

Álgebra Lineal

Volumen I

Transformaciones lineales

Jorge Luis Arocha

versión compilada el
17 de mayo de 2015

Jorge Luis Arocha Pérez
Instituto de Matemáticas
Universidad Nacional Autónoma de México
México D.F. 04510

BibTeX:

```
@textbook{ArochaLA
  AUTHOR  = {Arocha, Jorge L.}
  TITLE   = {\'Algebra Lineal}
  YEAR    = {2014}
  NOTE    = {Available at combinatoria.matem.unam.mx}
}
```

Mathematics Subject Classification 2010: 00–01, 12–01, 15–01

Introducción

Este libro lo escribí como texto básico para el curso de “Álgebra Lineal” para estudiantes de Licenciatura que he impartido por muchos años en la Facultad de Ciencias de la Universidad Nacional Autónoma de México.

Se presupone que los estudiantes hayan cursado ya las materias de “Álgebra superior” y “Geometría Analítica” o sea tengan ciertos conocimientos sobre matrices, vectores, polinomios, números complejos, etc.

El objetivo es desarrollar el álgebra lineal sobre campos arbitrarios pero se hace énfasis en los reales, los complejos y los residuos módulo un número primo. Después de una introducción corta a la teoría de campos se estudian los espacios vectoriales, las transformaciones lineales, los determinantes y finalmente los teoremas de descomposición de operadores.

Por ahora, aquí no hay prácticamente nada de transformaciones bilineales, productos escalares, espacios duales, ortogonalidad, tensores etc. En mis planes está escribir un segundo volumen o aumentar este libro con capítulos sobre estos temas.

El material desarrollado es demasiado para un semestre y usualmente yo imparto en un semestre los capítulos I—IV (aquellas secciones que no están marcadas como avanzadas). Un segundo semestre podría comenzar con las secciones de polinomios sobre campos, continuar con la descomposición de operadores lineales y terminar con aquellos temas que ya señalé, faltan aquí. Otras secciones las escribí porque me parecieron un buen material de lectura complementaria para los alumnos curiosos.

Una particularidad de la exposición es que para mí, las matrices no tienen orden, por ejemplo, las matrices de cambio de base están naturalmente indexadas por conjuntos de vectores y los conjuntos de vectores no tienen un orden natural. Como resultado de esto este libro es fundamentalmente conceptual y no computacional.

El libro es inusualmente colorido y visualmente agresivo. La razón de esto es que cuando estaba en el papel de alumno yo prefería estudiar por las libretas de notas de mis compañeras de clase. Se me hacía muy fácil memorizar la imagen completa de una página llena de rayas, flores, corazones etc. y dentro de todo esto, las matemáticas. La idea es que cada página luzca visualmente diferente. He tratado dentro de lo posible, lograr esto.

Los caracteres de matemáticas están en un color diferente. El texto y las secciones avanzadas están marcados con un birrete. Uso unos lentes para marcar aquello que el lector no debe dejar pasar. Errores comunes que cometen los que no están familiarizados con el material están marcados con calaveras. Los teoremas están resaltados visualmente, etc.

Se incluyen más de un centenar de ejercicios. Los ejercicios más notables consisten en material adicional que un estudiante de matemáticas o física debe conocer tarde o temprano. Las soluciones de los ejercicios no rutinarios se dan al final del libro.

He compilado un glosario de términos que es a la vez un índice del libro y un diccionario de los conceptos introducidos y/o usados.

Incluyo una guía de estudio. De esta guía yoescojo las preguntas para los exámenes. Esto representa una ventaja enorme para los estudiantes ya que una de las dificultades

más importantes de ser estudiante es comprender que es lo que se espera de ellos.

Finalmente, quiero agradecer a mis colegas Javier Bracho, Francisco Larrión y Omar Antolín que han influenciado de una u otra manera a esclarecer mis ideas sobre el tema y que contribuyeron substancialmente a hacer de este, un libro mejor. Muchos de mis estudiantes más notables han encontrado errores en mis notas de clase, mi agradecimiento a todos ellos.

Contenido

Capítulo 1 Campos	1
1.1 Operaciones binarias	1
Commutatividad (3). Asociatividad (3). Elementos neutros (4). Elementos inversos (4). Distributividad (5). El álgebra “abstracta”(5).	
1.2 Números	6
Naturales (6). Enteros (6). Grupos (7). Anillos (7). Racionales (8). Reales (8). Complejos (9).	
1.3 Morfismos	10
Morfismos de grupos (10). Morfismos de anillos (11). Isomorfismos (12). Composición de morfismos (13).	
1.4 Campos de restos	14
El anillo de los enteros módulo n (14). Dominios de integridad (15). El campo de los enteros módulo p (16).	
1.5 Campos primos. Característica	17
Campos primos (17). Característica (19).	
1.6 Aritmética de campos	19
Múltiplos y exponentes enteros (19). Asociatividad general (19). Distributividad general (20). Fórmula multinomial (20). La expansión de $\prod \sum \alpha_{ij}$ (21).	
*1.7 Anillos con división	22
Quaterniones (23). Caso finito (24).	
Capítulo 2 Espacios vectoriales	27
2.1 El plano cartesiano	27
2.2 Definición y ejemplos	28
El espacio de n -adas \mathbb{K}^n (29). El espacio de polinomios $\mathbb{K}[x]$ (30). El espacio de sucesiones $\mathbb{K}^{\mathbb{N}}$ (30). El espacio de series $\mathbb{K}[[x]]$ (30). El espacio de funciones $\mathbb{K}^{\mathbb{N}}$ (30). El espacio de N -adas \mathbb{K}^N (31). El espacio de N -adas finitas $\mathbb{K}^{[N]}$ (31). Subcampos (32). El espacio de N -adas de vectores \mathbb{E}^N (32). El espacio de NM -matrices \mathbb{K}^{NM} (32). El espacio de tensores (33).	
2.3 Subespacios	34
Unión e intersección de subespacios (34). Combinaciones lineales (35). Cerradura lineal (36).	
2.4 Bases	37
Conjuntos generadores (38). Conjuntos linealmente independientes (38). Bases (39). Dimensión (41). Bases canónicas (43).	
2.5 Clasificación de espacios vectoriales	44
Isomorfismos lineales (44). Coordinatización (45). Clasificación (46). Campos de Galois (46). Como pensar en espacios vectoriales (47).	
2.6 Suma de subespacios	48

Subespacios de \mathbb{R}^n (48). Suma de conjuntos y subespacios (49). La igualdad modular (49). Suma directa (50). Isomorfismo canónico entre la suma y la suma directa. (51). Subespacios complementarios (52). Espacios vectoriales versus conjuntos (53).

2.7 Espacios cocientes	54
Subespacios afines (54). El espacio cociente (56). El isomorfismo con los complementarios (56).	
*2.8 El espacio afín	57
La regla del paralelogramo (58). Cerradura afín (58). Generadores e independencia (59). Bases afines (59).	
*2.9 El caso de dimensión infinita	61
El Lema de Zorn (61). Existencia de bases (61). Cardinales (62). Equicardinalidad de las bases (63).	

Capítulo 3 Transformaciones lineales

3.1 Definición y ejemplos	65
Imágenes de subespacios (65). Homotecias (66). Inmersiones (67). Proyecciones (67).	
3.2 Operaciones entre transformaciones lineales	68
El espacio vectorial de las transformaciones lineales (68). Composición de transformaciones lineales (69). El álgebra de operadores lineales (70). El grupo general lineal (71).	
3.3 Extensiones lineales	71
Extensiones y restricciones (71). El isomorfismo entre \mathfrak{F}^N y $\text{Mor}(\mathfrak{E}, \mathfrak{F})$ (73). Un criterio de isomorfismo (73).	
3.4 Coordinatización de transformaciones lineales	74
El producto escalar canónico (75). El producto de matrices (76). Productos de matrices y vectores (76). La transformación lineal de una matriz (77). La matriz de una transformación lineal (77). Composición de TLs y producto de matrices (78). Matrices inversas (79).	
3.5 Cambios de base	80
Cambios de base en un espacio vectorial (80). Cambios de base en el espacio de transformaciones lineales (81). Cambios de base en el espacio de operadores lineales (82).	
3.6 El núcleo y la imagen de una TL	82
Definiciones (82). Transformaciones lineales con núcleo trivial (83). Descomposición de transformaciones lineales (83). Un criterio de isomorfismo (84). Descomposición canónica de transformaciones lineales (85).	
*3.7 Trasformaciones semilineales y coalineaciones	86
Trasformaciones semilineales reales (86). Propiedades de las transformaciones semilineales (87). Automorfismos semilineales. (87). Coalineaciones (88). Estructura de las coalineaciones (90).	

Capítulo 4 Determinantes

4.1 Permutaciones	93
El grupo simétrico (93). Ciclos y órbitas (94). El grupo alternante (95). El signo de una permutación (97).	

4.2 Determinantes. Propiedades básicas	97
Definición de los determinantes (98). Determinantes de matrices pequeñas (98). El determinante de la identidad (99). Matrices con filas nulas (100). El determinante de la transpuesta (100). El determinante del producto (100). Matrices con filas iguales (101). Matrices de permutaciones (102). Permutaciones de columnas y renglones (103).	
4.3 Expansión de Laplace	104
Cambios de índices (104). Complementos algebraicos (106). La expansión de un determinante por sus renglones (106). La expansión de Laplace en forma gráfica (107). Multinearidad de los determinantes (108). La inversa de una matriz (110). El determinante de un operador lineal (111).	
*4.4 La expansión generalizada de Laplace	112
Matrices diagonales y triangulares por bloques (113). La expansión generalizada de Laplace en forma gráfica (114).	
4.5 El rango de una matriz	116
Matrices no singulares (116). Espacios de columnas y renglones (116). Lema de aumento de matrices no singulares (117). Bases de una matriz (118).	
4.6 Sistemas de ecuaciones lineales	119
Regla de Cramer (120). Existencia de soluciones (121). Eliminación de ecuaciones dependientes (121). El núcleo y la imagen de una matriz (122). Bases del subespacio afín de soluciones (122).	
4.7 Método de eliminación de Gauss	123
Transformaciones elementales (124). Ejemplo (125). El caso general (125). Solución de ecuaciones matriciales, matriz inversa (126).	
Capítulo 5 Polinomios	129
5.1 Polinomios sobre campos	129
Suma y producto de polinomios (129). La función de evaluación (130). División de polinomios (131). Divisibilidad (131). Factores y raíces (133). Ideales de polinomios (134). Unicidad de la factorización en irreducibles. (135). El conjunto ordenado de polinomios mónicos (136). Desarrollo de Taylor (137).	
*5.2 Polinomios complejos. Teorema de Gauss	138
Forma polar. Igualdad de Moivre (139). Continuidad (140). Límite de sucesiones complejas (141). Teorema de Gauss (142).	
5.3 Factorización de polinomios complejos y reales	143
Caso Complejo (144). Caso real (144).	
*5.4 Campos de fracciones. Funciones racionales	145
Campos de fracciones (146). Funciones racionales (147).	
Capítulo 6 Descomposición de operadores lineales	149
6.1 Suma directa de operadores lineales	149
Subespacios invariantes, componentes irreducibles (150). Ejemplos en dimensión 2 (152). Las matrices y los subespacios invariantes (152).	
6.2 Polinomios de operadores lineales	153

El morfismo de $\mathbb{K}[x]$ en $\text{End}(\mathfrak{E})$ (153). La subálgebra $\mathbb{K}[h]$ (155). El polinomio mínimo (155). El período de un vector (156). Anuladores (157). Propiedades del período. (157).

6.3 Subespacios radicales	158
Núcleos de polinomios de operadores lineales (158). Operadores lineales radicales (159). Componentes radicales (160). Existencia de un vector de período máximo (161).	
6.4 Subespacios cíclicos	162
h -combinaciones (162). Conjuntos h -generadores (162). Subespacios cíclicos (163). Conjuntos h -independientes (164). h -bases (165).	
6.5 Descomposición en subespacios cíclicos radicales.	166
El espacio cociente por un subespacio invariante (166). Polinomios y el espacio cociente (167). El período en el espacio cociente (168). Existencia de h -bases (169). Unicidad de la descomposición (170). Estructura de los operadores cíclico-radicales (173).	
6.6 Polinomio característico	174
Rectas invariantes (174). El polinomio característico de un operador lineal (174). El polinomio característico y el polinomio mínimo (176).	
6.7 Formas normales	178
Forma normal de Jordán (179). Forma normal real (180). Forma normal canónica (182).	
Soluciones de ejercicios selectos	185
Glosario	201
Notaciones	213
Guía de estudio	217

El álgebra es la oferta hecha por el Diablo a los matemáticos.

El Diablo dice:

*“Yo te daré a ti esta poderosa maquinaria
que responderá cualquier pregunta que tu quieras.*

*Todo lo que se necesita que tu hagas es entregarme tu alma:
dame la geometría y tendrás esta maravillosa máquina”*

...el daño a nuestra alma está ahí,
porque cuando usted pasa a hacer cálculos algebraicos,
esencialmente usted deja de pensar...

Sir Michael Atiyah

La forma correcta de leer las matemáticas consiste en primero leer las definiciones de los conceptos y las afirmaciones de los teoremas; luego, poner el libro a un lado y tratar de descubrir por si mismo las pruebas adecuadas.

*Si los teoremas no son triviales, el intento puede fallar,
pero es probable que de todas maneras sea instructivo.*

Para el lector pasivo un cómputo rutinario y un milagro de creatividad, se lean con la misma facilidad, y más tarde, cuando deba depender de sí mismo, se dará cuenta de que se fueron tan fácilmente como vinieron.

El lector activo, que se ha enterado de lo que no funciona, entiende mejor la razón del éxito del método del autor,

y después encontrará las respuestas que no están en los libros...

Paul Halmos

Capítulo primero

Campos

 El objeto del álgebra lineal es el estudio de los **espacios vectoriales**. Estos espacios son estructuras algebraicas cuyos objetos son de dos tipos los **vectores** y los **escalares**. Las operaciones definidas en los espacios vectoriales son la suma y resta de vectores, la suma resta multiplicación y división de escalares y la multiplicación de escalares por vectores. La mayoría de los temas estudiados en este libro no dependen del conjunto de escalares y el lector puede casi siempre considerar que los escalares son los reales \mathbb{R} y que el espacio vectorial es el espacio “geométrico” común \mathbb{R}^n .

Sin embargo, como esta teoría no depende (al menos en gran parte) del conjunto de escalares (y teniendo en cuenta diferentes aplicaciones a otras áreas de las matemáticas y las ciencias naturales) es conveniente elevar un paso el nivel de abstracción y pensar que el conjunto de escalares es un **campo** arbitrario \mathbb{K} .

El primer objetivo de este capítulo es dar al lector un conocimiento básico de lo que es un campo. Esto se pretende lograr por tres medios: dando la definición formal, estudiando algunas propiedades (como la característica) de los mismos, viendo que las reglas usuales de manipulación de fórmulas en un campo no se diferencian esencialmente de las fórmulas en los reales y sobre todo, dando los ejemplos fundamentales de campos.

1.1 Operaciones binarias

Sea A un conjunto. Una **operación binaria** es una función del producto cartesiano $A \times A$ en A . O sea, es una regla mediante la cual a cualesquiera dos elementos de A se le hace corresponder un tercer elemento de A . Demos algunos ejemplos sencillos:

1)	$a + b$	suma	5)	a^b	exponenciación
2)	$a - b$	resta	6)	$\log_a b$	logaritmo
3)	ab	producto	7)	$\text{mcd}(a, b)$	máx común divisor
4)	$\frac{a}{b}$	división	8)	$\text{mcm}(a, b)$	mín común múltiplo

Lo primero que observamos de los ejemplos anteriores es que no hemos definido en cual conjunto está definida la operación. Esto no es correcto formalmente, así por

ejemplo la división es una operación que no está definida en el conjunto de los números enteros. Sin embargo el lector podrá fácilmente encontrar los conjuntos en los cuales estos ejemplos son operaciones binarias.

Ejercicio 1 ¿En cuales conjuntos las operaciones 1-8 están correctamente definidas?

Ejercicio 2 ¿Que es una operación unaria? De ejemplos. [185]

Ejercicio 3 Dados tres números reales a, b, c definamos $\mathcal{A}(a, b, c)$ como el area del triángulo con lados a, b y c . ¿Es esta una operación ternaria en \mathbb{R} ? [185]

Lo segundo que debemos de observar, es la variedad de notaciones usadas para representar las operaciones binarias. Sobre todo, son complicadas las notaciones de la operaciones 4-6. Lo que tienen en común, es que no nos alcanza una línea de símbolos para escribirlas. Necesitamos subir y/o bajar además de movernos de derecha a izquierda. O sea, necesitamos dos dimensiones para escribirlas.

$$\int_0^{\pi/4} \left(a \sin x + b \sin \frac{x}{2} \right) dx$$

Quizá sea más ilustrativo, poner un ejemplo más complejo de **notación dos-dimensional**. La integral en el recuadro a la izquierda está bien definida para cualesquiera valores reales a, b y por lo tanto es una operación binaria en \mathbb{R} .

Más sencillos son los ejemplos de notaciones lineales 1-3,7-8. En realidad, para las notaciones lineales solo hay tres posibilidades:

- (a, b) notación prefija o funcional
- $a \circ b$ notación operacional
- $(a, b) \circ$ notación sufija

Las operaciones 1-3 están en notación operacional y las operaciones 7-8 están en notación prejija. La notación sufija es útil sobre todo en la programación de compiladores para lenguajes de computadoras (tales como pascal o C++) ya que frecuentemente lo más fácil es decirle a una computadora “toma el número a ”, “toma el número b ”, “súmalos” y no hacerlo de otra manera.

Ejercicio 4 La notación sufija para $a(b+c)/2$ es $bc+a\times 2\div$. ¿Cual será la notación sufija para la expresión $(a+b)(x+y)$? [185]

Cualquier intento, de tratar de unificar las notaciones usadas en la comunicación entre humanos, solo llevaría a confusiones mucho peores. Sin embargo, tenemos la libertad de escoger una notación unificada para las operaciones binarias abstractas que definamos. De una vez, postularemos que siempre usaremos la notación operacional para definir operaciones binarias abstractas.

Recalquemos que una operación “abstracta” no significa nada más que es una operación que puede ser una de muchas. Primero aprendemos lo que quiere decir $3+2$.

Después, tempranamente en el estudio de las matemáticas, la expresión $a + b$ significa que a un número a (no se sabe cual) le sumamos un número b (tampoco se sabe cual). Ahora, la expresión $a + b$ significará que a un objeto a (número, polinomio, matriz, quien sabe que) le “sumamos” (no se sabe lo que quiere decir “suma”) un objeto b (del que tampoco se sabe mucho).

Commutatividad

$$\forall a, b \in A \\ a \circ b = b \circ a$$

¿Es $3+2$ igual a $2+3$? Sí. ¿Es 3^2 igual a 2^3 ? No. Una operación binaria denotada por \circ y definida en el conjunto A se dice que es **commutativa** si se cumple la propiedad en el recuadro a la izquierda. Ser o no commutativa es la propiedad más sencilla que diferencia las operaciones binarias.

Ejercicio 5 ¿Cuales de las operaciones 1-9 son commutativas? [185]

Asociatividad

¿Que quiere decir $2+3+5$? ¿Acaso debemos sumar $2+3$ y al resultado sumarle 5 ? ¿No será que debemos sumar 2 al resultado de la suma $3+5$? Claro, no hay ninguna diferencia entre los dos procedimientos. Este hecho se expresa como $(2+3)+5 = 2+(3+5)$.

Una operación binaria denotada por \circ y definida en el conjunto A se dice que es **asociativa** si se cumple la propiedad en el recuadro de la derecha.

$$\forall a, b, c \in A \\ a \circ (b \circ c) = (a \circ b) \circ c$$

Los estudiantes preuniversitarios se encuentran por primera vez con la dificultad de una operación no asociativa en el caso de la operación de exponentiación. A esta temprana edad es muy necesario insistir que la expresión 2^{2^3} es ambigua porque $2^{(2^3)} = 256 \neq 64 = (2^2)^3$.

Ejercicio 6 ¿Cuales de las operaciones 1-9 son asociativas? [185]

La asociatividad de una operación es una propiedad crucial. Sin esta propiedad, el manejo algebraico de una operación se complica bastante.

Es más, gracias a ella podemos introducir la notación de la operación “repetida”. Si tenemos 11 elementos a_1, a_2, \dots, a_{11} entonces, para denotar la suma $a_1 + a_2 + \dots + a_{11}$ podemos usar la notación (mucho más cómoda) que se muestra en el recuadro a la derecha. Esta notación no requiere de la commutatividad de la operación “suma” gracias a que los índices tienen un orden y sabemos cual elemento debe ir primero y cual después.

$$\sum_{n=1}^{11} a_n$$

Si tenemos que la operación es no solamente asociativa sino también commutativa entonces podemos ser más generosos con esta notación.

Supongamos que a_ρ , a_ℓ , a_\varkappa , a_9 y a_∇ son elementos de un conjunto con una suma asociativa y conmutativa. Entonces la suma de estos (¡no importa el orden!) la podemos denotar por la expresión en el recuadro a la izquierda, donde \mathbf{N} es el conjunto de índices $\{\rho, \varkappa, \ell, \nabla, 9\}$.

Si la operación binaria definida no se llama “suma” sino “producto”, entonces es usual, en lugar de usar la letra griega \sum (sigma mayúscula), usar la letra griega \prod (pi mayúscula). Podemos, en este caso, usar la primera o segunda notación dependiendo de si nuestro producto es conmutativo o no.

Elementos neutros

La suma de números naturales tiene un elemento especial y único: el cero. Su propiedad definitoria es que cualquier número sumado con cero da el mismo número. La misma propiedad la tiene el uno con respecto al producto de números naturales.

Para una operación binaria denotada por \circ y definida en el conjunto A se dice que $e \in A$ es un **elemento neutro** si este cumple la propiedad en el recuadro.

$$\forall a \in A \\ a \circ e = e \circ a = a$$

Una operación binaria no puede tener más de un elemento neutro. Efectivamente, sean e y e' elementos neutros. Por ser e neutro, tenemos $e \circ e' = e'$. Por ser e' neutro, tenemos $e \circ e' = e$. De estas dos igualdades obtenemos $e = e'$.

Ejercicio 7 ¿Cuales de las operaciones 1-9 tienen neutro? [185]

Los elementos neutros juegan un papel importante en las notaciones para operaciones repetidas. Supongamos que tenemos un producto asociativo y conmutativo. Sean

$$\prod_{i \in N} a_i \bullet \prod_{i \in M} a_i = \prod_{i \in NUM} a_i$$

además \mathbf{N} y \mathbf{M} dos conjuntos finitos y disjuntos de índices. Naturalmente, de la definición se sigue la propiedad del recuadro a la izquierda.

Pero ¿que pasa si alguno de los conjuntos de índices (digamos M) es vacío? Si queremos que esta propiedad se conserve entonces observamos que

$$\prod_{i \in N} a_i \bullet \prod_{i \in \emptyset} a_i = \prod_{i \in N \cup \emptyset} a_i = \prod_{i \in N} a_i$$

por lo que necesariamente $\prod_{i \in \emptyset} a_i$ tiene que ser el elemento neutro de nuestra operación (si no hay neutro entonces estamos en problemas).

Es por esto, como el lector seguramente ya sabe, que la suma vacía de números es igual a cero y el producto vacío de números es igual a uno.

Elementos inversos

Para cada número entero a hay un único número $-a$ tal que sumado con a da cero. Generalizemos esta propiedad a operaciones binarias arbitrarias. Sea \circ una operación binaria en el conjunto A con elemento neutro. Se dice que $a \in A$ tiene **elemento inverso** b si se cumple la propiedad en el recuadro a la izquierda.

