuación $ax = a$, la cual tiene solución para todo $a \in G$, según nuestra suposición. Llamemos e a la solución de $ax = a$ y probemos que $be = b$ para todo $b \in G$.

Sea, entonces, la ecuación $xa = b$, la cual también tendrá solución única en G :

$$xa = b \Rightarrow x(ae) = b \Rightarrow (xa)e = b \Rightarrow be = b \text{ para todo } b \in G.$$

Tomando la solución de $xa = a$, obtenemos que existe $e' \in G$ tal que $e'b = b$ para todo $b \in G$.

Como $e'e = e = e'$, resulta que $e = e'$.

Nótese que en el último paso hemos demostrado que si para una ley interna existe elemento neutro a la izquierda y a la derecha, estos tienen que ser iguales; luego existe neutro.

Probemos ahora que existe el simétrico para cada $a \in G$.

Consideremos las ecuaciones $ax = e$ y $xa = e$, que deben tener solución única. Llamemos a estas soluciones a' y a'' , respectivamente.

Como

$$a''aa' = a''(aa') = a''e = a''$$

$$\text{y } a''aa' = (a''a)a' = ea' = a',$$

se obtiene $a' = a''$.

Al demostrar que $a' = a''$ hemos demostrado que si para una ley interna asociativa y que posee neutro, un elemento posee simétrico a la izquierda y a la derecha, entonces estos son iguales. De aquí que dicho elemento posee simétrico para esta ley.

De las propiedades (4) y (5) resulta que podemos dar la siguiente definición equivalente de grupo:

DEFINICIÓN 3.4.1

Un conjunto G dotado de una ley interna $\cdot$ es un grupo si y solo si

(g1) es asociativa,

(g2) las ecuaciones $ax = b$ y $xa = b$ tienen solución única para toda $a, b \in G$.

Pueden encontrarse otras definiciones equivalentes de grupo y algunas específicas para los grupos finitos. Algunas de ellas aparecerán como ejercicio.

(6) Para todo $a, b \in G, (ab)^{-1} = b^{-1}a^{-1}$.

En efecto,

$$abb^{-1}a^{-1} = aea^{-1} = aa^{-1} = e \text{ y } b^{-1}a^{-1}ab = b^{-1}eb = b^{-1}b = e.$$

(7) En los conjuntos numéricos estamos acostumbrados a trabajar con potencias y múltiplos. También es posible definirlos en un grupo.

Si $(G, \cdot)$ es un grupo, para todo $a \in G$ y $m \in \mathbb{Z}$ podemos definir a^m como sigue:

$$a^m = \underbrace{a \cdot a \cdot \dots \cdot a}_{m \text{ veces}} \text{ para } m \in \mathbb{Z}^+$$

m veces.

$$a^0 = 1$$

$$a^m = (a^{-1})^{-m} \text{ para } m \in \mathbb{Z}^+$$

Las potencias definidas cumplen las propiedades:

$$a^{n+m} = a^n \cdot a^m$$

$$(a^m)^n = a^{nm}$$

Recomendamos al estudiante que verifique estas propiedades.

Si el grupo G es un grupo aditivo, podemos definir *múltiplo* así:

$$ma = \underbrace{a + a + \dots + a}_{m \text{ veces}} \text{ para } m \in \mathbb{Z}^+$$

m veces

$$0 \cdot a = 0 \text{ y } ma = -m(-a) \text{ para } m \in \mathbb{Z}^+$$

Los múltiplos definidos cumplen las siguientes propiedades:

$$ma + na = (m+n)a$$

$$m(na) = (mn)a$$

Ejercicios

1. Si G es un grupo tal que $(a \cdot b)^2 = a^2 \cdot b^2 \forall a, b \in G$, pruebe que G debe ser abeliano.

2. Pruebe que en un grupo finito de orden par existe al menos un elemento distinto de la identidad, el cual es su propio inverso.

3. Pruebe que un grupo finito de orden menor o igual a 4 es necesariamente abeliano.

Sugerencia: ba es igual a uno de los elementos e, a, b, ab .

4. Pruebe que si en un grupo $xx = x$, entonces $x = e$.

5. Pruebe que si $x^2 = e$ para todos los elementos de G , entonces G es conmutativo.

6. Pruebe que si $(ab)^n = a^n b^n$ para todo $a, b \in G$ y todos los enteros positivos n , entonces G es conmutativo, y viceversa.

7. Sea G un conjunto donde hay definida una ley interna asociativa tal que:

(1) Existe $e \in G: ae = a \forall a \in G$.

(2) Para todo $a \in G$ existe $a' \in G$ tal que $a \cdot a' = e$.

a) Pruebe que $(G, \cdot)$ es un grupo.

b) ¿Se puede afirmar lo mismo si en lugar de (2) se considera:

(2) Para todo $a \in G$ existe $a'' \in G$ tal que $a''a = e$?

Sugerencia: Analice el ejercicio 3 del epígrafe 3.3.

8. Sea G un conjunto finito provisto de una ley interna asociativa $\cdot$ y tal que:

$$(1) ax = ay \Rightarrow x = y \quad \forall a, x, y \in G$$

$$(2) xa = ya \Rightarrow x = y \quad \forall a, x, y \in G$$

a) Pruebe que $(G, \cdot)$, es un grupo.

Sugerencia: Analice la inyectividad y la sobreyectividad de la aplicación $\phi_a: G \rightarrow G$ dada por $x \mapsto xa$ para cada $a \in G$ fijo.

b) ¿Qué ocurrirá si se supone la validez de (1) o de (2) solamente?

3.5 Grupos finitos expresados mediante tablas

Ya hemos señalado anteriormente la importancia de los grupos finitos desde el punto de vista histórico y práctico. Por otra parte, vimos la posibilidad de describir una ley interna en un conjunto finito mediante una tabla de multiplicar.

A partir de los axiomas de grupo, veamos qué posibilidades dan dichos axiomas en la construcción de las tablas para grupos de dos, tres y cuatro elementos.

Está claro que el número mínimo de elementos que puede tener un grupo es uno, ya que por el axioma (g2) debe existir un elemento neutro. En este caso el grupo $G = \{e\}$ constará de un solo elemento, que será a la vez el neutro y el inverso de sí mismo.

Sea ahora $g = \{e, a\}$ un conjunto de dos elementos. Denotamos uno de los dos elementos por e , ya que sabemos que uno de los dos va a ser el elemento neutro del grupo. Completamos la tabla de modo que se cumplan los axiomas de grupo:

	e	a
e	e	a
a	a	e

La operación interna así definida cumple evidentemente (g2) y (g3). En cuanto a la asociatividad, también se cumple, lo cual puede comprobar el estudiante en los distintos casos.

Dados los axiomas de grupo, la anterior es la única posibilidad para una tabla de un grupo de dos elementos.

Observemos detenidamente las características de la tabla para ver qué condiciones han resultado de la exigencia de los axiomas y las propiedades de los grupos.

El hecho de que las ecuaciones $ax=b$ y $xa=b$ tengan solución para todo a y b de G , implica que cada elemento del grupo debe aparecer una y solo una vez en cada fila y en cada columna de la tabla. En particular, el elemento neutro e cumple esta propiedad. Esto nos facilita eliminar posibilidades, al ser una condición necesaria que debe cumplir la tabla de un grupo.

La descripción del grupo mediante la tabla permite analizar todas las propiedades menos la asociatividad; esta debe ser comprobada casuísticamente.

De la tabla de orden 2 podemos deducir que existe un único grupo de dos elementos, en el sentido de que estructuralmente todo grupo de orden 2 responde a la tabla dada anteriormente.

¿Qué ocurrirá para un grupo de tres elementos?

Analicemos las posibilidades de la tabla según las propiedades que vimos en el caso de orden 2. Para $G=\{e, a, b\}$ tendremos

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Obsérvese que se cumplen las características señaladas en el caso de orden 2.

La posibilidad anterior es la única para una ley interna en un conjunto de tres elementos, que da lugar a una ley de grupo. Recomendamos al estudiante que escriba las otras posibilidades de una tabla de tres elementos y compruebe que solo la dada anteriormente sirve para representar una ley de grupo.

Podemos, pues, decir que en el caso de orden 3 también existe un único grupo.

Posteriormente estudiaremos el concepto de *isomorfismo de grupos*, que formalizará las ideas expresadas anteriormente sobre la existencia o no de un único grupo de un orden dado.

Dejamos al estudiante el estudio de las posibilidades en el caso de cuatro elementos, para el cual existen solo las dos siguientes:

1.

	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

2.

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Al grupo descrito por la tabla 1 lo denotaremos por Z_4 ; el de la tabla 2 es conocido por V_4 o *grupo de Klein*.

Los grupos finitos también pueden describirse por otras vías en las que no se utilizan tablas. Por ejemplo, definamos el grupo siguiente: sea Z_m el conjunto $\{0, 1, 2, \dots, m-1\}$ formado por los m primeros números naturales, y definamos en Z_m la ley:

$$\begin{aligned} a \cdot b &= a+b & \text{si } a+b < m \\ a \cdot b &= r & \text{si } a+b = m+r, \text{ donde } 0 \leq r < m \end{aligned}$$

Esta ley interna es una ley de grupo en el conjunto de m elementos Z_m , luego da lugar a un grupo finito de m elementos.

Esta segunda forma de describir un grupo finito, mediante la caracterización de la ley interna, tiene la ventaja con respecto a las tablas de que es fácil comprobar la asociatividad, ya que se conoce la forma explícita del producto de dos elementos.

Construyamos las tablas de Z_m para algunos valores pequeños de m .

	0	1
0	0	1
1	1	0

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Si identificamos e con 0, a con 1, b con 2 y c con 3, encontramos de nuevo las tablas de los únicos grupos de orden 2 y 3, y en el caso de $\mathbb{Z}_4$, la posibilidad dada por la tabla 1 presentada anteriormente; de ahí la notación de $\mathbb{Z}_4$ que le habíamos dado a este grupo. Los grupos $\mathbb{Z}_m$ desempeñarán un importante papel en lo que sigue.

Producto de grupos

Nuestro propósito ahora es seguir aumentando el número de ejemplos de grupos. Estudiaremos un proceso constructivo de formación de grupos a partir de grupos conocidos. Mediante este proceso podremos construir gran cantidad de grupos, entre los que quedarán incluidos todos los grupos abelianos finitos, según veremos posteriormente.

Comencemos por dar algunas definiciones:

DEFINICIÓN 3.5.1

Se llama *producto directo externo* de los grupos $G_1, G_2, \dots, G_n$, al grupo dado por el producto cartesiano $G_1 \times G_2 \times \dots \times G_n = \prod_{i=1}^n G_i$ dotado de la operación

$$(a_1, a_2, \dots, a_n)(b_1, b_2, \dots, b_n) = (a_1 b_1, \dots, a_n b_n).$$

(El estudiante puede comprobar que se cumple la afirmación hecha en la definición, de que $\prod_{i=1}^n G_i$ con la operación de producto componente a componente es un grupo.)

Si los grupos G_i son abelianos, entonces $\prod_{i=1}^n G_i$ es abeliano. También es fácil comprobar que si G_i tiene r_i elementos, entonces $\prod_{i=1}^n G_i$ tiene

$$r_1 \cdot r_2 \cdot \dots \cdot r_n \text{ elementos.}$$

Ejemplos

1. Consideremos $\mathbb{Z}_2 \times \mathbb{Z}_3$, que tiene $2 \cdot 3 = 6$ elementos.

$$\mathbb{Z}_2 \times \mathbb{Z}_3 = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2)\}$$

Hallemos, por ejemplo, las siguientes sumas en $\mathbb{Z}_2 \times \mathbb{Z}_3$:

$$(0,2) + (1,1) = (1,0)$$

$$(1,2) + (1,2) = (0,1)$$

$$(0,3) + (0,3) = (0,0)$$

Observemos que las operaciones utilizadas son las de $\mathbb{Z}_2$ para las primeras componentes y las de $\mathbb{Z}_3$ para las segundas.

2. $\mathbb{Z}_3 \times \mathbb{Z}_3 = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2)\}$ tiene nueve elementos.

3. $\mathbb{Z} \times \mathbb{Z}_2 = \{(m,a), m \in \mathbb{Z}, a=0, 1\}$ y es un grupo infinito.

4. $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0), (0,1), (1,0), (1,1)\}$. Si construimos la tabla correspondiente, obtenemos:

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,1)	(1,0)	(1,1)
(0,1)	(0,1)	(0,0)	(1,1)	(1,0)
(1,0)	(1,0)	(1,1)	(0,0)	(0,1)
(1,1)	(1,1)	(1,0)	(0,1)	(0,0)

Identificando e con (0,0), a con (0,1), b con (1,0) y c con (1,1), la tabla corresponde a la estructura del grupo de Klein V_4 .

Con este último ejemplo vemos que mediante el producto de grupos podemos obtener grupos ya conocidos.

Queremos señalar, además, que en los grupos de orden 2, 3 o 4, no es necesario construir la tabla en cada caso para verificar que tenemos uno de estos grupos. Dichas tablas conllevan ciertas relaciones definitorias. Así, en el caso de orden 2, tendremos un neutro y un elemento que es su propio simétrico, en el caso de orden 3, un neutro y dos elementos mutuamente simétricos, en V_4 , un neutro y tres elementos que son su propio simétrico, y en $\mathbb{Z}_4$, un neutro, un elemento que es su propio simétrico y otros dos elementos mutuamente simétricos.

Por consiguiente, en el ejemplo anterior para verificar que $\mathbb{Z}_2 \times \mathbb{Z}_2$ poseía la estructura de V_4 , hubiera bastado comprobar que (0,0) es neutro y que

$$(0,1) + (0,1) = (0,0)$$

$$(1,0) + (1,0) = (0,0)$$

$$(1,1) + (1,1) = (0,0)$$

Ejercicios

1. Demuestre que el conjunto de funciones $f_1(x) = x$, $f_2(x) = 1 - x$, $f_3(x) = \frac{1}{x}$, $f_4(x) = \frac{1}{1-x}$, $f_5(x) = 1 - \frac{1}{x}$, $f_6(x) = \frac{x}{x-1}$, es un grupo con la ley de composición.

2. Probar que las matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} w & 0 \\ 0 & w^2 \end{pmatrix}, \begin{pmatrix} w^2 & 0 \\ 0 & w \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & w^2 \\ w & 0 \end{pmatrix}, \begin{pmatrix} 0 & w \\ w^2 & 0 \end{pmatrix}$$

en las que $w^3 = 1$ ($w \neq 1$) forman un grupo multiplicativo.

Buscar todos los grupos de cinco y de seis elementos.

Sugerencia: Seguir el método descrito en la búsqueda de los grupos de dos, tres y cuatro elementos.

4. Sea G un conjunto finito provisto de una ley interna que posee elemento neutro. Pruebe que G es un grupo si y solo si la tabla de multiplicar de la ley de G tiene las propiedades siguientes:

(1) Cada fila y cada columna contiene todos los elementos de G exactamente una vez.

(2) Si para cada par de elementos x , y de G , distintos ambos del neutro, consideramos un rectángulo que tenga como vértices el neutro, x en la misma fila que el neutro y y en la misma columna que el neutro, entonces el cuarto vértice del rectángulo depende solamente del par x , y y no de la posición del neutro.

5. Construya las tablas de multiplicar de $\mathbb{Z}_5$ y $\mathbb{Z}_6$.

3.6 Subgrupos

Al estudiar el concepto de grupo vimos los ejemplos: $GL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A \neq 0\}$, que era un grupo para el producto de matrices, y $SL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A = 1\}$, que también era un grupo para dicho producto. Es evidente que $SL(n, \mathbb{R}) \subset GL(n, \mathbb{R})$ y es un grupo para la operación definida en $GL(n, \mathbb{R})$.

De la misma forma, si tomamos una de las dos tablas de grupo de orden 4, por ejemplo:

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	a	e
c	c	b	e	a

Observemos que los elementos e y a son tales, que sus productos forman una tabla de grupo de orden 2 para la misma ley del grupo de orden 4 al cual pertenecen.

Los ejemplos anteriores nos llevan a considerar el concepto de *subgrupo* que definimos a continuación.

DEFINICIÓN 3.6.1

Un subconjunto H no vacío de un grupo G se llama *subgrupo* de G , si H es un grupo para la ley definida sobre G .

En la práctica se requieren criterios para decidir si un subconjunto dado de un grupo es o no un subgrupo. Estos criterios son los siguientes:

PROPOSICIÓN 3.6.1

Un subconjunto no vacío H de un grupo G es un subgrupo de G si y solo si

(1) $a, b \in H \Rightarrow a \cdot b \in H$ (esto es, la ley de G es una ley interna en H)

(2) $a \in H \Rightarrow a^{-1} \in H$

Demostración

Si H es un subgrupo de G , es evidente por la propia definición de subgrupo que se cumplen (1) y (2).

Recíprocamente, supongamos que H es un conjunto no vacío de G para el cual se verifican (1) y (2), y comprobemos que entonces H es un grupo para la ley de G . Para ello bastará verificar que la ley de G es asociativa para los elementos de H y que $e \in H$.

Está claro que la ley es asociativa en H , ya que es asociativa para todos los elementos de G y como H es no vacío, de (1) y (2) resulta que tomando un elemento cualquiera x de H , tenemos $x \cdot x^{-1} = e \in H$.

Debe observarse que la suposición de que H es no vacío es fundamental para probar que si se cumplen (1) y (2), entonces H es un subgrupo.

Otra definición equivalente de subgrupo es la siguiente:

PROPOSICIÓN 3.6.2

Un subconjunto no vacío H de un grupo G es un subgrupo de G si y solo si $a, b \in H \Rightarrow a \cdot b^{-1} \in H$.

Demostración

Si H es un subgrupo, por la propia definición es fácil comprobar que $a, b \in H \Rightarrow a \cdot b^{-1} \in H$, pues $a \cdot b \in H$ para todo $a, b \in H$ y $b \in H \Rightarrow b^{-1} \in H$.

Recíprocamente, la asociatividad se cumple por ser la ley asociativa en G . Además, si $a \in H$, $a \cdot a^{-1} = e \in H$ y $e \cdot a^{-1} = a^{-1} \in H$.

Por último, $a, b \in H \Rightarrow a(b^{-1})^{-1} = a \cdot b \in H$.

En el caso de los subconjuntos finitos, tenemos una caracterización de subgrupo aún más simple.

PROPOSICIÓN 3.6.3

Si H es un subconjunto finito no vacío de un grupo G , entonces H es un subgrupo de G si y solo si $a, b \in H \Rightarrow a \cdot b \in H$.

Demostración

Si H es un subgrupo, es obvio que se cumple la condición dada en la proposición.

Recíprocamente, bastará probar, por la proposición 3.5.1, que $a \in H \Rightarrow a^{-1} \in H$.

En efecto, si $a \in H$ y se cumple que la ley es interna en H , tenemos que las potencias a^n de a deben estar en H para todo n entero; pero como H es finito, estas potencias no pueden ser todas distintas, luego debe existir m entero positivo tal que $a^m = e$. De aquí que a^{m-1} sea el inverso de a y al ser una potencia de a pertenezca a H .

Veamos ahora algunos ejemplos de subgrupos.

Ejemplos

1. En un grupo $(G, \cdot)$, G y $\{e\}$ son subgrupos de G que se llaman *triviales*. Los subgrupos no triviales de G , o sea, los subgrupos distintos de G y de $\{e\}$ se llaman *propios* y son los de mayor interés.
2. En los conjuntos numéricos $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, cada uno es un subgrupo de los siguientes. Lo mismo ocurre con $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}^*, \cdot)$, $(\mathbb{C}^*, \cdot)$.
3. El grupo estudiado en los ejemplos $SL(n, \mathbb{R})$ es, en realidad, un subgrupo del grupo de las matrices inversibles de orden n sobre $\mathbb{R}$, el cual denotamos por $GL(n, \mathbb{R})$, según vimos en la introducción de este epígrafe.
4. En los ejemplos de grupos dados mediante tablas vimos que solo existían dos posibilidades de grupo de orden 4. El grupo dado por la primera tabla corresponde al grupo $\mathbb{Z}_4$ (llamando $0=e$, $1=a$, $2=b$, $3=c$), y el grupo descrito por la segunda tabla es el grupo de Klein, V_4 . Los subgrupos de $\mathbb{Z}_4$ y V_4 se pueden obtener observando las tablas, ya que para que un subconjunto sea subgrupo es necesario que las posiciones de la tabla correspondientes a los elementos del subconjunto, constituyan a su vez una tabla de grupo. Así de $\mathbb{Z}_4$ el único subgrupo propio es $\{0, 2\}$ y en V_4 hay tres subgrupos propios: $\{e, a\}$, $\{e, b\}$, $\{e, c\}$.

5. Los subgrupos de $(\mathbb{Z}, +)$ tienen una forma especial que nos permite determinarlos completamente, ya que son los conjuntos de la forma $n\mathbb{Z} = \{nx : x \in \mathbb{Z}\}$ para cada $n \in \mathbb{Z}$ fijo.

En efecto, $n\mathbb{Z}$ es un subgrupo de $\mathbb{Z}$ para todo $n \in \mathbb{Z}$ fijo, pues aplicando la proposición 3.6.2 a $n\mathbb{Z}$, tenemos que $nx - ny = n(x - y) \in n\mathbb{Z}$ para $nx, ny \in n\mathbb{Z}$ cualesquiera.

Esto nos demuestra que todo conjunto de la forma $n\mathbb{Z}$ es un subgrupo de $\mathbb{Z}$. Probemos ahora que los subgrupos de $\mathbb{Z}$ son exactamente los $n\mathbb{Z}$, para lo cual usaremos una técnica de demostración a la que debe prestarse atención pues se reiterará a menudo.

Sea H un subgrupo de $\mathbb{Z}$. Si $H = \{0\}$, entonces $H = 0\mathbb{Z}$. Tomemos, pues, $H \neq \{0\}$.

Si $H \neq \{0\}$, existen en H elementos no nulos, lo que nos permite decir que hay en H elementos positivos y negativos, pues al ser un subgrupo, cada elemento de H debe tener su simétrico en H . Si consideramos los elementos positivos de H , entre ellos tiene que existir un elemento mínimo, pues todo subconjunto no vacío de $\mathbb{Z}_+$ tiene un primer elemento. Sea m el mínimo de los elementos positivos de H . Probemos que $H = m\mathbb{Z}$.

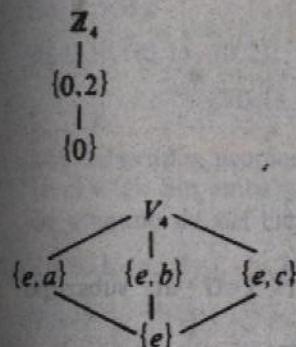
Hay que probar $H \subset m\mathbb{Z}$ y $m\mathbb{Z} \subset H$. Al ser $m \in H$ y H un subgrupo, $mx = \underbrace{m + \dots + m}_{x \text{ veces}} \in H$; de aquí que $m\mathbb{Z} \subset H$.

Sea ahora $n \in H$ y probemos que $n \in m\mathbb{Z}$, o sea, $n = mq$ para algún $q \in \mathbb{Z}$.

Tenemos que $n = mq + r$ con $r = 0$ o $0 < r < m$ por el algoritmo de la división en $\mathbb{Z}$. Luego $r = n - mq$ y de aquí $r \in H$, lo que indica que $r = 0$ pues $r \leq m$ y m es el menor entero positivo tal que $m \in H$. Se deduce que $n = mq$, o sea, $n \in m\mathbb{Z}$ según queríamos probar.

Para representar los subgrupos de un grupo G es usual hacerlo como aparece en el diagrama 3.1

Para $\mathbb{Z}_4$:

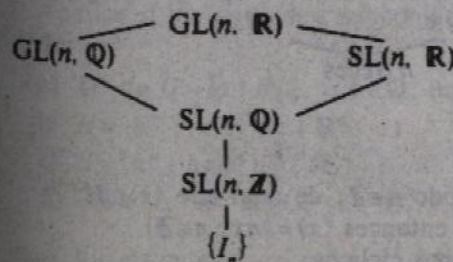


Para V_4 :

$\{0\}$ es subgrupo de $\{0, 2\}$ y de $\mathbb{Z}_4$

$\{0, 2\}$ es subgrupo de $\mathbb{Z}_4$

En el ejemplo 3 vimos que $SL(n, \mathbb{R})$ es un subgrupo de $GL(n, \mathbb{R})$, pero además, existen otras relaciones entre los subgrupos de $GL(n, \mathbb{R})$ que resumimos en el diagrama 3.2.



Aquí $GL(n, \mathbb{Q})$, $SL(n, \mathbb{Q})$, $SL(n, \mathbb{Z})$ representan los grupos de matrices sobre $\mathbb{Q}$ y $\mathbb{Z}$ similares a los considerados sobre $\mathbb{R}$.

Subgrupos cíclicos

En la búsqueda de los subgrupos de $\mathbb{Z}$ encontramos un tipo muy importante de subgrupo: los dados por los múltiplos de un elemento fijo del grupo. Es evidente que si la ley del grupo se denotara multiplicativamente, en vez de múltiplos tendríamos potencias de un elemento fijo. Veamos si lo encontrado en $\mathbb{Z}$ puede demostrarse para un caso general.

PROPOSICIÓN 3.6.4

Sean $(G, \cdot)$ un grupo y $x \in G$ fijo. Entonces el conjunto $H = \{x^n : n \in \mathbb{Z}\}$ es un subgrupo de $(G, \cdot)$.

Demostración

$H \neq \emptyset$ pues $x^0 = e \in H$.

Sean $a, b \in H$; entonces $a = x^n$, $b = x^m$.

$a \cdot b^{-1} = x^n (x^m)^{-1} = x^n \cdot x^{-m} = x^{n-m} \in H$.

Luego H es un subgrupo por la proposición 3.6.2.

DEFINICIÓN 3.6.2

Se llama *subgrupo cíclico* de G generado por $x \in G$ al subgrupo $H = \{x^n : n \in \mathbb{Z}\}$, y se denota por $\langle x \rangle$.

Algunas propiedades de los subgrupos cíclicos son:

(1) Todo subgrupo cíclico es abeliano.

Demostración

Sean $(G, \cdot)$ un grupo y $\langle x \rangle$ un subgrupo cíclico de G . Debemos probar que $a \cdot b = b \cdot a$ para todo $a, b \in \langle x \rangle$. Pero $a = x^n$, $b = x^m$, de donde

$$a \cdot b = x^n \cdot x^m = x^m \cdot x^n = b \cdot a$$

(2) $\langle x \rangle$ es el "menor" subgrupo de G que contiene a x ; esto es, si H es un subgrupo de G que contiene a x , entonces $\langle x \rangle \subset H$.

Demostración

Está claro que $x \in \langle x \rangle$, pues $x = x^1$.

Sea H un subgrupo de G tal que $x \in H$. Al ser H subgrupo, $x^n = \underbrace{x \cdot \dots \cdot x}_{n \text{ veces}} \in H$ para todo $n \in \mathbb{Z}_+$ y $x^{-n} = \underbrace{x^{-1} \cdot \dots \cdot x^{-1}}_{n \text{ veces}} \in H$ para todo $n \in \mathbb{Z}_-$.

Por otra parte, $x^0 = e \in H$.

Podemos, pues, decir que $x^n \in H$ para todo $n \in \mathbb{Z}$; de aquí que $\langle x \rangle \subset H$. Si la ley de G se denota aditivamente, entonces $\langle x \rangle = \{nx : n \in \mathbb{Z}\}$.

Veamos algunos ejemplos de subgrupos cíclicos.

Ejemplos

1. En $\mathbb{Z}$ todos los $n\mathbb{Z}$ son cíclicos. En este caso los subgrupos cíclicos son los únicos subgrupos de $\mathbb{Z}$, pues $\{0\}$ y $\mathbb{Z}$ son de la forma $n\mathbb{Z}$.

2. $H = \left\{ \dots -\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \right\}$ es un subgrupo cíclico de $(\mathbb{Q}, +)$, exactamente $H = \left\langle \frac{1}{2} \right\rangle$.

3. $\mathbb{Z}$ es un subgrupo cíclico de $(\mathbb{Q}, +)$ generado por 1 o por -1 .

4. En $\mathbb{Z}_6$ los únicos subgrupos propios, $\{0, 2, 4\}$ y $\{0, 3\}$, son cíclicos: $\{0, 2, 4\} = \langle 2 \rangle$, $\{0, 3\} = \langle 3 \rangle$. En el caso de $\mathbb{Z}_6$ todos los subgrupos son cíclicos, pues $\mathbb{Z}_6 = \langle 1 \rangle$ y $\{0\} = \langle 0 \rangle$.

5. En $\mathbb{Z}_2 \times \mathbb{Z}_4$,

$$\langle (0, 1) \rangle = \{ (0, 1), (0, 2), (0, 3), (0, 0) \}.$$

$$2(0, 1) \quad 3(0, 1) \quad 4(0, 1)$$

6. Los subgrupos propios de V_4 son todos cíclicos: $\langle e, a \rangle = \langle a \rangle$, $\langle e, b \rangle = \langle b \rangle$, $\langle e, c \rangle = \langle c \rangle$. Sin embargo, V_4 no coincide con el conjunto de las potencias de ninguno de sus elementos.

7. En $(GL(n, \mathbb{Z}), \cdot)$, $H = \left\{ \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}$ es cíclico, ya que $H = \left\langle \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle$.

En los ejemplos de $\mathbb{Z}$ y $\mathbb{Z}_6$ vimos que es posible que el grupo G coincida con el conjunto de las potencias (o de los múltiplos, según el caso) de alguno de sus elementos; esto es, que $G = \langle x \rangle$ para algún $x \in G$. En este caso decimos que G es cíclico.

Los grupos cíclicos son una clase importante de grupos y serán estudiados en detalle en el epígrafe 3.13.

Un grupo G posee siempre subgrupos cíclicos, pero no necesariamente todos sus subgrupos son cíclicos; por ejemplo, en $\mathbb{Z}_2 \times \mathbb{Z}_4$, $H = \{(0, 0), (0, 2), (1, 0), (1, 2)\}$ no es cíclico.

Ejercicios

1. Analice si los siguientes subconjuntos son subgrupos de los grupos dados:

a) $\mathbb{Q}^*$ de $G = \{a + b\sqrt{2} : a, b \in \mathbb{Q}, \text{ no nulos ambos}\}$ con la ley $\cdot$.

b) $H = \{6^n : n \in \mathbb{Z}\}$ de $(\mathbb{R}^*, \cdot)$.

c) $H = \{\pi^n : n \in \mathbb{Z}\}$ de $(\mathbb{R}, +)$.

d) $i\mathbb{R} = \{ix : x \in \mathbb{R}\}$ de $(\mathbb{C}, +)$ y de $(\mathbb{C}^*, \cdot)$.

e) En $(SL(2, \mathbb{Z}), \cdot)$, $H = \left\{ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^n : n \in \mathbb{Z} \right\}$.

f) En $A_1(\mathbb{R}) = \{\varphi_{ab} : \varphi_{ab}(x) = ax + b, a \neq 0, a, b \in \mathbb{R}\}$ con la composición de aplicaciones, el subconjunto $H = \{\varphi_{1,b} : b \in \mathbb{R}\}$.

2. Construya la tabla de los grupos $\mathbb{Z}_2$ y $\mathbb{Z}_{12}$. Busque los subgrupos y haga el esquema de estos. Realice lo mismo para $\mathbb{Z}_2 \times \mathbb{Z}_4$.
3. Busque los subgrupos del grupo $(S_3, \circ)$ de las permutaciones de tres elementos, estudiado como ejemplo al inicio de este capítulo.
4. Pruebe que en un grupo abeliano $(G, \cdot)$ con neutro e , $H = \{x \in G : x^2 = e\}$ es un subgrupo de G .
5. Pruebe que si $(G, \cdot)$ es un grupo finito con neutro e y $a \in G$, entonces existe $n \in \mathbb{N} : a^n = e$.
Sugerencia: Considere el subgrupo $\langle a \rangle$ de G .
6. Pruebe que en un grupo abeliano $(G, \cdot)$ con neutro e , el conjunto $H_n = \{x \in G : x^n = e\}$ para el entero $n \geq 1$ fijo es un subgrupo de G .
7. Pruebe que dados un grupo G y dos subgrupos H y H' de G , entonces $H \cap H'$ es un subgrupo de G . Generalice el resultado al caso de una familia (finita o infinita) de subgrupos de G .
8. Dados un grupo G y dos subgrupos H y H' de G , será $H \cup H'$ un subgrupo de G ? (Analice, por ejemplo, el caso de $\mathbb{Z}_6$). Busque una condición suficiente para que la unión de dos subgrupos de G sea un subgrupo de G .
9. Sea A un subconjunto de un grupo G . Se denomina *centralizador* de A en G y se denota por $C(A)$, el conjunto de los elementos $x \in G$ tales que $ax = xa, \forall a \in A$. En particular, $C(G)$ se denomina *centro* del grupo G . Demuestre que $C(A)$ es un subgrupo de G .
10. Pruebe que si G es un grupo abeliano y H un subgrupo de G , entonces $S(H) = \{x \in G : x^2 \in H\}$ es un subgrupo de G .
11. Sean G un grupo, A un subconjunto de G , $x \in G$, y llamemos $Ax = \{ax : a \in A\}$ y $xA = \{xa : a \in A\}$.
Considere $N(A) = \{x \in G : Ax = xA\}$ y pruebe que $N(A)$ es un subgrupo de G y que $C(A)$, definido en el ejercicio 9, es un subgrupo de $N(A)$.
12. Pruebe que si G_1, G_2 son grupos y H_1, H_2 son subgrupos de G_1 y G_2 respectivamente, entonces $H_1 \times H_2$ es un subgrupo del grupo producto $G_1 \times G_2$.
Generalice el resultado al caso de n grupos $G_1, G_2, \dots, G_n$ y n subgrupos $H_1, H_2, \dots, H_n$ de $G_1, G_2, \dots, G_n$ respectivamente.

3.7 Grupos de transformaciones

Al inicio de este capítulo, en la breve reseña del desarrollo histórico del concepto *grupo*, señalamos cómo, sobre la base de los trabajos de numerosos matemáticos de fines del siglo XVIII y del siglo XIX, Félix Klein tomó este concepto como fundamento para clasificar los distintos tipos de geometrías.

Los grupos utilizados por Klein son los llamados *grupos de transformaciones*, que constituyen subgrupos del grupo de las biyecciones de un conjunto en sí mismo, el cual constituye un grupo para la composición de aplicaciones.

Muchos de los grupos de transformaciones más interesantes surgen al considerar en el grupo de las biyecciones de un conjunto en sí mismo, aquellas que conservan alguna propiedad de dicho conjunto. Por ejemplo, en el conjunto de las biyecciones del plano, al considerar las biyecciones que conservan la distancia, dicho conjunto es un subgrupo del grupo de las biyecciones del plano.

Por otra parte, al considerar los grupos de transformaciones, estudiamos completamente la clase de los grupos, ya que puede demostrarse que todo grupo es estructuralmente idéntico a un cierto grupo de transformaciones.

Estudiemos, pues, estos grupos y sus subgrupos con más detalle.

DEFINICIÓN 3.7.1

Se llama *permutación* de un conjunto E a una biyección de E en E .

Tendremos, entonces, el siguiente resultado, según señalamos anteriormente.

DEFINICIÓN 3.7.2

Sea E un conjunto y sea S_E el conjunto de las permutaciones de E . Entonces S_E es un grupo para la composición de aplicaciones, que se llama *grupo simétrico* de E .

DEFINICIÓN 3.7.3

Se llama *grupo de transformaciones* de E a un subgrupo de S_E .

Pero nos interesan particularmente los conjuntos finitos, donde la definición de permutación toma la forma usualmente conocida, y el estudiante puede observar que la composición de aplicaciones coincide con el producto de permutaciones. En este caso es obvio que si E y F tienen el mismo número de elementos, el grupo de las permutaciones de E es estructuralmente idéntico al grupo de las permutaciones de F , ya que un grupo puede obtenerse del otro sustituyendo cada elemento de E por un y solo un elemento de F . Esta idea la formalizaremos más adelante. Ahora daremos la definición siguiente:

DEFINICIÓN 3.7.4

Sea E un conjunto finito de n elementos. El grupo de permutaciones de E se llama *grupo simétrico de orden n* y se denota por S_n .

PROPOSICIÓN 3.7.1

S_n tiene $n!$ elementos.

Demostración

Sea $E = \{1, 2, \dots, n\}$ —podemos tomar también $E = \{a_1, a_2, \dots, a_n\}$ y llegaremos al mismo resultado— y sea $\sigma \in S_n$.
 $\sigma(1)$ puede ser escogido de n formas; fijemos una de ellas. Entonces a

le quedan $(n-1)$ posibilidades (ya que σ es una biyección); a $\alpha(3)$, $n-3$ posibilidades, y así sucesivamente. Concluimos que S_n tiene $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1 = n!$ elementos.

Ejemplos

1. Habíamos dicho que $S_3 = \{\rho_0, \rho_1, \rho_2, \mu_1, \mu_2, \mu_3\}$, donde

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \mu_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$$\rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \mu_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

La tabla de multiplicar de S_3 es la siguiente:

	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_1	ρ_1	ρ_2	ρ_0	μ_3	μ_1	μ_2
ρ_2	ρ_2	ρ_0	ρ_1	μ_2	μ_3	μ_1
μ_1	μ_1	μ_2	μ_3	ρ_0	ρ_1	ρ_2
μ_2	μ_2	μ_3	μ_1	ρ_2	ρ_0	ρ_1
μ_3	μ_3	μ_1	μ_2	ρ_1	ρ_2	ρ_0

Observemos que este grupo no es abeliano. Realmente es a partir del orden 6 que comenzaremos a encontrar ejemplos de grupos no abelianos, ya que según hemos visto todo grupo de orden 2, 3 o 4 es abeliano y posteriormente veremos que los de orden 5 también lo son.

De la tabla también resulta que $\{\rho_0, \rho_1, \rho_2\}$ es un subgrupo de S_3 que se llama *grupo alternado de orden 3* y se denota por A_3 .

S_3 tiene, además tres subgrupos de orden 2. (En el ejercicio 3 del epígrafe 3.6 se pidió al estudiante que los buscara.)

Existe una correspondencia natural entre los elementos de S_3 y la forma en la cual dos "copias" de un triángulo equilátero de vértices 1, 2, 3 pueden ser situadas para que una "copia" cubra a la otra (figura 3.1). Esto es, exis-

te una correspondencia con transformaciones del plano, tales, que dejan invariantes el triángulo; estas son las rotaciones con centro en el origen y ángulos de 0° , 120° y 240° , y las simetrías con respecto a los ejes E_1 , E_2 y E_3 , que corresponden respectivamente a las permutaciones ρ_0 , ρ_1 , ρ_2 y μ_1 , μ_2 , μ_3 .

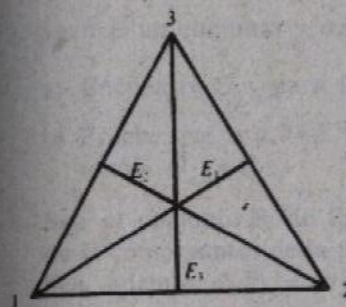


Fig. 3.1

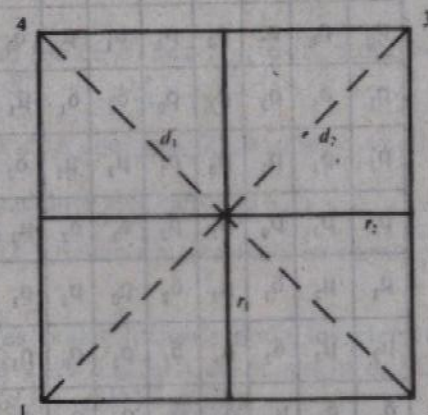


Fig. 3.2

Este conjunto de biyecciones del plano que dejan invariantes el triángulo equilátero, forman un grupo para la composición de aplicaciones que se llama *grupo de las simetrías del triángulo equilátero* o *grupo dihedral*, y se denota por D_3 .

De manera similar se puede encontrar el grupo de simetrías de un polígono regular de n lados, que se denota por D_n , y puede demostrarse que tiene $2n$ elementos. Cada uno de estos grupos de simetrías se corresponde estructuralmente con un subgrupo del grupo de permutaciones S_n .

2. Construyamos el grupo D_4 correspondiente a las formas en que dos "copias" de un cuadrado de vértices 1, 2, 3, 4 pueden ser colocadas una encima de la otra. D_4 será el grupo de las simetrías del cuadrado (figura 3.2).

Usaremos de nuevo ρ_i para denotar las rotaciones, μ_i para las simetrías según los segmentos r_i , y δ_i para las simetrías respecto a las diagonales d_i . Las permutaciones correspondientes son:

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \quad \mu_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

$$\rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \quad \mu_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

$$\rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} \quad \delta_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

$$\rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix} \quad \delta_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}$$

En este caso la tabla de multiplicar es:

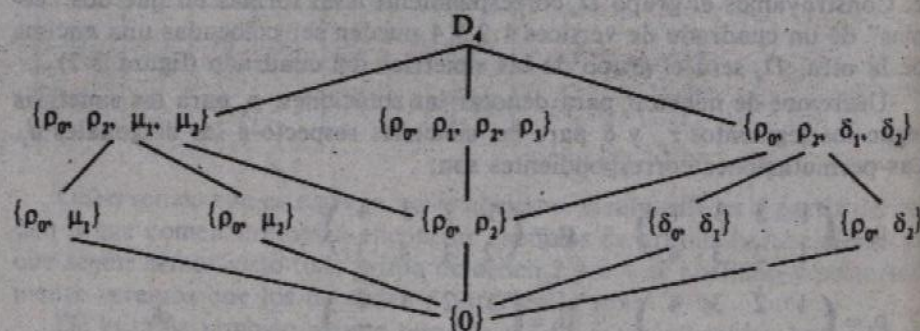
	ρ_0	ρ_1	ρ_2	ρ_3	μ_1	μ_2	δ_1	δ_2
ρ_0	ρ_0	ρ_1	ρ_2	ρ_3	μ_1	μ_2	δ_1	δ_2
ρ_1	ρ_1	ρ_2	ρ_3	ρ_0	δ_2	δ_1	μ_1	μ_2
ρ_2	ρ_2	ρ_3	ρ_0	ρ_1	μ_2	μ_1	δ_2	δ_1
ρ_3	ρ_3	ρ_0	ρ_1	ρ_2	δ_1	δ_2	μ_2	μ_1
μ_1	μ_1	δ_1	μ_2	δ_2	ρ_0	ρ_2	ρ_1	ρ_3
μ_2	μ_2	δ_2	μ_1	δ_1	ρ_2	ρ_0	ρ_3	ρ_1
δ_1	δ_1	μ_2	δ_2	μ_1	ρ_3	ρ_1	ρ_0	ρ_2
δ_2	δ_2	μ_1	δ_1	μ_2	ρ_1	ρ_3	ρ_2	ρ_0



El grupo D_4 tampoco es abeliano.

El conjunto de las rotaciones $\{\rho_0, \rho_1, \rho_2, \rho_3\}$ forma un subgrupo, y también lo forma $\{\rho_0, \rho_2, \delta_1, \delta_2\}$.

En el diagrama 3.3 se dan todos los subgrupos de D_4 :



Hemos dado solamente algunos ejemplos de grupos de transformaciones. Este es un tema amplio y que el estudiante que desee profundizarlo puede utilizar los libros indicados en la bibliografía. De todas formas en los ejercicios daremos algunos ejemplos más de grupos de transformaciones.

Ejercicios

1. Construya la tabla de multiplicar para el subgrupo cíclico de S_5 generado por $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$.
2. Sean E un conjunto y a un elemento fijo de E . Sea $T_a = \{\sigma \in S_E : \sigma(a) = a\}$. Pruebe que T_a es un subgrupo de S_E .
3. Sean E un conjunto y $\sigma \in S_E$. Para $a \in E$ fijo, sea $O_{a,\sigma} = \{\sigma^n(a) : n \in \mathbb{Z}\}$
 - a) Encuentre $O_{1,\sigma}$ para la permutación σ del ejercicio 1.
 - b) Pruebe que si $a, b \in E$ y $O_{a,\sigma}$ y $O_{b,\sigma}$ tienen un elemento en común, entonces $O_{a,\sigma} = O_{b,\sigma}$.
4. Sea el conjunto $\mathbb{R}$ de los números reales (considerando los elementos de $\mathbb{R}$ como puntos de la recta real) y sea $d(a, b) = |a - b|$, dados $a, b \in \mathbb{R}$. Sea, además, $I(\mathbb{R}) = \{\sigma \in S_{\mathbb{R}} : d(\sigma(a), \sigma(b)) = d(a, b)\}$. Pruebe que $I(\mathbb{R})$ es un grupo de transformaciones de $\mathbb{R}$. Este grupo recibe el nombre de grupo de las isometrías de $\mathbb{R}$.
5. Al igual que en el problema 4, consideremos $\mathbb{R}^2$ como puntos del plano y, dados $A = (x_A, y_A)$, $B = (x_B, y_B)$, sea

$$d(A, B) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}.$$

- a) Pruebe que $I(\mathbb{R}^2) = \{\sigma \in S : d(\sigma(A), \sigma(B)) = d(A, B)\}$ es un subgrupo de S . Este subgrupo de S se llama grupo de las isometrías del plano.
- b) Pruebe que $T = \{\tau_{a,b} \text{ dado por } \tau_{a,b}(x, y) = (x + a, y + b), a, b \in \mathbb{R}\}$ es un subgrupo de $I(\mathbb{R}^2)$. Este subgrupo de $I(\mathbb{R}^2)$ se llama grupo de las traslaciones de $\mathbb{R}^2$.
- c) Pruebe que $R = \{\rho_\theta, \rho_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)\}$ es un subgrupo de $I(\mathbb{R}^2)$. Este subgrupo de $I(\mathbb{R}^2)$ se llama subgrupo de las rotaciones con centro en el origen.

3.8 Isomorfismos de grupos y teorema de Cayley

En nuestra exposición anterior hemos trabajado con la idea de grupos G y G' estructuralmente idénticos, es decir, los grupos eran idénticos con excepción de los nombres de los elementos y de las operaciones, por lo que po-

diamos obtener G' de G llamando a cada elemento x de G con el nombre de un cierto elemento x' de G' .

