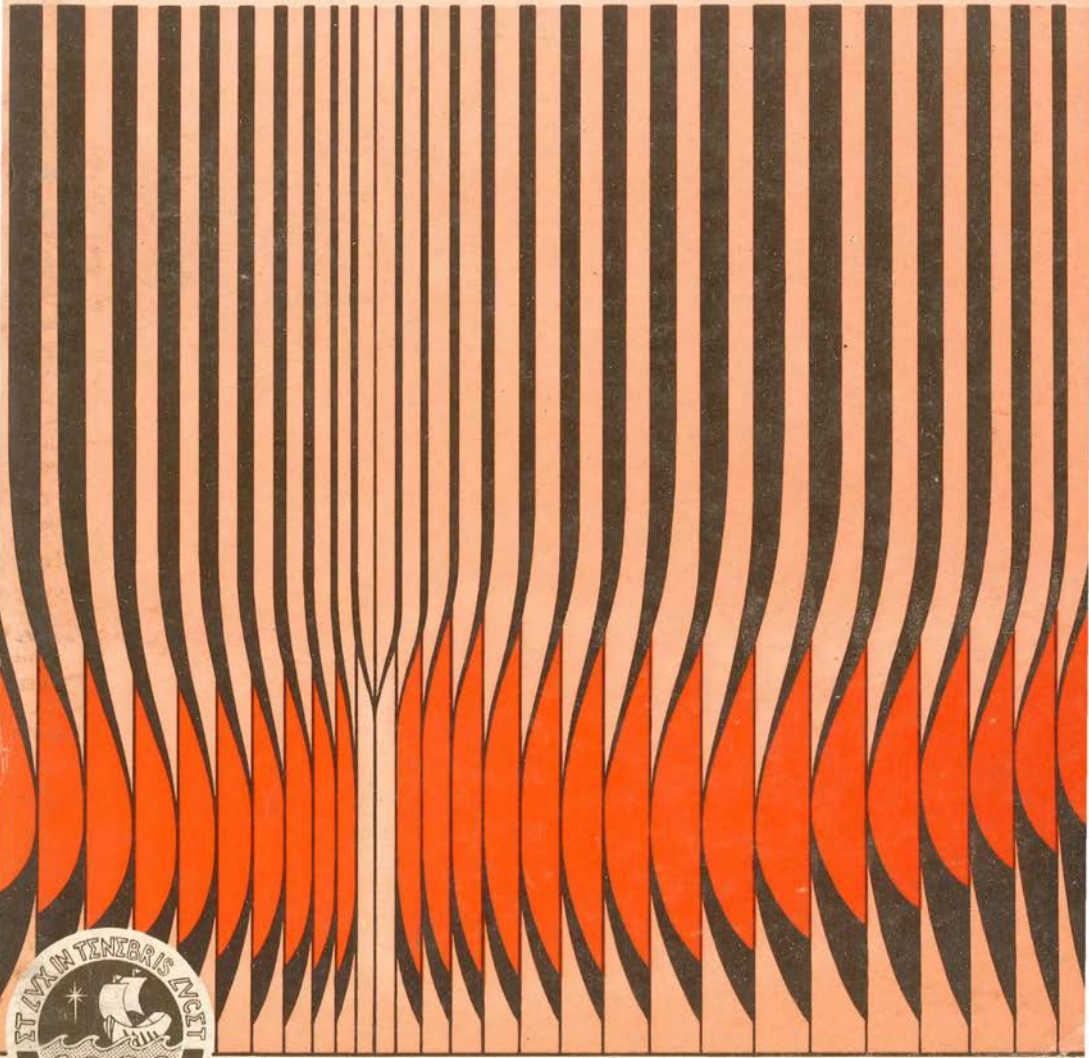
ALGEBRA LINEAL Y MULTILINEAL

PRIMERA PARTE

José Tola Pasquel

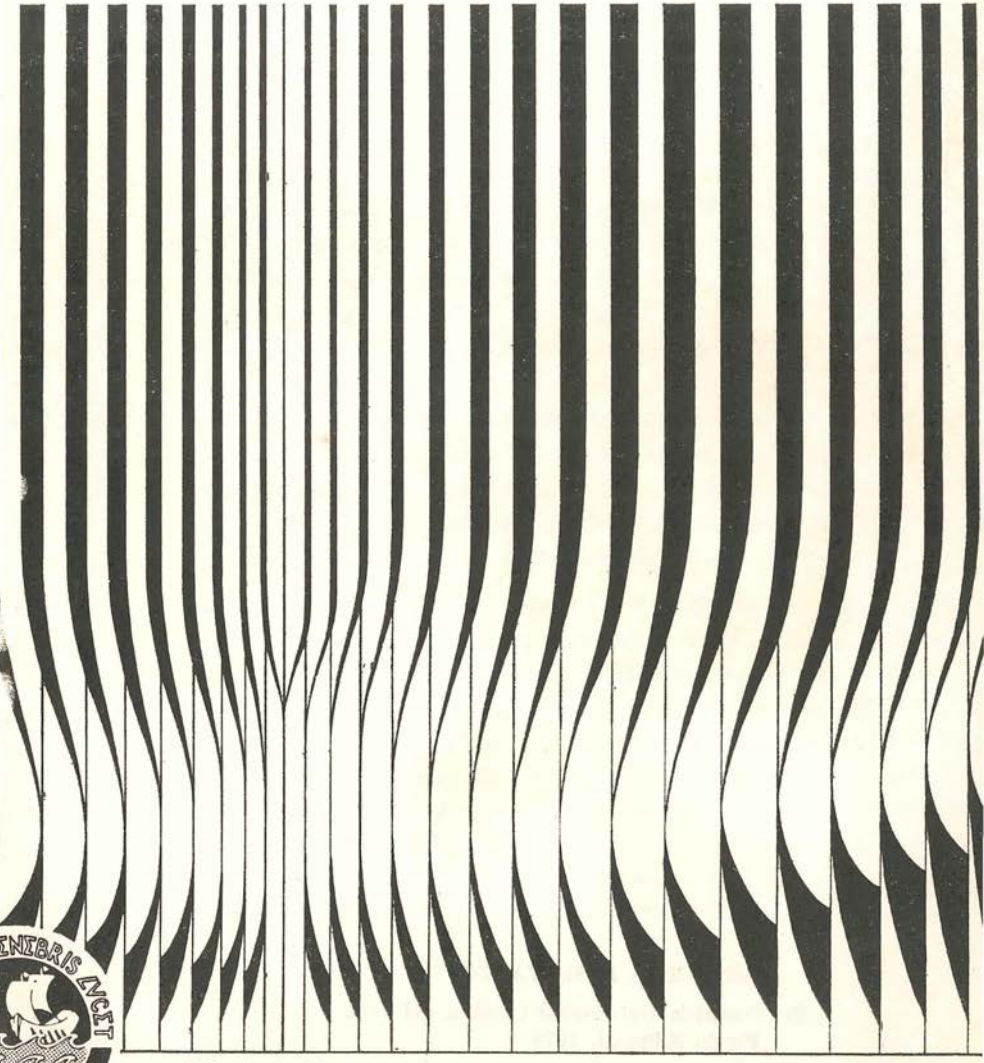


PONTIFICIA UNIVERSIDAD CATOLICA DEL PERU
FONDO EDITORIAL 1978

ALGEBRA LINEAL Y MULTILINEAL

PRIMERA PARTE

José Tola Pasquel



PONTIFICIA UNIVERSIDAD CATOLICA DEL PERU
FONDO EDITORIAL 1978

DERECHOS RESERVADOS

© Pontificia Universidad Católica del Perú
Fondo Editorial, 1978

PREFACIO

Creo que un paso importante en el esfuerzo por impulsar entre nosotros el conocimiento y la utilización de las matemáticas consiste en proporcionar los medios necesarios para acceder a ellas en la forma más directa que sea posible, eliminando innecesarias dificultades, entre las cuales puede señalarse el uso de textos que no se acomodan bien a las exigencias de nuestros estudiantes y de nuestros programas y que, casi siempre, fueron escritos en idiomas distintos del nuestro. No significa esto, desde luego, que no conside re recomendable y aun necesario que sean consultadas y estudiadas las numerosas obras excelentes escritas sobre el tema, con otros propósitos y en otros idiomas. Considero, sin embargo, que esa recomendación complementa y no se opone al serio esfuerzo que deba hacerse para redactar textos básicos en nuestro propio idioma y con arreglo a nuestras propias demandas.

Este volumen, dedicado al Algebra Lineal, constituye la pri mera parte del curso de Algebra Lineal y Multilineal que, con duración de dos semestres he ofrecido en la Pontificia Universidad Católica del Perú. He incluido algunos temas que no fueron tratados en clase y cuyo estudio fué confiado al cuidado de los estudiantes. He tratado de que la exposición sea suficientemente clara como pa ra que su estudio pueda hacerse, si fuera necesario, sin el concur so de un profesor. Supone, sin embargo, una cierta formación matemática, en particular, un conocimiento adecuado de las nociones fundamentales acerca de las estructuras algebraicas.

Es necesario señalar que este libro debe mucho a varios otros que sería largo enumerar.

Debo recomendar en forma especial la resolución de los ejercicios; en particular, los que están marcados con un asterisco (*) , los cuales complementan el texto y ocasionalmente son empleados en él.

La segunda parte está ya en preparación y se refiere al Algebra Multilineal, que también ha sido objeto de un curso semestral.

Agradezco al Dr. Maynard Kong, quien ha revisado el texto original y sugerió valiosas modificaciones.

Lima, noviembre de 1976.

José Tola .

ALGEBRA LINEAL

INDICE GENERAL

CAPITULO I

ESPACIOS VECTORIALES, TRANSFORMACIONES LINEALES Y MATRICES

§ 1. Campos

1.1	Definición de campo	1
1.2	Subcampo	3
1.3	Característica de un campo	4

§ 2. Espacios vectoriales

2.1	Definición de espacio vectorial	5
2.2	Dependencia e independencia lineal	7
2.3	Subespacios vectoriales	12
2.4	Espacio vectorial factor o cociente	12
2.5	Base de un espacio vectorial. Dimensión	15

§ 3. Transformaciones lineales

3.1	Definiciones	26
3.2	Núcleo e imagen de una transformación lineal	29
3.3	Comosición de transformaciones lineales. Transformaciones lineales inversas	31

3.4 Isomorfismos de espacios vectoriales 32
3.5 Isomorfismo inducido por una transformación lineal 34

§4. Dualidad de espacios vectoriales

4.1 Aplicaciones bilineales 36
4.2 Producto escalar y espacios duales 36
4.3 Relación de ortogonalidad entre vectores de espacios duales 40
4.4 Transformaciones duales 41

§5. Dualidad de los espacios de dimensión finita

5.1 Dimensión de los espacios duales 43
5.2 Base dual 44
5.3 El espacio conjugado $\mathcal{L}(V)$ de un espacio de dimensión finita 45
5.4 La ortogonalidad en el caso de espacios de dimensión finita 48
5.5 Transformaciones duales entre espacios de dimensión finita 50

§6. Transformaciones lineales y matrices

6.1 Algebras 54
6.2 Algebra de los endomorfismos de un espacio vectorial 55
6.3 Matrices 58
6.4 Matriz de una transformación lineal 59
6.5 Vectores-columna y vectores-fila de matrices cualesquiera 62
6.6 Matriz de la transformación dual de una transformación lineal 63
6.7 Rango de una matriz 66
6.8 Producto de matrices 69
6.9 Algebra de las matrices cuadradas. Grupo de las matrices regulares 72
6.10 Cambio de base 74
6.11 Matrices semejantes 77
6.12 Endomorfismos semejantes 79

CAPITULO II

DETERMINANTES Y SISTEMAS DE ECUACIONES LINEALES

§7. Determinantes

7.1 Definición de un determinante de orden n 81
7.2 Matriz inversa de una matriz cuyo determinante es diferente de cero 86
7.3 Aplicación de los determinantes a la resolución de los sistemas de n ecuaciones lineales con n incógnitas: regla de Cramer 88
7.4 Dependencia lineal de las columnas o filas de una matriz. Rango de una matriz 91
7.5 Determinante del producto de matrices 99
7.6 Determinante de un endomorfismo 105
7.7 Desarrollo de un determinante por la fórmula de Laplace 106

§8. Sistema de ecuaciones lineales

8.1 Consideraciones y teoremas generales 112
8.2 Criterio del rango de la matriz aumentada 119
8.3. Resolución de los sistemas de ecuaciones linealmente independientes 121
8.4 Resolución de sistemas en el caso más general. Condiciones de compatibilidad 124

CAPITULO III

POLINOMIOS SOBRE CAMPOS. EXTENSIONES DE CAMPOS Y DE ESPACIOS VECTORIALES

§9. Anillos

9.1 Anillos 133
9.2 Ideales 134
9.3 Homomorfismos de anillos 136
9.4 Anillos cocientes 137

9.5 Dominios de integridad. Campo de cocientes 140

§ 10. Polinomios

10.1 Polinomios sobre un campo 143
10.2 División de polinomios 146
10.3 Ideales de $\mathbb{F}[x]$ 149
10.4 Máximo común divisor y mínimo común múltiplo de polinomios 150
10.5 Factorización y divisibilidad en $\mathbb{F}[x]$ 153
10.6 Funciones polinomiales 161
10.7 Factorización en $\mathbb{R}[x]$ y en $C[x]$ 163

§ 11. Extensiones de campos

11.1 Extensiones de un campo 165
11.2 Construcción de extensiones simples algebraicas 171
11.3 Campos algebraicamente cerrados 174
11.4 Cosntrucción de extensiones simples trascendentes 174

§ 12. Extensión de un espacio vectorial por ampliación del campo de escalares

12.1 Introducción. Transformaciones $\mathbb{F}$-lineales 180
12.2 Extensión de un espacio vectorial por ampliación del campo de escalares 181
12.3 Una nueva definición de la extensión de un espacio vectorial 184
12.4 Extensión de un espacio de dimensión finita 188
12.5 Extensión de un endomorfismo de V, por extensión de V 189
12.6 Extensión de un espacio vectorial real, correspondiente a la ampliación C de $\mathbb{R}$ 191

CAPITULO IV

TEORIA DE LOS ENDOMORFISMOS DE UN ESPACIO VECTORIAL DE DIMENSION FINITA

§ 13. Valores propios, polinomio característico y polinomio minimal de un endomorfismo

13.1 Valores propios y vectores propios de un endomorfismo 196
13.2 Polinomio característico de un endomorfismo 199
13.3 Polinomio minimal. Teorema de Cayley-Hamilton cuando $\mathbb{F}$ es algebraicamente cerrado 208
13.4 Suma directa de transformaciones lineales 225
13.5 Reducción del estudio de los endomorfismos al de los endomorfismos nilpotentes 226
13.6 El teorema de Cayley-Hamilton en el caso de campos no algebraicamente cerrados 227
13.7 Representación de un endomorfismo mediante una matriz diagonal 231

§ 14. Endomorfismos nilpotentes

14.1 Consideraciones preliminares 242
14.2 Forma canónica de Jordan de la matriz de un endomorfismo nilpotente 243

§ 15. Representación de un endomorfismo mediante una matriz de Jordan. Reducción de una matriz a la forma canónica de Jordan

15.1 Matrices de Jordan 255
15.2 Proposiciones preliminares 257
15.3 Representación de un endomorfismo de un espacio vectorial sobre un campo algebraicamente cerrado, mediante una matriz de Jordan 259
15.4 Forma canónica de Jordan de una matriz 269

§ 16. Divisores elementales de un endomorfismo

16.1 Los polinomios $D_j(x)$ de una matriz y de un endomorfismo 275
16.2 Cálculo de los polinomios $D_j(x)$ 278

16.3 Divisores elementales y cálculo de la forma canónica de Jordan 285

CAPITULO V
ESPACIOS DE PRODUCTO INTERNO

§ 17. El producto interno

17.1 Definiciones fundamentales 289
17.2 Norma o longitud de un vector de un espacio de producto interno 292
17.3 Bases ortonormales de los espacios de producto interno 298
17.4 Cambios de bases ortonormales de un espacio de producto interno. Matrices unitarias y ortogonales 304
17.5 Conjugación en los espacios unitarios 310
17.6 Dualidad de un espacio de producto interno consigo mismo 311
17.7 Ortogonalidad en los espacios de producto interno 312
17.8 El teorema de Riesz. Base recíproca 313
17.9 Distancia de un vector a un subespacio 316
17.10 Coeficientes de Fourier 318
17.11 Desigualdad de Bessel e igualdad de Parseval 319

§ 18. Endomorfismos normales y autoadjuntos e isometrias de los espacios de producto interno de dimensión finita

18.1 Transformaciones lineales adjuntas de espacios de producto interno 322
18.2 Endomorfismos normales de los espacios de producto interno 326
18.3 Endomorfismos autoadjuntos de los espacios de producto interno 333
18.4 Isometrias de espacios de producto interno 334

§ 19. Endomorfismos de los espacios unitarios

19.1 Endomorfismos normales de los espacios unitarios 340

19.2 Endomorfismos hermitianos de los espacios unitarios 342
19.3 Endomorfismos unitarios de los espacios unitarios 343

§ 20. Endomorfismos de los espacios euclideanos

20.1 Endomorfismos normales de los espacios euclideanos 346
20.2 Endomorfismos simétricos de los espacios euclideanos 346
20.3 Endomorfismos ortogonales de los espacios euclideanos 351

CAPITULO VI

ESPACIOS PUNTUALES AFINES Y EUCLIDEANOS

§ 21. El espacio puntual afín

21.1 Espacio puntual afín 361
21.2 Sistemas de coordenadas cartesianas en un espacio afín 362
21.3 Cambio de sistemas de coordenadas cartesianas 365
21.4 Aplicaciones afines 369
21.5 Subespacios afines 383
21.6 Paralelismo en los espacios afines 388
21.7 Representación de susbespacios. Coordenadas baricéntricas. Convexidad 389

§ 22. Espacios puntuales euclideanos

22.1 Conceptos fundamentales 399
22.2 Isometrías de espacios puntuales euclideanos 406
22.3 Clasificación de las isometrías de un espacio euclideano 412

Indice.
Indice de simbolos.

CAPITULO I

ESPACIOS VECTORIALES, TRANSFORMACIONES LINEALES Y MATRICES

Las teorias que vamos a estudiar en el presente libro suponen que son dados ciertos conjuntos de elementos dotados de una estructura algebraica particular, que se llaman campos o cuerpos conmutativos. Los ejemplos más importantes son los campos de números: los números racionales, los números reales, los números complejos, etc. En el parágrafo § 1 trataremos acerca de la definición y de las propiedades fundamentales de los campos. Nuestra exposición tiene el caracter de una revisión suscinta por cuanto el estudio de los campos pertenece al Algebra General.

§ 1. Campos

1.1. Definición de campo. Un conjunto $\mathbb{F}$ se llama *campo* si cumple las siguientes condiciones:

a) Es un grupo aditivo, es decir que es definida en él una operación de *adición*, tal que a cada par de elementos $x, y \in \mathbb{F}$ hace corresponder un tercer elemento de $\mathbb{F}$ que se designa por $x + y$, de manera que:

a1) $x + y = y + x$ *(propiedad conmutativa).*

a2) $(x + y) + z = x + (y + z)$, $\forall\, x, y, z \in \mathbb{F}$ *(propiedad asociativa).*

a3) Existe un elemento, que se designa por 0, y se llama *cero* o *elemento nulo*, tal que $x + 0 = x$ para todo $x \in \mathbb{F}$.

a4) Para cada $x \in \mathbb{F}$ existe un elemento de $\mathbb{F}$, que se llama *opuesto* de x o *inversa aditiva* de x, que se designa por $-x$, y es tal que $x + (-x) = 0$.

En adelante, dados los elementos x y $-y$ escribiremos $x + (-y) = x - y$, y al elemento que así resulta lo llamaremos *resta* o *diferencia* de x menos y.

b) Es definida una operación de *multiplicación*, que a cada par de elementos $x, y \in \mathbb{F}$ hace corresponder un tercer elemento de $\mathbb{F}$ que se designa por xy o, $x.y$, de manera que:

b1) $xy = yx$ *(propiedad conmutativa).*

b2) $(xy)z = x(yz)$ *(propiedad asociativa).*

b3) $x(y + z) = xy + xz$ *(propiedad distributiva).*

b4) Existe un elemento que se designa con 1, diferente de 0, que se llama *elemento idéntico* o *identidad* y es tal que $\forall\ x \in \mathbb{F}$, es $x.1 = x$.

b5) Para cada $x \neq 0$ de $\mathbb{F}$ existe un elemento $x \in \mathbb{F}$, tal que $xx^{-1} = 1.x^{-1}$ se llama *inversa* de x.

El elemento xy se llama *producto* de x por y.

La demostración de las propiedades contenidas en las dos proposiciones siguientes resulta inmediata a partir de la definición y constituye un ejercicio para el lector.

Teorema 1.1.1. Sea $\mathbb{F}$ un campo. Se cumplen las siguientes propiedades:

1. El elemento 1 que cumple la propiedad b4 es único.
2. Dado el elemento $x \neq 0$ de $\mathbb{F}$, el elemento x^{-1} es único.
3. $x.0 = 0$, $\forall\ x \in \mathbb{F}$.
4. Si a y b son elementos de $\mathbb{F}$ y $a \neq 0$, entonces existe uno y solo un elemento de $\mathbb{F}$ tal que $ax = b$. Ese elemento es $ba^{-1} = a^{-1}b$

Corolario. Si $ab = 0$, entonces es o bien $a = 0$ o bien $b = 0$. (Se expresa esta propiedad diciendo que $\mathbb{F}$ no contiene divisores de cero).

El producto de un número finito de elementos $x_1, x_2, \ldots, x_k$

de un campo $\mathbb{F}$ se define inductivamente mediante la fórmula

$$x_1 x_2 \ldots x_k = (x_1 x_2 \ldots x_{k-1}) x_k \ .$$

Cualquiera que sea j, $1 \leqslant j \leqslant k-1$, se tiene entonces

$$x_1 x_2 \ldots x_k = (x_1 x_2 \ldots x_j)(x_{j+1} \ldots x_k).$$

Ejemplos. Los siguientes son ejemplos de campos:

- El conjunto de los números reales $\mathbb{R}$, el conjunto de los números racionales $\mathbb{Q}$ y el conjunto de los números complejos $\mathbb{C}$, con las operaciones ordinarias de suma y producto en cada uno de ellos.

- El conjunto $\{0,1\}$, con las operaciones definidas de la manera siguiente:

$$0 + 0 = 0 \ , \ 0 + 1 = 1 + 0 = 1 \ , \ 1 + 1 = 0,$$
$$0\,0 = 0 \ , \ 0\,1 = 1\,0 = 0 \ , \ 1\,1 = 1.$$

- Si p es un número entero primo y (a) designa a la clase de restos módulo p que contiene al número a, entonces dichas clases de restos, con las operaciones

$$(a) + (b) = (a + b)$$
$$(a)(b) = (ab)$$

constituye un campo que consta de p elementos.

- Los números de la forma $a + b\sqrt{2}$, donde a y b son números racionales, constituyen un campo con las operaciones

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$$
$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2} \ ,$$

que no son otras que las operaciones de suma y producto ordinarios.

1.2. Subcampos. Un subconjunto K de un campo $\mathbb{F}$ se llama *subcampo* de $\mathbb{F}$ si:

1. 0 y 1 pertenecen a K.
2. Si a y b pertenecen a K, entonces $a + b$, $a - b$ y ab pertenecen a K.

3. Si $a \neq 0$ está en K, entonces a^{-1} está en K.

K es evidentemente un campo con las mismas operaciones que $\mathbb{F}$. $\mathbb{F}$ y $\{0\}$ son subcampos de $\mathbb{F}$. Los demás subcampos de $\mathbb{F}$ se llaman *subcampos propios.*

Ejemplo. $\mathbb{Q}$ es un subcampo de $\mathbb{R}$ y $\mathbb{R}$ es un subcampo de $\mathbb{C}$.

1.3. Característica de un campo.

Definición 1.3.1. Dado un elemento cualquiera a de un campo, y un entero positivo arbitrario n se define el elemento na del campo, mediante la fórmula

$$na = a + \ldots + a \quad (n \text{ sumandos}).$$

Si 0 designa al número entero cero, se define $0a = 0$ (el 0 del segundo miembro es el cero del campo!)
si n es un entero negativo, se define

$$(-n)a = (-a) + \ldots + (-a) \quad (n \text{ sumandos}),$$

y por tanto se tiene que $(-n)a = -na$.

Definición 1.3.2. Se dice que un campo $\mathbb{F}$ tiene *característica* p, si p es el menor entero positivo tal que $p1 = 0$. Si $n1 \neq 0$ para todo entero positivo n, entonces se dice que $\mathbb{F}$ tiene característica 0.

Teorema 1.3.1. Si p es la característica de un campo, entonces p es un número primo.

Demostración. Si p fuera descomponible en la forma $p = qr$, donde q y r son enteros positivos diferentes de 1, entonces, como es fácil comprobar, se tendrá la relación

$$p1 = (qr)1 = (q1)(r1)$$

entre los elementos $p1$, $q1$ y $r1$ del campo. Por definición de característica se tendrá que $(q1)(r1) = 0$; y en virtud del corolario del teorema 1.1.1 sería o bien $q1 = 0$ o bien $r1 = 0$, y en ambos casos se llegaría a una contradicción pues q y r son menores que p. ■

Ejercicios

1.3.1. El campo de las clases de restos módulo el número primo p

tiene característica p.

1.3.2. Los campos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ tienen característica 0.

1.3.3. Sea $\mathbb{F}$ un campo. Para que $K \subset \mathbb{F}$ sea subcampo basta que:

1) Contenga un elemento no nulo,

2) Con a y b contenga a $a-b$ y ab^{-1}.

1.3.4. Se llama *campo primo* a un campo que no posee ningun sub campo propio. Demostrar que cada campo $\mathbb{F}$ contiene uno y solo un subcampo primo. (Considerar la intersección de todos los subcampos de $\mathbb{F}$).

1.3.5. Demostrar que si la característica de $\mathbb{F}$ es el número primo p el subcampo primo es isomorfo con el campo de clases de números enteros congruentes módulo p, y si la característica de $\mathbb{F}$ es 0 el subcampo primo es isomorfo con el campo $\mathbb{Q}$ de los números racionales.

1.3.6. En un cuerpo de característica p se cumplen la relaciones

$$(a + b)^p = a^p + b^p$$

$$(a - b)^p = a^p - b^p \quad ,$$

y también

$$(a_1 + a_2 + \ldots + a_n)^p = a_1^p + a_2^p + \ldots + a_n^p \; .$$

1.3.7. Aplicar la última relación del ejercicio anterior a la suma $1 + 1 + \ldots + 1$ módulo p, y obtener el teorema de Fermat: $a^{p-1} \equiv 1(p)$, para $a \not\equiv 0(p)$.

1.3.8. ¿De cuantos elementos consta un campo de característica p que es de grado n respecto de su campo primo?

§ 2. Espacios vectoriales

2.1. **Definición de espacio vectorial.** Un *espacio vectorial* V sobre un campo $\mathbb{F}$ cualquiera, es un grupo aditivo (abeliano) en el cual es definida además de la *adición*, la operación de *multiplicación* de los elementos del campo $\mathbb{F}$ por elementos del espacio V, de modo

que a cada par (α, x), en que $\alpha \in \mathbb{F}$ y $x \in V$, le corresponde un elemento $\alpha x \in V$, y se cumplen las siguientes propiedades:

1. *Propiedad asociativa:*

$(\lambda\mu)x = \lambda(\mu x)$, para $\lambda, \mu \in \mathbb{F}$ y $x \in V$

2. *Propiedades distributivas:*

$$(\lambda + \mu)x = \lambda x + \mu x$$
$$\lambda(x + y) = \lambda x + \lambda y \qquad \text{para } \lambda, \mu \in \mathbb{F} \text{ y } x, y \in V.$$

3. $1x = x$, donde 1 designa a la identidad de $\mathbb{F}$.

Los elementos de V se llaman *vectores* y los de $\mathbb{F}$, *escalares*. El vector nulo se representa por 0.

Aun cuando, salvo indicación especial, nuestras consideraciones son válidas para cualquier campo $\mathbb{F}$, puede pensarse que $\mathbb{F}$ coincide con el campo $\mathbb{R}$ de los números reales o con el campo $\mathbb{C}$ de de los números complejos.

Los espacios vectoriales sobre $\mathbb{R}$ se llaman *espacios vectoriales reales* y los espacios vectoriales sobre $\mathbb{C}$ se llaman *espacios vectoriales complejos*.

1) Ejercicios [1]

*2.1.1. Probar que la ecuación $\lambda x = 0$, donde $\lambda \in \mathbb{F}$ y $x \in V$, se cumple únicamente si $\lambda = 0$ o bien si $x = 0$.

*2.1.2. Demostrar que $\lambda(-x) = (-\lambda)x = -\lambda x$, para $\lambda \in \mathbb{F}$ y $x \in V$.

*2.1.3. Demostrar que cualesquiera que sean los escalres $\lambda_1, \ldots, \lambda_r$ y los vectores $x_1, \ldots, x_s$ se cumple:

$$(\lambda_1 + \ldots + \lambda_r)(x_1 + \ldots + x_s) = \sum_{i=1}^{r} \sum_{j=1}^{s} \lambda_i x_j .$$

*2.1.4. El conjunto $\mathbb{F}^n = \mathbb{F} \times \ldots \times \mathbb{F}$ (n factores) de la n tuplas

$$x = (\xi^1, \ldots, \xi^n) \quad , \quad \xi^i \in \mathbb{F} \ ,$$

con las operaciones

---

1) Los ejercicios que están marcados con asterisco deben ser necesariamente resueltos, porque contienen resultados complementarios, algunos de los cuales se utilizarán después.

$$(\xi^1, \ldots, \xi^n) + (\eta^1, \ldots, \eta^n) = (\xi^1 + \eta^1, \ldots, \xi^n + \eta^n)$$

y $\quad \lambda(\xi^1, \ldots, \xi^n) = (\lambda\xi^1, \ldots, \lambda\xi^n)$

constituye un espacio vectorial de suma importancia, que se llama *espacio cartesiano sobre $\mathbb{F}$ de dimensión n.* Los elementos $\xi^1, \ldots, \xi^n$ se llaman *componentes* del vector $(\xi^1, \ldots, \xi^n)$.

En particular consideremos los espacios $\mathbb{R}^n$ y $\mathbb{C}^n$ que son los *espacios cartesianos real y complejo de dimensión n*, el primero de los cuales tiene una conocida interpretación geométrica en los casos en que n tiene uno de los valores reales 1,2,3.

***2.1.5.** Se considera el conjunto $C[0,1]$ de las funciones reales continuas f definidas en el intervalo $I = [0,1]$.
Demostrar que $C[0,1]$ es un espacio vectorial con las operaciones $(f,g) \to f + g$ y $(\lambda, f) \to \lambda f$ $(\lambda \in \mathbb{R})$ definidas por

$$(f + g)(t) = f(t) + g(t) \qquad t \in I$$
$$(\lambda f)(t) = \lambda f(t)$$

***2.1.6.** Se considera un conjunto cualesquiera S. Dado el espacio vectorial V, el conjunto de todas las funciones $f: S \to V$ constituye un espacio vectorial con las operaciones

$$(f, g) \to f + g \qquad \text{y} \qquad (\lambda, f) \to \lambda f$$

definidas por

$$(f + g)(s) = f(s) + g(s), \quad (\lambda f)(s) = \lambda f(s) \quad (\forall s \in S)$$

***2.1.7.** Si λ es un escalar y x e y son vectores cualesquiera, probar que $\lambda(x - y) = \lambda x - \lambda y$.

***2.1.8.** Demostrar que todos los polinomios en la variable x, con coeficientes reales, forman un espacio vectorial con la adición y la multiplicación por un escalar definidas en la forma usual.

2.2. Dependencia e independencia lineal.

Definición 2.2.1. Sea $(x_\alpha)_{\alpha \in A}$ una familia de vectores de un espa-

cio vectorial V. Se llama *combinación lineal* de vectores de la familia a cada expresión de la forma

$$\lambda_{\alpha_1} x_{\alpha_1} + \ldots + \lambda_{\alpha_n} x_{\alpha_n} = \sum_{i=1}^{n} \lambda_{\alpha_i} x_{\alpha_i} ,$$

donde $(\alpha_1, \ldots, \alpha_n)$ es un subconjunto cualquiera de A y $\lambda_{\alpha_1}, \ldots, \lambda_{\alpha_n}$ son escalares arbitrarios.

Si x es tal que

$$x = \sum_{i=1}^{n} \lambda_{\alpha_i} x_{\alpha_i} ;$$

decimos que x es una *combinación lineal de vectores* de la familia. Es esencial observar que las sumas anteriores tienen un número finito de términos.

Algunas veces representaremos una combinación lineal en la forma simplificada $\sum_{\alpha \in A} \lambda_\alpha x_\alpha$, y aún por $\sum \lambda_\alpha x_\alpha$.

Resulta de la precedente definición que el vector nulo es combinación lineal de cualquier familia de vectores. Además es claro que si x es combinación lineal de una familia de vectores también es combinación lineal de alguna de sus subfamilias finitas.

Si la familia considerada es finita, podemos expresarla en la forma $(x_1, \ldots, x_p)$, y entonces una combinación lineal de los vectores de esa familia podrá escribirse en la forma

$$\sum_{i=1}^{p} \lambda_i x_i$$

Definición 2.2.2. Un subconjunto S de un espacio vectorial V, se llama *sistema de generadores* de V, si cada vector $x \in V$ es una combinación lineal de vectores de S. Evidentemente, V constituye un sistema de generadores de sí mismo.

Definición 2.2.3. Una combinación lineal $\sum_{\alpha \in A} \lambda_\alpha x_\alpha$ de vectores de una familia $(x_\alpha)_{\alpha \in A}$ se dice que es *no trivial* si alguno de los coeficientes λ_α es diferente de cero. En caso contrario se dice que es *tri-*

vial. Debe observarse que una combinación trivial es igual al vector nulo y que una combinación no trivial puede o no ser igual al vector nulo.

Ejemplo A. En $\mathbb{R}^3$, dados los vectores $x_1 = (-2,3,4)$, $x_2 = (-3,1,2)$ $x_3 = (5,3,2)$, su combinación lineal $4x_1 - 6x_2 - 2x_3$ es no trivial, y es igual al vector nulo $(0,0,0)$.

Definición 2.2.4. Se dice que una familia (x_α) es *linealmente dependiente*, o bien que sus vectores son *linealmente dependientes* (entre sí), si el vector nulo es una combinación lineal no trivial de de sus elementos; es decir si existe el sistema de escalres (λ_α) tales que

$$\sum_\alpha \lambda_\alpha x_\alpha = 0 \quad , \qquad (2.2.1)$$

y que por lo menos uno de los escalares λ_α es diferente de cero. Se dice en ese caso que (2.2.1) es una *relación lineal no trivial* entre los elementos de (x_α).

Ejemplo B. Entre los vectores (x_1, x_2, x_3) de $\mathbb{R}^3$, considerados en el ejemplo A, se cumple la relación lineal no trivial

$$4x_1 - 6x_2 - 2x_3 = 0$$

Definición 2.2.5. Una familia $(x_\alpha)_{\alpha \in A}$ de vectores, se llama *linealmente independiente*, y se dice entonces que los vectores de la familia *son linealmente independientes* (entre sí), si dicha familia no es linealmente dependiente, o sea si y solo si la ecuación

$$\sum_\alpha \lambda_\alpha x_\alpha = 0$$

implica que $\lambda_\alpha = 0$ para cada $\alpha \in A$; es decir que la única combinación lineal igual al vector nulo es la combinación lineal trvial, o bien que entre los vectores de la familia no tiene lugar ninguna relación no trivial.

En adelante emplearemos las iniciales *l.d.* y *l.i.* en lugar de las expresiones *linealmente dependiente(s)* y *linealmente independiente(s)*.

Ejercicios

*2.2.1. Si x es una combinación lineal de la familia $(x_\alpha)_{\alpha\in A}$ y cada x_α es una combinación lineal de una familia $(y_{\alpha\beta})_{\beta\in B}$, entonces x es una combinación lineal de los $y_{\alpha\beta}$.

*2.2.2. Sea S un sistema de generadores de V. Si cada vector de S es una combinación lineal de vectores de un subconjunto $T \subset S$, entonces T es también un sistema de generadores de V.

*2.2.3. Si la familia de vectores $(x_\alpha)_{\alpha\in A}$ es *l.i.*, ninguno de esos vectores es nulo, y entre ellos no hay dos iguales.

*2.2.4. Una familia de vectores $(x_\alpha)_{\alpha\in A}$ es *l.d.*, si y solo si, para algún $\beta \in A$, el vector x_β es una combinación lineal de los vectores x_α $\alpha \neq \beta$. En particular, dos vectores x e y son *l.d.* si y solo si se cumple una de la relaciones $x = \lambda y$ o bien $y = \lambda x$, para algún escalar λ.

*2.2.5. Una subfamilia cualquiera de una familia *l.i.* de vectores es también *l.i.*

*2.2.6. Una familia $(x_\alpha)_{\alpha\in A}$ de vectores es *l.i.* si y solo si cada vector x puede escribirse a lo más de una manera como combinación lineal de los x_α, es decir si y solo si la relación

$$\Sigma\, \lambda_\alpha x_\alpha = \Sigma\, \mu_\alpha x_\alpha$$

implica que $\lambda_\alpha = \mu_\alpha$ para cada $\alpha \in A$. (Tener presente que los λ_α son nulos, salvo a lo más para índices α pertenecientes a un subconjunto finito de A).

2.2.7. Probar que los vectores $(1,0)$, $(0,1)$ y $(2,4)$ de $\mathbb{R}^2$ son *l.d.* y que dos cualesquiera de ellos son *l.i.*

2.2.8. Averiguar si los tres vectores $(2,-1,3,2)$, $(-1,1,1,-3)$ y $(1,1,9,-5)$ de $\mathbb{R}^4$ son o no *l.i.*

2.2.9. Obtener una relación de dependencia lineal no trivial entre

los vectores $v_1 = (1,0,1)$, $v_2 = (0,1,1)$, $v_3 = (1,1,0)$, $v_4 = (2,-1,3)$ de $\mathbb{R}^3$.

2.2.10. Mostrar que cada solución (x_1, x_2, x_3, x_4) del sistema de ecuaciones

$$x_1 - x_3 + 2x_4 = 0$$
$$x_2 + 3x_3 + x_4 = 0 ,$$

considerada como vector de $\mathbb{R}^4$ es *l.d.* de las soluciones $(-2,-1,0,1)$ y $(1,-3,1,0)$ del sistema.

***2.2.11.** Probar que si S es un sistema de vectores *l.i.*, v es un vector, y el sistema $\{S,v\}$ es *l.d.*, entonces v es *l.d.* de S, es decir que es combinación lineal de S.

2.2.12. Dado el sistema *l.i.* de vectores $(x_1, \ldots, x_p)$ probar que el sistema $(x_1, \ldots, x_{i-1}, x_i + \lambda x_j, x_{i+1}, \ldots, x_p)$ $i \neq j$, con λ un escalar arbitrario, es tambien *l.i.*

2.2.13. Probar que el conjunto de todas las soluciones de la ecuación diferencial lineal homogénea

$$\frac{d^2y}{dx^2} + p\frac{dy}{dx} + qy = 0 ,$$

donde p y q son funciones dadas de x, es un espacio vectorial.

2.2.14. ¿Cuáles de los siguientes conjuntos de funciones del espacio del ejercicio 2.1.5 son *l.i.*?

a) $f_1 = 3t$, $f_2 = t + 5$, $f_3 = 2t^2$, $f_4 = (t+1)^2$.

b) $f_1 = (t+1)^2$, $f_2 = t^2 - 1$, $f_3 = 2t^2 + 2t - 3$.

c) $f_1 = 1$, $f_2 = e^t$, $f_3 = e^{-t}$.

d) $f_1 = t^2$, $f_2 = t$, $f_3 = 1$

e) $f_1 = 1 - t$, $f_2 = t(1-t)$, $f_3 = 1 - t^2$

2.3. Subespacios vectoriales. Un subconjunto no vacío V_1 de un espacio vectorial V es un *subespacio* (vectorial) de V si

a) $x \in V_1$, $y \in V_1$ $\Rightarrow$ $x + y \in V_1$

b) $x \in V_1$, $\alpha \in F$ $\Rightarrow$ $\alpha x \in V_1$.

En tal caso V_1 es también un espacio vectorial respecto a las mismas operaciones vectoriales que son dadas en V.

Si S es un subconjunto de V, el conjunto de todos los vectores de este espacio que son combinaciones lineales de los vectores de S constituye un subespacio que se llama *subespacio generado por* S. Evidentemente S es un sistema de generadores de ese subespacio (Ejercicio 2.3.1).

Definición 2.3.1. Decimos que dos subespacios V_1 y V_2 de un espacio vectorial V son *disjuntos*, si se cumple que

$$V_1 \cap V_2 = \{0\} ,$$

es decir si V_1 y V_2 no tienen otro elemento común que el vector cero.

Ejercicios

*2.3.1. Demostrar que si S es un subconjunto del espacio vectorial V, el conjunto V_1 de todos los vectores de V que son combinaciones lineales de vectores de S, es un subespacio de V.

*2.3.2. Sean U y V dos subespacios de un mismo espacio vectorial. Demostrar que

a) La intersección $U \cap V$ es un subespacio

b) La suma $U + V = \{x \mid x = u + v \ , \ u \in V\}$ es un subespacio.

2.4. Espacio vectorial factor o cociente. Si V_1 es un subespacio del espacio vectorial V, la relación

$$x \sim x' \Leftrightarrow x - x' \in V_1 ,$$

entre los vectores de V, es una relación de equivalencia. (Demostrar lo!).

El conjunto de las clases de equivalencia relativas a la relación

que acabamos de definir es un espacio vectorial si las operaciones se definen de la manera siguiente: Dada una clase de equivalencia k, designemos por v_k a un vector cualquiera perteneciente a la clase k; y dado un vector $v \in V$, designemos por $\bar{v}$ a la clase de equivalencia que lo contiene. Definiremos entonces la suma de dos clases y el producto de un escalar por una clase mediante las fórmulas

$$k_1 + k_2 = \overline{v_{k_1} + v_{k_2}}$$
$$\alpha k = \overline{\alpha v_k}$$

Estas definiciones son adecuadas por cuanto definen en el conjunto de clases de equivalencia operaciones que son independientes de los vectores particulares que se elijan en las clases que intervienen. En efecto, si

$$v'_{k_1} - v_{k_1} \in V_1 \quad \text{y} \quad v'_{k_2} - v_{k_2} \in V_1$$

entonces se tiene

$$(v'_{k_1} + v'_{k_2}) - (v_{k_1} + v_{k_2}) = (v'_{k_1} - v_{k_1}) + (v'_{k_2} - v_{k_2}) \in V_1$$

y por tanto

$$\overline{v'_{k_1} + v'_{k_2}} = \overline{v_{k_1} + v_{k_2}} \ ;$$

y semejantemente, si

$$v'_k - v_k \in V_1$$

se tiene

$$\alpha v'_k - \alpha v_k = \alpha(v'_k - v_k) \in V_1$$

de donde

$$\overline{\alpha v'_k} = \overline{\alpha v_k} \quad .$$

(Demostrar que las operaciones que hemos definido con las clases de equivalencia cumplen los axiomas de los espacios vectoriales).

El espacio vectorial constituido por las clases de equivalencias

relativas al subespacio V_1, se designa por V/V_1 y se llama *espacio vectorial factor de V respecto de V_1* o bien *espacio cociente de V por V_1*.

La aplicación $v \mapsto \overline{v}$ es un homomorfismo del grupo aditivo V sobre el grupo aditivo V/V_1, pero además goza de la propiedad de que al vector αv le corresponde el vector $\alpha\overline{v}$, por lo cual podemos decir que es un *homomorfismo de los espacios vectoriales* V y V/V_1 al que llamaremos *homomorfismo canónico* de V sobre V/V_1.

Ejercicios

***2.4.1.** Demostrar que la aplicación $v \mapsto \overline{v}$ de V sobre el espacio factor V/V_1 tiene las propiedades

$$v_1 + v_2 \mapsto \overline{v_1} + \overline{v_2}\ , \quad \alpha v \mapsto \alpha\overline{v}\ ,$$

y por tanto es un homomorfismo del espacio vectorial V sobre el espacio vectorial V/V_1.

2.4.2. ¿Cuáles de los siguientes subconjuntos de $\mathbb{R}^4$ son subespacios?

a) el conjunto de los vectores (a_1, a_2, a_3, a_4) tales que $a_1 = 0$

b) el conjunto de los vectores para los que $a_1 = 2$,

c) el conjunto de los vectores para los que $a_1 = 2a_2$

d) el conjunto de los vectores para los cuales los números a_1, a_2, a_3, a_4 son soluciones del sistema de ecuaciones

$$2y_1 - y_4 = 0$$
$$y_2 = 0$$

e) el conjunto de los vectores para los cuales es:

$$a_1 + a_2 + a_3 = 1.$$

2.4.3. Describir los espacios factores de $\mathbb{R}^4$ respecto de los subespacios que resultan del ejercicio anterior.

2.4.4. Demostrar que un subconjunto no vacío V_1 de un espacio vectorial V es un subespacio si y solo si las combinaciones lineales $\alpha x + \beta y$, donde α y β son escalares cualesquiera y x e

y son vectores arbitrarios de V_1, pertenecen también a V_1.

2.4.5. Demostrar que un subconjunto de un espacio vectorial es un subespacio si y solo si coincide con el subespacio generado por sus elementos.

*2.4.6. Demostrar que el subespacio generado por un subconjunto S de un espacio vectorial es la intersección de todos los subespacios de V que contienen a S.

*2.4.7. Dados dos subespacios U y V, de un espacio vectorial, demostrar que $U + V$ es el espacio generado por $U \cup V$.

2.4.8. Demostrar que los siguientes subconjuntos de vectores (x_1, x_2, x_3) de $\mathbb{R}^3$ constituyen subespacios

a) El conjunto de los vectores cuyas componentes satisfacen a la ecuación

$$c_1 x_1 + c_2 x_2 + c_3 x_3 = 0$$

donde c_1, c_2 y c_3 son escalares dados.

b) El conjunto de aquellos cuyas componentes satisfacen al sistema de ecuaciones

$$c_1 x_1 + c_2 x_2 + c_3 x_3 = 0$$

$$d_1 x_1 + d_2 x_2 + d_3 x_3 = 0,$$

donde los c_i y los d_i son escalares dados.

Representados los vectores de $\mathbb{R}^3$ en la forma conocida, mediante puntos del espacio geométrico ordinario, describir los subespacios aquí considerados.

2.4.9. Caracterizar los siguientes espacios factores de $\mathbb{R}^3$ respecto de los subespacios del ejercicio anterior.

2.5. Base de un espacio vectorial. Dimensión. Una familia de vectores $(x_\alpha)_{\alpha \in A}$ de un espacio vectorial V se llama *base* del espacio si es un sistema de generadores del espacio y al mismo tiempo es *l.i.*

La definición 2.2.2 de un sistema de generadores y la propo-

sición del ejercicio 2.2.6 permite afirmar que el conjunto $(x_\alpha)_{\alpha \in A}$ es una base si y solo si cada vector $x \in V$ puede ser escrito de una y solo una manera en la forma

$$x = \sum_{\alpha} \xi^\alpha x_\alpha \quad , \quad (\xi^\alpha \in \mathbb{F}),$$

donde los escalares ξ^α (Observar aquí que los α son índices y no exponentes; una potencia de ξ^α se expresará en la forma $(\xi^\alpha)^n$), que se llaman *componentes* de x respecto de la base $(x_\alpha)_{\alpha \in A}$ son todos nulos, salvo a lo más un número finito de ellos [1].

Lema 2.5.1. Si los vectores $(x_\alpha)_{\alpha \in A}$ son *l.i.* , y $a = \sum_{\alpha \in A} \xi^\alpha x_\alpha$ es un vector para el cual es $\xi^\beta \neq 0$ para algún $\beta \in A$, entonces los vectores

$$\{(x_\alpha)_{\alpha \in A - \{\beta\}} , \ a \} \qquad (2.5.1)$$

son *l.i.*

Demostración. Supongamos que entre los vectores (*2.5.1*) tiene lugar una relación lineal

$$\sum_{\alpha \neq \beta} \lambda^\alpha x_\alpha + \lambda a = 0 \qquad (2.5.2)$$

sustituyendo aquí a por su valor dado en el enunciado se obtiene

$$\sum_{\alpha \neq \beta} (\lambda^\alpha + \lambda \xi^\alpha) x_\alpha + \lambda \xi^\beta x_\beta = 0$$

Puesto que los $(x_\alpha)_{\alpha \in A}$ son *l.i.* se tiene entonces

$$\lambda^\alpha + \lambda \xi^\alpha = 0 \quad \text{para} \quad \alpha \neq \beta \quad \text{y} \quad \lambda \xi^\beta = 0 \ .$$

Por hipótesis es $\xi^\beta \neq 0$; luego se tiene $\lambda = 0$ y $\lambda^\alpha = 0$ para $\alpha \neq \beta$. Por consiguiente la relación (2.5.2) es trivial. ∎

---

1) Algunas veces se llama *componentes* de x respecto de la base (x_α) a los vectores $\xi^\alpha x_\alpha$ que no son nulos, cuya suma es x, y a los escalares ξ^α se les llama *coordenadas* de x.

Corolario. Si los vectores $(x_\alpha)_{\alpha \in A}$ del espacio V constituyen una base y

$$a = \sum_{\alpha} \xi^\alpha x_\alpha$$

es un vector tal que $\xi^\beta \neq 0$ para algún $\beta \in A$, entonces

$$\{(x_\alpha)_{\alpha \in A - \{\beta\}}, a\} \qquad (2.5.3)$$

constituye también una base de V.

Demostración. Según el lema anterior los vectores (2.5.3) son *l.i.*. Solo falta probar que generan a V. Para eso haremos uso de la proposición del ejercicio 2.2.2. Puesto que $(x_\alpha)_{\alpha \in A}$ es un sistema de generadores, el conjunto

$$S = \{(x_\alpha)_{\alpha \in A}, a\}$$

también lo será. El subconjunto

$$T = \{(x_\alpha)_{\alpha \in A - \{\beta\}}, a\}$$

de S es también un sistema de generadores porque cada vector de S es combinación lineal de vectores de T. (En particular es

$$x_\beta = -\sum_{\alpha \neq \beta} \frac{\xi^\alpha}{\xi^\beta} x_\alpha + \frac{1}{\xi^\beta} a \,. \qquad \blacksquare$$

Teorema 2.5.1. *(Teorema de intercambio de Steinitz).*

Sea $(x_\alpha)_{\alpha \in A}$ una base del espacio vectorial V, y sea $(a_1, \ldots, a_p)$ un subconjunto de vectores *l.i.* del mismo espacio. Entonces es posible sustituir a p de los vectores de la base dada por los vectores $a_1, \ldots, a_p$ de modo que el conjunto que resulta sea una nueva base de V. Si el número de vectores de la base dada es finito e igual a m, entonces es necesariamente $p \leqslant m$, es decir que el espacio no puede contener más de m vectores *l.i.*

Demostración. En la expresión $a_1 = \sum_{\alpha \in A} \xi^\alpha x_\alpha$ del vector a_1 en

términos de la base (x_α), uno por lo menos de los coeficientes ξ^α es diferente de cero, pues en caso contrario sería $a_1 = 0$ y los vectores $a_1, \ldots, a_p$ no serían *l.i.* Supongamos que sea $\xi^{\alpha_1} \neq 0$, donde α_1 es un elemento de A. Del corolario del lema 2.5.1 resulta que el conjunto $\{(x_\alpha)_{\alpha \neq \alpha_1}, a_1\}$ es una base de V. Por tanto existen los escalares η' y η^α $(\alpha \neq \alpha_1)$ tales que

$$a_2 = \eta^1 a_1 + \sum_{\alpha \neq \alpha_1} \eta^\alpha x_\alpha .$$

En esta ecuación uno por lo menos de los coeficientes η^α $(\alpha \neq \alpha_1)$ debe ser diferente de cero, pues si no, los vectores a_1, a_2 serían *l.d.* y los vectores $a_1, \ldots, a_p$ no serían *l.i.* Sea $\eta^{\alpha_2} \neq 0$. El corolario del lema 2.5.1. permite establecer entonces que $\{(x_\alpha)_{\alpha \neq \alpha_1, \alpha_2}, a_1, a_2\}$ es una base. Si la base $(x_\alpha)_{\alpha \in A}$ es infinita, el procedimiento descrito podrá ser aplicado hasta que se agoten los vectores a_i.

Si la base $(x_\alpha)_{\alpha \in A}$ fuera finita y m fuera el número de sus elementos, entonces debería ser necesariamente $p \leqslant m$, pues si fuera $p > m$, el procedimiento que hemos descrito podría seguir siendo empleado m veces consecutivas hasta obtener que $(a_1, \ldots, a_m)$ resultara siendo una base, lo que no es posible porque entonces los vectores $a_{m+1}, \ldots, a_p$ dependerían linealmente de los vectores a_1, $a_2, \ldots, a_m$ y los vectores $(a_1, \ldots, a_p)$ no serían *l.i.* Se sigue de aquí que $p \leqslant m$; y por tanto, si el espacio tiene una base formada por m elementos, no puede contener más de m vectores *l.i.*

De lo dicho resulta que el procedimiento que antes hemos empleado puede ser aplicado p veces consecutivas y se obtendrá así que el conjunto

$$\{(x_\alpha)_{\alpha \neq \alpha_1, \ldots, \alpha_p}, a_1, a_2, \ldots, a_p\},$$

donde eventualmente el conjunto $(x_\alpha)_{\alpha \neq \alpha_1, \ldots, \alpha_p}$ podría ser vacío,

constituye una nueva base de V. ■

Corolario 1. Si un espacio vectorial tiene una base finita, entonces todas las bases de ese espacio tienen el mismo número de elementos.

Demostración. Si $(u_1, \ldots, u_m)$ y $(v_1, \ldots, v_{m'})$ son dos bases de un espacio vectorial, los vectores de cada una de ellas son *l.i.* , y por tanto, aplicando dos veces el teorema anterior, resultara que $m \leqslant m'$ y que $m' \leqslant m$; de donde se sigue que $m = m'$. ■

La siguiente proposición nos va a permitir demostrar luego que todo espacio vectorial posee una base.

Teorema 2.5.2. Sea V un espacio vectorial. Supongamos que S es un sistema de generadores de V, y que R es un subconjunto *l. i.* de vectores de S. Entonces existe una base T de V, tal que $R \subset T \subset S$.

Demostración. En el conjunto $\mathbb{P}(S)$ de todos los subconjuntos *l. i.* de S que contienen a R, la relación de inclusión permite introducir un orden parcial[1]. Es evidente que $R \in \mathbb{P}(S)$. Si $\{S_\alpha\}_{\alpha \in A}$ es

1) Sea P un conjunto no vacío. Se dice que P es *parcialmente ordenado* por la relación $\leqslant$ entre elementos de P si se cumplen las propiedades siguientes:

(1) $x \leqslant y$, para cada $x \in P$ *(reflexividad)*

(2) $x \leqslant y$ e $y \leqslant x \Rightarrow x = y$ *(antisimetría)*

(3) $x \leqslant y$ e $y \leqslant z \Rightarrow x \leqslant z$ *(transitividad)*

Si P es un conjunto de partes de un conjunto dado, la relación de inclusión expresada con el signo $\subset$, es una relación de orden parcial.

Dado un subconjunto S de un conjunto parcialmente ordenado P, se dice que $a \in P$ es una *cota superior de* S, si se cumple que $x \leqslant a$, $\forall\, x \in S$.

Se dice que un subconjunto S de un conjunto parcialmente ordenado P es una *cadena* si cualesquiera que sean $x, y \in S$ es o bien $x \leqslant y$ o bien $y \leqslant x$.

Lema de Zorn. Si cada cadena de un conjunto parcialmente ordenado P tiene una cota superior, entonces P posee un elemento *maximal*, es decir un elemento x tal que si $y \geqslant x$ entonces es $y = x$.

(Ver por ejemplo; Simmons. Introduction to Topology and Modern Analysis, pág. 43. También; Holman, Lineare und Multilineare Algebra I, Hochschuletaschenbücher, Bibliographisches Institut, pág 18).

una cadena de tales subconjuntos, entonces

$$\bigcup_{\alpha \in A} S_\alpha \in \mathbb{P}(S) \qquad (2.5.3)$$

En efecto, es claro en primer lugar, que $R \subset \bigcup_\alpha S_\alpha \subset S$, y si suponemos que entre los elementos $x_1, \ldots, x_n$ de $\bigcup_\alpha S_\alpha$ se cumple la relación lineal

$$\sum_{i=1}^{n} \lambda^i x_i = 0 \qquad (2.5.4)$$

entonces, para cada valor de i es $x_i \in S_{\alpha_i}$ para algún $\alpha_i \in A$. Puesto que $\{S_\alpha\}$ es una cadena, podemos suponer que

$$S_{\alpha_i} \subset S_{\alpha_1} \qquad (i = 1,2,\ldots,n) ,$$

es decir que S_{α_1} es aquél de los subconjuntos $S_{\alpha_1}, \ldots, S_{\alpha_n}$ que contiene a todos los demás. Se tiene por tanto que

$$x_i \in S_{\alpha_1} \qquad (i = 1, \ldots, n).$$

Ahora bien, S_{α_1} es un conjunto *l.i.*; luego, de (2.5.4) se sigue que $\lambda^i = 0$ $(i = 1, \ldots, n)$, y por consiguiente el conjunto $\bigcup_\alpha S_\alpha$ es *l.i.* y se cumple (2.5.3). Dado que para todo $\alpha \in A$ es

$$S_\alpha \subset \bigcup_\alpha S_\alpha ,$$

resultará que $\bigcup_\alpha S_\alpha$ es una cota superior de la cadena $\{S_\alpha\}$ del conjunto parcialmente ordenado $\mathbb{P}(S)$.

Podemos aplicar ahora a $\mathbb{P}(S)$ el lema de Zorn, queda así probada la existencia de un elemento maximal T de $\mathbb{P}(S)$, es decir de un subconjunto T de S que es *l.i.*, y que no está contenido propiamente en ningún subconjunto de S que contenga a R, y que sea *l.i.*

Para probar que T es una base de V basta únicamente probar que T es un sistema de generadores de V. Puesto que S es un sistema de generadores de V y que $T \subset S$ será suficiente que pro-

bemos que cada elemento de S es una combinación lineal de vectores de T (ver la proposición del ejercicio 2.2.2). Sea x un elemento cualquiera de S. Si $x \in T$ nuestra afirmación es obvia. Supongamos pues que sea $x \notin T$. El conjunto $\{x\} \cup T$ contiene a R y está contenido en S, pues T cumple esas dos condiciones y $x \in S$. Luego, dicho conjunto es *l.d.*, pues si no fuera así pertenecería a $\mathbb{P}(S)$ y T no sería maximal. Existe por consiguiente una relación no trivial

$$\lambda x + \sum_{i=1}^{n} \lambda^i x_i = 0 \quad , \quad (\lambda, \lambda^i \in \mathbb{F}, \; x_i \in T). \quad (2.5.5)$$

Necesariamente debe ser $\lambda \neq 0$, por que si fuera $\lambda = 0$ la independencia lineal de T implicaría que $\lambda^i = 0$ para todo i, y entonces la precedente relación sería trivial. Multiplicando la ecuación (2.5.5.) por λ^{-1} se obtiene

$$x = - \sum_{i=1}^{n} \lambda^{-1} \lambda^i x_i \; ;$$

con lo cual queda probado que cada elemento de S es una combinación lineal de elementos de T. Se deduce que T genera a V, y por tanto es una base de V que, por pertenecer a $\mathbb{P}(S)$, cumple la condición $R \subset T \subset S$. ■

Corolario 1. Cada sistema de generadores de un espacio vectorial no trivial contiene a una base del mismo.

Demostración. Puesto que el espacio vectorial considerado V no es trivial, y que S es un conjunto de generadores, existe en S por lo menos un vector no nulo x. Este vector constituye un sistema linealmente independiente $R = \{x\}$. Luego, en virtud del teorema, existe una base T de V tal que es $R \subset T \subset S$. ■

Corolario 2. Cada conjunto *l.i.*, R, de V puede ser completado hasta formar una base de V.

Demostración. Basta con suponer $S = V$ y aplicar el teorema 2.5.2. ■

Corolario 3. Cada espacio vectorial no trivial tiene una base.

Demostración. Basta considerar $R = \{x\}$ donde x es un vector no nulo de V, y aplicar el corolario 2. ∎

Definición 2.5.1. Se dice que un espacio es de *dimensión finita* si posee una base finita. En este caso el número n de elementos de cada una de sus bases (corolario 1 del teorema 2.5.1) se llama *dimensión* del espacio, y se dice que el espacio tiene dimensión n o que es *n-dimensional*. Es evidente que todo subespacio de un espacio de dimensión finita es también de dimensión finita no mayor que la dimensión del espacio; y que es igual a ella unicamente si el subespacio es idéntico al espacio. Si el espacio no tiene base finita se dice que es de *dimensión infinita*.

De las proposiciones precedentemente demostradas resulta que para que un espacio sea de dimensión finita es necesario y suficiente que admita un sistema finito de generadores. La dimensión es entonces igual al número mínimo de elementos generadores o al número máximo de vectores *l.i.*

Ejercicios

***2.5.1.** Si $(x_\alpha)_{\alpha \in A}$ es una *familia minimal de generadores* del espacio vectorial V, es decir que no existe ninguna familia de generadores de V contenida propiamente en la familia dada, entonces $(x_\alpha)_{\alpha \in A}$ es una base de V. En particular, si V contiene a la familia minimal finita de generadores $(x_1, \ldots, x_n)$ entonces esta familia es una base de V y la dimensión es n.

***2.5.2.** Si $(x_\alpha)_{\alpha \in A}$ es una *familia maximal l.i.* del espacio vectorial V, es decir que no existe ninguna otra familia *l.i.* de V que contenga propiamente a la familia dada, entonces $(x_\alpha)_{\alpha \in A}$ es una base de V. En particular, si V contiene a la familia maximal finita *l.i.* $(x_1, \ldots, x_n)$, entonces esta familia es una base, y n es la dimensión de V.

***2.5.3.** Sea $S = (x_1, \ldots, x_m)$ un sistema finito de generadores de V, y supongamos que los vectores $x_1, \ldots, x_r$ $(r \leqslant m)$ son

linealmente independientes. Entonces existe una base de V contenida en S y que contiene a los vectores $x_1, \ldots, x_r$. (Esta proposición es el caso particular del teorema 2.5.2, en que S es finito. Para demostrarla no es necesario hacer uso del lema de Zorn).

*2.5.4. En un espacio vectorial de dimensión n cada conjunto de n vectores *l.i.* constituye una base.

*2.5.5. Si $(a_1, \ldots, a_p)$ es un conjunto de vectores *l.i.* del espacio V, pueden elegirse los vectores $(a_\alpha)_{\alpha \in A}$ de manera que los vectores $a_1, \ldots, a_p, (a_\alpha)_{\alpha \in A}$ constituyan una base de V.

* 2.5.6. Probar que el conjunto $S = \{(1,0,0),(0,1,0),(0,0,1)\}$ de vectores de $\mathbb{F}^3$, es una base de este espacio y por tanto

$$\dim \mathbb{F}^3 = 3 .$$

(En particular, $\mathbb{R}^3$ y $\mathbb{C}^3$ son tridimensionales).

*2.5.7. Generalizar el resultado del ejercicio anterior para los espacios cartesianos $\mathbb{F}^n$ ($n = 1,2, \ldots$). (En particular, $\mathbb{R}^n$ y $\mathbb{C}^n$ son n-dimensionales).

*2.5.8. Consideremos el sistema de ecuaciones

$$\begin{aligned}
\alpha_1^1 \xi^1 + \alpha_2^1 \xi^2 + \ldots + \alpha_m^1 \xi^m &= 0 \\
\alpha_1^2 \xi^1 + \alpha_2^2 \xi^2 + \ldots + \alpha_m^2 \xi^m &= 0 \\
\vdots \qquad\quad \vdots \qquad\qquad & \vdots \\
\alpha_1^n \xi^1 + \alpha_2^n \xi^2 + \ldots + \alpha_m^n \xi^m &= 0 ,
\end{aligned}$$

donde suponemos que todos los coeficientes son elementos dados, α_j^i, del campo $\mathbb{F}$. Decimos entonces que esas ecuaciones constituyen un *sistema homogéneo de n ecuaciones lineales con m incógnitas* sobre el campo $\mathbb{F}$ y cada vector $(\xi^1, \ldots, \xi^m)$ de $\mathbb{F}^n$, cuyas componentes satisfacen a esas ecuaciones constituyen una *solución* del sistema. Se trata de demostrar que las soluciones del sistema homogéneo considerado constituyen un subespacio de $\mathbb{F}^m$.

2.5.9. Hallar la dimensión del espacio vectorial (ver el ejercicio anterior) de todas las soluciones contenidas en $\mathbb{R}^4$, del sistema homogéneo

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

¿Cambiará el resultado si consideramos a este sistema sobre el campo $\mathbb{C}^4$ en vez de $\mathbb{R}^4$?

2.5.10. Probar que el espacio de polinomios del ejercicio 2.1.3 no tiene dimensión finita. Lo mismo para el espacio de funciones del ejercicio 2.1.5.

2.5.11. Probar que el espacio $\mathbb{R}^3$ no puede tener una base que contenga a los vectores $(1,2,3)$ y $(2,4,6)$.

2.5.12. Hallar una base del subespacio de $\mathbb{R}^3$ generado por los vectores $(2,-3,4)$, $(4,-6,8)$, $(6,-8,12)$ y $(9,-11,18)$.

2.5.13. Sea U_1 el subespacio de $\mathbb{R}^3$ generado por el conjunto

$$S_1 = \{(1,0,0),(0,1,0)\} ,$$

y sea U_2 el subespacio generado por

$$S_2 = \{(0,1,2) , (2,1,3)\} .$$

Obtener una base del subespacio $U_1 \cap U_2$ y una base del subespacio $U_1 + U_2$.

***2.5.14.** Dados dos espacios U y V sobre el campo $\mathbb{F}$ el producto cartesiano $U \times V$ constituye un espacio vectorial con las operaciones

$$(u_1 , v_1) + (u_2 , v_2) = (u_1 + u_2 , v_1 + v_2)$$
$$\alpha(u , v) = (\alpha u , \alpha v).$$

Si se identifica el par $(u , 0)$ con el vector $u \in U$ y el par $(0 , v)$ con el vector $v \in V$, entonces puede verse que el espacio que hemos construido contiene como subespacios a U y a V. Ese espacio se llama entonces *suma directa* de U

y V y se representa por $U \oplus V$. Cada elemento de $U \oplus V$ puede escribirse de una sola manera en la forma $u + v$ $(u \in U\ ,\ v \in V)$. Las dimensiones cumplen la relación

$$dim\,(U \oplus V) = dim\,U + dim\,V.$$

*2.5.15. Demostrar que si V es de dimensión finita y V_1 es un subespacio de V entonces

$$dim\,V/V_1 = dim\,V - dim\,V_1$$

La dimensión de V/V_1 se llama *codimensión* de V_1 y se designa por $codim\,V_1$. Luego:

$$dim\,V_1 + codim\,V_1 = dim\,V.$$

*2.5.16. Si V_1 y V_2 son subespacios de dimensión finita de un espacio vectorial, entonces

$$dim\,V_1 + dim\,V_2 = dim\,(V_1 + V_2) + dim\,(V_1 \cap V_2),$$

es decir que la suma de las dimensiones de dos subespacios es igual a la dimensión de la suma de esos subespacios más la dimensión de su intersección. Si en particular es $V_1 \cap V_2 = \{0\}$ (en cuyo caso decimos, como ya vimos, que los subespacios son disjuntos) resultará que

$$dim\,(V_1 + V_2) = dim\,V_1 + dim\,V_2.$$

2.5.17. ¿Para qué valores de α los vectores $x = (1,-2)$ e $y = (\alpha,-1)$ forman una base de $\mathbb{R}^2$?

*2.5.18. Sean U y V dos espacios n-dimensionales sobre $\mathbb{F}$, y sea ϕ una aplicación biyectiva de U sobre V tal que

$$\phi(u_1 + u_2) = \phi(u_1) + \phi(u_2)$$

$$\phi(\alpha u) = \alpha\phi(u)$$

Probar que si $(e_1, \ldots, e_n)$ es una base de U, entonces $(\phi(e_1), \ldots, \phi(e_n))$ es una base de V; y que si $(f_1, \ldots, f_n)$ es una base de V, entonces $(\phi^{-1}(f_1), \ldots, \phi^{-1}(f_n))$ es una base de U. Es decir que ϕ determina una correspondencia biunívoca entre las bases de los espacios U y V.

§ 3. Transformaciones lineales

3.1. Definiciones. Dados los espacios vectoriales U y V sobre el campo $\mathbb{F}$, una aplicación T: $U \to V$ es una *transformación lineal* u *homomorfismo* de U en V si se cumplen las condiciones:

a) $T(u_1 + u_2) = T(u_1) + T(u_2)$, para $u_1, u_2 \in U$,

b) $T(\alpha u) = \alpha T(u)$, para $\alpha \in \mathbb{F}$ y $u \in U$.

Estas dos condiciones son equivalentes a la sola condición:

$$T(\alpha u_1 + \beta u_2) = \alpha T(u_1) + \beta T(u_2) \text{ , para } \alpha, \beta \in \mathbb{F} \text{ y } u_1, u_2 \in U.$$

El conjunto de todas las transformaciones lineales de U en V es un espacio vectorial respecto de las operaciones de adición y de producto por un escalar definidas por

$$(T_1 + T_2)(x) = T_1(x) + T_2(x)$$
$$(\alpha T)(x) = \alpha T(x) \text{ , para } x \in U \text{ y } \alpha \in \mathbb{F}$$

Ese espacio se designa por $\mathcal{L}(U,V)$ o bien por $Hom\ (U,V)$.
Las transformaciones lineales biyectivas se llaman transformaciones *regulares*; y las biyectivas , se llaman *isomorfismos*. Si existe un isomorfismo de U sobre V se dice que dichos espacios son *isomorfos*.

Las transformaciones lineales de un espacio U en sí mismo se llaman *endomorfismos*, y el espacio vectorial de los endomorfismos de U se designa por $\mathcal{E}(U)$. Las transformaciones lineales biyectivas de U sobre U se llaman *automorfismos*.

Las transformaciones lineales de U en $\mathbb{F}$ llaman *funciones* o *formas lineales*, y el espacio de las funciones lineales definidas en U se designa por $\mathcal{L}(U)$.

Ejercicios

***3.1.1.** Demostrar que si T: $U \to V$ es una transformación lineal, entonces $T(0) = 0$ y $T(-x) = -(T(x)) = -Tx$.

***3.1.2.** Demostrar que si T: $U \to V$ es una transformación lineal, entonces

$$T(\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_r u_r) = \alpha_1 Tu_1 + \alpha_2 Tu_2 + \ldots + \alpha_r Tu_r,$$

o bien, en forma abreviada,

$$T\Big(\sum_{i=1}^{r} \alpha_i u_i\Big) = \sum_{i=1}^{r} \alpha_i T u_i .$$

*3.1.3. Si $(u_1, \ldots, u_n)$ es un sistema *l.d.*, los vectores $(Tu_1, \ldots, Tu_n)$ son tambien *l.d.* ¿Podría decirse algo semejante si los vectores $u_1, \ldots, u_n$ son *l.i.* ?

*3.1.4. Si *dim* $U = m$ y *dim* $V = n$, demostrar que

$$dim\ \mathcal{L}(U,V) = mn .$$

En particular, $\mathcal{L}(U)$ tendrá dimensión m.
(Sugerencia: Las transformaciones T_j^i $(i = 1,\ldots,m\ ;\ j = 1,\ldots, n)$ que, para las bases $(u_1,\ldots,u_m)$ de U y $(v_1, \ldots, v_n)$ de V son definidas por las fórmulas

$$T_j^i(u_k) = \delta_k^i\, v_j \quad , \quad (k = 1, \ldots, m) \ ,$$

(donde δ_k^i es el *símbolo de Kronecker* que toma el valor *0* cuando $i \neq k$ y el valor *1* cuando $i = k$), constituyen una base de $\mathcal{L}(U,V)$. Esa base se llama *base canónica* de este espacio *asociada a las bases* $\{u_i\}$ y $\{v_j\}$.

*3.1.5. Sea $T: U \to V$ una transformación lineal. Demostrar que si $(Tu_1, \ldots, Tu_r)$ es un sistema *l.i.*, entonces el sistema $(u_1, \ldots, u_r)$ es *l.i.*

3.1.6. Probar que la aplicación $P : \mathbb{R}^3 \to \mathbb{R}^3$ definida por la fórmula

$$P(a_1, a_2, a_3) = (a_1, a_2, 0)$$

es una transformación lineal. Hallar los vectores de $\mathbb{R}^3$ que tienen por imagen al vector nulo, y caracterizar a los vectores de $\mathbb{R}^3$ que son imágenes por la transformación P.

3.1.7. Demostrar que la aplicación $T : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$T(a_1, a_2, a_3) = (2a_1, 3a_2, -a_3)$$

es una transformación lineal.

3.1.8. La misma cuestión anterior para $T : \mathbb{R}^3 \to \mathbb{R}^2$ donde

$$T(a_1, a_2, a_3) = (2a_1 + a_3, -a_2 + 2a_3).$$

3.1.9. Probar que la aplicación $A : \mathbb{R}^3 \to \mathbb{R}^3$ dada por

$$A(a_1, a_2, a_3) = (a_1 + 1, a_2, a_3)$$

no es una transformación lineal.

3.1.10. La transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$ aplica a $(2, -3)$ en $(1,2)$, y a $(1,4)$ en $(2, -1)$. ¿Cuál es la imagen del vector (x,y) por T?

3.1.11. ¿Existe alguna transformación lineal T de $\mathbb{R}^2$ en $\mathbb{R}^2$ tal que $T(2,3) = (-1,2)$, $T(1,-1) = (3,1)$ y $T(-1,9) = (11,2)$?

***3.1.12.** Sea V un espacio bidimensional que tiene la base (e_1, e_2). Los endomorfismos L y M transforman a cada vector $\xi_1 e_1 + \xi_2 e_2$ en un vector $\eta_1 e_1 + \eta_2 e_2$ según las fórmulas siguientes:

$$L \begin{cases} \eta_1 = 2\xi_1 + 3\xi_2 \\ \eta_2 = \xi_1 + 2\xi_2 \end{cases} \qquad M \begin{cases} \eta_1 = -3\xi_1 + 2\xi_2 \\ \eta_2 = -6\xi_1 + 4\xi_2 \end{cases}$$

(a) Hallar las imágenes de e_1, e_2 y $(2,3)$ por esas transformaciones.

(b) ¿qué vectores tienen por imágenes a e_1 y a e_2?

(c) Hallar ecuaciones análogas a las que hemos dado para L y M, para las transformaciones $3L$, $-2M$, $3L - 2M$.

(d) Se dice que un vector es *fijo o invariante* bajo un endomorfismo T si $Tx = x$. ¿Tienen los endomorfismos L y M puntos fijos aparte de $x = 0$?

(e) Un subespacio S es invariante bajo T si $T(S) \subset S$. ¿Tienen L o M subespacios invariantes unidimensionales? (Sugerencia: buscar un escalar λ y un vector $x \neq 0$ tales que, respectivamente, $Lx = \lambda x$ o bien $Mx = \lambda x$).

3.1.13. ¿Cuáles de las siguientes aplicaciones de $\mathbb{F}^4$ en sí mismo son transformaciones lineales ?

a) $(x^1, x^2, x^3, x^4) \mapsto (x^1 x^2, x^2 - x^1, x^3, x^4)$

b) $(x^1, x^2, x^3, x^4) \mapsto (\lambda x^2, x^2 - x^1, x^3, x^4)$

c) $(x^1, x^2, x^3, x^4) \mapsto (0, x^2, x^3, x^1 + x^2 + x^3 + x^4)$

3.2. Núcleo e imagen de una transformación lineal. Se llama *núcleo* de una transformación lineal $T : U \to V$, y se designa por *Ker* T [1], al subespacio de U

$$Ker\ T = \{u \in U \ / \ Tu = 0\} = T^{-1}(0).$$

Así por ejemplo, si U_1 es un subespacio de U, el núcleo del homomorfismo canónico (sección 2.4) $U \to U/U_1$, que aplica a cada vector de U en la clase de U/U_1 a que pertenece es justamente U_1 (Demostrarlo!).

Como ya vimos, si T es *inyectiva*, es decir si la relación $Tx = Ty$ implica que $x = y$, se dice que es *regular*.

Teorema 3.2.1. Una transformación lineal es regular si y solo si es *Ker* $T = \{0\}$, es decir si su núcleo se reduce al vector nulo.

Demostración. Sea T regular. Si $u \in Ker\ T$, entonces es $Tu = 0$, es decir $Tu = T0$. Por ser T regular será $u = 0$, es decir que *Ker* $T = \{0\}$.

Invsersamente, sea *Ker* $T = \{0\}$. Si $Tx = Ty$, entonces es $T(x - y) = 0$, es decir $(x - y) \in Ker\ T$. Luego tiene que ser $x-y=0$ o sea $x = y$. T es por tanto regular. ∎

Teorema 3.2.2. Cada transformación lineal regular $T : U \to V$ aplica a los vectores de un sistema *l.i.* en vectores de un sistema *l.i.*

Demostración. Sea $(u_\alpha)_{\alpha \in A}$, un sistema *l.i.* de U. Debemos probar que el sistema $(Tu_\alpha)_{\alpha \in A}$ de vectores de V es también *l.i.*, es decir que si $(Tu_{\alpha_1}, \ldots, Tu_{\alpha_r})$ es un subconjunto finito cualquiera del

1) Del alemán : Kern, o del inglés Kernel que significan núcleo.

sistema $(Tu_\alpha)_{\alpha \in A}$, entonces una relación de la forma

$$c_1 Tu_{\alpha_1} + c_2 Tu_{\alpha_2} + \ldots + c_r Tu_{\alpha_r} = 0 , \qquad (3.2.1)$$

implica que $c_1 = c_2 = \cdots = c_r = 0$.

En efecto, supongamos que tiene lugar una relación de la forma (3.2.1). Se tendrá entonces

$$T(c_1 u_{\alpha_1} + \ldots + c_r u_{\alpha_r}) = 0 ,$$

y por tanto,

$$c_1 u_{\alpha_1} + \ldots + c_r u_{\alpha_r} \in Ker\ T.$$

Pero T es regular, luego, en virtud del teorema anterior, es $Ker\ T = \{0\}$, y por consiguiente

$$c_1 u_{\alpha_1} + \ldots + c_r u_{\alpha_r} = 0 \qquad (3.2.2)$$

Ahora bien, $(u_\alpha)_{\alpha \in A}$ es un sistema *l.i.* . Luego de (3.2.2) se sigue que $c_1 = c_2 = \cdots = c_r = 0$. ■

Se llama *imagen* de la transformación lineal $T : U \to V$ y se designa por $Im\ T$ al subconjunto de V

$$Im\ T = T(U) = \{v \in V \mid \exists\ u \in U,\ Tu = v\} ,$$

el cual constituye un subespacio como puede ser fácilmente comprobado.

Cuando $Im\ T = V$ entonces T es una transformación de U *sobre* V o sea una *sobreyección* (se dice también epiyección o suryección).

Si U y V tienen dimensión finita, $Ker\ T$ e $Im\ T$ tienen dimensión finita.

La dimensión de $Im\ T$ se llama *rango* de T y se designa por $r(T)$. Evidentemente se tiene

$$r(T) \leqslant dim\ V ,$$

y la igualdad tiene lugar si y solo si T es *sobre* V. Por otra parte

puesto que vectores *l.d.* de U se aplican por T en vectores *l.d.* de V (ejercicio 3.1.3), es claro que

$$r(T) \leqslant dim\ U,$$

por cuanto si $Tu_1, \ldots, Tu_p$ son p vectores *l.i.* de *Im* T, los vectores $u_1, \ldots, u_p$ deberán ser vectores *l.i.* de U. (ejercicio 3.1.5).

Se ha demostrado así la proposición siguiente

Teorema 3.2.3. Si T es una transformación lineal de U en V, y U y V son espacios de dimensión finita entonces el rango de T no excede a *dim* U ni a *dim* V.

Ejercicio

3.2.1. Determinar el núcleo y la imagen de las transformaciones lineales de los ejercicios 3.1.6 , 3.1.7 , 3.1.8 , 3.1.10, y 3.1.12.

3.3. Composición de transformaciones lineales. Transformaciones lineales inversas.

Si $S : U \to V$ y $T : V \to W$ son dos transformaciones lineales, se comprueba inmediatamente que la transformación compuesta

$$T \circ S : U \to W \quad , \text{ definida por}$$

$$(T \circ S)(u) = T[S(u)] \quad , \quad \forall\ u \in U,$$

es también una transformación lineal. En efecto,

$$(T \circ S)(\alpha u_1 + \beta u_2) = T[S(\alpha u_1 + \beta u_2)] = T[\alpha S u_1 + \beta S u_2]$$

$$= \alpha T[S u_1] + \beta T[S u_2] = \alpha (T \circ S) u_1 + \beta (T \circ S) u_2.$$

En adelante, la transformación lineal $T \circ S$ se representará simplemente por TS. En particular, se podrán definir las potencias de una transformación lineal T:

$$T = TT \ , \quad T^3 = (TT)T = T(TT) = TTT \ , \text{ etc.}$$

La transformación $I_U : U \to U$, tal que $Tu = u$, $\forall\ u \in U$ se

llama *transformación identica de* U, y es evidentemente lineal. Convendremos entonces en que $T^0 = I_U$.

Dada una transformación lineal $T : U \to V$, se dice que la tranformación lineal $T^{-1} : V \to U$ es *inversa* de T si se cumplen las relaciones

$$TT^{-1} = I_V \qquad \text{y} \qquad T^{-1}T = I_U .$$

Se deduce de esta definición que la transformación T^{-1} es única.

En virtud de la primera relación, y puesto que I_V es sobreyectiva, deberá serlo también T. En virtud de la segunda relación, por cuanto I_U es inyectiva, deberá serlo T. Luego si T tiene inversa es necesariamente biyectiva. Obviamente ocurrirá lo mismo con T^{-1} .

Ejercicios

3.3.1. Consideremos las transformaciones L y M del ejercicio 3.1.12. Obtener las ecuaciones que expresan a las componentes del vector (η_1 , η_2), imagen del vector (ξ_1 , ξ_2), para cada una de las siguientes transformaciones

$$L^2 , L^2 - 4L + I , M^2 , M^2 - M , LM , ML ,$$

donde I es la transformación idéntica de V. Observar que

$$LM \neq ML.$$

***3.3.2.** Si $R : U \to V$, $S : V \to W$ y $T : W \to Z$ son tres aplicaciones de conjuntos, demostrar que se cumple la propiedad asociativa $(TS)R = T(SR)$.

***3.3.3.** Probar las siguientes propiedades distributivas de las tranformaciones lineales:

a) Si $R : U \to V$, $S : V \to W$ y $T : V \to W$, entonces

$$(S + T)R = SR + TR$$

b) Si $R : U \to V$, $S : U \to V$ y $T : V \to W$, entonces

$$T(R + S) = TR + TS.$$

3.4. Isomorfismos de espacios vectoriales. Como hemos dicho prece-

dentemente, (sección 3.1), una transformación lineal T de U *sobre* V que es *regular* ,es decir una transformación lineal biyectiva $T: U \to V$ se denomina *isomorfismo* de U sobre V.

Dados los espacios vectoriales U y V, si existe un isomorfismo de U sobre V, diremos que U y V son *isomorfos* y escribiremos $U \cong V$ La relación de isomorfismo entre espacios vectoriales es una relación de equivalencia (Demostrarlo !).

Si la transformación lineal $T: U \to V$ es un isomorfismo, entonces existe la transformación inversa $T^{-1}: V \to U$ y también es un isomorfismo (Demostrarlo !). En ese caso se tiene, según sabemos,

$$TT^{-1} = I_V \quad \text{y} \quad T^{-1}T = I_U \ ,$$

donde I_U e I_V son las aplicaciones idénticas de U y de V respectivamente.

Teorema 3.4.1. La condición necesaria y suficiente para que dos espacios de dimensión finita, U y V, sean isomorfos es que tengan la misma dimensión: $dim\ U = dim\ V$.

Demostración. Si U y V son isomorfos y $T: U \to V$ es un isomorfismo, entonces se tiene

$$dim\ V = dim\ Im\ T = r(T) \leqslant dim\ U$$

y

$$dim\ U = dim\ Im\ T^{-1} = r(T^{-1}) \leqslant dim\ V,$$

luego

$$dim\ U = dim\ V.$$

Supongamos ahora que se cumple esta última relación. Se demostrará entonces facilmente que si $(u_1, \ldots, u_n)$ es una base de U y $(v_1, \ldots, v_n)$ es una base de V, la transformación lineal $T: U \to V$ definida por

$$T(u_i) = v_i \quad , \quad (i = 1, \ldots, n)$$

es un isomorfismo de U sobre V (Demostrar que T existe y que es un isomorfismo). ■

Ejercicios

*3.4.1. Si $T: U \to V$ es una transformación lineal regular, entonces U y $T(U) = Im\ T$ son isomorfos.

*3.4.2. Una transformación lineal T de un espacio n-dimensional U en un espacio n-dimensional V es biyectiva si y solo si es regular. Dicho de otra manera, la condición necesaria y suficiente para que una transformación lineal de un espacio n-dimencional en otro de la misma dimensión sea un isomorfismo es que sea regular y por tanto que su núcleo sea el subespacio $\{0\}$.

3.5. Isomorfismo inducido por una transformación lineal.

Sea $T: U \to V$ una transformación lineal. Puesto que su núcleo, $Ker\ T$, es un subespacio de U, existe un espacio factor $U/_{Ker\ T}$ y un isomorfismo canónico

$$\overline{T} : U/_{Ker\ T} \to Im\ T ,$$

definido por la fórmula

$$\overline{T}\,\overline{x} = Tx$$

que hace corresponder a cada clase $\overline{x} \in U/_{Ker\ T}$, la imagen Tx de cualquiera de sus elementos x. El lector debe comprobar que Tx tiene el mismo valor, cualquiera que sea el elemento x de $\overline{x}$ que se elija. Además deberá comprobar que $\overline{T}$ es un isomorfismo de $U/_{Ker\ T}$ sobre $Im\ T$. La relación que se establece en el ejercicio 2.5.15. permite escribir, cuando U y V tienen dimensión finita,

$$dim\ Im\ T = dim\ U/_{Ker\ T} = dim\ U - dim\ Ker\ T , \quad (3.5.1)$$

es decir que *la dimensión de la imagen de una transformación lineal $T: U \to V$ es igual a la diferencia de la dimensión de U menos la dimensión del núcleo de T.*

La presente relación puede expresarse en la forma

$$r(T) = dim\ U - dim\ Ker\ T . \quad (3.5.2)$$

Teorema 3.5.1. La condición necesaria y suficiente para que la transformación lineal $T: U \to V$ sea regular es que

$$r(T) = dim\ Im\ T = dim\ U\ .$$

Demostración. Basta aplicar la relación (3.5.2) teniendo presente el teorema 3.2.1. ∎

Ejercicios

3.5.1. Para las transformaciones lineales de los ejercicios 3.1.6 , 3.1.7 , 3.1.8 y 3.1.10, determinar el núcleo y estudiar el isomorfismo inducido ¿Cuáles de esas transformaciones son isomorfismos?

*3.5.2. El núcleo de una función lineal $f \in \mathcal{L}(V)$, que no idénticamente nula, definida en un espacio vectorial V de dimensión n, es un subespacio de V de dimensión $n-1$.

3.5.3. U y V son dos espacios bidimensionales reales en los que son dadas las bases (e_1 , e_2) y (f_1 , f_2) respectivamente. Se consideran las transformaciones lineales $T: U \to V$ tales que

$$T(\xi e_1 + \eta e_2) = \lambda f_1 + \mu f_2 ,$$

que son definidas por las ecuaciones

(a) $\begin{cases} \lambda = \alpha\xi \\ \mu = \beta\eta \end{cases}$ (b) $\begin{cases} \lambda = \xi \\ \mu = \gamma\xi + \eta \end{cases}$

(c) $\begin{cases} \lambda = (\xi - \eta)/\sqrt{2} \\ \mu = (\xi + \eta)/\sqrt{2} \end{cases}$ (d) $\begin{cases} \lambda = \xi \\ \mu = -\eta \end{cases}$

donde α, β y γ son números reales dados. ¿Cuáles de estas transformaciones son isomorfismos? Para aquellas que lo sean, hallar las transformaciones inversas.

§4. Dualidad de espacios vectoriales

4.1. Aplicaciones bilineales. Dados los espacios vectoriales U, V y W sobre un campo $\mathbb{F}$, se dice que una aplicación $\phi: U \times V \to W$ es *bilineal* si se cumplen las condiciones

$$\phi(\alpha u_1 + \beta u_2 , v) = \alpha\phi(u_1 , v) + \beta\phi(u_2 , v) ,$$
$$\phi(u , \alpha v_1 + \beta v_2) = \alpha\phi(u , v_1) + \beta\phi(u , v_2) ,$$

cualesquiera que sean $u , u_1 , u_2 \in U$, $v , v_1 , v_2 \in V$ y $\alpha,\beta \in \mathbb{F}$.

Si $W = \mathbb{F}$, decimos que ϕ es una *función bilineal.*

4.2. Producto escalar y espacios duales. Los espacios que vamos a considerar en esta sección pueden ser de dimensión finita o infinita.

Sean V y V^* dos espacios vectoriales sobre un campo $\mathbb{F}$. Supongamos que a cada par (x^*, x) en que $x^* \in V^*$ y $x \in V$ le corresponde un escalar bien determinado que designaremos por $\langle x^* , x\rangle$, de manera que se cumplen las condiciones siguientes:

E1. $\langle \alpha x^* + \beta y^* , x\rangle = \alpha\langle x^* , x\rangle + \beta\langle y^* , x\rangle$

$\langle x^* , \alpha x + \beta y\rangle = \alpha\langle x^* , x\rangle + \beta\langle x^* , y\rangle$.

E2. Si $\langle x^* , x\rangle = 0$ para x^* fijo en V^* y para cualquier x de V, entonces es $x^* = 0$; y si es $\langle x^* , x\rangle = 0$ para x fijo en V y para cualquier x^* de V^*, entonces $x = 0$.

La primera condición expresa que la aplicación $(x^* , x) \to \langle x^* , x\rangle$ de $V^* \times V$ en $\mathbb{F}$ es una función bilineal. La segunda condición se expresa diciendo que dicha función bilineal es *no degenerada.*

Se sigue de E1 que el producto escalar $\langle x^* , x\rangle$ se anula si uno de los factores es igual a 0. Puede observarse además que la condición E2 es equivalente a la condición

E2'. Si tiene lugar la relación

$$\langle x^* , x\rangle = \langle x_1^* , x\rangle , \quad \forall\ x \in V ,$$

entonces es $x^* = x_1^*$; y si es

$$\langle x^* , x\rangle = \langle x^* , x_1\rangle , \quad \forall\ x^* \in V^* ,$$

entonces es $x = x_1$.

La función bilineal $(x^*, x) \to \langle x^*, x\rangle$, que algunas veces se designa con el símbolo $\langle\ ,\ \rangle$, y que cumple las dos condiciones precedentes, se llama *producto escalar* de los espacios V^* y V . $\langle x^*, x\rangle$ es el producto escalar de los vectores x^* y x, o el producto escalar de x^* por x.

Dos espacios vectoriales V y V^* para los cuales es definido un producto escalar, se llaman espacios *duales*.

Si V^* es dual de V, con el producto escalar $\langle\ ,\ \rangle$, entonces V es dual de V^* con el producto escalar $\langle\ ,\ \rangle^*$ definido por

$$\langle x, x^*\rangle^* = \langle x^*, x\rangle \quad , \quad x \in V \ , \ x^* \in V^*$$

Por tanto la relación de dualidad es simétrica en este sentido. En vez de $\langle x, x^*\rangle^*$ escribiremos simplemente $\langle x^*, x\rangle$.

Se dice que los vectores $x \in V$ y $x^* \in V^*$ son *ortogonales* , si es $\langle x^*, x\rangle = 0$. En particular, la precedente condición E2 expresa que un vector de uno de los espacios es ortogonal a todos los vectores del otro espacio solamente si es el vector nulo.

Teorema 4.2.1. Dado un espacio vectorial cualquiera V, el espacio $\mathcal{L}(V)$ de las transformaciones lineales de V en $\mathbb{F}$ es dual de V si se define el producto escalar mediante la función lineal $\langle\ ,\ \rangle$ tal que

$$\langle f, x\rangle = f(x) \quad , \quad f \in \mathcal{L}(V) \quad , \quad x \in V.$$

Demostración. La condición E1 se cumple porque

$$\langle \alpha f_1 + \beta f_2 , x\rangle = (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x)$$
$$= \alpha\langle f_1 , x\rangle + \beta\langle f_2 , x\rangle \ ,$$

y

$$\langle f, \alpha x_1 + \beta x_2\rangle = f(\alpha x_1 + \beta x_2) = \alpha f(x_1) + \beta f(x_2)$$
$$= \alpha\langle f, x_1\rangle + \beta\langle f, x_2\rangle \ .$$

Probemos ahora que se cumple tanbién la condición E2. En efecto, si es $\langle f , x\rangle = 0$ para f dado y para todo $x \in V$, entonces es $f(x) = 0$ para todo $x \in V$, y por tanto f es la función idénti

camente nula, es decir el vector nulo del espacio vectorial $\mathcal{L}(V)$.

Por otra parte, sea a un vector de V tal que $<f, a> = 0$ para todo $f \in \mathcal{L}(V)$. Supongamos que fuera $a \neq 0$. Vamos a probar que entonces es posible hallar un elemento $g \in \mathcal{L}(V)$ tal que

$$g(a) = <g , a> \neq 0 ,$$

y por tanto se llegaría a una contradicción. En efecto, en virtud de la proposición del ejercicio 2.5.3 ,podrían elegirse los vectores $(a_\alpha)_{\alpha \in A}$ de V de manera que $\{a , (a_\alpha)_{\alpha \in A}\}$ fuera una base de V. Según el ejercicio 1.2.6 cada vector $x \in V$ podría expresarse exactamente de una manera en la forma

$$x = \lambda a + \sum_{\alpha \in A} \lambda^\alpha a_\alpha . \qquad (4.2.1)$$

En particular, para $x = a$ se tendría que $\lambda = 1$ y $\lambda^\alpha = 0$ para todo $\alpha \in A$. Se podría definir entonces la función lineal

$$g(x) = \lambda$$

que haría corresponder a cada vector $x \in V$ el elemento λ de $\mathbb{F}$ que permite expresarlo en la forma (4.2.1). Esa función g sería aquella que buscamos por cuanto $g(a) = 1 \neq 0$. ∎

El espacio $\mathcal{L}(V)$ que, según acabamos de ver, es dual de V, se llama algunas veces *espacio conjugado* de V, y es importante porque, como estableceremos en el siguiente teorema, cada espacio dual de V, es isomorfo de una manera natural con un subespacio de $\mathcal{L}(V)$. En el caso particular en que V es de dimensión finita, puede considerarse a $\mathcal{L}(V)$ como *el* espacio dual de V, porque es isomorfo de manera natural con cada uno de los espacios duales de V, como veremos en la sección 5.1.

Teorema 4.2.2. Sean V y V^* dos espacios duales respecto del producto escalar $< , >$. Existe entonces una transformación lineal inyectiva

$$\varphi: V^* \rightarrow \mathcal{L}(V)$$

tal que

$$(\varphi x^*)(x) = \langle x^* , x \rangle , \qquad (4.2.2)$$

y por tanto $\varphi : V^* \to \varphi(V^*)$ es un isomorfismo de V^* sobre un subespacio de $\mathcal{L}(V)$. Además φ es unívocamente determinado por la relación (4.2.2).

Demostración. Dado $x^* \in V^*$, la aplicación φx^* de V en $\mathbb{F}$ definida por (4.2.2) es evidentemente lineal en virtud de la linealidad del producto escalar, luego pertenece a $\mathcal{L}(V)$.

Consideremos ahora la aplicación $\varphi : V^* \to \mathcal{L}(V)$, tal que $\varphi(x^*) = \varphi x^*$. Cualesquiera que sean los vectores $x_1^*, x_2^* \in V^*$, los elementos $\alpha, \beta \in \mathbb{F}$ y el vector $x \in V$, podemos escribir

$$\begin{aligned}[\varphi(\alpha x_1^* + \beta x_2^*)](x) &= \langle \alpha x_1^* + \beta x_2^* , x \rangle = \alpha \langle x_1^* , x \rangle + \beta \langle x_2^* , x \rangle \\ &= \alpha[(\varphi x_1^*)(x)] + \beta[(\varphi x_2^*)(x)] \\ &= [\alpha(\varphi x_1^*) + \beta(\varphi x_2^*)](x).\end{aligned}$$

Por consiguiente,

$$\varphi(\alpha x_1^* + \beta x_2^*) = \alpha\varphi(x_1^*) + \beta\varphi(x_2^*) ,$$

lo cual demuestra que φ es lineal.

Probemos ahora que φ es inyectiva. Si es $\varphi x_1^* = \varphi x_2^*$ entonces es $\langle x_1^* , x \rangle = \langle x_2^* , x \rangle$ para todo $x \in V$. Resulta entonces de la condición E2′ que $x_1^* = x_2^*$. Que φ es determinada por (4.2.2) es obvio porque esta fórmula define completamente a la función φx^*, para cada $x^* \in V^*$. ∎

Ejercicios

*4.2.1. Demostrar que $\mathbb{R}^n$ $(n = 1, 2, \ldots)$ es dual de sí mismo respecto del producto escalar definido por

$$\langle (x^1, \ldots, x^n), (y^1, \ldots, y^n) \rangle = x^1 y^1 + \ldots + x^n y^n$$

¿Cuál es la función $\varphi(x^*)$ que corresponde a $x^* \in \mathbb{R}^n$ en virtud del teorema 4.2.2.?

4.2.2. Si $V = V^* = \mathbb{F}$ y se define $\langle \alpha ; \beta \rangle = \alpha\beta$, $\alpha, \beta \in \mathbb{F}$,

F puede considerarse dual consigo mismo.

4.3. Relación de ortogonalidad entre vectores de espacios duales.

Hemos visto (sección 4.2) que si V y V^* son espacios duales, los vectores $x \in V$ y $x^* \in V^*$ son *ortogonales* si $\langle x^* , x\rangle = 0$.

Si V_1 es un subespacio de V, los vectores de V^* que son ortogonales a todos los vectores de V_1 constituyen evidentemente un subespacio de V^*. Ese subespacio se llama *complemento ortogonal* de V_1 y se designa por $V_1^\perp$. Semejantemente, los vectores de V que son ortogonales a todos los vectores de un subespacio V_1^* de V^* constituyen un subespacio de V que se llama complemento ortogonal de V_1^* y se designa por $(V_1^*)^\perp$.

Ejercicios

*4.3.1. Si V y V^* son dos espacios duales, se tiene $V^\perp = \{0\}$ y $(V^*)^\perp = \{0\}$.

(Obsérvese que la notación $V^\perp$ no es enteramente clara en tanto no se precise a qué relación de dualidad se refiere la ortogonalidad. Así, si V es un subespacio de W y la dualidad es la que existe entre W y su dual W^*, entonces $V^\perp$ es un subespacio de W^* que no es necesariamente trivial).

*4.3.2. Sean V y V^* dos espacios duales. Si V_1 es un subespacio de V entonces se tiene que

$$V_1 \subset (V_1^\perp)^\perp = V_1^{\perp\perp} .$$

En particular, si $V = V_1$, entonces se tiene

$$V_1 = (V_1^\perp)^\perp = V_1^{\perp\perp} .$$

*4.3.3. Sean V y V^* dos espacios duales. Si V_1 es un subespacio de V, los espacios V_1 y $V^*/V_1^\perp$ son duales respecto del producto escalar definido por

$$\langle \overline{x^*} , y\rangle = \langle x^* , y\rangle \quad , (\overline{x^*} \in V^*/V_1^\perp \quad , y \in V_1),$$

donde x^* designa a un representante cualquiera de la clase $\overline{x^*}$ de $V^*/V_1^{\perp}$.

*4.3.4. Sea V un espacio vectorial de dimensión infinita. Demostrar que existe un espacio dual V^* de V tal que la inyección $\varphi^* : V^* \to \mathcal{L}(V)$, que fué definida en el teorema 4.2.2, no es sobreyectiva.

4.4. **Transformaciones duales.** Sean U y V dos espacios vectoriales, y U^* y V^* sus respectivos duales.

Definición. 4.4.1. Las transformaciones lineales $\varphi : U \to V$ y $\varphi^* : V^* \to U^*$ se llaman *duales*, si satisfacen a la condición

$$\langle y^* , \varphi x \rangle = \langle \varphi^* y^* , x \rangle ,$$

cualesquiera que sean los vectores $x \in U$, $y^* \in V^*$.

Dada la transformación φ, no podemos asegurar, en general, que exista la transformación dual. En la sección 5.5 demostraremos tal existencia en el caso en que los espacios U y V son de dimensión finita. Podemos probar ahora que si existe la transformación dual de φ, dicha transformación dual es única. En efecto, si φ_1^* y φ_2^* son duales de φ, entonces se tiene

$$\langle \varphi_1^* y^*, x \rangle = \langle \varphi_2^* y^*, x \rangle \quad , \quad \forall\, x \in U \quad \text{y} \quad \forall\, y^* \in V^* .$$

La propiedad E2' del producto escalar implica entonces que

$$\varphi_1^* y^* = \varphi_2^* y^* \quad , \quad \forall\, y^* \in V^* ,$$

de donde se deduce que

$$\varphi_1^* = \varphi_2^* \quad ,$$

es decir que φ no posee sino un solo dual.

Hasta ahora hemos tratado acerca de la dualidad de espacios vectoriales cualesquiera. Algunos otros resultados acerca de la dualidad entre espacios de dimensión finita serán estudiados en las secciones siguientes.

Ejercicios

4.4.1. Consideremos los pares duales V, V^* y V_1 , $V^*/V_1^{\perp}$, donde

V_1 es un subespacio de V, y el producto escalar se define como en el ejercicio 4.3.3. Probar que la proyección canónica de V^* sobre $V^*/V_1^\perp$, que aplica a cada elemento de V^* en la clase de $V^*/V_1^\perp$ que lo contiene:

$$\pi : V^* \to V^*/V_1^\perp ,$$

es la transformación dual de la inyección canónica

$$i : V_1 \to V$$

que aplica a cada vector de V_1 en sí mismo. Es decir que $\pi = i^*$.

4.4.2. Sean U,U^* y V,V^* dos pares de espacios duales. Sean además $\varphi : U \to V$ y $\psi : U \to V$ dos transformaciones lineales que admiten las duales $\varphi^* : V^* \to U^*$ y $\psi^* : V^* \to U^*$. Probar entonces que $\varphi + \psi$ y $\lambda\varphi$ tienen las duales

$$(\varphi + \psi)^* = \varphi^* + \psi^* \quad \text{y} \quad (\lambda\varphi)^* = \lambda\varphi^* .$$

Si, además, W, W^* son duales y $\chi : V \to W$ y $\chi^* : W^* \to V^*$ son lineales, entonces

$$(\chi \circ \varphi)^* = \varphi^* \circ \chi^* .$$

4.4.3. Si i_U es una aplicación idéntica de U e i_{U^*} es la aplicación idéntica de U^*, entonces

$$i_{U^*} = (i_U)^* .$$

Si $\varphi : U \to V$ y $\varphi_1 : V \to U$ son tales que

$$\varphi_1 \circ \varphi = i_U ,$$

y si existen las duales $\varphi^* : V^* \to U^*$ y $\varphi_1^* : U^* \to V^*$, entonces es

$$\varphi^* \circ \varphi_1^* = (\varphi \circ \varphi_1)^* = (i_U)^* = i_{U^*} .$$

En particular, si φ y φ_1 son isomorfismos inversos entre sí, entonces φ^* y φ_1^* también son isomorfismos inversos entre sí. Se cumple por tanto la relación $(\varphi^{-1})^* = (\varphi^*)^{-1}$.

§ 5. Dualidad de los espacios de dimensión finita.

Los espacios vectoriales que vamos a considerar en este parágrafo son todos de dimensión finita.

5.1. Dimensión de los espacios duales.

Teorema 5.1.1. Los espacios duales de un espacio V de dimensión finita n tienen todos la misma dimensión n.

Demostración. Sea V^* un espacio dual del espacio vectorial V de dimensión n. Supongamos que $(e_1, \ldots, e_n)$ es una base de V. Definamos la transformación lineal $\phi : V^* \to \mathbb{F}^n$ por la fórmula

$$\phi(v^*) = (\langle v^*, e_1\rangle, \langle v^*, e_2\rangle, \ldots, \langle v^*, e_n\rangle).$$

La aplicación ϕ es inyectiva. En efecto, si se supone que

$$\phi v_1^* = \phi v_2^* \quad ,$$

entonces se deduce de la definición de ϕ que

$$\langle v_1^* - v_2^*, e_i\rangle = 0 \qquad (i = 1, \ldots, n).$$

De aquí se deduce que $\langle v_1^* - v_2^*, v\rangle = 0$, $\forall\, v \in V$.

De la propiedad E2 del producto escalar se sigue entonces que

$$v_1^* = v_2^* \quad .$$

Por consiguiente ϕ es regular, es decir que V^* es isomorfo con $\phi(V^*)$ (ejercicio 3.4.2), y por tanto

$$\dim V^* = \dim \phi(V^*) \leqslant \dim \mathbb{F}^n = n = \dim V.$$

Si se invierten los papeles de V y V^* en el razonamiento anterior, lo que el lector debe hacer en detalle, se obtiene que

$$\dim V \leqslant \dim V^* \ .$$

De las dos desigualdades que hemos establecido resulta

$$\dim V = \dim V^* \ . \qquad \blacksquare$$

Corolario. Dados n escalares cualesquiera ξ_ν $(\nu = 1, \ldots, n)$, existe uno y solo un vector $v^* \in V^*$ tal que

$$<v^*, e_\nu> = \xi_\nu \quad , \quad (\nu = 1,2, \ldots, n) \ ,$$

donde $(e_1 , \ldots , e_\nu)$ es una base dada de V.