Para cualquier operación binaria asociativa el elemento inverso de otro es único. Efectivamente si b y c son inversos de a entonces $b = b \circ e = b \circ (a \circ c) = (b \circ a) \circ c = e \circ c = c$ o sea que b y c tienen que ser el mismo.

Ejercicio 8 Describa los inversos en las operaciones 1-9. [185]

Distributividad

Frecuentemente nos encontramos con conjuntos en los cuales hay más de una operación binaria definida. El ejemplo más sencillo son los naturales en los que sabemos sumar y sabemos multiplicar. Estas dos operaciones están relacionadas con la propiedad de que podemos sacar factor común o sea $ax + ay = a(x + y)$.

Sean \circ y \diamond dos operaciones binarias definidas en el conjunto A . Se dice que la operación \diamond es **distributiva** respecto a la operación \circ si se cumplen las dos propiedades en el recuadro a la derecha.

$$\begin{aligned} \forall a, b, c \in A \\ a \diamond (b \circ c) = (a \diamond b) \circ (a \diamond c) \\ (b \circ c) \diamond a = (b \diamond a) \circ (c \diamond a) \end{aligned}$$

 Que \diamond sea distributiva respecto a \circ no es lo mismo que \circ sea distributiva respecto a \diamond . Por ejemplo, en los naturales el producto es distributivo con respecto a la suma: $a(b + c) = (ab) + (ac)$ y sin embargo, la suma de naturales no es distributiva respecto al producto: $a + (bc) \neq (a + b)(a + c)$.

Ejercicio 9 De un ejemplo de dos operaciones binarias tales que ambas son distributivas una con respecto a la otra. [185]

El álgebra “abstracta”

Filosóficamente, el concepto de “abstracción” es la propiedad, que tiene el pensamiento humano, de que podemos fijarnos solamente en ciertas propiedades “esenciales” de un objeto o fenómeno, y olvidarnos de las restantes.

La abstracción es imprescindible para el lenguaje. El concepto “silla” nos permite reconocer una silla, independientemente si esta es de madera, de hierro, plástica, grande, cómoda, con tres, cuatro o cinco patas etc. Casi cada palabra del español (y de cualquier idioma) representa un concepto abstracto, sea este verbo, sustantivo o adjetivo.

La ciencia lleva este nivel de abstracción a un nivel aún mayor. Parte de este conocimiento científico, pasa al conocimiento público. Baste recordar conceptos como: velocidad, volumen, higiene, ADN, penicilina, electrón, metal, colesterol, triángulo, etc. Algunos de los mencionados, son muy antiguos, otros surgieron hace muy poco. Sin embargo, la mayoría de estos conocimientos queda solamente en manos de los especialistas en la materia.

Con las matemáticas pasa igual. No hace falta saber que la suma de naturales es una operación binaria commutativa para saber que $2 + 3 = 3 + 2$. Sin embargo, el concepto de “operación” y que estas operaciones pueden cumplir o no ciertas propiedades es relativamente “nuevo”.

En la primera mitad del siglo XX, progresivamente, la comunidad matemática se fué dando cuenta de las ventajas del pensamiento algebraico en el lenguaje de operaciones abstractas. Tanto fué el entusiasmo, que muchos, en un principio, le llamaron a esta forma de pensar “Algebra moderna”. Otros aún más entusiastas le llamaron “Matemática moderna”. En la actualidad este lenguaje es parte intrínseca e indivisible del pensamiento en matemáticas y cualquier calificación de “moderna” suena muy tonta.

Otros, por otro lado, prefierieron referirse a esta forma de pensar como “Algebra abstracta”. Esto, en mi opinión, aunque más moderado, tampoco tiene ningún sentido. Toda álgebra es abstracta, de hecho, todas las matemáticas son abstractas. Estoy convencido de que, el tiempo se encargará de acabar con todos estos calificativos.

1.2 Números

En esta sección repasaremos los principales tipos de números que el lector ya conoce: naturales, enteros, racionales, reales y complejos. Esto nos dará la posibilidad de introducir las definiciones más básicas del álgebra: grupos, anillos y campos.

Naturales

$$ab = \underbrace{a + \dots + a}_{b \text{ veces}} = \underbrace{b + \dots + b}_{a \text{ veces}}$$

Hay una frase famosa que dice “Dios hizo los naturales y el hombre todo lo demás”. El conjunto de los números naturales $\mathbb{N} = \{0, 1, 2, \dots\}$ es el conjunto de

los cardinales de los conjuntos finitos. En \mathbb{N} hay dos operaciones binarias bien definidas: la suma y el producto. De hecho, el producto es una operación derivada de la suma y la suma solo se puede definir en términos de conjuntos. Por ejemplo, $a + b$ es el cardinal de la unión de dos conjuntos finitos y disjuntos uno de cardinal a y otro de cardinal b .

Como la unión de conjuntos es asociativa también lo es la suma de naturales. De la definición se obtiene que el producto de naturales también es asociativo. Tanto la suma como el producto son commutativos. La suma tiene elemento neutro 0 y el producto tiene elemento neutro 1 . El producto es distributivo respecto a la suma.

Enteros

Ningún elemento de \mathbb{N} salvo el cero tiene inverso para la suma. Para lograr la existencia de inversos inventamos los números negativos $\mathbb{Z}^- = \{-1, -2, -3, \dots\}$ y en el conjunto $\mathbb{Z} = \mathbb{N} \cup \mathbb{Z}^-$ de los números enteros definimos la suma como la operación conmutativa definida por las propiedades en el recuadro a la derecha.

$$\begin{aligned}-b + a &= c \Leftrightarrow a = b + c \\ -b + (-a) &= -(a + b)\end{aligned}$$

Grupos

Nuevamente la suma de enteros es asociativa con neutro cero pero ahora, cada elemento tiene inverso. O sea, los enteros dan el primer ejemplo de grupo. A un conjunto no vacío con una operación binaria se le llama **grupo** si se

- G1) la operación es asociativa
- G2) tiene elemento neutro
- G3) todo elemento tiene inverso



cumplen los tres axiomas G1-G3. A los grupos cuya operación es conmutativa se les llama **abelianos** en honor al matemático noruego Niels Henrik Abel (1802-1829). Abel fué el que resolvió el problema algebraico más importante de su época. Demostró, que no existen fórmulas en radicales para resolver las ecuaciones polinomiales de grado 5 o mayor (a diferencia de las ecuaciones de grado ≤ 4 para las cuales si hay fórmulas generales). Al momento de encontrar esta demostración, el problema ya duraba varios siglos sin resolverse.

Abel murió a los 26 años a causa de una neumonía.

Anillos

La operación de producto de naturales se extiende fácilmente al conjunto de los enteros mediante las reglas en el recuadro. Nuevamente el producto es asociativo, conmutativo y distributivo con respecto a la suma. O sea los enteros también dan el primer ejemplo de anillo.

$$\begin{aligned}a(-b) &= -(ab) \\ (-a)b &= -(ab) \\ (-a)(-b) &= ab\end{aligned}$$

Un conjunto **A** no vacío con dos operaciones binarias $+$ y \bullet se le llama **anillo** si se cumplen los axiomas en el recuadro a la derecha. Si el anillo es tal que la operación \bullet es conmutativa entonces se dice que tenemos un **anillo conmutativo**.

- A1) $(A, +)$ es un grupo abeliano
- A2) \bullet es asociativa
- A3) \bullet es distributiva con respecto a $+$
- A4) \bullet tiene elemento neutro

En un anillo al neutro para la suma se le llama **cero** y se denota por 0. Al neutro para el producto se le llama **uno** y se denota por 1. Al inverso de un elemento con respecto a la suma de un elemento se le llama su **opuesto**. Al inverso con respecto al producto de un elemento se le llama **inverso multiplicativo** o simplemente inverso a secas.

Observemos que si a es un elemento de un anillo entonces $a \bullet 0 = a \bullet 0 + a - a =$

$a \bullet (0 + 1) - a = a - a = 0$. De la misma manera vemos que $0 \bullet a = 0$. Si $1 = 0$ entonces, $a = 1 \bullet a = 0 \bullet a = 0$ por lo que el anillo consta de un solo elemento. Para descartar esta trivialidad supondremos siempre que $1 \neq 0$. De aquí se desprende que 0 no puede tener inverso ya que $0 = a \bullet 0 = 1$ es una contradicción.

En cualquier anillo $-a$ denota al opuesto de a y a^{-1} (si existe) denota al inverso de a . Como $1 \times 1 = 1$ tenemos $1^{-1} = 1$. También $(-1)^{-1} = -1$ ya que si $a = -1$ entonces $0 = a(a + 1) = aa + a$ por lo que $aa = -a$ o sea, $(-1)(-1) = 1$. Luego, en todo anillo 1 y -1 tienen inversos. En \mathbb{Z} ningún elemento salvo 1 y -1 tiene inverso.



Normalmente en la definición de anillo no se pide el axioma **A4**. En este caso, a los anillos que tienen elemento neutro para el producto se le llaman anillos unitarios. Un ejemplo de anillo no unitario es el conjunto de todos los enteros pares. Con el objetivo de simplificar, para nosotros todos los anillos son unitarios.

Racionales

Para lograr que cada elemento diferente de cero tenga inverso inventamos las fracciones y con ellas el conjunto de números racionales \mathbb{Q} . Una fracción es un par ordenado de números enteros denotado por a/b donde $b \neq 0$. Dos fracciones son iguales cuando se cumple la igualdad en el recuadro. Los números racionales \mathbb{Q} son las fracciones con la relación de igualdad así definida. Los enteros son parte de los racionales por cuanto podemos identificar cada número entero $a \in \mathbb{Z}$ con la fracción $a/1$.

La suma y el producto de números racionales se definen por las igualdades en el recuadro. Nuevamente los racionales con la suma forman un grupo abeliano y otra vez el producto es asociativo, conmutativo, tiene elemento neutro y es distributivo con respecto a la suma. Sin embargo, ahora todo elemento diferente de cero tiene inverso multiplicativo. O sea los racionales nos dan el primer ejemplo de campo.

- C1) $(\mathbb{K}, +)$ es un grupo abeliano
- C2) $(\mathbb{K} \setminus 0, \bullet)$ es un grupo abeliano
- C3) \bullet es distributiva con respecto a $+$

$$\frac{a}{b} + \frac{c}{d} = \frac{a \bullet d + c \bullet b}{b \bullet d}$$

$$\frac{a}{b} \bullet \frac{c}{d} = \frac{a \bullet c}{b \bullet d}$$

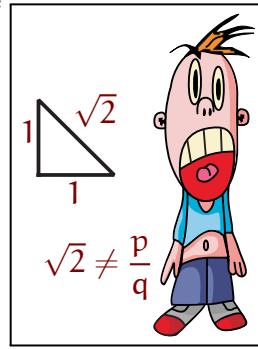
Un conjunto \mathbb{K} no vacío con dos operaciones binarias $+$ y \bullet se le llama **campo** si se cumplen los tres axiomas C1-C3. Un campo es un anillo conmutativo en la cual todo elemento diferente de cero tiene inverso multiplicativo.

Ejercicio 10 Si \circ es una operación en A entonces, en A^2 está definida la operación por coordenadas $(x, y) \circ (x', y') = (x \circ x', y \circ y')$. Pruebe que si (A, \circ) es un grupo entonces (A^2, \circ) es un grupo, si $(A, +, \bullet)$ es un anillo entonces, $(A^2, +, \bullet)$ es también un anillo.

Ejercicio 11 Sea $(\mathbb{K}, +, \bullet)$ un campo. Como \mathbb{K} es un anillo entonces, por el ejercicio anterior, $(\mathbb{K}^2, +, \bullet)$ también es un anillo. ¿Será \mathbb{K}^2 un campo? [186]

Reales

Dicen que cuando Pitágoras (Samos 569-475 A.C.) descubrió que la longitud de la hipotenusa de un triángulo rectángulo con catetos de longitud uno no es un número racional quedó horrorizado. A nosotros nos parece esto una exageración. Sin embargo, si nos ponemos en el lugar de Pitágoras comprenderemos que en aquel momento era inconcebible que existan números que no sean cociente de dos enteros. La pintura a la izquierda es un detalle del fresco de Rafael “La escuela de Atenas” en la cual supuestamente, se muestra a Pitágoras.



Sigamos a Pitágoras y probemos que efectivamente $\sqrt{2}$ no es un racional. Para esto denotemos por $\|a\|_2$ el número de veces que el natural a se divide entre 2. Tenemos $\sqrt{2} = \frac{p}{q} \Rightarrow \|2q^2\|_2 = \|p^2\|_2 \Rightarrow 1 + 2\|q\|_2 = 2\|p\|_2$ lo que es una contradicción ya que un número impar no puede ser igual a uno par.

Ejercicio 12 Sea n un natural. Pruebe que \sqrt{n} es un natural o no es racional. [186]

Ejercicio 13 Basándose en el anterior de otra prueba de que $\sqrt{2}$ no es racional. [186]

Esto motiva la construcción de los números reales \mathbb{R} . La construcción de los reales es un proceso complicado y se han descubierto muchas formas de formalizar esta construcción siendo la más popular la de las cortaduras de Dedekind. Para nuestros propósitos basta una definición menos formal y más intuitiva: un número real es simplemente un límite de racionales. Las propiedades de la suma y producto de racionales se traspasan fácilmente a los reales usando las propiedades del límite de sucesiones. De esta manera obtenemos nuestro campo principal $(\mathbb{R}, +, \cdot)$. El campo de los reales se destaca porque es **ordenado** (siempre podemos decidir si un número real es mayor, menor o igual a cero) y porque es **cerrado** (el límite de reales si existe es un real). Por otro lado, no es un factor a despreciar el hecho de que el espacio en que vivimos es (o al menos nos parece que es) \mathbb{R}^3 .

Ejercicio 14 Pruebe que la longitud de un segmento de recta es un número real. [186]

Complejos

En 1546 Gerolamo Cardano publicó su libro “Ars Magna” en el cual dió métodos (basados en parte en el trabajo de otros matemáticos) para el cálculo de las raíces de los polinomios de grado 3 y 4. Estos métodos, a veces requerían el extraer raíces

cuadradas de números negativos, incluso cuando el resultado final era un número real. Rafael Bombelli estudió este asunto en detalle y es considerado como el descubridor de los números complejos.

Para lograr que todos los polinomios tengan raíces inventamos el imaginario $i = \sqrt{-1}$ y definimos que un número complejo es algo de la forma $a + bi$ donde $a, b \in \mathbb{R}$. La suma y el producto de complejos se definen por las fórmulas en los recuadros a la derecha y abajo a la izquierda.

$$(a + bi) + (a' + b'i) = \\ (a + a') + (b + b')i$$

Las propiedades de la suma y el producto se desprenden inmediatamente de sus definiciones y es fácil comprobar que $(\mathbb{C}, +, \bullet)$ es un anillo conmutativo. Para ver que es un campo, observamos que $(a + bi)^{-1} =$

$$(a + bi)^{-1} = \frac{(a - bi)}{(a^2 + b^2)}$$
. La principal propiedad que hace que para muchas cosas el campo \mathbb{C} sea el más simple es que el (a diferencia de \mathbb{R}) es **algebraicamente cerrado**, o sea que todo polinomio de grado $n > 0$ con coeficientes en complejos tiene n raíces complejas.

1.3 Morfismos

En las matemáticas cada vez que se estudian ciertos objetos, es necesario también estudiar las funciones entre ellos, que “preservan” las propiedades de dichos objetos. En esta sección estudiaremos las funciones entre conjuntos con operaciones binarias.

Morfismos de grupos

Sean \circ y \bullet operaciones binarias definidas en los conjuntos A y B respectivamente. Una función $f : A \rightarrow B$ se le llama **morfismo** si para cualesquiera a_1 y a_2 elementos de A se cumple que $f(a_1 \circ a_2) = f(a_1) \bullet f(a_2)$.

 Todo morfismo conserva las propiedades fundamentales de las operaciones binarias. Más precisamente, si $f : (A, \circ) \rightarrow (B, \bullet)$ es un morfismo entonces,

1. \bullet es una operación binaria dentro de la imagen de f .
2. Si \circ es conmutativa entonces \bullet es conmutativa en la imagen de f .
3. Si \circ es asociativa entonces \bullet es asociativa en la imagen de f .
4. Si e es neutro de \circ entonces $f(e)$ es neutro de \bullet en la imagen de f .
5. Si a' es inverso de a en A entonces $f(a')$ es inverso de $f(a)$ en B .

Prueba. Sean b_1, b_2 y b_3 elementos cualesquiera en la imagen de f . Existen a_1, a_2 y a_3 en A tales que $f(a_i) = b_i$ para $i \in \{1, 2, 3\}$. Como f es un morfismo, obtenemos la igualdad

$$b_1 \bullet b_2 = f(a_1) \bullet f(a_2) = f(a_1 \circ a_2) \quad (*)$$

que prueba la primera afirmación. Si \circ es conmutativa entonces, usando (*) obtenemos

$$b_1 \bullet b_2 = f(a_1 \circ a_2) = f(a_2 \circ a_1) = b_2 \bullet b_1$$

por lo que \bullet es también conmutativa. Si \circ es asociativa entonces, usando (*) obtenemos

$$\begin{aligned} (b_1 \bullet b_2) \bullet b_3 &= f(a_1 \circ a_2) \bullet f(a_3) = f((a_1 \circ a_2) \circ a_3) = \\ &= f(a_1 \circ (a_2 \circ a_3)) = f(a_1) \bullet f(a_2 \circ a_3) = b_1 \bullet (b_2 \bullet b_3) \end{aligned}$$

y por lo tanto \bullet es asociativa en la imagen de f . Si e es neutro de la operación \circ entonces,

$$\begin{aligned} b_1 \bullet f(e) &= f(a_1) \bullet f(e) = f(a_1 \circ e) = f(a_1) = b_1 \\ f(e) \bullet b_1 &= f(e) \bullet f(a_1) = f(e \circ a_1) = f(a_1) = b_1 \end{aligned}$$

por lo que $f(e)$ es el neutro de \bullet en la imagen de f . Sea a' el inverso de a en A entonces,

$$\begin{aligned} f(a) \bullet f(a') &= f(a \circ a') = f(e) \\ f(a') \bullet f(a) &= f(a' \circ a) = f(e) \end{aligned}$$

de lo que concluimos que $f(a')$ es el inverso de $f(a)$. ■

Ejercicio 15

Justifique todas las igualdades utilizadas en la prueba de 1.1.

¿Y porqué siempre dentro de la imagen de f y no en todo B ? La respuesta es que lo único que sabemos de B está dado por el morfismo. Aquellos elementos de B que no tienen preimagen no los podemos enlazar con los de A y por lo tanto no podemos decir nada de ellos. De aquí en lo adelante a la imagen de cualquier función f (y en particular de un morfismo) la denotaremos por $\text{Im } f$.

1.2

Si (A, \circ) es un grupo entonces $(\text{Im } f, \bullet)$ es un grupo.

Prueba. Por 1.1.1 \bullet es una operación binaria en $\text{Im } f$. Por 1.1.3 esta operación es asociativa. Por 1.1.4 esta operación tiene elemento neutro. Por 1.1.5 cada elemento $b = f(a) \in \text{Im } f$ tiene su inverso $f(a')$ donde a' es el inverso de a en A . Esto completa la prueba de todos los axiomas de grupo. ■

Recordemos que si $f : A \rightarrow B$ es una función entonces al conjunto A se le llama **dominio** de f y al conjunto B **codominio** de f . Si el dominio y el codominio de un morfismo son grupos entonces se dice que este es un **morfismo de grupos**.

Ejercicio 16 Construya un morfismo inyectivo de $(\mathbb{R}, +)$ en (\mathbb{R}, \bullet) . ¿Cuál es la imagen de este morfismo? ¿Es esta imagen un grupo?

Morfismos de anillos

¿Y que pasa con la distributividad? ¿También se conserva? El primer problema que tenemos que resolver es que en la distributividad están involucradas dos operaciones. Sean $(A, +, \bullet)$ y $(B, +, \bullet)$ dos conjuntos cada uno con dos operaciones binarias. Ob-

servese que estas son cuatro operaciones distintas pero hemos usado estas notaciones porque el trabajar con cuatro símbolos diferentes ya es demasiada confusión.

Una función $f : A \rightarrow B$ se le llama **morfismo** si para cualesquiera a_1 y a_2 elementos de A se cumple que $f(a_1 + a_2) = f(a_1) + f(a_2)$ y $f(a_1 \bullet a_2) = f(a_1) \bullet f(a_2)$. Recalquemos que el “y” quiere decir que se tienen que cumplir las dos propiedades. O sea, si hay dos operaciones entonces, se requiere que la función sea morfismo para cada una de ellas.

4.3

- Si \bullet es distributiva con $+$ en A entonces,*
- \bullet es distributiva con $+$ en la imagen de f .*

Prueba. Sean $x, y, z \in A$ tales que $f(x) = a$, $f(y) = b$ y $f(z) = c$. Tenemos

$$\begin{aligned} a \bullet (b + c) &= f(x) \bullet (f(y) + f(z)) = f(x) \bullet f(y + z) = f(x \bullet (y + z)) = \\ &= f(x \bullet y + x \bullet z) = f(x \bullet y) + f(x \bullet z) = f(x) \bullet f(y) + f(x) \bullet f(z) = a \bullet b + a \bullet c \end{aligned}$$

y esto prueba la tesis. ■

Si el dominio y el codominio de un morfismo son anillos entonces se dice que este es un **morfismo de anillos**. Si el dominio y el codominio de un morfismo son campos entonces se dice que este es un **morfismo de campos**.

Ejercicio 17 Demuestre que si $(A, +, \bullet)$ es un anillo y $f : A \rightarrow B$ es un morfismo entonces, $(\text{Im } f, +, \bullet)$ es un anillo. Demuestre que lo mismo ocurre para los campos.

Ejercicio 18 Pruebe que si $f : A \rightarrow B$ es un morfismo de anillos y A es un campo entonces f es inyectivo. En particular todo morfismo de campos es inyectivo. [186]

Isomorfismos

A los morfismos biyectivos se les llama **isomorfismos**. Esto se aplica tanto para conjuntos con una como también con dos operaciones binarias. Así que tenemos isomorfismos de grupos, de anillos y de campos. Para cada isomorfismo f existe una función inversa f^{-1} . ¿Cuando será f^{-1} un morfismo? La respuesta es que siempre.

4.4

La inversa de un isomorfismo es un isomorfismo.

Prueba. Sea $f : (A, \circ) \rightarrow (B, \bullet)$ un isomorfismo. Sean b_1, b_2 cualesquiera elementos de B . Denotemos $a_1 = f^{-1}(b_1)$ y $a_2 = f^{-1}(b_2)$. Tenemos

$$f^{-1}(b_1 \bullet b_2) = f^{-1}(f(a_1) \bullet f(a_2)) = f^{-1}(f(a_1 \circ a_2)) = a_1 \circ a_2 = f^{-1}(b_1) \circ f^{-1}(b_2)$$

que es lo que se requería demostrar. Si el isomorfismo involucra dos operaciones binarias entonces el mismo argumento aplicado a las dos operaciones, nos da la prueba de la tesis. ■

Ahora podemos aplicar 1.1 en los dos sentidos. Si $f : (A, \circ) \rightarrow (B, \bullet)$ es un isomorfismo entonces de que \bullet es commutativa implica que \circ es commutativa y en conclusión \circ es commutativa si y solo si \bullet es commutativa. Lo mismo ocurre con la asociatividad, con la existencia de neutros e inversos y para el caso de dos operaciones con la distributividad. O sea que \circ tiene exáctamente las mismas propiedades de \bullet .