En el caso de los grupos finitos, lo anterior se traduciría en que las tablas de multiplicar de ambos grupos eran idénticas, una vez establecido qué elemento de G' sustituimos por cada elemento de G . Esto se vio de manera natural en la correspondencia existente entre los elementos de S_3 y las transformaciones que dejan invariante el triángulo equilátero y que constituyen el grupo D_3 de las simetrías del triángulo equilátero.

Es natural, pues, formalizar estas ideas mediante la definición siguiente:

DEFINICIÓN 3.8.1

Sean G y G' grupos. Se llama *isomorfismo de G en G'* a una aplicación $f: G \rightarrow G'$ que cumple las condiciones siguientes:

- (1) f es biyectiva.
- (2) $f(x \cdot y) = f(x) \cdot f(y)$ para todo $x, y \in G$.

Decimos, entonces, que los grupos G y G' son *isomorfos* y escribimos $G \cong G'$. En el caso particular en que $G = G'$ y coincidan las leyes internas en ambos conjuntos (esto es, que los grupos G y G' sean idénticos), decimos que tenemos un *automorfismo de G* .

En los siguientes ejemplos, en lugar de los conjuntos escribiremos el par $(G, \cdot)$ para indicar a qué ley interna nos estamos refiriendo en cada caso. Quede claro que las aplicaciones se establecen entre los conjuntos.

Ejemplos de isomorfismos de grupos.

1. $f: (\mathbb{R}^+, \cdot) \rightarrow (\mathbb{R}, +)$ dada por $x \mapsto \ln x$
2. $g: (\mathbb{R}, +) \rightarrow (\mathbb{R}^+, \cdot)$ dada por $x \mapsto e^x$ (es la inversa de f del ejemplo 1).
3. $h: (\mathbb{Z}, +) \rightarrow (n\mathbb{Z}, +)$ dada por $m \mapsto nm$.

Estos son ejemplos de isomorfismos. Comprobémoslo en el caso del ejemplo 1.

En efecto, sabemos de los cursos de Análisis matemático que $\ln x$ es una aplicación biyectiva de $\mathbb{R}^+$ sobre $\mathbb{R}$:

$$f(x \cdot y) = \ln x \cdot y = \ln x + \ln y = f(x) + f(y).$$

4. $\varphi: (\mathbb{Z}, +) \rightarrow (\mathbb{Z}, +)$ dada por $n \mapsto n+1$ no es un isomorfismo, ya que aunque es biyectiva, pues $n+1=m+1 \Rightarrow n=m$ (inyectiva) y $\forall n \in \mathbb{Z}, \exists n-1$ tal que $\varphi(n-1)=n-1+1=n$ (sobreyectiva), no se cumple que $\varphi(n+m)=\varphi(n)+\varphi(m)$; por ejemplo, $\varphi(1+1)=\varphi(2)=3$, $\varphi(1)+\varphi(1)=2+2=4$.

Al estudiar los grupos de transformaciones, vimos cómo entre S_3 y D_3 aparecía de manera natural una correspondencia biyectiva. Sin embargo, hay que comprobar de todas formas que se cumple la condición (2) de la definición 3.8.1, para lo cual es necesario comparar las tablas de multiplicar de dichos grupos, pues aunque intuitivamente vemos que estos tie-

nen la misma estructura, debemos formalizarlo. Esta forma de trabajo para los grupos finitos vamos a mostrarla con un ejemplo sencillo.

5. Sean los grupos G y H dados por las tablas:

	-1	1
-1	1	-1
1	-1	1

+	0	1
0	0	1
1	1	0

Debemos encontrar una aplicación $f: G \rightarrow H$ que cumpla las condiciones de isomorfismo. La que aparece de forma más natural es la siguiente:

$$1 \mapsto 0$$

$$-1 \mapsto 1$$

Evidentemente esta aplicación es biyectiva, con lo que comprobamos la condición (1) de la definición. Para comprobar la (2), debemos considerar todas las combinaciones posibles. Así:

$$f(1 \cdot 1) = f(1) = 0 = 0 + 0 = f(1) + f(1)$$

$$f((-1) \cdot 1) = f(-1) = 1 = 1 + 0 = f(-1) + f(1)$$

$$f((-1) \cdot (-1)) = f(1) = 0 = 1 + 1 = f(-1) + f(-1)$$

$$f(1 \cdot (-1)) = f(-1) = 1 = 0 + 1 = f(1) + f(-1)$$

Luego f es un isomorfismo de G en H .

Como vemos, la comprobación de la condición (2) es muy laboriosa, mas aún si el orden del grupo crece. Es por ello que en la práctica se acude a otros métodos. En este caso si observamos lo que hicimos, podemos definir intuitivamente la correspondencia asignando al neutro de G el de H , más adelante, al estudiar las propiedades de los isomorfismos, dispondremos otros recursos que nos ayudarán a determinar si los dos grupos dados son isomorfos o no.

Un ejemplo más teórico y de importancia futura es el siguiente:

6. En un grupo $(G, \cdot)$, dado $a \in G$ fijo, la aplicación $i_a: G \rightarrow G$ dada por $i_a(x) = axa^{-1} \forall x \in G$, es un automorfismo de G llamado *automorfismo interior*.

- (1) i_a es biyectiva.

$$i_a(x) = i_a(y) \iff axa^{-1} = aya^{-1} \Rightarrow x = y \text{ (inyectividad)}$$

$$\forall y \in G \text{ existe } x = a^{-1}ya \in G \text{ tal que}$$

$$i_a(x) = i_a(a^{-1}ya) = a(a^{-1}ya)a^{-1} = y \text{ (sobreyectividad)}$$

$$(2) i_a(x \cdot y) = axya^{-1} = axa^{-1} \cdot aya^{-1} = (axa^{-1})(aya^{-1}) = i_a(x) \cdot i_a(y)$$

Propiedades de los isomorfismos

PROPOSICIÓN 3.8.1

(1) Si G y G' son grupos, f un isomorfismo de G en G' y 1 el elemento neutro de G , entonces $f(1)$ es el elemento neutro de G' . Además, $f(a^{-1}) = f(a)^{-1}$ para todo $a \in G$.

(Es muy natural, dada la idea de isomorfismo, que se cumpla la propiedad anterior, que plantea que un isomorfismo transforma neutro en neutro e inverso en inverso.)

Demostración

Sea $x' \in G'$. Como f es sobreyectiva, existe $x \in G$ tal que $f(x) = x'$, pero $x' = f(x) = f(x \cdot 1) = f(x) \cdot f(1) = x' \cdot f(1)$ para todo $x' \in G'$. Similarmente probamos $f(1) \cdot x' = x'$, de donde $f(1)$ es el elemento neutro de G' .

(2) Para el inverso de un elemento a , se cumple $f(a^{-1}) = [f(a)]^{-1}$.

Demostración

$a \cdot a^{-1} = 1 \Rightarrow f(a \cdot a^{-1}) = f(1)$, pero $f(a \cdot a^{-1}) = f(a) \cdot f(a^{-1}) = f(1)$. Similarmente probamos $f(a^{-1}) \cdot f(a) = f(1)$, de donde resulta que $f(a^{-1}) = [f(a)]^{-1}$. (Obsérvese que hemos usado la unicidad del neutro y del inverso.)

(3) Si $f: G \rightarrow G'$ es un isomorfismo, entonces $f^{-1}: G' \rightarrow G$ es un isomorfismo.

Demostración

a) f^{-1} es una biyección, ya que la inversa de una biyección es una biyección.

b) Sea $f^{-1}(x' \cdot y')$. Como f es sobreyectiva, existen $x, y \in G$ tales que $x' = f(x)$, $y' = f(y)$.

Entonces

$$f^{-1}(x' \cdot y') = f^{-1}(f(x) \cdot f(y)) = f^{-1}(f(x \cdot y)) = x \cdot y = f^{-1}(x') \cdot f^{-1}(y').$$

(Hemos utilizado que f y f^{-1} son biyecciones.)

(4) Si $f: G \rightarrow G'$ es un isomorfismo y $g: G' \rightarrow G''$ es un isomorfismo, entonces $g \circ f: G \rightarrow G''$ es un isomorfismo.

Demostración

a) Como f y g son biyecciones, $g \circ f$ será una biyección de G sobre G'' .

$$\begin{aligned} b) (g \circ f)(x \cdot y) &= g(f(x \cdot y)) = g(f(x) \cdot f(y)) = g(f(x)) \cdot g(f(y)) = \\ &= (g \circ f)(x) \cdot (g \circ f)(y). \end{aligned}$$

(5) Si H es un subgrupo de G , entonces $f(H)$ es un subgrupo de G' , y si H' es un subgrupo de G' , entonces $f^{-1}(H')$ es un subgrupo de G .

En las demostraciones de (1), (2) y (3) hemos dado la técnica a seguir, por lo que dejamos la demostración de la propiedad (5) como un ejercicio para el estudiante.

Las propiedades de los isomorfismos formalizan la idea de que los grupos isomorfos son "copias" distintas de un conjunto con una operación dada. De aquí que podamos considerar como lo importante a estudiar las ope-

raciones y sus propiedades, independientemente de la naturaleza de los elementos del conjunto de que se trate.

Hemos dado anteriormente otra idea que es necesario formalizar. Nos referimos a las afirmaciones de este tipo: "Hay un único grupo de orden 3." Estas afirmaciones llevan implícita la idea de una clasificación dentro de la clase de los grupos, en la cual se consideran los grupos que son estructuralmente idénticos y los que no lo son. Esta clasificación está basada en las siguientes propiedades de los isomorfismos:

- (1) Todo grupo G es isomorfo a sí mismo (id_G es un isomorfismo).
- (2) Si G es isomorfo a G' , G' es isomorfo a G (el inverso de un isomorfismo es un isomorfismo).
- (3) Si G es isomorfo a G' y G' a G'' , entonces G es isomorfo a G'' (la composta de dos isomorfismos es un isomorfismo).

Estas propiedades de los isomorfismos permiten dividir los grupos en clases disjuntas dos a dos, de modo que dos grupos pertenecen a una misma clase si y solo si son isomorfos, lo cual probaremos posteriormente. Esta división según isomorfismo clasifica totalmente los grupos.

En un epígrafe posterior introduciremos el concepto *relación de equivalencia*, y el estudiante podrá comprobar que la relación entre grupos dada por: *dos grupos están relacionadas si y sólo si son isomorfos*, es una relación de equivalencia.

Las propiedades de las relaciones de equivalencia justificarán las afirmaciones que acabamos de hacer con respecto a la división de los grupos en clases de grupos isomorfos. Así, cuando decimos: "existe un único grupo de orden 3, salvo isomorfismo", queremos decir que todos los grupos de orden 3 existentes son isomorfos entre sí. De la misma forma podemos decir que existen dos únicos grupos de orden 4, salvo isomorfismo: $\mathbb{Z}_4$ y el grupo de Klein.

Nos queda un problema pendiente: en caso de encontrar grupos no isomorfos, "cómo probarlo?"

Hay que ser cuidadoso con este punto pues, por ejemplo, a primera vista parece que $\mathbb{Z}$ y $n\mathbb{Z}$ no son isomorfos y, sin embargo, hemos comprobado que lo son: de modo que el hecho de ser uno subconjunto del otro no es fundamental en el caso infinito. En cuanto al caso finito, el hecho de tener o no el mismo número de elementos es una condición necesaria para la existencia de un isomorfismo, pero no es suficiente. Para ilustrar esto hemos dado el ejemplo de $\mathbb{Z}_4$ y V_4 .

¿Qué determina, en el caso de $\mathbb{Z}_4$ y V_4 , el que no sean isomorfos, es decir, estructuralmente idénticos?

Lo determina el hecho de que en V_4 todo elemento es su propio inverso y en $\mathbb{Z}_4$ no; esto es, las propiedades algebraicas de V_4 no se trasladan exactamente a $\mathbb{Z}_4$.

Este tipo de situación de no conservación de una propiedad inherente a la estructura algebraica del grupo, es la que nos permitirá decidir si dos grupos dados no son isomorfos.

Ejemplos

1. $(\mathbb{Q}^*, \cdot)$ no es isomorfo a $(\mathbb{R}^*, \cdot)$ ya que, por ejemplo, la ecuación $x^2 = 2$ tiene solución en $\mathbb{R}^*$ y no en $\mathbb{Q}^*$.
2. $(\mathbb{R}^*, \cdot)$ no es isomorfo a $(\mathbb{C}^*, \cdot)$ pues, por ejemplo, $x^2 = -1$ tiene solución en $\mathbb{C}^*$ y no en $\mathbb{R}^*$.

Teorema de representación de Cayley

Como ya hemos señalado, históricamente los grupos aparecieron en su mayoría como grupos de transformaciones de algún ente matemático. El siguiente teorema, debido al matemático inglés Cayley, muestra que *salvo isomorfismo* todo grupo es un grupo de transformaciones de un cierto conjunto adecuado (que según veremos en la demostración resultará ser el mismo G). Si G es finito será, salvo isomorfismo, un subgrupo de un S_n .

TEOREMA 3.8.1 (Teorema de representación de Cayley)

Cada grupo G es isomorfo a un cierto subgrupo de S_X para algún X adecuado.

Demostración

Antes de pasar a la demostración, consideremos la tabla de multiplicar de un grupo finito, por ejemplo, $\mathbb{Z}_3$:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Observemos que a cada elemento a de $\mathbb{Z}_3$, le corresponde de forma natural un elemento de $S_{\{0,1,2\}}$ dado por la fila (o la columna) correspondiente al elemento a . Por ejemplo:

$$0 \rightarrow (0, 1, 2) = e_1$$

$$1 \rightarrow (1, 2, 0) = e_2$$

$$2 \rightarrow (2, 0, 1) = e_3$$

Si construimos la tabla de multiplicar de las permutaciones e_1, e_2, e_3 , comprobamos que forman un subgrupo de orden 3 de $S_{\{0,1,2\}}$ isomorfo a $\mathbb{Z}_3$.

Por otra parte, $(0, 1, 2)$, por ejemplo, resulta de hallar los productos $0+0$, $0+1$ y $0+2$, o sea, las imágenes de los elementos de $\mathbb{Z}_3$ por la aplicación $x \rightarrow 0+x$. Esto nos da una idea de cómo proceder en el caso general, ya que en el ejemplo anterior está claro que el X buscado es el propio $\mathbb{Z}_3$, y el subconjunto de S_X está dado por aplicaciones de una forma especial.

Así que debemos buscar un subconjunto de las biyecciones de G en sí mismo, que esté en correspondencia biyectiva con G y que además constituya un subgrupo de S_G . Este problema lo resolveremos con el siguiente tipo de biyecciones.

Sea $a \in G$. Definamos $i_a : G \rightarrow G$ dada por $i_a(x) = ax$, para todo $x \in G$.

i_a es una biyección de G , ya que si $i_a(x) = i_a(y)$, entonces $ax = ay \Rightarrow x = y$, o sea, i_a es inyectiva, y también es sobreyectiva, pues para cada $y \in G$ podemos tomar $x = a^{-1}y$ y $i_a(a^{-1}y) = a \cdot a^{-1}y = y$.

(A i_a se le llama *multiplicación por a a la izquierda* o también *traslación por a a la izquierda*.)

Consideramos ahora el siguiente subconjunto G' de S_G :

$$G' = \{i_a : a \in G\}.$$

es decir, el conjunto de las multiplicaciones a la izquierda por todos los elementos de G .

Probemos que G' es un subgrupo de S_G .

Para ello sean i_a y i_b en G' y probemos que $i_a \circ i_b \in G'$, que G' posee elemento neutro y que todo elemento de G' posee simétrico en G' . No es necesario probar la asociatividad, según vimos en las caracterizaciones de subgrupo.

Hallemos $i_a \circ i_b$.

$$i_a \circ i_b(x) = i_a(i_b(x)) = i_a(bx) = a(bx) = (ab)x, \text{ o sea } i_a \circ i_b = i_{ab}.$$

Y como $i_{ab} \in G'$, queda probado que G' está cerrado para el producto de permutaciones.

En cuanto al elemento neutro, este es i_1 , donde 1 es el elemento neutro de $(G, \cdot)$.

Por último, si $i_a \in G'$, tomando $i_{a^{-1}}$ obtenemos:

$$i_a \circ i_{a^{-1}} = i_{a \cdot a^{-1}} = i_1, \quad i_{a^{-1}} \circ i_a = i_{a^{-1} \cdot a} = i_1,$$

lo que prueba que $i_{a^{-1}}$ es el simétrico de i_a .

Para tener demostrado el teorema de Cayley falta encontrar un isomorfismo entre G y G' , y es natural tomar la siguiente correspondencia:

$$\varphi: G \rightarrow G' \text{ dada por } a \mapsto i_a.$$

φ es una biyección pues $\varphi(a) = \varphi(b) \Rightarrow i_a = i_b$.

En efecto, $i_a = i_b$ quiere decir que $i_a(x) = i_b(x)$ para todo $x \in G$, o sea, $ax = bx$ para todo $x \in G$.

Tomando $x=1$ obtenemos $a=b$ y φ será inyectiva. Además, por la definición de φ , es obvio que φ es sobreyectiva. Falta probar que $\varphi(ab) = \varphi(a) \circ \varphi(b)$; pero realmente ya lo probamos, pues

$$\varphi(ab) = i_{ab} = i_a \circ i_b = \varphi(a) \circ \varphi(b).$$

Así queda demostrado el teorema de Cayley.

El grupo G' obtenido en el teorema de representación de Cayley se llama *representación regular izquierda de G* . Podemos obtener una representación regular derecha G'' usando las aplicaciones d_a dadas por $d_a(x) = xa$ y haciendo una demostración similar del teorema, pero en este caso habrá que definir el isomorfismo así:

$$\psi: G \longrightarrow G'' \text{ dado por } a \sim d_a.$$

La palabra *representación* en el nombre del teorema de Cayley corresponde al hecho de que representamos el grupo mediante permutaciones, o sea, podemos sustituir en un momento dado los elementos del grupo por permutaciones.

Veamos un ejemplo de la aplicación de este teorema.

Ejemplo:

Sea el grupo G dado por la siguiente tabla y busquemos una representación regular izquierda de G .

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Busquemos el subgrupo de S_3 que lo representa. Para ello debemos encontrar i_e, i_a, i_b .

i_e es la aplicación tal que $i_e(x) = ax$ para todo $x \in G$.

Luego

$$i_e(e) = a, i_e(a) = b, i_e(b) = e.$$

$$\text{Así, } i_e = \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix} = p_1.$$

De modo similar,

$$i_a = \begin{pmatrix} e & a & b \\ e & a & b \end{pmatrix} = p_0, \quad i_b = \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix} = p_2.$$

De modo que la representación regular izquierda de G es A_1 .

Ejercicios

1. Divida los siguientes grupos en clases según la relación de isomorfismo.

$$(\mathbb{Z}, +), \mathbb{Z}_6, \mathbb{Z}_2, S_3, S_4, (17\mathbb{Z}, +), (\mathbb{Q}, -),$$

$$(\mathbb{R}, +), (\mathbb{R}^*, \cdot), (\mathbb{Q}^*, \cdot), (\mathbb{C}^*, \cdot).$$

2. Sea $(G, \cdot)$ un grupo. Se define en G la siguiente ley $a * b = b \cdot a$. Pruebe que $(G, *)$ es un grupo isomorfo a $(G, \cdot)$.

Sugerencia: Considere la aplicación $\varphi: G \longrightarrow G$ dada por $\varphi(a) = a^{-1}$.

3. Pruebe que si G es un grupo finito y H es un grupo infinito, entonces H y G no son isomorfos.

4. Pruebe que si G es un grupo finito y H es un subgrupo de G , $H \neq G$, entonces G y H no son isomorfos.

5. Diga dos razones por las cuales $\mathbb{Z}_4$ no es isomorfo al grupo de Klein.

6. Pruebe que un grupo G es abeliano si y solo si la correspondencia $a \rightarrow a^{-1}$ de G en G es un automorfismo.

7. ¿Cuántos automorfismos tiene:

$$a) \mathbb{Z}_2, b) \mathbb{Z}_6, c) \mathbb{Z}, d) \mathbb{Z}_7?$$

8. Sea $S = \mathbb{R} \setminus \{-1\}$ con la ley $a * b = a + b + ab$ (que sabemos por el ejercicio 2 del epígrafe 3.3 que es un grupo). Pruebe que $(S, *)$ es isomorfo a $(\mathbb{R}^*, \cdot)$.

9. Busque la representación regular izquierda de $\mathbb{Z}_4$ y la representación regular derecha de S_3 .

10. Sea G un grupo abeliano finito de orden m y sea n un entero primo relativo con m . Pruebe que todo $g \in G$ puede escribirse como $g = x^n$ con $x \in G$.

Sugerencia: Considere la aplicación $\phi: G \longrightarrow G$ definida por $\phi(y) = y^n$ y pruebe que es un automorfismo de G .

11. Pruebe que en un grupo G el conjunto $\text{Aut } G = \{f: G \longrightarrow G \text{ automorfismo}\}$ es un grupo para la composición de aplicaciones y que $\text{Inn } G = \{i_a: a \in G\}$ donde i_a es un automorfismo interior, es un subgrupo de $\text{Aut } G$.

12. Sean G un grupo finito y φ un automorfismo de G con la propiedad $\varphi(x) = x \iff x = e$ y tal que $\varphi^2 = \varphi \circ \varphi = \text{id}_G$. Pruebe que G debe ser abeliano.

Sugerencia: a) Pruebe que $\forall g \in G$ puede escribirse como $g = x^{-1} \cdot \varphi(x)$ para algún x .

b) Pruebe que φ coincide con la aplicación $g \mapsto g^{-1}$ de G en G .

3.9 Relaciones de equivalencia

El problema que estudiaremos ahora consiste en construir, a partir de un grupo dado, un grupo relacionado con este, al que llamaremos *grupo cociente*. Veremos que si conocemos la estructura del grupo cociente, conoceremos la de G en gran medida, esto es, podremos sacar del cociente información sobre G . Este grupo cociente lo formaremos a partir de cierta clasificación de los elementos de G y para ello debemos estudiar previamente el concepto de *relación de equivalencia*.

Las relaciones de equivalencia aparecen en todas las ramas de la Matemática y se usan con significados diversos. Por ejemplo, en Álgebra lineal cuando hablamos de sistemas de ecuaciones lineales equivalentes, estamos clasificando los sistemas según tengan o no el mismo conjunto solución, y en Aritmética cuando decimos que $\frac{1}{2}$ y $\frac{2}{4}$ representan el mismo número racional, estamos estableciendo una clasificación en el conjunto de los pares (m, n) con $m, n \in \mathbb{Z}$ y $n \neq 0$.

DEFINICIÓN 3.9.1

Se dice que existe una *relación binaria* en un conjunto E , si para cada par ordenado (a, b) de elementos de E se puede determinar si a está relacionado o no con b .

Las relaciones se denotan por las letras R, S, T , etc. Se escribe, por ejemplo, $a R b$ y se lee: a está relacionado con b .

DEFINICIÓN 3.9.2

La relación binaria R en un conjunto E se llama *relación de equivalencia* si y solo si se cumple:

- (1) $a R a$ para todo $a \in E$ (reflexividad).
- (2) $a R b \Rightarrow b R a$ (simetría)
- (3) $a R b$ y $b R c \Rightarrow a R c$ (transitividad)

La definición anterior se puede expresar en un lenguaje más formal diciendo que una relación de equivalencia en un conjunto E es un subconjunto R de $E \times E$ tal que:

- (1) Para todo $a \in E$, $(a, a) \in R$

- (2) $(a, b) \in R \Rightarrow (b, a) \in R$.

- (3) $(a, b) \in R$ y $(b, c) \in R \Rightarrow (a, c) \in R$

Esto es válido ya que en definitiva una relación R determina un subconjunto de $E \times E$: el de los pares cuyos elementos están relacionados por ella, y reciprocamente: dado un subconjunto R de $E \times E$, existe una regla para decidir si $a, b \in E$ están relacionados, esto es, si $(a, b) \in R$.

Ejemplos

Son relaciones de equivalencia:

1. La relación idéntica en E dada por $x R y \iff x = y$
2. La relación trivial en un conjunto E cualquiera: $x R y \forall x, y \in E$.
3. En el conjunto E de los puntos del plano $a R b \iff a$ y b pertenecen a una misma recta horizontal. En efecto:

$a R a$	porque por todo punto puede pasar una recta horizontal (paralela a OX).
$a R b \Rightarrow b R a$	ya que la recta determinada por dos puntos es única.
$a R b$ y $b R c \Rightarrow a R c$	pues si a y b pertenecen a la misma recta horizontal, como esta en definitiva está determinada por b, c pertenecerá a esta misma recta, con lo cual $a R c$.
4. En el conjunto E de los puntos del plano, $a R b$ si y solo si a y b equidistan de un mismo punto O .
5. Sea el conjunto $\mathbb{Z}$ de los números enteros y $n > 1$ un entero fijo. Definamos en $\mathbb{Z}$ la relación

$$a R b \iff a - b \text{ es múltiplo de } n \text{ (esto es, existe } q \in \mathbb{Z} : a - b = nq).$$

Esta relación es de equivalencia porque

R es reflexiva pues $a - a = 0 = 0 \cdot q$.

R es simétrica ya que $a R b \iff a - b = nq \Rightarrow b - a = (-q) \cdot n \Rightarrow b R a$.

R es transitiva porque

$$\left. \begin{array}{l} a R b \iff a - b = q \cdot n \\ b R c \iff b - c = q' \cdot n \end{array} \right\} \Rightarrow a - c = (q + q') \cdot n \iff a R c$$

La relación definida recibe el nombre de *congruencia módulo n* , se denota frecuentemente por $a \equiv b \pmod{n}$ y se lee: a es congruente b módulo n . Esta relación tendrá gran importancia en lo que sigue.

6. En $\mathbb{N}^*$ la relación $a R b \iff \exists q \in \mathbb{N}^* : a = bq$ no es de equivalencia, ya que por ejemplo $6 R 3$ pero $3 \not R 6$, o sea, que R no es simétrica.

7. En $\mathbb{R}$ la relación $aRb \iff \exists c \in \mathbb{R}^+ : a+c=b$ no es de equivalencia, pues no es reflexiva ni simétrica.

Un ejemplo más teórico de relación de equivalencia, con el cual tendremos que trabajar próximamente, es el siguiente:

8. Sean E y F conjuntos y $f: E \rightarrow F$ una aplicación. Podemos definir sobre E la siguiente relación de equivalencia, que llamaremos *relación de equivalencia asociada a f* :

Dados $x, y \in E$, se cumple:

$$xRy \iff f(x) = f(y).$$

Comprobemos que R es una relación de equivalencia.

En efecto, xRx ya que la imagen por una aplicación de un punto del dominio es única.

$$xRy \Rightarrow yRx \text{ pues } f(x) = f(y) \Rightarrow f(y) = f(x)$$

$$xRy, yRz \Rightarrow xRz \text{ pues } f(x) = f(y), f(y) = f(z) \Rightarrow f(x) = f(z)$$

Dejamos al estudiante que compruebe las afirmaciones hechas en los ejemplos sobre si las relaciones dadas son de equivalencia o no, en los casos en que esto no está realizado en forma explícita.

Hemos introducido las relaciones de equivalencia con el objetivo de obtener una clasificación de los elementos de un conjunto dado, veamos ahora cómo logramos esto.

Según vimos, la relación da una regla para determinar, dado $a \in E$, cuáles elementos de E están relacionados con a y cuáles no. Consideremos la definición siguiente:

DEFINICIÓN 3.9.3

Sean R una relación de equivalencia en un conjunto E y $a \in E$. El conjunto $C_a = \{b \in E : bRa\}$ se llama *clase de equivalencia de a según R* . Se denota también por $\bar{a}$ o $\bar{a}$.

Observemos que dada la reflexividad de R , $a \in C_a$; luego ninguna clase de equivalencia según R es vacía.

Veamos qué forma toman las clases en algunos de los ejemplos de relaciones dados anteriormente. Dejamos al estudiante que haga lo mismo para el resto de los ejemplos.

En el ejemplo 1 las clases consisten en conjuntos unitarios.

En el ejemplo 3 la clase de equivalencia de un punto a del plano es la recta horizontal que pasa por a , y en el ejemplo 4 es la circunferencia con centro O y que pasa por a .

En el ejemplo de la congruencia módulo n hay que realizar algunos cálculos para hallar la clase $n \in \mathbb{Z}$.

$$C_n = \{m \in \mathbb{Z} : m-r=nq \text{ para algún } q \in \mathbb{Z}\},$$

luego $m=r+nq$, que no es más que la expresión del algoritmo de la división

con resto para m y n . Vemos que m es tal que su resto en la división por n debe ser r .

Como por el algoritmo antes citado sabemos que $0 \leq r < n$, tenemos que en este caso hay exactamente n clases:

$$C_0, C_1, \dots, C_{n-1}$$

Por ejemplo, para $m=2$ hay dos clases:

$$C_0 = \{0, 2, \dots, 2n, \dots\} \text{ y } C_1 = \{1, 3, \dots, 2n+1, \dots\}.$$

Por último, en el ejemplo 9 y si recordamos de la teoría de conjuntos que dadas una aplicación $f: E \rightarrow F$ y $z \in F$, se tiene que $f^{-1}(z) = \{x \in E : f(x) = z\}$ es un subconjunto de E que recibe el nombre de *preimagen de z por f* , podemos ver que la clase de $x \in E$ según la relación dada en el ejemplo 9, es $C_x = f^{-1}(f(x)) = \{y \in E : f(y) = f(x)\}$. (Recuérdese que $f^{-1}(z)$ es simplemente una notación, ya que para una aplicación cualquiera f , no tiene por qué existir f^{-1} .)

En todos los ejemplos anteriores podemos notar que si consideramos el conjunto de las clases de equivalencia de todos los puntos de E según R , dichas clases constituyen familias de subconjuntos no vacíos de E tales que su unión es E y los subconjuntos distintos son disjuntos dos a dos. Representemos lo anterior gráficamente en el caso de las clases de equivalencia de los ejemplos 3 y 4 (figuras 3.3 y 3.4).

De lo que hemos explicado intuimos que las clases de equivalencia nos darán la clasificación de los elementos de E buscada. Formalicemos estas ideas.

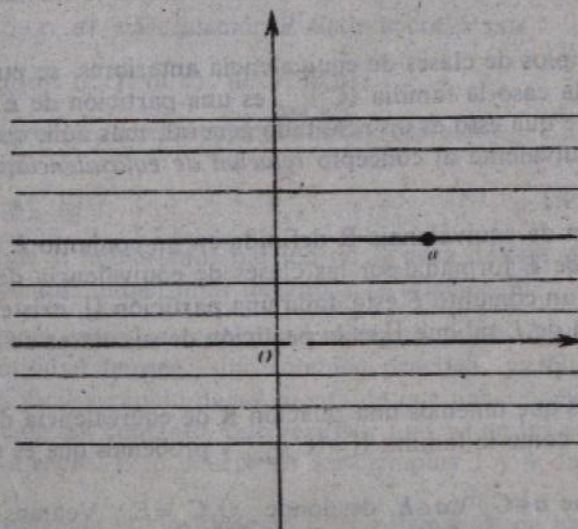


Fig. 3.3

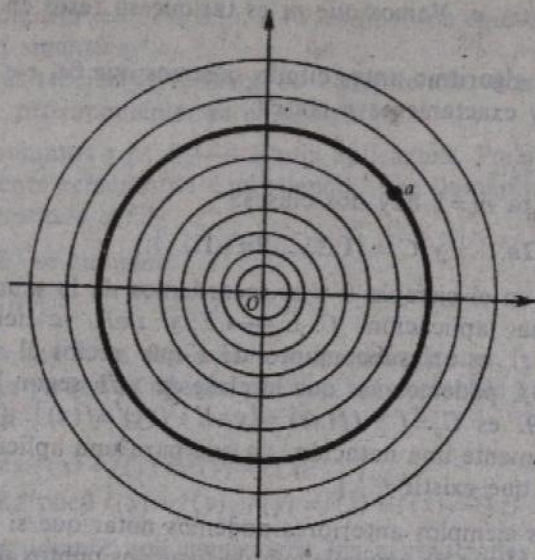


Fig. 3.4

Relaciones de equivalencia y partición

DEFINICIÓN 3.9.4

Se llama *partición de un conjunto E* a una familia de subconjuntos no vacíos de *E* tales que la unión de dichos subconjuntos es *E* y los subconjuntos distintos son disjuntos. Los subconjuntos de la familia se llaman *clases de la partición*.

En los ejemplos de clases de equivalencia anteriores, se puso en evidencia que en cada caso la familia $\{C_a\}_{a \in E}$ es una partición de *E*. Probaremos inmediatamente que esto es un resultado general, más aún, que el concepto *partición* es equivalente al concepto *relación de equivalencia*.

TEOREMA 3.9.1

A toda relación de equivalencia *R* definida en un conjunto *E* está asociada una partición de *E* formada por las clases de equivalencia de *R*. Recíprocamente, si en un conjunto *E* está dada una partición Π , existe una relación de equivalencia de *E* tal que Π es la partición de sus clases de equivalencia.

Demostración

Supongamos que tenemos una relación *R* de equivalencia definida sobre *E*. Tomemos Π como la familia $\Pi = \{C_a\}_{a \in E}$ y probemos que es una partición de *E*.

Sabemos que $a \in C_a \forall a \in E$, de donde $\bigcup_{a \in E} C_a = E$. Veamos ahora que $C_a = C_b \iff aRb$.

En efecto, sea aRb y $c \in C_a$. Entonces cRa y por la transitividad de *R*, cRb ; luego $c \in C_b$ y $C_a \subset C_b$.

De manera similar probamos $C_b \subset C_a$.

Inversamente, supongamos $C_a = C_b$. De $a \in C_a$, tenemos que $a \in C_b$ y aRb por definición de C_b .

Queda probar, pues, que dos clases diferentes son disjuntas. Para ello supongamos que $C_a \cap C_b \neq \emptyset$ y sea $c \in C_a \cap C_b$. Entonces cRa y cRb , de donde $aRb \iff C_a = C_b$. Luego $C_a \neq C_b \iff C_a \cap C_b = \emptyset$.

Observemos que como un paso intermedio en la demostración hemos probado un hecho importante: en una relación de equivalencia *R*, $aRb \iff C_a = C_b$.

Sea ahora Π una partición en el conjunto *E*. Definamos en *E* la relación $aRb \iff a$ y b pertenecen a la misma clase de la partición, y probemos que *R* es una relación de equivalencia sobre *E*.

Esta relación está bien definida, ya que cada $a \in E$ está contenida en una y solo una clase de la partición y, evidentemente, es reflexiva, simétrica y transitiva, luego *R* es una relación de equivalencia.

Está claro que el teorema anterior nos dice que dada una relación *R*, está determinada una partición Π cuyas clases son exactamente las clases de equivalencia de la relación dada. Inversamente, si conocemos una partición de un conjunto *E*, la relación que se establece entre los elementos del conjunto, según pertenezcan o no a la misma clase de la partición, es una relación de equivalencia cuyas clases son exactamente las clases de la partición. De ahí nuestra afirmación de que los conceptos *relación de equivalencia* y *partición* son equivalentes. Concretemos esto mediante un ejemplo en un conjunto finito.

Ejemplo

Sea $E = \{a, b, c, d\}$ y la relación *R* dada sobre *E* por

$$R = \{(a, a), (b, b), (c, c), (d, d), (a, c), (c, a)\}.$$

Es fácil comprobar que esta relación es de equivalencia. La partición correspondiente es

$$\Pi = \{\{a, c\}, \{b\}, \{d\}\}.$$

Esta claro que si hubiéramos partido de Π para decidir qué elementos estaban relacionados, hubiéramos llegado a la relación *R*.

La equivalencia entre los conceptos *relación de equivalencia* y *partición* no solo tiene utilidad teórica, sino también práctica, ya que en algunos ejemplos es más fácil y rápido darse cuenta de que una relación dada establece una partición del conjunto, que verificar uno a uno los axiomas de relación de equivalencia. Esto ocurre en los ejemplos 3 y 4 dados anteriormente.

Una vez establecida la clasificación de los elementos de un conjunto según una relación de equivalencia, cada clase puede ser considerada como

un elemento único de un nuevo conjunto. Por ejemplo, a los efectos del trabajo con los números fraccionarios, el conjunto

$$\left\{ \frac{1}{2}, \frac{-1}{-2}, \frac{2}{4}, \frac{-2}{-4}, \frac{4}{8}, \frac{-4}{-8}, \dots \right\}$$

constituye un solo elemento, y a veces ocurre que en algunos problemas basta trabajar con uno de los representantes de esta clase para obtener la información requerida. Así, en el trabajo con los números racionales, para sumar el número representado por

$$\left\{ \frac{1}{2}, \frac{-1}{-2}, \frac{2}{4}, \frac{-2}{-4}, \dots \right\}$$

y el representado por

$$\left\{ \frac{2}{3}, \frac{-2}{-3}, \frac{4}{6}, \frac{-4}{-6}, \dots \right\}$$

basta sumar $\frac{1}{2}$ y $\frac{2}{3}$.

Esto da lugar a la siguiente definición:

DEFINICIÓN 3.9.5

Dada una relación de equivalencia R sobre E , se llama *conjunto cociente de E según R* al conjunto de las clases de equivalencia según R de los elementos de E , y se denota por E/R .

Así $E/R = \{C_a : a \in E\}$.

Debemos enfatizar que E/R no es un subconjunto de E ; sus elementos son clases de equivalencia, o sea, subconjuntos de E y, por tanto, E/R es un subconjunto de $P(E)$.

Ejemplos

En los ejemplos dados de relaciones de equivalencia, los conjuntos cocientes son los siguientes:

1. $E/R = \{\{x\} : x \in E\}$
2. $E/R = \{E\}$
3. La familia de todas las rectas paralelas a OX .
4. La familia de todas las circunferencias con centro O .
5. $\mathbb{Z}/R = \{C_0, \dots, C_{n-1}\}$, donde $C_r = \{r + nq : q \in \mathbb{Z}\}$, $0 \leq r \leq n-1$.

Aunque dada una clase de equivalencia de una relación de equivalencia R en un conjunto E , cualquier elemento perteneciente a la clase puede representarla, es usual escoger un elemento con ciertas características espe-

ciales para tomarlo como representante de la clase. El conjunto de estos representantes es un subconjunto de E con el cual podemos "identificar" al conjunto cociente.

Representemos el cociente de algunos de los ejemplos de relaciones de equivalencia:

En el ejemplo 3, por las características de las clases es posible hacer una representación geométrica, ya que evidentemente un punto natural a tomar como representante de cada recta horizontal es el punto de intersección de dicha recta con el eje OY (figura 3.5). En este caso podemos "identificar" el cociente con el eje OY , esto es, con $\mathbb{R}$.

(Cuando relacionemos más adelante los conceptos de *cociente* y de *grupo*, veremos que como base de esta "identificación" está el concepto de *isomorfismo*.)

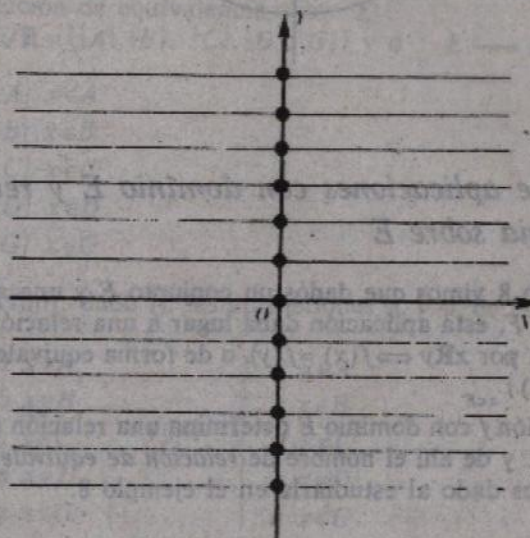


Fig. 3.5

En el ejemplo 4, para representar las circunferencias podemos tomar un radio común a todas y para cada circunferencia su punto de intersección con dicho radio. Así, en este caso el cociente se "identifica" con una semirrecta (figura 3.6).

En el ejemplo 5 podemos identificar el cociente con el conjunto $\{0, 1, \dots, n-1\}$.

Obsérvese que en todos los casos la correspondencia entre cada clase y su representante es biyectiva.

En lo que sigue tendrá gran importancia la aplicación $\phi : E \rightarrow E/R$, dada por $a \mapsto C_a$. Esta aplicación es sobreyectiva y se llama *sobreyección canónica de E en E/R* .

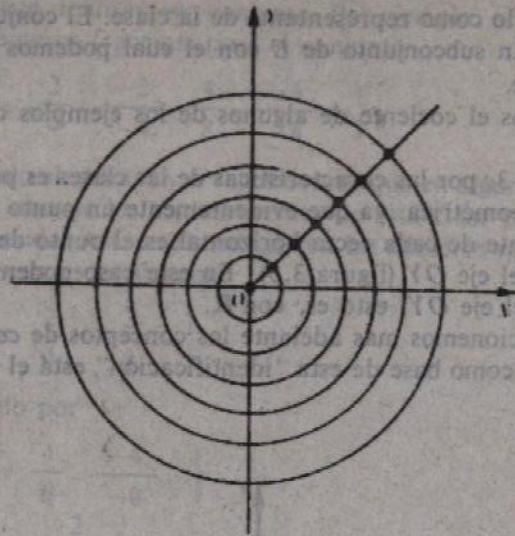


Fig. 3.6

Relación entre aplicaciones con dominio E y relaciones de equivalencia sobre E

En el ejemplo 8 vimos que dados un conjunto E y una aplicación de E en otro conjunto F , esta aplicación daba lugar a una relación de equivalencia sobre E dada por $xRy \iff f(x) = f(y)$ o de forma equivalente por la partición, $\{f^{-1}(f(x))\}_{x \in E}$.

Cada aplicación f con dominio E determina una relación de equivalencia definida sobre E , y de ahí el nombre de *relación de equivalencia asociada a f* que le habíamos dado al estudiarla en el ejemplo 8.

Ejemplo

Si consideramos en el plano la aplicación ortogonal sobre el eje OY y buscamos la relación de equivalencia asociada a dicha aplicación, entonces dos puntos del plano están relacionados si y solo si tienen la misma proyección sobre OY , esto es, si pertenecen a una misma recta paralela a OX . Encontramos, pues, la relación de equivalencia del ejemplo 3, de la cual conocemos la partición equivalente y una representación del cociente.

La expresión analítica de la proyección ortogonal es $p: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, dada por $(x, y) \mapsto (0, y)$. Si consideramos la aplicación $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $(x, y) \mapsto y$ y buscamos la relación de equivalencia en $\mathbb{R}^2$ asociada a f , encontramos que dicha relación es de nuevo la dada por el ejemplo 3.

Este ejemplo nos dice que puede haber más de una aplicación que tenga la misma relación de equivalencia asociada. Analicemos este problema.

Ya establecimos que para cada aplicación con dominio E existe una relación de equivalencia R que le corresponde de forma natural. Ahora bien, a partir de una relación de equivalencia R sobre E (y por tanto, una partición de E), es posible definir una aplicación con dominio E tal que si buscamos la relación de equivalencia asociada a dicha aplicación obtenemos R . Esto lo podemos hacer asignando a todos los elementos de cada una de las clases de equivalencia de R (o de las clases de la partición) la misma imagen.

El ejemplo dado anteriormente nos dice que esta aplicación no tiene por qué ser única, pero que la relación de equivalencia (partición) sobre E determina de forma esencial la aplicación. Veamos el siguiente ejemplo que nos aclarará este problema.

Ejemplo

Sean E y F dos conjuntos y $\Pi = \{A, B, C, D, G\}$ una partición de E (que define una relación de equivalencia sobre E).

Tenemos $E/R = \{\{A\}, \{B\}, \{C\}, \{D\}, \{G\}\}$ y $\varphi: E \rightarrow E/R$ dada por

$$\varphi(x) = \begin{cases} \{A\} & x \in A \\ \{B\} & x \in B \\ \{C\} & x \in C \\ \{D\} & x \in D \\ \{G\} & x \in G \end{cases}$$

Podemos definir, dado Π , las aplicaciones φ_1 y φ_2 de E en F , dadas por:

$$\varphi_1(x) = \begin{cases} a & x \in A \\ b & x \in B \\ c & x \in C \\ d & x \in D \\ g & x \in G \end{cases} \quad \varphi_2(x) = \begin{cases} b & x \in A \\ c & x \in B \\ d & x \in C \\ g & x \in D \\ a & x \in G \end{cases}$$

Denotamos los elementos de E/R con notación de conjunto para recordar que estos son puntos considerados como elementos de $P(E)$.

Representemos gráficamente las aplicaciones φ_1 y φ_2 (figura 3.7).

Con este ejemplo podemos darnos cuenta que aunque desde el punto de vista de la teoría de conjuntos las aplicaciones φ , φ_1 , φ_2 son diferentes, a los efectos de la relación de equivalencia (partición) que definen en E pudieran ser consideradas como una sola aplicación. Vemos también que aunque cambiáramos el codominio, si mantenemos la partición en E la diferencia entre las aplicaciones obtenidas está en la denominación de las imágenes.

Vamos a formalizar las ideas anteriores, esclareciendo con ello cuál es la relación existente entre las aplicaciones con dominio E y las relaciones de equivalencia definidas sobre E .

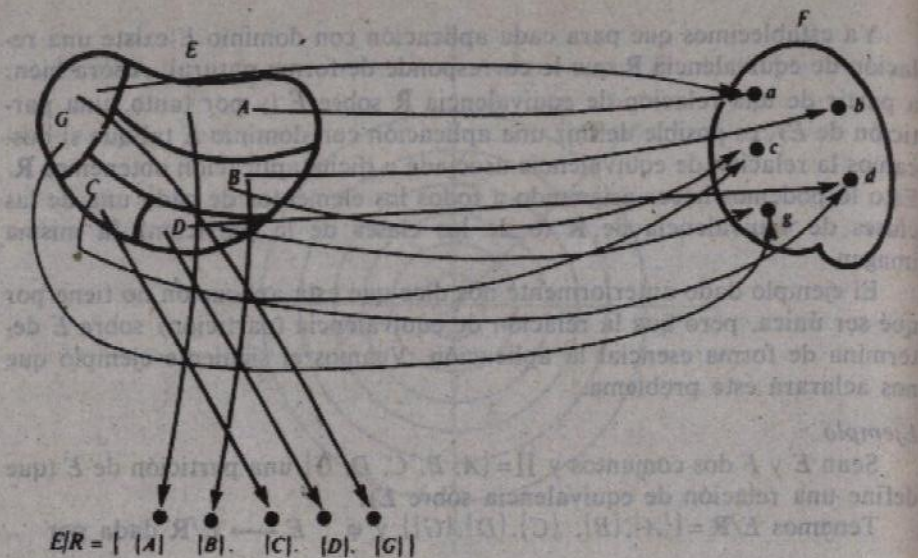


Fig. 3.7

TEOREMA 3.9.2

Dada una relación de equivalencia R sobre un conjunto E , la aplicación más simple definida sobre E a la cual R está asociada es precisamente la sobreyección canónica $\varphi : E \longrightarrow E/R$.