Demostración. La transformación lineal $\phi : V^* \to \mathbb{F}^n$, definida en la demostración del teorema anterior, es regular, según vimos entonces, y además es

$$dim\ V^* = dim\ \mathbb{F}^n \quad ;$$

por consiguiente (ejercicio 3.4.2) ϕ es un isomorfismo de V^* sobre $\mathbb{F}^n$. Significa esto que a cada vector $(\xi_1 , \ldots , \xi_n)$ de $\mathbb{F}^n$ le corresponde uno y solo un vector $v^* \in V^*$ tal que

$$\phi v^* = (\xi_1 , \ldots , \xi_n) \ ,$$

lo cual según la definición de ϕ, prueba que

$$<v^*, e_\nu> = \xi_\nu \quad , (\nu = 1, \ldots , n) \ . \ \blacksquare$$

La proposición dual de la enunciada en el corolario anterior expresa que *dados n escalares cualesquiera,* ξ^ν $(\nu = 1, \ldots, n)$, *existe uno y solo un vector* $v \in V$ *tal que*

$$<e^{*\nu} , v> = \xi^\nu \quad , \quad (\nu = 1, \ldots , n)$$

donde $(e^{*1} , \ldots , e^{*n})$ *es una base de* V^*. Proposiciones duales, como ésta no expresan nada esencialmente nuevo y por consiguiente no serán explícitamente enunciadas.

5.2. Base dual. El corolario del teorema 5.1.1 nos permite hacer una importante observación, a saber que dada una base $(e_1, \ldots , e_n)$ de V existe un conjunto bien determinado de vectores $(e^{*1}, \ldots , e^{*n})$ de V^* tales que

$$<e^{*i} , e_j> = \delta^i_j \quad , \quad (i,j = 1, \ldots , n) \ , \quad (5.2.1)$$

donde δ^i_j es el símbolo de Kronecker. Los n vectores e^{*j} son *l.i.*, pues si

$$\alpha_1 e^{*1} + \ldots + \alpha_n e^{*n} = 0$$

entonces

$$0 = <\alpha_1 e^{*1} + \ldots + \alpha_n e^{*n} , e_j> = \alpha_1 <e^{*1} , e_j> + \ldots + \alpha_n <e^{*n} , e_j>$$

$$= \alpha_j \quad , \quad (j = 1, \ldots, n) \quad ,$$

Dado que *dim* $V^* = n$, los vectores $(e^{*1}, \ldots, e^{*n})$ forman una base de V^* que se llama *base dual* de la base $(e_1, \ldots, e_n)$. Deberá recordarse en adelante que esa base es determinada por las relaciones (5.2.1).

En términos de las bases duales v^{*i} y v_j dos vectores cualesquiera $y^* \in V^*$ y $x \in V$ pueden expresarse de la manera siguiente

$$y^* = \sum_{i=1}^{n} \eta_i \, v^{*i} \quad , \qquad x = \sum_{j=1}^{n} \xi^j \, v_j \quad .$$

Luego, puesto que

$$\langle v^{*i}, v_j \rangle = \delta^i_j \quad , \quad (i,j = 1, \ldots, n)$$

el producto escalar de dichos vectores puede escribirse en la forma

$$\langle y^*, x \rangle = \Big\langle \sum_{i=1}^{n} \eta_i \, v^{*i} \, , \sum_{j=1}^{n} \xi^j \, v_j \Big\rangle = \sum_{i,j=1}^{n} \eta_i \, \xi^j \quad , \qquad (5.2.2)$$

que es la expresión del producto escalar en términos de las componentes de los vectores factores, relativas a bases duales.

Ejercicio

***5.2.1.** Dadas una base de V, (x_ν), y la base dual $(x^{*\nu})$ de V^*, probar que las bases

$$(x_1, x_2 - \lambda_2 x_1, \ldots, x_n - \lambda_n x_1) \quad \text{y} \quad \Big(x^{*i} + \sum_{\nu=2}^{n} \lambda_\nu \, x^{*\nu}, x^{*2}, \ldots, x^{*n}\Big)$$

de V y de V^* son también duales.

5.3. El espacio conjugado $\mathcal{L}(V)$ de un espacio V de dimensión finita. Vamos a referirnos ahora al espacio conjugado $\mathcal{L}(V)$ en el caso que V es de dimensión finita.

Según ya vimos en el caso general, $\mathcal{L}(V)$ es dual de V (teorema 4.2.1). En el teorema 4.2.2 se demostró que cada dual de V es isomorfo con un subespacio de $\mathcal{L}(V)$; y en el ejercicio 4.3.4 se

establece que cuando V es de dimensión infinita existe un dual que es isomorfo con un subespacio propio de $\mathcal{L}(V)$. En el caso que nos ocupa, es decir cuando V tiene dimensión finita, probaremos en el teorema 5.3.1 que todos los duales de V son isomorfos con $\mathcal{L}(V)$ de manera natural, o sea mediante la misma aplicación φ que se ha definido en el teorema 4.2.2.

Debemos observar que se deduce del teorema 5.1.1 que $\mathcal{L}(V)$, como dual de V, tiene dimensión n, lo cual además ya fué establecido en el ejercicio 3.1.4. Resultaría así innecesario demostrar que $\mathcal{L}(V)$ es isomorfo con los duales de V, que son también espacios n-dimensionales. El hecho importante que vamos a probar a continuación es que ese isomorfismo puede establecerse de manera ligada estrechamente al caracter de $\mathcal{L}(V)$ y al producto escalar $< , >$ relativo al cual V^* es dual de V, y por eso dicho isomorfismo se denomina *natural*.

Teorema 5.3.1. El espacio $\mathcal{L}(V)$ conjugado de V es isomorfo de manera natural con cualquier otro dual V^* de V.

Demostración. La transformación lineal inyectiva

$$\varphi : V^* \to \mathcal{L}(V)$$

definida en el teorema 4.2.2 por

$$(\varphi x^*)(x) = <x^* , x> \quad , \quad x^* \in V^* , \; x \in V,$$

es una transformación regular de un espacio n-dimensional en otro de la misma dimensión, por tanto es un isomorfismo de V^* sobre $\mathcal{L}(V)$ (ejercicio 3.4.2). ∎

Corolario. Dado el producto escalar $< , >$ de los espacios duales V^* y V, y dada una función lineal cualquiera $f \in \mathcal{L}(V)$, existe uno y solo un vector $v^* \in V^*$ tal que

$$f(x) = <v^* , x> \quad , \quad \forall \; x \in V.$$

Demostración. Basta elegir v^* de manera que

$$\varphi v^* = f ,$$

donde φ es el isomorfismo de V^* sobre $\mathcal{L}(V)$ definido en el teorema anterior. ■

De lo expuesto precedentemente acerca de $\mathcal{L}(V)$, y en particular del teorema 5.3.1, resulta que, en el caso en que V es de dimensión finita, todos los espacios duales pueden ser identificados canónicamente con el espacio conjugado $\mathcal{L}(V)$. Por este motivo se dice que $\mathcal{L}(V)$ es *el espacio dual de* V. Por consiguiente, *en adelante, al referirnos al dual de un espacio* V *de dimensión finita, entenderemos que se trata del espacio* $\mathcal{L}(V)$ *y que la dualidad es relativa al producto escalar* $< \ , \ >$ *definido por*

$$<f \ , \ x> = f(x) \quad , \quad (x \in V \ , \ f \in \mathcal{L}(V)).$$

Ejercicios

5.3.1. Dada una base de V, describir la base dual de $\mathcal{L}(V)$.

***5.3.2.** Haciendo uso del resultado enunciado en el ejercicio 4.2.1. y del teorema 5.3.1, establecer el isomorfismo natural entre $\mathcal{L}(\mathbb{F}^n)$ y $\mathbb{F}^n$. Este isomorfismo justifica la identificación que haremos usualmente de $\mathbb{F}^n$ con su dual. (Ver el capítulo sobre los espacios vectoriales de producto interno).

***5.3.3.** Haciendo uso de la identificación de $\mathbb{F}^n$ con su dual , a que se refiere el ejercicio anterior, describir la base (e^i) de $\mathbb{F}^n$ que es dual de la base (e_i) del mismo espacio (e^i) se llama *base recíproca de* (e_i).
Obtener además la base recíproca de la base canónica de $\mathbb{F}^n$

***5.3.4.** Si V es un espacio de dimensión finita y $V^{**} = (V^*)^*$ es un dual del dual de V, entonces existe el isomorfismo natural $\varphi : V \to V^{**}$ que hace corresponder a x el elemento $\varphi x \in V^{**}$ tal que

$$<\varphi x \ , \ x^*> = <x^* \ , \ x> \quad , \quad \forall \ x^* \in V^*$$

mediante ese isomorfismo se identifican en forma natural

los espacios V y V^{**} y puede decirse que V es el dual de V^* (V es un dual de V^*, según se observó en la sección 4.2. Según el teorema 5.3.1 tanto V como V^{**} son isomorfos con $\mathcal{L}(V^*)$, luego son isomorfos entre sí. Debe expresarse explícitamente ese isomorfismo).

5.4. La ortogonalidad en el caso de espacios de dimensión finita.

Sea V un espacio de dimensión n y sea V^* su dual. Podemos aplicar sin más las definiciones y los resultados de la sección 4.3. Por tratarse ahora de espacios de dimensión finita podemos probar algunas nuevas proposiciones.

Teorema 5.4.1. Si V_1 es un subespacio de V, entonces se tiene que

$$dim\ V_1 + dim\ V_1^{\perp} = dim\ V \qquad (5.4.1)$$

Demostración. La base $(v_1, \ldots, v_p)$ de V_1 puede extenderse hasta construir la base $(v_1, \ldots, v_p, \ldots, v_n)$ de V. Sea

$$(v^{*1}, v^{*2}, \ldots, v^{*p}, \ldots, v^{*n})$$

la base dual de V^*. Vamos a demostrar que los $n - p$ vectores

$$(v^{*p+1}, \ldots, v^{*n})$$

que evidentemente pertenecen a $V_1^{\perp}$, constituye una base de ese espacio. En primer término podemos probar que esos vectores son li por ser vectores de una base de V^*. Solo nos falta probar que generan a $V_1^{\perp}$. Un vector cualquiera x^* de $V_1^{\perp}$ se puede expresar en la forma

$$x^* = \xi_1 v^{*1} + \ldots + \xi_p v^{*p} + \ldots + \xi_n v^{*n} \qquad (5.4.2)$$

De aquí se deduce que

$$\langle x^*, v_i \rangle = \xi_1 \langle v^{*1}, v_i \rangle + \ldots + \xi_p \langle v^{*p}, v_i \rangle + \ldots + \xi_n \langle v^{*n}, v_i \rangle$$

$$= \xi_1 \delta_i^1 + \ldots + \xi_p \delta_i^p + \ldots + \xi_n \delta_i^n = \xi_i$$

Si suponemos ahora que i toma sucesivamente los valores $1, \ldots, p$,

en cuyo caso es $v_i \in V_1$, y por tanto $\langle x^* , v_i \rangle = 0$, se obtiene que

$$\xi_i = 0 \quad \text{para} \quad i = 1, \ldots, p$$

La fórmula 5.4.2 se reduce entonces a

$$x^* = \xi_{p+1} v^{*p+1} + \cdots + \xi_n v^{*n} ,$$

lo cual prueba que los vectores $(v^{*p+1}, \ldots, v^{*n})$ generan a $V_1^\perp$, y por tanto constituye una base de ese espacio. Se tiene así que

$$dim\ V_1^\perp = n - p = dim\ V - dim\ V_1 . \quad \blacksquare$$

Corolario. Si V_1 es un subespacio cualquiera del espacio V de dimensión finita, se tiene

$$V_1 = (V_1^\perp)^\perp = V_1^{\perp\perp} .$$

Demostración. La aplicación del teorema anterior al subespacio $V_1^\perp$ de V^* permite escribir

$$dim\ V_1^\perp + dim\ V_1^{\perp\perp} = dim\ V^* = dim\ V.$$

Comparando esta ecuación con (5.4.1) se deduce que

$$dim\ V_1 = dim\ V_1^{\perp\perp} .$$

Pero según lo establece la proposición del ejercicio 4.3.2, V_1 es un subespacio de $V_1^{\perp\perp}$. Luego debe ser $V_1 = V_1^{\perp\perp}$. ■

La correspondencia $V_1 \leftrightarrow V_1^\perp$, que es inyectiva en virtud del corolario anterior, es una biyección del conjunto de los subespacios de V sobre el conjunto de los subespacios de V^*, por cuanto, si V_1^* es un subespacio cualquiera de V^*, se tiene

$$V_1^* = (V_1^*)^{\perp\perp} = [(V_1^*)^\perp]^\perp ,$$

y por tanto V_1^* es el correspondiente al subespacio $(V_1^*)^\perp$ de V.

Ejercicio

5.4.1. Teniendo en cuenta que el dual de $\mathbb{F}^n$ es el mismo $\mathbb{F}^n$ (ejercicios 4.2.1 y 5.3.2), interpretar la ortogonalidad en $\mathbb{F}^n$, y en

particular en $\mathbb{R}^3$.

5.5. Transformaciones duales entre espacios de dimensión finita.

Los espacios vectoriales que se consideran en esta sección son de dimensión finita.

Si U, U^* y V, V^* son dos pares de espacios duales de dimensión finita, cada transformación lineal $T: U \to V$ tiene una transformación dual o sea una transformación lineal $T^*: V^* \to U^*$ tal que

$$\langle y^*, Tx\rangle = \langle T^*y^*, x\rangle \quad \forall\ x \in U, \quad \forall\ y^* \in V^*$$

Además, la aplicación $\mathcal{L}(U,V) \to \mathcal{L}(V^*,U^*)$ que hace corresponder a cada $T \in \mathcal{L}(U,V)$ su dual $T^* \in \mathcal{L}(V^*, U^*)$, es un isomorfismo.

Demostración 1. *Existencia de la transformación dual de cada transformación lineal* $T: U \to V$.

Dado un elemento $y^* \in V^*$, la aplicación $x \mapsto \langle y^*, Tx\rangle$ es evidentemente una función lineal definida sobre U. En virtud del corolario del teorema 5.3.1 existirá un vector de U^* y uno solo, que designaremos por T^*y^*, tal que

$$\langle T^*y^*, x\rangle = \langle y^*, Tx\rangle \quad , \quad \forall\ x \in U. \qquad (5.5.1)$$

La transformación $T^*: y^* \mapsto T^*y^*$ de V^* en U^* es lineal, pues para todo $x \in U$ se tiene

$$\langle T^*(\alpha y_1^* + \beta y_2^*), x\rangle = \langle \alpha y_1^* + \beta y_2^*, Tx\rangle = \alpha\langle y_1^*, Tx\rangle + \beta\langle y_2^*, Tx\rangle$$

$$= \alpha\langle T^*y_1^*, x\rangle + \beta\langle T^*y_2^*, x\rangle$$

$$= \langle \alpha T^*y_1^* + \beta T^*y_2^*, x\rangle ,$$

y por tanto

$$T^*(\alpha y_1^* + \beta y_2^*) = \alpha T^*y_1^* + \beta T^*y_2^* \quad ,$$

lo cual prueba que T^* es lineal, De (5.5.1) se sigue, además, que es la transformación dual de T.

2. *La aplicación* $T \mapsto T^*$ *es una biyección de* $\mathcal{L}(U,V)$ *sobre* $\mathcal{L}(V^*,U^*)$.

La aplicación $T \mapsto T^*$ es inyectiva, pues si las transformaciones

lineales $T: U \to V$ y $T_1: U \to V$ tuvieran la misma dual T^*, la definición 4.4.1 implicaría que

$$\langle y^*, Tx\rangle = \langle y^*, T_1 x\rangle \quad , \quad \forall\, x \in U \quad \text{y} \quad \forall\, y^* \in V^* \ .$$

Por consiguiente sería $T = T_1$.

Probemos ahora que $T \mapsto T^*$ es sobreyectiva. Sea $T^*: V^* \to U^*$ un elemento cualquiera de $\mathcal{L}(V^*, U^*)$. Dado un elemento $x \in U$, la aplicación $y^* \mapsto \langle T^* y^*, x\rangle$ es una función lineal definida sobre V^*. En virtud del corolario del teorema 5.3.1 existe un elemento y solo uno de $(V^*)^*$, que designaremos por Tx, tal que

$$\langle T^* y^*, x\rangle = \langle Tx, y^*\rangle$$

El isomorfismo natural $\phi: V \to (V^*)^*$, que existe en la forma expresada en el ejercicio 5.3.4, permite identificar a Tx con el elemento $\phi^{-1}(Tx)$ de V, de modo que podemos escribir

$$\langle Tx, y^*\rangle = \langle \phi(\phi^{-1}(Tx)), y^*\rangle = \langle y^*, \phi^{-1}(Tx)\rangle = \langle y^*, Tx\rangle \ .$$

Por tanto se tiene

$$\langle T^* y^*, x\rangle = \langle y^*, Tx\rangle \quad , \quad \forall\, x \in U \ \text{y} \ \forall\, y^* \in V^* \ ,$$

es decir que $T^*: V^* \to U^*$ es la transformación dual de la transformación $T: U \to V$ que aplica a x en Tx. Resulta así que la aplicación $T \mapsto T^*$ es sobreyectiva.

La aplicación $T \mapsto T^*$ es una transformación lineal de $\mathcal{L}(U,V)$ sobre $\mathcal{L}(V^*,U^*)$. En efecto, cualesquiera que sean $x \in U$ e $y^* \in V^*$ se tiene

$$\begin{aligned}\langle(\alpha T_1 + \beta T_2)^* y^*, x\rangle &= \langle y^*, (\alpha T_1 + \beta T_2)x\rangle = \langle y^*, \alpha T_1 x + \beta T_2 x\rangle \\ &= \alpha\langle y^*, T_1 x\rangle + \beta\langle y^*, T_2 x\rangle \\ &= \alpha\langle T_1^* y^*, x\rangle + \beta\langle T_2^* y^*, x\rangle \\ &= \langle(\alpha T_1^* + \beta T_2^*) y^*, x\rangle \ .\end{aligned}$$

Puesto que esta relación se cumple, dado y^*, para todo $x \in U$, se

sigue de la propiedad E2 del producto escalar que

$$(\alpha T_1 + \beta T_2)^* y^* = (\alpha T_1^* + \beta T_2^*) y^* \quad , \quad \forall\ y^* \quad ;$$

y por tanto

$$(\alpha T_1 + \beta T_2)^* = \alpha T_1^* + \beta T_2^* \quad ,$$

lo que demuestra que la aplicación $T \mapsto T^*$ es lineal. Antes habíamos establecido que es una biyección. Luego es un isomorfismo de $\mathcal{L}(U,V)$ sobre $\mathcal{L}(V^* , U^*)$. ∎

El siguiente teorema establece relaciones entre los vectores que pertenecen al núcleo y a la imagen de transformaciones que son duales.

Teorema 5.5.2. **a)** Un vector pertenece al núcleo de una transformación lineal si y solo si es ortogonal a todos los vectores de la imagen de la transformación dual.

b) Un vector pertenece a la imagen de una transformación lineal si y solo si es ortogonal a todos los vectores del núcleo de la transformación dual.

Demostración. **a)** Podemos escribir la siguiente sucesión de equivalencias

$$u \in Ker\ T \Leftrightarrow Tu = 0 \Leftrightarrow \langle v^* , Tu \rangle = 0 \ , \quad \forall\ v^* \in V^*$$
$$\Leftrightarrow \langle T^* v^* , u \rangle = 0 \ , \quad \forall\ v^* \in V^* \Leftrightarrow u \in (Im\ T^*)^{\perp} \ .$$

Por consiguiente se tiene que

$$Ker\ T = (Im\ T^*)^{\perp} \qquad (5.5.2)$$

En forma análoga se prueba la relación dual

$$Ker\ T^* = (Im\ T)^{\perp} \qquad (5.5.3)$$

b) Tomando los complementos ortogonales de los dos miembros de las relaciones (5.5.2) y (5.5.3) (corolario del teorema 5.4.1) resultan las relaciones

$$Im\ T^* = (Ker\ T)^{\perp} \qquad (5.5.4)$$
$$Im\ T = (Ker\ T^*)^{\perp} \qquad (5.5.5) \ . \quad ∎$$

Teorema 5.5.3. Las imágenes $T(U)$ y $T^*(V^*)$ de las transformaciones lineales duales $T: U \to V$ y $T^*: V^* \to U^*$ tienen la misma dimensión, es decir que

$$dim\ Im\ T = dim\ Im\ T^* \qquad (5.5.6)$$

Puede expresarse esta relación diciendo que *dos transformaciones duales tienen el mismo rango.*

Demostración. De la relación (5.5.2), teniendo presente la fórmula (5.4.1) y el teorema 5.1.1, se deduce

$$dim\ Ker\ T = dim\ (Im\ T^*)^{\perp} = dim\ U^* - dim\ Im\ T^*$$
$$= dim\ U - dim\ Im\ T^* .$$

Aplicando ahora la relación (3.5.1) resulta inmediatamente

$$dim\ Im\ T = dim\ Im\ T^* . \quad \blacksquare$$

Ejercicios

5.5.1. Demostrar que si a^* y b son vectores de U^* y de V respectivamente, y la transformación $T: U \to V$ se define por

$$Tx = \langle a^* , x \rangle b \quad , \quad x \in U ,$$

entonces es

$$T^* y^* = \langle y^* , b \rangle a^* \quad , \quad y^* \in V^* .$$

***5.5.2.** Probar que si T es un isomorfismo de U sobre V, entonces T^* es un isomorfismo de V^* sobre U^* y que

$$(T^*)^{-1} = (T^{-1})^*$$

***5.5.3.** (*Espacios vectoriales de funciones bilineales*). Dados los espacios V de dimensión n y W de dimensión p, las funciones bilineales $B: V \times W \to F$, constituyen un espacio vectorial, $B(V,W)$, de dimensión np, con las operaciones

$$(B_1 + B_2)(v , w) = B_1(v , w) + B_2(v , w) ,$$
$$(\alpha B)(v , w) = \alpha B(v , w)$$

(Sugerencia: Dadas las bases (e_i) de V y (f_j) de W, se puede escribir

$$v = \sum_{i=1}^{n} v^i e_i \qquad \text{y} \qquad w = \sum_{j=1}^{n} w^j f_j \quad .$$

Se puede probar entonces que las np forman bilinelaes

$$B^{ij} \qquad (i = 1, \ldots, n \ ; \ j = 1, \ldots, p) \ ,$$

tales que

$$B^{ij}(u, v) = v^i w^j \quad ,$$

forman una base de $B(V, W)$.

5.5.4. Dados los dos pares de espacios duales U y U^*, V y V^*, de dimensión finita, y las transformaciones lineales $T: U \to V$ y $T_1 : V^* \to U^*$, se definen las aplicaciones $\overline{T}: V^* \times U \to \mathbb{F}$ y $\overline{T}_1 : V^* \times U \to \mathbb{F}$, tales que

$$\overline{T}(v^* , u) = \langle v^* , Tu \rangle \ , \quad \forall \ v^* \in V^* , \ u \in U \ ,$$

$$\overline{T}_1(v^* , u) = \langle T_1 v^* , u \rangle \ , \quad \forall \ v^* \in V^* , \ u \in U \ .$$

Demostrar: a) que $\overline{T}$ y $\overline{T}_1$ son funciones bilineales.

b) que las aplicaciones $T \mapsto \overline{T}$ y $T_1 \mapsto \overline{T}_1$ de $\mathcal{L}(U , V)$ y de $\mathcal{L}(V^* , U^*)$, respectivamente, en $B(V^* , U)$ son isomorfismos.

c) que los precedentes isomorfismos permiten establecer un isomorfismo de $\mathcal{L}(U , V)$ sobre $\mathcal{L}(V^* , U^*)$, por el cual cada transformación $T \in \mathcal{L}(U, V)$ se aplica sobre su dual $T^* \in \mathcal{L}(V^*, U^*)$.

§ 6. Transformaciones lineales y matrices

Todos los espacios vectoriales que se consideran en este parágrafo son de dimensión finita.

6.1. Algebras.

Definición 6.1.1. Se llama *álgebra* sobre un campo $\mathbb{F}$ a un espacio

vectorial A al cual está asociado una aplicación bilineal

$$m: A \times A \rightarrow A$$

que se denomina *producto.*

Dados $x,y \in A$, el vector $m(x,y) \in A$ se llama producto de x por y, y se representa por xy.

Si el producto cumple la propiedad asociativa

$$x(yz) = (xy)z \quad ,$$

se dice que el álgebra A es una *álgebra asociativa.*

Si existe en A un elemento e, tal que $xe = ex = x$ para todo $x \in A$, se dice que e es el *elemento unidad* (o unielemento) y que A es un *álgebra con elemento unidad.* Si el espacio vetorial A es dimensión finita, esa dimensión se llama *dimensión* del álgebra.

Un elemento x de álgebra A se llama *invertible* si existe otro elemento $x^{-1} \in A$, que se llama *inversa* de x tal que

$$xx^{-1} = x^{-1}x = e \ .$$

Se puede probar fácilmente que cada elemento de A tiene a lo más una inversa.

6.2. Algebra de los endomorfismos de un espacio vectorial.

Según vimos en la sección 3.1 el conjunto $\mathcal{L}(U,V)$ de las transformaciones lineales de U en V admite una estructura de espacio vectorial. En la presente sección vamos a tratar acerca del espacio $\mathcal{L}(V,V)$ de las transformaciones lineales de V, es decir de los endomorfismos de V, espacios que representaremos, como ya dijimos en la sección 3.1 por $\mathcal{E}(V)$.

Vamos a demostrar que en $\mathcal{E}(V)$ es posible introducir una nueva ley de composición interna, a saber el *producto de endomorfismos.*

Recordemos primero que si U,V,W son tres espacios vectoriales y $T_1: U \rightarrow V$ y $T_2: V \rightarrow W$ son transformaciones lineales, entonces $T_2 T_1$ es la transformación lineal de U en W tal que

$$(T_2 T_1)x = T_2(T_1 x) \quad , \quad x \in U.$$

La transformación $T_2 T_1$ es entonces el *producto* de T_2 por T_1 .

Vamos a enunciar algunas propiedades de esta operación, cuya demostración es un ejercicio sencillo. Sean

$$T_1 : U \to V \ , \quad T_2 : V \to W \ , \quad T_3 : W \to Z \ ,$$

transformaciones lineales cualesquiera y α, β escalares arbitrarios. Se tiene entonces

$$T_3(T_2 T_1) = (T_3 T_2)T_1 \qquad \textit{(asociatividad)}$$

$$\left.\begin{aligned} (\alpha T_1 + \beta T_2)T_3 &= \alpha T_1 T_3 + \beta T_2 T_3 \\ T_1(\alpha T_2 + \beta T_3) &= \alpha T_1 T_2 + \beta T_1 T_3 \end{aligned}\right\} \quad \textit{(bilinealidad)}$$

Si a las operaciones vectoriales definidas en $\mathfrak{L}(V)$ añadimos la operación de producto que acabamos de introducir resultará que $\mathfrak{L}(V)$ será un álgebra asociativa cuyo elemento identidad es evidentemente la transformación idéntica I_V. Obsérvese que esa álgebra no es conmutativa, es decir no siempre es $T_1 T_2 = T_2 T_1$ (Dar un ejemplo).

El espacio $\mathfrak{L}(V)$ tiene un subconjunto particularmente importante, a saber el subconjunto formado por los automorfismos de V, es decir por los endomorfismos regulares de V sobre V. Dicho subconjunto, que se designa por $GL(V)$, *no constituye un subespacio* de $\mathfrak{L}(V)$, pues la suma de dos automorfismos no es necesariamente un automorfismo. Así, por ejemplo, si T es un automorfismo, entonces $-T$ también lo es y la suma de ambos es $T + (-T) = 0$, que no es un automorfismo porque obviamente no es sobreyectiva.

El producto $T_2 T_1$ de dos automorfismos es también un automorfismo. Es fácil observar que respecto de esa operación $GL(V)$ es un grupo no conmutativo cuando $\dim V \geqslant 2$. (Por ejemplo, para los automorfismos f y g de $\mathbb{R}^3$, precedentemente definidos, se tiene $(fg)e_1 = e_1$ y $(gf)e_1 = -4e_3$ y por tanto $fg \neq gf$). Su elemento neutro es la transformación idéntica I_V. $GL(V)$ se llama *grupo lineal* de V.

Ejercicios

*6.2.1. Un endomorfismo $\pi: V \to V$ se llama *proyección* si satisface a la condición

$$\pi^2 = \pi\pi = \pi.$$

Demostrar que si π es una proyección, entonces

$$Im\ \pi \oplus Ker\ \pi = V.$$

Además, si π es una proyección de V, π^* es una proyección de V^*, y los espacios $\pi(V)$ y $\pi^*(V^*)$ son duales.

*6.2.2. Se llama *involución* a un endomorfismo $E: V \to V$ tal que $E^2 = I_V$. Probar que si π es una proyección, $2\pi - I$ es una involución, y que cada involución puede expresarse en la forma $2\pi - I_V$, donde π es una proyección.

*6.2.3. Si $T_1: U \to V$ y $T_2: V \to W$, entonces $(T_2 T_1)^* = T_1^* T_2^*$.

*6.2.4. Si W es un espacio $(n + m)$-dimensional y si U es un subespacio m-dimensional de W, entonces existe un subespacio n-dimensional V de W tal que

$$W = U \oplus V.$$

Además existe una proyección π de W, tal que para cada $w \in W$, si $w = u + v$, $u \in U$, $v \in V$, entonces es

$$\pi(w) = v.$$

Diremos que π *es una proyección de* W *en* V *según* U. Se prueba además que $I - \pi$ es una proyección de W en U, según V.

*6.2.5. Sea T un endomorfismo del espacio de dimensión finita V. Demostrar que si cualquiera que sea el endomorfismo S de V es $TS = ST$, entonces existe un escalar λ tal que

$$T = \lambda I_V.$$

6.2.6. Sea V un espacio de dimensión 2 sobre el campo J_2 de las clases de restos de enteros módulo 2. Probar que entonces

$GL(V)$ es isomorfo con S_3 , el grupo simétrico de orden 3.

6.3. **Matrices.** Se llama *matriz* de tipo $n \times m$ o simplemente matriz $n \times m$, a un conjunto de nm escalares

$$\alpha_j^i \quad , \qquad (i = 1, \ldots, n \ , \ j = 1, \ldots, m) \ ,$$

dispuestos en un cuadro rectangular, de la manera siguiente

$$A = \begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \ldots & \alpha_m^1 \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_m^2 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ \alpha_1^n & \alpha_2^n & \ldots & \alpha_m^n \end{pmatrix} \qquad (6.3.1)$$

Los escalares α_j^i se llaman *términos* de la matriz. Los términos $\alpha_1^i, \alpha_2^i, \ldots, \alpha_m^i$ constituyen la *fila* i de la matriz, y los términos $\alpha_j^1, \alpha_j^2, \ldots, \alpha_j^n$ constituyen la *columna* j . Por tanto, la matriz tiene n filas, de m términos cada una, y m columnas, de n términos cada una. La matriz (6.3.1) se representa en forma abreviada por (α_j^i) . Sin embargo, no siempre dispondremos los índices que caracterizan a los términos uno arriba y otro abajo, ni tampoco mantendremos siempre el convenio de que el índice superior indique el orden de la fila en que encuentra el término y el índice inferior el orden de la columna. En alguna ocación llamaremos término (i,j) al término situado en la fila de orden i y en la columna de orden j.

El conjunto $\mathbb{M}(n \times m)$ de todas las matrices $n \times m$ constituye un espacio vectorial respecto de las operaciones siguientes:

$$A + B = (\alpha_j^i) + (\beta_j^i) = (\alpha_j^i + \beta_j^i)$$

$$\lambda A = \lambda(\alpha_j^i) = (\lambda \alpha_j^i) . \qquad (\lambda \in \mathbb{F}) \ ,$$

donde λ es un escalar cualquiera. En particular, $(\lambda \alpha_j^i)$ designa a la matriz cuyo término de la fila i y de la columna j es $\lambda \alpha_j^i$. Se comprueba fácilmente que esas operaciones cumplen los axiomas de es-

pacio vectorial, que la matriz cero es aquella cuyos términos son todos nulos y que $M(n \times m)$ es un espacio de dimensión nm (De mostrarlo!).

6.4. Matriz de una transformación lineal. Sean U y V dos espacios vectoriales de dimensión m y n respectivamente, y sea $T : U \to V$ una transformación lineal cualquiera. Si suponemos que $\mathfrak{E} = \{e_1, \ldots, e_m\}$ es una base de U y $\mathfrak{F} = \{f_1, \ldots, f_n\}$ es una base de V, entonces es fácil observar que T queda bien definida cuando son conocidos los vectores Te_j, que son las imágenes de los vectores de la base de U, las cuales pueden ser expresadas en la forma

$$Te_j = \sum_{i=1}^{n} \alpha_j^i f_i \quad , \quad (j = 1, \ldots, m). \qquad (6.4.1)$$

donde $\alpha_j^1, \alpha_j^2, \ldots, \alpha_j^n$ son las componentes de Te_j respecto de la base $\mathfrak{F}$. Resulta de esta manera que la matriz

$$M = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_m^1 \\ \vdots & \vdots\vdots & \vdots \\ \alpha_1^n & \ldots & \alpha_m^n \end{pmatrix} \qquad (6.4.2)$$

cuyas columnas son las componentes de los vectores Te_j, basta para definir completamente la transformación T cuando son dadas las bases $\mathfrak{E}$ y $\mathfrak{F}$ de U y V. Por eso, dicha matriz se llama *matriz de la transformación T respecto de las bases $\mathfrak{E}$ y $\mathfrak{F}$* y se le representa por $M(T, \mathfrak{E}, \mathfrak{F})$, o bien, si no hay ambigüedad respecto de las bases, por $M(T)$.

Es importante señalar que la matriz que representa a una transformación lineal respecto de bases dadas dependen del orden en que se suponen ordenados los vectores de las bases. Por este motivo, en particular *cuando se trata de la representación matricial de transformaciones lineales es esencial considerar a las bases como conjuntos ordenados.*

Llamaremos en adelante *vector-columna* de la matriz $M = M(T)$,

y designaremos por a_j , a cada uno de los m vectores

$$a_j = \sum_{i=1}^{n} \alpha_j^i f_i \quad , \quad (j = 1, \ldots, m) \ , \qquad (6.4.3)$$

del espacio V. Semejantemente, llamaremos *vector-fila* de la matriz $M = M(T)$, y designaremos por a a cada uno de los n vectores

$$a^i = \sum_{j=1}^{m} \alpha_j^i e^j \quad , \quad (i = 1, \ldots, n) \ , \qquad (6.4.4)$$

de U^*, en donde $\{e^1, \ldots, e^m\}$ es la base dual de la base $\{e_j\}$ de U.

Las relaciones (6.4.1) pueden entonces ser escritas en la forma

$$Te_j = a_j \quad , \quad (j = 1, \ldots, m) \ , \qquad (6.4.5)$$

En la observación del final de la sección 6.6 veremos que si

$$T^* : V^* \to U^*$$

es la transformación dual de T, entonces se puede escribir también la relación

$$T^* f^i = a^i \ , \ (i = 1, \ldots, n) \ , \qquad (6.4.6)$$

donde $\{f^1, \ldots, f^n\}$ es la base dual de (f_i).

Por otra parte, dados los espacios vectoriales U y V, y sus bases $\mathfrak{E}$ y $\mathfrak{F}$, cada matriz (α_j^i) de tipo $n \times m$ determina a una y solo a una transformación lineal $T: U \to V$, que es definida por las ecuaciones (6.4.1).

Existe por tanto una biyección

$$M : \mathcal{L}(U,V) \to \mathcal{M}(n \times m)$$

que aplica a cada transformación lineal $T: U \to V$ en la matriz $M(T, \mathfrak{E}, \mathfrak{F})$; y es fácil comprobar que M es un isomorfismo del espacio vectorial $\mathcal{L}(U,V)$ sobre el espacio vectorial $\mathcal{M}(n \times m)$, pero que ese isomorfismo varia con las bases $\mathfrak{E}$ y $\mathfrak{F}$ que se consideren. (Demostrarlo !) .

Dado un vector cualquiera $u = \sum_{j=1}^{m} \xi^j e_j$ de U, se tiene

$$Tu = \sum_{j=1}^{m} \xi^j Te_j = \sum_{j=1}^{m} \xi^j a_j = \sum_{j=1}^{m} \sum_{i=1}^{n} \alpha_j^i \xi^j f_i \quad ,$$

de modo que si $\eta^i \quad (i = 1, \ldots, n)$ designan a las componentes de Tu respecto de la base (f_i), se tiene

$$\eta^i = \sum_{j=1}^{m} \alpha_j^i \xi^j \quad , \ (i = 1, \ldots, n), \qquad (6.4.7)$$

que son las ecuaciones *(ecuaciones de componentes)* que expresan a las componentes de Tu en términos de las componentes de u. En forma explícita esas ecuaciones se escriben

$$\left.\begin{aligned} \eta^1 &= \alpha_1^1 \xi^1 + \alpha_2^1 \xi^2 + \cdots + \alpha_m^1 \xi^m \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \eta^n &= \alpha_1^n \xi^1 + \alpha_2^n \xi^2 + \cdots + \alpha_m^n \xi^m \end{aligned}\right\} \qquad (6.4.8)$$

en donde se ve que la matriz de los coeficientes de los segundos miembros,

$$\begin{pmatrix} \alpha_1^1 & \alpha_2^1 & \cdots & \alpha_m^1 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^n & \alpha_2^n & \cdots & \alpha_m^n \end{pmatrix} \qquad (6.4.9)$$

es justamente la matriz α que hemos denominado matriz de T. (Notese bien que la matriz de los coeficientes de los segundos miembros de las ecuaciones (6.4.1) que expresan a las imágenes de los vectores de base es

$$\begin{pmatrix} \alpha_1^1 & \alpha_1^2 & \cdots & \alpha_1^n \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_m^1 & \alpha_m^2 & \cdots & \alpha_m^n \end{pmatrix}$$

que resulta de la matriz (6.4.9) mediante una simetría respecto de la diagonal que baja de izquierda a derecha *(diagonal principal)*.

Ejercicio

6.4.1. Sea $T: \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal cuya matriz respecto de la base canónica es

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

Calcular las matrices de T respecto de las dos bases siguientes

a) $(1,1,1)$, $(0,1,1)$, $(0,0,1)$

b) $(1,1,0)$, $(1,2,0)$, $(1,2,1)$.

6.5. Vectores-columna y vectores-fila de matrices cualesquiera. Es conveniente extender la idea de vector-columna y de vector-fila a matrices cualesquiera, aun cuando no sean consideradas como representativas de transformaciones lineales.

Dada la matriz

$$M = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_m^1 \\ \vdots & \vdots\ \vdots & \vdots \\ \alpha_1^n & \dots & \alpha_m^n \end{pmatrix}$$

de tipo $n \times m$, supongamos que U y V son espacios vectoriales, arbitrariamente fijados, de dimensión m y n respectivamente. Sean además $\mathcal{E} = \{e_1, \dots, e_m\}$ y $\mathcal{F} = \{f_1, \dots, f_n\}$ bases que suponemos dadas en ellos.

Llamaremos entonces *vectores-columna* de M, o simplemente *columnas* a los vectores

$$a_j = \sum_{i=1}^{n} \alpha_j^i f_i \quad , \quad (j = 1, \dots, m)$$

de V, y llamaremos *vectores-fila* de M, o simplemente *filas*, a los vectores

$$a^i = \sum_{j=1}^{m} \alpha^i_j e^j \quad , \quad (i = 1, \ldots, n)$$

de U^*, donde $\{e^1, \ldots, e^m\}$ es la base dual de la base (e_j). No hay inconveniente en suponer que U es $\mathbb{F}^m$, que V es $\mathbb{F}^n$, y que las bases $\mathcal{E}$ y $\mathcal{F}$ son las respectivas bases canónicas. En ese caso U^* se identifica con $\mathbb{F}^m$ y V^* con $\mathbb{F}^n$ (ver los ejercicios 4.2. 1, 5.3.2 y 5.3.3). La base (e^j) resulta ser entonces la base recíproca de la base (e_j), y ambas son idénticas a la base canónica.

6.6. Matriz de la transformación dual de una transformación lineal.

Definición 6.6.1. Dada una matriz cualquiera

$$M = \begin{pmatrix} \alpha^1_1 & \ldots & \alpha^1_m \\ \vdots & & \vdots \\ \alpha^n_1 & \ldots & \alpha^n_m \end{pmatrix}$$

se llama *matriz transpuesta* de M a la matriz

$$M^t = \begin{pmatrix} \alpha^1_1 & \ldots & \alpha^n_1 \\ \vdots & & \vdots \\ \alpha^1_m & \ldots & \alpha^n_m \end{pmatrix}$$

que resulta mediante la simetría de la matriz M respecto de su diagonal principal.

Si la matriz M^t se expresa en la forma

$$M^t = \begin{pmatrix} \beta^1_1 & \ldots & \beta^1_m \\ \vdots & & \vdots \\ \beta^n_1 & \ldots & \beta^n_m \end{pmatrix}$$

entonces se cumplen las relaciones

$$\beta_i^j = \alpha_j^i \quad , \quad (i = 1, \ldots, n \; ; \; j = 1, \ldots, m) .$$

Teorema 6.6.1. Sean U, U^* y V, V^* dos pares de espacios duales. (e_j) y (e^j), $(j = 1, \ldots, m)$, sean base duales de los primeros, y (f_i) y (f^i), $(i = 1, \ldots, n)$, sean bases duales de los dos últimos. Supongamos que las transformaciones $T: U \to V$ y $T^*: V^* \to U^*$ son duales. Nos proponemos demostrar que si

$$M = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_m^1 \\ \vdots & & \vdots \\ \alpha_1^n & \ldots & \alpha_m^n \end{pmatrix}$$

es la matriz $n \times m$ que coresponde a la transformación T, respecto de las bases (e_j) y (f_i) ;

$$M^* = \begin{pmatrix} \beta_1^1 & \ldots & \beta_n^1 \\ \vdots & & \vdots \\ \beta_1^m & \ldots & \beta_n^m \end{pmatrix}$$

es la matriz $m \times n$ que corresponde a la transformación T^* respecto de las bases (f^i) y (e^j) , entonces se cumplen las relaciones

$$\beta_i^j = \alpha_j^i \quad , \quad (i = 1, \ldots, n \; ; j = 1, \ldots, m) .$$

Es decir que la matriz M^* es idéntica a la transpuesta M^t de M.

Demostración. Por definición de las matrices de las transformaciones, se tiene

$$Te_j = \sum_{i=1}^{n} \alpha_j^i f_i \quad , \quad (j = 1, \ldots, m) , \qquad (6.6.1)$$

$$T^* f^i = \sum_{j=1}^{m} \beta_i^j e^j \quad , \quad (i = 1, \ldots, n) , \qquad (6.6.2)$$

Por la propiedad fundamental de las transformaciones duales podemos

escribir

$$\langle f^i , Te_j \rangle = \langle T^* f^i , e_j \rangle .$$

Empleando (6.6.1) y (6.6.2) tendremos entonces

$$\langle f^i , \sum_{k=1}^{n} \alpha_j^k f_k \rangle = \langle \sum_{l=1}^{m} \beta_i^l e^l , e_j \rangle .$$

Desarrollando ambos miembros, teniendo en cuenta la bilinealidad, y recordando que

$$\langle f^i , f_k \rangle = \delta_k^i \quad \text{y} \quad \langle e^l , e_j \rangle = \delta_j^l ,$$

se obtiene

$$\beta_i^j = \alpha_j^i \quad (i = 1, \ldots, n \; ; j = 1, \ldots, m). \quad \blacksquare$$

Observación. En virtud de la relación (6.6.2) y del teorema anterior, podemos escribir

$$T^* f^i = \sum_{j=1}^{m} \alpha_j^i e^j \qquad (i = 1, \ldots, n).$$

Teniendo en cuenta entonces las fórmulas (6.4.4) que definen a los vectores-filas, tendremos las fórmulas

$$T^* f^i = a^i \quad , \quad (i = 1, \ldots, n) ,$$

que son las fórmulas (6.4.6) a que ya hicimos referencia en la sección 6.4.

Ejercicios

***6.6.1.** Demostrar que los vectores-columna de la matriz de T^* son los vectores-fila de la matriz de T, y que los vectores-fila de la matriz de T^* son los vectores-columna de la matriz de T. (Deberá tenerse presente el ejercicio 5.1.1).

6.6.2. Sea $\pi : V \to V$ una proyección (ver el ejercicio 6.2.1). Demostrar que existe en V una base respecto de la cual la matriz de V tiene la forma

$$\left.\begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \cdot & & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \cdot & & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}\right\} \begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} m \\ \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} n-m \end{matrix}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{n}$$

donde n es la dimensión de V y $m \leqslant n$ es un entero.

6.7. Rango de una matriz. Consideremos una matriz cualquiera

$$M = (\alpha_j^i) \quad ,$$

de tipo $n \times m$. Refiramos los vectores-columna y los vectores-fila de M a los espacios U y V cualesquiera de bases $\mathfrak{E} = \{e_1, \dots, e_m\}$ y $\mathfrak{F} = \{f_1, \dots, f_n\}$, respectivamente, en la forma que fué descrita en la sección 6.5. (Recordemos que, como hemos visto en la sección 6.5, puede suponerse que $U = \mathbb{F}^m$ y que $V = \mathbb{F}^n$. Nuestras consideraciones son independientes de los espacios U y V y de las bases $\mathfrak{E}$ y $\mathfrak{F}$ que se considere en ellas).

Vamos a designar por r_1 al número máximo de vectores-fila de la matriz M, que son *l.i.*; y por r_2 al número máximo de vectores-columna de la misma matriz, que son *l.i.* Puesto que los vectores-fila pertenecen al espacio U^* de dimensión m, y los vesctores-columna pertenecen al espacio V de dimensión n, se tendrá $r_1 \leqslant m$ y $r_2 \leqslant n$. *Vamos a demostrar que*

$$r_1 = r_2$$

Definamos una transformación lineal $T : U \to V$, mediante las fórmulas

$$Te_j = a_j \qquad (j = 1, \dots, m) \ ,$$

donde los a_j son los vectores-columna de la matriz α. La matriz de T referida a las bases $\mathcal{E}$ y $\mathcal{F}$ es evidentemente M. El subespacio $Im\ T$, como es fácil reconocer, es generado por los vectores columna a_j. Por consiguiente, el rango de T (ver la sección 3.2) es igual al número máximo de vectores a_j que son *l.i.* (corolario 1 del teorema 2.5.2), es decir que

$$r(T) = r_2 \ . \qquad (6.7.1)$$

La transformación $T^* : V^* \to U^*$, dual de T, es definida (ver la observación de la sección 6.6) por las fórmulas

$$T^* f^i = a^i \ , \quad (i = 1, \ldots, n) \ ,$$

donde los a^i son los vectores-fila de la matriz α. El subespacio

$$Im\ T^* \subset U^*$$

es entonces generado por los vectores-fila a^i, y por consiguiente el rango de T^* (o sea la dimensión de $Im\ T^*$) es igual al número máximo de vectores a^i que son *l.i.*, es decir que

$$r(T^*) = r_1 \ . \qquad (6.7.2)$$

Ahora bien, el teorema 5.5.3 establece que dos transformaciones duales tienen el mismo rango. Luego es $r(T) = r(T^*)$ y por consiguiente, resulta de (6.7.1) y (6.7.2) que

$$r_1 = r_2 = r(T) = r(T^*) \ .$$

Hemos demostrado de esta manera una proposición que tiene mucha importancia en la teoría que estamos desarrollando, y que puede enunciarse así:

Teorema 6.7.1. A cada matriz M de tipo $n \times m$ le corresponde un número natural $r = r(M) \leqslant m, n$, que se llama *rango de la matriz* y que es igual a los siguientes números:

a) Al número máximo de vectores-fila de M que son *l.i.* ,
b) Al número máximo de vectores-columna de M que son *l.i.* ,
c) Al rango $r(T)$ de cada transformación lineal $T : U \to V$ cuya matriz respecto de bases dadas de U y V es la matriz M.

d) Al rango $r(T^*)$ de cada transformación lineal $T^*: V^* \to U^*$ cuya matriz respecto de bases dadas de V^* y U^* es la matriz M^t transpuesta de M.

Ejercicios

*6.7.1. Demostrar que el rango de una matriz es independiente de los espacios vectoriales y de las bases que se empleen para expresar sus vectores-fila y sus vectores-columna.

6.7.2. ¿Cuáles son los rangos de las transformaciones lineales T_1 y T_2 de $\mathbb{R}^3$ en $\mathbb{R}^3$ dadas por las fórmulas:

$$T_1(\xi^1, \xi^2, \xi^3) = (\xi^1 + 3\xi^2 - 2\xi^3, -2\xi^1 + \xi^2 + \xi^3, 7\xi^2 - 3\xi^3)$$

$$T_2(\xi^1, \xi^2, \xi^3) = (\xi^1 - 2\xi^2, 3\xi^1 + \xi^2 + 7\xi^3, -2\xi^1 + \xi^2 - 3\xi^3)\ ?$$

(Comparar las matrices de estas dos transformaciones).

6.7.3. ¿Cuáles son los valores de los números η^1, η^2, η^3 para los cuales existen soluciones ξ^1, ξ^2, ξ^3 del sistema de ecuaciones

$$\begin{aligned} \xi^1 + 3\xi^2 - 2\xi^3 &= \eta^1 \\ -2\xi^1 + \xi^2 + \xi^3 &= \eta^2 \\ 7\xi^2 - 3\xi^3 &= \eta^3\ ? \end{aligned}$$

(Hacer suo de los resultados del ejercicio anterior).

6.7.4. Calcular el rango de la matriz

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

6.8. Producto de matrices. En la sección 6.3 hemos visto que las matrices de tipo $n \times m$ constituyen el espacio vectorial $\mathcal{M}(n \times m)$. En la sección 6.4 establecimos un isomorfismo M entre el espacio $\mathcal{M}(n \times m)$ y el espacio $\mathcal{L}(U, V)$ de las transformaciones lineales del espacio U de dimensión m en el espacio V de dimensión n. Ese isomorfismo depende de las bases que se elijan en U y V; pero una vez elegidas éstas, es bien definido y permite representar cada transformación lineal mediante una matriz determinada. El producto de las transformaciones de $\mathcal{L}(U, V)$ por las de $\mathcal{L}(V, W)$, que fué introducido en la sección 3.3, sugiere la posibilidad de definir, en base al isomorfismo a que nos hemos referido en el comienzo de la presente sección, un producto de las matrices de $\mathcal{M}(p \times n)$ por las matrices de $\mathcal{M}(n \times m)$, de manera que la matriz del producto $T_2 T_1$ de dos transformaciones lineales $T_1 : U \to V$ y $T_2 : V \to W$ resulte ser igual al producto $M(T_2)M(T_1)$ de las matrices de esas transformaciones.

Sean

$$A = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_m^1 \\ \vdots & \vdots\vdots & \vdots \\ \alpha_1^n & \dots & \alpha_m^n \end{pmatrix} \qquad \text{y} \qquad B = \begin{pmatrix} \beta_1^1 & \dots & \beta_n^1 \\ \vdots & \vdots\vdots & \vdots \\ \beta_1^p & \dots & \beta_n^p \end{pmatrix}$$

dos matrices de tipos $n \times m$ y $p \times n$ respectivamente. Sean entonces U, V y W espacios vectoriales de dimensiones m, n y p respectivamente, y $\mathfrak{E} = \{e_1, \dots, e_m\}$, $\mathcal{F} = \{f_1, \dots, f_n\}$ y $\mathcal{G} = \{g_1, \dots, g_p\}$ bases dadas en dichos espacios. Supondremos ahora que

$$T_1 : U \to V \qquad \text{y} \qquad T_2 : V \to W$$

son las transformaciones lineales cuyas matrices relativas a las mencionadas bases son A y B respectivamente. Para calcular la matriz de $T_2 T_1$ basta, según sabemos, determinar las imágenes por esa transformación, de los vectores de la base $\mathfrak{E} = (e_j)$ de U en términos de los vectores de la base $\mathcal{G} = (g_k)$ de W. Con tal fin podemos

escribir

$$(T_2 T_1) e_j = T_2(T_1 e_j) = T_2\Big(\sum_{i=1}^{n} \alpha_j^i f_i\Big) = \sum_{i=1}^{n} \alpha_j^i T_2 f_i$$

$$= \sum_{i=1}^{n} \alpha_j^i \Big(\sum_{k=1}^{p} \beta_i^k g_k\Big) = \sum_{k=1}^{p} \Big(\sum_{i=1}^{n} \beta_i^k \alpha_j^i\Big) g_k$$

Resulta de esta manera que la matriz de $T_2 T_1$ es

$$C = \begin{pmatrix} \gamma_1^1 & \cdots & \gamma_m^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \gamma_1^p & \cdots & \gamma_m^p \end{pmatrix}$$

donde

$$\gamma_j^k = \sum_{i=1}^{n} \beta_i^k \alpha_j^i \quad .$$

Para que se cumpla la fórmula

$$M(T_2 T_1) = M(T_2) M(T_1) \quad ,$$

o sea

$$C = BA$$

es necesario establecer la siguiente definición .

Definición 6.8.1. Se llama *producto de las matrices*

$$B = \begin{pmatrix} \beta_1^1 & \cdots & \beta_n^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \beta_1^p & \cdots & \beta_n^p \end{pmatrix} \qquad \text{y} \qquad A = \begin{pmatrix} \alpha_1^1 & \cdots & \alpha_m^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_1^n & \cdots & \alpha_m^n \end{pmatrix}$$

de tipos $p \times n$ y $n \times m$, a la matriz

$$BA = \begin{pmatrix} \gamma_1^1 & \cdots & \gamma_m^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \gamma_1^p & \cdots & \gamma_m^p \end{pmatrix}$$

de tipo $p \times m$, cuyos términos son dados por la fórmula

$$\gamma_j^k = \sum_{i=1}^{n} \beta_i^k \alpha_j^i$$

La aplicación de la definición que acabamos de dar, a las matrices

$$(\beta_1^k \ \ldots \ \beta_n^k) \qquad \text{y} \qquad \begin{pmatrix} \alpha_j^1 \\ \vdots \\ \alpha_j^n \end{pmatrix},$$

formadas por los términos de la fila k de B y de la columna j de A, nos da

$$(\beta_1^k \ \ldots \ \beta_n^k) \begin{pmatrix} \alpha_j^1 \\ \vdots \\ \alpha_j^n \end{pmatrix} = \beta_1^k \alpha_j^1 + \beta_n^k \alpha_j^n = \sum_{i=1}^{n} \beta_i^k \alpha_j^i = \gamma_j^k .$$

Por tanto podemos establecer que *el elemento* γ_j^k *del producto* BA *es igual al producto de la matriz - fila de orden* k *de* B *por la matriz - columna de orden* j *de* A.

La definición 6.8.1 nos permite enunciar, en virtud de la precedente exposición, el siguiente teorema

Teorema 6.8.1. Dados los espacios U, V y W y las transformaciones $T_1 : U \to V$ y $T_2 : V \to W$, entre las matrices de T_1, T_2 y $T_2 T_1$ relativas a bases dadas de U, V y W, se cumple la relación

$$M(T_2 T_1) = M(T_2) M(T_1). \qquad (6.8.1)$$

Los isomorfismos a que hemos hecho referencia al comienzo de esta sección nos permiten introducir la inversa M^{-1} de la aplicación M. De esta manera, dadas las matrices

$$B \in \mathcal{M}(p \times n) \ , \quad A \in \mathcal{M}(n \times m) \quad \text{y} \quad BA \in \mathcal{M}(p \times m) \ ,$$

se determinan las transformaciones lineales

$$M^{-1}(B) : U \to V \ , \quad M^{-1}(A) : V \to W \quad \text{y} \quad M^{-1}(BA) : U \to W$$

cuyas matrices, para bases dadas de U, V y W, son respectivamente B, A y BA. Puede establecerse entonces el teorema siguiente:

Teorema 6.8.2. Dadas las matrices $B \in M(p \times n)$ y $A \in M(n \times m)$, entre las transformaciones $M^{-1}(B)$, $M^{-1}(A)$ y $M^{-1}(BA)$ se cumple la relación

$$M^{-1}(BA) = M^{-1}(B)\,M^{-1}(A) . \qquad (6.8.2)$$

Demostración. Si escribimos $T_1 = M^{-1}(A)$ y $T_2 = M^{-1}(B)$, entonces se tiene $M(T_1) = A$ y $M(T_2) = B$. La relación (6.8.1) puede escribirse entonces en la forma

$$M(M^{-1}(B)\,M^{-1}(A)) = BA.$$

Aplicando el isomorfismo M^{-1} a ambos miembros resulta inmediatamente la relción (6.8.2). ■

6.9. Algebra de las matrices cuadradas. Grupo de las matrices regulares.

En la sección 6.2 hemos establecido las propiedades fundamentales del producto de transformaciones lineales. Si aplicamos a los dos miembros de cada una de esas relaciones el isomorfismo M se obtienen las siguientes propiedades del producto de matrices:

$$C(BA) = (CB)A \qquad \textit{(asociatividad)}$$

$$\left.\begin{aligned} (\alpha A + \beta B)C &= \alpha AC + \beta BC \\ A(\alpha B + \beta C) &= \alpha AB + \beta BC \end{aligned}\right\} \quad \textit{(bilinealidad)}$$

relaciones en las cuales, naturalmente, las matrices deben ser de tipos tales que tengan sentido los productos y las adiciones que figuren en las fórmulas.

El isomorfismo M hace corresponder, en particular, a los endomorfismos del espacio n-dimensional V, las matrices $n \times n$ o *matrices cuadradas de orden n*, es decir matrices que tienen n filas y n columnas. Los resultados que ya hemos obtenido muestran que las matrices cuadradas de orden n constituyen una álgebra (ver la sección 6.2) asociativa, cuyo elemento unidad es la matriz

$$1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = (\delta^i_j) \ .$$

que algunas veces representaremos por E_n o simplemente por E. Esa álgebra será designada por $\mathbb{M}(n)$, y es evidentemente isomórfica con el álgebra $\mathfrak{E}(V)$ de los endomorfismos de V, mediante la correspondencia que aplica cada endomorfismo de V en la matriz que le corresponde cuando se elije una base fija de V.

Los automorfismos de V son los endomorfismos cuyo rango es igual a $dim\ V = n$, es decir los endomorfismos cuyas matrices tienen rango n, (teorema 6.7.1). Recíprocamente, cada matriz de $\mathbb{M}(n)$, de rango n, corresponde a un endomorfismo de V que tiene rango n, y que por tanto es un automorfismo. Resulta así que las matrices $n \times n$ de rango n, a las cuales llamaremos *matrices regulares*, corresponden biunívocamente a los automorfismos de V. Se puede comprobar fácilmente que el conjunto de las matrices regulares de orden n constituyen un grupo respecto de la operación de producto de matrices, y que ese grupo es isomorfo del grupo $GL(V)$ que es el grupo lineal de V que fué introducido en la sección 6.2. El grupo de las matrices regulares de orden n cuyos términos son elementos del campo $\mathbb{F}$ se representa por $GL(n, \mathbb{F})$ y se llama *grupo lineal general.* Cada matriz regular A tiene una matriz A^{-1}, que también es regular, de manera que

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

Si A es la matriz del automorfismo T, entonces A^{-1} es la matriz del automorfismo T^{-1}. (Demostrarlo !). Las matrices cuadradas (α^i_j) tales que $\alpha^i_j = 0$ para $i \neq j$ se llaman *matrices diagonales.*

Ejercicios

***6.9.1.** Mostrar en detalle que las matrices regulares de orden n constituyen un grupo que es isomorfo con el grupo lineal de un espacio vectorial de dimensión n.

*6.9.2. Demostrar las siguientes relaciones entre matrices.

$$(A + B)^t = A^t + B^t$$
$$(\alpha A)^t = \alpha A^t$$
$$(A^{-1})^t = (A^t)^{-1}$$
$$(AB)^t = B^t A^t$$

6.9.3. Una matriz cuadrada se llama *triangular superior (inferior)* si todos sus elementos situados debajo (encima) de la diagonal principal son iguales a cero. Probar que la suma y el producto de matrices triangulares son también matrices triangulares.

6.9.4. Denotemos por A_{ij} a la matriz cuadrada de orden n cuyo término (i,j) es igual a 1, y cuyos otros términos son ceros. Demostrar la fórmula

$$A_{ij}A_{jk} = A_{ik}$$

y probar que las matrices A_{ij} forman una base del espacio $M(n)$.

6.10. Cambio de base. Sea V un espacio de dimensión n.

$$\mathcal{E} = \{e_1, \ldots, e_n\} \quad \text{y} \quad \mathcal{E}' = \{e_{1'}, \ldots, e_{n'}\}$$

sean dos bases de ese espacio. Existirán entonces los escalares

$$A_i^{i'} \quad \text{y} \quad A_{i'}^{i}\ ,^{1)}$$

---

1) La notación que empleamos tiene ventajas pero requiere, una explicación, que vamos a dar brevemente. Escribimos $A_i^{i'}$, para representar a los símbolos de la forma $A_1^{1'}, A_2^{1'}, \ldots, A_1^{2'}, A_2^{2'}$, etc., los cuales, más correctamente deberíamos expresar en la forma $(A')_1^1, (A')_2^1, \ldots, (A')_1^2, (A')_2^2, \ldots$, etc., en cuyo caso su expresión genérica habría sido $(A')_j^i$. Semejantemente, $A_{i'}^{i}$ representa a símbolos de la forma $A_{1'}^{1}, A_{2'}^{2}, \ldots, A_{1'}^{2}, A_{2'}^{2}, \ldots$, etc., que deberían escribirse más bien en la forma $(A_{,})_{1}^{1}, (A_{,})_{2}^{1}, \ldots, (A_{,})_{1}^{2}, (A_{,})_{2}^{2}, \ldots$, etc., y que entonces en forma genérica deberían haberse representado por $(A_{,})_{j}^{i}$. La notación que estamos empleando tiene ventajas que se verán más claras cuando se estudie el cálculo tensorial. No debe interpretarse el símbolo $A_i^{i'}$ o el símbolo $A_{i'}^{i}$ como si i e i' fueran números naturales sino como si i e i' fueran índices, el primero de los cuales puede tomar los valores $1, \ldots, n$, y el segundo los valores $1', \ldots, n'$. El uso aclarará el empleo de esta notación. Nótese en especial que n' no designa aquí a un entero diferente de n sino al mismo entero n pero acompañado de un acento. Semejantemente, emplearemos los símbolos $e_{i'}$ y $e^{j'}$ y otros análogos, en vez de $(e_{,})_{i}$, $(e')^{j}$.

tales que se cumplan las relaciones siguientes:

$$e_i = \sum_{i'=1'}^{n'} A_i^{i'} e_{i'} \quad , \ (i = 1, \ldots, n) \qquad (6.10.1)$$

$$e_{i'} = \sum_{i=1}^{n} A_{i'}^{i} e_i \quad , \ (i' = 1', \ldots, n'). \qquad (6.10.2)$$

Si x^i $(i = 1, \ldots, n)$ son las componentes de un vector cualquiera $x \in V$ en la base $\mathfrak{E}$ y $x^{i'}$ son las componentes del mismo vector en la base $\mathfrak{E}'$, tendremos

$$x = \sum_{i=1}^{n} x^i e_i = \sum_{i=1}^{n} \sum_{i'=1'}^{n'} x^i A_i^{i'} e_{i'} = \sum_{i'=1'}^{n'} x^{i'} e_{i'} = \sum_{i'=1'}^{n'} x^{i'} A_{i'}^{i} e_i .$$

Se deduce de aquí que

$$x^{i'} = \sum_{i=1}^{n} A_i^{i'} x^i \quad , \quad x^i = \sum_{i'=1'}^{n'} A_{i'}^{i} x^{i'} \qquad (6.10.3)$$

Las primeras de estas ecuaciones permiten el cálculo de las componentes del vector x relativas a la base $\mathfrak{E}'$ en términos de las componentes respecto de la base $\mathfrak{E}$. Las segundas resuelven el problema inverso.

La matriz

$$A = \begin{pmatrix} A_1^{1'} & \ldots & A_n^{1'} \\ \vdots & \vdots\vdots\vdots & \vdots \\ A_1^{n'} & \ldots & A_n^{n'} \end{pmatrix} \qquad (6.10.4)$$

se llama *matriz de cambio de la base* $\mathfrak{E}$ *en la base* $\mathfrak{E}'$. En virtud de (6.10.1) cada una de sus columnas está formada por las componentes de uno de los vectores de la base $\mathfrak{E}$, relativas a la base $\mathfrak{E}'$.

La matriz

$$A' = \begin{pmatrix} A_{1'}^{1} & \ldots & A_{n'}^{1} \\ \vdots & \vdots\vdots\vdots & \vdots \\ A_{1'}^{n} & \ldots & A_{n'}^{n} \end{pmatrix} \qquad (6.10.5)$$

se llama *matriz de cambio de la base* $\mathfrak{E}'$ *en la base* $\mathfrak{E}$. En virtud de (6.10.2) cada una de sus columnas está formada por las componentes de uno de los vectores de la base $\mathfrak{E}'$, respecto de la base $\mathfrak{E}$.

Los vectores e_i $(i = 1, \ldots, n)$ pueden ser considerados, en virtud de 6.10.1, como los vectores-columna de la matriz A, y los vectores $e_{i'}$ $(i' = 1', \ldots, n')$ como los vectores columna de la matriz A'. Dado que tanto los vectores e_i como los vectores $e_{i'}$ son *l.i.*, resultará que *las matrices A y A' tienen rango n, y por tanto son regulares*. Vamos a probar que son inversas entre sí.

En efecto, si reemplazamos (6.10.2) en (6.10.1) se obtiene

$$e_i = \sum_{i'=1'}^{n'} \sum_{i=1}^{n} A_i^{i'} A_{i'}^{k} e_k \qquad (i = 1, \ldots, n),$$

de donde

$$\sum_{k=1}^{n} (\delta_i^k - \sum_{i'=1}^{n'} A_i^{i'} A_{i'}^{k}) e_k = 0 \quad , \ (i = 1, \ldots, n);$$

y puesto que los vectores e_k son *l.i.*, resulta

$$\sum_{i'=1'}^{n'} A_i^{i'} A_{i'}^{k} = \delta_i^k \qquad (i, k = 1, \ldots, n);$$

Es decir que

$$A'A = 1.$$

En forma semejante, sustituyendo (6.10.1) en (6.10.2) se obtiene que

$$AA' = 1.$$

Por consiguiente,

$$A^{-1} = A' \quad , \qquad (6.10.6).$$

lo cual demuestra que las matrices A y A' son inversas entre sí. Podemos, en resumen, establecer el

Teorema 6.10.1. Si $\mathfrak{E} = (e_1, \ldots, e_n)$ y $\mathfrak{E}' = (e_{1'}, \ldots, e_{n'})$ son dos bases del espacio V entre las cuales se cumplen las relaciones

$$e_i = \sum_{i'=1}^{n'} A_i^{i'} e_{i'} \quad , \qquad e_{i'} = \sum_{i=1}^{n} A_{i'}^{i} e_i \quad ,$$

entonces las matrices $A = (A_i^{i'})$ y $A' = (A_{i'}^{i})$ de cambio de base son regulares e inversas entre sí.

Ejercicios

6.10.1. Dadas las bases $\{(1,3,2), (2,2,-1), (-1,0,3)\}$ y $\{(-1,2,-2), (0,2,-1), (-3,1,2)\}$ de $\mathbb{R}^3$, calcular las matrices de cambio y comprobar que son inversas.

6.10.2. Sea

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

la matriz de cambio de la base $\mathcal{E} = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$ en la base $\mathcal{E}' = \{e_{1'}, e_{2'}, e_{3'}\}$. Expresar los vectores de la base $\mathcal{E}'$ en términos de los vectores de la base $\mathcal{E}$.

6.11. Matrices semejantes. En esta sección nos proponemos en primer lugar estudiar cómo varía la matriz de un endomorfismo T de V cuando se cambia la base $\mathcal{E} = (e_1, \ldots, e_n)$ por la base $\mathcal{E}' = (e_{1'}, \ldots, e_{n'})$. Sea $M = (\alpha_i^j)$ la matriz de T relativa a la base $\mathcal{E}$ y sea $M' = (\alpha_{j'}^{i'})$ la matriz de la misma transformación relativa a la base $\mathcal{E}'$. Podemos escribir evidentemente

$$Te_{j'} = T\Big(\sum_{i=1}^{n} A_{j'}^{i} e_i\Big) = \sum_{i=1}^{n} A_{j'}^{i} Te_i = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{j'}^{i} \alpha_i^j e_j$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i'=1'}^{n'} A_{j'}^{i} \alpha_i^j A_j^{i'} e_{i'} = \sum_{i'=1'}^{n'} \Big(\sum_{i=1}^{n}\sum_{j=1}^{n} A_{j'}^{i} \alpha_i^j A_j^{i'}\Big) e_{i'} \quad .$$

Resulta de esta manera que la matriz M' de T respecto de la base

$\mathcal{E}'$ es dada por

$$\alpha^{i'}_{j'} = \sum_{i=1}^{n} \sum_{j=1}^{n} A^{i}_{j'} \alpha^{j}_{i} A^{i'}_{j} \quad .$$

Se reconoce inmediatamente , teniendo presente la definición del producto de matrices y recordando que la matriz $A' = (A^{i}_{i'})$ es inversa de la matriz $A = (A^{i'}_{i})$, que

$$M' = AMA^{-1} \qquad (6.11.1)$$

Podemos observar ahora que si $\mathcal{E}$ es una base de V, respecto de la cual la matriz del endomorfismo T es M, entonces si $A = (A^{i'}_{i})$ es una matriz regular cualquiera de $M(n)$, donde n es la dimensión de V, las ecuaciones (6.10.2) definen una nueva base $\mathcal{E}'$ de V, tal que la matriz de cambio de la base $\mathcal{E}$ en la base $\mathcal{E}'$ es A, y que la matriz de T respecto de la nueva base $\mathcal{E}'$ es $M' = AMA^{-1}$. Esta observación nos induce a establecer la siguiente definición.

Definición 6.11.1. Dos matrices M y M' de $M(n)$ se llaman *semejantes* si existe una matriz regular A de $M(n)$ tal que .

$$M' = AMA^{-1}$$

Se demuestra fácilmente que la relación así establecida entre las matrices de $M(n)$ es efectivamente una relación de equivalencia. (Demostrarlo !).

El resultado (6.11.1) puede ser expresado entonces en el siguiente teorema.

Teorema 6.11.1. Las matrices que representan a un endomorfismo de un espacio vectorial de n dimensiones respecto de todas las posibles bases de dicho espacio constituyen una clase de equivalencia de la relación de equivalencia entre las matrices de $M(n)$, definida por

$$M \sim M' \Leftrightarrow M' = AMA^{-1} \quad \text{para} \quad A \in M(n) \ ,$$

regular.

6.12. Endomorfismos semejantes.

Definición 6.12.1. Dados los espacios V y W de la misma dimensión n, se dice que los endomorfismos T de V y S de W son *semejantes* si existe un isomorfismo $A : V \to W$ tal que se cumple la relación

$$S = ATA^{-1} . \qquad (6.12.1)$$

Es fácil comprobar que la relación de semejanza así definida es una relación de equivalencia entre los endomorfismos de espacios de la misma dimensión.

Teorema 6.12.1. La condición necesaria y suficiente para que los endomorfismos T de V y S de W sean semejantes es que existan en V y W bases respecto de las cuales las matrices de T y S sean iguales.

Demostración. Si S y T son semejantes la relación (6.12.1) implica la relación

$$M(S) = M(A)M(T)M(A)^{-1}$$

entre las matrices de S,T y A relativas a bases cualesquiera $\mathfrak{E}$ y $\mathfrak{F}$ de V y W respectivamente. Sea $\mathfrak{F}'$ una nueva base de W tal que la matriz de cambio de la base $\mathfrak{F}$ en la base $\mathfrak{F}'$ sea la matriz regular $M(A)^{-1}$. La matriz $M'(S)$ de S respecto de la nueva base será

$$M'(S) = M(A)^{-1} M(S)M(A) = M(T).$$

Recíprocamente, supongamos que respecto de las bases $\mathfrak{E} = \{e_i\}$ de V y $\mathfrak{F} = \{f_i\}$ de W las matrices $M(S)$ y $M(T)$ sean iguales a la matriz (α_j^i). Definamos el isomorfismo $A : V \to W$ por las relaciones

$$A(e_i) = f_i$$

Se tiene entonces que

$$S(f_i) = \sum_{j=1}^{n} \alpha_i^j f_i \quad \text{y} \quad T(e_i) = \sum_{j=1}^{n} \alpha_i^j e_j$$

Por consiguiente

$$S(f_i) = \sum_{j=1}^{n} \alpha_i^j A(e_i) = A(\sum_{j=1}^{n} \alpha_i^j e_i) = AT(e_i) = ATA^{-1}(f_i),$$

de donde se deduce que

$$S = ATA^{-1} . \qquad \blacksquare$$

En esta sección y en la sección precedente hemos introducido una relación de equivalencia entre matrices y otra entre endomorfismos. Entre las correspondientes clases de equivalencia puede establecerse una biyección, de manera que los endomorfismos que pertenecen a una clase pueden representarse mediante bases adecuadas por las matrices de la clase correspondiente (Comprobar esta afirmación!).

Ejercicio

6.12.1. Las bases mediante las cuales un endomorfismo dado T puede representarse por una matriz dada M son tales que las matrices de cambio entre ellas son las matrices regulares de cambio que conmutan con la matriz M.

CAPITULO II

DETERMINANTES Y SISTEMAS DE ECUACIONES LINEALES

§ 7. Determinantes

No obstante que suponemos que el lector tiene un conocimiento de la noción de determinante y de sus propiedades, expondremos aquí la teoría pero sin detenernos mucho en detalles y demostraciones que son conocidas. Insistiremos más bien en los aspectos que serán para nosotros de mayor interés en lo que sigue. Conviene observar aquí que los términos de los determinantes que consideraremos son elementos de un campo $\mathbb{F}$ de característica 0, que en particular puede ser $\mathbb{R}$ o $\mathbb{C}$.

7.1. Definición de un determinante de orden n.

Consideremos una matriz cuadrada de orden n con términos pertenecientes al campo $\mathbb{F}$, es decir una matriz del conjunto que hemos designado por $\mathbb{M}(n)$:

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

A partir de esta matriz podemos obtener $n!$ (factorial de n) escalares de la manera siguiente:

Elijamos en cada columna j de la matriz un término $\alpha_j^{i_j}$, pero de manera que los números $i_1, i_2, \ldots, i_n$ sean los números $1, \ldots, n$ escritos en un orden arbitrario. Así pues, los términos $\alpha_1^{i_1}, \ldots, \alpha_n^{i_n}$ no solamente pertenecerán a columnas diferentes, sino también a filas diferentes. Multiplicando los términos elegidos como se ha dicho, resulta el escalar

$$\alpha_1^{i_1} \; \alpha_2^{i_2} \; \ldots \; \alpha_n^{i_n} \quad , \qquad (7.1.1)$$

que es uno de los escalares que nos hemos propuesto construir. Este escalar puede expresarse también en la forma

$$\alpha_1^{\sigma(1)} \; \alpha_2^{\sigma(2)} \; \ldots \; \alpha_n^{\sigma(n)} \quad , \qquad (7.1.2)$$

donde σ representa a la permutación de los números $(1,2, \ldots, n)$ tal que $\sigma(j) = i_j$. Si σ recorre a las $n!$ permutaciones de esos escalares resultarán $n!$ números de la forma (7.1.2).

Ahora bien, cada permutación σ puede ser par o impar según sea producto de un número par o impar de trasposiciones. Si ϵ_σ designa a un número que es igual a 1 si σ es permutación par y a -1 si σ es permutación impar, el *determinante* de la matriz M dada se define por la fórmula

$$D = \det M = \begin{vmatrix} \alpha_1^1 & \alpha_2^1 & \ldots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_1^n & \alpha_2^n & \ldots & \alpha_n^n \end{vmatrix} = \sum_{\sigma \in S_n} \epsilon_\sigma \, \alpha_1^{\sigma(1)} \; \ldots \; \alpha_n^{\sigma(n)} \quad ,$$

donde S_n designa al grupo simétrico de orden n es decir al grupo de las permutaciones de los n primeros números.

Corrientemente se denomina filas y columnas de un determinan-

te a las filas y columnas de la matriz que les da origen.

De la precedente definición resultan propiedades conocidas que nos limitaremos aquí a enunciar.

1. El determinante de una matriz M es igual al determinante de su traspuesta.
2. Si se permutan dos columnas o dos filas de una matriz, su determinante cambia de signo. *(antisimetría)*.
3. El determinante de una matriz que tiene dos filas o dos columnas iguales, es igual a cero.
4. Si la columna j de un determinante

$$D = \begin{vmatrix} \alpha_1^1 & \dots & \alpha_n^1 \\ \vdots & \vdots & \vdots \\ \alpha_1^n & \dots & \alpha_n^n \end{vmatrix}$$

es una combinación lineal de dos columnas de escalares (β^i) y (γ^i), es decir si

$$\alpha_j^i = \lambda\beta^i + \mu\gamma^i$$

donde λ y μ son elementos de $\mathbb{F}$, entonces el determinante D es igual a la combinación lineal

$$D = \lambda D_1 + \mu D_2$$

donde D_1 es el determinante que resulta sustituyendo en D la columna j por la columna de los números (β^i) y D_2 es el determinante que resulta sustituyendo en D la columna j por la columna de los números (γ^i).

La precedente propiedad se cumple también cuando se consideran filas en vez de columnas.

Se deducen inmediatamente de la propiedad 4 las consecuencias siguientes:

Si todos los términos de una columna (fila) de un determinante tienen un factor común λ, este puede sacarse fuera del de

terminante como un factor.

Si todos los términos de una columna (fila) de un determinante son nulos, el determinante es nulo.

5. El valor de un determinante no varía si a los términos de una columna (fila) se suma una combinación lineal arbitraria de los términos de las demás columnas (filas).

Definición 7.1.1. El determinante que resulta cuando se suprimen de un determinante dado D la fila i y la columna j se llama *menor* de D correspondiente a α_j^i y se designa por D_i^j (Notar la inversión de los índices).

El número

$$A_i^j = (-1)^{i+j} D_i^j \qquad (7.1.3)$$

se llama *cofactor* del elemento α_j^i de D.

Se pueden establecer entonces las siguientes *fórmulas de expansión del determinante D según los términos de su columna j y de su fila i:*

$$\begin{aligned} D &= \alpha_j^1 A_1^j + \alpha_j^2 A_2^j + \cdots + \alpha_j^n A_n^j = \sum_{i=1}^{n} \alpha_j^i A_i^j \\ D &= \alpha_1^i A_i^1 + \alpha_2^i A_i^2 + \cdots + \alpha_n^i A_i^n = \sum_{j=1}^{n} \alpha_j^i A_i^j \end{aligned} \qquad (7.1.4)$$

Estos resultados pueden expresarse en el teorema siguiente:

Teorema 7.1.1. La suma de los productos de los términos de una columna (fila) del determinante D por los correspondientes cofactores es igual a D.

Corolario. La suma de los productos de los términos de una columna (fila) del determinante D por los cofactores de los correspondientes elementos de cualquier otra columna (fila), es igual a cero.

Demostración. La suma a que se refiere el corolario resulta ser, en virtud de la primera (segunda) fórmula (7.1.4), el determinante de una matriz que tiene dos columnas (filas) iguales. ■

Ejercicios

7.1.1. Calcular los determinantes siguientes:

a) $\begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix}$; b) $\begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix}$; c) $\begin{vmatrix} \operatorname{sen}\alpha & \cos\alpha \\ -\cos\alpha & \operatorname{sen}\alpha \end{vmatrix}$

d) $\begin{vmatrix} a & c + di \\ c - di & b \end{vmatrix}$; e) $\begin{vmatrix} \alpha + \beta i & \gamma + \delta i \\ \gamma - \delta i & \alpha - \beta i \end{vmatrix}$

f) $\begin{vmatrix} tg\,\alpha & -1 \\ 1 & tg\,\alpha \end{vmatrix}$; g) $\begin{vmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{vmatrix}$

h) $\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix}$; i) $\begin{vmatrix} a & a & a \\ -a & a & x \\ -a & -a & x \end{vmatrix}$

j) $\begin{vmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w \end{vmatrix}$, donde $w = \cos\frac{2\pi}{3} + i\,\operatorname{sen}\frac{2\pi}{3}$

k) $\begin{vmatrix} 1 & 1 & w \\ 1 & 1 & w^2 \\ w^2 & w & 1 \end{vmatrix}$, donde $w = \cos\frac{2\pi}{3} + i\,\operatorname{sen}\frac{2\pi}{3}$

7.1.2. En el desarrollo del determinante de quinto orden (a_{ij}) el término $a_{1r}\, a_{32}\, a_{4s}\, a_{25}\, a_{53}$ aparece con signo positivo. Encontrar los valores de r y s.

7.1.3. Usar únicamente la definición del determinante para calcular los coeficientes de x^4 y de x^3 en el desarrollo de

$$\begin{vmatrix} 2x & x & 1 & 2 \\ 1 & x & 1 & -1 \\ 3 & 2 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix}$$

7.2. Matriz inversa de una matriz cuyo determinante es diferente de cero.

Teorema 7.2.1. Si el determinante D de una matriz $A = (\alpha^i_j)$ de orden n es diferente de cero, entonces la matriz es invertible y su inversa es la matriz

$$A^{-1} = \begin{pmatrix} \beta^1_1 & \dots & \beta^1_n \\ \vdots & \vdots\vdots\vdots & \vdots \\ \beta^n_1 & \dots & \beta^n_n \end{pmatrix}$$

cuyos términos son dados por las fórmulas

$$\beta^i_j = \frac{A^i_j}{D} \quad . \qquad (7.1.5)$$

(Nótese bien que A^i_j es el cofactor de α^j_i) .

Demostración. El producto de las matrices (α^i_j) y (β^i_j) es la matriz (γ^i_j) cuyos términos son

$$\gamma^i_j = \sum_{k=1}^{n} \alpha^i_k \beta^k_j = \sum_{k=1}^{n} \alpha^i_k \frac{A^k_j}{D} = \frac{1}{D} \sum_{k=1}^{n} \alpha^i_k A^k_j = \delta^i_j \quad ,$$

por cuanto, en virtud del teorema 7.1.1 y su corolario, se tiene

$$\sum_{k=1}^{n} \alpha^i_k A^k_j = \begin{cases} D & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \qquad (7.1.6)$$

Resulta así que (γ_j^i) es la matriz unidad y por tanto (β_j^i) es la inversa de la matriz (α_j^i). ■

Ejercicios

7.2.1. Hallar la inversa de cada una de las siguientes matrices:

a) $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ b) $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

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

7.2.2. Probar que si M es una matriz de tipo $m \times n$ y A es la matriz cuadrada de orden m (resp. n) de la forma

$$A = \begin{array}{c} \\ \\ \\ i \rightarrow \\ \\ j \rightarrow \\ \\ \end{array} \begin{array}{c} \begin{matrix} & & & i\downarrow & & j\downarrow & \\ \end{matrix} \\ \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}$$

entonces $M_1 = AM$ (resp. $M_2 = MA$) es la matriz que resulta intercambiando las filas (resp. las columnas) de la matriz M.

7.2.3. Probar que si M es una matriz de tipo $m \times n$ y B es la matriz cuadrada de orden m (resp. n) de la forma

$$B = \begin{array}{c} \\ \\ \\ i \to \\ \\ j \to \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccccc} & & & i & & j & & \\ & & & \downarrow & & \downarrow & & \end{array} \\ \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & -c & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{array}$$

entonces $M_1 = BM$ (resp. $M_2 = MB$) es la matriz que resulta restando de la fila i (resp. columna j) de M el múltiplo c de la fila j (resp. columna i).

7.2.4. Calcular las inversas de las matrices siguientes, si las tienen:

a) $\begin{pmatrix} 0 & -1 & -5 \\ -1 & 2 & 9 \\ -3 & 6 & 28 \end{pmatrix}$ b) $\begin{pmatrix} 1 & -1 & 2 \\ 2 & 3 & 4 \\ 0 & 2 & 0 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 6 \\ 3 & -2 & 0 \end{pmatrix}$ d) $\begin{pmatrix} -1 & 0 & 0 \\ 2 & 3 & 6 \\ -5 & 1 & -2 \end{pmatrix}$

e) $\begin{pmatrix} 3 & 4 & 5 \\ 2 & 4 & 5 \\ 3 & 3 & 4 \end{pmatrix}$ f) $\begin{pmatrix} 1 & 3 & 5 \\ 6 & 4 & 7 \\ 4 & 1 & 2 \end{pmatrix}$

7.3. Aplicación de los determinantes a la resolución de los sistemas de n ecuaciones lineales con n incógnitas: Regla de Cramer.

Sea el sistema de ecuaciones

$$\sum_{j=1}^{n} \alpha_j^i \xi^j = \eta^i \qquad (i = 1,2,\dots,n) \qquad (7.3.1)$$

donde los escalares α_j^i así como los η^i son dados. Los primeros se llaman coeficientes del sistema y los últimos se llaman los segundos miembros. Resolver el sistema es determinar escalares $\xi^1, \ldots, \xi^n$ que lo satisfagan. Cada conjunto $(\xi^1, \ldots, \xi^n)$ que satisface al sistema es una solución.

Supongamos que el sistema tiene una solución que designaremos por $(\xi^1, \ldots, \xi^n)$. Si introducimos las matrices

$$M = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_n^1 \\ \vdots & \vdots & \vdots \\ \alpha_1^n & \ldots & \alpha_n^n \end{pmatrix}, \quad X = \begin{pmatrix} \xi^1 \\ \vdots \\ \xi^n \end{pmatrix}, \quad Y = \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^n \end{pmatrix}$$

el sistema (7.3.1) puede expresarse en la forma matricial

$$MX = Y \qquad (7.3.2)$$

Si la matriz M es invertible, es decir regular, lo cual ocurre en particular *si se cumple la condición*

$$D = \det M \neq 0 ,$$

según vimos en la sección anterior, entonces podemos multiplicar ambos miembros de (7.3.2) y obtenemos

$$X = M^{-1} Y \qquad (7.3.3)$$

La matriz M^{-1} es la matriz (β_j^i) cuyos términos son dados por las fórmulas (7.1.5) en que A_j^i son los cofactores de los términos de D. Por consiguiente, teniendo en cuenta la definición del producto de matrices, de (7.3.3) se deducen las expresiones

$$\xi^j = \sum_{k=1}^{n} \beta_k^j \eta^k = \sum_{k=1}^{n} \eta^k \frac{A_k^j}{D} = \frac{1}{D} \sum_{k=1}^{n} \eta^k A_k^j .$$

Ahora bien, de la primera de las fórmulas 7.1.4 es deduce que

$$\sum_{k=1}^{n} \eta^k A_k^j$$

es el determinante de la matriz que resulta de la matriz M cuando se sustituye la columna j por la columna (η^k). Si designamos a ese determinante por D^j, obtendremos las fórmulas

$$\xi^j = \frac{D^j}{D} \qquad (j = 1, \ldots, n) \qquad (7.3.4)$$

que dan los valores de las soluciones, supuesto que éstas existen. Además se cumplirá, siempre que sea $D \neq 0$, que si las soluciones existen, son únicas. Las fórmulas (7.3.4) se llaman *fórmulas de Cramer*, y podemos demostrar inmediatamente que los valores que proporcionan son efectivamente soluciones del sistema (7.3.1). En efecto, sustituyendo en el primer miembro de (7.3.1) los ξ^j por los valores (7.3.4) resulta

$$\sum_{j=1}^{n} \alpha_j^i \frac{D^j}{D} = \sum_{j=1}^{n} \alpha_j^i \frac{\sum_{k=1}^{n} \eta^k A_k^j}{D} = \frac{1}{D} \sum_{k=1}^{n} \eta^k \sum_{j=1}^{n} \alpha_j^i A_k^j = \eta^i$$

por cuanto, en virtud de (7.1.6) es $\sum_{j=1}^{n} \alpha_j^i A_k^j = D\delta_k^i$. Queda así comprobado que si $D \neq 0$, las fórmulas (7.3.4) proporcionan las soluciones del sistema (7.3.1). El resultado obtenido permite establecer la siguiente proposición.

Teorema 7.3.1. *(Regla de Cramer).* Si el determinante del sistema (7.3.1), es decir el determinante cuyos términos son los coeficientes del sistema, es diferente de cero, entonces el sistema tiene una solución única $(\xi^1, \ldots, \xi^n)$ y el valor de ξ^j es dado por una fracción cuyo denominador es el determinante D del sistema y cuyo numerador es el determinante obtenido reemplazando la columna j de D por la columna formada por los segundos miembros del sistema.

Corolario. Un sistema de ecuaciones lineales que es homogéneo, es decir cuyos segundos miembros son nulos, y que tiene determinante diferente de cero solo admite la solución trivial, para la cual todas

las incógnitas $\xi^1, \ldots, \xi^n$ son nulas.

Demostración. Basta con aplicar la regla de Cramer. ∎

Ejercicios

Resolver los siguientes sistemas:

7.3.1.
$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 &= 13 \\ 2x_1 + x_2 + 2x_3 + 3x_4 + 4x_5 &= 10 \\ 2x_1 + 2x_2 + x_3 + 2x_4 + 3x_5 &= 11 \\ 2x_1 + 2x_2 + 2x_3 + x_4 + 2x_5 &= 6 \\ 2x_1 + 2x_2 + 2x_3 + 2x_4 + x_5 &= 3 . \end{aligned}$$

7.3.2.
$$\begin{pmatrix} 2 & -1 & -1 \\ 3 & 4 & -2 \\ 3 & -2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \\ 11 \end{pmatrix}$$

7.3.3.
$$\begin{pmatrix} 1 & 1 & 2 \\ 2 & -1 & 2 \\ 4 & 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \\ -2 \end{pmatrix}$$

7.3.4.
$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 11 \end{pmatrix}$$

7.4. Dependencia lineal de las columnas y filas de una matriz. Rango de una matriz.

Los vectores-columna y los vectores-fila se llaman aquí columnas y filas respectivamente.

Teorema 7.4.1. Si una de las columnas de una matriz $n \times n$ es una

combinación lineal de las restantes, entonces su determinante es igual a cero.

Demostración. Este teorema es consecuencia inmediata de la propiedad 5 de la sección 7.1. ∎

Definición 7.4.1. Dada una matriz $n \times m$

$$M = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_m^1 \\ \vdots & \vdots\ \vdots & \vdots \\ \alpha_1^n & \dots & \alpha_m^n \end{pmatrix},$$

si se eligen k filas y k columnas cualesquiera y se eliminan todos los términos que no están situados en las intersecciones de esas filas y columnas, se obtiene una matriz cuadrada de orden k cuyo determinante se llama *menor de orden k* de la matriz M.

Definición 7.4.2. Dada una matriz cualquiera cuyos términos no son todos nulos, se llama *menor de base* a un menor no nulo de orden ρ, tal que todo menor de orden mayor que ρ, si es que existe, es igual a cero.

Las ρ columnas que contienen a un menor de base se llama *columnas de base*.

A cada matriz M le podemos hacer corresponder por tanto un número $\rho(M)$ que es igual a 0 si todos los términos de la matriz son nulos, y que es igual al número ρ de la definición 7.4.2 en caso contrario.

Teorema 7.4.2. *(Teorema de las columnas de base).* Cualquier columna de una matriz es una combinación lineal de sus columnas de base.

Demostración. Es evidente que si la propiedad enunciada tiene lugar para una matriz, tiene lugar también para cualquier otra que resulte de ella por una permutación arbitraria de sus columnas o de sus filas. Podemos suponer por eso que el menor de base de la matriz M considerada está situado en las $\rho = \rho(M)$ primeras columnas y en las ρ primeras filas.

Cualesquiera que sean los números s $(1 \leqslant s \leqslant m)$ y k $(1 \leqslant k \leqslant n)$, consideremos el determinante

$$D_s^k = \begin{vmatrix} \alpha_1^1 & \alpha_2^1 & \dots & \alpha_\rho^1 & \alpha_s^1 \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_\rho^2 & \alpha_s^2 \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots & \vdots \\ \alpha_1^\rho & \alpha_2^\rho & \dots & \alpha_\rho^\rho & \alpha_s^\rho \\ \alpha_1^k & \alpha_2^k & \dots & \alpha_\rho^k & \alpha_s^k \end{vmatrix}$$

que se obtiene orlando el menor de base con los términos de la fila k y de la columna s.

Si es $k \leqslant \rho$ o bien si $s \leqslant \rho$, este determinante es evidentemente nulo, porque tiene entonces dos filas o dos columnas idénticas. Si $k > \rho$ y si $s > \rho$, el mismo determinante es nulo, porque es un menor de orden $\rho + 1$ de una matriz para la que $\rho(M) = \rho$. Quiere decir que $D_s^k = 0$ para k y s cualesquiera.

Desarrollemos ahora a D_s^k según los términos de su última fila. Obtendremos

$$\alpha_1^k A_k^1 + \alpha_2^k A_k^2 + \dots + \alpha_\rho^k A_k^\rho + \alpha_s^k A_k^s = 0 \qquad (7.4.1)$$

donde $A_k^1, A_k^2, \dots, A_k^\rho, A_k^s$ designan a los cofactores de los elementos de la última fila de D_s^k. Estos cofactores no dependen del número k sino únicamente del menor de base. Podemos introducir por tanto la notación

$$C^j = A_k^j \quad , \qquad (j = 1, \dots, \rho, s).$$

Sustituyendo en (7.4.1) y atribuyendo luego a k los valores $1, \dots, n$ se obtienen las n relaciones

$$\sum_{j=1, \dots, \rho, s} C^j \alpha_j^i = 0 \qquad (i = 1, \dots, n) \qquad (7.4.2)$$

El número $C^s = A_k^s$ es diferente de cero, pues A_k^s es el menor de

base de la matriz dada M. Dividiendo (7.4.2) por C^s y escribiendo

$$-\frac{C^j}{C^s} = \lambda^j \quad , \qquad (j = 1, \ldots, \rho),$$

se obtiene

$$\alpha^i_j = \sum_{j=1}^{\rho} \lambda^j \alpha^i_j \quad , \qquad (i = 1, \ldots, n). \qquad (7.4.3)$$

Estas ecuaciones muestran que la columna s de la matriz M es una combinación lineal de las ρ primeras columnas de la matriz. Puesto que s es cualquiera de los números $1, \ldots, m$, el teorema queda demostrado. ■

Teorema 7.4.3. Si el determinante D de la matriz (cuadrada) M tiene el valor cero, entonces M tiene por lo menos una columna que es combinación lineal de las restantes. En otros términos: si un determinante se anula, sus columnas son linealmente dependientes.

Demostración. Si M es la matriz nula la proposición es trivial. Supongamos pues que M no es la matriz nula. Puesto que su determinante es nulo, sus menores de base serán de orden $\rho < n$. Por consiguiente, seleccionadas ρ columnas de base, resta por lo menos una columna de la matriz. Según el teorema anterior esa columna es combinación lineal de las bases y por tanto el teorema se cumple. ■

De los teoremas 7.4.1 y 7.4.3 se deduce la proposición siguiente

Teorema 7.4.4. Un determinante se anula si y solo si sus columnas son linealmente dependientes.

En la sección 6.5. vimos cómo cada columna de una matriz $n \times m$ puede imaginarse que está formada por las componentes de un vector (vector columna) de un espacio V de dimensión n, referido a una cierta base de ese espacio (Puede elegirse $V = \mathbb{F}^n$ y la base canónica).

Teorema 7.4.5. Los vectores columna de una matriz generan un subespacio de V cuya dimensión es igual a $\rho(M) = \rho$.

Demostración. Observemos primero que , en virtud del teorema 7.4.2, cada vector columna de la matriz es combinación lineal de las ρ columnas de base. Por tanto dichos vectores-columna de base generan un espacio que contiene a todos los vectores-columna. Para probar el teorema basta con demostrar que las columnas de base son *l.i.* Supongamos por el contrario que las columnas de base fueran *l.d.* Dichas columnas constituyen una matriz $n \times \rho$, cuyas submatrices cuadradadas de orden ρ tendrían sus columnas linealmente dependientes, y por tanto, por el teorema 7.4.4, los determinantes de todas esas submatrices serían nulos. Se llegaría así a una contradicción por cuanto una de dichas submatrices sería el menor de base y su determinante debería ser diferente de cero. Se concluye por tanto que las columnas de base son *l.i.* ■

Corolario 1. El rango $r(M)$ de una matriz M es igual al orden del menor de mayor orden que tiene determinante diferente de cero, es decir que $r(M) = \rho(M)$.

Demostración. Del teorema anterior y de la definición del rango de una matriz se deduce que $r(M) = \rho(M)$. Basta entonces tener presente la definición de $\rho(M)$. ■

Corolario 2. Las matrices regulares de orden n (sección 6.9) son aquellas cuyo determinante es diferente de cero.

Demostración. Las matrices regulares de orden n son las de rango $r(M) = n$. Según el teorema 7.4.5 será entonces $\rho(M) = n$, y por tanto el determinante de la matriz debe ser diferente de cero. ■

Observación. En todos los enunciados de la presente sección pueden ser sustitudas las columnas por las filas; y los razonamientos, salvo modificaciones obvias son válidos.

Definición 7.4.3. Dados p vectores $v_1, \ldots, v_p$ de un espacio vectorial de dimensión n, se llama *rango* de ese conjunto de vectores al número máximo de vectores *l.i.* contenido en él. Dicho rango es evidentemente igual al rango de la matriz $n \times p$ cuyas columnas (filas) están formadas por las componentes de los vectores $v_1, \ldots, v_p$

respecto de una base cualquiera. En particular si $p = n$, los vectores serán *l.i.* si y solo si la matriz de sus componentes tienen rango n, es decir si su determinante es diferente de cero.

Cálculo del rango de una matriz. Consideremos una matriz de la forma

$$\begin{pmatrix} D & M \\ 0 & 0 \end{pmatrix} \qquad (7.4.4)$$

en que D es un bloque $r \times r$ cuyos términos diagonales son diferentes de cero, y cuyos términos situados debajo de la diagonal principal son todos nulos, M es un bloque arbitrario de r filas, y los otros dos bloque son nulos. El rango de la matriz (7.4.4) es r, pues el determinante de la submatriz D es diferente de cero y todo menor de orden mayor que r es nulo.

Cada matriz puede reducirse a la forma (7.4.4) mediante un número finito de las siguientes operaciones que, como es fácil observar, dejan invariable el rango:

1° Supuesto que la matriz no es nula, en cuyo caso su rango sería cero, se lleva a cabo, si es necesario, permutaciones de filas y columnas para obtener una matriz cuyo término $(1,1)$ sea distinto de cero;

2° A cada fila, salvo la *primera*, se le suma un múltiplo de esta última tal que su *primer término* resulte igual a cero;

3° Si la matriz obtenida no tiene aún la forma (7.4.4), se lleva a cabo, si es necesario, permutaciones de las $n - 1$ últimas filas y columnas de manera que se obtenga una matriz cuyo término $(2,2)$ sea distinto de cero;

4° A cada una de las filas, salvo las *dos primeras*, se le suma un multiplo de la segunda, tal que su *segundo término* resulte igual a cero

5° Si la matriz obtenida no tiene todavía la forma (7.4.4), se lleva a cabo, si es necesario, permutaciones de las $n - 2$ últimas filas y columnas de manera que se obtenga una matriz cuyo término $(3,3)$

sea distinto de cero;

6° A cada una de las filas, salvo las *tres primeras*, se le suma un múltiplo de la tercera, tal que su *tercer término* resulte igual a cero; y se prosigue así sucesivamente hasta reducir la matriz a la forma (7.4.4) y determinar por consiguiente su rango r.

Ejemplo. Con el objeto de simplificar la exposición, en cada una de las matrices que obtendremos en este ejemplo indicamos las operaciones que hay que ejecutar en ella para obtener la que le sigue de la manera siguiente: los extremos de un simple arco señalan las filas o columnas que deben ser permutadas; y un arco orientado mediante una flecha y acompañado de un número indica que la fila correspondiente al origen del arco debe multiplicarse por dicho número y el resultado debe sumarse a la fila que corresponde al extremo del arco.

La primera de las matrices que aparecen a continuación es aquella cuyo rango se trata de obtener.

$$\begin{pmatrix} 0 & 1 & 2 & 1 & 3 & 0 \\ 0 & 2 & 4 & 2 & 0 & 1 \\ 1 & 1 & 2 & 1 & 9 & 0 \\ 2 & 0 & 0 & 0 & -2 & 2 \\ 3 & -2 & -4 & -2 & 4 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 & 1 & 9 & 0 \\ 0 & 2 & 4 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 & 3 & 0 \\ 2 & 0 & 0 & 0 & -2 & 2 \\ 3 & -2 & -4 & -2 & 4 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 & 1 & 9 & 0 \\ 0 & 2 & 4 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 & 3 & 0 \\ 0 & -2 & -4 & -2 & -20 & 2 \\ 0 & -5 & -10 & -5 & -23 & 1 \end{pmatrix}$$

(-2) (-3)

En la última matriz llevamos a cabo una permutación auxiliar que permite tener el término (2,2) igual a uno, lo que facilita las operaciones que siguen:

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 9 & 0 \\ 0 & 1 & 2 & 1 & 3 & 0 \\ 0 & 2 & 4 & 2 & 0 & 1 \\ 0 & -2 & -4 & -2 & -20 & 2 \\ 0 & -5 & -10 & -5 & -23 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 & 1 & 9 & 0 \\ 0 & 1 & 2 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & -6 & 1 \\ 0 & 0 & 0 & 0 & -14 & 2 \\ 0 & 0 & 0 & 0 & -8 & 1 \end{pmatrix}$$

(-2) (2) 5

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 9 & 2 \\ 0 & 1 & 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 0 & -6 & 0 \\ 0 & 0 & 2 & 0 & -14 & 0 \\ 0 & 0 & 1 & 0 & -8 & 0 \end{pmatrix} \quad \begin{matrix} \\ \\ \\ -2 \\ -1 \end{matrix} \qquad \begin{pmatrix} 1 & 1 & 0 & 1 & 9 & 2 \\ 0 & 1 & 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 0 & -6 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 & 9 & 1 & 2 \\ 0 & 1 & 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & -6 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \end{pmatrix} \quad \begin{matrix} \\ \\ \\ \\ -1 \end{matrix} \qquad \begin{pmatrix} 1 & 1 & 0 & 9 & 1 & 2 \\ 0 & 1 & 0 & 3 & 1 & 2 \\ 0 & 0 & 1 & -6 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

La última matriz tiene la forma (7.4.4). Por tanto el rango de la matriz propuesta es igual a 4.

Ejercicios

7.4.1. Determinar el rango de las matrices siguientes:

(a) $\begin{pmatrix} 0 & 4 & 10 & 1 \\ 4 & 8 & 18 & 7 \\ 10 & 18 & 40 & 17 \\ 1 & 7 & 17 & 3 \end{pmatrix}$ (b) $\begin{pmatrix} 75 & 0 & 116 & -39 & 0 \\ 171 & -69 & 402 & 123 & 45 \\ 301 & 0 & 87 & -417 & -169 \\ 114 & -46 & 268 & 82 & 30 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & 1 & 11 & 2 \\ 1 & 0 & 4 & -1 \\ 11 & 4 & 56 & 5 \\ 2 & -1 & 5 & -6 \end{pmatrix}$ (d) $\begin{pmatrix} 14 & 12 & 6 & 8 & 2 \\ 6 & 104 & 21 & 9 & 17 \\ 7 & 6 & 3 & 4 & 1 \\ 35 & 30 & 15 & 20 & 5 \end{pmatrix}$

(e) $\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 4 & 1 \\ 1 & 1 & 1 & 5 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 0 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 & 6 \\ 1 & 2 & 3 & 14 & 32 \\ 4 & 5 & 6 & 32 & 77 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & -2 & 3 & -1 & -1 & -2 \\ 2 & -1 & 1 & 0 & -2 & -2 \\ -2 & -5 & 8 & -4 & 3 & -1 \\ 6 & 0 & -1 & 2 & -7 & -5 \\ -1 & -1 & 1 & -1 & 2 & 1 \end{pmatrix}$ (h) $\begin{pmatrix} 2 & 1 & 3 & -1 \\ 3 & -1 & 2 & 0 \\ 1 & 3 & 4 & -2 \\ 4 & -3 & 1 & 1 \end{pmatrix}$

7.4.2. Averiguar si son *l.i.* los siguientes vectores de $\mathbb{R}^3$:

a) $(1,3,7)$, $(2,-4,5)$, $(1,2,3)$.

b) $(-2,1,0)$, $(3,1,-1)$, $(-13, -1,3)$.

7.5. Determinante del producto de matrices.

Con el objeto de facilitar la exposición que sigue haremos en primer lugar algunas observaciones importantes.

Dados n vectores de $\mathbb{F}^n$.

$$a^i = (\alpha_1^i , \ldots , \alpha_n^i) , \quad i = 1 , \ldots , n ,$$

vamos a designar por

$$d(a^1 , \ldots , a^n)$$

al determinante de la matriz

$$A = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_n^1 \\ \vdots & \vdots\vdots & \vdots \\ \alpha_1^n & \ldots & \alpha_n^n \end{pmatrix}$$

cuyas filas están formadas por las componentes de los vectores dados.

Las propiedades de los determinantes, acerca de las cuales hemos tratado en la sección 7.1, se traducen en propiedades de la función

$$d : (a^1 , \ldots , a^n) \to \det A ,$$

que pueden resumirse así:

1. $d(a^1 , \ldots , a^{i-1} , a^i + b^i , a^{i+1} , \ldots , a^n) = d(a^1 , \ldots , a^{i-1} , a^i , a^{i+1} , \ldots , a^n)$

$$+ \; d(a^1 , \ldots , a^{i-1} , b^i , a^{i+1} , \ldots , a^n) ,$$

2. $d(a^1, \ldots, a^{i-1}, \lambda a^i, a^{i+1}, \ldots, a^n) = \lambda d(a^1, \ldots, a^n)$ para todo $\lambda \in \mathbb{F}$,

3. Si dos cualesquiera de los vectores a^i son iguales, entonces

$$d(a^1, \ldots, a^n) = 0 ,$$

4. $d(a^1, \ldots, a^{r-1}, a^r, a^{r+1}, \ldots, a^{s-1}, a^s, a^{s+1}, \ldots, a^n) =$

$$= -d(a^1, \ldots, a^{r-1}, a^s, a^{r+1}, \ldots, a^{s-1}, a^r, a^{s+1}, \ldots, a^n) ,$$

es decir que d cambia de signo cuando se permutan dos cualesquiera de los vectores a^i.

Las dos primeras propiedades se expresan diciendo que la función d es *multilineal*; más precisamente *n-lineal*. La propiedad 4 se expresa diciendo que d es una función *alternada*. Es facil ver que la propiedad 3 es una consecuencia de la propiedad 4.