Pero no solo son operaciones parecidas sino que son en cierto sentido la misma. Para convencernos de esto supongamos que conocemos la operación \circ y conocemos el isomorfismo f pero no sabemos nada de la operación \bullet . ¿Podremos calcular $b_1 \bullet b_2$? La respuesta es sí, lo podemos calcular por la identidad $b_1 \bullet b_2 = f(f^{-1}(b_1) \circ f^{-1}(b_2))$. Recíprocamente, \circ se define de forma única por la operación \bullet y el isomorfismo f mediante la identidad $a_1 \bullet a_2 = f^{-1}(f(a_1) \circ f(a_2))$. En conclusión ambas operaciones se definen una a otra.

Para que el lector comprenda mejor eso de que \circ y \bullet son la misma operación veamos un ejemplo. Sea A el conjunto de letras $\{u, v\}$ y B el conjunto de los números $\{1, -1\}$. Definamos las operaciones \circ y \bullet mediante las tablas del recuadro a la derecha.

\circ	u	v	\bullet	1	-1
u	u	v	1	1	-1
v	v	u	-1	-1	1

El lector debe observar que la segunda tabla es la tabla usual de multiplicación de enteros. Además, para obtener la segunda tabla de la primera lo único que necesitamos es cambiar \circ por \bullet , u por 1 y v por -1 . Esto lo que quiere decir, es que la función $u \mapsto 1, v \mapsto -1$ es un isomorfismo de (A, \circ) en (B, \bullet) . El lector puede ver que ambas tablas son en esencia la misma, solamente que las notaciones para los elementos y la operación están cambiadas.

Si para dos grupos (o anillos o campos) existe un isomorfismo entre ellos entonces se dice que ellos son **isomorfos**. Etimológicamente, la palabra “isomorfo” significa que “tienen la misma forma”. En forma intuitiva, que ellos sean isomorfos quiere decir que los dos son iguales con la salvedad de que podemos cambiar las notaciones de los elementos y las operaciones.

Ciertos tipos de morfismos tienen nombres especiales. A los morfismos sobreyectivos se les llama **epimorfismos**, a los injectivos se les llama **monomorfismos**. A los morfismos de un conjunto en si mismo se les llama **endomorfismos** y a los endomorfismos biyectivos se les llama **automorfismos**.



En otras ramas de las matemáticas también se definen morfismos e isomorfismos. Sin embargo no siempre es suficiente la biyectividad para definir los isomorfismos. Por ejemplo, en topología los morfismos son las funciones continuas. Pero la inversa de una biyección continua no siempre es continua. Por esto, un isomorfismo de espacios topológicos hay que definirlo como una biyección continua cuya inversa es continua.

Composición de morfismos

Sean A , B y C tres conjuntos y $f : A \rightarrow B$, $g : B \rightarrow C$ dos funciones. A la función $g \circ f : A \rightarrow C$ definida por $(g \circ f)(a) = g(f(a))$ se le llama la **composición** de f con g . A partir de ahora el símbolo \circ solo lo utilizaremos para denotar la composición de

funciones. Observese el orden en que escribimos las funciones ya que la composición de funciones no es conmutativa.

1.5

La composición de funciones es asociativa.

Prueba. Sean $f : A \rightarrow B$, $g : B \rightarrow C$ y $h : C \rightarrow D$ tres funciones. Por definición de composición para cualquier $a \in A$ tenemos

$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a))) = (h \circ g)(f(a)) = ((h \circ g) \circ f)(a)$$

que es lo que se quería probar ■

Ahora, supongamos que en A , B y C hay definidas operaciones binarias. Entonces f , g y $g \circ f$ pueden ser morfismos o no. Sin embargo, si f y g lo son entonces $f \circ g$ también lo es.

1.6

Las composiciones de morfismos son morfismos.

Prueba. Denotemos las operaciones en A , B y C con el mismo símbolo \bullet . Como f y g son morfismos tenemos $(g \circ f)(a \bullet b) = g(f(a \bullet b)) = g(f(a) \bullet f(b)) = g(f(a)) \bullet g(f(b)) = (g \circ f)(a) \bullet (g \circ f)(b)$ que es lo que se necesitaba probar. ■

Ejercicio 19 Pruebe que el conjunto de los automorfismos de un conjunto con una o dos operaciones binarias es un grupo con respecto a la composición de funciones.

Ejercicio 20 Sean f y g dos funciones. Pruebe que si $g \circ f$ es la identidad entonces, f es inyectiva y g es sobreyectiva. [186]

1.4 Campos de restos

Hasta ahora los campos que conocemos son \mathbb{Q} , \mathbb{R} y \mathbb{C} que se supone que ya son muy conocidos por el lector. Es imprescindible, para dar una intuición saludable de lo que es un campo, introducir otros que no sean tan usuales. En esta sección presentaremos ciertos campos que tienen un número finito de elementos. Para construirlos, usaremos las propiedades de los morfismos de la sección anterior.

El anillo de los enteros módulo n

Sea n un número natural mayor que 1. Para un entero a la notación $a \bmod n$ significa el resto de la división de a entre n . O sea, el menor natural k tal que existe un entero t para los cuales $a = k + tn$. Por definición $a \bmod n \in \mathbb{Z}_n = \{0, 1, \dots, n - 1\}$ y en \mathbb{Z}_n hay naturalmente definidas dos operaciones binarias como se muestra en el recuadro.

$$a \boxplus b = (a + b) \bmod n$$

$$a \boxdot b = (ab) \bmod n$$

Ejercicio 21 Construya las tablas de sumar y multiplicar en \mathbb{Z}_2 , \mathbb{Z}_3 y \mathbb{Z}_4 .

1.7 $(\mathbb{Z}_n, \boxplus, \boxdot)$ es un anillo conmutativo.

Prueba. Denotemos $f : \mathbb{Z} \ni a \longmapsto a \bmod n \in \mathbb{Z}_n$ o sea $f(a) \stackrel{\text{def}}{=} a \bmod n$. Observemos que por definición de las operaciones \boxplus y \boxdot se tiene que $a \boxplus b = f(a+b)$ y $a \boxdot b = f(ab)$. Además, para cualesquiera $x, y \in \mathbb{Z}$ existen enteros p, q tales que $x = f(x) + qn$, $y = f(y) + pn$ y por lo tanto $f(x) = f(y)$ si y solo si $x - y$ es múltiplo de n .

Probaremos que $f : (\mathbb{Z}, +) \rightarrow (\mathbb{Z}_n, \boxplus)$ es un morfismo. Tenemos que $f(x) + f(y) = x + y - (q + p)n$ y por lo tanto $f(x) + f(y) - (x + y)$ es múltiplo de n . Luego, $f(x+y) = f(f(x) + f(y))$ o lo que es lo mismo, $f(x+y) = f(x) \boxplus f(y)$ y esto prueba que f es morfismo para la suma.

Ahora probaremos que $f : (\mathbb{Z}, \cdot) \rightarrow (\mathbb{Z}_n, \boxdot)$ es un morfismo. Tenemos que $f(x)f(y) = (x - qn)(y - pn) = xy + (nqp - yq - xp)n$ y por lo tanto $f(x)f(y) - xy$ es múltiplo de n . Luego, $f(xy) = f(f(x)f(y))$ o lo que es lo mismo, $f(xy) = f(x) \boxdot f(y)$ y esto prueba que f es morfismo para el producto.

Como f es sobreyectiva, $(\mathbb{Z}, +, \cdot)$ es anillo conmutativo y los morfismos preservan las propiedades de las operaciones, concluimos que $(\mathbb{Z}_n, \boxplus, \boxdot)$ es también un anillo conmutativo. ■



Hemos denotado la suma y el producto en \mathbb{Z}_n con los símbolos extraños \boxplus y \boxdot . El objetivo de esto fué el asegurarnos que en la demostración del resultado anterior el lector no se confundiera con la suma y el producto habitual de números enteros. De ahora en lo adelante no haremos más esto. La suma en \mathbb{Z}_n se denotará con el símbolo $+$ y el producto, con la ausencia de símbolo alguno, o a lo más, con un punto. Para poder hacer esto es necesario que el lector comprenda (muchas veces solo del contexto) en qué sentido estamos utilizando estas notaciones. Así por ejemplo, $2 + 3 = 5$ si la suma es la habitual de enteros o es la de \mathbb{Z}_{11} pero $2 + 3 = 1$ si la suma es la de \mathbb{Z}_4 .

Dominios de integridad

Después de saber que $(\mathbb{Z}_n, +, \cdot)$ es un anillo conmutativo, lo natural es preguntarnos si este es un campo. Lo único que le falta a un anillo conmutativo para ser campo, es la existencia de inversos para el producto. Veamos por ejemplo el caso de \mathbb{Z}_6 . Aquí tenemos $2 \cdot 3 = 0$. Que raro, el producto de dos números diferentes de cero es igual a cero. ¿Es posible eso en un campo? Veremos que no.

Un anillo conmutativo se le llama **dominio de integridad** si el producto de elementos distintos de cero es siempre diferente de cero. Sabemos que \mathbb{Z} , \mathbb{Q} , \mathbb{R} y \mathbb{C} son dominios de integridad. También, ya vimos que \mathbb{Z}_6 no es dominio de integridad.

8. Todo campo es un dominio de integridad.

Prueba. Supongamos $pq = 0$ y $p \neq 0$ entonces multiplicando la primera igualdad por el inverso multiplicativo de p obtenemos $0 = p^{-1}0 = p^{-1}pq = q$. Luego, $q = 0$. ■

Luego, \mathbb{Z}_6 no es un campo. Este ejemplo se generaliza fácilmente. Sea $n = pq$ una descomposición en factores no triviales (ambos diferentes a 1) de n . Sabemos que p y q están en \mathbb{Z}_n y que $pq = n = 0 \bmod n$. Luego, si n es un número compuesto (no primo) entonces, \mathbb{Z}_n no es un dominio de integridad y por lo tanto no es un campo.

El campo de los enteros módulo p

Y ¿que pasa cuando p es primo?

9. \mathbb{Z}_p es un dominio de integridad.

Prueba. Sean $x, y \in \{1, \dots, p-1\}$. Si $xy = 0$ en \mathbb{Z}_p entonces $xy = 0 \bmod p$. Luego, en \mathbb{Z} tenemos que $xy = kp$. Como p es primo entonces, p divide a x o a y pero esto no puede ser ya que ambos son menores que p . ■

Este resultado no es suficiente para probar que \mathbb{Z}_p es un campo ya que hay dominios de integridad que no son campos (por ejemplo \mathbb{Z}). Nos hace falta el siguiente resultado.

10. Todo dominio de integridad finito es un campo.

Prueba. Sea A un dominio de integridad finito. Para ver que A es un campo solo hay que demostrar que todo elemento no nulo tiene inverso. Sea a un elemento arbitrario no nulo de A . Denotemos por f_a la función $A \ni x \mapsto ax \in A$. Esta función es inyectiva ya que $(ax = ay) \Rightarrow (a(x - y) = 0) \Rightarrow (x - y = 0) \Rightarrow x = y$.

Como f_a es una función inyectiva de un conjunto finito en si mismo es también sobreyectiva. Luego tiene que existir b tal que $f_a(b) = ab = 1$. Como el producto es commutativo, esto demuestra que a tiene inverso. ■



Como \mathbb{Z}_p es finito, concluimos inmediatamente que \mathbb{Z}_p es un campo si y solo si p es un número primo. Los campos \mathbb{Z}_p son los primeros ejemplos de campos que tienen un número finito de elementos. A los campos finitos se les llama campos de Galois en honor al matemático francés Evariste Galois (1811-1832). Al resolver el problema de encontrar cuales ecuaciones polinomiales son solubles en radicales y cuales no, Galois de facto inventó la Teoría de Grupos. Galois murió a los 20 años en un duelo provocado por asuntos amorosos y/o políticos. El apellido Galois se pronuncia en español como "galuá"

Ejercicio 22 Halle los inversos de los elementos no nulos de \mathbb{Z}_5 . [186]

Ejercicio 23 Demuestre que \mathbb{Z}_{11}^2 con las operaciones $(a, b) + (x, y) = (a+x, b+y)$ y $(a, b)(x, y) = (ax + 7by, ay + xb)$ es un campo de 121 elementos. [187]

1.5 Campos primos. Característica

Sea \mathbb{K} un campo. Un **subcampo** de \mathbb{K} es sencillamente un subconjunto de \mathbb{K} que es campo para las mismas operaciones. Si \mathbb{L} es un subcampo de \mathbb{K} entonces $\forall a, b \in \mathbb{L}$ $\forall c \in \mathbb{L} \setminus \{0\}$ se tienen que cumplir las siguientes propiedades

1. $a + b \in \mathbb{L}, -a \in \mathbb{L}$, (\mathbb{L} es un subgrupo aditivo)
2. $ab \in \mathbb{L}, c^{-1} \in \mathbb{L}$, ($\mathbb{L} \setminus \{0\}$ es un subgrupo multiplicativo)

Recíprocamente, si se cumplen las propiedades 1 y 2 entonces, las operaciones de suma, opuesto, producto e inverso están correctamente definidas dentro de \mathbb{L} y el tiene que contener a 0 y a 1. Como los axiomas de campo se cumplen dentro de todo \mathbb{K} , con más razón se cumplen dentro de \mathbb{L} . Esto indica que para comprobar si \mathbb{L} es un subcampo basta comprobar las propiedades 1 y 2. El ejemplo más sencillo de esto es \mathbb{Q} , que es subcampo \mathbb{R} , que a su vez es subcampo de \mathbb{C} .



El concepto de subcampo incluye las operaciones. Si por un lado $\{0, \dots, p-1\} = \mathbb{Z}_p$ es un subconjunto de \mathbb{Q} , por el otro, \mathbb{Z}_p NO es un subcampo de \mathbb{Q} (ya que por ejemplo $2(p-1) = p-2$ en \mathbb{Z}_p lo que no es cierto en \mathbb{Q}). De la misma manera, ningún \mathbb{Z}_p es subcampo de \mathbb{Z}_q para $p \neq q$.

Campos primos

Todo campo es subcampo de si mismo. A los campos que no tienen ningún subcampo distinto de si mismo se les llama **campos primos**. Los campos primos son los más sencillos y deduciremos cuales son todos ellos.

1.1 *Todo campo \mathbb{K} contiene un único subcampo primo que está contenido en cualquier subcampo de \mathbb{K} .*

Prueba. La intersección de una colección arbitraria de subcampos de \mathbb{K} es un subcampo. Para ver esto observamos que si a y b pertenecen a todos los subcampos de la colección entonces $a+b$ también. Por lo tanto, $a+b$ está en la intersección de ellos. Lo mismo ocurre para el producto, para los neutros y los inversos. En particular, la intersección de todos los subcampos de \mathbb{K} es un subcampo que no contiene subcampos y está contenida en cualquier subcampo de \mathbb{K} . ■

1.12 *El campo \mathbb{Q} de los números racionales es primo.*

Prueba. Sea \mathbb{K} un subcampo de \mathbb{Q} . Tenemos $1 \in \mathbb{K}$ y por lo tanto todas las sumas $1 + \dots + 1$ y sus opuestos aditivos tienen que estar en \mathbb{K} . Luego, todos los enteros están en \mathbb{K} . También los inversos multiplicativos de los números enteros y sus productos tienen que estar todos en \mathbb{K} . Luego, todos los racionales están en \mathbb{K} . ■

1.13 *Los campos \mathbb{Z}_p de los restos módulo un número primo son campos primos.*

Prueba. Sea \mathbb{K} un subcampo de \mathbb{Z}_p . Tenemos $1 \in \mathbb{K}$ y por lo tanto todas las sumas $1 + \dots + 1$ están en \mathbb{K} . Como cualquier elemento de \mathbb{Z}_p es suma de unos obtenemos que $\mathbb{Z}_p \subseteq \mathbb{K}$. ■

Teorema de Clasificación de Campos Primos

1.14 *Los campos \mathbb{Z}_p y el campo \mathbb{Q} son los únicos campos primos.*

Prueba. Sea un \mathbb{K} campo primo. Para un número natural n denotemos por $\overline{n} = \underbrace{1 + \dots + 1}_{n \text{ veces}}$ donde 1 es el neutro multiplicativo de \mathbb{K} . Observese que $\overline{0} = 0$. Obviamente, \overline{n} es un elemento del campo. Denotemos por $P = \{\overline{n} \in \mathbb{K} \mid n \in \mathbb{N}\}$. Hay dos posibilidades excluyentes

1. La aplicación $n \mapsto \overline{n}$ es una biyección de \mathbb{N} en P .
2. Existen dos naturales distintos n y m tales que $\overline{n} = \overline{m}$.

En el primer caso \mathbb{K} contiene a los naturales. Como \mathbb{K} es un campo también tiene que contener a los opuestos de los naturales, o sea a los enteros. Por la misma razón, \mathbb{K} tiene que contener a los inversos de los enteros con lo que se prueba que los racionales son un subcampo de \mathbb{K} . Como \mathbb{K} es primo obtenemos que $\mathbb{K} = \mathbb{Q}$.

En el segundo caso, sea p el natural más pequeño para el cual existe $n < p$ tal que $\overline{n} = \overline{p}$. Tenemos $\overline{n} = \overline{p} \Rightarrow \overline{p - n} = \overline{p - n} = 0$. Si $n > 0$ entonces, $p - n < p$ y además $\overline{p - n} = \overline{0}$ lo que contradice la minimalidad de p . Luego, $n = 0$ y por lo tanto $\overline{p} = 0$.

Sea ahora $x > p$. Como sabemos dividir los enteros con resto entonces, existen naturales a, k tales que $x = a + kp$ y $a < p$. De aquí

$$\overline{x} = \overline{a + kp} = \overline{a} + \overline{kp} = \overline{a} + \underbrace{\overline{1 + \dots + 1}}_{p \text{ veces}} + \dots + \underbrace{\overline{1 + \dots + 1}}_{p \text{ veces}} = \overline{a} + \underbrace{\overline{0 + \dots + 0}}_{k \text{ veces}} = \overline{a}$$

lo que muestra que P es el anillo \mathbb{Z}_p de los restos módulo p . Si p no es primo entonces, en $\mathbb{Z}_p \subseteq \mathbb{K}$ hay dos elementos a, b no cero tales que $ab = 0$. Como en un campo esto no es posible entonces, deducimos que p es primo. Luego, \mathbb{Z}_p es un subcampo de \mathbb{K} . Como \mathbb{K} es primo obtenemos que $\mathbb{K} = \mathbb{Z}_p$. ■

Característica

Por el teorema anterior cualquier campo o contiene a \mathbb{Q} o contiene a \mathbb{Z}_p . Se dice que un campo es de **característica 0** si este contiene a \mathbb{Q} . Se dice que un campo es de **característica p** si este contiene a \mathbb{Z}_p . La característica de un campo es un número primo o es cero. La propiedad fundamental de la característica de un campo es la siguiente:

1.15

Si \mathbb{K} es un campo de característica t entonces, $ta = 0$ para cualquier $a \in \mathbb{K}$.

Prueba. Si t es un número primo entonces, el campo contiene a \mathbb{Z}_t y

$$tx \stackrel{\text{def}}{=} \underbrace{x + \dots + x}_{t \text{ veces}} = 1x + \dots + 1x = (t1)x$$

Como $1 \in \mathbb{Z}_t$ entonces también $t1 \in \mathbb{Z}_t$. En \mathbb{Z}_t se tiene que $t1 = 0$. Si $t = 0$ la afirmación es trivial. ■

Ejercicio 24 ¿Contradice o no 1.15 que todo campo es un dominio de integridad? [187]

Ejercicio 25 Pruebe el recíproco de 1.15: Si t es el menor natural tal que para todo $a \in \mathbb{K}$ se tiene que $ta = 0$ entonces la característica de \mathbb{K} es igual a t .

Ejercicio 26 ¿Es cierto o no que en todo campo $(a = -a) \Rightarrow (a = 0)$? [187]

1.6 Aritmética de campos

Los campos se comportan en la mayoría de las cosas importantes como los números reales. Por más que tratemos construir un campo raro pero muy raro (lo que es posible) no lograremos que se dejen de cumplir todas las propiedades de la aritmética las cuales nos son familiares desde temprana edad. Pasemos a describir en toda su generalidad algunas consecuencias simples y otras un poco más complicadas de los axiomas de campo lo que nos convencerá de lo afirmado.

Múltiplos y exponentes enteros

En todo campo para cualquier número entero n y cualquier elemento del campo a se usan las siguientes notaciones

$$na = \begin{cases} \overbrace{a + a + \dots + a}^{n \text{ veces}} & \text{si } n > 0 \\ 0 & \text{si } n = 0 \\ (-n)(-a) & \text{si } n < 0 \end{cases} \quad a^n = \begin{cases} \overbrace{aa\dots a}^{n \text{ veces}} & \text{si } n > 0 \\ 1 & \text{si } n = 0 \\ \frac{1}{a^{-n}} & \text{si } n < 0 \end{cases}$$

y se cumplen las propiedades usuales: $(n + m) a = na + ma$ y $a^{n+m} = a^n a^m$.

Asociatividad general

$$\sum_{i=1}^n a_i$$

Recordemos nuevamente el uso del símbolo de sumatoria Σ . Si A es un conjunto finito de elementos del campo entonces podemos escribir la expresión en el recuadro a la izquierda para expresar que queremos sumar todos los elementos del conjunto $A = \{a_1, \dots, a_n\}$.

La asociatividad de la operación de suma nos dice que esta expresión tiene sentido único ya que no es necesario explicitar las sumas hay que realizar primero y cuales después.

En realidad incluso esta notación es redundante, más consisa es esta otra notación en el recuadro a la derecha que podemos usar gracias a la commutatividad de la suma. O sea no importa si en la suma a_1 está delante de a_2 o al revéz. Solo es necesario especificar cual es el conjunto A de elementos que se están sumando.

$$\sum_{a \in A} a$$

$$\prod_{i=1}^n a_i = \prod_{a \in A} a$$

Como tenemos que el producto de elementos de un campo es también asociativo y commutativo podemos usar las expresiones equivalentes de la izquierda para denotar el producto de todos los elementos del conjunto $A = \{a_1, \dots, a_n\}$.

Distributividad general

Más difícil es dar una forma general de la ley distributiva. Usando las leyes de los campos obtenemos $(a + b)(c + d) = a(c + d) + b(c + d) = ac + ad + bc + bd$

$$\left(\sum_{a \in A} a \right) \left(\sum_{b \in B} b \right) = \sum_{a \in A} \sum_{b \in B} ab$$

y el lector podrá convencerse fácilmente, haciendo el cálculo para conjuntos pequeños A y B , que en general se cumple la fórmula de la izquierda.

Más general, para muchos factores tenemos

$$\left(\sum_{a \in A} a \right) \left(\sum_{b \in B} b \right) \cdots \left(\sum_{c \in C} c \right) = \sum_{a \in A} \sum_{b \in B} \cdots \sum_{c \in C} ab \cdots c$$

A esta igualdad la llamaremos **forma general de la ley distributiva** y tendremos muchas ocasiones en que la usaremos.