Demostración

Busquemos la relación de equivalencia S asociada a φ :

$$xSy \iff \varphi(x) = \varphi(y) \iff C_x = C_y \iff xRy, \text{ o sea, } S=R.$$

Con esto sabemos que no solo a cada aplicación con dominio E le corresponde una relación de equivalencia sobre E , sino que para cada relación de equivalencia R sobre E existe al menos una aplicación con dominio E a la cual R está asociada. También vimos que en general no es solamente la sobreyección canónica la aplicación a la que R está asociada, por lo que daremos un resultado que muestra cómo están relacionadas con φ las aplicaciones con dominio E a las cuales está asociada la relación dada R .

En el ejemplo gráfico que vimos (figura 3.7) se evidenció la correspondencia biunívoca existente entre los elementos de E/R y los del conjunto $\text{Im } f$. De ahí el teorema siguiente.

TEOREMA 3.9.3 (Descomposición de una aplicación).

Sean E, F conjuntos, $f: E \longrightarrow F$ una aplicación, R la relación de equivalencia asociada a f , $\varphi : E \longrightarrow E/R$ la sobreyección canónica, entonces existe una única aplicación inyectiva $\bar{f}: E/R \longrightarrow F$ tal que $\bar{f} \circ \varphi = f$.

Demostración

Está claro que la forma más natural de definir $\bar{f}$ es

$$\bar{f}: E/R \longrightarrow F \text{ dada por } \bar{f}(C_a) = f(a).$$

$\bar{f}$ está bien definida y es inyectiva ya que

$$C_a = C_b \iff f(a) = f(b) \iff \bar{f}(C_a) = \bar{f}(C_b).$$

Observemos que en general cuando definimos una aplicación es obvio, por la propia definición, que esta es correcta. Sin embargo, no es este el caso de $\bar{f}$, la cual hemos definido para una clase de equivalencia basándonos solo en un representante de dicha clase.

Es por ello que el paso que acabamos de dar es imprescindible, y en general ocurrirá así cuando el dominio de la aplicación a definir sea un cociente.

$$\bar{f} \circ \varphi = f \text{ ya que } \bar{f} \circ \varphi(a) = \bar{f}(C_a) = f(a) \quad \forall a \in E.$$

$\bar{f}$ es única, pues si suponemos que existe $g = E/R \longrightarrow F$ tal que $g \circ \varphi = f$, entonces

$$g(C_a) = g(\varphi(a)) = f(a) = \bar{f}(\varphi(a)) = \bar{f}(C_a) \quad \forall C_a \in E/R.$$

Observemos que $\bar{f}$ será biyectiva si y solo si f es sobreyectiva, pero que esto siempre puede resolverse considerando como codominio de f el conjunto $\text{Im } f$.

Con el teorema anterior hemos demostrado que toda aplicación con dominio E cuyas preimágenes estén determinadas por una relación de equivalencia (partición) R , "difiere" de la aplicación canónica φ de E en E/R en una biyección. Esto es, podemos cambiar las imágenes, pero estas están en correspondencia biyectiva con E/R .

Cuando estudiemos este problema en el caso en que el conjunto E posee una cierta estructura algebraica, por ejemplo la de grupo, esto se expresará en términos de isomorfismo.

El teorema anterior ofrece también la posibilidad de demostrar la existencia de biyecciones entre un cociente y otro conjunto, sin tener que definir estas directamente, ya que esto último en general resulta más difícil.

Ejemplo

En $\mathbb{R}^2 \setminus \{0\} = \mathbb{R}^2$ la relación R dada por

$$(x_1, x_2) R (y_1, y_2) \iff \exists \lambda \in \mathbb{R}^*: \lambda(x_1, x_2) = (y_1, y_2).$$

es de equivalencia, lo que podrá fácilmente comprobar el estudiante, bien por la verificación de las propiedades o por la búsqueda de la correspondiente partición de $\mathbb{R}^2 \setminus \{0\}$.

Si buscamos $\mathbb{R}^2/R$, obtenemos que está formado por rectas que pasan por el origen (0,0) excluyendo dicho punto. Este cociente se representa geométricamente como se muestra en la figura 3.8.

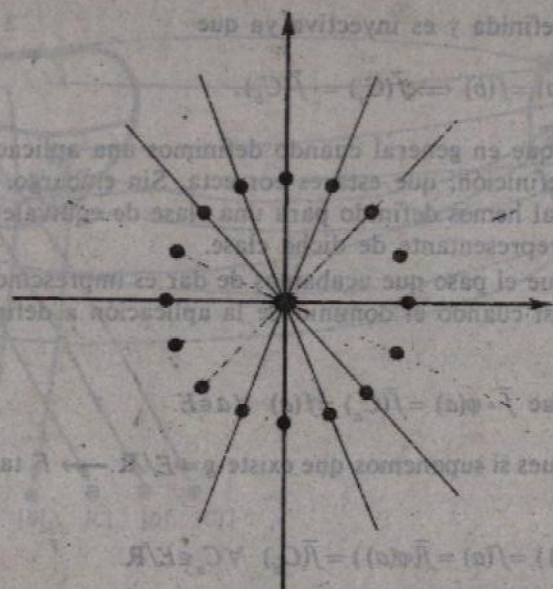


Fig. 3.8

Ahora bien, al buscar una representación geométrica del cociente nos damos cuenta de que podemos identificarlo con $\mathbb{R} \cup \infty$, ya que a cada recta le podemos hacer corresponder su pendiente y al eje OY el ∞ . Formalizar esto resulta engorroso porque habría que probar que la aplicación que a la recta que pasa por (x, y) , $x \neq 0$ le asigna su pendiente y a $(0, y)$ el ∞ , es independiente de la selección del punto (x, y) .

Sin embargo, si consideramos la aplicación $f: \mathbb{R}^2 \rightarrow \mathbb{R} \cup \infty$ dada por

$$(x, y) \mapsto \frac{y}{x}, \quad x \neq 0.$$

$$(0, y) \mapsto \infty$$

es evidente que esta es una aplicación bien definida, cuya relación de equivalencia asociada S es la propia R , pues

$$(x, y) S (x', y') \iff \frac{y}{x} = \frac{y'}{x'} \iff \exists \lambda \in \mathbb{R}^*: (x', y') = \lambda(x, y)$$

y todos los puntos $(0, y)$ pertenecen a una misma clase.

Además, $\text{Im } f = \mathbb{R} \cup \infty$; por tanto, aplicando el teorema de descomposición, podemos afirmar que existe una biyección entre $\mathbb{R}^2/R$ y $\mathbb{R} \cup \infty$.

Al estudiante puede parecerle que es innecesario utilizar este teorema, ya que en este caso sabíamos cómo definir la biyección entre el cociente y el otro conjunto dado, y solo hubiéramos tenido que verificar que la aplicación estaba bien definida y era una biyección. Dejamos, pues, que realice ese trabajo y compare ambas vías. Más adelante podrá comprobar que en

muchos casos no es tan evidente cuál es la biyección asociada y no es necesario conocerla de forma explícita por lo cual la aplicación del teorema de descomposición de una aplicación tendrá un gran valor práctico.

Ejercicios

- Analice si las relaciones dadas son de equivalencia.
 - En el conjunto de los habitantes de la Tierra, tenemos:
 - aRb si y solo si a y b tienen un antepasado común.
 - aRb si y solo si a y b viven a menos de 100 km.
 - aRb si y solo si a y b tienen el mismo padre.
 - En $S = \{0, 1, 2, 3\}$ tenemos:
 - $E = \{(0, 0), (1, 1), (2, 2), (3, 3), (0, 2), (1, 3), (2, 0), (3, 1)\}$
 - $F = \{(0, 0), (1, 1), (2, 2), (3, 3), (1, 2), (2, 3), (2, 1), (3, 2)\}$
- Analice si la siguiente proposición es verdadera o falsa y justifique su respuesta:
"Toda relación binaria simétrica y transitiva es reflexiva, o sea, es una relación de equivalencia."
- Determine cuáles de las siguientes relaciones son de equivalencia y halle sus clases de equivalencia.
 - nRm en $\mathbb{Z}$ si y solo si $mn > 0$
 - xRy en $\mathbb{R}$ si y solo si $|x| = |y|$
 - xRy en $\mathbb{R}$ si y solo si $x \geq y$
 - xRy en $\mathbb{R}$ si y solo si $|x - y| \leq 3$
 - En $\mathbb{Z}$, $aRb \iff a^2 + a = b^2 + b$
 - En $\mathbb{R}^2$, $(x, y)R(x', y') \iff xy = x'y'$
- Sea $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ y $E = \mathbb{Z} \times \mathbb{Z}^*$. La relación R se define en E como $(a, b)R(c, d) \iff ad = cb$.
 - Pruebe que R es una relación de equivalencia.
 - Halle el conjunto cociente E/R .
 - Utilice el teorema de descomposición de una aplicación para demostrar que E/R está en biyección con $\mathbb{Q}$.
- Demuestre que un conjunto X es infinito si y solo si existen infinitas relaciones de equivalencias en X .
- Sea E el plano (considerado como un conjunto de puntos) y sea r una recta fija en E . Se dice que los puntos p, q de E están relacionados (pRq) si y solo si p y q determinan una recta paralela a r .
 - Pruebe que R es una relación de equivalencia en E .
 - Determine sus clases de equivalencia, un conjunto representativo de estas y el conjunto cociente E/R .

7. Sea E el espacio ordinario (considerado como un conjunto de puntos) y sea O un punto fijo en E . Dados dos puntos p, q se considera la relación R tal que $(pRq) \iff (O, p, q \text{ son colineales})$.
- ¿Es R una relación de equivalencia en E ?
 - Pruebe que R es una relación de equivalencia en E^* y determine sus clases de equivalencia. ($E^* = E - \{O\}$)
8. Sea R la relación definida en $\mathbb{R}^2$ por $(x, y) R (x_1, y_1) \iff \text{existe } \alpha \in \mathbb{R} \text{ tal que } x - x_1 = \alpha \text{ y } y - y_1 = 2\alpha$.
- Pruebe que R es una relación de equivalencia en $\mathbb{R}^2$.
 - Halle las clases de equivalencia y representélas geométricamente.
 - Halle $\mathbb{R}^2/R$.
9. Se define en $\mathbb{R}$ la siguiente relación: $xRy \iff x - y = 2k\pi$ para algún $k \in \mathbb{Z}$.
- Pruebe que R es una relación de equivalencia en $\mathbb{R}$.
 - Halle las clases de equivalencia.
10. Sea $E = GL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A \neq 0\}$.
- Pruebe que la relación definida sobre E , $ARB \iff \det A = \det B$ es de equivalencia.
 - Halle E/R y pruebe que está en biyección con $\mathbb{R}^*$.
11. Considere $E = P(A)$ donde A es un conjunto cualquiera y $B \in P(A)$ fijo, y sea la relación R definida sobre $P(A)$ por $XRY \iff X \cap B = Y \cap B$.
- Pruebe que R es de equivalencia.
 - Halle $P(A)/R$ y pruebe que está en biyección con $P(B)$.
12. Pruebe que si G es un grupo, la relación R definida sobre G por $aRb \iff a = gb g^{-1}$ para algún $g \in G$, es una relación de equivalencia en G .
Algunas clases de equivalencia contienen un único elemento c . Caracterice estos elementos c .
13. En la clase de los grupos, pruebe que la relación $GRG' \iff G$ es isomorfo a G' es una relación de equivalencia.
14. Considere las siguientes aplicaciones, escriba su diagrama de descomposición y represente geométricamente las clases de equivalencia.
- $P_1 : \mathbb{R}^2 \longrightarrow \mathbb{R} \times \{0\}$ dada por $(x, y) \mapsto (x, 0)$
 - $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $(x, y) \mapsto |x| = \sqrt{x^2 + y^2}$
 - $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $(x, y) \mapsto y - mx$, $m \in \mathbb{R}$ fijo
 - $h : \mathbb{R}^3 \longrightarrow \mathbb{R}$ dada por $(x, y, z) \mapsto |x, y, z| = \sqrt{x^2 + y^2 + z^2}$
 - $e : C^* \longrightarrow C^*$ dada por $z \mapsto |z|^{-1} \cdot z$
15. Sean X, Y dos conjuntos y R y S relaciones de equivalencia sobre X y

Y respectivamente. Sea $f : X \longrightarrow Y$ una aplicación. Considere el siguiente diagrama, donde p, q son las sobreyecciones canónicas de X en X/R y de Y en Y/S respectivamente.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ p \downarrow & & \downarrow q \\ X/R & & Y/S \end{array}$$

a) Pruebe que las siguientes propiedades son equivalentes:

(1) Existe $\bar{f} : X/R \longrightarrow Y/S$ tal que $f \circ p = q \circ \bar{f}$.

(2) $xRy \implies f(x) S f(y)$

b) Pruebe que si se cumple (1), $\bar{f}$ es única.

3.10 Subgrupos invariantes

Los conceptos de *relación de equivalencia*, *partición*, etc., que acabamos de estudiar, nos interesan en su relación con las estructuras algebraicas, en primer lugar con la estructura de grupo. Por tanto, dado un grupo G , analizaremos qué condiciones debe cumplir una relación de equivalencia R definida sobre G para poder, a partir de la ley interna de G , definir una ley de grupo en G/R .

Como lo primero que tenemos que resolver es la definición de la ley interna en el cociente G/R y esta situación se repetirá para otras estructuras algebraicas, trataremos el problema general siguiente: dado un conjunto E en el cual esté definida una ley interna y una relación de equivalencia R , cuáles son las condiciones para poder definir una ley interna en E/R .

Compatibilidad de una relación de equivalencia con una ley interna

A modo de introducción recordemos el problema de la construcción de $\mathbb{Q}$ a partir de $\mathbb{Z}$. Construimos $\mathbb{Q}$ para tratar de resolver el problema de la solución de ecuaciones del tipo $ax = b$ para todo a y b en $\mathbb{Z}$. Como sabemos, estas ecuaciones no siempre tienen solución en $\mathbb{Z}$, pero siempre tienen solución en $\mathbb{Q}$ si $a \neq 0$.

Pero, ¿qué es un número racional? Un número racional es una clase de equivalencia de pares de enteros. Por ejemplo, la clase $\left\{ \frac{1}{2}, \frac{2}{4}, \frac{4}{8}, \dots, \frac{m}{2m}, \dots \right\}$ puede ser representada por cualquiera de los pares de la forma $\frac{m}{2m}$, pero se suele tomar $\frac{1}{2}$ como representante principal.

Para definir operaciones en $\mathbb{Q}$ debe hacerse de forma tal que el representante de la clase que se escoja no intervenga en el resultado, lo cual se logra con estas clases y las operaciones de suma y producto usuales en $\mathbb{Q}$. Por ejemplo, si tomamos

$$\frac{1}{2} \sim \frac{1}{2} \sim \frac{2}{4} \sim \dots \sim \frac{m}{2m}$$

$$\frac{1}{3} \sim \frac{1}{3} \sim \frac{2}{6} \sim \dots \sim \frac{m}{3m}$$

tenemos que

$$\frac{1}{2} + \frac{1}{3} = \frac{3+2}{6} = \frac{5}{6}, \quad \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$$

$$\frac{2}{4} + \frac{3}{9} = \frac{18+12}{36} = \frac{30}{36} \sim \frac{5}{6}, \quad \frac{2}{4} \cdot \frac{3}{9} = \frac{6}{36} \sim \frac{1}{6}$$

La construcción de $\mathbb{Q}$ permite resolver el problema de obtener solución a las ecuaciones que no tenían solución en $\mathbb{Z}$, lo que es equivalente a resolver el problema de la existencia del simétrico en $\mathbb{Q}$ para el producto. Observemos, además, que las operaciones en $\mathbb{Q}$ se definen a partir de las de $\mathbb{Z}$. Esto es, la relación de equivalencia debe ser tal, que el resultado de la operación entre clases sea independiente de los representantes que escojamos.

Dados un conjunto E con una ley interna $\perp$, R relación de equivalencia definida sobre E , se define

$$\perp : E/R \times E/R \rightarrow E/R, \text{ por } Cx \perp Cy = Cx \perp y.$$

Para que esta ley esté bien definida es evidente que la condición necesaria y suficiente es que el resultado de la operación no dependa de los representantes que escojamos en las clases.

DEFINICIÓN 3.10.1

Se dice que la relación R es compatible con la ley $\perp$ si y solo si se cumple

$$\left. \begin{array}{l} xRx' \\ yRy' \end{array} \right\} \Rightarrow xTy \text{ y } Rx'Ty'$$

o lo que es equivalente:

$$\left. \begin{array}{l} Cx = Cx' \\ Cy = Cy' \end{array} \right\} \Rightarrow CxTy = Cx'Ty'.$$

La ley así definida será conmutativa y asociativa si lo es T .

Si T tiene elemento neutro e , $\perp$ tendrá elemento neutro Ce , y es posible

que para $\perp$ ningún elemento posea simétrico y, sin embargo, los elementos de E/R posean simétrico para $\perp$.

Las propiedades de la ley definida en E/R las trataremos en detalle en cada una de las estructuras algebraicas.

Es conveniente observar que la definición de relación de equivalencia compatible con una ley $\perp$ es equivalente a la siguiente:

$$xRy \Rightarrow \left\{ \begin{array}{l} x \perp x' \text{ y } y \perp y' \quad \forall x' \in E \\ x' \perp x \text{ y } R x' \perp y' \quad \forall x' \in E \end{array} \right. \quad \begin{array}{l} (1) \\ (2) \end{array}$$

Si se cumple (1) solamente, decimos que R es compatible a la derecha con la ley $\perp$, y si se cumple (2), que es compatible a la izquierda. Esta definición es de gran utilidad cuando la ley interna no es conmutativa; y nos dice que R es compatible con $\perp$ si y solo si es compatible con $\perp$ a la izquierda y a la derecha.

Relaciones de equivalencia compatibles con una ley de grupo

Analicemos desde el punto de vista de que toda relación de equivalencia en G no es más que una partición de G , el problema de la compatibilidad de una relación de equivalencia con la ley de un grupo finito dado por una tabla de multiplicar.

Consideremos el grupo S_3 , cuya tabla de multiplicar es:

	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_1	ρ_1	ρ_2	ρ_0	μ_1	μ_1	μ_2
ρ_2	ρ_2	ρ_0	ρ_1	μ_2	μ_3	μ_1
μ_1	μ_1	μ_2	μ_3	ρ_0	ρ_1	ρ_2
μ_2	μ_2	μ_1	μ_1	ρ_2	ρ_0	ρ_1
μ_3	μ_3	μ_1	μ_2	ρ_1	ρ_2	ρ_0

La tabla de S_3 quedará dividida en bloques $B_0 = \{\rho_0, \rho_1, \rho_2\}$ y $B_\mu = \{\mu_1, \mu_2, \mu_3\}$ que constituyen una partición de S_3 . Considerando la tabla cuyos elementos son B_0 y B_μ , que surgen de manera natural de la tabla de

S_3 , obtenemos la tabla de un grupo de dos elementos B_p, B_μ , que llamaremos G .

	B_p	B_μ
B_p	B_p	B_μ
B_μ	B_μ	B_p

Dicho grupo es el cociente de S_3 por cierta relación de equivalencia y el elemento neutro de G es $B_p = C_p$, o sea, el elemento neutro de G es la clase de equivalencia del elemento neutro de S_3 .

Para seguir obteniendo información del ejemplo de S_3 , observemos que B_p es un subgrupo de S_3 .

Ahora haremos una deducción teórica de la forma que deben tener las relaciones de equivalencia compatibles con una ley de grupo. Dicha deducción puede ser realizada de varias maneras, pero por cualquier vía que se realice resultará un tanto abstracta, ya que estamos resolviendo un problema de existencia. Después concretaremos esta deducción teórica en ejemplos.

Vamos a suponer la existencia de una relación de equivalencia compatible con la ley de un grupo G y veamos qué forma toma dicha relación. Como nuestro grupo no es necesariamente conmutativo, separemos el análisis de la compatibilidad a la derecha y a la izquierda. Sabemos que

R es compatible si y solo si es compatible a la derecha y a la izquierda
 $(R_d \text{ es compatible a la derecha}) \iff (xR_dy \iff x \cdot x'R_d y \cdot x' \text{ para todo } x' \in G)$

$(R_l \text{ es compatible a la izquierda}) \iff (xR_l y \iff x' \cdot xR_l x' \cdot y \text{ para todo } x' \in G)$

(Por comodidad usaremos R_d y R_l en vez de R solamente.)

En el ejemplo de S_3 vimos que la clase del elemento neutro parecía desempeñar un papel importante y que era un subgrupo de S_3 . Veamos si podemos llegar de forma general a una propiedad similar para la clase del elemento neutro de G .

Sea, pues, R_d una relación de equivalencia compatible a la derecha, y denotemos el elemento neutro de G por e . Dicha compatibilidad a la derecha implica:

Si xR_dy , entonces $xy^{-1}R_dyy^{-1}=e$.

Luego la relación R_d es del tipo $xR_dy \iff xy^{-1} \in Ce$.

Ce es un subgrupo de G pues xR_de , $yR_de \implies xR_dy \iff xy^{-1}R_de$, o sea, $x \in Ce$, $y \in Ce \implies xy^{-1} \in Ce$.

Es decir, que si R_d es compatible a la derecha, encontramos un subgrupo de G , el subgrupo Ce , tal que la forma de la relación R_d es $xR_dy \iff xy^{-1} \in Ce$.

Por otra parte, si H es un subgrupo de G , la relación $xR_dy \iff xy^{-1} \in H$ es de equivalencia compatible a la derecha con la ley de G .

En efecto, R_d es una relación de equivalencia, ya que se cumplen las propiedades:

(1) Reflexiva: $xx^{-1}=e \in H \iff xR_dx$

(2) Simétrica: $xR_dy \iff xy^{-1} \in H \implies (xy^{-1})^{-1}=yx^{-1} \in H \iff yR_dx$

(3) Transitiva: a) $xR_dy \iff xy^{-1} \in H$, b) $xR_dz \iff yz^{-1} \in H$

De (3a) y (3b) se deduce que $(xy^{-1}) \cdot (yz^{-1}) = xz^{-1} \in H \iff xR_dz$. Además, R_d es compatible a la derecha con la ley de G , pues si $x, y, z \in G$, tenemos

$$xR_dy \iff xy^{-1} \in H$$

$$xy^{-1} = xzz^{-1}y^{-1} = (xz)(yz)^{-1}$$

$$(xz)(yz)^{-1} \in H \implies xzR_dyz$$

Con un razonamiento similar llegamos a que si R_l es una relación de equivalencia compatible a la izquierda, R_l toma la forma

$xR_ly \iff y^{-1}x \in Ce$ donde Ce es un subgrupo de G , y recíprocamente, la relación $xR_ly \iff y^{-1}x \in H$ donde H es un subgrupo de G es una relación de equivalencia compatible a la izquierda con la ley de G .

El estudiante no debe perderse en los detalles técnicos de la demostración anterior. Lo esencial es que conozca que existen métodos teóricos para determinar las condiciones que permiten definir en el cociente de un grupo una ley interna a partir de la ley de dicho grupo.

Esto lleva a considerar relaciones del tipo $xR_ly \iff xy^{-1} \in H$ o $xR_ly \iff y^{-1}x \in H$ donde H es un subgrupo arbitrario de G , que en los detalles de la demostración se concreta mediante la clase del elemento neutro del grupo por la relación compatible. Pero como la compatibilidad en un solo sentido (esto es, a la derecha o a la izquierda) no basta para la construcción del cociente, debemos imponer algunas condiciones al subgrupo H , que es en definitiva el que determinará la relación.

Hemos visto, pues, que para que una relación de equivalencia sea compatible a la derecha con la ley de grupo, es necesario y suficiente que sea de la forma $xR_dy \iff xy^{-1} \in H$ donde H es un subgrupo de G , y para que sea compatible a la izquierda deberá ser de la forma $xR_ly \iff y^{-1}x \in H$ donde H es un subgrupo de G .

Por tanto, para que R sea compatible tendrán que cumplirse ambas condiciones, por lo cual nos interesarán las relaciones dadas por los subgrupos H de G tales que la relación R_d coincida con la R_l , y la mejor forma de lograr que esto se cumpla es exigiendo que las particiones de G dadas por ambas relaciones coincidan, o sea, que las clases de equivalencia de $x \in G$ por R_l y por R_d sean idénticas. Busquemos dichas clases para un subgrupo H de G .

Llamaremos $C_{x,d}$ a la clase de x por R_d y $C_{x,i}$ a la clase de x por R_i .
 $C_{x,d} = \{y \in G : y R_d x\} = \{y \in G : \text{existe } h \in H \text{ tal que } yx^{-1} = h\} =$
 $= \{y \in G : \text{existe } h \in H \text{ tal que } y = hx\} = \{hx : h \in H\}$

A la clase $C_{x,d} = \{hx : h \in H\}$ la denotamos por Hx y la llamamos *coseto derecho de x según H* . Acabamos, pues, de encontrar que la clase de x según R_d no es más que el coseto derecho de x según H .

De forma similar llegamos a que $C_{x,i} = \{xh : h \in H\}$, a la cual denotamos por xH y la llamamos *coseto izquierdo de x según H* .

La denominación coseto la utilizamos para simplificar la notación, ya que en inglés se utiliza *coset* y no existe una traducción al español, también podemos llamarlas *clases según H a la derecha o a la izquierda*.

Los cosetos izquierdos y derechos según un subgrupo H de G son una vía para obtener información sobre H y sobre el grupo G .

Veamos el siguiente resultado que desempeña un papel fundamental en el estudio de los grupos y que recibe el nombre de *teorema de Lagrange*, ya que este matemático se dio cuenta de su validez cuando se encontraba estudiando el problema de la solubilidad de ecuaciones algebraicas.

TEOREMA 3.10.1 (Teorema de Lagrange)

Sean G un grupo finito de orden n y H un subgrupo de G de orden m . Entonces m divide a n .

Demostración

Esta demostración se basa en una técnica simple que encontraremos reiteradamente en el estudio de las estructuras algebraicas: la de contar los elementos.

En efecto, sea $H = \{a_1, a_2, \dots, a_m\}$, donde los a_i son todos distintos. Entonces cada coseto xH tiene exactamente m elementos, pues si por ejemplo $xa_i = xa_j$, se infiere que $a_i = a_j$ por la propiedad cancelativa de G , lo que contradice que los a_i sean todos distintos.

Por otra parte, si tomamos dos cosetos distintos xH y $x'H$, la aplicación de xH en $x'H$ dada por $xh \rightarrow x'h$ es una biyección, lo que indica que todos los cosetos tienen la misma cantidad de elementos.

Sabemos, además, que los cosetos constituyen una partición de G , de aquí que los n elementos de G quedarán divididos por dicha partición en q partes iguales de m elementos cada una. Por tanto, $n = mq$, con lo que queda demostrado el teorema.

Este teorema es de gran utilidad en la búsqueda de los subgrupos de un grupo ya que, por ejemplo, nos dice que en un grupo de orden 6 no puede haber subgrupos de orden 4 o 5.

Sin embargo, hay que tener cuidado con la aplicación del teorema de Lagrange, pues este no asegura la existencia de subgrupos de un orden dado, por ejemplo, que en un grupo de orden 6 tengan que existir subgrupos de orden 2 o 3. Esto pudiéramos afirmarlo si se cumpliera el recíproco del teorema de Lagrange.

TEOREMA 3.10.2 (Recíproco del teorema de Lagrange)

Si G es un grupo finito de orden n y m es un divisor de n , entonces G posee un subgrupo de orden m .

Este recíproco no se cumple en general. Por ejemplo, el subgrupo A_4 formado por las permutaciones pares de S_4 es de orden 12 y no tiene subgrupos de orden 6. Sin embargo, más adelante veremos que el recíproco de Lagrange es válido para grupos abelianos finitos.

El teorema de Lagrange tiene algunos corolarios muy importantes que dan información sobre el grupo.

COROLARIO 1°

Un grupo de orden primo p no tiene subgrupos propios.

Demostración

El orden de los subgrupos de G deberá ser 1 o p .

COROLARIO 2°

Un grupo de orden primo es cíclico (esto es, coincide con alguno de sus subgrupos cíclicos).

Demostración

En efecto, si $G = \{0\}$ esto es evidente. Sea $G \neq \{0\}$ y $a \in G$, $a \neq 0$; entonces $\langle a \rangle = G$.

Una vez obtenida la forma de las clases según R_d y R_i para H y atendiendo a lo señalado anteriormente sobre las condiciones para que una relación sea compatible con la ley del grupo, está claro que dado un subgrupo H , para que la relación $x R_d y \iff xy^{-1} \in H$ sea compatible, H debe cumplir la condición siguiente:

$$xH = Hx \text{ para todo } x \in G.$$

De esta forma la relación R_d coincidirá con la R_i y será compatible, y reciprocamente. Esto da lugar a la siguiente definición:

DEFINICIÓN 3.10.2

Se dice que un subgrupo H de un grupo G es *normal*, *invariante* o *distinguido* si y solo si $xH = Hx$ para todo $x \in G$, es decir, si cada coseto derecho coincide con el correspondiente coseto izquierdo.

El nombre de *invariante* dado también al subgrupo normal procede de la siguiente definición equivalente a la dada:

$$H \text{ es normal si y solo si } x H x^{-1} = H \text{ para todo } x \in G \text{ donde } x H x^{-1} = \{x h x^{-1} : h \in H\}.$$

Esta definición es equivalente a la anterior, lo cual puede ser verificado fácilmente por el estudiante (una definición muy útil en la práctica se da en el ejercicio 2 de este epígrafe).

Podemos, entonces, considerar las aplicaciones $i_x: G \rightarrow G$ dadas por

$$i_x(g) = x g x^{-1} \text{ para todo } g \in G.$$

Para cada $x \in G$ tenemos una tal aplicación y, según vimos anteriormente, dichas aplicaciones son automorfismos de G llamados automorfismos interiores de G . Luego H es normal si y solo si queda invariante bajo la acción de todos los automorfismos interiores de G , esto es, $i_x(H) = H \quad \forall x \in G$.

Ejemplos

1. En un grupo G , $\{0\}$ y G son subgrupos invariantes.
2. Si el grupo G es conmutativo, es obvio que todos sus subgrupos son invariantes. Por esto en los grupos conmutativos tenemos un gran número de ejemplos de subgrupos invariantes.
3. En el grupo simétrico S_n , el subgrupo alternado A_n es un subgrupo invariante, como puede comprobarse por la tabla de multiplicar o por demostración directa.
4. El subgrupo $SL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A = 1\}$ es un subgrupo invariante de $GL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : \det A \neq 0\}$.
5. El subgrupo $H = \{\rho_0, \mu_1\}$ de S_3 no es invariante, ya que los cosetos derechos son

$$\begin{aligned} H &= \{\rho_0, \mu_1\} & \mu_1 H &= \{\mu_1, \rho_0\} \\ \rho_1 H &= \{\rho_1, \mu_2\} & \mu_2 H &= \{\mu_2, \rho_0\} \\ \rho_2 H &= \{\rho_2, \mu_3\} & \mu_3 H &= \{\mu_3, \rho_0\} \end{aligned}$$

mientras que los izquierdos son

$$\begin{aligned} H &= \{\rho_0, \mu_1\} & H\mu_1 &= \{\mu_1, \rho_0\} \\ H\rho_1 &= \{\rho_1, \mu_1\} & H\mu_2 &= \{\mu_2, \rho_2\} \\ H\rho_2 &= \{\rho_0, \mu_2\} & H\mu_3 &= \{\mu_3, \rho_1\} \end{aligned}$$

En general, para verificar en los grupos finitos que un subgrupo H es invariante, deben buscarse todos los cosetos izquierdos y derechos correspondientes a H y comprobar que cada xH coincide con Hx .

No debe confundirse $xH = Hx$ con $xh = hx \quad \forall x \in G$ y $\forall h \in H$.

Un ejemplo más teórico es el siguiente:

6. Dado un grupo G , consideremos el centro de G :

$$C(G) = \{z \in G : gz = zg \quad \forall g \in G\}$$

Sabemos que $C(G)$ es un subgrupo de G (ver ejercicio 9 del epígrafe 3.6). Veamos que es invariante.

En efecto, debemos probar que

$$gC(G) = C(G)g \quad \forall g \in G,$$

o sea, que todo producto del tipo gc con $c \in C(G)$ se escribe como $c'g$ con $c' \in C(G)$ y viceversa. Pero

$$gc = cg \quad \forall g \in G, \quad \forall c \in C(G);$$

Luego, en el caso del centro podemos tomar c como el c' buscado.

Queremos llamar la atención sobre el hecho de que hemos dado varias caracterizaciones de subgrupo normal, invariante o distinguido. Dependerá del problema que estemos tratando, que sea más conveniente usar una u otra caracterización. En lo que sigue el estudiante encontrará ejemplos de ello.

Ejercicios

1. Pruebe que en un grupo G hay el mismo número de cosetos izquierdos que de cosetos derechos según un subgrupo H . En el caso de ser G un grupo finito, este número se llama *índice de H en G* .
2. Pruebe que un subgrupo H de G es invariante si y solo si $xhx^{-1} \in H \quad \forall x \in G, \forall h \in H$.
3. Pruebe que si H es un subgrupo de índice 2 de un grupo finito G , entonces H es invariante. Como consecuencia pruebe que A_n es invariante en S_n . (Recuerde que A_n es el subgrupo de las permutaciones pares de S_n .)
4. Pruebe que si un grupo G es finito de orden n y e es su elemento neutro, entonces $a^n = e$ para todo $a \in G$.
5. Pruebe que si A y B son subgrupos invariantes de G , entonces $A \cdot B$ es un subgrupo invariante de G ($A \cdot B = \{a \cdot b : a \in A, b \in B\}$).
6. Sean G un grupo y H y K subgrupos invariantes de G tales que $H \cap K = \{e\}$. Pruebe que $hk = kh$ para todo $h \in H$ y para todo $k \in K$.
7. Demuestre que si G tiene un solo subgrupo H de un orden dado, entonces H es invariante.
8. Demuestre que la intersección de dos subgrupos invariantes es invariante.
9. Pruebe que $\text{Aut } G = \{f: G \rightarrow G \mid f \text{ automorfismo}\}$ es un grupo para la composición de aplicaciones y que $\text{Inn } G = \{i_a : a \in G\}$ (automorfismos interiores de G) es un subgrupo invariante de $\text{Aut } G$.
10. Sea H un subgrupo normal de un grupo finito G y sea m el índice de H en G (ver ejercicio 1). Pruebe que $a^m \in H$ para todo $a \in G$.
11. Pruebe que el subconjunto $H = \{\varphi_{1,b} : b \in \mathbb{R}\}$ es un subgrupo normal de $A_1(\mathbb{R}) = \{\varphi_{a,b} : a \neq 0, a, b \in \mathbb{R}\}$. (Ver ejercicio 1.f del epígrafe 3.6.)
12. En los ejercicios 9 y 11 del epígrafe 5, dado un grupo G y $A \in G$ definimos los conjuntos $C(A)$ y $N(A)$ e indicamos probar que $C(A)$ y $N(A)$ son subgrupos de G y que $C(A)$ es un subgrupo de $N(A)$. Pruebe que
 - a) Si A es un subgrupo abeliano de G , A es invariante en $C(A)$.
 - b) Si A es un subgrupo de G , A es normal en G si y solo si $N(A) = G$.

13. Pruebe mediante un ejemplo que pueden encontrarse tres grupos E , F y G tales que $E \subset F \subset G$, donde E sea normal en F , F sea normal en G , pero E no sea normal en G .
14. Dados dos grupos G_1 y G_2 simples (que no tengan subgrupos normales propios), determine los subgrupos normales de $G_1 \times G_2$.
Sugerencia: Diferencie los casos $G_1 \cong G_2$ y $G_1 \not\cong G_2$.

3.11 Grupo cociente

De todo lo analizado en el epígrafe anterior, llegamos a la conclusión de que la búsqueda de las relaciones de equivalencia compatibles con la ley de un grupo G está determinada por la búsqueda de los subgrupos invariantes de dicho grupo. Una vez encontrado un subgrupo invariante H de G , la relación $x \sim y$ si y solo si $xy^{-1} \in H$ (o $x \sim y$ si y solo si $yx^{-1} \in H$) será compatible con la ley de grupo. Podemos, entonces, definir en G/H la operación $Cx \cdot Cy = Cxy$ que será una ley de grupo en G/H y tendrá las propiedades siguientes:

- (1) Ce será el elemento neutro de la ley definida en G/H , pues $Cx \cdot Ce = Cx \cdot e = Cx$ y $Ce \cdot Cx = Ce \cdot x = Cx \forall Cx \in G/H$.
- (2) La ley del cociente será asociativa, pues lo es la ley de G .
 $Cx \cdot (Cy \cdot Cz) = Cx \cdot (Cyz) = Cx(yz) = C(xy) \cdot z = Cxy \cdot Cz = (Cx \cdot Cy) Cz$
- (3) Para Cx el simétrico será Cx^{-1} , ya que $Cx \cdot Cx^{-1} = Cxx^{-1} = Ce$ y $Cx^{-1} \cdot Cx = Cx^{-1}x = Ce$.
- (4) Al grupo cociente dado por G/H con la relación $x \sim y \iff xy^{-1} \in H$ y con la ley de clases definida anteriormente, lo denotamos simplemente por G/H (ya que está claro que es precisamente H el que determina la relación) y lo llamamos *grupo cociente de G según H* .
Si G es conmutativo, G/H lo será, lo cual dejamos para que sea verificado por el estudiante.

Ejemplos

1. Como $\mathbb{Z}$ es abeliano, todos los subgrupos $n\mathbb{Z}$ son invariantes y podemos considerar los grupos cocientes $\mathbb{Z}/n\mathbb{Z}$: $\mathbb{Z}/n\mathbb{Z}$ es de orden n e isomorfo a $\mathbb{Z}_n$ según el isomorfismo
 $\phi: \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}_n$
dado por $m+n\mathbb{Z} \sim m$.
2. S_3/A_3 es el grupo cociente dado por $\{B_\sigma, B_\mu\}$ donde $B_\sigma = \{\rho_\sigma, \rho_1, \rho_2\} = A_3$, $B_\mu = \{\mu_1, \mu_2, \mu_3\}$.
Como existe un único grupo de orden 2 y S_3/A_3 tiene dos elementos, $S_3/A_3 \cong \mathbb{Z}_2$.

3. Hallemos $\mathbb{Z}_4 \times \mathbb{Z}_6 / \langle (0,1) \rangle$ donde $\langle (0,1) \rangle$ indica el subgrupo cíclico de $\mathbb{Z}_4 \times \mathbb{Z}_6$ generado por $(0,1)$.

$$\langle (0,1) \rangle = \{(0,0), (0,1), (0,2), (0,3), (0,4), (0,5)\}$$

Cada coseto según $\langle (0,1) \rangle$ tiene seis elementos. Luego, como los cosetos dan una partición de $\mathbb{Z}_4 \times \mathbb{Z}_6$, el cociente $\mathbb{Z}_4 \times \mathbb{Z}_6 / \langle (0,1) \rangle$ será de orden 4 y es

$$\mathbb{Z}_4 \times \mathbb{Z}_6 / \langle (0,1) \rangle = \{(0,0) + \langle (0,1) \rangle, (1,0) + \langle (0,1) \rangle, (2,0) + \langle (0,1) \rangle, (3,0) + \langle (0,1) \rangle\}.$$

(Las clases que hemos escogido son disjuntas; por eso podemos tomar los representantes $(0,0), (1,0), (2,0), (3,0)$.)

Este grupo es obviamente isomorfo a $\mathbb{Z}_4$, lo cual podemos comprobar comparando las tablas de multiplicar de ambos.

4. Busquemos $\mathbb{Z} \times \mathbb{Z} / \langle (0,1) \rangle$.

$$\langle (0,1) \rangle = \{(0,n) : n \in \mathbb{Z}\}$$

Luego, si buscamos $(a,b) + \langle (0,1) \rangle$, obtenemos

$$(a,b) + \langle (0,1) \rangle = \{(a,b+n) : n \in \mathbb{Z}\} = \{(a,m) : m \in \mathbb{Z}\} \text{ para cada } (a,b) \in \mathbb{Z} \times \mathbb{Z}.$$

Podemos representar gráficamente este cociente como se muestra en la figura 3.9.

Como los puntos de $\mathbb{Z} \times \mathbb{Z}$ son los de coordenadas enteras del plano, cada clase de equivalencia estará dada por una recta paralela al eje vertical. Y como para cada una de ellas podemos encontrar un representante natural en los puntos enteros del eje horizontal, intuitivamente vemos que el cociente es isomorfo al grupo $(\mathbb{Z}, +)$. (El estudiante puede hallar un isomorfismo entre $\mathbb{Z} \times \mathbb{Z} / \langle (0,1) \rangle$ y $(\mathbb{Z}, +)$ y probarlo rigurosamente.)

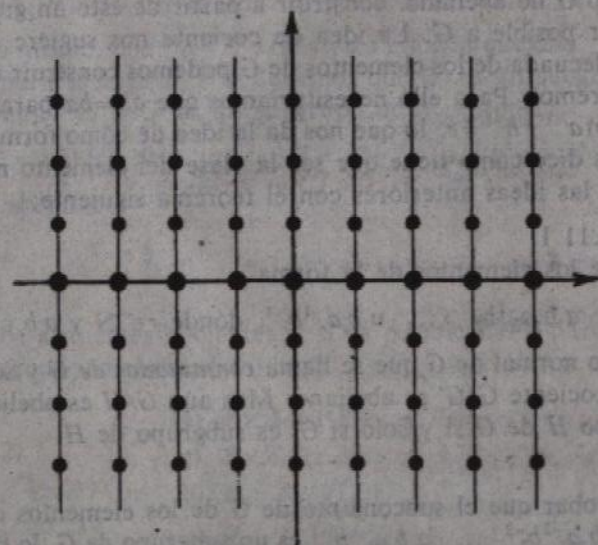


Fig. 3.9

Propiedades y aplicaciones del cociente

(1) Como consecuencia inmediata del teorema de Lagrange, tenemos que si G es un grupo finito y H un subgrupo de G , entonces $\alpha(G/H) = \alpha(G)/\alpha(H)$. (Este $\alpha(G/H)$ coincide con la definición de índice de H en G dada en el ejercicio 1 del epígrafe 3.9.)

(2) Como consecuencia del teorema de Lagrange vimos que un grupo finito de orden primo G no posee subgrupos propios y, por tanto, podemos decir que no posee subgrupos normales propios.

Se puede demostrar fácilmente, y lo dejamos al estudiante, que los únicos grupos abelianos que no poseen subgrupos normales propios son los grupos de orden primo. Para estos grupos el grupo cociente por un subgrupo no da ninguna información, ya que los únicos grupos cocientes posibles son $G/\langle e \rangle = G$ y $G/G = \langle e \rangle$. Es por esto que daremos la siguiente definición, ya que es de importancia en las aplicaciones de los grupos.

DEFINICIÓN 3.11.1

Un grupo se llama *grupo simple* si no tiene subgrupos normales propios.

En el caso abeliano finito, según señalamos, los únicos grupos simples son los de orden primo. En el caso de los grupos no abelianos de orden finito se ha demostrado que un grupo simple tiene que tener un número par de elementos.

Un resultado muy importante en la teoría de Galois es que el grupo A_n es simple para $n \geq 5$.

Estudiaremos resultados que dan una información completa sobre la estructura de una gran cantidad de grupos abelianos. Será, pues, interesante, dado un grupo G no abeliano, construir a partir de este un grupo abeliano lo más similar posible a G . La idea de cociente nos sugiere que con una agrupación adecuada de los elementos de G podemos construir el grupo abeliano que queremos. Para ello necesitaríamos que $ab=ba$ para todo a, b de G , esto es, $(ab)a^{-1} \cdot b^{-1} = e$, lo que nos da la idea de cómo formar el cociente, ya que nos dice cómo tiene que ser la clase del elemento neutro de G . Formalicemos las ideas anteriores con el teorema siguiente:

TEOREMA 3.11.1

El conjunto de los elementos de la forma

$$a_1 b_1 a_1^{-1} b_1^{-1} \cdot a_2 b_2 a_2^{-1} b_2^{-1} \cdot \dots \cdot a_r b_r a_r^{-1} b_r^{-1}, \text{ donde } r \in \mathbb{N} \text{ y } a_i b_i \in G,$$

es un subgrupo normal de G que se llama *conmutador de G* y se denota por G' . El grupo cociente G/G' es abeliano. Más aún G/H es abeliano para un cierto subgrupo H de G si y solo si G' es subgrupo de H .

Demostración

Es fácil probar que el subconjunto de G de los elementos de la forma $a_1 b_1 a_1^{-1} b_1^{-1} \cdot a_2 b_2 a_2^{-1} b_2^{-1} \cdot \dots \cdot a_r b_r a_r^{-1} b_r^{-1}$ es un subgrupo de G , lo cual dejamos al estudiante.

A un elemento de la forma $aba^{-1}b^{-1}$ lo llamaremos conmutador.

Probemos ahora que G' es invariante en G . Para ello debemos probar que si $x \in G'$, entonces $gxg^{-1} \in G'$ para todo $g \in G$.

x es un producto finito de conmutadores, pero basta tomar para la demostración un conmutador, pues si, por ejemplo, $x = aba^{-1}b^{-1}cdc^{-1}d^{-1}$, tomamos $aba^{-1}b^{-1}g^{-1}gcdc^{-1}d^{-1}$, y si

$$gaba^{-1}b^{-1}g^{-1} \in G' \text{ y } gcdc^{-1}d^{-1}g^{-1} \in G',$$

entonces

$$gaba^{-1}b^{-1}cdc^{-1}d^{-1}g^{-1} \in G'.$$

Así, insertando entre cada par de conmutadores $e = g^{-1}g$, quedará resuelto el problema para un producto.