Teorema 7.5.1. Dadas las matrices de orden n, $A = (\alpha^i_j)$ y $B = (\beta^i_j)$, entonces se cumple que

$$det\ (BA) = det\ B \cdot det\ A.$$

Demostración. Sean $a^1, \ldots, a^n$ los vectores-fila de la matriz A. Introduzcamos los n vectores

$$\begin{aligned} c^1 &= \beta^1_1 a^1 + \beta^1_2 a^2 + \cdots + \beta^1_n a^n \\ c^2 &= \beta^2_1 a^1 + \beta^2_2 a^2 + \cdots + \beta^2_n a^n \\ &\vdots \quad\ \vdots \quad\ \vdots \quad\ \vdots \\ c^n &= \beta^n_1 a^1 + \beta^n_2 a^2 + \cdots + \beta^n_n a^n \end{aligned} \qquad (7.5.1)$$

Vamos a considerar ahora el valor $d(c^1, c^2, \ldots, c^n)$, donde d es la función que fue definida anteriormente: Si sustituimos los c^i por sus expresiones (7.5.1) y aplicamos las propiedades de la función d que fueron enunciadas precedentemente, podremos escribir

$$d(c^1, \ldots, c^n) = \sum_{j_1, j_2, \ldots, j_n} \beta^1_{j_1} \beta^2_{j_2} \cdots \beta^n_{j_n}\ d(a^{j_1}, a^{j_2}, \ldots, a^{j_n})$$

donde cada término de la suma corresponde a un conjunto de números $(j_1, \ldots, j_n)$, cada uno de los cuales puede tomar los valores $1, \ldots, n$ independiente de los demás, de modo que esa suma consta de n^n términos. Sin embargo, es claro en virtud de la propiedad 3 de la función d, que si dos cualesquiera de los números, j_r son iguales, entonces el valor del correspondiente término se anula. Por tanto los únicos términos no nulos corresponden a conjuntos $(j_1, \ldots, j_n)$ en que los j_r son todos distintos. Podemos escribir por consiguiente

$$d(c^1, \ldots, c^n) = \sum_\sigma \beta^1_{\sigma(1)} \beta^2_{\sigma(2)} \cdots \beta^n_{\sigma(n)} d(a^{\sigma(1)}, \ldots, a^{\sigma(n)}),$$

donde σ es una permutación de los números $(1, \ldots, n)$ y la suma se extiende a todas las permutaciones σ.

Ahora bien, de la propiedad de antisimetría de la función d se deduce que

$$d(a^1, \ldots, a^n) = \epsilon_\sigma d(a^{\sigma(1)}, \ldots, a^{\sigma(n)}),$$

donde ϵ_σ es el símbolo que hemos introducido en la sección 7.1. La relación anterior puede escribirse

$$d(a^{\sigma(1)}, \ldots, a^{\sigma(n)}) = \epsilon_\sigma d(a^1, \ldots, a^n),$$

y por consiguiente tendremos

$$\begin{aligned} d(c^1, \ldots, c^n) &= \left(\sum_\sigma \epsilon_\sigma \beta^1_{\sigma(1)} \cdots \beta^n_{\sigma(n)}\right) det\, A \\ &= det\, B \cdot det\, A \qquad (7.5.2) \end{aligned}$$

Solo nos queda por probar que $det\,(BA) = d(c^1, \ldots, c^n)$. Sea C la matriz cuyas filas son las componentes de los vectores c^i, es decir, en virtud de (7.5.1), los números

$$\sum_{k=1}^n \beta^i_k \alpha^k_1 \,,\ \sum_{k=1}^n \beta^i_k \alpha^k_2 \,,\ \ldots,\ \sum_{k=1}^n \beta^i_k \alpha^k_n \,.$$

La definición del producto de matrices permite conocer inme-

diatamente que $C = BA$. Luego

$$d(c^1, \ldots, c^n) = det\, C = det\,(BA)$$

De donde resulta finalmente, comparando con (7.5.2),

$$det\,(BA) = det\, B.\, det\, A. \quad \blacksquare$$

Corolario 1. Si A es una matriz regular y A^{-1} es su inversa, entonces es

$$det\ A = \frac{1}{det\ A^{-1}} \qquad (7.5.3)$$

Demostración. Puesto que $AA^{-1} = 1$, donde 1 representa aquí a la matriz unidad de orden n, el teorema anterior permite escribir

$$det\ A\ det\ A^{-1} = 1$$

Se deduce de aquí que $det\ A \neq 0$, lo que confirma al corolario 2 del teorema 7.4.5, y resulta por último la relación (7.5.3). ■

Corolario 2. Dos matrices semejantes (sección 6.11) tienen el mismo determinante.

Demostración. Si M y M' son semejantes existe una matriz regular A tal que

$$M' = A M A^{-1}$$

El teorema 7.5.1 y su corolario 1 permiten escribir

$$det\ M' = det\ A.\, det\ M.\, det\ A^{-1} = det\ M\ . \quad \blacksquare$$

Si llamamos *invertibles* a las matrices A que tienen una inversa A^{-1}, de manera que

$$AA^{-1} = 1\ ,$$

entonces se deduce que $det\ A \neq 0$, y por tanto A es regular. Por consiguiente, podemos establecer el

Corolario 3. Las matrices regulares son las matrices invertibles.

Para concluir, estableceremos una fórmula que proporciona el de

sarrollo del determinante de una matriz cuadrada que es el producto de dos matrices rectangulares. Previamente introduciremos algunas nociones.

Dada una matriz $n \times m$ cualquiera

$$\alpha = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_m^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_1^n & \dots & \alpha_m^n \end{pmatrix} ,$$

y los conjuntos de números $J = \{j_1, \dots, j_p\}$ y $K = \{k_1, \dots, k_p\}$ tales que $1 \leqslant j_1 < \dots < j_p \leqslant n$ y $1 \leqslant k_1 < \dots < k_p \leqslant m$, designaremos por α_K^J a la submatriz

$$\alpha_K^J = \begin{pmatrix} \alpha_{k_1}^{j_1} & \dots & \alpha_{k_p}^{j_1} \\ \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_{k_1}^{j_p} & \dots & \alpha_{k_p}^{j_p} \end{pmatrix}$$

de α, de orden p, que resulta cuando se suprimen en α todas las filas cuyos ordenes no pertenecen a J y todas las columnas cuyos ordenes no pertenecen a K. El determinante de α_K^J se llama *menor de α de orden p relativo al par (J,K)*, y su valor es (ver la sección 7.1).

$$det\ \alpha_K^J = \sum_\sigma \epsilon_\sigma\, \alpha_{k_1}^{\sigma j_1} \dots \alpha_{k_p}^{\sigma j_p} = \sum_\sigma \epsilon_\sigma\, \alpha_{\sigma k_1}^{j_1} \dots \alpha_{\sigma k_p}^{j_p} ,$$

donde las sumas se extienden a todas las permutaciones de α de los números $\{1,2,\dots,p\}$, y hemos escrito además $j_{\sigma(i)} = \sigma j_i$ y $k_{\sigma(i)} = \sigma k_i$.

Si, en particular, es $n \leqslant m$ y $p = n$, en cuyo caso es $J = \{1,\dots,n\}$, escribiremos α_K en vez de α_K^J, y tendremos

$$det\ \alpha_K = \sum_\sigma \epsilon_\sigma\, \alpha_{\sigma k_1}^1 \dots \alpha_{\sigma k_n}^n .$$

Análogamente, si es $m \leqslant n$ y $p = m$, en cuyo caso es $K = \{1,\dots,m\}$, escribiremos α^J en vez de α_K^J, y tendremos

$$det\ \alpha^J = \sum_{\sigma} \epsilon_{\sigma}\ \alpha_{i}^{\sigma j_1} \ \ldots\ \alpha_{m}^{\sigma j_m}$$

La fórmula a que hemos hecho referencia está contenida en el siguiente teorema.

Teorema 7.5.2. Dadas las matrices rectangulares

$$\alpha = (\alpha_j^i) \quad \text{y} \quad \beta = (\beta_l^k)$$

de tipos $p \times n$ y $n \times p$ respectivamente, el determinante de la matriz cuadrada $\alpha\beta$ de orden p es dado por la fórmula

$$det\ \alpha\beta = \sum_{J} det\ \alpha_J . det\ \beta^J \quad ,$$

donde la suma se extiende a todos los conjuntos $J = \{j_1, \ldots, j_p\}$ tales que $1 \leqslant j_1 < \cdots < j_p \leqslant n$.

Demostración. Podemos escribir

$$det\ \alpha\beta = \begin{vmatrix} \sum_{j_1=1}^{n} \alpha_{j_1}^1 \beta_1^{j_1} & \ldots & \sum_{j_p=1}^{n} \alpha_{j_p}^1 \beta_p^{j_p} \\ \vdots & & \vdots \\ \sum_{j_1=1}^{n} \alpha_{j_1}^p \beta_1^{j_1} & \ldots & \sum_{j_p=1}^{n} \alpha_{j_p}^p \beta_p^{j_p} \end{vmatrix}$$

$$= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \ldots \sum_{j_p=1}^{n} \beta_1^{j_1} \beta_2^{j_2} \ldots \beta_p^{j_p} \begin{vmatrix} \alpha_{j_1}^1 & \ldots & \alpha_{j_p}^1 \\ \vdots & & \vdots \\ \alpha_{j_1}^p & \ldots & \alpha_{j_p}^p \end{vmatrix} .$$

Aquellos términos de la última sumación, que corresponden a valores no distintos entre sí de los índices $j_1, \ldots, j_p$, se anulan pues para

ellos el determinante tiene dos columnas iguales. Puede escribirse entonces

$$det\ \alpha\beta = \sum_J \sum_\sigma \beta_1^{\sigma j_1} \ldots \beta_p^{\sigma j_p} \begin{vmatrix} \alpha^1_{\sigma j_1} & \ldots & \alpha^1_{\sigma j_p} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha^p_{\sigma j_1} & \ldots & \alpha^p_{\sigma j_p} \end{vmatrix},$$

donde la primera suma se extiende a todos los conjuntos $J = \{j_1, \ldots, j_p\}$ tales que $1 \leqslant j_1 < \cdots < j_p \leqslant n$, y la segunda se extiende a todas las permutaciones σ de los números $\{1, \ldots, p\}$. Se sigue de allí que

$$det\ \alpha\beta = \sum_J \sum_\sigma \epsilon_\sigma \beta_1^{\sigma j_1} \ldots \beta_p^{\sigma j_p} \begin{vmatrix} \alpha^1_{j_1} & \ldots & \alpha^1_{j_p} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha^p_{j_1} & \ldots & \alpha^p_{j_p} \end{vmatrix}$$

$$= \sum_J det\ \alpha_J \sum_\sigma \epsilon_\sigma \beta_1^{\sigma j_1} \ldots \beta_p^{\sigma j_p} = \sum_J det\ \alpha_J . det\ \beta^J . \quad \blacksquare$$

7.6. Determinante de un endomorfismo.

Vamos a aplicar ahora los resultados del presente parágrafo a la teoría de los endomorfismos de un espacio vectorial V de n dimensiones.

En la sección 6.4, vimos que cada endomorfismo puede representarse, dada una base de V, por una matriz de orden n. En el teorema 6.11.1 se estableció que las matrices que representan a un endomorfismo son semejantes. El corolario 2 del teorema 7.5.1 permite establecer entonces el

Teorema 7.6.1. Las matrices que representan a un endomorfismo de V respecto de bases cualesquiera tienen todas el mismo determinante.

Definición 7.6.1. Se llama *determinante de un endomorfismo* al determinante de cualquiera de las matrices que lo representan.

Teorema 7.6.2. Los automorfismos de V son los endomorfismos cuyos determinantes son diferentes de cero.

Demostración. Las matrices que representan a los automorfismos son las matrices regulares (sección 6.9). Se aplica entonces el corolario 2 del teorema 7.4.5. ■

Ejercicio

7.6.1. Sea V el espacio vectorial de los polinomios de grado no mayor que 3 sobre $\mathbb{F}$. Definamos en V el endomorfismo T tal que $T(\alpha_o + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3) = \alpha_o + \alpha_1(x + 1) + \alpha_2(x + 1)^2 + \alpha_3(x + 1)^3$. Calcular la matriz de T en las bases:

(a) $1, x, x^2, x^3$.

(b) $1, 1 + x, 1 + x^2, 1 + x^3$.

Si la matriz que resulta en el caso (a) es A, y en el caso (b) es B, hallar la matriz C tal que

$$CAC^{-1} = B.$$

Comprobar que A y B tienen el mismo determinante.

7.7. Desarrollo de un determinante por la fórmula de Laplace.

La presente sección tiene por objeto complementar nuestro estudio de los determinantes con algunas consideraciones adicionales que permitirán establecer la fórmula de Laplace.

Siguiendo una práctica corriente, llamaremos términos, filas, columnas y menores de un determinante

$$\Delta = \begin{vmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & \vdots\ \vdots & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{vmatrix}$$

a los correspondientes conceptos relativos a la matriz

$$\begin{pmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{pmatrix}$$

En consecuencia, recordaremos (definición 7.4.1) que un *menor de orden k* de un determinante Δ de orden n es el determinante que resulta suprimiendo $n-k$ filas y $n-k$ columnas del determinante dado. Para determinar un menor δ basta dar los números de orden $f_1 < f_2 < \cdots < f_k$ que corresponden a las filas que contienen a sus términos y los números de orden $c_1 < c_2 < \cdots < c_k$ que corresponden a las columnas que contienen a dichos términos. El menor δ' que resulta suprimiendo del determinante las filas $f_1, \ldots, f_k$ y las columnas $c_1, \ldots, c_k$ se llama *menor complementario*, o simplemente *complementario de* δ.

Se dice que un menor δ es de *clase par* si la suma de los números de orden de sus filas y columnas

$$\Sigma f + \Sigma c = f_1 + \cdots + f_k + c_1 + \cdots + c_k$$

es par. En caso contrario se dice que δ es de *clase impar*. Para hallar la clase de su complementario δ', formaremos la suma correspondiente.

$$\Sigma f' + \Sigma c' = f_{k+1} + \cdots + f_n + c_{k+1} + \cdots + c_n ,$$

donde $f_{k+1}, \ldots, f_n$ designan a los órdenes de las filas que fueron excluidas para formar el mismo menor. Por tanto

$$\Sigma f + \Sigma f' = 1 + 2 + \cdots + n = N$$
$$\Sigma c + \Sigma c' = 1 + 2 + \cdots + n = N ,$$

de donde

$$\Sigma f + \Sigma c + \Sigma f' + \Sigma c' = 2N$$

y por consiguiente $\Sigma f + \Sigma c$ y $\Sigma f' + \Sigma c'$ tienen la misma pa-

ridad ; es decir que: *dos menores complementarios son de la misma clase.*

Se llama *complemento algebraico* o *cofactor de un menor* δ al menor complementario δ' precedido del signo + o del signo - , según que sea de clase par o de clase impar, respectivamente. En particular, si el menor dado se reduce a un solo elemento, el complemento algebraico coincide con el cofactor de dicho elemento (definición 7.1.1).

Un menor de orden k se llama *principal* cuando está formado por las primeras filas y las k primeras columnas del determinante. Su complemento algebraico coincide con su menor complementario porque es de clase par.

Ejemplo. Para la matriz (α_j^i) de orden 5, el complementario del menor

$$\begin{vmatrix} \alpha_2^2 & \alpha_5^2 \\ \alpha_2^4 & \alpha_5^4 \end{vmatrix}$$

es el menor

$$\begin{vmatrix} \alpha_1^1 & \alpha_3^1 & \alpha_4^1 \\ \alpha_1^3 & \alpha_3^3 & \alpha_4^3 \\ \alpha_1^5 & \alpha_3^5 & \alpha_4^5 \end{vmatrix}$$

Puesto que tanto la suma $(2 + 4) + (2 + 5)$, que corresponde al primero, como la suma $(1 + 3 + 5) + (1 + 3 + 4)$ que corresponde al segundo, son impares, ambos menores son de clase impar. Por tanto, el complemento algebraico de uno cualquiera de ellos es el otro cambiado de signo.

Lema 7.7.1. Los términos del producto del desarrollo de un menor por el desarrollo de su cofactor, son términos del desarrollo del determinante total.

Demostración. 1°. Supongamos primero que el menor δ del determi

nante Δ dado esté formado por las k primeras filas y las k primeras columnas como se ve en el siguiente esquema

$$\Delta = \left|\begin{array}{ccc:ccc} \alpha_1^1 & \alpha_2^1 & \dots & \alpha_k^1 & \alpha_{k+1}^1 & \dots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_k^2 & \alpha_{k+1}^2 & \dots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_1^k & \alpha_2^k & \dots & \alpha_k^k & \alpha_{k+1}^k & \dots & \alpha_n^k \\ \hdashline \alpha_1^{k+1} & \dots & & \alpha_k^{k+1} & \alpha_{k+1}^{k+1} & \dots & \alpha_n^{k+1} \\ \vdots & & \vdots & & \vdots & \vdots & \vdots \\ \alpha_1^n & \dots & & \alpha_k^n & \alpha_{k+1}^n & \dots & \alpha_n^n \end{array}\right|$$

En este caso, δ es de clase par y su cofactor coincide con el menor complementario δ'. Si se supone desarrollado cada uno de esos menores, y se lleva a cabo el producto de esos desarrollos, se obtiene una suma, un término cualquiera de la cual será

$$\epsilon_{\sigma_1} \alpha^1_{\sigma_1(1)}\, \alpha^2_{\sigma_1(2)} \cdots \alpha^k_{\sigma_1(k)} \cdot \epsilon_{\sigma_2} \cdot \alpha^{k+1}_{\sigma_2(k+1)} \cdots \alpha^n_{\sigma_2(n)} \quad ,$$

donde σ_1 es una permutación de los números $(1, \dots, k)$ y σ_2 es una permutación de los números $(k+1, \dots, n)$, y ϵ_{σ_1} y ϵ_{σ_2} son los respectivos signos de esas permutaciones. El signo ϵ_σ de la permutación de los números $(1,2, \dots, n)$ cuya restricción de los k primeros coincide con σ_1 y cuya restricción a los restantes coincide con σ_2 entonces es

$$\epsilon_\sigma = \epsilon_{\sigma_1} \cdot \epsilon_{\sigma_2} \quad ^{1)}$$

Por consiguiente el término a que hemos hecho referencia puede escribirse en la forma

1) Conviene recordar que ϵ_σ es igual a $(-1)^\nu$ donde ν es igual al número de inversiones de σ o al número de trasposiciones en que σ puede descomponerse.

$$\epsilon_\sigma \alpha^1_{\sigma(1)} \cdots \alpha^n_{\sigma(n)} ,$$

y es por tanto uno de los términos del desarrollo de Δ.

2° Si el menor δ no es principal, sino que está formada por las filas de ordenes $f_1 < f_2 < \cdots < f_k$, y por las columnas de orden $c_1 < c_2 < \cdots < c_k$, puede ser convertido en principal por transposiciones sucesivas de las filas y de las columnas, por cada una de las cuales el determinante cambia de signo. Basta permutar la fila f_1 sucesivamente con cada una de las anteriores (que son $f_1 - 1$) hasta que ocupe el primer lugar; la fila f_2 con las $f_2 - 2$ que le preceden hasta llegar al segundo lugar; la fila f_k con las $f_k - k$ que hay desde ella hasta la fila k. Haciendo lo mismo con las columnas, habremos llevado el menor δ a la posición principal, y este caso quedará reducido al anterior.

En el desarrollo del nuevo determinante Δ', el cofactor de δ es el complementario δ' de δ en el determinante dado Δ; luego se tiene que

$$\Delta' = \delta\delta' + \cdots ,$$

y como Δ' se deduce del determinante dado Δ mediante un número de transposiciones igual a la suma

$$(c_1 - 1) + (c_2 - 2) + \cdots + (c_k - k) + (f_1 - 1) + (f_2 - 2) + \cdots + (f_k - k) =$$

$$= \Sigma c + \Sigma f - 2(1 + 2 + \cdots + k) ,$$

será

$$\Delta = (-1)^{\Sigma c + \Sigma f} \Delta' = (-1)^{\Sigma c + \Sigma f} \delta\delta' + \cdots$$

Puestos que $(-1)^{\Sigma c + \Sigma f} \delta'$ es, por definición, el cofactor de δ en el determinante Δ, queda demostrado el lema en el caso general. ■

Teorema 7.7.1. *(Fórmula del desarrollo de Laplace).* Todo determinante es igual a la suma de los productos que se obtienen multiplicando cada uno de los menores de orden k que pertenecen a k filas dadas cualesquiera, por sus correspondientes cofactores.

Demostración. Todos los términos de los productos a que se refiere el enunciado pertenecen al desarrollo de Δ, en virtud del lema anterior; y todos son distintos entre sí porque en su formación intervienen términos distintos del determinante Δ. Solo falta ver que en el desarrollo de Δ no hay más términos que ellos. Basta para esto comprobar que mediante el método descrito en el enunciado se obtienen todos los $n!$ términos del desarrollo del determinante. En efecto, el número de menores de orden k que pueden formarse con k filas es $\binom{n}{k}$; luego el número de términos que se obtiene por el procedimiento del enunciado es

$$\binom{n}{k} k!\,(n-k)! = \frac{n!}{k!\,(n-k)!}\,k!\,(n-k)! = n! \quad ,$$

por cuanto el desarrollo de un menor de orden k tiene $k!$ términos y el de cofactor tiene $(n-k)!$ términos. ■

Observación. Un enunciado enteramente análogo al precedente es válido para las columnas.

Ejemplo. El cálculo del determinante

$$\Delta = \begin{vmatrix} -1 & 0 & 5 & 0 & 4 \\ 3 & 0 & 2 & 0 & 1 \\ -4 & 7 & 1 & -5 & 0 \\ 1/2 & 3 & -1 & 2 & 0 \\ 0 & 0 & -3/2 & 1 & 5 \end{vmatrix}$$

se facilita observando que la submatriz formada por las dos primeras filas tiene dos columnas formadas por ceros. Esta circunstancia facilita el cálculo de Δ mediante el dasarrollo de Laplace respecto de las dos primeras filas. Se tendrá

$$\Delta = - \begin{vmatrix} -1 & 5 \\ 3 & 2 \end{vmatrix} \begin{vmatrix} 7 & -5 & 0 \\ 3 & 2 & 0 \\ 0 & 1 & 5 \end{vmatrix} - \begin{vmatrix} -1 & 4 \\ 3 & 1 \end{vmatrix} \begin{vmatrix} 7 & 1 & 5 \\ 3 & -1 & 2 \\ 0 & -3/2 & 1 \end{vmatrix}$$

$$- \begin{vmatrix} 5 & 4 \\ 2 & 1 \end{vmatrix} \begin{vmatrix} -4 & 7 & -5 \\ 1/2 & 3 & 2 \\ 0 & 0 & 1 \end{vmatrix} = -(-17)145 - (-13)\frac{67}{2} - (-3)(-\frac{31}{2})$$

$$= 2854.$$

Ejercicio

7.7.1. Desarrollar mediante la fórmula de Laplace el determinante de orden $2n$.

$$\begin{vmatrix} a & 0 & 0 & \dots & 0 & 0 & b \\ 0 & a & 0 & \dots & 0 & b & 0 \\ 0 & 0 & a & \dots & b & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & b & \dots & a & 0 & 0 \\ 0 & b & 0 & \dots & 0 & a & 0 \\ b & 0 & 0 & \dots & 0 & 0 & a \end{vmatrix}$$

§ 8. Sistemas de ecuaciones lineales

8.1. Consideraciones y teoremas generales. Un *sistema de ecuaciones lineales* es un conjunto de n ecuaciones

$$\begin{aligned} \alpha_1^1 \xi^1 + \alpha_2^1 \xi^2 + \dots + \alpha_m^1 \xi^m &= \eta^1 \\ \alpha_1^2 \xi^1 + \alpha_2^2 \xi^2 + \dots + \alpha_m^2 \xi^m &= \eta^2 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots\vdots\vdots \qquad\quad \vdots &\qquad\quad \vdots \\ \alpha_1^n \xi^1 + \alpha_2^n \xi^2 + \dots + \alpha_m^n \xi^m &= \eta^n \end{aligned} \qquad (8.1.1)$$

que en forma abreviada podemos escribir

$$\sum_{j=1}^{m} \alpha_j^i \xi^j = \eta^i \,, \quad (i = 1, \dots, n) \qquad (8.1.2)$$

en que los *coeficientes* α_j^i y los *segundos miembros* η^i son escalares dados, es decir elementos dados de $\mathbb{F}$, y se trata de determinar los diversos conjuntos de valores del mismo campo, que puede atribuirse a las *incógnitas* ξ^j $(j = 1, \ldots, m)$ para que el sistema de ecuaciones sea satisfecho. Cada uno de esos conjuntos de valores se llama *solución* del sistema. Si todos los elementos η^i son nulos se dice que el sistema es *homogéneo*. En caso contrario se llama *inhomogéneo*. A continuación vamos a introducir algunas definiciones y notaciones que emplearemos en el resto del parágrafo sin nuevas indicaciones. Esas definiciones y notaciones tienen sus antecedentes en nuestras consideraciones del parágrafo 6, especialmente en la sección 6.5.

U y V designarán a dos espacios vectoriales de dimensión m y n respectivamente. U^* y V^* son sus duales. [1]

$\mathfrak{E} = \{e_1, \ldots, e_m\}$ es una base de U a la cual referiremos todos los vectores de ese espacio. Semejantemente los vectores de V serán referidos a la base $\mathfrak{F} = \{f_1, \ldots, f_n\}$ de dicho espacio.

$$\mathfrak{E}^* = \{e^1, \ldots, e^m\} \quad \text{y} \quad \mathfrak{F}^* = \{f^1, \ldots, f^n\}$$

son las bases duales de $\mathfrak{E}$ y $\mathfrak{F}$.

La matriz

$$A = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_m^1 \\ \vdots & \vdots\ \vdots\ \vdots & \vdots \\ \alpha_1^n & \ldots & \alpha_m^n \end{pmatrix}$$

se llama *matriz del sistema* y está constituido por el cuadro de los coeficientes de las ecuaciones (8.1.1).

Los vectores

$$a^i = \sum_{j=1}^{m} \alpha_j^i e^j \qquad (i = 1, \ldots, n)$$

1) Como ya hemos explicado anteriormente U y V pueden suponerse idénticos a $\mathbb{F}^m$ y $\mathbb{F}^n$, cuyos duales pueden identificarse con ellos mismos.

de U^* y los vectores

$$a_j = \sum_{i=1}^{n} \alpha_j^i f_i \qquad (j = 1, \ldots, m)$$

de V se llaman vectores-fila y vectores-columna del sistema respectivamente, y son los vectores-fila y los vectores-columna de la matriz A.

Asociaremos al sistema (8.1.2) la transformación lineal $T: U \to V$ definida por

$$Te_j = \sum_{i=1}^{n} \alpha_j^i f_i = a_j \quad , \quad (j = 1, \ldots, m)$$

cuya matriz es la matriz del sistema. La transformación dual correspondiente $T^*: V^* \to U^*$ es dada por

$$T^* f^i = \sum_{j=1}^{m} \alpha_j^i e^j = a^i \quad , \quad (i = 1, \ldots, n),$$

y su matriz es la transpuesta A^T de A.

Teorema 8.1.1. La condición necesaria y suficiente para que los números $\xi^1, \ldots, \xi^m$ constituyan una solución del sistema es que los vectores $x = \xi^1 e_1 + \cdots + \xi^m e_m$ *(vector solución)* y el vector $y = \eta^1 f_1 + \cdots + \eta^n f_n$ *(vector segundo miembro* satisfagan a la relación

$$Tx = y \quad , \qquad (8.1.3)$$

es decir que y sea la imagen de x por la transformación T.

Demostración. Se tiene

$$Tx = T.(\sum_{j=1}^{m} \xi^j e_j) = \sum_{j=1}^{m} (Te_j)\,\xi^j = \sum_{j=1}^{m}\sum_{i=1}^{n} \alpha_j^i \xi^j f_i = \sum_{i=1}^{n} (\sum_{j=1}^{m} \alpha_j^i \xi^j) f_i$$

Por consiguiente las n ecuaciones (8.1.1) expresan que las componentes del vector Tx son iguales a las del vector y; es decir que las ecuaciones (8.1.1) son equivalentes a la ecuación vectorial (8.1.3). ∎

Observación. La introducción de la noción de vector-solución del sis-

tema permite hacer uso del lenguaje vectorial. Es así como podemos hablar de combinación lineal de soluciones, de conjunto linealmente independiente de soluciones, etc.

El sistema de ecuaciones

$$\sum_{j=1}^{m} \alpha_j^i \xi^j = 0 \qquad (i = 1, \ldots, n) \qquad (8.1.4)$$

que resulta del sistema (8.1.1) cuando se sustituyen por cero los segundos miembros se llama *sistema homogéneo asociado* al sistema(8.1.1) y puede escribirse, recurriendo a la transformación T, en la forma

$$Tx = 0.$$

y se ve que sus soluciones (más precisamente, sus vectores soluciones) constituyen el núcleo de T.

Teorema 8.1.2. *(Teorema de existencia).* **a)** Las soluciones del sistema homogéneo (8.1.4) constituyen el subespacio $Ker\ T$ de U, de dimensión $m - r(A)$, donde $r(A)$ es el rango de la matriz A. Si $r(A) = m$, entonces esas soluciones se reducen a la solución trivial, o sea que las incógnitas deben ser todas iguales a cero.

b) El sistema (8.1.1) solo admite soluciones si se cumple la condición $y \in Im\ T$. En tal caso, las soluciones constituyen una clase del espacio factor $U/_{Ker\ T}$,el cual es isomorfo con el subespacio $Im\ T$ de V.

c) si $x_1 \in U$ es una solución particular del sistema (8.1.1), todas las demás soluciones del sistema tienen la forma $x_1 + x_0$, donde x_0 es una solución cualquiera del sistema homogéneo asociado, es decir un vector de $Ker\ T$.

Demostración. Ya hemos visto antes del enunciado del teorema que las soluciones del sistema homogéneo (8.1.4) constituyen el subespacio $Ker\ T$ de U.

El isomorfismo $U/_{Ker\ T} \cong Im\ T$ fué establecido en la sección 3.5, en donde se vió que cada clase del espacio factor $U/_{Ker\ T}$ está

formado por todos los vectores de U cuya imagen por T es un mismo vector. Si, en particular $y \in Im\ T$ y $x_1 \in U$ son tales que $Tx_1 = y$, entonces los vectores x que satisfacen a la relación $Tx = y$, pertenecen a la misma clase de $U/Ker\ T$ que x_1, y por tanto serán de la forma

$$x_1 + x_0 \quad , \quad x_0 \in Ker\ T \ .$$

Dichos vectores serán las soluciones del sistema (8.1.1).

Además, en virtud de la ecuación (3.5.2) se tiene

$$dim\ Ker\ T = dim\ U - r(T).$$

Puesto que $dim\ U = m$ y que $r(T) = r(A)$ por el teorema 6.7.1, se tendrá finalmente

$$dim\ Ker\ T = m - r(A) \ . \quad \blacksquare$$

Como una consecuencia del teorema anterior, la resolución de un sistema (8.1.1) se reduce a obtener una solución particular x_1 de dicho sistema y todas las soluciones x_0 del sistema homogéneo asociado. La solución general del sistema (8.1.1) es de la forma $x_1 + x_0$ donde x_0 es cualquiera de las soluciones del sistema homogéneo. Desde luego x_1 existirá únicamente si $y \in Im\ T$, y x_0 se reducirá al vector cero en el caso en que el rango de A sea igual a m, caso en el cual el sistema (8.1.1) tendrá una sola solución. Observemos además que la determinanción del rango $r(A)$ puede hacerse en la práctica con el auxilio de los determinantes (parágrafo § 7.)

Teorema 8.1.3. *(Teorema de unicidad).* Si el sistema (8.1.1) posee solución (es decir si $y \in Im\ T$), entonces es condición necesaria y suficiente para que la solución sea única que la transformación T sea regular, es decir que $r(T) = r(A) = dim\ U = m$; o bien, lo que es equivalente, que los m vectores-columna de la matriz del sistema sean *l.i.*, lo cual implica desde luego que sea $m \leqslant n$, es decir que haya a lo más tantas incógnitas como ecuaciones.

Demostración. Supuesto que $y \in Im\ T$ y que por tanto existe una solución x_1 del sistema (8.1.1), se deduce del teorema 8.1.2 que pa

ra que esa solución sea única es necesario y suficiente que el sistema homogéneo asociado no tenga otra solución que la solución trivial. Para ello es necesario y suficiente que el rango $r(A)$ de la matriz del sistema sea igual al número m de incógnitas. ∎

Observación. Del teorema anterior se sigue que si $m > n$, es decir si el sistema tiene más incógnitas que ecuaciones entonces, o bien el sistema no tiene solución alguna, lo cual ocurre si $y \notin Im\ T$, o bien tiene más de una. En este último caso, como veremos luego, el sistema tiene infinitas soluciones.

La condición $y \in Im\ T$, que es necesaria y suficiente para la existencia de soluciones de las ecuaciones (8.1.1), unida a la relación

$$Im\ T = (Ker\ T^*)^{\perp}$$

que se cumple en virtud de la parte b) del Teorema 5.5.2 nos permite establecer la última proposición de la presente sección en la cual se expresa de manera diferente esa condición de existencia.

Teorema 8.1.4. La condición necesaria y suficiente para que el sistema (8.1.1) tenga solución es que cada solución $(\zeta_1, \ldots, \zeta_n)$ del sistema homogéneo

$$\begin{aligned}
&\alpha_1^1 \zeta_1 + \alpha_1^2 \zeta_2 + \cdots + \alpha_1^n \zeta_n = 0 \\
&\alpha_2^1 \zeta_1 + \alpha_2^2 \zeta_2 + \cdots + \alpha_2^n \zeta_n = 0 \\
&\quad\vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \\
&\alpha_m^1 \zeta_1 + \alpha_m^2 \zeta_2 + \cdots + \alpha_m^n \zeta_n = 0
\end{aligned} \tag{8.1.5}$$

cuya matriz es la transpuesta de la matriz A del sistema (8.1.1), satisfaga a la relación

$$\sum_{i=1}^{n} \eta^i \zeta_i = 0. \tag{8.1.6}$$

Demostración. En virtud del teorema 5.5.2 la condición $y \in Im\ T$ se puede expresar en la forma $y \in (Ker\ T^*)^{\perp}$, es decir que

$$<z^*, y> = 0 \text{ para todo vector } z^* \in Ker\ T^* .$$

Puesto que $y = \sum_{i=1}^{n} \eta^i f_i$ y $z^* = \sum_{j=1}^{n} \zeta_j f^j$, la condición precedente puede expresarse en la forma

$$< \sum_{j=1}^{n} \zeta_j f^j , \sum_{i=1}^{n} \eta^i f_i > = \sum_{j=1}^{n} \sum_{i=1}^{n} \zeta_j \eta^i < f^j , f_i >$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \eta^i \zeta_j \delta_i^j = \sum_{i=1}^{n} \eta^i \zeta_i = 0 ,$$

para toda solución $(\zeta_1 , \ldots , \zeta_n)$ del sistema (8.1.5). ■

El hecho de que, como se observa en la demostración, la condición (8.1.6) pueda escribirse en la forma $< z^* , y > = 0$, que expresa que y debe ser ortogonal a cada solución z^* del sistema transpuesto homogéneo (8.1.5), explica la denominación de *relación de ortogonalidad* que se da a dicha condición (8.1.6).

Observación. Vamos a observar finalmente que las ecuaciones del sistema (8.1.1) pueden escribirse también en la forma

$$< a^i , x > = \eta^i , \quad (i = 1 , \ldots , n), \qquad (8.1.7)$$

que resultan teniendo en cuenta la expresión del producto escalar de dos vectores $a^i \in U^*$ y $x \in U$ en función de sus componentes relativas a las bases duales (e^i) y (e_i). Es claramente evidente que x satisfaga a las ecuaciones (8.1.7) y que sus componentes $(\xi^1 , \ldots , \xi^n)$ satisfagan al sistema (8.1.1). Algunas veces resulta conveniente escribir el sistema (8.1.1) en la forma (8.1.7). Lo haremos así en adelante algunas veces, sin nuevos comentarios.

Ejercicios

8.1.1. Consideremos el sistema siguiente de dos ecuaciones con dos incógnitas

$$\sum_{i=1}^{n} \alpha_i \xi^i = \alpha , \quad \sum_{i=1}^{n} \beta_i \xi^i = \beta .$$

Hallar las soluciones del correspondiente sistema transpuesto homogéneo.

8.1.2. a) Demostrar que el sistema

$$\begin{aligned} x + y + z &= 3 \\ x - y - z &= 4 \\ x + 3y + 3z &= 1 \ , \end{aligned}$$

no posee solución alguna. Encontrar para ello una solución del sistema transpuesto homogéneo que no sea ortogonal al vector de componentes $(3,4,1)$.

b) Caracterizar el núcleo de la transformación lineal asociada al sistema dado.

c) Reemplazar el segundo miembro 1 de la tercera ecuación por un nuevo número de modo que el sistema que resulte tenga solución.

8.2. Criterio del rango de la matriz aumentada.

Si $u = \sum_{j=1}^{m} u^j e_j$ es un vector cualquiera del espacio U, podemos escribir

$$T u = \sum_{j=1}^{m} u^j a_j \ ,$$

Por consiguiente, como ya habíamos visto antes, el subespacio imagen $Im\ T$ es generado por los vectores-columna a_j. Resulta así que $y \in Im\ T$ si y solo si el vector y es combinación lineal de los vectores-columna. Por tanto la condición necesaria y suficiente para que el sistema (8.1.1) tenga solución es que los dos conjuntos de vectores

$$(a_1, \ldots, a_m) \qquad \text{y} \qquad (a_1, \ldots, a_m, y)$$

tengan el mismo número máximo de vectores linealmente independientes, es decir que ambos conjuntos de vectores tengan el mismo rango (ver la definición 7.4.3). Ahora bien, el número máximo de vectores $l.i.$ en el conjunto $(a_1, \ldots, a_m)$ es el rango de la matriz de las componentes de los vectores de dicho conjunto, o sea la

matriz

$$A = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_m^1 \\ \vdots & \vdots\ \vdots\ \vdots & \vdots \\ \alpha_1^n & \dots & \alpha_m^n \end{pmatrix}$$

que es la matriz del sistema (8.1.1); y el número máximo de vectores $l\,i.$ en el conjunto $(a_1, \dots, a_m, y)$ es el rango de la matriz

$$A_a = \begin{pmatrix} \alpha_1^1 & \dots & \alpha_m^1 & \eta^1 \\ \vdots & \vdots\ \vdots\ \vdots & \vdots & \vdots \\ \alpha_1^n & \dots & \alpha_m^n & \eta^n \end{pmatrix}$$

que se llama *matriz aumentada* del sistema (8.1.1), la cual resulta agregando a la matriz A de dicho sistema una columna formada por los segundos miembros de las ecuaciones del sistema. Podemos por tanto enunciar el siguiente teorema.

Teorema 8.2.1. La condición necesaria y suficiente para que el sistema de ecuaciones lineales (8.1.1) tenga solución es que la matriz aumentada tenga el mismo rango que la matriz del sistema.

Observación 1. La condición del teorema anterior se cumple evidentemente para todo sistema homogéneo. Efectivamente, como ya hemos observado antes, dichos sistemas tienen por lo menos la solución trivial.

2. Un sistema inhomogéneo, cuya matriz es de rango igual al número de ecuaciones, tiene soluciones, por cuanto su matriz aumentada tiene evidentemente el mismo rango, y por tanto se cumple la condición del teorema 8.2.1. En particular, un sistema inhomogéneo con el mismo número de incognitas que ecuaciones, y cuyo determinante es diferente de cero, tiene una solución que puede ser calculada con la fórmula de Cramer (Teorema 7.3.1). Un sistema que cumple esas condiciones se llama *sistema de Cramer.*

8.3. Resolución de los sistemas de ecuaciones linealmente independientes. Decimos que un sistema de r ecuaciones con m incógnitas

$$\sum_{j=1}^{m} \alpha_j^i \xi^j = \eta^i \quad , \quad (i = 1, \ldots, r), \qquad (8.3.1)$$

es un *sistema de ecuaciones linealmente independientes*, o más brevemente, un *sistema l.i.*, si los vectores fila $a^1, a^2, \ldots, a^r$ son *l.i.* Para ello es necesario y suficiente (ver la definición 7.4.3) que sea $r(A) = r$, es decir que el rango de la matriz A de los coeficientes sea igual al número de ecuaciones. Como se anotó en la observación 2, al final de la sección precedente, los sistemas que estamos considerando tienen siempre soluciones. Nuestro propósito aquí es obtener la expresión explícita de esas soluciones.

Notemos ante todo que se debe tener $r \leqslant m$, es decir que en un sistema *l.i.* el número de ecuaciones es no mayor que el de incógnitas. Según sabemos el número máximo de vectores-columna *l. i.* es también r, y podemos suponer, reordenando si fuera necesario las incógnitas, que los r primeros vectores-columna $a_1, \ldots, a_r$ son *l.i.*, y por tanto que la matriz

$$\begin{pmatrix} \alpha_1^1 & \ldots & \alpha_r^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_1^r & \ldots & \alpha_r^r \end{pmatrix}$$

es un menor de base de la matriz A y que, por consiguiente, su determinante, que designaremos por D, es diferente de cero.

En primer lugar supondremos que es $r < m$. El sistema (8.3.1) puede entonces ser escrito en la forma

$$\sum_{j=1}^{r} \alpha_j^i \xi^j = \theta^i \quad , \quad (i = 1, \ldots, r), \qquad (8.3.2)$$

donde

$$\theta^i = \eta^i - \sum_{j=r+1}^{m} \alpha_j^i \xi^j \quad , \quad (i = 1, \ldots, r) \qquad (8.3.3)$$

Para cada sistema arbitrario de valores escalares que se atribuya a $\xi^{r+1}, \ldots, \xi^m$, resultará un sistema de valores $\theta^1, \ldots, \theta^r$; y se obtendrá un correspondiente sistema de la forma (8.3.2), de r ecuaciones con r incógnitas $\xi^1, \ldots, \xi^r$, cuyo determinante es $D \neq 0$. Cada uno de esos sistemas es por tanto un sistema de Cramer y puede ser resuelto mediante las correspondientes fórmulas (Teorema 7.3.1). Esas soluciones serán por consiguiente

$$\xi^j = \frac{D^j}{D} \quad , \quad (j = 1, \ldots, r) \quad , \qquad (8.3.4)$$

donde D^j es el determinante que resulta de D cuando se sustituye la columna de orden j de este último determinante por la columna formada por los segundos miembros $\theta^1, \ldots, \theta^r$ del sistema (8.3.2) que se estudia. Se puede observar de inmediato que los valores $\xi^1, \ldots, \xi^r$ dados por las fórmulas (8.3.4), completados con los valores $\xi^{r+1}, \ldots, \xi^m$, que fueron arbitrariamente elegidos, y con los cuales se calcularon los segundos miembros $\theta^1, \ldots, \theta^r$, constituyen una solución del sistema (8.3.1). El procedimiento de cálculo empleado prueba al mismo tiempo que, mediante él, se pueden obtener todas las soluciones del sistema (8.3.1). Las soluciones del sistema (8.3.1) son pues de la forma

$$\left(\frac{D^1}{D}, \frac{D^2}{D}, \ldots, \frac{D^r}{D}, \xi^{r+1}, \ldots, \xi^m\right)$$

donde los escalares $\xi^{r+1}, \ldots, \xi^m$ son arbitrarios, y los determinantes D^j $(j = 1 \ldots r)$ depende de ellos; por intermedio de las expresiones (8.3.3). Para expresar D^j, teniendo presente los valores (8.3.3), vamos a introducir la notación D_o^j para representar al determinante que resulta de D cuando se sustituye la columna de orden j por la columna formada por $\eta^1, \ldots, \eta^r$; y la notación D_s^j, donde $r + 1 \leqslant s \leqslant m$, para representar al determinante que resulta de D cuando se sustituye la columna de orden j por la columna formada por los coeficientes $\alpha_s^1, \ldots, \alpha_s^r$. Se obtiene entonces, haciendo uso de las conocidas propiedades de los determinantes:

$$D^j = D_o^j - \sum_{s=r+1}^{m} D_o^j \, \xi^s \quad , \quad (j = 1, \ldots, r).$$

Por consiguiente las soluciones (8.3.4) del sistema (8.3.2) pueden escribirse en la forma

$$\xi^j = \frac{D_o^j}{D} + \sum_{s=r+1}^{m} \left(-\frac{D_s^j}{D}\right) \xi^s \quad , \quad (j = 1, \ldots, r)$$

Las soluciones del sistema (8.3.1) serán entonces

$$\left[\frac{D_o^1}{D} + \sum_{s=r+1}^{m} \left(-\frac{D_s^1}{D}\right) \xi^s \, , \, \frac{D_o^2}{D} + \sum_{s=r+1}^{m} \left(-\frac{D_s^2}{D}\right) \xi^s, \ldots, \frac{D_o^r}{D} + \right.$$

$$\left. \sum_{s=r+1}^{m} \left(-\frac{D_s^r}{D}\right) \xi^s \, , \, \xi^{r+1} \, , \ldots, \, \xi^m \right] ,$$

donde los $\xi^{r+1}, \ldots, \xi^m$ son escalares arbitrarios. En particular, si se supone $\eta^1 = \eta^2 = \ldots = \eta^r = 0$, el sistema (8.3.1) se reduce a su sistema homogéneo asociado, y sus soluciones son

$$\left[\sum_{s=r+1}^{m} \left(-\frac{D_s^1}{D}\right) \xi^s \, , \, \sum_{s=r+1}^{m} \left(-\frac{D_s^2}{D}\right) \xi^s \, , \ldots, \, \sum_{s=r+1}^{m} \left(-\frac{D_s^r}{D}\right) \xi^s, \xi^{r+1}, \ldots, \xi^m\right].$$

Si por otra parte, se supone $\xi^{r+1} = \xi^{r+2} = \ldots = \xi^m = 0$, se obtiene la solución particular del sistema (8.3.1)

$$\left[\frac{D_o^1}{D} \, , \, \frac{D_o^2}{D} \, , \ldots, \, \frac{D_o^r}{D} \, , \underbrace{0, 0, \ldots, 0}_{m-r}\right]$$

Se comprueba así que la solución del sistema (8.3.1) es la suma de una solución particular del mismo y de la solución general del sistema homogéneo asociado (Teorema 8.1.2).

Observaremos finalmente que las soluciones del sistema homogéneo asociado son las combinaciones lineales con coeficientes

$$\xi^{r+1} \, , \, \xi^{r+2} \, , \ldots, \, \xi^m$$

de las siguientes $m - r$ soluciones linealmente independientes, del mis

mo sistema:

$$\left[-\frac{D^1_{r+1}}{D}, -\frac{D^2_{r+1}}{D}, \ldots, -\frac{D^r_{r+1}}{D}, \underbrace{1, 0, \ldots, 0}_{m-r}\right]$$

$$\left[-\frac{D^1_{r+2}}{D}, -\frac{D^2_{r+2}}{D}, \ldots, -\frac{D^r_{r+2}}{D}, 0, 1, 0, \ldots, 0\right]$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\left[-\frac{D^1_m}{D}, -\frac{D^2_m}{D}, \ldots, -\frac{D^r_m}{D}, 0, 0, 0, \ldots, 1\right],$$

lo cual concuerda con lo establecido en la parte a) del teorema 8.1.2.

Hasta ahora hemos supuesto que $r < m$. Si suponemos ahora que $r = m$, entonces el sistema (8.3.1) se reduce a un sistema de Cramer que tiene una sola solución la cual se obtiene mediante las fórmulas de Cramer y es

$$\left[\frac{D^1_o}{D}, \frac{D^2_o}{D}, \ldots, \frac{D^r_o}{D}\right],$$

o bien, simplemente

$$\left[\frac{D^1}{D}, \frac{D^2}{D}, \ldots, \frac{D^r}{D}\right],$$

pues en ese caso es $D^j = D^j_o$ para $j = 1, \ldots, r$.

8.4. Resolución de sistemas en el caso más general. Condiciones de compatibilidad.

Ya hemos observado precedentemente que el sistema general de n ecuaciones lineales con m incógnitas que en el presente estudio vamos a escribir en la forma (ver sección 8.1)

$$<a^i\ ,\ x> = \eta^i\ ,\ (i = 1, \ldots, n)\ , \qquad (8.4.1)$$

no siempre tiene solución. Para que posea soluciones deben cumplirse condiciones que hemos expresado ya de diferentes maneras (teorema 8.1.1, teorema 8.1.4 y teorema 8.2.1). En la presente sección, al mismo tiempo que indicaremos la manera de resolver un sistema (8.4.1) que tiene soluciones, deduciremos condiciones necesarias y suficientes para que tales soluciones existan o, como también se dice, para que las ecuaciones del sistema sean *compatibles*. Las condiciones que vamos a deducir se llaman por eso *condiciones de compatibilidad*.

Supongamos que r es el rango de la matriz A del sistema (8.4.1). Entre los vectores-fila $a^1, \ldots, a^n$ habrá un conjunto maximal de vectores *l.i.*, que podemos suponer que son $a^1, \ldots, a^r$. Todos los demás vectores-fila dependen linealmente de ellos, de modo que se cumplen relaciones de la forma

$$a^k = \sum_{i=1}^{r} \rho_i^k\, a^i\ , \quad (k = r+1, \ldots, n) \qquad (8.4.2)$$

donde los ρ_i^k son escalares determinados.

Las r primeras ecuaciones del sistema (8.4.1),

$$<a^i\ ,\ x> = \eta^i\ ,\ (i = 1, \ldots, r)\ , \qquad (8.4.3)$$

constituyen el sistema que llamaremos *reducido* de (8.4.1).

Las demás ecuaciones

$$<a^k\ ,\ x> = \eta^k\ , \quad (k = r+1, \ldots, n)\ , \qquad (8.4.4)$$

constituyen el sistema *complementario*.

El sistema reducido, por cuanto los vectores $a^1, \ldots, a^r$ son *l.i.*, es un sistema *l.i.* de la forma (8.3.1) al que son aplicables las consideraciones de la sección anterior. Es por tanto resoluble y todas sus soluciones pueden obtenerse de la manera que allí se explicó. Las soluciones del sistema (8.4.1) son las que satisfacen tanto al sistema re

ducido como al complementario.

Sea x una solución cualquiera del sistema reducido (8.4.3). Se tendrá

$$<a^k\ ,\ x> = <\sum_{i=1}^{r} \rho_i^k a^i\ ,\ x> = \sum_{i=1}^{r} \rho_i^k <a^i\ ,\ x> = \sum_{i=1}^{r} \rho_i^k \eta^i\ ,$$

$(k = r+1, \ldots, n)$.

Luego, las condiciones necesarias y suficientes para que x sea también solución del sistema complementario (8.4.4), y por consiguiente para que x sea solución del sistema propuesto (8.4.1), son

$$\sum_{i=1}^{r} \rho_i^k \eta^i = \eta^k \qquad (k = r+1, \ldots, n)$$

El resultado a que hemos llegado puede expresarse de la manera siguiente:

Teorema 8.4.1. Las condiciones necesaria y suficientes para que un sistema lineal de n ecuaciones con m incógnitas

$$<a^i\ ,\ x> = \eta^i \quad , \quad (i = 1, \ldots, n)$$

tenga solución es que si $a^1, \ldots, a^r$ son vectores-fila *l.i.* de la matriz A del sistema, las ecuaciones

$$a^k = \sum_{i=1}^{r} \rho_i^k a^i \quad , \quad (k = r+1, \ldots, n)\ ,$$

impliquen a las condiciones de compatibilidad

$$\eta^k = \sum_{i=1}^{r} \rho_i^k \eta^i \quad , \quad (k = r+1, \ldots, n)\ ,$$

que deben ser satisfechas por los segundos miembros.

Además, supuesto que la condición precedente se cumple, las soluciones del sistema propuesto son justamente las soluciones del sistema reducido

$$\sum_{j=1}^{m} \alpha_j^i \xi^j = \eta^i \quad , \quad (i = 1, \ldots, r)\ ,$$

que es linealmente independiente y por tanto puede ser resuelto por

el método descrito en la sección 8.3.

El método de cálculo de las soluciones de un sistema cualquiera puede resumirse en la siguiente

Regla. 1°. Se calcula el rango r de la matriz del sistema. La dimensión del subespacio de las soluciones del sistema homogéneo asociado será $m - r$.

2°. Se analiza la compatibilidad del sistema empleando de preferencia las condiciones de compatibilidad o bien el criterio de la matriz aumentada. Si el sistema es compatible, se procede a resolver el sistema reducido, reordenando previamente las incógnitas de manera que el determinante de los coeficientes de las r primeras sea diferente de cero.

3°. Se calcula la solución general del sistema homogéneo asociado al sistema reducido. Esa solución general se expresa en términos de las últimas $m - r$ incógnitas, a las cuales puede atribuirse valores arbitrarios.

4°. Se calcula una solución particular del sistema reducido, atribuyendo el valor 0 a las $m - r$ últimas incógnitas y haciendo uso de las fórmulas de Cramer.

5°. La solución general del sistema propuesto será la suma de las soluciones que fueron obtenidas en los pasos 3° y 4° precedentes.

El método de cálculo descrito es suceptible de muchas variantes y simplificaciones que son objeto de un capítulo especial del Análisis Numérico. Cabe señalar en particular el método de Gauss. [1)]

Con el objeto de ilustrar el método vamos a exponer un ejemplo, en detalle.

Ejemplo. Consideremos el sistema inhomogéneo de $n = 5$ ecuaciones y $m = 4$ incógnitas

$$\ell_1 = x_1 + 2x_2 + 3x_3 - x_4 = 1$$
$$\ell_2 = 3x_1 + 2x_2 + x_3 - x_4 = 1$$

1) Ver por ejemplo: F.B. Hildebrand, Introduction to Numerical Analysis, McGraw-Hill Book Co. Inc. (1956), pág 428.

$$\ell_3 = 2x_1 + 3x_2 + x_3 + x_4 = 1$$
$$\ell_4 = 2x_1 + 2x_2 + 2x_3 - x_4 = 1$$
$$\ell_5 = 5x_1 + 5x_2 + 2x_3 = 2.$$

Es fácil observar, o bien deducir calculando, que entre los primeros miembros de estas ecuaciones se cumplen las relaciones

$$\ell_4 = \frac{1}{2}\ell_1 + \frac{1}{2}\ell_2$$
$$\ell_5 = \ell_2 + \ell_3 \qquad (8.4.5)$$

Se deduce de aquí que las dos últimas filas de la matriz del sistema son combinaciones lineales de las tres primeras, las cuales son *l. i.* por que

$$\begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 3 & 1 \end{vmatrix} = 12 \neq 0 .$$

Por consiguiente, el rango de la matriz del sistema es $r = 3$, y la dimensión del espacio de las soluciones del sistema homogéneo asociado es $4 - 3 = 1$.

Las relaciones (8.4.5) equivalen a las siguientes relaciones entre los vectores-fila:

$$a^4 = \frac{1}{2}a^1 + \frac{1}{2}a^2$$
$$a^5 = a^2 + a^3 .$$

Las condiciones de compatibilidad se cumplen, pues

$$\frac{1}{2}\eta^1 + \frac{1}{2}\eta^2 = \frac{1}{2} + \frac{1}{2} = 1 = \eta^4$$

y

$$\eta^2 + \eta^3 = 1 + 1 = 2 = \eta^5 .$$

El sistema es por tanto compatible. El sistema reducido es

$$\begin{aligned} x_1 + 2x_2 + 3x_3 - x_4 &= 1 \\ 3x_1 + 2x_2 + x_3 - x_4 &= 1 \\ 2x_1 + 2x_2 + 2x_3 + x_4 &= 1 \,. \end{aligned} \qquad (8.4.6)$$

El sistema homogéneo asociado

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= x_4 \\ 3x_1 + 2x_2 + x_3 &= x_4 \\ 2x_1 + 2x_2 + 2x_3 &= -x_4 \end{aligned}$$

admite la solución general

$$\left(\frac{5}{6} x_4 \,,\, -\frac{7}{6} x_4 \,,\, \frac{5}{6} x_4 \,,\, x_4 \right) .$$

Haciendo $x_4 = 0$ en el sistema reducido resulta un sistema de Cramer cuya solución es $(1/6 \,,\, 1/6 \,,\, 1/6)$. Se obtiene así la solución $(1/6 \,,\, 1/6 \,,\, 1/6 \,,\, 0)$ del sistema propuesto. La solución genaral será entonces

$$\left(\frac{1 + 5x_4}{6} \,,\, \frac{1 - 7x_4}{6} \,,\, \frac{1 + 5x_4}{6} \,,\, x_4 \right)$$

donde x_4 es un número arbitrario.

Ejercicios

8.4.1. Calcular el rango de la matriz

$$\begin{pmatrix} 75 & 0 & 116 & -39 & 0 \\ 171 & -69 & 402 & 123 & 45 \\ 301 & 0 & 87 & -417 & -169 \\ 114 & -46 & 268 & 82 & 30 \end{pmatrix}$$

8.4.2. Calcular el rango de la matriz

$$\begin{pmatrix} 2 & -1 & 1 & 3 & 4 \\ 2 & -1 & 2 & 1 & -2 \\ 2 & -3 & 1 & 2 & -2 \\ 1 & 0 & 1 & -2 & -6 \\ 1 & 2 & 1 & -1 & 0 \\ 4 & -1 & 3 & -1 & -8 \end{pmatrix}$$

8.4.3. ¿Para qué valores de λ tiene solución el sistema

$$\begin{aligned} 2x_1 - x_2 + x_3 + x_4 &= 1 \\ x_1 + 2x_2 - x_3 + 4x_4 &= 2 \\ x_1 + 7x_2 - 4x_3 + 11x_4 &= \lambda. \quad ? \end{aligned}$$

Emplear sucesivamente las condiciones de compatibilidad (teorema 8.4.1), el criterio de la matriz aumentada (teorema 8.2.1) y las relaciones de ortogonalidad (teorema 8.1.4).

8.4.4. Resolver el sistema

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

8.4.5. Resolver el sistema

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 2 \\ x_1 + 3x_2 + x_3 &= 5 \\ x_1 + x_2 + 5x_3 &= -7 \\ 2x_1 + 3x_2 - 3x_3 &= 14 \end{aligned}$$

8.4.6. Resolver el sistema

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

$$4x_1 - x_2 + x_3 = 3$$
$$x_1 + 3x_2 - 13x_3 = -6.$$

8.4.7. Resolver el sistema

$$x_1 + 3x_2 + 2x_3 = 0$$
$$2x_1 - x_2 + 3x_3 = 0$$
$$3x_1 - 5x_2 + 4x_3 = 0$$
$$x_1 + 17x_2 + 4x_3 = 0 .$$

8.4.8. Resolver el sistema

$$x_1 + 3x_2 + 5x_3 - 4x_4 = 1$$
$$x_1 + 3x_2 + 2x_3 - 2x_4 + x_5 = -1$$
$$x_1 - 2x_2 + x_3 - x_4 - x_5 = 3$$
$$x_1 - 4x_2 + x_3 + x_4 - x_5 = 3$$
$$x_1 + 2x_2 + x_3 - x_4 + x_5 = 1$$

8.4.9. Estudiar el sistema siguiente para todos los valores reales de a:

$$(3a - 1)x + 2ay + (3a + 1)z = 1$$
$$2ax + 2ay + (3a + 1)z = a$$
$$(a + 1)x + (a+ 1)y + 2(a + 1)z = a^2$$

8.4.10. Averiguar si los vectores

$$(1 , -2 , 1 , 0 , 0)$$
$$(1 , -2 , 0 , 1 , 0)$$
$$(0 , 0 , 1 , -1 , 0)$$
$$(1 , -2 , 3 , -2 , 0)$$

constituyen un *sistema fundamental de soluciones* (es decir una base del espacio de las soluciones) del sistema

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 0 \\ 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 &= 0 \\ x_2 + 2x_3 + 2x_4 - 6x_5 &= 0 \\ 5x_1 + 4x_2 + 3x_3 + 3x_4 - x_5 &= 0 \end{aligned}$$

3.4.11. Hallar un sistema fundamental de soluciones del sistema del ejercicio anterior.

3.4.12. ¿Es

$$(1\ ,\ -2\ ,\ 1\ ,\ 0\ ,\ 0)$$
$$(0\ ,\ 0\ ,\ -1\ ,\ 1\ ,\ 0)$$
$$(4\ ,\ 0\ ,\ 0\ ,\ -6\ ,\ 2)$$

un sistema fundamental del sistema de ecuaciones lineales del ejercicio 8.4.10 ?

CAPITULO III

POLINOMIOS SOBRE CAMPOS. EXTENSIONES DE CAMPOS Y DE ESPACIOS VECTORIALES

En el primer parágrafo del presente capítulo vamos a tratar acerca de la noción de anillos y de las propiedades de estos que utilizaremos a continuación.

§ 9. Anillos

9.1. Anillos.

Definición 9.1.1. Un conjunto A se llama *anillo* si es un grupo aditivo, es decir si es dada en él una operación de *adición* que es asociativa y conmutativa y cuyo elemento neutro, *cero*, se designa por 0; y existe además una aplicación

$$(x,y) \in A \times A \to xy \in A,$$

que se llama *multiplicación* que tiene las propiedades siguientes:

1. $(xy)z = x(yz)$ *(asociatividad)*
2. $(x + y)z = xz + yz$, $z(x + y) = zx + zy$ *(distributividad)*

Si A contiene a un elemento 1 *(identidad o unielemento)* tal que $x1 = 1x = x$ para todo $x \in A$, se dice que es un *anillo con identidad* o con unielemento. La identidad, cuando existe, es única.

Un anillo se dice que es conmutativo si tiene lugar la propiedad

3. $xy = yx$ para todos $x,y \in A$.

Ejemplos. Las clases de restos de enteros módulo m (m entero positivo cualquiera) constituyen un anillo conmutativo con identidad. Si m no es primo este anillo no es un campo. Así en el anillo de clases de restos módulo 6 la clase (3) no tiene inversa.

Los campos son anillos conmutativos con identidad en que cada elemento diferente de cero tiene una inversa.

9.2. Ideales.

Definición 9.2.1. Sea A un anillo. Un subconjunto no vacío I de A se llama *ideal izquierdo* de A si tiene las propiedades siguientes:

1. $a, b \in I \Rightarrow a + b \in I$ y $a - b \in I$

2. $a \in I$, $x \in A \Rightarrow xa \in I$

Si en esta definición se sustituye la propiedad 2 por

2'. $a \in I$, $x \in A \Rightarrow ax \in I$,

se dice que I es un ideal derecho de A.

Si I satisface a 1., 2. y 2'. se llama *ideal por ambos lados*, o simplemente *ideal*.

En los anillos conmutativos no cabe la distinción entre ideales derechos e izquierdos.

Conviene observar que todo ideal contiene al cero. El ideal $\{0\}$ se llama *ideal cero o ideal nulo*.

Ejemplo. El conjunto $\mathbb{Z}$ de los números enteros es un anillo conmutativo con unidad. Si a y b son dos enteros fijos cualesquiera, los números de la forma

$$xa + yb$$

donde x e y son enteros arbitrarios constituyen un ideal que se llama *ideal generado por a y b*. En particular, los múltiplos de un entero constituyen un ideal. El concepto de ideal generado puede extenderse a más de dos números. También puede ser aplicado a un anillo cualquiera.

Teorema 9.2.1. Si $I_1, \ldots, I_n$ son n ideales por la izquierda (o por

la derecha, o por amb lados) de un anillo A, entonces el conjunto

$$I_1 + \ldots + I_n = \{a_1 + \ldots + a_n \ / \ a_i \in I_i \ , \ i = 1, \ldots, n\}$$

es un ideal del mismo tipo, que se llama suma de los ideales dados

Así mismo $I_1 \cap I_2 \cap \ldots \cap I_n$ es un ideal del mismo tipo.

Demostración. Se propone como ejercicio.

Definición 9.2.2. Sea A un anillo conmutativo con identidad. Dado un elemento cualquiera $a \in A$, el conjunto

$$I(a) = \{xa \ / \ x \in A\}$$

es un ideal que se llama *ideal principal* generado por a, que es evidentemente el ideal minimal que contiene al elemento a. En particular el ideal $I(0) = \{0\}$ es el ideal cero.

Ejemplo. Vamos a demostrar que en el anillo $\mathbb{Z}$ de los enteros cada ideal I es principal.

Si $I = \{0\}$ entonces se tiene evidentemente $I = I(0)$.

Si $I \neq \{0\}$, existe en I un entero positivo mínimo d, y es claro que $I(d) \subset I$. Para demostrar que $I = I(d)$ basta con probar que $I \subset I(d)$.

Supongamos primero que $m \in I$ y que $m > 0$. Puede escribirse entonces $m = dq + r$, donde $q \geqslant 0$ y r, $(0 \leqslant r < d)$ son enteros no negativos bien determinados. Se tiene que

$$r = m - dq \in I \ , \ 0 \leqslant r < d.$$

Si fuera $0 < r < d$ resultaría una contradicción porque d es el menor entero positivo contenido en I. Luego debe ser $r = 0$ y por tanto $m = dq \in I(d)$.

Supongamos ahora que $m \in I$ y que $m < 0$. Puesto que $-m > 0$, por lo antes demostrado se tendrá $-m \in I(d)$. Se deduce de aquí que $m \in I(d)$. En todo caso pues $m \in I$ implica que $m \in I(d)$, y por tanto $I \subset I(d)$.

Teorema 9.2.2. Un anillo conmutativo con identidad, A, es un campo si y solo si no tiene más ideales que él mismo y el ideal cero.

Demostración. Supongamos que A es un campo y que $I \neq \{0\}$ es uno de sus ideales. En I existirá un elemento $a \neq 0$. En A existirá la inversa a^{-1} de a y por la definición de ideal se tendrá $aa^{-1} = 1 \in I$. Por la misma razón, cualquiera que sea $x \in A$ se tendrá entonces $x = 1.x \in I$, y por tanto $A = I$.

Supongamos ahora que A es un anillo conmutativo con identidad que no tiene otros ideales que A y $\{0\}$. Para demostrar que A es un campo basta con probar que, cualquiera que sea el elemento $a \neq 0$ de A, existe en A el elemento a' tal que $aa' = 1$. Con ese fin consideremos el ideal principal $I(a)$ constituido por los múltiplos de a. En particular, ese ideal contiene al elemento a, pues $a = a.1$. Por hipótesis, y puesto que $I(a) \neq \{0\}$ (contiene a a !), se tendrá que $I(a) = A$. Se ve de esta manera que todo elemento de A es múltiplo de a. En particular la identidad será un múltiplo de a, y por tanto existe $a' \in A$ tal que $aa' = 1$. ■

9.3 Homomorfismos de anillos.

Definición 9.3.1. Una aplicación φ de un anillo A en otro anillo B se llama *homomorfismo* de A en B si se cumplen las siguientes condiciones cualesquiera que sean los elementos a y b de A:

1. $\varphi(a + b) = \varphi(a) + \varphi(b)$
2. $\varphi(ab) = \varphi(a)\varphi(b)$.

El conjunto de los elementos $a \in A$ tales que $\varphi(a) = 0$ se llama *núcleo* de φ.

Si φ es una aplicación biyectiva, entonces se dice que es un *isomorfismo* de A sobre B y que A y B son isomorfos.

Teorema 9.3.1. Sea φ un homomorfismo de un anillo A en un anillo B. Se cumple entonces:

1. $\varphi(0) = 0$
2. $\varphi(-a) = -\varphi(a)$, para cada $a \in A$
3. El núcleo de φ es un ideal (por los dos lados) de A, es decir que los elemento de A tales que $\varphi(a) = 0$ constituyen un ideal de A

4. φ es inyectiva si y solo si $\varphi(a) = 0$ implica que $a = 0$, es decir si el núcleo de A es $\{0\}$.

Demostración.

1. $\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0) \Rightarrow \varphi(0) = 0$.

2. $\varphi(a) + \varphi(-a) = \varphi(a + (-a)) = \varphi(0) = 0 \Rightarrow \varphi(-a) = -\varphi(a)$.

3. Sea I el núcleo de φ. Demostraremos que es un ideal:
   Si $a, b \in I$, entonces
   $$\varphi(a + b) = \varphi(a) + \varphi(b) = 0 \Rightarrow a + b \in I$$
   y
   $$\varphi(a - b) = \varphi(a + (-b)) = \varphi(a) - \varphi(b) = 0 \Rightarrow a - b \in I.$$
   Si $a \in I$ y $x \in A$, entonces
   $$\varphi(xa) = \varphi(x)\varphi(a) = 0 \Rightarrow xa \in I$$
   $$\varphi(ax) = \varphi(a)\varphi(x) = 0 \Rightarrow ax \in I.$$

4. Supongamos que φ es inyectiva. Puesto que $\varphi(0) = 0$, el núcleo de φ se reduce a $\{0\}$.
   Si suponemos ahora que el núcleo de φ es $\{0\}$, entonces, si $\varphi(a) = \varphi(b)$, se sigue que
   $$\varphi(b - a) = \varphi(b) - \varphi(a) = 0$$
   y por tanto $b - a$ pertenece al núcleo de φ, es decir que
   $$b - a \in \{0\}, \text{ o sea } b - a = 0.$$
   Resulta así que $a = b$ y por consiguiente φ es inyectiva. ■

9.4. Anillos cocientes.

Definición 9.4.1. Sea A un anillo conmutativo con identidad. Dado un ideal I de A, diremos que dos elementos a y b de A son *congruentes módulo* I (o bien que a es congruente con b módulo I) si $a - b \in I$. Escribiremos en ese caso

$$a = b \ mod \ I, \quad \text{o simplemente} \quad a \equiv b$$

Es fácil observar que la relación así definida es reflexiva, simé-

trica y transitiva, y por tanto es una relación de equivalencia.

El conjunto de las clases de equivalencia relativas a la relación que acabamos de definir se representa por $A/_I$, y constituye un anillo conmutativo con identidad cuando se definen en él las operaciones de adición y de multiplicación, de la manera que se expone a continuación.

Designemos por π a la aplicación de A sobre $A/_I$, que aplica a cada elemento $a \in A$ en la clase $\pi(a) \in A/_I$ que lo contiene. Resulta así que

$$\pi(a) = \pi(a_1) \Leftrightarrow a - a_1 \in I .$$

La adición en $A/_I$ puede entonces ser definida así:

$$\pi(a) + \pi(b) = \pi(a + b) ,$$

donde hemos tenido en cuenta que cada clase de $A/_I$ puede ser representada en la forma $\pi(a)$ donde a es un elemento de A que pertenece a ella. Esta definición es consistente porque si $\pi(a)$ y $\pi(b)$ se expresan en la forma $\pi(a_1)$ y $\pi(b_1)$ respectivamente, entonces se tendría $a - a_1 \in I$, $b - b_1 \in I$, y por tanto

$$(a + b) - (a_1 + b_1) = (a - a_1) + (b - b_1) \in I ,$$

es decir que

$$\pi(a + b) = \pi(a_1 + b_1) ,$$

y por tanto la suma de clases, que hemos definido, es independiente de los representantes que se elijan en las clases.

Semejantemente el producto de dos clases $\pi(a)$ y $\pi(b)$ de $A/_I$ se define por

$$\pi(a)\,\pi(b) = \pi(a\,b) .$$

También ahora debemos comprobar que esta definición es independiente de los representantes a y b elegidos en las clases consideradas. En efecto, si $\pi(a) = \pi(a_1)$ y $\pi(b) = \pi(b_1)$, entonces es $a - a_1 \in I$ y $b - b_1 \in I$, y por tanto

$$ab - a_1 b_1 = (a - a_1)b + a_1(b - b_1) \in I ,$$

y por consiguiente

$$\pi(ab) = \pi(a_1 b_1).$$

La comprobación detallada de que, con las operaciones definidas como hemos visto, $A/_I$ es un anillo conmutativo con elemento identidad es un ejercicio que debe realizar el lector. Se probará que el elemento nulo de $A/_I$ es una clase que es idéntica a I, y que π es un homomorfismo de A sobre $A/_I$, cuyo núcleo es el ideal I.

Definición 9.4.2. El anillo $A/_I$ se llama *anillo cociente* de A respecto del ideal I. π se llama *proyección* de A sobre $A/_I$. Se dice también que π es el *homomorfismo canónico o natural* de A sobre $A/_I$.

Ejemplo. El anillo $\mathbb{Z}/_{I(m)}$ donde m es un entero cualquiera es el anillo de las clases de restos módulo m.

Definición 9.4.3. Un ideal I del anillo A, que es diferente de A, se dice que es *maximal* si no existen en A otros ideales que contengan a I que I mismo y A. (Según el teorema 9.2.2, el ideal nulo es maximal si y solo si A es un campo).

Teorema 9.4.1. Sea I un ideal del anillo A. Entre el conjunto de los ideales de A que contienen a I y el conjunto de los ideales de $A/_I$, existe una biyección que a cada ideal J de A, que contiene a I, le hace corresponder el ideal $\pi(J)$ de $A/_I$, formado por los elementos $\pi(a)$, $a \in A$, de $A/_I$; y que a cada ideal $\overline{J}$ de $A/_I$ le hace corresponder el ideal $\pi^{-1}(\overline{J})$ de A formado por los elementos $a \in A$ tales que $\pi(a) \in \overline{J}$. En particular se tiene que

$$\pi(A) = A/_I \quad \text{y} \quad \pi(I) = \{I\}.$$

Demostración. Es fácil comprobar que si J es un ideal de A que contiene a I, $\pi(J)$ es un ideal de $A/_I$, y que si $\overline{J}$ es un ideal de $A/_I$, $\pi^{-1}(\overline{J})$ es un ideal de A que contiene a I.

La aplicación $J \mapsto \pi(J)$ es una biyección del conjunto de los ideales de A que contienen a I sobre el conjunto de los ideales de $A/_I$. Es una inyección porque si se tiene $\pi(J) = \pi(J')$ y $a \in J$ en-

tonces existe $a' \in J'$ tal que $\pi(a) = \pi(a')$. Luego se tiene

$$a - a' \in I \subset J' ,$$

y por tanto $a = a' + (a - a') \in J'$. Es decir que $J \subset J'$. Semejantemente se prueba que $J' \subset J$ y por consiguiente es $J = J'$.

La aplicación considerada es también una sobreyección por que si $\overline{J}$ es un ideal de A/I, $\pi^{-1}(\overline{J})$ es un ideal de A que contiene a I, tal que

$$\pi(\pi^{-1}(\overline{J})) = \overline{J} . \qquad \blacksquare$$

Corolario. Sea A un anillo conmutativo con identidad, y sea I un ideal de A. El anillo cociente A/I es un campo si y solo si I es maximal.

Demostración. El anillo conmutativo con identidad A/I es un campo si y solo si no tiene más ideales que él mismo y el ideal cero (teorema 9.2.2). En virtud del teorema 9.4.1 esa condición equivale a que en A no existen otros ideales que contienen a I que I mismo y A, es decir que I es ideal maximal de A. $\blacksquare$

Ejemplo. Los ideales $I(m)$ de $\mathbb{Z}$ son maximales si y solo si m es primo. Solo en ese caso $\mathbb{Z}/I(m)$ será un campo. Así $\mathbb{Z}/I(3)$ es un campo, pero $\mathbb{Z}/I(6)$ no lo es. A este respecto ver el primer ejemplo de la sección 9.1.

Ejercicio

*9.4.1. Si I es un ideal del anillo A y π es la proyección de A sobre A/I, entonces:

1. Si J es un ideal cualquiera de A, $\pi(J)$ es un ideal de A/I.
2. Si $\overline{J}$ es un ideal cualquiera de A/I, $\pi^{-1}(J)$ es un ideal de A.

9.5. Dominios de integridad. Campo de cocientes.

Definición 9.5.1. Un anillo conmutativo con identidad, A, se llama *dominio de integridad*, si dados en A dos elementos a y b de A tales que $ab = 0$, necesariamente uno por lo menos de los dos es igual

a 0. Se expresa esta condición diciendo que A no tiene divisores de cero. Se dice también que A es un *dominio integral*.

Si A es un dominio integral que no es un campo puede construirse un campo $\mathbb{F}$ que contiene a A como subconjunto, de manera que tanto la suma como el producto de dos elementos de A son los mismos en el dominio integral dado como en el campo $\mathbb{F}$. Expresado de otra manera, el dominio integral A puede considerarse como un subanillo del campo $\mathbb{F}$. El método de construcción de $\mathbb{F}$ a partir de A es calcado exactamente del que se emplea para construir el campo $\mathbb{Q}$ de los números racionales a partir del dominio de integridad $\mathbb{Z}$ de los enteros. Nos limitaremos por tanto a exponer el procedimiento en forma muy suscinta, dejando al lector el cuidado de detallar las pruebas.

Designemos por S al conjunto de todos los pares (a,b) tales que $a,b \in A$ y que $b \neq 0$.

Lema. La relación

$$E : (a,b) \approx (c,d) \Leftrightarrow ad = bc$$

es una relación de equivalencia

Designemos por $(a,b)^*$ a la clase de equivalencia que contiene al par (a,b). Se demuestra entonces que el conjunto $S/_E$ de las clases de equivalencia de S respecto de la relación E constituyen un campo $\mathbb{F}$ si se definen las operaciones de la manera siguiente:

$$(a,b)^* + (c,d)^* = (ad + bc\ ,\ bd)^*$$

$$(a,b)^*(c,d)^* = (ac\ ,\ bd)^*$$

En especial se prueba que los resultados de estas operaciones no dependen de los pares elegidos para representar a las clases con que se opera.

Puede hacerse las observaciones siguientes:

- El elemento nulo de $\mathbb{F}$ es $(0,1)^*$
- La inversa aditiva de $(a,b)^*$ es $(-a,b)^*$
- La identidad de $\mathbb{F}$ es $(1,1)^*$

- Si $(a,b)^* \neq (0,1)^*$, es decir si $a \neq 0$, entonces la inversa de $(a,b)^*$ es $(b,a)^*$.

La identificación de los elementos de A con los elementos de un subconjunto de $\mathbb{F}$ puede hacerse mediante la aplicación $\phi: A \to \mathbb{F}$ definida por

$$\phi(a) = (a,1)^* \text{ , para cada } a \in A.$$

Se tiene entonces:

1. $\phi(a) = \phi(b) \Leftrightarrow a = b$ *(inyectividad)*
2. $\phi(a + b) = \phi(a) + \phi(b)$, $\forall\, a,b \in A$
3. $\phi(ab) = \phi(a)\phi(b)$, $\forall\, a,b \in A$
4. Cada elemento $(a,b)^*$ puede expresarse en la forma

$$(a,1)^*(1,b)^* = (a,1)^*[(b,1)^*]^{-1} = \phi(a)[\phi(b)]^{-1} = \phi(a)/\phi(b) .$$

Resulta de lo dicho que la aplicación $a \in A \mapsto \phi(a) \in \mathbb{F}$ es un isomorfismo de A sobre $\phi(A) \subset \mathbb{F}$. Por tal motivo podemos identificar al elemento $a \in A$ con el elemento $\phi(a) \in \mathbb{F}$, y denotar a este último elemento por a. La propiedad 4 anterior permite entonces escribir cada elemento de $\mathbb{F}$ en la forma

$$ab^{-1} = a/b \quad , \quad a,b \in A \ , \ b \neq 0.$$

Definición 9.5.2. El campo $\mathbb{F}$ construido como se acaba de ver, a partir de un dominio de integridad A, se llama *campo de cocientes* de dicho dominio.

Ejercicios

9.5.1. El conjunto de todas las funciones $f(x)$ continuas en el intervalo $[0,1]$ es un anillo con las operaciones

$$(f + g)(x) = f(x) + g(x)$$

$$(fg)(x) = f(x)g(x) .$$

¿Es este anillo un dominio de integridad?

***9.5.2.** Probar que en un dominio de integridad se cumple la ley de

cancelación, es decir que si $ax = bx$ y $x \neq 0$ entonces es $a = b$.

*9.5.3. Sea A un anillo conmutativo con identidad. Un ideal I de A se llama *ideal primo* si $ab \in I$ implica que $a \in I$ o que $b \in I$. Demostrar que un ideal I es primo si y solo si el anillo cociente $A/_I$ es un dominio de integridad.

*9.5.4. Si A es un anillo de integridad y $\mathbb{F}$ es un campo de cocientes, no existe ningún subcampo de $\mathbb{F}$, distinto de $\mathbb{F}$, que contenga a A.

§ 10. Polinomios

10.1. Polinomios sobre un campo

Definición 10.1.1. Un *polinomio sobre un campo* $\mathbb{F}$ es una sucesión infinita

$$f = (a_0 , a_1 , a_2 , \ldots)$$

de elementos de $\mathbb{F}$ que son todos nulos, salvo, a lo más, un número finito de ellos. Los elementos $a_0 , a_1 , a_2 , \ldots$ se llama *coeficientes* del polinomio. Si todos los coeficientes son nulos se dice que f es el *polinomio cero*. En caso contrario, el mayor número n tal que $a_n \neq 0$ se llama *grado* del polinomio y se representa por *grado* f. El coeficiente a_n se llama en tal caso *coeficiente director* del polinomio.

Dos polinomios son iguales si y solo si son iguales los coeficientes del mismo orden, es decir que

$$(a_0 , a_1 , \ldots) = (b_0 , b_1 , \ldots) \Leftrightarrow a_0 = b_0 , a_1 = b_1 , \ldots .$$

Definiremos ahora las operaciones de adición y de multiplicación de polinomios.

Dados los polinomios

$$f = (a_0 , a_1 , \ldots) \quad \text{y} \quad g = (b_0 , b_1 , \ldots)$$

se llama *suma* de estos dos polinomios, al polinomio

$$f + g = (a_o + b_o , a_1 + b_1 , \ldots) ;$$

y se llama *producto* de los mismos polinomios, al polinomio

$$fg = (c_o , c_1 , \ldots)$$

tal que $c = a_o b_o$, $c_1 = a_o b_1 + a_1 b_o$, $c_2 = a_o b_2 + a_1 b_1 + a_2 b_o$, y en general,

$$c_k = \sum_{i+j=k} a_i b_j \qquad , \qquad (k = 0,1,2,\ldots) .$$

Es muy fácil comprobar que el conjunto de todos los polinomios sobre $\mathbb{F}$, con las operaciones que acabamos de definir, es un *anillo conmutativo* . El elemento cero de ese anillo es el polinomio $(0 , 0 , \ldots)$, y su elemento identidad es el polinomio $(1 , 0 , \ldots)$. El anillo de los polinomios sobre $\mathbb{F}$ se representará en adelante por $\mathbb{F}[x]$.

El anillo $\mathbb{F}[x]$ no es un campo. En efecto, el ideal generado por x no es el ideal cero, pues contiene a $x \neq 0$; y tampoco es el ideal $\mathbb{F}[x]$ pues no contiene a los polinomios constantes. Se aplica entonces el teorema 9.2.2.

Las consideraciones que siguen tienen por objeto justificar la manera usual de representar a los polinomios.

Sea $\varphi : \mathbb{F} \to \mathbb{F}[x]$ una aplicación tal que

$$\varphi(a) = (a , 0 , 0 , \ldots) \qquad , \qquad \forall\, a \in \mathbb{F} .$$

φ es evidentemente una biyección de $\mathbb{F}$ sobre el conjunto de todos los polinomios de la forma $(a , 0 , 0 , \ldots)$ en que a es un elemento cualquiera de $\mathbb{F}$. Esa biyección tiene las propiedades

$$\begin{aligned} \varphi(a + b) &= \varphi(a) + \varphi(b) \\ \varphi(ab) &= \varphi(a)\varphi(b) \end{aligned} \qquad (10.1.1)$$

y por tanto es un isomorfismo de $\mathbb{F}$ sobre dicho conjunto de polinomios.

Sustituyamos ahora en $\mathbb{F}[x]$ cada polinomio $(a , 0 , 0 , \ldots)$ por el escalar a correspondiente, pero sin modificar las reglas operativas. Se obtendrá así un nuevo anillo que seguiremos representando por $\mathbb{F}[x]$ Este anillo contendrá a $\mathbb{F}$, con cuyos elementos, en virtud de las re-

laciones (10.1.1), se operará como si fueran elementos del campo . Además tendrán lugar las reglas operatorias.

$$a + (b_o, b_1, \ldots) = (a + b_o, b_1, \ldots)$$
$$a(b_o, b_1, \ldots) = (ab_o, ab_1, \ldots).$$

En adelante, cuando nos refiramos al anillo de polinomios $\mathbb{F}[x]$ entenderemos que en él está contenido el campo $\mathbb{F}$, cada uno de cuyos elementos, a, representa a un polinomio constante $(a, 0, 0, \ldots)$. En particular, el polinomio cero será representado por el escalar 0, y el polinomio identidad, por 1.

En virtud de todo lo precedente un polinomio cualquiera de $\mathbb{F}[x]$ puede expresarse de la manera siguiente

$$\begin{aligned}(a_o a_1, a_2, \ldots) &= (a_o, 0, 0, \ldots) + (0, a_1, 0, 0, \ldots) + \\ &+ (0, 0, a_2, 0, \ldots) + \cdots \\ &= a_o(1, 0, \ldots) + a_1(0, 1, 0, \ldots) + \\ &+ a_2(0, 0, 1, 0, \ldots) + \cdots\end{aligned}$$

Si sustituimos $(1, 0, \ldots)$ por su representante 1 e introducimos la notación

$$x = (0, 1, 0, \ldots)$$

de la cual se deduce

$$x^2 = (0, 0, 1, 0, \ldots)$$
$$x^3 = (0, 0, 0, 1, 0, \ldots), \text{ etc.},$$

se tendrá finalmente que

$$(a_o, a_1, a_2, \ldots) = a_o + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

donde n es el grado del polinomio considerado. De esta manera ha quedado establecido que los polinomios de $\mathbb{F}[x]$ pueden escribirse en la forma

$$f = a_o + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum_{i=0}^{n} a_i x^i, \quad (x^o = 1).$$

donde n es un número natural cualquiera, y a_0, a_1, a_2, ... son escalares arbitrarios. x se llama *indeterminada*.

Se demuestra inmediatamente, de conformidad con las operaciones con polinomios, las reglas bien conocidas:

$$(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) + (b_0 + b_1x + b_2x^2 + \cdots + b_n x^m)$$
$$= (a + b) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$$

$$(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n)(b_0 + b_1x + b_2x^2 + \cdots + b_nx^m)$$
$$= a_0b_0 + (a_0b_1 + a_1b_0)x + \cdots + a_nb_nx^{n+m} .$$

De la última de las dos relaciones precedentes se deduce inmediatamente la proposición siguiente:

Teorema 10.1.1. Para los polinomios f, g de $\mathbb{F}[x]$ se cumplen las propiedades siguientes:

1. *grado* (fg) = *grado* f + *grado* g, si tanto f como g son diferentes de 0
2. Si $fg = 0$, entonces o bien $f = 0$ o bien $g = 0$.

Esta última propiedad puede expresarse diciendo que *el anillo* $\mathbb{F}[x]$ *es un dominio de integridad*, de donde resulta una analogía entre este anillo y el anillo de los enteros (Ninguno de los dos es un campo) que será más destacada aún en las próximas secciones en que estableceremos la teoría de la divisibilidad de los polinomios según el módelo de la teoría de la divisibilidad de los números enteros.

10.2. División de polinomios. En esta sección veremos que en el anillo de los polinomios existe una división semejante a la que existe entre los números enteros.

Lema 10.2.1. Si f y d son dos polinomios no nulos de $\mathbb{F}[x]$, tales que *grado* $d \leqslant$ *grado* f, entonces existe un polinomio $g \in \mathbb{F}[x]$ tal que, o bien es

$$f - dg = 0 \quad \text{o bien} \quad \textit{grado}\ (f - dg) < \textit{grado}\ f .$$

Demostración. Si convenimos en definir $x^0 = 1$ y

$$\sum_{i=0}^{-1} a_i x^i = \sum_{i=0}^{-1} b_i x^i = 0 ,$$

entonces los polinomios no nulos f y d, de grados n y k respectivamente, pueden escribirse en las formas

$$f = a_n x^n + \sum_{i=0}^{n-1} a_i x^i \quad , \quad a_n \neq 0 ,$$

$$d = b_k x^k + \sum_{j=0}^{k-1} b_j x^j \quad , \quad b_k \neq 0 , \ k \leqslant n.$$

Se ve inmediatamente que el polinomio

$$g = (a_n b_k^{-1}) x^{n-k}$$

satisface las condiciones del enunciado. ■

Teorema 10.2.1. *(Algoritmo de la división).* Sean $d \neq 0$ y f polinomios de $\mathbb{F}[x]$. Existen entonces los polinomios q y r de $\mathbb{F}[x]$ tales que

1. $f = dq + r$
2. o bien $r = 0$ o bien *grado* $r <$ *grado* d.

Además, los polinomios q y r son los únicos que satisfacen a las condiciones anteriores.

Demostración. Probemos en primer lugar que existen los polinomios q y r que cumplen las condiciones del teorema.
Si *grado* $f <$ *grado* d, la condición 1. implica que $q = 0$ y que $r = f$; y para estos valores de q y de r se cumplen obviamente las dos condiciones del teorema.

Supongamos ahora que sea *grado* $f \geqslant$ *grado* d. En virtud del lema precedente existe el polinomio $g \in \mathbb{F}[x]$ tal que es o bien

$$f - dg = 0 \quad \text{o bien} \quad \textit{grado}\ (f - dg) < \textit{grado}\ f .$$

Si $f - dg = 0$, o si *grado* $(f - dg) <$ *grado* d, entonces es evidente que puede suponerse $q = g$ y $r = f - dg$.

Si $f - dg \neq 0$ y *grado* $(f - dg) \geqslant$ *grado* d, entonces podemos aplicar el lema precedente a los polinomios $f - dg$ y d y hallaremos el polinomio $h \in \mathbb{F}[x]$ tal que o bien $f - dg - dh = 0$ o bien

$$grado\ (f - dg - dh) < grado\ (f - dg)\ ,$$

es decir

$$grado\ (f - d(g + h)) < grado\ (f - dg)\ .$$

Si $f - dg - dh = 0$, o si *grado* $(f - d(g + h)) <$ *grado* d, entonces puede suponerse

$$q = g + h \quad \text{y} \quad r = f - d(g + h)\ .$$

Si en cambio, es $f - dg - dh \neq 0$ y además

$$grado\ (f - d(g + h)) \geqslant grado\ d\ ,$$

podemos aolicar nuevamente el lema a los polinomios $f - d(g + h)$ y d. Continuando el proceso de la misma manera llegaremos, en un número finito de pasos, a obtener los polinomios q y r que satisfacen las condiciones del enunciado.

Probemos ahora que los polinomios q y r, que cumplen las condiciones del teorema, son únicos.

Supongamos que no fuera así, y que q' y r' cumplieran las mismas condiciones pero que $q' \neq q$. Por la condición 1. del enunciado se tendría

$$f = dq + r = dq' + r'\ ,$$

de donde

$$d(q - q') = r' - r\ , \qquad (10.2.1)$$

y por tanto $r' - r \neq 0$ y

$$grado\ d(q - q') = grado\ (r' - r)$$

La propiedad 1. del teorema 10.1.1 permitiría deducir de esta última igualdad que

$$grado\ (r' - r) = grado\ d + grado\ (q - q'). \qquad (10.2.2)$$

Ahora bien, la condición 2. del enunciado implicaría que

$$grado\ (r' - r) < grado\ d$$

(recordar que r' y r no se pueden anular simultaneamente porque $(r' - r) \neq 0$); luego, puesto que

$$grado\ d \leqslant grado\ d + grado\ (q - q') \ ,$$

resultaría

$$grado\ (r' - r) < grado\ d + grado\ (q - q') \ ,$$

en contradicción con (10.2.2). Se concluye entonces que debe ser $q' = q$ y por consiguiente, en virtud de (10.2.1), que

$$r' = r \ . \quad \blacksquare$$

Definición 10.2.1. Dados los polinomios $d \neq 0$ y f, los polinomios q y r, cuya existencia ha sido establecida en el teorema 10.2.1, se llaman *cocientes y resto*, respectivamente, de la *división de* f por d.

En el caso en que $r = 0$ se dice que f es un *multiplo* de d o que f es *divisible* por f. d se llama entonces *divisor* de f, y se escribe $d \mid f$.

10.3. Ideales de $\mathbb{F}[x]$. Al anillo $\mathbb{F}[x]$ son aplicables las consideraciones del parágrafo 9, y en espacial lo dicho en la sección 9.2 acerca de los ideales. Podemos entonces demostrar que $\mathbb{F}[x]$, así como antes se demostró para $\mathbb{Z}$ (sección 9.2, Ejemplo), es un *anillo de ideales principales*, es decir que cada ideal es principal.

Definición 10.3.1. Se dice que un polinomio no nulo es *mónico* si su coeficiente director es igual a 1.

Teorema 10.3.1. Cada ideal I de $\mathbb{F}[x]$ es un ideal principal. Además si I no es el ideal cero, existe un único polinomio mónico d tal que I es el ideal principal $I(d)$ generado por d. d es entonces el polinomio mónico de grado mínimo contenido en I.

Demostración. Si I es el ideal cero se tiene $I = I(0)$.

Supongamos por tanto que $I \neq \{0\}$. En el caso en que I contiene a un polinomio constante $a \neq 0$, se tiene

$$a^{-1}a = 1 \in I ,$$

y por tanto, cualquiera que sea $g \in \mathbb{F}[x]$, se tiene $g = 1.g \in I$. Resulta así que en ese caso es $I = \mathbb{F}[x] = I(1)$, y 1 es el polinomio mónico a que se refiere el enunciado.

Supongamos ahora que todos los polinomios no nulos de I tienen *grado* $\geqslant 1$. Entre ellos habrá alguno, que designaremos por d, cuyo *grado* k es mínimo. Podemos suponer que es mónico. Si no lo fuera, podríamos sustituirlo por el polinomio $a_k^{-1} d \in I$, donde a_k es el coeficiente director de d. Puesto que $d \in I$ es claro que $I(d) \subset I$. Si f es un polinomio cualquiera de I, podemos escribir (teorema 10.2.1)

$$f = dq + r \quad , \quad (r = 0, \text{ o bien } grad\ r < grad\ d).$$

Dado que tanto d como f pertenecen a I, se tendrá que

$$r = f - dq \in I.$$

Si fuera $r \neq 0$, entonces sería *grado* $r <$ *grado* d, lo cual constituiría una contradicción con la manera como hemos elegido d. Luego debe ser $r = 0$ y por lo tanto $f = dq$, lo cual prueba que

$$f \in I(d) \ , \text{ y por consiguiente } I \subset I(d).$$

En resumen resulta $I = I(d)$

Vamos a probar finalmente que el polinomio mónico d tal que $I = I(d)$ es único. Supongamos que d' fuera un polinomio mónico tal que $I = I(d')$. Puesto que $d \in I(d')$ existiría un polinomio

$$q' \in \mathbb{F}[x] \quad \text{para el cual} \quad d = d'q' \ .$$

Si q' no fuera constante se tendría *grado* $d >$ *grado* d', y el polinomio d' sería un polinomio de I de grado menor que el de d, lo cual contradiría la forma como fué definido el polinomio d. Por consiguiente debe ser q' constante. Pero d y d' son mónicos, luego en virtud de la igualdad $d = d'q'$, tendrá que ser $q' = 1$ y $d = d'$. ■

10.4. Máximo común divisor y mínimo común múltiplo de polinomios. Consideraremos polinomios no nulos.

Teorema 10.4.1. Dados los polinomios $f_1, \ldots, f_n$ de $\mathbb{F}[x]$ existe un

único mónico $d(x)$ con las siguientes propiedades:

1. d divide a $f_1, \ldots, f_n$
2. Si d' divide a $f_1, \ldots, f_n$, entonces d' divide a d. Además existen $g_1, \ldots, g_n \in \mathbb{F}[x]$ tales que

$$d = f_1 g_1 + \ldots + f_n g_n$$

Demostración. Consideremos el conjunto I de todos los polinomios de $\mathbb{F}[x]$ de la forma

$$\sum_{i=1}^{n} f_i g_i \ ,$$

donde $g_1, \ldots, g_n$ son polinomios cualesquiera de $\mathbb{F}[x]$. Evidentemente I es el ideal $I(f_1) + \ldots + I(f_n)$ de $\mathbb{F}[x]$ (ver el teorema 9.2.1). En virtud del teorema 10.3.1 existe un polinomio mónico único d tal que $I(d) = I$. Puesto que $f_i \in I$ para $i = 1, \ldots, n$, se sigue que d divide a f_i para cada i. Además, por definición de I, puede esribirse

$$d = \sum_{i=1}^{n} f_i g_i$$

para polinomios $g_1, \ldots, g_n \in \mathbb{F}[x]$ convenientemente elegidos. Si d' es un polinomio que divide a todos los f_i entonces d' divide a d, porque de $f_i = h_i d'$, $i = 1, \ldots, n$, se sigue

$$d = \sum_{i=1}^{n} f_i g_i = \sum_{i=1}^{n} d' h_i g_i = d' \left(\sum_{i=1}^{n} h_i g_i \right) .$$

Resulta así que el polinomio mónico d cumple las propopiedades del enunciado. Solo falta probar que es el único polinomio mónico que tiene esa propiedad. En efecto, si $\overline{d} \in \mathbb{F}[x]$ fuera un polinomio mónico que tuviera las mismas propiedades, se tendría que $d \mid \overline{d}$ y que $\overline{d} \mid d$. Sería por consiguiente $I(\overline{d}) \subset I(d)$ e $I(d) \subset I(\overline{d})$ o sea $I(\overline{d}) = I(d) = I$. Dado que d es el único polinomio mónico tal que $I(d) = I$, se concluye que $\overline{d} = d$. ∎

El teorema que acabamos de demostrar permite dar la definición siguiente:

Definición 10.4.1. Dados $f_1, \ldots, f_n \in \mathbb{F}[x]$, el único polinomio mónico generador del ideal $I(f_1) + \ldots + I(f_n)$ se llama *máximo común divisor (M.C.D.)* de $f_1, \ldots, f_n$, y es el único polinomio mónico que goza de las propiedades siguientes:

1. divide a $f_1, \ldots, f_n$
2. es dividido por cada polinomio que divide a $f_1, \ldots, f_n$
3. Existen $g_1, \ldots, g_n \in \mathbb{F}[x]$ tales que dicho polinomio puede expresarse en la forma $g_1 f_1 + \ldots + g_n f_n$

El *M.C.D.* de $f_1, \ldots, f_n$ se designa por $[f_1, \ldots, f_n]$. Se tiene

$$\begin{aligned} I([f_1, \ldots, f_n]) &= I(f_1) + \ldots + I(f_n) \\ &= I([f_1, \ldots, f_k]) + I([f_{k+1}, \ldots, f_n]) \\ &= I([f_1, \ldots, f_k], [f_{k+1}, \ldots, f_n]) \end{aligned}$$

De la unicidad se deduce:

$$[f_1, \ldots, f_n] = [[f_1, \ldots, f_k], [f_{k+1}, \ldots, f_n]].$$

Además se tiene evidentemente: $[f_1, f_2] = [f_2, f_1]$.

Las relaciones *asociativa y conmutativa* que acabamos de obtener se generalizan facilmente.

Definición 10.4.2. Se dice que los polinomios $f_1, \ldots, f_n$ son *primos relativos o primos entre sí*, si $[f_1, \ldots, f_n] = 1$.

De esta definición y del teorema 10.4.1 se sigue que si $f_1, \ldots, f_n$ son primos entre sí, existen los polinomios $g_1, \ldots, g_n$ tales que

$$f_1 g_1 + \ldots + f_n g_n = 1.$$

Esta relación es también suficiente para que $f_1, \ldots, f_n$ sean primos entre sí. En efecto, se deduce de ella que

$$1 \in I(f_1) + \ldots + I(f_n) = I([f_1, \ldots, f_n])$$

Pero un ideal que contiene al polinomio 1 es idéntico con el ideal $I(1) = \mathbb{F}[x]$. Luego $I([f_1, \ldots, f_n]) = I(1)$, y por tanto $[f_1, \ldots, f_n] = 1$.

10. Polinomios.

La analogía que podemos observar hasta ahora con el caso de los números enteros nos permite ahora establecer el *algoritmo de Euclides* para obtener el *M.C.D.* de dos polinomios cualesquiera. Convengamos primero en designar por $\overline{g}$ al polinomio mónico que está asociado a un polinomio g; es decir el que se obtiene multiplicando a g por la inversa de sus coeficiente director.

Sean f y g dos polinomios cualesquiera (no nulos). Supongamos que sea *grado* $f \geqslant$ *grado* g. Sea entonces

$$f = gq + g_1 \ , \ g_1 = 0 \text{ o bien } \textit{grado } g_1 < \textit{grado } g \ .$$

Si $g_1 = 0$, entonces $f = gq$ y por tanto

$$I([f, g]) = I(f) + I(g) = I(gq) + I(g) = I(g) = I(\overline{g}) \ ,$$

pues $I(gq) \subset I(g)$. Resulta por consiguiente que $[f, g] = \overline{g}$.

Si $g_1 \neq 0$, entonces

$$I([f, g]) = I(f) + I(g) = I(gq + g_1) + I(g) = I(g_1) + I(g)$$

$$= I([g, g_1]) \ .$$

(Demostrar que $I(gq + g_1) + I(g) = I(g_1) + I(g)$) .

Por consiguiente

$$[f, g] = [g, g_1] \ .$$

Sea ahora

$$g = g_1 q_1 + g_2 \ , \ g_2 = 0 \text{ o bien } \textit{grado } g_2 < \textit{grado } g_1 \ .$$

Si $g_2 = 0$, entonces $g = g_1 q_1$ y por tanto

$$I([f, g]) = I([g, g_1]) = I(g) + I(g_1) = I(g_1 q_1) + I(g_1)$$

$$= I(g_1) = I(\overline{g_1}) \ ,$$

es decir que $[f, g] = [g, g_1] = \overline{g_1}$.

Si $g_2 \neq 0$, entonces

$$I([f, g]) = I([g, g_1]) = I(g) + I(g_1) = I(g_1 q_1 + g_2) + I(g_1)$$

$$= I(g_2) + I(g_1) = I([g_1 , g_2]) ,$$

y por consiguiente

$$[f , g] = [g_1 , g_2] .$$

Continuando de esta manera obtendremos los polinomios no nulos

$$g , g_1 , \ldots , g_n$$

de grados decrecientes, tales que

$$g_i = g_{i+1} q_{i+1} + g_{i+2}$$

para $n = 0,1, \ldots , n-2$, de donde resulta

$$[f , g] = [g_{i+1} , g_{i+2}] ;$$

y

$$g_{n-1} = g_n q_n .$$

Se deduce facilmente que

$$[f , g] = [g_{n-1} , g_n] = \overline{g_n} . \quad \blacksquare$$

Definición 10.4.3. Dados los polinomios $f_1 , \ldots , f_n$ se llama *mínimo común múltiplo* (*M.C.M*) de $f_1 , \ldots , f_n$ al polinomio mónico generador del ideal

$$I(f_1) \cap \ldots \cap I(f_n).$$

Ese polinomio es el único polinomio mónico que posee las propiedades siguientes:

1. es múltiplo de $f_1 , \ldots , f_n$
2. es divisor de cada polinomio que es múltiplo de $f_1 , \ldots , f_n$.

Representemos por $\{f_1 , \ldots , f_n\}$ al mínimo común múltiplo de $f_1, \ldots , f_n$. Nos limitaremos a señalar las relaciones asociativas y conmutativas:

$$\begin{aligned} I(\{f_1 , \ldots , f_n\}) &= I(f_1) \cap \ldots \cap I(f_n) \\ &= [I(f_1) \cap \ldots \cap I(f_k)] \cap [I(f_{k+1}) \cap \ldots \cap I(f_n)] \\ &= I(\{f_1 , \ldots , f_k\}) \cap I(\{f_{k+1} , \ldots , f_n\}) \end{aligned}$$

$$= I(\{f_1, \ldots, f_k\}, \{f_{k+1}, \ldots, f_n\}) \quad ,$$

de donde resulta

$$\{f_1, \ldots, f_n\} = \{\{f_1, \ldots, f_k\}, \{f_{k+1}, \ldots, f_n\}\} \ .$$

Además se tiene

$$\{f_1, f_2\} = \{f_2, f_1\} \quad .$$

Ejercicios

***10.4.1.** Demostrar que si f y g son polinomios tales que $I(f) = I(g)$ entonces existe $a \in \mathbb{F}$ de modo que $f = ag$. Si f y g son mónicos deberá ser $f = g$.

***10.4.2.** Probar que cualesquiera que sean los polinomios f_1, f_2 y g se tiene

$$I(f_1) + I(gf_1 + f_2) = I(f_1) + I(f_2) \quad .$$

Ejemplo. Consideremos los polinomios

$$f = 3x^6 - x^5 - 9x^4 - 14x^3 - 11x^2 - 3x - 1 \ ,$$

$$g = 3x^5 + 8x^4 + 9x^3 + 15x^2 + 10x + 9.$$

Las divisiones sucesivas que corresponden al algoritmo de Euclides, proporcionan las relaciones:

$$f = g.(x - 3) + 6x^4 - 2x^3 + 24x^2 + 18x + 26 \qquad (10.4.1)$$

$$g = g_1(\tfrac{1}{2}x + \tfrac{3}{2}) - 30x^2 - 30x - 30 \quad , \qquad (10.4.2)$$

donde

$$g_1 = 6x^4 - 2x^3 + 24x^2 + 18x + 26,$$

y por último

$$g_1 = (-30x^2 - 30x - 30)(-\tfrac{1}{5}x^2 + \tfrac{4}{15}x - \tfrac{13}{15}) \quad .$$

El último resto diferente de cero es $-30x^2 - 30x - 30$.

Luego el *M.C.D.* de f y g es el polinomio mónico asociado con ese resto, es decir $x^2 + x + 1$.

Valiéndonos de (10.4.2) y (10.4.1) podemos obtener una expresión del *M.C.D.* como combinación lineal de f y g. En efecto, de (10.4.2) se deduce

$$-30x^2 - 30x - 30 = g - g_1.(\frac{1}{2}x + \frac{3}{2}) \quad .$$

Sustituyendo aquí el valor de g_1, deducido de (10.4.1), se obtiene

$$-30x^2 - 30x - 30 = g - (f - g.(x-3))(\frac{1}{2}x + \frac{3}{2})$$

$$= (-\frac{1}{2}x - \frac{3}{2})f + (\frac{1}{2}x^2 - \frac{7}{2})g \ .$$

Dividiendo por -30 resulta por último

$$x^2 + x + 1 = (\frac{1}{60}x + \frac{1}{20})f - (\frac{1}{60}x^2 - \frac{7}{60})g$$

que es la expresión buscada.