Fórmula multinomial

Aplicando la forma general de la ley distributiva al caso en que todos los conjuntos sean iguales obtenemos la siguiente fórmula:

$$\left(\sum_{a \in A} a \right)^n = \sum_{a_1 \in A} \cdots \sum_{a_n \in A} a_1 \cdots a_n \quad (*)$$

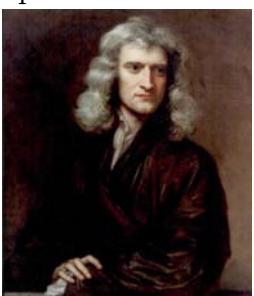
Esta fórmula aunque relativamente sencilla tiene un gran defecto. Por ejemplo, el producto **aabaccccba** que pudiera aparecer como sumando a la derecha de la igualdad tiene (gracias a la asociatividad y commutatividad del producto) una manera mucho más sencilla de expresarse como $a^4b^3c^3$. Para arreglar este defecto, démosle nombres a los elementos de **A** y pongamos $A = \{x_1, \dots, x_m\}$. Ahora, si n_1, \dots, n_m son naturales entonces, un monomio $x_1^{n_1} \cdots x_m^{n_m}$ aparece como sumando a la derecha en la fórmula (*) si y solo si $\sum n_i = n$ (ya que los sumandos en (*) son productos de n elementos de **A**). Supongamos que $\sum n_i = n$. Si todos los n_i son uno o cero entonces en (*) hay $n!$ sumandos iguales al monomio $x_1^{n_1} \cdots x_m^{n_m}$ (ya que podemos ordenarlos de ese número de maneras). Si digamos n_7 es mayor que 1 entonces tenemos que dividir por $n_7!$ ya que al permutar $x_7 \dots x_7$ no obtenemos nuevos sumandos en (*). Lo mismo sucede con los otros n_i . Como por definición $0! = 1! = 1$, finalmente obtenemos la siguiente expresión conocida como **fórmula multinomial**.

$$\left(\sum_{i=1}^m x_i \right)^n = \sum \frac{n!}{n_1! \cdots n_m!} x_1^{n_1} \cdots x_m^{n_m}$$

donde la suma a la derecha de la igualdad recorre todas las soluciones en números naturales de la ecuación $\sum n_i = n$. En el caso particular $m = 2$, haciendo el cambio de variable $n_1 = k$ obtenemos

$$(x + y)^n = \sum_{n_1+n_2=n} \frac{n!}{n_1! n_2!} x^{n_1} y^{n_2} = \sum_{k=0}^n \frac{n!}{k! (n-k)!} x^k y^{n-k}$$

que es la famosa fórmula del **binomio de Newton**.



Si bien las fórmulas que hemos demostrado parecen ser complicadas los argumentos que llevan a ellas son muy sencillos. Es importante que el estudiante se familiarize bien con estos argumentos ya que las formas multilineales y en particular los determinantes son muy parecidos a la parte derecha de la igualdad (*).

Sir Isaac Newton (Inglaterra 1643-1727) es probablemente el científico más renombrado de todos los tiempos. Fundador de la mecánica, la óptica y el cálculo diferencial. Sus tres leyes de la mecánica fundaron la base de la ingeniería que llevó a la revolución industrial.

Ejercicio 27 Sea \mathbb{K} un campo de característica $p > 0$. Demuestre que la función $\mathbb{K} \ni x \mapsto x^p \in \mathbb{K}$ es un morfismo de campos. Demuestre que si \mathbb{K} es un campo finito entonces esta función es un automorfismo de \mathbb{K} . A este automorfismo se le llama **automorfismo de Frobenius**. [187]

Ahora deduciremos otra consecuencia de la forma general de la ley distributiva que usaremos mucho más adelante. Supongamos que \mathbf{N} es un conjunto finito de índices y que para cada pareja de índices (i, j) tenemos un elemento del campo \mathbb{K} que denotaremos por α_{ij} . Nuestro objetivo es usar la ley distributiva para expresar el elemento del campo del recuadro a la derecha como una suma de productos.

$$\prod_{i \in \mathbf{N}} \sum_{j \in \mathbf{N}} \alpha_{ij}$$

Para esto lo más cómodo es pensar que el conjunto \mathbf{N} es $\{1, \dots, n\}$ y expresar la forma general de la ley distributiva en nuestro caso de la siguiente manera

$$\left(\sum_{j_1 \in \mathbf{N}} \alpha_{1j_1} \right) \cdots \left(\sum_{j_n \in \mathbf{N}} \alpha_{nj_n} \right) = \sum_{j_1 \in \mathbf{N}} \cdots \sum_{j_n \in \mathbf{N}} \alpha_{1j_1} \cdots \alpha_{nj_n} = \sum \prod_{i \in \mathbf{N}} \alpha_{ij_i}$$

donde la suma más a la derecha en esta igualdad recorre todos los elementos (j_1, \dots, j_n) del producto cartesiano $\mathbf{N} \times \cdots \times \mathbf{N}$ de n copias de \mathbf{N} o sea, \mathbf{N}^n . Otra manera de pensar a (j_1, \dots, j_n) es que tenemos una función $f : \mathbf{N} = \{1, \dots, n\} \ni i \mapsto j_i = f(i) \in \mathbf{N}$ y en nuestra suma tenemos que recorrer todas estas posibles funciones o sea, el conjunto $\mathbf{N}^{\mathbf{N}}$ de todas las funciones de \mathbf{N} en \mathbf{N} . Luego, en estas notaciones finalmente obtenemos

$$\prod_{i \in \mathbf{N}} \sum_{j \in \mathbf{N}} \alpha_{ij} = \sum_{f \in \mathbf{N}^{\mathbf{N}}} \prod_{i \in \mathbf{N}} \alpha_{if(i)}$$

que es una fórmula que ya no depende de cual es el conjunto \mathbf{N} .

Debemos recalcar una vez más que todo lo dicho en esta sección es válido para cualquier campo, sea este \mathbb{R} , \mathbb{C} , \mathbb{Q} , \mathbb{Z}_p o cualquier otro campo que aún no conocemos. Esta es la ventaja intrínseca de la abstracción. No tenemos que demostrar el teorema de Pitágoras para triángulos de acero, madera, etc. Estas propiedades no tienen nada que ver con que el cuadrado de la hipotenusa sea igual a la suma de los cuadrados de los catetos. De la misma manera el binomio de Newton no tiene nada que ver con que si los números son reales o complejos u otros. Solo basta que nuestros números formen un campo.



Un lector atento, podría observar que en las pruebas de todas las fórmulas en ningún momento usamos la existencia de inversos en el campo. Luego, estas son válidas en cualquier anillo conmutativo.



A diferencia de las matemáticas elementales en matemáticas superiores, por aritmética se entiende el estudio de las propiedades de divisibilidad en anillos. En este sentido, la aritmética de campos es trivial ya que todo elemento se divide entre cualquier otro no nulo.

1.7 Anillos con división



Si en los axiomas de campo eliminamos la condición de que el producto es commutativo obtenemos el concepto de **anillo con división** (también se le llama **cuerpo**). O

- AD1) $(D, +)$ es un grupo abeliano
- AD2) $(D \setminus \{0\}, \bullet)$ es un grupo
- AD3) \bullet es distributivo con respecto a $+$

sea, un anillo con división es un conjunto D con una suma y un producto tal que se cumplen los axiomas del recuadro. En particular, todo campo es anillo con división.

Ejercicio 28 Pruebe que si $(D, +, \bullet)$ es tal que $(D, +)$ es un grupo, $(D \setminus \{0\}, \bullet)$ es un grupo y el producto es distributivo respecto a la suma entonces, la suma es commutativa. En otras palabras en los axiomas de anillo con división la commutatividad de la suma es consecuencia del resto de los axiomas. [188]

Quaterniones

No es muy fácil construir anillos con división que no sean campos. Para esto, supongamos que en lugar de un solo número imaginario i tenemos tres diferentes imaginarios i, j y k que cumplen que $i^2 = j^2 = k^2 = -1$. Un **quaternion** es un número de la forma $a + bi + cj + dk$ donde a, b, c, d son números reales. El lector debe observar la analogía con los números complejos. Podemos sumar quaterniones por coordenadas

$$\begin{aligned} & (a + bi + cj + dk) + (a' + b'i + c'j + d'k) \\ &= ((a + a') + (b + b')i + (c + c')j + (d + d')k) \end{aligned}$$

y es fácil comprobar que el conjunto de todos los quaterniones \mathbb{H} es un grupo abeliano respecto a la suma.

Para poder multiplicar quaterniones postulamos que si a es un real y x es un imaginario entonces $ax = xa$. También postulamos que si x y y son dos imaginarios distintos entonces $xy = -yx$. Esto nos dice que nuestra multiplicación de quaterniones no es commutativa.

Ahora, si $a + bi + cj + dk$ y $a' + b'i + c'j + d'k$ son dos quaterniones arbitrarios entonces los multiplicamos como si fueran polinomios en las variables no commutativas i, j, k y usando los postulados obtenemos que su producto es igual a

$$\begin{aligned} & aa' - bb' - cc' - dd' + \\ & + (ab' + ba')i + (ac' + ca')j + (da' + ad')k + \\ & + (bc' - cb')ij + (cd' - dc')jk + (db' - bd')ki. \end{aligned}$$

Para poder concluir la definición de la multiplicación de quaterniones postulamos que $ij = k$, $jk = i$ y $ki = j$. Así definitivamente, obtenemos que el producto de nuestros quaterniones es $(a'' + b''i + c''j + d''k)$ donde los coeficientes a'', b'', c'' y d'' son los definidos en el recuadro a la derecha. O sea, el producto de quaterniones es un quaternion. No es difícil (pero si laborioso) probar directamente que este producto es asociativo tiene elemento neutro $1 = 1 + 0i + 0j + 0k$ y que es distributivo respecto a la suma.

$$\begin{aligned} a'' &= aa' - bb' - cc' - dd' \\ b'' &= ab' + ba' + cd' - dc' \\ c'' &= ac' + ca' + db' - bd' \\ d'' &= da' + ad' + bc' - cb' \end{aligned}$$

Para comprobar la existencia de inversos multiplicativos definimos para un quaternion no nulo $x = a + bi + cj + dk$ su **quaternion conjugado** $\bar{x} = a - bi - cj - dk$. De la definición de producto de quaterniones tenemos que $x\bar{x} = \bar{x}x = a^2 + b^2 + c^2 + d^2$ es un número real que tiene inverso multiplicativo. De esto se deduce que $\bar{x}(x\bar{x})^{-1}$ es el inverso multiplicativo de x . En resumen, el conjunto de los quaterniones \mathbb{H} son un anillo con división pero no son un campo.



Los quaterniones fueron descubiertos en 1843 por el físico, matemático y astrónomo irlandés Sir William Rowan Hamilton (1805 – 1865). De aquí la notación \mathbb{H} para denotar el anillo de quaterniones. Los quaterniones jugaron un papel fundamental en la matemática y la física del siglo XIX. Por ejemplo, James Clerk Maxwell usó los quaterniones para desarrollar sus ecuaciones del electromagnetismo. En la actualidad son importantes por ejemplo, para las aplicaciones en las cuales se requiere describir en forma eficiente rotaciones espaciales (robótica, control de naves espaciales, gráficos por computadoras, etc.).

Ejercicio 29 Muestre que la tabla de multiplicar de los imaginarios i, j, k es consecuencia de las igualdades de Hamilton $i^2 = j^2 = k^2 = ijk = -1$. [188]

Ejercicio 30 Muestre que los quaterniones que son raíces cuadradas de -1 forman naturalmente una esfera en \mathbb{R}^3 . Más precisamente $(a + bi + cj + dk)^2 = -1$ si y solo si se cumple que $b^2 + c^2 + d^2 = 1$. [188]

Ejercicio 31 Constrained 5.5 con el hecho de que hay un infinito número de quaterniones que son raíces de -1 (ejercicio 30). ¿Qué falla en la prueba de 5.5 para el caso de los quaterniones? [188]

Caso finito

Los quaterniones son un anillo con división infinito y es natural preguntarse si se pueden construir anillos con división finitos que no sean campos. El siguiente teorema responde esta pregunta en forma negativa.

Teorema de Wedderburn

1.16

Todo anillo con división finito es un campo.

La prueba de este teorema involucra un análisis del grupo multiplicativo del anillo y usa técnicas de teoría de grupos más allá de los objetivos de este libro.



Capítulo segundo

Espacios Vectoriales

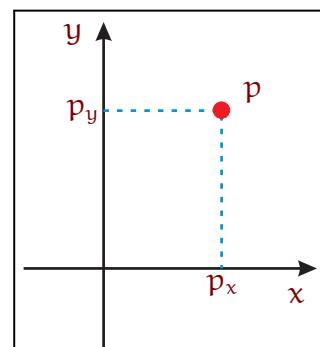
Este es el objeto central del álgebra lineal. Motivaremos la introducción de este concepto en el ejemplo geométrico del plano cartesiano. Daremos las definición de espacio vectorial complementandola con los ejemplos más fundamentales. Profundizaremos en la estructura de estos estudiando sus subespacios, sus bases, su dimensión etc. lo que nos llevará a entender todos los espacios vectoriales (al menos los de dimensión finita). Finalizaremos este capítulo con el estudio de las operaciones entre subespacios y subespacios afines lo que nos llevará a entender los espacios cocientes.

2.1 El plano cartesiano



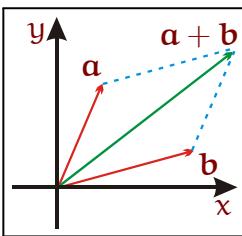
Hubo algún tiempo, en que el álgebra y la geometría eran dos cosas totalmente aparte. Los algebristas trabajaban con números, polinomios, raíces, fórmulas, etc. y los geómetras con puntos, líneas, polígonos, etc. René Descartes (Francia 1596-1650) fué el que tuvo la brillante idea de introducir los ejes de coordenadas. Tomamos dos rectas perpendiculares, a la horizontal se le llama “eje de las equis” y a la vertical se le llama “eje de las yes”.

Para cada punto del plano p trazamos la perpendicular al eje x y al eje y y de esta manera obtenemos los puntos p_x en el eje x y p_y en el eje y . Por cuento, el eje de las x lo podemos identificar (escogiendo una unidad de medida) con \mathbb{R} donde el cero es el origen de coordenadas (la intersección de los dos ejes) por tanto, p_x es simplemente un número real. De la misma manera p_y es otro número real. Así, a cada punto del plano p se le hace corresponder biunívocamente una pareja de números reales (p_x, p_y) . Además, es conveniente representar cada punto del plano como el segmento dirigido desde el origen de coordenadas hasta el punto o sea como **vectores**. A los elementos de \mathbb{R} (para diferenciarlos de los vectores) los llamaremos **escalares**.



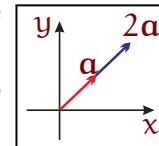


Denotaremos los vectores por letras latinas en negritas. A diferencia de estos, los escalares se denotarán por letras latinas y griegas normales.



Si $\mathbf{a} = (a_x, a_y)$ y $\mathbf{b} = (b_x, b_y)$ son dos vectores la suma de los mismos se define como $\mathbf{a} + \mathbf{b} = (a_x + b_x, a_y + b_y)$. La suma, geométricamente no es nada más que la diagonal del paralelogramo generado por \mathbf{a} y \mathbf{b} . Del hecho que $(\mathbb{R}, +)$ es un grupo abeliano se desprende fácilmente que nuestro plano cartesiano \mathbb{R}^2 es también un grupo con respecto a la suma de vectores.

Si $\mathbf{a} = (a_x, a_y)$ es un vector y α un escalar el producto $\alpha\mathbf{a}$ se define como $(\alpha a_x, \alpha a_y)$. Geométricamente este producto es aumentar (o reducir) en un factor α el vector \mathbf{a} . Si el factor es negativo entonces el resultado apunta en la dirección opuesta. Si el factor es cero entonces el vector se degenera al origen de coordenadas.



$$\begin{aligned}\alpha(\mathbf{a} + \mathbf{b}) &= \alpha\mathbf{a} + \alpha\mathbf{b} \\ (\alpha + \beta)\mathbf{a} &= \alpha\mathbf{a} + \beta\mathbf{a}\end{aligned}$$

El producto por escalares es distributivo con respecto a la suma de vectores y también con respecto a la suma de escalares o sea se cumplen las igualdades en el recuadro. Estas dos leyes distributivas (¡son diferentes!) se cumplen porque son ciertas en cada coordenada. Además, obviamente tenemos que $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$. Todo esto nos dice que el plano cartesiano es nuestro primer ejemplo de espacio vectorial sobre los reales.

2.2 Definición y ejemplos



La primera definición de espacio vectorial la dio Giuseppe Peano (Italia, 1858 -1932), en su libro “Cálculo Geométrico” publicado en 1888. Peano es más conocido por su axiomática de los números naturales, o por la “Curva de Peano” que es una inmersión continua sobreyectiva del intervalo en el cuadrado.

Sea \mathbb{K} un campo cuyos elementos los llamaremos **escalares** y $(\mathfrak{E}, +)$ un grupo abeliano cuyos elementos los llamaremos **vectores**. Diremos que es un **espacio vectorial sobre \mathbb{K}** si está definida una operación de producto de escalares por vectores $\mathbb{K} \times \mathfrak{E} \rightarrow \mathfrak{E}$ que cumple las propiedades **E1-E3**. A los axiomas **E1** y **E2** se les llama **distributividad** y **asociatividad** respectivamente del producto por escalares.

- E1**) $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$
 $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$
- E2**) $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$
- E3**) $1\mathbf{a} = \mathbf{a}$

Como $(\mathfrak{E}, +)$ es un grupo abeliano entonces, tiene que tener un vector neutro que denotaremos por $\mathbf{0}$. El opuesto del vector \mathbf{a} se denota por $-\mathbf{a}$. Por otro lado el campo de escalares tiene un neutro para la suma: el 0 , un neutro para el producto: el 1 , opuestos para la suma $-\alpha$ e inversos multiplicativos α^{-1} . Las relaciones que cumplen estos con respecto al producto por escalares nos las da el siguiente resultado básico.

2.1

En todo espacio vectorial para cualquier vector \mathbf{a} y cualquier escalar α se cumple que $0\mathbf{a} = \mathbf{0}$, $(-1)\mathbf{a} = -\mathbf{a}$ y $\alpha\mathbf{0} = \mathbf{0}$.

Prueba. La demostración se realiza mediante las siguientes tres cadenas de igualdades

$$0\mathbf{a} \stackrel{\text{E3}}{=} 0\mathbf{a} + 1\mathbf{a} - \mathbf{a} \stackrel{\text{E1}}{=} (0+1)\mathbf{a} - \mathbf{a} = \mathbf{a} - \mathbf{a} = \mathbf{0}$$

$$(-1)\mathbf{a} \stackrel{\text{E3}}{=} (-1)\mathbf{a} + 1\mathbf{a} - \mathbf{a} \stackrel{\text{E1}}{=} (-1+1)\mathbf{a} - \mathbf{a} = 0\mathbf{a} - \mathbf{a} = -\mathbf{a}$$

$$\alpha\mathbf{0} \stackrel{\text{E3}}{=} \alpha\mathbf{0} + 1 \cdot \mathbf{0} \stackrel{\text{E1}}{=} \alpha(0 + \alpha^{-1}\mathbf{0}) = \alpha(\alpha^{-1}\mathbf{0}) \stackrel{\text{E2}}{=} 1 \cdot \mathbf{0} \stackrel{\text{E3}}{=} \mathbf{0}$$

donde signos “=” están marcados con los axiomas por los que son válidos. ■

Ejercicio 32 Demuestre geométricamente que la diagonal del paralelogramo generado por \mathbf{a} y \mathbf{b} tiene coordenadas $(\mathbf{a}_x + \mathbf{b}_x, \mathbf{a}_y + \mathbf{b}_y)$. [188]

Ejercicio 33 ¿Cuál es la interpretación geométrica de la resta de vectores?

Ejercicio 34 ¿Cuantas diferentes operaciones hay en $\alpha(\beta\mathbf{a})$ y $(\alpha\beta)\mathbf{a}$? [188]

Ejercicio 35 Demuestre que $\alpha\mathbf{a} = \mathbf{0} \Rightarrow \alpha = 0$ o $\mathbf{a} = \mathbf{0}$. [188]

Ejercicio 36 ¿Cuál es el mínimo número de elementos que puede tener un espacio vectorial? [188]

Veamos ahora algunos ejemplos de espacios vectoriales. Es muy importante que el lector no se pierda en la cantidad de “ejemplos” que sigue. La mayoría de ellos son fundamentales para entender este libro. En particular, introduciremos notaciones básicas que serán usadas constantemente.

El espacio de n -adas \mathbb{K}^n

Consideremos el producto cartesiano de n copias del campo \mathbb{K} . Este producto se denota por \mathbb{K}^n y está formado por todas las **n -adas** $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$. Al escalar \mathbf{a}_i se le llama la i -ésima **coordenada** de la n -ada $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$. Luego, una n -ada tiene n coordenadas.

$$\begin{array}{r} (\mathbf{a}_1, \dots, \mathbf{a}_n) \\ + \quad (\mathbf{b}_1, \dots, \mathbf{b}_n) \\ \hline (\mathbf{a}_1 + \mathbf{b}_1, \dots, \mathbf{a}_n + \mathbf{b}_n) \end{array}$$

Los vectores serán las n -adas, los escalares serán los elementos de \mathbb{K} . En \mathbb{K}^n se introduce fácilmente la suma por coordenadas como se muestra en el recuadro a la izquierda. Del hecho de que las propiedades necesarias se cumplen en cada coordenada se desprende que $(\mathbb{K}^n, +)$ es un grupo abeliano.

La multiplicación por escalares también se introduce por coordenadas. El axioma E1 se reduce en cada coordenada a la distributividad del producto respecto a la suma en el campo \mathbb{K} . El axioma E2 a la asociatividad del producto en \mathbb{K} . Finalmente, el axioma E3 se reduce en cada coordenada a que 1 es el neutro para el producto en el campo \mathbb{K} . De esta manera obtenemos que \mathbb{K}^n es un espacio vectorial sobre \mathbb{K} .

$$\begin{array}{r} (\mathbf{a}_1, \dots, \mathbf{a}_n) \\ \times \quad \alpha \\ \hline (\alpha\mathbf{a}_1, \alpha\mathbf{a}_2, \dots, \alpha\mathbf{a}_n) \end{array}$$

El espacio de polinomios $\mathbb{K}[x]$

Podemos sumar polinomios, esta suma se hace coeficiente por coeficiente. También sabemos multiplicar un elemento de \mathbb{K} por un polinomio. Es muy conocido que todos los axiomas de espacio vectorial se cumplen. En cierto sentido este espacio vectorial es parecido al anterior. Podemos asociar a cada polinomio $\sum_{i=0}^n a_i x^i$ la $(n+1)$ -ada (a_0, a_1, \dots, a_n) y la multiplicación por escalar es la misma que en \mathbb{K}^{n+1} . Para la suma podemos tener un problema si son de grados diferentes pero esto se resuelve agregando suficientes ceros o sea si $\sum_{i=0}^m b_i x^i$ es otro polinomio con $n > m$ entonces podemos asociarle la n -ada $(b_0, b_1, \dots, b_m, 0, \dots, 0)$ con suficientes ceros para completar las n coordenadas y entonces la suma es la misma que en \mathbb{K}^{n+1} . Aunque sepamos multiplicar polinomios, esto no tiene ningún papel en el concepto de espacio vectorial.

$$\mathbb{K}[x] = \left\{ \sum_{i=0}^n a_i x^i \mid a_i \in \mathbb{K}, n \in \mathbb{N} \right\}$$

El espacio de sucesiones $\mathbb{K}^\mathbb{N}$

$$\begin{array}{r} (a_1, a_2, \dots, a_n, \dots) \\ + (b_1, b_2, \dots, b_n, \dots) \\ \hline (a_1 + b_1, \dots, a_n + b_n, \dots) \end{array}$$

Dos sucesiones se suman por coordenadas como se muestra en el recuadro. La multiplicación un escalar es también por coordenadas. Los axiomas de espacio vectorial se comprueban fácilmente. El lector debe observar que los elementos de la sucesión pueden estar en un campo arbitrario y no necesariamente en \mathbb{R} como estamos acostumbrados. La noción de convergencia de sucesiones no tiene nada que ver con el concepto de espacio vectorial.