Pasemos a probar que si $x \in G'$, entonces $gxg^{-1} \in G'$ para todo $g \in G$.

En efecto, sea $x = aba^{-1}b^{-1}$. Entonces

$$\begin{aligned} gxg^{-1} &= gaba^{-1}b^{-1}g^{-1} = gaba^{-1}(g^{-1}b^{-1}bg)b^{-1}g^{-1} \\ &= (ga)b(ga)^{-1}b^{-1}bg b^{-1}g^{-1} \in G'. \end{aligned}$$

Queda ahora probar que G/G' es abeliano, esto es, que $aG' \cdot bG' = bG' \cdot aG'$ para todo $a, b \in G$.

$$aG'bG' = abG' = ab(b^{-1}a^{-1}ba)G' = (abb^{-1}a^{-1})baG' = baG' = (bG')(aG')$$

Por último, probemos que G/H es abeliano si y solo si G' es un subgrupo de H .

Si G/H es abeliano, entonces

$$(a^{-1}H) \cdot (b^{-1}H) = (b^{-1}H)(a^{-1}H), \text{ es decir, } a^{-1}b^{-1}H = b^{-1}a^{-1}H,$$

o lo que es lo mismo, $aba^{-1}b^{-1}H = H$, y esto implica que $aba^{-1}b^{-1} \in H$. Como esto ocurre para todo $aba^{-1}b^{-1}$, tenemos que G' es un subgrupo de H .

Recíprocamente, si G' es un subgrupo de H , entonces

$$(aH)(bH) = (ab)H = ab(b^{-1}a^{-1}ba)H = (abb^{-1}a^{-1})baH = baH = bH \cdot aH$$

y G/H será abeliano.

Ejercicios

1. Calcule los siguientes cocientes y si es posible halle para cada uno de ellos un grupo isomorfo.

- | | |
|---|--|
| a) $\mathbb{Z}_{12}/\langle 4 \rangle$ | f) $\mathbb{Z} \times \mathbb{Z}/\langle (1,1) \rangle$ |
| b) $\mathbb{Z}_{30}/\langle 12 \rangle$ | g) $\mathbb{Z} \times \mathbb{Z}/\langle (1,2) \rangle$ |
| c) $\mathbb{Z}_2 \times \mathbb{Z}_4/\langle (1,0) \rangle$ | h) $\mathbb{Z}_4 \times \mathbb{Z}_4/\langle (1,0) \rangle$ |
| d) $\mathbb{Z}_2 \times \mathbb{Z}_4/\langle (0,2) \rangle$ | i) $\mathbb{Z}_4 \times \mathbb{Z}_4/\langle 2 \rangle \times \langle 2 \rangle$ |
| e) $\mathbb{Z}_2 \times \mathbb{Z}_4/\langle (1,2) \rangle$ | |

- Encuentre el grupo G' para el grupo D_4 de las simetrías del cuadrado.
- Pruebe que si dos grupos G y H son isomorfos, entonces sus conmutadores G' y H' son isomorfos y $C(G) \cong C(H)$.
Sugerencia: Considere, dado un isomorfismo ϕ de G en H , la restricción de ϕ a G' y a $C(G)$.
- Pruebe que si H es un subgrupo de G que contiene a G' y H es de índice 2 en G , entonces H es invariante en G y G/H es cíclico de orden 2.
Sugerencia: Observe que $G = H \cup Hg$; por tanto, si $k \in G$, entonces $k = h_1$ o $k = h_2 g$ ($h_1, h_2 \in H$), y utilice la definición de invariante.

3.12 Homomorfismos de grupos y teorema fundamental de homomorfismo

Ya estudiamos un tipo de aplicaciones entre grupos relacionadas con la estructura de estos: los isomorfismos. Vimos que estas aplicaciones daban "copias" de un grupo dado. Sin embargo, hay aplicaciones entre grupos G y G' , llamemos f a una de ellas, que sin ser isomorfismos poseen la propiedad $f(a \cdot b) = f(a) \cdot f(b)$ para todo $a, b \in G$. Como es lógico pensar, al no ser biyectiva f su imagen no será una copia del grupo dominio, pero de todas formas aportará importante información, como veremos en lo que sigue.

DEFINICIÓN 3.12.1

Dados dos grupos G y G' , se llama *homomorfismo de G en G'* a una aplicación $f: G \rightarrow G'$ tal que $f(x \cdot y) = f(x) \cdot f(y) \quad \forall x, y \in G$.

Si f es biyectiva, la definición anterior coincide con la de isomorfismo de grupos.

Algunas propiedades elementales de los homomorfismos son:

Sean G y G' grupos y denotemos por e y e' los neutros de G y G' , respectivamente. Sea $f: G \rightarrow G'$ un homomorfismo. Entonces:

- $f(e)$ es el neutro en G' .
- Si $a \in G$ y a^{-1} es su inverso, entonces $f(a^{-1})$ es el inverso de $f(a)$ en G' .
- Si H es un subgrupo de G , entonces $f(H)$ es un subgrupo de G' . Si H es normal en G , entonces $f(H)$ es normal en $f(G)$.
- Si K es un subgrupo de G' , entonces $f^{-1}(K)$ es un subgrupo de G . Si K es normal en G' , entonces $f^{-1}(K)$ es normal en G .

Todas las propiedades anteriores se demuestran muy fácilmente. Para ejemplificar la técnica de demostración, demostraremos la propiedad 3).

Demostración

Sean H un subgrupo de G y $z, z' \in f(H)$. Debemos probar que $z \cdot z^{-1} \in f(H)$. Pero $z = f(x)$ y $z' = f(x')$ para ciertos $x, x' \in H$. Entonces

$$z(z')^{-1} = f(x) \cdot f(x')^{-1} = f(x) \cdot f(x'^{-1}) = f(x \cdot x'^{-1}), \quad x \cdot x'^{-1} \in H$$

y queda probado que $f(H)$ es un subgrupo de G' .

En cuanto a que si H es invariante o normal en G , entonces $f(H)$ lo es en $f(G)$, habría que probar que $bzb^{-1} \in f(H)$ para todo $z \in f(H)$ y $b \in f(G)$. Tomemos $b = f(a)$, $z = f(x)$ y tendremos

$$f(a)f(x)f(a)^{-1} = f(axa^{-1}).$$

y como $axa^{-1} \in H$, queda probado que $f(H)$ es normal en $f(G)$.

Como se observa, la técnica de demostración es similar a la utilizada para demostrar las propiedades de los isomorfismos; pero se debe ser muy cuidadoso, ya que en este caso no es posible usar la condición de biyectividad.

La siguiente definición nos permitirá relacionar el concepto de homomorfismo de grupo con el de grupo cociente.

DEFINICIÓN 3.12.2 (Núcleo o kernel de un homomorfismo de grupos)

Sea $f: G \rightarrow G'$ un homomorfismo de grupos. Se llama *núcleo o kernel de f* y se denota por $\text{Ker } f$ a la preimagen por f del elemento neutro e' de G' . Esto es:

$$\text{Ker } f = f^{-1}(e') = \{x \in G : f(x) = e'\}.$$

PROPOSICIÓN 3.12.1

Dado $f: G \rightarrow G'$ homomorfismo, $\text{Ker } f$ es un subgrupo invariante de G .

Demostración

Según las propiedades anteriores, el núcleo de un homomorfismo $f: G \rightarrow G'$ es un subgrupo invariante de G , ya que es la preimagen de e' que es un subgrupo invariante de G' .

Mediante la definición de núcleo podemos dar una caracterización de la inyectividad de un homomorfismo.

PROPOSICIÓN 3.12.2

Un homomorfismo $f: G \rightarrow G'$ es inyectivo si y solo si $\text{Ker } f = \{e\}$, donde e es el elemento neutro de G .

Demostración

Si f es inyectivo, sea $x \in \text{Ker } f$, entonces $f(x) = e' = f(e)$. Luego $x = e$ y de aquí $\text{Ker } f = \{e\}$.

Recíprocamente, si $\text{Ker } f = \{e\}$, sean $x, x' \in G$, entonces $f(x) = f(x') \iff f(x)f(x')^{-1} = e' \iff f(x \cdot x'^{-1}) = e' \iff x \cdot x'^{-1} \in \text{Ker } f \iff x \cdot x'^{-1} = e \iff x = x'$.

Veamos ahora ejemplos de homomorfismos y de núcleos de homomorfismos.

Ejemplos

1. Dado un grupo G , la aplicación $\text{id}_G : G \rightarrow G$ dada por $\text{id}_G(x) = x$ para todo $x \in G$ es el caso más sencillo de homomorfismo de grupos.
2. Dado el grupo $(\mathbb{Z}, +)$, la aplicación $f : \mathbb{Z} \rightarrow \mathbb{Z}_n$ dada por $m \mapsto r$ donde $m = nq + r$ (esto es, el resto de la división de m entre n), es un homomorfismo del grupo $(\mathbb{Z}, +)$ en el grupo $(\mathbb{Z}_n, +)$ cuyo núcleo es $n\mathbb{Z}$.
3. La aplicación $\varphi : (GL(n, \mathbb{R}), \cdot) \rightarrow (\mathbb{R}^*, \cdot)$ dada por $\varphi(A) = \det A$, es un homomorfismo de grupos cuyo núcleo es $SL(n, \mathbb{R})$.
4. Si consideramos el grupo cíclico $C_2 = \{1, -1\}$, subgrupo de $(\mathbb{R}^*, \cdot)$, y la aplicación $g : S_3 \rightarrow C_2$ dada por $g(\sigma) = s\sigma s$, g será un homomorfismo de grupos cuyo núcleo es A_3 .
5. Si G es un grupo y H un subgrupo normal de G , la aplicación $\varphi : G \rightarrow G/H$ dada por $a \mapsto aH$ es un homomorfismo de grupos al que llamamos *homomorfismo canónico de G sobre G/H* . Este homomorfismo canónico no es más que la sobrección canónica en el caso de un grupo y su cociente.

Ahora estamos en condiciones de relacionar el concepto de *homomorfismo de grupos* con el concepto de *grupo cociente*.

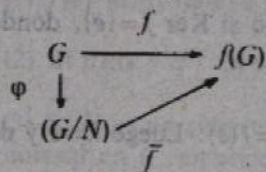
En efecto, el ejemplo 5 permite ver que para cada subgrupo normal H de G , existe un homomorfismo con dominio G del cual H es núcleo, pues $\text{Ker } \varphi = \{x \in G : xH = H\} = H$.

Por otra parte, habíamos señalado que para cada homomorfismo f de G en G' existe un subgrupo normal de G que corresponde de manera natural a f , este es $\text{Ker } f$. Al igual que en el caso de las aplicaciones entre conjuntos, no es difícil ver que el homomorfismo que tiene como núcleo un subgrupo normal dado no es único, de ahí que estudiemos el teorema siguiente:

TEOREMA 3.12.1 (Teorema fundamental de homomorfismo.)

Sea f un homomorfismo de G en G' y sea $N = \text{Ker } f$. Entonces existe un único homomorfismo biyectivo $\bar{f}$ de G/N en $f(G)$ tal que $\bar{f} \circ \varphi = f$, donde $\varphi : G \rightarrow G/N$ es el homomorfismo canónico.

Este resultado puede representarse mediante el diagrama siguiente:



Demostración

Al estudiar las relaciones de equivalencia demostramos el teorema de descomposición de una aplicación f , que se refería a un problema similar al que tenemos ahora. En aquella ocasión consideramos la relación de equivalencia R asociada a f^* en lugar de la relación $xRy \Leftrightarrow xy^{-1} \in \text{Ker } f$.

La relación de equivalencia asociada a f^* está dada por

$$xRy \Leftrightarrow f(x) = f(y) \Leftrightarrow f(x) f(y)^{-1} = f(xy^{-1}) = e' \Leftrightarrow xy^{-1} \in \text{Ker } f.$$

Luego, en este caso coincide la relación dada por el núcleo de f con la asociada a f , y podemos aplicar el teorema de descomposición de una aplicación.

Existirá, pues, según dicho teorema una única aplicación biyectiva $\bar{f}$ de G/N sobre $f(G)$ dada por $\bar{f}(Cx) = \bar{f}(xN) = f(x)$ y que cumple $\bar{f} = f \circ \varphi$. Quedará solo probar que $\bar{f}$ es un homomorfismo de grupos. En efecto,

$$\begin{aligned} \bar{f}[(x \cdot N) \cdot (y \cdot N)] &= \bar{f}(xy \cdot N) = f(xy) = f(x) \cdot f(y) = \bar{f} \circ \varphi(x) \cdot \bar{f} \circ \varphi(y) \\ &= \bar{f}(x \cdot N) \cdot \bar{f}(y \cdot N) \end{aligned}$$

y, por tanto, es un isomorfismo de grupos.

Hemos demostrado el teorema aplicando el resultado que conocíamos anteriormente, pero el estudiante puede comprobar que también puede demostrarse directamente definiendo $\bar{f} : G/N \rightarrow f(G)$ por $\bar{f}(x \cdot N) = f(x)$ y comprobando que es un isomorfismo y que cumple $f = \bar{f} \circ \varphi$.

¿Qué conclusiones podemos sacar del teorema 3.12.1?

El teorema nos dice que para un homomorfismo de grupo la imagen, salvo el nombre de los elementos, es sencillamente $G/\text{Ker } f$ y el homomorfismo $\bar{f}$ es esencialmente el homomorfismo canónico φ , ya que difiere de φ en un isomorfismo.

Tenemos, pues, que cada homomorfismo de un grupo G determina un subgrupo invariante $\text{Ker } f$ y, por tanto, una estructura cociente $G/\text{Ker } f$.

Recíprocamente, dado un subgrupo invariante N , existe un homomorfismo del cual N es núcleo; este es el homomorfismo canónico $\varphi : G \rightarrow G/N$.

Los otros homomorfismos que tienen a N como núcleo difieren de φ en un isomorfismo, y en ese sentido podemos decir que existe un único homomorfismo que tiene a N como núcleo.

Hemos establecido así la correspondencia:

Subgrupos invariantes-grupos cocientes-homomorfismo de grupo.

Podemos hablar de uno y otro indistintamente, lo cual nos sirve para sustituir un concepto por otro en la búsqueda de resultados. Por ejemplo, si queremos investigar sobre los homomorfismos de un grupo de orden primo, podemos hacerlo buscando los subgrupos invariantes de dicho grupo.

En los ejemplos de homomorfismo de grupos que vimos, tenemos mediante este teorema los isomorfismos siguientes:

1. $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n$
2. $GL(n, \mathbb{R})/SL(n, \mathbb{R}) = \mathbb{R}^*$
3. $S_3/A_3 = \{1, -1\} = \mathbb{Z}_2$

Como ejemplo de la aplicación del teorema fundamental de homomorfismo de grupos, estudiaremos dos de los llamados teoremas de isomorfismo.

La técnica utilizada en la demostración de estos teoremas debe ser estudiada con atención, ya que es útil en la demostración de resultados similares. Los nombres de primer y segundo teorema que les daremos varían según los autores.

TEOREMA 3.12.2 (Primer teorema de isomorfismo)

Sean H y K subgrupos invariantes de G , y K subgrupo de H . Entonces existe un isomorfismo natural de G/H sobre $G/K/H/K$.

Demostración

La idea fundamental de la demostración de este teorema consiste en buscar un homomorfismo sobreectivo de G sobre $G/K/H/K$ tal que su núcleo sea H . El isomorfismo resulta entonces directamente del teorema fundamental de homomorfismo.

Primero debemos comprobar que tiene sentido considerar $G/K/H/K$.

Tomando el homomorfismo canónico $p: G \rightarrow G/K$ vemos que $p_K(H)$ es precisamente H/K , y como $p(H)$ es la imagen por un homomorfismo de un subgrupo invariante de G , H/K será invariante en G/K . Luego, tiene sentido hablar de $G/K/H/K$.

Ahora definamos

$\Phi: G \rightarrow G/K/H/K$ dada por $\Phi(a) = (aK)(H/K)$ para todo $a \in G$.

Está claro que

$$\Phi(a \cdot b) = (abK)(H/K) = [(aK)(bK)](H/K) = [(aK)(H/K)] \cdot [(bK)(H/K)] = \Phi(a) \cdot \Phi(b)$$

Así, Φ es un homomorfismo cuya imagen es $G/K/H/K$.

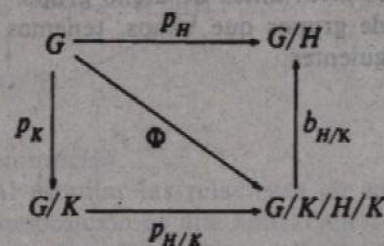
El núcleo de Φ está formado por los $x \in G$ tales que $\Phi(x) = H/K$; estos son exactamente los elementos de H .

Entonces el teorema fundamental prueba que G/H es naturalmente isomorfo a $(G/K)/(H/K)$.

El diagrama siguiente resume la situación:

p_H, p_K y $p_{H/K}$ indican los homomorfismos canónicos.

$b_{H/K}$ indica el isomorfismo de $G/K/H/K$ en G/H dado por el teorema fundamental de homomorfismo.



TEOREMA 3.12.3 (Segundo teorema de isomorfismo)

Sean G un grupo y N y H subgrupos invariantes de G . Entonces $H \cap N$ es invariante en H . $HN = \{hn : h \in H, n \in N\}$ es un subgrupo de G y existe un isomorfismo natural de $H/H \cap N$ en HN/N .

Demostración

1) Probemos que $H \cap N$ es invariante en H .

Sea hah^{-1} con $h \in H$ y $a \in H \cap N$. Entonces $hah^{-1} \in N$ pues N es invariante en G . Pero $hah^{-1} \in H$ ya que $h \in H$ y $a \in H$. Luego $hah^{-1} \in H \cap N$.

2) Probemos ahora que HN es un subgrupo de G . (En general HN no es un subgrupo de G ; en este caso interviene el hecho de que N es invariante en G .)

Sean hn y $h'n'$ elementos de HN . Entonces

$$(hn)(h'n')^{-1} = \underbrace{hnn^{-1}}_{n''} h'^{-1} = hn''h'^{-1} = \underbrace{hh'^{-1}}_{h''} h'n''h'^{-1} = h''h'n''h'^{-1},$$

pero $h''h'n''h'^{-1} \in N$ ya que N es invariante en G .

Luego $h''h'n''h'^{-1} \in HN$ ya que $h'n''h'^{-1} \in N$ y queda probado que HN es un subgrupo de G .

Tiene, pues, sentido hablar de $H/H \cap N$ y también de HN/N , ya que al ser N invariante en G es obvio que es invariante en el subgrupo HN de G .

Ahora aplicaremos la técnica usada en el teorema 3.12.2 de buscar un homomorfismo Φ , en este caso de H en HN/N y tal que su núcleo sea $H \cap N$.

Consideremos $\Phi: H \rightarrow HN/N$ dado por $\Phi(h) = hN$.

Dados $h, h' \in H$, $\Phi(hh') = hh'N = (hN)(h'N) = \Phi(h)\Phi(h')$, luego Φ es un homomorfismo.

Φ es sobreectivo ya que para $hnN \in HN/N$ tomamos $h \in H$ y entonces $\Phi(h) = hn \cdot N$.

El núcleo de Φ es $H \cap N$, ya que

$$\text{Ker } \Phi = \{x \in N : hN = N\} = \{x \in H : x \in N\} = H \cap N.$$

Aplicando el teorema fundamental de homomorfismo existirá un isomorfismo $\psi: H/H \cap N \rightarrow HN/N$ según queríamos probar.

Ejemplos

1. La técnica de aplicación del teorema fundamental de homomorfismo de grupos puede ser usada también en ejercicios prácticos. El siguiente ejemplo fundamenta, desde el punto de vista de la teoría de grupos, la representación de los ángulos mediante el círculo unitario.

Probaremos que si consideramos el grupo $(\mathbb{R}, +)$, entonces el cociente $\mathbb{R}/\langle 2\pi \rangle$ es isomorfo a $(\mathbb{C}_1, \cdot)$ donde $\mathbb{C}_1 = \{z \in \mathbb{C} : |z| = 1\}$.

En efecto, consideremos

$$\phi: (\mathbb{R}, +) \rightarrow (\mathbb{C}_1, \cdot) \text{ dada por}$$

$$\phi(x) = \cos x + i \sin x.$$

Se puede comprobar fácilmente que φ es un homomorfismo.
 $\text{Ker } \varphi = \{x : \cos x + i \sin x = 1\} = \{2n\pi : n \in \mathbb{Z}\} = \langle 2\pi \rangle$.

$$\varphi(\mathbb{R}) = \{z \in \mathbb{C} : z = \cos x + i \sin x, x \in \mathbb{R}\} = \{z \in \mathbb{C} : |z| = 1\} = C_1.$$

Según el teorema fundamental de homomorfismo de grupos.

$$\mathbb{R}/\langle 2\pi \rangle \cong C_1.$$

En este caso la aplicación de la técnica consiste en buscar el φ apropiado, de forma que el núcleo y la imagen den lugar al isomorfismo buscado.

Los ángulos en radianes se definen como clases de equivalencia módulo 2π , y está claro que los números complejos de módulo 1 recorren el círculo unitario.

2. Analicemos si se cumple la siguiente propiedad:

Dados $n, m, d \in \mathbb{N}$ y $n = m \cdot d$, entonces $\mathbb{Z}_d = \mathbb{Z}_n / \mathbb{Z}_m$.

En efecto, podemos demostrar que esta propiedad se cumple aplicando el teorema fundamental y el segundo teorema de isomorfismo.

Observemos que $n\mathbb{Z} \subset d\mathbb{Z}$ y si definimos $\varphi : \mathbb{Z} \rightarrow d\mathbb{Z}/n\mathbb{Z}$ dada por $x \mapsto dx + n\mathbb{Z}$, φ es un homomorfismo sobreyectivo cuyo núcleo es $\text{Ker } \varphi = m\mathbb{Z}$. Entonces por el teorema fundamental, $\mathbb{Z}/m\mathbb{Z} \cong d\mathbb{Z}/n\mathbb{Z}$.

Aplicando el primer teorema de isomorfismo, tenemos

$$\mathbb{Z}/d\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z} / d\mathbb{Z}/n\mathbb{Z}, \text{ o sea, } \mathbb{Z}_d \cong \mathbb{Z}_n / \mathbb{Z}_m.$$

(Dejamos al estudiante que compruebe la validez de las hipótesis.)

3. Veamos ahora una aplicación práctica del segundo teorema de isomorfismo.

Consideremos el grupo simétrico S_4 y sus subgrupos

$$\tilde{V}_4 = \{(1,2,3,4), (2,1,4,3), (3,4,1,2), (4,3,2,1)\}$$

$$\tilde{S}_3 = \{(1,2,3,4), (1,3,2,4), (2,1,3,4), (2,3,1,4), (3,1,2,4), (3,2,1,4)\}.$$

Se puede comprobar que $\tilde{V}_4$ es invariante en S_4 .

Evidentemente $\tilde{S}_3 \cap \tilde{V}_4 = \{(1,2,3,4)\}$ y $\{(1,2,3,4)\}$ es el neutro de S_4 .

Según el segundo teorema de isomorfismo:

$$\tilde{S}_3 \tilde{V}_4 / \tilde{V}_4 \cong \tilde{S}_3 / \tilde{S}_3 \cap \tilde{V}_4 \cong \tilde{S}_3,$$

pero como $o(\tilde{S}_3 \tilde{V}_4) = 24$ resulta que $\tilde{S}_3 \tilde{V}_4 = S_4$, ya que según el ejercicio 5 del epígrafe 3.10, $\tilde{S}_3 \tilde{V}_4$ es un subgrupo de S_4 . De aquí que $S_4 / \tilde{V}_4 \cong \tilde{S}_3$.

Ejercicios

1. Determine cuáles de las siguientes aplicaciones son homomorfismos de grupo. Si lo son halle su núcleo y su imagen.

a) $f : (\mathbb{R}, +) \rightarrow (\mathbb{Z}, +)$ dada por $f(x) =$ menor entero, menor o igual que x .

b) $\varphi : (\mathbb{R}^*, \cdot) \rightarrow (\mathbb{R}^*, \cdot)$ dada por $\varphi(x) = |x|$.

c) $g : \mathbb{Z}_6 \rightarrow \mathbb{Z}_2$ dada por $g(x) =$ resto de x al ser dividido por 2.

2. Pruebe que dados $(\mathbb{Q}, +)$ y su subgrupo $\mathbb{Z}$, $\mathbb{Q}/\mathbb{Z}$ es isomorfo al grupo formado por $\{e^{2\pi i q}, 0 \in \mathbb{Q}\}$ con la ley $\cdot$.

3. Pruebe que dado el grupo $A_1(\mathbb{R}) = \{\varphi_{ab} : a \neq 0\}$ del ejercicio 1.f del epígrafe 3.6 y su subgrupo $H = \{\varphi_{1,b} : b \in \mathbb{R}\}$, entonces $A_1(\mathbb{R})/H$ es isomorfo a $(\mathbb{R}^*, \cdot)$.

4. Pruebe que si $0 : G \rightarrow G'$ es un homomorfismo y $o(G)$ es finito, entonces $o(0(G))/o(G)$ (o sea, el orden de G es un múltiplo del orden de la imagen de 0).

5. ¿Qué se puede decir acerca de los homomorfismos de un grupo simple?

6. Diga si las siguientes proposiciones son verdaderas o falsas y justifique sus respuestas.

a) La imagen de un grupo de 6 elementos por un homomorfismo puede tener 4 elementos.

b) La imagen de un grupo de 6 elementos por un homomorfismo puede tener 12 elementos.

c) Existe un homomorfismo de un grupo de 6 elementos en un grupo de 12 elementos.

d) Existe un homomorfismo de un grupo de 6 elementos en un grupo de 10 elementos.

7. Sean dos grupos G_1 y G_2 y la aplicación $\pi_1 : G_1 \times G_2 \rightarrow G_1$ dada por $\pi_1(x, y) = x$. Pruebe que π_1 es un homomorfismo y halle su núcleo. ¿A qué grupo es isomorfo, el núcleo de π_1 ?

8. Sean G un grupo, a un elemento de G y $\phi : \mathbb{Z} \rightarrow G$ definida por $\phi(n) = a^n$. Pruebe que ϕ es un homomorfismo. Describa la imagen de ϕ y las posibilidades para su núcleo.

9. Sean G un grupo abeliano finito de orden n y r un entero positivo primo relativo con n .

a) Pruebe que $\phi_r : G \rightarrow G$ dada por $\phi_r(a) = a^r$ es un automorfismo de G .

b) Deduzca que la ecuación $x^r = a$ siempre tiene una solución única en G si r es relativamente primo con el orden de G . ¿Qué puede pasar si r no es primo relativo con el orden de G ?

10. Sean G un grupo e $\text{Inn}(G)$ el grupo de los automorfismos interiores de G (ver ejercicio 11 del epígrafe 3.8).

a) Pruebe que $\phi : G \rightarrow \text{Inn}(G)$ dada por $\phi(a) = i_a^-$ es un homomorfismo de G sobre $\text{Inn}(G)$.

b) Halle el núcleo de ϕ y describa cuándo este es un isomorfismo.

11. Sean G y G' grupos y H y H' subgrupos normales de G y G' , respectivamente. Sea ϕ un homomorfismo de G en G' . Pruebe que ϕ induce un homomorfismo natural $\bar{\phi} : G/H \rightarrow G'/H'$ si $\phi(H) \subset H'$.
12. Sean G y G' grupos y $\phi : G \rightarrow G'$ un homomorfismo sobreyectivo. Sean N' un subgrupo normal de G' y $N = \{x \in G : \phi(x) \in N'\}$. Pruebe que $G/N \cong G'/N'$.
13. Sea ϕ un automorfismo de un grupo G que deja fijo solamente el neutro del grupo, esto es, $\phi(a) = a \Rightarrow a = e$.
 - a) Pruebe que la aplicación $a \mapsto \phi(a) \cdot a^{-1}$ es inyectiva y de aquí que si G es finito, entonces todo elemento de G tiene la forma $\phi(a) \cdot a^{-1}$.
 - b) Pruebe usando a) que si G es finito y $\phi^2 = \text{id}_G$, entonces G es abeliano de orden par.
14. Sean $|(Q^*, \cdot)|$ y $N = \{1, -1\}$. Sea $H = \langle \frac{1}{2} \rangle$. Encuentre $H \cdot N$, $H \cdot N/N$ y verifique que $H \cdot N/N \cong H/H \cap N$.
15. Sean G un grupo y M y N subgrupos normales de G . Sea H un subgrupo de G . Pruebe que $H \cap M = H \cap N \Rightarrow HM/M \cong H \cdot N/N$.
16. Sean H y K subgrupos de G , N un subgrupo invariante de G y $HN = K \cdot N$. Pruebe que $H/H \cap N \cong K/K \cap N$.

3.13 Grupos cíclicos

Al estudiar los subgrupos de un grupo G , lo cual es en definitiva una forma de estudiar diferentes clases de grupos, encontramos un tipo especial de subgrupos: los cíclicos, esto es, los subgrupos de la forma $H = \langle a^n : n \in \mathbb{Z} \rangle = \langle a \rangle$, $a \in G$. Vimos también en los ejemplos que a veces ocurría que para algún subgrupo cíclico, H de G , $G = H$, es decir, $G = \langle a^n : n \in \mathbb{Z} \rangle$.

El hecho de que los elementos de un grupo G sean de una forma conocida (potencias de un elemento fijo) permite obtener mucha información sobre dicho grupo. Vamos, pues, a estudiar los grupos cíclicos, que en cierto sentido pueden considerarse los grupos más simples.

DEFINICIÓN 3.13.1

Se dice que un grupo G es *cíclico* si y solo si existe $a \in G$ tal que $G = \langle a \rangle = \langle a^n : n \in \mathbb{Z} \rangle$. Al elemento a lo llamamos *generador* de G .

Ejemplos:

1. $\mathbb{Z}$ es cíclico, pues $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$.
2. $n\mathbb{Z} = \langle n \rangle = \langle -n \rangle$, luego es cíclico.
3. $\mathbb{Z}_n = \langle 1 \rangle = \langle \bar{1} \rangle$, por lo que es cíclico. En general es fácil comprobar que $\mathbb{Z}_n$ es cíclico para todo n .
4. S_3 no es cíclico ya que no existe elemento x de S_3 tal que todo elemento de S_3 pueda escribirse como x^n . Esto puede comprobarse fácilmente en la

tabla de S_3 mediante la búsqueda de los subgrupos cíclicos generados por todos los elementos de S_3 .

5. V_4 no es cíclico. Para comprobarlo se utiliza el método descrito en el ejemplo 4.
6. $\mathbb{Z}_2 \times \mathbb{Z}_2$ es cíclico, pues $\mathbb{Z}_2 \times \mathbb{Z}_2 = \langle (1, 1) \rangle = \langle (1, 2) \rangle$.
7. En $\mathbb{C}$, consideremos el grupo C_n de las raíces de la unidad. Recordemos que dichas raíces están caracterizadas por la propiedad de que ξ es raíz n -sima de la unidad si y solo si $\xi^n = 1$. C_n está formado por n elementos $\xi_0, \xi_1, \dots, \xi_{n-1}$, donde $\xi_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$, $0 \leq k \leq n-1$.

Además, $C_n = \{\xi_1^n : n \in \mathbb{Z}\}$. Luego, C_n es un grupo cíclico de orden n , ya que tenemos $C_n = \langle \xi_1 \rangle$.

Recordemos también que C_n puede expresarse como las potencias de cualquier raíz primitiva de la unidad. Las raíces primitivas (que están dadas por los ξ_k tales que k es primo relativo con n) son los otros generadores del grupo cíclico C_n .

En los ejemplos anteriores podemos observar que los grupos cíclicos pueden tener más de un generador.

Podemos ver que todos los ejemplos dados de grupos cíclicos son grupos abelianos. Esta es una propiedad general de los grupos cíclicos que puede demostrarse muy fácilmente.

PROPOSICIÓN 3.13.1

Todo grupo cíclico es abeliano.

Demostración

En efecto, sea $G = \langle a \rangle$. Entonces dos elementos cualesquiera de G son, por ejemplo, $x = a^n$, $y = a^m$ y $x \cdot y = a^n \cdot a^m = a^{n+m} = a^{m+n} = a^m \cdot a^n = y \cdot x$; de aquí que G sea conmutativo.

Evidentemente, el recíproco de la proposición anterior no se cumple. Por ejemplo, V_4 es abeliano, pero no cíclico. Sin embargo, veremos más adelante que una amplia clase de grupos abelianos puede construirse a partir de grupos cíclicos.

Otra propiedad que podemos intuir de los ejemplos es la referente a la forma de los subgrupos de un grupo cíclico.

En el epígrafe 3.6 hemos dado la demostración de que los únicos subgrupos de $\mathbb{Z}$ son los $n\mathbb{Z}$, que son cíclicos. Los subgrupos propios de $\mathbb{Z}_n$ son exactamente $\{0, 2, 4\}$ y $\{0, 3\}$, que son cíclicos. Además, es evidente que en un grupo cíclico G , los subgrupos triviales $\{0\}$ y G son cíclicos.

En el caso de un grupo cíclico es posible obtener el siguiente resultado referente a la forma de los subgrupos.

TEOREMA 3.13.1 Los subgrupos de un grupo cíclico son cíclicos.

Demostración

Efectivamente, sea $G = \langle a \rangle$ un grupo cíclico y sea H un subgrupo de G . Tenemos que encontrar $b \in H$ tal que $H = \langle b \rangle$.

Está claro que si $H = \{e\}$ o $H = G$ se cumple el teorema; así que supongamos que H es un subgrupo propio de G .

Si H es un subgrupo propio, podemos asegurar que existe un $s \in \mathbb{Z}_+^*$ tal que $a^s \in H$, o sea, que existen en H potencias positivas de a . Entonces podemos tomar el menor entero positivo m tal que $a^m \in H$. El generador será precisamente $b = a^m$. Por tanto, debemos probar que $H = \langle b \rangle = \langle a^m \rangle$.

Como evidentemente $\langle a^m \rangle \subset H$ por ser H un subgrupo y $a^m \in H$, queda probar que $H \subset \langle a^m \rangle$. En otras palabras, debemos probar que todo $h \in H$ es igual a alguna potencia de a^m , o sea, $h = (a^m)^q$ para algún q .

Como $h \in H$, $h = a^n$ para algún $n \in \mathbb{Z}$, podemos escribir por el algoritmo de la división con resto para los números enteros:

$$n = mq + r \text{ con } 0 < r < m \text{ o } r = 0.$$

Entonces

$$h = a^n = a^{mq+r} = a^{mq} \cdot a^r,$$

de donde

$$a^r = a^{n-mq} = a^n \cdot a^{-mq} \text{ y } a^r \in H.$$

Pero para que esto no implique una contradicción, ya que hemos supuesto a m como el menor entero positivo tal que $a^m \in H$, es necesario que $r = 0$, de donde se infiere que $h = a^{mq} = (a^m)^q$ según queríamos demostrar.

La demostración anterior es clásica dentro del Álgebra, ya que en ella utilizamos técnicas que el estudiante volverá a encontrar con frecuencia. Por ejemplo, el uso de la existencia de un menor entero positivo en los conjuntos que estén en correspondencia con un subconjunto de $\mathbb{Z}_+^*$ o el uso del algoritmo de la división con resto en situaciones similares a la dada.

Del teorema podemos inferir algunos resultados, como son:

- 1) Los generadores de los subgrupos de un grupo cíclico G son de la forma a^m , donde a es un generador de G .
- 2) Los subgrupos de $\mathbb{Z}$ son exactamente los $n\mathbb{Z}$, según habíamos demostrado anteriormente, ya que los únicos generadores de $\mathbb{Z}$ son 1 y -1 , lo que implica que los generadores de los subgrupos sean del tipo n o $-n$.
- 3) Un grupo puede tener todos sus subgrupos propios cíclicos y no ser cíclico, como podemos ver en el ejemplo de S_4 .

Hemos visto ejemplos de grupos cíclicos finitos e infinitos. Más aún, los ejemplos de $\mathbb{Z}_n$ y $\mathbb{C}_n$ indican que para cada $n \in \mathbb{Z}$ existe al menos un grupo cíclico finito de orden n . Ahora nos proponemos investigar cuántos grupos cíclicos distintos existen, esto es, cómo mediante el concepto *isomorfismo* podemos hacer una clasificación de los grupos cíclicos.

Clasificación de los grupos cíclicos

Sabemos que un grupo finito y uno infinito no pueden ser isomorfos, luego es natural separar el análisis para la clasificación de los grupos cíclicos en dos casos:

1ro. Caso infinito

Sea $G = \langle a \rangle$ infinito.

Intuitivamente y guiándonos por los ejemplos vemos que todas las potencias de a deben ser distintas. En efecto, demostremos que si aparecen potencias repetidas, el grupo tiene que ser de orden finito.

Sea, pues, $n, m \in \mathbb{Z}$, $n \neq m$ y $a^n = a^m \Rightarrow a^{n-m} = e$.

Sin pérdida de generalidad podemos suponer que $n > m$, con lo que habremos encontrado que existen potencias positivas k de a tales que $a^k = e$.

Aplicaremos de nuevo las técnicas utilizadas en la demostración de que los subgrupos de un grupo cíclico son cíclicos.

Tomemos m el menor entero positivo tal que $a^m = e$ y probemos que $G = \{a^0, a^1, \dots, a^{m-1}\}$.

Sea $a^n \in G$.

Podemos escribir $n = mq + r$ con $0 \leq r < m$, y entonces

$$a^n = a^{mq+r} = a^{mq} \cdot a^r = (a^m)^q \cdot a^r = a^r,$$

con lo que queda probada nuestra afirmación, pues toda potencia a^n de a coincide con una de la forma a^r con $0 \leq r < m$.

Lo que demostramos equivale a decir que si G es infinito, entonces

$$n \neq m \Rightarrow a^n \neq a^m.$$

De manera natural surge la siguiente comparación entre el grupo cíclico $(\mathbb{Z}, +)$ y un grupo cíclico infinito G .

Sea la aplicación: $\varphi: \mathbb{Z} \longrightarrow G = \langle a \rangle$ dada por $n \mapsto a^n$.

Evidentemente es un homomorfismo de grupos, ya que

$$\varphi(n+m) = a^{n+m} = a^n \cdot a^m.$$

φ es inyectiva porque al ser G infinito hemos probado que $n \neq m \Rightarrow a^n \neq a^m$, y es sobreyectiva pues por la definición de G todo $x \in G$ es de la forma a^n , y evidentemente $\varphi(n) = x$.

Hemos demostrado el resultado siguiente:

TEOREMA 3.13.2

Todo grupo cíclico infinito es isomorfo a $\mathbb{Z}$.

Podemos decir que, salvo isomorfismo, existe un único grupo cíclico infinito.

2do. Caso finito

Sea $G = \langle a \rangle$ finito de orden m .

En este caso es evidente, por los razonamientos hechos en el caso infinito, que no todas las potencias de a pueden ser distintas. Más aún, tendremos que $G = \{a^0, a^1, \dots, a^{m-1}\}$, donde m es el menor entero positivo tal que

$$a^m = e.$$

Si consideramos de nuevo el homomorfismo $\varphi: \mathbb{Z} \rightarrow G = \langle a \rangle$ dado por $n \mapsto a^n$, será un homomorfismo sobreyectivo aunque no inyectivo, pues su núcleo $\text{Ker } \varphi = \{n \in \mathbb{Z} : a^n = e\}$ está formado exactamente por los múltiplos de m , es decir, $\text{Ker } \varphi = m\mathbb{Z}$.

Aplicando el teorema fundamental de homomorfismo de grupos, tenemos

$$\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z} = G = \langle a \rangle.$$

De lo anterior podemos deducir:

Todo grupo cíclico finito de orden m es isomorfo a $\mathbb{Z}_m$.

Dos grupos cíclicos finitos de orden m son isomorfos entre sí.

Existe, salvo isomorfismo, un único grupo cíclico finito de orden m para cada m entero positivo.

Podemos tomar como "modelos" de los grupos cíclicos finitos los grupos $\mathbb{Z}_m$.

El hecho de que dos grupos cíclicos finitos $G = \langle a \rangle$ y $G' = \langle b \rangle$ del mismo orden sean isomorfos, puede ser demostrado directamente sin el uso del teorema fundamental de homomorfismo mediante la aplicación $a^n \mapsto b^n$, lo que dejamos como ejercicio para el estudiante, ya que preferimos destacar una vez más la utilidad del teorema fundamental de homomorfismo en nuestra demostración.

Orden de un elemento

Mediante la noción de grupo cíclico es posible obtener una clasificación de los elementos de un grupo G cualquiera, la cual es de gran utilidad en la obtención de información sobre el grupo.

DEFINICIÓN 3.13.2

Sea G un grupo. Se dice que un elemento a de G es de orden infinito o de orden finito m , si el subgrupo cíclico $\langle a \rangle$ de G es de orden infinito o de orden finito m , respectivamente. El orden de a se denota por $\alpha(a)$.

Está claro que si $\alpha(a)$ es infinito, $a^n \neq e$ para $n \neq 0$, mientras que si $\alpha(a) = m$, $a^m = e$ y m es el menor entero positivo que posee esta propiedad. Más aún, si $a^n = e$, entonces n es un múltiplo de m .

Para grupos finitos son fáciles de comprobar las siguientes propiedades, que pueden ser consideradas corolarios del teorema de Lagrange.

PROPOSICIÓN 3.13.2

Si G es un grupo finito y $a \in G$, entonces $\alpha(a)$ divide a $\alpha(G)$.

Demostración

En efecto, si consideramos el subgrupo $\langle a \rangle$ de G , por el teorema de Lagrange $\alpha(a) = \alpha(\langle a \rangle)$ debe ser un divisor del orden de G .

PROPOSICIÓN 3.13.3

Si G es un grupo finito y $a \in G$, entonces $a^{\alpha(G)} = e$.

Demostración

De acuerdo con la proposición anterior, $m = \alpha(a)$ divide a $\alpha(G) = n$. Luego $n = m \cdot q$ y de aquí:

$$a^n = a^{mq} = (a^m)^q = e.$$

Subgrupos de un grupo cíclico

Señalamos anteriormente que los grupos cíclicos pueden considerarse, en cierto sentido, como los grupos más sencillos. En el estudio de la teoría de grupos existen problemas importantes y no fáciles de resolver para un grupo cualquiera que, sin embargo, es posible solucionar para la clase de los grupos cíclicos. Por ejemplo, podemos determinar todos los subgrupos de un grupo cíclico, de lo cual nos ocuparemos inmediatamente. Estos resultados servirán de ayuda en la búsqueda de la solución al mismo problema en clases más amplias de grupos.

En la búsqueda de los subgrupos de un grupo cíclico ya encontramos que dichos subgrupos deben ser cíclicos. Sin embargo, quedan por contestar algunas preguntas importantes con respecto a este problema, como son:

Si $G = \langle a \rangle$ y H es un subgrupo de G , sea $H = \langle a^m \rangle$, ¿Cuál es el orden de H ?

En el caso finito:

¿Se cumple el recíproco de Lagrange? Esto es, para cada divisor del orden del grupo, ¿existe un subgrupo de dicho orden?

¿Se puede determinar la cantidad de subgrupos distintos que tiene G ?

Pasemos a responder dichas preguntas mediante el teorema siguiente:

TEOREMA 3.13.3

(1) Si $G = \langle a \rangle$ es infinito, los subgrupos no nulos de G son infinitos y la aplicación $s \mapsto \langle a^s \rangle$ es una aplicación biyectiva de $\mathbb{N}$ en el conjunto de los subgrupos de $G = \langle a \rangle$.

(2) Si $G = \langle a \rangle$ es finito de orden n y q es un divisor de n , existe un único subgrupo de G de orden q .

Demostración

En la demostración de que los subgrupos de un grupo cíclico son cíclicos, vimos que un subgrupo H de G es de la forma $H = \langle a^m \rangle$, donde m es el menor entero positivo tal que $a^m \in H$.

1) Si $G = \langle a \rangle$ es infinito sabemos que $r \neq s \Rightarrow a^r \neq a^s$. Luego $a^{mr} \neq a^{ms}$ para todo $r, s \in \mathbb{Z}$; esto es, los elementos $(a^m)^n$, $n \in \mathbb{Z}$, son todos distintos y H será un grupo infinito.

Si consideramos la correspondencia entre $\mathbb{N}$ y el conjunto de los subgrupos de G dada por $m \mapsto \langle a^m \rangle$, esta es una biyección, ya que para cada subgrupo H de G existe un entero positivo m (el menor entero positivo tal que $a^m \in H$) y a m le corresponde precisamente $\langle a^m \rangle$, con lo cual queda demostrado (1).

2) Sea ahora $G = \langle a \rangle$ finito de orden n , esto es, $G = \{a^0, a^1, \dots, a^{n-1}\}$. De nuevo, si H es subgrupo de G , entonces $H = \langle a^m \rangle$, donde m es el menor entero positivo tal que $a^m \in H$.

Probemos que $n = mq$.

Sea $n = mq + r$ con $0 < r < m$ o $r = 0$, de donde

$$a^n = a^{mq+r} \text{ y } a^n \cdot a^{-mq} = a^r.$$

Luego, como $a^n = e$, tenemos

$$a^r = a^{-mq} \in H.$$

Por la forma en que tomamos m se infiere que tenemos solo la opción $r = 0$. De esto resulta que $n = mq$ y $H = \{e, a^m, a^{2m}, \dots, a^{(q-1)m}\}$, ya que $a^{mq} = a^n = e$.

Podemos decir, pues, que $o(\langle a^m \rangle) = q = \frac{n}{m}$.

Obsérvese que en la demostración de 2) se ha repetido la técnica usada en la demostración de que los subgrupos de un grupo cíclico son cíclicos (teorema 3.13.1). El resultado anterior demuestra de nuevo el teorema de Lagrange para el caso particular de los grupos cíclicos.

Sea ahora, sea q divisor del orden n de $G = \langle a \rangle$. Entonces $o(\langle a^n \rangle) = q$, con lo que se demuestra que para cada divisor q del orden de G , existe un subgrupo de G de dicho orden, lo que sería en este caso el recíproco del teorema de Lagrange.

En otras palabras, en el caso finito la correspondencia entre el conjunto de los divisores positivos del orden de G y los subgrupos de G dada por $q \mapsto \langle a^{n/q} \rangle$ es biyectiva. De ahí que podamos afirmar que para cada divisor positivo q del orden de G , existe un y solo un subgrupo de G de dicho orden.