Ejercicios

10.4.3. Hallar el *M.C.D.* de los siguientes pares de polinomios

a) $x^4 + x^3 - 3x^2 - 4x - 1 \ ; \ x^3 + x^2 - x - 1$

b) $x^5 + x^4 - x^3 - 2x - 1 \ ; \ 3x^4 + 2x^3 + x^2 + 2x - 2$

c) $x^6 - 7x^4 + 8x^3 - 7x + 7 \ ; \ 3x^5 - 7x^3 + 3x^2 - 7$

d) $x^5 - 2x^4 + x^3 + 7x^2 - 12x + 10; \ 3x^4 - 6x^3 + 5x^2 + 2x - 2$

e) $x^6 + 2x^4 - 4x^3 - 3x^2 + 8x - 5 \ ; \ x^5 + x^2 - x + 1$

f) $x^5 + 3x^4 - 12x^3 - 52x^2 - 52x - 12; \ x^4 + 3x^3 - 6x^2 - 22x - 12$

g) $x^5 + x^4 - x^3 - 3x^2 - 3x - 1 \ ; \ x^4 - 2x^3 - x^2 - 2x + 1$.

10.4.4. Mediante el algoritmo de Euclides obtener polinomios M_1 y M_2 tales que

$$M_1 f_1 + M_2 f_2 = [f_1 \ , \ f_2]$$

a) $f_1 = x^4 + 2x^3 - x^2 - 4x - 2$
$f_2 = x^4 + x^3 - x^2 - 2x - 2$

b) $f_1 = x^5 + 3x^4 + x^3 + x^2 + 3x + 1$
$f_2 = x^4 + 2x^3 + x + 2$

c) $f_1 = x^6 - 4x^5 + 11x^4 - 27x^3 + 37x^2 - 35x + 35$
$f_2 = x^5 - 3x^4 + 7x^3 - 20x^2 + 10x - 25.$

10.4.5. Mediante el algoritmo de Euclides obtener polinomios M_1 y M_2 tales que

$$M_1 f_1 + M_2 f_2 = 1.$$

a) $f_1 = 3x^3 - 2x^2 + x + 2$
$f_2 = x^2 - x + 1$

b) $f_1 = x^4 - x^3 - 4x^2 + 4x + 1$
$f_2 = x^2 - x - 1$

c) $f_1 = x^5 - 5x^4 - 2x^3 + 12x^2 - 2x + 12$
$f_2 = x^3 - 5x^2 - 3x + 17.$

*10.4.6. Demostrar que el polinomio mónico generador del ideal

$$I(f_1) \cap \cdots \cap I(f_n)$$

es el único polinomio mónico que tiene las propiedades de la definición 10.4.3.

*10.4.7. Demostrar que si d es el *M.C.D.* de los polinomios f_1 y f_2, entonces

$$I(f_1 f_2/d) = I(f_1) \cap I(f_2),$$

y por tanto el *M.C.M.* de f_1 y f_2 es el polinomio mónico asociado al polinomio $f_1 f_2/d = f_1(f_2/d) = f_2(f_1/d)$.

10.4.8. Calcular el *M.C.M.* de cada uno de los pares de polinomios de los ejercicios 10.4.3. y 10.4.5.

10.5 Factorización y divisibilidad en $\mathbb{F}[x]$

Definición 10.5.1. Un polinomio *no constante* $f \in \mathbb{F}[x]$ se llama *irreducible en* $\mathbb{F}[x]$ *(o irreducible sobre* $\mathbb{F}[x]$*)*, si la relación $f = gh$, donde $g,h \in \mathbb{F}[x]$ implica que uno de los polinomios g o h es una constante no nula $a \in \mathbb{F}$ y que, por lo tanto, el otro es igual a $a^{-1}f$. En particular los polinomios de primer grado son irreducibles.

Ejemplos. A. Sea $\mathbb{F}$ el campo $\{0, 1\}$ de los enteros módulo 2. Puesto que en $\mathbb{F}[x]$ es

$$x^2 + 1 = (x + 1)^2 ,$$

$x^2 + 1$ no es irreducible. Por otra parte, $x^2 + 2x + 1$ es irreducible sobre $\mathbb{F}$ porque los únicos polinomios de grado 1 de $\mathbb{F}[x]$ son x y $x + 1$, y ninguno de los productos de dos de ellos es igual a $x^2 + 2x + 1$.

B. El polinomio $x^2 + 1$ es irreducible sobre $\mathbb{R}$ pero no sobre $\mathbb{C}$, pues $x^2 + 1 = (x + i)(x - i)$ en $\mathbb{C}[x]$.

Teorema 10.5.1. Si $f \in \mathbb{F}[x]$ es un polinomio irreducible, el ideal $I(f)$ es maximal y por tanto $\mathbb{F}[x]/_{I(f)}$ es un campo.

Demostración. Sea I_1 un ideal de $\mathbb{F}[x]$ que contiene a $I(f)$. Existirá entonces un polinomio $g \in I_1$ que no pertenece a $I(f)$. Puesto que f es irreducible y que g no es múltiplo de f, dichos polinomios son primos entre sí. Existen por tanto los polinomios r y s de $\mathbb{F}[x]$ tales que

$$rf + sg = 1 .$$

Puesto que I_1 contiene a f y a g, contendrá al polinomio constante 1, y por consiguiente coincide con $\mathbb{F}[x]$. $I(f)$ es por tanto maximal. ∎

Teorema 10.5.2. Sean $f,g \in \mathbb{F}[x]$. Si el producto fg es divisible por un polinomio p irreducible sobre $\mathbb{F}$, entonces o bien f o bien g es divisible por p.

Demostración. Es evidente que basta considerar el caso en que p es un polinomio mónico irreducible que divide a fg. Sea $d = [p, f]$ el

M.C.D. de p y f. Puesto que p es mónico irreducible, necesariamente es o bien $d = p$ o bien $d = 1$. En el primer caso $p = [p,f]$ dividirá a f. En el segundo caso existirán, en virtud del teorema 10.4.1, los polinomios h y k tales que

$$fh + pk = 1 ,$$

de donde, multiplicando ambos miembros por g,

$$hfg + pkg = g .$$

Puesto que p divide a fg por hipótesis, dividirá a los dos sumandos del primer miembro, y por tanto dividirá a la suma g. ■

Corolario. Si $p \in \mathbb{F}[x]$ es irreducible y divide a $f_1 \ldots f_k$ donde $f_i \in \mathbb{F}[x]$, $i = 1, \ldots, k$, entonces p divide por lo menos a uno de los f_i.

Demostración. Según el teorema anterior p divide a f_1 o a $f_2 \ldots f_k$. Si no divide a f_1, entonces divide a f_2 o a $f_3 \ldots f_k$. Prosiguiendo de esta manera se encontrará un polinomio f_i que es divisible por p. ■

Teorema 10.5.3. Cada polinomio mónico no constante f de $\mathbb{F}[x]$ puede expresarse en la forma

$$f = p_1 \ldots p_m \qquad (m \geqslant 1)$$

donde los polinomios $p_1, \ldots, p_m$ son mónicos irreducibles.

La representación precedente es única salvo el orden de los factores.

La propiedad enunciada la expresaremos abreviadamente diciendo que cada polinomios admite una descomposición única en sus factores primos.

Demostración. Si f es irreducible, el teorema se cumple obviamente y la descomposición se reduce a un factor. Procederemos ahora por inducción respecto del grado de f. Si el grado de f es 1, es irreducible y admite la descomosición. Supongamos ahora que el teo-

rema se cumple para todos los polinomios de grado menor que n, y sea f un polinomio de grado n. Si fuera irreducible, sería descomponible en el sentido del enunciado. Si no es irreducible puede escribirse $f = gh$ donde g y h no son constantes. Se tendrá por tanto que *grado* $g < n$ y *grado* $h < n$. Por la hipótesis de inducción, podemos expresar g y h en las formas

$$g = g_1 \cdot \cdot \cdot g_k \quad , \quad h = h_1 \cdot \cdot \cdot h_m \quad ,$$

donde los g_i y los h_j son polinomios mónicos irreducibles. Sustituyendo esas expresiones de g y de h en la ecuación $f = gh$ se obtendrá la expresión de f como producto de polinomios mónicos irreducibles, lo cual demuestra la existencia de tal descomposición para todo polinomio de $\mathbb{F}[x]$.

Demostraremos ahora la unicidad de la factorización. Si *grado* $f = 1$, la unicidad es evidente pues f solo admite como factor mónico irreducible a sí mismo. Aplicaremos la inducción respecto del grado de f. Supongamos que la unicidad de la factorización se cumple para todos los polinomios mónicos de grado menor que n. Sea entonces f un polinomio mónico de grado $n > 1$ que admite dos factorizaciones:

$$f = p_1 \cdot \cdot \cdot p_m = q_1 \cdot \cdot \cdot q_k \qquad (10.5.1)$$

donde los p_i y los q_j son polinomios mónicos irreducibles. El polinomio irreducible p_1 divide al producto $q_1 \cdot \cdot \cdot q_k$, luego, en virtud del corolario del teorema 10.5.1, divide a uno de los polinomios $q_1, \ldots, q_k$. Modificando, si fuera necesario, la ordenación de estos polinomios, podemos suponer que p_1 divide a q_1. Puesto que ambos son mónicos e irreducibles, se deduce que $p_1 = q_1$, y por consiguiente, aplicando a (10.5.1) la ley de cancelación [1] resulta que

$$\overline{f} = p_2 \cdot \cdot \cdot p_m = q_2 \cdot \cdot \cdot q_k \quad ,$$

donde $\overline{f}$ es un polinomio mónico de grado menor que n. La hipótesis de inducción permite deducir que $m = k$ y que $p_2 = q_2, \ldots, p_m = q_m$ para una ordenación adecuada de los índices. Con lo cual queda demostrada la unicidad de la descomposición. ∎

1) Si $f \neq 0$ y $fg = fh$, entonces $f(g - h) = 0$, de donde: $g = h$.

Ejemplo. Sean $a_1, \ldots, a_n$, $n > 1$ elementos de $\mathbb{F}$ distintos entre sí. Vamos a probar que cualesquiera que sean los enteros positivos $k_1, \ldots, k_n$, los n polinomios

$$(x - a_1)^{k_1}, \ldots, (x - a_n)^{k_n}, \qquad (10.5.2)$$

son primos entre sí. Para ello basta demostrar que el *M.C.D.* de estos polinomios es *1*. Supongamos que d es ese *M.C.D.* y que $d \neq 1$. El polinomio d sería mónico y no constante; luego, por el teorema anterior, tendría por lo menos un factor p mónico irreducible. El polinomio p dividiría a $(x - a_1)^{k_1}$; pero los factores $(x - a_i)$ son irreducibles, luego p dividiría a $x - a_1$. Puesto que p y $x - a_1$ son mónicos irreducibles debería ser $p = x - a_1$. De la misma manera se podría probar que $p = x - a_2$ y resultaría que

$$x - a_1 = x - a_2 ,$$

lo cual contradice la hipótesis según la cual es $a_1 \neq a_2$. Esta contradicción permite deducir que $d = 1$, y por tanto que los polinomios (10.5.2) son primos entre sí.

10.6. Funciones polinomiales. Dado un polinomio cualquiera

$$f = \sum_{i=0}^{n} a_i x^i \in \mathbb{F}[x]$$

la aplicación de $\mathbb{F}$ en que hace corresponder a cada elemento $c \in \mathbb{F}$ el elemento

$$\sum_{i=0}^{n} a_i c^i \in \mathbb{F} ,$$

se llama *función polinomial o función racional entera* correspondiente a f. Se designa por f o por $f(x)$ y se escribe

$$f(c) = \sum_{i=0}^{n} a_i c^i .$$

$f(c)$ se llama *valor del polinomio f en c.*

Debe sin embargo tenerse cuidado de no confundir el polinomio f,

elemento de $\mathbb{F}[x]$ con la función $f(x)$, elemento del conjunto de las funciones de $\mathbb{F}$ en $\mathbb{F}$.

Teorema 10.6.1. Si c es un elemento dado cualquiera de $\mathbb{F}$, la aplicación de $\mathbb{F}[x]$ en $\mathbb{F}$ que aplica a f en $f(c)$ es un homomorfismo de anillos, es decir que

$$(f + g)(c) = f(c) + g(c) \quad \text{y} \quad (fg)(c) = f(c)g(c).$$

Demostración. Sean $f = \sum_{i=0}^{n} a_i x^i$ y $g = \sum_{j=0}^{m} b_j x^j$ dos polinomios cualesquiera. Nos limitaremos a demostrar la segunda relación del enunciado, pues la primera es inmediata.
Se tiene

$$(fg)(x) = \sum_{k=0}^{m+n} \Big(\sum_{i+j=0} a_i b_j \Big) c^k = \sum_{k=0}^{m+n} \Big(\sum_{i=0}^{n} a_i b_{k-i} c^k \Big)$$

$$= \sum_{i=1}^{n} \Big(\sum_{k=0}^{m+n} a_i b_{k-i} c^k \Big) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j c^{i+j}$$

$$= \Big(\sum_{i=0}^{n} a_i c^i \Big) \Big(\sum_{i=0}^{m} b_j c^j \Big) = f(c)\, g(c) \ . \qquad \blacksquare$$

Definición 10.6.1. Se dice que $c \in \mathbb{F}$ es una *raiz* o un *cero de la función polinomial* f *(o del polinomio* f*)* si se cumple la condición $f(c) = 0$. También se dice en ese caso que c es una *raiz de la ecuación* $f(x) = 0$.

No todo polinomio de $\mathbb{F}[x]$ tiene una raiz. Así por ejemplo el polinomio $x^2 + 1 = 0$ no tiene raiz cuando se le considera en $\mathbb{R}[x]$ por cuanto no hay número real que sea raiz de esa ecuación. Sí tiene, en cambio, raiz en $\mathbb{C}$. Veremos más tarde que, dado un campo $\mathbb{F}$ y un polinomio $f \in \mathbb{F}[x]$, siempre es posible la existencia de un nuevo campo $\mathbb{F}'$ que contiene a $\mathbb{F}$ de manera que $f \in \mathbb{F}'[x]$ y que contenga a una raiz de f. Un polinomio constante diferente de cero, es decir un polinomio de grado 0, no tiene raices.

Teorema 10.6.2. La condición necesaria y suficiente para que el polinomio $f \in \mathbb{F}[x]$ sea divisible por $x-c$, donde $c \in \mathbb{F}$, es que c sea un cero de f, es decir que $f(c) = 0$.

Demostración. Sea $f = (x - c)\, q + r$. Puesto que r es cero o bien es un polinomio constante que designaremos por r:

$$f = (x - c)\, q + r \ .$$

Si aplicamos el teorema 10.6.1, podremos escribir

$$f(c) = r \ .$$

Resulta así evidente que para que f sea divisible por $x - c$, es decir para que sea $r = 0$, es necesario y suficiente que sea

$$f(c) = 0. \qquad \blacksquare$$

Ejercicio

10.6.1. Demostrar que cada clase del anillo cociente $\mathbb{F}[x]\,/I(x - c)$, donde $c \in \mathbb{F}$, está formada por todos los polinomios $f \in \mathbb{F}[x]$ tales que las funciones polinomiales correspondientes $f(x)$ toman el mismo valor en c.

Expresado de otra manera,

$$f_1 - f_2 \in I(x - c) \Leftrightarrow f_1(c) = f_2(c)$$

Resulta así un isomorfismo entre los grupos $\mathbb{F}[x] \,/\, I(x - c)$ y $\mathbb{F}$ que debe ser explicitado por el lector.

10.7. Factorización en $\mathbb{R}[x]$ y en $\mathbb{C}[x]$.

Suponemos conocido el teorema fundamental del álgebra, según el cual todo polinomio de $\mathbb{C}[x]$ tiene por lo menos una raiz compleja.

Teorema 10.7.1. En $\mathbb{R}[x]$ cada polinomio irreducible es de grado 1 o 2.

Demostración. Sea el polinomio irreducible $f \in \mathbb{R}[x]$. Si tiene una raiz real a, deberá ser de la forma $c(x - a)$, $c \in \mathbb{R}$, en virtud del teorema 10.6.2. Si no tiene raiz real, tendrá necesariamente la raíz compleja $a + bi$, $b \neq 0$.

El polinomio de segundo grado con coeficientes reales

$$(x - (a + bi))(x - (a - bi)) = x^2 - 2ax + (a^2 + b^2)$$

tiene la raiz $a + bi$. Sea entonces

$$f = (x^2 - 2ax + (a^2 + b^2))\,q + r \qquad (r, q \in \mathbb{R}[x]) \ ,$$

donde r es el polinomio cero o bien un polinomio de primer grado. Reemplazando x por $a + bi$ en la ecuación precedente, empleando el teorema 10.6.1, se obtiene

$$r(a + bi) = 0.$$

Ahora bien, un polinomio de $\mathbb{R}[x]$ de primer grado o de grado cero no puede tener una raiz compleja. Luego r debe ser el polinomio cero, y q se reduce a una constante porque f es irreducible.

De lo expuesto resulta que un polinomio irreducible de $\mathbb{R}[x]$ es de grado *1* o de grado *2*. ■

Teorema 10.7.2. En $\mathbb{C}[x]$ cada polinomio irreducible es de grado *1*.

Demostración. Cada polinomio $f \in \mathbb{C}[x]$ tiene por lo menos una raiz $c \in \mathbb{C}$. Según el teorema 10.6.2, f será divisible por $x - c$, y existirá el polinomio $q \in \mathbb{C}[x]$ tal que

$$f = (x - c)\,q$$

Si *grado* $f > 1$, entonces $q(x)$ es un polinomio no constante y por consiguiente f no es irreducible. Por tanto los únicos polinomios irreducibles de $\mathbb{C}[x]$ son los polinomios de primer grado. ■

Corolario. Cada polinomio $f \in \mathbb{C}[x]$ puede ser factorizado en la forma

$$c(x - a_1)^{k_1} \cdots (x - a_m)^{k_m}$$

donde $a_1, \ldots, a_m$ son elementos de $\mathbb{C}$, distintos entre sí, $c \in \mathbb{C}$ y $k_1, \ldots, k_m$ son números enteros positivos.

Demostración. Emplearemos la inducción relativa al grado de f. Si f es de grado *1* la proposición es evidente. Supongamos que f es de grado n y que el teorema se cumple para los polinomios de grado menor que n. Sea $a \in \mathbb{C}$ una raiz de f. Se tendrá

$$f = (x - a)q \quad ,$$

donde q es de grado $n - 1$. Si se reemplaza q por su factorización en la forma del enunciado, se obtendrá para f una factorización de la misma clase. ■

Ejercicios

*10.7.1. Demostrar que los polinomios mónicos irreducibles de $\mathbb{R}[x]$ de segundo grado, son de la forma $(x + a)^2 + b^2$ donde $a,b \in \mathbb{R}$ y $b \neq 0$.

10.7.2. Obtener la factorización en $\mathbb{R}[x]$ y en $\mathbb{C}[x]$ de los polinomios $f = x^4 - 2x^3 + 2x^2 + 2x + 1$ y

$$g = x^4 - 2x^3 - 2x - 1 \ .$$

10.7.3. Hallar los polinomios irreducibles de tercer grado de $\mathbb{F}[x]$, donde $\mathbb{F}$ es el campo $\{0 , 1\}$ de clase de restos módulo 2

§ 11. Extensiones de campos

11.1. Extensiones de un campo.

Definición 11.1.1. Se llama, *extensión de un campo* $\mathbb{F}$ a otro campo $\mathbb{F}'$ tal que $\mathbb{F} \subset \mathbb{F}'$ y que $\mathbb{F}$ es subcampo de $\mathbb{F}'$.

Así, por ejemplo, el campo de los números complejos es una extensión del campo de los reales y éste lo es del de los números racionales.

Definición 11.1.2. Si $\mathbb{F}'$ es una extensión de $\mathbb{F}$ y θ es un elemento cualquiera de $\mathbb{F}'$, entonces, dado el polinomio

$$f = \sum_{i=0}^{n} a_i x^i$$

de $\mathbb{F}[x]$, se llama *valor de* f *en el elemento* θ *de* $\mathbb{F}'$ al elemento

$$f(\theta) = \sum_{i=0}^{n} a_i \theta^i$$

de $\mathbb{F}'$.

Teorema 11.1.1. Si $\mathbb{F}'$ es una extensión de $\mathbb{F}$ y θ es un elemento cualquiera de $\mathbb{F}'$, la aplicación $f \mapsto f(\theta)$ es un homomorfismo de $\mathbb{F}[x]$ en $\mathbb{F}'$, es decir que

$$(f + g)(\theta) = f(\theta) + g(\theta) \quad \text{y} \quad (fg)(\theta) = f(\theta)g(\theta) .$$

Demostración. Es análoga a la del teorema 10.6.1. ∎

Definición 11.1.3. Sea $\mathbb{F}'$ una extensión de $\mathbb{F}$. Se dice que un elemento $\theta \in \mathbb{F}'$ es *trascendente sobre* $\mathbb{F}$ si no existe ningún polinomio no nulo $f \in \mathbb{F}[x]$ tal que $f(\theta) = 0$, es decir, si θ no es raiz de ningún polinomio de $\mathbb{F}[x]$ salvo el polinomio cero. Esta condición es equivalente a que el núcleo I del homomorfismo $f \mapsto f(\theta)$ se reduce a $\{0\}$. También puede expresarse la condición diciendo que el homomorfismo $f \mapsto f(\theta)$ es inyectivo.

Definición 11.1.4. Sea $\mathbb{F}'$ una extensión de $\mathbb{F}$. Se dice que un elemento $\theta \in \mathbb{F}'$ *es algebraico sobre* $\mathbb{F}$ si existe un polinomio no nulo $f \in \mathbb{F}[x]$ tal que $f(\theta) = 0$, es decir si θ es la raiz de un polinomio no nulo de $\mathbb{F}[x]$. En ese caso el núcleo I del homomorfismo

$$f \mapsto f(\theta)$$

es diferente del ideal $\{0\}$ y dicho homomorfismo no es inyectivo. Si θ es algebraico sobre $\mathbb{F}$, el polinomio mónico de menor grado, f_o, tal que θ es raiz de f_o se llama *polinomio minimal de* θ. Es claro que f_o es el polinomio mónico generador del ideal I constituido por los polinomios de $\mathbb{F}[x]$ que se anula en el elemento θ de $\mathbb{F}'$.

Teorema 11.1.2. Los polinomios minimales de los elementos algebraicos sobre un campo $\mathbb{F}$ son irreducibles en $\mathbb{F}[x]$.

Demostración. Supongamos que, contrariamente a lo que se quiere demostrar, el polinomio minimal f_o del elemento algebraico θ sobre $\mathbb{F}$ puede expresaese en la forma

$$f_o = gh \qquad (11.1.1)$$

donde g y h son polinomios no constantes. Los dos polinomios g y h serán de grado menor que f_o ; luego se tendrá que $g(\theta) \neq 0$ y $h(\theta) \neq 0$, por la definición del polinomio minimal. Ahora bien, de

(11.1.1), aplicando el teorema 11.1.1, se deduce

$$g(\theta)h(\theta) = f_o(\theta) = 0.$$

Resulta así una contradicción pues $g(\theta)$ y $h(\theta)$ serían elementos no nulos de $\mathbb{F}$ cuyo producto es sin embargo nulo. ∎

Ejemplos. Como ejemplos de números trascendentes sobre el campo $\mathbb{Q}$ de los números racionales tenemos los números e y π. La trascendencia del primero fué demostrado por Charles Hermite (1822 - 1905), y la del segundo fué establecida por F. Lindemann en 1882. Como es sabido la demostración de la trascendencia de π resolvió definitivamente el problema de la cuadratura del círculo.

El número $\sqrt{2} \in \mathbb{R}$ es algebraico sobre $\mathbb{Q}$, por cuanto es raiz del polinomio $x^2 - 2$. Los polinomios de primer grado sobre $\mathbb{Q}$ son de la forma $x + a$ donde $a \in \mathbb{Q}$ y por tanto sus raices son necesariamente números racionales. Luego $\sqrt{2}$ no es raiz de ningún polinomio mónico de primer grado sobre $\mathbb{Q}$. Resulta así que el polinomio minimal de $\sqrt{2}$ es $x^2 - 2$.

Una extensión $\mathbb{F}'$ de un campo $\mathbb{F}$ puede ser considerada como un espacio vectorial sobre $\mathbb{F}$. Tanto la suma de dos elementos de $\mathbb{F}'$ como el producto de un elemento de $\mathbb{F}$ por un elemento de $\mathbb{F}'$ son dados por las respectivas operaciones en el campo $\mathbb{F}'$.

Definición 11.1.5. Una extensión $\mathbb{F}'$ de $\mathbb{F}$ se llama *extensión finita de* $\mathbb{F}$ si el espacio vectorial $\mathbb{F}'$ sobre $\mathbb{F}$ es de dimensión finita. Esa dimensión se llama *grado de* $\mathbb{F}'$ sobre $\mathbb{F}$ y se designa por $(\mathbb{F}' : \mathbb{F})$.

Teorema 11.1.3. Todo elemento perteneciente a una extensión finita $\mathbb{F}'$ de un campo $\mathbb{F}$ es algebraico sobre $\mathbb{F}$.

Demostración. Supongamos que el grado de $\mathbb{F}'$ sobre $\mathbb{F}$ es igual a n. Si θ es un elemento cualquiera de $\mathbb{F}'$, los $n + 1$ elementos

$$1 , \theta , \theta^2, \ldots , \theta^n$$

del espacio vectorial $\mathbb{F}'$ sobre $\mathbb{F}$ serán *l.d.* Por consiguiente existen elementos $a_o , a_1 , \ldots , a_n$ de $\mathbb{F}$, no todos nulos, tales que

$$\sum_{i=0}^{n} a_i \theta^i = 0.$$

Resulta así que θ es raiz del polinomio $\sum_{i=0}^{n} a_i x^i$ de $\mathbb{F}[x]$ y por tanto es algebraico sobre $\mathbb{F}$. ∎

Definición 11.1.6. Se dice que una *extensión* $\mathbb{F}'$ del campo $\mathbb{F}$ es *algebraica* si cada elemento de $\mathbb{F}'$ es algebraico sobre $\mathbb{F}$.

Según el teorema 11.1.3. toda extensión finita es algebraica.

Definición 11.1.7. La intersección de todos los subcampos del campo $\mathbb{F}$ que contienen a un subconjunto S de $\mathbb{F}$ es un subcampo de $\mathbb{F}$ que contiene a S, que se llama *subcampo de* $\mathbb{F}$ *generado por* S.

Definición 11.1.8. Sea $\mathbb{F}'$ una extensión del campo $\mathbb{F}$ y sea θ un elemento cualquiera de $\mathbb{F}'$. El subconjunto $\mathbb{F} \cup \{\theta\}$ de $\mathbb{F}'$ genera a un subcampo de $\mathbb{F}'$, que se designa por $\mathbb{F}(\theta)$, del cual se dice que resulta por *adjunción* de θ al campo $\mathbb{F}$. Se dice que θ es *generador de* $\mathbb{F}(\theta)$ sobre $\mathbb{F}$.

Definición 11.1.9. Se dice que un campo $\mathbb{F}'$ es *extensión simple de* $\mathbb{F}$ si resulta por adjunción a $\mathbb{F}$ de un elemento θ de $\mathbb{F}'$, es decir si $\mathbb{F}' = \mathbb{F}(\theta)$. Si θ es algebraico sobre $\mathbb{F}$, entonces se dice que $\mathbb{F}'$ es *extensión simple algebraica de* $\mathbb{F}$. Si θ es trascendente sobre $\mathbb{F}$, entonces se dice que $\mathbb{F}$ es *extensión simple trascendente de* $\mathbb{F}$.

El teorema siguiente justifica la denominación de extensión simple algebraica que hemos dado a la extensión que resulta por adjunción de un elemento algebraico, pues tal extensión es algebraica en el sentido de la definición 11.1.4.

Teorema 11.1.4. Sea $\mathbb{F}'$ una extensión del campo $\mathbb{F}$ y $\mathbb{F}(\theta)$ la extensión de $\mathbb{F}$ que resulta por adjunción del elemento $\theta \in \mathbb{F}'$ algebraico sobre $\mathbb{F}$. Entonces $\mathbb{F}(\theta)$ está formado por todos los elementos de la forma $f(\theta)$ donde f es un polinomio cualquiera de $\mathbb{F}[x]$, cuyo grado es menor que el grado n del polinomio minimal de θ. Además $\mathbb{F}(\theta)$ es una extensión finita de $\mathbb{F}$ que posee la base

$$(1, \theta, \ldots, \theta^{n-1}),$$

y por tanto es una extensión algebraica cuyo grado es $(\mathbb{F}(\theta) : \mathbb{F}) = n$.

Demostración. Sea f el polinomio minimal de θ, cuyo grado es n.

Vamos a demostrar que el conjunto

$$\mathbb{F}^* = \{f(\theta) \mid f \in \mathbb{F}[x] \text{ , grado } f < n\} \cup \{0\}$$

coincide con $\mathbb{F}(\theta)$, como lo afirma la primera parte del enunciado. Es evidente que $\mathbb{F}^* \subset \mathbb{F}(\theta)$ y que $\mathbb{F} \subset \mathbb{F}^*$. Esto último porque todo elemento a de $\mathbb{F}$ puede expresarse en la forma $f(\theta)$ si f es el polinomio constante a.

Vamos a probar ahora que $\mathbb{F}^*$ es un subcampo de $\mathbb{F}'$.

Los elementos 0 y 1 pertenecen a $\mathbb{F}^*$ porque $\mathbb{F} \subset \mathbb{F}^*$

Dados los elementos cualesquiera $f(\theta)$ y $g(\theta)$ de $\mathbb{F}^*$ se tiene

$$f(\theta) + g(\theta) = (f + g)(\theta)$$
$$f(\theta) - g(\theta) = (f - g)(\theta)$$

donde $f + g$ y $f - g$ son polinomios de grado $< n$. Luego tanto $f(\theta) + g(\theta)$ como $f(\theta) - g(\theta)$ son elementos de $\mathbb{F}^*$.

Por otra parte puede escribirse

$$f(x)g(x) = f_o(x)\,q(x) + r(x)$$

donde es $r(x) \equiv 0$, o bien *grado* $r <$ *grado* $f_o = n$. Luego

$$f(\theta)g(\theta) = r(\theta) \text{ ,}$$

pues $f_o(\theta) = 0$. Si $r = 0$, entonces $r(\theta) = 0$ y por tanto

$$f(\theta)g(\theta) = 0 \in \mathbb{F}^* \text{ ;}$$

y si $r \neq 0$ y *grado* $r < n$, entonces

$$f(\theta)g(\theta) = r(\theta)$$

pertenecerá también a $\mathbb{F}^*$.

En todo caso, pues, el producto de dos elementos de $\mathbb{F}^*$ es igual a un elemento de $\mathbb{F}^*$.

Solo falta probar que la inversa de un elemento de $\mathbb{F}^*$ que no es nulo, pertenece a $\mathbb{F}^*$ Sea $f(\theta) \neq 0$, *grado* $f < n$, un elemento no nulo de $\mathbb{F}^*$. Puesto que el polinomio minimal f_o de θ es irreducible (teorema 11.1.2), se tiene que $[f_o, f] = 1$, pues f_o no tiene más divisores mónicos que f_o y 1 y f no es divisible por f_o ,por

que *grado* $f <$ *grado* f_o .

Por tanto existen los polinomios $g, g_o \in \mathbb{F}[x]$ tales que

$$fg + f_o g_o = 1 \ ;$$

de donde, sustituyendo x por θ, resulta

$$f(\theta)g(\theta) = 1 \ . \qquad (11.1.2)$$

Si *grado* $g < n$, entonces es claro que $g(\theta) \in \mathbb{F}^*$ y que $g(\theta)$ es la inversa de $f(\theta)$ en $\mathbb{F}^*$. Si *grado* $g \geqslant n$, entonces sea en $\mathbb{F}[x]$,

$$g = f_o q + r \ , \qquad (11.1.3)$$

donde necesariamente debe ser $r \neq 0$, pues de lo contrario sería $g(\theta) = 0$ en contradicción con (11.1.2). Además es

$$\text{grado } r < \text{grado } f_o = n \ .$$

De (11.1.3) se deduce

$$g(\theta) = r(\theta) \in \mathbb{F}^* \ .$$

Por tanto, también en este caso, resulta de (11.1.2) que $g(\theta)$ es la inversa de $f(\theta)$ en $\mathbb{F}^*$.

Con lo demostrado anteriormente queda establecido que $\mathbb{F}^*$ es un subcampo de $\mathbb{F}'$. Ahora bien, este subcampo contiene a θ y está contenido en $\mathbb{F}(\theta)$ como vimos anteriormente. Luego, por definición de $\mathbb{F}(\theta)$ debera ser

$$\mathbb{F}^* = \mathbb{F}(\theta) \ ,$$

lo cual prueba la primera parte del teorema.

La segunda parte del teorema resulta fácilmente de la primera. Puesto que cada elemento de $\mathbb{F}(\theta)$ es de la forma $f(\theta)$, donde $f \in \mathbb{F}[x]$ y *grado* $f < n$, se sigue que los n elementos $1, \theta, \theta^2, \ldots, \theta^{n-1}$ generan al espacio vectorial $\mathbb{F}(\theta)$ sobre $\mathbb{F}'$. Para probar que

$$\dim \mathbb{F}(\theta) = n \ ,$$

basta probar que esos elementos son *l.i.* En efecto, en caso contrario existirían los elementos $a_1, \ldots, a_{n-1}$ de $\mathbb{F}$, no todos nulos, tales que

$$a_o + a_1\theta + \cdots + a_n \theta^{n-1} = 0.$$

θ sería entonces raiz del polinomio no trivial.

$$a_0 + a_1 x + \cdots + a_n x^{n-1} \in \mathbb{F}[x] ,$$

de *grado* $< n$. Esto contradice la suposición de que el polinomio minimal f_0 de θ es de *grado* n. Se tiene por tanto que

$$(1, \theta, \ldots, \theta^{n-1})$$

es una base de $\mathbb{F}(\theta)$ y por tanto $(\mathbb{F}(\theta) : \mathbb{F}) = n$. ■

Corolario. Si $\mathbb{F}'$ es una extensión de $\mathbb{F}$ y $\theta \in \mathbb{F}'$ es algebraico sobre $\mathbb{F}$, el campo $\mathbb{F}(\theta)$ es isomorfo con el campo $\mathbb{F}[x] / I(f_0)$, donde f_0 es el polinomio minimal de θ.

Demostración. La imagen de $\mathbb{F}[x]$ por el homomorfismo $f \mapsto f(\theta)$ definido en el teorema 11.1.4 es, en virtud del teorema anterior, el campo $\mathbb{F}(\theta)$. El núcleo de ese homomorfismo es el ideal de los polinomios que tienen la raiz θ, o sea el ideal principal generado por el polinomio minimal f_0 de θ. Tiene lugar por tanto un isomorfismo canónico del campo $\mathbb{F}(\theta)$ sobre el campo $\mathbb{F}[x] / I(f_0)$. Por ese isomorfismo, a cada elemento $\alpha \in \mathbb{F}(\theta)$ le corresponde una clase de polinomios congruentes entre sí módulo f_0. ■

11.2. Construcción de extensiones simples algebraicas.

En esta sección vamos a demostrar que dado un campo $\mathbb{F}$ y un polinomio $f \in \mathbb{F}[x]$ que no posee una raiz perteneciente a $\mathbb{F}$, es posible construir una extensión $\mathbb{F}^*$ de $\mathbb{F}$ que contenga por lo menos una raiz θ del polinomio f. Bastará que llevemos a cabo esa construcción en la hipótesis de que f es irreducible, pues en caso contrario, si f_1 fuera uno de sus divisores irreducibles, entonces una raiz θ de f_1 que perteneciera a una extensión $\mathbb{F}'$ de $\mathbb{F}$ que pudiera construirse, sería también una raiz de f.

En la presente sección desempeñarán un papel principal las consideraciones del parágrafo § 9 y, en especial, de la sección 9.4.

Teorema 11.2.1. Sea f un polinomio mónico irreducible de $\mathbb{F}[x]$. Entonces $\mathbb{F}^* = \mathbb{F}[x] / I(f)$ es un campo. Además, $\mathbb{F}^*$ es una extensión simple algebraica de $\mathbb{F}$, $\mathbb{F}(\theta)$, donde θ es una raiz de la ecuación $f(x) = 0$, y f es el polinomio minimal de θ. Además, cual-

quier extensión simple algebraica $\mathbb{F}(\theta')$, donde θ' es raiz de f es isomorfa de $\mathbb{F}^*$.

Demostración. $\mathbb{F}^* = \mathbb{F}[x] / I(f)$ es un campo en virtud del teorema 10.5.1. Sea π la proyección (definición 9.4.2) de $\mathbb{F}[x]$ sobre $\mathbb{F}^*$ y designemos por θ a la imagen de $\pi(x)$ del polinomio x, es decir a la clase de $\mathbb{F}^*$ que contiene a x.

Vamos a observar en primer lugar que $\mathbb{F}^*$ puede ser considerado como una extensión de $\mathbb{F}$. En efecto, la restricción del homomorfismo π al campo $\mathbb{F}$ contenido en $\mathbb{F}[x]$ es una inyección pues

$$\pi(a) = \pi(b) \text{ , para } a,b \in \mathbb{F} \text{ ,}$$

implica que $a - b \in I(f)$, y por tanto $a = b$, porque $I(f)$ no contiene otro polinomio constante que el polinomio cero (los polinomios de $I(f)$ son múltiplos de f, y f no es constante porque es irreducible). Resulta que la restricción de π al campo $\mathbb{F}$ es un isomorfismo de $\mathbb{F}$ sobre el subcampo $\pi(\mathbb{F})$ de $\mathbb{F}^*$. Por consiguiente podemos identificar a cada elemento $\pi(a) \in \mathbb{F}^*$, donde $a \in \mathbb{F}$, con el escalar a (notar que $\pi(a)$ es el conjunto de los polinomios de la forma

$$gf + a \text{ ,}$$

donde g es un polinomio cualquiera de $\mathbb{F}[x]$), y resulta así que puede suponerse que $\mathbb{F}^*$ contiene al subcampo $\mathbb{F}$, y por tanto que $\mathbb{F}^*$ es una extensión de $\mathbb{F}$.

Si $g = \sum_{i=0}^{n} a_i x^i$ es un polinomio cualquiera de $\mathbb{F}[x]$, y se tiene presente que $\pi(x) = \theta$ y que $\pi(a_i) = a_i$, se obtiene

$$\pi(g) = \sum_{i=0}^{n} a_i \theta^i = g(\theta) \text{ .}$$

En particular, y puesto que $f \in I(f)$, y por tanto que $\pi(f) = 0 \in \mathbb{F}^*$ se tendrá que

$$f(\theta) = 0,$$

es decir que θ es raiz de $f(x)$, o sea que es un elemento de $\mathbb{F}^*$, algebraico sobre $\mathbb{F}$. Además podemos comprobar que el polinomio minimal de θ es f. En efecto, supongamos que fuera $g(\theta) = 0$ para un polinomio $g \in \mathbb{F}[x]$ de grado menor que el de f. Se tendrá en

tonces $\pi(g) = g(\theta) = 0$, y por tanto se tendría que $g \in I(f)$, es decir que f dividiría a g, en contradicción con la suposición de que *grado* $g <$ *grado* f.

Solo nos falta probar que $\mathbb{F}^* = \mathbb{F}(\theta)$. Puesto que $\pi : \mathbb{F}[x] \to \mathbb{F}^*$ es una sobreyección, cada elemento de $\mathbb{F}^*$ es de la forma $\pi(h)$ donde $h \in \mathbb{F}[x]$. Se tiene entonces

$$\pi(h) = h(\theta) \in \mathbb{F}(\theta).$$

Luego $\mathbb{F}^* \subset \mathbb{F}(\theta)$. Pero $\mathbb{F}(\theta) \subset \mathbb{F}^*$ porque $\mathbb{F} \subset \mathbb{F}^*$ y $\theta \in \mathbb{F}^*$; por consiguiente $\mathbb{F}^* = \mathbb{F}(\theta)$.

Si $\mathbb{F}(\theta')$ es cualquier otra extensión simple algebraica de $\mathbb{F}$, tal que θ' es raiz de f, el corolario del teorema 11.1.4 implica que $\mathbb{F}(\theta')$ es isomorfo con $\mathbb{F}^* = \mathbb{F}[x] / I(f)$. ∎

Ejemplo. El polinomio $f = x^2 + 1$ de $\mathbb{R}[x]$ es irreducible porque no posee ninguna raiz real. Por tanto $\mathbb{F}^* = \mathbb{R}[x] / I(f)$ será un campo, extensión de $\mathbb{R}$, que es generado por el elemento θ, cuyo polinomio minimal es $x^2 + 1$. Vamos a demostrar que $\mathbb{F}^*$ es isomorfo con el campo $\mathbb{C}$ de los números complejos.

Según el teorema 11.1.4, los elementos de $\mathbb{F}^*$ son de la forma $a + b\theta$ donde a y b son números reales unívocamente determinados.

Podemos establecer ahora entre $\mathbb{F}^*$ y $\mathbb{C}$ la biyección

$$\phi : a + b\theta \to a + bi \quad , \quad (a,b \in \mathbb{R}) .$$

Esta biyección es un isomorfismo, pues

$$\begin{aligned} \phi[(a + b\theta) + (a_1 + b_1\theta)] &= \phi[(a + a_1) + (b + b_1)\theta] \\ &= (a + a_1) + (b + b_1)i \\ &= (a + bi) + (a_1 + b_1 i) \\ &= \phi[a + b\theta] + \phi[a_1 + b_1\theta] , \end{aligned}$$

y además, teniendo presente que $\theta^2 = -1$,

$$\phi[(a + b\theta)(a_1 + b_1\theta)] = \phi[aa_1 + bb_1\theta^2 + ab_1\theta + ba_1\theta]$$

$$= \phi[(aa_1 - bb_1) + (ab_1 + ba_1)\theta]$$
$$= (aa_1 - bb_1) + (ab_1 + ba_1)i$$
$$= (a + bi)(a_1 + b_1 i)$$
$$= \phi[a + b\theta]\phi[a_1 + b_1\theta] .$$

11.3. Campos algebraicamente cerrados.

Definición 11.3.1. Un campo se llama *algebraicamente cerrado* si cada polinomio $f(x)$ de $\mathbb{F}[x]$ tiene una raiz en $\mathbb{F}$.

El teorema fundamental del álgebra establece que el campo $\mathbb{C}$ de los números complejos es algebraicamente cerrado. En ese hecho se fundan el teorema 10.7.2 y su corolario. Haciendo uso de un procedimiento idéntico al que fué empleado para probar esas proposiciones, y teniendo en cuenta el teorema 10.6.2, se prueba en general que un campo $\mathbb{F}'$ es algebraicamente cerrado si y solamente si cada polinomio irreducible de $\mathbb{F}[x]$ es de grado *1*, o bien si cada polinomio de $\mathbb{F}[x]$ puede expresarse como un producto de polinomios de grado *1*.

Un teorema importante, que aquí nos limitaremos a enunciar, establece que cada campo $\mathbb{F}$ tiene una extensión algebraica $\mathbb{F}^*$ que es algebraicamente cerrada, la cual es única, salvo un isomorfismo. $\mathbb{F}^*$ se llama entonces *cerradura algebraica de* $\mathbb{F}$.[1] Así, por ejemplo, la cerradura algebraica del campo $\mathbb{R}$ de los números reales es el campo $\mathbb{C}$ de los números complejos, o un campo isomorfo con él.

11.4. Construcción de extensiones simples trascendentes.

En esta sección también emplearemos las consideraciones del parágrafo § 9, principalmente las de la sección 9.5.

Recordemos que una extensión trascendente de un campo $\mathbb{F}$ es una extensión $\mathbb{F}(\theta)$ que se obtiene de $\mathbb{F}$ por adjunción de un elemento θ, trascendente sobre $\mathbb{F}$, es decir de un elemento θ perteneciente

1) La demostración del teorema cuando $\mathbb{F}$ es enumerable puede verse en Van der Waerden, Algebra I, Springer (1937), pág 203. Para el caso general, la demostración hace uso del lema de Zorn. Puede verse esa demostración en Bourbaki, Algèbre, cap. 5, Corps Commutatifs pág 91, (1959), o Lang, Algebra, pág. 169.

a una extensión $\mathbb{F}'$ de $\mathbb{F}$, que no es raiz de ningún polinomio no nulo de $\mathbb{F}[x]$.

$\mathbb{F}[x]$ es un anillo de integridad (teorema 10.1.1), por consiguiente existe su *campo de cocientes* $\mathbb{F}'$ (definición 9.5.2), el cual contiene a $\mathbb{F}[x]$, y cuyos elementos pueden expresarse en la forma

$$fg^{-1} = g^{-1}f = f/g \qquad \text{(ver sección 9.5)}$$

donde f y g, $g \neq 0$, son polinomios cualesquiera de $\mathbb{F}[x]$.

Ahora bien, el subcampo, $\mathbb{F}(x)$ (nótese aquí que empleamos parentesis en vez de corchetes) de $\mathbb{F}'$, que resulta por adjunción de $x \in \mathbb{F}'$ al campo $\mathbb{F}$, contiene a los polinomios f de $\mathbb{F}[x]$ y a las inversas g^{-1} de los polinomios no nulos del mismo anillo; luego contiene a todos los elementos de la forma fg^{-1}, y por tanto es

$$\mathbb{F}' \subset \mathbb{F}(x).$$

Pero $\mathbb{F}(x) \subset \mathbb{F}'$ porque $\mathbb{F}(x)$ es subcampo de $\mathbb{F}'$ luego se tiene que $\mathbb{F}' = \mathbb{F}(x)$; es decir que el campo de cocientes $\mathbb{F}'$ de $\mathbb{F}[x]$ coincide con el subcampo $\mathbb{F}(x)$ de $\mathbb{F}'$ que resulta por adjunción de x al campo $\mathbb{F}$. Podemos, por tanto, dar la definición siguiente.

Definición 11.4.1. El campo de cocientes del anillo de polinomio $\mathbb{F}[x]$ se llama *campo de las funciones racionales sobre* $\mathbb{F}$ y coincide con el campo $\mathbb{F}(x)$ que resulta por adjunción del polinomio $x \in \mathbb{F}'$ al campo $\mathbb{F}$.

Teorema 11.4.1. El campo de las funciones racionales sobre $\mathbb{F}$ es una extensión simple trascendente del campo $\mathbb{F}$. Además si

$$\mathbb{F}^* = \mathbb{F}(\theta)$$

es una extensión simple trascendente de $\mathbb{F}$, entonces $\mathbb{F}^*$ es isomorfo con el campo $\mathbb{F}(x)$ de las funciones racionales sobre $\mathbb{F}$; mas precisamente, existe un isomorfismo ϕ de $\mathbb{F}^*$ sobre $\mathbb{F}(x)$ tal que

$$\phi(\theta) = x \quad \text{y} \quad \phi(a) = a \text{, para cada } a \in \mathbb{F}.$$

Demostración. Hemos visto antes de establecer la definición 11.4.1

que el campo de las funciones racionales sobre $\mathbb{F}$ es la extensión simple de $\mathbb{F}$ que resulta por adjunción a este último campo del polinomio x perteneciente a la extensión de $\mathbb{F}$ constituida por el campo de cocientes $\mathbb{F}'$ de $\mathbb{F}[x]$ ' Es decir que ese campo de funciones racionales es el campo $\mathbb{F}(x)$, que además coincide con $\mathbb{F}'$. Para probar que $\mathbb{F}(x)$ es extensión simple trascendente de $\mathbb{F}$ solo falta probar, según la definición 11.1.3, que ningún polinomio de $\mathbb{F}[x]$, salvo el polinomio nulo, se anula en el elemento x de $\mathbb{F}(x)$. Ahora bien el homomorfismo de $\mathbb{F}[x]$ en $\mathbb{F}(x)$ que hace corresponder a cada polinomio $f \in \mathbb{F}[x]$ el valor de f en el elemento x, (ver el teorema 11.1.1) es evidentemente la aplicación idéntica. Luego el único polinomio de $\mathbb{F}[x]$ que se aplica sobre cero por ese homomorfismo, es decir el único polinomio que tiene la raiz x, es el polinomio cero. Por consiguiente x es trascendente sobre $\mathbb{F}$.

Supongamos ahora que $\mathbb{F}^* = \mathbb{F}(\theta)$ es una extensión simple trascendente de $\mathbb{F}$; es decir (definiciones 11.1.5 y 11.1.6) que $\mathbb{F}^*$ es un campo que es extensión de $\mathbb{F}$, que contiene al elemento θ trascendente sobre $\mathbb{F}$, y que es generado por $\mathbb{F} \cup \{\theta\}$, de manera que no contiene ningún subcampo propio que contenga a $\mathbb{F}$ y a θ.

Vamos a demostrar en primer lugar que $\mathbb{F}^*$ coincide con el campo $\overline{\mathbb{F}^*}$ cuyos elementos se expresan en la forma

$$f(\theta)[g(\theta)]^{-1} = \frac{f(\theta)}{g(\theta)} \quad , \text{ donde } \quad f,g \in \mathbb{F}[x] \quad \text{y} \quad g \neq 0.$$

(Notar que $g(\theta) \neq 0$ pues θ es trascendente sobre $\mathbb{F}$),

Probaremos que $\overline{\mathbb{F}^*}$ es un subcampo de $\mathbb{F}^*$. En efecto:
0 y 1 pertenecen a $\overline{\mathbb{F}^*}$, porque es claro que $\mathbb{F} \subset \overline{\mathbb{F}^*}$

Si $f(\theta)[g(\theta)]^{-1}$ es un elemento cualquiera de $\overline{\mathbb{F}^*}$, diferente de cero, entonces es $f(\theta) \neq 0$, y la inversa de dicho elemento es

$$g(\theta)[f(\theta)]^{-1} \in \overline{\mathbb{F}^*}$$

Por último,

$$f(\theta)[g(\theta)]^{-1} \pm f_1(\theta)[g_1(\theta)]^{-1} = [f(\theta)g_1(\theta) \pm f_1(\theta)g(\theta)][g(\theta)g_1(\theta)]^{-1} \in \mathbb{F}^*$$

y

$$\{f(\theta)[g(\theta)]^{-1}\}\{f_1(\theta)[g_1(\theta)]^{-1}\} = [f(\theta)f_1(\theta)][g(\theta)g_1(\theta)]^{-1} \in \mathbb{F}^*$$

Puesto que $\overline{\mathbb{F}^*}$ es un subcampo de $\mathbb{F}^*$ que contiene a $\mathbb{F}$ y a θ, y que $\mathbb{F}^*$ es el menor campo que cumple esas condiciones se tiene que $\mathbb{F}^* = \overline{\mathbb{F}^*}$.

Consideremos ahora la aplicación $\psi : \mathbb{F}(x) \to \mathbb{F}^*$ que aplica a cada elemento $r = f/g$ $(g \neq 0)$ de $\mathbb{F}(x)$ sobre el elemento

$$f(\theta)/g(\theta) \quad \text{de } \mathbb{F}^*.$$

Es facil ver que ψ es un isomorfismo de $\mathbb{F}(x)$ sobre $\mathbb{F}^*$ que aplica a x en θ y a cada elemento de $\mathbb{F}^*$ sobre si mosmo[1]. La inversa $\phi = \psi^{-1}$ de ese isomorfismo es el isomorfismo de $\mathbb{F}^*$ sobre $\mathbb{F}[x]$ que tiene las propiedades establecidas en el enunciado. ■

Ejercicios

***11.4.1.** Sea $\mathbb{F}'$ una extensión de $\mathbb{F}$ y S un subconjunto de $\mathbb{F}'$. Demostrar que el subcampo de $\mathbb{F}'$ generado por $\mathbb{F} \cup S$ es el conjunto $\mathbb{F}(S)$ de todas las expresiones recionales de elementos de $\mathbb{F}$ y de S, es decir de todos los elementos de $\mathbb{F}'$ que se obtienen por adición, sustracción, multiplicación y división de elementos de $\mathbb{F}$ y de S. $\mathbb{F}(S)$ se dice que resulta de $\mathbb{F}$ por *adjunción de* S. Los elementos de $\mathbb{F}(S)$ son por tanto de la forma

$$\frac{P_1(s_1, \ldots, s_m)}{P_2(s'_1, \ldots, s'_n)}$$

donde P_1 y P_2 son polinomios cualesquiera sobre $\mathbb{F}$, $s_1, \ldots, s_m, s'_1, \ldots, s'_n$ son elementos arbitrarios de S y cumple la condición $P_2(s'_1, \ldots, s'_n) \neq 0$. Si S consta de un único elemento θ, entonces $\mathbb{F}(S)$ se reduce a $\mathbb{F}(\theta)$ y se llama extensión simple (definición 11.1.7).

---

1) Para demostrar que la aplicación ψ es bien definida y que es un isomorfismo deberá tenerse presente que la relaciones $f(\theta)g_1(\theta) - f_1(\theta)g(\theta) = 0$ y $fg_1 - f_1g = 0$ son equivalentes debido a que θ es trascendente sobre $\mathbb{F}$.

*11.4.2. Con las mismas notaciones del ejercicio anterior, si

$$S = \{a_1, \ldots, a_n\}$$

es finito, entonces $\mathbb{F}(S) = \mathbb{F}(a_1, \ldots, a_n)$ es el campo

$$\mathbb{F}(a_1)(a_2) \ldots (a_n)$$

que se obtiene por la adjunción sucesiva de $a_1, \ldots, a_n$.

En general, si $S = S_1 \cup S_2$, entonces

$$\mathbb{F}(S) = \mathbb{F}(S_1)(S_2),$$

es decir que $\mathbb{F}(S)$ resulta por adjunción sucesiva de S_1 y S_2

11.4.3. Hallar un elemento generador y su polinomio minimal

a) del campo de los números complejos sobre el campo de los números reales

b) del campo $\mathbb{Q}(\sqrt[5]{3})$ sobre el campo $\mathbb{Q}$ de los números racionales

c) del campo $\mathbb{Q}(e^{\frac{2\pi i}{5}})$ sobre el campo $\mathbb{Q}$ de los números racionales.

11.4.4. Los números complejos $a + bi$ donde a y b son números racionales enteros se llaman *enteros de Gauss* y constituyen un anillo que designaremos por $\mathbb{Z}[i]$, el cual contiene al anillo $\mathbb{Z}$ de los enteros.

Demostrar que $\mathbb{Z}[i]/(7)$, donde (7) es el ideal principal generado por 7, es un campo cuyo campo primo (ver 11.4.4) es isomorfo con $\mathbb{Z}_{/(7)}$, de modo que puede suponerse que $\mathbb{Z}_{/(7)}$ está contenido en $\mathbb{Z}[i]_{/(7)}$. Determinar un elemento generador de $\mathbb{Z}[i]_{/(7)}$ sobre $\mathbb{Z}_{/(7)}$ y el polinomio minimal correspondiente.

11.4.5. Sea $\mathbb{F}$ el campo $\{0, 1, 2\}$ de clases de restos módulo 3. Determinar la función polinomial que corresponda al polinomio $x^2 + 2x \in \mathbb{F}[x]$, y obtener la factorización de este polinomio en factores irreducibles.

11.4.6. $\mathbb{Q}$ es el campo de los racionales. Probar que $\sqrt{3}$ es algebraico sobre $\mathbb{Q}$. ¿Cuál es el grado de $\mathbb{Q}(\sqrt{3})$ sobre $\mathbb{Q}$?

11.4.7. Probar que $\sqrt{a} + \sqrt{b}$ donde a y b son enteros positivos que no son cuadrados perfectos, es algebraico sobre $\mathbb{Q}$. Hallar además su polinomio minimal.

*11.4.8. Probar que si $\mathbb{F}''$ es extensión finita de $\mathbb{F}'$ y $\mathbb{F}'$ es extensión finita de $\mathbb{F}$, entonces $\mathbb{F}''$ es extensión finita de $\mathbb{F}$ y

$$(\mathbb{F}'' : \mathbb{F}) = (\mathbb{F}'' : \mathbb{F}')(\mathbb{F}' : \mathbb{F}).$$

(Se sugiere considerar una base $\{u_1, \ldots, u_r\}$ de $\mathbb{F}'$ respecto de $\mathbb{F}$ y una base $\{v_1, \ldots, v_s\}$ de $\mathbb{F}''$ respecto de $\mathbb{F}'$ y demostrar que los elementos $v_i u_j$ $(i = 1,\ldots,s\ ;\ j = 1,\ldots,r)$ constituyen una base de $\mathbb{F}''$ respecto de $\mathbb{F}$).

*11.4.9. Sea $\mathbb{F}'$ una extensión de $\mathbb{F}$. Probar que el conjunto de todos los elementos de $\mathbb{F}'$ que son algebraicos sobre $\mathbb{F}$ es un subcampo de $\mathbb{F}'$ que contiene a $\mathbb{F}$.

Ejemplo: el subcampo de $\mathbb{C}$ formado por los números algebraicos sobre $\mathbb{Q}$, *es el campo de los números algebraicos.* El subcampo de $\mathbb{R}$ formado por los números algebraicos sobre $\mathbb{Q}$ es el *campo de los números reales algebraicos.*

*11.4.10. Probar que si $\mathbb{F}''$ es extensión algebraica de $\mathbb{F}'$ y $\mathbb{F}'$ es extensión algebraica de $\mathbb{F}$, entonces $\mathbb{F}''$ es extensión algebraica de $\mathbb{F}$.

*11.4.11. Sea $\mathbb{F}'$ una extensión del campo $\mathbb{F}$. Probar que si un sistema de ecuaciones lineales sobre $\mathbb{F}$

$$\sum_{j=1}^{m} \alpha_j^i \xi^j = \eta^i \quad , \quad (i = 1, \ldots, n)$$

tiene una solución en $\mathbb{F}'$, entonces tiene una solución en $\mathbb{F}$.

11.4.12. ¿Cuál es el grado del campo $\mathbb{Q}(i\ ,\ \sqrt{2})$ respecto del campo $\mathbb{Q}$ de los números racionales?

§ 12. Extensión de un espacio vectorial por ampliación del campo de escalares

12.1. Introducción. Transformaciones $\mathbb{F}$-lineales.

Si $\mathbb{F}'$ es una extensión de un campo $\mathbb{F}$ y W' es un espacio vectorial sobre $\mathbb{F}'$, es evidente que el mismo conjunto W' puede considerarse como un espacio vectorial sobre $\mathbb{F}$, respecto de las operaciones siguientes:

1. La misma adición que en el espacio vectorial dado W',
2. un producto de los escalares de $\mathbb{F}$ por elementos de W' que es el mismo que tiene lugar en el espacio dado W' por cuanto $\mathbb{F} \subset \mathbb{F}'$.

Cuando sea necesario distinguir el espacio W' sobre $\mathbb{F}'$ del espacio W' sobre $\mathbb{F}$ que acabamos de definir, designaremos al primero por $W'_{\mathbb{F}'}$, y al segundo por $W'_{\mathbb{F}}$, y diremos que $W'_{\mathbb{F}}$ resulta de $W'_{\mathbb{F}'}$ por *restricción* del campo de escalares de $\mathbb{F}'$ a $\mathbb{F}$.

Ejemplo. El conjunto $\mathbb{C}$ de los números complejos puede considerarse como un espacio vectorial sobre $\mathbb{C}$ dimensión *1*. $\mathbb{C}_{\mathbb{R}}$ que resulta por restricción del campo de escalares al campo $\mathbb{R}$ de los números reales, es en cambio un espacio vectorial sobre $\mathbb{R}$ de dimensión 2.

Vamos a considerar ahora la situación inversa de aquella que hemos considerado al comienzo de esta sección. Sea V un espacio vectorial sobre $\mathbb{F}$ y sea $\mathbb{F}'$ una extensión del campo $\mathbb{F}$. Nos proponemos construir un espacio vectorial V' sobre $\mathbb{F}'$ que pueda ser considerado, en algún sentido, como la extensión natural de V cuando el campo escalar se extiende de $\mathbb{F}$ a $\mathbb{F}'$. Con tal fin trataremos de definir el espacio V' sobre $\mathbb{F}'$ de manera que:

a) contenga a V como subespacio vectorial de $V'_{\mathbb{F}}$, es decir que resulta de V' por restricción del campo de escalares de $\mathbb{F}'$ a $\mathbb{F}$;

b) que V' sea generado por V, es decir que cada elemento x' de V' pueda expresarse en la forma

$$x' = \sum_{i=1}^{p} \lambda^i x_i \quad , \quad x_i \in V \, , \, \lambda^i \in \mathbb{F}' \, ,$$

de manera que no exista un subespacio propio de V' que contenga a V;

c) que dadas dos extensiones V' y V'_1 de V correspondientes a la extensión $\mathbb{F}'$ de $\mathbb{F}$, exista entre ellas un isomorfismo bien determinado.

Antes de tratar el problema que hemos enunciado, y con el objeto de facilitar la exposición, introduciremos una nueva noción.

Definición 12.1.1. Una aplicación f de un espacio vectorial V sobre $\mathbb{F}$ en un espacio vectorial W' sobre $\mathbb{F}'$, donde $\mathbb{F}'$ es una extensión de $\mathbb{F}$, se dice que es una *trasformación $\mathbb{F}$-lineal* o que es un *$\mathbb{F}$-homomorfismo* si cumple las condiciones siguientes:

$$f(x + y) = f(x) + f(y) \quad , \quad \text{para } x,y \in V$$

$$f(\alpha x) = \alpha f(x) \quad , \quad \text{para } \alpha \in \mathbb{F} \text{ y } x \in V.$$

Esta definición puede expresarse abreviadamente diciendo que f es *$\mathbb{F}$-lineal* si es una transformación lineal de V en $W'_{\mathbb{F}}$.

En el presente contexto puede decirse para mayor claridad que una trasformación lineal de un espacio V' sobre $\mathbb{F}'$ en un espacio W' sobre $\mathbb{F}'$ es una trasformación *$\mathbb{F}'$-lineal.*

En correspondencia con la dimensión de *$\mathbb{F}$-homomorfismo* que se da a una transformación *$\mathbb{F}$-lineal*, podremos también emplear las denominaciones de *$\mathbb{F}$-isomorfismo*, *$\mathbb{F}$-endomorfismo* y *$\mathbb{F}$-automorfismo*, cuya significación es obvia.

12.2. Extensión de un espacio vectorial por ampliación del campo de escalares.

En adelante supondremos que V es un espacio vectorial cualquiera dado sobre $\mathbb{F}$, y que $\mathbb{F}'$ es una extensión de $\mathbb{F}$. Cuando debamos limitarnos al caso en que V es de dimensión finita, haremos la respectiva aclaración.

Trataremos ahora del problema a que nos hemos referido en la sección 12.1.

Definición 12.2.1. Un par (V', θ) se dice que es una *extensión del*

espacio vectorial V sobre $\mathbb{F}$, correspondiente a la extensión $\mathbb{F}'$ del campo $\mathbb{F}$ si :

1. V' es un espacio vectorial sobre $\mathbb{F}'$
2. $\theta : V \to V'$ es una transformación *$\mathbb{F}$-lineal*
3. Si $(a_\alpha)_{\alpha \in A}$ es una base de V, entonces $\{\theta(a_\alpha)\}_{\alpha \in A}$ es una base de V' .

Corrientemente diremos, si se cumplen estas condiciones, que V' es *extensión* de V, y entonces estará sobreentendida la existencia de la transformación *$\mathbb{F}$-lineal* θ de V en V', la cual se llama *transformación de extensión.*

Teorema 12.2.1. Sea (V' , θ) una extensión del espacio vectorial V sobre $\mathbb{F}$, correspondiente a la extensión $\mathbb{F}'$ del campo $\mathbb{F}$. Se cumple entonces las propiedades siguientes:

a) θ es inyectiva
b) $\theta(V)$ genera a V'
c) Si (V'' , θ') es otra extensión de V correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$, entonces existe un *$\mathbb{F}'$-isomorfismo* bien determinado $\varphi : V' \to V''$ tal que

$$\theta' = \varphi \theta .$$

Esta última propiedad constituye la *unicidad de la extensión.*

Demostración. a) Para probar que θ es inyectiva basta demostrar que $\theta x = 0$ para $x \in V$, implica que $x = 0$. Si fuera $x \neq 0$ existiría una base de V uno de cuyos elementos sería x (ver el corolario 2 del teorema 2.5.2). Según la condición 3 de la definición 12.2.1, θx formaría parte de una base de V', y por tanto se tendría que $\theta x \neq 0$ contra lo supuesto.

b) $\theta(V)$ genera a V'. En efecto, en virtud de la condición 3 de la definición 12.2.1, $\theta(V)$ contiene bases de V'.

c) Sea (V'' , θ') otra extensión de V correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$. Dada la base $(a_\alpha)_{\alpha \in A}$ de V definamos la aplicación

$$\varphi : V' \to V''$$

por

$$\varphi(x') = \varphi(\sum_{\alpha\in A} \lambda^\alpha \theta(a_\alpha)) = \sum_{\alpha\in A} \lambda^\alpha \theta'(a_\alpha) \quad , x' \in V' \ , \ \lambda^\alpha \in \mathbb{F}'$$

Esta aplicación es bien definida porque $\{\theta(a_\alpha)\}_{\alpha\in A}$ es una base de V'; y obviamente es *$\mathbb{F}'$-lineal.* Es además un *$\mathbb{F}'$-isomorfismo*: es sobreyectiva porque $\{\theta'(a_\alpha)\}_{\alpha\in A}$ es una base de V'', y es inyectiva porque si $\varphi(x') = 0$ entonces es $\lambda^\alpha = 0$ para todo $\alpha \in A$ y por tanto

$$x' = \sum_{\alpha\in A} \lambda^\alpha \theta(a_\alpha) = 0 \ .$$

Finalmente, cualquiera que sea el elemento

$$x = \sum_{\alpha\in A} x^\alpha a_\alpha \quad , x^\alpha \in \mathbb{F} \ ,$$

de V, se tiene, en virtud de la definición de φ ,

$$\theta'(x) = \theta'(\sum_{\alpha\in A} x^\alpha a_\alpha) = \sum_{\alpha\in A} x^\alpha \theta'(a_\alpha) = \varphi(\sum_{\alpha\in A} x^\alpha \theta(a_\alpha))$$

$$= \varphi[\theta(\sum_{\alpha\in A} x^\alpha a_\alpha)] = (\varphi\theta)(x) \ ,$$

y por consiguiente se cumple la relación

$$\theta' = \varphi\theta \ .$$

Solo falta probar que si $\varphi_1 : V' \to V''$ es otro isomorfismo tal que

$$\theta' = \varphi_1 \theta$$

entonces es $\varphi = \varphi_1$. En efecto, se tiene

$$\varphi\theta = \varphi_1 \theta \ ;$$

y puesto que $\theta(V)$ genera a V', se tendrá $\varphi = \varphi_1$. ■

Corolario. Si $V_1' = \theta(V)$ es la imagen por θ del espacio V, entonces V_1' es un subespacio vectorial de $V'_{\mathbb{F}}$, y θ es un isomorfismo de V sobre V_1' .

Demostración. La aplicación θ de V en $V'_{\mathbb{F}}$ es una transformación

IF-lineal inyectiva en virtud de la propiedad a) del teorema anterior. Luego $\theta(V) = V_1'$ es un subespacio de V_{IF}', y θ es un isomorfismo de V sobre V_1' . ■

En virtud del precedente corolario podemos identificar a cada vector x de V con el vector $\theta(x)$ de V', y por tanto podemos suponer que V' contiene a V como subespacio de V_{IF}' .

Observación. Notese que si bien V_1' es un subespacio de V_{IF}' no lo es en general de $V' = V_{IF'}'$, pues para $x' \in V_1'$ y $\lambda \in IF'$ puede ocurrir que $\lambda x' \notin V_1'$. Esto ocurre en el ejemplo que exponemos a continuación.

Ejemplo. Sea IR el conjunto de los números reales, considerado como un espacio vectorial sobre el mismo IR como campo de escalares. El conjunto $\mathbb{C}$ de los números complejos, considerado como un espacio vectorial sobre el mismo $\mathbb{C}$, como campo de escalares, es la extensión de IR correspondiente a la extensión $\mathbb{C}$ del campo IR. En efecto:

1. $\mathbb{C}$ es un espacio sobre $\mathbb{C}$
2. θ es la inclusión de IR en $\mathbb{C}$
3. Una base de IR está formada por un número real no nulo. Su imagen por θ es el mismo número real, el cual constituye evidentemente una base de $\mathbb{C}$ sobre $\mathbb{C}$.

Obsérvese que IR no es un subespacio de $\mathbb{C} = \mathbb{C}_{\mathbb{C}}$, es decir que IR no es un espacio sobre el campo $\mathbb{C}$ pues si r es un número real ir, por ejemplo, no lo es. IR es en cambio un subespacio de $\mathbb{C}_{IR}$, como es facil observar.

El teorema 12.2.1, su corolario y la identificación que hemos hecho de cada elemento x de V con el elemento $\theta(x)$ de V', prueban que la extensión V' descrita en la definición 12.2.1 cumplen las demandas enunciadas en la sección 12.1.

12.3. Una nueva definición de la extensión de un espacio vectorial.

Dedicaremos la presente sección a una nueva forma de la definición de la extensión de un espacio vectorial, que es equivalente a la

definición 12.2.1.

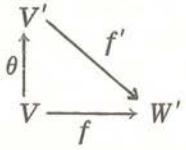
Teorema 12.3.1. Sea (V', θ) una extensión del espacio vectorial V sobre $\mathbb{F}$ correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$. Entonces, si

$$f : V \to W'$$

es una transformación $\mathbb{F}$-*lineal* de V en un espacio cualquiera W' sobre $\mathbb{F}'$, existe *una y solo una* trasformación $\mathbb{F}'$-*lineal* f' de V' en W' tal que

$$f = f'\theta \ ,$$

es decir tal que el diagrama



es conmutativo.

Demostración. Si $(a_\alpha)_{\alpha \in A}$ es una base de V entonces $(\theta(a_\alpha))_{\alpha \in A}$ es una base de V'. Podemos por tanto definir la transformación $\mathbb{F}'$-*lineal* $f' : V' \to W'$ mediante las fórmulas

$$f'(\theta(a_\alpha)) = f(a_\alpha) \quad , \quad \forall\ \alpha \in A.$$

Se tiene evidentemente

$$f = f'\theta$$

y por tanto f' cumple las condiciones del enunciado.

Probaremos ahora que f' es única. Si existiera otra transformación $\mathbb{F}'$-*lineal* $f_1' : V' \to W'$ tal que

$$f' = f_1'\theta \ ,$$

entonces, dado que $\theta(V)$ genera a V' (teorema 12.2.1), se tendrá

$$f = f_1' \ . \ \blacksquare$$

El teorema siguiente es recíproco del anterior.

Teorema 12.3.2. Si V' es un espacio vectorial sobre $\mathbb{F}'$ y θ es una

transformación *IF-lineal* del espacio vectorial V sobre $\mathbb{F}$ en V', tal que para cada transformación *IF-lineal* f de V en un espacio vectorial cualquiera W' sobre $\mathbb{F}'$ existe *una y solo una* transformación $\mathbb{F}'$-*lineal* $f' : V' \to W'$ tal que $f = f'\theta$, entonces (V', θ) es una extensión de V correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$.

Demostración. Por la definición 12.2.1, solo falta demostrar que si $(a_\alpha)_{\alpha \in A}$ es una base de V, entonces $(\theta(a_\alpha))_{\alpha \in A}$ es una base de V'. Llevaremos a cabo esta demostración en dos partes:

1. $(\theta(a_\alpha))_{\alpha \in A}$ *es un sistema de generadores de* V'.

Sea $f_o : V \to \mathbb{F}'$ la función *IF-lineal* idénticamente nula. A ella le corresponde, de conformidad con el enunciado, una única función $\mathbb{F}'$-*lineal* $f'_o : V' \to \mathbb{F}'$ tal que

$$f_o = f'_o \theta$$

la cual evidentemente no puede ser otra que la función idénticamente nula.

Designemos por V'_1 al subespacio de V' generado por los elementos del conjunto $(\theta(a_\alpha))_{\alpha \in A}$. Nuestro propósito se reduce a demostrar que $V'_1 = V'$. Probaremos, con tal fin, que si se supone que es $V'_1 \neq V'$ es posible definir una función $\mathbb{F}'$- lineal

$$f''_o : V' \to \mathbb{F}'$$

diferente de f'_o, tal que

$$f_o = f''_o \theta \quad ;$$

con lo cual habremos llegado a una contradicción con la condición de unicidad supuesta en el enunciado del teorema. En efecto, supongamos que sea $V'_1 \neq V'$, es decir que V'_1 sea un subespacio propio de V'. Existiría entonces un elemento $b \neq 0$ de V', no contenido en V'_1; y habría una base $(b, (b_\beta)_{\beta \in B})$ de V' que contendría a b entre sus elementos. Para cada $y \in V'$ designemos por η a la componente de y respecto de esa base, relativa al vector b. Sea $\alpha' \neq 0$ un elemento cualquiera de $\mathbb{F}'$. Definamos la función $\mathbb{F}'$-*lineal*

$$f''_o : V' \to \mathbb{F}'$$

tal que

$$f_o''(y) = \eta\,\alpha'$$

Se tiene que $f_o''(b) = \alpha' \neq 0$, y por tanto es $f_o'' \neq f_o'$. Además es $f_o''(y) = 0$ para $y \in V_1'$, pues en este caso es $\eta = 0$; luego es

$$f_o''(\theta x) = 0 \quad \text{para} \quad x \in V\,,$$

o sea

$$f_o''\,\theta = 0.$$

Se llega así a la contradicción anunciada. Por tanto es $V_1' = V'$.

2. $(\theta(a_\alpha))_{\alpha\in A}$ *es un sistema l.i. de vectores de* V'.

Supongamos que se tenga

$$\sum_{\alpha\in A} \lambda^\alpha \theta(a_\alpha) = 0\,, \quad (\lambda^\alpha \in \mathbb{F}'\,,\ \forall\ \alpha \in A). \quad (12.3.1)$$

Definamos la familia $(f_\alpha)_{\alpha\in A}$ de funciones $\mathbb{F}$*-lineales*

$$f_\alpha : V \to \mathbb{F}'\,, \qquad \forall\ \alpha \in A$$

de manera que sea

$$f_\alpha(a_\alpha) = 1 \quad \text{y} \quad f_\alpha(a_{\alpha'}) = 0 \quad \text{para} \quad \alpha' \neq \alpha\,.$$

Por la hipótesis del teorema, existirá la familia $(g_\alpha)_{\alpha\in A}$ de transformaciones $\mathbb{F}'$*-lineales*

$$g_\alpha : V' \to \mathbb{F}'\,, \qquad \forall\ \alpha \in A\,,$$

tal que

$$f_\alpha = g_\alpha\,\theta\,, \qquad \forall\ \alpha \in A.$$

Para cada $\alpha_1 \in A$ se tendrá entonces

$$\lambda^{\alpha_1} = \sum_{\alpha\in A} \lambda^\alpha f_{\alpha_1}(a_\alpha) = \sum_{\alpha\in A} \lambda^\alpha g_{\alpha_1}(\theta(a_\alpha)) = g_{\alpha_1}\Big(\sum_{\alpha\in A} \lambda^\alpha \theta(a_\alpha)\Big)$$

$$= g_{\alpha_1}(0) = 0.$$

Resulta así que la relación lineal (12.3.1) es trivial, y por tanto el sistema $(\theta(a_\alpha))_{\alpha\in A}$ es *l.i.*

En conclusión $(\theta(a_\alpha))_{\alpha\in A}$ será una base de V'. ∎

Los dos teoremas precedentes permiten dar una nueva definición de la extensión de un espacio vectorial, que es enteramente equivalente a la definición 12.2.1.

Definición 12.3.1. Se dice que el par (V', θ) es una extensión del espacio vectorial V sobre $\mathbb{F}$, correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$, si V' es un espacio vectorial sobre $\mathbb{F}'$, y θ es una transformación *$\mathbb{F}$-lineal* de V en V', tal que para cada transformación *$\mathbb{F}$-lineal* f de V en un espacio cualquiera W' sobre $\mathbb{F}'$, existe *una y solo una* transformación *$\mathbb{F}'$-lineal* f' de V' en W' tal que

$$f = f'\theta \ .$$

Observación. Esta nueva definición muestra que la extensión (V', θ) es la solución del problema de aplicación universal para un espacio V sobre el campo $\mathbb{F}$, respecto de la especie de estructuras de espacios vectoriales sobre la extensión $\mathbb{F}'$ de $\mathbb{F}$ (cuyos morfismos son las transformaciones *$\mathbb{F}'$-lineales*) y de las transformaciones *$\mathbb{F}$-lineales* de V en un espacio vectorial cualquiera sobre $\mathbb{F}'$. (Ver a este repecto: *Bourbaki, Théorie des ensembles*, ch. *4*, Structures, Hermann (1957) , pág 43).

12.4. Extensiones de un espacio de dimensión finita.

Hasta ahora hemos supuesto que el espacio V, cuya extensión hemos considerado, es un espacio enteramente cualquiera. En esta sección vamos a suponer que V es de dimensión finita: $dim\ V = n$.

Teorema 12.4.1. Las condiciones necesarias y suficientes para que el par (V', θ) sea una extensión del espacio V de dimensión n sobre $\mathbb{F}$ correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$ son que V' sea un espacio sobre $\mathbb{F}'$ de dimensión n y que $\theta : V \to V'$ sea una transformación *$\mathbb{F}$-lineal* tal que V' sea generado por $\theta(V)$.

Demostración. Sea (V', θ) una extensión de V. La primera condición del enunciado es consecuencia inmediata de la definición 12.2.1, pues si $(a_1, \ldots, a_n)$ es una base de V, entonces $(\theta a_1, \ldots, \theta a_n)$

será una base de V' y por tanto $dim\ V' = n$.

La segunda condición del enunciado fue establecida en el teorema 12.2.1, b).

Supongamos ahora, recíprocamente, que se cumplan las condiciones del enunciado. Para probar que (V', θ) es una extensión de V bastará demostrar, por la definición 12.2.1, que si $(a_1, \ldots, a_n)$ es una base de V, entonces $(\theta a_1, \ldots, \theta a_n)$ es una base de V'. Ahora bien, puesto que por hipótesis $\theta(V)$ genera a V', y que $(a_1, \ldots, a_n)$ es una base de V, los elementos $(\theta a_1, \ldots, \theta a_n)$ generarán a V'. Los n elementos $(\theta a_1, \ldots, \theta a_n)$ constituyen entonces una base de V', por cuanto la dimensión de este espacio es n. ■

Teorema 12.4.2. *(Teorema de existencia)* Dados el espacio V sobre $\mathbb{F}$ de dimensión finita n, y la extensión $\mathbb{F}'$ del campo $\mathbb{F}$, existe una extensión (V', θ) de V correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$. Si $(a_1, \ldots, a_n)$ es una base de V, $(e_1, \ldots, e_n)$ es la base canónica de $\mathbb{F}'^n$, y $\theta : V \to \mathbb{F}'^n$ es la transformación $\mathbb{F}$-*lineal* tal que $\theta(a_i) = e_i$, entonces $(\mathbb{F}'^n, \theta)$ es una de esas extensiones.

Demostración. Vamos a probar que el par $(\mathbb{F}'^n, \theta)$, a que se refiere el enunciado, es una extensión de V correspondiente a la extención $\mathbb{F}'$ de $\mathbb{F}$. De conformidad con el teorema 12.4.1, y dado que $\mathbb{F}'^n$ es de dimensión n, basta con probar que $\mathbb{F}'^n$ es generado por $\theta(V)$. Ahora bien, cada elemento de $\mathbb{F}'^n$ se expresa en la forma

$$\sum_{i=1}^{n} \lambda^i e_i = \sum_{i=1}^{n} \lambda^i \theta(a_i) \quad , \quad \lambda^i \in \mathbb{F}' \ .$$

Resulta de aquí inmediatamente que $\theta(V)$ genera a V'. ■

12.5. Extensión de un endomorfismo de V por extensión de V.

En esta sección V es un espacio vectorial cualquiera.

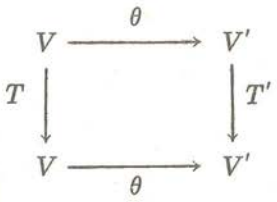
Teorema 12.5.1. Sea (V', θ) una extensión de V correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$. Dado un $\mathbb{F}$-*endomorfismo* cualquiera

$$T : V \to V$$

existe un único $\mathbb{F}'$-*endomorfismo* $T' : V' \to V'$ tal que

$$\theta T = T'\theta \ ,$$

es decir tal que hace conmutativo al diagrama



Demostración. $f = \theta T$ es una transformación *IF-lineal* de V en V'. En virtud de la definición 12.3.1 existirá uno y solo un *IF'-endomorfismo* $T' : V' \to V'$ tal que $\theta T = T'\theta$. ■

Definición 12.5.1. Si (V' , θ) es una extensión de V sobre *IF* correspondiente a una extensión *IF'* de *IF*, entonces, dado un *IF-endomorfismo* cualquiera $T : V \to V$, se llama *extensión* de T correspondiente a la extensión *IF'* de *IF* al *IF'-endomorfismo* $T' : V' \to V'$ tal que

$$\theta T = T'\theta \ .$$

Observación. Supongamos ahora, en particular, que V es de dimensión finita n y que $(a_1 , \ldots , a_n)$ es una de sus bases. $(\theta a_1 , \ldots , \theta a_n)$ será una base de V'.

Si $T' : V' \to V'$ es la extensión del *IF-endomorfismo* $T : V \to V$, correspondiente a la extensión *IF'* de *IF*, entonces la matriz que representa a T' en la base $\{\theta(a_i)\}$ de V' es la misma que representa a T en la base $\{a_i\}$. En efecto, se tiene

$$T'[\theta(a_i)] = \theta[T(a_i)] \ .$$

Por tanto, si

$$T(a_i) = \alpha_i^1 a_1 + \cdots + \alpha_i^n a_n \ ,$$

entonces se tiene

$$T'[\theta(a_i)] = \alpha_i^1 \theta(a_1) + \cdots + \alpha_i^n \theta(a_n) \ .$$

Se ve así que las componentes de $T'[\theta(a_i)]$ respecto de la base $\{\theta(a_i)\}$ son idénticas a las componentes de $T(a_i)$ respecto de la base $\{a_i\}$ lo cual prueba nuestra precedente afirmación.

De lo dicho resulta que una matriz $n \times n$ sobre $\mathbb{F}$ puede interpretarse como representativa de un *$\mathbb{F}$-endomorfismo* T de un espacio V sobre $\mathbb{F}$ de dimensión n o bien del *$\mathbb{F}'$-endomorfismo* T' que es extensión de T correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$.

12.6. Extensión de un espacio vectorial real, correspondiente a la extensión $\mathbb{C}$ de $\mathbb{R}$.

Vamos a considerar aquí el caso particular en que $\mathbb{F} = \mathbb{R}$ y $\mathbb{F}' = \mathbb{C}$. V será entonces un campo vectorial real, y su extensión será un espacio vectorial complejo, V' que se llama *complejificación* de V y que se designa por $V^{\mathbb{C}}$. Convendremos en *identificar* cada vector x de V con su imagen $\theta(x)$ por la transformación de extensión. Podemos admitir por tanto que $V \subset V^{\mathbb{C}}$.

Teorema 12.6.1. Si $V^{\mathbb{C}}$ es la complejificación de un espacio vectorial real *n-dimensional* V. Entonces:

1. Cada elemento z de $V^{\mathbb{C}}$, puede representarse de una sola manera en la *forma canónica.*

$$z = x + iy \quad , \quad x,y \in V$$

2. Para cada $z \in V^{\mathbb{C}}$, sea $z = x + iy$, y definamos el conjugado de z como el vector $\overline{z} \in V^{\mathbb{C}}$ tal que $\overline{z} = x - iy$. Entonces se cumplen las siguientes relaciones:

i) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

ii) $\overline{cz} = \overline{c}\,\overline{z}$, donde $c \in \mathbb{C}$ y $\overline{c}$ es el conjugado de c.

iii) $\overline{\overline{z}} = z$.

iv) $z \in V^{\mathbb{C}}$ pertenece a V si y solo si $\overline{z} = z$.

Demostración 1. Puesto que V genera a $V^{\mathbb{C}}$ (teorema 12.2.1), cada elemento z de $V^{\mathbb{C}}$ puede expresarse en la forma

$$z = \sum_{k=1}^{m} c^k x_k \quad , \quad x_k \in V \quad , \quad c^k \in \mathbb{C}.$$

Si $c^k = a^k + ib^k$, donde a^k y b^k son reales, podremos escribir

$$z = x + yi \qquad (12.6.1)$$

donde

$$x = \sum_{k=1}^{m} a^k x_k \quad , \quad y = \sum_{k=1}^{m} b^k x$$

son elementos de V.

Para probar que la representación (12.6.1) es única, supongamos que x_1 , y_1 son vectores de V tales que

$$x + yi = x_1 + y_1 i \ .$$

Si escribimos $u = x - x_1$, $v = y - y_1$,
tendremos

$$u = iv \quad , \qquad (12.6.2)$$

y por tanto u y v serán vectores de V que son *l.d.* en $V^{\mathbb{C}}$. Ahora bien, si u y v fuerán *l.i.* en V, lo serían en $V^{\mathbb{C}}$, por cuanto podrían incluirse en una base de V que por θ se aplicaría sobre una base de $V^{\mathbb{C}}$ (definición 12.2.1, propiedad 3), a la cual pertenecerían $\theta u = u$ y $\theta v = v$, que serían *l.i.* Luego u y v son *l.d.* en V y existe por tanto un número real a tal que $u = av$. De (12.6.2) se deduce entonces que

$$(a - i)\, v = 0.$$

Si fuera $v \neq 0$ se tendría que $a = i$ contra la suposición de que a es real. Luego debe ser $v = 0$, y por tanto $u = av = 0$. Es decir que $x = x_1$ e $y = y_1$, y por consiguiente la representación (12.6.1) de los vectores de $V^{\mathbb{C}}$ es única.

2.i) $$\overline{z_1 + z_2} = \overline{(x_1 + y_1 i) + (x_2 + y_2 i)}$$
$$= \overline{(x_1 + x_2) + (y_1 + y_2) i}$$
$$= (x_1 + x_2) - (y_1 + y_2) i$$

$$= (x_1 - y_1 i) + (x_2 - y_2 i)$$
$$= \overline{z}_1 + \overline{z}_2$$

2.ii) $\overline{cz} = \overline{c(x + yi)} = \overline{(a + ib)(x + yi)}$

$$= \overline{(ax - by) + (ay + bx)i}$$
$$= (ax - by) - (ay + bx)i$$
$$= (a - ib)(x - yi) = \overline{c}\,\overline{z} \quad .$$

2.iii) $\overline{\overline{z}} = \overline{\overline{(x + yi)}} = \overline{x - yi} = x + yi.$

2. iv) Si $z \in V$, entonces su escritura canónica es $z = z + 0i$. Por tanto $\overline{z} = z - 0i = z$.

Si se cumple ahora la condición $\overline{z} = z$, entonces si $z = x + yi$, $x,y \in V$, es la escritura canónica de z, se tiene $x + yi = x - yi$

La unicidad de la escritura canónica implica que $y = -y$, es decir que $y = 0$, y por consiguiente $z = x \in V$. ■

Ejercicios

*12.6.1. Si se considera a $\mathbb{C}^n$ como la complejificación de $\mathbb{R}^n$, entonces el conjugado del vector

$$z = (\zeta_1, \ldots, \zeta_n) \in \mathbb{C}^n$$

donde los ζ_i son números complejos, es

$$\overline{z} = (\overline{\zeta}_1, \ldots, \overline{\zeta}_n)$$

donde $\overline{\zeta}_i$ es el número complejo conjugado de ζ_i .

12.6.2. Una matriz compleja A, $n \times n$, es una matriz real si y solo si $A\zeta \in \mathbb{R}^n$ para cada $\zeta \in \mathbb{R}^n$.

CAPITULO IV

TEORIA DE LOS ENDOMORFISMOS DE UN ESPACIO VECTORIAL DE DIMENSION FINITA

En este capítulo vamos a suponer que V es un espacio vectorial de dimensión finita n sobre el campo $\mathbb{F}$. Si $\mathfrak{B} = \{e_1, \ldots, e_n\}$ es una base dada de V, entonces, según (sección 6.4) al *endomorfismo*

$$T : V \longrightarrow V$$

le corresponde una matriz determinada de orden n

$$M(T, \mathfrak{B}) = \begin{pmatrix} \alpha_1^1 & \ldots & \alpha_n^1 \\ \vdots & \vdots\vdots\vdots & \vdots \\ \alpha_1^n & \ldots & \alpha_n^n \end{pmatrix} \qquad \alpha_j^i \in \mathbb{F} \quad ,$$

tal que

$$Te_j = \sum_{i=1}^{n} \alpha_j^i e_i \quad , \qquad i = 1, 2, \ldots, n .$$

Estudiaremos diversos conceptos fundamentales ligados a un endomorfismo dado T de V. En particular veremos cómo, mediante la elección adecuada de la base, es posible representar a T mediante una matriz de forma particularmente sencilla.

En el presente capítulo desempeña un papel importante el concepto de subespacio invariante respecto de un endomorfismo. Se dice que el subespacio V_1 del espacio vectorial V es *invariante o estable res-*

pecto del endomorfismo T de V si se cumple la condición

$$T(V_1) \subset V_1 \quad ,$$

es decir si la imagen por T de cada elemento de V_1 pertenece a V_1

§ 13. Valores propios, polinomio característico y polinomio minimal de un endomorfismo.

13.1. Valores propios y vectores propios de un endomorfismo.

Definición 13.1.1. Un escalar $\lambda \in \mathbb{F}$ se llama *valor propio o valor característico del endomorfismo* T del espacio V sobre $\mathbb{F}$, si existe en V un vector no nulo a, tal que

$$T(a) = \lambda a \ .$$

Cada vector *no nulo* a de V que satisface a esa condición se llama *vector propio o vector característico de* T correspondiente al valor propio λ.

Ejemplo A. Supongamos que V es el espacio cartesiano $\mathbb{R}^3$ y que T aplica al vector $x = (\xi^1 , \xi^2 , \xi^3)$ en el vector $y = (\eta^1 , \eta^2 , \eta^3)$ tal que

$$\begin{aligned} \eta^1 &= \xi^1 + \xi^2 \\ \eta^2 &= \xi^2 \\ \eta^3 &= 2\xi^3 \ . \end{aligned}$$

T tiene los valores propios $\lambda_1 = 1$ y $\lambda_2 = 2$. Al primero le corresponden los vectores propios $(\alpha , 0 , 0)$ donde $\alpha \neq 0$ es un número real cualquiera; y al segundo le corresponden los valores propios $(0 , 0 , \beta)$ donde $\beta \neq 0$ es un número real cualquiera. Se tiene en efecto

$$\begin{aligned} T(\alpha , 0 , 0) &= (\alpha , 0 , 0) \\ T(0 , 0 , \beta) &= 2(0 , 0 , \beta) \ . \end{aligned}$$

Cada uno de esos vectores propios se caracteriza porque se transforma por T en un vector que le es paralelo. Observemos además que para cada uno de los dos valores propios 1 y 2 hemos hallado un

subespacio de $\mathbb{R}^3$ formado por vectores propios a excepción del vector 0.

B. Sea $\{i, j, k\}$ una base de $\mathbb{R}^3$ y sea T el endomorfismo tal que

$$Ti = i \quad , \quad Tj = j \quad , \quad Tk = -k$$

Los valores propios de T son 1 y -1. Los vectores propios correspondientes al primero son los vectores no nulos del plano (i, j); los vectores propios correspondientes a -1 son los vectores no nulos paralelos al vector k.

C. Un endomorfismo puede no tener vectores propios. Así una rotación de $\mathbb{R}^2$ sobre sí mismo en un ángulo que no es múltiplo de π no los tiene.

D. Cero es valor propio de un endomorfismo si y solo si ese endomorfismo es singular, porque solo entonces hay vectores no nulos que se aplican en cero.

Teorema 13.1.1. Si λ es un valor propio de T, entonces los vectores propios asociados a λ son los vectores no nulos de un subespacio de V de dimensión $\geqslant 1$.

Demostración. La proposición resulta de que la suma de dos vectores propios correspondientes a λ y que el producto de un escalar por un vector propio asociado a λ son vectores propios correspondientes al mismo valor propio; y de que, además, existe por lo menos un vector propio correspondiente a λ, el cuál por definición no es nulo. ∎

Teorema 13.1.2. Si $x_1, \ldots, x_k$ son vectores propios del endomorfismo T que corresponden, respectivamente, a valores propios $\lambda_1, \lambda_2, \ldots, \lambda_k$, distintos entre sí, entonces los vectores $x_1, \ldots, x_k$ son linealmente independientes.

Demostración. Inducción respecto a k. Si $k = 1$ la proposición resulta obvia pues entonces solo tenemos un vector propio que, por definición, es diferente de cero, y por tanto es *l.i.* Supongamos pues que el teorema se cumple para $k-1$ vectores propios y lo demos-

traremos en el caso de k vectores propios. Si en el caso de los k-vectores propios no se cumpliera el teorema, existiría entre ellos una relación lineal no trivial

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k = 0 \quad , \qquad (13.1.1)$$

donde podemos suponer, por ejemplo que es $\alpha_1 \neq 0$. Aplicando el endomorfismo T a los miembros de la precedente ecuación resultaría

$$\alpha_1 T x_1 + \alpha_2 T x_2 + \cdots + \alpha_k T x_k = 0 \ ,$$

y teniendo presente que por hipótesis es $Tx_i = \lambda_i x_i$ para $i = 1, \ldots, k$, se tendría

$$\alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \cdots + \alpha_k \lambda_k x_k = 0$$

Si restamos de esta ecuación la ecuación (13.1.1) multiplicada por λ_k resultaría

$$\alpha_1(\lambda_1 - \lambda_k)x_1 + \alpha_2(\lambda_2 - \lambda_k)x_2 + \cdots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)x_{k-1} = 0$$

donde $\alpha_1(\lambda_1 - \lambda_k) \neq 0$ (recordar que por hipótesis es $\lambda_1 \neq \lambda_k$). La relación no trivial que acabamos de hallar para los vectores

$$x_1, \ldots, x_{k-1}$$

estaría en contradicción con la hipótesis de inducción, según el cual esos $k - 1$ vectores propios que corresponden a $k - 1$ valores propios diferentes entre sí deberían ser *l.i.* ■

Ejercicios

13.1.1. ¿Cuáles son los valores propios y los vectores propios del endomorfismo $T = \alpha I$ donde $\alpha \in \mathbb{F}$ e I es el endomorfismo identidad de V?

13.1.2. Consideremos el endomorfismo de $\mathbb{R}^2$

$$T : (\xi^1, \xi^2) \mapsto (\xi^1 - \xi^2 \ , \ 2\xi^1 - \xi^2)$$

¿Tiene esta transformación valores propios?

13.1.3. Considerar el mismo ejercicio anterior pero suponiendo que T es un endomorfismo de $\mathbb{C}^2$.

13.1.4. Sea V un espacio sobre $\mathbb{R}$ de dimensión 2 y sea (e_1, e_2) una base de ese espacio. ¿Tiene vectores propios el endomorfismo T, tal que

$$Te_1 = e_2 \quad , \quad Te_2 = -e_1 \ ?$$

13.2. Polinomio característico de un endomorfismo.

Teorema 13.2.1. La condición necesaria y suficiente para que el escalar $\lambda_o \in \mathbb{F}$ sea un valor propio del endomorfismo T del espacio vectorial V sobre $\mathbb{F}$, de n dimensiones, es que si $M = (\alpha_j^i)$ es la matriz de T relativa a una base cualquiera de V, entonces λ_o sea raiz de la ecuación en λ:

$$\begin{vmatrix} (\lambda - \alpha_1^1) & -\alpha_2^1 & \dots & -\alpha_n^1 \\ -\alpha_1^2 & (\lambda - \alpha_2^2) & \dots & -\alpha_n^2 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ -\alpha_1^n & -\alpha_2^n & \dots & (\lambda - \alpha_n^n) \end{vmatrix} = 0 \qquad (13.2.1)$$

Demostración. Supongamos que $\lambda_o \in \mathbb{F}$ sea un valor propio de T, y que x sea un vector propio correspondiente a λ_o.
Se tiene por consiguiente

$$Tx = \lambda_o x \ ,$$

o sea

$$(\lambda_o I - T)(x) = 0 \quad , \quad x \neq 0,$$

donde I es el automorfismo identidad de V. La transformación

$$\lambda_o I - T$$

es por tanto no regular. Si $\{e_i\}$ es una base cualquiera de V y M es la matriz de T en esa base, la matriz de la transformación

$$\lambda_o I - T$$

es $\lambda_o E_n - M$, donde E_n es la matriz unidad. Deberá por tanto cumplirse la relación

$$det\ (\lambda_o E_n - M) = 0 \qquad (13.2.2)$$

que expresa que λ_o es raiz de la ecuación (13.2.1).

Si, recíprocamente, suponemos que se cumple (13.2.2), entonces la transformación $\lambda I - T$ no será regular y existirá un vector $x \neq 0$ tal que

$$(\lambda_o I - T)(x) = 0$$

o sea, que

$$Tx = \lambda_o x .$$

x será por tanto un vector propio correspondiente al valor propio λ_o. ■

La ecuación (13.2.1) puede escribirse en la forma

$$det\ (\lambda E_n - M) = 0 . \qquad (13.2.3)$$

Sus raices pertenecientes a $\mathbb{F}$ son los valores propios de T. Los vectores propios correspondientes a cada valor propio $\lambda_o \in \mathbb{F}$ son las soluciones $x \neq 0$ de la ecuación vectorial $(\lambda_o I - T)(x) = 0$, o sea los vectores diferentes de cero, del núcleo del endomorfismo no regular

$$\lambda_o I - T$$

cuya matriz en la base (e_i) es

$$\lambda_o E_n - M .$$

El núcleo de ese endomorfismo se llama *espacio propio* correspondiente a λ_o ; y su dimensión (ver la sección 3.5 y el teorema 6.7.1) es igual a

$$n - r(\lambda_o I - T) = n - r(\lambda_o E_n - M).$$

Puesto que los valores propios de una transformación han sido definidos con completa independencia de las bases que se puedan elegir en V, podemos estar seguros de que las raices del polinomio

$$det\ (xE_n - M) , \qquad (13.2.4)$$

pertenecientes al campo $\mathbb{F}$, son las mismas cualquiera que sea la base elegida en V, y por tanto, cualquiera que sea la matriz M que representa al endomorfismo considerado T. A continuación vamos a

probar un resultado mucho más fuerte, a saber que el polinomio (13.2.4) es independiente de la base y solo depende de T. Pero antes de tratar esa cuestión nos detendremos brevemente en probar algunos resultados que son consecuencias inmediatas de lo dicho hasta aquí.

Teorema 13.2.2. Si un endomorfismo T del espacio V de dimensión n tiene n vectores propios *l.i.*, entonces estos vectores constituyen una base de V, respecto de la cual T se representa por una matriz diagonal, cuya diagonal principal está formada por los valores propios que corresponden a los valores propios dados.

Inversamente, si la matriz de T respecto de la base $\{e_1, \ldots, e_n\}$ es una matriz diagonal

$$\begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix} ,$$

los escalares $\lambda_1, \ldots, \lambda_n$ son los valores propios de T, y los vectores $e_1, \ldots, e_n$ son vectores propios correspondientes, respectivamente, a esos valores propios.

Demostración. Puesto que $\dim V = n$, es claro que los n vectores propios *l.i.*: $(e_1, \ldots, e_n)$, cuya existencia se ha supuesto, constituyen una base de V. Ahora bien, por definición se tiene

$$Te_i = \lambda_i e_i \quad ,$$

donde λ_i es el valor propio al que corresponde el valor propio e_i. Luego la matriz que representa a T en la base $\{e_i\}$ es

$$\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

Inversamente, si suponemos que la matriz precedente es la ma-

triz de T respecto de la base $(e_1, \ldots, e_n)$, se tendrá que $Te_i = \lambda_i e_i$ para $i = 1, \ldots, n$ y por tanto e_i es un vector propio correspondiente al valor propio λ_i. Además, si λ es un valor propio de T y $\sum_i \xi^i e_i$ es un correspondiente valor propio, se deberá tener

$$T\left(\sum_i \xi^i e_i\right) = \lambda \sum_i \xi^i e_i \quad ,$$

y por tanto

$$\sum_i \xi^i(\lambda_i - \lambda)e_i = 0 \quad ,$$

de donde

$$\xi^i(\lambda_i - \lambda) = 0 \quad , \quad (i = 1, \ldots, n) .$$

Puesto que los ξ^i no son todos nulos, existe por lo menos un valor de i tal que $\lambda = \lambda_i$; es decir que los λ_i son todos los valores propios de T. ■

Corolario. Si el endomorfismo T tiene n valores propios diferentes entre sí, entonces la base de V puede elegirse de manera que la matriz de T sea diagonal.

Demostración. Si $e_1, \ldots, e_n$ son vectores propios correspondientes a los n valores propios, diferentes entre sí, cuya existencia ha sido supuesta, entonces, según el teorema 13.1.2, dichos vectores son *l.i.* Basta aplicar ahora el teorema anterior. ■

Definición 13.2.1. Dada una matriz M de orden n sobre $\mathbb{F}$, se llama *polinomio característico de* M al polinomio (13.2.4) de $\mathbb{F}[x]$, al que designaremos por $f_M(x)$. Tendremos por consiguiente

$$f_M(x) = det\,(x E_n - M) = det \begin{pmatrix} x - \alpha_1^1 & -\alpha_2^1 & \ldots & -\alpha_n^1 \\ \alpha_1^2 & x - \alpha_2^2 & \ldots & -\alpha_n^2 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ \alpha_1^n & -\alpha_2^n & \ldots & x - \alpha_n^n \end{pmatrix}$$

El teorema 13.2.1 ha establecido, por tanto, que los valores propios de T son las raices del polinomio característico de cualquiera de las

matrices que lo representan.

Teorema 13.2.3. Si M' es semejante a la matriz M (sección 6.11); es decir, si existe una matriz regular A tal que

$$M' = AMA^{-1} ,$$

entonces es

$$f_{M'}(x) = f_M(x).$$

Demostración.

$$xE_n - M' = xE_n - AMA^{-1} = A(xE_n - M)A^{-1}$$

Luego ,

$$det\ (xE_n - M') = det\ (xE_n - M) ,$$

o sea

$$f_{M'}(x) = f_M(x) . \quad \blacksquare$$

Corolario. Si las matrices M y M' representan a un mismo endomorfismo T, respecto de dos bases cualesquiera, entonces se tiene que

$$f_{M'}(x) = f_M(x) \quad ;$$

es decir que ambos polinomios tienen los mismos coeficientes.

Demostración. Según sabemos (sección 6.11), las matrices M y M' son semejantes. Se aplica entonces el teorema 13.2.3. ■

El corolario anterior asegura que, dado un endomorfismo, todos los polinomios característicos de las matrices que lo representan respecto de todas las bases posibles, son idénticas. En consecuencia , podemos dar la definición siguiente:

Definición 13.2.2. Dado un endomorfismo T de V, se llama *polinomio característico de* T, y se designa por $f_T(x)$ al polinomio característico de cualquiera de las matrices que representan a T. Según el corolario del teorema anterior, los coeficientes de ese polinomio son independientes de la base considerada.

Cualquiera que sea la matriz $M = (\alpha^i_j)$ que representa a T, el polinomio característico $f_T(x)$ es dado por

$$f_T(x) = f_M(x) = det\ (xE_n - M) = \begin{vmatrix} (x-\alpha_1^1) & -\alpha_2^1 & \dots & -\alpha_n^1 \\ -\alpha_1^2 & (x-\alpha_2^2) & \dots & -\alpha_n^2 \\ \vdots & \vdots & & \vdots \\ -\alpha_1^n & -\alpha_2^n & \dots & (x-\alpha_n^n) \end{vmatrix}$$

$$= x^n - p_1 x^{n-1} + p_2 x^{n-2} - \dots + (-1)^n p_n \ .$$

Los coeficientes $p_1, \dots, p_n$ son, como ya hemos visto, independientes de la base que se haya adoptado, y por tanto son *invariantes de T*. Es fácil observar que p_1 es igual a la suma de los términos de la diagonal principal de la matriz de M, suma que se llama *traza* de T; p_2 es la suma de los menores principales de orden 3 de M, etc. Finalmente p_n es igual al *determinante de T*, cuya naturaleza invariante respecto de los cambios de base ya nos era conocida.

En virtud del Teorema 13.2.1 podemos decir que los valores de T son las raíces del polinomio característico $f_T(x)$, que están contenidas en $\mathbb{F}$. En general, $f_T(x)$ puede no tener raíces en $\mathbb{F}$. En efecto, según el teorema 11.2.1, si $f_T(x)$ tiene el factor irreducible $g(x)$, entonces este polinomio, y por consiguiente $f_T(x)$ tiene una raiz en el campo $\mathbb{F}^* = \mathbb{F}[x]/_{I(g)}$. Sin embargo, las raíces de $f_T(x)$ que no pertenecen a $\mathbb{F}$, no son valores propios de T. (Comparar por ejemplo, los resultados de los ejercicios 13.1.2 y 13.1.3).

Es claro que los problemas relativos a los valores propios resultan más fáciles de estudiar en el caso en que $\mathbb{F}$ es un campo algebraicamente cerrado, pues entonces todas las raíces de $f_T(x)$ son valores propios de T. Tal es, en particular, el caso cuando $\mathbb{F}$ es el campo algebraicamentes cerrado $\mathbb{C}$ de los números complejos.

Las consideraciones que siguen se refieren al caso en que T es un endomorfismo de un espacio V sobre un campo algebraicamente cerrado. El teorema 13.2.4 establece la propiedad de los endomor

fismos de un espacio vectorial sobre un campo $\mathbb{F}$ algebraicamente cerrado, de poder ser representados por una matriz triangular. Previamente probaremos el siguiente lema.

Lema 13.2.1. Si T es un endomorfismo de un espacio vectorial V de dimensión n sobre el campo $\mathbb{F}$ algebraicamente cerrado, existen $n + 1$ subespacios $M_0, M_1, \ldots, M_n$ de V, con las siguientes propiedades:

(i) Los espacios M_j $(j = 0,1, \ldots, n)$ son invariantes respecto de T

(ii) La dimensión de M_j es j.

(iii) ($\{0\}$ =) $M_0 \subset M_1 \subset \cdots \subset M_{n-1} \subset M_n$ $(= V)$.