El espacio de series $\mathbb{K}[[x]]$

Las series se suman coeficiente por coeficiente. Para multiplicar por un escalar se multiplican todos los coeficientes por el escalar. Los axiomas de espacio vectorial se cumplen porque se cumplen para cada coeficiente. De hecho, este ejemplo es el mismo que el del espacio de sucesiones. Cada serie $\sum_{i=0}^{\infty} a_i x^i$ se determina únicamente por la sucesión $(a_0, a_1, \dots, a_n, \dots)$ de sus coeficientes y no hay diferencia entre sumar series y sumar las correspondientes sucesiones. Lo mismo pasa con la multiplicación por un escalar. Al igual que con los polinomios, el concepto de producto de series no juega ningún papel en el hecho de que $\mathbb{K}[[x]]$ sea un espacio vectorial.

$$\mathbb{K}[[x]] = \left\{ \sum_{i=0}^{\infty} a_i x^i \mid a_i \in \mathbb{K} \right\}$$

El espacio de funciones \mathbb{K}^N

Hay una biyección natural entre las n -adas $(a_1, a_2, \dots, a_n) \in \mathbb{K}^n$ y las funciones $\{1, \dots, n\} \ni i \mapsto a_i \in \mathbb{K}$. Igualmente las sucesiones $(a_0, a_1, \dots, a_n, \dots)$ se corresponden biunívocamente con las funciones $\mathbb{N} \ni i \mapsto a_i \in \mathbb{K}$. Ahora, generalicemos estos ejemplos. Si N es un conjunto arbitrario (no necesitamos de operación alguna en N) entonces el conjunto de *todas las funciones* de N en \mathbb{K} se denota por \mathbb{K}^N . Dadas dos funciones

$f, g \in \mathbb{K}^N$ la suma de ellas se define como es habitual por $(f + g)(i) = f(i) + g(i)$ para cualquier $i \in N$. El producto por un escalar se define por $(\lambda f)(i) = \lambda f(i)$. Los axiomas de espacio vectorial se comprueban fácilmente. Por ejemplo, la suma de funciones es conmutativa porque para cada i en N se cumple que $f(i) + g(i) = g(i) + f(i)$ gracias a la comutatividad de la suma en el campo. Como hemos visto, el espacio \mathbb{K}^n y el espacio de sucesiones son un caso particular de este ejemplo cuando N es $\{1, \dots, n\}$ y \mathbb{N} respectivamente. Además, ya observamos que $\mathbb{K}[[x]] = \mathbb{K}^N$.

El espacio de N -adas \mathbb{K}^N



Muy frecuentemente una función en \mathbb{K}^N se denotará por α_N . Más precisamente, α_N es la función que a cada $i \in N$ le hace corresponder el escalar α_i . Por ejemplo, si $N = \{1, \dots, n\}$, entonces $\alpha_N = (\alpha_1, \dots, \alpha_n)$.

La bondad de esta notación es que podemos pensar los elementos de \mathbb{K}^N (funciones) como si fueran n -adas. Para poder seguir pensando en α_N como si fuera en una n -ada necesitamos las palabras adecuadas. A las funciones α_N de \mathbb{K}^N las llamaremos **N -adas**. Al conjunto N lo llamaremos **conjunto de índices** de α_N y a los elementos de N los llamaremos **índices**. Si i es un índice, entonces diremos que α_i es la **i -ésima coordenada** de α_N . En estas notaciones la suma de dos N -adas es por coordenadas. O sea, la i -ésima coordenada de $\alpha_N + \beta_N$ es $\alpha_i + \beta_i$. El producto por escalares también es por coordenadas. O sea, la i -ésima coordenada de $\lambda \alpha_N$ es $\lambda \alpha_i$.

Si el conjunto N es finito y tiene n elementos entonces, la diferencia fundamental entre una N -ada y una n -ada es que las coordenadas de la N -ada no necesariamente están ordenadas. Por ejemplo, si el conjunto N es un conjunto de tres vectores entonces, estos no tienen un orden natural. Para poder identificar una N -ada de estos con una 3 -ada, necesitamos definir artificialmente cual es el primer vector cual es el segundo y cual es el tercero.

El espacio de N -adas finitas $\mathbb{K}^{\{N\}}$

Sea $\alpha_N \in \mathbb{K}^N$ una N -ada. Diremos que α_N es **finita** si el conjunto de índices i tales que $\alpha_i \neq 0$ es finito. Si el conjunto de índices N es finito entonces, cualquier N -ada es finita (porque un subconjunto de un conjunto finito es finito). Sin embargo, si el conjunto de índices es infinito entonces habrá N -adas infinitas.

Si sumamos dos N -adas finitas el resultado será una N -ada de finita (porque $0+0=0$). Si multiplicamos una N -ada finita por un elemento del campo el resultado será una N -ada de finita (porque $\lambda 0 = 0$). Los axiomas de espacio vectorial se cumplen porque ya sabemos que se cumplen para cualesquiera N -adas. Al espacio de N -adas finitas se le denota por $\mathbb{K}^{\{N\}}$. Si N es finito $\mathbb{K}^{\{N\}} = \mathbb{K}^N$. Si N es infinito $\mathbb{K}^{\{N\}} \neq \mathbb{K}^N$.

Como ejemplo de espacio de N -adas finitas veamos el siguiente. Cada polinomio $\sum_{i=0}^n a_i x^i$ determina únicamente una N -ada a_N donde $a_i = 0$ si $i > n$. Esta N -ada es finita. Recíprocamente, si a_N es una N -ada finita entonces, necesariamente, hay un

natural n tal que $a_i = 0$ para cualquier $i > n$. Así, le podemos hacer corresponder a a_N el polinomio $\sum_{i=0}^n a_i x^i$. Esta correspondencia es biunívoca y las operaciones son las mismas en ambos espacios. Esto nos muestra que, el espacio de polinomios es el espacio de N -adas finitas. En otras palabras $\mathbb{K}^{[N]} = \mathbb{K}[x]$.

Subcampos

Sea L un subcampo de \mathbb{K} . Podemos sumar los elementos de \mathbb{K} y al multiplicar un elemento de L por uno de \mathbb{K} obtenemos un elemento de \mathbb{K} . Esto significa que tenemos una operación binaria en \mathbb{K} (la suma) y una multiplicación por escalares $L \times \mathbb{K} \rightarrow \mathbb{K}$. En este caso, los axiomas de espacio vectorial se desprenden directamente de las definiciones de campo y subcampo y obtenemos que \mathbb{K} es un espacio vectorial sobre L . Un caso muy particular es que todo campo es espacio vectorial sobre si mismo. Otro ejemplo importante es que \mathbb{R} es subcampo de \mathbb{C} . Como cada complejo se puede escribir como una pareja (a, b) con coordenadas en \mathbb{R} y la suma de complejos y el producto por un real son las mismas operaciones que en \mathbb{R}^2 tenemos que \mathbb{C} como espacio vectorial sobre \mathbb{R} es lo mismo que \mathbb{R}^2 . Otro ejemplo más es que \mathbb{R} es un espacio vectorial sobre \mathbb{Q} . Este caso es suficientemente complicado para que no digamos nada más sobre él.

El espacio de N -adas de vectores \mathbb{E}^N

Ahora, nos debemos preguntar, cual es el mínimo de condiciones que le debemos pedir a las coordenadas de una N -ada, para poder definir un espacio vectorial. Ya vimos que, si las coordenadas están en un campo entonces, obtenemos un espacio vectorial. Pero esto es pedir mucho. En realidad, solo necesitamos saber sumar las coordenadas y multiplicar cada coordenada por un escalar, o sea, necesitamos que las coordenadas sean vectores.

Más precisamente. Sea E un espacio vectorial sobre el campo \mathbb{K} . Denotaremos por \mathbb{E}^N el conjunto de todas las N -adas de E . Si a_N y b_N son dos elementos de \mathbb{E}^N entonces, la i -ésima coordenada de $a_N + b_N$ es $a_i + b_i$. Esto lo podemos hacer porque sabemos sumar los elementos de E . Ahora, si λ es un elemento de \mathbb{K} entonces, la i -ésima coordenada de λa_N es λa_i . Esto lo podemos hacer porque sabemos multiplicar los escalares en \mathbb{K} por los vectores en E . Los axiomas de espacio vectorial se demuestran fácilmente debido a que se cumplen en cada coordenada.

El espacio de NM -matrices \mathbb{K}^{NM}

Un caso particular de N -adas de vectores es cuando estos vectores son M -adas de escalares. Este espacio es $(\mathbb{K}^M)^N$. Para aclarar un poco que sucede en este caso, supongamos que $N = \{1, 2\}$ y que $M = \{1, 2, 3\}$. Entonces, un elemento del espacio $(\mathbb{K}^M)^N$ es una 2-ada (a_1, a_2) de vectores y cada uno de ellos es una 3-ada de escalares.

$$\begin{aligned}\mathbf{a}_1 &= (\alpha_{11}, \alpha_{12}, \alpha_{13}) \\ \mathbf{a}_2 &= (\alpha_{21}, \alpha_{22}, \alpha_{23})\end{aligned}$$

Los dos vectores los podemos representar como en el recuadro a la izquierda y para simplificar nos quedamos con la tabla que vemos a la derecha.

$$\left(\begin{array}{ccc} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \end{array} \right)$$

Esta tabla es un elemento del espacio de $(N \times M)$ -adas $\mathbb{K}^{N \times M}$, o sea, los índices son parejas en el producto cartesiano $N \times M$. Esto nos permite identificar $(\mathbb{K}^M)^N$ con $\mathbb{K}^{N \times M}$ y como las operaciones en ambos espacios son las mismas estos espacios son en esencia el mismo. Cambiando el orden en que ponemos los indices obtenemos las igualdades

$$(\mathbb{K}^M)^N = \mathbb{K}^{N \times M} = \mathbb{K}^{M \times N} = (\mathbb{K}^N)^M$$

que el lector debe comparar con $(a^x)^y = a^{xy} = a^{yx} = (a^y)^x$ para números naturales a, x, y .

Desde ahora, para simplificar la notación, a una $(N \times M)$ -ada cualquiera la llamaremos **NM-ada**. También, en lugar de usar la notación $\alpha_{(N \times M)}$ para denotar una **NM-ada** concreta, usaremos la notación α_{NM} . A una coordenada de esta **NM-ada** la denotaremos α_{ij} en lugar de usar la más complicada $\alpha_{(i,j)}$. En esta notación, es más cómodo pensar que hay dos conjuntos de índices (**N** y **M**) en lugar de uno solo ($N \times M$). O sea, una **NM-ada** es un conjunto de elementos del campo indexado por dos conjuntos de índices. Cuando $N = \{1, \dots, n\}$ y $M = \{1, \dots, m\}$ entonces obtenemos una **matriz** con **n** renglones y **m** columnas. En el ejemplo anterior obtuvimos una matriz con **2** renglones y **3** columnas.

Para conjuntos **N** y **M** arbitrarios (por ejemplo, conjuntos de jeroglíficos chinos), la diferencia es que no hay un orden preestablecido entre los elementos de los conjuntos de índices por lo que no sabríamos cual columna o renglón poner primero y cual después. Como esta diferencia no es tan importante y para no formar confusión en la terminología desde ahora, a las **NM-adas** los llamaremos **NM-matrices**. Escribiremos \mathbb{K}^{NM} para denotar el espacio vectorial de todas las **NM-matrices**. Como ya sabemos, en este espacio las matrices se suman por coordenadas y se multiplican por escalares también por coordenadas.

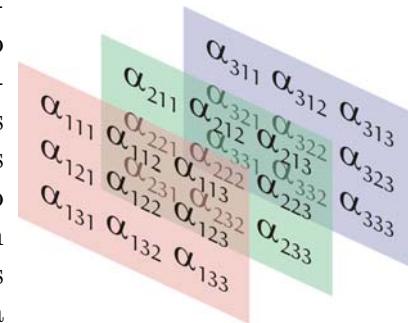
Sea α_{NM} una **NM-matriz**. Es costumbre llamarle a las coordenadas α_{ij} de α_{NM} **entradas** de la matriz. Si $i \in N$ entonces la **M-ada** $\alpha_{iM} \in \mathbb{K}^M$ es el **i-ésimo renglón** de la matriz α_{NM} . Si $j \in M$ entonces la **N-ada** $\alpha_{Nj} \in \mathbb{K}^N$ es la **j-ésima columna**. Al conjunto **N** se le llama **conjunto de índices de los renglones**. Análogamente, al conjunto **M** se le llama **conjunto de índices de las columnas**.

Si N' , M' son subconjuntos no vacíos de **N** y **M** respectivamente entonces a la matriz $\alpha_{N'M'}$ se le llama **submatriz** de α_{NM} . Si $|N'| = |M'| = 1$ entonces la submatriz es una entrada. Un **renglón** es una submatriz donde $|N'| = 1$ y $M' = M$ o sea una submatriz del tipo α_{iM} . Una **columna** es una submatriz donde $|M'| = 1$ y $N = N'$ o sea una submatriz del tipo α_{Nj} . Una última observación acerca de las matrices, es que toda esta terminología de “columnas” y “renglones” viene de la manera usual en forma de tabla en que escribimos una matriz concreta.

El espacio de tensores

Bueno, ¿y si tenemos más de dos conjuntos de índices? Pues es lo mismo. Una **NML**-ada es una $(N \times M \times L)$ -ada. Al igual que en las matrices denotaremos por α_{NML} a una **NML**-ada con tres conjuntos de índices o sea un **tensor de exponente 3**. Las matrices son tensores de exponente 2 y las **N**-adas son tensores de exponente 1. Los tensores de exponente 0 son los elementos del campo.

Si bien podemos pensar un tensor de exponente 1, como una serie de elementos del campo uno al lado de otro y pensar los de exponente 2, como una tabla rectangular de elementos del campo, también podemos pensar los tensores de exponente 3, como un conjunto de elementos del campo dispuestos en un arreglo que llenan un cubo en el espacio. Para cada $i \in N$ tenemos que α_{iML} es un tensor de exponente 2 o sea una matriz. Todo α_{NML} nos lo podemos imaginar como muchas matrices puestas una encima de la otra.



Sin embargo cuando el exponente del tensor es grande ya nuestra imaginación no da para visualizar geométricamente un arreglo de los elementos del campo dispuestos en un cubo de dimensión grande. Es mucho más útil el manejo algebraico de estos, que tratar de visualizarlos. En este contexto, la terminología de “renglones” y “columnas” se hace inutilizable. Como es lógico, los tensores con conjuntos de índices fijos forman un espacio vectorial. La suma se hace por coordenadas y también la multiplicación por un escalar.

2.3 Subespacios

Sea \mathfrak{E} un espacio vectorial sobre \mathbb{K} . Un **subespacio** es un conjunto no vacío de vectores que es un espacio vectorial para las mismas operaciones.

2.9 *Un conjunto de vectores \mathfrak{F} no vacío es un subespacio si y solo si para cualesquiera $\mathbf{a}, \mathbf{b} \in \mathfrak{F}$ y $\lambda \in \mathbb{K}$ los vectores $\mathbf{a} + \mathbf{b}$ y $\lambda\mathbf{a}$ están en \mathfrak{F} .*

Prueba. Si \mathfrak{F} es un subespacio de \mathfrak{E} entonces, por definición $\mathbf{a} + \mathbf{b}$ y $\lambda\mathbf{a}$ están en \mathfrak{F} . Recíprocamente, sea \mathfrak{F} un conjunto de vectores de \mathfrak{E} que cumple las hipótesis. Como la suma de vectores es asociativa y commutativa en \mathfrak{E} , también lo es en \mathfrak{F} . Sea $\mathbf{a} \in \mathfrak{F}$ (existe por ser \mathfrak{F} no vacío). Tenemos $0\mathbf{a} = \mathbf{0} \in \mathfrak{F}$. Además $\forall \mathbf{b} \in \mathfrak{F} \quad (-1)\mathbf{b} = -\mathbf{b} \in \mathfrak{F}$. Con esto se comprueba que $(\mathfrak{F}, +)$ es grupo abeliano. Los axiomas de espacio vectorial se cumplen en \mathfrak{F} por ser este un subconjunto de \mathfrak{E} . ■

Unión e intersección de subespacios

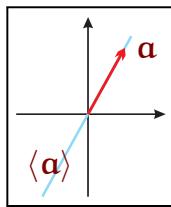
¿Serán la intersección (y la unión) de subespacios un subespacio? La unión de subespacios no es un subespacio. Por ejemplo, el eje **x** y el eje **y** en el plano cartesiano son subespacios sin embargo, la unión de los mismos no es un subespacio ya que $(0, 1) + (1, 0) = (1, 1)$ que no pertenece a la unión de los dos ejes. Por el contrario, la intersección de subespacios sí es un subespacio.

2.3 *La intersección de un conjunto de subespacios es un subespacio.*

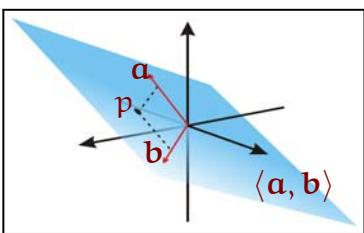
Prueba. La intersección de subespacios nunca es vacía porque cada subespacio contiene al **0**. Si **a** y **b** son dos vectores en *cada uno* de los subespacios de un conjunto entonces, $\mathbf{a} + \mathbf{b}$ también está en *cada uno* de los subespacios del conjunto. Lo mismo sucede para $\alpha\mathbf{a}$. ■

Combinaciones lineales

Veamos unos ejemplos importantes de subespacios. Sea **a** un vector no nulo en \mathbb{R}^2 . El conjunto $\langle \mathbf{a} \rangle = \{\alpha\mathbf{a} \mid \alpha \in \mathbb{R}\}$ de todos los múltiplos del vector **a** es un subespacio de \mathbb{R}^2 ya que $\alpha\mathbf{a} + \beta\mathbf{a} = (\alpha + \beta)\mathbf{a}$, y $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$. Geométricamente, teniendo en cuenta la identificación de vectores con los puntos del plano cartesiano, vemos que los vectores en este espacio son los puntos de la recta por el origen que pasa por **a**.



Sean ahora **a** y **b** dos vectores en \mathbb{R}^3 no colineales o sea $\mathbf{a} \neq \beta\mathbf{b}$. Consideremos el conjunto de vectores $\{\alpha\mathbf{a} + \beta\mathbf{b} \mid \alpha, \beta \in \mathbb{R}\}$. Este conjunto de vectores se denota por $\langle \mathbf{a}, \mathbf{b} \rangle$ y es un subespacio ya que $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{a} + \delta\mathbf{b} = (\alpha + \gamma)\mathbf{a} + (\beta + \delta)\mathbf{b}$ y $\lambda(\alpha\mathbf{a} + \beta\mathbf{b}) = \lambda\alpha\mathbf{a} + \lambda\beta\mathbf{b}$. Geométricamente, vemos que este conjunto contiene a la recta $\langle \mathbf{a} \rangle$ que pasa por **a** y a la línea $\langle \mathbf{b} \rangle$ que pasa por **b**.



En \mathbb{R}^3 hay un solo plano que contiene estas dos rectas. ¿Será este plano el subespacio $\langle \mathbf{a}, \mathbf{b} \rangle$? Claro que sí. Para cada punto **p** de este plano dibujemos la paralela a la recta $\langle \mathbf{a} \rangle$ que pasa por **p** obteniendo un punto de intersección con la recta $\langle \mathbf{b} \rangle$ el cual es $\beta\mathbf{b}$ para algún $\beta \in \mathbb{R}$. Análogamente, dibujando la paralela a $\langle \mathbf{b} \rangle$ obtenemos el punto $\alpha\mathbf{a}$ en la recta $\langle \mathbf{a} \rangle$ y vemos que $\mathbf{p} = \alpha\mathbf{a} + \beta\mathbf{b}$. Luego, $\mathbf{p} \in \langle \mathbf{a}, \mathbf{b} \rangle$.

Estos ejemplos nos motivan a la siguiente definición. Sea **N** un conjunto de vectores. Una **combinación lineal** de **N** es cualquier vector de la forma que se muestra en el recuadro a la derecha. A los escalares α_i se les llama **coeficientes** de la combinación lineal.

$$\sum_{i \in N} \alpha_i \mathbf{i}$$

Es posible que no se entienda esta notación. El conjunto de índices es el conjunto de vectores **N** por lo que en la suma hay tantos sumandos como vectores hay en **N**. Así por ejemplo si **N** = {**a**, **b**, **c**, **d**} entonces una combinación lineal de **N** es un

vector de la forma $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} + \delta\mathbf{d}$. En otras palabras, la notación $\sum_{i \in N} \alpha_i \mathbf{i}$ debe interpretarse así: “tómese los vectores de N , cada uno de ellos multiplíquese por un escalar y símese los resultados”. No importa el orden de los sumandos, la suma de vectores es commutativa

Todo esto está muy bien si el conjunto N es finito. Si N es infinito tenemos un problema. ¿Qué es una suma infinita de vectores? La manera de evitar este problema es que le pediremos a los coeficientes α_i que *todos salvo un número finito sean cero* o sea, que la N -ada α_N cuyas coordenadas son los coeficientes de la combinación lineal es finita. Como podemos despreciar los sumandos iguales a cero, tenemos que una combinación lineal es siempre una suma de un número finito de vectores.

2.4

El conjunto de todas las combinaciones lineales de N es un subespacio.

Prueba. Sean $\sum_{i \in N} \alpha_i \mathbf{i}$ y $\sum_{i \in N} \beta_i \mathbf{i}$ dos combinaciones lineales de N entonces, de la asociatividad y commutatividad de la suma de vectores y de la distributividad del producto por escalares obtenemos:

$$\sum_{i \in N} \alpha_i \mathbf{i} + \sum_{i \in N} \beta_i \mathbf{i} = \sum_{i \in N} (\alpha_i + \beta_i) \mathbf{i}$$

que es una combinación lineal de N ya que todos los $(\alpha_i + \beta_i)$ salvo un número finito tienen que ser cero. Lo mismo ocurre con $\gamma (\sum_{i \in N} \alpha_i \mathbf{i}) = \sum_{i \in N} \gamma \alpha_i \mathbf{i}$. ■

Cerradura lineal

Denotaremos por $\langle N \rangle$ al conjunto de todas las combinaciones lineales de N .

2.5

Si \mathfrak{F} es un subespacio que contiene a N entonces, \mathfrak{F} contiene a $\langle N \rangle$.

Prueba. Si $N \subseteq \mathfrak{F}$ entonces, \mathfrak{F} contiene a todos los múltiplos de los vectores en N y a todas sus sumas finitas. Luego, cualquier combinación lineal de N está en \mathfrak{F} . ■

2.6

Los siguientes tres conjuntos de vectores coinciden

1. *El conjunto de todas las combinaciones lineales de N .*
2. *La intersección de todos los subespacios que contienen a N .*
3. *El subespacio más pequeño que contiene a N .*

Prueba. Denotemos por \mathfrak{F}_1 , \mathfrak{F}_2 y \mathfrak{F}_3 a los tres conjuntos del enunciado de la proposición. Por definición $\mathfrak{F}_1 = \langle N \rangle$. Por la proposición 2.3 el conjunto \mathfrak{F}_2 es un subespacio que contiene a N y de 2.5 obtenemos que $\mathfrak{F}_1 \subseteq \mathfrak{F}_2$. Por definición de intersección tenemos que $\mathfrak{F}_2 \subseteq \mathfrak{F}_3$. Por definición, \mathfrak{F}_3 está contenido en cualquier subespacio que contiene a N y en particular $\mathfrak{F}_3 \subseteq \mathfrak{F}_1$. Resumiendo, $\mathfrak{F}_1 \subseteq \mathfrak{F}_2 \subseteq \mathfrak{F}_3 \subseteq \mathfrak{F}_1$. ■

A cualquiera de estos tres conjuntos (¡son iguales!) se le denota por $\langle \mathbf{N} \rangle$ y se le llama **cerradura lineal** de \mathbf{N} o también el **subespacio generado** por \mathbf{N} .

Propiedades de la cerradura lineal

2.7

La cerradura lineal cumple las siguientes propiedades:

- ◆ $\mathbf{N} \subseteq \langle \mathbf{N} \rangle$ (incremento),
- ◆ $\mathbf{N} \subseteq \mathbf{M} \Rightarrow \langle \mathbf{N} \rangle \subseteq \langle \mathbf{M} \rangle$ (monotonía),
- ◆ $\langle \langle \mathbf{N} \rangle \rangle = \langle \mathbf{N} \rangle$ (idempotencia).