Queremos destacar que en el caso finito (2) el resultado anterior nos dice que si $q_1, q_2, \dots, q_k$ son los divisores de n , entonces un grupo cíclico $G = \langle a \rangle$ de orden n tiene exactamente k subgrupos: $\langle a^{q_1} \rangle, \langle a^{q_2} \rangle, \dots, \langle a^{q_k} \rangle$.

En el caso infinito puede obtenerse un resultado similar al anterior. En efecto, si en el caso finito consideramos en lugar de los grupos $\langle a_i \rangle$ los cocientes $G / \langle a_i \rangle$, $1 \leq i \leq k$, podemos enunciar el resultado anterior diciendo que en un grupo cíclico finito G existe un subgrupo para cada índice que sea un divisor del orden de G , lo que da una idea de cómo generalizar el resultado al caso infinito.

Exactamente en el caso infinito puede demostrarse que existe un y solo un subgrupo finito de índice n para cada $n > 0$ dado, lo cual veremos más adelante.

Método de búsqueda de los generadores y subgrupos de un grupo cíclico finito

Hemos visto una serie de resultados teóricos acerca de los subgrupos de un grupo cíclico. Nos proponemos ahora tratar la solución de un problema más práctico: la búsqueda de los subgrupos de un grupo cíclico finito.

Habíamos probado que si $G = \langle a \rangle$ y $H = \langle a^m \rangle$, entonces $o(H) = \frac{n}{m}$ donde m es el menor entero positivo tal que $a^m \in H$ y n es el orden de G .

Ahora bien, en la práctica el problema de la búsqueda de los subgrupos se nos presenta de otra forma, pues aunque sabemos que para cada q divisor del orden n de G existe exactamente un subgrupo de orden q , muchas veces es necesario saber cuál es el orden del subgrupo generado por un elemento dado, lo que nos permitirá saber cuáles elementos generan el mismo subgrupo e incluso cuáles son generadores del grupo completo.

Para aclarar las ideas tomemos el ejemplo siguiente:

Ejemplo

Consideremos el grupo cíclico $\mathbb{Z}_6$ de orden 6. Sabemos que $\mathbb{Z}_6 = \langle 1 \rangle$ y que $\mathbb{Z}_6$ tiene como subgrupos los triviales, un subgrupo H_1 de orden 2 y otro H_2 de orden 3.

Es fácil determinar que H_1 y H_2 están dados por

$$H_1 = \langle 3 \rangle = \{0, 3\} \text{ y } H_2 = \langle 2 \rangle = \{0, 2, 4\}.$$

Ahora bien, con la información que tenemos hasta el momento, ¿qué podemos afirmar sobre los subgrupos $\langle 4 \rangle$ y $\langle 5 \rangle$? Sabemos que siempre se pueden calcular dichos subgrupos, con lo cual obtendríamos que $\langle 4 \rangle = \langle 2 \rangle$ y que $\langle 5 \rangle = \mathbb{Z}_6$, pero esto resulta engorroso cuando el orden del grupo cíclico es mayor, por ejemplo, en $\mathbb{Z}_{12}$ o $\mathbb{Z}_{18}$. Es por ello que se impone la búsqueda de un criterio práctico que nos permita resolver el problema planteado. Este criterio está dado por el siguiente teorema:

TEOREMA 3.13.4

Sea $G = \langle a \rangle$ un grupo cíclico de n elementos y sea $b \in G$, $b = a^s$. Entonces

$H = \langle b \rangle$ tiene $\frac{n}{d}$ elementos donde $d = \text{mcd}(n, s)$.

Demostración

Sabemos que el número de elementos de H está determinado por el menor entero positivo m tal que $a^m = e$. El problema consiste, pues, en hallar m , esto es, el menor entero positivo tal que $b^m = (a^s)^m = e$.

Como de $a^m = e$ se deduce que n debe dividir a a^m , debemos determinar el m mínimo tal que n divida a sm . Este número está dado precisamente por $m = \frac{n}{d}$, ya que al dividir por el máximo común divisor de n y s aseguramos los factores mínimos para que ms sea divisible por n .

Por esta vía encontramos de nuevo el teorema de Lagrange para grupos cíclicos.

La búsqueda de los generadores de un grupo cíclico se resuelve ahora muy fácilmente gracias al siguiente corolario:

COROLARIO

Si G es un grupo cíclico finito de orden n y $G = \langle a \rangle$, entonces los otros generadores de G son los elementos de la forma a^r donde $\text{mcd}(r, n) = 1$.

La demostración es obvia a partir del teorema.

Existe una fórmula que da la cantidad de números primos relativos con un número entero n y menores que n . Dicha cantidad está dada por $\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$ donde p_i son los factores primos de n . A $\varphi(n)$ se le llama *función de Euler*.

Con los datos anteriores nos podemos proponer la búsqueda de todos los subgrupos de un grupo cíclico de orden finito n .

Ejemplo

Busquemos los subgrupos de $\mathbb{Z}_{12}$ y hagamos el correspondiente diagrama de los subgrupos.

Hallemos primero los generadores de $\mathbb{Z}_{12}$. Estos serán los primos relativos con 12, o sea, 1, 5, 7, 11, ya que todo elemento de $\mathbb{Z}_{12}$ se escribe como $n=1$. Esto quiere decir que

$$\mathbb{Z}_{12} = \langle 1 \rangle = \langle 5 \rangle = \langle 7 \rangle = \langle 11 \rangle.$$

Veamos el subgrupo $\langle 2 \rangle$.

$\text{mcd}(2, 12) = 2$; luego $\langle 2 \rangle$ tiene seis elementos:

$$\langle 2 \rangle = \{0, 2, 4, 6, 8, 10\}$$

Ahora busquemos los generadores de $\langle 2 \rangle$. Estos serán los elementos de la forma 2^r con r primo relativo con 6, esto es, $2^5 = 10$; así que obtenemos $\langle 2 \rangle = \langle 10 \rangle$.

Busquemos los otros subgrupos de $\langle 2 \rangle$.

$\langle 4 \rangle$ tiene tres elementos ya que $\text{mcd}(2, 6) = 2$ y $\frac{6}{2} = 3$.

$$\langle 4 \rangle = \{0, 4, 8\} \text{ y } 8 = 4^2 \text{ y } \text{mcd}(2, 3) = 1.$$

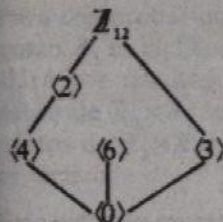
Luego $\langle 4 \rangle = \langle 8 \rangle$.

De los subgrupos de $\langle 2 \rangle$ falta el $\langle 6 \rangle$. $6 = 2^3$, $\text{mcd}(3, 6) = 3$; luego $\langle 6 \rangle$ tiene dos elementos y $\langle 6 \rangle = \{0, 6\}$.

Ya sabemos la forma de $\langle 0 \rangle$, $\langle 1 \rangle$, $\langle 2 \rangle$, $\langle 4 \rangle$, $\langle 5 \rangle$, $\langle 6 \rangle$, $\langle 7 \rangle$, $\langle 8 \rangle$, $\langle 10 \rangle$, $\langle 11 \rangle$. Quedan $\langle 3 \rangle$ y $\langle 9 \rangle$.

Sabemos que $\langle 3 \rangle$ es de orden 4; $\langle 3 \rangle = \{0, 3, 6, 9\}$. En cuanto a $\langle 9 \rangle$, $9 = 3^2$ y $\text{mcd}(3, 4) = 1$; luego $\langle 9 \rangle = \langle 3 \rangle$.

Hagamos a modo de resumen el diagrama de los subgrupos de $\mathbb{Z}_{12}$:



La técnica descrita en el ejemplo anterior es la usada para la búsqueda de todos los subgrupos de un grupo cíclico. El estudiante puede también buscar directamente todos los subgrupos, pero esto sería muy tedioso. Además, la construcción del diagrama resulta de modo natural del método seguido arriba.

A modo de recordatorio señalamos que lo estudiado en el primer curso de Álgebra sobre las raíces n -simas de la unidad en $\mathbb{C}$ y, en particular sobre las raíces primitivas, no es más que un caso particular de los resultados que acabamos de estudiar.

Llamamos raíz primitiva n -sima de la unidad a una raíz $\xi_k = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}$ tal que $\text{mcd}(k, n) = 1$, pero como el grupo $\mathbb{C}_n$ de las raíces de la unidad es $\mathbb{C}_n = \langle \xi_1 \rangle$ y $\xi_k = (\xi_1)^k$, la definición de raíz primitiva responde completamente al corolario anterior. En otras palabras, una raíz primitiva no es más que un generador del grupo de las raíces n -simas de la unidad, como señalamos en los ejemplos de grupos cíclicos.

En algunos casos sencillos el método anterior permite determinar completamente los subgrupos de un grupo finito.

Ejemplo

Busquemos los subgrupos de $\mathbb{Z}_2 \times \mathbb{Z}_4$.

Según el teorema de Lagrange, los subgrupos propios de $\mathbb{Z}_2 \times \mathbb{Z}_4$ pueden ser de orden 2 o 4. Luego, en este caso las posibilidades son: subgrupos cíclicos de órdenes 2 y 4 y subgrupos isomorfos a V_4 .

$$\mathbb{Z}_2 \times \mathbb{Z}_4 = \{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (1,3)\}$$

Comencemos buscando los subgrupos cíclicos, por ejemplo, de

$$\langle (0,1) \rangle = \{(0,1), (0,2), (0,3), (0,0)\}.$$

Aplicando a $\langle (0,1) \rangle$ el método de búsqueda de generadores y subgrupos, obtenemos

$$\langle (0,1) \rangle = \langle (0,3) \rangle.$$

y son subgrupos de $\langle 0,1 \rangle$:

$$\langle 0,2 \rangle = \{ (0,2), (0,0) \}$$

$$\langle 1,0 \rangle = \{ (1,0), (0,0) \}.$$

Busquemos $\langle 1,1 \rangle = \{ (1,1), (0,2), (1,3), (0,0) \}$.

Aplicando de nuevo el método, obtenemos

$$\langle 1,1 \rangle = \langle 1,3 \rangle,$$

cuyo único subgrupo es $\langle 0,2 \rangle$.

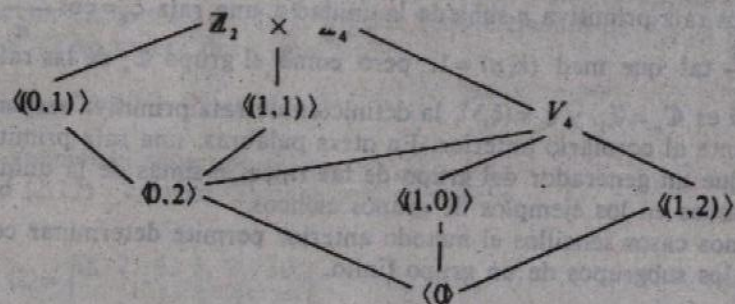
Por último, $\langle 1,2 \rangle = \{ (1,2), (0,0) \}$.

Así quedan analizados todos los subgrupos cíclicos de $\mathbb{Z}_2 \times \mathbb{Z}_4$.

En cuanto a un posible subgrupo de orden 4 isomorfo a V_4 , su estructura está caracterizada, como sabemos, por poseer un neutro y tres elementos que sean su propio simétrico, esto es, tres elementos de orden 2, con lo que obtenemos

$$V_4 = \{ (0,0), (0,2), (1,0), (1,2) \}.$$

El diagrama de los subgrupos de $\mathbb{Z}_2 \times \mathbb{Z}_4$ es el siguiente:



Con la búsqueda de todos los subgrupos cíclicos de $\mathbb{Z}_2 \times \mathbb{Z}_4$ probamos que este no es cíclico (ya que ningún subgrupo cíclico de $\mathbb{Z}_2 \times \mathbb{Z}_4$ coincide con el grupo completo).

El método anterior de búsqueda de subgrupos solo es válido, con los conocimientos que tenemos hasta el momento, para ciertos grupos muy sencillos. Al estudiar los grupos abelianos finitos, podemos ampliar la clase de grupos a la cual es posible aplicar dicho método; sin embargo, los resultados que hallaremos inmediatamente con respecto a grupos del tipo $\mathbb{Z}_n \times \mathbb{Z}_m$ nos ayudarán en ese sentido.

Veamos primero otro ejemplo:

Ejemplo

Tomemos $\mathbb{Z}_3 \times \mathbb{Z}_5$.

$$\mathbb{Z}_3 \times \mathbb{Z}_5 = \{ (0,0), (0,1), (1,0), (1,1), (0,2), (1,2), (2,2), (2,0), (2,1) \}$$

En este caso $\mathbb{Z}_3 \times \mathbb{Z}_5$ no es cíclico y, por tanto, aunque tiene nueve elementos, no es isomorfo a $\mathbb{Z}_9$.

Para comprobar que $\mathbb{Z}_3 \times \mathbb{Z}_5$ no es cíclico usaremos el siguiente argumento: como $\mathbb{Z}_5$ es cíclico, cada elemento de $\mathbb{Z}_5$, sumado tres veces consigo mismo dará por resultado el neutro, y el subgrupo cíclico generado por cada elemento de $\mathbb{Z}_3 \times \mathbb{Z}_5$ será de orden 3 y no de orden 9 que es el número de elementos de $\mathbb{Z}_3 \times \mathbb{Z}_5$. Así queda probada nuestra afirmación de que $\mathbb{Z}_3 \times \mathbb{Z}_5$ no es cíclico.

Este ejemplo muestra cómo obtener un nuevo grupo de orden 9 mediante el producto cartesiano de dos grupos conocidos y plantea el problema siguiente: *¿Cuándo un producto de grupos cíclicos es a su vez cíclico?*

La respuesta está dada en el siguiente teorema, cuya demostración se basa en las técnicas expuestas en los ejemplos anteriores.

TEOREMA 3.13.5

Si m y n son primos relativos, entonces $\mathbb{Z}_{mn}$ es isomorfo a $\mathbb{Z}_m \times \mathbb{Z}_n$.

Demostración

Tomemos el subgrupo cíclico de $\mathbb{Z}_m \times \mathbb{Z}_n$ generado por $(1,1)$ y probemos que $\langle (1,1) \rangle = \mathbb{Z}_m \times \mathbb{Z}_n$. Como $\mathbb{Z}_m \times \mathbb{Z}_n$ tiene mn elementos, quedaría probada nuestra afirmación. Veamos, pues, que $\langle (1,1) \rangle = \mathbb{Z}_m \times \mathbb{Z}_n$. Para que $\langle (1,1) \rangle = \mathbb{Z}_m \times \mathbb{Z}_n$ es necesario que $\langle (1,1) \rangle$ tenga mn elementos, lo cual ocurrirá si mn es el menor entero positivo s tal que

$$(1,1)^s = (0,0), \quad (1,1)^s = (0,0) \Rightarrow s = k_1 n \text{ y } s = k_2 m.$$

Y como $(n,m)=1$, el menor múltiplo de m y n es mn , con lo que queda probado que el s buscado es mn .

El siguiente corolario se puede demostrar por el método de inducción.

COROLARIO:

Si los números m_i , $i=1, \dots, n$, son primos relativos dos a dos, esto es, $\text{med}(m_i, m_j)=1$ para $i \neq j$, entonces $\prod_{i=1}^n \mathbb{Z}_{m_i}$ es cíclico y es isomorfo a $\mathbb{Z}_{m_1 m_2 \dots m_n}$.

El teorema 3.13.5 y este corolario son muy útiles, ya que permiten resolver problemas como el siguiente:

$\mathbb{Z}_{45}$ es isomorfo a $\mathbb{Z}_9 \times \mathbb{Z}_5$, ya que $45=3^2 \cdot 5$.

En general, dado un número entero positivo n , podemos descomponerlo en factores primos:

$$n = (p_1)^{r_1} (p_2)^{r_2} \dots (p_k)^{r_k}.$$

Entonces $\mathbb{Z}_n$ será isomorfo a $\mathbb{Z}_{p_1^{r_1}} \times \dots \times \mathbb{Z}_{p_k^{r_k}}$.

Así, $\mathbb{Z}_{60} = \mathbb{Z}_{15} \times \mathbb{Z}_4 = \mathbb{Z}_{12} \times \mathbb{Z}_5 = \mathbb{Z}_{20} \times \mathbb{Z}_3$.

Existe otro resultado que se demuestra a partir de la técnica descrita en el teorema 3.13.5 y en los ejemplos dados. Este resultado consiste en lo siguiente:

TEOREMA 3.13.6

Si tenemos un grupo $G = \prod_{i=1}^n G_i$, producto directo externo de los grupos G_i , y $(a_1, a_2, \dots, a_n) \in G$ donde el orden de a_i es r_i , entonces el orden de $(a_1, a_2, \dots, a_n)$ es igual al mínimo común múltiplo de los r_i .

Demostración

Podemos usar un argumento similar al del teorema 3.13.5. El orden de un elemento $(a_1, a_2, \dots, a_n)$ será el menor entero positivo s tal que

$$(a_1, a_2, \dots, a_n)^s = (0, 0, \dots, 0).$$

s está dado por el menor múltiplo común de $r_1, r_2, \dots, r_n$, esto es, $\text{mcm}(r_1, r_2, \dots, r_n)$.

Cocientes y homomorfismos de grupos cíclicos

Por último, estudiemos los cocientes y homomorfismos de grupos cíclicos.

Es bastante natural que al pasar al cociente se mantenga la estructura de grupo cíclico. Observemos que en los grupos cíclicos, por ser abelianos, podemos hallar el cociente por cualquier subgrupo, pues en un grupo abeliano todo subgrupo es invariante.

PROPOSICIÓN 3.13.4

El grupo cociente de un grupo cíclico es cíclico.

Demostración

Sean $G = \langle a \rangle$ y H un subgrupo de G . Probaremos que aH genera a G/H . En efecto, si $x \in G$, entonces $x = a^n$ para algún $n \in \mathbb{Z}$. Si tomamos xH , $xH = a^n H = (aH)^n$. De aquí que $G/H = \{(aH)^n : n \in \mathbb{Z}\}$; luego $G/H = \langle aH \rangle$.

Probemos ahora el siguiente resultado referente a los subgrupos de un grupo cíclico infinito y al cual ya hicimos referencia anteriormente.

TEOREMA 3.13.7

En un grupo cíclico infinito G para cada entero positivo n existe un único subgrupo de G de índice n .

Demostración

Sea $G = \langle a \rangle$, donde G es un grupo cíclico infinito. Llamemos $H_n = \langle a^n \rangle$. Entonces $H_n = \{a^{nr} : r \in \mathbb{Z}\}$. De aquí que los cosetos $H_n, aH_n, \dots, a^{n-1}H_n$ sean todos distintos y que sean los cosetos de H_n en G . Luego H_n es de índice n .

Recíprocamente, sea H un subgrupo de índice n . Sabemos que H está generado por la menor potencia positiva a^m tal que $a^m \in H$ y esto significa que $H = H_m$. Pero H_m es de índice m , de aquí $m = n$ y $H = H_n$.

Como ya habíamos señalado el resultado anterior generaliza mediante el concepto de *índice*, la propiedad de los grupos cíclicos finitos de que existe un subgrupo para cada orden divisor del orden del grupo.

En los homomorfismos de grupos cíclicos, también la estructura de estos permite obtener gran información, hasta el punto de poder encontrar todos los homomorfismos de un grupo cíclico.

TEOREMA 3.13.8

Sean $G = \langle a \rangle$ un grupo cíclico y φ un homomorfismo con dominio G .

- (1) Entonces $\varphi(G)$ es cíclico y si G es finito, $\alpha(\varphi(G))/\alpha(G)$ (o sea el orden de G es un múltiplo del orden de $\varphi(G)$).
- (2) Si H es un grupo cíclico tal que $\alpha(H)/\alpha(G)$, entonces existe un homomorfismo de G sobre H .
- (3) Si G es infinito, existe un homomorfismo de G sobre cualquier grupo cíclico.

Demostración

- (1) Si $G = \langle a \rangle = \{a^n : n \in \mathbb{Z}\}$, entonces

$$\varphi(G) = \{\varphi(a^n) = (\varphi(a))^n : n \in \mathbb{Z}\} = \langle \varphi(a) \rangle,$$

de donde tenemos que $\varphi(G)$ es cíclico.

Que $\alpha(\varphi(G))/\alpha(G)$ resulta como consecuencia inmediata del teorema fundamental de homomorfismo y el teorema de Lagrange si G es finito, ya que

$$\alpha(\varphi(G)) = \alpha(G/\text{Ker } \varphi) = \alpha(G) : \alpha(\text{Ker } \varphi).$$

- (2) Supongamos ahora que H es cíclico, $H = \langle b \rangle$ y que el orden m de H divide al orden n de G . Esto es, $n = mq$. Consideremos $N = \{e, a^m, \dots, (a^m)^{q-1}\} = \langle a^m \rangle$. Entonces $\alpha(N) = q$ y N es de índice m . Como G es cíclico, tenemos que G/N tiene sentido y es cíclico de orden m , de donde $G/N = H$, por lo que podemos deducir que H es la imagen de G por un homomorfismo cuyo núcleo es N .

- (3) Por último, si $G = \langle a \rangle$ es cíclico infinito, sabemos que si $G_n = \langle a^n \rangle$, entonces el índice de G_n en G es n , esto es, $\alpha(G/G_n) = n$ y G/G_n es la imagen homomorfa de G por el homomorfismo canónico. Luego cualquier grupo cíclico de orden n es imagen homomorfa de G , pues podemos construir G/G_n para todo $n \geq 0$.

Por otra parte, como todos los grupos cíclicos infinitos son isomorfos entre sí, es obvio que G tiene como imagen homomorfa a cualquier grupo cíclico infinito.

Ejercicios

1. ¿Cuáles de los siguientes grupos son cíclicos? En caso de serlo, busque todos los generadores del grupo; en caso de no serlo, justifique su respuesta.

a) $(\mathbb{Q}, +)$ b) $(6^n : n \in \mathbb{Z}, \cdot)$ c) $(\{a + b\sqrt{2} : a, b \in \mathbb{Z}\}, +)$ d) D_4 .

2. Encuentre el número de generadores de los grupos de orden 6, 8, 12 y 60.
3. Pruebe que un grupo que tiene solo un número finito de subgrupos debe ser finito.
4. Encuentre el número de elementos en cada uno de los subgrupos cíclicos siguientes:
- a) $\langle 25 \rangle$ en $\mathbb{Z}_{30}$
 b) $\langle 30 \rangle$ en $\mathbb{Z}_{42}$
 c) $\langle i \rangle$ en $(\mathbb{C}^*, \cdot)$ y $\langle \frac{1+i}{2} \rangle$ en $\langle i \rangle$.
5. Para cada uno de los siguientes grupos halle todos los subgrupos y haga el diagrama correspondiente.
- a) $\mathbb{Z}_{24}$ b) $\mathbb{Z}_{12}$ c) $\mathbb{Z}_{18}$ d) $\mathbb{Z}_{60}$ e) $\mathbb{Z}_3 \times \mathbb{Z}_4$ f) $\mathbb{Z}_2 \times \mathbb{Z}_4$
6. Pruebe que en un grupo cíclico finito de orden n , la ecuación $x^m = e$ tiene exactamente m soluciones x en G para cada entero positivo m que divida a n . ¿Cuál es la situación si $1 < m < n$ y m no divide a n ?
7. Pruebe que si G es un grupo finito y $a \in G$, entonces $o(a)/o(G)$
8. Pruebe que en un grupo finito G de orden n y elemento neutro e , se tiene $a^n = e \quad \forall a \in G$.
9. Si G es un grupo, $a \in G$ y $a^m = e$, pruebe que $o(a)/m$.
10. Utilice el hecho de que un grupo cíclico finito de orden n tiene exactamente un subgrupo de cada orden d que divida a n para probar que
- $$n = \sum_{d|n} \phi(d),$$
- donde $\phi(d)$ es la función de Euler que da el número de enteros positivos menores que d y primos relativos con d .
11. Sea G un grupo finito de orden n . Pruebe que si para cada entero positivo m el número de soluciones x de la ecuación $x^m = e$ es a lo sumo m , entonces G es cíclico.
- Sugerencia:* Use el ejercicio 8 para probar que G debe contener un elemento de orden n .
12. Diga cuantos homomorfismos hay de
- a) $\mathbb{Z}$ sobre $\mathbb{Z}$ f) $\mathbb{Z}_{12}$ sobre $\mathbb{Z}_3$
 b) $\mathbb{Z}$ en $\mathbb{Z}_2$ g) $\mathbb{Z}_{12}$ en $\mathbb{Z}_6$
 c) $\mathbb{Z}$ sobre $\mathbb{Z}_2$ h) $\mathbb{Z}_{12}$ sobre $\mathbb{Z}_4$
 d) $\mathbb{Z}$ en $\mathbb{Z}_4$ i) $\mathbb{Z}_{12}$ en $\mathbb{Z}_{14}$
 e) $\mathbb{Z}$ sobre $\mathbb{Z}_4$ j) $\mathbb{Z}_{12}$ en $\mathbb{Z}_{16}$.

13. Pruebe que si n es un entero positivo y a es primo relativo con n , entonces $a^{o(n)} = 1 \pmod{n} \iff (n \text{ divide a } a^{o(n)} - 1)$.

Sugerencia: Pruebe que los números enteros positivos primos con n y menores que n forman un grupo respecto al producto definido en $\mathbb{Z}/n\mathbb{Z}$ por $(r+n\mathbb{Z})(s+n\mathbb{Z}) = rs+n\mathbb{Z}$, y aplique a dicho grupo el ejercicio 8.

14. Pruebe que si p es un número primo y a es un entero cualquiera, entonces $a^p \equiv a \pmod{p}$, o sea, $(a^p - a)$ es divisible por p .

Sugerencia: Si a es primo relativo con p , utilice el ejercicio 13. Si a no es primo relativo con p , como p es primo, use que p debe dividir a a .

15. a) Pruebe que si G es un grupo finito de orden primo p , entonces G es cíclico.

b) Pruebe que un grupo que tiene al menos dos elementos y no posee subgrupos propios es finito y de orden primo.

16. Si N es un subgrupo normal de un grupo G y $a \in G$ es de orden n , pruebe que el orden de aN en G/N es un divisor de n .

17. Sean a y b elementos de un grupo abeliano G tales que $o(a) = n$ y $o(b) = m$ son primos relativos. Pruebe que

a) $\langle a \rangle \cap \langle b \rangle = \{e\}$ (e elemento neutro de G).

b) $\langle ab \rangle$ tiene orden mn .

Sugerencia: Considere para a) un elemento d de $\langle a \rangle \cap \langle b \rangle$ y pruebe que $d = e$. Para b) pruebe que $o(ab)/mn$ y que $mn/o(ab)$.

18. Pruebe que si G es un grupo abeliano finito, G contiene un elemento x cuyo orden es divisible por el orden de cada elemento de G .

Sugerencia: Pruebe que es posible construir un elemento cuyo orden sea el mínimo común múltiplo de todos los elementos de G , usando el ejercicio 17.

19. Pruebe que si G es un grupo abeliano finito, G es cíclico si y solo si $o(G)$ es el menor entero positivo n tal que $a^n = e$ (e neutro de G) para todo a de G .

Sugerencia: Use el ejercicio 18.

20. Pruebe que si G es cíclico de orden n y p divide a n , entonces existe un homomorfismo de G sobre un grupo cíclico de orden p . ¿Cuál es el núcleo de este homomorfismo?

Sugerencia: Considere $G = \langle a \rangle$ de orden n y $H = \langle g \rangle$ de orden p , y defina el homomorfismo de la forma más natural posible dado que $G = \{a^n : n \in \mathbb{Z}\}$ y $H = \{g^m : m \in \mathbb{Z}\}$.

21. Sea $\phi: G \rightarrow G'$ un homomorfismo de grupos sobreyectivo. Sea G' cíclico de orden 10. Pruebe que G tiene subgrupos normales de índices 2, 5 y 10.
22. Sean G un grupo y N un subgrupo normal de G . Suponga que G/N es cíclico infinito. Pruebe que G tiene un subgrupo normal de índice n para cada entero positivo n .
23. Sean N y M subgrupos normales de G tales que $N \supset M$. Suponga que G/N es cíclico y $o(N/M) = 2$. Pruebe que G/M es abeliano.

3.14 Nociones sobre la teoría de grupos abelianos

Al estudiar los grupos cíclicos vimos cómo el hecho de conocer la forma de los elementos permitía obtener gran cantidad de información sobre el grupo. Ahora nos proponemos ampliar un poco nuestro campo de estudio a grupos más complejos que los cíclicos.

Nos limitaremos, sin embargo, a cierta clase de grupos abelianos, ya que la suposición de conmutatividad simplifica considerablemente la estructura del grupo, y para la clase de grupos abelianos que estudiaremos (los grupos abelianos finitamente generados), permite demostrar que dichos grupos se expresan mediante grupos cíclicos, cuya estructura ya conocemos.

Será necesario introducir nuevos conceptos como el de *producto (suma) directa de grupos* y utilizar otros ya conocidos como el de *orden de un elemento de un grupo*.

Nos proponemos estudiar una clase de grupos que constituya una generalización natural de los grupos cíclicos.

En un grupo cíclico existe un elemento del grupo que "genera" el grupo completo. Por tanto, lo más natural parece ser considerar que el grupo pueda estar generado por dos, tres o un número finito de elementos, es decir, sustituir el generador a de un grupo cíclico por un conjunto "generador".

Recordemos que en el caso de los grupos cíclicos comenzamos por estudiar el subgrupo cíclico generado por un elemento de un grupo, y después vimos que en algunos casos el grupo completo coincidía con dicho subgrupo. Además, notamos que el subgrupo cíclico generado por un elemento x de G es el menor subgrupo de G que contiene a x , esto es, dicho subgrupo da la forma mínima de encontrar un grupo a partir del elemento x .

Vamos a extender este razonamiento al caso de más de un generador. Consideraremos el problema siguiente:

Subgrupo generado por una familia de elementos

Sean G un grupo y $(a_i)_{i \in I}$ una familia de elementos de G . Al igual que en el caso de los subgrupos cíclicos, existirá un subgrupo mínimo de G que

contenga a $S = \{a_i : i \in I\}$, en el sentido de que cualquier otro subgrupo de G que contenga a S , contendrá a dicho subgrupo mínimo.

Hemos señalado anteriormente que la intersección de una familia no vacía de subgrupos de G es de nuevo un subgrupo de G , y obviamente si consideramos la intersección I de la familia de los subgrupos de G que contienen a S , dicha familia es no vacía (G pertenece a ella) y es el menor subgrupo de G que contiene a S , ya que:

- (1) I es subgrupo de G .
- (2) $S = \{a_i : i \in I\} \subset I$
- (3) Si G' es un subgrupo de G y $S \subset G'$, entonces $I \subset G'$.

Una vez demostrada la existencia del menor subgrupo de G que contiene a S , podemos dar la definición siguiente:

DEFINICIÓN 3.14.1

Se llama *subgrupo de G generado por S* , y se denota por $\langle S \rangle$ al menor subgrupo de G que contiene a S , el cual está dado por la intersección de todos los subgrupos de G que contienen a S .

La anterior caracterización sirve para fines teóricos, pero desde el punto de vista del trabajo con el subgrupo generado, es obvio que no es adecuada. Por esto es que caracterizamos el subgrupo por un subconjunto que cumpla las propiedades (1), (2) y (3) dadas anteriormente y ahora trataremos de darle una forma más explícita a dicho subgrupo. Estudiemos el teorema siguiente:

TEOREMA 3.14.1

Sea G un grupo y $S \subset G$. Entonces

$$\langle S \rangle = \{x_1 \cdot x_2 \cdot \dots \cdot x_n : n \in \mathbb{N}\},$$

donde $x_i \in S$ o $x_i^{-1} \in S$, $1 \leq i \leq n$.

En otras palabras, $\langle S \rangle$ está formado por los productos finitos de elementos de S o de inversos de elementos de S . (Recordemos que en el caso de subespacio vectorial, el subespacio generado estaba formado por las combinaciones lineales finitas de elementos de S , y observemos la analogía de dicho resultado con el presente.)

Demostración

Llamemos $\pi = \{x_1 \cdot x_2 \cdot \dots \cdot x_n : n \in \mathbb{N}, x_i \in S \text{ o } x_i^{-1} \in S\}$ y probemos que $\langle S \rangle = \pi$.

En efecto, si $x_1 \cdot x_2 \cdot \dots \cdot x_n \in \pi$ y $y_1 \cdot y_2 \cdot \dots \cdot y_m \in \pi$, entonces

$$(x_1 \cdot x_2 \cdot \dots \cdot x_n) (y_1 \cdot y_2 \cdot \dots \cdot y_m)^{-1} = (x_1 \cdot \dots \cdot x_n) (y_m^{-1} \cdot \dots \cdot y_1^{-1}) \in \pi.$$

Luego π es un subgrupo de G y se cumple la propiedad (1) de la caracterización de subgrupo generado por un subconjunto.

Por otra parte, $S \subset \pi$ ya que para cada $x_i \in S$, x_i es un producto finito de elementos de S .

Por último, si G' es un subgrupo de G tal que $S \subset G'$, por la definición de subgrupo, G' debe contener a los elementos de π . Luego se cumple $\pi \subset G'$ para cada G' tal que $S \subset G'$. De aquí que $\langle S \rangle = \{x_1 \cdot x_2 \cdots x_n : n \in \mathbb{N} \text{ y } x_i \in S \text{ o } x_i^{-1} \in S\}$.

Ejemplos

1. En $\mathbb{Z}_6$, $\langle \{4, 6, 2\} \rangle = \{0, 2, 4, 6\} = \langle 4 \rangle$.
2. En $\mathbb{Z}_4 \times \mathbb{Z}_4$, $\langle \{(0, 1), (2, 2)\} \rangle = \{(0, 0), (2, 2), (0, 1), (0, 2), (0, 3), (2, 0), (2, 1), (2, 3)\}$.
3. En $\mathbb{Z} \times \mathbb{Z}_2$, $\langle \{(1, 0), (0, 1)\} \rangle = \mathbb{Z} \times \mathbb{Z}_2$.

Volvemos a encontrarnos en $\mathbb{Z} \times \mathbb{Z}_2$ con la situación vista en grupos cíclicos, de que un grupo coincide con el subgrupo generado por uno de sus subconjuntos.

Nos interesan particularmente los grupos generados por un subconjunto con un número finito de elementos, esto es, los grupos que poseen un conjunto generador finito.

DEFINICIÓN 3.14.2

Se dice que un grupo es un *grupo finitamente generado* si y solo si posee un sistema de generadores finito, esto es, si existe $S \subset G$, S finito, tal que $\langle S \rangle = G$.

Cuando el grupo posee un sistema de generadores finito, podemos plantearnos la búsqueda de un sistema generador minimal, es decir, un sistema H tal que $\langle H \rangle = G$, y si $H' \subset H$ entonces $\langle H' \rangle \neq G$. Esta construcción es también similar a la que hacemos en los espacios vectoriales de dimensión finita, pues precisamente la búsqueda del sistema generador minimal se realiza eliminando los elementos del sistema generador que se expresen como productos finitos de los restantes elementos del sistema y de sus inversos. Análogamente en los espacios vectoriales buscamos un sistema generador minimal eliminando los elementos del sistema que sean combinación lineal de los restantes.

Antes de seguir adelante queremos señalar que todo lo dicho con respecto a un subgrupo generado por un subconjunto, es válido para un grupo cualquiera no necesariamente abeliano; pero como ya dijimos, nos limitaremos al estudio del caso abeliano.

La estructura de un grupo abeliano puede ser muy compleja y en su estudio desempeña un importante papel el concepto *orden* de los elementos.

En los grupos cíclicos, que son los grupos que ya conocemos, observamos que en los finitos todo elemento del grupo es de orden finito —esto ocurre en general para los grupos finitos aunque no sean cíclicos—, y en los infi-

nitos el único elemento que es de orden finito es el neutro. Pero si consideramos, por ejemplo, $\mathbb{Z} \times \mathbb{Z}_2$, existen elementos distintos del neutro que son de orden finito, por ejemplo, $o\langle(0, 1)\rangle = 2$, y elementos de orden infinito, el $(1, 0)$ es uno de ellos. Más aún, si consideramos el grupo multiplicativo $(\mathbb{C}^*, \cdot)$, existen elementos de orden finito n para cada n entero positivo dado, a la vez que existen elementos de orden infinito.

Daremos ahora algunas definiciones y resultados que nos permitirán obtener una clasificación de los grupos abelianos.

DEFINICIÓN 3.14.3

Se dice que un grupo G es un *grupo con torsión*, si todo elemento de G es de orden finito, y que un grupo G es *libre de torsión* si el único elemento de orden finito es el neutro. G es mixto si no es con torsión ni libre de torsión.

Ejemplos

1. Todos los grupos finitos son grupos con torsión.
2. $(\mathbb{Z}, +)$ es un grupo libre de torsión.
3. $\mathbb{Z} \times \mathbb{Z}$, no es un grupo con torsión, pues

$$\langle(1, 0)\rangle = \{(n, 0) : n \in \mathbb{Z}\};$$

y tampoco es libre de torsión, pues

$$\langle(0, 1)\rangle = \{(0, 0), (0, 1)\}.$$

Sin embargo, estudiando los elementos de orden finito de un grupo podemos obtener el siguiente resultado:

TEOREMA 3.14.2

En un grupo abeliano G , el conjunto de los elementos de orden finito es un subgrupo T de G llamado el *subgrupo de torsión* de G .

Demostración

En efecto, si a y b están en T es decir, son de orden finito, entonces existen m, n tales que $a^m = b^n = e$.

Probemos que $ab \in T$. Para ello basta tomar $(ab)^{mn} = (a^m)^n (b^n)^m = e$.

Por último, si a es de orden finito, $a^m = e$ para algún m . De aquí

$$(a \cdot a^{-1})^m = e^m = e$$

y

$$(a \cdot a^{-1})^m = a^m \cdot (a^{-1})^m = (a^{-1})^m = e,$$

con lo cual queda probado que a^{-1} es de orden m .

Está claro que el subgrupo de torsión de un grupo es un grupo con torsión.

Ejemplo

El subgrupo de torsión de $\mathbb{Z} \times \mathbb{Z}_2 = \{(0,0), (0,1)\} \cong \mathbb{Z}_2$.
 En $\mathbb{Z} \times \mathbb{Z}_2$ todo elemento (n,m) puede ser escrito en forma única como $(n,m) = (n,0) + (0,m)$ donde $n \in \mathbb{Z}$, $m \in \mathbb{Z}_2$, lo que puede ser verificado muy fácilmente. Los elementos de la forma $(n,0)$, $n \in \mathbb{Z}$, son de orden infinito en $\mathbb{Z} \times \mathbb{Z}_2$ y $\mathbb{Z} \times \{0\}$ constituye un subgrupo libre de torsión de $\mathbb{Z} \times \mathbb{Z}_2$.

Lo anterior nos da la idea de considerar la posibilidad de "separar" de forma adecuada todo grupo mixto en un grupo libre de torsión y en un grupo con torsión, con lo cual el estudio de los grupos abelianos se reduciría a:

1. Los grupos con torsión.
2. Los grupos libres de torsión.
3. Cómo obtener los grupos mixtos mediante grupos con torsión y grupos libres de torsión.

El tercer punto puede ser planteado de otra manera mediante el resultado siguiente:

TEOREMA 3.14.3

Si G es un grupo abeliano y T su subgrupo de torsión, entonces G/T es libre de torsión.

Demostración

En efecto, supongamos que existan en G/T elementos de orden finito, esto es, gT tal que $(gT)^m = T$ para algún entero positivo m .

$(gT)^m = g^m T = T \Rightarrow g^m \in T \Rightarrow (g^m)^n = e$ para algún entero positivo n . Pero $(g^m)^n = g^{mn} = e \Rightarrow g \in T \Rightarrow gT = T$. Esto significa que el único elemento de G/T que es de orden finito es el neutro T de G/T , y que por tanto, G/T es libre de torsión.

Podemos, pues, plantear el tercer punto de la forma siguiente:

Cómo construir grupos abelianos partiendo de un subgrupo de torsión dado T y de forma tal que A/T sea isomorfo a un grupo libre de torsión dado.

Hemos querido dar una idea del planteamiento del problema para grupos abelianos en general. Sin embargo, el estudio de este problema no corresponde al objetivo de este libro. Nos limitaremos, pues, a dar para el caso de los grupos abelianos finitamente generados, los resultados que prueban que estos pueden construirse a partir de grupos con torsión y grupos libres de torsión. Veremos que, al igual que en el caso de los grupos cíclicos, en el cual obtuvimos una clasificación de estos mediante $\mathbb{Z}$ y los $\mathbb{Z}_m$, ahora podemos clasificar en forma adecuada los grupos abelianos finitamente generados mediante productos de $\mathbb{Z}$ y $\mathbb{Z}_m$.

Muchos de los resultados los daremos sin demostrarlos, ya que aunque el nivel de las demostraciones no difiere del encontrado hasta el momento

en este texto, son laboriosas y alargarían mucho el presente epígrafe, en el cual según señalamos al principio solo queremos dar nociones del tema.

Antes de pasar a dar los resultados que caracterizan los grupos abelianos finitamente generados, es necesario tratar un concepto no conocido: el *producto directo interno*.

El estudiante encontrará en lo que sigue afirmaciones de este tipo: "El grupo G es igual al producto directo de sus subgrupos F y H ". Veamos cómo interpretar esta afirmación, ya que $F \times H$ es un conjunto de pares de elementos de G y no es comparable con G en sentido estricto.

Realmente lo que ocurre es que al definir el producto directo externo de los grupos G_i , o sea, $\prod_{i=1}^n G_i$, podemos considerar los subconjuntos

$$\bar{G}_i = \{(e_1, e_2, \dots, e_{i-1}, a_i, e_{i+1}, \dots, e_n), a_i \in G_i\}$$

(e_i es el elemento neutro de G_i).

Estos subconjuntos son subgrupos de G isomorfos a los G_i , es decir, $\bar{g}_i \cong G_i$ para cada $1 \leq i \leq n$ y $\prod_{i=1}^n \bar{G}_i$ es el *producto directo interno* de los subgrupos $\bar{G}_i$. Aunque ya señalamos que no daremos las demostraciones de muchos de los resultados donde aparece el concepto de *producto directo interno*, formalizaremos dicho concepto debido a su gran importancia.

DEFINICIÓN 3.14.4

Se dice que un grupo G es el *producto directo interno* de sus subgrupos G_i , $i=1, \dots, n$ si y solo si la aplicación

$$\Phi: \prod_{i=1}^n G_i \longrightarrow G \text{ dada por } (g_1, g_2, \dots, g_n) \mapsto g_1 \cdot g_2 \cdot \dots \cdot g_n,$$

es un isomorfo de grupos.

Observemos que en la definición anterior se afirma que la imagen isomorfa de $\bar{G}_i$ es G_i . Esto justifica que, por abuso de lenguaje, podamos usar el término *producto de grupos* sin especificar si es externo o interno.

Veamos ahora los resultados que nos permitirán dar la caracterización de los grupos abelianos finitamente generados.

LEMA 1

Si G es un grupo finitamente generado y T su subgrupo de torsión, entonces G es el producto directo de $T \times F$, donde F es un subgrupo de G libre de torsión.

LEMA 2

Un grupo libre de torsión finitamente generado F es isomorfo a un producto $\underbrace{\mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z}}_{m \text{ veces}}$ para un cierto número de factores m . El número m , llamado

número de Betti de F , es único para cada F .

LEMA 3

Un grupo abeliano finito T es isomorfo a dos diferentes tipos de productos de grupos cíclicos:

(1) T es isomorfo a un producto

$$\mathbb{Z}_{(p_1)^{r_1}} \times \mathbb{Z}_{(p_2)^{r_2}} \times \dots \times \mathbb{Z}_{(p_n)^{r_n}},$$

donde los p_i son primos y no necesariamente distintos. Este producto es único salvo el orden de los factores.

(2) T es isomorfo a un producto

$$\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \dots \times \mathbb{Z}_{m_r},$$

donde m_i divide a m_{i+1} . Los números m_i se llaman *coeficientes de torsión* de T y son únicos.

Las denominaciones de *número de Betti* y *coeficiente de torsión* provienen de la topología algebraica y son invariantes topológicos. El primero se debe al matemático italiano Enrico Betti (1823-1892), el cual observó las propiedades de dichos invariantes.

LEMA 4

Los subgrupos de un grupo finitamente generado son grupos finitamente generados.

Los resultados anteriores conducen a las siguientes observaciones y conclusiones:

El señalamiento de que los primos p_i que aparecen en la proposición (1) del lema 3 pueden no ser distintos, es decir, que combinando en distintas formas potencias de un número primo dado, podemos obtener grupos distintos; por ejemplo, $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$ no es isomorfo a $\mathbb{Z}_9 \times \mathbb{Z}_5$, ya que en $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5$ los elementos tienen a lo sumo orden 15, mientras que en $\mathbb{Z}_9 \times \mathbb{Z}_5 = \mathbb{Z}_{45}$ pueden tener orden 45.

Realmente el lema 3 ofrece la posibilidad de encontrar, salvo isomorfismo, todos los grupos finitos de un orden dado, para lo cual describimos inmediatamente un método.

Dicho método se basa en el lema 3, que dice que todo grupo finito es en definitiva, salvo isomorfismo, igual a un producto directo en grupos cíclicos finitos, y también en los resultados que habíamos obtenido anteriormente sobre el producto directo de grupos cíclicos.

Ejemplo

Si queremos hallar todos los grupos finitos de orden 600, comenzamos por descomponer 600 en factores primos:

$$600 = 2^3 \cdot 3 \cdot 5^2$$

Según la parte (1) del lema 3, tenemos las siguientes posibilidades para grupos de orden 600.

1) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$

2) $\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$

3) $\mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$

4) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_{25}$

5) $\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_{25}$

6) $\mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_{25}$

Recordemos que en el lema 3 se afirma que los grupos de orden 600 son, salvo isomorfismo, los seis que acabamos de hallar, los cuales pueden ser expresados también mediante los coeficientes de torsión. Veamos cómo obtener esta última expresión de los grupos de orden 600.

Tomemos la posibilidad 1):

$$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$$

En este caso ordenamos los coeficientes primos de los subíndices como sigue:

$$\begin{array}{ccc} 2 & 2 & 2 \\ & & 3 \\ & & 5 & 5 \end{array}$$

Esto es, alineamos los correspondientes a cada primo por filas, en orden creciente, y a su vez alineamos los lados derechos de las filas.