Demostración. Para $n = 0$ y $n = 1$ el resultado es trivial. Procedamos ahora por inducción respecto de n, a cuyo fin supondremos que el enunciado se cumple para los espacios de dimensión $n - 1$. Consideremos la transformación $T^* : V^* \to V^*$ dual de T. Puesto que $\mathbb{F}$ es algebraicamente cerrado, T^* tiene por lo menos un valor propio y por tanto un vector propio x^*. El espacio M de dimensión 1, generado por x^*, es evidentemente invariante respecto de T^*. Sea $M_{n-1} = M^{\perp}$ el complemento ortogonal de M, y por tanto un subespacio de V de dimensión $n - 1$. M_{n-1} es invariante respecto de T, pues para $x \in M_{n-1}$ se tiene que $\langle x , y \rangle = 0$ para todo $y \in M$; luego es $\langle Tx , y \rangle = \langle x , T^*y \rangle = 0$ para todo $y \in M$. Se deduce de aquí que $Tx \in M^{\perp} = M_{n-1}$, y por tanto M_{n-1} es invariante respecto de T. Podemos por consiguiente considerar la restricción de T a M_{n-1}, la cual será un endomorfismo de este último espacio. Por la hipótesis de inducción podemos hallar los subespacios M_0 ($= \{0\}$), $M_1, \ldots, M_{n-2}, M_{n-1}$ que satisfacen a las condiciones (i) , (ii) y (iii) del lema. Si agregamos a ellos el subespacio $M_n = V$, el teorema quedará demostrado. ■

Teorema 13.2.4. Sea T un endomorfismo de un espacio vectorial *n-dimensional* V sobre el campo $\mathbb{F}$ algebraicamente cerrado. Existe entonces una base $\{e_1, \ldots, e_n\}$ de V, con respecto a la cual la matriz de T es triangular superior, es decir de la forma

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \dots & \alpha_{2n} \\ 0 & 0 & \alpha_{33} & \dots & \alpha_{3n} \\ \vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{nn} \end{pmatrix}$$

que se caracteriza porque todos sus términos situados debajo de la diagonal principal son iguales a cero. Los escalares $\alpha_{11}, \dots, \alpha_{nn}$ son los valores propios, no necesariamente diferentes entre sí, y el polinomio característico de T es $f_T(x) = (x - \alpha_{11}) \dots (x - \alpha_{nn})$.

Demostración. Con las mismas notaciones del lema 13.2.1, consideremos una base $\{e_1\}$ de M_1. Esa base puede ser aplicada hasta formar una base $\{e_1, e_2\}$ de M_2 y luego una base $\{e_1, e_2, e_3\}$ de M_3, y puede seguirse así hasta obtener una base $\{e_1, \dots, e_n\}$ de V. Calculemos ahora la matriz de T respecto de esta base. Puesto que

$$T(M_j) \subset M_j ,$$

la imagen Te_j depende solamente de $e_1, \dots, e_j$. Por tanto la matriz es triangular superior y tendrá la forma dada en el enunciado; de donde resulta inmediatamente el resto del enunciado. ∎

Corolario. Si M es una matriz $n \times n$ sobre un campo $\mathbb{F}$, existe una matriz regular A sobre una cerradura algebraica $\mathbb{F}'$ de $\mathbb{F}$, tal que AMA^{-1} es una matriz triangular superior cuyos términos diagonales son todas las raices características de M. En particular esto tiene lugar cuando $\mathbb{F} = \mathbb{R}$ y $\mathbb{F}' = \mathbb{C}$.

Demostración. Podemos suponer que M es la matriz de un endomorfismo de $\mathbb{F}'^n$. Puesto que $\mathbb{F}'$ es algebraicamente cerrado por hipótesis, puede aplicarse el teorema 13.2.4, y por tanto existe una base de $\mathbb{F}'^n$, respecto de la cuál la matriz T, que es de la forma AMA^{-1} donde A es regular, cumple las condiciones del enunciado. ∎

Definición 13.2.3. Dada una matriz M de tipo $n \times n$ sobre un campo $\mathbb{F}$ se llaman *raices características de* M a las raices $\lambda_1, \dots, \lambda_n$ del

polinomio característico $f_M(x)$ de M, considerado como polinomio sobre una cerradura algebraica $\mathbb{F}'$ de $\mathbb{F}$. $\lambda_1, \ldots, \lambda_n$ son elementos de $\mathbb{F}'$, y en $\mathbb{F}'[x]$ tiene lugar la descomposición

$$f_M(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_n) .$$

Las raices características de M son evidentemente las mismas que las de cada matriz equivalente a ella. Los valores propios de un endomorfismo T de un espacio vectorial V sobre $\mathbb{F}$ son aquellas raices características de una matriz que representa a T, que pertenece a $\mathbb{F}$.

Multiplicidad de una raiz característica es su multiplicidad como raiz del polinomio característico.

Evidentemente puede hablarse también sin ambigüedad de las *raices carateristicas de un endomorfismo T.*

Teorema 13.2.5. Sea T un endomorfismo del espacio *n-dimensional* V sobre $\mathbb{F}$. Si (V', θ) es una extensión de V que corresponde a la extensión $\mathbb{F}'$ de $\mathbb{F}$ (ver el parágrafo § 12), entonces la extensión T' de T a V' (definición 12.5.1) tiene el mismo polinomio característico que T, es decir que $f_{T'} = f_T$.

Demostración. De la observación hecha al final de la sección 12.5 resulta que si $\{a_1, \ldots, a_n\}$ es una base cualquiera de V, la matriz $M = (\alpha_j^i)$ que representa a T respecto de esa base es idéntica a la matriz que representa a T' respecto de la base $\{\theta a_1, \ldots, \theta a_n\}$ de V'. La identidad de esas matrices implica inmediatamente que

$$f_{T'} = f_T \quad . \quad \blacksquare$$

Ejercicios

13.2.1. Hallar el polinomio característico de la matriz

$$\begin{pmatrix} \lambda_o & 0 & 0 & \ldots & 0 & 0 \\ 1 & \lambda_o & 0 & \ldots & 0 & 0 \\ 0 & 1 & \lambda_o & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & \lambda_o \end{pmatrix}$$

*13.2.2. Probar que si M es una matriz cualquiera, su transpuesta M^T tiene el mismo polinomio característico que la matriz M.

13.2.3. Hallar el polinomio característico de la matriz

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

13.2.4. Hallar los polinomios característicos de las siguientes matrices.

a) $\begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}$ b) $\begin{pmatrix} 3 & -1 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix}$ c) $\begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & -3 \end{pmatrix}$

13.2.5. Hallar matrices triangulares superiores semejantes a las matrices

a) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 4 & 5 \\ 2 & 5 & 7 \\ 3 & 6 & 9 \end{pmatrix}$

13.3. Polinomio minimal. Teorema de Cayley–Hamilton cuando $\mathbb{F}$ es algebraicamente cerrado. Sean V un espacio *n-dimensional* sobre un campo $\mathbb{F}$ y T un endomorfismo de V. Dado un polinomio cualquiera de $\mathbb{F}[x]$,

$$g(x) = a_o + a_1 x + \ldots + a_n x^n$$

se llama *valor* de ese polinomio en T, o para $x = T$, al endomorfismo

$$g(T) = a_o I + a_1 T + \ldots + a_n T^n \quad ,$$

donde I designa al automorfismo idéntico de V. Se tiene evidentemente

$$(g + h)(T) = g(T) + h(T) \quad , \quad (gh)(T) = g(T)h(T) \ ,$$

cualesquiera que sean los polinomios g y h de $\mathbb{F}[x]$. En otras palabras, la aplicación de $\mathbb{F}[x]$ es el anillo $\mathfrak{E}(V)$ de los endomorfismos de V, que aplica a cada $g(x) \in \mathbb{F}[x]$ en $g(T)$, es un homomorfismo de $\mathbb{F}[x]$ en $\mathfrak{E}(V)$. El núcleo K de ese homomorfismo es el ideal formado por los polinomios g de $\mathbb{F}[x]$ tales que $g(T) = 0$, es decir por los polinomios que se anulan en T. Puesto que $\mathbb{F}[x]$ es un anillo de ideales principales, existe un polinomio mónico único $\phi(x) \in \mathbb{F}[x]$ tal que $K = I(\phi)$ es el ideal principal generado por ϕ; $\phi(x)$ es por tanto el único polinomio mónico de grado mínimo tal que $\phi(T) = 0$; y todo otro polinomio que se anula en T es un mútiplo de ϕ.

Definición 13.3.1. El polinomio mónico de $\mathbb{F}[x]$ de grado mínimo que se anula para $x = T$ se llama *polinomio minimal de* T y se representa por $\phi_T(x)$.

Teorema 13.3.1. Todos los endomorfismos de la forma $g(T)$, donde $g \in \mathbb{F}[x]$, son conmutables entre sí. En particular son conmutables con T.

Demostración. $g(T)h(T) = (gh)(T) = (hg)(T) = h(T)g(T)$. ∎

Todo lo que en la presente sección hemos expresado hasta ahora acerca del endomorfismo T es válido, palabra por palabra, cuando se sustituyen el endomorfismo T por una matriz cualquiera M de tipo $n \times n$ sobre $\mathbb{F}$, el automorfismo identidad I de V por la matriz unidad E_n, y el anillo $\mathfrak{E}(V)$ de los endomorfismos de V por el anillo $\mathbb{M}(n)$ de las matrices $n \times n$ sobre $\mathbb{F}$. Se cumple además la proposición siguiente:

Teorema 13.3.2. Dado un polinomio $g \in \mathbb{F}[x]$, la relación $g(T)=0$ donde T es un endomorfismo de V es equivalente a la relación $g(M) = 0$, donde M es cualquiera de las matrices que representan a T.

Demostración. Para cada base que se elija en V existe un determi

nado isomorfismo de $\mathcal{L}(V)$ sobre $\mathbb{M}(n)$ (ver pág 73) que aplica a T en la matriz M que lo representa en esa base. Si suponemos que

$$g(T) = a_0 I + a_1 T + \ldots + a_n T^n = 0,$$

entonces, aplicando dicho isomorfismo resultará

$$g(M) = a_0 E_n + a_1 M + \ldots + a_n M^n = 0.$$

La recíproca se demuestra de manera análoga. ■

Corolario. Los polinomios minimales $\phi_T(x)$ y $\phi_M(x)$ son idénticos si M es una matriz que representa al endomorfismo T en alguna base de V.

Ejemplos. A. Sea

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

una matriz sobre $\mathbb{R}$. Se comprueba por cálculo directo que el polinomio

$$\phi(x) = x^2 - 3x + 2 = (x - 2)(x - 1)$$

se anula para $x = M$. Ahora bien $\phi(x)$ no tiene más divisores de primer grado que $x - 2$ y $x - 1$; pero $M - 2$ y $M - 1$ son diferentes de cero, luego $\phi(x)$ es el polinomio minimal de M.

B. Sean

$$M = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \quad \text{y} \quad M_1 = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}, \quad b \neq 0,$$

dos matrices sobre un campo cualquiera $\mathbb{F}$. Los polinomios característicos de esa matriz son

$$f_M(x) = f_{M_1}(x) = (x - a)^2$$

En cambio sus polinomios minimales son

$$\phi_M(x) = x - a \quad \text{y} \quad \phi_{M_1}(x) = (x - a)^2,$$

es decir diferentes entre sí, El hecho de que ϕ_M y ϕ_{M_1} son divisores de f_M y f_{M_1} respectivamente, no es una simple coincidencia co-

mo veremos luego.

La siguiente proposición será empleada en la demostración de los teoremas que siguen.

Lema 13.3.1. Sea T un endomorfismo de V y designemos por $\overline{T}$ a la restricción $T|_{\overline{V}}$ de T a un subespacio invariante $\overline{V}$ respecto de T, es decir a un subespacio tal que

$$T(\overline{V}) \subset \overline{V} \ .$$

Entonces, cualquiera que sea el polinomio $f(x) \in \mathbb{F}[x]$, se tiene que

$$f(T)(\bar{v}) = f(\overline{T})(\bar{v}) \quad , \quad \forall\, \bar{v} \in \overline{V} \ ;$$

o sea que la restricción de $f(T)$ a $\overline{V}$ es $f(\overline{T})$.

Demostración. Es evidente que si se cumple la propiedad enunciada para dos polinomios $f_1(x)$ y $f_2(x)$, se cumple también para su suma $f_1 + f_2$, o bien para el producto de un escalar por cualquiera de ellos. Resulta de esta manera que es suficiente demostrar el lema para los polinomios de La forma $f(x) = x^n$. Emplearemos para ello la inducción respecto de n.

Para $n = 1$, la relación que deseamos demostrar se reduce a

$$T(\bar{v}) = \overline{T}(\bar{v}) \quad ,$$

que se cumple evidentemente para todo $\bar{v} \in \overline{V}$, por definición de T.

Supongamos ahora que sea

$$T^k(\bar{v}) = \overline{T}^k(\bar{v}) \ , \text{ para } \quad k < m \text{ y } \bar{v} \in \overline{V} \ ;$$

es decir que se cumpla la propiedad del lema cuando

$$f(x) = x^k \quad , \quad k < m.$$

Demostraremos que se cumple para $k = m$.

Tendremos

$$T^m(\bar{v}) = T^{m-1}(T(\bar{v})) \ ,$$

y puesto que $T(\bar{v}) \in \overline{V}$ y que $T(\bar{v}) = \overline{T}(\bar{v})$, la hipótesis de inducción nos permite escribir

$$T^m(\bar{v}) = \overline{T}^{m-1}(\overline{T}(\bar{v})) = \overline{T}^m(\bar{v}) \ . \qquad \blacksquare$$

Probaremos a continuación (teorema 13.3.3) un resultado fundamental de la teoría que estamos estudiando. Según este resultado, a una descomposición del polinomio minimal de T en el producto de dos factores mónicos primos entre sí, ϕ_1 y ϕ_2, le corresponde una expresión del espacio V como suma directa de dos subespacios de V invariantes respecto de T, de manera que esos factores son justamente los polinomios minimales de las restricciones de T a esos subespacios. Además, cada uno de los subespacios es el núcleo de uno de los endomorfismos $\phi_1(T)$, $\phi_2(T)$, y es imagen del otro.

Previamente precisaremos lo que se entiende al decir que un espacio es *suma directa* de dos de sus subespacios. Haremos uso para ello de la noción de suma directa de dos espacios vectoriales que fué introducida en el ejercicio 2.5.14.

Definición 13.3.2. Dados un espacio vectorial V y dos de sus subespacios V_1 y V_2, se dice que V es suma directa de V_1 y V_2 y se escribe

$$V = V_1 \oplus V_2$$

si existe un isomorfismo $V_1 \oplus V_2 \to V$ tal que a $(v_1, 0)$ le hace corresponder v_1 y a $(0, v_2)$ le hace corresponder v_2, de modo que a (v_1, v_2) le corresponde $v_1 + v_2$.

Queda al lector como ejercicio demostrar que para que sea

$$V = V_1 \oplus V_2$$

es necesario y suficiente que se cumpla una de las dos condiciones siguientes:

D1. $V = V_1 + V_2$ y $V_1 \cap V_2 = \{0\}$ (ver el ejercicio 2.3.2)

D2. Cada vector v de V puede ser expresado de una y solo una manera en la forma $v = v_1 + v_2$ donde $v_1 \in V_1$ y $v_2 \in V_2$.

Es fácil demostrar además, que si $(a_\alpha)_{\alpha \in A}$ es una base de V_1 y $(b_\beta)_{\beta \in B}$ es una base de V_2 entonces la reunión de ambas bases es una base de V. En particular, si V es de dimensión finita entonces se tiene

$$dim\ V = dim\ V_1 + dim\ V_2$$

Por último, podemos extender las consideraciones precedentes a la suma directa de un número finito de subespacios. En especial, se tendrá que la relación

$$V = V_1 \oplus \ldots \oplus V_k$$

tiene lugar cuando cada vector $v \in V$ puede expresarse de una y solo una manera en la forma

$$v = v_1 + \ldots + v_k \quad , v_i \in V_i \quad (i = 1, \ldots, k) ;$$

o bien cuando

$$V = V_1 + \ldots + V_k \quad \text{y} \quad V_i \cap (V_1 + \ldots + V_{i-1} + V_{i+1} + \ldots + V_k) = \{0\} \quad \text{para} \quad v = 1,2, \ldots, k .$$

(Dejamos al cuidado del lector comprobar que estas dos últimas condiciones son equivalentes).

Lema 13.3.2. Si ϕ es un polinomio de $\mathbb{F}[x]$ y T es un endomorfismo de V, el subespacio $Ker\ \phi(T)$ es invariante respecto de T.

Además si ϕ_1 y ϕ_2 son polinomios primos entre sí, entonces :

1. Cualquiera que sea el vector $v \in V$ existen los vectores v_1 y v_2 del mismo espacio tales que

$$v = \phi_1(T)(v_1) + \phi_2(T)(v_2) ,$$

es decir que

$$V = Im\ \phi_1(T) + Im\ \phi_2(T) .$$

2. Si $\phi_1(T)(v) = \phi_2(T)(v) = 0$, entonces es $v = 0$. Es decir que

$$Ker\ \phi_1(T) \cap Ker\ \phi_2(T) = \{0\} .$$

3. $Ker\ \phi_1(T) \subset Im\ \phi_2(T)$ y $Ker\ \phi_2(T) \subset Im\ \phi_1(T)$.

Demostración. $Ker\ \phi(T)$ es un subespacio invariante respecto de T porque, en virtud del teorema 13.3.1, si v es un vector cualquiera de $Ker\ \phi(T)$, tiene lugar la relación

$\phi(T)(Tv) = (\phi(T)T)v = (T\phi(T))v = T(\phi(T))v = 0$; y por tanto

$$Tv \in Ker\ \phi(T).$$

1. El M.C.D. de ϕ_1 y ϕ_2 es el polinomio 1 (ver la definición 10.4.2); luego, en virtud del teorema 10.4.1, existen los polinomios g y h de $\mathbb{F}[x]$ tales que

$$\phi_1 g + \phi_2 h = 1.$$

Tomando los valores de ambos miembros de la ecuación precedente para $x = T$ se obtiene

$$\phi_1(T)g(T) + \phi_2(T)h(T) = I \quad , \qquad (13.3.1)$$

donde I es el automorfismo identidad de V. Si v es un vector cualquiera de V y definimos los vectores

$$v_1 = g(T)(v)$$
$$v_2 = h(T)(v) \quad ,$$

entonces, si aplicamos a v los dos miembros de (13.3.1), resulta

$$v = \phi_1(T)(v_1) + \phi_2(T)(v_2),$$

como queríamos demostrar.

2. La relación (13.3.1) puede ser escrita, si se tiene en cuenta el teorema 13.3.1, en la forma

$$I = g(T)\phi_1(T) + h(T)\phi_2(T). \qquad (13.2.2)$$

si suponemos que $\phi_1(T)(v) = \phi_2(T)(v) = 0$, y aplicamos a v los dos miembros de (13.3.2), resulta que $v = 0$, lo cual prueba que

$$Ker\ \phi_1(T) \cap Ker\ \phi_2(T) = \{0\}$$

3. Si v es un elemento cualquiera de $Ker\ \phi_1(T)$, es decir que

$$\phi_1(T)(v) = 0 \quad ,$$

y aplicamos a v los dos miembros de (13.3.2) se obtiene (teorema 13.3.1),

$$v = h(T)\phi_2(T)(v) = \phi_2(T)(h(T)(v)) \in Im\ \phi_2(T)$$

Luego es

$$Ker\ \phi_1(T) \subset Im\ \phi_2(T).$$

De manera análoga se prueba que

$$Ker\ \phi_2(T) \subset Im\ \phi_1(T)\ . \quad \blacksquare$$

Teorema 13.3.3. Sea T un endomorfismo de un espacio vectorial *n-dimensional* V sobre un campo $\mathbb{F}$. Supongamos que el polinomio minimal ϕ_T de T es el producto de dos polinomios mónicos ϕ_1 y ϕ_2, sobre $\mathbb{F}$, primos entre sí. Designemos por V_1 y V_2 a los núcleos de $\phi_1(T)$ y $\phi_2(T)$:

$$V_1 = Ker\ \phi_1(T) \quad , \quad V_2 = Ker\ \phi_2(T)\ .$$

Se cumple entonces:

1. V_1 y V_2 son invariantes respecto de T, es decir
$$T(V_1) \subset V_1 \quad , \quad T(V_2) \subset V_2$$
2. $V_1 = Im\ \phi_2(T)$ y $V_2 = Im\ \phi_1(T)$
3. $V = V_1 \oplus V_2$
4. La restricción T_1 (resp. T_2) de T a V_1 (resp. V_2) tiene a $\phi_1(x)(\phi_2(x))$ como polinomio minimal [1)]

Demostración. 1. La invariancia de V_1 y V_2 es consecuencia inmediata de la primera parte del lema 13.3.2.

2. La propiedad 3 del lema 13.3.2 establece que

$$V_1 \subset Im\ \phi_2(T) \quad \text{y} \quad V_2 \subset Im\ \phi_1(T).$$

Por consiguiente bastará con probar las inclusiones inversas de estas. Un elemento cualquiera de $Im\ \phi_2(T)$ puede expresarse en la forma $\phi_2(T)(v)$. Se tiene

$$\phi_1(T)(\phi_2(T)(v)) = [\phi_1(T)\phi_2(T)](v) = \phi_T(T)(v) = 0\ ;$$

por tanto

$$\phi_2(T)(v) \in V_1 \quad ,$$

1) Hemos adoptado aquí el recurso de resumir dos enunciados en uno solo. Entre paréntesis y precedidos de la abreviatura resp. (respectivamente) se dan los términos que deben reemplazarse en el primer enunciado para obtener el segundo. En algunas ocaciones eliminaremos la abreviatura resp.

es decir que

$$Im\ \phi_2(T) \subset V_1$$

Resulta en conclusión que

$$V_1 = Im\ \phi_2(T).$$

Análogamente se puede probar que

$$V_2 = Im\ \phi_1(T).$$

3. Los resultados que acabamos de obtener y la propiedad 1 del lema 13.3.2. permiten escribir $V = V_1 + V_2$; y la propiedad 2 del mismo lema establece que $V_1 \cap V_2 = \{0\}$.

Se tiene por consiguiente que

$$V = V_1 \oplus V_2 .$$

4. Sean T_1 y T_2 las restricciones de T a V_1 y a V_2 respectivamente. T_1 es un endomorfismo de V_1 y T_2 es un endomorfismo de V_2, en virtud de que V_1 y V_2 son subespacios de V, invariantes respecto de T. Designemos por ψ_1 y ψ_2 a los polinomios minimales de T_1 y T_2 respectivamente. Puesto que $\psi_1(T)\psi_2(T)$ coincide con $\psi_1(T_1)\psi_2(T_1)$ sobre V_1 y con $\psi_1(T_2)\psi_2(T_2)$ sobre V_2, se anula sobre ambos subespacios y por tanto se anulará también sobre

$$V = V_1 \oplus V_2$$

Por consiguiente $\psi_1 \psi_2$ pertenece al ideal principal de los polinomios que se anulan en T, ideal que es generado por el polinomio minimal $\phi_T = \phi_1 \phi_2$ de T. Resulta así que $\psi_1 \psi_2$ es divisible por $\phi_1\phi_2$, y existe por tanto el polinomio θ tal que

$$\psi_1\psi_2 = \theta\phi_1\phi_2 . \qquad (13.3.4)$$

Por otra parte, cualquiera que sea $v_1 \in V_1$, el lema 13.3.1 nos permite escribir

$$\phi_1(T_1)(v_1) = \phi_1(T)(v_1) = 0.$$

Por tanto es

$$\phi_1(T_1) = 0.$$

Semejantemente se demuestra que

$$\phi_2(T_2) = 0.$$

Puesto que ψ_1 y ψ_2 son los polinomios minimales de T_1 y T_2, se deduce de las dos últimas ecuaciones que ψ_1 divide a ϕ_1 y ψ_2 divide a ϕ_2; es decir que

$$\phi_1 = \chi_1 \psi_1 \quad , \quad \phi_2 = \chi_2 \psi_2 \, . \tag{13.3.5}$$

Sustituyendo estos valores en (13.3.4) se obtiene

$$\psi_1 \psi_2 = \theta \chi_1 \chi_2 \psi_1 \psi_2 \, .$$

Se deduce de aquí que $\theta \chi_1 \chi_2 = 1$, y por tanto θ, χ_1 y χ_2 son escalares. Dado que ϕ_1, ϕ_2, ψ_1, y ψ_2 son polinomios mónicos, se deberá tener $\theta = \chi_1 = \chi_2 = 1$, y por tanto

$$\phi_1 = \psi_1 \quad , \quad \phi_2 = \psi_2 \, . \qquad \blacksquare$$

Corolario. Supongamos que el polinomio minimal $\phi_T(x)$ del endomorfismo T de V es un producto

$$\phi_1 \, \phi_2 \, \cdots \, \phi_\mu \quad ,$$

donde $\phi_1, \ldots, \phi_\mu$ son polinomios mónicos de $F[x]$, primos entre sí; y consideremos los subespacios

$$V_i = Ker \ \phi_i(T) \qquad i = 1, \ldots, \mu$$

de V.

Se cumple entonces,

1. $T(V_i) \subset V_i$ para $i = 1, \ldots, \mu$, y
$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_\mu \, ,$$
2. El polinomio minimal de la restricción T_i de T a V_i es igual a ϕ_i, para $i = 1, \ldots, \mu$.

Demostración. Emplearemos la inducción respecto de μ. Para $\mu = 2$ el

corolario se reduce al teorema 13.3.3

Supongamos ahora que la proposición se cumple cuando

$$\phi_T = \phi_1 , \ldots , \phi_k \quad , \text{ donde } \quad k < \mu$$

Vamos a probar que se cumple para $k = \mu$.

Escribamos

$$\phi_1 \cdot \phi_2 \cdots \phi_{\mu-1} = \bar{\phi} .$$

En virtud de la hipótesis, el polinomio minimal $\phi_T(x)$ de T es entonces el producto $\bar{\phi}\phi_\mu$ de dos polinomios mónicos primos entre sí. Podemos aplicar entonces el teorema 13.3.3 y establecer que

$$V = \overline{V} \oplus V_\mu \qquad (13.3.6)$$

donde

$$\overline{V} = Ker \ \bar{\phi}(T) ,$$

y

$$V_\mu = Ker \ \phi_\mu(T) \quad , \qquad (13.3.7)$$

$$T(\overline{V}) \subset \overline{V} \qquad (13.3.8)$$

y

$$T(V_\mu) \subset V_\mu \qquad (13.3.9)$$

Además:

$$\phi_{\overline{T}} = \bar{\phi} \quad , \qquad (13.3.10)$$

donde $\overline{T}$ es la restricción de T a $\overline{V}$, y

$$\phi_{T_\mu} = \phi_\mu \quad . \qquad (13.3.11)$$

La restricción $\overline{T}$ de T a $\overline{V}$ es, según (13.3.8), un endomorfismo de $\overline{V}$ cuyo polinomio minimal, en virtud de (13.3.10), es

$$\bar{\phi} = \phi_1 \ \phi_2 \cdots \phi_{\mu-1} \quad .$$

La hipótesis de inducción aplicada a $\overline{T}$ implica que

$$\overline{V} = \overline{V}_1 \oplus \overline{V}_2 \cdots \overline{V}_{\mu-1} \quad , \qquad (13.3.12)$$

donde

$$\overline{V}_i = Ker\ (\phi_i(\overline{T}))$$

y

$$\overline{T}(\overline{V}_i) \subset \overline{V}_i \quad ; \qquad (13.3.13)$$

y el polinomio minimal de la restricción $\overline{T}_i$ de $\overline{T}$ a $\overline{V}_i$ (que es idéntica a la restricción de T a $\overline{V}_i$) es ϕ_i , para $i = 1, \ldots, \mu-1$.

Probaremos ahora que

$$\overline{V}_i = V_i \quad \text{para} \quad i = 1, \ldots, \mu - 1 \ ,$$

o sea que

$$Ker\ (\phi_i(\overline{T})) = Ker\ (\phi_i(T)) \ , \ i = 1, \ldots, \mu-1.$$

Puesto que $\overline{V}$ es invariante respecto de T, el lema 13.3.1 permite escribir

$$\phi_i(\overline{T}) = (\phi_i(T))\ |_{\overline{V}}$$

Por tanto bastará con demostrar que

$$Ker\ (\phi_i(T))\ |_{\overline{V}} = Ker\ (\phi_i(T)) \quad , \quad (i = 1, \ldots, \mu-1) .$$

Ahora bien $\overline{V}$ es el conjunto de los vectores en que se anula el endomorfismo $\phi(T) = (\phi_1(T)) \cdots (\phi_{\mu-1}(T))$. Se tiene por consiguiente que

$$Ker\ (\phi_i(T)) \subset \overline{V} \quad , \qquad (i = 1, \ldots, \mu - 1) ,$$

de donde resulta

$$Ker\ (\phi_i(T))\ |_{\overline{V}} = Ker\ (\phi_i(T))$$

y, según lo antes dicho,

$$\overline{V}_i = V_i \ , \qquad (13.3.14)$$

que es lo que queríamos probar.

Se deduce de (13.3.12), que

$$\overline{V} = V_1 \oplus V_2 \oplus \ldots \oplus V_{\mu-1} \ ,$$

y llevando a (13.3.6) se obtiene que

$$\overline{V} = V_1 \oplus V_2 \oplus \ldots \oplus V_\mu \ .$$

Además, (13.3.13) permite escribir, teniendo en cuenta (13.3.14),

$$T(V_i) = T(\overline{V}_i) = \overline{T}(\overline{V}_i) \subset \overline{V}_i = V_i \quad , \quad (i = 1, \ldots, \mu - 1)$$

Si a este resultado se agrega (13.3.9) quedará establecida la primera parte del teorema.

La relación (13.3.14) implica que la restricción T_i de T a V_i es idéntica a la restricción de T a $\overline{V}_i$, y por tanto, como ya hemos observado antes, a la restricción $\overline{T}_i$ de $\overline{T}$ a $\overline{V}_i$. Por tanto su polinomio minimal será ϕ_i; y esto para $i = 1, \ldots, \mu - 1$. Teniendo en cuenta la relación (13.3.11) quedará establecida la segunda parte del teorema. ■

El teorema siguiente se refiere al caso en que el polinomio minimal del endomorfismo T se descompone en la forma

$$\phi_T(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_\mu)^{r_\mu}$$

donde $\lambda_1, \ldots, \lambda_\mu$ son escalares de $\mathbb{F}$, distintos entre sí, y los r_i son números enteros positivos. Esa situación se presenta en particular, para todo endomorfismo cuando $\mathbb{F}$ es un campo algebraicamente cerrado. Por ejemplo, cuando $\mathbb{F}$ es el campo C de los números complejos.

Teorema 13.3.4. Sea T un endomorfismo de un espacio *n-dimensional* V sobre un campo $\mathbb{F}$

A) Si el polinomio minimal de T puede descomponerse en la forma

$$\phi_T(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_\mu)^{r_\mu} \quad , \qquad (13.3.15)$$

donde los escalares $\lambda_1, \ldots, \lambda_\mu$ son distintos entre sí y los $r_1, \ldots, r_\mu$ son números enteros positivos, entonces se tiene que

1. $V = V_1 \oplus V_2 \oplus \cdots \oplus V_\mu \qquad (13.3.16)$

donde $V_i = Ker\ (T - \lambda_i)^{r_i}$,

y $T(V_i) \subset V_i \qquad (13.3.17)$

2. El polinomio minimal de la restricción de T a V_i es $(x - \lambda_i)^{r_i}$ para cada $i = 1, \ldots, \mu$.

3. $\{\lambda_1, \ldots, \lambda_\mu\}$ es el conjunto de todos los valores propios de T.

4. $n_i = dim\ V_i \geqslant 1$.

B) Si $\mathbb{F}$ es algebraicamente cerrado, entonces tiene lugar la descomposición de la forma (13.3.15) y por tanto se cumplen las cuatro propiedades anteriores y también:

5. El polinomio característico de la restricción T_i de T a V_i es

$$(x - \lambda_i)^{n_i} \quad , \text{ donde } \quad n_i \geqslant r_i$$

Además el polinomio característico de T es

$$f_T(x) = (x - \lambda_1)^{n_1} \ldots (x - \lambda_\mu)^{n_\mu} \qquad (13.3.18)$$

y por tanto el polinomio minimal es un divisor del polinomio característico.

Demostración. A). Las propiedades 1. y 2. son consecuencia inmediata del corolario del teorema 13.3.3.

Vamos a demostrar las propiedades 3. y 4. En primer término probaremos que cada uno de los escalares $\lambda_1, \ldots, \lambda_\mu$ es valor propio de T. Probemos que λ_i es un valor propio de T. Basta demostrar para ello que

$$Ker\ (T - \lambda_i) \neq \{0\} \ .$$

Ahora bien, puesto que

$$\phi_T(T) = (T - \lambda_i)\ \overline{\phi}(T) = 0 \quad ,$$

donde $\overline{\phi}$ es un polinomio de grado menor que el de ϕ_T, se tiene

$$Im\ \overline{\phi}(T) \subset Ker\ (T - \lambda_i) \ .$$

Pero, puesto que ϕ_T es el polinomio minimal de T, debe ser $\overline{\phi}(T) \neq 0$, y por tanto $Ker\ (T - \lambda_i) \neq 0$.

Por otra parte,

$$\phi_T(T) = (T - \lambda_i)^{r_i}\ \overline{\phi}(T) = 0 \ ,$$

donde $\bar{\phi}_i(T)$ es tal que

$$\bar{\phi}(T) = (T - \lambda_i)^{r_i - 1} \bar{\phi}_i(T).$$

Por tanto debe ser también $\bar{\phi}_i(T) \neq 0$, y se tiene

$$Im\ \bar{\phi}_i(T) \subset Ker\ (T - \lambda_i)^{r_i} = V_i \ .$$

Se deduce de aquí que $V_i \neq \{0\}$, y por tanto que

$$n_i = dim\ V_i \geqslant 1 \ .$$

Vamos a probar ahora, para dejar totalmente establecida la propiedad 3, que si λ es un valor propio de T, entonces es igual a uno de los escalares $\lambda_1, \ldots, \lambda_\mu$. En efecto, sea $v \neq 0$ tal que

$$Tv = \lambda v \ .$$

Según la propiedad 1 del enunciado, que ya fué demostrada, puede escribirse

$$v = v_1 + \ldots + v_\mu \quad , \quad v_i \in V_i \qquad (i = 1, \ldots, \mu).$$

Puesto que $v \neq 0$ es $v_k \neq 0$ para algún índice k. Probaremos que $\lambda = \lambda_k$.

Se tiene

$$\lambda v_1 + \ldots + \lambda v_\mu = \lambda v = Tv = Tv_1 + \ldots + Tv_\mu \ .$$

Puesto que $Tv_i \in V_i$, la propiedad 1 implica que $\lambda v_i = Tv_i$, $i = 1, \ldots \mu$, y por tanto

$$\lambda v_k = Tv_k \qquad (13.3.19)$$

Ahora bien, por la definción (13.3.16) de V_k es

$$(T - \lambda_k)^{r_k} (v_k) = 0 \quad ;$$

o bien, desarrollando por la fórmula del binomio

$$\sum_{s=0}^{r_k} \binom{r_k}{s} (-\lambda_k)^s\ T^{r_k - s} (v_k) = 0 \ .$$

Teniendo en cuenta (13.3.19) puede escribirse

$$\sum_{s=0}^{r_k} \binom{r_k}{s} (-\lambda_k)^s \lambda^{r_k - s} v_k = 0 ,$$

es decir

$$(\lambda - \lambda_k)^{r_k} v_k = 0 .$$

Puesto que $v_k \neq 0$, se sigue que $\lambda = \lambda_k$.

B). Vamos a suponer ahora que $\mathbb{F}$ es algebraicamente cerrado. En este caso $\phi_T(x)$ puede descomponerse en la forma que fue supuesta en la parte A). del enunciado. Por tanto se cumplen las propiedades 1. a 4. Demostraremos ahora la propiedad 5.

El polinomio minimal de T_i, que es la restricción de T a V_i, es como ya vimos $(x - \lambda_i)^{r_i}$. Luego, por la propiedad A3, ya demostrada, el único valor propio de T_i es λ_i. De conformidad con el teorema 13.2.4 puede elegirse en V_i una base respecto de la cual la matriz de T_i, que es de orden n_i, sea triangular superior con todos sus términos diagonales iguales a λ_i .

Se deduce de aquí inmediatamente que

$$f_{T_i}(x) = (x - \lambda_i)^{n_i} .$$

Además, la matriz de $T_i - \lambda_i$ respecto de la misma base es también de orden n_i y triangular superior, pero con todos los términos diagonales iguales a cero. La potencia n_i-ésima de esa matriz es igual a cero, como es fácil comprobar. Luego se tiene

$$(T_i - \lambda_i)^{n_i} = 0 ,$$

es decir que el polinomio $(x - \lambda_i)^{n_i}$ se anula en T_i, y por consiguiente debe ser divisible por $(x - \lambda_i)^{r_i}$ que es el polinomio minimal de T_i. Debe ser por tanto $n_i \geqslant r_i$.

Finalmente, si tomamos como base de V a la reunión de las bases de los subespacios V_i, respecto de las cuales las matrices M_i

de las restricciones T_i son triangulares superiores, entonces la matriz de T es

$$\begin{pmatrix} M_1 & 0 & \cdots & 0 \\ 0 & M_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_\mu \end{pmatrix} ,$$

que también es triangular superior, y en cuya diagonal principal aparece n_i veces cada valor propio λ_i de T. Se deduce inmediatemente que

$$f_T(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_\mu)^{n_\mu} . \quad \blacksquare$$

Teorema 13.3.5. Si T es un endomorfismo de un espacio vectorial sobre un campo $\mathbb{F}$ algebraicamente cerrado ϕ_T divide a f_T. Además f_T divide a $(\phi_T(x))^n$.

Demostración. Puesto que $\mathbb{F}$ es algebraicamente cerrado, ϕ_T puede expresarse en la forma (13.3.15), donde $\lambda_1, \ldots, \lambda_\mu$ son sus distintas raices. Se deduce entonces de la propiedad 5 de ese teorema que f_T es divisible por ϕ_T.

Puesto que es $r_i \geqslant 1$ y $n \geqslant n_i$, se tiene $nr_i \geqslant n_i$, y se deduce de (13.3.15) y (13.3.18) que f_T divide a

$$(\phi_T(x))^n . \quad \blacksquare$$

Teorema 13.3.6. (*Teorema de Cayley-Hamilton en el caso en que el campo $\mathbb{F}$ es algebraicamente cerrado*) Si T es un endomorfismo de un espacio vectorial sobre un campo algebraicamente cerrado, entonces es

$$f_T(T) = 0 ,$$

es decir que T anula a su polinomio característico.

Demostración. Puesto que ϕ_T divide a f_T y que $\phi_T(T) = 0$, se tendrá que

$$f_T(T) = 0 . \quad \blacksquare$$

13.4. Suma directa de transformaciones lineales.

Definición 13.4.1. Sean V_1, V_2, W_1 y W_2 cuatro espacios vectoriales sobre $\mathbb{F}$. Dadas las transformaciones lineales $T_1: V_1 \to W_1$ y $T_2: V_2 \to W_2$, se llama *suma directa de* T_1 *y* T_2, y se designa por $T_1 \oplus T_2$ a la transformación lineal de $V_1 \oplus V_2$ en $W_1 \oplus W_2$ definida por

$$(T_1 \oplus T_2)(v_1, v_2) = (T_1 v_1, T_2 v_2), \text{ para } (v_1, v_2) \in V_1 \oplus V_2$$

En el caso particular en que V_1 y V_2 son subespacios de un espacio V y W_1 y W_2 son subespacios de un espacio W, entonces $V_1 \oplus V_2$, según la definición 13.3.2, es el subespacio de V formado por los vectores $v_1 + v_2$, en que $v_1 \in V_1$ y $v_2 \in V_2$. Análogamente $W_1 \oplus W_2$ es un subespacio de W. Se tiene entonces que

$$(T_1 \oplus T_2)(v_1 + v_2) = T_1 v_1 + T_2 v_2 .$$

Las definiciones precedentes se extienden sin dificultad a la suma directa de un número finito cualquiera de espacios.

Teorema 13.4.1. Si $T = T_1 \oplus T_2$ es un endomorfismo de un espacio V de dimensión finita entonces el conjunto de los valores propios de T es la reunión de los conjuntos de valores propios de T_1 y T_2.

Demostración. Por definición de $T_1 \oplus T_2$ existen los subespacios V_1 y V_2 de V tales que $V = V_1 \oplus V_2$, que T_1 es un endomorfismo de V_1, que T_2 es un endomorfismo de V_2 y que

$$T(v_1 + v_2) = T_1 v_1 + T_2 v_2$$

para $v_1 \in V_1$ y $v_2 \in V_2$. Sea λ un valor propio de T y $v = v_1 + v_2$ sea uno de los vectores propios correspondientes. Se tendrá por consiguiente $T_1 v_1 + T_2 v_2 = \lambda v_1 + \lambda v_2$; y puesto que $T_1 v_1 \in V_1$ y $T_2 v_2 \in V_2$, deberá ser $T v_1 = \lambda v_1$ y $T v_2 = \lambda v_2$. Puesto que $v \neq 0$,

v_1 y v_2 no pueden ser ambos nulos, luego λ es valor propio de T_1 o de T_2 .

Recíprocamente es evidente que si λ es valor propio de T_1 o de T_2 lo será de $T_1 \oplus T_2$. ∎

13.5. Reducción del estudio de los endomorfismos al de los endomorfismos nilpotentes.

Con las hipótesis y las notaciones del teorema 13.3.4 resulta que el endomorfismo $T_1 \oplus \cdots \oplus T_\mu$ de $V_1 \oplus \cdots \oplus V_\mu = V$ es idéntico al endomorfismo T , pues para $v = v_1 + \cdots + v_k$

$$\begin{aligned}(T_1 \oplus \cdots \oplus T_\mu)(v_1 + \cdots + v_k) &= T_1 v_1 + \cdots + T_\mu v_\mu \\ &= Tv_1 + \cdots + Tv_\mu \\ &= T(v_1 + \cdots + v_\mu) = Tv\end{aligned}$$

Se tiene por consiguiente

$$T = T_1 \oplus \cdots \oplus T_\mu \quad .$$

De esta manera, T se descompone en la suma de endomorfismos T_i cada uno de los cuales satisface a una relación

$$(T_i - \lambda_i)^{r_i} = 0 \quad .$$

Si escribimos

$$B_i = T_i - \lambda_i \quad ,$$

entonces T_i puede expresarse en la forma

$$T_i = \lambda_i + B_i \quad ,$$

donde $B_i^{r_i} = 0$. Esta última relación se expresa diciendo que B_i es *nilpotente*. Resulta así, en resumen, que T es suma directa de los endomorfismos $T_1, \ldots, T_\mu$, cada uno de los cuales es suma de un endomorfismo de la forma λI más un endomorfismo nilpotente. Puede reconocerse así que el estudio de un endomorfismo cuyo polinomio minimal es producto de potencias de polinomios mónicos de primer

grado se reduce en último término al estudio de los endomorfismos nilpotentes. Esto es cierto, en particular, para los endomorfismos de un espacio vectorial sobre un campo algebraicamente cerrado.

La obtención de una representación material canónica de un endomorfismo nilpotente nos permitirá, en particular, construir una base de V_i respecto de la cual sea canónica la representación de B_i. También tendrá entonces una cierta forma canónica la representación del endomorfismo $T_i = \lambda_i + B_i$. Ahora bien, respecto de la base de V formada por las bases de V_i elegidas como acabamos de ver, la matriz de $T = T_1 \oplus \ldots \oplus T_\mu$ es diagonal de bloques, con bloques diagonales que son las matrices canónicas de los endomorfismos T_i. Se obtendrá así una representación canónica de T. Es este el procedimiento que más tarde seguiremos para obtener la matriz de Jordan de un endomorfismo T. Los dos parágrafos siguientes serán dedicados al estudio de los endomorfismos nilpotentes y al de la representación matricial canónica.

13.6. El teorema de Cayley-Hamilton en el caso de campos no algebraicamente cerrados.

Los teoremas siguientes constituyen la generalización de los teoremas 13.3.5 y 13.3.6, en el caso en que el espacio V es sobre un campo $\mathbb{F}$ no algebraicamente cerrado.

Teorema 13.6.1. El polinomio minimal ϕ_T de un endomorfismo T de un espacio *n-dimensional* V, sobre un campo cualquiera $\mathbb{F}$, divide al polinomio característico f_T de ese endomorfismo. Además, f_T divide a $(\phi_T(x))^n$.

Demostración. Sea $\mathbb{F}'$ la cerradura algebraica de $\mathbb{F}$ (sección 11.3), y sea V' la extensión de V correspondiente a la extensión $\mathbb{F}'$ de $\mathbb{F}$ (sección 12.2). Si T' es la extensión de T al espacio V', entonces, en virtud del teorema 13.3.5, los polinomios $f_{T'}$ y $\phi_{T'}$ tienen las propiedades de nuestro enunciado. Además, según el teorema 13.2.5, se tiene que $f_T = f_{T'}$. Luego se tiene

$$f_{T'}(T') = f_T(T') = 0 \quad ;$$

y puesto que T es la restricción de T' a V (lo que es consecuencia de la definición 12.5.1), se tiene que $f_T(T) = 0$ (esta relación constituye el teorema de Cayley-Hamilton que es el teorema 13.6.2 que se enuncia a continuación). Resulta así que $f_T(x)$ pertenece al ideal de los polinomios de $\mathbb{F}[x]$ que tienen la raiz T, y por consiguiente $f_T(x)$ *será divisible por el polinomio minimal* $\phi_T(x)$, lo que prueba la primera parte del teorema.

Por otra parte, puesto que $\phi_T(T) = 0$, podemos demostrar que $\phi_T(T') = 0$. En efecto, si $\{v_1, \ldots, v_n\}$ es una base de V y θ es la aplicación de extensión de V a V', entonces $\{\theta v_1, \ldots, \theta v_n\}$ es una base de V', y por la observación de la sección 12.4, las matrices M_T de T respecto de la base $\{v_i\}$, y $M_{T'}$ de T', respecto de la base $\{\theta v_i\}$ son idénticas.

La relación $\phi_T(T) = 0$ implica $\phi_T(M_T) = 0$ y por tanto

$$\phi_T(M_{T'}) = 0.$$

Se sigue de aquí que

$$\phi_T(T') = 0 \ ,$$

como deseabamos establecer. Se deduce de aquí que $\phi_{T'}$, polinomio minimal de T', perteneciente a $\mathbb{F}'[x]$, divide a ϕ_T, polinomio de

$$\mathbb{F}[x] \subset \mathbb{F}'[x] \ .$$

Por consiguiente $(\phi_{T'})^n \in \mathbb{F}'[x]$ divide a $(\phi_T)^n \in \mathbb{F}[x]$, cuando ambos se consideran como polinomios de $\mathbb{F}'[x]$.

En virtud del teorema 13.3.5 $f_{T'}$ divide a $(\phi_{T'})^n$ sobre $\mathbb{F}'$. Puesto que $f_{T'} = f_T$, podemos decir que f_T divide a $(\phi_{T'})^n$ sobre $\mathbb{F}'$; luego f_T dividirá a $(\phi_T)^n$ sobre $\mathbb{F}$. [1] ■

---

1) Empleamos aquí la proposición siguiente: Si $\mathbb{F}'$ es extensión del campo $\mathbb{F}$, y si el polinomio $f \in \mathbb{F}[x]$ divide al polinomio $g \in \mathbb{F}[x]$, considerados ambos como polinomios de $\mathbb{F}'[x]$, entonces lo divide también en $\mathbb{F}[x]$.

Demostración. La división de g entre f, sobre $\mathbb{F}$, da lugar a la relación

$$g = qf + r \qquad (r = 0 \ , \text{ o } \ grad\ r < grad\ f) \ .$$

Los polinomios que aquí figuran pertenecen también a $\mathbb{F}'[x]$, y por tanto esa misma relación expresa la división de g entre f, sobre $\mathbb{F}'$. Puesto que, por hipótesis, g es divisible por f sobre $\mathbb{F}'$, se deberá tener que $r = 0$. ■

Teorema 13.6.2. (*Teorema de Cayley-Hamilton*).

El polinomio característico $f_T(x)$ de un endomorfismo T se anula para $x = T$, es decir que

$$f_T(T) = 0.$$

Demostración. La relación $f_T(T) = 0$ fué establecida en el curso de de la demostración del teorema anterior, como hicimos notar en su oportunidad. ■

Ejercicios

***13.6.1.** Si T_1 y T_2 son dos endomorfismos conmutables es decir tales que $T_1T_2 = T_2T_1$, entonces el núcleo de cualquiera de ellos es invariante respecto del otro.

13.6.2. Si $T: V \to V'$, $T_1: V' \to V'$ y $T_2: V \to V$ son tres transformaciones lineales, y si $T_1T = TT_2$, entonces cualquiera que sea el polinomio f se tiene

$$f(T_1)T = Tf(T_2) .$$

13.6.3. Hallar los polinomios minimales de las siguientes matrices

(a) $\begin{pmatrix} -1 & 1/2 & -1/2 \\ 3 & -1/2 & -3/2 \\ -5 & -5/2 & 1/2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} a & b & c \\ 0 & a & b \\ 0 & 0 & a \end{pmatrix}$

13.6.4. Demostrar que una matriz M de tipo $n \times n$ y su traspuesta M^t, tienen el mismo polinomio minimal.

(Sugerencia: probar que $(AB)^t = B^tA^t$ y que

$$(A + B)^t = A^t + B^t .$$

Si $f(M) = 0$, donde f es un polinomio, entonces es $f(M^t) = 0$, y viceversa).

*13.6.5. Si

$$M = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \dots & 0 \end{pmatrix},$$

es una matriz superdiagonal de orden n cuya diagonal principal está formada por ceros, se cumple que

$$M^n = 0.$$

13.6.6. Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ un endomorfismo cuya matriz respecto de la base canónica es

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

¿Cuáles son sus valores propios?

13.6.7. El mismo problema suponiendo que es $T: \mathbb{C}^2 \to \mathbb{C}^2$.

13.6.8. a) Hallar los valores propios λ y μ y dos correspondientes vectores propios z_λ y z_μ del endomorfismo cuya matriz es

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

Determinar la matriz de ese endomorfismo referido a la base $\{z_\lambda, z_\mu\}$.

b) Lo mismo para la matriz $\begin{pmatrix} 2 & -3 \\ -1 & 4 \end{pmatrix}$

13.6.9. Hallar los valores propios y los correspondientes vectores propios de los endomorfismos de espacios complejos cuyas matrices son:

(a) $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

13.6.10. Probar que si λ es un valor propio de T, $\lambda + \mu$ es un valor propio de $T + \mu I$.

13.6.11. ¿Cuál es la relación entre los valores propios de T y los de αT?

13.6.12. Hallar la relación entre los valores propios de una matriz M sobre $\mathbb{C}$ y los de su inversa M^{-1}.

13.6.13. Hallar la relación entre los valores propios de una matriz M sobre $\mathbb{C}$ y los de su cuadrado M^2.

13.7. Representación de un endomorfismo mediante una matriz diagonal.

Teorema 13.7.1. Sea V un espacio vectorial *n-dimensional* sobre un campo cualquiera $\mathbb{F}$. Un endomorfismo T de V puede ser representado por una matriz diagonal, respecto de una base apropiada de V, si y solo si $\phi_T(x)$, polinomio minimal de T, es un producto de factores mónicos de primer grado distintos entre sí, es decir si, $\phi_T(x) = (x - \lambda_1) \ldots . (x - \lambda_\mu)$, donde $\lambda_1, \ldots, \lambda_\mu$ son todos los valores propios de T distintos entre sí. Entonces los términos de la diagonal principal de la matriz diagonal de T son dichos valores propios.

Demostración. Supongamos, en primer término, que existe una base de V con respecto a la cual T se representa por una matriz diagonal. Reordenando, si fuera necesario, el orden de los vectores de base, la matriz diagonal puede suponerse que tiene la forma siguiente

$$M = \begin{pmatrix} \lambda_1 E_{n_1} & & & 0 \\ & \lambda_2 E_{n_2} & & \\ & & \ddots & \\ 0 & & & \lambda_\mu E_{n_\mu} \end{pmatrix}.$$

donde $\lambda_1, \ldots, \lambda_\mu$ son los valores propios de T distintos entre sí, E_{n_i} denota a la matriz unidad de tipo $n_i \times n_i$, y

$$n_1 + \cdots + n_\mu = \dim V = n \ .$$

Tendremos entonces

$$\lambda_1 E_n - M = \begin{pmatrix} 0_{n_1} & & & 0 \\ & (\lambda_1 - \lambda_2)E_{n_2} & & \\ & & \ddots & \\ 0 & & & (\lambda_1 - \lambda_\mu)E_{n_\mu} \end{pmatrix}$$

donde 0_{n_i} es la matriz cero de tipo $n_i \times n_i$. Una representación análoga puede obtenerse para cada matriz $\lambda_i E_n - M$, para $2 \leqslant i \leqslant \mu$; de allí resulta entonces que

$$(\lambda_1 E_n - M)(\lambda_2 E_n - M) \cdots (\lambda_\mu E_n - M) = 0 \qquad 1)$$

M es por consiguiente una raiz del polinomio

$$(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_\mu) \ ,$$

el cual, en virtud del teorema 13.6.1, será un múltiplo del polinomio minimal de M, es decir del de T. Se tendrá por tanto

$$\phi_T(x) = (x - \lambda_{i_1})(x - \lambda_{i_2}) \ldots (x - \lambda_{i_r})$$

donde $\lambda_{i_1}, \ldots, \lambda_{i_r}$ son distintos entre sí.

1) Recordaremos aquí que para calcular el producto AB de la matriz A de tipo $m \times n$ por la matriz B de tipo $n \times p$, las matrices A y B pueden ser descompuestas en submatrices por medio de lineas verticales y horizontales. La partición horizontal de A así como la partición vertical de B pueden ser arbitrarias; pero la partición vertical de A y la partición horizontal de B deben ser idénticas. Hechas esas particiones, el producto AB puede formarse como si las submatrices fueran elementos ordinarios. Puede consultarse a este respecto cualquier obra sobre matrices; por ejemplo: *Frazer, Duncan y Collar* Elementary Matrices, Cambridge, 1938 pág 13; o bien *Bodewig*, Matrix Calculus, North-Holland Publishing Co., (1956), pág 14.

Ahora bien, si aplicamos la parte 3 del teorema 13.3.4, los escalares $\lambda_{i_1}, \ldots, \lambda_{i_r}$ deben ser los valores propios de T. Pero estos son

$$\lambda_1, \ldots, \lambda_\mu .$$

Resulta así que $r = \mu$ y que $\lambda_{i_1}, \ldots, \lambda_{i_r}$ son los $\lambda_1, \ldots, \lambda_\mu$ aunque quizás en otro orden. Resulta por tanto que

$$\phi_T(x) = (x - \lambda_1) \cdots (x - \lambda_\mu) \qquad (13.7.1)$$

Supongamos ahora, inversamente, que ϕ_T tiene la forma (13.7.1). Si aplicamos la parte A) del teorema 13.3.4, tendremos que

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_\mu \quad ,$$

donde

$$V_i = Ker\ (T - \lambda_i) \ , \qquad 1 \leqslant i \leqslant \mu \quad ,$$

Si tomamos una base de V formada por la reunión de bases de $V_1, \ldots, V_\mu$, entonces, puesto que para cada vector v_{ir} de la base de V_i se tiene

$$T(v_{ir}) = \lambda_i v_{ir} \ ,$$

la matriz que representa a T en dicha base de V será de la forma

$$\begin{pmatrix} \lambda_1 E_{n_1} & & & 0 \\ & \lambda_2 E_{n_2} & & \\ & & \ddots & \\ 0 & & & \lambda_\mu E_{n_\mu} \end{pmatrix}$$

donde $n_i = dim\ V_i$, $1 \leqslant i \leqslant \mu$. ∎

Corolario. Si M es una matriz $n \times n$ sobre el campo $\mathbb{F}$, existe una matriz regular $n \times n$, A, tal que AMA^{-1} es una matriz diagonal si y solo si el polinomio minimal ϕ_M es un producto de factores mónicos de primer grado, pertenecientes a $\mathbb{F}[x]$, distintos entre sí.

Observación. Resumiendo lo dicho, si

$$\phi_M = (x - \lambda_1) \cdots (x - \lambda_\mu)$$

donde $\lambda_1, \ldots, \lambda_\mu$ son distintos entre sí, entonces y solo entonces existe, de conformidad con el corolario anterior, una matriz regular A tal que AMA^{-1} es diagonal. Según el teorema 13.6.1, f_M divide a una potencia de ϕ_M, y por tanto

$$f_M = (x - \lambda_1)^{n_1} \cdots (x - \lambda_\mu)^{n_k}$$

donde $n_1 + \cdots + n_k = n$. Luego la matriz diagonal AMA^{-1} es de la forma

$$\begin{pmatrix} \lambda_1 E_{n_1} & & 0 \\ & \lambda_2 E_{n_2} & \\ & & \ddots \\ 0 & & \lambda_\mu E_{n_\mu} \end{pmatrix},$$

pues debe tener por elementos diagonales a los λ_i, cada uno con el orden de multiplicidad que le corresponde. Además n_i es la dimensión del subespacio V_i, núcleo de la transformación cuya matriz es $M - \lambda_i E_n$

Definición 13.7.1. Si el polinomio $g(x) \in \mathbb{F}[x]$ es un producto de factores mónicos de primer grado, distintos entre sí:

$$g(x) = (x - \lambda_1) \cdots (x - \lambda_\mu) \quad , \quad \lambda_1, \ldots, \lambda_\mu \in \mathbb{F},$$

se dice que g tiene *raíces simples* en $\mathbb{F}$. Es decir que $\lambda \in \mathbb{F}$ es *raíz simple* de $g(x)$ si $g(x) = (x - \lambda) g_1(x)$ donde $g_1(x) \in \mathbb{F}[x]$ no es divisible por $x - \lambda$.

Por tanto, la condición expresada en el teorema 13.7.1 y en su corolario puede enunciarse diciendo que ϕ_T o ϕ_M tienen raíces simples en $\mathbb{F}$.

Ejemplos. A. Consideremos la matriz sobre $\mathbb{R}$

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

Vamos a determinar su polinomio minimal.

El polinomio característico es

$$f_M = \begin{vmatrix} (x-1) & 0 & -1 \\ 0 & (x-1) & 0 \\ -1 & 0 & (x-1) \end{vmatrix} = (x-1)^3 - (x-1)$$

$$= (x-1)x(x-2).$$

Del teorema 13.3.5 se sigue que el polinomio minimal coincide con f_M, es decir que

$$\phi_M = (x-1)x(x-2).$$

Existirá una matriz regular A tal que

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

Representemos por T al endomorfismo de $\mathbb{R}^3$ cuya matriz, respecto de la base canónica, es M. Empleando las notaciones del teorema 13.3.4, tendremos que

$$V_1 = Ker\ (T-2)$$
$$V_2 = Ker\ (T-1)$$
$$V_3 = Ker\ T.$$

De la parte final de la demostración del teorema 13.7.1 resulta que una base respecto de la cual puede representarse T en forma diagonal puede formarse por reunión de bases de V_1, V_2 y V_3. Tratemos por tanto de obtener bases de estos espacios.

V_1 es el conjunto de los vectores (x_1, x_2, x_3) tales que

$$\begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,$$

o sea

$$-x_1 + x_3 = 0$$
$$- x_2 = 0$$
$$x_1 - x_3 = 0 .$$

Tales vectores son por tanto de la forma $(x_1 , 0 , x_1)$, donde x_1 es arbitrario. La base de V_1 que podemos elegir está formada por el so lo vector $v_1 = (1 , 0, 1)$.

V_2 es el conjunto de los vectores (x_1 , x_2 , x_3) tales que

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 ,$$

o sea

$$x_3 = 0$$
$$x_1 = 0 .$$

Esos vectores son por tanto de la forma $(0 , x_2 , 0)$ donde x_2 es ar bitrario. La base de V_2 está formada por $v_2 = (0 , 1 , 0)$.

V_3 es el conjunto de los vectores (x_1 , x_2 , x_3) tales que

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 ,$$

o sea

$$x_1 + x_3 = 0$$
$$x_2 = 0$$
$$x_1 + x_3 = 0 .$$

Son por tanto los vectores de la forma $(x_1 , 0 , -x_1)$. La base de V_3 será $v_3 = (1 , 0 , -1)$.

Quiere decir que la base buscada está formada por

$$v_1 = e_1 + e_3$$
$$v_2 = e_2$$

$$v_3 = e_1 - e_3$$

Según sabemos la matriz de cambio de la base (v_1 , v_2 , v_3) en la base (e_1 , e_2 , e_3) (ver la sección 6.10) es

$$A^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} ,$$

cuya inversa A, que es la matriz de cambio de la base (e_1 , e_2 , e_3) es la base (v_1 , v_2 , v_3), la cual puede deducirse también de las relaciones

$$e_1 = \frac{1}{2} v_1 + \frac{1}{2} v_3 \quad , \quad e_2 = v_2 \quad , \quad e_3 = \frac{1}{2} v_1 - \frac{1}{2} v_3 \quad ,$$

es

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

Se comprueba por cálculo directo que

$$AMA^{-1} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Ejemplo B. Consideremos ahora la matriz sobre $\mathbb{R}$

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

y designemos por T al endomorfismo de $\mathbb{R}^3$ cuya matriz respecto de la base canónica (e_1 , e_2 , e_3) es M.

El polinomio característico de M es

$$f_M(x) = \begin{vmatrix} (x - 5) & 6 & 6 \\ 1 & (x - 4) & -2 \\ -3 & 6 & (x + 4) \end{vmatrix} = (x - 1)(x - 2)^2 .$$

Por el teorema 13.3.5, $\phi_M(x)$ es o bien $(x - 1)(x - 2)$ o bien $f_M(x)$. Comprobemos si el primero se anula para $x = M$. Se tiene

$$(M - E_3)(M - 2E_3) = \begin{pmatrix} 4 & -6 & -6 & 3 & -6 & -6 \\ -1 & 3 & 2 & -1 & 2 & 2 \\ 3 & -6 & -5 & 3 & -6 & -6 \end{pmatrix} = 0.$$

Por consiguiente se tiene

$$\phi_M(x) = (x - 1)(x - 2) .$$

En virtud de nuestras consideraciones precedentes, partcularmente de la observación que antecede a la definición 13.7.1. existe una matriz regular A tal que

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

Vamos a determinar la matriz A.

Según el teorema 13.3.3

$$V_1 = Ker\ (T - 1) = Im\ (T - 2)$$
$$V_2 = Ker\ (T - 2) = Im\ (T - 1).$$

La matriz de $T_2 = T - 2$ es

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

Luego la imagen de T_2 es generada por los vectores

$$T_2 e_1 = (3 , -1 , 3)$$
$$T_2 e_2 = (-6 , 2 , -6)$$
$$T_2 e_3 = (-6 , 2 , -6).$$

Es evidente que los dos últimos dependen linealmente del primero.

Luego V_1 admite como base a $v_1 = (3, -1, 3)$.

Análogamente, la matriz de $T_1 = T - I$ es

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

Luego la imagen de T_1 es generada por

$$T_1 e_1 = (4, -1, 3), \quad T_1 e_2 = (-6, 3, -6), \quad T_1 e_3 = (-6, 2, -5).$$

Estos tres vectores son linealmente dependientes, pues sus componentes constituyen una matriz cuyo determinante resulta ser igual a cero. En cambio los dos primeros vectores son linelamente independientes porque la matriz de sus componentes

$$\begin{pmatrix} 4 & -6 \\ -1 & 3 \\ 3 & -6 \end{pmatrix}$$

tiene rango 2 ya que su submatriz 2×2

$$\begin{pmatrix} 4 & -6 \\ -1 & 3 \end{pmatrix}$$

tiene determinate diferente de cero. Así pues los vectores

$$v_2 = (4, -1, 3), \quad v_3 = (-6, 3, -6)$$

forman una base de V_2.

De esa manera hemos obtenido la base de V

$$v_1 = (3, -1, 3)$$
$$v_2 = (4, -1, 3)$$
$$v_3 = (-6, 3, -6)$$

respecto de la cual, según demostramos al probar el teorema 13.7.1, T tiene una matriz diagonal.

La matriz de cambio de la base (v_1, v_2, v_3) en la base canónica (e_1, e_2, e_3) es

$$A^{-1} = \begin{pmatrix} 3 & 4 & -6 \\ -1 & -1 & 3 \\ 3 & 3 & -6 \end{pmatrix} ,$$

cuya inversa, que es la matriz de cambio de la base (e_1 , e_2 , e_3) en la base (v_1 , v_2 , v_3), es

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

En efecto, mediante el cálculo comprobamos finalmente que, referida la transformación T a la nueva base (v_1 , v_2 , v_3), tiene la matriz

$$AMA^{-1} = \begin{pmatrix} -1 & 2 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 1/3 \end{pmatrix} \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix} \begin{pmatrix} 3 & 4 & -6 \\ -1 & -1 & 3 \\ 3 & 3 & -6 \end{pmatrix}$$

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

Ejemplo C. Sea M una matriz cualquiera real, simétrica y de tipo 2×2. Vamos a probar que existe una matriz regular A tal que AMA^{-1} es diagonal.

Sea

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix} .$$

Se tiene

$$f_M(x) = \begin{pmatrix} (x - a) & b \\ b & (x - c) \end{pmatrix} = x^2 - (a + c)x + (ac - b^2) .$$

El discriminante de este trinomio es

$$(a + c)^2 - 4(ac - b^2) = (a - c)^2 + 4b^2 \geqslant 0.$$

Por consiguiente $f_M(x)$ tiene raices reales, y pueden ocurrir los dos casos siguientes:

1. $f_M(x)$ tiene dos raices reales distintas λ_1 y λ_2, de manera que

$$f_M(x) = (x - \lambda_1)(x - \lambda_2).$$

En este caso se tiene evidentemente

$$\phi_M(x) = (x - \lambda_1)(x - \lambda_2) \ ,$$

y existe A ,regular ,tal que

$$AMA^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} .$$

2. $f_M(x)$ tiene la raiz doble λ. Tal cosa ocurre cuando

$$(a - c)^2 + 4b^2 = 0 \ ,$$

es decir cuando $a = c$ y $b = 0$. En tal caso la matriz M ya tiene la forma diagonal.

Ejercicios

13.7.1. Para las siguientes matrices M sobre $\mathbb{C}$, calcular una matriz regular A, si es que existe, tal que AMA^{-1} sea diagonal

(a) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 0 & 0 & -1 \end{pmatrix}$ (e) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 0 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 0 & 0 \end{pmatrix}$

13.7.2. Sean A y B dos endomorfismos de un espacio V de dimensión finita sobre $\mathbb{F}$. Si $AB = BA$ y si los polinomios

minimales ϕ_A y ϕ_B son productos de factores mónicos de primer grado distintos entre sí, probar que existe una base de V respecto de la cual las matrices de A y B son ambas diagonales.

§ 14. Endomorfismos nilpotentes

14.1. Consideraciones preliminares.

Definición 14.1.1. Un endomorfismo T de un espacio vectorial de dimensión n se dice que es *nilpotente de índice k* (*k es el índice de nilpotencia*) si k es un entero positivo tal que se cumplen las relaciones

$$T^k = 0 \quad , \quad T^{k-1} \neq 0 \ .$$

La última relación implica que existe un vector x para el cual es $T^{k-1}x \neq 0$. Si V_1 es un subespacio de V, y la restricción a V_1 de un endomorfismo T de V es nilpotente de índice k', se dice que T' *es nilpotente de índice k' en* V_1. Si además T es nilpotente de índice k entonces se tiene evidentemente que $k' \leqslant k$.

Teorema 14.1.1. El polinomio minimal de un endomorfismo nilpotente T de índice k es $\phi_T(x) = x^k$, y su polinomio característico es $f_T(x) = x^n$. Por consiguiente el índice k de un endomorfismo nilpotente es no mayor que n y su único valor propio es 0.

Demostración. Puesto que $T^k = 0$, el polinomio minimal de T divide a x^k. Será por tanto de la forma $x^{k'}$ donde $k' \leqslant k$. Pero si fuera $k' < k$ se tendría $T^{k'} = 0$, y k no sería el índice de T. Por tanto es $k' = k$ y $\phi_T(x) = x^k$.

Por el teorema 13.6.1 (o por el corolario 2 del teorema 13.3.4, cuando el campo $\mathbb{F}$ es algebraicamente cerrado), el polinomio característico de T debe dividir a x^{nk}. Pero f_T es de grado n; luego debe ser $f_T(x) = x^n$, y puesto que ϕ_T divide a f_T es claro que

$$k \leqslant n. \quad \blacksquare$$

14.2. Forma canónica de Jordan de la matriz de un endomorfismo nilpotente.

Lema 14.2.1. Sea T un endomorfismo nilpotente de índice k de un espacio V de dimensión n, y sea x_o un vector de V para el cual es $T^{k-1} x_o \neq 0$. Entonces los vectores

$$x_o \ , \ Tx_o \ , \ T^2 x_o \ , \ \ldots \ , \ T^{k-1} x_o$$

son linealmente independientes, el subespacio H generado por esos vectores es invariante respecto de T, y existe otro subespacio K, que también es invariante respecto de T, tal que tiene lugar la descomposición de V en la suma directa

$$V = H \oplus K \ .$$

Además, T es nilpotente de índice k en H y de índice $\leqslant k$ en K.

Demostración. Probaremos en primer lugar que los vectores

$$x_o \ , \ Tx_o \ , \ \ldots \ , \ T^{k-1} x_o \quad \text{son } l.i.$$

Sea $\sum_{i=0}^{k-1} \alpha_i T^i x_o = 0$, y j sea el menor índice tal que $\alpha_j \neq 0$ (eventualmente podría ser $j = 0$). Si dividimos por α_j y modificamos la notación de una manera que es evidente, resulta

$$T^j x_o = \sum_{i=j+1}^{k-1} \alpha_i T^i x_o = T^{j+1} \Big(\sum_{i=j+1}^{k-1} \alpha_i T^{i-j-1} x_o \Big) = T^{j+1} y.$$

De la definición del índice se sigue que

$$T^{k-1} x_o = T^{k-1-j} T^j x_o = T^{k-1-j} T^{j+1} y = T^k y = 0 \ ;$$

y puesto que este último resultado contradice a la elección de x_o, podemos concluir que $\alpha_j = 0$ para todo j.

Es evidente que el subespacio H, generado por los vectores

$$x_o, \ Tx_o \ , \ \ldots \ , \ T^{k-1} x_o \ ,$$

es invariante respecto de T. Para obtener K, procederemos por inducción respecto de k. Si $k = 1$ el resultado es trivial, pues entonces es $T^k = 0$, $T^{k-1} = I$, H es el subespacio generado por

cualquier vector $x_o \neq 0$ y K es cualquier complemento en ese subespacio.

Supongamos pues que el teorema se cumple cuando el índice es $k-1$. El subespacio $R = Im\ T$ es evidentemente invariante respecto de T, y la restricción $\overline{T}$ de T a R es nilpotente de índice $k-1$ (recordar que se tiene $\overline{T}^{k-2}(\overline{T}\,x_o) = \overline{T}^{k-1}x_o \neq 0$). Por tanto puede aplicarse a R y a esa restricción la propiedad del teorema. El vector $y_o = Tx_o$ pertenece a R y es $\overline{T}^{k-2}y_o \neq 0$. Luego, por la hipótesis de inducción, los vectores $y_o, \overline{T}y_o, \ldots, \overline{T}^{k-2}y_o$, que son *l.i.*, generan a un subespacio H_o de R, y existe el subespacio K_o de R, invariante respecto de $\overline{T}$, y por tanto de T, tal que

$$R = H_o \oplus K_o \qquad (14.2.1)$$

Sea $K_1 = T^{-1}(K_o)$ el subespacio de los vectores de V que por T se aplica en K_o. Se ve que la imagen de H por T es H_o y la de K_1 es K_o, y podemos comprobar que $V = H + K_1$. En efecto, cualquiera que sea $x \in V$ se tiene que $Tx \in R$; luego en virtud de (14.2.1) es $Tx = y + z$, $y \in H_o$, $z \in K_o$. Puede escribirse

$$y = \sum_{i=1}^{k-1} \alpha_i T^i x_o = T\Big(\sum_{i=0}^{k-2} \alpha_{i+1} T^i x_o\Big) = Ty_1 \quad , \quad y_1 \in H.$$

Se sigue que $Tx = Ty_1 + z$, o bien $T(x - y_1) = z$; de modo que $T(x - y_1) \in K_o$. Esto significa que $x - y_1 \in K_1$, y por consiguiente, x es la suma de $y_1 \in H$ y $x - y_1 \in K_1$.

El hecho demostrado de que $V = H + K_1$ y de que

$$T(H) \oplus T(K_1) = H_o \oplus K_o = R = T(V) \ ,$$

podría inducir a pensar que K_1 es el espacio K que buscamos; pero no es así en general, pues puede ocurrir que sea $H \cap K_1 \neq \{0\}$ (Todo lo que podemos probar es que $H \cap K_1$ pertenece al núcleo de T. En efecto, sea $x \in H \cap K_1$. Será entonces $x \in H$ y $Tx \in K_o$ Por lo primero será $Tx \in H_o$; luego $Tx \in H_o \cap K_o$, o sea $Tx = 0$).

Si bien K_1 no es en general el subespacio de K buscado, vamos

a valernos de él para obtenerlo. La idea es elegir a K entre los complementos de $H \cap K_1$ en K_1. Puede observarse que, cualesquiera que sea el subespacio K, tal que $K_1 = K \oplus (H \cap K_1)$, que se considere, se cumple que $V = H \oplus K$. En efecto, debido a la manera como se supone K introducido, se deberá tener

$$K \subset K_1 \qquad \text{y} \qquad K \cap H \cap K_1 = \{0\} \ ;$$

y por consiguiente $K \cap H = K \cap H \cap K_1 = \{0\}$. Además,

$$H + K \supset H + K + (H \cap K_1) = H + K_1 = V.$$

Luego, puesto que $H + K \subset V$, resultará $H + K = V$; y por consiguiente $V = H \oplus K$, relación que se cumplirá cualquiera que sea el complemento de K de $H \cap K_1$ en K_1 que se considere. Sin embargo, para que K cumpla todas las condiciones del enunciado debe elegirse de modo que sea, además, invariante respecto de T. Con ese fin probaremos que existe el subespacio K_0' de V tal que

$$K_1 = K_0' \oplus K_0 \oplus (H \cap K_1) \ ,$$

y que $K_0' \oplus K_0$ es un complemento de $H \cap K_1$ en K_1, invariante respecto de T, y por tanto puede considerarse como el subespacio K que buscamos. Observemos ante todo que

$$K_0 \cap (H \cap K_1) = \{0\} \ .$$

En efecto, supongamos que sea $x \in K_0 \cap (H \cap K_1)$. Se tendrá entonces que $x \in K_0 \cap H$ y por tanto $x \in \mathcal{R} \cap H$, de donde se sigue que

$$x = \alpha_0 x_0 + \alpha_1 T x_0 + \cdots + \alpha_{k-1} T^{k-1} x_0 = Ty \ ,$$

y $0 = T^k y = \alpha_0 T^{k-1} x_0$; lo que, puesto que $T^{k-1} x_0 \neq 0$, implica que $\alpha_0 = 0$, o sea $x \in H_0$. Se tiene entonces que $x \in H_0 \cap K_0$, y (14.2.1) permite deducir que $x = 0$. (Observemos de paso que hemos probado que $\mathcal{R} \cap H = H_0$).

Los espacios K_0 y $H \cap K_1$, que acabamos de probar que son disjuntos, pertenecen a K_1; el primero porque $TK_0 \subset K_0$, de donde $K_0 \subset T^{-1}(K_0) = K_1$, y el segundo obviamente. Resulta así que

$$K_o \oplus (H \cap K_1)$$

es un subespacio de K_1. Designemos por K_o' a uno de sus complementos en K_1. Se tendrá

$$K_1 = K_o' \oplus K_o \oplus (H \cap K_1) ,$$

y por tanto $K_o' \oplus K_o$ es un complemento de $(H \cap K_1)$ en K_1. La invariancia de K respecto de T resulta porque, dado que $K \subset K_1$, se tiene $TK \subset TK_1 = K_o \subset K$.

Por último, T es nilpotente de índice k, luego lo será, de índice no mayor que k, en los subespacios H y K de V; y puesto que $T^{k-1}x_o \neq 0$, será nilpotente de índice k en H. ■

Teorema 14.2.1. Dado un endomorfismo nilpotente T de índice k del espacio V de dimensión n, existen los enteros positivos

$$p , k_1 , \ldots , k_p \quad \text{tales que}$$

$$k_1 = k \geqslant k_2 \geqslant \cdots \geqslant k_p \quad \text{y} \quad k_1 + k_2 + \cdots + k_p = n,$$

y los vectores $x_1 , \ldots , x_p$ tales que

$$T^{k_1}x_1 = 0 , \ldots , T^{k_p}x_p = 0 ,$$

y que

$$\mathcal{B} = \{x_i , Tx_i , T^2x_i , \ldots , T^{k_i-1}x_i ; \quad 1 \leqslant i \leqslant p\}$$

constituye una base de V.

Se tiene, además, que V se descompone en la forma

$$V = H_1 \oplus \cdots \oplus H_p$$

donde H_i es el subespacio de dimensión k_i, que tiene la base

$$\{x_i , Tx_i , \ldots , T^{k_i-1}x_i\}$$

y es invariante respecto de T; y T es nilpotente de índice k_i en H_i.

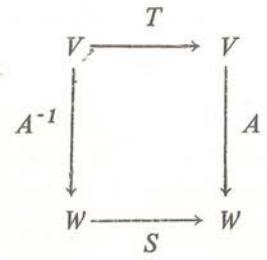
Los enteros $p , k_1 , \ldots , k_p$ forman un *sistema de invariantes de* T en el sentido de que si S es un endomorfismo nilpotente del espa

cio W de dimensión n, entonces una condición necesaria y suficiente para que S y T sean *semejantes*, es decir para que exista el isomorfismo $A: V \to W$ tal que sea $S = ATA^{-1}$, es que S sea nilpotente de índice k y que los enteros

$$p', k_1', \ldots, k_p',$$

asociado a S en virtud de la primera parte del teorema, sean respectivamente iguales a

$$p, k_1, \ldots, k_p .$$



En particular, los invariantes $p, k_1, \ldots, k_p$ son únicos para cada endomorfismo nilpotente.

Demostración. Definamos $k_1 = k$ y escojamos un vector cualquiera $x_1 \in V$ tal que

$$T^{k_1-1} x_1 = 0.$$

En virtud del lema 14.2.1 el subespacio H_1 de dimensión k_1, con base $\{x_1, Tx_1, \ldots, T^{k_1-1} x_1\}$ es invariante respecto de T, y en él T es nilpotente de indice k_1; además, H_1 posee un complemento K_1 en V que también es invariante respecto de T y cuya dimensión es $n - k_1$. En ese complemento, T es nilpotente de índice

$$k_2 \leqslant k_1 .$$

Luego podemos aplicar a K_1 el procedimiento de descomposición en suma directa del lema 14.2.1, a cuyo fín elegimos el vector $x_2 \in K_1$ tal que $T^{k_2-1} x_2 \neq 0$, y obtendremos $K_1 = H_2 \oplus K_2$, donde H_2 es el subespacio de dimensión k_2, con base

$$\{x_2, Tx_2, \ldots, T^{k_2-1} x_2\},$$

invariante respecto de T, y en el cual T es nilpotente de índice k_2; y K_2 es invariante respecto de T, su dimensión es $n - k_1 - k_2$ y en él T es nilpotente de índice $k_3 \leqslant k_2$. Por consiguiente se tiene

$$V = H_1 \oplus H_2 \oplus K_2 \ .$$

Prosiguiendo de la misma manera a la descomposición de K_2, y así sucesivamente hasta agotar el espacio V resulta la descomposición

$$V = H_1 \oplus H_2 \oplus \ldots \oplus H_p$$

a que se refiere el enunciado. La base $\mathcal{B}$ es evidentemente la reunión de las bases que hemos utilizado para definir los subespacios H_i. Con esto queda demostrada toda la primera parte del teorema y vamos a probar la segunda parte.

Sea S un endomorfismo del espacio W de dimensión n, y sea $A: V \to W$ un isomorfismo tal que $S = ATA^{-1}$. S es nilpotente con el mismo índice k que T, pues cualquiera que sea el entero positivo m se tiene

$$S^m = AT^mA^{-1} \quad \text{o bien} \quad T^m = A^{-1}S^mA \ ;$$

luego

$$(T^m x = 0 \ , \ \forall \ x \in V) \Leftrightarrow (S^m x' = 0 \ , \ \forall \ x' \in W) \ ,$$

o sea

$$(T^m = 0) \Leftrightarrow (S^m = 0) \ ;$$

y por tanto T y S son nilpotentes del mismo índice k.

Sean p', k_1', $\ldots$, $k_{p'}'$ números que tienen respecto de S las propiedades que, de acuerdo con el enunciado, tienen $p, k_1, \ldots, k_p$ respecto de T. La descomposición que corresponde a W será

$$W = H_1' \oplus H_2' \oplus \ldots \oplus H_{p'}' \ .$$

Sabemos, en virtud de lo establecido precedentemente, que

$$k_1' = k_1 = k.$$

Sea entonces r el mayor número entero positivo tal que $1 \leqslant r \leqslant p$ y que $k_i = k_i'$ para $i = 1, 2, \ldots, r$. Probaremos ahora que si suponemos que $r < p$ se llega a una contradicción. Sea pues $1 \leqslant r < p$.

Será entonces $r < p'$, pues si fuera $r = p'$ resultaría

$$k_1 + \cdots + k_r = k'_1 + \cdots + k'_{p'} = n = k_1 + \cdots + k_p$$

y por tanto $r = p$ contra lo supuesto. Podemos escribir

$$V = H \oplus K \ , \qquad (14.2.2)$$

donde $H = H_1 \oplus \cdots \oplus H_r$ y $K = H_{r+1} \oplus \cdots \oplus H_p$, y ambos son invariantes respecto de T. Además

$$W = H' \oplus K' \ , \qquad (14.2.3)$$

donde $H' = H'_1 \oplus \cdots \oplus H'_r$ y $K' = H'_{r+1} \oplus \cdots \oplus H'_{p'}$, ambos invariantes respecto de S. T es nilpotente en K de índice k_{r+1} y S es nilpotente en K' de índice k'_{r+1}. De (14.2.2) y (14.2.3) se deduce que

$$V = H + K \quad \text{y} \quad W = H' + K'.$$

Luego

$$T^{k_{r+1}}(V) = T^{k_{r+1}}(H) + T^{k_{r+1}}(K) \qquad (14.2.4)$$

y

$$S^{k_{r+1}}(W) = S^{k_{r+1}}(H') + S^{k_{r+1}}(K') \ . \qquad (14.2.5)$$

Pero $T^{k_{r+1}}(H) \subset H$ y $T^{k_{r+1}}(K) \subset K$; luego, puesto que

$$H \cap K = \{0\} \ ,$$

se tiene

$$T^{k_{r+1}}(H) \cap T^{k_{r+1}}(K) = \{0\} \ ,$$

y por tanto

$$T^{k_{r+1}}(V) = T^{k_{r+1}}(H) \oplus T^{k_{r+1}}(K) \qquad (14.2.6)$$

y semejantemente

$$S^{k_{r+1}}(W) = S^{k_{r+1}}(H') \oplus S^{k_{r+1}}(K').$$

De esta última se deduce, puesto que A es un isomorfismo,

$$A^{-1}S^{k_{r+1}}(W) = A^{-1}S^{k_{r+1}}(H') \oplus A^{-1}S^{k_{r+1}}(K') \ ,$$

o bien

$$T^{k_r+1}A^{-1}(W) = T^{k_r+1}A^{-1}(H') \oplus T^{k_r+1}A^{-1}(K') ,$$

y por consiguiente

$$T^{k_r+1}(V) = T^{k_r+1}(A^{-1}H') \oplus T^{k_r+1}(A^{-1}K') .$$

De esta relación y de (14.2.6) se deduce

$$dim\ T^{k_r+1}(H) + dim\ T^{k_r+1}(K) = dim\ T^{k_r+1}(A^{-1}H') + dim\ T^{k_r+1}(A^{-1}K'). \qquad (14.2.7)$$

Puesto que H tiene la base $\{x_i, Tx_i, \ldots, T^{k_i-1}x_i;\ 1 \leqslant i \leqslant r\}$ y que $T^{k_i}x_i = 0$, $T^{k_r+1}(H)$ tiene la base :

$$\{T^{k_r+1}x_i ,\ T^{k_r+1+1}x_i ;\ \ 1 \leqslant i \leqslant r\} ;$$

y se tiene

$$dim\ T^{k_r+1}(H) = (k_1 - k_{r+1}) + (k_2 - k_{r+1}) + \ldots + (k_r - k_{r+1})$$

H' tiene la base $\{x'_i, Sx'_i, \ldots, S^{k'_i-1}x'_i\ ;\ 1 \leqslant i \leqslant r\}$; luego $A^{-1}H'$ tiene la base

$$\{A^{-1}x'_i ,\ A^{-1}Sx'_i, \ldots, A^{-1}S^{k'_i-1}x'_i ;\ \ 1 \leqslant i \leqslant r\} .$$

Se sigue de aquí que $T^{k_r+1}(A^{-1}H')$ es generado por los vectores

$$T^{k_r+1}A^{-1}x'_i ,\ T^{k_r+1}A^{-1}Sx'_i, \ldots, T^{k_r+1}A^{-1}S^{k'_i-1}x'_i ,\ (1 \leqslant i \leqslant r)$$

o sea, puesto que $S^{k'_i}x'_i = 0$, por los vectores

$$A^{-1}S^{k_r+1}x'_i ,\ A^{-1}S^{k_r+1+1}x'_i, \ldots, A^{-1}S^{k'_i-1}x'_i ,\ (1 \leqslant i \leqslant r),$$

los cuales son $l.i.$ y por tanto constituyen una base de

$$T^{k_r+1}(A^{-1}H') ,$$

de manera que, si se tiene presente que $k'_i = k_i$ para $i = 1, \ldots, r$,

$$\dim T^{k_{r+1}}(A^{-1}H') = (k_1 - k_{r+1}) + \cdots + (k_r - k_{r+1}) = \dim T^{k_{r+1}}(H) .$$

Se sigue entonces de (14.2.7) que

$$\begin{aligned} \dim T^{k_{r+1}}(K) &= \dim T^{k_{r+1}}(A^{-1}K') \\ &= \dim A^{-1}(S^{k_{r+1}}K') \\ &= \dim S^{k_{r+1}}(K') \end{aligned}$$

Ahora bien, T es nilpotente de índice k_{r+1} en K; luego

$$T^{k_{r+1}}(K) = \{0\} \quad ,$$

y por consiguiente es

$$S^{k_{r+1}}(K') = 0 .$$

Pero S es nilpotente de índice k'_{r+1} en K'. Luego debe ser

$$k'_{r+1} \leqslant k_{r+1} \quad .$$

Modificaciones sencillas en la demostración anterior, que consisten esencialmente en el reemplazo de k_{r+1} por k'_{r+1} en las ecuaciones (14.2.4) y (14.2.5), y cuyos detalles dejamos al lector, permiten prober que $k_{r+1} \leqslant k'_{r+1}$ y por tanto que $k_{r+1} = k'_{r+1}$, lo cual contradice la forma como hemos definido a r. Se deduce de aquí que $r = p$, es decir que $k_i = k'_i$ para $i = 1,2,\ldots,p$. Por consiguiente $k'_1 + \cdots + k'_p = k_1 + \cdots + k_p = n$. Puesto que los k'_i son enteros positivos y que, según sabemos, debe ser

$$k'_1 + \cdots + k'_{p'} = n \quad ,$$

resultará que $p = p'$, con lo cual los números p', $k'_1, \ldots, k'_{p'}$, serán respectivamente iguales a los números p, $k_1, \ldots, k_p$.

En particular, si suponemos que $W = V$ y $S = T$, podremos suponer que $A = I_V$ y deduciremos de allí la unicidad de los inva-

riantes $p, k_1, \ldots, k_p$ de cada endomorfismo.

Recíprocamente, supongamos ahora que S es un endomorfismo nilpotente de W, cuyos invariantes son $p, k_1, \ldots, k_p$. Si aplicamos a S la parte ya probada del teorema obtendremos una base de W,

$$\mathcal{B}' = \{x_i', Sx_i', \ldots, S^{k_i-1}x_i' ; \ 1 \leqslant i \leqslant p\}.$$

Podemos definir entonces el isomorfismo $A: V \to W$ tal que

$$A(T^j x_i) = S^j(x_i'). \qquad (0 \leqslant j \leqslant k_i - 1 \ ; \ 1 \leqslant i \leqslant p).$$

Si $u = \sum_{j,i} \alpha_{ji} T^j x_i$ es un elemento cualquiera de V referido a la base $\mathcal{B}$, se tendrá

$$AT(u) = \sum_{j,i} \alpha_{ji} AT^{j+1} x_i = \sum_{j,i} \alpha_{ji} S^{j+1} A x_i = \sum_{j,i} \alpha_{ji} S^{j+1} x_i'$$

$$= S \sum_{j,i} \alpha_{ji} S^j x_i' = S\Big(\sum_{j,i} \alpha_{ji} AT^j x_i\Big) = SA(u)$$

Por consiguiente es $AT = SA$, o sea $S = ATA^{-1}$. ∎

Observación 1. Nótese que dos endomorfismos T_1 y T_2 de V, nilpotentes que tienen los mismos invariantes, no son idénticos sino semejantes, es decir que existe un automorfismo A de V tal que

$$T_2 = AT_1A^{-1}$$

Definición 14.2.1. El conjunto de los números enteros positivos $k_1, \ldots, k_p$ tales que $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_p$ y que $k_1 + k_2 + \cdots + k_p = n$, que en virtud del teorema anterior corresponde a cada endomorfismo nilpotente de índice k_1 de un espacio V de dimensión n, se llama *sistema inavariante de T.*

El lector puede comprobar fácilmente que, en virtud del teorema anterior, los invariantes $k_1, \ldots, k_p$ se caracterizan por ser los números enteros positivos tales que $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_p$, y que V pue-

de expresarse como suma directa de p subespacios invariantes respecto de T

$$V = H_1 \oplus H_2 \oplus \ldots \oplus H_p \quad ,$$

de modo que $dim\ H_i = k_i$, para $i = 1, \ldots, p$, y que T es nilpotente de índice k_i en H_i.

Teorema 14.2.2. Sea T un endomorfismo nilpotente de un espacio V de dimensión n y sea $\{k_1, \ldots, k_p\}$ el sistema invariante de T. Existe entonces una base de V respecto de la cual T se representa por una matriz de bloques de la forma

$$J = \begin{pmatrix} J_1 & 0 & \ldots & 0 \\ 0 & J_2 & \ldots & 0 \\ \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & \ldots & J_p \end{pmatrix} \qquad (14.2.8)$$

donde cada bloque J_i es de orden k_i y tiene la forma

$$J_i = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix} \qquad (14.2.9)$$

en que los términos son nulos, salvo los que están situados inmediatamente encima de la diagonal principal, los cuales son todos iguales a *1*. La matriz (14.2.8) se llama *forma canónica de Jordan de la matriz del endomorfismo nilpotente T,* es única para cada endomorfismo nilpotente y es la misma para dos endomorfismos nilpotentes semejantes.

Demostración. Basta considerar la base $\mathcal{B}$ cuya existencia fué establecida en el teorema 14.2.1, y calcular la matriz de T respecto de esa base, pero suponiendo que sus vectores están ordenados en

la forma siguiente:

$$\{T^{k_1-1}x_1, T^{k_1-2}x_1, \ldots, Tx_1, x_1; T^{k_2-1}x_2, \ldots, Tx_2, x_2; \ldots; T^{k_p-1}x_p, T^{k_p-2}x_p, \ldots, Tx_p, x_p\} \quad (14.2.10).$$

Se obtiene así una matriz de la forma (14.2.8)

La forma canónica de Jordan de un endomorfismo nilpotente T es totalmente determinada por los invariantes de T. Puesto que estos son únicos, también será única dicha forma canónica.

Para dos endomorfismos semejantes, la igualdad de sus conjuntos de invariantes, establecida en el teorema 14.2.1, implica la igualdad de sus formas canónicas. ■

Observación 1. Conviene observar que hasta ahora no hemos hecho uso de la suposición de que el campo $\mathbb{F}$ es algebraicamente cerrado, suposición que emplearemos en el parágrafo siguiente para asegurar la existencia de valores propios de un endomorfismo.

Ejercicios

14.2.1. Clasificar los endomorfismos nilpotentes de $\mathbb{R}^3$ de conformidad con los casos que pueden presentarse según sean los valores de los sistemas de invariantes. (Se deberán tomar en cuenta en especial los teoremas 14.1.1 y 14.2.1 y la precedente observación 4).

14.2.2. Tratar la misma cuestión anterior para los endomorfismos de $\mathbb{R}^5$ de índice 3.

14.2.3. Id. para todos los endomorfismos de $\mathbb{R}^6$.

14.2.4. Demostrar que un endomorfismo de un espacio vectorial sobre C que no tiene otro valor propio que 0 es nilpotente.

§ 15. Representación de un endomorfismo mediante una matriz de Jordan. Reducción de una matriz a la forma canónica de Jordan.

15.1. Matrices de Jordan. Introduciremos primero algunas notaciones. Cualesquiera que sean el escalar $\lambda \in \mathbb{F}$ y el numero positivo k, denotaremos por $J(\lambda\ ;\ k)$ a la matriz $k \times k$ sobre $\mathbb{F}$

$$J(\lambda\ ;\ k) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

Dados los enteros positivos $k_1 \geqslant k_2 \geqslant \dots \geqslant k_p$, denotaremos por $J(\lambda; k_1, \dots, k_p)$ a la matriz diagonal de bloques, de orden

$$\sum_{i=1}^{p} k_i$$

$$J(\lambda ; k_1, \dots, k_p) = \begin{pmatrix} J(\lambda;k_1) & 0 & 0 \\ 0 & J(\lambda\ ;\ k_2) & 0 \\ & & \ddots \\ 0 & 0 & J(\lambda\ ;\ k_p) \end{pmatrix},$$

en que todos los bloques no diagonales son nulos.

Definición 15.1.1. Una matriz J de orden n sobre $\mathbb{F}$ se llama *matriz de Jordan* si es una matriz diagonal de bloques de la forma

$$J = \begin{matrix} J(\lambda_1 ; k_1^{(1)}, \ldots, k_{p_1}^{(1)}) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_\mu ; k_1^{(\mu)}, \ldots, k_{p_\mu}^{(\mu)}) \end{matrix} \qquad (15.1.1)$$

en que todos los bloques no diagonales son nulos, los λ_i $(i = 1,\ldots,\mu)$ son escalares de $\mathbb{F}$, distintos entre sí, y para cada $i = 1, \ldots, \mu$ es

$$k_1^{(i)} \geqslant k_2^{(i)} \geqslant \ldots \geqslant k_{p_i}^{(i)} .$$

Teorema 15.1.1. Si $k_1 \geqslant \ldots \geqslant k_p$ son enteros positivos, y J es una matriz de la forma

$$J = J(0;k_1, \ldots, k_p) = \begin{matrix} J(0 ; k_1) & & 0 \\ & \ddots & \\ 0 & & J(0 ; k_p) \end{matrix} \qquad (15.1.2)$$

entonces se tiene que

$$J^{k_1} = 0 \quad \text{y} \quad J^{k_1 - 1} \neq 0.$$

Demostración. Designemos por T al endomorfismo de un espacio V de dimensión $k_1 + \ldots + k_p$ cuya matriz relativa a una base cualquiera que representamos por

$$\{e_{11}, \ldots, e_{1k_1} ; e_{21}, \ldots, e_{2k_2} ; e_{p1}, \ldots, e_{pk_p}\} \qquad (15.1.3)$$

es la matriz J. Se tiene entonces para $j = 1, \ldots, p$

$$Te_{j1} = 0 , \; Te_{j2} = e_{j1} , \ldots , \; Te_{jl} = e_{j(l-1)} , \ldots , \; Te_{jk_j} = e_{j(k_j - 1)} . \qquad (15.1.4)$$

Se deduce de aquí que $T^l e_{jl} = Te_{jl} = 0$ y $T^{l-1} e_{jl} = e_{j1}$. Luego

$$T^{k_1} e_{jl} = 0 \quad (j = 1, \ldots, p\,;\, l = 1, \ldots, k_j) \text{ y } T^{k_1 - 1} e_{1k_1} = e_{11} \neq 0.$$

Luego es $T^{k_1} = 0$ y $T^{k_1 - 1} \neq 0$, y por tanto $J^{k_1} = 0$ y $J^{k_1 - 1} \neq 0$. ∎

Observación. De las relaciones (15.1.4) resulta que si el endomorfismo T se representa mediante una matriz de Jordan J de la forma (15.1.2), respecto de la base (15.1.3), entonces, si escribimos

$$e_{1k_1} = x_1 \,,\; e_{2k_2} = x_2 \,,\; \ldots,\; e_{pk_p} = x_p \,,$$

dicha base puede expresarse en la forma

$$\{T^{k_1 - 1} x_1 \,,\, T^{k_1 - 2} x_1 \,,\, \ldots \,,\, Tx_1 \,,\, x_1 \,;\, T^{k_2 - 1} x_2 \,,\, T^{k_2 - 2} x_2 \,,\, \ldots \,,\, Tx_2 \,,\, x_2 \,;$$
$$\ldots \,;\, T^{k_p - 1} x_p \,,\, T^{k_p - 2} x_p \,,\, \ldots \,,\, Tx_p \,,\, x_p\} \,,$$

y el endomorfismo T es nilpotente de índice k_1, con los invariantes $k_1, \ldots, k_p$.

15.2. Proposiciones preliminares. La demostración del teorema de la sección siguiente, que asegura la existencia de una base respecto de la cual un endomorfismo cualquiera de un espacio *n-dimensional* sobre un campo $\mathbb{F}$ algebraicamente cerrado se representa por una matriz de Jordan que es esencialmente única, requiere establecer las siguientes proposiciones previas.

Teorema 15.2.1. Sea T un endomorfismo del espacio V de dimensión n sobre $\mathbb{F}$, y sea

$$V = V_1 + V_2$$

donde V_1 y V_2 son subespacios invariantes respecto de T. Sean T_1 y T_2 las restricciones de T a V_1 y V_2 respectivamente. Si el polinomio minimal de T_1 es $p_1(x)$ y el de T_2 es $p_2(x)$, entonces el polinomio minimal de T es el mínimo común multiplo de $p_1(x)$ y $p_2(x)$.

Demostración. Si $p(x)$ designa al polinomio minimal de T, entonces

es $p(T) = 0$. Ahora bien, según el lema 13.3.1, la restricción de $p(T)$ a V_1 es $p(T_1)$; por tanto, es $p(T_1) = 0$. Semejantemente puede observarse que $p(T_2) = 0$. Puesto que T_1 y T_2 anulan a $p(x)$, se tendrá que $p_1 \mid p$ y $p_2 \mid p$. Por consiguiente el *M.C.M.*, $q(x)$, de $p_1(x)$ y $p_2(x)$ divide a $p(x)$, o sea

$$q \mid p \,. \qquad (15.2.1)$$

Por otra parte, puesto que $p_1 \mid q$ y $p_2 \mid q$ y que

$$p_1(T_1) = 0 \quad \text{y} \quad p_2(T_2) = 0, \text{ será } \quad q(T_1) = 0 \quad \text{y} \quad q(T_2) = 0.$$

Ahora bien ,por hipótesis, dado $v \in V$ existen $v_1 \in V_1$ y $v_2 \in V_2$ tales que $v = v_1 + v_2$. Luego

$$q(T)v = q(T)v_1 + q(T)v_2 = q(T_1)v_1 + q(T_2)v = 0 \,,$$

o sea $q(T) = 0$, y por tanto

$$p \mid q \,. \qquad (15.2.2)$$

De (15.2.1) y (15.2.2) se sigue, por cuanto p y q son mónicos,que

$$p = q \,. \quad \blacksquare$$

Corolario. Sea T un endomorfismo de V , y sea

$$V = V_1 + V_2 + \cdots + V_k$$

donde $V_1, \ldots, V_k$ son subespacios de V invariantes respecto de T. Sea T_i, $i = 1,2, \ldots, k$, la restricción de T a V_i . Si el polinomio minimal de T_i es $p_i(x)$, entonces el polinomio minimal de T es el *M.C.M.* de los polinomios $p_1(x), \ldots, p_k(x)$.

Demostración. Inducción respecto de k. Si escribimos

$$V_1 + V_2 + \cdots + V_{k-1} = U,$$

entonces

$$V = U + V_k \,.$$

U es evidentemente invariante respecto de T, y por el teorema anterior el polinomio minimal de la restricción T_U de T a U es el *M.C.M.* de los polinomios $p_1, \ldots, p_{k-1}$. El polinomio minimal de T será el *M.C.M.* de los polinomios minimales de T_U y T_k, es decir el *M.C.M.* de $p_1, \ldots, p_k$. ■

15.3. Representación de un endomorfismo de un espacio vectorial sobre un campo algebraicamente cerrado, mediante una matriz de Jordan.

Lema 15.3.1. Cada endomorfismo T de un espacio V de dimensión finita es suma directa (definición 13.3.3) de una transformación nilpotente y de una transformación invertible. Es decir que V es suma directa de dos subespacios N y R, invariantes respecto de T, tales que las restricciones de T a N y R son respectivamente nilpotente e invertible. La referida descomposición es única: si q es el menor entero positivo tal que

$$Ker\ T^q = Ker\ T^{q+1} ,$$

entonces se tiene

$$N = Ker\ T^q \quad \text{y} \quad R = Im\ T^q ,$$

y T es nilpotente de índice q en N.

Demostración. Designemos por $N_k = Ker\ T^k$ al núcleo de la k-ésima potencia de T. Se tiene evidentemente

$$N_1 \subset N_2 \subset N_3 \subset \ldots$$

Afirmamos primero que si $N_k = N_{k+1}$ entonces es $N_k = N_{k+j}$ para todo entero positivo j. En efecto, si $T^{k+j}x = 0$, entonces es $T^{k+1}T^{j-1}x = 0$, de donde se sigue, puesto que $N_k = N_{k+1}$, que $T^k T^{j-1}x = 0$, y por tanto $T^{k+j-1}x = 0$; o sea $N_{k+j} \subset N_{k+j-1}$, y por consiguiente $N_{k+j} = N_{k+j-1}$. La inducción respecto de j prueba nuestra aserción.

Puesto que V es de dimensión finita, las dimensiones de los sub espacios N_k constituyen una sucesión no decreciente que alcanzará un valor máximo, y por tanto existirá un entero positivo mínimo q para el cual es $N_{q-1} \neq N_q = N_{q+1} = N_{q+2} = \cdots$ Para todo $x \in N_q$ es $T^q x = 0$ y existe por lo menos un $y \in N_q$ tal que

$$T^{q-1} y \neq 0.$$

Luego q es el índice de nilpotencia de T en N_q .

Es fácil observar que tanto $N_q = Ker\ T^q$ como $R_q = Im\ T^q$ son invariantes respecto de T. Probaremos que

$$V = N_q \oplus R_q$$

y que T es invariante sobre R_q. Puesto que ya hemos probado que T es nilpotente sobre N_q, quedará de esa manera demostrado el lema .

Si $x \in N_q \cap R_q$ entonces es $T^q x = 0$ y $x = T^q y$ para algún $y \in V$. Se sigue de aquí de $T^{2q} y = 0$, y por tanto $y \in N_{2q}$; y puesto que $N_q = N_{2q}$, será $y \in N_q$, o sea $x = T^q y = 0$. De esta manera hemos probado que $N_q \cap R_q = \{0\}$. Además, sabemos que

$$dim\ N_q + dim\ R_q = dim\ Ker\ T^q + dim\ Im\ T^q = n \ ,$$

luego

$$dim\ (N_q \oplus R_q) = dim\ V \ ,$$

y por consiguiente

$$N_q \oplus R_q = V \ .$$

Probaremos por último que T es inversible en R_q. Para ello bastará que probemos que el núcleo de la restricción de T a R_q es $\{0\}$. Sea pues $Tx = 0$ para $x \in R_q$, es decir para $x = T^q y$. Se tendrá por tanto $T^{q+1} y = 0$ o sea $y \in N_{q+1}$; y puesto que $N_q = N_{q+1}$

será $y \in N^q$. Por tanto $x = T^q y = 0$, es decir que el núcleo de la restricción de T a R_q se reduce a $\{0\}$.

Vamos a probar que la descomposición de T en una parte nilpotente y otra invertible, en la forma ya demostrada, es única. Supongamos que $V = N' \oplus R'$ es una descomposición de V tal que T es nilpotente sobre N' e invertible sobre R'. Puesto que

$$N' \subset Ker\, T^k = N_k \quad \text{para algún } k,$$

se sigue que $N' \subset N_q$ (porque o bien es $k < q$, y entonces es $N_k \subset N_q$; o bien es $k \geqslant q$ y entonces es $N_k = N_q$); y puesto que $R' \subset Im\, T^k$ para todo k (porque T es inversible sobre R'), se sigue que $R' \subset Im\, T^q = R_q$; y de los dos hechos mencionados se deduce que $N' = N_q$ y $R' = R_q$. Se tiene efectivamente:

$$V = N_q \oplus R_q = N' \oplus R' \;,\; N' \subset N_q \;,\; R' \subset R_q \;;$$

y por tanto

$$dim\, V = dim\, N_q + dim\, R_q = dim\, N' + dim\, R' \;,$$

$$dim\, N' \leqslant dim\, N_q \;,\; dim\, R' \leqslant dim\, R_q \;.$$

Luego

$$N' = N_q \quad \text{y} \quad R' = R_q \;. \;\blacksquare$$

Teorema 15.3.1. Si T es un endomorfismo de un espacio V de dimensión finita sobre un campo algebraicamente cerrado y si λ_1, $\lambda_2, \ldots, \lambda_\mu$ son los valores propios de T, distintos entre sí, con multiplicidades $n_1, n_2, \ldots, n_\mu$ respectivamente, y por tanto es

$$f_T(x) = (x - \lambda_1)^{n_1} (x - \lambda_2)^{n_2} \cdots (x - \lambda_\mu)^{n_\mu} \;,$$

entonces V es la suma directa de μ subespacios $N_1, \ldots, N_\mu$, de dimensiones $n_1, \ldots, n_\mu$, respectivamente; tales que N_j es invariante

respecto de T y que $T - \lambda_j (= T - \lambda_j I)$ es nilpotente sobre N_j.