Prueba. El incremento es inmediato de 2.6.3. Supongamos ahora que $\mathbf{N} \subseteq \mathbf{M}$. Cualquier combinación lineal de \mathbf{N} es también combinación lineal de \mathbf{M} (haciendo cero los coeficientes de los vectores en $\mathbf{M} \setminus \mathbf{N}$). Finalmente, por 2.6.3 $\langle \langle \mathbf{N} \rangle \rangle$ es el subespacio más pequeño que contiene a $\langle \mathbf{N} \rangle$. Como $\langle \mathbf{N} \rangle$ es un subespacio entonces, $\langle \langle \mathbf{N} \rangle \rangle = \langle \mathbf{N} \rangle$. ■

En matemáticas las palabras “clausura”, “cerradura” y “envoltura” son sinónimos. Por esto con el mismo éxito, otros autores le llaman “clausura lineal” o “envoltura lineal” a nuestro concepto de cerradura lineal. Además, es bueno que el lector encuentre el parecido entre el concepto de cerradura lineal y el concepto de “cerradura” en análisis (la intersección de todos los cerrados que contienen a un conjunto).



En general una función que a un conjunto \mathbf{N} le hace corresponder otro conjunto $\langle \mathbf{N} \rangle$ y que cumple las propiedades 1-3 se le llama operador de cerradura (clausura, envoltura).

En este caso a los conjuntos tales que $\mathbf{N} = \langle \mathbf{N} \rangle$ se le llaman cerrados. Que se cumplan las propiedades 1-3 es, en general, equivalente a que el operador de cerradura se defina como la intersección de todos los cerrados que contienen al conjunto. En todas las ramas de las matemáticas se encuentran operadores de cerradura.

Ejercicio 37 Demuestre que \mathbf{N} es un subespacio si y solo si $\mathbf{N} = \langle \mathbf{N} \rangle$.

Ejercicio 38 Demuestre que $(\mathbf{x} \in \langle \mathbf{N} \cup \mathbf{y} \rangle \setminus \langle \mathbf{N} \rangle) \Rightarrow (\mathbf{y} \in \langle \mathbf{N} \cup \mathbf{x} \rangle)$. [189]

2.4 Bases

$$\mathbb{K}^{\{\mathbf{N}\}} \ni \alpha_{\mathbf{N}} \mapsto \sum_{\mathbf{i} \in \mathbf{N}} \alpha_{\mathbf{i}} \mathbf{i} \in \mathfrak{E}$$

Observemos que la definición de combinación lineal de \mathbf{N} determina la función $f_{\mathbf{N}}$ del recuadro a la izquierda que a cada \mathbf{N} -ada finita $\alpha_{\mathbf{N}}$ le hace corresponder el vector $\sum_{\mathbf{i} \in \mathbf{N}} \alpha_{\mathbf{i}} \mathbf{i}$.

Esta función no tiene porque ser sobreyectiva ni inyectiva, depende del conjunto de vectores \mathbf{N} . Por ejemplo, sea $\mathbf{N} = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} \subseteq \mathbb{R}^3$ donde $\mathbf{x} = (0, 0, 1)$, $\mathbf{y} = (0, 1, 0)$ y $\mathbf{z} = (0, 1, 1)$. En este caso $f_{\mathbf{N}}(2, 2, 0) = f_{\mathbf{N}}(0, 0, 2) = (0, 2, 2)$ por lo que $f_{\mathbf{N}}$ no es inyectiva. Tampoco es sobreyectiva porque $(1, 0, 0)$ no tiene preimagen.

En esta sección estableceremos que en todo espacio vectorial existen conjuntos \mathbf{N} para los cuales la función $f_{\mathbf{N}}$ es biyectiva. Este hecho es fundamental, porque en este caso existe la función inversa $f_{\mathbf{N}}^{-1} : \mathfrak{E} \rightarrow \mathbb{K}^{|\mathbf{N}|}$ que nos permitirá introducir coordenadas en cualquier espacio vectorial. Es más, veremos que los conjuntos de vectores para los cuales $f_{\mathbf{N}}$ es biyectiva tienen la misma cantidad de vectores. Esto quiere decir que cada vector de \mathfrak{E} se define por un cierto número de elementos del campo (coordenadas) y que este número es independiente del vector a definir. Solo depende del espacio. Así, en \mathbb{R}^2 nos hacen falta siempre dos reales y en \mathbb{R}^7 nos hacen falta siete.

Conjuntos generadores

Primero, lo más fácil. La imagen de la función $f_{\mathbf{N}}$ es $\langle \mathbf{N} \rangle$, o sea, la cerradura lineal de \mathbf{N} . Diremos que \mathbf{N} es un **conjunto generador** si $\langle \mathbf{N} \rangle$ es todo el espacio o sea, si $f_{\mathbf{N}}$ es sobreyectiva.

Los generadores son un filtro

2.8

Todo sobreconjunto de un conjunto generador es generador.

Prueba. Supongamos $\mathbf{M} \supseteq \mathbf{N}$. Por monotonía de la cerradura lineal tenemos $\langle \mathbf{N} \rangle \subseteq \langle \mathbf{M} \rangle$. Si $\langle \mathbf{N} \rangle$ es todo el espacio entonces, necesariamente $\langle \mathbf{M} \rangle$ también lo es. ■

Conjuntos linealmente independientes

Ahora, veamos cuando $f_{\mathbf{N}}$ es inyectiva. Observese que en un lenguaje más descriptivo, $f_{\mathbf{N}}$ es inyectiva si y solo si se cumple la propiedad 3 del siguiente resultado.

Teorema de Caracterización de Conjuntos Linealmente Independientes

2.9

Sea \mathbf{N} un conjunto de vectores. Las siguientes afirmaciones son equivalentes:

1. *Cualquier subconjunto propio de \mathbf{N} genera un subespacio más pequeño que $\langle \mathbf{N} \rangle$.*
2. *Cualquier vector en \mathbf{N} no es combinación lineal de los restantes.*
3. *Si dos combinaciones lineales de \mathbf{N} son iguales entonces, todos sus coeficientes son iguales.*
4. *Si una combinación lineal de \mathbf{N} es cero entonces, todos sus coeficientes son cero.*

Prueba. (1 \Leftrightarrow 2) Si 1 no es cierto entonces existe $\mathbf{a} \in \mathbf{N}$ tal que $\langle \mathbf{N} \rangle = \langle \mathbf{N} \setminus \mathbf{a} \rangle$ pero entonces $\mathbf{a} \in \langle \mathbf{N} \setminus \mathbf{a} \rangle$ lo que contradice 2. Recíprocamente, si 2 no es cierto entonces existe $\mathbf{a} \in \mathbf{N}$ tal que $\mathbf{a} \in \langle \mathbf{N} \setminus \mathbf{a} \rangle$. Por incremento $\mathbf{N} \setminus \mathbf{a} \subseteq \langle \mathbf{N} \setminus \mathbf{a} \rangle$ y como además $\mathbf{a} \in \langle \mathbf{N} \setminus \mathbf{a} \rangle$ entonces $\mathbf{N} \subseteq \langle \mathbf{N} \setminus \mathbf{a} \rangle$. Por monotonía e idempotencia $\langle \mathbf{N} \rangle \subseteq \langle \mathbf{N} \setminus \mathbf{a} \rangle$ lo que contradice 1.

(3 \Leftrightarrow 4) Observese que $(\sum_{i \in N} \alpha_i i = \sum_{i \in N} \beta_i i) \Leftrightarrow (\sum_{i \in N} (\alpha_i - \beta_i) i = 0)$ y además $(\alpha_i = \beta_i) \Leftrightarrow (\alpha_i - \beta_i = 0)$. Luego, la existencia de combinaciones lineales iguales con algunos coeficientes diferentes es equivalente a la existencia de combinaciones lineales iguales a cero con algunos coeficientes no cero.

(2 \Leftrightarrow 4) Si no se cumple 2 entonces hay un vector $a \in N$ que es combinación lineal de $N \setminus a$. Pasando todos los sumandos hacia un lado de la igualdad obtenemos una combinación lineal de N igual a cero con un coeficiente distinto de cero. Esto contradice 4. Recíprocamente, si no se cumple 4 entonces hay un vector $a \in N$ y una combinación lineal de N igual a cero y con $\alpha_a \neq 0$. Despejando a , obtenemos que $a \in \langle N \setminus a \rangle$. Esto contradice 2. ■

Ejercicio 39 Busque en la prueba del Teorema de Caracterización de Conjuntos Linealmente Independientes (2.9) donde se usan los inversos multiplicativos. [189]

Diremos que un conjunto de vectores N es **linealmente independiente** si se cumple alguna (y por lo tanto todas) las afirmaciones de la proposición anterior. Los conjuntos de vectores que no las cumplen se les llama **linealmente dependientes**.



A partir de ahora para no repetir constantemente frases largas, a los conjuntos de vectores linealmente independientes los llamaremos **conjuntos LI**. A los conjuntos de vectores linealmente dependientes los llamaremos **conjuntos LD**. Estas notaciones se deben pronunciar “ele i” y “ele de”.

Los independientes son un ideal

2.10

Todo subconjunto de un conjunto LI es LI.

Prueba. Sea $N \subseteq M$ tal que M es LI. Cualquier combinación lineal de N es combinación lineal de M . Luego toda combinación lineal de N igual a cero tiene todos sus coeficientes iguales a cero ■

Antes de pasar a la definición de base, demostremos un pequeño resultado que usaremos repetidamente en lo adelante.

Lema de Aumento de un Conjunto LI

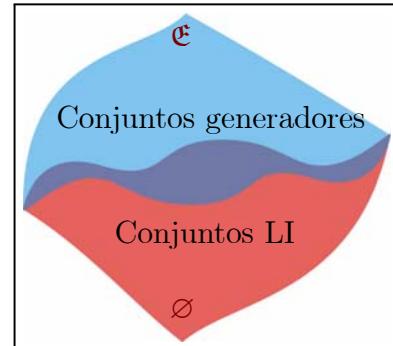
2.11

Si N es independiente y $N \cup a$ es dependiente entonces $a \in \langle N \rangle$.

Prueba. Si $N \cup a$ es LD entonces, por la caracterización 2.9.4 de los conjuntos LI, hay una combinación de $N \cup a$ igual a cero cuyos coeficientes no todos son igual a cero. Si el coeficiente en a fuera igual a cero tendríamos una contradicción con que N es LI. Luego, el coeficiente en a es diferente a cero y despejando a obtenemos $a \in \langle N \rangle$. ■

Bases

La relación entre los conjuntos generadores y los conjuntos LI la podemos ver intuitivamente en la figura a la derecha. Aquí están representados los subconjuntos del espacio \mathbb{E} que son generadores o que son LI. Mientras más arriba más grande es el subconjunto. Nuestro interés ahora se va a concentrar en la franja del centro donde hay subconjuntos que a la vez son generadores y LI. La siguiente proposición nos dice que “esta franja no puede ser muy ancha”.

**Teorema de Caracterización de Bases**

2.12

Sea \mathbf{N} un conjunto de vectores. Las siguientes afirmaciones son equivalentes

1. \mathbf{N} es generador y \mathbf{N} es LI,
2. \mathbf{N} es LI y cualquier sobreconjunto propio de \mathbf{N} es LD,
3. \mathbf{N} es generador y cualquier subconjunto propio de \mathbf{N} no es generador.

Prueba. (1 \Rightarrow 2) Sea \mathbf{N} independiente y generador. Sea $\mathbf{a} \notin \mathbf{N}$. Como \mathbf{N} es generador $\mathbf{a} \in \langle \mathbf{N} \rangle$. De la caracterización 2.9.2 de los conjuntos LI se sigue que $\mathbf{N} \cup \mathbf{a}$ es LD. Luego todo sobreconjunto propio de \mathbf{N} es LD.

(2 \Rightarrow 3) Sea \mathbf{N} independiente maximal. Por incremento $\mathbf{N} \subseteq \langle \mathbf{N} \rangle$. Si $\mathbf{a} \notin \mathbf{N}$ entonces $\mathbf{N} \cup \mathbf{a}$ es LD. Por el Lema de Aumento de un Conjunto LI (2.11) tenemos $\mathbf{a} \in \langle \mathbf{N} \rangle$. Luego, \mathbf{N} es generador. Si algún subconjunto propio de \mathbf{N} fuera generador entonces existiría $\mathbf{a} \in \mathbf{N}$ tal que $\mathbf{a} \in \langle \mathbf{N} \setminus \mathbf{a} \rangle$ y por la caracterización 2.9.2 de los conjuntos LI esto contradice la suposición de que \mathbf{N} es LI.

(3 \Rightarrow 1) Sea \mathbf{N} generador minimal. Si \mathbf{N} es LD entonces, por la caracterización 2.9.1 de los conjuntos LI existe subconjunto propio \mathbf{M} de \mathbf{N} tal que $\langle \mathbf{M} \rangle = \langle \mathbf{N} \rangle$. Como \mathbf{N} es generador entonces \mathbf{M} también lo es. Esto contradice que \mathbf{N} es minimal. ■

Sea \mathfrak{F} un subespacio. Un conjunto de vectores que es generador de \mathfrak{F} y que es LI se le llama **base** de \mathfrak{F} . Las bases de todo el espacio se llaman simplemente **bases**. El ser base es equivalente a ser un conjunto LI lo más grande posible o ser un conjunto generador lo más pequeño posible. También, el ser base es equivalente a que nuestra función $f_{\mathbf{N}}$ del principio de esta sección sea biyectiva.

Lema de Incomparabilidad de las Bases

2.13

Dos bases diferentes no están contenidas una dentro de la otra.

Prueba. Si $\mathbf{N} \subseteq \mathbf{M}$ son dos bases diferentes entonces \mathbf{N} es un subconjunto propio de \mathbf{M} y por el Teorema de Caracterización de Bases (2.12) esto es una contradicción. ■

Teorema de Existencia de Bases

2.14

Sea N un conjunto generador y $L \subseteq N$ un conjunto LI. Entonces, existe una base M tal que $L \subseteq M \subseteq N$.

Prueba. Sean L y N como en las hipótesis del teorema. Sea M un conjunto LI tal que $L \subseteq M \subseteq N$. Si M es maximal con estas propiedades ($\forall a \in N \setminus M \quad M \cup a$ es LD) entonces, por el Lema de Aumento de un Conjunto LI (2.11) tenemos $N \subseteq \langle M \rangle$. Como N es generador, esto significa que M también lo es y por lo tanto M es una base.

Nuestra tarea es encontrar un M que es maximal con estas propiedades. Esto es fácil si el conjunto N es finito. Primero ponemos M igual a L . Si M es maximal terminamos, si no, entonces existe $a \in N \setminus M$ tal que $M \cup a$ es LI. Agregamos a al conjunto M y repetimos este proceso. Como N es finito este proceso tiene que terminar. ■

Ejercicio 40 Usando el Teorema de Existencia de Bases pruebe que: 1. Todo espacio vectorial tiene base. 2. Cualquier conjunto LI esta contenido en una base. 3. Cualquier conjunto generador contiene a una base. [189]

Dimensión

Ahora lo que queremos ver es que todas las bases tienen el mismo número de vectores. Para probar esto se necesita ver primero una propiedad clásica de las bases.

Propiedad del Cambio de las Bases

2.15

Si M y N son dos bases entonces, para cualquier $a \in M$ existe $b \in N$ tal que $(M \setminus a) \cup b$ es base.

Prueba. Sea a un vector fijo pero arbitrario en M . Para un vector $b \in N$ denotemos $M_b = (M \setminus a) \cup b$. Si para todos los $b \in N \setminus (M \setminus a)$ el conjunto M_b fuera LD entonces, por el Lema de Aumento de un Conjunto LI (2.11) todo $b \in N$ sería combinación lineal de $M \setminus a$. O sea, $N \subset \langle M \setminus a \rangle$ y por lo tanto $\langle N \rangle \subseteq \langle M \setminus a \rangle$. Como $\langle N \rangle$ es todo el espacio tendríamos $a \in \langle M \setminus a \rangle$ lo que no es posible ya que M es LI.

Hemos demostrado que existe $b \in N \setminus (M \setminus a)$ tal que M_b es LI. Demostremos que M_b es generador. Sea v un vector arbitrario. Por ser M base tenemos que b y v son combinaciones lineales de M o sea existen escalares apropiados tales que

$$b = \alpha_a a + \sum_{x \in M \setminus a} \alpha_x x \quad v = \lambda_a a + \sum_{x \in M \setminus a} \lambda_x x$$

Como M_b es LI entonces, b no es una combinación lineal de $M \setminus a$ y por lo tanto $\alpha_a \neq 0$. Luego, podemos despejar a de la primera ecuación, substituirla en la segunda y así obtener que $v \in \langle M_b \rangle$. Esto significa que M_b es generador. ■

Equivocación de las Bases

2.16

Dos bases cualesquiera de un espacio vectorial tienen el mismo cardinal.

Prueba. Sean $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ y B dos bases. Por la Propiedad del Cambio de las Bases (2.15) existe otra base $A_1 = \{\mathbf{b}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ donde $\mathbf{b}_1 \in B$. Observese que $|A_1| = n$ ya que en otro caso $A_1 \subsetneq A$ lo que contradice el Lema de Incomparabilidad de las Bases (2.13). Aplicando la Propiedad del Cambio de las Bases (2.15) a A_1 obtenemos otra base $A_2 = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{a}_3, \dots, \mathbf{a}_n\}$ tal que $\mathbf{b}_2 \in B$. Observese que $|A_2| = n$ ya que en otro caso $A_2 \subsetneq A_1$ lo que contradice el Lema de Incomparabilidad de las Bases (2.13). Repitiendo este proceso obtenemos bases A_3, A_4, \dots, A_n todas con n vectores y además $A_n \subseteq B$. Como B es base $A_n = B$ y por lo tanto B también tiene n vectores. ■

Al cardinal de una base (cualesquiera) se le llama **dimensión** del espacio vectorial. La dimensión de un espacio vectorial E se denotará por $\dim E$. Los espacios vectoriales que tienen una base finita se les llama **finito dimensionales** o de **dimensión finita**. Por el contrario, si sus bases son infinitas entonces, el espacio vectorial es de **dimensión infinita**. La teoría de los espacios vectoriales de dimensión finita es más sencilla pero más completa que la de los espacios de dimensión infinita. Para darnos cuenta de que hay muchas cosas que no se cumplen en el caso infinito dimensional veamos como ejemplo el siguiente resultado que tendremos múltiples ocasiones para utilizarlo.

2.17

Sea \mathfrak{F} un espacio vectorial finito dimensional y E un subespacio de \mathfrak{F} . Si $\dim E = \dim \mathfrak{F}$ entonces $E = \mathfrak{F}$.

Prueba. Sea N una base de E . Por el Teorema de Existencia de Bases (2.14) existe una base M del espacio \mathfrak{F} que contiene a N . Como M es finita entonces, N también es finita. Como las dimensiones coinciden, el número de vectores en N es igual al número de vectores en M . Luego $N = M$ y por lo tanto $E = \langle N \rangle = \langle M \rangle = \mathfrak{F}$. ■



La prueba de la Equicardinalidad de las Bases la hemos hecho solamente para el caso que el espacio tiene una base finita. Nuestra prueba del Teorema de Existencia de Bases es solo para el caso que el conjunto generador es finito. Finalmente, solo probamos que los espacios vectoriales que tienen un conjunto generador finito tienen base. Es importante que el lector sepa que estas tres afirmaciones son válidas en general. Las pruebas generales dependen de un conocimiento más profundo de la Teoría de Conjuntos que no presuponemos que lo posea el lector. A los interesados en esto les recomiendo leer ahora la última sección de este capítulo.

Ejercicio 41 Pruebe que si E un subespacio de \mathfrak{F} entonces $\dim E \leq \dim \mathfrak{F}$. [189]

Ejercicio 42 Encuentre un ejemplo donde no se cumple 2.17. [189]

Bases canónicas

Ya sabemos que todo espacio vectorial tiene bases. Sin embargo no solamente tienen una base sino muchas. Por esto es conveniente construir bases que son las más “sencillas” para los ejemplos de espacios vectoriales que construimos en la Sección 2.2.

\mathbb{R}^2 Comenzemos con \mathbb{R}^2 . Sean $e_1 = (1, 0)$ y $e_2 = (0, 1)$. Cualquier vector (a, b) en \mathbb{R}^2 tiene una manera de expresarse como combinación lineal de $N = \{e_1, e_2\}$. Efectivamente, tenemos la igualdad $(a, b) = ae_1 + be_2$. Esto quiere decir que el conjunto N es generador. Por otro lado, si $\alpha e_1 + \beta e_2 = (\alpha, \beta) = 0 = (0, 0)$ entonces, necesariamente, α y β son ambos cero. De aquí obtenemos que conjunto N es una base que la llamamos **base canónica** de \mathbb{R}^2 . Los vectores e_1 y e_2 se pueden visualizar geométricamente como los dos vectores de longitud 1 en ambos ejes de coordenadas. Luego, la dimensión de \mathbb{R}^2 es 2.

\mathbb{K}^n Pasemos ahora a \mathbb{K}^n . Para cada i en el conjunto de índices $\{1, \dots, n\}$ denotemos por e_i el vector cuya i -esima coordenada es 1 y todas las demás son cero (recuerde que todo campo tiene uno y cero). Sea N el conjunto de todos esos vectores. Tenemos $(\alpha_1, \dots, \alpha_n) = \sum_{i=1}^n \alpha_i e_i$ y esto significa que cada vector en \mathbb{K}^n se expresa de forma única como combinación lineal de N . O sea, N es generador y por la caracterización 2.9.3 de los conjuntos LI el conjunto N es LI. A la base N se le llama **base canónica** de \mathbb{K}^n . La dimensión de \mathbb{K}^n es n .

$\mathbb{K}[x]$ Veamos el espacio de polinomios $\mathbb{K}[x]$. Sea N el conjunto $\{x^i \mid i \in \mathbb{N}\}$. Es claro que todo polinomio se puede escribir de forma única como combinación lineal finita de N . A la base N se le llama **base canónica** de $\mathbb{K}[x]$. Luego, $\mathbb{K}[x]$ es de dimensión infinita contable. Desafortunadamente, para el espacio de series no se puede dar explicitamente una base.

$\mathbb{K}^{[N]}$ Para el espacio $\mathbb{K}^{[N]}$ de todas las N -adas finitas ya hicimos casi todo el trabajo (véase \mathbb{K}^n). Para cada i en el conjunto de índices N denotemos por e_i el vector cuya i -esima coordenada es 1 y las demás son cero. Sea M el conjunto de todos esos vectores. Tenemos $\alpha_N = \sum_{e_i \in M} \alpha_i e_i$ donde los coeficientes son distintos de cero solamente en un subconjunto finito de indices. Luego, M es una base de este espacio la cual es su **base canónica**. La dimensión de $\mathbb{K}^{[N]}$ es el cardinal de N ya que hay una biyección entre N y M .

\mathbb{K}^N Recordemos al lector que lo de N -ada finita es no trivial solo para cuando el conjunto N es infinito. Si el conjunto de índices es finito entonces estamos hablando de todas las N -adas o sea de \mathbb{K}^N . Luego, en el caso de que N es finito $\mathbb{K}^{[N]} = \mathbb{K}^N$ y la **base canónica** de \mathbb{K}^N es la construida en el párrafo anterior. En el caso de que el conjunto N sea infinito entonces el espacio \mathbb{K}^N no tiene una base canónica.

El campo de los números complejos como espacio vectorial sobre los reales tienen como **base canónica** al conjunto $\{1, i\}$ y por lo tanto tiene dimensión 2. Los reales como espacio vectorial sobre los racionales no tienen una base canónica.