Para mostrar esto con más claridad, hagamos el arreglo correspondiente a 2):

$$\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$$

Esto es,

$$\begin{array}{ccc} 2 & & 4 \\ & & 3 \\ & & 5 & 5 \end{array}$$

Aunque no daremos una demostración formal del método, el estudiante puede observar que está basado en el siguiente hecho: mediante el ordenamiento vamos agrupando los factores para formar grupos cíclicos isomorfos a los dados y que cumplan las condiciones de los coeficientes de torsión. Así, en $\mathbb{Z}_2 \times \mathbb{Z}_{10} \times \mathbb{Z}_{30}$ tenemos:

$$\mathbb{Z}_{10} = \mathbb{Z}_2 \times \mathbb{Z}_5, \quad \mathbb{Z}_{30} = \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5$$

y además, 2 divide a 10 y este a 30.

Observemos que si tomamos ahora los productos por columna en el ordenamiento de los coeficientes correspondientes a la posibilidad 1), obtenemos: 2 10 30. Entonces

$$\mathbb{Z}_2 \times \mathbb{Z}_{10} \times \mathbb{Z}_{30} = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5.$$

En la posibilidad (2) obtenemos: 10 60, y de aquí

$$\mathbb{Z}_{10} \times \mathbb{Z}_{60} = \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5.$$

En total, después de aplicado el método tendremos para $\mathbb{Z}_{600}$ las siguientes posibilidades expresadas en la forma (1) y (2) del lema:

$$1) \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5 = \mathbb{Z}_2 \times \mathbb{Z}_{10} \times \mathbb{Z}_{30}$$

$$2) \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5 = \mathbb{Z}_{10} \times \mathbb{Z}_{60}$$

$$3) \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 = \mathbb{Z}_2 \times \mathbb{Z}_{120}$$

$$4) \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_{25} = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{150}$$

$$5) \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_{25} = \mathbb{Z}_2 \times \mathbb{Z}_{300}$$

$$6) \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_{25} = \mathbb{Z}_{600}$$

En la justificación del método descrito se utiliza el siguiente resultado, que será necesario en el siguiente teorema y puede ser fácilmente verificado por el estudiante:

Si un grupo G es producto directo de los grupos H y K , y H es producto directo de L y M , entonces $G = L \times M \times K$.

Ahora, de los resultados anteriores podemos obtener el siguiente teorema que indica la forma de todo grupo abeliano finitamente generado.

TEOREMA 3.14.4 (Caracterización de los grupos abelianos finitamente generados)

Todo grupo abeliano finitamente generado G es isomorfo a un producto directo de grupos cíclicos de la forma

$$(1) \mathbb{Z}_{p_1^{r_1}} \times \mathbb{Z}_{p_2^{r_2}} \times \dots \times \mathbb{Z}_{p_n^{r_n}} \times \mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z},$$

donde los p_i son primos no necesariamente distintos, y también de la forma

$$(2) \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \dots \times \mathbb{Z}_{m_r} \times \mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z},$$

donde m_i divide a m_{i+1} .

En ambos casos el producto directo es único, salvo isomorfismo. El número de Betti de G (o sea, el número de factores $\mathbb{Z}$) es único, los coeficientes de torsión m_i de G son únicos y las potencias $p_i^{r_i}$ son únicas.

Demostración

Según el lema 1, $G = T \times F$ donde T es el grupo de torsión de G y F es un grupo libre de torsión.

Como F es un grupo abeliano libre de torsión y subgrupo de un grupo abeliano finitamente generado, será a su vez finitamente generado, según el lema 4.

Por otra parte, por ser T el grupo de torsión de un grupo finitamente generado, será un grupo finito. (El estudiante puede comprobar que un grupo de torsión finitamente generado es finito, mediante la aplicación de la definición de grupo finitamente generado.)

Haciendo, entonces, uso del lema 3 para T y del lema 2 para F , queda demostrado el teorema.

Para concluir, veamos algunas aplicaciones de los resultados anteriormente expuestos.

Aplicaciones

TEOREMA 3.14.5

Si m divide el orden de un grupo abeliano finito G , entonces G tiene un subgrupo de orden m .

Demostración

Podemos suponer por el lema 3 que

$$G = \mathbb{Z}_{p_1^{r_1}} \times \mathbb{Z}_{p_1^{r_2}} \times \dots \times \mathbb{Z}_{p_n^{r_n}}.$$

Entonces

$$m = p_1^{s_1} \cdot p_2^{s_2} \cdot \dots \cdot p_n^{s_n} \text{ con } 1 \leq s_i \leq r_i.$$

ya que el orden n de G podía escribirse como $n = p_1^{r_1} \cdot p_2^{r_2} \cdot \dots \cdot p_n^{r_n}$.

Entonces $(p_i)^{r_i - s_i}$ genera un subgrupo cíclico de $\mathbb{Z}_{p_i^{r_i}}$ de orden igual al cociente de $(p_i)^{r_i}$ entre el máximo común divisor de $p_i^{r_i}$ y $p_i^{r_i - s_i}$. Como $\text{mcd}(p_i^{r_i}, p_i^{r_i - s_i}) = p_i^{r_i - s_i}$, tenemos que $(p_i)^{r_i - s_i}$ genera un subgrupo cíclico de $\mathbb{Z}_{p_i^{r_i}}$ de orden $(p_i)^{s_i}$. Entonces

$$\langle (p_1)^{r_1 - s_1} \rangle \times \langle (p_2)^{r_2 - s_2} \rangle \times \dots \times \langle (p_n)^{r_n - s_n} \rangle$$

es el subgrupo cíclico de G de orden m que queríamos encontrar.

Acabamos de demostrar que el recíproco del teorema de Lagrange se cumple para los grupos abelianos finitos. Esto es de gran utilidad, pues permite afirmar, por ejemplo, que en un grupo abeliano de orden 24 existen subgrupos propios de orden 2, 3, 4, 6, 8 y 12.

Veamos con un ejemplo otro problema que podemos resolver mediante el teorema de caracterización de los grupos abelianos finitamente generados y finitos que acabamos de estudiar.

Sabemos que un grupo de orden primo es cíclico y también tenemos un criterio para determinar cuándo un grupo del tipo $Z_n \times Z_m$ es cíclico. Vamos ahora a encontrar un criterio suficiente para que un grupo abeliano finito de orden m sea cíclico.

TEOREMA 3.14.6

Si m no es divisible por el cuadrado de ningún número primo, entonces todo grupo abeliano de orden m es cíclico.

Demostración

Sea G un grupo abeliano de orden m , donde m no es divisible por el cuadrado de ningún número primo.

Sabemos que G es isomorfo a un grupo de la forma

$$Z_{p_1^{r_1}} \times Z_{p_2^{r_2}} \times \dots \times Z_{p_n^{r_n}}.$$

Luego podemos identificarlo con este.

Pero como $m = p_1^{r_1} \cdot p_2^{r_2} \cdot \dots \cdot p_n^{r_n}$ y m no es divisible por el cuadrado de ningún número primo, tenemos que

$$r_1 = r_2 = \dots = r_n = 1.$$

De aquí

$$G = Z_{p_1} \times Z_{p_2} \times \dots \times Z_{p_n},$$

y como los p_i son primos relativos, tenemos

$$G \cong Z_{p_1 \cdot p_2 \cdot \dots \cdot p_n}.$$

Por tanto, G es cíclico.

Ejemplo

Tomemos un grupo de orden 42, que no es divisible por el cuadrado de ningún número primo ya que sus únicos divisores son 2, 3 y 7. Según el teorema 3.14.6, podemos decir que todo grupo de orden 42 es cíclico; en otras palabras, existe un único grupo de orden 42, Z_{42} .

Observemos que con una aplicación directa del método de búsqueda de los grupos abelianos finitos de un orden dado, llegaríamos al mismo resultado ya que $42 = 2 \cdot 3 \cdot 7$ nos da la única posibilidad para un grupo de orden 42: $Z_2 \times Z_3 \times Z_7 = Z_{42}$; pero la propiedad dada por el teorema 3.14.6 nos permite ganar tiempo en la búsqueda de los grupos finitos de un orden dado.

Destaquemos que la condición dada por el teorema anterior es una condición suficiente pero no necesaria, ya que para los grupos de orden primo su orden no es divisible por el cuadrado de un primo y, sin embargo, son cíclicos.

Hemos dicho en varias ocasiones que los grupos cíclicos son los grupos más sencillos que podemos considerar. El siguiente teorema confirma esa afirmación en el caso de los grupos abelianos finitos.

TEOREMA 3.14.7

Los grupos abelianos finitos que no son isomorfos a un producto directo de dos de sus subgrupos propios, son exactamente los grupos cíclicos de orden potencia de un primo.

(Un tal grupo se llama *indescomponible*. Esto significa, intuitivamente, que lo hemos llevado a la mínima expresión.)

Demostración

Supongamos que G es un grupo abeliano finito que no es posible expresarlo como producto directo de dos de sus subgrupos propios. Entonces G debe ser isomorfo a un producto del tipo $Z_{p_1^{r_1}} \times \dots \times Z_{p_k^{r_k}}$ y debe constar de un solo factor. Esto es,

$$G = Z_{p^r}.$$

Recíprocamente, sea un grupo Z_{p^r} . Si Z_{p^r} pudiera ser expresado como producto directo de dos de sus subgrupos propios, entonces $Z_{p^r} = Z_{p^i} \times Z_{p^j}$ con $i+j=r$, lo cual es imposible ya que el orden máximo posible de los elementos de $Z_{p^i} \times Z_{p^j}$ es: $p^{\max\{i, j\}} < p^r$.

Los resultados hallados para grupos abelianos finitos permiten extender el método de búsqueda de los subgrupos, expuesto en el epígrafe 3.13, a todos los grupos abelianos finitos, ya que sabemos de antemano que estos poseen subgrupos de orden m para cada m divisor del orden del grupo (no necesariamente uno de cada orden como en el caso de los cíclicos) y, por otra parte, conocemos los modelos para clasificar un grupo abeliano finito cualquiera.

Ejemplo

Busquemos los subgrupos de $Z_4 \times Z_4$.

$o(Z_4 \times Z_4) = 16$; luego tiene subgrupos propios de órdenes 2, 4 y 8.

$$Z_4 \times Z_4 = \{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (1,3), (2,0), (2,1), (2,2), (2,3), (3,0), (3,1), (3,2), (3,3)\}.$$

Los subgrupos de orden 2 son cíclicos:

$$\langle (0,2) \rangle = \{(0,2), (0,0)\}$$

$$\langle (2,0) \rangle = \{(2,0), (0,0)\}$$

$$\langle (2,2) \rangle = \{(2,2), (0,0)\}$$

Los subgrupos cíclicos de orden 4 son:

$$\langle (0,1) \rangle = \{(0,1), (0,2), (0,3), (0,0)\}$$

$$\langle (1,0) \rangle = \{(1,0), (2,0), (3,0), (0,0)\}$$

$$\langle (1,2) \rangle = \{(1,2), (2,0), (3,2), (0,0)\}$$

$$\langle (1,3) \rangle = \{(1,3), (2,2), (3,1), (0,0)\}$$

$$\langle (2,1) \rangle = \{(2,1), (0,2), (2,3), (0,0)\}$$

$$\langle (3,3) \rangle = \{(3,3), (2,2), (1,1), (0,0)\}$$

Hay un subgrupo de orden 4 isomorfo a V_4 :

$$\tilde{V}_4 = \{(0,0), (0,2), (2,0), (2,2)\}$$

No hay subgrupo cíclico de orden 8. Por tanto, si existe un subgrupo de orden 8 tiene que ser isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_4$ o $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$. Como en $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ no existen elementos de orden 4 y en $\mathbb{Z}_2 \times \mathbb{Z}_4$ sí, podemos decir que en $\mathbb{Z}_8 \times \mathbb{Z}_4$ hay al menos un subgrupo de orden 8 isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_4$.

Realmente podemos buscar los subgrupos de orden 8 pues están dados por grupos isomorfos a los productos de los grupos cíclicos generados por elementos de orden 2 y de orden 4, respectivamente; pero esto resulta un tanto engorroso y con los resultados hallados tenemos ya una información bastante completa sobre los subgrupos.

Ejercicios

- Encuentre todos los grupos abelianos (salvo isomorfismo) de orden: a) 60 b) 720 c) 1 800 d) 24 e) 25 f) 24 · 25.
- Sean m y n primos relativos. Pruebe que si existen (salvo isomorfismo) r grupos abelianos de orden m , y s de orden n , entonces existen (salvo isomorfismo) rs grupos abelianos de orden $m \times n$.
- Encuentre el subgrupo de $\mathbb{Z}_{12}$ generado por $\{2,3\}$, el generado por $\{4,6\}$ y el generado por $\{8,6,10\}$.
- Sea G un grupo abeliano de orden 72.
 - Diga cuántos subgrupos de orden 8 tiene. ¿Por qué?
 - Diga cuántos subgrupos de orden 4 tiene. ¿Por qué?
- Pruebe que si un grupo abeliano finito tiene orden una potencia de un primo p , entonces el orden de cada elemento del grupo es una potencia de p .
- Encuentre el subgrupo de torsión T de $(\mathbb{C}^*, \cdot)$.
- Sean G , H y K grupos abelianos finitamente generados. Pruebe que si $G \times K = H \times K$, entonces $G = H$.
- Pruebe que si G es finitamente generado y φ es un homomorfismo de G en un grupo G' , el valor de φ en cada elemento de G está completamente determinado por los valores $\varphi(a_i)$, ($1 \leq i \leq n$) de los generadores a_i de G .
- Halle cuántos homomorfismos existen
 - de $\mathbb{Z}_2 \times \mathbb{Z}_2$ en $\mathbb{Z}_2$
 - de $\mathbb{Z}_2 \times \mathbb{Z}_2$ sobre $\mathbb{Z}_2$
 - de $\mathbb{Z}_2 \times \mathbb{Z}_2$ en $\mathbb{Z}_4$
- Pruebe que si G es un grupo y H un subgrupo de G tal que G/H es libre de torsión, entonces H contiene el subgrupo de torsión de G .

Sugerencia: Considere un elemento de orden finito de G y pruebe que está contenido en H .

- Analizando el orden de los elementos de los grupos $\mathbb{Q}$, $\mathbb{R}/\mathbb{Z}$ y $\mathbb{C}$, clasifíquelos en: libre de torsión, con torsión y mixtos. Deduzca de aquí que estos tres grupos no son isomorfos.
- Pruebe que si T es el subgrupo de torsión de G , entonces $T \cap H$ es el subgrupo de torsión de cualquier subgrupo H dado en G .
- Pruebe que el subgrupo de torsión de $\mathbb{R}/\mathbb{Z}$ es $\mathbb{Q}/\mathbb{Z}$.

4 Anillos y cuerpos

4.1 Introducción

En este capítulo estudiaremos otro tipo de estructura algebraica llamada anillo. El modelo para esta estructura lo constituye el conjunto $\mathbb{Z}$ de los números enteros con las leyes de suma y producto y sus propiedades.

Mientras que la teoría de grupos partía fundamentalmente de las transformaciones biyectivas y la estructura que de ellas resulta con la ley de composición de aplicaciones, la teoría de anillos surge de diversos problemas, por lo cual se presenta menos ordenada y unificada que la teoría de grupos.

Sin embargo, veremos que los problemas fundamentales estudiados en los grupos tienen sus homólogos en los anillos. Así, estudiaremos las nociones de *ideal*, *anillo cociente* y *homomorfismo de anillos*, que son los homólogos de los de subgrupos invariante, *grupo cociente* y *homomorfismo de grupos*, respectivamente. También estudiaremos varios tipos de anillos, en especial los llamados *dominios de integridad* y los *cuerpos*. A partir de estos últimos haremos una breve introducción a la *teoría de extensión de cuerpos*, que de hecho permite formalizar las ideas que dieron impulso al álgebra: las de la solubilidad de ecuaciones algebraicas.

Durante la segunda mitad del siglo XIX, el desarrollo de la idea de estructura algebraica llevó a amplias generalizaciones de los conjuntos numéricos y de la aritmética. Gauss generalizó el concepto *número entero* mediante el estudio de los "enteros de Gauss", esto es, números complejos de la forma $a+bi$, con a y b enteros. Dedekind fue más lejos en la generalización con los "enteros algebraicos", que son números que satisfacen una ecuación polinomial de coeficientes enteros y coeficiente principal igual a uno. Tales sistemas de "enteros" no poseen todas las propiedades de los números racionales o reales, ya que no se cumple que todo número perteneciente a los enteros de Gauss o a los enteros algebraicos y que sea no nulo, posea un inverso para la multiplicación, pero poseen todas las otras propiedades de $\mathbb{Q}$ y $\mathbb{R}$, más aún, todas las propiedades de $\mathbb{Z}$.

El estudio de estas generalizaciones y los problemas relacionados con ellas, tiene como punto de partida el del estudio de los conceptos *anillo* e *ideal* introducidos por Kummer a mediados del siglo XIX (1846 aproximadamente) y que habían tenido su antecedente en los trabajos de Gauss y Dedekind.

4.2 Definiciones, ejemplos y primeras propiedades

DEFINICIÓN 4.2.1

Se dice que un conjunto A no vacío es un *anillo* si y solo si están definidas sobre A dos operaciones, que se denotan por $+$ y $\cdot$, respectivamente, y que cumplen las propiedades siguientes:

(A1) $(A, +)$ es un grupo abeliano.

El elemento neutro para $+$ lo denotaremos por 0.

(A2) La operación $\cdot$ es distributiva respecto a $+$, esto es,

$$\left. \begin{aligned} a \cdot (b+c) &= ab+ac \\ (b+c) \cdot a &= ba+ca \end{aligned} \right\} \text{ para todo } a, b, c \in A.$$

(A3) La operación $\cdot$ es asociativa, o sea, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ para todo $a, b, c \in A$.

El grupo $(A, +)$ se llama el *grupo aditivo de A*.

Si existe $1 \in A$ tal que $a \cdot 1 = 1 \cdot a = a$ para todo $a \in A$, decimos que A es un *anillo unitario* o con *identidad* y a 1 lo llamamos *elemento unitario de A*. En este caso el elemento unitario 1 es único.

Cuando hablemos de anillo unitario no estaremos considerando el caso trivial $\{0\}$, que es el único en que es posible que coincidan los elementos neutros para la suma y el producto.

La definición que hemos adoptado para anillo es la más general, esto es, no se asume la existencia de un elemento neutro, ni la conmutatividad para $\cdot$. Sin embargo, veremos que muchos de los ejemplos importantes que aparecen cumplen una de estas propiedades o ambas. En particular, cuando expliquemos el concepto de *isomorfismo de anillos* señalaremos cómo es posible, dado un anillo no unitario, encontrar un anillo unitario que contenga un subconjunto estructuralmente idéntico al primero.

Ya señalamos en la introducción cómo el esquema seguido en el estudio de los grupos se repite en lo fundamental en los anillos. De ahí que pasemos inmediatamente a estudiar el concepto de *subestructura* para los anillos.

DEFINICIÓN 4.2.2

Se dice que $A' \subset A$, $A' \neq \emptyset$, es un *subanillo de A* si y solo si A' es un anillo para las mismas operaciones de A , esto es, si y solo si

$$x, y \in A' \Rightarrow \left\{ \begin{aligned} a) & x+y \in A' \\ b) & x \cdot y \in A' \end{aligned} \right.$$

Con a) aseguramos que $(A', +)$ sea un subgrupo de $(A, +)$.

Al ser interno en A por b , se cumplen las propiedades asociativa de y distributiva de con respecto a $+$.

Al igual que hicimos en el capítulo de grupos, podemos hablar del subanillo de A generado por un subconjunto B de A , al cual denotamos por $\langle B \rangle$ y es la intersección de todos los subanillos de A que contienen a B como subconjunto; pero no insistiremos más en esta idea ya que el concepto de subanillo no desempeña en la teoría de anillos la función que el concepto de subgrupo tiene en la teoría de grupos. Realmente nos interesan solo cierto tipo de subanillos: los llamados ideales, que estudiaremos más adelante.

Cuando la ley del anillo es conmutativa, decimos que A es un anillo conmutativo. El estudio de los anillos conmutativos es muy amplio, constituye un tema de estudio especial y es de gran importancia para la Geometría algebraica, disciplina matemática que relaciona el Álgebra, la Geometría y la Topología.

Ejemplos

1. El conjunto $\mathbb{Z}$ con las operaciones de suma y producto usuales es un anillo conmutativo y unitario. Los subgrupos $m\mathbb{Z}$ de $\mathbb{Z}$ son también subanillos de $\mathbb{Z}$ como anillo.
2. Los conjuntos $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ son anillos conmutativos y unitarios para las operaciones de suma y producto usuales, y las inclusiones $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ indican que cada uno de ellos es un subanillo del siguiente.
3. En el grupo cíclico $(\mathbb{Z}_n, +)$ podemos definir la operación $a \cdot b = r$, donde r es el resto de la división de $a \cdot b$ entre n . Con esta operación $(\mathbb{Z}_n, +, \cdot)$ se convierte en un anillo.
4. El conjunto $\mathbb{Z}(i) = \{a+bi : a, b \in \mathbb{Z}\}$ es un anillo para la suma y el producto de $\mathbb{C}$ y es un subanillo de $(\mathbb{C}, +, \cdot)$.
5. $M_n(\mathbb{R})$ con la suma y producto de matrices es un anillo unitario, pero no conmutativo. El elemento unitario es la matriz unitaria $I_n \in M_n(\mathbb{Z})$ y $M_n(\mathbb{Q})$ son subanillos de $M_n(\mathbb{R})$.
6. $\mathbb{R}^{\mathbb{R}} = \{f : \mathbb{R} \rightarrow \mathbb{R}\}$, o sea, el conjunto de las funciones reales de variable real con la suma y el producto usuales de funciones, es un anillo conmutativo y unitario.

El ejemplo anterior admite la generalización siguiente:

7. Sea A un anillo arbitrario y X un conjunto cualquiera. Consideremos $A^X = \{f : X \rightarrow A\}$, esto es, el conjunto de las funciones de X en A , y sobre A^X las operaciones o leyes internas siguientes:

$$(f+g)(x) = f(x) + g(x) \quad \text{para todo } x \in X$$

$$(f \cdot g)(x) = f(x) \cdot g(x) \quad \text{para todo } x \in X$$

(Por comodidad denotamos con los mismos símbolos $(+ \text{ y } \cdot)$ las operaciones del anillo A y las de A^X .)

Al comprobar el estudiante que $(A^X, +, \cdot)$ es un anillo, notará que solo es necesario utilizar las propiedades de las operaciones en A , ya que $f(x) \in A$, por lo cual el conjunto X no requiere de ninguna condición adicional.

El conjunto A^X con la suma y el producto de funciones definidas arriba es un anillo. Si el anillo A es conmutativo o unitario, también lo será A^X .

El anillo A^X generaliza una gran cantidad de ejemplos de anillos de funciones, como son:

$$\mathbb{R}^{[0,1]} = \{\text{funciones reales definidas en } [0,1]\}, \mathbb{R}^{\mathbb{R}}, \text{ etcétera.}$$

En $\mathbb{R}^{[0,1]}$, $C[0,1] = \{\text{funciones reales continuas definidas en } [0,1]\}$ es un subanillo, lo que puede verificarse fácilmente mediante las propiedades de las funciones continuas.

8. Si $(A, +)$ es un grupo abeliano arbitrario, podemos convertirlo en un anillo definiendo el siguiente producto en A :

$$x \cdot y = 0 \text{ para todo } x, y \in A.$$

$$(A, +, \cdot) \text{ es un anillo.}$$

9. $(\mathbb{Z}_n, +)$ es un anillo con el producto definido por $a \cdot b = r$, entonces $a \cdot b = r$, donde r es el resto de $a \cdot b$ al ser dividido por n . Como ejemplo presentamos las tablas de multiplicar de $\mathbb{Z}_2$, $\mathbb{Z}_3$ y $\mathbb{Z}_4$:

	0	1
0	0	0
1	0	1

	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

Primeras propiedades (reglas de cálculo)

Las siguientes propiedades se conocen como reglas de cálculo, ya que permiten saber cómo "operar" en un anillo. Sea A un anillo cualquiera.

En primer lugar, no debemos olvidar que un anillo A es un grupo abeliano; por tanto, podemos utilizar en el trabajo con un anillo todas las propiedades que conocemos para los grupos abelianos para la operación de suma del anillo.

$$(1) a \cdot 0 = 0 \cdot a = 0 \text{ para todo } a \in A.$$

En efecto,

$$a+0=a \Rightarrow a(a+0)=a \cdot a \Rightarrow a^2+a \cdot 0=a^2 \Rightarrow a \cdot 0=0$$

ya que el neutro es único.

Análogamente $0 \cdot a=0$.

(2) Si $A \neq \{0\}$ y A es unitario, $0 \neq 1$, pues si suponemos $0=1$ obtenemos $a=a \cdot 1=a \cdot 0=0$ para todo $a \in A$, o sea, A sería el anillo nulo.

$$(3) (-a) \cdot b = a(-b) = -(ab)$$

Para probarlo consideremos

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

Podemos obtener otras propiedades similares, como son:

$$(4) (-a)(-b)=a \cdot b$$

$$(5) -a=(-1) \cdot a$$

$$(6) n(ab)=(na)b=a(nb) \text{ para todo } n \in \mathbb{Z} \text{ y } a, b \in A.$$

(7) Si $a, b \in A$ y a y b conmutan, entonces

$$(a+b)^n = \sum_{i=0}^n \binom{n}{i} a^i b^{n-i},$$

lo que se conoce como *fórmula del binomio*.

Por supuesto, si A es conmutativo la propiedad (7) se cumple para todo $a, b \in A$. Esta propiedad se demuestra por inducción y la dejamos como ejercicio al estudiante, ya que será de gran utilidad en la obtención de resultados futuros.

Observemos en el ejemplo $\mathbb{Z}$ y sus subanillos $n\mathbb{Z}$ que un anillo puede ser unitario y tener subanillos no unitarios. Más aún, puede ocurrir que un anillo unitario posea un subanillo unitario con elemento unitario distinto al del anillo. Este es el caso del anillo $\mathbb{Z} \times \mathbb{Z} = \{(a, b) : a, b \in \mathbb{Z}\}$ con la suma y el producto componente a componente, el cual tiene como elemento unitario a $(1, 1)$; mientras que su subanillo $\mathbb{Z} \times \{0\} = \{(n, 0) : n \in \mathbb{Z}\}$ tiene como elemento unitario a $(1, 0)$. Nótese que $(1, 1) \notin \mathbb{Z} \times \{0\}$.

Isomorfismo de anillos

De la misma forma que en los grupos, se presenta en los anillos la situación de dos anillos que son idénticos, salvo el nombre de los elementos. Demos, pues, la definición siguiente:

DEFINICIÓN 4.2.3

Se dice que el anillo A es *isomorfo* al anillo A' si y solo si existe una aplicación biyectiva de A en A' tal que:

- (1) $\varphi(a+b) = \varphi(a) + \varphi(b)$
 - (2) $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$
- para todo $a, b \in A$.

También para los anillos usaremos la notación $A=A'$, que se lee: A es isomorfo a A' , y denotaremos con los mismos símbolos las leyes de A y A' .

Según (1), los grupos $(A, +)$ y $(A', +)$ son isomorfos.

Ejemplos

1. El anillo $M_n(\mathbb{R})$ de las matrices cuadradas de orden n con coeficientes reales y el anillo $\text{End}(\mathbb{R}^n)$ de los endomorfismos u operadores del espacio vectorial $\mathbb{R}^n$ son isomorfos.
2. Los grupos $(\mathbb{Z}, +)$ y $(2\mathbb{Z}, +)$ son isomorfos mediante el isomorfismo $\varphi(n)=2n$. Sin embargo, este isomorfismo de grupos no es un isomorfismo de los anillos $(\mathbb{Z}, +, \cdot)$ y $(2\mathbb{Z}, +, \cdot)$, ya que en general $\varphi(xy) \neq \varphi(x) \varphi(y)$ para $x, y \in \mathbb{Z}$.

Los isomorfismos de anillos tienen propiedades similares a los isomorfismos de grupos, como son:

- (1) La imagen por φ de un subanillo de A es un subanillo de A' .
- (2) La preimagen por φ de un subanillo de A' es un subanillo de A .
- (3) Si A es unitario con elemento unitario 1, entonces $\varphi(1)$ es unitario y su elemento unitario es $\varphi(1)$.

Son válidas también las aclaraciones hechas en el caso de grupos sobre cómo decidir si dos anillos son isomorfos o no atendiendo a las propiedades estructurales de ambos, teniendo en cuenta las dos operaciones en el caso de los anillos. Por ejemplo, para decidir que $(\mathbb{Z}, +, \cdot)$ y $(2\mathbb{Z}, +, \cdot)$ no son isomorfos podíamos haber atendido a que $(\mathbb{Z}, +, \cdot)$ es unitario y, sin embargo $(2\mathbb{Z}, +, \cdot)$ no lo es.

Dijimos anteriormente que una vez estudiado el concepto de *isomorfismo de anillos* era posible demostrar que todo anillo no unitario podía ser "sumergido" en un anillo unitario, en el sentido de que siempre es posible encontrar, dado un anillo A , un anillo unitario A' que posea un subanillo isomorfo a A . Este hecho permite que muchos problemas sobre los anillos puedan estudiarse solamente en el caso de los anillos unitarios. Veamos el teorema siguiente:

TEOREMA 4.2.1

Dado un anillo A , siempre es posible encontrar un anillo unitario A' que contenga un subanillo isomorfo a A .

Demostración

Consideremos el siguiente conjunto A' :

$A' = \mathbb{Z} \times A = \{(n, a) : n \in \mathbb{Z}, a \in A\}$ y definamos las operaciones

$$(m, a) + (n, b) = (m+n, a+b)$$

$$(m, a) \cdot (n, b) = (mn, mb+na+ab),$$

considerando las operaciones entre los elementos de $\mathbb{Z}$ y de A como las usuales de esos anillos y $na = \underbrace{a + \dots + a}_n$.

Sabemos que $(\mathbb{Z} \times A, +)$ es un grupo abeliano, ya que es el producto directo de los grupos $(\mathbb{Z}, +)$ y $(A, +)$. Veamos las propiedades del producto:

$$\begin{aligned} ((m, a) \cdot (n, b)) \cdot (q, c) &= (mn, mb + na + ab) \cdot (q, c) \\ &= ((mn)q, (mn)c + q(mb + na + ab) + (mb + na + ab)c) \\ &= ((mn)q, (mn)c + q(mb) + q(na) + q(ab) + (mb)c + \\ &\quad + (na)c + (ab)c) \\ (m, a) \cdot ((n, b) \cdot (q, c)) &= (m, a) \cdot (nq, nc + qb + bc) = \\ &= (m(nq), m(nc) + m(qb) + m(bc) + (nq)a + \\ &\quad + a(nc + qb + bc)) \end{aligned}$$

Aplicando las propiedades asociativas para $\cdot$ en $\mathbb{Z}$ y A , la distributividad de $\cdot$ con respecto a $+$ en A y las propiedades de los múltiplos, es fácil ver que se cumple la asociatividad del producto en $\mathbb{Z} \times A$.

El elemento $(1, 0)$ es el elemento unitario para $\mathbb{Z} \times A$, pues

$$(1, 0) \cdot (m, a) = (m, a) = (m, a) \cdot (1, 0) \text{ para todo } (m, a) \in \mathbb{Z} \times A.$$

También se cumple la distributividad, ya que

$$\begin{aligned} (m, a) \cdot ((n, b) + (q, c)) &= (m, a) \cdot (n + q, b + c) \\ &= (m(n + q), m(b + c) + (n + q)a + a(b + c)) \\ &= (mn + mq, mb + mc + na + qa + ab + ac) \\ (m, a) \cdot (n, b) + (m, a) \cdot (q, c) &= (mn, mb + na + ab) + (mq, mc + qa + ac) \\ &= (mn + mq, mb + mc + na + qa + ab + ac). \end{aligned}$$

De modo análogo se prueba la distributividad a la derecha. Por tanto, $(A, +, \cdot)$ es un anillo unitario con elemento unitario $(1, 0)$.

Si tomamos $\tilde{A} = \{(0, a) : a \in A\}$, $\tilde{A}$ es evidentemente un subanillo de A' pues

$$\begin{aligned} (0, a) - (0, b) &= (0, a - b) \in \tilde{A} \\ (0, a) \cdot (0, b) &= (0, ab) \in \tilde{A}. \end{aligned}$$

De manera natural podemos definir

$$\varphi : A \longrightarrow \tilde{A} \text{ dado por } a \longmapsto (0, a),$$

que es un isomorfismo de anillos, ya que para todo $a, b \in A$ tenemos:

$$\begin{aligned} (1) \quad \varphi(a + b) &= (0, a + b) = (0, a) + (0, b) = \varphi(a) + \varphi(b) \\ (2) \quad \varphi(a \cdot b) &= (0, ab) = (0, a) \cdot (0, b) = \varphi(a) \cdot \varphi(b) \end{aligned}$$

$\text{Im } \varphi = \tilde{A}$ por la propia definición de φ .

φ es inyectiva pues $(0, a) = (0, b) \Rightarrow a = b$ por la definición de igualdad entre pares.

Ejercicios

1. Analice cuáles de los siguientes conjuntos con las operaciones indicadas

son anillos. En los casos en que lo sean, analice si son conmutativos y/o unitarios.

- a) $A = \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$ con la suma y el producto usuales en $\mathbb{R}$.
- b) $R = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ con la suma y el producto usuales en $\mathbb{R}$.
- c) $\mathbb{R}_i = \{ri : r \in \mathbb{R}\}$ con la suma y el producto usuales en $\mathbb{C}$.
- d) $\mathbb{Q}$ con respecto a las operaciones $\odot$ y $\circ$ (en ese orden):

$$a \odot b = a \cdot b \text{ y } a \circ b = a + b \quad \forall a, b \in \mathbb{Q},$$

y donde $+$ y $\cdot$ son la suma y el producto usuales en $\mathbb{Q}$.

- e) El conjunto $M = \{(a, b, c, d) : a, b, c, d \in \mathbb{Q}\}$ con $+$ y definidas por

$$\begin{aligned} (a, b, c, d) + (e, f, g, h) &= (a + e, b + f, c + g, d + h) \\ (a, b, c, d) \cdot (e, f, g, h) &= (ae + bg, af + bh, ce + dg, cf + dh). \end{aligned}$$
- f) $T = \{2n + 1 : n \in \mathbb{Z}\}$ con la suma y el producto de $\mathbb{Z}$.
- g) En $\mathbb{Z}$ con $a \odot b = a + b + 1$ y $a \circ b = a + b + ab$.

2. Un anillo A es un *anillo de Boole* si y solo si $a^2 = a$ para todo $a \in A$. Pruebe que un anillo de Boole es conmutativo.
3. Pruebe que la intersección de una familia de subanillos de un anillo A es un subanillo de A .
4. Sea E un conjunto y $\mathbf{P}(E)$ el conjunto de todos los subconjuntos de E (potencia de E). Considere en $\mathbf{P}(E)$ las operaciones:

$$\begin{aligned} A + B &= (A \cup B) - (A \cap B) = \{x : (x \in A \text{ o } x \in B) \wedge x \notin A \cap B\} \\ A \cdot B &= A \cap B, \end{aligned}$$
 definidas para todo $A, B \in \mathbf{P}(E)$.
 - a) Pruebe que $(\mathbf{P}(E), +, \cdot)$ es un anillo.
 - b) Dé las tablas para $+$ y $\cdot$ cuando $E = \{a, b\}$.
 - c) Pruebe que para todo conjunto E , $(\mathbf{P}(E), +, \cdot)$ es un anillo de Boole.
5. Sean G un grupo abeliano y $\text{End}(G) = \{f : G \longrightarrow G \text{ endomorfismo}\}$.
 - a) Pruebe que $\text{End}(G)$ con las operaciones:

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \quad \forall x \in G \\ (f \cdot g)(x) &= f(g(x)) \quad \forall f, g \in \text{End}(G) \end{aligned}$$
 es un anillo
 - b) ¿Es este anillo conmutativo? ¿Es unitario?

6. Dados un anillo A y $n \in \mathbb{N}$, llamaremos $M_n(A)$ al anillo formado por las matrices cuadradas de orden n cuyos coeficientes pertenecen a A . Pruebe que si consideramos en $M_n(A)$ la suma y el producto usuales de matrices (considerando que los coeficientes se suman y se multiplican en A), entonces $M_n(A)$ es un anillo para dichas operaciones.

7. Sea A un conjunto dotado de dos operaciones $+$ y $\cdot$ que satisface todos los axiomas de los anillos unitarios con excepción del axioma de la conmutatividad de la suma. Pruebe que dicho axioma puede deducirse de los restantes y, por tanto, $(A, +, \cdot)$ es un anillo unitario.

Sugerencia: Desarrolle en dos formas $(a+b)(1-1)$.

8. Pruebe que el conjunto $A_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$ es un subanillo de $M_2(\mathbb{R})$ y que la aplicación $\varphi: \mathbb{C} \rightarrow A_2(\mathbb{R})$ dada por $a+bi \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ es un isomorfismo de anillos.

9. Pruebe que en un anillo A el conjunto de los elementos que conmutan con cada elemento de un subconjunto S dado de A es un subanillo de A llamado $C(S)$. En el caso particular $S=A$, este subanillo se llama el *centro* de A y es un subanillo conmutativo.

Halle $C(S)$ en $M_n(\mathbb{R})$ para $S = \{E_{ij} : i, j = 1, \dots, n\}$ $E_{ij} = (a_{km})$

$$\text{con } a_{km} = \delta_{ik} \delta_{jm} \text{ y } \delta_{ik} = \begin{cases} 1, & \text{para } k=i \\ 0, & \text{para } k \neq i \end{cases}$$

Determine $C(M_n(\mathbb{R}))$.

10. Pruebe que si un anillo A tiene n elementos, entonces para todo $a \in A$, $n \cdot a = 0$.

11. a) Pruebe que $P = \{(a, b, -b, a) : a, b \in \mathbb{Z}\}$ con las operaciones:

$$(a, b, -b, a) + (c, d, -d, c) = (a+c, b+d, -b-d, a+c)$$

$$(a, b, -b, a) \cdot (c, d, -d, c) = (ac-bd, ad+bc, -ad-bc, ac-bd)$$

es un subanillo conmutativo del anillo M del ejercicio 1e, y analice si la aplicación $(a, b, -b, a) \mapsto a$ de P en $\mathbb{Z}$ es un isomorfismo de anillos.

- b) Pruebe que la aplicación

$$(a, b, -b, a) \mapsto a+bi \text{ de } P \text{ en } \mathbb{Z}(i) = \{a+bi : a, b \in \mathbb{Z}\} \text{ (enteros de Gauss)}$$

es un isomorfismo de anillos.

12. En un anillo A , se dice que $x \in A$ es *nilpotente* si y solo si existe un entero $n > 0$ tal que $x^n = 0$. Pruebe que si x y y son nilpotentes y $x \cdot y = y \cdot x$, entonces $x+y$, $x \cdot y$ son nilpotentes.

Sugerencia: Para $x+y$ utilice la fórmula del binomio.

13. Un elemento a de un anillo A se llama *casi inversible a la izquierda* (derecha) si existe $b \in A$ tal que $a+b-ba=0$ ($a+b-ab=0$). Pruebe que a casi inversible a la izquierda es equivalente a que $1-a$ tenga un inverso a la izquierda (inverso a la derecha) $1-b$ en $\mathbb{Z} \times A$, dotado este de la estructura de anillo unitario dada en el teorema 4.2.1.

14. Sea A un anillo unitario. Pruebe que A con las operaciones $a \oplus b = a+b+1$ y $a \odot b = ab+a+b$ (donde $+$ y $\cdot$ son las operaciones de A) es un anillo isomorfo a A .

4.3 Distintos tipos de anillos

Veamos algunos conceptos que nos permitirán distinguir entre varios tipos de anillos.

Divisores de cero

Al introducir el concepto de anillo dijimos que dicho concepto podía ser considerado una generalización del conjunto numérico $\mathbb{Z}$. En dicho conjunto estamos acostumbrados a hacer uso de la propiedad:

$$x \cdot y = 0 \Rightarrow x = 0 \text{ o } y = 0.$$

Sobre esta propiedad sabemos de los cursos de Álgebra lineal que no la conservan anillos más generales. Por ejemplo, es clásico en $M_2(\mathbb{Z})$ considerar casos como el siguiente:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq 0 \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \neq 0 \quad \text{y} \quad A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

que nos indican que dicha propiedad, usual en los conjuntos numéricos, no se cumple en general en los anillos.

DEFINICIÓN 4.3.1

Sea A un anillo. Se dice que $a \in A$, $a \neq 0$, es un *divisor de cero a la izquierda*, si existe $b \in A$, $b \neq 0$, tal que $a \cdot b = 0$. Si dado $a \neq 0$ existe $b \neq 0$ tal que $b \cdot a = 0$, se dice que a es un *divisor de cero a la derecha*.

Si el anillo es conmutativo no será necesario distinguir entre divisor izquierdo o derecho y hablaremos simplemente de divisores de cero.

Ejemplos

1. En $\mathbb{Z}_{12}$, son divisores de cero: 2, 3, 4, 6, 8, 9 y 10. Aprovechemos este ejemplo para mostrar una consecuencia que puede traer la existencia de divisores de cero en un anillo.

Es conocido que para resolver en $\mathbb{Z}$ la ecuación $x^2 - 5x + 6 = 0$, tomamos $x^2 - 5x + 6 = (x-2)(x-3) = 0$, de donde se obtienen las soluciones $x=2$ o $x=3$.

De hecho estamos asumiendo que $\mathbb{Z}$ no tiene divisores de cero.

Pero si planteamos este problema en $\mathbb{Z}_{12}$, o sea, si queremos resolver la ecuación $x^2 - 5x + 6 = 0$ (considerando $1, 5, 6, 0 \in \mathbb{Z}_{12}$), tendremos también $(x-2)(x-3) = 0$. Sin embargo, en este caso las soluciones serán no solo 2 y 3, sino muchas más; por ejemplo, 6 y 11 son soluciones ya que

$$(6-2)(6-3) = 4 \cdot 3 = 0$$

$$(11-2)(11-3) = 9 \cdot 8 = 0.$$

Observemos que los elementos de $\mathbb{Z}_{12}$ que no son divisores de cero, son 1, 5, 7, 11, esto es, los primos relativos con 12 (al hablar de primos relativos nos referimos a los enteros representantes de las clases en $\mathbb{Z}_{12}$).

2. $\mathbb{Z} \times \mathbb{Z} = \{(a, b) : a, b \in \mathbb{Z}\}$ con la suma y el producto componente a componente, es un anillo conmutativo y unitario. Tenemos que $(1, 0) \cdot (0, 1) = (0, 0)$ y, según habíamos visto anteriormente, esto significa que $(1, 0)$ y $(0, 1)$ son divisores de cero en $\mathbb{Z} \times \mathbb{Z}$.

3. En $\mathbb{R}^{\mathbb{R}} = \{f : \mathbb{R} \rightarrow \mathbb{R}\}$ con la suma y el producto usuales de funciones, consideremos f dada por $x \mapsto |x| + x$ y g dada por $x \mapsto |x| - x$. Es evidente que f y g son no nulas y, sin embargo, como $f(x) = 0$ para $x \leq 0$ y $g(x) = 0$ para $x \geq 0$, tenemos que $f \cdot g = 0$.

4. Es clásico el siguiente ejemplo en $A^X = \{f : X \rightarrow A\}$ con A anillo, X conjunto arbitrario y la suma y producto usuales de funciones. Tomamos $X = X' \cup CX'$, donde $X' \subset X$ y CX' es el complemento de X' en X , y definimos

$$f(x) = \begin{cases} a \in A \text{ para } x \in X' \ (a \neq 0) \\ 0 \text{ para } x \in CX' \end{cases}$$

$$g(x) = \begin{cases} 0 \text{ para } x \in X' \\ b \text{ para } x \in CX' \ (b \neq 0) \end{cases}$$

es evidente que f y g así definidas son no nulos y, sin embargo, $f \cdot g = 0$.

Es interesante la siguiente propiedad que encontramos presente en el ejemplo de $\mathbb{Z}_n$ y dejamos como ejercicio al estudiante.

PROPOSICIÓN 4.3.1

En el anillo $\mathbb{Z}_n$ los divisores de cero son precisamente los elementos que no son primos relativos con n .

Esta propiedad tiene el importante corolario siguiente:

COROLARIO

Si p es primo, entonces $\mathbb{Z}_p$ no tiene divisores de cero.

Está claro que la existencia o no de divisores de cero en un anillo está íntimamente ligada con la posibilidad de "cancelar" en una igualdad del tipo $a \cdot b = a \cdot c$, pero sabemos por experiencia que es imposible hacer esto para la multiplicación de matrices. Formalicemos esta idea:

Dado el anillo A , $a, b, c \in A$, se llama propiedad cancelativa para el producto a:

$$ab = ac \text{ con } a \neq 0 \Rightarrow b = c \text{ y}$$

$$ba = ca \text{ con } a \neq 0 \Rightarrow b = c.$$

PROPOSICIÓN 4.3.2

La propiedad cancelativa para el producto se cumple en un anillo A si y solo si A no posee divisores de cero.

Demostración

Sea A un anillo en el cual se cumple la propiedad cancelativa y supon-

gamos que $a \cdot b = 0$ para algún $a, b \in A$. Debemos probar que al menos uno de los dos, a o b , es cero.

Si $a \neq 0$, entonces $a \cdot b = a \cdot 0$ implica que $b = 0$ por la propiedad cancelativa, y análogamente, $b \neq 0$ implica que $a = 0$. Por tanto, no existen divisores de cero.

Recíprocamente, supongamos que A no tiene divisores de cero y que $ab = ac$ con $a \neq 0$. Entonces

$$ab - ac = a(b - c) = 0$$

Como $a \neq 0$ y A no tiene divisores de cero, tenemos $b - c = 0$ de donde $b = c$. Un argumento similar prueba que $ba = ca$ con $a \neq 0$ implica que $b = c$.

Unidades en un anillo unitario

En $\mathbb{Z}$ vimos que existe un elemento no nulo distinto de 1, este es -1 , que cumple $(-1) \cdot (-1) = 1$. Una situación similar encontramos en las matrices, donde las de determinante no nulo son inversibles.