Demostración. Tomemos un número $j \in \{1, \ldots, \mu\}$ cualquiera, y consideremos el endomorfismo $T - \lambda_j$. A $T - \lambda_j$ podemos aplicarle el lema anterior para obtener subespacios N_j y R_j tales que

$$V = N_j \oplus R_j$$

y que $T - \lambda_j$ es nilpotente sobre N_j e invertible sobre R_j. Puesto que N_j y R_j son invariantes respecto de $T - \lambda_j$, también son invariantes respecto de $T = (T - \lambda_j) + \lambda_j$. Por tanto puede elegirse la base de V de modo que, si designamos por T' y T'' a las restricciones de T a N_j y R_j, y con las mismas letras entre corchetes designamos a las matrices de esas restricciones respecto de esa base, la matriz de T sea de la forma

$$[T] = \begin{matrix} [T'] & 0 \\ 0 & [T''] \end{matrix}$$

Luego el polinomio característico de T será

$$det\ (T - x) = det\ (T' - x)\, det\ (T'' - x) \qquad (15.3.1)$$

es decir el producto del polinomio característico de T' por el polinomio característico de T''.

Supongamos que λ sea un valor propio de T'; se tendrá

$$T'x = \lambda x$$

para un vector x no nulo de N_j. Para ese valor se tendrá

$$det\ (T' - \lambda) = 0, \text{ y por tanto } det\ (T - \lambda) = 0\ ;$$

luego λ será valor propio de T. Para determinarlo observaremos que

$$(T' - \lambda_j)x = (\lambda - \lambda_j)x\ ,$$

es decir que $\lambda - \lambda_j$ es un valor propio de $T' - \lambda_j$. Ahora bien

$T' - \lambda_j$; es nilpotente,

por tanto su único valor propio es cero. Luego es $\lambda = \lambda_j$. Es decir que λ_j es el único valor propio de la restricción T' de T a N_j. Por otra parte, puesto que $T - \lambda_j$ es invertible en R_j se tiene

$$det\ (T'' - \lambda) \neq 0 \quad ,$$

y por tanto T no tiene el valor propio λ_j sobre R_j. De la relación (15.3.1) se sigue que la multiplicidad del valor propio λ_j de T es la suma de las multiplicidades del mismo para T' y para T''. Pero para T'' esa multiplicidad es cero, luego la multiplicidad de λ_j para T es la misma que para T', la cual es igual a la dimensión del subespacio N_j. Por consiguiente, $dim\ N_j = n_j$.

Si fuera $N_1 \cap N_2 \neq 0$, la restricción de T a esa intersección tendría un valor propio que lo sería para las restricciones de T a N_1 y a N_2. Pero los únicos valores propios de estas últimas son λ_1 y λ_2 respectivamente; luego se tendría que $\lambda_1 = \lambda_2$ contra lo supuesto en el enunciado. Por consiguiente debe ser $N_1 \cap N_2 = \{0\}$. Consideremos ahora el subespacio $N_1 \oplus N_2$ de V. La restricción de T a él tiene los valores propios λ_1 y λ_2 únicamente, de conformidad con el teorema 13.4.1, y el único valor propio de T en N_3 es λ_3. Por un razonamiento semejante al precedente resulta que

$$(N_1 \oplus N_2) \cap N_3 = \{0\} \quad ,$$

y puede construirse el subespacio $N_1 \oplus N_2 \oplus N_3$ de V, en el cual los valores propios de T son λ_1 , λ_2 y λ_3. Prosiguiendo de esta manera se construye el subespacio $N_1 \oplus N_2 \oplus \cdots \oplus N_\mu$ de V, cuya dimensión es $n_1 + n_2 + \cdots + n_\mu$. Pero puesto que $\lambda_1, \ldots, \lambda_\mu$ son todos los valores propios de T y $n_1, n_2, \ldots, n_\mu$ son sus respectivas multiplicidades , se tiene que

$$n_1 + n_2 + \cdots + n_\mu = dim\ V.$$

Por consiguiente debe ser

$$N_1 \oplus N_2 \oplus \cdots \oplus N_\mu = V \ . \quad \blacksquare$$

Teorema 15.3.2. Si T es un endomorfismo de un espacio V de di-

mensión n sobre un campo algebraicamente cerrado, la matriz de T respecto de una base apropiada es una matriz de Jordan (definición 15.1.1), en la cual, por consiguiente, cada término no situado en la diagonal principal o inmediatamente encima de ella es igual a cero. En la diagonal principal figuran los valores propios de T y cada uno de ellos un número de veces consecutivas igual a su multiplicidad n_i. Encima de cada valor propio particular aparecen únicamente unos y ceros de la manera siguiente: sucesiones de unos seguidos de un cero, y los números de elementos de esas sucesiones decrecen de arriba hacia abajo. La matriz de Jordan

$$M(T) = \begin{pmatrix} J(\lambda_1 ; k_1^{(1)}, \ldots, k_{p_1}^{(1)}) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_\mu ; k_1^{(\mu)}, \ldots, k_{p_\mu}^{(\mu)}) \end{pmatrix},$$

así obtenida es la *forma canónica de Jordan del endomorfismo T,* el polinomio minimal de la cual es

$$\phi_T(x) = (x - \lambda_1)^{k_1^{(1)}} (x - \lambda_2)^{k_1^{(2)}} \ldots (x - \lambda_\mu)^{k_1^{(\mu)}} .$$

Además, los endomorfismos T de V y S de W son semejantes, es decir que existe un isomorfismo $A: V \to W$ tal que $S = ATA^{-1}$, si y solo si sus formas canónicas de Jordan son idénticas salvo el orden en que aparecen los bloques diagonales.

Demostración. Con referencia al teorema 15.3.1 y a las notaciones allí empleadas, podemos elegir en cada N_i una base respecto de la cual, como lo establece el teorema 14.2.2, la matriz de la restricción $T_i - \lambda_i$ de $T - \lambda_i$ a N_i, que es un endomorfismo nilpotente cuyo índice designaremos por $k_1^{(i)}$, tiene la forma

$$J(0 ; k_1^{(i)}, \ldots, k_{p_i}^{(i)}) \quad ;$$

y la restricción $T_i = (T_i - \lambda_i) + \lambda_i$ de T a N_i tiene la matriz

$$J(\lambda_i ; k_1^{(i)}, \ldots, k_{p_i}^{(i)}) .$$

Con respecto a la reunión de las bases elegidas de esa manera en los N_i, la matriz de T tiene evidentemente la forma descrita en el enunciado.

Por ser $T_i - \lambda_i$ nilpotente de índice $k_1^{(i)}$ se cumplen las relaciones

$$(T_i - \lambda_i)^{k_1^{(i)}} = 0 \qquad \text{y} \qquad (T_i - \lambda_i)^{k_1^{(i)}-1} \neq 0.$$

La primera de estas ecuaciones prueba que $(x - \lambda_i)^{k_1^{(i)}}$ es un múltiplo del polinomio minimal de T_i. La segunda ecuación demuestra que ese polinomio minimal es justamente

$$(x - \lambda_i)^{k_1^{(i)}} .$$

Puesto que los λ_i son diferentes entre sí, del corolario del teorema 15.2.1 se sigue que el polinomio minimal del endomorfismo T es

$$(x - \lambda_1)^{k_1^{(1)}} (x - \lambda_2)^{k_1^{(2)}} \cdots (x - \lambda_\mu)^{k_1^{(\mu)}} .$$

Vamos a probar ahora la última parte del teorema. Supongamos que exista el isomorfismo $A: V \to W$ tal que $S = ATA^{-1}$. Entre las matrices que representan a S, T y A respecto de bases cualesquiera dadas en V y W se cumple la relación

$$M(S) = M(A)M(T)M(A)^{-1} \quad \text{(ver la sección 6.8).}$$

Luego las matrices de T y S son semejantes y sus polinomios característicos, que son los de T y S, son idénticos (teorema 13.2.3) Resulta de aquí que ambos endomorfismos tienen los mismos valores propios con las mismas multiplicidades.

Supongamos que en V y W son elegidas bases respecto de las cuales las matrices de T y S tienen la forma canónica de Jordan. En cada una de ellas figurarán en la diagonal principal los mismos

valores propios $\lambda_1, \ldots, \lambda_\mu$, el mismo número de veces cada uno de ellos. Siempre será posible evidentemente llevar a cabo un cambio de base en W de manera que el orden en que aparecen los valores propios en la diagonal principal de la matriz de S sea el mismo en que aparecen en la matriz de T, de manera que si esta última tiene la forma dada en el enunciado, la matriz de S sea

$$M(S) = \begin{pmatrix} J(\lambda_1 ; l_1^{(1)}, \ldots, l_{q_1}^{(1)}) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_\mu ; l_1^{(\mu)}, \ldots, l_{q_\mu}^{(\mu)}) \end{pmatrix}$$

y se cumplen la relaciones

$$k_1^{(i)} + k_2^{(i)} + \ldots + k_{p_i}^{(i)} = l_1^{(i)} + l_2^{(i)} + \ldots + l_{q_i}^{(i)} = n_i \quad ,$$

$$n_1 + n_2 + \ldots + n_\mu = n \quad .$$

Según hemos señalado, el subespacio de V generado por los primeros n vectores de base se designa por N_1, el formado por los siguientes n_2 vectores de base se designa por N_2, y así sucesivamente, de manera que $V = N_1 \oplus \ldots \oplus N_\mu$. En forma enteramente se definen los subespacios $N'_1, \ldots, N'_\mu$ de W de modo que

$$W = N'_1 \oplus \ldots \oplus N'_\mu \quad .$$

Vamos a demostrar que

$$N'_i = A(N_i) \quad , \quad i = 1, \ldots, \mu \qquad (15.3.2)$$

Con ese objeto consideremos al subespacio N_i de V, y sea R_i la suma directa de los restantes subespacios N_j $(j \neq i)$. Se tendrá por tanto

$$V = N_i \oplus R_i \quad .$$

Dada la forma de la matriz $M(T)$ puede observarse que N_i y R_i

son invariantes respecto de T. La matriz de la restricción de $T - \lambda_i$ a N_i es $J(0\,; k_1^{(i)}, \ldots, k_{p_i}^{(i)})$, como hemos observado anteriormente; luego, por el teorema 15.1.1, dicha restricción es nilpotente de índice $k_1^{(i)} = k^{(i)}$. La matriz de la restricción de $T - \lambda_i$ a R_i es una matriz de Jordan en que los elementos de la diagonal principal son de la forma $\lambda_j - \lambda_i$, donde $j \neq i$, y por tanto son todos diferentes de cero; luego el determinante de esa matriz es también diferente de cero, y la restricción de $T - \lambda_i$ a R_i es un endomorfismo invertible. Podemos aplicar entonces el lema 15.3.1 al endomorfismo $T - \lambda_i$ de V, que es la suma directa de sus restricciones a N_i y a R_i, nilpotente la primera de índice $k^{(i)}$ e invertible la segunda. Según dicho lema se tiene

$$N_i = Ker\ (T - \lambda_i)^{k^{(i)}}$$

y $k^{(i)}$ es, además, el menor entero positivo tal que

$$Ker\ (T - \lambda_i)^{k^{(i)}} = Ker\ (T - \lambda_i)^{k^{(i)}+1} . \qquad (15.3.3)$$

En forma enteramente análoga se establecerá que

$$N_i' = Ker\ (S - \lambda_i)^{l^{(i)}}$$

donde $l^{(i)} = l_1^{(i)}$ es el índice de nilpotencia de la restricción de

$$S - \lambda_i \quad \text{a} \quad N_i' \ ,$$

y también es el menor entero positivo tal que

$$Ker\ (S - \lambda_i)^{l^{(i)}} = Ker\,(S - \lambda_i)^{l^{(i)}+1} \qquad (15.3.4)$$

Ahora bien, cualquiera que sean el vector $x \in V$ y el entero positivo r, se tiene

$$(S - \lambda_i)^r (Ax) = (ATA^{-1} - \lambda_i)^r (Ax) = A(T - \lambda_i)^r (x) \ ,$$

y por tanto tiene lugar la relación

$$Ker\ (S - \lambda_i)^r = A[Ker(T - \lambda_i)^r] \quad . \qquad (15.3.5)$$

Puesto que $A: V - W$ es un isomorfismo, se deduce de (15.3.3) que $k^{(i)}$ es el menor entero positivo para el cual es

$$A[Ker\ (T - \lambda_i)^{k^{(i)}}] = A[Ker\ (T - \lambda_i)^{k^{(i)}+1}] \quad ,$$

o sea, teniendo en cuenta (15.3.5) ,

$$Ker\ (S - \lambda_i)^{k^{(i)}} = Ker\ (S - \lambda_i)^{k^{(i)}+1} \quad .$$

Se deduce entonces de la relación (15.3.4) que $k^{(i)} = l^{(i)}$, y por consiguiente

$$N_i' = Ker\ (S - \lambda_i)^{l^{(i)}} = Ker\ (S - \lambda_i)^{k^{(i)}} = A[Ker\ (T - \lambda_i)^{k^{(i)}}] = A(N_i),$$

que es la relación (15.3.2). De aquí y de la relación

$$S - \lambda_i = A(T - \lambda_i)A^{-1} \quad ,$$

que tiene lugar en virtud de la hipótesis, se deduce que

$$S_i - \lambda_i = A_i(T_i - \lambda_i)A_i^{-1} \quad ,$$

donde S_i es la restricción de S a N_i' y A_i y T_i son, respectivamente, las restricciones de A y T a N_i; de manera que $S_i - \lambda_i$ y $T_i - \lambda_i$; son las restricciones de $S - \lambda_i$ y $T - \lambda_i$ a N_i' y N_i respectivamente, y $A_i: N_i \to N_i'$ es un isomorfismo. Puesto que $T_i - \lambda_i$ y $S_i - \lambda_i$; son nilpotentes de índice $k^{(i)}$, según ya hemos probado anteriormente, podemos aplicarles la segunda parte del teorema 14.2.2, de la cual resulta que sus matrices canónicas, que son

$$J(0\ ;\ k_1^{(i)}, \ldots, k_{p_i}^{(i)}) \quad \text{y} \quad J(0\ ;\ l_1^{(i)}, \ldots, l_{q_i}^{(i)}) \quad ,$$

son idénticas; de donde resulta inmediatamente la identidad de las matrices $M(T)$ y $M(S)$.

Por último, supongamos recíprocamente que las matrices canón

cas de T y S son iguales, salvo eventualmente el orden en que aparecen los bloques diagonales. Se observará fácilmente que un cambio de la ordenación de los vectores de la base de W dará lugar a una matriz canónica de S idéntica a la de T. Se deduce entonces del teorema 6.12.1 que T y S son endomorfismos semejantes. ■

15.4 Forma canónica de Jordan de una matriz.

Definición 15.4.1. Decimos que una matriz M de orden n puede ser reducida a la *forma canónica de Jordan* J, si J es una matriz de Jordan semejante a M; es decir si existe una matriz regular R de orden n tal que

$$J = RMR^{-1}$$

es una matriz de Jordan.

Las proposiciones precedentemente demostradas para los endomorfismos tienen sus análogas para las matrices, por cuanto cada matriz de orden n sobre $\mathbb{F}$ puede considerarse como la matriz de un endomorfismo de $\mathbb{F}^n$ referido a la base canónica. Nos limitaremos aquí a enunciar explícitamente una de ellas; pero, como veremos en uno de los ejemplos siguientes, puede ser útil algunas veces emplear alguna otra.

Teorema 15.4.1. Si M es una matriz $n \times n$ sobre un campo algebraicamente cerrado $\mathbb{F}$, entonces M puede ser reducida a la forma canónica de Jordan, la cual es única, salvo la ordenación de los bloques diagonales.

Además si M y M' son matrices de orden n sobre el campo $\mathbb{F}$ algebraicamente cerrado, entonces sus formas canónicas de Jordan son iguales, salvo eventualmente el orden en que aparecen los bloques diagonales, si y solo M y M' son semejantes.

Observación. En las proposiciones de la presente sección, así como en las de la sección anterior puede dejar de suponerse que el campo $\mathbb{F}$ es algebraicamente cerrado, *siempre que se admita en cambio que todas las raíces características de las matrices consideradas* (ver la definición 13.2.3), *pertenecen a* $\mathbb{F}$.

Ejemplo A. Reducir a la forma canónica de Jordan a la matriz sobre $\mathbb{R}$:

$$M = \begin{pmatrix} 3 & -4 & 0 \\ 3 & -2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \qquad (15.4.1)$$

El polinomio característico de esta matriz es

$$f_M(x) = \begin{vmatrix} x-3 & 4 & 0 \\ -3 & x+2 & -1 \\ 0 & 0 & x-2 \end{vmatrix} = (x-2)(x^2 - x + 6) .$$

La ecuación $f_M(x) = 0$ no tiene todas sus raices en $\mathbb{R}$ y por tanto M no puede ser reducido a la forma canónica. Si consideramos, en cambio, que M es una matriz sobre $\mathbb{C}$, $f_M(x)$ puede descomponerse en la forma

$$f_M(x) = (x-2)(x - \tfrac{1}{2}(1 + i\sqrt{23}))(x - \tfrac{1}{2}(1 - i\sqrt{23})) .$$

el polinomio minimal $\phi_M(x)$ es idéntico a $f_M(x)$ y la forma canónica de la matriz será

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2}(1 + i\sqrt{23}) & 0 \\ 0 & 0 & \frac{1}{2}(1 - i\sqrt{23}) \end{pmatrix}$$

Ejemplo B. Estudiar la reductibilidad a la forma canónica de Jordan de las matrices sobre $\mathbb{R}$, de la forma

$$M = \begin{pmatrix} 3 & 0 & 0 \\ a & 3 & 0 \\ b & c & -2 \end{pmatrix}$$

El polinomio característico de esta matriz es

$$f_M(x) = (x - 3)^2(x + 2) ,$$

cuyas raices $\lambda_1 = 3$, $\lambda_2 = -2$ son reales y constituyen los dos valores propios de la matriz. Es posible por tanto la reducción a la forma canónica, y se tiene $\mu = 2$, $n_1 = 2$ y $n_2 = 1$.

El polinomio minimal de M depende del valor de a, como vamos a comprobar a continuación. En efecto, $\phi_M(x)$ solo puede tener una de las dos formas siguientes:

$$\phi_M^{(1)}(x) = (x - 3)(x + 2)$$

$$\phi_M^{(2)}(x) = (x - 3)^2(x + 2) .$$

Primer caso. $\phi_M^{(1)}(x) = (x - 3)(x + 2)$. Este será el polinomio minimal si se cumple la condición

$$\phi_M^{(1)}(M) = 0 \quad ,$$

o sea si

$$(M - 3E)(M + 2E) = 0 .$$

Ahora bien,

$$(M - 3E)(M + 2E) = \begin{pmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ b & c & -5 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ a & 5 & 0 \\ b & c & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 5a & 0 & 0 \\ ac & 0 & 0 \end{pmatrix}$$

por consiguiente este primer caso solo puede ocurrir si $a = 0$, o sea si M tiene la forma

$$M = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ b & c & -2 \end{pmatrix}$$

Dada la forma del polinomio minimal se tiene que $k_1^{(1)} = 1$ y $k_1^{(2)} = 1$. Las condiciones

$$\sum_{j=1}^{p_i} k_j^{(i)} = n_i \qquad \text{y} \qquad k_1^{(i)} \geqslant \ldots \geqslant k_{p_i}^{(i)}$$

implica entonces que

$$k_2^{(1)} = 1 \quad , \quad p_1 = 2 \quad , \quad \text{y} \quad p_2 = 1 \ .$$

Por consiguiente la forma canónica de la matriz M es

$$J^{(1)} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad ,$$

como se deduce del teorema 15.3.2.

Resulta así que todas las matrices de la forma (15.4.1) en que $a = 0$ son semejantes entre sí porque se reducen a la misma forma canónica $J^{(1)}$.

Segundo caso. $\phi_M^{(2)}(x) = (x - 3)^2(x + 2)$. Para que éste sea el polinomio minimal de M se requiere que sea $\phi_M^{(1)}(M) \neq 0$, o sea que $a \neq 0$. La forma del polinomio minimal implica ahora que

$$k_1^{(1)} = 2 \quad \text{y} \quad k_1^{(1)} = 1 \ ; \text{ luego } \quad p_1 = 1 \quad \text{y} \quad p_2 = 1.$$

La forma canónica será

$$J^{(2)} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Resulta así que todas las matrices (15.4.1) en que $a \neq 0$ son semejantes entre sí porque se reducen a la misma forma canónica $J^{(2)}$

Ejemplo C. Consideremos ahora las matrices de la forma

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

sobre C. El polinomio característico es

$$f_M(x) = \begin{vmatrix} (x - a) & -b \\ -c & (x - d) \end{vmatrix} = x^2 - (a + d)x + (ad - bc) \ .$$

Distinguiremos dos casos según sea diferente de cero o igual a cero el discriminante $(a - d)^2 + 4bc$ del trinomio $f_M(x)$.

Primer caso. $(a - d)^2 + 4bc \neq 0$. La ecuación $f_M(x) = 0$ tiene entonces dos raices diferentes λ_1 y λ_2. Por tanto

$$f_M(x) = (x - \lambda_1)(x - \lambda_2) ,$$

de donde resulta $\mu = 2$, $n_1 = 1$, $n_2 = 1$. El polinomio minimal es necesariamente

$$\phi_M^{(1)}(x) = (x - \lambda_1)(x - \lambda_2) .$$

Luego $k_1^{(1)} = 1$ y $k_1^{(2)} = 1$, y por consiguiente $p_1 = 1$ y $p_2 = 1$. La forma canónica es

$$M^{(1)} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} ,$$

y por tanto todas las matrices de orden 2 sobre $\mathbb{C}$ para las cuales las raices λ_1 y λ_2 $(\lambda_1 \neq \lambda_2)$ son las mismas, son semejantes entre sí.

Segundo caso. $(a - d)^2 + 4bc = 0$. La ecuación $f_M(x) = 0$ tiene una raiz doble que designaremos por λ. Por consiguiente

$$f_M(x) = (x - \lambda)^2 ,$$

de donde se sigue que $\mu = 1$ y $n_1 = 2$.

El polinomio minimal puede tomar una de las dos formas

$$\phi_M^{(2)} = x - \lambda$$

$$\phi_M^{(3)} = (x - \lambda)^2 ,$$

a cada una de las cuales corresponde uno de dos subcasos que vamos a considerar separadamente.

Primer subcaso. $\phi_M^{(2)} = x - \lambda$. Para que sea así debe cumplirse la condición $\phi_M^{(2)}(M) = 0$, o sea

$$M - \lambda E = \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = 0.$$

Por tanto debe ser $a = d = \lambda$ y $b = c = 0$; es decir que la matriz debe tener la forma

$$M^{(2)} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} ,$$

y por tanto debe coincidir con su propia forma canónica. Efectivamente $M^{(2)}$ no es semejante sino consigo misma, por cuanto, puesto que $M^{(2)} = \lambda E_2$, cualquiera que sea la matriz regular A, se tiene

$$M' = AM^{(2)}A^{-1} = \lambda A E_2 A^{-1} = \lambda E_2 = M^{(2)} .$$

Segundo subcaso. $\phi_M^{(3)} = (x - \lambda)^2$. Este caso ocurre cuando $\phi_M^{(2)}(M) \neq 0$ es decir cuando M no es de la forma λE_2. Se tiene entonces $\mu = 1$ $k_1^{(1)} = 2$, $p_1 = 1$. Por consiguiente la forma canónica es

$$M^{(3)} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} .$$

Resulta así que todas las matrices de orden 2 sobre $\mathbb{C}$ cuyo polinomio característico tienen la raiz doble λ, y que no se reducen a λE_2 , son semejantes entre sí.

Ejercicios

15.4.1. Escribir en forma explícita las siguientes matrices sobre $\mathbb{C}$;
(a) $J(2\,;3)$, (b) $J(1+i\,;4)$ (c) $J(i\,;3\,,2)$,
(d) $J(-1\,;3\,,2\,,1)$ (e) una matriz de Jordan con los bloques diagonales $J(2\,;3)$ y $J(1\,;4\,,3)$.

15.4.2. Hallar los polinomios característicos y minimales de las ma-

trices del ejercicio anterior.

15.4.3. Hallar las formas canónicas de Jordan de las matrices siguientes:

(a) $\begin{pmatrix} 3 & -2 \\ 1 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 3 & 0 & -3 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

15.4.4. Averiguar si las matrices siguientes sobre $\mathbb{R}$ son nilpotentes y hallar sus formas canónicas.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

15.4.5. Caracterizar las diferentes clases de equivalencia de las matrices complejas de orden 3, en términos de las correspondientes formas canónicas.

§ 16. Divisores elementales de un endomorfismo.

En el presente parágrafo se supone que el campo $\mathbb{F}$ de escalares es algebraicamente cerrado. Puede suponerse que es igual a $\mathbb{C}$.

En la sección 16.3, en particular, se deduce un procedimiento para obtener la forma canónica de Jordan de un endomorfismo.

16.1. Los polinomios $D_j(x)$ de una matriz y de un endomorfismo

Sea M una matriz de orden n sobre un campo algebraicamente cerrado $\mathbb{F}$. Los menores de orden j de la matriz $xE - M$ son polinomios en x.

Definición 16.1.1. Se designa por $D_j(x)$ al polinomio que es el *M.C.D.* de los menores de orden j de la matriz $xE - M$. Convendre-

mos, además, en que $D_0(x) = 1$.

En particular, $D_n(x)$ es el determinante de la matriz $xE - M$ o sea el polinomio característico $f_M(x)$ de la matriz M. Uno de nuestros propósitos consistirá en probar que todos los polinomios $D_j(x)$ son invariantes, es decir que no varian cuando la matriz M se cambia por otra semejante.

Observemos que $D_{n-1}(x)$ divide a $D_n(x)$. En efecto, la definición de $D_{n-1}(x)$ implica que todos los menores de orden $n-1$ son divisibles por $D_{n-1}(x)$. Si se desarrolla el determinante $D_n(x)$ según los elementos de una fila cualquiera, se obtiene una suma cada uno de cuyos sumandos es un producto de un elemento de la fila en cuestión por su cofactor. Se sigue que $D_n(x)$ es divisible por $D_{n-1}(x)$. Semejantemente se prueba ,en general, que cada menor de orden j, y por consiguiente $D_j(x)$, es divisible por $D_{j-1}(x)$.

Ejemplo. Consideremos la matriz

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

Se tiene entonces $D_3(x) = (x - \lambda_o)^3$. Entre los menores de orden 2 está 1, por tanto $D_2(x) = 1$. Lo mismo ocurre con los menores de orden 1 y por tanto $D_1(x) = 1$.

Lema 16.1.1. Si M es una matriz cualquiera $n \times n$ cuyos términos son polinomios en x, y A es una matriz regular escalar del mismo orden, entonces los *M.C.D.* de los menores de orden k de las matrices M, AM y MA son los mismos.

Demostración. Consideremos la matriz M y MA. Si designamos por a_j^i a los términos de la primera y por b_j^i a los de la segunda, tendremos

$$b_j^i = \sum_{r=1}^{n} a_r^i \, \alpha_j^r$$

donde α_j^r son los términos de la matriz A, independientes por tan-

to de x. Consideremos un menor cualquiera m de MA de orden k (definición 7.4.1) formado por los elementos que pertenecen tanto a las filas $i_1, \ldots, i_k$ como a las columnas $j_1, \ldots, j_k$. Empleando propiedades conocidas de los determinantes, y en particular la propiedad 4. de la sección 7.1, podemos escribir

$$m = \begin{vmatrix} b^{i_1}_{j_1} & \cdots & b^{i_1}_{j_k} \\ \vdots & & \vdots \\ b^{i_k}_{j_1} & \cdots & b^{i_k}_{j_k} \end{vmatrix} = \begin{vmatrix} \sum_{r=1}^{n} \alpha^{r}_{j_1} a^{i_1}_{r} & \cdots & \sum_{r=1}^{n} \alpha^{r}_{j_k} a^{i_1}_{r} \\ \vdots & & \vdots \\ \sum_{r=1}^{n} \alpha^{r}_{j_1} a^{i_k}_{r} & \cdots & \sum_{r=1}^{n} \alpha^{r}_{j_k} a^{i_k}_{r} \end{vmatrix} =$$

$$= \sum_{(r_1, \ldots, r_k)} \alpha^{r_1}_{j_1} \alpha^{r_2}_{j_2} \cdots \alpha^{r_k}_{j_k} \begin{vmatrix} a^{i_1}_{r_1} & a^{i_1}_{r_2} & \cdots & a^{i_1}_{r_k} \\ \vdots & \vdots & & \vdots \\ a^{i_k}_{r_1} & a^{i_k}_{r_2} & \cdots & a^{i_k}_{r_k} \end{vmatrix},$$

donde la suma se extiende a todas las sucesiones $(r_1, \ldots, r_k)$ de k escalares que pueden tomar los valores $1, \ldots, n$. La última suma consta por tanto de n^k sumandos, cada uno de los cuales es el producto de un escalar por un polinomio que es un menor de orden k de M. Resulta así que cada divisor de todos los menores de orden k de M es divisor de cada menor de orden k de MA.

Para probar la recíproca podemos aplicar exactamente el mismo razonamiento, reemplazando simplemente las matrices M y MA antes consideradas por las matrices MA y $(MA)A^{-1} = M$, respectivamente. De esa manera queda probado que el *M.C.D.* de los menores de orden k de M es idéntico al *M.C.D.* de los menores de orden k de MA.

Un prodecimiento análogo puede emplearse para las matrices M y AM. ■

Lema 16.1.2. Las matrices semejantes tienen idénticos pilonomios $D_j(x)$.

Demostración. Sean M y $M' = AMA^{-1}$ dos matrices semejantes. En virtud del lema 16.1.1, el *M.C.D.* de los menores de orden j de $xE - M$ es el mismo que el de los menores de orden j de

$$(xE - M)A^{-1} ,$$

y por tanto de $A(xE - M)A^{-1} = xE - M'$. Por tanto los polinomios $D_j(x)$ de M son idénticos a los de M'. ∎

El teorema siguiente es consecuencia inmediata del lema anterior.

Teorema 16.1.1. Sea T un endomorfismo del espacio vectorial *n-dimensional* V. Entonces el *M.C.D.*, $D_j(x)$, de los menores de orden j de la matriz $xE - M$, donde M es una matriz que representa a T en una base cualquiera, no depende de la elección de esa base.

Definición 16.1.2. Los polinomios $D_j(x)$ de un endomorfismo T, son los polinomios $D_j(x)$ de cualquiera de las matrices que representan a T.

16.2. Cálculo de los polinomios $D_j(x)$.

Trataremos ahora acerca del cálculo de los polinomios $D_j(x)$ de una matriz que corresponde a un endomorfismo T. Según la definición 16.1.2., esos polinomios serán los $D_j(x)$ de T, y no dependen de la base elegida. En particular podemos suponer que la base elegida es tal que, respecto de ella, la matriz que corresponde a T es una matriz de Jordan. Nuestro problema consiste por tanto en calcular los polinomios $D_j(x)$ de una matriz de Jordan J.

En primer lugar vamos a calcular los $D_j(x)$ para una matriz $k \times k$ de la forma

$$J(\lambda\ ;\ k) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \qquad (16.2.1)$$

es decir para uno de los bloques diagonales de una matriz de Jordan. Evidentemente se tiene que

$$D_n(x) = (x - \lambda)^k$$

Si en la matriz $B = xE - J(\lambda ; k)$ eliminamos la primera columna y la última fila obtendremos una matriz B_1 cuya diagonal está formada por términos iguales a -1, y cuyos términos situados encima de esa diagonal son todos ceros. Luego se tiene que

$$D_{n-1}(x) = 1,$$

por cuanto $D_{n-1}(x)$ es un polinomio mónico que debe dividir al determinante de B_1 que es $+1$ o -1. Puesto que $D_i(x)$, para $1 \leqslant i < n - 1$, es un polinomio mónico divisor de $D_{n-1}(x)$, debera ser $D_i(x) = 1$ para $i = 1, \ldots, n$. En resumen pues, los polinomios $D_j(x)$ de una matriz $J(\lambda ; k)$ son

$$(x - \lambda)^k , 1 , \ldots , 1 .$$

Lema 16.2.1. Si M es una matriz $n \times n$ que puede descomponerse en bloques en la forma

$$M = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix}$$

donde M_1 es un bloque $n_1 \times n_1$ y M_2 es un bloque $n_2 \times n_2$, $(n_1 + n_2 = n)$, entonces los menores de M de orden m, que no son nulos son de la forma

$$\Delta_m = \pm \Delta^{(1)}_{m_1} \Delta^{(2)}_{m_2} \quad (m_1 , m_2 = 0,1,\ldots,m; m_1 + m_2 = m),$$

donde $\Delta^{(1)}_{m_1}$, para $m_1 \neq 0$, es uno de los menores de orden m_1 de M_1 y $\Delta^{(2)}_{m_2}$, para $m_2 \neq 0$, es uno de los menores de orden m_2 de M_2 y además

$$\Delta^{(1)}_0 = \Delta^{(2)}_0 = 1.$$

Demostración. Supongamos que Δ_m es un menor no nulo de orden m. Consideremos sucesivamente los tres casos posibles

1° Cada fila del menor Δ_m que se considera está contenida en una de las n_1 primeras filas de M. Si el menor no estuviera contenido en M_1, una de sus columnas estaría formada por ceros, y por tanto ese menor sería igual a cero contra lo supuesto. Por tanto Δ_m debe estar contenido en M_1, y coincide evidentemente con un menor no nulo $\Delta_m^{(1)}$ de M_1. Se tiene por tanto

$$\Delta_m = \Delta_m^{(1)} \; \Delta_o^{(2)} \quad ,$$

como lo prescribe el enunciado.

2° Cada fila del menor Δ_m que se considera está contenida en una de las n_2 últimas filas de M. Este caso es análogo al anterior y se tiene entonces

$$\Delta_m = \Delta_o^{(1)} \; \Delta_m^{(2)}$$

3° Cada una de las m_1 $(0 < m_1 \leqslant n_1)$ primeras filas de Δ_m está contenida en una de las n_1 primeras filas de M, y cada una de las m_2 $(0 < m_2 \leqslant n_2)$ últimas filas de Δ_m está contenida en una de las n_2 filas de M. En este caso, si se tiene presente el desarrollo del menor Δ_m, que por hipótesis es diferente de cero, mediante la fórmula de Laplace relativa a sus m_1 primeras filas, se ve de inmediato que es indispensable que sus m_1 primeras columnas estén entre las n_1 primeras columnas de M y que sus m_2 últimas columnas estén entre las n_2 últimas columnas de M; y entonces dicho desarrollo de Laplace se reduce a un solo término $\pm\Delta_{m_1}^{(1)} \; \Delta_{m_2}^{(2)}$ donde $\Delta_{m_1}^{(1)}$ es un menor de M_1 y $\pm\Delta_{m_2}^{(2)}$ es su cofactor en Δ_m, es decir que $\Delta_m^{(2)}$ es un menor de M_2.

Corolario. El teorema anterior se generaliza en forma obvia al caso en que M es una matriz de la forma

$$\Delta = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_p \end{pmatrix}$$

Demostración. Puede hacerse por inducción respecto de p. ∎

En el teorema siguiente se establece la forma de los polinomios $D_j(x)$ para una matriz de Jordan J, cuyos valores propios son $\lambda_1, \dots, \lambda_\mu$ y cuyos bloques diagonales son

$$J(\lambda_i \; ; \; k_1^{(i)}, \dots, k_{p_i}^{(i)}) \quad , \quad i = 1, \dots, \mu$$

Recordemos que $k_1^{(i)} \geqslant k_2^{(i)} \geqslant \dots \geqslant k_{p_i}^{(i)}$, y

$$k_1^{(i)} + \dots + k_{p_i}^{(i)} = n_i$$

donde n_i es la multiplicidad de la raíz λ_i de la ecuación característica.

Teorema 16.2.1. Los polinomios $D_j(x)$ de la matriz de Jordan J son

$$D_n(x) = \prod_{i=1}^{\mu} (x - \lambda_i)^{k_1^{(i)} + \dots + k_{p_i}^{(i)}}$$

$$D_{n-1}(x) = \prod_{i=1}^{\mu} (x - \lambda_i)^{k_2^{(i)} + \dots + k_{p_i}^{(i)}}$$

$$D_{n-2}(x) = \prod_{i=1}^{\mu} (x - \lambda_i)^{k_3^{(i)} + \dots + k_{p_i}^{(i)}}$$

. . . . . .

A partir de $D_{n-p_1}(x)$ el factor en $(x - \lambda_1)$ es reemplazado por 1; a partir de $D_{n-p_2}(x)$ el factor en $(x - \lambda_2)$ es sustituido por 1, etc.

Demostración. Cada bloque diagonal $J(\lambda_i \; ; \; k_1^{(i)}, k_2^{(i)}, \dots, k_{p_i}^{(i)})$ puede descomponerse a su vez en bloques en la forma

$$\begin{pmatrix} J(\lambda_i\ ;\ k_1^{(i)}) & 0 & \cdots & 0 \\ 0 & J(\lambda_i\ ;\ k_2^{(i)}) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & J(\lambda_i\ ;\ k_{p_i}^{(i)}) \end{pmatrix}$$

Los $J(\lambda_i\ ;\ k_j^{(i)})$ se llaman sub-bloques diagonales de J.

A cada sub-bloque $J(\lambda_i\ ;\ k_j^{(i)})$ le corresponde en la matriz

$$B = xE - J$$

un sub-bloque

$$B_j^{(i)} = \begin{pmatrix} x-\lambda_1 & -1 & 0 & \cdots & 0 \\ 0 & x-\lambda_1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & x-\lambda_1 \end{pmatrix}$$

Todos estos sub-bloques, ordenados en la forma

$$B_1^{(1)},\ B_2^{(1)},\ \ldots,\ B_{p_1}^{(1)},\ B_1^{(2)},\ B_2^{(2)},\ \ldots,\ B_{p_2}^{(2)},\ \ldots,\ B_1^{(\mu)},\ \ldots,\ B_{p_\mu}^{(\mu)},$$

los designaremos en forma más abreviada por

$$B_1,\ B_2,\ \ldots,\ B_p \qquad (p = p_1 + \ldots p_\mu).$$

Calculemos primero $D_n(x)$, es decir el determinante de

$$B = xE - J\ .$$

Este determinante es evidentemente el producto de los determinantes de las matrices B_i, o sea

$$D_n(x) = \prod_{i=1}^{\mu} (x - \lambda_i)^{k_1^{(i)} + k_2^{(i)} + \ldots + k_{p_i}^{(i)}}$$

Calculemos ahora $D_{n-1}(x)$. Puesto que $D_{n-1}(x)$ es un factor de $D_n(x)$,

debe ser un producto de potencias de

$$(x - \lambda_1) , (x - \lambda_2) , \ldots , (x - \lambda_\mu)$$

Debemos calcular los exponentes de esas potencias. Calculemos por ejemplo el grado de $x - \lambda_1$ en $D_{n-1}(x)$. Observemos que cada menor no nulo de $B = xE - J$ de orden $n - 1$ es de la forma

$$\Delta_{n-1} = \pm \Delta^{(1)}_{t_1} \quad \Delta^{(2)}_{t_2} \ldots \Delta^{(p)}_{t_p}$$

donde $t_1 + t_2 + \ldots + t_p = n - 1$, y $\Delta^{(i)}_{t_i}$ denota a un menor de orden t_i de la matriz B_i. Puesto que la suma de los órdenes de los menores $\Delta^{(i)}_{t_i}$ es $n - 1$, esos órdenes serán iguales a los de los correspondientes sub-bloques B_i, salvo exactamente el orden de uno de los menores que deberá ser menor en 1 que el orden de la correspondiente matriz B_i, y que por tanto se obtiene eliminando una fila y una columna de dicha B_i. Ahora bien, si de un sub bloque B_i se eliminan una fila y una columna apropiadas se puede obtener un menor igual a ± 1. Luego es posible elegir Δ_{n-1} de manera que uno de los $\Delta^{(i)}_{t_i}$ sea igual a ± 1, y que los otros sean iguales a los determinantes de los correspondientes sub-bloques. Se sigue de aquí que para obtener un menor Δ_{n-1} para el cual

$$x - \lambda_1$$

tenga el menor grado posible, basta elegirlo de manera que aquel $\Delta^{(i)}_{t_i}$ cuyo orden es menor en una unidad que el orden del correspondiente sub-bloque B_i, y que por tanto puede suponerse igual a $+1$ o -1, sea tal que dicho B_i corresponda al sub-bloque

$$J(\lambda_1 \; ; \; k^{(1)}_j)$$

para el que $k^{(1)}_j$ tiene el mayor valor posible. Es decir que B_i debe corresponder a $J(\lambda_1 \; ; \; k^{(1)}_1)$. Cumplida esa condición, los restantes $\Delta^{(i)}_{t_i}$ que corresponden al menor Δ_{n-1} elegido, serán idénticos

a los determinantes de los correspondientes sub-bloques B_i, y por consiguiente la potencia de $x - \lambda_1$ que figura en Δ_{n-1} es

$$(x - \lambda_1)^{k_2^{(1)} + \ldots + k_{p_1}^{(1)}} ,$$

y es la mínima posible. Por tanto el *M.C.D.*, $D_{n-1}(x)$, de los menores de orden $n - 1$ de J deben contener al factor

$$(x - \lambda_1)^{k_2^{(1)} + \ldots + k_{p_1}^{(1)}} .$$

Semejantemente, para obtener un menor Δ_{n-2} de orden $n - 2$ con la menor potencia de $x - \lambda_1$ basta con que en la descomposición

$$\Delta_{n-2} = \pm \Delta_{t_1}^{(1)} \Delta_{t_2}^{(2)} \ldots \Delta_{t_p}^{(p)} \qquad (\Sigma\, t = n - 2)$$

que corresponde a ese menor, dos de los factores sean el resultado de suprimir una columna y una fila apropiadas de los sub-bloques de órdenes $k_1^{(1)}$ y $k_2^{(1)}$ que corresponden a λ_1. Así, $D_{n-2}(x)$ contiene a $(x - \lambda_1)$ a la potencia $k_3^{(1)} + \ldots + k_{p_1}^{(1)}$. Puede procederse así sucesivamente y establecer que $D_{n-p}(x)$ contiene a $(x - \lambda_1)$ a la potencia de exponente $k_{\nu+1}^{(1)} + \ldots + k_{p_1}^{(1)}$, pero que los polinomios $D_{n-p_1}(x)$, $D_{n-p_1-1}(x)$, . . . , $D_1(x)$ no contiene al factor $x - \lambda_1$.

Un razonamiento exactamente análogo es aplicable a la determinación de los grados de $x - \lambda_2$, $x - \lambda_3$, etc. en $D_j(x)$. Con lo cual queda probado el teorema. ∎

Corolario. Si la forma canónica J posee los bloques diagonales

$$J(\lambda_i\,;\, k^{(i)}) \quad , \quad i = 1, \ldots, \mu$$

entonces $D_n(x) = (x - \lambda_1)^{k^{(1)}} (x - \lambda_2)^{k^{(2)}} \ldots (x - \lambda_\mu)^{k^{(\mu)}}$

$$D_{n-1}(x) = D_{n-2}(x) = \ldots = 1.$$

Demostración. Este corolario resulta inmediatamente del teorema anterior. ∎

16.3. Divisores elementales y cálculo de la forma canónica de Jordan.

Supondremos que T es un endomorfismo de un espacio vectorial *n-dimensional* sobre un campo $\mathbb{F}$ algebraicamente cerrado, y que J es la matriz de Jordan de ese endomorfismo. Consideraremos los polinomios $D_j(x)$ de dicha matriz, o sea los polinomios $D_j(x)$ del endomorfismo T.

Definición 16.3.1. Los polinomios

$$E_j(x) = \frac{D_j(x)}{D_{j-1}(x)}$$

se llaman *divisores elementales* del endomorfismo T.

Si la forma canónica J posee los bloques diagonales

$$J(\lambda_i\ ;\ k_1^{(i)}, \ldots, k_{p_i}^{(i)}) \qquad , \qquad i = 1, \ldots, \mu$$

entonces resulta del teorema 16.3.1. que los divisores elementales son

$$E_n(x) = (x - \lambda_1)^{k_1^{(1)}} (x - \lambda_2)^{k_1^{(2)}} \ldots$$

$$E_{n-1}(x) = (x - \lambda_1)^{k_2^{(1)}} (x - \lambda_2)^{k_2^{(2)}} \ldots$$

$$E_{n-2}(x) = (x - \lambda_1)^{k_3^{(1)}} (x - \lambda_2)^{k_3^{(2)}} \ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

En las expresiones anteriores, a partir de $E_{n-p_i}(x)$, para $i = 1, 2, \ldots, \mu$ el factor en $(x - \lambda_i)$ se debe sustituir por 1. El número p_i es aquél para el cual es $k_1^{(i)} + k_2^{(i)} + \ldots + k_{p_i}^{(i)} = n_i$ don-

de n_i es la multiplicidad de la raiz λ_i del polinomio característico.

Es claro que, dados los divisores elementales $E_n(x), E_{n-1}(x), \ldots$ la forma canónica de Jordan de la matriz de T queda bien determinada. En efecto, los valores propios son las raices de la ecuación $E_n(x) = 0$, los invariantes $k_1^{(i)}$, $k_2^{(i)}$, . . . , $k_{p_i}^{(i)}$ son los exponentes de $(x - \lambda_i)$ en los divisores elementales $E_n(x)$, $E_{n-1}(x), \ldots, E_{n-p_i-1}(x)$ respectivamente. Además, el polinomio característico de T es

$$f_T(x) = D_n(x) = E_n(x)\ E_{n-1}(x) \ . \ . \ . \ ,$$

y su polinomio minimal es

$$\phi_T(x) = E_n(x) \ ,$$

como resulta de la última parte del enunciado del teorema 15.3.1.

Resulta de esta manera un procedimiento para obtener la forma canónica de la matriz de un endomorfismo. Basta para ello hallar los divisores elementales de la matriz de la transformación, relativa a una base cualquiera. Cuando esos divisores se escriben en la forma $(x - \lambda_1)^{k_1}\ (x - \lambda_2)^{k_2}$ . . . , se deducen, como acabamos de mostrar, los valores propios y los sistemas de invariantes de la transformación.

Ejemplo. Calcular la forma canónica de Jordan de la matriz

$$M = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}$$

Consideremos la matriz

$$xE - M = \begin{pmatrix} x & -3 & -3 \\ 1 & x - 8 & -6 \\ -2 & 14 & x + 10 \end{pmatrix}$$

(c)
$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

(d)
$$\begin{pmatrix} 1 & 1 & -1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix}$$

16. Divisores elementales.

El polinomio característico es $D_3(x) = (x + 1)^2 x$, por tanto los valores propios son -1 con multiplicidad $n_1 = 2$ y 0 con multiplicidad 1. Entre los menores de orden 2 se encuentran los polinomios $\begin{vmatrix} x & -3 \\ 1 & -6 \end{vmatrix} = -6(x - \frac{1}{2})$ y $\begin{vmatrix} x & -3 \\ -2 & 14 \end{vmatrix} = 14(x - \frac{3}{7})$, que son primos entre sí. Luego $D_2(x) = 1$. Es claro además que $D_1(x) = 1$. Los divisores elementales son

$$E_3(x) = \frac{D_3(x)}{D_2(x)} = (x + 1)^2 x ,$$

y los restantes son iguales a 1. se tiene entonces $k_1^{(1)} = 2 = n_1$ y $k_1^{(2)} = 1 = n_2$. La matriz de Jordan es por tanto

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

Ejercicios

16.3.1. Calcular la forma de Jordan de las matrices

(a) $\begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

CAPITULO V

ESPACIOS DE PRODUCTO INTERNO

En el presente capítulo emplearemos la letra $\mathbb{F}$, exclusivamente para designar indiferentemente al campo $\mathbb{R}$ de los números reales o al campo $\mathbb{C}$ de los números complejos. Los campos de escalares que consideraremos serán exclusivamente uno de esos dos.

§ 17. El producto interno

17.1. Definiciones fundamentales.

Se dice que un espacio vectorial V sobre $\mathbb{F}$, donde $\mathbb{F}$ es uno de los campos $\mathbb{R}$ o $\mathbb{C}$, es un espacio de producto interno, si es definida en él una función

$$(\ , \ ) : V \times V \to \mathbb{F} \ ,$$

cuyo valor para el par de vectores $x , y \in V$ se representa por (x , y) , se llama *producto interno*, y goza de las siguientes propiedades:

A. En el caso en que $\mathbb{F} = \mathbb{R}$,

A.1. $(x , y) = (y , x)$ *(simetría)*

A.2. $(\alpha x + \beta y , z) = \alpha(x , z) + \beta(y , z)$, $\alpha , \beta \in \mathbb{R}$
(linealidad respecto del primer factor)

A.3. $(x , x) \geq 0$ y $(x , x) = 0$ si y solo si $x = 0$.
(positividad definida)

De A.1. y A.2. se deduce la propiedad

A.4. $(x\,,\,\alpha y + \beta z) = \alpha(x\,,\,y) + \beta(x\,,\,z) \quad , \quad \alpha\,,\beta \in \mathbb{R}$

(linealidad respecto del segundo factor) [1]

B. En el caso en que $\mathbb{F} = \mathbb{C}$,

B.1. $(x\,,\,y) = \overline{(y\,,\,x)}$ *(propiedad hermitiana)*

B.2. $(\alpha x + \beta y\,,\,z) = \alpha(x\,,\,z) + \beta(y\,,\,z) \quad , \quad \alpha\,,\beta \in \mathbb{C}$

(linealidad respecto del primer factor)

B.3. $(x\,,\,x) \geqslant 0$ y $(x\,,\,x) = 0$ si y solo si $x = 0$

(propiedad definida)

De B.1. y B.2. se deduce la propiedad B.4.

B.4. $(x\,,\,\alpha y + \beta z) = \overline{\alpha}(x\,,\,y) + \overline{\beta}(x\,,\,z) \quad , \quad \alpha\,,\beta \in \mathbb{C}$

(linealidad conjugada respecto del segundo factor)

Las propiedades B.2. y B.4. se expresan diciendo que cuando $\mathbb{F} = \mathbb{C}$ la función $(\ ,\ )$ es *sesquilineal*.

En el caso en que $\mathbb{F} = \mathbb{R}$ el espacio de producto interno se llama *espacio real de producto interno o bien espacio euclideano*. En el caso en que $\mathbb{F} = \mathbb{C}$ se llama *espacio complejo de producto interno o espacio unitario.*

Observaciones. 1. La condición B.3. supone que $(x\,,\,x)$ es real, lo cual es consecuencia inmediata de B.1. porque cualquiera que sea $x \in V$ se debe cumplir que $(x\,,\,x) = \overline{(x\,,\,x)}$.

1) Las propiedades A.2. y A.4. se expresan diciendo que el producto interno es una *función bilineal*. Dicha función es *simétrica* en virtud de la propiedad A.1. La propiedad A.3. expresa que es una función *positiva definida*. En resumen, el producto interno es una *función bilineal, simétrica, positiva definida*. Como consecuencia, se cumple que

$$[(x\,,\,y) = 0 \quad , \quad \forall\, y \in V] \Rightarrow (x\,,\,x) = 0 \quad \Rightarrow \quad x = 0 \quad ,$$

o sea

$$[(x\,,\,y) = 0 \quad , \quad \forall\, y \in V] \Rightarrow x = 0 \,,$$

relación que se expresa diciendo que es también *no degenerada*.

2. Podrámos haber considerado únicamente las condiciones B por cuanto estas se reducen evidentemente a las condiciones A cuando $\mathbb{F} = \mathbb{R}$.

3. Es evidente que un subespacio V_1 de un espacio de producto interno es también un espacio de producto interno respecto de la restricción de la función $(\ ,\ )$ a $V_1 \times V_1$.

Ejemplos. A. Si en $\mathbb{R}^n$ $(\mathbb{C}^n)$ consideremos los vectores

$$x = (\xi^1, \ldots, \xi^n) \quad \text{e} \quad y = (\eta^1, \ldots, \eta^n),$$

puede definirse un producto interno mediante la fórmula

$$(x, y) = \xi^1 \eta^1 + \cdots + \xi^n \eta^n \qquad (17.1.2)$$

El producto interno así definido se llama *producto interno ordinario de* $\mathbb{R}^n$ $(\mathbb{C}^n)$.

B. Si el espacio $C[0, 1]$ de las funciones complejas continuas en el intervalo $[0, 1]$, el producto interno de las funciones $x = x(t)$ e $y = y(t)$ se define por la fórmula

$$(x, y) = \int_0^1 x(t)\, \overline{y(t)}\, dt.$$

Se prueba, haciendo uso de las propiedades de la integral, que se cumplen las propiedades del producto interno.

Ejercicios

***17.1.1.** Demostrar que si x e y son vectores de un espacio de producto interno V de dimensión n, y $(\xi^1, \ldots, \xi^n)$ y $(\eta^1, \ldots, \eta^n)$ son las componentes de esos vectores respecto de una base $\{e_1, \ldots, e_n\}$ de V, entonces

$$(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}\, \xi^i\, \overline{\eta^j} \qquad (17.1.3)$$

donde

$$g_{ij} = (e_i, e_j)$$

17.1.2. Si V es el espacio $\mathbb{R}^n$ con el producto interno ordinario definido por (17.1.1) ¿Cómo concuerda la fórmula (17.1.3) con la fórmula (17.1.1)?

17.2. Norma o longitud de un vector de un espacio de producto interno.

Definición 17.2.1. Se llama *norma o longitud* de un vector x de un espacio de producto interno al número real no negativo

$$\|x\| = \sqrt{(x\ ,\ x)}\ .$$

Es evidente que el único vector de longitud nula es el vector nulo. Un vector cuya longitud es igual a *1* se llama *vector unitario*.

Ejemplos. Si $x = (\xi^i) \in \mathbb{R}^n$, entonces es

$$\|x\| = \sqrt{\sum_{i=1}^{n} (\xi^i)^2}\ ;$$

y si $x = (\xi^i) \in \mathbb{C}^n$,

$$\|x\| = \sqrt{\sum_{i=1}^{n} \xi^i\ \overline{\xi^i}} = \sqrt{\sum_{i=1}^{n} |\xi^i|^2}$$

Teorema 17.2.1. Para dos vectores cualesquiera x , y de un espacio de producto interno V se cumple la *desigualdad de Schwarz* [1)]

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

La igualdad solamente tiene lugar si x e y son linealmente dependientes.

Demostración. En virtud de la positividad definida, cualesquiera que sean los escalares α , $\beta \in \mathbb{F}$, se tiene

$$(\alpha x - \beta y\ ,\ \alpha x - \beta y) \geqslant 0 \qquad (17.2.1)$$

Por otra parte podemos escribir

1) Algunas veces se designa a esta desigualdad con el nombre de desigualdad de Cauchy - Schwarz o de Buniakowsky.

$$(\alpha x - \beta y\ ,\ \alpha x - \beta y) = \alpha\overline{\alpha}(x\ ,\ x) - \alpha\overline{\beta}(x\ ,\ y) - \beta\overline{\alpha}(y\ ,\ x) + \beta\overline{\beta}(y\ ,\ y)$$
$$= |\alpha|^2\|x\|^2 - (\alpha\overline{\beta}(x\ ,\ y) + \overline{\alpha\overline{\beta}(x\ ,\ y)}) + |\beta|^2\|y\|^2$$
$$= |\alpha|^2\|x\|^2 - 2\mathcal{R}(\alpha\overline{\beta}(x\ ,\ y)) + |\beta|^2\|y\|^2. \qquad (17.2.2)$$

Esta última relación es válida cualesquiera que sean los escalares α y β. En particular podemos suponer que

$$\alpha = \|y\|^2 \qquad \text{y} \qquad \beta = (x\ ,\ y)\ .$$

En ese caso se tiene

$$\alpha\overline{\beta}(x\ ,\ y) = \|y\|^2\overline{(x\ ,\ y)}(x\ ,\ y) = \|y\|^2|(x\ ,\ y)|^2\ ,$$

y llevando a (17.2.2) se obtiene

$$(\alpha x - \beta y\ ,\ \alpha x - \beta y) = \|y\|^4\|x\|^2 - 2\|y\|^2|(x\ ,\ y)|^2 + \|y\|^2|(x\ ,\ y)|^2$$
$$= \|y\|^2(\|x\|^2\|y\|^2 - |(x\ ,\ y)|^2). \qquad (17.2.3)$$

De la desigualdad (17.2.1) se sigue entonces

$$\|x\|^2\|y\|^2 - |(x\ ,\ y)|^2 \geqslant 0\ ;$$

y por consiguiente, cualesquiera que sean $x\ ,\ y \in V$, se tiene

$$\|x\|^2\|y\|^2 \geqslant |(x\ ,\ y)|^2\ ,$$

de donde resulta

$$\|x\|\cdot\|y\| \geqslant |(x\ ,\ y)|$$

que es la desigualdad de Schwarz.

Vamos a probar ahora la segunda parte del teorema. Probaremos primero que si x e y son *l.d.* la desigualdad de Schwarz se reduce a una igualdad. Tal cosa es evidente si x o y fueran nulos. Supongamos pues que x e y son diferentes de cero. Existirá entonces un escalar α tal que $y = \alpha x$, y se tiene por tanto

$$|(x\ ,\ y)| = |(x\ ,\ \alpha x)| = |\overline{\alpha}(x\ ,\ x)| = |\alpha|(x\ ,\ x)$$

y

$$\|x\|\cdot\|y\| = \sqrt{(x\ ,\ x)}\ \sqrt{(\alpha x\ ,\ \alpha x)} = \sqrt{(x\ ,\ x)}\ \sqrt{\alpha\overline{\alpha}(x\ ,\ x)} = |\alpha|(x\ ,\ x).$$

Luego la desigualdad de Schwarz a la igualdad

$$|(x\ ,\ y)| = \|x\| \cdot \|y\| \ . \qquad (17.2.4)$$

Inversamente, supongamos que se cumple la igualdad (17.2.4). Vamos a probar que en tal caso los vectores x e y son *l.d.* Podemos suponer que x e y son diferentes de cero, pues en caso contrario la afirmación es obvia. Hemos visto que si $\alpha = \|y\|^2$ y $\beta = (x\ ,\ y)$ se cumple la igualdad (17.2.3) cualesquiera que sean $x\ ,\ y \in V$. La relación (17.2.4) que suponemos cumplida, implica entonces que

$$(\alpha x - \beta y\ ,\ \alpha x - \beta y) = 0,$$

y por tanto

$$\alpha x - \beta y = 0.$$

Sustituyendo α y β por los valores que les hemos atribuido resulta

$$x = \frac{(x\ ,\ y)}{\|y\|^2} y \qquad , \qquad (17.2.5)$$

lo cual prueba que x e y son *l.d.* ■

Observación 1. Si x e y son dos vectores no nulos *l.d.* y designamos por e al vector unitario $y/\|y\|$, la relación (17.2.5) puede escribirse en la forma

$$x = (x\ ,\ e)\, e$$

Observación 2. Es importante observar que la desigualdad de Schwarz se demuestra sin hacer uso de la segunda parte del axioma A.3 (o B.3), es decir utilizando únicamente el hecho de que el producto interno es una función *no negativa* y no necesariamente positiva definida.

Corolario. Dados dos conjuntos de n números complejos (reales) arbitrarios $(\xi^1, \ldots, \xi^n)$ y $(\eta^1, \ldots, \eta^n)$, se cumple la *desigualdad de Cauchy.*

$$\left| \sum_{i=1}^{n} \xi^i \overline{\eta^i} \right| \leqslant \sqrt{\sum_{i=1}^{n} |\xi^i|^2} \quad \sqrt{\sum_{i=1}^{n} |\eta^i|^2} \qquad (17.2.6)$$

$$\left(\text{resp.} \quad \left|\sum_{i=1}^{n} \xi^i \eta^i\right| \leqslant \sqrt{\sum_{i=1}^{n} (\xi^i)^2}\ \sqrt{\sum_{i=1}^{n} (\eta^i)^2}\ \right) \qquad (17.2.7)$$

Demostración. Si se supone que $x = (\xi^1, \ldots, \xi^n)$ e $y = (\eta^1, \ldots, \eta^n)$ son dos vectores de $\mathbb{C}^n$ ($\mathbb{R}^n$) el corolario resulta aplicando la desigualdad de Schwarz en los respectivos espacios provistos del producto interno ordinario. ■

En el teorema siguiente se establecen algunas propiedades fundamentales de la norma de los espacios vectoriales de producto interno.

Teorema 17.2.2. La norma de un espacio de producto interno tiene las propiedades siguientes

1. $\|x\| \geqslant 0$ y $\|x\| = 0$ si y solo si $x = 0$
2. $\|\alpha x\| = |\alpha|\,\|x\|$, cualesquiera que sean $x \in V$ y $\alpha \in \mathbb{F}$.
3. $\|x\| - \|y\| \leqslant \|x + y\| \leqslant \|x\| + \|y\|$ *(desigualdad triangular)*

Demostración 1. Estas propiedades son consecuencia de la positividad definida.

2. $\|\alpha x\| = \sqrt{(\alpha x\ ,\ \alpha x)} = \sqrt{\alpha\bar{\alpha}(x\ ,\ x)} = |\alpha| \cdot \|x\|$.

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

$$= \|x\|^2 + (x\ ,\ y) + \overline{(x\ ,\ y)} + \|y\|^2$$

$$\begin{cases} \leqslant \|x\|^2 + 2|(x\ ,\ y)| + \|y\|^2 \leqslant \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\ \qquad = (\|x\| + \|y\|)^2 \\ \geqslant \|x\|^2 - 2|(x\ ,\ y)| + \|y\|^2 \geqslant \|x\|^2 - 2\|x\| \cdot \|y\| + \|y\|^2 \\ \qquad = (\|x\| - \|y\|)^2 \end{cases}$$

Por tanto

$$\|x\| - \|y\| \leqslant \|x + y\| \leqslant \|x\| + \|y\| \ . \quad \blacksquare$$

Corolario. Dados los dos conjuntos de números complejos cualesquiera $\{\xi^1, \ldots, \xi^n\}$ y $\{\eta^1, \ldots, \eta^n\}$ se cumple la relación

$$\sqrt{\sum_{i=1}^{n} |\xi^i|^2} - \sqrt{\sum_{i=1}^{n} |\eta^i|^2} \leqslant \sqrt{\sum_{i=1}^{n} |\xi^i + \eta^i|} \leqslant \sqrt{\sum_{i=1}^{n} |\xi^i|^2} + \sqrt{\sum_{i=1}^{n} |\eta^i|^2}$$

Demostración. Basta con aplicar la relación 3 del teorema anterior a los vectores $x = (\xi^1, \ldots, \xi^n)$, $y = (\eta^1, \ldots, \eta^n)$ de $\mathbb{C}^n$ teniendo en cuenta que

$$\|x\|^2 = (x \ , \ x) = \sum_{i=1}^{n} |\xi^i|^2$$

y análogamente

$$\|y\|^2 = \sum_{i=1}^{n} |\eta^i|^2 \quad \text{y} \quad \|x + y\|^2 = \sum_{i=1}^{n} |\xi^i + \eta^i|^2 \quad \blacksquare$$

Aplicación. *El espacio vectorial de producto interno de las sucesiones de números de cuadrados sumables.*

Consideramos aquí a todas las sucesiones $x = \{\xi_n\}$ de números reales (complejos) tales que

$$\Sigma |\xi_n|^2 = \sum_{n=1}^{\infty} |\xi_n|^2 < \infty$$

es decir tales que la serie cuyo término general es $|\xi_n|^2$ es convergente. A tales sucesiones las llamaremos abreviadamente *sucesiones de cuadrados sumables.*

Las sucesiones de cuadrados sumables constituyen un espacio vectorial sobre $\mathbb{R}(\mathbb{C})$ si se adpotan las definiciones siguientes

$$\{\xi_n\} + \{\eta_n\} = \{\xi_n + \eta_n\}$$

$$\alpha\{\xi_n\} = \{\alpha\xi_n\} \quad , \qquad \alpha \in \mathbb{R}(\mathbb{C}).$$

Estas definiciones requieren justificación: debe comprobarse que si

$$a \ A = \Sigma |\xi_n|^2 < \infty \quad \text{y} \quad B = \Sigma |\eta_n|^2 < \infty \ ,$$

entonces también es

$$\Sigma|\xi_n + \eta_n|^2 < \infty \quad \text{y} \quad \Sigma|\alpha\,\xi_n|^2 < \infty.$$

Se tiene en efecto, en virtud del corolario del teorema 17.2.2,

$$\sqrt{\sum_{i=1}^{n}|\xi_i + \eta_i|^2} \leqslant \sqrt{\sum_{i=1}^{n}|\xi_i|^2} + \sqrt{\sum_{i=1}^{n}|\eta_i|^2} \leqslant \sqrt{A} + \sqrt{B}\ ,$$

de donde

$$\sum_{i=1}^{n}|\xi_i + \eta_i|^2 \leqslant (\sqrt{A} + \sqrt{B})^2\ .$$

Esta relación, que se cumple para todo n prueba que $\{\xi_n + \eta_n\}$ es una sucesión de cuadrados sumables. Además,

$$\sum_{i=1}^{n}|\alpha\,\xi_i|^2 = |\alpha|^2 \sum_{i=1}^{n}|\xi_i|^2 \leqslant |\alpha|^2 A$$

y por tanto la sucesión $\{\alpha\,\xi_n\}$ es de cuadrados sumables.

Probaremos por último que en el espacio vectorial que acabamos de construir se puede definir el producto interno mediante la fórmula

$$(x\ ,\ y) = \sum_{n=1}^{\infty}\xi\overline{\eta}_n = \Sigma\xi_n\ \overline{\eta}_n$$

Para ello bastará probar que si $A = \Sigma|\xi_n|^2 < \infty$ y $B = \Sigma|\eta_n|^2 < \infty$, entonces es $\Sigma\xi_n\ \overline{\eta}_n < \infty$. En efecto, la desigualdad de Cauchy (17.2.6) nos permite escribir

$$|\sum_{i=1}^{n}\xi_i\ \overline{\eta}_i| \leqslant \sqrt{\sum_{i=1}^{n}|\xi_i|^2}\ \sqrt{\sum_{i=1}^{n}|\eta_i|^2} \leqslant \sqrt{A}\ \sqrt{B}\ ,$$

y como esta relación se cumple para todo n, la serie $\Sigma\xi_n\ \overline{\eta}_n$ será convergente.

La demostración de que el producto interno que acabamos de definir cumple las propiedades de la sección 17.1 se deja al lector.

El espacio de producto interno constituido por las sucesiones de cuadrados sumables, de números complejos, es un espacio de Hilbert. [1)]

17.3. Bases ortonormales de los espacios de producto interno.

Definición 17.3.1. Un conjunto de vectores de un espacio de producto interno se llama *conjunto ortonormal* si todos sus elementos son vectores unitarios y para cada par de ellos x , y $(x \neq y)$ se cumple la condición

$$(x \ , \ y) = 0,$$

donde $(x \ , \ y)$ designa al producto interno de x e y.

Las condiciones para que el conjunto $\{x_1 , \ldots , x_n\}$ sea ortonormal puede escribirse en la forma

$$(x_i \ , \ x_j) = \delta_{ij}$$

donde δ_{ij} es el simbolo de Kronecker.

Definición 17.3.2. Una base de un espacio de producto interno se llama *base ortonormal* si el conjunto de sus vectores es ortonormal.

Ejemplo. En $\mathbb{R}^2$, provisto del producto interno ordinario, los vectores $(\cos\theta \ , \ \operatorname{sen}\theta)$ y $(-\operatorname{sen}\theta \ , \ \cos\theta)$ constituyen una base ortonormal, cualesquiera que sea θ.

Teorema 17.3.1. Cada conjunto ortonormal de un espacio de producto interno es linealmente independiente.

Demostración. Supongamos que $x_1 , \ldots , x_k$ son vectores del conjunto dado, distintos entre sí, tales que

$$\sum_{i=1}^{k} \alpha^i x_i = 0 \quad , \qquad (\alpha^i \in \mathbb{F}).$$

Si se lleva a cabo el producto interno de ambos miembros por el vector x_j, y tenemos en cuenta la definición de un conjunto ortonormal, se obtiene

$$\alpha^j = 0 \quad , \qquad (j = 1 , \ldots , k) .$$

1) Ver por ejemplo, Horáth, Topological Vector Spaces and Distributions I , Addison-Wesley (1966), pág. 16, Ex. 11.

Por tanto los vectores $x_1, \ldots, x_k$ son *l.i.* ■

Teorema 17.3.2. Sea V un espacio de producto interno de dimensión n. Dada una base cualquiera $\{x_i\}$ de V, existe una base ortonormal $\{y_1, \ldots, y_n\}$.

Demostración. Consideremos en primer lugar el vector unitario

$$y_1 = \frac{x_1}{\|x_1\|}$$

El conjunto $\{y_1\}$ es ortonormal y el vector y_1 depende linealmente de x_1. Si $n = 1$, ese conjunto constituye la base ortonormal que cumple la condición del teorema. Si $n > 1$ podemos construir la base ortonormal por inducción a partir de $\{y_1\}$. Supongamos que se ha hallado un conjunto ortonormal $\{y_1, \ldots, y_r\}$ donde $r < n$, tal que cada y_i es una combinación lineal de $x_1, \ldots, x_r$. Probaremos que puede encontrarse un vector y_{r+1} tal que

$$\{y_1, \ldots, y_{r+1}\}$$

es ortonormal, y por tanto *l.i.* y que y_{r+1} es una combinación lineal de $x_1, \ldots, x_{r+1}$. En efecto, definamos el vector

$$z_{r+1} = x_{r+1} - \sum_{i=1}^{r} \alpha_{r+1}^{i} \, y_i$$

donde los escalares α_{r+1}^{i}, aun no definidos, serán determinados a continuación. Debido a la ortonormalidad de los vectores $y_1, \ldots, y_r$ tendremos

$$(z_{r+1}, y_j) = (x_{r+1}, y_j) - \sum_{i=1}^{n} \alpha_{r+1}^{i} (y_i, y_j)$$

$$= (x_{r+1}, y_j) - \alpha_{r+1}^{j} \quad \text{para } j = 1, \ldots, r.$$

Para que z_{r+1} cumpla las condiciones

$$(z_{r+1}, y_j) = 0 \text{ , para } j = 1, \ldots, r,$$

basta definir α_{r+1}^{j} por las ecuaciones

$$\alpha^{j}_{r+1} = (x_{r+1} \cdot y_j) \qquad j = 1, \ldots, r.$$

Esto supuesto, se tiene

$$z_{r+1} = x_{r+1} - \sum_{i=1}^{n} (x_{r+1} \cdot y_i)\, y_i .$$

z_{r+1} es entonces diferente de cero, pues en caso contrario x_{r+1} sería *l.d.* de $y_1, \ldots, y_r$, lo cual no es posible porque, dependiendo estos últimos vectores de $x_1, \ldots, x_r$, lo mismo ocurriría con x_{r+1}, lo cual sería una contradicción por cuanto $(x_1, \ldots, x_n)$ es una base. Además z_{r+1} cumple las condiciones $(z_{r+1} \cdot y_j) = 0$ para $j = 1, 2, \ldots, r$.

El vector

$$y_{r+1} = \frac{z_{r+1}}{\|z_{r+1}\|}$$

es entonces un vector unitario tal que $(y_{r+1} \cdot y_j) = 0$. Los vectores $\{y_1, \ldots, y_{r+1}\}$ forman por tanto un conjunto ortonormal; y puesto que y_{r+1} es combinación lineal de los vectores $y_1, \ldots, y_r$, x_{r+1} y que cada y_i, $1 \leq i \leq r$, es combinación lineal de $x_1, \ldots, x_r$, se ve que y_{r+1} es combinación lineal de $x_1, \ldots, x_{r+1}$. En resumen $\{y_1, \ldots, y_{r+1}\}$ es un conjunto ortonormal tal que y_i depende linealmente de $x_1, \ldots, x_i$. De esta manera se obtendrá en un número finito de pasos una base $\{y_1, \ldots, y_n\}$ que cumple las condiciones del teorema. ■

El proceso descrito en la demostración del teorema, mediante el cual es posible, a partir de una base cualquiera de un espacio de producto interno de dimensión finita, obtener una base ortonormal, se conoce con el nombre de *proceso de ortogonalización de Gram-Schmidt*. Puede escribirse como sigue: Se calcula

$$y_1 = \frac{x_1}{\|x_1\|} \qquad (17.3.1)$$

Una vez obtenidos los vectores $y_1, \ldots, y_r$, donde r toma sucesivamente los valores $1, 2, \ldots, n-1$, se calcula y_{r+1} mediante las fórmulas siguientes:

$$\alpha^j_{r+1} = (x_{r+1}, y_j) \quad , \quad j = 1, \ldots, r, \qquad (17.3.2)$$

$$z_{r+1} = x_{r+1} - \sum_{i=1}^{r} \alpha^i_{r+1}\, y_i \qquad (17.3.3)$$

$$y_{r+1} = \frac{z_{r+1}}{\|z_{r+1}\|} \qquad (17.3.4)$$

Corolario. En cada espacio de producto interno de dimensión finita existe una base ortonormal.

Ejemplo. A partir de la base

$$x_1 = (1, 0, 1) \quad , \quad x_2 = (1, 2, 0) \quad , \quad x_3 = (0, 2, 3)$$

del espacio euclideano $\mathbb{R}^3$, con el producto interno ordinario, vamos a construir una base ortonormal, mediante el proceso de ortogonalización de Gram-Schmidt.

Se tiene

$$\|x_1\| = \sqrt{1 + 1} = \sqrt{2},$$

y por tanto, la fórmula (17.3.1) nos da

$$y_1 = \left(\frac{1}{\sqrt{2}} \; , \; 0 \; , \; \frac{1}{\sqrt{2}}\right) .$$

Las fórmulas (17.3.2) se reducen ahora a

$$\alpha^1_1 = (x_2, y_1) = \frac{1}{\sqrt{2}}$$

Por consiguiente de (17.3.3) se obtiene

$$z_2 = x_2 - \frac{1}{\sqrt{2}} = \left(\frac{1}{2} \; , \; 2 \; , \; -\frac{1}{2}\right) ,$$

y

$$\|z_2\| = \sqrt{\frac{1}{4} + 4 + \frac{1}{4}} = \frac{3}{2}\sqrt{2} .$$

Luego

$$y_2 = \frac{z_2}{\|z_2\|} = (\frac{\sqrt{2}}{6} , \frac{2\sqrt{2}}{3} , -\frac{\sqrt{2}}{6}) ,$$

$$\alpha_2^1 = (x_3 , y_1) = \frac{3\sqrt{2}}{2} , \qquad \alpha_2^2 = (x_3 , y_2) = \frac{5\sqrt{2}}{6}$$

$$z_3 = x_3 - \frac{3\sqrt{2}}{2} y_1 - \frac{5\sqrt{2}}{6} y_2 = (-\frac{16}{9} , \frac{8}{9} , \frac{16}{9}) ,$$

$$\|z_3\| = \frac{8}{3} , \qquad y_3 = \frac{z_3}{\|z_3\|} = (-\frac{2}{3} , \frac{1}{3} , \frac{2}{3}) .$$

La base ortonormal buscada es

$$\{(\frac{\sqrt{2}}{2} , 0 , \frac{\sqrt{2}}{2}) , (\frac{\sqrt{2}}{6} , \frac{2\sqrt{2}}{3} , -\frac{\sqrt{2}}{6}) , (-\frac{2}{3} , \frac{1}{3} , \frac{2}{3})\} .$$

Teorema 17.3.3. Si $\{e_1 , \ldots , e_n\}$ es una base ortonormal de un espacio V de producto interno, las componentes $\xi^1 , \ldots , \xi^n$ de un vector cualquiera x respecto de esa base son dadas por

$$\xi^i = (x , e_i) , \quad i = 1 , \ldots , n , \qquad (17.3.5)$$

y por tanto

$$x = \sum_{i=1}^{n} (x , e_i) e_i . \qquad (17.3.6)$$

Además, el producto interno de dos vectores

$$x = \sum_{i=1}^{n} \xi^i e_i \qquad y = \sum_{i=1}^{n} \eta^i e_i$$

es dado por la fórmula

$$(x , y) = \sum_{i=1}^{n} \xi^i \overline{\eta^i} . \qquad (17.3.7)$$

que en el caso en que V es un espacio real se reduce a

$$(x , y) = \sum_{i=1}^{n} \xi^i \eta^i . \qquad (17.3.8)$$

Demostración. Se tiene

$$(x \ , \ e_j) = \sum_{i=0}^{n} \xi^i (e_i \ , \ e_j) = \xi^j \quad ,$$

lo cual prueba las fórmulas (17.3.5) y por tanto la fórmula (17.3.6).

Además

$$(x \ , \ y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \xi^i \overline{\eta^j} \ (e_i \ , \ e_j) = \sum_{i=1}^{n} \xi^i \overline{\eta^i}$$

En el caso real debe sustituirse $\overline{\eta^i}$ por η^i en esta fórmula. ■

Si se tiene en cuenta la fórmula (17.3.5), las ecuaciones (17.3.7) y (17.3.8) pueden escribirse en las formas

$$(x \ , \ y) = \sum_{i=1}^{n} (x \ , \ e_i)\overline{(y \ , \ e_i)} \ , \text{ en el caso complejo}$$

$$(x \ , \ y) = \sum_{i=1}^{n} (x \ , \ e_i)(y \ , \ e_i) \ , \text{ en el caso real.}$$

Estas igualdades, en que $\{e_i\}$ es una base ortonormal y x e y son vectores cualesquiera se llaman *identidades de Parseval.*

Ejercicios.

17.3.1. A partir de la base $((1 \ , \ 0 \ , \ 1) \ , (2, \ 1 \ , \ -3) \ , (-1 \ , \ 1 \ , \ 0))$ del espacio $\mathbb{R}^3$, construir una base ortonormal mediante el proceso de ortogonalización de Gram-Schmidt.

17.3.2. Demostrar que un espacio V real de dimensión n es dual consigo mismo sis se define el producto escalar en la forma

$$<x \ , \ y> = (x \ , \ y) \quad \text{para} \quad x \ , y \in V$$

Probar luego que las bases ortonormales de V son precisamente las bases que son duales consigo mismas.

17.3.3. Determinar la condición para que los vectores

$$(a + bi\ ,\ c + di) \quad \text{y} \quad (-c + di\ ,\ a - bi)$$

constituyan una base ortonormal de $\mathbb{C}^2$.

17.4. Cambio de bases ortonormales en un espacio de producto interno. Matrices unitarias y ortogonales.

Teorema 17.4.1. Dadas dos bases ortonormales cualesquiera $\{e_i\}$ y $\{e_{i'}\}$ de un espacio V de producto interno, sean

$$e_i = \sum_{j'=1'}^{n'} A_i^{j'}\, e_{j'} \quad , \quad e_{j'} = \sum_{i=1}^{n} A_{j'}^{i}\, e_i \qquad (17.4.1)$$

las correspondientes ecuaciones de cambio de base. Se cumplen entonces las relaciones

$$\overline{A_{j'}^{k}} = A_k^{j'} \qquad (k\ ,\ j = 1\ ,\ \ldots\ ,\ n) \qquad (17.4.2)$$

Demostración. Según hemos probado en la sección 6.10, las matrices

$$A = (A_i^{j'}) \qquad \text{y} \qquad A' = (A_{j'}^{i})$$

son inversas una de la otra.

Por hipótesis se tiene

$$\Big(\sum_{k=1}^{n} A_{i'}^{k}\, e_k\ ,\ \sum_{l=1}^{n} A_{j'}^{l}\, e_l\Big) = (e_{i'}\ ,\ e_{j'}) = \delta_{i'j'} \quad .$$

Pero, por las propiedades del producto interno, resulta

$$\Big(\sum_{k=1}^{n} A_{i'}^{k} e_k\ ,\ \sum_{l=1}^{n} A_{j'}^{l} e_l\Big) = \sum_{k=1}^{n} \sum_{l=1}^{n} A_{i'}^{k} \overline{A_{j'}^{l}}\, (e_k\ ,\ e_l) = \sum_{k=1}^{n} \sum_{l=1}^{n} A_{i'}^{k} \overline{A_{j'}^{l}}\, \delta_{kl}$$

$$= \sum_{k=1}^{n} A_{i'}^{k}\, \overline{A_{j'}^{k}}$$

Por consiguiente: $$\sum_{k=1}^{n} A_{i'}^{k}\, A_{j'}^{k} = \delta_{i'j'} \qquad (17.4.2)$$

o sea: $(A')^t\, \overline{A'} = 1$,

y por tanto $\overline{(A')^t}\, A' = 1$.

Por otra parte es

$$AA' = I.$$

Se deduce de las dos últimas relaciones que

$$A = \overline{(A')^t}$$

o

$$A' = \overline{A^t}$$

Por consiguiente

$$\overline{A^k_{j'}} = A^{j'}_k \qquad (k = 1, \ldots, n\,;\, j' = 1, \ldots, n)\,. \quad \blacksquare$$

Si designamos por $\overline{M}$ a la matriz cuyos términos son los conjugados de los correspondientes términos de la matriz M de términos complejos, $\overline{M}$ se llama la *conjugada* de M. En virtud de las relaciones (17.4.4) podemos decir entonces que las matrices A y A' de cambio de bases ortonormales cumplen la condición

$$A = \overline{(A')^t} = \overline{(A^{-1})^t} \quad , \qquad (17.4.5)$$

o equivalentemente

$$A' = \overline{((A')^{-1})^t} \quad , \qquad (17.4.6)$$

y por consiguiente cada una de esas matrices es igual a la conjugada de la traspuesta de su inversa. Es fácil observar que las relaciones matriciales (17.4.5) y (17.4.6) son iguales respectivamente a

$$A^{-1} = \overline{A^t} \qquad (17.4.7)$$

y

$$(A')^{-1} = \overline{(A')^t} \qquad (17.4.8)$$

Introduciremos ahora algunas nuevas definiciones para simplificar nuestros enunciados.

Definición 17.4.1. Dada una matriz M sobre $\mathbb{F}$, es decir cuyos términos pertenecen a $\mathbb{F}$, se llama *matriz adjunta de M* y se designa por $\widetilde{M}$, a la matriz

$$\widetilde{M} = \overline{M^t} \quad ,$$

conjugada de la traspuesta de M.

Cuando M es una matriz real, es decir cuyos términos son reales, entonces se tiene

$$\widetilde{M} = M^t \ ;$$

y por tanto la dajunta de una matriz real coincide con la traspuesta.

Evidentemente se cumple que $det\ \widetilde{M} = \overline{det\ M}$ en el caso complejo, y $det\ \widetilde{M} = det\ M$ en el caso real.

Definición 17.4.2. Las matrices complejas M tales que

$$\widetilde{M}M = M\widetilde{M} = 1. \qquad (17.4.9)$$

donde 1 representa a la matriz unidad del mismo orden n que M, se llaman *matrices unitarias.*

Las matrices reales M que gozan de la misma propiedad (17.4.9) se llaman *matrices ortogonales.* Tanto para las matrices unitarias como para las ortonormales se tiene que

$$M^{-1} = \widetilde{M} \ .$$

Podemos establecer entonces la proposición siguiente que ya está en parte demostrada.

Teorema 17.4.2. Las matrices A de cambio de bases ortonormales son aquellas matrices que cumplen la condición

$$\widetilde{A}A = A\widetilde{A} = 1. \qquad (17.4.10)$$

Son por tanto unitarias si el espacio es unitario, y son ortogonales si el espacio es euclideano.

Demostración. Las relaciones (17.4.7) y (17.4.8) que hemos demostrado anteriormente, muestran que las matrices A y A' de cambio de bases ortonormales cumplen la relación (17.4.10).

Supongamos inversamente que $A = (A^{i'}_{i})$ es una matriz que satisface a la relación (17.4.10). Designemos a su inversa por

$$A' = (A^{i}_{j'})$$

De (17.4.10) se sigue que $A^{-1} = A' = \widetilde{A} = \overline{A^t}$. Por tanto se cumplirán las relaciones

$$\overline{A^{i}_{j'}} = A^{j'}_{i} \qquad (17.4.11)$$

Si $\{e_i\}$ es una base ortonormal de V, entonces, puesto que la matriz A' es regular, las fórmulas

$$e_{j'} = \sum_{i=1}^{n} A^{i}_{j'} e_i$$

definen a los vectores de una base, y se tiene además

$$e_i = \sum_{i=1}^{n} A^{j'}_{i} e_{j'} \quad .$$

Probaremos que la base $\{e_{j'}\}$ es también una base ortonotmal y por tanto A y A' son las matrices de cambio de las bases ortonormales $\{e_i\}$ y $\{e_{j'}\}$.

En efecto,

$$(e_{i'} , e_{j'}) = \Big(\sum_{k=1}^{n} A^{k}_{i'} e_k , \sum_{i=1}^{n} A^{i}_{j'} e_i\Big)$$

$$= \sum_{k=1}^{n} \sum_{i=1}^{n} A^{k}_{i'} \overline{A^{i}_{j'}} (e_k , e_i) = \sum_{k=1}^{n} A^{k}_{i'} \overline{A^{k}_{j'}} \quad .$$

En virtud de (17.4.11) tendremos

$$(e_{i'} , e_{j'}) = \sum_{k=1}^{n} A^{j'}_{k} A^{k}_{i'} = \delta^{j'}_{i'} \quad ,$$

y por tanto $\{e_{j'}\}$ es también una base ortonormal de V. ■

Las siguientes proposiciones dan condiciones necesarias y suficientes para que una matriz compleja sea unitaria y para que una matriz real sea ortogonal.

Teorema 17.4.3. Para que una matriz M de orden n cumpla la condición

$$\widetilde{M}M = M\widetilde{M} = I$$

es necesario y suficiente que se cumpla una de las dos siguientes condiciones

$$\widetilde{M}M = 1 \qquad (17.4.12)$$

$$M\widetilde{M} = 1 \qquad (17.4.13)$$

Demostración. Supongamos que sea $\widetilde{M}M = 1$. Se tendrá entonces

$$det\ \widetilde{M}\ .\ det\ M = 1 \quad \text{y por tanto} \quad det\ \widetilde{M} \neq 0 \quad \text{y} \quad det\ M \neq 0.$$

Tanto M como $\widetilde{M}$ son no singulares y por consiguiente tienen inversas M^{-1} y $\widetilde{M}^{-1}$. De $\widetilde{M}M = 1$ se deduce $M = \widetilde{M}^{-1}$ y por tanto

$$M\widetilde{M} = 1 \ .$$

Análogamente se prueba que (17.4.13) implica a (17.4.12). ■

Teorema 14.4.4. Para que la matriz compleja (real) $M = (\alpha^i_j)$ sea unitaria (resp. ortogonal) es necesario y suficiente que se cumpla una de las dos condiciones siguientes:

$$\left.\begin{array}{l} \sum_{i=1}^{n} \alpha^i_j\ \overline{\alpha^i_k} = \delta_{jk} \quad , \text{ para } j,k = 1, \ldots, n \\ \\ \sum_{i=1}^{n} \alpha^j_i\ \overline{\alpha^k_i} = \delta^{jk} \quad , \text{ para } j,k = 1, \ldots, n \end{array}\right\} \qquad (17.4.14)$$

respectivamente, una de las dos condiciones siguientes

$$\left.\begin{array}{l} \sum_{i=1}^{n} \alpha^i_j\ \alpha^i_k = \delta_{jk} \quad , \text{ para } j,k = 1, \ldots, n \\ \\ \sum_{i=1}^{n} \alpha^j_i\ \alpha^k_i = \delta^{jk} \quad , \text{ para } j,k = 1, \ldots, n \end{array}\right\} \qquad (17.4.15)$$

Demostración. Si $\widetilde{M} = (\beta^j_i)$, entonces es $\beta^j_i = \overline{\alpha^i_j}$ ($\beta^j_i = \alpha^i_j$ en el caso real). Luego las relaciónes $\widetilde{M}M = 1$ y $M\widetilde{M} = 1$ equivalen cada una a una de las relaciones (17.4.14) (resp. (17.4.15)). Basta aplicar entonces el teorema anterior. ■

Observación. La primera ecuación (17.4.14) (resp. (17.4.15)) expresa que los vectores $(\alpha_i^1, \alpha_i^2, \ldots, \alpha_i^n)$ de $\mathbb{F}^n$, constituidos por los términos de cada columna de la matriz M, forman una base ortonormal de dicho espacio. Semejantemente la segunda ecuación (17.4.14) (resp. (17.4.15)) muestra que los vectores fila $(\alpha_1^i, \ldots, \alpha_n^i)$ forman una base ortonormal de $\mathbb{F}^n$.

Teorema 17.4.5. El valor absoluto del determinante de una matriz unitaria u ortogonal M es igual a 1.

Demostración. Puesto que $M\widetilde{M} = I$, se tiene que

$$det\ M.\ det\ \widetilde{M} = det\ M.\ \overline{det\ M} = |det\ M|^2 = 1$$

Resulta así que $|det\ M| = 1$. ∎

Observación. Si M es ortonormal, entonces es $det\ M = \pm 1$.

Ejercicio

17.4.1. Comprobar mediante la definición 17.4.2 que las siguientes matrices son unitarias u ortogonales según sea el caso:

a) $\begin{pmatrix} i & \sqrt{2} \\ \sqrt{2} & -i \end{pmatrix}$ b) $\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$ c) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}$

e) $\begin{pmatrix} \operatorname{sen}\theta\cos\varphi & \cos\theta\cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\theta\operatorname{sen}\varphi & \cos\theta\operatorname{sen}\varphi & \cos\varphi \\ \cos\theta & -\operatorname{sen}\theta & 0 \end{pmatrix}$

Comporbar, además, el cumplimiento de las relaciones del teorema 17.4.4, y calcular los valores de los determinantes de las matrices dadas.

17.5. Conjugación en los espacios unitarios.

Definición 17.5.1. Se llama *conjugación* en un espacio vectorial complejo U a una aplicación $x \mapsto \overline{x}$ del espacio U en sí mismo, que tiene las propiedades siguientes:

1. $\overline{x_1 + x_2} = \overline{x}_1 + \overline{x}_2$
2. $\overline{\alpha x} = \overline{\alpha}\,\overline{x}$ para cada $\alpha \in \mathbb{C}$
3. $\overline{(\overline{x})} = \overline{\overline{x}} = x$

Si el espacio es unitario se debe cumplir además la condición

4. $(\overline{x}\ ,\ \overline{y}) = \overline{(x\ ,\ y)}$.

La conjugación constituye una *biyección* de U pues si $\overline{x_1} = \overline{x_2}$ entonces es $\overline{\overline{x_1}} = \overline{\overline{x_2}}$, y de la propiedad 3. se sigue que $x_1 = x_2$. Por tanto es una aplicación inyectiva. Es sobreyectiva porque dado x cualquiera de U se tiene $\overline{(\overline{x})} = \overline{\overline{x}} = x$. La propiedad 2. muestra, además, que la congugación no es lineal por cuanto $\overline{\alpha}\,\overline{x}$ es en general diferente de $\alpha\,\overline{x}$.

La conjugación definida en el teorema 12.6.1. en la complejificación $V^{\mathbb{C}}$ de un espacio real V, cumple las propiedades 1.,2. y 3 de la definición 17.5.1. En el presente parágrafo deberemos considerar conjugaciones en espacios unitarios, y por consiguiente deben cumplirse las cuatro propiedades.

En un espacio unitario U de dimensión n siempre es posible definir una conjugación con las cuatro propiedades mencionadas. En efcto, dada una base ortonormal $\{e_1\ ,\ \ldots\ ,\ e_n\}$ de U, se llama *conjugación correspondiente a la base* $\{e_i\}$ a la aplicación $x \mapsto \overline{x}$, de U sobre U que aplica a cada vector $x = \sum_{i=1}^{n} \xi^i e_i$ de U en el vector $\overline{x} = \sum_{i=1}^{n} \overline{\xi^i}\, e_i$, que se llama *conjugado de* x . La conjugación así definida tiene evidentemente las tres primeras propiedades

de la definición 17.5.1. La demostración de la propiedad 4. puede llevarse a cabo como sigue:

$$(\bar{x} \ , \ \bar{y}) = (\sum_{i=1}^{n} \overline{\xi^i} e_i \ , \ \sum_{j=1}^{n} \overline{\eta^j} e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{\xi^i \eta^j} (e_i \ , \ e_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{\xi^i} \ \eta^j \ \delta_{ij} = \overline{\sum_{i=1}^{n} \sum_{j=1}^{n} \xi^i \ \overline{\eta^j} \ \delta_{ij}}$$

$$= \overline{\sum_{i=1}^{n} \sum_{j=1}^{n} \xi^i \ \overline{\eta^j} \ (e_i \ , \ e_j)} = \overline{(\sum_{i=1}^{n} \xi^i e_i \ , \ \sum_{j=1}^{n} \eta^j e_j)}$$

$$= \overline{(x \ , \ y)} .$$

Observación. La conjugación aplicada a un espacio euclideano se reduce a la aplicación idéntica, cualquiera que sea la base ortonormal que se considere. Algunas veces puede ser útil hacer uso de esta observación en un razonamiento que puede ser válido tanto para espacios unitarios como para espacios euclideanos, cuando se tiene presente que para estos últimos es $\bar{\alpha} = \alpha$, si α es escalar, y $\bar{x} = x$ si x es un vector.

Ejercicio

17.5.1. Calcular los conjugados del vector $(i \ , \ 1)$ de $\mathbb{C}^2$ para la conjugación relativa a la base canónica y para aquella relativa a la base normal constituida por los vectores

$$(\tfrac{1}{2} + \tfrac{1}{2}i \ , \ \tfrac{1}{2} + \tfrac{1}{2}i) \ , \ (-\tfrac{1}{2} + \tfrac{1}{2}i \ , \ \tfrac{1}{2} - \tfrac{1}{2}i) \ .$$

17.6. Dualidad de un espacio de producto interno consigo mismo.

Un espacio *euclideano* V es dual consigo mismo, si se define el producto escalar por

$$\langle x \ , \ y \rangle = (x \ , \ y) \ , \quad x,y \in V \qquad (17.6.1)$$

Un espacio *unitario* U, en el que suponemos definida una conjugación $x \mapsto \bar{x}$, es dual consigo mismo, si se define el produc-

to escalar por

$$\langle x \ , \ y \rangle = (x \ , \ \overline{y}) \quad , \qquad x \ , y \ \epsilon \ U. \qquad (17.6.2)$$

En este capítulo no vamos a hacer uso del concepto de ortogonalidad que resultaría en base a las nociones que fueron introducidas en la sección 4.2 , cuando se considera en los espacios unitarios la dualidad a que acabamos de referirnos. Las razones de esta decisión se expondrán en la sección siguiente.

17.7. Ortogonalidad en los espacios de producto interno.

Definición 17.7.1. Se dice que dos vectores x e y de un espacio de producto interno son *ortogonales*, si se cumple la condición

$$(x \ , \ y) = 0.$$

En virtud de esta definición puede decirse que cada dos vectores diferentes de un conjunto ortonormal (en particular de una base ortonormal) son ortogonales.

Debe observarse que la ortogonalidad que acabamos de definir coincide, en el caso de un espacio euclideano, con la que resultaría de la dualidad del espacio consigo mismo dada por (17.6.1); pero no coincide, en el caso de un espacio unitario, con la que resultaría de la dualidad del espacio consigo mismo dada por (17.6.2). En efecto, en este último caso las relaciones

$$(x \ , \ y) = 0 \quad \text{y} \quad \langle x \ , \ y \rangle = (x \ , \ \overline{y}) = 0$$

no son equivalentes. Además la definición (17.6.2) depende de la base ortonormal que se haya elegido en el espacio. Es por estas razones que *la ortogonalidad a que nos vamos a referir en todo el resto de este capítulo es exclusivamente la dada en la definición 17.7.1.*

Definición 17.7.2. Dado un subespacio V_1 del espacio V de producto interno, se llama *complemento ortogonal de* V_1 y se designa por $V_1^{\perp}$ al subespacio de V formado por los vectores que son ortogonales a todos los elementos de V_1 .

Teorema 17.7.1. Dado un subespacio V_1 del espacio V de producto interno y de dimensión n, se cumple la relación (definición 13.3.2)

$$V = V_1 \oplus V_1^{\perp}$$

y por tanto es

$$dim\ V = dim\ V_1 + dim\ V_1^{\perp}$$

Demostración. Según sabemos, V_1 puede considerarse como un espacio de producto interno. Tiene por tanto una base ortonormal

$$\{e_1 , \ldots , e_k\} ,$$

la cual puede ser ampliada hasta obtener una base

$$\{e_1 , \ldots , e_k , e'_{k+1} , \ldots , e'_n\}$$

de V, de la que, a su vez, puede obtenerse, por el proceso de ortogonalización de Gram-Schmidt, una base ortonormal $\{e_1 , \ldots , e_n\}$ Los vectores $e_{k+1} , \ldots , e_n$ son ortogonales a V_1 y pertenecen por tanto a $V_1^{\perp}$. Además, si $x = \sum_{i=1}^{n} \xi^i e_i$ es un vector de $V_1^{\perp}$, se cumplen las relaciones $(x , e_j) = 0$ para $j = 1 , \ldots , k$. Por consiguiente se deduce que $\xi^j = 0$ para $j = 1 , \ldots , k$, y resulta que

$$x = \sum_{i=k+1}^{n} \xi^i e_i \quad .$$

Por tanto $\{e_{k+1} , \ldots , e_n\}$ es una base ortonormal de $V_1^{\perp}$. Se deduce inmediatamente que $V = V_1 \oplus V_1^{\perp}$ ∎

17.8. El teorema de Riesz. Base recíproca.

Teorema 17.8.1. *(Teorema de Riesz).* Si V es un espacio de producto interno de dimensión finita y f es una función lineal definida en V, es decir un elemento del espacio $\mathcal{L}(V) = V^*$, entonces existe un vector único $v \in V$ tal que

$$f(x) = (x , v) \quad , \text{ para todo } x \in V. \qquad (17.8.1)$$

Demostración. Vamos a considerar aquí el producto escalar definido en $V \times V$ por (17.6.1) en el caso real, y por (17.6.2) en el caso complejo. Podemos limitarnos a la consideración de la relación (17.6.2) si tenemos presente que en el caso real es $\overline{x} = x$.

La aplicación $f_1 : V \to \mathbb{F}$ tal que $f_1(x) = \overline{f(x)}$ es una función lineal como es fácil comprobar. Luego, del corolario del teorema 5.3.1, aplicado al espacio V, que es dual de sí mismo respecto del producto escalar a que hemos hecho referencia, resulta la existencia de un vector $v \in V$ tal que

$$\overline{f(x)} = \langle v , x \rangle \qquad \forall\, x \in V.$$

Se tiene por tanto

$$\overline{f(x)} = (v , \overline{x}) = \overline{(\overline{x} , v)} \quad , \quad \forall\, x \in V,$$

relación que equivale a

$$f(x) = (x , v) \quad , \qquad \forall\, x \in V.$$

El vector v que cumple esta última relación es único, pues si fuera $f(x) = (x , v_1)$, para otro vector v_1, se tendría $(x , v) = (x , v_1)$ para todo x, y sería por tanto $v = v_1$. v no depende por consiguiente de la conjugación supuesta en V. ∎

Corolario. Si V es un espacio *euclideano* de dimensión finita, la aplicación $\phi : V^* \to V$ que a cada función $f \in V^* = \mathcal{L}(V)$ le hace corresponder el vector $v \in V$ tal que $f(x) = (x , v)$ para todo $x \in V$ es un isomorfismo de V^* sobre V.

Demostración. La aplicación $\phi : V^* \to V$ del enunciado, la cual existe en virtud del teorema anterior, es justamente el isomorfismo existente entre $\mathcal{L}(V)$ y V cuando este último se considera como dual de sí mismo mediante el producto escalar definido por (17.6.2). También puede comprobarse de manera directa que ϕ es un isomorfismo. ∎

Observación. Si V es un espacio unitario, la aplicación $\phi : V^* \to V$ a que se refiere el corolario precedente, no es un isomorfismo. En

efecto, sea $f \in \mathcal{L}(V)$ y $\phi(f) = v$. Se tendrá

$$f(x) = (x \ , \ v) \ , \text{ para todo } x \in V.$$

Por tanto, cualquiera que sea $\alpha \in \mathbb{C}$, se tiene

$$(\alpha f)(x) = \alpha f(x) = \alpha(x \ , \ v) = (x \ , \ \bar{\alpha} v) \text{ para todo } x \in V.$$

Por consiguiente es

$$\phi(\alpha f) = \bar{\alpha} v = \bar{\alpha} \phi(f) \ ,$$

que, en general es diferente de $\alpha\phi(f)$.

Definición 17.8.1. Dada una base $\{e_1 , \ldots , e_n\}$ de un espacio euclideano V de dimensión finita, sean $f^1, \ldots , f^n$ las funciones pertenecientes a V^* definidas por las relaciones

$$f^j(e_i) = \delta_i^j \quad , \quad j,i = 1 , \ldots , n.$$

Es fácil ver que esas funciones forman una base de V^*. Puesto que ϕ es un isomorfismo, los vectores $\phi(f^j) = e^j$ de V, constituyen también una base de V que se llama *base recíproca* de la base $\{e_i\}$, y se cumplen las relaciones

$$(e_i \ , \ e^j) = (e_i \ , \ \phi(f^j)) = f^j(e_i) = \delta_i^j \ .$$

para $i,j = 1 , \ldots , n$.

Dada una base $\{e_i\}$ de V y la base recíproca $\{e^i\}$, las componentes de un vector de V respecto de la primera se llaman *componentes contravariantes* del vector y las componentes respecto de la segunda se llaman *componentes covariantes.*

Ejercicios

17.8.1. Sea V un espacio euclideano de dimensión finita. Dada la base $\{e_i\}$ de V y la base recíproca $\{e^i\}$, obtener las fórmulas que relacionan las componentes contravariantes x^1, $x^2 , \ldots , x^n$ de un vector x de V con las componentes covariantes $x_1 , \ldots , x_n$ del mismo vector. (Se usarán los sím-

bolos $(e_i \ , \ e_j)$ y $(e^i \ , \ e^j)$, respectivamente).

17.8.2. Definidos los escalares g_{ij} y g^{ij} como se indicó en el ejercicio anterior, demostrar la fórmula siguiente

$$\sum_{j=1}^{n} g_{ij} \, g^{jk} = \delta_i^k \quad .$$

17.8.3. Calcular las expresiones de los g_{ij} en términos de los g^{ij}, e inversamente. (Observar que las matrices (g_{ij}) y (g^{ij}) son inversas una de otra, según el ejercicio anterior.

17.8.4. Expresar en términos de loss vectores de la base $\{e_i\}$ a los vectores de la base recíproca $\{e^i\}$ e inversamente.

17.9. Distancia de un vector a un subespacio.

En esta sección y en las dos siguientes consideraremos un espacio V euclideano de *dimensión infinita*. Los resultados son válidos cuando V es de dimensión finita, con algunas simplificaciones.

Sea V_1 un subespacio de V de dimensión finita r.

Teorema 17.9.1. Cada vector $x \in V$ puede expresarse de una y solo de una manera en la forma

$$x = p + d \quad , \quad p \in V_1 \quad , \quad d \in V_1^{\perp} \qquad (17.9.1)$$

Demostración. Sea $\{e_1 \ , \ \ldots \ , \ e_r\}$ una base ortonormal de V_1. Supongamos que existen los vectores $p \in V_1$ y $d \in V_1^{\perp}$ que satisfacen a la relación (17.9.1). Si escribimos

$$p = \alpha_1 e_1 + \ \ldots \ + \alpha_r e_r \quad ,$$

la relación (17.9.1) implica que

$$(x \ , \ e_i) = (p \ , \ e_i) = \alpha_i$$

Es decir que p es único y sus componentes respecto de una base ortonormal cualquiera de V_1 son idénticas a las componentes de x respecto de los vectores de esa base.

Sea ahora

$$p = \sum_{i=1}^{r} \alpha_i e_i \quad , \qquad \alpha_i = (x \ , \ e_i) \ .$$

Si $x \in V_1$ se tiene

$$x = \sum_{i=1}^{r} \alpha_i e_i = p \ ,$$

y por tanto $x - p = 0 \in V_1^{\perp}$, lo cual prueba que se cumple (17.9 1). Si $x \notin V_1$, $\{e_1 , \ldots , e_r , x\}$ es una base del espacio generado por V_1 y x. Puede entonces construirse la base ortonormal $\{e_1 , e_2, e_3, \ldots , e_r , e_{r+1}\}$ de ese espacio, haciendo uso del proceso de Gram-Schmidt, y se tendrá

$$x = \alpha_1 e_1 + \ldots + \alpha_r e_r + \alpha_{r+1} e_{r+1} \quad .$$

Por tanto se tiene

$$d = x - p = \alpha_{r+1} e_{r+1} \in V_1^{\perp} . \quad \blacksquare$$

El vector $d \in V_1$ definido por (17.9.1) se llama *componente de x perpendicular a V_1*, y el vector p se llama *proyección de x sobre el espacio V_1* .

Teorema 17.9.2. La componente de un vector x de V, perpendicular a V_1, es el vector de menor longitud entre todos los vectores z de V tales que $x - z \in V_1$.

Demostración. Sea $z \neq d$ un vector cualquiera de V tal que $x - z \in V_1$. Puede escribirse

$$\|z\|^2 = (z , z) = (d + (z - d) , d + (z - d))$$

$$= \|d\|^2 + 2(d , z - d) + \|z - d\|^2$$

Pero

$$z - d = (z - x) + (x - d) \in V_1 \ ;$$

luego $(d , z - d) = 0$. Además es $\|z - d\| > 0$. Luego

$$\|z\| > \|d\| \ . \quad \blacksquare$$

Definición. 17.9.1. Se llama *distancia* de un vector x al subespacio V_1 a la longitud $\|d\|$ de la componente de x perpendicular a V_1 .

17.10. Coeficientes de Fourier.

A continuación vamos a introducir una definición de base de un espacio euclideano de dimensión infinita, que difiere de la noción general de base que dimos en la sección 2.5. En lo que sigue, cuando se trata de un espacio euclideano de dimensión infinita emplearemos esta noción.