Todos los espacios que hemos visto que no tienen base canónica son de dimensión infinita. Pero, *no todos los espacios de dimensión finita tienen una base canónica*. El

ejemplo más sencillo es tomar un subespacio de un espacio de dimensión finita. Si este subespacio es suficientemente general, entonces no hay manera de construir una base canónica del mismo incluso en el caso de que todo el espacio tenga una base canónica.

Ejercicio 43 ¿Cuál es la base canónica del espacio de las **NM-matrices**? [189]

2.5 Clasificación de espacios vectoriales

En esta sección veremos que todo espacio vectorial es isomorfo a un espacio de **N-adas** finitas. Para el caso finito dimensional esto quiere decir que el único (salvo isomorfismos) espacio vectorial de dimensión **n** sobre un campo **K** que existe es el espacio de las **n-adas** **Kⁿ**.

Isomorfismos lineales

En el Capítulo 1 vimos los morfismos de operaciones binarias, grupos, anillos y campos. Para los espacios vectoriales es igual, la única diferencia es que en ellos hay definida una operación que no es interna del espacio: el producto de un escalar por un vector. Sin embargo, esto no presenta mayores dificultades.

$$\begin{aligned} f(\mathbf{a} + \mathbf{b}) &= f(\mathbf{a}) + f(\mathbf{b}) \\ f(\alpha\mathbf{a}) &= \alpha f(\mathbf{a}) \end{aligned}$$

Sean \mathfrak{E} y \mathfrak{F} dos espacios vectoriales sobre *el mismo* campo \mathbb{K} . Una función $f : \mathfrak{E} \rightarrow \mathfrak{F}$ se le llama **morfismo de espacios vectoriales** si para cualesquiera $\mathbf{a}, \mathbf{b} \in \mathfrak{E}$ y cualquier $\alpha \in \mathbb{K}$ se cumplen las propiedades del recuadro.

A los morfismos de espacios vectoriales también se les llama **transformaciones lineales**. Esta última expresión será la que usemos porque tiene dos importantes ventajas. Primero, es más corta que la primera. Segundo es mucho más antigua, hay mucha más cantidad de personas que la conoce y por lo tanto facilita la comunicación con ingenieros y científicos de diversas áreas.

Ejercicio 44 Demuestre que la composición de morfismos es un morfismo. [189]

El estudio de las transformaciones lineales lo pospondremos hasta el siguiente capítulo. Aquí estaremos interesados en los isomorfismos de espacios vectoriales. Una transformación lineal $f : \mathfrak{E} \rightarrow \mathfrak{F}$ es un **isomorfismo de espacios vectoriales** o **isomorfismo lineal** si esta es biyectiva. Análogamente al caso de operaciones binarias tenemos el siguiente resultado:

2.18 La inversa de cualquier isomorfismo lineal es un isomorfismo lineal.

Prueba. Sea $\alpha \in \mathbb{K}$ y $x, y \in \mathfrak{F}$. Como f es una biyección existen $a, b \in \mathfrak{E}$ tales que $f(a) = x$, $f(b) = y$. Como f es isomorfismo tenemos

$$\begin{aligned} f^{-1}(x+y) &= f^{-1}(f(a)+f(b)) = f^{-1}(f(a+b)) = a+b = f^{-1}(x)+f^{-1}(y) \\ f^{-1}(\alpha x) &= f^{-1}(\alpha f(a)) = f^{-1}(f(\alpha a)) = \alpha a = \alpha f^{-1}(x) \end{aligned}$$

que es lo que se quería demostrar. ■

Ya observamos, que los isomorfismos son nada más que un cambio de los nombres de los elementos y las operaciones. Esto quiere decir que “cualquier propiedad” debe conservarse al aplicar un isomorfismo. Lo siguiente es un ejemplo de esta afirmación.

2.19

Un isomorfismo transforma conjuntos LI en conjuntos LI, conjuntos generadores en conjuntos generadores y bases en bases.

Prueba. Sea $f : \mathfrak{E} \rightarrow \mathfrak{F}$ un isomorfismo de espacios vectoriales. Sean $M \subseteq \mathfrak{E}$ y $N \stackrel{\text{def}}{=} f(M) \subseteq \mathfrak{F}$. Sean además $x \in \mathfrak{E}$, $y = f(x) \in \mathfrak{F}$. Como f es isomorfismo tenemos

$$\left(\sum_{i \in M} \alpha_i i = x \right) \Leftrightarrow \left(\sum_{i \in M} \alpha_i f(i) = y \right) \Leftrightarrow \left(\sum_{j \in N} \alpha_j j = y \right)$$

donde el conjunto de coeficientes $\{\alpha_i \mid i \in M\}$ es exactamente el mismo que $\{\alpha_j \mid j \in N\}$.

Si M es generador entonces, para cualquier $y \in \mathfrak{F}$ el vector $x = f^{-1}(y)$ es combinación lineal de M y por la equivalencia anterior el vector y es combinación lineal de N . Luego, si M es generador entonces N también lo es.

Supongamos que M es LI entonces, poniendo $x = 0$ obtenemos

$$\left(\sum_{i \in M} \alpha_i i = 0 \right) \Leftrightarrow \left(\sum_{j \in N} \alpha_j j = 0 \right)$$

luego, cualquier combinación lineal de N que es nula también es combinación lineal nula de M y por lo tanto todos sus coeficientes son cero. Luego, N es LI. ■

Ejercicio 45 Demuestre que los isomorfismos comutan con el operador de cerradura lineal y que trasforman subespacios en subespacios de la misma dimensión.

Coordinatización

Sea N un conjunto de vectores del espacio vectorial \mathfrak{F} . Al principio de la sección anterior introducimos la función f_N que a cada N -ada finita le hace corresponder la combinación lineal cuyos coeficientes son dicha N -ada. Esta función es siempre una transformación lineal ya que

$$\begin{aligned} f_N(\alpha_N + \beta_N) &= \sum_{i \in N} (\alpha_i + \beta_i) i = \sum_{i \in N} \alpha_i i + \sum_{i \in N} \beta_i i = f_N(\alpha_N) + f_N(\beta_N) \\ f_N(\lambda \alpha_N) &= \sum_{i \in N} (\lambda \alpha_i) i = \lambda \sum_{i \in N} \alpha_i i = \lambda f_N(\alpha_N). \end{aligned}$$

Por otro lado, ya vimos que f_N es sobreyectiva si y solo si N es generador, que f_N es inyectiva si y solo si N es LI y por lo tanto f_N es un isomorfismo si y solo si N es una base. En el caso de que N sea una base, la función inversa $f_N^{-1} : \mathfrak{F} \rightarrow \mathbb{K}^{\{N\}}$ es un isomorfismo lineal llamado **coordinatización** de \mathfrak{F} mediante la base N . En otras palabras, si N es una base de \mathfrak{F} , y x es un vector de \mathfrak{F} entonces, existe una única N -ada $\alpha_N \in \mathbb{K}^{\{N\}}$ tal que $x = \sum_{i \in N} \alpha_i i$. En este caso, a los coeficientes α_i se les llama **coordenadas** de x en la base N .

Clasificación

Se dice que dos espacios son **isomorfos** si existe una función que es un isomorfismo entre ellos. No hay ninguna razón (por ahora) para que deseemos distinguir dos espacios vectoriales que son isomorfos. Si esto sucede, uno es igual a otro salvo los “nombres” de los vectores y las operaciones. La clave para ver que los espacios vectoriales son isomorfos es que estos tengan la misma dimensión.

2.20

Dos espacios vectoriales sobre el mismo campo son isomorfos si y solo si tienen la misma dimensión.

Prueba. Sean E y \mathfrak{F} dos espacios vectoriales. Si E y \mathfrak{F} son isomorfos entonces por 2.19 un isomorfismo entre ellos transforma biyectivamente una base de uno en una base del otro por lo que tienen la misma dimensión. Recíprocamente, sean N y M bases de E y \mathfrak{F} respectivamente. Mediante los isomorfismos de coordinatización podemos pensar que $E = \mathbb{K}^{\{N\}}$ y $\mathfrak{F} = \mathbb{K}^{\{M\}}$. Si los dos tienen la misma dimensión entonces, hay una biyección $f : M \rightarrow N$. Sea $g : \mathbb{K}^{\{N\}} \ni \alpha_N \mapsto \beta_M \in \mathbb{K}^{\{M\}}$ la función tal que $\beta_i = \alpha_{f(i)}$. Podemos pensar que la función g es la que le cambia el nombre a los índices de una N -ada. Esta es claramente un isomorfismo de espacios vectoriales (ya que la suma de N -adas es por coordenadas y lo mismo con el producto por un escalar). ■

Esta proposición nos permite saber como son TODOS los espacios vectoriales. Ya vimos (mediante la base canónica) que el espacio vectorial de todas las N -adas finitas tiene dimensión $|N|$. Escogiendo el conjunto N adecuadamente obtenemos todas las dimensiones posibles. En el caso de que N sea finito con n elementos , este espacio es \mathbb{K}^n por lo que es válido el siguiente teorema.

Teorema de Clasificación de Espacios Vectoriales

2.21

*Todo espacio vectorial es isomorfo a un espacio de N -adas finitas.
Todo espacio vectorial de dimensión finita es isomorfo a \mathbb{K}^n .*

Campos de Galois

Una aplicación sencilla del Teorema de Clasificación de Espacios Vectoriales (2.21) es la siguiente.

2.22

El número de elementos en un campo finito es potencia de un número primo.

Prueba. Sea \mathbb{K} un campo. Ya vimos que siempre que se tenga un subcampo \mathbb{L} de \mathbb{K} entonces, \mathbb{K} es un espacio vectorial sobre \mathbb{L} . Esto es en esencia porque podemos sumar vectores (los elementos de \mathbb{K}) y podemos multiplicar escalares (los elementos de \mathbb{L}) por vectores.

Si \mathbb{K} es finito entonces su subcampo primo no puede ser \mathbb{Q} . Esto quiere decir que \mathbb{K} contiene como subcampo a \mathbb{Z}_p . La dimensión de \mathbb{K} sobre \mathbb{Z}_p tiene que ser finita ya que si no, entonces \mathbb{K} sería infinito. Luego existe un natural n tal que el espacio vectorial \mathbb{K} es isomorfo a \mathbb{Z}_p^n que tiene exactamente p^n elementos. ■

Como pensar en espacios vectoriales

A partir de ahora el lector debe siempre tener en mente el Teorema de Clasificación de Espacios Vectoriales (2.21). Al hablar de un espacio vectorial en primer lugar, se debe pensar en \mathbb{R}^2 y \mathbb{R}^3 . La interpretación geométrica de estos dos espacios como los segmentos dirigidos con origen en el cero da una intuición saludable acerca de lo que es cierto y lo que no.

En segundo lugar el lector debe pensar en el ejemplo \mathbb{R}^n . Si el lector lo prefiere, el número n puede ser un número fijo suficientemente grande digamos $n = 11$. Ya en este caso, para entender es necesario usar varios métodos: las analogías geométricas en dimensiones pequeñas, el cálculo algebraico con símbolos y los razonamientos lógicos. Casi todo lo que se puede demostrar para espacios vectoriales finito dimensionales se demuestra en el caso particular de \mathbb{R}^n con la misma complejidad. Las demostraciones de muchos hechos válidos en \mathbb{R}^n se copian tal cual para espacios vectoriales de dimensión finita sobre cualquier campo.

En tercer lugar se debe pensar en \mathbb{C}^n . El espacio vectorial \mathbb{C}^n es extremadamente importante dentro de las matemáticas y la física. Además, el hecho de que \mathbb{C} es algebraicamente cerrado hace que para \mathbb{C}^n algunos problemas sean más fáciles que en \mathbb{R}^n . Hay una manera de pensar en \mathbb{C}^n que ayuda un poco para tener intuición geométrica. Como ya vimos $\{1, i\}$ es una base de \mathbb{C} como espacio vectorial sobre \mathbb{R} . De la misma manera \mathbb{C}^n es un espacio vectorial sobre \mathbb{R} de dimensión $2n$ o sea, hay una biyección natural entre \mathbb{C}^n y \mathbb{R}^{2n} (la biyección es $(a_1 + b_1i, a_2 + b_2i, \dots, a_n + b_ni) \mapsto (a_1, b_1, a_2, b_2, \dots, a_n, b_n)$). Sin embargo esta biyección no es un isomorfismo. Si E es un subespacio de \mathbb{C}^n sobre \mathbb{R} entonces no necesariamente E es un subespacio de \mathbb{C}^n sobre \mathbb{C} . Desde este punto de vista podemos pensar (no rigurosamente) a \mathbb{C}^n como un \mathbb{R}^{2n} en el que hay menos subespacios.

Los que se interesan en las ciencias de la computación deben también pensar en \mathbb{Z}_p^n y en general en cualquier \mathbb{K}^n donde \mathbb{K} es un campo finito. Las aplicaciones más relevantes incluyen la codificación de información con recuperación de errores y la criptografía, que es la ciencia de cifrar mensajes.

Ahora, si el lector no le tiene miedo al concepto de campo (que es uno de los objetivos de este libro) entonces, lo más fácil es pensar en \mathbb{K}^n . Esto tiene una gran ventaja en el sentido de que no hay que pensar en los detalles particulares que se cumplen en uno u otro campo.

Sin embargo, en el caso infinito-dimensional la situación es más fea. Nuestros teoremas afirman que hay un isomorfismo entre \mathbb{R} como espacio vectorial sobre \mathbb{Q} y $\mathbb{R}^{(\mathbb{Q})}$.

En otras palabras, existe un conjunto de números reales (la base) tal que cualquier número real se puede expresar de forma *única* como combinación lineal *finita* de este conjunto usando coeficientes racionales.

El problema es que nadie conoce (ni conocerá nunca) una base de este espacio, así que realmente, estos teoremas no dicen mucho para espacios vectoriales de dimensión mayor que el cardinal de los naturales \aleph_0 .

Ejercicio 46 Sean $x, y \in \mathbb{R}$ y $\mathcal{E} = \{ax + by \mid a, b \in \mathbb{Q}\}$. ¿Es \mathcal{E} un espacio vectorial sobre \mathbb{Q} ? ¿Cuál es su dimensión? ¿Es \mathcal{E} un espacio vectorial sobre \mathbb{R} ? [189]

2.6 Suma de subespacios

En esta sección introduciremos las operaciones más básicas entre subespacios. Pero antes, es saludable dar una interpretación geométrica de los subespacios para que el lector pueda comparar el caso general con el caso de los espacios \mathbb{R}^2 y \mathbb{R}^3 .

Subespacios de \mathbb{R}^n

¿Cuáles son todos los subespacios de \mathbb{R}^n ? Del Teorema de Clasificación de Espacios Vectoriales (2.21) sabemos que estos tienen que ser isomorfos a \mathbb{R}^i para $i \in \{0, 1, \dots, n\}$ según sea su dimensión. Si $i = 0$ entonces todo el subespacio es el vector $\mathbf{0}$ (el origen de coordenadas). Si $i = n$ entonces, el subespacio es por 2.17 todo el espacio. Luego, los casos interesantes son los que $0 < i < n$.

Si $i = 1$ entonces, el subespacio es isomorfo a \mathbb{R} y tiene que tener una base de un vector. O sea, existe un vector \mathbf{a} tal que el subespacio es igual a $\langle \mathbf{a} \rangle$. Ya vimos que $\langle \mathbf{a} \rangle$ es la recta que pasa por el origen y por el punto \mathbf{a} que es el final del vector \mathbf{a} . Luego, los subespacios de dimensión 1 de \mathbb{R}^n son las rectas que pasan por el origen.

Si $i = 2$ entonces, el subespacio tiene que ser isomorfo a \mathbb{R}^2 y tiene que tener una base $\{\mathbf{a}, \mathbf{b}\}$ de dos vectores. La base $\{\mathbf{a}, \mathbf{b}\}$ tiene que ser LI y en este caso eso quiere decir que \mathbf{b} no es un múltiplo escalar de \mathbf{a} . En otras palabras, los puntos finales de

los vectores \mathbf{a} , \mathbf{b} y el origen de coordenadas no están alineados. En este caso hay un único plano (\mathbb{R}^2) que pasa por estos tres puntos y también ya vimos que este plano es $\langle \mathbf{a}, \mathbf{b} \rangle$. Luego, los subespacios de dimensión 2 de \mathbb{R}^n son los planos por el origen. Para \mathbb{R}^3 este análisis termina con todas las posibilidades: sus subespacios son: el origen, las rectas por el origen, los planos por el origen y todo el espacio.

Ahora tenemos que pensar por lo menos en los subespacios de \mathbb{R}^4 y ya se nos acaba la intuición y la terminología geométrica. Por esto, pasemos al caso general. Un subespacio de dimensión i en \mathbb{R}^n esta generado por una base $\{\mathbf{a}_1, \dots, \mathbf{a}_i\}$ de i vectores LI. Que sean LI lo que quiere decir es que sus puntos y el origen no están contenidos en un subespacio de dimensión menor que i . Si este es el caso entonces hay un único \mathbb{R}^i que contiene a todos estos puntos y este es el subespacio $\langle \mathbf{a}_1, \dots, \mathbf{a}_i \rangle$.

Suma de conjuntos y subespacios

Sean \mathfrak{E} y \mathfrak{F} dos subespacios sobre un campo \mathbb{K} . En la proposición 2.3 vimos que $\mathfrak{E} \cap \mathfrak{F}$ es un subespacio. Por definición de intersección $\mathfrak{E} \cap \mathfrak{F}$ es el subespacio más grande contenido en \mathfrak{E} y en \mathfrak{F} . Ya observamos además que $\mathfrak{E} \cup \mathfrak{F}$ no es un subespacio. Sin embargo, hay un subespacio que es el más pequeño que contiene a los dos y este es $\langle \mathfrak{E} \cup \mathfrak{F} \rangle$. El subespacio $\langle \mathfrak{E} \cup \mathfrak{F} \rangle$ tiene una estructura muy simple:

$$\langle \mathfrak{E} \cup \mathfrak{F} \rangle = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathfrak{E}, \mathbf{b} \in \mathfrak{F}\}$$

Prueba. Todo vector $\mathbf{a} + \mathbf{b}$ es una combinación lineal de $\mathfrak{E} \cup \mathfrak{F}$. Recíprocamente, toda combinación lineal de $\mathfrak{E} \cup \mathfrak{F}$ se puede escribir como en el recuadro. Como \mathfrak{E} y \mathfrak{F} son subespacios entonces, el primer sumando está en \mathfrak{E} y el segundo sumando está en \mathfrak{F} . Esto quiere decir que todo elemento de $\langle \mathfrak{E} \cup \mathfrak{F} \rangle$ es de la forma $\mathbf{a} + \mathbf{b}$ con \mathbf{a} en \mathfrak{E} y \mathbf{b} en \mathfrak{F} . ■

$$\sum_{x \in \mathfrak{E}} \alpha_x x + \sum_{y \in \mathfrak{F}} \beta_y y$$

2.23 $A + B = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in A, \mathbf{b} \in B\}$ Dados dos conjuntos cualesquiera de vectores A y B definimos la **suma** de estos como en el recuadro a la izquierda. Después de esta definición el resultado 2.23 se puede reformular como $\langle \mathfrak{E} \cup \mathfrak{F} \rangle = \mathfrak{E} + \mathfrak{F}$. Es por esto que al subespacio $\langle \mathfrak{E} \cup \mathfrak{F} \rangle$ se le llama la **suma de los subespacios** \mathfrak{E} y \mathfrak{F} y desde ahora se le denominará por $\mathfrak{E} + \mathfrak{F}$.

La igualdad modular

Sean \mathfrak{E} y \mathfrak{F} dos subespacios. Ahora trataremos de calcular la dimensión de $\mathfrak{E} + \mathfrak{F}$. Para esto necesitaremos primero encontrar bases de $\mathfrak{E} \cap \mathfrak{F}$ y $\mathfrak{E} + \mathfrak{F}$.

2.24 Existen E una base de \mathfrak{E} y F una base de \mathfrak{F} tales que $E \cap F$ es base de $\mathfrak{E} \cap \mathfrak{F}$ y $E \cup F$ es base de $\mathfrak{E} + \mathfrak{F}$.

Prueba. Sea N una base de $\mathfrak{E} \cap \mathfrak{F}$. Como N es LI y está contenida en \mathfrak{E} entonces,

por el Teorema de Existencia de Bases (2.14) existe una base E de \mathfrak{E} que contiene a N . Análogamente, existe una base F de \mathfrak{F} que contiene a N .

Demostremos primero que $E \cap F = N$. Efectivamente, por la forma en que construimos E y F tenemos $N \subseteq E \cap F$. Para la prueba de la otra inclusión sea $a \in E \cap F$. Como $N \cup a \subseteq E$ entonces, tenemos que $N \cup a$ es LI. Como N es base de $\mathfrak{E} \cap \mathfrak{F}$ y las bases son los conjuntos LI más grandes, obtenemos que $a \in N$. Luego, $E \cap F \subseteq N$.

Solo nos queda probar que $E \cup F$ es una base de $\mathfrak{E} + \mathfrak{F}$. En primer lugar, cualquier $a + b \in \mathfrak{E} + \mathfrak{F}$ es combinación lineal de $E \cup F$ por lo que $E \cup F$ es generador de $\mathfrak{E} + \mathfrak{F}$. Necesitamos probar que $E \cup F$ es LI. Para esto supongamos que

$$\mathbf{0} = \sum_{i \in E \cup F} \alpha_i i = \sum_{i \in E \setminus N} \alpha_i i + \sum_{i \in N} \alpha_i i + \sum_{i \in F \setminus N} \alpha_i i$$

y demostremos que todos los α_i son cero. Sean x, y, z el primer, segundo y tercer sumando respectivamente en la igualdad anterior. Por construcción, tenemos $x \in \mathfrak{E}$, $y \in \mathfrak{E} \cap \mathfrak{F}$ y $z \in \mathfrak{F}$. Además, como $z = -(y + x)$ y \mathfrak{E} es un subespacio, obtenemos $z \in \mathfrak{E} \cap \mathfrak{F}$. Como N es una base de $\mathfrak{E} \cap \mathfrak{F}$ el vector z es combinación lineal de N , o sea $z = \sum_{i \in N} \lambda_i i$ para ciertos coeficientes λ_i . De aquí obtenemos que

$$\mathbf{0} = x + y + z = \sum_{i \in E \setminus N} \alpha_i i + \sum_{i \in N} (\alpha_i + \lambda_i) i$$

Como E es LI, todos los coeficientes de esta combinación lineal son cero. En particular, $\alpha_i = 0$ para todo $i \in E \setminus N$ y por lo tanto $x = \mathbf{0}$. Substituyendo x obtenemos

$$\mathbf{0} = y + z = \sum_{i \in N} \alpha_i i + \sum_{i \in F \setminus N} \alpha_i i$$

y como F es LI deducimos que los restantes coeficientes también son cero. ■

Igualdad modular

2.25 Si \mathfrak{E} y \mathfrak{F} son dos subespacios entonces,
 $\dim \mathfrak{E} + \dim \mathfrak{F} = \dim (\mathfrak{E} + \mathfrak{F}) + \dim (\mathfrak{E} \cap \mathfrak{F})$.

Prueba. Por 2.24 existen bases E y F de \mathfrak{E} y \mathfrak{F} tales que $E \cup F$ es base de $\mathfrak{E} + \mathfrak{F}$ y $E \cap F$ es base de $\mathfrak{E} \cap \mathfrak{F}$. Luego, la fórmula se reduce a la conocida igualdad entre cardinales de conjuntos $|E| + |F| = |E \cup F| + |E \cap F|$. ■

Ejercicio 47 Construya ejemplos de subespacios reales \mathfrak{E} y \mathfrak{F} tales que ellos y $\mathfrak{E} + \mathfrak{F}$ tienen las dimensiones definidas en la tabla del recuadro a la derecha. ¿Puede tener $\mathfrak{E} + \mathfrak{F}$ dimensión diferente a las de la tabla?