Precisamente sabemos de los conjuntos numéricos que la diferencia estructural entre $\mathbb{Z}$ y $\mathbb{Q}$, por ejemplo, es la no existencia en $\mathbb{Z}$ de un inverso multiplicativo para cada entero no nulo. El concepto *inverso* para el producto tiene sentido solamente en un anillo unitario.

Estudiaremos ahora una definición que generaliza las situaciones anteriores.

DEFINICIÓN 4.3.2

Sea A un anillo unitario. Se dice que un elemento u de A es una *unidad* si y solo si u tiene un inverso multiplicativo en A , esto es, si y solo si existe $u' \in A$ tal que $u \cdot u' = u' \cdot u = 1$.

Está claro que el elemento unitario es siempre una unidad del anillo.

Ejemplos

1. En $\mathbb{Z}$ las unidades son 1, -1 ; en $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ todos los elementos no nulos.
2. En $\mathbb{R}[X]$ anillo, de los polinomios con coeficientes reales, las unidades son los polinomios constantes no nulos.
3. En $M_n(\mathbb{R})$ las unidades son las matrices de determinante no nulo.
4. En el anillo $\mathbb{Z}(i) = \{a + bi : a, b \in \mathbb{Z}\}$ de los enteros de Gauss, busquemos las unidades. Estas serán los elementos que tengan inverso, esto es, debe existir

$$(a + bi)^{-1} = \frac{1}{a + bi} = \frac{a - bi}{a^2 + b^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i,$$

y no es difícil verificar que esto ocurre si y solo si $a^2 + b^2 = 1$ ya que, en efecto, si $a^2 + b^2 = 1$, entonces $\frac{a}{a^2 + b^2}$ y $\frac{b}{a^2 + b^2}$ son enteros.

Recíprocamente, si $\frac{a}{a^2+b^2} = n \in \mathbb{Z}$ y $\frac{b}{a^2+b^2} = m \in \mathbb{Z}$, entonces

$$\frac{a^2+b^2}{(a^2+b^2)^2} = \frac{1}{a^2+b^2} = m^2+n^2 \in \mathbb{Z} \Rightarrow a^2+b^2=1. \quad m^2+n^2=1 \Rightarrow$$

De lo anterior se deduce que las unidades de $\mathbb{Z}(i)$ son ± 1 y $\pm i$ ya que $a, b \in \mathbb{Z}$.

El conjunto de las unidades de un anillo unitario tiene la importante propiedad siguiente:

PROPOSICIÓN 4.3.3

El conjunto de las unidades de un anillo unitario A es un grupo multiplicativo, que se llama *grupo de las unidades del anillo* A y se denota por $U(A)$.

Demostración

Basta probar que el producto de dos unidades es una unidad y que el inverso de una unidad es una unidad, ya que el producto es asociativo en el anillo y $1 \in U(A)$.

Es obvio que el inverso u^{-1} de una unidad u es también una unidad. Probemos, pues, que si u y u' son unidades, entonces $u \cdot u'$ es una unidad.

En efecto,

$$(u \cdot u')^{-1} = u'^{-1} \cdot u^{-1}.$$

Luego $u \cdot u'$ tiene inverso y es, por tanto, una unidad.

Hemos visto que existen anillos $(\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}_p$ con p primo) en los cuales el grupo de las unidades está formado por todos los elementos no nulos del anillo. Esto da lugar a la definición de un nuevo tipo de estructura.

DEFINICIÓN 4.3.3

Se llama *cuerpo* a un anillo K conmutativo unitario y en el cual todo elemento no nulo es una unidad.

Ejemplos

1. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$

2. $\mathbb{Q}(\sqrt{2}) = \{a+b\sqrt{2} : a, b \in \mathbb{Q}\}$ con la suma y producto usuales, es un anillo conmutativo y unitario.

Es fácil verificar que todo elemento no nulo es inversible, luego $\mathbb{Q}(\sqrt{2})$ es un cuerpo.

3. $\mathbb{Z}, \mathbb{R}[x], M_n(\mathbb{R})$ y $\mathbb{Z}(i)$ no son cuerpos.

En un cuerpo $(K, +)$ y $(K^*, \cdot)$ son grupos abelianos y la ecuación $ax=b$ tiene la solución única $x=a^{-1}b$ para cada $a \neq 0$. Esta solución suele denotarse por $\frac{b}{a}$ y se llama *cociente de b por a*.

Esta claro que en un cuerpo no existen divisores de cero, ya que si K es un cuerpo y $a, b \in K$, entonces $a \cdot b=0 \Rightarrow a^{-1}ab=b=0$.

DEFINICIÓN 4.3.4

Se llama *subcuerpo* de un cuerpo K a todo subconjunto no vacío de K que sea un cuerpo para las mismas operaciones de K .

En los ejemplos de cuerpos dados anteriormente, $\mathbb{Q}$ es un subcuerpo de $\mathbb{R}$ y $\mathbb{R}$ de $\mathbb{C}$, y $\mathbb{Q}(\sqrt{2})$ es un subcuerpo de $\mathbb{R}$.

Existen ejemplos importantes de anillos unitarios en los cuales todo elemento no nulo es inversible, pero que no son conmutativos.

DEFINICIÓN 4.3.5

Se llama *anillo de división* a un anillo unitario en el cual todo elemento no nulo es inversible.

Ejemplo

$$\text{Sea } H = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} : a, b \in \mathbb{C} \right\} \subset M_2(\mathbb{C}).$$

Es fácil comprobar que H es un subgrupo de $(M_2(\mathbb{C})^*, +)$.

Por otra parte, tenemos que

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \cdot \begin{pmatrix} c & d \\ -\bar{d} & \bar{c} \end{pmatrix} = \begin{pmatrix} ac-b\bar{d} & ad+b\bar{c} \\ -\bar{b}c-\bar{a}\bar{d} & -\bar{b}d+\bar{a}\bar{c} \end{pmatrix} \in H.$$

Luego H es cerrado para $\cdot$ y es un subanillo de $M_2(\mathbb{C})$.

H es un subanillo unitario pues para $a=1, b=0$, se cumple $I_2 \in H$. Además, es un anillo de división pues

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}^{-1} = \frac{1}{|a|^2+|b|^2} \begin{pmatrix} \bar{a} & -\bar{b} \\ \bar{b} & a \end{pmatrix}$$

y $\cdot$ no es conmutativa como puede comprobarse fácilmente.

Si llamamos

$$i = \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix}, j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, k = \begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}$$

puede verificarse que toda matriz $x = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix}$ de H puede ser escrita en forma única

como $x = a_0 I_2 + a_1 i + a_2 j + a_3 k$, donde $a = a_0 + a_1 \sqrt{-1}$ y $b = a_2 + a_3 \sqrt{-1}$,

y que el producto de dos elementos de H expresados en la forma anterior está determinado por el producto y la suma en $\mathbb{R}$, la propiedad distributiva y las relaciones

$$\begin{cases} i^2=j^2=k^2=-1 \\ ij=-ji=k, jk=-kj=i, ki=-ik=j \end{cases}$$

En 1843 Hamilton construyó este ejemplo, que es una ampliación del cuerpo de los números complejos y está formado por cuádruplos (a_0, a_1, a_2, a_3) de números reales para los cuales la suma y el producto se definen de forma tal, que se satisfagan las relaciones

nes expuestas, donde $i=(0,1,0,0)$, $j=(0,0,1,0)$, $k=(0,0,0,1)$. Hamilton llamó *cuaterniones* a estos cuádruplos y su descubrimiento constituyó un punto de partida para importantes investigaciones en Álgebra.

Existen muchos otros ejemplos de anillos de división cuya investigación constituye una parte importante del Álgebra.

Conocemos ejemplos importantes de anillos, como son: $\mathbb{Z}$ y $\mathbb{R}[x]$, que sin llegar a ser cuerpos, poseen la propiedad de no tener divisores de cero. En estos anillos se puede dar solución a problemas que no la tienen en un anillo común, por ejemplo, resolver ecuaciones algebraicas igualando a cero cada factor; pero no se pueden resolver otros problemas, como la existencia del cociente de dos elementos, que sí tienen solución en un cuerpo.

DEFINICIÓN 4.3.6

Se llama *dominio de integridad* a un anillo conmutativo unitario y sin divisores de cero.

Ejemplos

$\mathbb{Z}$, $\mathbb{R}[x]$, $\mathbb{Z}_p$ cuando p es primo.

Según la proposición 4.3.2, en un dominio de integridad se cumple la propiedad cancelativa para el producto. Todos los ejemplos de cuerpo que conocemos son también dominios de integridad, y el siguiente resultado prueba que esto no es casual.

TEOREMA 4.3.1

Todo cuerpo K es un dominio de integridad.

Demostración

Hemos probado que si a es una unidad, a no es divisor de cero. Como todo elemento no nulo de K es una unidad, queda probado que K no tiene divisores de cero, y como K es conmutativo y unitario por ser cuerpo, el teorema queda demostrado.

Sabemos por los ejemplos dados que el recíproco del teorema anterior no se cumple en general; sin embargo, es posible demostrar que en el caso de un dominio de integridad finito los conceptos *cuerpo* y *dominio de integridad* son equivalentes.

TEOREMA 4.3.2

Todo dominio de integridad finito D es un cuerpo.

Demostración

Sea $D=\{0,1,a_1,\dots,a_n\}$. Debemos probar que si $a \in D$ y $a \neq 0$, entonces existe $b \in D$ tal que $a \cdot b = 1$.

Consideremos los productos $aa_1, aa_2, \dots, aa_n$.

Todos estos elementos son distintos, pues si suponemos $aa_i = aa_j$, entonces $a_i = a_j$ por la propiedad cancelativa. Además, como D no tiene divisores de cero, ninguno de los aa_i es cero, así que alguno de los elementos $aa_1, \dots, aa_n$ debe ser igual a 1, esto es, $aa_i = 1$ para algún i , luego a tiene un inverso multiplicativo.

Del resultado anterior resultan ejemplos de cuerpos finitos. En particular, tenemos el siguiente:

COROLARIO

Si p es primo, $\mathbb{Z}_p$ es un cuerpo.

Vamos a resumir las características de los dominios de integridad, los anillos de división y los cuerpos como sigue:

	Conmutatividad	Existencia de elemento unitario	Existencia de divisores de cero	Existencia de unidades	Propiedad cancelativa
Dominios de integridad	sí	sí	no existen	Pueden existir	Se cumple
Anillos de división	no	sí	no existen	Todo elemento no nulo	Se cumple
Cuerpos	sí	sí	no existen	Todo elemento no nulo es unidad	Se cumple

Las relaciones entre los conceptos *dominio de integridad*, *anillo de división* y *cuerpo* pueden representarse así:

cuerpo $\Rightarrow$ anillo de división

cuerpo $\Rightarrow$ dominio de integridad

dominio de integridad finito $\Rightarrow$ cuerpo

De manera similar a como se obtuvo el resultado de que todo dominio de integridad finito es un cuerpo, puede probarse que todo dominio de integridad finito es un anillo de división. Esto lo plantearémos en los ejercicios.

Un resultado más profundo y muy importante es el debido a Wedderburn (1905), referente a que un anillo de división con un número finito de elementos es un cuerpo. Este resultado es de gran aplicación en ciertas ramas del Análisis funcional. Su demostración aparece, por ejemplo, en el libro de I.N. Herstein: *Topics in Algebra*.

Ejercicios

- a) Considere el anillo $\mathbb{Z}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$. ¿Es $\mathbb{Z}(\sqrt{2})$ un dominio de integridad? Busque las unidades de $\mathbb{Z}(\sqrt{2})$. ¿Es $\mathbb{Z}(\sqrt{2})$ un cuerpo?

b) Responda las preguntas anteriores para $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$.
- Sea $\bar{Q} = \{z_1, z_2, -\bar{z}_1, -\bar{z}_2\} : z_1, z_2 \in \mathbb{C}\}$ con la suma y el producto definidos como en el ejercicio 1e del epígrafe 4.2. ¿Es $\bar{Q}$ con dichas operaciones un dominio de integridad? ¿Es un cuerpo?

3. Busque las unidades en $M_2(\mathbb{Z}_2)$.
4. Considere el anillo del ejercicio 11a del epígrafe 4.2. Busque las unidades de dicho anillo. ¿Es este anillo un dominio de integridad? ¿Es un cuerpo?
5. Sea el anillo $C[0,1] = \{f: [0,1] \rightarrow \mathbb{R} \mid f \text{ continua}\}$ con las operaciones usuales de suma y producto de funciones. Pruebe que $f \in C[0,1]$ es divisor de cero si y solo si $\{x: f(x)=0\}$ contiene un intervalo abierto. Busque las unidades de este anillo.
6. Pruebe que un anillo A tal que para cada $a \neq 0$ y $b \in A$, existe $r \in A$ tal que $ar=b$ es un anillo de división.
7. Se llama *subdominio de un dominio de integridad* D a un subanillo de D que sea dominio de integridad para las mismas operaciones de D . Pruebe que la intersección de una familia de subdominios de D es un subdominio de D .
8. Pruebe que un anillo finito unitario y sin divisores de cero es un anillo de división.
Sugerencia: Guíese por la demostración de que todo dominio finito D es un cuerpo. Observe que debe demostrar la existencia de inverso a la izquierda y a la derecha.
9. Pruebe que todo anillo conmutativo finito y sin divisores de cero D es un cuerpo.
Sugerencia: Usando el método de trabajo indicado en el ejercicio 8, pruebe la existencia de un elemento unitario en D .
10. Pruebe que en un dominio de integridad $X^2=X \Rightarrow X=0$ o $X=1$ y $X^n=0$ para algún $n \in \mathbb{N} \Rightarrow X=0$.
11. Sea A un anillo y $x \in A$ tal que existe $y \neq 0$ tal que $xy=x=0$. Pruebe que x es un divisor de cero a la izquierda o a la derecha.
12. Pruebe que en un anillo unitario si $1-ab$ es inversible, entonces $1-ba$ también lo es.
13. Pruebe que un subcuerpo K' de un cuerpo K tiene el mismo elemento unitario que K y que $K' \subseteq K$ es un subcuerpo de K si y solo si $a, b \in K' \Rightarrow a-b \in K'$ y $a \cdot b \in K'$, $b \neq 0 \Rightarrow ab^{-1} \in K'$.
14. Sea A un anillo unitario y u un elemento de A que tiene un inverso a la derecha. Pruebe que las siguientes condiciones son equivalentes:
 - a) u tiene más de un inverso a la derecha.
 - b) u no es una unidad.
 - c) u es un divisor de cero a la izquierda.
15. En el ejercicio 13 del epígrafe 4.2 definimos los elementos casi inversibles a la izquierda o a la derecha. Pruebe que en un anillo A en el cual

todo elemento menos uno es casi inversible a la derecha, tiene un elemento unitario y es un anillo de división.

16. Sea A un anillo sin divisores de cero y asuma que A contiene elementos a y b no nulos tales que $ab+kb=0$ para algún entero positivo k . Pruebe que $ca+kc=0=ac+kc$ para todo $c \in A$.
17. Pruebe que $K = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{Z}_8 \right\}$ es un cuerpo de nueve elementos y que el grupo multiplicativo del cuerpo K es cíclico de ocho elementos.

4.4 Extensión de un dominio de integridad a un cuerpo

Hemos visto que el anillo $\mathbb{Z}$ de los números enteros es un dominio de integridad, pero no un cuerpo. Sin embargo, es conocido que a partir de $\mathbb{Z}$ puede construirse $\mathbb{Q}$, que es un cuerpo y que "incluye" a $\mathbb{Z}$ en el sentido de que contiene un subanillo isomorfo a $\mathbb{Z}$.

El objetivo del presente epígrafe es analizar si esta propiedad de $\mathbb{Z}$ puede ser generalizada a un dominio de integridad cualquiera, esto es, si dado un dominio D podemos encontrar un cuerpo K que contenga un subanillo isomorfo a D .

Para comprender mejor el método a seguir, estudiaremos la relación entre un subanillo D de un cuerpo F y un subcuerpo K de F generado por D . ¿Cuáles son los elementos de K ? K deberá contener, dados $a, b \in D$ y $b \neq 0$, los elementos de la forma ab^{-1} .

Es fácil verificar que $\{ab^{-1} : a, b \in D, b \neq 0\}$ es un subcuerpo de F . De aquí que

$$K = \{ab^{-1} : a, b \in D, b \neq 0\}.$$

Ahora bien, es posible que ocurra que $ab^{-1} = cd^{-1}$, con lo cual tendríamos elementos repetidos, pero esto ocurre si y solo si $ad=bc$.

Volvamos al problema planteado, esto es, dado un dominio D encontrar un cuerpo K que contenga un subanillo isomorfo a D .

Lo que acabamos de explicar nos indica que si esto puede hacerse, los elementos de K deben obtenerse de los pares (a, b) , $a, b \in D$ y $b \neq 0$, y que existen pares distintos que dan por resultado el mismo elemento. De aquí que la posible no unicidad de la representación de un elemento de K conduzca a la idea de relación de equivalencia.

De todo lo anteriormente expuesto se obtiene el resultado siguiente:

TEOREMA 4.4.1

Para todo dominio de integridad D puede ser construido un cuerpo K que contenga un subanillo isomorfo a D .

(Se dice, entonces, que K es una *extensión* de D).

Demostración

Sea, pues, $D \times D^* = \{(a, b) : a, b \in D, b \neq 0\}$ y definamos en $D \times D^*$ la relación

$$(a, b) R (a', b') \iff ab' = ba'$$

Si consideramos que un número racional está dado por un par de enteros, lo anterior significa que dos pares distintos, por ejemplo (1,2) y (2,4) pueden representar el mismo número racional.

El estudiante puede comprobar fácilmente que la relación dada es una relación de equivalencia definida sobre $D \times D^*$ y que, por tanto, clasifica a los pares de elementos de $D \times D^*$ en clases, lo cual permite tomar un representante para cada clase y trabajar con este.

Tomemos ahora $K = D \times D^* / R$ esto es, el conjunto cociente de $D \times D^*$ por la relación de equivalencia R . La clase de (a, b) se representa por (a, b) .

Sobre la base de las operaciones de D podemos definir en $D \times D^* / R$ las operaciones siguientes:

$$\text{Suma: } (a, b) + (c, d) = (ad + bc, bd)$$

Notemos que si expresamos (a, b) en la forma $\frac{a}{b}$, la regla anterior se convertiría en

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$

que es la conocida regla de adición de fracciones.

$$\text{Producto: } (a, b) \cdot (c, d) = (ac, bd),$$

el cual en la notación de fracciones sería

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Observemos que si $b, d \in D^*$, entonces $bd \in D^*$ por ser D íntegro.

Este par de operaciones cumplen las siguientes propiedades que convierten a $(K, +, \cdot)$ en un cuerpo.

- (1) La suma es asociativa.
- (2) $(0, 1)$ es el elemento neutro para la suma.
- (3) $(-a, 0)$ es el opuesto de $(a, 0)$.
- (4) La suma es conmutativa.
- (5) El producto es asociativo.
- (6) El producto es distributivo con respecto a la suma.

(7) $(1, 1)$ es el elemento unitario (neutro para el producto).

(8) Si $(a, b) \neq (0, 1)$, entonces (b, a) es el inverso multiplicativo de (a, b) .

Para comprobar que se cumplen estas propiedades y que las operaciones definidas son verdaderas leyes internas, se utilizan técnicas bien conocidas. Por ejemplo, la suma está bien definida, pues si $(a, b) = (a', b')$ y $(c, d) = (c', d')$, se tiene que

$$ab' = ba' \text{ y } cd' = dc' \implies ab'dd' = ba'dd' \text{ y } cd'bb' = dc'bb',$$

de donde $ab'dd' + cd'bb' = ba'dd' + dc'bb'$.

Luego

$$(ad + bc) b'd' = (a'd' + b'c') bd \text{ y}$$

$$(ad + bc, bd) = (a'd' + b'c', b'd').$$

Con esto comprobamos que el resultado es independiente del representante de la clase que tomemos.

Dejamos al estudiante que verifique que el producto está bien definido, así como las propiedades (1) a (7).

Veamos la propiedad (8).

$$(a, b) \cdot (b, a) = (ab, ba) = (1, 1), \text{ pues } ab = ba \text{ ya que } D \text{ es conmutativo.}$$

Luego hemos encontrado un cuerpo K . Queda probar que efectivamente dicho cuerpo K contiene un subanillo isomorfo a D ; para ello demostraremos el resultado siguiente:

TEOREMA 4.4.2

El subconjunto D' de K dado por $D' = \{(a, 1) : a \in D\}$ es un subanillo de K y la aplicación $i : D \rightarrow D'$ dada por $i(a) = (a, 1)$ es un isomorfismo de anillos.

Demostración

D' es un subanillo de K pues $(a, 1) - (b, 1) = (a - b, 1)$ y $(a, 1) \cdot (b, 1) = (ab, 1)$.

Además,

$$i(a + b) = (a + b, 1) = (a, 1) + (b, 1) = i(a) + i(b)$$

$$i(a \cdot b) = (a \cdot b, 1) = (a, 1) \cdot (b, 1) = i(a) \cdot i(b)$$

i es una aplicación sobreyectiva, pues por definición $i(D) = D'$. Probemos ahora que i es inyectiva.

Sea $i(a) = i(b)$; entonces $(a, 1) = (b, 1)$. Esto quiere decir que $(a, 1) R (b, 1)$, esto es, $a \cdot 1 = b \cdot 1$, de donde $a = b$.

Vimos que D' es un subanillo de K ; probemos que no solo es un subanillo, sino que es un dominio de integridad.

$i(1) = (1, 1)$ es el elemento unitario de K y pertenece a $i(D)$, y como

$$(a, 1) \cdot (b, 1) = (0, 1) \text{ tenemos que}$$

$$(ab, 1) = (0, 1).$$

Luego $ab=0$, lo cual implica que $a=0$ o $b=0$ por ser D dominio y, por tanto, $(a,1)=(0,1)$ o $(b,1)=(0,1)$.

Hemos demostrado, pues, que K contiene un subanillo, D' , que es un dominio de integridad e isomorfo a D .

En general se cumple que la imagen por un isomorfismo de anillos de un dominio de integridad es un dominio de integridad. Dejamos la demostración como ejercicio al estudiante.

El teorema 4.4.2 permite representar $(a,b) \in K$ de la forma siguiente:

$$(a,b) = (a,1) (1,b) = (a,1) [(b,1)]^{-1} = i(a)/i(b).$$

Si identificamos $i(a)$ con a , $i(b)$ con b , podemos decir que todo elemento de K puede ser representado mediante un cociente de elementos de D , por lo que al cuerpo K se le denomina *cuerpo de cocientes de D* .

Si tomamos $D=\mathbb{Z}$, el método anterior nos da la manera formal de construir $\mathbb{Q}$ a partir de $\mathbb{Z}$. Cada número racional será una clase de equivalencia de pares de enteros.

Por las consideraciones que hicimos al principio del epígrafe, está claro que el cuerpo de cocientes de D es el "menor" cuerpo que contiene a D , ya que si un cuerpo K' contiene a D , debe contener a los elementos de la forma $\frac{a}{b}$ con $a,b \in D$, $b \neq 0$, y de aquí que deba contener a un subcuerpo isomorfo al cuerpo de cocientes.

El siguiente resultado generaliza la idea anterior y nos permite llegar a un resultado de unicidad del cuerpo de cociente.

TEOREMA 4.4.3

Sean D y D' dominios de integridad y $\phi: D \rightarrow D'$ un isomorfismo y sean K el cuerpo de cocientes de D y K' un cuerpo cualquiera que contenga a D . Entonces existe un isomorfismo φ de K en un subcuerpo de K' tal que $\varphi(a) = \phi(a) \quad \forall a \in D$.

(Se dice en este caso que ϕ es extendido a φ .)

Demostración

Definimos φ de la forma más natural. Ya vimos que todo elemento de un cuerpo de cocientes de un dominio de integridad puede ser representado por un cociente de elementos de este último. De aquí que si $\frac{a}{b} \in K$, definamos

$$\varphi\left(\frac{a}{b}\right) = \frac{\phi(a)}{\phi(b)}$$

Debemos probar que φ está bien definida y que es un isomorfismo de K en K' .

1) φ está bien definida, pues al ser $b \neq 0$ y ϕ un isomorfismo, $\phi(b) \neq 0$; luc-

go $\frac{\phi(a)}{\phi(b)}$ tiene sentido. Además, si $\frac{a}{b} = \frac{a'}{b'}$, se tiene $ab' = ba'$, de donde

$$\phi(a)\phi(b') = \phi(b)\phi(a') \quad \text{y} \quad \frac{\phi(a)}{\phi(b)} = \frac{\phi(a')}{\phi(b')}.$$

2) $\varphi(x+y) = \varphi(x) + \varphi(y)$ y $\varphi(xy) = \varphi(x)\varphi(y) \quad \forall x,y \in K$ pueden ser comprobadas fácilmente a partir de la definición de φ y del hecho de que ϕ es un isomorfismo. Por ejemplo, si $x = \frac{a}{b}$, $y = \frac{c}{d}$, tenemos

$$\begin{aligned} \varphi(xy) &= \varphi\left(\frac{a}{b} \cdot \frac{c}{d}\right) = \varphi\left(\frac{ac}{bd}\right) = \frac{\phi(ac)}{\phi(bd)} = \\ &= \frac{\phi(a)\phi(c)}{\phi(b)\phi(d)} = \frac{\phi(a)}{\phi(b)} \cdot \frac{\phi(c)}{\phi(d)} = \\ &= \varphi\left(\frac{a}{b}\right) \varphi\left(\frac{c}{d}\right) = \varphi(x) \varphi(y). \end{aligned}$$

3) φ es inyectiva, pues si $\varphi\left(\frac{a}{b}\right) = \varphi\left(\frac{c}{d}\right)$, tenemos $\frac{\phi(a)}{\phi(b)} = \frac{\phi(c)}{\phi(d)}$, de donde $\phi(a)\phi(d) = \phi(c)\phi(b)$. Luego $\phi(ad) = \phi(cb)$, y por tanto, $ad = cb$, lo que implica $\frac{a}{b} = \frac{c}{d}$.

Finalmente, si $a \in D$, entonces a se identifica con el elemento $\frac{a}{1}$ de K . Luego podemos escribir $\varphi(a) = \varphi\left(\frac{a}{1}\right) = \phi(a)$, ya que $\phi(1)$ es el elemento unitario de K' por ser ϕ un isomorfismo.

El resultado anterior tiene los siguientes importantes corolarios de los cuales haremos uso más adelante.

COROLARIO 1

Todo cuerpo K' que contenga a un dominio de integridad D , contiene al cuerpo de cocientes de D .

Demostración

Para demostrarlo tomamos en el teorema $D=D'$ y $\phi = \text{id}_D$.

COROLARIO 2

Dos cuerpos de cocientes de un dominio de integridad son isomorfos entre sí. (Esto permite afirmar que el cuerpo de cocientes de un dominio de integridad es único salvo isomorfismo.)

Demostración

Tomemos en el teorema K' como el cuerpo de cocientes de D' . Entonces todo elemento de K' puede ser expresado como $\frac{a'}{b'}$ con $a', b' \in D'$, $b' \neq 0$.

El hecho de que ϕ sea un isomorfismo permite afirmar que existen $a, b \in D$ tales que $a' = \phi(a)$, $b' = \phi(b)$. De aquí que $\frac{a'}{b'} = \frac{\phi(a)}{\phi(b)} = \phi\left(\frac{a}{b}\right)$ y, por tanto, ϕ es sobreyectiva.

Tomando $D = D'$, suponiendo $\phi = \text{id}_D$, y K y K' dos cuerpos de cocientes de D , obtenemos el resultado deseado.

Ejercicios

1. a) Describa el cuerpo de cocientes del subdominio de $\mathbb{Q}$ $\mathbb{Z}(i) = \{n + mi : n, m \in \mathbb{Z}\}$, esto es, indique la forma de los elementos de dicho cuerpo de cocientes.

b) Haga lo mismo para el dominio $D = \{m + n\sqrt{2} : m, n \in \mathbb{Z}\}$.

2. Dé un ejemplo de un cuerpo de cocientes F' de un subdominio propio D' de un dominio de integridad D , que a la vez sea el cuerpo de cocientes de D .

3. Pruebe que en un dominio de integridad D , si $a, b \in D$ tales que $a^m = b^m$ y $a^n = b^n$ para $m, n \in \mathbb{N}$ (m, n primos relativos), entonces $a = b$.

4. Sea D un dominio de integridad que posea la propiedad siguiente:

Dados $a, b \in D$, existe $m \neq 0$, $m \in D$ y $a_1, b_1 \in D : m = ab_1 = ba_1$.

a) Defina en $D \times D^*$ la relación R dada por

$(a, b) R (c, d) \iff (\text{para } b_1 \neq 0 \text{ y } d_1 \neq 0 \text{ tales que } bd_1 = db_1 \text{ se tiene } ad_1 = cb_1)$.

Pruebe que esta definición es independiente de la selección de b_1 y d_1 , y que R es una relación de equivalencia en $D \times D^*$.

b) Sea $F = D \times D^* / R$ y llamemos $\frac{a}{b}$ a la clase de equivalencia de (a, b) .

Pruebe que F es un anillo de división para las operaciones:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad_1 + cb_1}{m}, \text{ donde } m = bd_1 = db_1 \neq 0$$

$$0 = \frac{0}{1}, 1 = \frac{1}{1}$$

$$\left(\frac{a}{b}\right) \left(\frac{c}{d}\right) = \frac{ac_1}{db_1}, \text{ donde } b_1 \neq 0 \text{ y } cb_1 = bc_1$$

Pruebe que la aplicación $a \mapsto \frac{a}{1}$ es un isomorfismo de D en F y que

$$F = \left\{ \left(\frac{a}{1}\right) \left(\frac{b}{1}\right)^{-1} : a, b \in D, b \neq 0 \right\}$$

4.5 Ideales

Nos proponemos ahora realizar con los anillos un estudio similar al que hicimos en los grupos, relativo a los homomorfismos y la estructura cociente.

En el caso del anillo cociente, como todo anillo A es un grupo abeliano para la suma, la parte aditiva de la teoría que desarrollaremos ya la conocemos del capítulo 3, así que trataremos solamente el aspecto multiplicativo. Seguiremos el método de razonamiento que usamos en los grupos, o sea, la búsqueda de las relaciones de equivalencia definidas sobre A que sean compatibles con las operaciones del anillo, las cuales nos permitirán definir una estructura de anillo en el cociente.

Como señalamos anteriormente, en este análisis ya tenemos la ventaja de conocer que para que A/R sea un grupo aditivo, es condición necesaria y suficiente que R sea de la forma $xRy \iff x - y \in I$, donde I es un subgrupo cualquiera de A . (Al ser $(A, +)$ un grupo abeliano, todo subgrupo es invariante, de aquí que no tengamos que exigir a I la condición de invariante.)

Así, partiremos de relaciones R del tipo $xRy \iff x - y \in I$, donde I es un subgrupo de $(A, +)$, y analizaremos cuáles de estas relaciones son, además, compatibles con el producto del anillo.

Dada la relación R que cumple $xRy \iff x - y \in I$, donde I es un subgrupo de $(A, +)$, R debe cumplir, además, que

$$aRb \iff \begin{cases} axRbx \text{ para todo } x \in A \\ xaRxb \end{cases}$$

(compatibilidad con el producto del anillo).

La propiedad anterior no la cumple cualquier subgrupo de A , ni siquiera cualquier subanillo. Por ejemplo, en $M_2(\mathbb{Z})$ el conjunto

$$A' = \left\{ \begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} : \alpha, \beta, \delta \in \mathbb{Z} \right\} \text{ es un subgrupo de } (M_2(\mathbb{Z}), +);$$

sin embargo,

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \alpha & \beta + \delta \end{pmatrix} \notin A',$$

aunque A' es un subanillo de A , pues

$$\begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ 0 & \delta' \end{pmatrix} = \begin{pmatrix} \alpha\alpha' & \alpha\beta' + \beta\delta' \\ 0 & \delta\delta' \end{pmatrix} \in A'.$$

De este ejemplo podemos sacar dos conclusiones:

No todos los subgrupos I de un anillo dan lugar a relaciones del tipo $xRy \iff x - y \in I$ tales que R sea compatible con el producto del anillo.

Dentro de los subgrupos podemos suponer que aquellos que son a la vez subanillos, serán los que resolverán el problema de la compatibilidad para el producto, pero el ejemplo anterior contradice esta suposición.

Por tanto, seguimos en la búsqueda de cierto tipo de subgrupos I del grupo aditivo del anillo A tales, que la relación $xRy \Leftrightarrow x-y \in I$ sea compatible con el producto del anillo.

Recordemos una vez más la condición de compatibilidad con el producto para R .

R es compatible con el producto en A si y solo si

$$aRb \Leftrightarrow \begin{cases} xaRxb \Leftrightarrow x(a-b) \in I \\ axRxb \Leftrightarrow (a-b)x \in I \end{cases} \text{ para todo } x \in A.$$

La condición anterior nos sugiere la idea de considerar subgrupos I de $(A, +)$ que cumplan la condición adicional

$$x \in A, a \in I \Rightarrow \begin{cases} xa \in I \\ ax \in I \end{cases}$$

Dichos subgrupos reciben el nombre de *ideales*, y daremos su definición inmediatamente.

DEFINICIÓN 4.5.1

Se dice que $I \subset A$, $I \neq \emptyset$, es un *ideal* de A si y sólo si I es un subgrupo de $(A, +)$ y si para todo $x \in A$ y todo $a \in I$ se tiene que $ax \in I$ y $xa \in I$.

(Como se observa en la definición, el ideal es un tipo especial de subgrupo de $(A, +)$ y de subanillo de A).

Si I es un subgrupo de $(A, +)$ y para todo $x \in A$ y todo $a \in I$ se tiene que $xa \in I$, decimos que I es un *ideal izquierdo* de A . Si I es un subgrupo de $(A, +)$ y para todo $x \in A$ y todo $a \in I$ se tiene que $ax \in I$, decimos que I es un *ideal derecho* de A . I es un ideal de A si y sólo si es a la vez ideal izquierdo y derecho. Nosotros trabajaremos con ideales que a veces se llaman *biláteros* para indicar que son izquierdos y derechos, aunque daremos ejemplos de ideales izquierdos o derechos solamente.

Si el anillo es conmutativo, todo ideal izquierdo es a la vez derecho.

El origen del concepto de *ideal* lo encontramos en los trabajos de Kummer, quien tratando de demostrar el famoso último teorema de Fermat —este teorema establece que para todo entero $n > 2$ no existen valores enteros positivos de x, y, z tales que $x^n + y^n = z^n$ — pudo probar que se cumple el teorema para una gran cantidad de exponentes, pero no logró dar una prueba general. La dificultad que encontró Kummer para probar este teorema radicaba en el hecho de que, al factorar $x^n + y^n$ mediante la solución de $x^n + y^n = 0$ para x en términos de y , los enteros algebraicos (raíces de la ecuación) no satisfacen necesariamente el teorema de factorización única. Sin embargo aunque Kummer no pudo demostrar el último teorema de Fermat, los resultados obtenidos dieron origen a la teoría de ideales, y con dicha teoría y mediante

la construcción del anillo cociente se pudo resolver el problema de la no unicidad de la factorización en un anillo dado.

Esto no nos debe sorprender pues, como vimos en el capítulo 3, la introducción del grupo cociente permite resolver problemas similares. Es por ello que introducimos el concepto de *ideal* en este momento del curso y así no nos alejamos del origen histórico de dicho concepto.

Ejemplos

1. Son ejemplos triviales de ideales en un anillo A , el $\{0\}$ y el propio A .
2. En $\mathbb{Z}$, todos los $n\mathbb{Z}$ son ideales de $\mathbb{Z}$ ya que $m(nx) = n(mx) \in n\mathbb{Z}$ y $(nx)m = n(mx) \in n\mathbb{Z}$.
3. En $\mathbb{Z}_6$, cuya tabla para el producto es la siguiente:

	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

es evidente que $I_1 = \{0, 3\}$, $I_2 = \{0, 2, 4\}$ son ideales de $\mathbb{Z}_6$. Para comprobarlo deben buscarse los productos de cada elemento del ideal por todos los elementos del anillo en la tabla de multiplicar.

4. $\mathbb{Z} \times \mathbb{Z}$ es un anillo para la suma y el producto componente a componente, definidos por

$$(n, m) + (n', m') = (n+n', m'+m')$$

$$(n, m) \cdot (n', m') = (nn', nm').$$

El conjunto $I = \{(0, n) : n \in \mathbb{Z}\}$ es un ideal de $\mathbb{Z} \times \mathbb{Z}$, mientras que $A' = \{(n, n) : n \in \mathbb{Z}\}$ es un subanillo de $\mathbb{Z} \times \mathbb{Z}$, pero no un ideal, ya que $(a, b) \cdot (n, n) = (na, nb)$ no pertenece a A' para $a \neq b$.

5. En $M_2(\mathbb{R})$ si consideramos

$$I = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\}.$$

I es un ideal derecho, pero no izquierdo de $M_2(\mathbb{R})$.

6. En el anillo $R[X]$ de los polinomios con coeficientes reales,

$I = \{p(x) \cdot q(x) \mid q(x) \text{ es un polinomio fijo y } p(x) \text{ recorre } R[X]\}$ es un ideal de $R[X]$.

7. Sean A un anillo cualquiera y $a \in A$. Consideremos los conjuntos

$$aA = \{ax : x \in A\}$$

$$Aa = \{xa : x \in A\}$$

Entonces aA es un ideal derecho y Aa es un ideal izquierdo de A . Analicemos $aA = \{ax : x \in A\}$.

$ax - ax' = a(x - x') \in aA$. Luego aA es un subgrupo de $(A, +)$.

$(ax)x' = a(xx') \in aA$ para todo $x' \in A$, es un ideal derecho.

Similarmente probamos que Aa es un ideal izquierdo.

8. En el anillo de las funciones acotadas definidas sobre $[0,1]$, que denotaremos por M , esto es,

$$M = \{f : [0,1] \rightarrow \mathbb{R} : f \text{ acotada}\},$$

M es un subanillo de $\mathbb{R}^{[0,1]}$, por tanto, un anillo para la suma y el producto usuales de funciones.

El conjunto M' de las funciones de M que tienen límite igual a cero en el punto $\frac{1}{2}$, constituye un ideal de M .

En efecto,

$$\text{si } f, g \in M, \lim_{x \rightarrow \frac{1}{2}} (f-g)(x) = 0$$

$$\text{y si } h \in M \text{ y } f \in M', \lim_{x \rightarrow \frac{1}{2}} (h \cdot f)(x) = 0, \text{ ya que } h \text{ está acotada en } [0,1].$$

Ideal generado por un subconjunto X del anillo A

Igual que hicimos en los casos de subgrupo y subanillo, podemos considerar el ideal I de un anillo A generado por un subconjunto X de A , el cual desde el punto de vista teórico será el menor ideal de A que contiene a X , o lo que es equivalente, la intersección de todos los ideales de A que contienen a X .

En el capítulo 3 vimos que era complicado resolver este problema de forma explícita para un conjunto X cualquiera, por lo que nos vamos a limitar al caso $X = \{a\}$.

Como A es un anillo no necesariamente conmutativo, separaremos el análisis en la búsqueda del ideal derecho y la del ideal izquierdo generado por a . Buscaremos primeramente el menor ideal derecho I_a de A tal que

Al ser I_a subgrupo, debe contener los múltiplos de a , es decir, los elementos del tipo na con $n \in \mathbb{Z}$ y también los múltiplos derechos de A , o sea, los ax con $x \in A$. Esto conduce a examinar el conjunto

$$I_a = \{na + ax : n \in \mathbb{Z}, x \in A\}.$$

I_a es un ideal derecho de A pues $(na + ax) - (ma + ax') = (n - m)a + a(x - x') \in I_a$.

Si $x' \in A$, entonces,

$$(na + ax)x' = nax' + axx' = a(nx') + a(xx') = a(nx + xx') \in I_a \text{ ya que}$$

$$nx + xx' \in A.$$

Además, $a \in I_a$ pues $a = 1 \cdot a + a \cdot 0$.

Por último, si I' es un ideal de A y $a \in I'$, es obvio que $I' \supset I_a$ según lo analizado anteriormente.

Podemos afirmar, entonces, que $I_a = \{ax + na : x \in A, n \in \mathbb{Z}\}$ es el menor ideal derecho de A que contiene a a y, por tanto, es el ideal derecho generado por a .

Una búsqueda similar para el ideal izquierdo generado por a llevará a que este es $I_l = \{xa + na : x \in A, n \in \mathbb{Z}\}$.

La búsqueda del ideal generado adquiere particular interés en el caso de los anillos unitarios, en los cuales toma una forma más sencilla.

Si el anillo A es unitario, sea 1 el elemento unitario del anillo A , tenemos

$$I_a = \{xa + na : x \in A, n \in \mathbb{Z}\}.$$

Pero $xa + na = xa + n(1)a = (x + n(1))a = ya$, donde $y = x + n(1) \in A$,

así que podemos escribir $I_a = \{ya : y \in A\} = Aa$.

De forma similar, $I_l = aA$.

Si el anillo A es conmutativo y unitario, el ideal Aa coincidirá con el ideal aA , y podremos decir que el ideal generado por un elemento a será aA o Aa . En este caso al ideal generado por a lo denotamos por $\langle a \rangle$ y lo llamamos ideal principal generado por a . Demos la definición.

DEFINICIÓN 4.5.2

Si A es un anillo conmutativo y unitario y $a \in A$, se llama ideal principal generado por a y se denota por $\langle a \rangle$ al ideal aA , o sea, $\langle a \rangle = aA = \{ax : x \in A\}$.

(De igual modo podríamos definir $\langle a \rangle = Aa$ ya que A es conmutativo.)

Ejemplos

1. Sabemos que en $\mathbb{Z}$ los subgrupos cíclicos $m\mathbb{Z} = \{mx : x \in \mathbb{Z}\}$ son ideales. En este caso $m\mathbb{Z}$ no es más que el ideal principal $\langle m \rangle$ generado por m . Pero como todo ideal de $\mathbb{Z}$ es un subgrupo y los subgrupos de $\mathbb{Z}$ son los $m\mathbb{Z}$, tendremos que en $\mathbb{Z}$ todos los ideales son principales de la forma $m\mathbb{Z} = \langle m \rangle$.

2. Una situación similar se presenta en $R[X]$, donde puede demostrarse que todos los ideales son principales.

En efecto, sea I un ideal de $R[X]$. Si $I = \{0\}$ o $I = R[X]$, es trivial que sean principales ya que $\{0\} = \langle 0 \rangle$ y $R[X] = \langle 1 \rangle$.

Supongamos, pues, que I es distinto de $\{0\}$ y $R[X]$. Entonces existirá en I al menos un polinomio no nulo, y de ahí que podamos deducir la existencia de un polinomio mónico (con coeficiente principal 1) de grado mínimo contenido en I . (Nótese la semejanza de este razonamiento con el seguido en el estudio de los subgrupos de un grupo cíclico.)

Sea $m(x)$ dicho polinomio. Probemos que $I = \langle m(x) \rangle$.

Sea $f(x) \in I$; existen $q(x), r(x)$ tales que

$f(x) = m(x)q(x) + r(x)$ con $\text{gr } r(x) < \text{gr } m(x)$ o $r(x) = 0$.

Entonces

$r(x) = f(x) - m(x)q(x) \in I \Rightarrow r(x) = 0$,

pues si no tendríamos una contradicción. Luego $f(x) = m(x)q(x)$.

Así, $I \subset \langle m(x) \rangle$, y como obviamente $\langle m(x) \rangle \supset I$, resulta que $I = \langle m(x) \rangle$.

Los ejemplos anteriores motivan la definición siguiente:

DEFINICIÓN 4.5.3

Un anillo conmutativo y unitario tal que todos sus ideales son principales, se llama *anillo principal*.

Veamos si un cuerpo K es un anillo principal.

Sean I un ideal de K e $I \neq \{0\}$, $a \in I$.

Como I es un ideal, $a^{-1}a = 1 \in I$ y $\forall x \in K$, $x \cdot 1 = x \in I$. De aquí que $I \supset K$ y, por tanto $I = K$.

Obtuvimos, pues, el importante resultado siguiente:

PROPOSICIÓN 4.5.1

Los ideales de un cuerpo K son $\{0\}$ y K .

Resulta de manera trivial que K es principal, pues $\{0\} = \langle 0 \rangle$ y $K = \langle 1 \rangle$.

El recíproco de la proposición anterior se cumple, esto es, si K es un anillo conmutativo y unitario no nulo cuyos únicos ideales son $\{0\}$ y K , entonces K es un cuerpo, ya que si $a \neq 0$, $a \in K$, tenemos que $\langle a \rangle = K$. De aquí resulta que $1 \in \langle a \rangle$ y entonces debe existir $x \in K$ tal que $ax = 1$ y K es un cuerpo.

Obsérvese en la demostración que en un anillo conmutativo y unitario A , si $1 \in I$, I ideal, entonces $I = A$.

De la proposición y su recíproco obtenemos la siguiente caracterización de los cuerpos.

TEOREMA 4.5.1

Un anillo conmutativo y unitario K es un cuerpo si y solo si sus únicos ideales son $\{0\}$ y K .

El hecho de que un anillo sea principal permite utilizar determinadas técnicas de trabajo y obtener resultados que no podemos obtener en un anillo arbitrario.

En particular, en $\mathbb{Z}$ observamos que la definición clásica de divisibilidad: dados $m, n \in \mathbb{Z}$, se dice que m divide a n si y solo si existe $q \in \mathbb{Z}$ tal que $n = mq$, se puede expresar mediante el lenguaje de los ideales.

En efecto, m divide a n es equivalente a decir que $n\mathbb{Z} \subset m\mathbb{Z}$, y en este lenguaje es posible definir también otros conceptos clásicos, como son: *máximo común divisor*, *mínimo común múltiplo* y *número primo*.

El concepto de divisibilidad que conocemos en $\mathbb{Z}$ puede ser generalizado por la vía de los ideales a un anillo cualquiera y tiene especial interés en los anillos principales.

Ejercicios

1. Halle todos los ideales de $\mathbb{Z}_{12}$. ¿Coinciden los ideales de $\mathbb{Z}_n$ con los subgrupos de $\mathbb{Z}_n$?