Definición 17.10.1. Se llama *base* del espacio V euclideano de dimensión infinita a un conjunto *ortogonal* de vectores $\{e_1, e_2, \ldots\}$ tal que cada vector $x \in V$ puede expresarse de una sola manera en la forma

$$x = \sum_{k=1}^{\infty} a_k \, e_k \quad , \qquad (17.10.1)$$

relación que es por definición equivalente a

$$\lim_{n\to\infty} \|x - \sum_{k=1}^{n} a_k \, e_k\| = 0 \, . \qquad (17.10.2)$$

Los coeficientes a_k se llaman *coeficientes generalizados de Fourier* de x con respecto a la base dada; y la serie se llama *desarrollo generalizado en serie de Fourier de x.*

Teorema 17.10.3. Los coeficientes generalizados de Fourier del vector x, son dados por las fórmulas

$$a_k = (x \, , \, e_k)$$

Demostración. Sea W_1 el subespacio de V generado por los vectores $\{e_1, \ldots, e_r\}$ donde r es arbitrario. La proyección de x sobre W_1 es

$$p = \sum_{k=1}^{r} (x \, , \, e_k) e_k \quad ,$$

y la componente de x perpendicular a V_1 es

$$d = x - \sum_{k=1}^{r} (x , e_k) e_k .$$

Puesto que $\sum_{k=1}^{r} a_k \cdot e_k$ es un vector de W_1, el teorema 17.9.2. implica que

$$\|x - \sum_{k=1}^{r} (x , e_k) e_k\| \leqslant \|x - \sum_{k=1}^{r} a_k e_k\|$$

Esta relación se cumple cualquiera que sea r. La relación (17.10.2) implica por consiguiente que

$$\lim_{r\to\infty} \sum_{k=1}^{r} (x , e_k) e_k = x ,$$

y por tanto

$$x = \sum_{k=1}^{\infty} (x , e_k) e_k .$$

Se deduce de aquí y de (17.10.1) que

$$\sum_{k=1}^{\infty} (a_k - (x , e_k)) e_k = 0$$

de donde

$$\lim_{r\to\infty} \| \sum_{k=1}^{r} (a_k - (x , e_k)) e_k \| ,$$

o sea

$$\lim (\sum_{k=1}^{r} [a_k - (x , e_k)])^{1/2} = 0 .$$

Pero todos los términos de la suma son no negativos. Luego debe ser

$$a_k = (x , e_k) . \quad \blacksquare$$

17.11. Desigualdad de Bessel e igualdad de Parseval.

Teorema 17.11.1. Sea $\{e_1 , e_2 , \ldots\}$ un conjunto ortonormal de

vectores de un espacio euclideano V de dimensión infinita, y sea x un vector cualquiera de V. Tiene lugar entonces la relación

$$\sum_{k=1}^{\infty} (x \ . \ e_k)^2 \leqslant \|x\|^2 \qquad \textit{(desigualdad de bessel)} \qquad (17.11.1)$$

Además, si $\{e_1 , e_2 , \ldots\}$ es una base ortonormal de V, y solo en este caso, es

$$\sum_{k=1}^{\infty} (x \ , \ e_k)^2 = \|x\|^2 \qquad \textit{(igualdad de Parseval)} \qquad (17.11.2)$$

Demostración. Sea x un vector cualquiera de V y n un entero arbitrario. Se tiene entonces

$$\|x - \sum_{k=1}^{n} (x \ , \ e_k)e_k\| = (x - \sum_{k=1}^{n} (x \ , \ e_k)e_k)(x - \sum_{k=1}^{n} (x \ , \ e_k)e_k)$$

$$= (x \ , \ x) - 2 \sum_{k=1}^{n} (x \ , \ e_k)(x \ , \ e_k) +$$

$$+ (\sum_{k=1}^{n} (x \ , \ e_k)e_k \ , \ \sum_{k=1}^{n} (x \ , \ e_k)e_k)$$

$$= (x \ , \ x) - 2 \sum_{k=1}^{n} (x \ . \ e_k)^2 + \sum_{k=1}^{n} (x \ , \ e_k)^2$$

$$= \|x\|^2 - \sum_{k=1}^{n} (x \ , \ e_k)^2 \ , \qquad (17.11.3)$$

y por consiguiente

$$\sum_{k=1}^{n} (x \ , \ e_k)^2 \leqslant \|x\|^2 \ .$$

Puesto que esta relación se cumple para todo valor de n, se cumplirá también la relación (17.11.1).

Supongamos ahora que $\{e_1 , e_2 , \ldots\}$ es una base ortonormal de V. En virtud de la definición 17.10.1 y del teorema 17.10.3 se tiene

$$x = \sum_{k=1}^{\infty} (x \ , \ e_k)e_k \ ,$$

y por tanto,

$$\lim_{n \to \infty} \| x - \sum_{k=1}^{n} (x , e_k) e_k \| = 0 .$$

Se sigue entonces de (17.11.3) que

$$\lim_{n \to \infty} [\| x \|^2 - \sum_{k=1}^{n} (x , e_k)^2] = 0$$

de donde,

$$\sum_{k=1}^{\infty} (x , e_k)^2 = \| x \|^2 ,$$

que es la igualdad (17.11.2) de Parseval. Este razonamiento puede llevarse a cabo en orden inverso para demostrar que si se cumple la igualdad de Parseval para todo vector x, el conjunto ortonormal $\{e_1 , e_2 , \dots \}$ es una base de V (ver el ejercicio 17.11.1)

La igualdad de Parseval puede considerarse como la generalización del teorema de Pitágoras para un espacio de dimensión infinita que posee una base ortonormal. Según el teorema anterior permite determinar si un conjunto ortonormal $\{e_1 , e_2 , \dots \}$ es una base del espacio euclideano V.

La desigualdad de Bessel admite también una interpretación geométrica, a saber que la suma de los cuadrados de las componentes de un vector en una infinidad de direcciones mutuamente ortogonales es menor que la longitud del mismo vector.

Ejercicios

17.11.1. Un conjunto ortonormal $\{e_1 , e_2 , \dots \}$ es una base de un espacio euclideano V si y solamente si cada vector x de V puede expresarse en la forma

$$x = \sum_{k=1}^{\infty} (x , e_k) e_k$$

17.11.2. Si $\{e_1 , e_2 , \dots \}$ es un conjunto ortonormal de vectores de un espacio euclideano de dimensión infinita V, enton

ces es

$$\lim_{k\to\infty} (x, e_k) = 0$$

para cada x de V. Es decir que los coeficientes de Fourier de x tienden a cero cuando $k \to \infty$, para cualquier conjunto ortonormal de vectores de V.

§ 18. Endomorfismos normales y autoadjuntos e isometrias de los espacios de producto interno de dimensión finita.

En este parágrafo vamos a considerar espacios U y V de producto interno, *de dimensión finita,* que pueden ser unitarios o euclideanos. Nuestras consideraciones, salvo indicación expresa, serán válidas para ambos, y solo tienen modificaciones obvias en el caso de los espacios euclideanos, es decir en el llamado caso real.

Esas modificaciones provienen del hecho de que en este caso se cumple que $\overline{\alpha} = \alpha$ para cada escalar α y $\overline{x} = x$ para cada vector x.

18.1. Transformaciones lineales adjuntas de espacios de producto interno.

Teorema 18.1.1. Dados los espacios U y V de producto interno, y la transformación lineal

$$T : U \longrightarrow V ,$$

existe una y solo una transformación lineal

$$\widetilde{T} : V \longrightarrow U$$

tal que

$$(Tx, y) = (x, \widetilde{T}y) \qquad \text{para } x \in U , y \in V.$$

Demostración. Dado el vector $y \in V$, la función

$$x \mapsto (Tx, y) ,$$

definida en U, es lineal. Por consiguiente, en virtud del teorema de Riesz, existe un único vector $\widetilde{T}$ y de U tal que

$$(Tx\ ,\ y) = (x\ ,\ \widetilde{T}y)$$

para cada $x \in U$. La aplicación $\widetilde{T} : V \to U$ es pues bien definida para todo y de V. Es además lineal porque, para todo x, se tiene

$$(x\ ,\ \widetilde{T}(y_1 + y_2)) = (Tx\ ,\ y_1 + y_2) = (Tx\ ,\ y_1) + (Tx\ ,\ y_2)$$

$$= (x\ ,\ \widetilde{T}y_1) + (x\ ,\ \widetilde{T}y_2) = (x\ ,\ \widetilde{T}y_1 + \widetilde{T}y_2)\ ,$$

y por tanto

$$\widetilde{T}(y_1 + y_2) = \widetilde{T}y_1 + \widetilde{T}y_2\ ;$$

y porque, para todo x, y todo $\alpha \in \mathbb{C}$,

$$(x\ ,\ \widetilde{T}(\alpha y)) = (Tx\ ,\ \alpha y) = \overline{\alpha}(Tx\ ,\ y) = \overline{\alpha}(x\ ,\ \widetilde{T}y) = (x\ ,\ \alpha\widetilde{T}y)\ ,$$

de donde resulta

$$\widetilde{T}(\alpha y) = \alpha\widetilde{T}y\ .\ \blacksquare$$

Definición 18.1.1. Dados los espacios de producto interno U y V y una transformación lineal cualquiera

$$T : U \to V\ ,$$

la transformación lineal $\widetilde{T} : V \to U$ tal que

$$(Tx\ ;\ y) = (x\ ,\ \widetilde{T}y)\quad ,\quad \forall\ x \in U\ ,\ y \in V\ , \qquad (18.1.1)$$

cuya existencia fué establecida en el teorema 18.1.1, se llama *adjunta de T.*

Teorema 18.1.2. Dada la transformación $T : U \to V$, las transformaciones adjunta $\widetilde{T} : V \to U$. y dual $T^* : V \to U$ (relativa esta última a la dualidad de los espacios U y V consigo mismos, que fué definida en la sección 17.6) están relacionadas por la fórmula

$$\widetilde{T}y = \overline{T^*\overline{y}}\quad ,\qquad \forall\ y \in V\ , \qquad (18.1.2)$$

supuesto que son dadas una conjugación $\chi_U : U \to U$ de U y una conjugación $\chi_V : V \to V$ de V, de manera que (18.1.2) equivale a

$$\widetilde{T}y = \chi_U[T^*(\chi_V y)] \quad , \quad \forall\, y \in V \ ,$$

o sea a

$$\widetilde{T} = \chi_U \circ T^* \circ \chi_V \ .$$

Demostración. Para cualesquiera $x \in U$ e $y \in V$ se tiene

$$(x\ ,\ \widetilde{T}y) = (Tx\ ,y) = \langle Tx\ ,\bar{y}\rangle = \langle x\ ,T^*\bar{y}\rangle = (x\ ,\overline{T^*\bar{y}}) \ ;$$

y por consiguiente

$$\widetilde{T}y = \overline{T^*\bar{y}} \quad . \quad \blacksquare$$

Corolario. Si U y V son espacios euclideanos, las transformaciones $\widetilde{T}$ y T^* son idénticas

Demostración. Resulta inmediatamente del teorema anterior. ■

Teorema 18.1.3. Dadas la transformación lineal $T : U \to V$ de los espacios U y V de producto interno y de dimensión finita y las bases ortonormales $\{e_i\}$ y $\{f_j\}$ de U y V, respectivamente; si $M = (\alpha_j^i)$ es la matriz de T respecto de esas bases y $M_1 = (\beta_1^k)$ es la matriz de $\widetilde{T}$ respecto de las mismas bases, entonces es

$$M_1 = \overline{M^t} = \widetilde{M} \ ,$$

es decir que la matriz de $\widetilde{T}$ es la adjunta $\widetilde{M}$ (definición 17.4.1) de la matriz M de T. Recíprocamente, si la matriz M_1 de T_1 respecto de bases ortonormales es igual a $\widetilde{M} = \overline{M^t}$, entonces es $T_1 = \widetilde{T}$.

Demostración. Por definición de $\widetilde{T}$ se tiene para valores cualesquiera de i y j,

$$(T\,e_i\ ,f_j) = (e_i\ ,\widetilde{T}\,f_j)$$

Por tanto,

$$\Big(\sum_{k=1}^{n} \alpha_i^k f_k\ ,\ f_j\Big) = \Big(e_i\ ,\ \sum_{l=1}^{m} \beta_j^l\, e_l\Big) \ ,$$

de donde

$$\sum_{k=1}^{n} \alpha_i^k (f_k \,,\, f_j) = \sum_{l=1}^{m} \overline{\beta_j^l}\, (e_i \,,\, e_l) \quad ,$$

y puesto que las bases $\{e_i\}$ y $\{f_j\}$ son ortonormales,

$$\alpha_i^j = \overline{\beta_j^i} \quad .$$

Resulta de esta manera que $M_1 = \overline{M^t} = \widetilde{M}$. La demostración de la reciproca es sencilla. ∎

Corolario. Si U y V son espacios *euclideanos* de dimensión finita y $T : U \to V$ es una transformación lineal cuya matriz respecto de bases ortonormales es $M = (\alpha_j^i)$. entonces, si M_1 es la matriz de $\widetilde{T}$ respecto de las mismas bases, se tiene

$$M_1 = M^t \,,$$

es decir que la matriz de $\widetilde{T}$ es la transpuesta de la matriz de T; y, reciprocamente, si la matriz de T_1 respecto de una base ortonormal es igual a M^t, entonces es $T_1 = \widetilde{T}$.

Demostración. El corolario es consecuencia directa del teorema anterior. ∎

La noción de transformación lineal adjunta puede aplicarse en particular a los endomorfismos de los espacios de producto interno. Por tanto, dado un endomorfismo T de un espacio U de producto interno, de dimensión finita, es definido el endomorfismo adjunto $\widetilde{T} : U \to U$, tal que

$$(T\,x \,,\, y) = (x \,,\, \widetilde{T}\,y) \qquad \forall\, x,y \in U \,.$$

La matriz de $\widetilde{T}$ respecto de una base ortonormal de U será la matriz adjunta de la matriz de T respecto de la misma base. Si U es euclideana entonces la matriz de $\widetilde{T}$ es la transpuesta de la matriz de T.

En adelante emplearemos los símbolos $(T_1 + T_2)^{\sim}$ y $(T_1\,T_2)^{\sim}$ para representar a los adjuntos de los endomorfismos $T_1 + T_2$ y $T_1\,T_2$ respectivamente.

Teorema 18.1.4. Para los endomorfismos de un espacio U de producto interno de dimensión finita se cumplen las propiedades siguientes:

1. $(T_1 + T_2)^{\sim} = \widetilde{T}_1 + \widetilde{T}_2$
2. $\lambda\widetilde{T} = \bar{\lambda}\,\widetilde{T}$, para $\lambda \in \mathbb{F}$
3. $\widetilde{\widetilde{T}} = T$, por tanto $(\widetilde{T}x, y) = (x, Ty)$ para x e y cualesquiera de U.
4. $(T_1 T_2)^{\sim} = \widetilde{T}_2 \widetilde{T}_1$.
5. $\widetilde{T^k} = (\widetilde{T})^k$ para k entero positivo.

Demostración. Se deja al cuidado del lector. ∎

18.2. Endomorfismos normales de los espacios de producto interno.

Definición 18.2.1. Un *endomorfismo* T del espacio U de producto interno de dimensión n se llama *normal* si es conmutable con su adjunto, es decir si se cumple que

$$\widetilde{T}\,T = T\,\widetilde{T} \ .$$

En virtud de la propiedad 3 del teorema 18.1.4, T y $\widetilde{T}$ son normales simultáneamente.

Definición 18.2.2. Una *matriz* M de orden n es *normal* si es conmutable con su adjunta, es decir si

$$\widetilde{M}\,M = M\,\widetilde{M} \ .$$

En particular, las matrices unitarias y ortogonales (definición 17, 4.2) son normales.

Teorema 18.2.1. La condición necesaria y suficiente para que un endomorfismo T de un espacio U de producto interno de dimensión n sea normal, es que sean normales las matrices que lo representan respecto de las bases ortonormales.

Demostración. Si T es normal, es decir, si $\widetilde{T}\,T = T\,\widetilde{T}$, y si M es la matriz de T respecto de una base ortonormal, en cuyo caso la ma

triz de $\tilde{T}$ respecto de la misma base es $\tilde{M}$, se cumple que

$$\tilde{M}\,M = M\,\tilde{M} \quad ,$$

y por tanto M es normal.

Recíprocamente, si las matrices de T y $\tilde{T}$, respecto de una base ortonormal, son M y $\tilde{M}$ y se tiene $\tilde{M}\,M = M\,\tilde{M}$, entonces, si $M(T_1)$ representa a la matriz de un endomorfismo cualesquiera T_1 respecto de dicha base, se tiene (ver (6.8.1))

$$M(\tilde{T}\,T) = M(\tilde{T})\,M(T) = \tilde{M}\,M = M\,\tilde{M} = M(T)\,M(\tilde{T})$$
$$= M(T\,\tilde{T}) \ ,$$

de donde se deduce que $\tilde{T}\,T = T\,\tilde{T}$. ∎

En la presente sección nuestro objetivo principal es probar que un endomorfismo normal T de un espacio U de producto interno, cuyas raíces características (es decir las raíces de su polinomio característico (definición 13.2.3)) pertenecen al campo $\mathbb{F}$, puede ser representado por una matriz diagonal si se elige en U una base ortonormal adecuada (teorema 18.2.4 de descomposición espectral de endomorfismos normales). Nos valdremos para ello del hecho ya establecido de que la representación diagonal es posible si y solo si el polinomio mínimal ϕ_T se puede descomponer en el producto de factores lineales distintos entre sí, pertenecientes a $\mathbb{F}[x]$ (teorema 13.3.7).

Teorema 18.2.2. Sea T un endomorfismo normal del espacio U de producto interno y de dimensión n. Se cumple entonces que:

1. Para cualquier vector x de U es

$$\|Tx\| = \|\tilde{T}\,x\| \quad ,$$

es decir que la norma de la imagen de un vector por un endomorfismo normal es igual a la norma de la imagen del mismo vector por el endomorfismo adjunto.

2. Si $T\,x_0 = \lambda\,x_0$, donde $\lambda \in \mathbb{F}$, entonces es

$$\tilde{T}\,x_0 = \overline{\lambda}\,x_0 \quad .$$

Por tanto, *si* λ *es un valor propio de* T, $\overline{\lambda}$ *lo es de* $\widetilde{T}$, *y recíprocamente.*

3. Si $Tx_o = \lambda_1 x_o$ y $Ty_o = \lambda_2 y_o$, donde λ_1 y λ_2 son elementos de $\mathbb{F}$ distintos entre sí, entonces es

$$(x_o , y_o) = 0 ;$$

en particular, *dos vectores propios de un endomorfismo normal, que corresponden a valores propios diferentes, son ortogonales.*

4. Si $T^p x_o = 0$ para un vector x_o de U y para algún entero $p \geqslant 1$ entonces es $Tx_o = 0$.

5. Cualquiera que sea el polinomio $g(x) \in \mathbb{F}[x]$, $g(T)$ es un endomorfismo normal. En particular, si T es normal, también lo es T^n, cualquiera que sea n.

Demostración.

1. $$\|Tx\|^2 = (Tx , Tx) = (x , \widetilde{T}Tx) = (x , T\widetilde{T}x) = (\widetilde{T}x , \widetilde{T}x) = \|\widetilde{T}x\|^2 .$$

Antes de probar la propiedad 2. vamos a establecer un lema

Lema: Si uno de los endomorfismos T y $T - \lambda I$ es normal, el otro también lo es.

En efecto, se tiene

$$(T - \lambda I)^{\sim}(T - \lambda I) - (T - \lambda I)(T - \lambda I)^{\sim} = (\widetilde{T} - \overline{\lambda} I)(T - \lambda I) - (T - \lambda I)(\widetilde{T} - \overline{\lambda} I) = \widetilde{T}T - T\widetilde{T} .$$

El primer miembro de esta igualdad se anula si y solo si el último miembro se anula. Lo cual prueba el lema.

2. La hipótesis de que T es normal y el lema anterior, del cual se sigue que $T - \lambda I$ es también normal, permiten aplicar la propiedad 1. ya demostrada y escribir,

$$\|(T - \lambda I)^{\sim}(x_o)\| = \|(T - \lambda I)(x_o)\| = 0 ,$$

por cuanto se supone que $Tx_o = \lambda x_o$. Se tiene por tanto

$$(T - \lambda I)^{\sim}(x_o) = (\widetilde{T} - \overline{\lambda})x_o = 0 \ ,$$

o sea

$$\widetilde{T}\,x_o = \overline{\lambda}x_o \quad .$$

3. Por la propiedad 2,

$$\lambda_1(x_o, y_o) = (\lambda_1 x_o, y_o) = (Tx_o, y_o) = (x_o, \widetilde{T}y_o) = (x_o, \overline{\lambda}_2 y_o)$$
$$= \lambda_2(x_o, y_o) \ .$$

Puesto que $\lambda_1 \neq \lambda_2$, se deduce que $(x_o, y_o) = 0$.

4. Sea p el menor entero positivo tal que $T^p x_o = 0$. Supongamos que sea $p > 1$. Escribamos $y_o = T^{p-1}x_o$. Puesto que

$$Ty_o = T^p x_o = 0 \ ,$$

se tiene $\| Ty_o\| = \| \widetilde{T}y_o\| = 0$, en virtud de 1. Luego es $\widetilde{T}(y_o) = 0$, y por tanto

$$0 = (\widetilde{T}y_o \ , \ T^{p-2}x_o) = (y_o \ , \ TT^{p-2}x_o) = (y_o \ , \ y_o) \ .$$

Se sigue de aquí que $y_o = 0$, o sea que $T^{p-1}x_o = 0$; lo cual contradice la suposición hecha respecto de p. Por consiguiente debe ser $p = 1$, es decir que $Tx_o = 0$.

5. Sea $g(x) = \sum_{i=1}^{r} \lambda_i x^i$ un polinomio de grado r de $\mathbb{F}[x]$. Se tiene

$$g(T) = \sum_{i=1}^{r} \lambda_i T^i \ ,$$

y en virtud del teorema 18.1.4,

$$[g(T)]^{\sim} = \sum_{i=1}^{r} \overline{\lambda}_i (\widetilde{T})^i \quad .$$

Ahora bien, puesto que $\widetilde{T}$ es conmutable con T, $\sum_{i=1}^{r} \overline{\lambda}_i (\widetilde{T})^i$ es conmutable con T, y por tanto con $g(T)$. Es decir que $[g(T)]^{\sim}$ es conmutable con $g(T)$, y por consiguiente es normal. ∎

Corolario 1. Si T es un endomorfismo normal, entonces

$$Ker\ T = Ker\ T^p \ , \ \forall \ p \geqslant 1 \ ;$$

y por tanto *rango* T = *rango* T^p .

Demostración. Resulta de la propiedad 4. del teorema anterior y de la relación (3.5.2).

Corolario. Si T es un endomorfismo normal del espacio U de producto interno, de dimensión n, entonces es

$$U = Ker\ T \oplus Im\ T.$$

Demostración. La propiedad 1. del teorema 18.2.2 implica, evidentemente, que $Ker\ T = Ker\ \widetilde{T}$. Ahora bien, por definición de $\widetilde{T}$, se tiene que

$$(Tx\ ,\ y) = (x\ ,\ \widetilde{T}y) \qquad ,\quad \forall\ x\ ,y\ \epsilon\ U. \qquad (18.2.1)$$

Se deduce de aquí que, para un vector dado y de U, las condiciones

$$(Tx\ ,\ y) = 0 \qquad ,\quad \forall\ x\ \epsilon\ U$$

y

$$(x\ ,\ \widetilde{T}y) = 0 \qquad ,\quad \forall\ x\ \epsilon\ U$$

son equivalentes. La primera condición expresa que y es ortogonal a los vectores de la imagen de T, o sea que y pertenece al complemento ortogonal de la imagen de T. La segunda condición implica que $\widetilde{T}y = 0$, o sea que y pertenece a $Ker\ \widetilde{T} = Ker\ T$. Por consiguiente podemos establecer que

$$Ker\ T = (Im\ T)^{\perp} .$$

Ahora bien, según hemos establecido en el teorema 17.7.1 ,

$$U = (Im\ T)^{\perp} \oplus Im\ T.$$

Por consiguiente se tiene

$$U = Ker\ T \oplus Im\ T. \qquad \blacksquare$$

Teorema 18.2.3. Sea T un endomorfismo normal de U. Si ϕ_1 es un factor irreducible del polinomio minimal ϕ_T , entonces es

$$\phi_T = \phi_1 \cdot \psi \ ,$$

donde ϕ_1 y ψ son primos entre sí.

Demostración. ϕ_T puede escribirse en la forma

$$\phi_T = \phi_1^p \psi \quad ,$$

donde ϕ_1 y ψ son primos entre sí y p es un entero positivo no menor que 1. Probaremos que $p = 1$. Sea x un vector cualquiera de U. Puesto que

$$\phi_T(T) = (\phi_1(T))^p \psi(T) = 0 \ ,$$

se tiene

$$(\phi_1(T))^p (\psi(T)(x)) = 0.$$

Y como $\phi_1(T)$ es normal al mismo tiempo que T, según la propiedad 5. del teorema 18.2.2, la propiedad 4. del mismo teorema nos permite deducir que

$$\phi_1(T)(\psi(T)(x)) = 0 \ .$$

Pero esta relación se cumple para todo $x \in U$, luego

$$\phi_1(T)\psi(T) = 0.$$

T es por tanto raiz de la ecuación $\phi_1(x)\psi(x) = 0$, y por tanto

$$\phi_1(x)\,\psi(x)$$

es el polinomio minimal ϕ_T de T. Es decir que $p = 1$. ∎

Corolario. Si T es un endomorfismo normal cuyas raices características pertenecen a $\mathbb{F}$, entonces ϕ_T se descompone en un producto de factores lineales pertenecientes a $\mathbb{F}[x]$, diferentes entre sí.

Demostración. Sabemos que ϕ_T puede expresarse en la forma

$$\phi_T(x) = (x - \lambda_1)^{p_1} \cdots (x - \lambda_k)^{p_k} \quad ,$$

donde $\lambda_1, \ldots, \lambda_k$ son elementos de $\mathbb{F}$ distintos entre sí. Si aplicamos el teorema 18.2.3 a cada uno de los factores irreducibles $(x - \lambda)$, resulta $p_i = 1$ para $i = 1, \ldots, k$. ∎

Teorema 18.2.4. *(Teorema de descomposición espectral de los endomorfismos normales cuyas raices características pertenecen a $\mathbb{F}$).*

Sea T un endomorfismo normal de un espacio de producto interno U, de dimensión n, cuyas raices características pertenecen a

IF . Existe entonces una base ortonormal de U respecto de la cual T puede representarse por una matriz diagonal; los términos diagonales de esa representación son los valores propios de T, cada uno de los cuales aparece un número de veces igual a su multiplicidad, cuando se le considera como raiz del polinomio característico f_T de T.

Demostración, Por el corolario del teorema 18.2.3 y por el teorema 13.3.4, puede verse que

$$U = U_1 \oplus U_2 \oplus \cdots \oplus U_k$$

donde

$$U_i = \{x \in U \mid Tx = \lambda_i x\} \quad , \quad 1 \leqslant i \leqslant k \ ,$$

y $\lambda_1 , \ldots , \lambda_k$ son las diferentes raices de ϕ_T ,las cuales son todos los distintos valores propios de T . Por la propiedad 3. del teorema 18. 2.2 se ve que U_i y U_j son ortogonales entre sí para $i \neq j$. Podemos elegir entonces una base ortonormal en cada uno de los subespacios U_i y construir con ellas una base ortonormal de U que tiene la propiedad del enunciado, lo cual resulta de la manera como se construye la matriz de un endomorfismo. ∎

Ejercicios

***18.2.1.** Demostrar que para que un endomorfismo T de un espacio U de producto interno, de dimensión finita sea normal es necesario y suficiente que se cumpla una de las condiciones siguientes:

1) $(Tx \ , \ Ty) = (\widetilde{T}x \ , \ \widetilde{T}y) \quad , \quad \forall \ x,y \in U$

2) $\| Tx \| = \| \widetilde{T} \, x \| \quad , \quad \forall \ x \in U.$

18.2.2. Probar que un endomorfismo normal T que es nilpotente es el endomorfismo 0.

18.2.3. Probar que una matriz superdiagonal compleja M es diagonal si y solo si es normal .

18.3. Endomorfismos autoadjuntos de los espacios de producto interno.

U representa como anteriormente un espacio de producto interno de dimensión n.

Definición 18.3.1. Un endomorfismo T del espacio U se llama *autoadjunto* si es idéntico a su adjunto, es decir si

$$T = \widetilde{T} \quad .$$

En el caso real, T se llama *simétrico* y en el caso complejo, *hermitiano.*

Es evidente que un endomorfismo autoadjunto es normal, por cuanto $\widetilde{T}\,T = T^2 = T\,\widetilde{T}$.

Definición 18.3.2. Una matriz $M = (\alpha_j^i)$ se llama *autoadjunta* si es igual a su adjunta, es decir si $M = \widetilde{M}$, o sea

$$\alpha_j^i = \overline{\alpha_i^j} \quad .$$

La matrices complejas autoadjuntas se llaman *hermitianas.* Las matrices reales satisfacen a la condición

$$M = \widetilde{M} = M^t \ ,$$

o sea que tienen lugar las relaciones

$$\alpha_j^i = \alpha_i^j \ ,$$

y se llaman matices *simetricas.*

Las matrices autoadjuntas, tanto hermitianas como simétricas, son evidentemente normales.

Teorema 18.3.1. Cada una de las dos condiciones siguientes es necesaria y suficiente para que el endomorfismo T sea autoadjunto:

(1) $(Tx\ ,\ y) = (x\ ,\ Ty)$ para x e y cualesquiera de U

(2) T se representa por una matriz autoadjunta, respecto de cada basa ortonormal de U.

Para que T sea autoadjunto basta con que se represente por una matriz autoadjunta respecto de una base ortonormal.

Demostración.

(1) Si T es autoadjunto, entonces se cumple (1) por definición de $\tilde{T}$. Reciprocamente, si se cumple (1) se tiene que

$$(x\ ,\ Ty) = (x\ ,\ \tilde{T}y)$$

para x e y cualesquiera de U. Por tanto es $T = \tilde{T}$

(2) En virtud del teorema 18.1.3, para que T sea autoadjunto es necesario que su matriz M respecto de cada base ortonormal cumpla la condición $M = \overline{M'} = \tilde{M}$, es decir que sea autoadjunta. Por el mismo teorema, si esta condición se cumple para una base ortonormal, entonces T es autoadjunta. ∎

18.4. Isometrías de espacios de producto interno.

Definición 18.4.1. Una transformación lineal T de un espacio de producto interno U sobre $\mathbb{F}$ en otro espacio de producto interno V sobre el mismo campo $\mathbb{F}$, se llama *isometría del espacio U en el espacio V* si se cumple que

$$(Tx\ ,\ Ty) = (x\ ,\ y) \quad , \qquad \forall\ x,y\ \epsilon\ U$$

En particular, si $U = V$ se dice que T es una *isometría de V*.

Teorema 18.4.1. Para que la transformación lineal $T : U \to V$ del espacio de producto interno U en el espacio de producto interno V sea una isometría es necesario y suficiente que se cumpla que

$$\| Tx \| = \| x \| \qquad \forall\ x\ \epsilon\ U\ ,$$

o lo que es lo mismo, que

$$(Tx\ ,\ Tx) = (x\ ,\ x) \quad , \quad \forall\ x\ \epsilon\ U.$$

Demostración. Sea T una isometría. Se tiene entonces

$$\| x \|^2 = (x\ ,x) = (Tx\ ,\ Tx) = \| Tx \|^2\ .$$

Inversamente, supongamos que se cumple la condición

$$\| x \| = \| Tx \|\ ,\ \text{o sea}\ (x\ ,x) = (Tx\ ,\ Tx)\ .$$

Cualesquiera que sean x e y de U se tendrá

$$(T(x + y)\ ,\ T(x + y)) = (x + y\ ,\ x + y)\ ,$$

es decir,

$$(Tx, Tx) + (Tx, Ty) + (Ty, Tx) + (Ty, Ty) = (x, x) + (x, y) + (y, x) + (y, y),$$

de donde

$$(Tx, Ty) + (Ty, Tx) = (x, y) + (y, x). \qquad (18.4.1)$$

Ahora bien,

$$(Ty, Tx) = \overline{(Tx, Ty)} \quad \text{y} \quad (y, x) = \overline{(x, y)}$$

De (18.4.1) se deduce entonces

$$\mathcal{R}(Tx, Ty) = \mathcal{R}(x, y) \ . \qquad (18.4.2)$$

Esta relación es válida cualesquiera que sean los vectores x e y de U. En particular, para x e y arbitrarios se tendrá

$$\mathcal{R}(Tx, T(iy)) = \mathcal{R}(x, iy) \ , \qquad (18.4.3)$$

Pero

$$(Tx, T(iy)) = (Tx, iTy) = -i(Tx, Ty)$$

y

$$(x, iy) = -i(x, y) .$$

Luego

$$\mathcal{R}(Tx, T(iy)) = \mathcal{J}(Tx, Ty) \quad \text{y} \quad \mathcal{R}(x, iy) = \mathcal{J}(x, y) .$$

Por consiguiente, de (18.4.3) se sigue

$$\mathcal{J}(Tx, Ty) = \mathcal{J}(x, y) .$$

De esta relación y de 18.4.2 resulta que

$$(Tx, Ty) = (x, y) . \qquad \blacksquare$$

Observación. La demostración de la segunda parte del teorema anterior se simplifica considerablemente en el caso particular en que $\mathbb{F} = \mathbb{R}$, es decir cuando U y V son euclideanos. En ese caso, la simetría del producto interno permite deducir inmediatamente de (18.4.1) que $(Tx, Ty) = (x, y)$.

Teorema 18.4.2. Si U y V son espacios de producto interno sobre un campo $\mathbb{F}$ y $T : U \to V$ es una isometría, entonces T es una transformación inyectiva. Si además es $dim\ U = dim\ V = n$, entonces T es un isomorfismo. Si, en especial, es $U = V$, T es un automorfismo.

Demostración. Sea $T : U \to V$ una isometría. Si se supone que

$$Tx = Ty \text{ , entonces es } \quad T(x - y) = 0. \text{ Por tanto}$$

$$(x - y\ ,\ x - y) = (T(x - y)\ ,\ T(x - y)) = 0,$$

de donde resulta que $x - y = 0$, o sea que $x = y$. T es por tanto inyectiva. Si $dim\ U = dim\ V = n$, entonces la imagen de T tiene dimensión n, y por tanto es sobreyectiva. Por consiguiente T es un isomorfismo. ∎

Teorema 18.4.3. Para que la transformación lineal $T: U \to V$, donde U y V son espacios de producto interno y de dimensión n, sea una isometría, es necesario y suficiente que

$$\widetilde{T}\,T = I_U \qquad \text{y} \qquad T\,\widetilde{T} = I_V\,, \qquad (18.4.4)$$

o sea que

$$T^{-1} = \widetilde{T}\ .$$

Además, para que se cumplan las condiciones (18.4.4) basta con que se cumpla una de ellas.

Demostración. Si $T : U \to V$ es una isometría, se tiene

$$(x\ ,\ y) = (Tx\ ,\ Ty) = (x\ ,\ \widetilde{T}\,Ty)\,. \qquad (18.4.5)$$

Ahora bien, dado que, en virtud del teorema 18.4.2, T es un isomorfismo de U sobre V, existe su inversa $T^{-1} : V \to U$, y se tiene

$$(x\ ,\ y) = (x\ ,\ T^{-1}\,Ty)\,. \qquad (18.4.6)$$

De (18.4.5) y (18.4.6) se deduce

$$(x\ ,\ \widetilde{T}\,Ty) = (x\ ,\ T^{-1}Ty)$$

para x e y cualesquiera de U. Por tanto,

$$\tilde{T}T = T^{-1}T ,$$

y por consiguiente

$$T^{-1} = \tilde{T} ,$$

lo que prueba el cumplimiento de (18.4.4).

Si suponemos ahora que se cumple $\tilde{T}T = I_U$, entonces se tiene

$$(x , y) = (x , \tilde{T}Ty) = (Tx , Ty) ,$$

cualesquiera que sean x e y. Por tanto T es una isometría, lo que, por lo ya demostrado, implica que se cumple también $T\tilde{T} = I_V$.

Si se hubiera supuesto que se cumple $T\tilde{T} = I_V$, entonces por lo que acabamos de probar, $\tilde{T}$ sería una isometría y por tanto se cumpliría también que $\tilde{T}T = I_U$. ∎

Corolario 1. Si T es una isometría, también lo es $\tilde{T}$ y recíprocamente.

Corolario 2. Si T es una isometría del espacio U de producto interno, se tiene que

$$T\tilde{T} = \tilde{T}T = I_U ,$$

y por tanto T es un endomorfismo normal.

En consecuencia, a las isometrías son aplicables las propiedades establecidas para los endomorfismos normales.

Teorema 18.4.4. Sea U un espacio *n-dimensional* de producto interno, y sean $\{e_i\}$ y $\{f_i\}$ dos de sus bases ortonormales. Existe entonces una y solo una isometría T de U tal que

$$Te_i = f_i \quad , \quad i = 1, \ldots , n . \qquad (18.4.7)$$

Demostración. Ya sabemos que las relaciones (18.4.7) definen uno y solo un isomorfismo T de U. Vamos a probar que es una isometría. Sea

$$x = \sum_{i=1}^{n} \xi^i e_i$$

un elemento cualquiera de U. Se tiene entonces debido a la orto-

normalidad de la base $\{e_i\}$,

$$(x \,.\, x) = \sum_{i=1}^{n} \xi^i \overline{\xi^i}$$

Semejantemente, debido a la ortonormalidad de la base $\{f_i\}$ se tiene

$$(Tx \,.\, Tx) = \sum_{i=1}^{n} \sum_{j=1}^{n} \xi^i \overline{\xi^j} (Te_i \,.\, Te_j) = \sum_{i=1}^{n} \xi^i \overline{\xi^i} \quad .$$

Por consiguiente es

$$(x \,.\, x) = (Tx \,.\, Tx),$$

lo cual prueba que T es una isometría. ∎

Corolario 1. Si $\{e_i\}$ es una base ortonormal de U, es condición necesaria y suficiente para que T sea una isometría que $\{Te_1, \ldots, Te_n\}$ sea una base ortonormal de U.

Demostración. La condición es suficiente en virtud del teorema anterior. Es necesaria porque si T es una isometría, entonces en virtud de la definición 18.4.1

$$(Te_i \,.\, Te_j) = (e_i \,.\, e_j) = \delta_{ij} \quad . \qquad \blacksquare$$

Corolario 2. Dada una base ortonormal $\{e_i\}$ de U, es definida una biyección entre las bases ortonormales de U y las isometrías del mismo espacio.

Demostración. La biyección es definida por

$$T \longleftrightarrow \{Te_1, \ldots, Te_n\} \quad . \qquad \blacksquare$$

Teorema 18.4.3. La condición necesaria y suficiente para que un endomorfismo T de un espacio *n-dimensional* U de producto interno sea una isometría es que las matrices M que lo representan *respecto de las bases ortonormales de* U cumplan la condición

$$\widetilde{M}\, M = M\, \widetilde{M} = I \quad , \qquad (18.4.8)$$

que es equivalente a

$$M^{-1} = \widetilde{M} \quad ;$$

es decir que M sea unitaria u ortogonal según que sea compleja o

real. Para que la precedente condición se cumpla es suficiente que la matriz M que representa a T respecto una base ortonormal dada cumpla una de las condiciones

$$\widetilde{M}\,M = I \qquad \text{ó} \qquad M\,\widetilde{M} = I.$$

Demostración. Si T es una isometría $\Rightarrow T\,\widetilde{T} = \widetilde{T}\,T = I \Rightarrow$

$$M(T)M(\widetilde{T})\,M(T) = I.$$

Por tanto, si la base elegida es ortonormal, se tendrá

$$M(\widetilde{T})M(T) = M(T)M(\widetilde{T}) = I\,.$$

Reciprocamente, si para una base ortonormal se cumple que

$$M(\widetilde{T})\,M(T) = I \Rightarrow \widetilde{T}\,T = I \Rightarrow T\,\widetilde{T} = I. \quad \blacksquare$$

En el cuadro siguiente se han considerado los endomorfismos que serán objeto particular de nuestro interés:

- **Endomorfismos normales** $T\,\widetilde{T} = \widetilde{T}\,T$
  - **autoadjuntos** $T = \widetilde{T}$
    - **hermitianos** $(\mathbb{C})$
    - **simétricos** $(\mathbb{R})$
  - **Isométricas** $T^{-1} = \widetilde{T}$
    - **unitarios** $(\mathbb{C})$
    - **ortogonales** $(\mathbb{R})$

Los endomorfismos isométricos o isometrías, se llaman *unitarios* **si son de espacios unitarios y** *ortogonales* **si son de espacios euclídeanos.**

§19. Endomorfismos de los espacios unitarios.

En el presente parágrafo, U designará a un espacio unitario de dimensión n.

19.1. Endomorfismos normales de los espacios unitarios.

Teorema 19.1.1. *(Teorema de descomposición espectral de los endomorfismos normales de un espacio unitario).*

Sea T un endomorfismo normal de un espacio unitario de dimensión n. Entonces existe una base ortonormal de U respecto de la cual T puede representarse por una matriz diagonal, cuyos elementos diagonales son todos los valores propios de T, cada uno de los cuales aparece repetido un número de veces igual a su multiplicidad cuando se le considera como raiz del polinomio característico f_T de T.

Demostración. Este teorema es consecuencia inmediata del teorema 18.2.4 pues todas las características de T pertenecen en este caso a $\mathbb{F} = \mathbb{C}$, por cuanto $\mathbb{C}$ es algebraicamente cerrado. ∎

Corolario. Para una matriz compleja normal M de orden n existe una matriz unitaria P del mismo orden, tal que PMP^{-1} es una matriz diagonal.

Demostración. Según el teorema 18.2.1, M puede considerarse como la matriz de un endomorfismo normal T del espacio unitario $\mathbb{C}^n$, referido a la base canónica de dicho espacio. Por el teorema 18.2.4 existe una base ortonormal $\{e_i\}$ de $\mathbb{C}^n$, respecto de la cual T se representa por una matriz diagonal PMP^{-1}, donde P es la matriz de cambio de la base canónica por la base $\{e_i\}$, y por tanto es una matriz unitaria, de conformidad con el teorema 17.4.1. ∎

La representación diagonal de un endomorfismo normal de un espacio unitario puede obtenerse por una construcción directa, como se muestra a continuación..

Lema. Si T es un endomorfismo normal de un espacio unitario, entonces T y $\widetilde{T}$ tienen los mismos vectores propios, y los correspon-

dientes valores propios son números complejos conjugados entre sí.

Demostración. Esta proposición es la parte 2. del teorema 18.2.2. ∎

Teorema 19.1.2. Si T es un endomorfismo normal de un espacio unitario entonces posee n vectores propios que son mutuamente ortogonales, y que constituyen una base de U referida a la cual la matriz de T es diagonal, y los elementos diagonales son los correspondientes valores propios. (Esta proposición está contenida en el teorema 18.2.4. Su demostración dada aquí proporciona la construcción directa a que nos hemos referido).

Demostración. Consideremos el polinomio característico f_T de T. Según el teorema fundamental del álgebra dicho polinomio tiene una raiz λ_1, la cual será valor propio de T. Sea x_1 un correspondiente vector propio y U_1 el complemento ortogonal de x_1. El subespacio U_1 es invariante respecto de T, porque cualquiera que sea el vector $y \in U_1$ se tiene

$$(Ty , x_1) = (y , \widetilde{T}x_1) = (y , \overline{\lambda}_1 x_1) = \lambda_1 (y , x_1) = 0 ,$$

y por tanto es $Ty \in U_1$. Además es $dim\ U_1 = n - 1$. La restricción de T a U_1 es evidentemente un endomorfismo normal de este espacio; por tanto existe un vector propio $x_2 \in U_1$ de esa restricción y por consiguiente de T, correspondiente a un valor propio λ_2.

Sea U_2 el complemento ortogonal del subespacio generado por x_1 y x_2. U_2 es invariante respecto de T, porque cualquiera que sea el vector $y \in U_2$ se tiene

$$(Ty , \alpha x_1 + \beta x_2) = (y , \widetilde{T}(\alpha x_1 + \beta x_2)) = (y , \alpha\bar{\lambda}_1 x_1 + \beta\bar{\lambda}_2 x_2)$$

$$= \bar{\alpha}\lambda_1 (y , x_1) + \bar{\beta}\lambda_2 (y , x_2) = 0 ,$$

y por tanto $Ty \in U_2$. Además $dim\ U_2 = n - 2$. La restricción de T a U_2 es un endomorfismo normal, y por tanto existe un vector propio $x_3 \in U_2$ de esa restricción, y por consiguiente de T, correspondiente a un valor propio λ_3. Continuando de esa manera se obtienen n vectores propios $x_1 , \ldots , x_n$ que son mutuamente ortogo

nales, es decir tales que

$$(x_i , x_j) = 0 , \qquad (i \neq j).$$

Puede suponerse que esos vectores constituyen una base ortonormal de U por cuanto, si fuera necesario, el vector x_i puede ser sustituido por el vector unitario

$$\frac{x_i}{\|x_i\|} .$$

Dado entonces que $Tx_i = \lambda_i x_i$, $i = 1, \ldots, n$, la matriz de T relativa a la base $\{x_i\}$ es diagonal y sus términos diagonales son los valores propios $\lambda_1, \ldots, \lambda_n$ de T. ∎

19.2. Endomorfismos hermitianos de un espacio unitario.

Definición 19.2.1. Se llama endomorfismos *hermitianos* a los endomorfismos autoadjuntos de los espacios unitarios.

Teorema 19.2.1. Cada valor propio de un endomorfismo hermitiano T del espacio unitario U es real.

Demostración. Sea λ un valor propio de T, y sea $Tx = \lambda x$ donde $x \neq o$ es un vector propio correspondiente a λ. Se tiene entonces

$$(Tx , x) = (\lambda x , x) = \lambda (x , x)$$
$$= (x , Tx) = (x , \lambda x) = \overline{\lambda}(x , x).$$

Puesto que $(x , x) \neq 0$, se deduce que $\lambda = \overline{\lambda}$, y por tanto λ es real. ∎

Teorema 19.2.2. *(Teorema de descomposición espectral de un endomorfismo hermitiano de un espacio unitario).*

Si T es un endomorfismo hermitiano de un espacio unitario n–*dimensional* U, existe una base ortonormal de U, respecto de la cual T puede representarse por una matriz diagonal; los elementos diagonales son los valores propios de T, todos ellos reales y repetidos un número de veces igual a sus respectivas multiplicidades.

Demostración. Basta aplicar los teoremas 19.1.1 y 19.2.1. ∎

Teorema 19.2.3. Para una matriz hermitiana M de orden n existe

una matriz unitaria P del mismo orden tal que PMP^{-1} es una matriz diagonal cuyos elementos diagonales son reales

Demostración. Esta proposición resulta inmediatamente del teorema 19.2.2 si se supone que M es la matriz de un endomorfismo hermitiano T del espacio unitario $\mathbb{C}^n$ referido a su base canónica. ∎

19.3. Endomorfismos unitarios de los espacios unitarios.

Definición 19.3.1. Un endomorfismo T del espacio unitario U se llama *endomorfismo unitario* si es una isometría de dicho espacio.

A los endomorfismos unitarios son por tanto aplicables las consideraciones de la sección 18.4. En particular serán representados, respecto de las bases ortonormales, por medio de matrices unitarias.

Teorema 19.3.1. Si U es un espacio unitario de dimensión n, se cumple que:

1. Cada valor propio de un endomorfismo unitario T de U tiene valor absoluto 1.
2. El determinante de un endomorfismo unitario T tiene valor absoluto 1.
3. El conjunto de todos los endomorfismos unitarios de U es un grupo respecto de la composición de endomorfismos. El conjunto de todos los endomorfismos unitarios cuyo determinante es igual a 1 constituye un subgrupo del anterior.

Demostración. 1. Si λ es un valor propio de T, existe $x \neq 0$ en U tal que $Tx = \lambda x$. Luego

$$\| x \| = \| Tx \| = |\lambda| \| x \| ,$$

y por tanto

$$|\lambda| = 1.$$

2. Puesto que $T\widetilde{T} = I$, se tiene

$$1 = det\ I = det\ T.\ det\ \widetilde{T} = det\ T.\ \overline{det\ T} = |det\ T|^2 ,$$

de donde

$$|det\ T| = 1 .$$

Este mismo resultado puede deducirse de la propiedad 1. y de la representación diagonal de T que es posible porque T es normal y porque todas sus raices características pertenecen a $\mathbb{F} = \mathbb{C}$ (teorema 18.2.4).

3. Si T_1 y T_2 son dos endomorfismos unitarios de U, entonces es

$$((T_1T_2)x\ ,\ (T_1T_2)y) = (T_1(T_2x)\ ,\ T_1(T_2y)) = (T_2x\ ,\ T_2y) = (x\ ,\ y)\ ,$$

y por tanto T_1T_2 es una isometría, es decir un endomorfismo unitario de U.

Además, cuando T es unitario, su inversa $T^{-1} = \widetilde{T}$ también lo es, en virtud del corolario 1 del teorema 18.4.3. Por tanto los endomorfismos unitarios constituyen un grupo.

Finalmente, si $det\ T_1 = det\ T_2 = 1$, entonces se tiene

$$det\ (T_1T_2) = 1.$$

Y si $det\ T = 1$, entonces $det\ T^{-1} = det\ \widetilde{T} = \overline{det\ T} = 1$. ∎

Corolario 1. Las raices características de las matrices unitarias tienen valor absoluto igual a *1*.

Demostración. Cada matriz unitaria puede considerarse como matriz de un endomorfismo unitario de $\mathbb{C}^n$ referido a la base canónica. Las raices características de la matriz, son los valores propios de ese endomorfismo y por tanto tienen valor absoluto *1*, según el teorema anterior. ∎

Corolario 2. Las matrices unitarias de orden n forman un grupo y las matrices unitarias cuyo determinante es igual a *1*, forman un subgrupo del anterior.

Demostración. Se deja al lector. ∎

Definición 19.3.2. El grupo de las matrices complejas unitarias de orden n se llama *grupo unitario de orden n* y se designa por $U(n)$. El subgrupo de $U(n)$ formado por las matrices unitarias de determinante igual a *1* se llama *grupo unitario especial de orden n* y se designa por $SU(n)$.

Ejemplo. La matriz

$$\begin{pmatrix} e^{iw} & 0 \\ 0 & e^{-iw} \end{pmatrix}$$

pertenece a $SU(2)$, cualquiera que sea el número real w. En efecto, es unitaria porque

$$\begin{pmatrix} e^{iw} & 0 \\ 0 & e^{-iw} \end{pmatrix} \begin{pmatrix} \overline{e^{iw}} & 0 \\ 0 & \overline{e^{-iw}} \end{pmatrix} = \begin{pmatrix} e^{iw}\,\overline{e^{iw}} & 0 \\ 0 & e^{-iw}\,\overline{e^{-iw}} \end{pmatrix}$$

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

Además, su determinante es igual a 1.

Ejercicios

19.3.1. Si M es una matriz compleja normal y U es una matriz compleja unitaria, ambas de orden n, probar que la matriz $UM\tilde{U}$ es también normal.

19.3.2. Si T_1 y T_2 son endomorfismos hermitianos del espacio unitario U, entonces T_1T_2 es hermitiano si y solo si

$$T_1T_2 = T_2T_1 .$$

19.3.3. Demostrar que si T es un endomorfismo cualquiera de un espacio unitario, entonces $\tilde{T}T$ es un endomorfismo hermitiano. Si T es hermitiano, entonces STS es hermitiano cualquiera que sea el endomorfismo S.

19.3.4. Demostrar que cada matriz compleja M puede expresarse de una sola manera en la forma $M_1 + iM_2$ donde M_1 y M_2 son hermitianas. Establecer la proposición análoga para los endomorfismos de un espacio unitario de dimensión n.

19.3.5. Demostrar que el conjunto de todas las matrices hermitianas de orden n forma un espacio vectorial sobre $\mathbb{R}$. Determinar luego la dimensión de ese espacio.

§ 20. Endomorfismos de espacios euclideanos.

En este parágrafo V designa a un espacio euclideano de dimensión n.

20.1. Endomorfismos normales de los espacios euclideanos.

Teorema 20.1.1. *(Teorema de descomposición espectral de los endomorfismos normales de un espacio euclideano, cuyas raices características son reales)*

Sea T un endomorfismo normal del espacio euclideano V. Si todas las raices características de T son reales, entonces existe una base ortonormal de V respecto de la cual T se representa por una matriz diagonal cuyos términos diagonales son los valores propios de T repetidos un número de veces igual a su multiplicidad como raices del polinomio característico.

Demostración. Este teorema es aplicación inmediata del teorema 18.2.4, pues todas las raices características de T pertenecen por hipótesis a $\mathbb{F} = \mathbb{R}$.

Corolario. Si M es una matriz real normal de orden n tal que todas sus raices características son reales, entonces existe una matriz ortogonal S tal que SMS^{-1} es una matriz diagonal.

Demostración. El razonamiento es enteramente análogo al empleado para probar el corolario del teorema 19.1.1; pero tratandose del cambio de bases ortonormales en $\mathbb{R}^n$, la matriz de cambio S es ortogonal y por tanto la matriz diagonal, cuya existencia se estableció en el teorema anterior, tiene la forma $D = SMS^{-1}$, donde S es ortogonal. ∎

20.2 Endomorfismos simétricos de un espacio euclideano.

Definición 20.2.1. Los endomorfismos autoadjuntos de los espacios euclideanos se llaman *endomorfismos simétricos.* Sus matrices, para bases ortonormales, son simétricas.

Teorema 20.2.1. Las raices características de un endomorfismo simétrico son números reales.

(Conviene tener presente que los valores propios de un endomor-

fismo de un espacio real, euclideano en particular, son números reales; a saber las raices características reales. Luego para un endomorfismo simétrico todas las raices características son valores propios).

Demostración. Sea T un endomorfismo simétrico y M la matriz real simétrica que lo representa respecto de una base ortonormal. Se tiene entonces $M = \widetilde{M}$. El polinomio característico f_T de T es igual al polinomio característico f_M de M. Ahora bien, M es autoadjunta y por tanto puede considerarse como una matriz hermitiana que representa a una transformación hermitiana de C^n referido a su base canónica. En virtud del teorema 19.2.1 , todos los valores propios de esa transformación, y por tanto todas sus raices características, que son las raices del polinomio f_M, son números reales. Lo mismo puede decirse entonces de las raices de f_T, es decir de las raices características de T. ■

Observación. Podemos dar otra demostración del teorema anterior, sin hacer uso de las propiedades de las matrices hermitianas y sin salir del campo real:

El polinomio minimal ϕ_T puede descomponerse sobre $\mathbb{R}$ en un producto de factores irreducibles que son lineales o bien de segundo grado (teorema 10.5.2 y 10.7.1). Los factores de segundo grado son de la forma $(x-a)^2 + b^2$, donde $b \neq 0$. Probaremos que ϕ_T no puede ser divisible por un polinomio de segundo grado de esa forma, y por tanto es un producto de factores lineales pertenecientes a $\mathbb{R}[x]$. Supongamos por el contrario que sea

$$\phi_T(x) = ((x-a)^2 + b^2)\psi(x) \qquad (20.2.1)$$

donde $\psi(x)$ es un polinomio sobre $\mathbb{R}$. Dado un vector cualquiera $u \in E$, sea $v = \psi(T)(u)$. Se tiene entonces

$$\phi_T(T)(u) = [(T - aI)^2 + b^2](v) = 0,$$

por cuanto $\phi_T(T) = 0$. Puesto que T es simétrico, es decir tal que $T = \widetilde{T}$, lo mismo ocurre evidentemente con $T - aI$ y por tanto es $(T - aI)^{\sim} = T - aI$. Podemos escribir entonces

$$0 = ([(T - aI)^2 + b^2](v), v) = ((T - aI)^2(v), v) + (b^2 v, v)$$
$$= ((T - aI)(v), (T - aI)(v)) + b^2(v, v)$$
$$= \| (T - aI)(v) \|^2 + b^2 \|v\|^2 .$$

Si fuera $v \neq 0$, resultaría una contradicción pues el último miembro de las igualdades anteriores sería diferente de cero. Debe ser por consiguiente $v = 0$, es decir que $\psi(T)(u) = 0$, y esto para todo $u \in E$. Por tanto sería $\psi(T) = 0$. $\psi(x)$ sería entonces divisible por $\phi_T(x)$, lo cual contradice a la relación (20.2.1). Se concluye que esta última relación no es posible, y por consiguiente $\phi_T(x)$ es un producto de factores lineales pertenecientes a $\mathbb{R}[x]$. Puesto que $\phi_T(x)$ y $f_T(x)$ tienen las mismas raices, resulta que todas las raices de f_T, es decir las raices características de T, son reales.

La demostración puede reducirse a términos elementales de la manera siguiente. Si (a_{ij}) es la matriz del endomorfismo simétrico T respecto de alguna base, las raices características son las reices de la ecuación

$$\begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{vmatrix} = 0 , \qquad (a_{ij} = a_{ji}) .$$

Se trata de probar que todas las raices son reales. Sea $\alpha + \beta i$ una cualquiera de ellas. El sistema de ecuaciones

$$(a_{11} - \alpha - \beta i)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$
$$a_{21}x_1 + (a_{22} - \alpha - \beta i)x_2 + \cdots + a_{2n}x_n = 0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \alpha - \beta i)x_n = 0$$

cuyo determinante es nulo por hipótesis, tendrá una solución no trivial que designaremos por $x_1, \ldots, x_n$. El sistema puede escribirse en la forma

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - (\alpha + \beta i)x_1 &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - (\alpha + \beta i)x_2 &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n - (\alpha + \beta i)x_n &= 0 \ .
\end{aligned}
$$

Escribamos

$$x_k = y_k + z_k i \qquad (k = 1, \ldots, n)\ ,\ y_k\ ,\ z_k \in \mathbb{R}\ .$$

El sistema de ecuaciones considerado es entonces equivalente a los dos sistemas

$$
\left.
\begin{aligned}
a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n - \alpha y_1 + \beta z_1 &= 0 \\
a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n - \alpha y_2 + \beta z_2 &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n - \alpha y_n + \beta z_n &= 0
\end{aligned}
\right\}
\qquad (20.2.2)
$$

$$
\left.
\begin{aligned}
a_{11}z_1 + a_{12}z_2 + \cdots + a_{1n}z_n - \alpha z_1 - \beta y_1 &= 0 \\
a_{21}z_1 + a_{22}z_2 + \cdots + a_{2n}z_n - \alpha z_2 - \beta y_2 &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{n1}z_1 + a_{n2}z_2 + \cdots + a_{nn}z_n - \alpha z_n - \beta y_n &= 0
\end{aligned}
\right\}
\qquad (20.2.3)
$$

Multipliquemos por $z_1, \ldots, z_n$, respectivamente las ecuaciones del sistema (20.2.2) y sumemos los resultados; y análogamente multipliquemos las ecuaciones (20.2.3) por $y_1, \ldots, y_n$ y sumemos los resultados, se obtendrán las ecuaciones

$$\sum_{r,s} a_{rs} z_r y_s - \alpha \sum_r y_r z_r + \beta \sum_r z_r^2 = 0$$

$$\sum_{r,s} a_{sr} z_r y_s - \alpha \sum_s y_s z_s - \beta \sum_s y_s^2 = 0\ .$$

Restando miembro a miembro, teniendo presente que $a_{rs} = a_{sr}$, resulta

$$\beta \sum_r (y_r^2 + z_r^2) = 0.$$

Puesto que no todos los λ_r son nulos, es $\sum_r (y_r^2 + z_r^2) \neq 0$. Luego debe ser $\beta = 0$, es decir que cada raiz característica es real.

Teorema 20.2.2. *(Teorema espectral de los endomorfismos simétricos).* Sea T un endomorfismo simétrico del espacio euclideano E de dimensión n. Existe entonces una base ortonormal de E respecto de la cual T se representa por una matriz diagonal, cuyos elementos diagonales son los valores propios de T repetidos tantas veces como sus respectivas multiplicidades como raices de f_T.

Demostración. Basta aplicar los teoremas 20.2.1. y 20.1.1, pues un endomorfismo simétrico es normal. ∎

Teorema 20.2.3. Dada una matriz real simétrica M de orden n, existe una matriz ortogonal S tal que SMS^{-1} es una matriz diagonal.

Demostración. Se aplica el corolario del teorema 20.1.1, pues M es normal. ∎

A continuación estableceremos que cada endomorfismo simétrico de un espacio euclideano *n-dimensional* V tiene n vectores propios que son mutuamente ortogonales. Este hecho se deduce del teorema 18. 2.4 pues un endomorfismo simétrico es normal. No obstante, es instructivo dar aquí una demostración directa.

Lema 20.2.1. Si T es un endomorfismo simétrico del espacio V y $V_1 \subset V$ es un subespacio de V_1, invariante respecto de T; entonces el complemento ortogonal $V_1^{\perp}$ es también invariante respecto de T. (Se expresa esta propiedad diciendo que los endomorfismos simétricos son *semisimples*).

Demostración. Consideremos un vector cualquiera $z \in V_1^{\perp}$. Cualquiera que sea $y \in V_1$ se tiene $Ty \in V_1$, y por tanto

$$(Tz\ ,\ y) = (z\ ,\ Ty) = 0.$$

Luego $Tz \in V_1^{\perp}$, y por consiguiente $TV_1^{\perp} \subset V_1^{\perp}$. ∎

Teorema 20.2.4. Si T es un endomorfismo simétrico del espacio E, existe un sistema ortonormal de n vectores propios de T que constituyen una base ortonormal de dicho espacio.

Demostración. Sea λ_1 una raiz cualquiera del polinomio característico de T. Por el teorema 20.2.1, λ_1 será real, y por el teorema 13.2.1, λ_1 será valor propio de T. Sea e_1 un vector propio unitario correspondiente a λ_1. Consideremos el subespacio V_1 de dimensión 1 generado por e_1. V_1 es invariante respecto de T pues

$$Te_1 = \lambda_1 e_1 \in V_1 .$$

Luego su complemento ortogonal $V_1^{\perp}$ es también invariante respecto de T. La restricción T_1 de T a $V_1^{\perp}$ será también simétrica y por tanto la precedente construcción puede ser aplicada a T_1. Existe por consiguiente un vector propio unitario e_2 de T_1, y por tanto de T, correspondiente al valor propio real λ_2, el cual, por pertenecer a $V_1^{\perp}$ cumple la condición $(e_1 , e_2) = 0$.

Continuando de esa manera obtendremos finalmente un sistema de n vectores propios e_ν $(\nu = 1, \ldots, n)$ que forman un sistema ortonormal que evidentemente es una base ortonormal de V. ∎

Observación. Los vectores propios $e_1, \ldots, e_n$, constituidos en la demostración del teorema anterior, corresponden a los valores propios reales $\lambda_1, \ldots, \lambda_n$, y satisfacen a las relaciones

$$Te_i = \lambda_i e_i .$$

Por tanto, referido T a la base constituida por esos vectores, su matriz es diagonal y los términos diagonales son los números

$$\lambda_1 , \ldots , \lambda_n .$$

20.3. Endomorfismos ortogonales de espacios euclideanos.

Definición 20.3.1. Un endomorfismo T de un espacio euclideano se llama *ortogonal* si es una isometría de dicho espacio. Respecto de bases ortonormales esos endomorfismos serán representados por ma-

trices ortogonales (Teorema 18.4.5).

Puesto que los endomorfismos ortogonales son normales, si las raices características de uno de ellos son reales, entonces existirá una base ortonormal respecto de la cual puede representarse mediante una matriz diagonal. Pero esto no ocurre en general. Así por ejemplo, la matriz ortogonal

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix},$$

cuyo polinomio característico es $x^2 - 2x\cos\theta + 1$, tiene raices características complejas, salvo el caso en que $\cos\theta = \pm 1$. Esa matriz no puede diagonalizarse porque en tal caso sus raices características serian reales. Tal como en el caso de las matrices unitarias, las matrices ortogonales de orden n constituyen un grupo $O(n)$ respecto de la composición de matrices; y las matrices ortogonales cuyo determinante es igual a 1 constituyen el subgrupo $SO(n)$. $O(n)$ se llama *grupo ortogonal de orden n* y $SO(n)$, *grupo ortogonal especial de orden n*.

Los teoremas siguientes tienen por objeto establecer una representación canónica de los endomorfismos ortogonales.

Teorema 10.3.1. Si V es un espacio euclideano de dimensión 2, y si T es un endomorfismo ortogonal de V, entonces existe una base ortonormal $\{e_1, e_2\}$ de V con respecto a la cual T se representa por una de las matrices

1) $$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \qquad 0 \leqslant \theta < 2\pi$$

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

Demostración. Sea $\{x_1, x_2\}$ una base ortonormal cualquiera de V y sea

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \qquad (20.3.1)$$

la matriz ortogonal que reprepresenta a T respecto de esa base. Debido a la ortogonalidad de la matriz (20.3.1) se cumplen las relaciones

$$a^2 + b^2 = c^2 + d^2 = 1 \quad , \quad ac + bd = 0.$$

Existe entonces un valor único θ , $0 \leqslant \theta < 2\pi$, tal que

$$a = \cos\theta \quad , \quad b = \operatorname{sen}\theta.$$

Por tanto debe ser

$$c.\cos\theta + d.\operatorname{sen}\theta = 0. \qquad (20.3.2)$$

Probaremos que, en virtud de (20.3.2) ,pueden suceder únicamente dos posibilidades, a saber:

I. $c = -\operatorname{sen}\theta \quad , \quad d = \cos\theta$.

II. $c = \operatorname{sen}\theta \quad , \quad d = -\cos\theta$.

En efecto, supongamos primero que $\operatorname{sen}\theta = 0$. En este caso es $\cos\theta = \pm 1$, y resulta de (20.3.2) que $c = 0$. Se tiene además $b = 0$ y a y d pueden tomar los valores $+1$ o -1. Los cuatro casos posibles corresponden a los casos I y II precedentes, con $\theta = 0$ o $\theta = \pi$.

Si $\operatorname{sen}\theta \neq 0$, entonces es $d = -c\,\dfrac{\cos\theta}{\operatorname{sen}\theta}$. De $c^2 + d^2 = 1$ se deduce entonces $c^2 = \operatorname{sen}^2\theta$, $d^2 = \cos^2\theta$. Si $c = -\operatorname{sen}\theta$, entonces es $d = \cos\theta$, y si $c = \operatorname{sen}\theta$ entonces es $d = -\cos\theta$. Resultan asi los casos I y II respectivamente.

Basta probar que las formas I y II se reducen a una de las formas 1) o 2) del enunciado del teorema.

I) $c = -\operatorname{sen}\theta$ y $d = \cos\theta$. La matriz es entonces de la forma 1). Su polinomio característico es $x^2 - (2\cos\theta)x + 1 = 0$, y no tiene raices reales salvo para $\theta = 0$ y $\theta = \pi$,en cuyo caso la matriz se reduce a

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

respectivamente.

II. $c = sen\ 0$, y $d = -cos\ \theta$.

La matriz es entonces

$$\begin{pmatrix} cos\ \theta & sen\ \theta \\ sen\ \theta & -cos\ \theta \end{pmatrix}$$

Su polinomio característico es $x^2 - 1$ y tiene las raices $+1$ y -1, que serán los valores propios de T. Según el teorema 13.2.1 existirán los vectores propios x e y, diferentes de cero, tales que $Tx = x$ y $Ty = -y$. Es claro que podemos suponer que dichos vectores son unitarios. Además, en virtud de la propiedad 3. del teorema 18.2.2, son ortogonales entre sí, luego $\{x , y\}$ es una base ortonormal. Referido a ella, la matriz de T es evidentemente

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

y por tanto de la forma 2). ∎

Teorema 20.3.2. Si T es un endomorfismo ortogonal de un espacio euclideano V de dimensión finita, entonces:

1. Cada raiz característica de T tiene valor absoluto 1, y si λ es una raiz característica real, en cuyo caso λ es un valor propio de T, entonces es o bien $\lambda = 1$ o bien $\lambda = -1$.
2. Si V_1 es un subespacio de V invariante respecto de T, entonces $V_1^{\perp}$ es también invariante respecto de T. Es decir que un endomorfismo ortogonal es semisimple (ver el lema 20.2.1)

Demostración. 1. Sea M una matriz ortogonal que representa a T. Esta matriz puede considerarse también como unitaria, luego, según el corolario 1 del teorema 19.3.1, sus raices características tienen valor absoluto 1.

Sea $Tx = \lambda x$ donde λ es real y $x \neq 0$. Se tiene entonces

$$(x\,,\,x) = (Tx\,,\,Tx) = (\lambda x\,,\,\lambda x) = \lambda^2(x\,,\,x).$$

Por consiguiente es $\lambda^2 = 1$.

2. Puesto que T es una isometría, su restricción a V_1 es una isometría de este subespacio. Por tanto, según el teorema 18.4.1, es un automorfismo de V_1. Luego, si y es un elemento cualquiera de V_1, también es $T^{-1}y \in V_1$. Por consiguiente, dado el elemento arbitrario $z \in V_1^{\perp}$, se tiene

$$(Tz\,,\,y) = (z\,,\,T^{-1}y) = 0\,.$$

Es decir que Tz es ortogonal a todo elemento $y \in V_1$, y por tanto es $Tz \in V_1^{\perp}$, es decir $TV_1^{\perp} \subset V_1^{\perp}$. ∎

Teorema 20.3.3. Sea T un endomorfismo ortogonal de un espacio euclideano V de dimensión n. Existe entonces una base ortonormal $\{e_1\,,\,\ldots\,,\,e_n\}$ de V, respecto de la cual T se expresa por una matriz de la forma

$$\begin{pmatrix}
1 & & & & & & & \\
 & \ddots & & & & & & \\
 & & 1 & & & & & \\
 & & & -1 & & & & \\
 & & & & \ddots & & & \\
 & & & & & -1 & & \\
 & & & & & & \cos\theta_1 & -\operatorname{sen}\theta_1 \\
 & & & & & & \operatorname{sen}\theta_1 & \cos\theta_1 \\
 & & & & & & & & \ddots \\
 & & & & & & & & & \cos\theta_k & -\operatorname{sen}\theta_k \\
 & & & & & & & & & \operatorname{sen}\theta_k & \cos\theta_k
\end{pmatrix}$$

donde 1 y -1 aparecen un número de veces igual a sus respectivas

multiplicidades como raíces características, y $\theta_1, \ldots, \theta_k$, donde $0 < \theta_j < 2\pi$ $(j = 1, \ldots, k)$, son tales que $\alpha_j = \cos \theta_j \pm i \operatorname{sen} \theta_j$ son las distintas raíces características de T que son diferentes de ± 1.

Cada uno de los bloques

$$\begin{pmatrix} \cos \theta_j & -\operatorname{sen} \theta_j \\ \operatorname{sen} \theta_j & \cos \theta_j \end{pmatrix}$$

aparece un número de veces igual a la multiplicidad de α_j como raíz característica. Los demás términos de la matriz son nulos.

Demostración. Llevaremos a cabo la demostración por inducción respecto de la dimensión del espacio V. Si la dimensión de V es 1, entonces el endomorfismo ortogonal T, que según el teorema 18.4.2 es un automorfismo de V, es tal que $Tx = \rho x$ donde ρ es un número real no nulo y x es un vector cualquiera de V.

En efecto, si $x_1 \neq 0$ es un vector cualquiera de V, dicho vector constituye una base de V, y por tanto existe el número real no nulo ρ tal que $Tx_1 = \rho x_1$. Para cualquier otro vector $x = \xi x_1$ de V se tiene entonces

$$Tx = T(\xi x_1) = \xi T x_1 = \xi \rho x_1 = \rho x.$$

Además, por ser T una siometría se debe cumplir que

$$(Tx\ ,\ Tx) = (x\ ,\ x) = (\rho x\ ,\ \rho x) = \rho^2 (x\ ,\ x).$$

Por consiguiente debe ser $\rho^2 = 1$, o sea $\rho = \pm 1$. T debe ser entonces idéntico a I o a $-I$, y por consiguiente el teorema se cumple evidentemente porque la matriz de T será (1) o (-1).

Supongamos ahora que el teorema se cumple cuando $\dim V < n$; y sea $\dim V = n$.

El polinomio minimal ϕ_T es un producto de factores de la forma $(x - \lambda)$ o de la forma $(x - a)^2 + b^2$, donde λ, a, b son números reales y $b \neq 0$. (Además, puesto que T es normal, se deduce del teorema 18.2.3 que esos factores son todos distintos entre sí. La demostración de este hecho puede hacerse como la del corolario del

teorema 18.2.3).

Podemos considerar los dos casos siguientes:

1. ϕ_T tiene un factor de la forma $x - \lambda$. En este caso λ es una raiz característica real, y por tanto debe ser $\lambda = 1$ o $\lambda = -1$, en virtud del teorema 20.3.2. Sea x_1 un vector propio unitario correspondiente a λ, es decir un vector unitario tal que $Tx_1 = \pm x_1$. El espacio V_1 generado por x_1 es pues invariante respecto de T. $V_1^{\perp}$ será también invariante respecto de T. Además es $V = V_1 \oplus V_1^{\perp}$. Por tanto, puesto que $dim\ V_1 = 1$, será $dim\ V_1^{\perp} = n - 1$. Podemos aplicar entonces la hipótesis de inducción a la restricción de T a $V_1^{\perp}$, la cual es evidentemente un endomorfismo ortogonal de este último subespacio. Obtendremos así una base ortonormal $\{x_2, ..., x_n\}$ de $V_1^{\perp}$, con respecto a la cual dicha restricción de T a $V_1^{\perp}$ es representada por una matriz de la forma descrita en el enunciado del teorema. Se comprueba entonces fácilmente que $\{x_1, x_2, ..., x_n\}$ es una base ortonormal de V respecto de la cual T se representa mediante una matriz de la forma deseada.

2. ϕ_T tiene únicamente factores de la forma $(x - a)^2 + b^2$, $b \neq 0$

Si se tiene presente el teorema 13.3.3, puede deducirse que para un factor $(x - a)^2 + b^2$ por lo menos, el subespacio

$$V_1 = Ker\ [(T - aI)^2 + b^2]$$

es diferente de $\{0\}$. Ese subespacio es invariante respecto de T, y la restricción de T a él tiene el polinomio minimal $(x - a)^2 + b^2$

Por tanto, si T_1 es la restricción de T al subespacio V_1, se tiene

$$(T_1 - aI)^2 + b^2 I = 0 \quad ,$$

de donde

$$\left(\frac{T_1 - aI}{b}\right)^2 = -I \ .$$

Si designamos por J a la transformación $\dfrac{T_1 - aI}{b}$ de V_1, entonces se podrá escribir

$$T_1 = aI + bJ$$

y

$$J^2 = -I \ .$$

Sea z_1 un vector arbitrario no nulo de V_1, y escribamos

$$z_2 = Jz_1 \quad . \qquad (z_2 \in V_1)$$

El subespacio V_2 de V_1, generado por z_1 y z_2, es invariante respecto de T porque

$$Tz_1 = T_1 z_1 = (aI + bJ)(z_1) = az_1 + bz_2 \ ,$$

$$Tz_2 = T_1 z_2 = (aI + bJ)(z_2) = az_2 + bJz_2 = az_2 - bz_1 \ .$$

Podemos comprobar ahora que z_1 y z_2 son *l.i.* En efecto supongamos que se cumple la relación

$$\mu z_1 + \nu z_2 = 0 \ , \qquad (20.3.3)$$

donde μ y ν son números reales. Aplicando a ambos miembros el endomorfismo J resulta

$$\mu z_2 - \nu z_1 = 0. \qquad (20.3.4)$$

De (20.3.3) y (20.3.4) se deduce que $(\mu^2 + \nu^2)z_1 = 0$, y puesto que $z_1 \neq 0$, debe ser $\mu^2 + \nu^2 = 0$, o sea $\mu = \nu = 0$. Por tanto z_1 y z_2 son *l.i.* Resulta así que $dim\ V_2 = 2$. Del teorema 20.3.1. se sigue que existe una base ortonormal $\{e_1 , e_2\}$ de V_2, con respecto a la cual la restricción de T a V_2 se representa por la matriz

$$\begin{pmatrix} cos\ \theta & -sen\ \theta \\ sen\ \theta & cos\ \theta \end{pmatrix} \ ,$$

donde $a = cos\ \theta$ y $b = sen\ \theta$.

(Obsérvese que el otro caso del teroema 20.3.1 no puede tener lugar pues T no tiene a $+1$ ni a -1 como varlores propios). Tenemos ahora que

$$V = V_2 \oplus V_2^{\perp} \ ,$$

de donde resulta que $dim\ V_2^{\perp} = n - 2$. Podemos aplicar entonces la hipótesis de inducción a la restricción de T a $V_2^{\perp}$, la cual es también ortogonal. Se obtiene así una base ortonormal de V que tiene la propiedad del enunciado. ∎

Corolario. Para cada matriz ortogonal M de orden n existe una matriz ortogonal S tal que SMS^{-1} tiene la forma de la matriz del enunciado del teorema 20.3.3. Esa forma se llama forma *canónica* de la matriz ortogonal M.

Ejercicios

20.3.1. Dado el espacio euclideano V de dimensión n, si $V^{\mathbb{C}}$ designa a la complejificación de V (sección 12.6), cada uno de cuyos elementos, por tanto, se expresa de una única manera en la forma $x + iy$, donde x e y son vectores de V, entonces $V^{\mathbb{C}}$ es un espacio unitario respecto del producto interno definido por la fórmula

$$(x_1 + iy_1 , x_2 + iy_2) = [(x_1 , x_2) + (y_1 , y_2)] + i[(y_1 , x_2) - (x_1 , y_2)]$$

$V^{\mathbb{C}}$ se llama *extensión unitaria del espacio euclideano V.*

20.3.2. Dado un endomorfismo T de V, probar que existe el endomorfismo $T^{\mathbb{C}}$ de $V^{\mathbb{C}}$ tal que

$$T^{\mathbb{C}}(x + iy) = Tx + iTy \quad .$$

$T^{\mathbb{C}}$ se llama extensión de T a $V^{\mathbb{C}}$ y coincide con la extensión que fué definida en la sección 12.5

Probar además las relaciones siguientes

1. $(T_1 T_2)^{\mathbb{C}} = T_1^{\mathbb{C}} T_2^{\mathbb{C}}$
2. $(T_1 + T_2)^{\mathbb{C}} = T_1^{\mathbb{C}} + T_2^{\mathbb{C}}$
3. Si $T_1^{\mathbb{C}} = T_2^{\mathbb{C}}$, entonces $T_1 = T_2$
4. $\widetilde{T^{\mathbb{C}}} = (\widetilde{T})^{\mathbb{C}}$
5. Si T es normal, entonces también lo es $T^{\mathbb{C}}$ y reciproca

mente

6. Si T es un automorfismo de V, $T^{\mathbb{C}}$ es un automorfismo de $V^{\mathbb{C}}$ e inversamente, y se tiene

$$(T^{-1})^{\mathbb{C}} = (T^{\mathbb{C}})^{-1} .$$

7. Si $p(x)$ es un polinomio con coeficientes reales, entonces es

$$p(T^{\mathbb{C}}) = (p(T))^{\mathbb{C}} .$$

20.3.3. Si $\{e_1, \ldots, e_n\}$ es una base ortonormal de V, entonces es también una base ortonormal de $V^{\mathbb{C}}$. La conjugación en $V^{\mathbb{C}}$, introducida mediante la definición dada en la sección 17.5, respecto de una base ortonormal $\{e_i\}$ de $V^{\mathbb{C}}$ perteneciente a V, coincide con la conjugación introducida en el teorema 12.6.1, y es independiente de la base ortonormal $\{e_i\}$ elegida.

20.3.4. Las matrices de T y de $T^{\mathbb{C}}$ respecto de la base ortonormada $\{e_i\}$ de V son reales e idénticas. Además es $f_T \equiv f_{T^{\mathbb{C}}}$

20.3.5. Los valores propios de T son los valores propios reales de $T^{\mathbb{C}}$. Si todos los valores propios de $T^{\mathbb{C}}$ son reales, entonces es

$$\phi_{T^{\mathbb{C}}} \equiv \phi_T .$$

20.3.6. Si T es simétrico (ortogonal) entonces $T^{\mathbb{C}}$ es hermitiano (unitario) e inversamente.

20.3.7. Supuesto que se ha demostrado ya que la matriz de un endomorfismo hermitiano (unitario) respecto de una base ortonormal cualquiera es hermitiana (unitaria) probar, valiéndose de los resultados de los ejercicios anteriores, que la matriz de un endomorfismo simétrico (ortogonal) respecto de una base ortonormal cualquiera es una matriz simétrica (ortogonal).

CAPITULO VI

ESPACIOS PUNTUALES AFINES Y EUCLIDEANOS

En el presente capítulo vamos a tratar acerca de conjuntos de elementos provistos de una estructura cuyo modelo es la del espacio ordinario de la geometría, considerado como conjunto de puntos. Un concepto esencial para nuestro estudio es el de espacio vectorial de dimensión finita, y en particular el de espacio vectorial euclideano.

§ 21. El espacio puntual afín

21.1. Espacio puntual afín.

Definición 21.1.1. Se llama *espacio puntual afín* (o simplemente *espacio afín*) sobre un campo dado $\mathbb{F}$, a una terna (E, V, Φ) constituida por un conjunto E, no vacío, cuyos elementos se llaman *puntos* del espacio, un espacio vectorial V *de dimensión finita* sobre $\mathbb{F}$, que se llama *espacio vectorial asociado al espacio afín o espacio de diferencias*, y una aplicación

$$\Phi : E \times E \longrightarrow V ,$$

que a cada par ordenado (A, B) de puntos le hace corresponder un vector de V, que se llama *vector diferencia* del par de puntos, y que se designa por $\Phi(A, B)$ o por $\overrightarrow{AB}$, de manera que se cumplen las condiciones siguientes:

a) Dados tres puntos cualesquiera A, B, C, tiene lugar que

$$\overrightarrow{AB} = \overrightarrow{AC} + \overrightarrow{CB} ,$$

b) Si A es un punto cualquiera, a cada vector $v \in V$ le corresponde un punto y solo uno, B, tal que

$$\overrightarrow{AB} = v.$$

Si la dimensión de V es n se dice que el espacio afín es de *dimensión* n.

Si $\mathbb{F} = \mathbb{R}$ el espacio afín se llama *espacio afín real* , y si $\mathbb{F} = \mathbb{C}$, se llama *espacio afín complejo*.

Ordinariamente designaremos al espacio afín (E, V, Φ) por el símbolo E, pero deberá sobreentenderse que V y Φ son bien determinados. Si la dimensión del espacio es n escribiremos

$$dim\ E = n\ .$$

De la precedente definición se deducen inmediatamente las siguientes consecuencias:

c) Si B es un punto cualquiera, entonces es $\overrightarrow{BB} = 0$.

Basta para probar esta afirmación suponer $C = B$ en la relación de la propiedad a).

d) Cualesquiera que sean los puntos A y C se cumple que

$$\overrightarrow{CA} = -\overrightarrow{AC}\ .$$

Resulta este hecho si se supone $B = A$ en a).

e) Dados cuatro puntos A, B, A', B', si se cumple la relación

$$\overrightarrow{AB} = \overrightarrow{A'B'}\ ,$$

entonces se tiene que

$$\overrightarrow{AA'} = \overrightarrow{BB'}\ .$$

En efecto,

$$\overrightarrow{AB'} = \overrightarrow{AB} + \overrightarrow{BB'} = \overrightarrow{AA'} + \overrightarrow{A'B'} = \overrightarrow{AA'} + \overrightarrow{AB}\ ,$$

de donde se sigue que $\overrightarrow{AA'} = \overrightarrow{BB'}$

21.2. Sistemas de coordenadas cartesianas en un espacio afín.

Definición 21.2.1. Se llama *Sistema de coordenadas cartesianas* en un espacio afín E de dimensión n a un conjunto

$$S = (O, \{e_1, \ldots, e_n\}) = (O, \{e_i\}),$$

formado por un punto O que se llama *origen* del sistema y por una base $\{e_1, \ldots, e_n\} = \{e_i\}$ del espacio vectorial asociado V. La base $\{e_i\}$ se llama *base del sistema*.

En lo que sigue los sistemas de coordenadas cartesianas serán llamados simplemente sistemas de coordenadas por cuanto no consideraremos otras clases de sistemas de coordenadas.

Si A es un punto cualquiera del espacio E y

$$\overrightarrow{OA} = \sum_{i=1}^{n} x^i e_i$$

es la representación del vector $\overrightarrow{OA}$ en la base $\{e_i\}$, los escalares x^i, es decir las componentes de $\overrightarrow{OA}$ respecto de esa base, se llaman *coordenadas* del punto A respecto del sistema S de coordenadas.

Corrientemente, dado un sistema de coordenadas $(O, \{e_i\})$ del espacio E, identificaremos a cada punto A con el correspondiente vector $\overrightarrow{OA}$. Así por ejemplo, si x designa al vector $\overrightarrow{OA}$, podremos designar también por x al punto A. De esta manera el punto O, origen del sistema de coordenadas se identifica con el vector nulo. Debe observarse que esa correspondencia cambia si se cambia el sistema de coordenadas.

Dado un sistema de coordenadas $(O, \{e_i\})$, el conjunto de los puntos λe_i, donde λ es un escalar arbitrario, se llama *eje de coordenadas* ligado al vector e_i. Por tanto hay n ejes de coordenadas que únicamente tienen en común al punto O. El eje de coordenadas ligado al vector e_i está formado por todos los puntos del espacio cuyas coordenadas son nulas salvo la coordenada x^i. Por ese motivo dicho eje se llama también eje Ox^i o eje x^i y el sistema de coordenadas se designa por (x^i).

Si se designa por P_i al punto del eje Ox^i tal que $\overrightarrow{OP_i} = e_i$ entonces es evidente que el conjunto de puntos $\{O, P_1, \ldots, P_n\}$ define al sistema de coordenadas $(O, \{e_i\})$ por cuanto el primero de esos puntos es el origen, y los vectores $\overrightarrow{OP_i}$ son los vectores de bases. En general, un conjunto $\{O, P_1, \ldots, P_n\}$ formado por $n + 1$ puntos,

tales que los vectores $\overrightarrow{OP_1}, \ldots, \overrightarrow{OP_n}$ son linealmente independientes, definen a un sistema de coordenadas cuyo origen es O y cuya base está formada por los vectores $e_i = \overrightarrow{OP_i}$. En adelante determinaremos algunas veces un sistema de coordenadas mediante un conjunto ordenado de $n + 1$ puntos, con arreglo a las consideraciones anteriores.

Cuando el espacio E es real y tiene 1, 2 o 3 dimensiones, es usual designar las coordenadas por x, (x, y) y (x, y, z) respectivamente.

Sean (x_1^i) las coordenadas de un punto $A = x_1$ y (x_2^i) las de un punto $B = x_2$. Las componentes del vector $\overrightarrow{AB}$ serán entonces $(x_2^i - x_1^i)$, en virtud de que

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = \overrightarrow{OB} - \overrightarrow{OA} = x_2 - x_1 .$$

Por consiguiente $x_2 - x_1$ designa al punto C tal que

$$\overrightarrow{OC} = x_2 - x_1 .$$

En la figura 21.2.1 se ilustra lo dicho anteriormente, en el caso en que el espacio afín que se considera es el espacio real de tres dimensiones interpretado geométricamente en la forma usual.

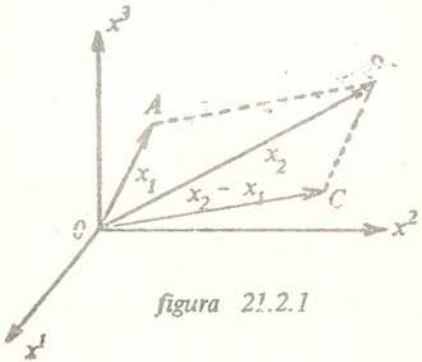


figura 21.2.1

Ejercicios

21.2.1. Calcular las coordenadas de los puntos $O, P_1, \ldots, P_n$ respecto del sistema de coordenadas $\{O, P_1, \ldots, P_n\}$.

***21.2.2.** [Se llama *espacio afín cartesiano real* de dimensión n al espacio afín definido por la terna $(\mathbb{R}^n, \mathbb{R}^n, \Phi)$, tal que

$$\Phi(x, y) = y - x \quad , \qquad x, y \in \mathbb{R}^n .$$

Los puntos de ese espacio son los vectores de $\mathbb{R}^n$ y el es-

pacio afín definido se designa también por $\mathbb{R}^n$. En forma semejante se define el *espacio afín cartesiano complejo* $\mathbb{C}^n$.

El sistema de coordenadas del espacio afín cartesiano, cuyo origen es el vector nulo y cuya base es la base canónica, se llama *sistema canónico de coordenadas*] .

Calcular las coordenadas del punto $P = (a^1, \ldots, a^n)$ del espacio afín $\mathbb{R}^n$ respecto del sistema de coordenadas $(P_o, \{e_i\})$ donde P_o es un punto cualquiera y $\{e_i\}$ es la base canónica. Considerar, en particular, el caso en que $(P_o, \{e_i\})$ es el sistema canónico de coordenadas.

Demostrarr luego que si $P_o = (p_o^1, \ldots, p_o^n)$ y si

$$\overrightarrow{P_o P_i} = (p_i^1 - p_o^1, \ldots, p_i^n - p_o^n), \quad i = 1, \ldots, n,$$

entonces la condición necesaria y suficiente para que

$$\{P_o, P_1, \ldots, P_n\}$$

sea un sistema de coordenadas, es decir para que los vectores $\overrightarrow{P_o P_1}, \ldots, \overrightarrow{P_o P_n}$ sean linealmente independientes es que

$$\begin{vmatrix} p_o^1 & p_1^1 & \ldots & p_n^1 \\ \cdots & \cdots & \cdots & \cdots \\ p_o^n & p_1^n & \ldots & p_n^n \\ 1 & 1 & \ldots & 1 \end{vmatrix} \neq 0 .$$

*21.2.3. Demostrar que si $P_1, \ldots, P_k$ son puntos cualesquiera de un espacio afín, se tiene

$$\overrightarrow{P_1 P_2} + \overrightarrow{P_2 P_3} + \ldots + \overrightarrow{P_{k-1} P_k} + \overrightarrow{P_k P_1} = 0.$$

21.3. Cambio de sistemas de coordenadas cartesianas.

Consideremos en el espacio afín E de dimensión n los sistemas de coordenadas $S = (O, \{e_i\})$ y $S' = (O, \{e_{i'}\})$, y sean (I,6.10)

$$e_i = \sum_{j'=1'}^{n'} A_i^{j'} e_{j'} \quad , \qquad e_{j'} = \sum_{i=1}^{n} A_{j'}^{i} e_i$$

las ecuaciones de cambio de las bases $\{e_i\}$ y $\{e_{i'}\}$ de V.

Supongamos que x designa al vector $\overrightarrow{OA}$, cuyas componentes respecto de la base $\{e_i\}$ son x^i, y respecto de la base $\{e_{i'}\}$ son $x^{i'}$. Designemos por y al vector $\overrightarrow{O'A}$ cuyas componentes respecto de la base $\{e_i\}$ son y^i, y respecto de la base $\{e_{i'}\}$ son $y^{i'}$. Es claro que x^i son las coordenadas del punto A respecto del sistema S, e $y^{i'}$ son sus coordenadas respecto del sistema S'. Designemos además por a^i a las componentes del vector $\overrightarrow{OO'}$ respecto de la base $\{e_i\}$, y por $a^{i'}$ a las componentes del vector $\overrightarrow{O'O}$ respecto de la base $\{e_{i'}\}$.

Las relaciones

$$\overrightarrow{OA} = \overrightarrow{OO'} + \overrightarrow{O'A} \quad , \quad \overrightarrow{O'A} = \overrightarrow{O'O} + \overrightarrow{OA} \tag{21.3.1}$$

implican respectivamente (sección I, 6.10)

$$\left.\begin{aligned} x^i &= a^i + y^i = a^i + \sum_{j'=1'}^{n'} A^i_{j'} y^{j'} \\ y^{j'} &= a^{j'} + x^{j'} = a^{j'} + \sum_{i=1}^{n} A^{j'}_i x^i \end{aligned}\right\} \tag{21.3.2}$$

que son las ecuaciones de cambio de coordenadas de los sistemas S y S'. Entre los escalares a^i y los escalares $a^{j'}$ se cumplen evidentemente las relaciones

$$a^{j'} = -\sum_{i=1}^{n} A^{j'}_i a^i \quad , \quad a^i = -\sum_{j'=1'}^{n'} A^i_{j'} a^{j'} \, . \tag{21.3.3}$$

Las escuaciones (21.3.2) pueden expresarse convenientemente en forma matricial, si se introducen las matrices siguientes:

$$A = (A^{j'}_i) \, , \; A' = (A^i_{j'}) \, , \; X = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \, , \; Y' = \begin{pmatrix} y^{1'} \\ \vdots \\ y^{n'} \end{pmatrix} \, , \; \alpha = \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$$

$$\alpha' = \begin{pmatrix} a^{1'} \\ \vdots \\ a^{n'} \end{pmatrix} \qquad (21.3.4)$$

Se tiene entonces

$$\left.\begin{aligned} Y' &= AX + \alpha' \\ X &= A'Y' + \alpha \end{aligned}\right\} , \qquad (21.3.5)$$

y las ecuaciones (21.3.3) toman la forma

$$\alpha' = -A\alpha \quad , \quad \alpha = A'\alpha' . \qquad (21.3.6)$$

Se ve que cada una de las ecuaciones (21.3.5) se deduce de la otra por cuanto se cumple la relación

$$AA' = A'A = 1.$$

Además, dichas ecuaciones pueden escribirse en la forma

$$\left.\begin{aligned} Y' &= A(X - \alpha) \\ X &= A'(Y' - \alpha') \end{aligned}\right\} \qquad (21.3.7)$$

Observaremos por último que si introducimos las matrices

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix} = \begin{pmatrix} X \\ 1 \end{pmatrix}, \quad \begin{pmatrix} y^{1'} \\ \vdots \\ y^{n'} \\ 1 \end{pmatrix} = \begin{pmatrix} Y' \\ 1 \end{pmatrix}, \quad \begin{pmatrix} A & \alpha' \\ 0 & 1 \end{pmatrix} \text{ y } \begin{pmatrix} A' & \alpha \\ 0 & 1 \end{pmatrix}$$

las relaciones (21.3.2) y (21.3.7) pueden escribirse en la forma

$$\begin{pmatrix} y^{1'} \\ \vdots \\ x^n \\ 1 \end{pmatrix} = \begin{pmatrix} A & \alpha' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix} = \begin{pmatrix} A' & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{1'} \\ \vdots \\ y^{n'} \\ 1 \end{pmatrix} \qquad (21.3.8)$$

o en la forma

$$\left.\begin{aligned}\begin{pmatrix} Y' \\ 1 \end{pmatrix} &= \begin{pmatrix} A & \alpha' \\ 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ 1 \end{pmatrix} \\ \begin{pmatrix} X \\ 1 \end{pmatrix} &= \begin{pmatrix} A' & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} Y' \\ 1 \end{pmatrix}\end{aligned}\right\} \qquad (21.3.9)$$

La matriz

$$\overline{A} = \begin{pmatrix} A & \alpha' \\ 0 & 1 \end{pmatrix}$$

se llama *matriz de cambio* del sistema S en el sistema S'

Ejercicios

*21.3.1. Sea $\mathbb{R}^3$ el espacio cartesiano real afín de tres dimensiones. Consideremos en él dos sistemas de coordenadas de orígenes O y O' respectivamente. Designemos por x,y,z a las coordenadas del primer sistema, y por x', y', z' a las del segundo sistema. Sea $\mathfrak{B}$ la base asociada al primer sistema y $\mathfrak{B}'$ a la asociada al segundo sistema. Sean (x_1 , y_1 , z_1) las coordenadas de O' respecto del primer sistema, y sean (p , q , r) , (p', q', r') y (p'', q'', r'') las componentes de los vectores de la base $\mathfrak{B}'$ respecto de la base $\mathfrak{B}$. Obtener las ecuaciones de cambio de coordenadas de los dos sistemas.

*21.3.2. Con la hipótesis del ejercicio anterior, obtener las ecuaciones de cambio de coordenadas en el caso en que las dos bases $\mathfrak{B}$ y $\mathfrak{B}'$ son idénticas.

21.3.3. Consideremos en el espacio afín cartesiano complejo $\mathbb{C}^2$ los puntos $P_o = (1 , i)$, $P_1 = (0 , 1)$ y $P_2 = (1 + i , 1)$.

21. Espacio puntual afín.

Demostrar que $\{P_0, P_1, P_2\}$ es un sistema de coordenadas. Hallar las relaciones entre las coordenadas de los puntos del espacio referidas a este sistema y las coordenadas de los mismos puntos referido al sistema canónico de coordenadas.

21.4. Aplicaciones afines.

Definición 21.4.1. Se dice que una aplicación $L: E_1 \to E_2$ de un espacio afín E_1 sobre $\mathbb{F}$, asociado al espacio vectorial V_1, en un espacio afín E_2 sobre el mismo campo, asociado al espacio vectorial V_2, es una *aplicación afín de E_1 en E_2*, si existe la transformación lineal $T: V_1 \to V_2$ tal que, cualesquiera que sean los puntos P y Q de E_1 se cumple que

$$\overrightarrow{L(P)L(Q)} = T(\overrightarrow{PQ}) . \qquad (21.4.1)$$

Se dice entonces que T es el *homomorfismo asociado* con L. Si L es además una biyección, se llama *isomorfismo afín de E_1 sobre E_2*.

Interés particular tendrá para nosotros el caso en que

$$E_1 = E_2 = E.$$

Si esto sucede se tendrá que $V_1 = V_2 = V$. La aplicación $L: E \to E$ asociada a $T: V \to V$ se llama *aplicación afín de E*, y T se dice que es el *endomorfismo asociado a L*. Si una aplicación afín L de E es una *biyección* se dice que es una *transformación afín de E*.

El teorema siguiente asegura la existencia de aplicaciones afines y prueba que una aplicación afín es bien definida cuando es dada la imagen, por ella, de uno de los puntos de E_1, y también es dado el homomorfismo asociado $T: V_1 \to V_2$.

Teorema 21.4.1. Dados los puntos $P_1 \in E_1$ y $P_2 \in E_2$, y el homomorfismo $T: V_1 \to V_2$, existe una y solo una aplicación afín $L: E_1 \to E_2$ tal que

$$L(P_1) = P_2 \qquad (21.4.2)$$

cuyo homomorfismo asociado es T.

Demostración. Dado un punto caulquiera $P_1' \in E_1$ es bien definido

el punto $P_2' \in E_2$ tal que

$$T(\overrightarrow{P_1P_1'}) = \overrightarrow{P_2P_2'} \quad . \tag{21.4.3}$$

Podemos definir entonces la aplicación $L: E_1 \to E_2$ tal que

$$L(P_1') = P_2' \quad . \tag{21.4.4}$$

L cumple por tanto la condición

$$T(\overrightarrow{P_1P_1'}) = \overrightarrow{P_2L(P_1')} \quad . \tag{21.4.5}$$

Demostraremos que L es una aplicación afín que cumple las condiciones del enunciado. En efecto, si $P_1' = P_1$, entonces es $\overrightarrow{P_1P_1'} = 0$, y por tanto, según (21.4.3), debe ser $\overrightarrow{P_2P_2'} = 0$, o sea $P_2' = P_2$. Luego $L(P_1) = P_2$, y se cumple (21.4.2). Además, si P y Q son puntos cualesquiera de E_1, se tiene, haciendo uso de (1.4.5),

$$\overrightarrow{L(P)L(Q)} = \overrightarrow{L(P)L(P_1)} + \overrightarrow{L(P_1)L(Q)} = \overrightarrow{L(P)P_2} + \overrightarrow{P_2L(Q)}$$

$$= \overrightarrow{P_2L(Q)} - \overrightarrow{P_2L(P)} = T(\overrightarrow{P_1Q}) - T(\overrightarrow{P_1P}) = T(\overrightarrow{PQ}).$$

Por consiguiente L es una aplicación afín que cumple las condiciones del enunciado.

Probemos ahora que L es la única aplicación afín que tiene las propiedades requeridas. Si $L': E_1 \to E_2$ también tuviera esas propiedades, entonces, cualquiera que fuera el punto P de E_1 se tendría

$$T(\overrightarrow{P_1P}) = \overrightarrow{L'(P_1)L'(P)} = \overrightarrow{P_2L'(P)} = \overrightarrow{P_2L(P)} \quad .$$

De la definición de espacio afín se deduciría que

$$L(P) = L'(P)$$

para todo $P \in E_1$, y por tanto sería $L = L'$. ∎

Corolario. La condición necesaria y suficiente para que la aplicación $L: E_1 \to E_2$ sea la aplicación afín que aplica al punto $P_1 \in E_1$ en el punto $P_2 \in E_2$, y que está asociada al homomorfismo $T: V_1 \to V_2$, es que cualquiera que sea el punto $P \in E_1$ se cumpla la condición

$$\overrightarrow{P_2L(P)} = T(\overrightarrow{P_1P})$$

Demostración. La condición es obviamente necesaria por la definición de aplicación afín. La condición es suficiente porque si L' designa a la aplicación afín de E_1 en E_2 que aplica a P_1 en P_2 y que está asociada a T, aplicación que existe en virtud del teorema anterior, entonces se tiene, cualquiera que sea $P \in E_1$,

$$T(\overrightarrow{P_1P}) = \overrightarrow{P_2L(P)} = \overrightarrow{P_2L'(P)} ,$$

de donde se deduce que $L(P) = L'(P)$, y por tanto $L = L'$. ∎

Teorema 21.4.2. Sean E_1 y E_2 dos espacios afines sobre el campo $\mathbb{F}$, asociado a los espacios vectoriales V_1 y V_2 respectivamente. La condición necesaria y suficiente para que la aplicación $L: E_1 \to E_2$ sea una aplicación afín asociada al homomorfismo $T: V_1 \to V_2$ es que dados los puntos arbitrarios $O \in E_1$ y $O' \in E_2$, exista un vector $b \in V_2$ tal que, si se escribe

$$x = \overrightarrow{OP} \quad , \quad y = \overrightarrow{O'L(P)} ,$$

donde P es un punto cualquiera de E_1, se cumpla la relación

$$y = T(x) + b .$$

se tiene entonces que $b = \overrightarrow{O'L(O)}$.

Demostración. La condición es necesaria porque si L es una aplicación afín asociada a T se tiene

$$y = \overrightarrow{O'L(P)} = \overrightarrow{L(O)L(P)} + \overrightarrow{O'L(O)} = T(\overrightarrow{OP}) + b = T(x) + b.$$

Probemos ahora que la condición es suficiente. Por hipótesis se cumple para todo $P \in E_1$ la relación

$$\overrightarrow{O'L(P)} = T(\overrightarrow{OP}) + b.$$

En particular, haciendo $P = O$ se tendrá

$$b = \overrightarrow{O'L(O)} ,$$

y por tanto, para cualquier $P \in E_1$,

$$T(\overrightarrow{OP}) = \overrightarrow{O'L(P)} - b = \overrightarrow{O'L(P)} - \overrightarrow{O'L(O)} = \overrightarrow{L(O)L(P)}.$$

En virtud del corolario del teorema anterior L será la aplicación afín asociada a T que aplica a O en $L(O)$. ∎

Corolario. Sea E un espacio afín asociado al espacio vectorial V. La condición necesaria y suficiente para que la aplicación $L : E \to E$ sea una aplicación afín de E asociada al endomorfismo $T : V \to V$ es que, dado un punto arbitrario $O \in E$, exista un vector $b \in V$ tal que, cualquiera que sea el punto $P \in E$, si se escribe

$$x = \overrightarrow{OP} \qquad , \qquad y = \overrightarrow{OL(P)} \ ,$$

se cumple la condición

$$y = T(x) + b \ .$$

Se tiene entonces $b = \overrightarrow{OL(O)}$.

Demostración. El corolario es una consecuencia inmediata del teorema anterior. ∎

Definición 21.4.2. Dada la aplicación afín $L : E \to E$ y el punto $O \in E$, la ecuación

$$y = T(x) + b \qquad (21.4.6)$$

que se cumple cualquiera que sea $P \in E$ cuando se escribe

$$x = \overrightarrow{OP} \ , \ y = \overrightarrow{OL(P)} \quad \text{y} \quad b = \overrightarrow{OL(O)} \ ,$$

se llama *ecuación vectorial de L referida al punto O.*

Teorema 21.4.3. La condición necesaria y suficiente para que la aplicación afín $L : E_1 \to E_2$ sea un isomorfismo afín es que el homomorfismo asociado $T : V_1 \to V_2$ sea un isomorfismo de los espacios vectoriales V_1 y V_2 .

Demostración. Supongamos que L es un isomorfismo afín es decir que es una biyección de E_1 sobre E_2 . Para probar que T es un isomorfismo basta probar que también es una biyección. Si fuera

$$T(x) = T(x') \ ,$$

entonces, elegido un punto cualquiera $O \in E_1$, si $\overrightarrow{OP} = x$ y $\overrightarrow{OP'} = x'$, se tendrá $T(\overrightarrow{OP}) = T(\overrightarrow{OP'})$, y por tanto $\overrightarrow{L(O)L(P)} = \overrightarrow{L(O)L(P')}$, de donde resulta $L(P) = L(P')$. Por ser L una biyección se tendrá que $P = P'$. Luego $x = \overrightarrow{OP} = \overrightarrow{OP'} = x'$. Resulta de esta manera que T es una inyección. Probemos ahora que es una sobreyección. Dado $y \in V_2$, y elegidos los puntos Q y Q' de E_2 de manera que sea $\overrightarrow{QQ'} = y$, existirán P y P' de E_1 tales que $L(P) = Q$ y $L(P') = Q'$. Para el vector $x = \overrightarrow{PP'}$ de V_1 se cumplirá entonces que

$$T(x) = T(\overrightarrow{PP'}) = \overrightarrow{L(P)L(P')} = \overrightarrow{QQ'} = y$$

y por tanto L es sobreyectiva.

Supongamos ahora que T es un isomorfismo. Probemos que L es biyectiva y por tanto es un isomorfismo afín. L es inyectiva porque si fuera $L(P) = L(P')$, se tendría

$$T(\overrightarrow{PP'}) = \overrightarrow{L(P)L(P')} = 0$$

y por tanto $\overrightarrow{PP'} = 0$, o sea $P = P'$. L es sobreyectiva porque dado $Q \in E_2$, si O es un punto cualquiera de E_1, y designamos por $y \in V_2$ al vector $\overrightarrow{L(O)Q}$, entonces existirá un vector $x \in V_1$ tal que $T(x) = y = \overrightarrow{L(O)Q}$. Sea $\overrightarrow{OP} = x$. Se tendrá entonces que

$$T(x) = \overrightarrow{L(O)L(P)} \quad ,$$

y por consiguiente $L(P) = Q$ ■

Definición 21.4.3. Una aplicación afín de un espacio afín E se llama *traslación* si el endomorfismo asociado es el isomorfismo identidad de V.

El teorema 21.4.3 implica que una traslación es una transformación afín, es decir que es biyectiva.