2. Analice si los siguientes subgrupos son ideales de $\mathbb{Z}(i)$.

a) $I = \{2x : x \in \mathbb{Z}(i)\}$, $J = \{5x : x \in \mathbb{Z}(i)\}$, $K = \{(1+i)x : x \in \mathbb{Z}(i)\}$.

3. a) Determine todos los ideales de $\mathbb{Z} \times \mathbb{Z}$.

b) Determine los ideales del anillo $T = \left\{ \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} : a, b, c \in \mathbb{Z} \right\}$ con la suma y el producto de matrices.

4. Sea A un anillo conmutativo y $a \in A$. Pruebe que $Ia = \{x \in A : ax = 0\}$ es un ideal de A .

5. Demuestre que el conjunto de los elementos nilpotentes de un anillo conmutativo A ($a \in A$ es nilpotente si y solo si existe $n \in \mathbb{Z}$ tal que $a^n = 0$) es un ideal de A , llamado *radical de A* . Encuentre los radicales de $\mathbb{Z}_{12}$, $\mathbb{Z}$ y $\mathbb{Z}_{12}$.

6. Muestre que $M_2(\mathbb{Z})$ no tiene ideales propios.

7. Sean A un anillo conmutativo e I un ideal de A .

a) Pruebe que el conjunto $\sqrt{I} = \{a \in A : a^n \in I \text{ para algún } n \in \mathbb{Z}^+\}$ es un ideal de A , llamado *radical de I* . Si $A = \mathbb{Z}$, halle $\sqrt{2\mathbb{Z}}$, $\sqrt{9\mathbb{Z}}$, $\sqrt{12\mathbb{Z}}$.

¿Tiene relación este concepto con el de radical de un anillo definido en el ejercicio anterior?

b) Si I, J son ideales de A , pruebe que $\sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$ y $\sqrt{\sqrt{I}} = \sqrt{I}$.

8. Sean I, J ideales de un anillo A . Definimos

$I + J = \{a + b : a \in I, b \in J\}$.

a) Pruebe que $I + J$ es un ideal.

b) Pruebe que $I \subset I+J$, $J \subset I+J$.

c) Pruebe que $\sqrt{I+J} = \sqrt{\sqrt{I} + \sqrt{J}}$.

9. Sean I, J ideales de un anillo A . Definimos

$$I \cdot J = \left\{ \sum_{i=1}^n a_i b_i : a_i \in I, b_i \in J, n \in \mathbb{Z}_+ \right\}.$$

a) Pruebe que $I \cdot J$ es un ideal de A .

b) Pruebe que $I \cdot J \subset I \cap J$.

c) Pruebe que $I \cdot (J \cdot K) = (I \cdot J) \cdot K$ para I, J, K ideales cualesquiera de A .

10. Sean I, J ideales de un anillo conmutativo A . Definimos

$$I : J = \{a \in A : ab \in I \text{ para todo } b \in J\}.$$

Pruebe que $I : J$ es un ideal de A .

11. a) Pruebe que en un anillo de división A los únicos ideales son $\{0\}$ y A .

b) Sea A un anillo unitario tal que sus únicos ideales derechos son $\{0\}$ y A . Pruebe que A es un anillo de división.

12. Sea A un anillo tal que sus únicos ideales son $\{0\}$ y A . Pruebe que A es un anillo de división o A es un anillo con un número primo de elementos en el que $ab=0 \forall a, b \in A$.

13. Pruebe que en un anillo conmutativo A el ideal generado por un subconjunto S de A está dado por

$$\left\{ \sum_{i=1}^k n_i a_i + \sum_{j=1}^l m_j a_j : a_i \in S; n_i \in \mathbb{Z}; A m_j \right\}.$$

14. Sean I un ideal de un anillo unitario A , U el grupo de las unidades de A y $U_1 = \{a \in U : a-1 \in I\}$. Pruebe que U_1 es un subgrupo normal de U .

15. En $\mathbb{Z}$ sean $m, n \in \mathbb{Z}$. Pruebe que $\langle m, n \rangle = \langle d \rangle$ donde $d = \text{mcd}(m, n)$ y que $\langle m \rangle \cap \langle n \rangle = \langle \text{mcm}(m, n) \rangle$.

16. Sea I un ideal de un anillo A y sea $M_n(I) = \{B \in M_n(A) : \text{los coeficientes de } B \text{ pertenecen a } I\}$. Pruebe que $M_n(I)$ es un ideal en $M_n(A)$.

17. Sea $T = \left\{ \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \text{ con } a, b, c \in \mathbb{Z} \right\}$. Verifique que T es un subanillo de $M_2(\mathbb{Z})$. Determine los ideales de T .

4.6 Anillo cociente

Profundizando en el estudio de los ideales nos hemos apartado de nuestro objetivo inmediato: la construcción del anillo cociente. Volvamos a dicho problema.

Anillo cociente por un ideal

Hemos establecido la posibilidad de que las relaciones de equivalencia compatibles con ambas leyes del anillo A sean de la forma

$$xRy \iff x-y \in I \text{ donde } I \text{ es un ideal de } A.$$

Probemos esta afirmación.

1) Las relaciones del tipo anterior son compatibles con las leyes del anillo. Al ser I un ideal, es un subgrupo, y como ya hemos señalado no tenemos que ocuparnos de la parte aditiva: veamos, pues, la multiplicativa. Debemos comprobar

$$aRb \implies \begin{cases} xaRxb \\ axRbx \end{cases} \text{ para todo } x \in A.$$

Pero $aRb \iff a-b \in I \implies x(a-b) \in I$ y $(a-b)x \in I$ para todo $x \in A$. De aquí que

$$x(a-b) = xa - xb \in I \iff xaRxb$$

para todo $x \in A$.

$$(a-b)x = ax - bx \in I \iff axRbx.$$

Con lo que queda demostrada la compatibilidad de R con

2) Recíprocamente, si R es compatible con $+$ y $\cdot$, debemos comprobar que existe un ideal I de A tal que $xRy \iff x-y \in I$.

Del capítulo de grupos ya sabemos que si R es compatible con $+$, entonces $C0$ es un subgrupo de A y $xRy \iff x-y \in C0$.

Si comprobamos ahora que para R compatible con $+$ y $\cdot$, $C0$ es no solo un subgrupo, sino un ideal de A , ya quedaría demostrada la existencia de un ideal I tal que $xRy \iff x-y \in I$ para toda R compatible con $+$ y $\cdot$. Probemos, pues, que $C0$ es un ideal de A . Queda solo probar

$$(a \in C0, x \in A) \implies \begin{cases} xa \in C0 \\ ax \in C0 \end{cases}$$

Pero

$$a \in C0 \iff aR0 \implies \begin{cases} xaR0 \iff xa \in C0 \\ axR0 \iff ax \in C0 \end{cases}$$

para todo $x \in A$

por la compatibilidad de R con $\cdot$. Resulta, entonces, que $C0$ es un ideal de A .

Resumiendo podemos decir:

Las relaciones de equivalencia compatibles con las leyes del anillo son exactamente las de la forma $xRy \iff x-y \in I$ donde I es un ideal de A .

Dada una relación R del tipo anterior, podemos definir en A/R las operaciones:

$$+ \text{ dada por } Cx + Cy = Cx + y$$

$$\cdot \text{ dada por } Cx \cdot Cy = Cx \cdot y$$

que convierten a A/R en un anillo.

(Como ejercicio el estudiante puede comprobar que efectivamente $(A/R, +, \cdot)$ es un anillo guiándose por la demostración hecha en el caso del grupo cociente.)

Si A es conmutativo, A/R será conmutativo; si A es unitario, A/R será unitario.

Como en definitiva la relación R , compatible con las leyes de A , está determinada por un cierto ideal I de A , usaremos, al igual que en los grupos, la notación A/I que indicará el anillo cociente dado por la relación de equivalencia $xRy \Leftrightarrow x - y \in I$ y la suma y el producto de clases definidas anteriormente. A dicho anillo cociente lo llamaremos *anillo cociente de A por el ideal I* y lo denotaremos por A/I .

Ejemplos

1. En $\mathbb{Z}$ teníamos los grupos cocientes $\mathbb{Z}/n\mathbb{Z}$ dados por las relaciones $mRm' \Leftrightarrow m - m' \in n\mathbb{Z}$. Estos grupos también son anillos cocientes, ya que $n\mathbb{Z} = \langle n \rangle$ es el ideal principal generado por n y $\mathbb{Z}/n\mathbb{Z}$ será ahora el anillo cociente de $\mathbb{Z}$ por el ideal principal $n\mathbb{Z}$. Aquí encontramos de nuevo el anillo $\mathbb{Z}_n$ que ya habíamos definido de otra forma en el epígrafe 4.2.

2. En $\mathbb{Z}_4$ vimos que $I = \{0, 2\}$ es un ideal.

$\mathbb{Z}_4/I$ tiene tres elementos. $\mathbb{Z}_4/I = \{0 + I, 1 + I, 2 + I\}$.

Nótese que las clases están determinadas por la estructura de grupo de $\mathbb{Z}_4/I$.

Podemos demostrar que $\mathbb{Z}_4/I \cong \mathbb{Z}_3$. Ya sabemos que como grupos abelianos son isomorfos; queda pues verificar lo correspondiente al producto. La tabla de multiplicar de $\mathbb{Z}_4/I$ es:

	$0 + I$	$1 + I$	$2 + I$
$0 + I$	$0 + I$	$0 + I$	$0 + I$
$1 + I$	$0 + I$	$1 + I$	$2 + I$
$2 + I$	$0 + I$	$2 + I$	$1 + I$

donde $0 + I = 0$, $1 + I = 1$, $2 + I = 2$.

Nótese que esta tabla se construye siguiendo la definición de $\cdot$ en $\mathbb{Z}_4/I$ y la clasificación de los elementos de $\mathbb{Z}_4$ en dicho cociente, y que coincide estructuralmente con la tabla de multiplicar de $\mathbb{Z}_3$.

En este caso vemos que cada elemento no nulo de $\mathbb{Z}_4/I$ tiene un inverso, mientras que en la tabla de $\mathbb{Z}_4$ podemos observar que existen elementos no nulos sin inverso, por ejemplo, 2, 3, 4. Luego el cociente ha "mejorado" la estructura del anillo, en el sentido de que hemos obtenido un cuerpo.

3. En $\mathbb{Z} \times \mathbb{Z}$ vimos que $N = \{(0, n) : n \in \mathbb{Z}\}$ es un ideal. Consideremos $\mathbb{Z} \times \mathbb{Z}/N$.

En este caso las clases de equivalencia son de la forma $(m, 0) + N$ con $m \in \mathbb{Z}$, por lo cual $\mathbb{Z} \times \mathbb{Z}/N$ es naturalmente isomorfo a $\mathbb{Z}$ mediante el isomorfismo $(m, 0) + N \mapsto m$.

Sabemos que en $\mathbb{Z}$, $ab = 0 \Rightarrow a = 0$ o $b = 0$; sin embargo, esto no ocurre en $\mathbb{Z} \times \mathbb{Z}$ donde, por ejemplo,

$$(0, 1)(1, 0) = (0, 0).$$

Este es otro ejemplo en el cual el cociente "mejora" la estructura del anillo, ya que el anillo tenía divisores de cero, mientras que el cociente no.

4. De la forma de los ideales de un cuerpo K podemos deducir los posibles cocientes de un cuerpo.

En efecto, si K es un cuerpo sabemos que sus únicos ideales son $\{0\}$ y K . Los posibles cocientes son $K/\{0\}$ y K/K , que evidentemente son isomorfos a K y a $\{0\}$, respectivamente. (Verifíquelo el estudiante.)

Podemos, pues, decir que mediante cocientes no podemos "mejorar" la estructura de un cuerpo.

Ejercicios

- Utilizando el ejercicio 1 del epígrafe 4.5, halle todos los posibles anillos cociente de $\mathbb{Z}_{12}$ buscando en cada caso un anillo al cual dicho cociente sea isomorfo.
- Construya las tablas de sumar y multiplicar de $2\mathbb{Z}/8\mathbb{Z}$. ¿Son $2\mathbb{Z}/8\mathbb{Z}$ y $\mathbb{Z}_4$ anillos isomorfos?
- Halle $\mathbb{Z} \times \mathbb{Z}/I$ donde $I = \{(0, n) : n \in \mathbb{Z}\}$.
- Pruebe que dado un anillo A y un ideal I de A , A/I es conmutativo si y solo si $rs - sr \in I$ para todo $r, s \in A$.
- Con referencia al ejercicio 5 del epígrafe 4.5, pruebe:
 - Si I es radical de un anillo conmutativo A , entonces el radical de A/I es $\{0 + I\}$.
 - Si A es un anillo conmutativo e I un ideal de A ; si todo elemento de I es nilpotente y el radical de A/I es A/I , entonces el radical de A es A .
- Con referencia a los ejercicios 5 y 6 del epígrafe 4.5, dado un ideal I de un anillo conmutativo A , ¿qué relación existe entre $\sqrt{I}$ y el radical de A/I ?

7. Pruebe que $\mathbb{Z}/k\mathbb{Z}$ contiene elementos nilpotentes (esto es, $z: z^n=0$) si y solo si k es divisible por el cuadrado de un número primo.

4.7 Homomorfismos de anillos

Para terminar lo relativo a la estructura cociente nos falta estudiar un último aspecto: la relación del anillo cociente con los homomorfismos de anillos.

Siguiendo la línea trazada en los homomorfismos de grupos, es natural considerar las definiciones siguientes:

DEFINICIÓN 4.7.1

Se dice que dados los anillos A y A' (cuyas leyes denotaremos por $+$ y $\cdot$ en ambos conjuntos), se tiene un homomorfismo f de A en A' si y solo si existe $f: A \rightarrow A'$ tal que

$$\begin{aligned} f(x+y) &= f(x) + f(y) \\ f(x \cdot y) &= f(x) \cdot f(y) \end{aligned} \quad \text{Para todo } x, y \in A.$$

Está claro que si f es biyectiva tenemos la definición de *isomorfismo*.

DEFINICIÓN 4.7.2

Sea $f: A \rightarrow A'$ un homomorfismo de anillos. Se llama *núcleo de f* y se denota por $\text{Ker } f$ a

$$\text{Ker } f = \{x \in A : f(x) = 0\} = f^{-1}(0).$$

Está claro que f es inyectiva si y solo si $\text{Ker } f = \{0\}$. Si pensamos el papel que desempeña el núcleo de los homomorfismos de grupos en la relación entre cociente y homomorfismo, resulta natural la definición anterior en el caso de los anillos.

El resultado dado por la siguiente proposición reafirma nuestra afirmación.

PROPOSICIÓN 4.7.1

El núcleo de un homomorfismo de anillo $f: A \rightarrow A'$ es un ideal del anillo.

Demostración

En efecto, $\text{Ker } f$ es un subgrupo del grupo aditivo del anillo A . Veamos que no es solamente un subgrupo, sino un ideal de A .

Sean $a \in A$ y $x \in \text{Ker } f$, entonces

$$f(xa) = f(x)f(a) = 0 \cdot f(a) = 0,$$

lo cual implica que $xa \in \text{Ker } f$.

De manera similar,

$$f(ax) = f(a) \cdot f(x) = f(a) \cdot 0 = 0,$$

con lo que queda probado que $\text{Ker } f$ es un ideal de A .

Ejemplos

1. La aplicación natural de $\mathbb{Z}$ en $\mathbb{Z}_n$ dada por $\varphi: \mathbb{Z} \rightarrow \mathbb{Z}_n$, donde $\varphi(m) = r$ y r es el resto de la división de m entre n , es un homomorfismo del anillo $\mathbb{Z}$ en el anillo $\mathbb{Z}_n$.

$\text{Ker } \varphi = \{m \in \mathbb{Z} : \text{el resto de la división entre } n \text{ es } 0\} = n\mathbb{Z}$.

2. $f: (\mathbb{Z}, +, \cdot) \rightarrow (2\mathbb{Z}, +, \cdot)$ dada por $x \mapsto 2x$ no es un homomorfismo de anillos, pues

$$f(xy) = 2xy \neq (2x)(2y) = f(x)f(y).$$

3. En el anillo $A = \{m + n\sqrt{2} : m, n \in \mathbb{Z}\}$, la aplicación

$$g: A \rightarrow A \text{ dada por } g(m + n\sqrt{2}) = m - n\sqrt{2}$$

es un homomorfismo de anillos. $\text{Ker } g = \{0\}$

4. Sea el anillo $C[0,1] = \{f: [0,1] \rightarrow \mathbb{R} \text{ continuas}\}$ con la suma y el producto de funciones. La aplicación

$$\varphi: C[0,1] \rightarrow \mathbb{R} \text{ dada por } \varphi(f) = f\left(\frac{1}{2}\right)$$

es un homomorfismo de anillos.

En efecto,

$$\varphi(f+g) = (f+g)\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right) + g\left(\frac{1}{2}\right) = \varphi(f) + \varphi(g)$$

$$\varphi(f \cdot g) = (f \cdot g)\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right) \cdot g\left(\frac{1}{2}\right) = \varphi(f) \cdot \varphi(g).$$

Como vimos anteriormente, $\text{Ker } \varphi = \{f \in C[0,1] : f\left(\frac{1}{2}\right) = 0\}$ es un ideal de $C[0,1]$.

Los homomorfismos de anillo tienen propiedades similares a las de los homomorfismos de grupos.

Sea $f: A \rightarrow A'$ un homomorfismo de anillos. Se cumple:

- (1) Si B es un subanillo de A , entonces $f(B)$ es un subanillo de A' .
- (2) Si I es un ideal de A , entonces $f(I)$ es un ideal de $f(A)$.
- (3) Si B' es un subanillo de A' , entonces $f^{-1}(B')$ es un subanillo de A .
- (4) Si I' es un ideal de $f(A)$, entonces $f^{-1}(I')$ es un ideal de A .
- (5) Si A es unitario, 1 es el elemento unitario y $1 \notin \text{Ker } f$, entonces $f(1)$ es el elemento unitario de $f(A)$.

La demostración de estas propiedades es análoga a la realizada para las propiedades similares en el caso de los grupos, por lo que las dejamos al estudiante como ejercicio. La propiedad del núcleo de un homomorfismo de anillos de ser un ideal del dominio, surge de nuevo como un caso particular de (3) para $B' = \{0\}$.

Por supuesto que al ser un homomorfismo de anillos, en primera instancia, un homomorfismo de grupos, todas las propiedades de homomorfismo de grupos para la adición estudiadas anteriormente, son válidas también en este caso.

Pasemos ahora a estudiar la relación entre los homomorfismos de anillo y los anillos cocientes.

En el caso de los grupos vimos que los grupos cocientes estaban determinados exactamente por los núcleos de homomorfismos de grupo. En los anillos ocurre exactamente lo mismo. Demostraremos ahora que los anillos cocientes están determinados exactamente por los núcleos de homomorfismos de anillos.

Ya sabemos que todo núcleo de homomorfismo de anillos $f: A \rightarrow A'$ es un ideal del anillo A y, por tanto, da lugar a un anillo cociente $A/\text{Ker } f$.

Además veremos que dado un ideal I del anillo A , existe un homomorfismo de anillos φ tal que $\text{Ker } \varphi = I$, este es exactamente el homomorfismo canónico

$$\varphi: A \rightarrow A/I$$

PROPOSICIÓN 4.7.2

Si I es un ideal de un anillo A , entonces la aplicación canónica $\varphi: A \rightarrow A/I$ dada por $\varphi(x) = x + I$ para todo $x \in A$ es un homomorfismo de anillos cuyo núcleo es I .

Demostración

En efecto, sabemos que φ es un homomorfismo de grupos (entre los grupos aditivos de A y A/I) cuyo núcleo es I .

Tomemos $\varphi(xx') = xx' + I = (x + I) \cdot (x' + I) = \varphi(x) \cdot \varphi(x')$, con lo que queda probado que φ es un homomorfismo de anillo.

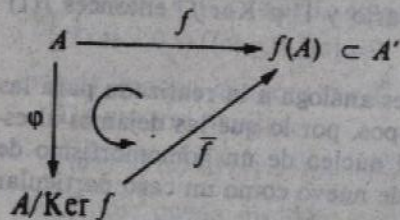
Sin embargo, como es conocido del caso de grupos y se mantiene en los anillos, para un mismo ideal I de A puede existir más de un homomorfismo del anillo A del cual I sea núcleo.

El siguiente teorema, similar al obtenido en el caso de los grupos, precisa la relación entre homomorfismo de anillos e ideales.

TEOREMA 4.7.1 (Teorema fundamental de homomorfismo de anillos)

Sea $f: A \rightarrow A'$ un homomorfismo de anillos. Entonces existe un isomorfismo natural $\bar{f}: A/\text{Ker } f \rightarrow f(A)$ tal que $f = \bar{f} \circ \varphi$ donde φ es el homomorfismo canónico de A en $A/\text{Ker } f$.

Antes de realizar la demostración hagamos un diagrama que describe la situación:



El único elemento que es de orden finito es el neutro. Pero si consideramos, por ejemplo, $\mathbb{Z} \times \mathbb{Z}$, existen elementos distintos del neutro que son de orden finito, por ejemplo, $\phi(0,1) = 2$, y elementos de orden infinito, el $(1,0)$ es uno de ellos. Más aún, si consideramos el grupo multiplicativo $(\mathbb{C}^*, \cdot)$, existen elementos de orden finito n para cada n entero positivo dado, a la vez que existen elementos de orden infinito.

Daremos ahora algunas definiciones y resultados que nos permitirán obtener una clasificación de los grupos abelianos.

DEFINICIÓN 3.14.3

Se dice que un grupo G es un grupo con torsión, si todo elemento de G es de orden finito, y que un grupo G es libre de torsión si el único elemento de orden finito es el neutro. G es mixto si no es con torsión ni libre de torsión.

Ejemplos

1. Todos los grupos finitos son grupos con torsión.

2. $(\mathbb{Z}, +)$ es un grupo libre de torsión.

3. $\mathbb{Z} \times \mathbb{Z}$ no es un grupo con torsión, pues

$$\langle (1,0) \rangle = \{(n,0) : n \in \mathbb{Z}\};$$

y tampoco es libre de torsión, pues

$$\langle (0,1) \rangle = \{(0,n) : n \in \mathbb{Z}\}.$$

Sin embargo, estudiando los elementos de orden finito de un grupo podemos obtener el siguiente resultado:

TEOREMA 3.14.2

En un grupo abeliano G , el conjunto de los elementos de orden finito es un subgrupo T de G llamado el subgrupo de torsión de G .

Demostración

En efecto, si a y b están en T es decir, son de orden finito, entonces existen m, n tales que $a^m = b^n = e$.

Probemos que $ab \in T$. Para ello basta tomar $(ab)^{mn} = (a^m)^n (b^n)^m = e$.

Por último, si a es de orden finito, $a^m = e$ para algún m . De aquí

$$(a \cdot a^{-1})^m = e^m = e$$

y

$$(a \cdot a^{-1})^m = a^m \cdot (a^{-1})^m = (a^{-1})^m = e,$$

con lo cual queda probado que a^{-1} es de orden m .

Está claro que el subgrupo de torsión de un grupo es un grupo con torsión.

Ejemplo

El subgrupo de torsión de $\mathbb{Z} \times \mathbb{Z}_2 = \{(0,0), (0,1)\} = \mathbb{Z}_2$.
 En $\mathbb{Z} \times \mathbb{Z}_2$ todo elemento (n,m) puede ser escrito en forma única como $(n,m) = (n,0) + (0,m)$ donde $n \in \mathbb{Z}$, $m \in \mathbb{Z}_2$, lo que puede ser verificado muy fácilmente. Los elementos de la forma $(n,0)$, $n \in \mathbb{Z}$, son de orden infinito en $\mathbb{Z} \times \mathbb{Z}_2$ y $\mathbb{Z} \times \{0\}$ constituye un subgrupo libre de torsión de $\mathbb{Z} \times \mathbb{Z}_2$.

Lo anterior nos da la idea de considerar la posibilidad de "separar" de forma adecuada todo grupo mixto en un grupo libre de torsión y en un grupo con torsión, con lo cual el estudio de los grupos abelianos se reduciría a:

1. Los grupos con torsión.
2. Los grupos libres de torsión.
3. Cómo obtener los grupos mixtos mediante grupos con torsión y grupos libres de torsión.

El tercer punto puede ser planteado de otra manera mediante el resultado siguiente:

TEOREMA 3.14.3

Si G es un grupo abeliano y T su subgrupo de torsión, entonces G/T es libre de torsión.

Demostración

En efecto, supongamos que existan en G/T elementos de orden finito, esto es, gT tal que $(gT)^m = T$ para algún entero positivo m .

$(gT)^m = g^m T = T \Rightarrow g^m \in T \Rightarrow (g^m)^n = e$ para algún entero positivo n . Pero $(g^m)^n = g^{mn} = e \Rightarrow g \in T \Rightarrow gT = T$. Esto significa que el único elemento de G/T que es de orden finito es el neutro T de G/T , y que por tanto, G/T es libre de torsión.

Podemos, pues, plantear el tercer punto de la forma siguiente:

Cómo construir grupos abelianos partiendo de un subgrupo de torsión dado T y de forma tal que A/T sea isomorfo a un grupo libre de torsión dado.

Hemos querido dar una idea del planteamiento del problema para grupos abelianos en general. Sin embargo, el estudio de este problema no corresponde al objetivo de este libro. Nos limitaremos, pues, a dar para el caso de los grupos abelianos finitamente generados, los resultados que prueban que estos pueden construirse a partir de grupos con torsión y grupos libres de torsión. Veremos que, al igual que en el caso de los grupos cíclicos, en el cual obtuvimos una clasificación de estos mediante $\mathbb{Z}$ y los $\mathbb{Z}_m$, ahora podemos clasificar en forma adecuada los grupos abelianos finitamente generados mediante productos de $\mathbb{Z}$ y $\mathbb{Z}_m$.

Muchos de los resultados los daremos sin demostrarlos, ya que aunque el nivel de las demostraciones no difiere del encontrado hasta el momento

en este texto, son laboriosas y alargarían mucho el presente epígrafe, en el cual según señalamos al principio solo queremos dar nociones del tema.

Antes de pasar a dar los resultados que caracterizan los grupos abelianos finitamente generados, es necesario tratar un concepto no conocido: el *producto directo interno*.

El estudiante encontrará en lo que sigue afirmaciones de este tipo: "El grupo G es igual al producto directo de sus subgrupos F y H ". Veamos cómo interpretar esta afirmación, ya que $F \times H$ es un conjunto de pares de elementos de G y no es comparable con G en sentido estricto.

Realmente lo que ocurre es que al definir el producto directo externo de los grupos G_i , o sea, $\prod_{i=1}^n G_i$, podemos considerar los subconjuntos

$$\bar{G}_i = \{(e_1, e_2, \dots, e_{i-1}, a_i, e_{i+1}, \dots, e_n) \mid a_i \in G_i\}$$

(e_i es el elemento neutro de G_i).

Estos subconjuntos son subgrupos de G isomorfos a los G_i , es decir,

$\bar{g}_i \cong \bar{G}_i$ para cada $1 \leq i \leq n$ y $\prod_{i=1}^n \bar{G}_i$ es el *producto directo interno* de los subgrupos $\bar{G}_i$.

Aunque ya señalamos que no daremos las demostraciones de muchos de los resultados donde aparece el concepto de *producto directo interno*, formalizaremos dicho concepto debido a su gran importancia.

DEFINICIÓN 3.14.4

Se dice que un grupo G es el *producto directo interno* de sus subgrupos G_i , $i=1, \dots, n$ si y solo si la aplicación

$$\Phi: \prod_{i=1}^n G_i \longrightarrow G \text{ dada por } (g_1, g_2, \dots, g_n) \mapsto g_1 \cdot g_2 \cdot \dots \cdot g_n$$

es un isomorfo de grupos.

Observemos que en la definición anterior se afirma que la imagen isomorfa de $\bar{G}_i$ es G_i . Esto justifica que, por abuso de lenguaje, podamos usar el término *producto de grupos* sin especificar si es externo o interno.

Veamos ahora los resultados que nos permitirán dar la caracterización de los grupos abelianos finitamente generados.

LEMA 1

Si G es un grupo finitamente generado y T su subgrupo de torsión, entonces G es el producto directo de $T \times F$, donde F es un subgrupo de G libre de torsión.

LEMA 2

Un grupo libre de torsión finitamente generado F es isomorfo a un producto $\underbrace{\mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z}}_{m \text{ veces}}$ para un cierto número de factores m . El número m , llamado

número de Betti de F , es único para cada F .

LEMA 3
Un grupo abeliano finito T es isomorfo a dos diferentes tipos de productos de grupos cíclicos:

(1) T es isomorfo a un producto

$$\mathbb{Z}_{(p_1)^{a_1}} \times \mathbb{Z}_{(p_2)^{a_2}} \times \dots \times \mathbb{Z}_{(p_r)^{a_r}},$$

donde los p_i son primos y no necesariamente distintos. Este producto es único salvo el orden de los factores.

(2) T es isomorfo a un producto

$$\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \dots \times \mathbb{Z}_{m_r},$$

donde m_i divide a m_{i+1} . Los números m_i se llaman *coeficientes de torsión* de T y son únicos.

Las denominaciones de *número de Betti* y *coeficiente de torsión* provienen de la topología algebraica y son invariantes topológicos. El primero se debe al matemático italiano Enrico Betti (1823-1892), el cual observó las propiedades de dichos invariantes.

LEMA 4
Los subgrupos de un grupo finitamente generado son grupos finitamente generados.

Los resultados anteriores conducen a las siguientes observaciones y conclusiones:

El señalamiento de que los primos p_i que aparecen en la proposición (1) del lema 3 pueden no ser distintos, es decir, que combinando en distintas formas potencias de un número primo dado, podemos obtener grupos distintos; por ejemplo, $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$ no es isomorfo a $\mathbb{Z}_9 \times \mathbb{Z}_3$, ya que en $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$ los elementos tienen a lo sumo orden 3, mientras que en $\mathbb{Z}_9 \times \mathbb{Z}_3 \cong \mathbb{Z}_{27}$ pueden tener orden 27.

Realmente el lema 3 ofrece la posibilidad de encontrar, salvo isomorfismo, todos los grupos finitos de un orden dado, para lo cual describimos inmediatamente un método.

Dicho método se basa en el lema 3, que dice que todo grupo finito es en definitiva, salvo isomorfismo, igual a un producto directo en grupos cíclicos finitos, y también en los resultados que habíamos obtenido anteriormente sobre el producto directo de grupos cíclicos.

Ejemplo

Si queremos hallar todos los grupos finitos de orden 600, comenzamos por descomponer 600 en factores primos:

$$600 = 2^3 \cdot 3 \cdot 5^2$$

Según la parte (1) del lema 3, tenemos las siguientes posibilidades para grupos de orden 600.

1) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$,

2) $\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5$,

3) $\mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$,

4) $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_{25}$,

5) $\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_{25}$,

6) $\mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_{25}$.

Recordemos que en el lema 3 se afirma que los grupos de orden 600 son, salvo isomorfismo, los seis que acabamos de hallar, los cuales pueden ser expresados también mediante los coeficientes de torsión. Veamos cómo obtener esta última expresión de los grupos de orden 600.

Tomemos la posibilidad 1):

$$\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_5$$

En este caso ordenamos los coeficientes primos de los subíndices como sigue:

$$\begin{array}{ccccccc} 2 & 2 & 2 & 3 & 5 & 5 \\ & & & 3 & & \\ & & & 5 & 5 & \end{array}$$

Esto es, alineamos los correspondientes a cada primo por filas, en orden creciente, y a su vez alineamos los lados derechos de las filas.

Para mostrar esto con más claridad, hagamos el arreglo correspondiente a 2):

$$\mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5$$

Esto es,

$$\begin{array}{ccccccc} 2 & 4 & & & & & \\ & 3 & & & & & \\ & 5 & 5 & & & & \end{array}$$

Aunque no daremos una demostración formal del método, el estudiante puede observar que está basado en el siguiente hecho: mediante el ordenamiento vamos agrupando los factores para formar grupos cíclicos isomorfos a los dados y que cumplan las condiciones de los coeficientes de torsión. Así, en $\mathbb{Z}_2 \times \mathbb{Z}_{10} \times \mathbb{Z}_{30}$ tenemos:

$$\mathbb{Z}_{10} = \mathbb{Z}_2 \times \mathbb{Z}_5, \quad \mathbb{Z}_{30} = \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5,$$

y además, 2 divide a 10 y este a 30.

Observemos que si tomamos ahora los productos por columna en el ordenamiento de los coeficientes correspondientes a la posibilidad 1), obtenemos: 2 10 30. Entonces

$$\mathbb{Z}_2 \times \mathbb{Z}_{10} \times \mathbb{Z}_{30} = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_3.$$

En la posibilidad (2) obtenemos: 10 60, y de aquí

$$\mathbb{Z}_{10} \times \mathbb{Z}_{60} = \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \times \mathbb{Z}_3.$$

En total, después de aplicado el método tendremos para $\mathbb{Z}_{600}$ las siguientes posibilidades expresadas en la forma (1) y (2) del lema:

$$1) \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5 = \mathbb{Z}_2 \times \mathbb{Z}_{10} \times \mathbb{Z}_{30}$$

$$2) \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_5 = \mathbb{Z}_{10} \times \mathbb{Z}_{60}$$

$$3) \mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_5 = \mathbb{Z}_2 \times \mathbb{Z}_{120}$$

$$4) \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_{15} = \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_{150}$$

$$5) \mathbb{Z}_2 \times \mathbb{Z}_4 \times \mathbb{Z}_3 \times \mathbb{Z}_{15} = \mathbb{Z}_2 \times \mathbb{Z}_{100}$$

$$6) \mathbb{Z}_8 \times \mathbb{Z}_3 \times \mathbb{Z}_{15} = \mathbb{Z}_{600}$$

En la justificación del método descrito se utiliza el siguiente resultado, que será necesario en el siguiente teorema y puede ser fácilmente verificado por el estudiante:

Si un grupo G es producto directo de los grupos H y K , y H es producto directo de L y M , entonces $G = L \times M \times K$.

Ahora, de los resultados anteriores podemos obtener el siguiente teorema que indica la forma de todo grupo abeliano finitamente generado.

TEOREMA 3.14.4 (Caracterización de los grupos abelianos finitamente generados)

Todo grupo abeliano finitamente generado G es isomorfo a un producto directo de grupos cíclicos de la forma

$$(1) \mathbb{Z}_{p_1^{r_1}} \times \mathbb{Z}_{p_2^{r_2}} \times \dots \times \mathbb{Z}_{p_n^{r_n}} \times \mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z},$$

donde los p_i son primos no necesariamente distintos, y también de la forma

$$(2) \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \dots \times \mathbb{Z}_{m_r} \times \mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z},$$

donde m_i divide a m_{i+1} .

En ambos casos el producto directo es único, salvo isomorfismo. El número de Betti de G (o sea, el número de factores $\mathbb{Z}$) es único, los coeficientes de torsión m_i de G son únicos y las potencias $p_i^{r_i}$ son únicas.

Demostración

Según el lema 1, $G = T \times F$ donde T es el grupo de torsión de G y F es un grupo libre de torsión.

Como F es un grupo abeliano libre de torsión y subgrupo de un grupo abeliano finitamente generado, será a su vez finitamente generado, según el lema 4.

Por otra parte, por ser T el grupo de torsión de un grupo finitamente generado, será un grupo finito. (El estudiante puede comprobar que un grupo de torsión finitamente generado es finito, mediante la aplicación de la definición de grupo finitamente generado.)

Haciendo, entonces, uso del lema 3 para T y del lema 2 para F , queda demostrado el teorema.

Para concluir, veamos algunas aplicaciones de los resultados anteriormente expuestos.

Aplicaciones

TEOREMA 3.14.5

Si m divide el orden de un grupo abeliano finito G , entonces G tiene un subgrupo de orden m .

Demostración

Podemos suponer por el lema 3 que

$$G = \mathbb{Z}_{p_1^{r_1}} \times \mathbb{Z}_{p_1^{r_2}} \times \dots \times \mathbb{Z}_{p_n^{r_n}}.$$

Entonces

$$m = p_1^{s_1} \cdot p_2^{s_2} \cdot \dots \cdot p_n^{s_n} \text{ con } 1 \leq s_i \leq r_i.$$

ya que el orden n de G podía escribirse como $n = p_1^{r_1} \cdot p_2^{r_2} \cdot \dots \cdot p_n^{r_n}$.

Entonces $(p_i)^{r_i - s_i}$ genera un subgrupo cíclico de $\mathbb{Z}_{p_i^{r_i}}$ de orden igual al cociente de $(p_i)^{r_i}$ entre el máximo común divisor de $p_i^{r_i}$ y $p_i^{s_i}$. Como $\text{mcd}(p_i^{r_i}, p_i^{s_i}) = p_i^{s_i}$, tenemos que $(p_i)^{r_i - s_i}$ genera un subgrupo cíclico de $\mathbb{Z}_{p_i^{r_i}}$ de orden $(p_i)^{s_i}$. Entonces

$$\langle (p_1)^{r_1 - s_1} \rangle \times \langle (p_2)^{r_2 - s_2} \rangle \times \dots \times \langle (p_n)^{r_n - s_n} \rangle$$

es el subgrupo cíclico de G de orden m que queríamos encontrar.

Acabamos de demostrar que el recíproco del teorema de Lagrange se cumple para los grupos abelianos finitos. Esto es de gran utilidad, pues permite afirmar, por ejemplo, que en un grupo abeliano de orden 24 existen subgrupos propios de orden 2, 3, 4, 6, 8 y 12.

Veamos con un ejemplo otro problema que podemos resolver mediante el teorema de caracterización de los grupos abelianos finitamente generados y finitos que acabamos de estudiar.

Sabemos que un grupo de orden primo es cíclico y también tenemos un criterio para determinar cuándo un grupo del tipo $Z_n \times Z_m$ es cíclico. Vamos ahora a encontrar un criterio suficiente para que un grupo abeliano finito de orden m sea cíclico.

TEOREMA 3.14.6

Si m no es divisible por el cuadrado de ningún número primo, entonces todo grupo abeliano de orden m es cíclico.

Demostración

Sea G un grupo abeliano de orden m , donde m no es divisible por el cuadrado de ningún número primo.

Sabemos que G es isomorfo a un grupo de la forma

$$Z_{p_1^{r_1}} \times Z_{p_2^{r_2}} \times \dots \times Z_{p_k^{r_k}}.$$

Luego podemos identificarlo con este.

Pero como $m = p_1^{r_1} \cdot p_2^{r_2} \cdot \dots \cdot p_k^{r_k}$ y m no es divisible por el cuadrado de ningún número primo, tenemos que

$$r_1 = r_2 = \dots = r_k = 1.$$

De aquí

$$G = Z_{p_1} \times Z_{p_2} \times \dots \times Z_{p_k}.$$

y como los p_i son primos relativos, tenemos

$$G = Z_{p_1 \cdot p_2 \cdot \dots \cdot p_k}.$$

Por tanto, G es cíclico.

Ejemplo

Tomemos un grupo de orden 42, que no es divisible por el cuadrado de ningún número primo ya que sus únicos divisores son 2, 3 y 7. Según el teorema 3.14.6, podemos decir que todo grupo de orden 42 es cíclico; en otras palabras, existe un único grupo de orden 42, Z_{42} .

Observemos que con una aplicación directa del método de búsqueda de los grupos abelianos finitos de un orden dado, llegaríamos al mismo resultado ya que $42 = 2 \cdot 3 \cdot 7$ nos da la única posibilidad para un grupo de orden 42: $Z_2 \times Z_3 \times Z_7 = Z_{42}$; pero la propiedad dada por el teorema 3.14.6 nos permite ganar tiempo en la búsqueda de los grupos finitos de un orden dado.

Destaquemos que la condición dada por el teorema anterior es una condición suficiente pero no necesaria, ya que para los grupos de orden primo su orden no es divisible por el cuadrado de un primo y, sin embargo, son cíclicos.

Hemos dicho en varias ocasiones que los grupos cíclicos son los grupos más sencillos que podemos considerar. El siguiente teorema confirma esa afirmación en el caso de los grupos abelianos finitos.

TEOREMA 3.14.7

Los grupos abelianos finitos que no son isomorfos a un producto directo de dos de sus subgrupos propios, son exactamente los grupos cíclicos de orden potencia de un primo.

(Un tal grupo se llama *indescomponible*. Esto significa, intuitivamente, que lo hemos llevado a la mínima expresión.)

Demostración

Supongamos que G es un grupo abeliano finito que no es posible expresarlo como producto directo de dos de sus subgrupos propios. Entonces G debe ser isomorfo a un producto del tipo $Z_{p_1^{r_1}} \times \dots \times Z_{p_k^{r_k}}$ y debe constar de un solo factor. Esto es,

$$G = Z_{p^r}.$$

Recíprocamente, sea un grupo Z_{p^r} . Si Z_{p^r} pudiera ser expresado como producto directo de dos de sus subgrupos propios, entonces $Z_{p^r} = Z_{p^i} \times Z_{p^j}$ con $i+j=r$, lo cual es imposible ya que el orden máximo posible de los elementos de $Z_{p^i} \times Z_{p^j}$ es, $p^{\max\{i,j\}} < p^r$.

Los resultados hallados para grupos abelianos finitos permiten extender el método de búsqueda de los subgrupos, expuesto en el epígrafe 3.13, a todos los grupos abelianos finitos, ya que sabemos de antemano que estos poseen subgrupos de orden m para cada m divisor del orden del grupo (no necesariamente uno de cada orden como en el caso de los cíclicos) y, por otra parte, conocemos los modelos para clasificar un grupo abeliano finito cualquiera.

Ejemplo

Busquemos los subgrupos de $Z_4 \times Z_4$.

$\phi(Z_4 \times Z_4) = 16$; luego tiene subgrupos propios de órdenes 2, 4 y 8.

$$Z_4 \times Z_4 = \{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (1,3), (2,0), (2,1), (2,2), (2,3), (3,0), (3,1), (3,2), (3,3)\}.$$

Los subgrupos de orden 2 son cíclicos:

$$\langle (0,2) \rangle = \{(0,2), (0,0)\}$$

$$\langle (2,0) \rangle = \{(2,0), (0,0)\}$$

$$\langle (2,2) \rangle = \{(2,2), (0,0)\}$$

Los subgrupos cíclicos de orden 4 son:

$$\langle (0,1) \rangle = \{(0,1), (0,2), (0,3), (0,0)\}$$

$$\langle (1,0) \rangle = \{(1,0), (2,0), (3,0), (0,0)\}$$

$$\langle (1,2) \rangle = \{(1,2), (2,0), (3,2), (0,0)\}$$

$$\langle (1,3) \rangle = \{(1,3), (2,2), (3,1), (0,0)\}$$

$$\langle (2,1) \rangle = \{(2,1), (0,2), (2,3), (0,0)\}$$

$$\langle (3,3) \rangle = \{(3,3), (2,2), (1,1), (0,0)\}$$

Hay un subgrupo de orden 4 isomorfo a V_4 :

$$\bar{V}_4 = \{(0,0), (0,2), (2,0), (2,2)\}$$

No hay subgrupo cíclico de orden 8. Por tanto, si existe un subgrupo de orden 8 tiene que ser isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_4$ o $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$. Como en $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ no existen elementos de orden 4 y en $\mathbb{Z}_2 \times \mathbb{Z}_4$ si, podemos decir que en $\mathbb{Z}_4 \times \mathbb{Z}_4$ hay al menos un subgrupo de orden 8 isomorfo a $\mathbb{Z}_2 \times \mathbb{Z}_4$.

Realmente podemos buscar los subgrupos de orden 8 pues están dados por grupos isomorfos a los productos de los grupos cíclicos generados por elementos de orden 2 y de orden 4, respectivamente; pero esto resulta un tanto engorroso y con los resultados hallados tenemos ya una información bastante completa sobre los subgrupos.

Ejercicios

- Encuentre todos los grupos abelianos (salvo isomorfismo) de orden: a) 60 b) 720 c) 1 800 d) 24 e) 25 f) 24 · 25.
- Sean m y n primos relativos. Pruebe que si existen (salvo isomorfismo) r grupos abelianos de orden m , y s de orden n , entonces existen (salvo isomorfismo) rs grupos abelianos de orden $m \times n$.
- Encuentre el subgrupo de $\mathbb{Z}_{12}$ generado por $\{2,3\}$, el generado por $\{4,6\}$ y el generado por $\{8,6,10\}$.
- Sea G un grupo abeliano de orden 72.
 - Diga cuántos subgrupos de orden 8 tiene. ¿Por qué?
 - Diga cuántos subgrupos de orden 4 tiene. ¿Por qué?
- Pruebe que si un grupo abeliano finito tiene orden una potencia de un primo p , entonces el orden de cada elemento del grupo es una potencia de p .
- Encuentre el subgrupo de torsión T de $(\mathbb{C}^*, \cdot)$.
- Sean G, H y K grupos abelianos finitamente generados. Pruebe que si $G \times K = H \times K$, entonces $G = H$.
- Pruebe que si G es finitamente generado y φ es un homomorfismo de G en un grupo G' , el valor de φ en cada elemento de G está completamente determinado por los valores $\varphi(a_i)$, ($1 \leq i \leq n$) de los generadores a_i de G .
- Halle cuántos homomorfismos existen
 - de $\mathbb{Z}_2 \times \mathbb{Z}_2$ en $\mathbb{Z}_2$
 - de $\mathbb{Z}_2 \times \mathbb{Z}_2$ sobre $\mathbb{Z}_2$
 - de $\mathbb{Z}_2 \times \mathbb{Z}_2$ en $\mathbb{Z}_6$.
- Pruebe que si G es un grupo y H un subgrupo de G tal que G/H es libre de torsión, entonces H contiene el subgrupo de torsión de G .

Sugerencia: Considere un elemento de orden finito de G y pruebe que está contenido en H .

- Analizando el orden de los elementos de los grupos $\mathbb{Q}$, $\mathbb{R}/\mathbb{Z}$ y $\mathbb{C}$, clasifíquelos en: libre de torsión, con torsión y mixtos. Deduzca de aquí que estos tres grupos no son isomorfos.
- Pruebe que si T es el subgrupo de torsión de G , entonces $T \cap H$ es el subgrupo de torsión de cualquier subgrupo H dado en G .
- Pruebe que el subgrupo de torsión de $\mathbb{R}/\mathbb{Z}$ es $\mathbb{Q}/\mathbb{Z}$.