En virtud del corolario del teorema 21.4.1, la condición necesaria y suficiente para que una aplicación $L : E \to E$ sea una traslación es que, dado un punto arbitrario $O \in E$, exista el vector $b \in V$ tal que, cualquiera que sea $P \in E$, si se designa por x al vector $\overrightarrow{OP}$ y por y al vector $\overrightarrow{OL(P)}$, se cumpla la relación.

$$y = x + b \quad , \quad b = \overrightarrow{OL(O)} \qquad (21.4.7)$$

Esta es la *ecuación vectorial de la traslación* referida al punto O. Sin embargo, es fácil comprobar que si se cambia el punto O por otro punto cualquiera O' la ecuación no varía, y el vector b permanece invariante. Dicho vector, que caracteriza a la traslación, se llama *vector de la traslación.*

Teorema 21.4.4. La condición necesaria y suficiente para que la aplicación $L : E \to E$ sea una traslación es que exista un vector $b \in V$ tal que, cualquiera que sea $P \in E$, se cumpla que

$$\overrightarrow{PL(P)} = b \qquad (21.4.8)$$

Demostración. Es fácil comprobar que la relación (21.4.8) es equivalente a la relación (21.4.7) la cual, como acabamos de ver, es necesaria y suficiente para que L sea una traslación de vector b. ∎

Definición 21.4.4. Una aplicación afín L de un espacio afín E se llama *aplicación centroafín de centro O,* si O es un punto fijo de L, es decir si $L(O) = O$.

Teorema 21.4.5. La condición necesaria y suficiente para que una aplicación $L : E \longrightarrow E$ sea centroafín es que exista un punto O tal que L admita la ecuación vectorial

$$y = T(x) \qquad , \qquad (x = \overrightarrow{OP} \ , \ y = \overrightarrow{OL(P)}) \ .$$

Demostración. Esta proposición resulta inmediatamente del corolario del teorema 21.4.2. ∎

Teorema 21.4.6. Dada una aplicación afín L del espacio afín E, L puede ser expresado de una sola manera en la forma

$$L = \theta \circ L_o = \theta L_o \qquad (21.4.9)$$

donde L_o es una aplicación centroafín de centro O y θ es una traslación.

Demostración. Demostraremos primero la unicidad. Supongamos que existen θ y L_o que cumplen las propiedades del enunciado. Puesto que

$$L(O) = \theta(L_o(O)) \ .$$

el vector de la traslación θ será necesariamente igual a

$$\overrightarrow{O\theta(O)} = \overrightarrow{OL(O)} ,$$

y por tanto la traslación θ es bien determinada. Además, puesto que θ es biyectiva, deberá ser

$$L_o = \theta^{-1} L ,$$

y por consiguiente L_o será también determinada.

Probemos ahora la existencia de las transformaciones θ y L_o que cumplen las condicones del enunciado. Sea $y = T(x) + b$ la ecuación vectorial de L referida al punto 0. Designemos por L_o a la aplicación centroafín cuya ecuación es $y = x + b$. Demostraremos entonces que se cumple la relación (21.4.9). Se tiene en efecto, cualquiera que sea P,

$$\overrightarrow{OL_o(P)} = \overrightarrow{T(OP)} = \overrightarrow{L(O)L(P)} ,$$

y por tanto

$$\overrightarrow{O\theta(L_o(P))} = \overrightarrow{OL_o(P)} + b = \overrightarrow{L(O)L(P)} + \overrightarrow{OL(O)} = \overrightarrow{OL(P)} ,$$

de donde se deduce

$$L(P) = \theta(L_o(P))$$

para todo P; lo que demuestra la relación (21.4.9). ■

Definición 21.4.5. Dado un sistema de coordenadas $(O, \{e_i\})$ en el espacio afín y una aplicación afín L de E, cuyo endomorfismo asociado $T : V \to V$ tiene la matriz M respecto de la base $\{e_i\}$ se llama matriz de L relativa a dicho sistema de coordenadas a la matriz

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

donde b representa aquí a la matriz columna cuyos términos son las componentes del vector $b = \overrightarrow{OL(O)}$ respecto de la base $\{e_i\}$.

Si designamos por $(x^1, \ldots, x^n)$ a las componentes del vector $x = \vec{OP}$ respecto de la base $\{e_i\}$ y por $(y^1, \ldots, y^{n'})$ a las compo-

nentes del vector $y = \overrightarrow{OL(P)}$ respecto de la misma base, entonces, la ecuación vectorial de L relativa a O,

$$y = T(x) + b$$

equivale a la ecuación matricial

$$\begin{pmatrix} y^1 \\ \vdots \\ y^n \\ 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix} \qquad (21.4.10)$$

en la que podemos considerar a $(x^1, \ldots, x^n)$ como las coordenadas de P en el sistema $(O, \{e_i\})$ y a $(y^1, \ldots, y^n)$ como las coordenadas de $L(P)$ en el mismo sistema. Se ve por consiguiente que, dada la matriz de L respecto de un sistema de coordenadas, L queda totalmente determinado.

La ecuación (21.4.10), que en forma más compacta puede escribirse

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}, \qquad (21.4.11)$$

se llama, *ecuación matricial* de L respecto del sistema de coordenadas $(O, \{e_i\})$, y muestra que dicha aplicación puede reducirse a la restricción del endomorfismo del espacio vectorial $\mathbb{R}^n$, cuya matriz es

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix},$$

al subconjunto $\{u \in \mathbb{R}^{n+1} \mid u^{n+1} = 1\}$.

Es conveniente notar por último la relación

$$det \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} = det\ M$$

que muestra que el determinante de la matriz de la transformación afín L es igual al determinante de la transformación afín asociada T, y por tanto no depende del sistema de coordenadas.

Teorema 21.4.7. Dadas las aplicaciones afines L_1 y L_2 del espacio E cuyas matrices respecto del sistema de coordenadas $(0, \{e_i\})$ son

$$\begin{pmatrix} M_1 & b_1 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} M_2 & b_2 \\ 0 & 1 \end{pmatrix},$$

respectivamente, la aplicación $L_2L_1 : E \to E$ es también una aplicación afín de E, cuya matriz respecto del sistema $(0, \{e_i\})$ es el producto

$$\begin{pmatrix} M_2 & b_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_1 & b_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} M_2M_1 & M_2b_1 + b_2 \\ 0 & 1 \end{pmatrix}$$

de las matrices de las aplicaciones dadas. Además si T_1 y T_2 son los endomorfismos asociados con L_1 y L_2, el endomorfismo asociado con L_2L_1 es T_2T_1 . (1)

Demostración. Sea

$$x = \overrightarrow{OP} \quad , \quad y = \overrightarrow{OL_1(P)} \quad , \quad z = O(L_2L_1)\overrightarrow{(P)}.$$

se tiene entonces

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} M_1 & b_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} , \begin{pmatrix} z \\ 1 \end{pmatrix} = \begin{pmatrix} M_2 & b_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} M_2 & b_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_1 & b_1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}$$

(1) Es facil probar de manera directa que T_2T_1 es el endomorfismo asociado a L_2L_1. Se tiene, en efecto, si T designa a ese endomorfismo y P, Q son puntos arbitrarios:

$$\begin{aligned} T(\overrightarrow{PQ}) &= \overrightarrow{(L_2L_1)P(L_2L_1)Q} = \overrightarrow{L_2(L_1(P))L_2(L_1(Q))} \\ &= T_2(\overrightarrow{L_1(P)L_1(Q)}) = T_2(T_1(\overrightarrow{PQ})) \\ &= (T_2T_1)(\overrightarrow{PQ}) \quad \blacksquare \end{aligned}$$

$$= \begin{pmatrix} M_2M_1 & M_2b_1 + b_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ 1 \end{pmatrix}$$

La matriz del endomorfismo asociado con L_2L_1 es M_2M_1, que es la matriz de T_2T_1. Luego este último es el endomorfismo asociado con L_2L_1. ■

Definición 21.4.6. Las transformaciones afines de un espacio E afín sobre el campo $\mathbb{F}$ constituyen un grupo, que designaremos por $A(E)$ el cual es isomorfo con el grupo de las matrices sobre $\mathbb{F}$ que tienen la forma

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

donde M es una matriz regular de orden n y b es una matriz columna de tipo $n \times 1$. La matriz de la transformación afín inversa de la transformación cuya matriz es

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

es

$$\begin{pmatrix} M^{-1} & -M^{-1}b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}^{-1}.$$

Dicho grupo de matrices se llama *grupo afín* y se designa por

$$A(n, \mathbb{F}).$$

Teorema 21.4.8. Cada matriz

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix}$$

puede escribirse de una sola manera como el producto de una matriz de la forma

$$\begin{pmatrix} I_n & \bar{b} \\ 0 & 1 \end{pmatrix}$$

donde I_n es la matriz unidad de orden n y $\bar{b}$ es una columna $n \times 1$, por una matriz de la forma

$$\begin{pmatrix} \overline{M} & 0 \\ 0 & 1 \end{pmatrix}$$

donde $\overline{M}$ es una matriz de orden n. Se tiene además que $\bar{b} = b$ y que $\overline{M} = M$.

Demostración. Se tiene en efecto que

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_n & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} \quad ;$$

y si

$$\begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_n & \bar{b} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \overline{M} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \overline{M} & \bar{b} \\ 0 & 1 \end{pmatrix},$$

entonces es $\overline{M} = M$ y $\bar{b} = b$. ∎

Observación. El teorema precedente es equivalente al teorema 21.4. 6 pues

$$\begin{pmatrix} I_n & b \\ 0 & 1 \end{pmatrix}$$

es la matriz de una traslación de vector b y

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

es la matriz de una aplicación centroafín de centro en el origen de

coordenadas.

Teorema 21.4.9. Una aplicación $L: E \to E$ es una aplicación afín de E si y solo si cualquiera que sea el escalar λ se cumple la propiedad siguiente:

Si $\overrightarrow{PR} = \lambda(\overrightarrow{PQ})$ entonces es $\overrightarrow{L(P)L(R)} = \lambda(\overrightarrow{L(P)L(Q)})$, cualesquiera que sean los puntos P, Q y R de E.

Demostración. Si L es una aplicación afín y T es el endomorfismo asociado, se tiene

$$T(\overrightarrow{PR}) = T(\lambda\overrightarrow{PQ}) = \lambda T(\overrightarrow{PQ})$$

y por tanto

$$\overrightarrow{L(P)L(R)} = \lambda\overrightarrow{L(P)L(Q)} .$$

Supongamos, inversamente, que se cumple esta última relación cualesquiera que sean $\lambda \in F$ y los puntos P, Q, R tales que $\overrightarrow{PR} = \lambda\overrightarrow{PQ}$. Definamos la aplicación $T: V \to V$ tal que

$$T(\overrightarrow{OP}) = \overrightarrow{L(O)L(P)} ,$$

donde O es un punto dado arbitrario de E y P es cualquiera. Probaremos que T es lineal. Dados los vectores arbitrarios

$$x = \overrightarrow{OP} , \quad y = \overrightarrow{OQ} ,$$

sea

$$\overrightarrow{OS} = \frac{x + y}{2} , \quad \overrightarrow{OR} = x + y .$$

Se tendrá entonces

$$\overrightarrow{PS} = \overrightarrow{OS} - \overrightarrow{OP} = \frac{y - x}{2} .$$

Las relaciones evidentes

$$\overrightarrow{OR} = 2\overrightarrow{OS} ,$$
$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = y - x = 2\overrightarrow{PS}$$

implican, en virtud de nuestra suposición que,

$$\overrightarrow{L(O)L(R)} = 2\overrightarrow{L(O)L(S)}$$

y

$$\overrightarrow{L(P)L(Q)} = 2\overrightarrow{L(P)L(S)} .$$

Luego

$$\begin{aligned} T(x + y) &= T(\overrightarrow{OR}) = \overrightarrow{L(O)L(R)} = 2\overrightarrow{L(O)L(S)} \\ &= 2[\overrightarrow{L(O)L(P)} + \overrightarrow{L(P)L(S)}] \\ &= \overrightarrow{L(O)L(P)} + \overrightarrow{L(O)L(P)} + 2\overrightarrow{L(P)L(S)} \\ &= \overrightarrow{L(O)L(P)} + \overrightarrow{L(O)L(P)} + \overrightarrow{L(P)L(Q)} \\ &= \overrightarrow{L(O)L(P)} + \overrightarrow{L(O)L(Q)} \\ &= T(\overrightarrow{OP}) + T(\overrightarrow{OQ}) \\ &= Tx + Ty. \end{aligned}$$

Además, si λ es un escalar cualquiera y $\lambda x = \lambda\overrightarrow{OP} = \overrightarrow{OP_1}$, se tiene

$$\overrightarrow{L(O)L(P_1)} = \lambda(\overrightarrow{L(O)L(P)}) = \lambda\overrightarrow{T(OP)} ,$$

y por tanto

$$T(\lambda x) = T(\overrightarrow{OP_1}) = \overrightarrow{L(O)L(P_1)} = \lambda\overrightarrow{T(OP)} ,$$

es decir que T es un endomorfismo de V.

Por otra parte, cualesquiera que sean los puntos P y Q se tiene

$$T(\overrightarrow{PQ}) = T(\overrightarrow{OQ}) - T(\overrightarrow{OP}) = \overrightarrow{L(O)L(Q)} - \overrightarrow{L(O)L(P)} = \overrightarrow{L(P)L(Q)}.$$

Por consiguiente L es una aplicación afín de E asociada a T. ■

Corolario. Una aplicación biyectiva L de un espacio afín E sobre sí mismo es una transformación afín de E si y solo si, cualquiera que sea el escalar λ, se cumple la propiedad del teorema 21.4.9.

Demostración. Este corolario es consecuencia del teorema anterior porque la condición necesaria y suficiente para que una aplicación biyectiva de E sobre sí mismo sea una transformación afín es que sea una aplicación afín (definición 21.4.1). ■

Vamos a dar término a la presente sección probando un teore

ma que establece la manera como cambia la matriz de una aplicación afín $L : E \to E$ cuando cambia el sistema de coordenadas. El resultado que se obtiene es análogo al que se refiere al cambio de la matriz de un endomorfismo de un espacio vectorial cuando cambia la base.

Teorema 21.4.10. Sea $\overline{M}$ la matriz que representa a la transformación afín $L : E \to E$ (definición 21.4.5) respecto de un sistema S de coordenadas. Si $\overline{A}$ es la matriz del cambio del sistema de coordenadas S en el sistema de coordenadas S', entonces la matriz de L respecto del sistema S' es

$$\overline{A}\,\overline{M}\,\overline{A}^{-1} .$$

Demostración. Dado un punto P cualquiera de E sean (x^i) sus coordenadas en el sistema S. Designemos por (x_1^i) a las coordenadas de $L(P)$ en el mismo sistema. Tendremos (fórmula (1.4.10))

$$\begin{pmatrix} X_1 \\ \\ 1 \end{pmatrix} = \overline{M} \begin{pmatrix} X \\ \\ 1 \end{pmatrix}$$

Designemos por (y^i) e (y_1^i) a las coordenadas en el sistema S' de los puntos P y $L(P)$ respectivamente. Tendremos

$$\begin{pmatrix} Y \\ \\ 1 \end{pmatrix} = \overline{A} \begin{pmatrix} X \\ \\ 1 \end{pmatrix}$$

e

$$\begin{pmatrix} Y_1 \\ \\ 1 \end{pmatrix} = \overline{A} \begin{pmatrix} X_1 \\ \\ 1 \end{pmatrix}$$

Por consiguiente se tendrá

$$\overline{A}^{-1} \begin{pmatrix} Y_1 \\ \\ 1 \end{pmatrix} = \overline{M}\,\overline{A}^{-1} \begin{pmatrix} Y \\ \\ 1 \end{pmatrix}$$

o sea

$$\begin{pmatrix} Y_1 \\ 1 \end{pmatrix} = \overline{A}\,\overline{M}\,\overline{A}^{-1} \begin{pmatrix} Y \\ 1 \end{pmatrix}$$

relación que demuestra el teorema. ■

Ejercicios

21.4.1. Hallar la ecuación matricial de la aplicación afín L tal que

$$L(0\,,\,0) = (1\,,\,2)$$
$$L(1\,,\,0) = (1\,,\,-1)$$
$$L(0\,,\,1) = (0\,,\,1)$$

referida al sistema canónico de coordenadas.

21.4.2. Resolver la misma cuestión anterior para la aplicación afín de $\mathbb{R}^3$, tal que

$$L(1\,,\,0\,,\,1) = (1\,,\,2\,,\,1)$$
$$L(0\,,\,1\,,\,1) = (-1\,,\,2\,,\,0)$$
$$L(1\,,\,0\,,\,0) = (0\,,\,3\,,\,4)$$
$$L(-1\,,\,1\,,\,2) = (1\,,\,-2\,,\,5)$$

21.4.3. Demostrar:

a) que las transformaciones centroafines de E de centro O constituyen un subgrupo del grupo $A(E)$ de los isomorfismos afines de E.

b) que las traslaciones constituyen un subgrupo de $A(E)$

c) que si θ es una traslación y L es una transformación afín cualquiera, entonces $L\theta L^{-1}$ es una traslación. Por con siguiente las traslaciones constituyen un subgrupo normal de $A(E)$

21.5. Subespacios afines. Consideremos aquí que E es un espacio afín asociado con un espacio vectorial V, y que $(E\,,\,V\,,\,\Phi)$ es la terna que le corresponde en virtud de la definición 21.1.1.

Definición 21.5.1. Se dice que un subconjunto E_1 del espacio afín E es un *subespacio (afín)* de E si existe un subespacio V_1 de V tal que la terna (E_1 , V_1 , Φ_1), donde Φ_1 es la restricción de Φ a $E_1 \times E_1$ es un espacio afín.

Si P es un punto de E_1 se dice que el subespacio E_1 pasa por el punto P.

Teorema 21.5.1. La condición necesaria y suficiente para que un subconjunto E_1 de E sea un subespacio afín de E es que, dado un punto cualquiera P de E_1, el conjunto de vectores

$$V_P = \{\overrightarrow{PQ} \mid Q \in E_1\}$$

sea un subespacio de V. V_P es entonces el espacio vectorial asociado al espacio afín E_1.

Demostración. Si E_1 es un subespacio de E, cuyo espacio vectorial asociado es V_1, entonces, en virtud de la definición 21.1.1, es claro que el conjunto V_P coincide con V_1 y es por tanto un subespacio de V.

Supongamos ahora que V_P es un subespacio de V. Comprobaremos que entonces (E_1 , V_P , Φ_1), donde Φ_1 es la restricción de Φ a $E_1 \times E_1$, es un espacio afín. Debemos probar primero que Φ_1 aplica a $E_1 \times E_1$ en V_P. En efecto, dados A y B en E_1 se tiene

$$\Phi_1(A , B) = \Phi(A , B) = \overrightarrow{AB} = \overrightarrow{PB} - \overrightarrow{PA} \in V_P .$$

Probaremos luego que se cumplen las propiedades a) y b) de la definición 21.1.1.

La propiedad a) se cumple evidentemente, pues E_1 es un subconjunto de E, conjunto para el cual se cumple esa propiedad. La propiedad b) se cumple porque si A es un punto cualquiera de E_1 y v es un vector cualquiera de V_P, entonces, dado que $v \in V$, existe un único punto $B \in E$ tal que $\overrightarrow{AB} = v$. Por tanto $\overrightarrow{PB} = \overrightarrow{PA} + v$. Ahora bien, tanto $\overrightarrow{PA}$ como v pertenecen a V_P. Luego $\overrightarrow{PB} \in V_P$. Por definición de V_P existe $B_1 \in E_1$ tal que $\overrightarrow{PB_1} = \overrightarrow{PB}$, por consiguiente debe ser $B = B_1$, es decir que B es u punto de E_1, tal que $\overrightarrow{AB} = v$.

Se cumple por tanto la propiedad b). ■

Corolario. Dado un punto cualquiera $P \in E$ y un subespacio V_1 de V, el conjunto

$$E_1 = \{Q \in E \mid \overrightarrow{PQ} \in V_1\}$$

constituye el único subespacio de E que pasa por P y que está asociado a V_1.

Demostración. Para probar que E_1 es un subespacio de E basta probar, según el teorema anterior, que el conjunto $\{\overrightarrow{PQ} \mid Q \in E_1\}$ es un subespacio de V; lo cual es cierto porque

$$\{\overrightarrow{PQ} \mid Q \in E_1\} = V_1$$

En efecto, si $Q \in E_1$, por definición de E_1 es $\overrightarrow{PQ} \in V_1$, y por tanto

$$\{\overrightarrow{PQ} \mid Q \in E_1\} \subset V_1$$

Recíprocamente, dado $v \in V_1$, existe $Q_1 \in E$ tal que $\overrightarrow{PQ_1} = v \in V_1$. Luego $Q_1 \in E_1$, y por tanto

$$V_1 \subset \{\overrightarrow{PQ} \mid Q \in E_1\}$$

La unicidad se cumple porque si E_1 y E_1' fueran dos subespacios de E que pasan por P, asociados a V_1, entonces, si Q pertenece a uno de ellos se tendría $\overrightarrow{PQ} \in V_1$ y por tanto Q pertenecería también al otro subespacio. ■

Teorema 21.5.2. Dado el subespacio V_1 de V, el espacio E se descompone en una familia de subespacios disjuntos cada uno de los cuales está asociado con V_1.

Demostración. Se define una relación de equivalencia entre los puntos de E:

$$P_1 \sim P_2 \Leftrightarrow \overrightarrow{P_1P_2} \in V_1 \quad .z$$

Cada una de las clases de equivalencia que así resulta es un subespacio asociado con V_1 . ■

Definición 21.5.2. Los subespacios de dimensión *1* y *2* de un espa

cio afín E se llaman *rectas y planos de* E respectivamente. Si $\dim E = n$ los subespacios de E de dimensión $n - 1$ se llaman *hiperplanos.*

Teorema 21.5.3. Sea $(E_\alpha)_{\alpha \in A}$ una familia de subespacios del espacio afín E asociado al espacio vectorial V. Si $\cap E_\alpha \neq \phi$ entonces $\cap E_\alpha$ es un subespacio de E. Si V_α es el subespacio de V asociado con E_α, $\cap V_\alpha$ es el subespacio de V asociado con $\cap E_\alpha$.

Demostración. Elijamos arbitrariamente un punto P de $\cap E_\alpha$. Si Q designa a otro punto cualquiera de $\cap E_\alpha$ se tendrá que $\overrightarrow{PQ} \in V_\alpha$ para todo $\alpha \in A$, y por tanto $\overrightarrow{PQ}$ pertenece a $\cap V_\alpha$; es decir que

$$\{\overrightarrow{PQ} \mid Q \in \cap E_\alpha\} \subset \cap V_\alpha .$$

Además, si v es un vector cualquiera de $\cap V_\alpha$, y $Q \in E$ es tal que $\overrightarrow{PQ} \in v$, entonces es claro que $Q \in E_\alpha$ para todo $\alpha \in A$, y por tanto

$$\cap V_\alpha \subset \{\overrightarrow{PQ} \mid Q \in \cap E_\alpha\}$$

Se tiene por consiguiente que

$$\{\overrightarrow{PQ} \mid Q \in \cap E_\alpha\} = \cap V_\alpha .$$

El teorema 21.5.1 implica que $\cap E_\alpha$ es un subespacio de E cuyo espacio vectorial asociado es $\cap V_\alpha$. ∎

Corolario. Dado un subconjunto S, no vacío de E, la intersección de todos los subespacios de E que contienen a S es un subespacio que contiene a S y cuyo espacio vectorial asociado es el subespacio V_1 de V generado por el conjunto de vectores

$$\{\overrightarrow{PQ} \mid P, Q \in S\} \qquad (21.5.1)$$

Demostración. La familia $(E_\alpha)_{\alpha \in A}$ de los subespacios de E que contienen a S es no vacía pues a ella pertenece el espacio E. La intersección E_o de esos subespacios es evidentemente no vacía porque contiene a S. El teorema 21.5.3 implica que E_o es un subespacio de E que contiene a S, cuyo espacio vectorial asociado es la intersección $\cap V_\alpha$ de los espacios asociados a los espacios afines E_α. Demostraremos entonces que el espacio vectorial $\cap V_\alpha$ coincide con el

espacio V_o generado por el conjunto (21.5.1). En primer lugar, y puesto que $S \subset E_\alpha$, es evidente que cada V_α contiene al conjunto (21.5.1), y por tanto contiene a V_o. Por consiguiente

$$V_o \subset \cap V_\alpha .$$

Por otra parte, si P es un punto cualquiera de S, se sigue del corolario del teorema 21.5.1 que el conjunto

$$\{Q \in E \mid \overrightarrow{PQ} \in V_o\}$$

es un subespacio de E que contiene a S y es asociado a V_o. Dicho subespacio pertenece entonces a la familia $\{E_\alpha\}_{\alpha \in A}$, y por tanto V_o es uno de los espacios V_α. Por consiguiente

$$\cap V_\alpha \subset V_o$$

Resulta así que

$$V_o = \cap V_\alpha \quad \blacksquare$$

Definición 21.5.3. Dado un subconjunto S no vacío del espacio afín E, se llama *subespacio subtendido por* S al subespacio intersección de todos los subespacios de E que contiene a S. En particular, dados los subespacios E_1 y E_2 de E, el subespacio subtendido por la reunión de ellos se designa por $E_1 \vee E_2$. Si V_1 y V_2 son los espa vectoriales asociados a E_1 y E_2 y si es $E_1 \cap E_2 \neq \phi$ el espacio vectorial asociado a $E_1 \vee E_2$ es, de conformidad con el corolario del teorema 21.5.3., el subespacio de V generado por V_1 y V_2, es decir el subespacio $V_1 + V_2$. Si $E_1 \cap E_2 = \phi$, entonces el espacio asociado a $E_1 \vee E_2$ contiene a $V_1 + V_2$.

Teorema 21.5.4. Sean E_1 y E_2 dos subespacios de E cuyos subespacios vectoriales asociados sean V_1 y V_2 respectivamente. La condición necesaria y suficiente para que sea $E_1 \supset E_2$ es que $E_1 \cap E_2 \neq \phi$ y que $V_1 \supset V_2$.

Demostración. La condición es evidentemente necesaria. Es suficiente, porque si P es un punto de $E_1 \cap E_2$ y si $Q \in E_2$ entonces es $\overrightarrow{PQ} \in V_2 \subset V_1$, y por tanto $Q \in E_1$. $\blacksquare$

Corolario. Sean E_1 , E_2 , E_3 subespacios de E asociado a los espacios vectoriales V_1 , V_2 , V_3. Si $E_3 \supset E_1 \cup E_2$, entonces es

$$V_3 \supset V_1 + V_2$$

Demostración. Si $E_3 \supset E_1 \cup E_2$, entonces se tiene evidentemente

$$E_3 \supset E_1 \vee E_2 ,$$

y por el teorema anterior se tendrá que $V_3 \supset V_1 + V_2$. ∎

Teorema 21.5.5. Si E_1 y E_2 son subespacios de E tales que

$$E_1 \cap E_2 \neq \emptyset ,$$

entonces se cumple que

$$dim\ (E_1 \vee E_2) + dim\ (E_1 \cap E_2) = dim\ E_1 + dim\ E_2$$

Demostración. Si V_1 y V_2 son los espacios vectoriales asociados con E_1 y E_2, entonces los espacios asociados con $E_1 \vee E_2$ y con $E_1 \cap E_2$ son $V_1 + V_2$ respectivamente. La relación conocida (I, ejercicio 2.5.16)

$$dim\ (V_1 + V_2) + dim\ (V_1 \cap V_2) = dim\ V_1 + dim\ V_2$$

equivale a la relación que se trata de demostrar. ∎

21.6. Paralelismo en los espacios afines

Definición 21.6.1. Se dice que dos subespacios E_1 y E_2 de un subespacio afín E son *paralelos* si el espacio vectorial asociado con uno de ellos está contenido en el espacio vectorial asociado con el otro. Si E_1 y E_2 tienen la misma dimensión, su paralelismo equivale a la identidad de los espacios vectoriales asociados.

Teorema 21.6.1. Si E_1 y E_2 son subespacios paralelos, entonces o bien uno está contenido en el otro, o bien son disjuntos.

Demostración. Si $E_1 \cap E_2 \neq \emptyset$, el teorema 21.5.4 implica que uno de los subespacios contiene al otro. ∎

Teorema 21.6.2. La relación de paralelismo entre los subespacios de una misma dimensión de un espacio afín, es una relación de equivalencia.

Demostración. Esta proposición es evidente. ∎

Teorema 21.6.3. Dado un subespacio E_1 de E, por cada punto P de E pasa uno y solo un subespacio de E paralelo a E_1 y de la misma dimensión que E_1

Demostración. Este teorema se deduce del corolario del teorema 21. 5.1. ∎

Ejercicios

21.6.1. Dada la aplicación afín $L: E \to E$, asociada al endomor - fismo T, y cuya ecuación vectorial es $y = Tx + b$, la con - dición necesaria y suficiente para que posea puntos fijos es que $b \in Im(I - T)$. Si x_0 es un punto fijo cualquiera de T el conjunto de todos los puntos fijos es el subespacio afín de E, que pasa por x_0 y que está asociado al subespacio $Ker\ (I - T)$.

21.6.2. Dados los subespacios E_1 y E_2 de dimensión r_1 y r_2 $(r_1 \geqslant r_2)$ del espacio afín E, para que sean paralelos es necesario y suficiente que se cumpla una de las condiciones siguientes:

1. $E_1 \supset E_2$
2. $E_1 \cap E_2 = \phi$ y $dim\ (E_1 \vee E_2) = r_1 + 1$.

21.7. Representación de subespacios. Coordenadas baricéntricas Convexidad.

Teorema 21.7.1. Dado un subespacio E_1 de dimensión r del espacio afín E de dimensión n, existe un sistema de coordenadas $(y^{i'})$ en E tal que E_1 es el conjunto de los puntos cuyas coordenadas respecto de ese sistema cumplen las condiciones

$$y^{(r+1)'} = \ldots = y^{n'} = 0 .$$

Demostración. Basta elegir el sistema $(y^{i'})$ de manera que los r primeros vectores de la base asociada formen una base del espacio vectorial asociado con E_1 y que el origen sea un punto cualquiera de

E_1 .

Teorema 21.7.2. Si (x^i) es un sistema de coordenadas cualquiera de E, y E_1 es un subespacio de E de dimensión r, existe una matriz de tipo $(n - r) \times n$,

$$B = (b_i^j) \quad , \quad 1 \leqslant j \leqslant n-r \quad , \quad 1 \leqslant i \leqslant n,$$

de rango $n-r$, y una matriz columna de tipo $(n-r) \times 1$,

$$b = (b^j) \ ,$$

tales que E_1 es el conjunto de los puntos de E cuyas coordenadas respecto del sistema (x^i) satisfacen al sistema de $n-r$ ecuaciones

$$\sum_{i=1}^{n} b_i^j x^i + b^j = 0 \quad , \qquad 1 \leqslant j \leqslant n-r \ .$$

Demostración. Sea $(y^{i'})$ el sistema de coordenadas cuya existencia fué establecida en el teorema anterior. Según hemos visto en la sección 21.3 la relación entre las coordenadas de un punto respecto de los sistemas (x^i) e $(y^{i'})$ puede escribirse en la forma (ver segunda ecuación (21.3.2))

$$y^{j'} = \sum_{i=1}^{n} A_i^{j'} x^i + a^{j'} \qquad , \ 1' \leqslant j' \leqslant n'$$

donde $A = (A_i^{j'})$ es una matriz regular de orden n y $a^{j'}$, $j = 1, 2, \ldots, n$, son n escalares determinados. El teorema 21.7.1 implica que E_1 es el conjunto de los puntos de E cuyas coordenadas respecto al sistema (x^i) satisfacen a las ecuaciones

$$\sum_{i=1}^{n} A_i^{j'} x^i + a^{j'} = 0, \qquad (r+1)' \leqslant j' \leqslant n' \ .$$

Si escribimos

$$b_i^j = A_i^{(j+r)'} \qquad 1 \leqslant j \leqslant n-r \ ,$$
$$b^j = a^{(j+r)'}$$

E_1 resulta ser el conjunto de los puntos de E que cumplen las condiciones

$$\sum_{i=1}^{n} b_i^j x^i + b^j = 0 \quad , \qquad 1 \leqslant j \leqslant n-r \ ,$$

donde la matriz $(n-r)\times n$, $B = (b_i^j)$, es de rango $n-r$ por cuanto sus filas lo son de la matriz regular (A_i^j) y son por tanto, linealmente independientes. ∎

Corolario. Si E_1 es un hiperplano de E, sus puntos satisfacen a una ecuación de la forma

$$\sum_{i=1}^{n} b_i x^i + b_o = 0 \ ,$$

donde los escalares $b_1, \ldots, b_n$ no son todos nulos.

Demostración. Resulta del teorema anterior. ∎

Teorema 21.7.3. Los puntos de un espacio afín E de dimensión n que satisfacen a un sistema de $n-r$ ecuaciones

$$\sum_{i=1}^{n} b_i^j x^i + b^j = 0, \qquad 1 \leqslant j \leqslant n-r \ ,$$

donde (b_i^j) es una matriz $(n-r) \times n$ de rango $n-r$, constituyen un subespacio E_1 de E de dimensión r.

Demostración. El sistema de ecuaciones del enunciado es linealmente independiente (sección 8.3), por consiguiente posee soluciones (sección 8.2, observación 2).
Sea $(p^1, \ldots, p^n)$ una solución de ese sistema. $P = (p^1, \ldots, p^n)$ pertenecerá por tanto al conjunto E_1 de los puntos cuyas coordenadas satisfacen al sistema. De inmediato se ve que los puntos Q de E_1 son aquellos cuyas coordenadas (x^i) satisfacen a las ecuaciones

$$\sum_{i=1}^{n} b_i^j (x^i - p^i) = 0 \quad , \qquad 1 \leqslant j \leqslant n-r \ .$$

Si $\{e_i\}$ es la base de V asociada al sistema de coordenadas (x^i) las componentes del vector $\overrightarrow{PQ}$ serán $(x^i - p^i) = \xi^i$.
Por consiguiente los vectores del conjunto

$$\{\overrightarrow{PQ} \mid Q \in E_1\} \qquad (21.7.1)$$

son justamente los vectores

$$\overrightarrow{PQ} = \sum_{i=1}^{n} \xi^i e_i \quad ,$$

cuyas componentes cumplen las relaciones

$$\sum_{i=1}^{n} b_i^j \xi^i = 0 \quad , \quad 1 \leqslant j \leqslant n - r \, .$$

Puesto que la matriz de este sistema homogéneo es de rango $n - r$, los vectores del conjunto (21.7.1) constituyen un espacio vectorial de dimensión r (teorema 8.1.2), lo cual demuestra, en virtud del teorema 21.5.1 que E_1 es un subespacio de dimensión r. ■

Los dos teoremas precedentes establecen la representación de un subespacio de dimensión r mediante un sistema de $n - r$ ecuaciones linealmente independientes entre las coordenadas de sus puntos relativas a un sistema cualquiera de coordenadas. A continuación veremos que los subespacios pueden ser caracterizados de otras maneras. Estudiaremos en espacial las representaciones paramétrica y baricéntrica.

Si E_1 es un subespacio de dimensión r del espacio E de dimensión n, $\{e_i\}$ es una base del espacio V_1 asociado a E_1, y P_o es un punto cualquiera dado de E_1, los puntos P de E_1 son evidentemente aquellos para los cuales existen escalares $t^1, \ldots, t^r$ tales que

$$\overrightarrow{OP} = \overrightarrow{OP_o} + \sum_{i=1}^{r} t^i e_i \quad , \qquad (21.7.2)$$

donde O es un punto arbitrario de E. Si (x^i) es un sistema cualquiera de coordenadas de E de origen O, la ecuación vectorial precedente, que caracteriza a los puntos de P de E_1, es equivalente al sistema

$$x^k = x_o^k + \sum_{i=1}^{r} \alpha_i^k t^i \quad , \quad 1 \leqslant k \leqslant n. \qquad (21.7.3)$$

donde x^k son las coordenadas del punto P, x_o^k son las coordenadas de P_o y α_i^k son las componentes del vector e_i respecto de la base aso-

cida con el sistema de coordenadas (x^i). La matriz (α_i^k) de tipo $n \times r$, es evidentemente de rango r por cuanto sus columnas están formadas por las componentes de los vectores e_i que son linealmente independientes.

Las n ecuaciones (21.7.3), donde las variables t^i, *o parámetros,* pueden tomar valores escalares cualesquiera, y que proporcionan las coordenadas de todos los puntos de E_1 se llaman *ecuaciones paramétricas* del subespacio E_1.

La ecuación (21.7.2) puede escribirse en la forma

$$\overrightarrow{OP} = (1 - \sum_{i=1}^{r} t^i)\overrightarrow{OP_o} + \sum_{i=1}^{r} t^i(\overrightarrow{OP_o} + e_i) . \qquad (21.7.4)$$

Si introducimos los puntos P_i tales que

$$\overrightarrow{OP_i} = \overrightarrow{OP_o} + e_i ,$$

entonces la ecuación (1.7.4) puede escribirse en la forma

$$\overrightarrow{OP} = \sum_{i=0}^{r} s^i \overrightarrow{OP_i} , \qquad (21.7.5)$$

en donde hemos introducido las notaciones

$$s^o = 1 - \sum_{i=1}^{r} t^i , \quad s^i = t^i \quad \text{para } i = 1, \ldots, r .$$

Se cumple por tanto la relación

$$\sum_{i=0}^{r} s^i = 1 . \qquad (21.7.6)$$

De esta manera ha quedado establecido que el subespacio E_1 está formado por los puntos P que satisfacen a la ecuación (21.7.5), en la cual O designa a un punto arbitrario de E, P_o es un punto cualquiera de E_1, y los vectores $\overrightarrow{P_oP_i} = e_i$ constituyen una base del espacio V_1 asociado a E_1, de manera que $(P_o, \ldots, P_n)$ es un sistema de coordenadas de E_1.
Los números s^i que permiten expresar a $\overrightarrow{OP}$ en la forma (21.7.5) son únicos, supuesto que

$$\sum_{i=0}^{r} s^i = 1.$$

En efecto, supongamos que sea $\sum_{i=0}^{r} \bar{s}^i = 1$ y que

$$\sum_{i=0}^{r} s^i \overrightarrow{OP_i} = \sum_{i=0}^{r} \bar{s}^i \overrightarrow{OP_i} .$$

Se tendrá

$$\sum_{i=0}^{r} (s^i - \bar{s}^i \overrightarrow{OP_i} = 0$$

o bien puesto que $\overrightarrow{OP_i} = \overrightarrow{OP_o} + \overrightarrow{P_oP_i}$,

$$(\sum_{i=0}^{r} (s^i - \bar{s}^i))\overrightarrow{OP_o} + \sum_{i=0}^{r} (s^i - \bar{s}^i)\overrightarrow{P_oP_i} = 0 .$$

Pero

$$\sum_{i=0}^{r} (s^i - \bar{s}^i) = \sum_{i=0}^{r} s^i - \sum_{i=0}^{r} \bar{s}^i = 1 - 1 = 0 \quad \text{y} \quad \overrightarrow{P_oP_i} = 0$$

para $i = 0$; luego

$$\sum_{i=1}^{r} (s^i - \bar{s}^i)\overrightarrow{P_oP_i} = 0$$

y puesto que $\overrightarrow{P_oP_1}, \ldots, \overrightarrow{P_oP_r}$ son l.i. , se deduce que $s^i = \bar{s}^i$ para $i = 1, \ldots, r$

De $s^o = 1 - \sum_{i=1}^{r} s^i$ y $\bar{s}^o = 1 - \sum_{i=1}^{r} \bar{s}^i$ se deduce finalmente

$$s^o = \bar{s}^o .$$

Probaremos ahora que los escalares $s^o, s^1, \ldots, s^r$, cuya suma es igual a 1, y que corresponden a un punto P de E_1, son independientes del punto O. En efecto, cualquiera que sea el punto O' de E , se tiene

$$\overrightarrow{O'P} = \overrightarrow{OP} - \overrightarrow{OO'} = \sum_{i=0}^{r} s^i\overrightarrow{OP_i} - \sum_{i=0}^{r} s^i\overrightarrow{OO'}$$

$$= \sum_{i=0}^{r} s^i(\overrightarrow{OP_i} - \overrightarrow{OO'}) = \sum_{i=0}^{r} s^i \overrightarrow{O'P_i} .$$

Lo dicho prueba que cuando en V_1 se ha elegido un sistema de coordenadas $(P_0, P_1, \ldots, P_r)$, a cada punto P se puede asociar $r+1$ escalares $s^0, s^1, \ldots, s^r$) tales que

$$\sum_{i=1}^{r} s^i = 1 ,$$

y que

$$\overrightarrow{OP} = \sum_{i=0}^{r} s^i \overrightarrow{OP_i} ,$$

donde O es un punto cualquiera de E del cual no dependen los escalares s^i. En particular, si suponemos que O coincide con P_0 se ve que, si $s^1, \ldots, s^r$ son las coordenadas de P en el sistema $(P_0, P_1, \ldots, P_r)$, es $s^0 = 1 - \sum_{i=1}^{r} s^i$.

Definición 21.7.1. Los escalares $s^0, s^1, \ldots, s^r$, cuya suma es igual a 1, y que corresponden a cada punto P del subespacio E_1 de dimensión r, mediante la fórmula (21.7.5) se llaman *coordenadas baricéntricas de* P respecto del conjunto de $r+1$ puntos $(P_0, P_1, \ldots, P_r)$ de E_1 tales que los vectores $\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_r}$ son *l.i.*, y constituyen por tanto una base de E_1. La ecuación (21.7.5) es entonces la correspondiente *ecuación baricéntrica* del subespacio E_1 .

Definición 21.7.2. Dados los puntos A y B del espacio afín E, se llama *segmento de extremos* A y B, y se designa por AB al conjunto

$$AB = \{c \in E \mid \overrightarrow{AC} = \lambda \overrightarrow{AB} , \ 0 \leqslant \lambda \leqslant 1\}$$

Es evidente que los extremos A y B pertenecen al segmento AB.

Definición 21.7.3. Un conjunto no vacío S de un espacio afín S se dice que es *convexo* si, dados dos cualesquiera de sus puntos A y B, el segmento AB pertenece íntegramente a S.

Ejemplos. A. Un espacio afín, así como cada uno de sus subespacios, son conjuntos convexos.

B. Sea E_1 un hiperplano de E definido por la ecuación

$$\sum_{i=1}^{n} b_i x^i + b_o = 0 \ ,$$

relativa al sistema de coordenadas (x^i).

Se llaman *semiespacios* definidos por el hiperplano E_1 a cada uno de los conjuntos de puntos (x^i) definidos por las desigualdades

$$\sum_{i=1}^{n} b_i x^i + b_o > 0 \qquad (21.7.7)$$

y

$$\sum_{i=1}^{n} b_i x^i + b_o < 0 \ . \qquad (21.7.8)$$

Los conjuntos definidos por las condiciones

$$\sum_{i=1}^{n} b_i x^i + b_o \geqslant 0$$

y

$$\sum_{i=1}^{n} b_i x^i + b_o \leqslant 0$$

se llaman *semiespacios cerrados* definidos por el hiperplano E_1 .

Los semiespacios y los semiespacios cerrados son conjuntos convexos. Demostrémoslo, por ejemplo, para el semiespacio definido por la desigualdad (21.7.7). Si $P = (p^1, \ldots, p^n)$ y $Q = (q^1, \ldots, q^n)$ son dos puntos de ese semiespacio, se tiene que

$$\sum_{i=1}^{n} b_i p^i + b_o > 0$$

y

$$\sum_{i=1}^{n} b_i q^i + b_o > 0$$

Un punto cualquiera del segmento PQ es dado por

$$C = (p^1 + \lambda(q^1 - p^1), \ldots, p^n + \lambda(q^n - p^n)) \ ,$$

$$0 \leqslant \lambda \leqslant 1.$$

Se tiene entonces que

$$\sum_{i=1}^{n} b_i(p^i + \lambda(q^i - p^i)) + b_o = \lambda(\sum_{i=1}^{n} b_i q^i + b_o) + (1 - \lambda)(\sum_{i=1}^{n} b_i p^i + b_o)$$
$$! \qquad > 0,$$

y por consiguiente el punto C pertenece también al semiespacio considerado, el cual será por tanto convexo.

Ejercicios

*21.7.1. Se dice que $r + 1$ puntos P_i $(i = 0, 1, \ldots, r)$ de un espacio afín E de dimensión n, están en *posición general,* si no están contenidos en un subespacio de dimensión $r - 1$. Probar que la condición necesaria y suficiente para que los puntos P_i $(i = 0, 1, \ldots, r)$ estén en posición general es que los vectores $\overrightarrow{P_o P_i}$ $(i = 1, \ldots, r)$ sean linealmente independientes.

*21.7.2. Demostrar que un *paralelepípedo r-dimensional,* es decir un conjunto de la forma

$$\{P \in E \setminus \overrightarrow{P_o P} = \sum_{i=1}^{r} t^i \overrightarrow{P_o P_i} \quad , \quad 0 \leqslant t^i \leqslant 1\},$$

donde los puntos $P_o, P_1, \ldots, P_r$ están en posición general, es un conjunto convexo.

*21.7.3. Demostrar que un *cono r-dimensional*

$$\{P \in E \setminus \overrightarrow{P_o P_1} = \sum_{i=1}^{r} t^i \overrightarrow{P_o P_1} \quad , \quad 0 \leqslant t^i\},$$

donde los puntos $P_o, P_1, \ldots, P_r$ están en posición general, es un conjunto convexo.

*21.7.4. Demostrar que dados los puntos $P_o, P_1, \ldots, P_r$ en posición general, el *simplejo r-dimensional* cuyos vértices son dichos puntos, y que es definido por

$$\{P \in E \setminus \overrightarrow{OP} = \sum_{i=0}^{r} s^i \overrightarrow{OP_i} \ , \ \sum_{i=0}^{r} s^i = 1 \ , \ s^i \geqslant 0\},$$

donde O es un punto cualquiera de E, es un conjunto con

vexo. El punto del simplejo cuyas coordenadas baricentricas $(s^0, \ldots, s^r)$ son dadas por

$$s^0 = s^1 = \ldots = \frac{1}{r+1}$$

se llama *centro de gravedad o baricentro* del simplejo.

21.7.5. Por dos puntos distintos pasa una y solo una recta.

21.7.6. Dos rectas que tienen un solo punto común pertenecen a un solo plano.

21.7.7. Dados una recta y un punto situado fuera de ella, solo hay un plano que contiene a ambos.

21.7.8. Dos rectas situadas en un plano son paralelos o tienen a lo más un punto común.

21.7.9. Dos ecuaciones lineales

$$\sum_{i=1}^{n} a_i x^i + b = 0 \qquad (a_1, \ldots, a_n) \neq (0, \ldots, 0)$$

y

$$\sum_{i=1}^{n} c_i x^i + d = 0 \qquad (c_1, \ldots, c_n) \neq (0, \ldots, 0)$$

definen al mismo hiperplano si y solo si existe un escalar no nulo k tal que

$$c_i = k\, a_i \quad (i = 1, \ldots, n) \quad \text{y} \quad d = kb.$$

21.7.10. Demostrar que las dos ecuaciones del ejercicio anterior definen a dos planos paralelos y no identicos si y solo si existe el escalar no nulo k tal que

$$c_i = ka_i \quad (i = 1, \ldots, n) \quad \text{y} \quad d \neq kb.$$

21.7.11. Consideremos la transformación afín de $\mathbb{R}^2$ definida por

$$\begin{pmatrix} y^1 \\ y^2 \\ 1 \end{pmatrix} = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix}$$

22. Espacio puntual euclideano.

Hallar las ecuaciones de las imágenes por esta transformación de las rectas dadas por sus ecuaciones

a) $x^1 = x^2$

b) $2x^1 - x^2 = 3.$

§ 22. Espacios puntuales euclideanos

22.1. Conceptos fundamentales.

Definición 22.1.1. Se dice que un espacio afín E es un *espacio puntual euclideano de dimensión n,* si su espacio vectorial asociado V es euclideano, es decir real, de producto interno y de dimensión n. (cap. 5).

En el presente parágrafo los espacios puntuales euclideanos serán llamados simplemente *espacios euclideanos*. En lo que sigue, E designará a un espacio euclideano de dimensión n y V designará a su espacio vectorial euclideano asociado.

Definición 22.1.2. Dados dos puntos A y B de E, se llama *distancia* entre esos puntos, y se designa por $d(A, B)$, al número real no negativo.

$$d(A, B) = \|\overrightarrow{AB}\| ,$$

es decir a la norma del vector $\overrightarrow{AB}$ relativa al producto interno definido en V; se tiene por consiguiente que

$$d(A, B) = \sqrt{(\overrightarrow{AB}, \overrightarrow{AB})} .$$

El teorema siguiente es consecuencia inmediata de las propiedades conocidas de la norma (sección 17.2, en especial el teorema 17.2.2).

Teorema 22.1.1. La distancia definida en un espacio euclideano posee las propiedades siguientes:

1. $d(A, B) \geqslant 0 \quad ; \quad d(A, B) = 0 \Leftrightarrow A = B.$
2. $d(B, A) = d(A, B)$.
3. $d(A, C) \leqslant d(A, B) + d(B, C)$ *(desigualdad triangular)*

Una función real definida en $E \times E$ se llama *función de distancia* si posee las propiedades que, según el teorema anterior, posee la función d. Se dice entonces que tal función define en E una *métrica* y que, por tanto E es un *espacio métrico*. En particular, la función d es una función de distancia, E es un espacio métrico, y la métrica en él definida por d es una *métrica euclideana*, llamada así para distinguirla de otras métricas posibles.

Definición 22.1.3. Dados tres puntos distintos P, Q y R del espacio euclideano E se llama *ángulo QPR*, y se designa por $\measuredangle PQR$, al número real θ, $0 \leqslant \theta \leqslant \pi$, tal que

$$\cos \theta = \frac{(\overrightarrow{PQ} , \overrightarrow{PR})}{\|\overrightarrow{PQ}\| \, \|\overrightarrow{PR}\|} , \qquad (22.1.1)$$

donde $(\overrightarrow{PQ} , \overrightarrow{PR})$ designa al producto interno de los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$.

Debe notarse que el valor del segundo miembro de la fórmula (22.1.1) está comprendido entre -1 y $+1$, en virtud de la desigualdad se Schwarz (teorema 17.2.1), y por tanto el ángulo θ es bien definido.

Teorema 22.1.2. Dados tres puntos distintos P, Q y R de un espacio euclideano, se cumples la relación

$$d(Q , R)^2 = d(P , Q)^2 + d(P , R)^2 - 2d(P , Q)d(P , R)\cos \measuredangle QPR. \qquad (22.1.2)$$

Demostración. Sea $u = \overrightarrow{PQ}$, $v = \overrightarrow{PR}$ y $w = \overrightarrow{QR}$, de manera que es $w = v - u$. Si escribimos $\measuredangle QPR = \theta$, se tiene

$$\|w\|^2 = (v - u , v - u) = (v , v) - 2(u , v) + (u , u)$$

$$= \|v\|^2 + \|u\|^2 - 2\|u\| \, \|v\| \cos \theta ,$$

que es la relación (22.1.2). ∎

Corolario. La condición necesaria y suficiente para que, dados los tres puntos distintos P, Q, R, de un espacio euclideano, se cumpla la relación

$$d(Q\ ,\ R)^2 = d(P\ ,\ Q)^2 + d(P\ ,\ R)^2 \qquad (22.1.3)$$

es que

$$(\overrightarrow{PQ}\ ,\ \overrightarrow{PR}) = 0$$

Demostración. Esta proposición resulta inmediatamente del teorema anterior, si se tiene en cuenta que $\|\overrightarrow{PQ}\| = d(P\ ,\ Q)$ y $\|\overrightarrow{PR}\| = d(P, R)$ son diferentes de cero, y que, por tanto, la relación (22.1.3) es equivalente a $cos \measuredangle QPR = 0$, o sea a $(\overrightarrow{PQ}\ ,\ \overrightarrow{PR}) = 0$. ∎

Definición 22.1.4. Un sistema de coordenadas $(O\ ,\ \{e_i\})$ se llama *rectangular u ortogonal* si los vectores $\{e_i\}$ constituyen una base ortonormal del espacio vectorial V.

Teorema 22.1.3. Un sistema de coordenadas $(O\ ,\ \{e_i\})$ de un espacio euclideano es rectangular si y solo si la distancia entre dos puntos cualesquiera $A = (a^i)$ y $B = (b^i)$ es dada por la fórmula.

$$d(A\ ,\ B) = \sqrt{\sum_{i=1}^{n} (b^i - a^i)^2}$$

Demostración. La condición es necesaria porque si $\{e_i\}$ es una base ortonormal, se tiene

$$d(A\ ,\ B) = \sqrt{(\overrightarrow{AB}\ ,\ \overrightarrow{AB})} = \sqrt{(\sum_{i=1}^{n} (b^i - a^i)e_i\ ,\ \sum_{i=1}^{n} (b^i - a^i)e_i)} = \sqrt{\sum_{i=1}^{n} (b^i - a^i)^2}$$

Si suponemos, recíprocamente, que se cumple la fórmula de enunciado y consideramos los puntos P_i $(i = 1, \ldots, n)$ tales que

$$\overrightarrow{OP_i} = e_i$$

tendremos, teniendo presente que las coordenadas de P_i son nulas, salvo la *i-ésima* que es igual a 1, y aplicando la fórmula para calcular $d(O\ ,P_i)$.

$$(e_i\ ,\ e_i) = d(O\ ,\ P_i)^2 = 1\ ;$$

y para $i \neq j$,

$$d(P_i\ ,\ P_j) = \sqrt{2}\ .$$

Por tanto

$$d(P_i , P_j)^2 = d(O , P_i)^2 + d(O , P_j)^2$$

Del corolario del teorema 2.1.2. se deduce que

$$(\overrightarrow{OP}_i , \overrightarrow{OP}_j) = 0 \quad ,$$

es decir

$$(e_i , e_j) = 0.$$

Por consiguiente la base $\{e_i\}$ es ortogonal. ∎

Teorema 22.1.4. Si (x^i) e $(y^{i'})$ son dos sistemas rectangulares de coordenadas del espacio euclideano E, entre las coordenadas de un mismo punto en esos dos sistemas se cumplen las relaciones

$$\begin{pmatrix} y^{1'} \\ \vdots \\ y^{n'} \\ 1 \end{pmatrix} = \begin{pmatrix} A & \alpha' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix} = \begin{pmatrix} A' & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{1'} \\ \vdots \\ y^{n'} \\ 1 \end{pmatrix}$$

donde A y A' son matrices ortogonales de orden n tales que

$$AA' = A'A = I \ ,$$

y α' y α son columnas de n elementos escalares.

Demostración. Las fórmulas que deseamos probar son las fórmulas (21.3.8). Al obtener éstas se vio que las matrices A y A' son las matrices de cambio de las bases de los sistemas (x^i) e $(y^{i'})$. Por hipótesis esas bases son ortonormales, por consiguiente, en virtud del teorema 17.4.1, dichas matrices son ortogonales e inversas una de otra. ∎

Definición 22.1.5. Se dice que un vector $v \in V$ es *normal* u *ortogonal*

al subespacio E_1 del espacio euclideano E asociado a V, si v es ortogonal (definición 17.7.1) a todos los vectores del subespacio V_1 asociado a E_1, es decir si

$$(v \ , \ x) = \quad \forall \ x \ \epsilon \ V_1 \ .$$

Teorema 22.1.5. Sea E_1 un subespacio *r-dimensional* del espacio euclideano *n-dimensional* E. A cada punto $A \notin E_1$ le corresponde un único punto $B \ \epsilon \ E_1$ tal que $\overrightarrow{AB}$ es normal a E_1. Se tiene entonces

$$d(A \ , \ B) < d(A \ , \ P)$$

para todo punto $P \ \epsilon \ E_1$, diferente de B.

Demostración. Sean V y V_1 los espacios vectoriales asociados con E y E_1 respectivamente. Si $V_1^{\perp}$ es el complemento ortogonal de V_1 (definición 17.7.2), entonces, en virtud del teorema 17.7.1, se tiene

$$V = V_1 \oplus V_1^{\perp} \ .$$

Si P_o designa a un punto dado arbitrario de E_1, existirán los vectores bien determinados $u \ \epsilon \ V_1$ y $v \ \epsilon \ V_1^{\perp}$ tales que

$$\overrightarrow{P_o A} = u + v \ .$$

Existirá entonces un único punto B de E_1 tal que $\overrightarrow{P_o B} = u$, y por tanto $\overrightarrow{AB} = -v$ es normal a E_1. Para probar la primera parte del teorema solo falta demostrar que para cualquier otro punto B' de E_1 el vector $\overrightarrow{AB'}$ no es normal a E_1. Supongamos que, por el contrario, fuera B' un punto de E_1 distinto de B y tal que $\overrightarrow{AB'}$ fuera normal a E_1. El vector $\overrightarrow{BB'} = \overrightarrow{AB'} - \overrightarrow{AB}$ sería también normal a E_1. Ahora bien $\overrightarrow{BB'}$ pertenece a V_1. Se tendría por tanto que $(\overrightarrow{BB'}, \ \overrightarrow{BB'}) = 0$, de donde se seguiría que $B = B'$ y resultaría una contradicción.

Demostraremos ahora la parte final del enunciado. Cualesquiera que sea el punto P de E_1 podemos escribir

$$\overrightarrow{AP} = \overrightarrow{AB} + \overrightarrow{BP}$$

donde $(\overrightarrow{AB}\ ,\ \overrightarrow{BP}) = 0$, porque $\overrightarrow{AB}$ es normal a E_1 y $\overrightarrow{BP}$ pertenece a V_1. Luego

$$\|\overrightarrow{AP}\|^2 = \|\overrightarrow{AB} + \overrightarrow{BP}\|^2 = (\overrightarrow{AB} + \overrightarrow{BP}\ ,\ \overrightarrow{AB} + \overrightarrow{BP})$$

$$= \|\overrightarrow{AB}\|^2 + \|\overrightarrow{BP}\|^2 \geqslant \|\overrightarrow{AB}\|^2 .$$

La igualdad solamente tiene lugar cuando $\|\overrightarrow{BP}\| = 0$, o sea cuando $P = B$. Por tanto, para todo P de E_1 diferente de B, es

$$\|\overrightarrow{AP}\| > \|\overrightarrow{AB}\| \ ,$$

o sea

$$d(A\ ,\ P) > d(A\ ,\ B) \ . \qquad \blacksquare$$

Definición 22.1.6. Dado un sibespacio E_1 de E y un punto A situado fuera de E_1, se llama *distancia* de A a E y se designa por

$$d(A\ ,\ E_1)$$

al número $d(A\ ,\ B)$ donde B es el único punto de E_1 tal que $\overrightarrow{AB}$ es normal a E_1.

Definición 22.1.7. Dados un espacio euclideano E y un subespacio E_1 asociado al espacio vectorial V_1, cada subespacio asociado con $V_1^{\perp}$ se llama *subespacio ortogonal a* E_1. Si la dimensión de E es n y la de E_1 es r, la de cada subespacio ortogonal a E_1 será $n - r$. En particular un subespacio ortogonal a un hiperplano será una linea recta.

Ejercicios

22.1.1. Consideremos en el espacio euclideano $\mathbb{R}^2$ el sistema de coordenadas $\{P_o\ ,\ P_1\ ,\ P_2\}$, donde

$$P_o = (1\ ,\ 1) \quad , \quad P_1 = (2\ ,\ 2) \quad , \quad P_2 = (-3\ ,\ 5).$$

Hallar la distancia entre los puntos $(x_1\ ,\ y_1)$ y $(x_2\ ,\ y_2)$ referida a ese sistema de coordenadas.

22.1.2. Determinar la relación entre las coordenadas de un punto de $\mathbb{R}^2$ referido al sistema canónico y las coordenadas del

mismo referido al sistema $\{P_o , P_1 , P_2\}$

$P_o = (a , b)$, $P_1 = (a + cos\ \theta\ ,\ b + sen\ \theta)$,

$P_2 = (a - sen\ \theta\ ,\ b + cos\ \theta)$.

***22.1.3.** Demostrar que la fórmula

$$d'(A , B) = \max_{1 \leqslant i \leqslant n} \{ | b^i - a^i | \}$$

donde $A = (a^1 , \ldots , a^n)$ y $B = (b^1 , \ldots , b^n)$ define una métrica en $\mathbb{R}^n$.

***22.1.4.** Dado un *simplejo* de vértices $P_o , P_1 , \ldots , P_r$ del espacio euclideano E, demostrar que el mínimo de la función

$$P \mapsto \sum_{i=0}^{r} d(P , P_i)^2$$

tiene lugar cuando P coincide con el baricentro del simplejo.

***22.1.5.** Demostrar que si E_1 es un hiperplano del espacio afín E de dimensión n, (x^i) es un sistema de coordenadas rectangulares y $(a^1 , \ldots , a^n)$ son las componentes del vector unitario normal a E_1, respecto de la base asociada al sistema (x^i), entonces la ecuación del hiperplano E_1 es

$$\sum_{i=1}^{n} a^i x^i = b \quad \text{donde} \quad \sum_{i=1}^{n} (a^i)^2 = 1.$$

Probar además que puede suponerse $b > 0$, si se elige convenientemente el sentido del vector normal. La ecuación se llama entonces *ecuación normal* del hiperplano E_1. Probar por último que si P_o es el origen del sistema de coordenadas se tiene

$$d(P_o , E_1) = b.$$

22.1.6. Dado el plano de $\mathbb{R}^3$ cuya ecuación referida al sistema canónico es

$$x^1 + 2x^2 - 2x^3 + 4 = 0 ,$$

obtener su ecuación normal y la distancia del origen a ese plano.

22.1.7. Dados los planos π_1 y π_2 de $\mathbb{R}^3$ de ecuaciones

$$\pi_1 : x^1 + x^2 - x^3 + 3 = 0$$

$$\pi_2 : 2x^1 - x^2 - 1 = 0 ,$$

referidas al sistema canónico, obtener las ecuaciones normales y calcular $d(A , \pi_1)$ y $d(A , \pi_2)$, supuesto que $A=(-1 , 1 , 1)$.

22.2. Isometrías de espacios puntuales euclideanos.

Definición 22.2.1. Dados los espacios euclideanos E_1 y E_2 de dimensión n, asociados a los espacios vectorialed V_1 y V_2, se dice que una aplicación afín $f : E_1 \to E_2$ es una *isometría* si el homomorfismo asociado $T : V_1 \to V_2$ es una isometría de espacios vectoriales euclideanos (sección 18.4), es decir si

$$\|Tx\| = \|x\| \quad , \qquad x \in V_1$$

Teorema 22.2.1. La condición necesaria y suficiente para que la aplicación afín $f : E_1 \to E_2$ sea una isometría es que

$$d(P , Q) = d(f(P) , f(Q)) ,$$

cualesquiera que sean los puntos P y Q de E_1 .

Demostración. La condición es necesaria porque dados los puntos P y Q de E_1 se tiene

$$d(P , Q) = \|\overrightarrow{PQ}\| = \|\overrightarrow{f(P)f(Q)}\| = d(f(P) , f(Q)).$$

La condición es suficiente porque dado $x \in V$, si $\overrightarrow{PQ} = x$, se tiene

$$\|x\| = \|\overrightarrow{PQ}\| = d(P , Q) = d(f(P) , f(Q)) = \|\overrightarrow{f(P)f(Q)}\|$$

$$= \|T(\overrightarrow{PQ})\| = \|Tx\| . \quad \blacksquare$$

Teorema 22.2.2. Toda isometría $f : E_1 \to E_2$, donde E_1 y E_2 son

espacios afines de la misma dimensión, tiene las siguientes propiedades:

1. Es una biyección

2. Si (x^i) e $(y^{i'})$ son sistemas rectangulares en E_1 y en E_2 respectivamente, entonces entre las coordenadas (x^i) de un punto P y las coordenadas $(y^{i'})$ del punto $f(P)$ tiene lugar una relación de la forma

$$\begin{pmatrix} y^{1'} \\ \vdots \\ y^{n'} \\ 1 \end{pmatrix} = \begin{pmatrix} a_1^{1'} & \dots & a_n^{1'} & a^{1'} \\ \vdots & \vdots\vdots\vdots & \vdots & \vdots \\ a_1^{n'} & \dots & a_n^{n'} & a^{n'} \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ 1 \end{pmatrix}$$

donde

$$A = \begin{pmatrix} a_1^{1'} & \dots & a_n^{1'} \\ \vdots & \vdots\vdots\vdots & \vdots \\ a_1^{n'} & \dots & a_n^{n'} \end{pmatrix}$$

es una matriz ortogonal.

Demostración. La propiedad 1. es consecuencia del teorema 18.4. 2, que asegura que T es un isomorfismo, ydel teorema 21.4.3.

Probemos ahora la propiedad 2. Sea $\{P_o, P_1, \dots, P_n\}$ el sistema de coordenadas (x^i). Puesto que este sistema es ortogonal, se tiene

$$(\overrightarrow{f(P_o)f(P_i)}\ ,\ \overrightarrow{f(P_o)f(P_j)}) = (\overrightarrow{T(P_oP_i)}\ ,\ \overrightarrow{T(P_oP_j)})$$

$$= (P_oP_i\ ,\ P_oP_j) = \delta_{ij}\ .$$

Luego $\{f(P_o)\ ,\ f(P_1), \dots, f(P_n)\}$ es un sistema rectangular de coordenadas en E_2. Sean $(\bar{x}^i)$ las coordenadas correspondientes a ese sistema. La relación entre las coordenadas de un punto de E_2

relativas a los sistemas $(y^{i'})$ y $(\bar{x}^i)$ es de la forma

$$y^{j'} = \sum_{i=1}^{n} a_i^{j'} \bar{x}^i + a^{j'} \qquad (22.2.1)$$

donde $A = (a_i^{j'})$ es una matriz ortogonal (teorema 22.1.4). Ahora bien, es fácil observar que las coordenadas de un punto $P = (x^1, x^2, \ldots, x^n)$ de E_1, respecto del sistema de coordenadas $(P_o, P_1, P_2, \ldots, P_n)$, son idénticas a las correspondientes coordenadas del punto $f(P) = (\bar{x}^1, \ldots, \bar{x}^n)$ respecto del sistema $\{f(P_o), f(P_1), f(P_2), f(P_3), \ldots, f(P_n)\}$. Luego, de las relaciones (2.2.1), resultan las relaciones

$$y^{j'} = \sum_{i=1}^{n} a_i^{j'} x^i + a^{j'} \qquad (22.2.2)$$

que demuestran la propiedad 2 del enunciado. ■

Corolario. Si f es una isometría de E_1 sobre E_2, f^{-1} es una isometría de E_2 sobre E_1.

Demostración. Inmediata. ■

El teorema 22.2.1 establece que una aplicación afín que conserva las distancias es una isometría. El teorema siguiente demuestra que una aplicación cualquiera que conserva las distancias es afín y por consiguiente es una isometría.

Teorema 22.2.3. Dados dos espacios afines E_1 y E_2 de la misma dimensión n, y la aplicación $f : E_1 \to E_2$, si esta aplicación conserva las distancias, es decir si

$$d(f(P), f(Q)) = d(P, Q),$$

entonces es una isometría.

Demostración. Para probar que f es una isometría basta demostrar que es una aplicación afín, en virtud del teorema 22.2.1.

Sean V_1 y V_2 los espacios vectoriales asociados con E_1 y E_2. Sea P un punto cualquiera dado de E_1, y definamos la aplicación

$$T_p : V_1 \to V_2$$

tal que

$$T_P(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}$$

Para probar que f es una aplicación afín deberemos probar que T_P es una transformación lineal de V_1 en V_2, pues entonces el corolario del teorema 1.4.1 permite reconocer que f es una aplicación afín asociada a T_P , y T_P es independiente de P.

Previamente vamos a probar que, en virtud de la hipótesis, T_P conserva el producto interno.

Si en la relación (22.1.2) sustituimos con $\cos \measuredangle QPR$ por su valor (22.1.1), podremos escribir, cualesquiera que sean los puntos P , Q y R de E_1

$$2(\overrightarrow{PQ} , \overrightarrow{PR}) = d(P , Q)^2 + d(P , R)^2 - d(Q , R)^2$$
$$= d(f(P) , f(Q))^2 + d(f(P) , f(R))^2 - d(f(Q) , f(R))^2 ,$$

en virtud de que f conserva las distancias. Ahora bien, la misma relación (22.1.2) aplicada a los puntos $f(P)$, $f(Q)$ y $f(R)$ en vez de P , Q y R, muestra que el último miembro de la igualdad precedente es igual a

$$2(\overrightarrow{f(P)f(Q)} , \overrightarrow{f(P)f(R)}) .$$

Por consiguiente resulta la relación

$$(\overrightarrow{PQ} , \overrightarrow{PR}) = (\overrightarrow{f(P)f(Q)} , \overrightarrow{f(P)f(R)}) ,$$

o sea

$$(x , y) = (T_P(x) , T_P(y)) \quad , \quad \forall\, x,y \in V_1 ,$$

y por tanto T_P conserva el producto interno. De esta propiedad vamos a valernos ahora para demostrar finalmente que T_P es lineal.

Sea $\{e_1 , \ldots , e_n\}$ una base ortonormal de V_1. Los vectores

$$T_P(e_1) , \ldots , T_P(e_j)$$

constituyen también una base ortonormal por cuanto

$$(T_P(e_i) , T_P(e_j)) = (e_i , e_j) = \delta_{ij}$$

Si x es un vector cualquiera de V_1 el teorema 17.3.3. permite escribir

$$x = \sum_{i=1}^{n} (x \ , \ e_i)e_i \ ,$$

y semejantemente

$$T_P(x) = \sum_{i=1}^{n} (T_P(x) \ , \ T_P(e_i))T_P(e_i)$$

$$= \sum_{i=1}^{n} (x \ , \ e_i)T_P(e_i) \ .$$

Se deduce de aquí que

$$T_P(x + y) = \sum_{i=1}^{n} (x + y \ , \ e_i)T_P(e_i) = \sum_{i=1}^{n} (x \ , \ e_i)T_P(e_i) + \sum_{i=1}^{n} (y \ , \ e_i)T_P(e_i)$$

$$= T_P(x) + T_P(y) \ ,$$

y

$$T_P(\alpha x) = \sum_{i=1}^{n} (\alpha x \ , \ e_i)T_P(e_i) = \alpha \sum_{i=1}^{n} (x \ , \ e_i)T_P(e_i) = \alpha T_P(x).$$

Por consiguiente T_P es lineal y f es una isometría. ∎

Definición 22.2.2. Una isometría $f: E \longrightarrow E$ se llama *isometría de E* El endomorfismo asociado será un automorfismo ortogonal del espacio vectorial asociado V (definición 20.3.1).

Las isometrías de un espacio afín E constituyen un grupo respecto de la operación de composición de automorfismos. Dicho grupo se designa por $I(E)$ y es isomorfo con el grupo de la matrices de la forma

$$M = \begin{pmatrix} A & \alpha \\ 0 & 1 \end{pmatrix} \ , \qquad (22.2.3)$$

donde A es una matriz ortogonal de orden n. Es claro que el isomorfismo entre esos grupos varía con el sistema de coordenadas que se considere, porque a cada isometría le corresponde la matriz que

la representa en el sistema de coordenadas elegido. En virtud del teorema 10.3.3 las matrices de las isometrías tienen determinante igual a $+1$ o a -1.

Definición 22.2.3. Una isometría $f: E \longrightarrow E$ se llama *movimiento (o movimiento rígido) de* E si la matriz (22.2.3) que la representa respecto de cualquier sistema de coordenadas tiene determinante igual a $+1$.

Ejemplo. Toda traslación es un movimiento.

Definición 22.2.4. Una isometría f de un espacio afín E se llama *rotación alrededor de un punto* O si cumple la condición $f(O) = O$, es decir si es centroafín de centro O. Si además es un movimiento entonces se llama *rotación propia*, y si no es un movimiento se llama *rotación impropia*.

Una rotación alrededor de O puede expresarse en la forma

$$\begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix} = A \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} ,$$

donde A es una matriz ortogonal, si se adopta un sistema de coordenadas rectangulares de origen O (Demostrarlo!).

Si la rotación es propia, entonces es $\det A = 1$.

Teorema 22.2.4. Sea f una isometría de E. Dado un punto cualquiera O, puede escribirse de una manera única

$$f = \theta L_o$$

donde L_o es una rotación alrededor de O y θ es una traslación. f es un movimiento si y solo si L_o es una rotación propia.

Demostración. La descomposición única de f en la forma θL_o, donde L_o es una rotación alrededor de O y θ es una traslación, resulta del sistema 21.4.6. Es fácil observar que si

$$M = \begin{pmatrix} A & \alpha \\ 0 & 1 \end{pmatrix}$$

es la matriz que representa a f respecto de un sistema cualquiera de coordenadas, entonces A es la matriz de la rotación L_o, referida a ese sistema, y α es una columna cuyos términos son las componentes del vector de la traslación θ, respecto de la base asociada a dicho sistema de coordenadas (ver teorema 21.4.8 y la observación que le sigue). La parte final del enunciado se cumple porque la condición para que f sea un movimiento es la misma que para que los sea L_o, a saber que el determinante de A tenga el valor $+1$. ∎

22.3. Clasificación de las isometrías de un espacio euclideano.

Sea E un espacio euclideano de dimensión n, asociado al espacio vectorial V. f sea una isometría de E y T sea el automorfismo ortogonal de V, asociado a f. Según el teorema 20.3.3 existe una base ortonormal de V con respecto a la cual T se representa por una matriz de la forma canónica siguiente, en que los términos no indicados son iguales a cero.

$$A = \begin{pmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & -1 & & & & & \\ & & & & -1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & \cos\theta_1 & -\operatorname{sen}\theta_1 & \\ & & & & & & \operatorname{sen}\theta_1 & \cos\theta_2 & \\ & & & & & & & & \ddots \\ & & & & & & & & \cos\theta_k & -\operatorname{sen}\theta_k \\ & & & & & & & & \operatorname{sen}\theta_k & \cos\theta_k \end{pmatrix},$$

$$0 < \theta_i < 2\pi \,. \qquad (22.3.1)$$

Cada uno de los bloques diagonales de la forma $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ puede

ser escrito en la forma

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

donde $\theta = \pi$. Por consiguiente cada matriz del tipo 22.3.1 puede ser escrita en una de las dos formas siguientes

$$A = \begin{pmatrix} \overbrace{1 \quad 1 \quad \ddots \quad 1}^{r} & & & \\ & \cos\theta_1 & -\operatorname{sen}\theta_1 & \\ & \operatorname{sen}\theta_1 & \cos\theta_1 & \\ & & & \ddots \\ & & & \cos\theta_p \quad -\operatorname{sen}\theta_p \\ & & & \operatorname{sen}\theta_p \quad \cos\theta_p \end{pmatrix}, \qquad (a)$$

$$A = \begin{pmatrix} \overbrace{1 \quad 1 \quad \ddots \quad 1}^{r-1} & & & \\ & -1 & & \\ & & \cos\theta_1 & -\operatorname{sen}\theta_1 \\ & & \operatorname{sen}\theta_1 & \cos\theta_1 \\ & & & \ddots \\ & & & \cos\theta_p \quad -\operatorname{sen}\theta_p \\ & & & \operatorname{sen}\theta_p \quad \cos\theta_p \end{pmatrix}, \qquad (b)$$

donde $0 < \theta_i < 2\pi$, $0 \leqslant r \leqslant n$ y $r + 2p = n$.

Las formas (a) y (b) corresponden a los casos en que la matriz (22.3.1) tiene un número par e impar, respectivamente, de términos iguales a -1.

La matriz de la isometría f respecto de un sistema de coordenadas rectangulares asociado a la base relativa a la cual la matriz de T tiene la forma (22.3.1), será (definición 21.4.5).

$$M = \begin{pmatrix} A & \alpha \\ 0 & 1 \end{pmatrix}$$

donde

$$\alpha = \begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix}$$

Consideremos en primer lugar el caso en que la matriz A es de la forma (a). Puede escribirse entonces

$$A = \begin{pmatrix} I_r & 0 \\ 0 & R \end{pmatrix}$$

donde I_r designa a la matriz unidad de orden r y R a la matriz de orden $2p = n - r$

$$R = \begin{pmatrix} \cos\theta_1 & -\operatorname{sen}\theta_1 & & \\ \operatorname{sen}\theta_1 & \cos\theta_1 & & \\ & & \ddots & \\ & & \cos\theta_p & -\operatorname{sen}\theta_p \\ & & \operatorname{sen}\theta_p & \cos\theta_p \end{pmatrix}$$

La matriz M tendrá entonces la forma $M = \begin{pmatrix} I_r & 0 & \alpha' \\ 0 & R & \alpha^n \\ 0 & 0 & 1 \end{pmatrix}$ en

donde es

$$\alpha = \begin{pmatrix} \alpha' \\ \alpha'' \end{pmatrix} ,$$

y en que α' es una columna de r términos y α'' es una columna de $2p = n - r$ términos.

Se tiene, en el caso que consideramos, que $det\ M = 1$, y por tanto f *es un movimiento.*

La matriz M se puede escribir en la forma

$$M = \begin{pmatrix} I_r & 0 & \alpha' \\ 0 & I_{n-r} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_r & 0 & 0 \\ 0 & R & \alpha'' \\ 0 & 0 & 1 \end{pmatrix}$$

En virtud del teorema 1.4.7 puede escribirse

$$f = gh$$

donde g y h son isometrías cuyas matrices son

$$M_g = \begin{pmatrix} I_r & 0 & \alpha' \\ 0 & I_{n-r} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

y

$$M_h = \begin{pmatrix} I_r & 0 & 0 \\ 0 & R & \alpha'' \\ 0 & 0 & 1 \end{pmatrix}$$

En este caso, en que la matriz A es de la forma (a), podemos distinguir dos subcasos, a saber:

Subcaso (a1). $r = n$. M_h se reduce a la matriz unidad I_{n+1}, y por consiguiente es la transformación afín idéntica. Se tiene por tanto que $f = g$ *es una traslación* cuyas componentes respecto de la base asociada al sistema de coordenadas que estamos empleando son

$$(a^1, \ldots, a^n).$$

En este subcaso se cumple entonces la relación $\overrightarrow{Pf(P)} = a$ para todo punto $P \in E$, es decir que $\overrightarrow{Pf(P)}$ es un vector constante.

Subcaso (a2): $0 \leqslant r < n$. M_b no se reduce a la matriz identidad y por tanto h no se reduce a la transformación idéntica de E. Vamos a determinar los puntos fijos de h. Las coordenadas $(x^1, \ldots, x^r)$ de un punto cualquiera pertenecen invariables por la isometría h. Por consiguiente para que un punto sea fijo sus coordenadas $(x^{r+1}, \ldots, x^n)$ deben satisfacer a las ecuaciones

$$\left.\begin{aligned} x^{r+2k-1}\cos\theta_k - x^{r+2k}\operatorname{sen}\theta_k + a^{r+2k-1} &= x^{r+2k-1} \\ x^{r+2k-1}\operatorname{sen}\theta_k + x^{r+2k}\cos\theta_k + a^{r+2k} &= x^{r+2k} \end{aligned}\right\} \quad (k = 1, \ldots, p),$$

que pueden escribirse en la forma

$$\left.\begin{aligned} (\cos\theta_k - 1)x^{r+2k-1} - (\operatorname{sen}\theta_k)x^{r+2k} &= -a^{r+2k-1} \\ (\operatorname{sen}\theta_k)x^{r+2k-1} + (\cos\theta_k - 1)x^{r+2k} &= -a^{r+2k} \end{aligned}\right\} \quad (22.3.2)$$

$$(k = 1, \ldots, p)$$

El determinante de cada uno de estos sistemas de dos ecuaciones con dos incógnitas es igual a

$$2(1 - \cos\theta_k) \neq 0;$$

por tanto cada uno de dichos sistemas tiene una sola solución

$$(x_o^{r+2k-1}, x_o^{r+2k}).$$

Resulta así que h tiene los puntos fijos

$$(x^1, \ldots, x^r, x_o^{r+1}, x_o^n),$$

donde $x^1, \ldots, x^r$ son arbitrarios. El conjunto de los puntos fijos constituye entonces un subespacio de dimensión r, que se llama *subespacio de rotación*, por analogía con el eje de rotación del espacio eu-

clideano de tres dimensiones. En el caso particular en que $r = 0$, h tiene un único punto fijo, a saber

$$(x_o^1, \ldots, x_o^n) \quad .$$

Resulta de lo dicho que en este subcaso (a2) existen puntos fijos respecto de la transformación h, y otros que no lo son. Supongamos que P_o es un punto fijo y que P_1 no es fijo. Si designamos por $\bar{a}$ al vector de la traslación g, tendremos

$$f(P_o) = g(h(P_o)) = g(P_o) \quad ,$$

y por consiguiente

$$\overrightarrow{P_o f(P_o)} = \overrightarrow{P_o g(P_o)} = \bar{a} \ ,$$

mientras que

$$f(P_1) = g(h(P_1)) \ , \text{ donde } h(P_1) \neq P_1 \quad ,$$

y por tanto

$$\begin{aligned} \overrightarrow{P_1 f(P_1)} &= \overrightarrow{P_1 h(P_1)} + \overrightarrow{h(P_1) f(P_1)} \\ &= \overrightarrow{P_1 h(P_1)} + \overrightarrow{h(P_1) g(h(P_1))} \\ &= \overrightarrow{P_1 h(P_1)} + \bar{a} \quad , \end{aligned}$$

de donde, por cuanto $\overrightarrow{P_1 h(P_1)} \neq 0$, se deduce que

$$\overrightarrow{P_1 f(P_1)} \neq \bar{a} \ ,$$

o sea

$$\overrightarrow{P_o f(P_o)} \neq \overrightarrow{P_1 f(P_1)} \ .$$

Se deduce así que en este subcaso, a diferencia del subcaso (a1) , el vector $\overrightarrow{Pf(P)}$ no es constante para todos los puntos P de E. Por consiguiente *f no es una traslación.*

Consideremos ahora el caso en que la matriz A es de la forma (b). La matriz M puede escribirse entonces en la forma

$$M = \begin{pmatrix} I_{r-1} & 0 & 0 & \alpha' \\ 0 & -1 & 0 & a^r \\ 0 & 0 & R & \alpha'' \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

donde hemos escrito a la columna α en la forma

$$\alpha = \begin{pmatrix} \alpha' \\ a^r \\ \alpha'' \end{pmatrix}$$

en que α' es una columna de $r-1$ términos y α'' es una columna de $2p = n - r$ términos. R es la misma matriz de orden $2p$ que hemos considerado anteriormente. Se ve inmediatamente que $det\ M = -1$, y por tanto f *no es un movimiento.*

M puede ser escrita en la forma

$$M = \begin{pmatrix} I_{r-1} & 0 & 0 & \alpha' \\ 0 & 1 & 0 & 0 \\ 0 & 0 & I_{n-r} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_{r-1} & 0 & 0 & 0 \\ 0 & 1 & 0 & a^r \\ 0 & 0 & I_{n-r} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_{r-1} & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & I_{n-1} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I_{r-1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & R & \alpha'' \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

(22.3.3)

y por consiguiente se tiene

$$f = g_2 g_1 k h$$

donde g_2 , g_1 , k y h son isometrías cuyas matrices son, respectivamente, los cuatro factores del segundo miembro de la relación (22.3.3). Se ve por tanto que:

g_2 es una traslación que cambia a las coordenadas x^i $(i = 1, ..., r)$ en $x^i + a^i$ y deja invariables a las demás coordenadas;

g_1 es una traslación que cambia a las coordenadas x^r en $x^r + a^r$,

y deja invariables a las demás coordenadas;

k es una reflexión respecto del hiperplano $x^r = 0$, que cambia a la coordenada x^r en $-x^r$ y deja invariables a las demás coordenadas;

h es una isometría idéntica a la que hemos considerado precedentemente bajo el mismo nombre; pero, en el caso presente, en que es necesariamente $r \geqslant 1$, tiene los puntos fijos

$$(x^1, \ldots, x^r, x_o^{r+1}, \ldots, x_o^n) ,$$

donde $x^1, \ldots, x^r$ puede tomar valores arbitrarios, y los números $x_o^{r+1}, \ldots, x_o^n$ son obtenidos como soluciones de los sistemas (22. 3.2).

En el caso que ahora consideramos, en el que suponemos que A tiene la forma (b), podemos distinguir también dos subcasos:

Subcaso (b1): *o bien* $r = 1$ *o bien* $\alpha' = 0$. En este subcaso f se reduce a

$$f = g_1 kh .$$

Es fácil comprobar que esta isometría *posee los puntos fijos*

$$(x^1, \ldots, x^{r-1}, \frac{a^r}{2}, x_o^{r+1}, \ldots, x_o^n) ,$$

donde $x^1, \ldots, x^{r-1}$ son arbitrarios, y esto ocurre para $\alpha' = 0$; o bien cuando $r = 1$, posee el único punto fijo

$$(\frac{a^1}{2}, x_o^2, \ldots, x_o^n) .$$

Subcaso (b2): $r \geqslant 2$ *y* $\alpha' \neq 0$. Se tiene que

$$f = g_2(g_1 kh) .$$

La isometría $g_1 kh$ deja invariantes a las coordenadas $x^1, \ldots, x^{r-1}$, y la traslación no trivial g_2 cambia necesariamente a alguna de ellas Por tanto f *no tiene puntos fijos.*

La discusión anterior se resume como sigue:

Caso (a)

A es del tipo (a)
f es un movimiento

- *Subcaso (a1):* $r = n$ — f es una traslación
- *Subcaso (a2):* $0 \leqslant r \leqslant n$ — f no es una traslación

Caso (b)

A es del tipo (b)
f no es un movimiento

- *Subcaso (b1):* $r = 1$ o bien $\alpha' = 0$ — f tiene puntos fijos
- *Subcaso (b2):* $r \geqslant 2$ y $\alpha' \neq 0$ — f no tiene puntos fijos

Definición 22.3.1. Dos isometrías f_1 y f_2 de un espacio euclideano E son equivalentes cuando existe una tercera isometría g de E tal que se cumple la relación

$$f_2 = g f_1 g^{-1} .$$

La relación así definida es una relación de equivalencia.

Teorema 22.3.1. Dadas dos isometrias equivalentes f_1 y f_2 si alguna de las propiedades siguientes se cumple para una de ellas, entonces esa propiedad se cumple también para la otra:

a) es un movimiento
b) es una traslación
c) es una rotación
d) transforma a una recta en sí misma

Demostración. a) Si M_1 y M_2 son las matrices de f_1 y f_2 respecto de un sistema cualquiera de coordenadas, se tiene evidentemente

$$det\ M_1 = det\ M_2 .$$

Por tanto, o bien f_1 y f_2 son movimientos, o bien ambas no lo son.

b) Supongamos que f_1 es una traslación de vector a y que f_2 es

equivalente a f_1, de manera que existe la isometría g tal que

$$f_2 = gf_1g^{-1} .$$

Cualquiera que sea el punto P de E se tiene que

$$\overrightarrow{Pf_1(P)} = a .$$

También se tendrá, por consiguiente, cualquiera que sea el punto P de E,

$$\overrightarrow{g^{-1}(P)f_1(g^{-1}(P))} = a \qquad (22.3.4)$$

Designemos por T al endomorfismo de V asociado con la isometría g de E. Si aplicamos dicho endomorfismo a los dos miembros de (22.3.4), y tenemos en cuenta que

$$T(\overrightarrow{g^{-1}(P)f_1[g^{-1}(P)]}) = T(\overrightarrow{g^{-1}(P)g^{-1}(f_2(P))})$$
$$= \overrightarrow{Pf_2(P)} ,$$

resultará que

$$\overrightarrow{Pf_2(P)} = Ta ,$$

es decir que f_2 es una traslación de vector Ta.

c) Si O es un punto fijo de f_1, entonces

$$f_2(g(O)) = (gf_1g^{-1})(g(O)) = gf_1(O) = g(O) ,$$

y por tanto $g(O)$ es un punto fijo de f_2. Los papeles de f_1 y f_2 pueden permutarse si se sustituye $g(O)$ por $g^{-1}(O)$.

d) Las condicioness necesarias y suficientes para que una isometría f transforme a sí misma a una recta asociada a un vector $x_o \in V$ son evidentemente las dos siguientes:

1. que exista un punto $P_o \in E$ tal que sea

$$\overrightarrow{P_of(P_o)} = \lambda_1 x_o \quad , \quad \lambda_1 \in \mathbb{R} ,$$

2. que si T es un automorfismo asociado a f sea

$$Tx_o = \lambda_o x_o \quad , \quad \lambda_o \in \mathbb{R} .$$

La recta que pasa por P_o y que es asociada con x_o se dice que es *recta fija de f* y se designa por $(P_o\ ,\ x_o)$.

Supuesto que $f_2 = gf_1g^{-1}$, sean T_1 , T_2 , T_3 los automorfismos asociados con f_1 , f_2 y g respectivamente.

Supongamos que f_1 admite a la recta fija $(P_o\ ,\ x_o)$. Puesto que, según el teorema 21.4.7., se cumple que

$$T_2 = T_3T_1T_3^{-1} \quad ,$$

si designamos por x'_o al vector T_3x_o , tendremos

$$T_2x'_o = T_2(T_3x_o) = T_3(T_1x_o) = T_3(\lambda_o x_o) = \lambda_o T_3x_o = \lambda_o x'_o \qquad (22.3.5)$$

Si designamos por P'_o al punto $g(P_o)$ podremos escribir

$$\overrightarrow{P'_of_2(P'_o)} = \overrightarrow{g(P_o)f_2(g(P_o))} = \overrightarrow{g(P_o)g(f_1(P_o))} = T_3(\overrightarrow{P_of_1(P_o)})$$

$$= T_3(\lambda_1 x_o) = \lambda T_3x_o = \lambda_1 x'_o \ . \qquad (22.3.6)$$

Las relaciones (22.3.5) y (22.3.6) prueban que f_2 tiene a la recta fija $(P'_o\ ,\ x'_o)$. ∎

Del teorema anterior se deducen las siguientes consecuencias:

Una isometría perteneciente al caso (a) no puede ser equivalente a otra que pertenece al caso (b), pues la primera es un movimiento y la segunda no.

Una isometría que pertenece al subcaso (a1) no puede ser equivalente a otra que pertenece al subcaso (a2) pues la primera es una traslación y la segunda no.

Una isometría que pertenece al subcaso (b1) no puede ser equivalente a otra que pertenece al subcaso (b2) pues la primera tiene puntos fijos y la segunda no.

De lo dicho resulta que para que dos isometrías sean equivalentes es necesario que ambas pertezcan al mismo de los cuatro subcasos que hemos descrito. Esta condición no es sin embargo suficiente para que dos isometrías sean equivalentes. La proposición si-

guiente proporciona una condición necesaria y suficiente para que dos isometrías sean equivalentes.

Teorema 22.3.2. Para que dos isometrías sean equivalentes es necesario y suficiente que existan dos sistemas de coordenadas rectangulares, S y S', tales que la matriz que representa a la primera isometía en el sistema S sea idéntica a la matriz que representa a la segunda isometría en el sistema S'.

Demostración. Sean f y g las isometrias dadas. Si M_f, es la matriz de la primera referida al sistema S y M'_g es la matriz de la segunda referida al sistema S', supongamos primero que tiene lugar la igualdad

$$M_f = M'_g \quad .$$

Sea h la isometría que aplica a cada punto P, cuyas coordenadas en el sistema S son $(a^1, \ldots, a^n)$, en el punto P' que tiene las mismas coordenadas en el sistema S'. Es fácil comprobar entonces que se cumple la relación

$$f = h^{-1}gh \quad . \qquad (22.3.7)$$

Por consiguiente f y g son equivalentes.

Supongamos ahora, inversamente, que existe una isometría h tal que se cumple la relación (22.3.4). Entre las matrices que representan a las isometrías f, g, h y h^{-1} respecto de un sistema cualquiera S de coordenadas rectangulares tiene lugar la relación

$$M_f = M_{h^{-1}} M_g M_h = M_h^{-1} M_g M_h$$

Sea S' un nuevo sistema de coordenadas rectangulares tal que la matriz de cambio del sistema S en el sistema S' es M_h^{-1}. El teorema 21.4.10 implica que la matriz de g en el sistema S' es

$$M_h^{-1} M_g M_h \quad .$$

Por tanto es igual a M_f o sea a la matriz de f en el sistema S. ∎

Ejercicios

22.3.1. Aplicar las consideraciones de la sección 22.3 a la clasificación de las isometrias del plano euclideano y del espacio euclideano tridimensional.

22.3.2. Determinar la clase a que pertenece cada una de la siguientes isometrias del plano $\mathbb{R}^2$.

a) $\begin{aligned} x' &= -x + a \\ y' &= -y + b \end{aligned}$

b) $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

c) $\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 2 \\ 1/\sqrt{2} & 1/\sqrt{2} & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$

22.3.3. La misma cuestión para las siguientes isometrias de $\mathbb{R}^3$.

a) $\begin{aligned} y^{1'} &= x^1 + a^1 \\ y^{2'} &= -x^2 + a^2 \\ y^{3'} &= x^3 \end{aligned}$

b) $\begin{pmatrix} y^{1'} \\ y^{2'} \\ y^{3'} \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 & a \\ \operatorname{sen}\theta & \cos\theta & 0 & b \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \\ 1 \end{pmatrix}$

c) $\begin{pmatrix} y^{1'} \\ y^{2'} \\ y^{3'} \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} & 1 \\ 0 & 1 & 0 & -1 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \\ 1 \end{pmatrix}$.

22.3.4. En el plano euclideano $\mathbb{R}^2$, referido al sistema canónico $(x\ ,\ y)$ de coordenadas, obtener las ecuaciones de la reflexión respecto de la recta $2x - y = 1$.

22.3.5. En el espacio $\mathbb{R}^3$ referido a las cordenadas canónicas $(x,\ y\ ,\ z)$, hallar las ecuaciones de la reflexión respecto del plano $x + y - z = 0$.

22.3.6. Obtener las condiciones para que una isometría perteneciente a uno de los cuatro subcasos considerados en la sección 22.3 sea equivalente a otra isometría perteneciente al mismo subcaso.

INDICE

Adjunción, 168, 177
Algebra, 54
Algebra asociativa, 55
Algebra de las matrices cuadradas, 72
Algoritmo de Euclides, 153
Algoritmo de la división, 147
Angulo, 400
Anillo, 133
Anillo cociente, 139
Anillo de Ideales principales, 149
Antisimetría, 83
Aplicación afín, 369
bilineal, 36
centroafín, 374
inyectiva, 29
regular, 29
Automorfismo, 26

Baricentro, 398
Base, 15
Base canónica, 27
dual, 45
ortonormal, 298
recíproca, 47, 315

Cadena, 19
Cambio de base, 74
Campo, 1
Campo algebraicamente cerrado, 174
de cocientes, 142
de funciones racionales, 175
de números algebraicos, 179
de números reales algebraicos, 179
Característica, 4
Cero, 2, 133
Cero de un polinomio, 162
Cerradura algebraica, 174
Codimensión, 25
Coeficiente director, 143
Cofactor, 84, 108
Columna, 58
Columnas de base, 92
Combinación lineal, 8
Complejificación, 191
Complemento algebraico, 108
ortogonal, 40, 312
Componentes, 7, 16
Componentes contravariantes, 315

covariantes, 315
Condiciones de compatibilidad, 125
Conjugación, 310
Conjunto convexo, 395
ortonormal, 298
parcialmente ordenado,19
Cono, 397
Coordenadas Baricéntricas, 395
Cota superior, 19
Criterio del rango de la matriz aumentada, 119

Dependencia lineal, 9, 91
Descomposición espectral, 331, 340 342, 346, 350
Desigualdad de Bessel, 320
de Buniakovsky, 292
de Cauchy, 294
de Cauchy–Schwarz,292
de Schwarz, 292
Determinante, 82
Determinante de un endomorfismo, 106
Diagonal principal, 63
Dimensión, 22, 55
Dimensión finita, 22
Dimensión infinita, 22
Disjuntos (espacios), 12
Distancia, 399
Divisibilidad de polinomios, 149
División de polinomios, 149
Divisor, 149
Divisores elementales, 285
Dominio de integridad, 140
Dualidad, 36, 43

Ecuación normal, 405
Elemento algebraico, 166
invertible, 55
maximal, 19
trascendente, 166
unidad, 55
Endomorfismo, 26
Endomorfismo autoadjunto, 333
hermitiano, 333, 342
normal, 326
nilpotente, 242
ortogonal, 339
simétrico, 333, 346
unitario, 339, 343
Endomorfismos semejantes, 79
Enteros de Gauss, 178
Escalar, 6
Espacio afín, 361
cartesiano, 7
conjugado, 38, 45
dual, 37
euclideano, 290, 399
métrico, 400
propio, 200
unitario, 290
vectorial, 5
vectorial complejo, 6
vectorial factor, 12
vectorial real, 6
Espacios disjuntos, 12
Extensión algebraica, 168
de un campo, 165
de un espacio vectorial, 181
finita, 167

simple, 168
simple algebraica, 168
simple trascendente, 168
unitaria, 359

IF , 1
Factorización de polinomios, 159
Familia maximal l.i. , 22
Familia minimal de generadores, 22
IF-automorfismo, 181
-endomorfismo, 181
-homomorfismo, 181
Fila, 58
IF-isomorfismo, 181
Forma canónica de Jordan, 253, 264, 269
lineal, 26
Fórmula de Laplace, 110
Función alternada, 100
bilineal, 36, 53
lineal, 26
multilineal, 100
no degenerada,36, 290
polinomial, 161
positiva definida, 290
racional entera, 161
simétrica, 290

Generadores, 8, 168
Grado de un polinomio, 143
de la extensión de un campo, 167
Grupo afín, 378
de las matrices regulares, 73
lineal, 58
lineal general, 73
ortogonal, 352
unitario, 344

Homomorfismo canónico, 14, 139
de espacios vectoriales, 26
de anillos, 136

Ideal, 134
Ideal cero, 134
Ideal maximal, 139
primo, 143
principal, 135
Identidad, 2 , 133
Igualdad de Parseval, 320
Im, 30
Imagen , 30
Identidad de Parseval, 303
Independencia lineal, 9
Indice, 242
Inversa, 2, 55
Involución, 57
Isometría, 334, 406, 410
Isomorfismo, 26, 33, 136
Isomorfismo inducido, 34

Ker, 29

Lema de Zorn, 19
l.d., 9
l.i., 9
Longitud, 292

Matrices semejantes, 78
Matriz, 58
Matriz adjunta, 305
aumentada, 120
autoadjunta, 333
conjugada, 305
de Jordan, 255
de una transformación lineal, 59
de un sistema de ecuaciones, 113
diagonal, 73
inversa, 86
invertible, 102
hermitiana, 333
normal, 326
ortogonal, 306
regular, 73
simétrica, 333
transpuesta, 63
triangular superior, 74
unitaria, 306
M.C.D. de polinomios, 152
M.C.M. de polinomios, 154
Menor complementario, 107
Menor de base, 92
de un determinante, 92, 107
principal, 108
Movimiento, 411
Multiplicidad, 207
Múltiplo, 149

Norma, 292
Núcleo, 29

Opuesto, 2
Ortogonalidad en espacios duales, 37, 40, 48.
en espacios de producto interno, 312

Paralelismo, 388
Paralelepípedo, 397
Parcialmente ordenado, 19
Polinomio, 143
Polinomio característico, 202, 203
irreducible, 158
minimal, 166, 209
mónico, 149
Posición general, 397
Primo (campo), 5
Proceso de ortogonalización, 300
Producto de endomorfismos, 55
de matrices, 69
de polinomios, 144
escalar, 37
interno, 289
interno ordinario, 291
Propiedad hermitiana, 290
Proyección, 57, 139, 317

Raices características, 206
simples, 234
Raiz de una ecuación, 162
Rango de una matriz, 67, 96
de una transformación, 30
de un conjunto de vectores, 95

Regla de Cramer, 90
Relación lineal, 9
Relación de ortogonalidad, 118
Restricción, 180
Rotación, 411

Segmento, 395
Semiespacio, 396
Sesquilineal, 290
Símbolo de Kronecker 27
Simplejo, 397
Sistema compatible, 125
complementario, 125
de coordenadas baricéntricas, 395
de coordenadas cartesianas, 362
de Cramer, 120
de ecuaciones lineales,112
de ecuaciones l.i., 121
fundamental de soluciones, 131
homogéneo, 23, 113
homogéneo asociado, 115
inhomogéneo, 113
invariante de un endomorfismo nilpotente, 252
reducido, 125
Sobreyección, 30
Subcampo, 3, 168
Subcampo propio, 4
Subespacio generado, 12
invariante, 195
vectorial, 12
Sucesión de cuadrados sumables,296
Suma de polinomios, 143
directa de espacios, 24, 212
directa de transformaciones lineales, 225

Teorema de Cayley-Hamilton, 224, 229
de existencia, 115, 189
de intercambio de Steinitz, 17.
de las columnas de base,92
de Riesz, 313
de unicidad, 116
Transformación de extensión, 182
dual, 41
IF- lineal, 181
idéntica, 32
inversa, 32
lineal, 26
lineal adjunta, 323
lineal regular, 26
Traslación, 373
Traza, 204
Trivial, 8

Valor propio, 196
Vector, 6
Vector-columna, 59
-fila, 60
fijo, 28
invariante, 28
normal, 402
propio, 196

Vector solución, 114
segundo miembro, 114
unitario, 292.

$\mathbb{F}$, 1
$\mathbb{R}$, 3
$\mathbb{Q}$, 3
$\mathbb{C}$, 3
l.i., 9
l.d., 9
$U + V$, 12
V/V_1 , 14
$U \oplus V$, 25, 212
dim, 25
Codim, 25
$\mathcal{L}(U, V)$, 26
Hom (U, V), 26
$\mathcal{L}(U)$, 26
$\mathcal{L}(U)$, 26
δ_j^i, 27
Ker T, 29
Im T, 30
$r(T)$, 30
T^{-1} , 32
$U \cong V$, 33
$\langle x^*, x \rangle$, 36
$\langle \ , \ \rangle$, 37
$V^{\perp}$, 40, 312
φ^* , 41
x^{-1} , 55
GL (V), 56
M(T), 59
$\mathbb{M}(n \times m)$, 60
M^t, 63
$r(M)$, 67
$GL(n, \mathbb{F})$, 73
$\mathbb{M}(n)$, 73
ϵ_σ , 82
det M, 82
A_i^j , 84
D_i^j , 84
α_K^J , 103
α_K , 103
α^J , 103
$I(a)$, 135
$\mathbb{F}[x]$, 144
$d|f$, 149
M.C.D., 152
$[f_1, \ldots, f_n]$, 152
M.C.M., 154
$\{f_1, \ldots, f_n\}$, 154
$(\mathbb{F}' : \mathbb{F})$, 167
$\mathbb{F}(\theta)$, 168
$\mathbb{F}(S)$, 177
$\mathbb{Z}(i)$, 178
$V^{\mathbb{C}}$, 191
$\bar{z}$, 191
$M(T, \mathcal{E})$, 195
$f_M(x)$, 202
$f_T(x)$, 203
$\phi_T(x)$, 209
$T_1 \oplus T_2$, 225
E_n , 232
$J(\lambda ; k)$, 255
$J(\lambda ; k_1, \ldots, k_p)$, 255
$D_j(x)$, 275
$E_j(x)$, 285
(,), 289
$\|x\|$, 292
$\widetilde{M}$, 305
$(T_1 + T_2)^{\sim}$, 325
$(T_1 T_2)^{\sim}$, 325
$U(n)$, 344
$SU(n)$, 344
$O(n)$, 352
$SO(n)$, 352
$\overrightarrow{AB}$, 361
$A(n, \mathbb{F})$, 378
$\vee$, 387
$d(A, B)$, 399.

ERRATAS NOTABLES

Las líneas cuyos numeros de orden aparecen subrayados deben ser contadas a partir de abajo.

Pág.	Línea	Dice	Debe decir
12	8	$\{x / x = u + v\ ,\ u \in V\}$	$\{x / x = u + v\ ,\ u \in U,\ v \in V\}$
118	1	$\sum_{i=1}^{n}$	$\sum_{i=1}^{2}$
122	1	$\sum_{s=r+1}^{m} D_{o}^{j}\ \xi^{s}$	$\sum_{s=r+1}^{m} D_{s}^{j}\ \xi^{s}$
139	10	$a \in A$	$a \in J$
161	9	pero los factores $(x - a_1)$ son irreducibles, luego	luego
162	11	$\sum_{i+j=0}$	$\sum_{i+j=k}$
162	12	$\sum_{i=1}^{n}$	$\sum_{i=0}^{n}$
206	7	aplicada	ampliada
209	5	es el anillo	en el anillo
212	2	En particular, si V es de dimensión finita, entonces se tiene	En particular, si V es de dimensión finita y V_1 es un subespacio, existen subespacios V_2 de V, que se llaman *Complementos de V_1 en V*, tales que es $V = V_1 \oplus V_2$, y se tiene
215	15	$\phi_1(x)(\phi_2(x))$	$\phi_1(x)$(resp. $\phi_2(x)$)
219	1	$\overline{V}$	V
227	4	material	matricial

Pág.	Línea	Dice	Debe decir
244	1	cualquier complemento en ese subespacio	cualquier complemento de H en V es decir tal que $V = H \oplus K$.
244	6	$\overline{T}^{k-2}(\overline{T}x_o) = \overline{T}^{k-1}x_o \neq 0$	$\overline{T}^{k-2}(Tx_o) = T^{k-1}x_o \neq 0$
246	6	K	$K = K'_o \oplus K_o$
250	8	$\{T^{k_r+1}x_i, T^{k_r+1+1}x_i; 1 \leqslant i \leqslant r\}$	$\{T^{k_r+1}x_i, T^{k_r+1}x_{i+1}, \ldots, T^{k_i-1}x_i, 1 \leqslant i \leqslant r\}$
282	6	λ_1	λ_i
296	1	$a A =$	$A =$
298	2	Horáth	Horváth
301	2	$\frac{1}{\sqrt{2}}$	$\frac{y_1}{\sqrt{2}}$
315	1	(Se usarán los símbolos	(Se usarán los símbolos g_{ij} y g^{ij} para representar a los productos internos
339	6	$M(T)M(\widetilde{T})M(T) = 1$	$M(T)M(\widetilde{T}) = M(\widetilde{T})M(T) = 1$
339	9	se cumple que $M(\widetilde{T})M(T) = 1 \Rightarrow T\widetilde{T} = I \Rightarrow T\widetilde{T} = I$. ■	se cumple que $M(\widetilde{T})M(T) = 1$, de donde $M(\widetilde{T})M(T) = 1$, y por tanto $\widetilde{T}T = T\widetilde{T} = I$. ■
369	12	$\overline{T(\overrightarrow{PQ})}$	$T(\overrightarrow{PQ})$
375	12	$\overline{T(\overrightarrow{OP})}$	$T(\overrightarrow{OP})$
381	12 y 14	$\lambda\overline{T(\overrightarrow{OP})}$	$\lambda T(\overrightarrow{OP})$
409	9	sustituimos con $\cos \measuredangle PQR$	sustituímos $\cos \measuredangle PQR$
411	2	sistema	teorema
416	7	pertenecen	permanecen
418	3	x^i $(i = 1, \ldots, r)$	x^i $(i = 1, \ldots, r-1)$
421	6	a sí misma	en sí misma.

Carátula: **Víctor Cumpa**

Impresión - Carátula:
Industrial Gráfica S.A.