\mathfrak{E}	\mathfrak{F}	$(\mathfrak{E} + \mathfrak{F})$	\mathfrak{E}	\mathfrak{F}	$(\mathfrak{E} + \mathfrak{F})$
1	1	1	1	2	3
1	1	2	2	2	3
1	2	2	2	2	4

Suma directa

Para dos espacios vectoriales \mathfrak{E} y \mathfrak{F} (no necesariamente contenidos en otro espacio) sobre un mismo campo \mathbb{K} definiremos la **suma directa** de ellos como el espacio vectorial

$$\begin{aligned}\mathfrak{E} \oplus \mathfrak{F} &= \{(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \in \mathfrak{E}, \mathbf{b} \in \mathfrak{F}\} \\ (\mathbf{a}, \mathbf{b}) + (\mathbf{a}', \mathbf{b}') &= (\mathbf{a} + \mathbf{a}', \mathbf{b} + \mathbf{b}') \\ \lambda(\mathbf{a}, \mathbf{b}) &= (\lambda\mathbf{a}, \lambda\mathbf{b})\end{aligned}$$

Observese que como conjunto la suma directa es el producto cartesiano de conjuntos. La diferencia está en que la suma directa ya trae en su definición las operaciones de suma de vectores y multiplicación por un escalar. Deberíamos probar que la suma directa es efectivamente un espacio vectorial, o sea que cumplen todos los axiomas. Nosotros omitiremos esta prueba por ser trivial y aburrida. Mejor nos concentraremos en algo más interesante.

2.26 $\dim(\mathfrak{E} \oplus \mathfrak{F}) = \dim \mathfrak{E} + \dim \mathfrak{F}$.

Prueba. Sea $\mathfrak{E}' = \{(\mathbf{a}, \mathbf{0}) \mid \mathbf{a} \in \mathfrak{E}\}$ y $\mathfrak{F}' = \{(\mathbf{0}, \mathbf{b}) \mid \mathbf{b} \in \mathfrak{F}\}$. Es claro que los espacios \mathfrak{E} y \mathfrak{E}' son isomorfos y por lo tanto tienen la misma dimensión. Otro tanto ocurre con \mathfrak{F} y \mathfrak{F}' . Por definición de suma de subespacios tenemos que $\mathfrak{E}' + \mathfrak{F}' = \mathfrak{E} \oplus \mathfrak{F}$. De la Igualdad modular (2.25) obtenemos $\dim(\mathfrak{E}' + \mathfrak{F}') = \dim \mathfrak{E}' + \dim \mathfrak{F}' - \dim(\mathfrak{E}' \cap \mathfrak{F}') = \dim \mathfrak{E} + \dim \mathfrak{F}$. ■

Isomorfismo canónico entre la suma y la suma directa.

De esta última proposición y de la igualdad modular se deduce que si dos subespacios \mathfrak{E} y \mathfrak{F} son tales que $\mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$ entonces $\dim(\mathfrak{E} \oplus \mathfrak{F}) = \dim(\mathfrak{E} + \mathfrak{F})$ o sea $\mathfrak{E} \oplus \mathfrak{F}$ es isomorfo a $\mathfrak{E} + \mathfrak{F}$. A continuación probaremos solo un poquito más. Sin embargo, este poquito nos llevará a profundas reflexiones.

Isomorfismo Canónico entre la Suma y la Suma directa

2.27 Si \mathfrak{E} y \mathfrak{F} son subespacios tales que $\mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$ entonces la función $\mathfrak{E} \oplus \mathfrak{F} \ni (\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} + \mathbf{b} \in \mathfrak{E} + \mathfrak{F}$ es un isomorfismo de espacios vectoriales.

Prueba. Sea f la función definida en el enunciado. Tenemos

$$f(\lambda(\mathbf{a}, \mathbf{b})) = f(\lambda\mathbf{a}, \lambda\mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b} = \lambda(\mathbf{a} + \mathbf{b}) = \lambda f(\mathbf{a}, \mathbf{b})$$

$$f((\mathbf{a}, \mathbf{b}) + (\mathbf{x}, \mathbf{y})) = f(\mathbf{a} + \mathbf{x}, \mathbf{b} + \mathbf{y}) = \mathbf{a} + \mathbf{x} + \mathbf{b} + \mathbf{y} = f(\mathbf{a}, \mathbf{b}) + f(\mathbf{x}, \mathbf{y})$$

por lo que solo queda demostrar que f es biyectiva. Por definición de suma de subespacios f es sobreyectiva. Para probar la inyección supongamos que $f(\mathbf{a}, \mathbf{b}) = f(\mathbf{x}, \mathbf{y})$

entonces, $\mathbf{a} - \mathbf{x} = \mathbf{y} - \mathbf{b}$. Como $\mathbf{a}, \mathbf{x} \in \mathfrak{E}$ entonces $\mathbf{a} - \mathbf{x} \in \mathfrak{E}$. Como $\mathbf{b}, \mathbf{y} \in \mathfrak{F}$ entonces $\mathbf{y} - \mathbf{b} \in \mathfrak{F}$. Luego $\mathbf{a} - \mathbf{x} = \mathbf{y} - \mathbf{b} \in \mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$ por lo que $\mathbf{a} = \mathbf{x}$ y $\mathbf{b} = \mathbf{y}$. ■

Observemos que el isomorfismo $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} + \mathbf{b}$ no depende de escoger base alguna en $\mathfrak{E} \oplus \mathfrak{F}$. Todos los isomorfismos que construimos antes de esta proposición se construyeron escogiendo cierta base en alguno de los espacios. Los isomorfismos que no dependen de escoger alguna base juegan un papel importante en el álgebra lineal y se les llama **isomorfismos canónicos**. Decimos que dos espacios vectoriales son **canómicamente isomorfos** si existe un isomorfismo canónico entre ellos. Así por ejemplo todo espacio vectorial \mathfrak{E} de dimensión 3 es isomorfo a \mathbb{K}^3 pero no es canómicamente isomorfo a \mathbb{K}^3 porque no se puede construir de cualquier espacio vectorial de dimensión 3 un isomorfismo con \mathbb{K}^3 que no dependa de escoger alguna base. Cuando veamos los productos escalares veremos que hay fuertes razones para que diferenciamos las bases. Los isomorfismos no canónicos no necesariamente preservan una u otra propiedad de las bases. Por otro lado, los isomorfismos canónicos si las preservan. Si por un lado, debe haber cierta resistencia a considerar iguales a dos espacios vectoriales isomorfos por el otro, los espacios canómicamente isomorfos no se diferencian en nada uno del otro, por lo que se puede pensar que es el mismo espacio.

- Ejercicio 481.** Demuestre que $\mathfrak{E} \oplus \mathfrak{F}$ y $\mathfrak{F} \oplus \mathfrak{E}$ son canómicamente isomorfos.
2. ¿Como se debe llamar esta propiedad de la suma directa de espacios?
 3. Demuestre que la suma directa de espacios es asociativa.
 4. ¿Tiene la suma directa elemento neutro? [189]

Subespacios complementarios

Sea \mathfrak{S} un espacio vectorial y \mathfrak{E} un subespacio de \mathfrak{S} . Diremos que el subespacio \mathfrak{F} es **complementario** de \mathfrak{E} en \mathfrak{S} si $\mathfrak{S} = \mathfrak{E} \oplus \mathfrak{F}$. Esta igualdad por el Isomorfismo Canónico entre la Suma y la Suma directa (2.27) lo que quiere decir es que $\mathfrak{S} = \mathfrak{E} + \mathfrak{F}$ y $\mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$. En \mathbb{R}^2 dos rectas por el origen diferentes son complementarias una a otra. En \mathbb{R}^3 un plano y una recta no contenida en el plano (ambos por el origen) son complementarios uno a otro.

2.28 *Todo subespacio tiene complementario.*

Prueba. Sea \mathfrak{E} un subespacio de \mathfrak{S} . Sea \mathbf{A} una base de \mathfrak{E} . Por el Teorema de Existencia de Bases (2.14) hay una base \mathbf{C} de \mathfrak{S} que contiene a \mathbf{A} . Sea \mathfrak{F} el espacio generado por $\mathbf{B} = \mathbf{C} \setminus \mathbf{A}$. Como \mathbf{C} es generador de \mathfrak{S} , todo elemento en \mathfrak{S} es combinación lineal de \mathbf{C} y en particular todo elemento de \mathfrak{S} se expresa como $\mathbf{a} + \mathbf{b}$ con $\mathbf{a} \in \mathfrak{E}$ y $\mathbf{b} \in \mathfrak{F}$. Luego $\mathfrak{S} = \mathfrak{E} + \mathfrak{F}$.

Por otro lado, si $\mathbf{x} \in \mathfrak{E} \cap \mathfrak{F}$ entonces, existen combinaciones lineales tales que

$$\left(\mathbf{x} = \sum_{\mathbf{a} \in A} \alpha_a \mathbf{a} = \sum_{\mathbf{a} \in B} \alpha_a \mathbf{a} \right) \Rightarrow \left(\sum_{\mathbf{a} \in A} \alpha_a \mathbf{a} - \sum_{\mathbf{a} \in B} \alpha_a \mathbf{a} = \mathbf{0} \right).$$

Como C es LI entonces, por la caracterización 2.9.4 de los conjuntos LI la última combinación lineal tiene todos sus coeficientes iguales a cero. Luego, $\mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$. ■

2.29

Si \mathfrak{E} y \mathfrak{F} son dos subespacios complementarios entonces cada vector \mathbf{x} se expresa de forma única como $\mathbf{x} = \mathbf{a} + \mathbf{b}$ donde $\mathbf{a} \in \mathfrak{E}$ y $\mathbf{b} \in \mathfrak{F}$.

Prueba. Si \mathfrak{E} y \mathfrak{F} son complementarios entonces $\mathfrak{E} + \mathfrak{F}$ es todo el espacio y por lo tanto todo vector es de la forma $\mathbf{a} + \mathbf{b}$. Si $\mathbf{a} + \mathbf{b} = \mathbf{a}' + \mathbf{b}'$ entonces $\mathbf{a} - \mathbf{a}' = \mathbf{b}' - \mathbf{b} \in \mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$ por lo que la descomposición $\mathbf{x} = \mathbf{a} + \mathbf{b}$ es única. ■

Ejercicio 49 Demuestre el recíproco de 2.29: Si cada vector se expresa de forma única como $\mathbf{x} = \mathbf{a} + \mathbf{b}$ con $\mathbf{a} \in \mathfrak{E}$ y $\mathbf{b} \in \mathfrak{F}$ entonces, \mathfrak{E} y \mathfrak{F} son complementarios. [189]

Ejercicio 50 Pruebe que $\mathfrak{E} = \mathfrak{E}_1 \oplus \mathfrak{E}_2 \oplus \dots \oplus \mathfrak{E}_n$ si y solo si cualquier vector $\mathbf{x} \in \mathfrak{E}$ se expresa de forma única como $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_n$ con $\mathbf{x}_i \in \mathfrak{E}_i$. **Sugerencia:** Usar 2.29, el ejercicio anterior y la asociatividad de la suma directa para aplicar inducción en el número de sumandos n .

Espacios vectoriales versus conjuntos

Hemos demostrado ya unos cuantos resultados que se parecen mucho a los de Teoría de Conjuntos y siempre es saludable establecer analogías con resultados ya conocidos. Para acentuar más esta similitud hagamos un diccionario de traducción

Conjunto	↔	Espacio vectorial
Subconjunto	↔	Subespacio vectorial
Cardinal	↔	Dimensión
Intersección de subconjuntos	↔	Intersección de subespacios
Unión de subconjuntos	↔	Suma de subespacios
Unión disjunta	↔	Suma directa
Complemento	↔	Subespacio complementario
Biyecciones	↔	Isomorfismos

Si nos fijamos atentamente muchos de los resultados que hemos demostrado para espacios vectoriales tienen su contraparte válida para conjuntos usando el diccionario que hemos construido. Probablemente, el ejemplo más notable de esto son las igualdades modulares para conjuntos y para espacios vectoriales.

Sin embargo, es preciso ser cuidadosos en tratar de llevar resultados de los conjuntos a los espacios vectoriales. Por ejemplo, el complemento de un subconjunto es único y no siempre hay un único subespacio complementario. Otro ejemplo, un poco más

substancial, es que la intersección de conjuntos distribuye con la unión o sea $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. Por otro lado la intersección de subespacios en general no distribuye con la suma o sea, la igualdad $A \cap (B + C) = (A \cap B) + (A \cap C)$ no siempre es verdadera. Para ver esto tómese en \mathbb{R}^3 los subespacios $A = \langle (0, 0, 1), (1, 1, 0) \rangle$, $B = \langle (1, 0, 0) \rangle$ y $C = \langle (0, 1, 0) \rangle$. Calculamos $A \cap (B + C) = \langle (1, 1, 0) \rangle$ y $(A \cap B) + (A \cap C) = \{(0, 0, 0)\}$ y vemos que son distintos.

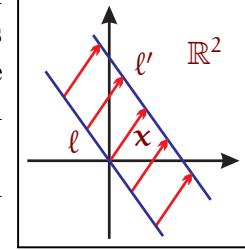
Ejercicio 51 Si A y B son dos conjuntos entonces la diferencia simétrica de los mismos es el conjunto $A +_2 B = (A \cup B) \setminus (A \cap B)$. Demuestre que todos los subconjuntos finitos de un conjunto cualquiera U forman un espacio vectorial de dimensión $|U|$ sobre el campo \mathbb{Z}_2 para la suma $_2+$ y el producto $1A = A$, $0A = \emptyset$. Vea, que los conceptos de nuestro diccionario de traducción se aplican a este caso en forma directa.

2.7 Espacios cocientes

Ya vimos que para un subespacio \mathfrak{E} del espacio \mathfrak{F} , siempre existen subespacios complementarios a él. Sin embargo hay muchos subespacios complementarios y no hay una manera canónica de escoger alguno de ellos. En esta sección nos dedicaremos a construir canónicamente el espacio cociente $\mathfrak{F}/\mathfrak{E}$ que es el espacio de todos los subespacios afines paralelos a \mathfrak{E} y que es isomorfo a cualquier subespacio complementario de \mathfrak{E} .

Subespacios afines

Ya vimos que en el plano cartesiano \mathbb{R}^2 los subespacios son el origen $\{0\}$, todo \mathbb{R}^2 y las rectas por el origen. Sin embargo, hay otras rectas en el plano que no pasan por el origen. Para obtener una de estas rectas lo que tenemos que hacer es trasladar una recta por el origen ℓ mediante un vector x (véase la figura). De esta manera, obtenemos la recta ℓ' que se obtiene sumandole x a cada vector en la recta ℓ . Observese que si $x \neq 0$ entonces ℓ y ℓ' no se intersectan.



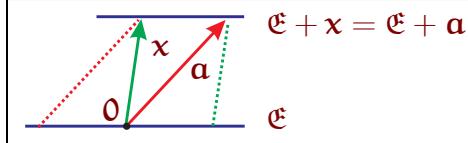
Esto nos motiva a la siguiente definición. Sea A un conjunto de vectores en un espacio vectorial (¡sobre cualquier campo!) y x un vector. Al conjunto $A+x = \{a+x \mid a \in A\}$ se le llama la **traslación** de A con el vector x . Observese que la operación de trasladar es un caso particular (cuando uno de los sumandos contiene a un solo vector) de la operación de suma de conjuntos de vectores introducida en la sección anterior.

Un **subespacio afín** es el trasladado de un subespacio. En otras palabras, un conjunto de vectores E es un subespacio afín si existen un subespacio \mathfrak{E} y un vector x tales que $E = \mathfrak{E} + x$. Observese que los subespacios son también subespacios afines. Basta trasladar con el vector cero. Esto causa una confusión lingüística: el adjetivo “afín” se usa para hacer el concepto de “subespacio” más general, no más específico.

Por esto, cuando se habla de subespacios afines es común referirse a los subespacios como **subespacios vectoriales** o como **subespacios por el origen**.

Las traslaciones de un subespacio cumplen una propiedad muy sencilla pero fundamental: que es posible trasladar el subespacio vectorial con diferentes vectores y obtener el mismo subespacio afín. Más precisamente:

2.30 Si $\mathbf{a} \in \mathfrak{E} + \mathbf{x}$ entonces, $\mathfrak{E} + \mathbf{x} = \mathfrak{E} + \mathbf{a}$.



Prueba. Probemos primero el caso que $\mathbf{x} = \mathbf{0}$. Como $\mathbf{a} \in \mathfrak{E}$ y \mathfrak{E} es un subespacio, tenemos $\mathbf{y} \in \mathfrak{E} \Leftrightarrow \mathbf{y} - \mathbf{a} \in \mathfrak{E} \Leftrightarrow \mathbf{y} \in \mathfrak{E} + \mathbf{a}$. Ahora por el caso general. Sabemos que, $\mathbf{z} = \mathbf{a} - \mathbf{x} \in \mathfrak{E}$ y del primer caso, $\mathfrak{E} = \mathfrak{E} + \mathbf{z}$. Luego, $\mathfrak{E} + \mathbf{x} = \mathfrak{E} + \mathbf{z} + \mathbf{x} = \mathfrak{E} + \mathbf{a}$. ■

Ejercicio 52 Pruebe el recíproco de 2.30: Si $\mathfrak{E} + \mathbf{x} = \mathfrak{E} + \mathbf{a}$ entonces $\mathbf{a} \in \mathfrak{E} + \mathbf{x}$.

Ahora observemos, que un trasladado de un subespacio afín es a su vez un subespacio afín ya que $(\mathfrak{E} + \mathbf{x}) + \mathbf{y} = \mathfrak{E} + (\mathbf{x} + \mathbf{y})$. Dos subespacios afines se le llaman **paralelos** si uno es un trasladado del otro. La relación de paralelismo entre subespacios afines es de equivalencia ya que si $\mathfrak{E} = \mathfrak{F} + \mathbf{x}$ y $\mathfrak{G} = \mathfrak{E} + \mathbf{y}$ entonces, $\mathfrak{G} = \mathfrak{F} + (\mathbf{x} + \mathbf{y})$. Esto significa que el conjunto de los subespacios afines se parte en **clases de paralelismo** y dos subespacios son paralelos si y solo si están en la misma clase de paralelismo.

Ahora veremos que en cada clase de paralelismo hay un solo subespacio afín que pasa por el origen.

2.31 Todo subespacio afín es paralelo a un solo subespacio vectorial.

Prueba. Si $\mathfrak{E} + \mathbf{x} = \mathfrak{F} + \mathbf{y}$ entonces, $\mathfrak{E} = \mathfrak{F} + \mathbf{y} - \mathbf{x}$. Como $\mathfrak{F} + \mathbf{y} - \mathbf{x}$ contiene al cero entonces, $\mathbf{x} - \mathbf{y} \in \mathfrak{F}$. Como \mathfrak{F} es un subespacio, $\mathbf{y} - \mathbf{x} \in \mathfrak{F}$ y 2.30 nos da $\mathfrak{F} = \mathfrak{F} + \mathbf{y} - \mathbf{x}$. ■

Este resultado nos permite definir la dimensión de un subespacio afín. Cada uno de ellos es paralelo a un único subespacio y su **dimensión** es la dimensión de ese subespacio. En otras palabras, si $\mathfrak{E} = \mathfrak{E} + \mathbf{x}$ entonces $\dim \mathfrak{E} = \dim \mathfrak{E}$. Es común utilizar la terminología geométrica para hablar de los subespacios afines de dimensión pequeña. Así, un **punto** es un subespacio afín de dimensión cero, una **recta** es un subespacio afín de dimensión uno, un **plano** es un subespacio afín de dimensión dos.

2.32 Dos subespacios afines paralelos o son el mismo o no se intersectan.

Prueba. Si $\mathbf{y} \in (\mathfrak{E} + \mathbf{x}) \cap (\mathfrak{E} + \mathbf{x}')$ entonces, por 2.30, $\mathfrak{E} + \mathbf{x} = \mathfrak{E} + \mathbf{y} = \mathfrak{E} + \mathbf{x}'$. ■

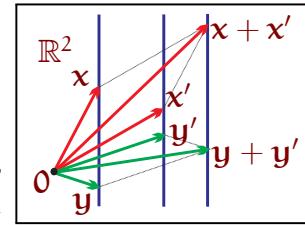


Es un error común (debido al caso de rectas \mathbb{R}^2) pensar que es equivalente que dos subespacios afines sean paralelos a que estos no se intersecten. Para convencerse de que esto no es cierto, el lector debe pensar en dos rectas no coplanares en \mathbb{R}^3 .

El espacio cociente

Sea \mathfrak{D} un espacio vectorial y \mathfrak{E} un subespacio vectorial de \mathfrak{D} .

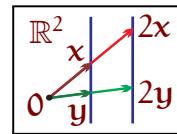
Si $E = \mathfrak{E} + x$ y $F = \mathfrak{E} + y$ son dos subespacios afines paralelos entonces $E + F = (\mathfrak{E} + x) + (\mathfrak{E} + y) = (\mathfrak{E} + \mathfrak{E}) + (x + y) = \mathfrak{E} + (x + y)$ lo que muestra que $E + F$ está en la misma clase de paralelismo que E y F . En otras palabras (véase la figura a la derecha) la suma de cualquier vector en E con cualquier otro en F está en un mismo espacio paralelo a E y F .



Denotemos por $\mathfrak{D}/\mathfrak{E}$ al conjunto de todos los subespacios afines de \mathfrak{D} paralelos a \mathfrak{E} . La observación anterior nos dice que la **suma de subespacios afines** es una operación en $\mathfrak{D}/\mathfrak{E}$. Esta suma es asociativa y conmutativa debido a que la suma de vectores cumple estas propiedades. El subespacio \mathfrak{E} es el neutro para esta operación y el opuesto de $\mathfrak{E} + x$ es $\mathfrak{E} - x$. Luego, $\mathfrak{D}/\mathfrak{E}$ es un grupo abeliano para la suma y para convertirlo en un espacio vectorial solo nos falta el producto por escalares..

$\lambda A = \{\lambda a \mid a \in A\}$ Sea A un conjunto arbitrario de vectores y λ un escalar. Definiremos al conjunto λA como en el recuadro a la izquierda.

Sean $E = \mathfrak{E} + x$ un subespacio afín paralelo a \mathfrak{E} y λ un escalar. Tenemos $\lambda E = \lambda(\mathfrak{E} + x) = \lambda\mathfrak{E} + \lambda x = \mathfrak{E} + \lambda x$ lo que muestra que λE está en la misma clase de paralelismo que E . En otras palabras (véase la figura a la derecha) el producto de un escalar por todos los vectores en E resulta en espacio paralelo a E . Este producto convierte a $\mathfrak{D}/\mathfrak{E}$ en un espacio vectorial llamado **espacio cociente** de \mathfrak{D} por \mathfrak{E} .



Ejercicio 53 Pruebe los axiomas de espacio vectorial para el espacio cociente.

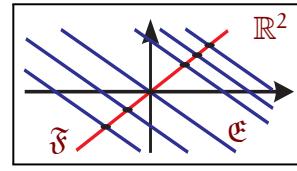
Ejercicio 54 ¿Cuál es el espacio cociente de \mathbb{R}^3 por el plano xy ?

Ejercicio 55 ¿Qué pasa si sumamos dos espacios afines no paralelos? [190]

El isomorfismo con los complementarios

2.33

Sea \mathfrak{F} un subespacio complementario a \mathfrak{E} . Entonces, cualquier subespacio afín paralelo a \mathfrak{E} intersecta a \mathfrak{F} en un y solo un vector.



Prueba. Sea $E = \mathfrak{E} + x$ cualquier subespacio afín paralelo a \mathfrak{E} . Como $\mathfrak{D} = \mathfrak{E} \oplus \mathfrak{F}$

existen unos únicos vectores $\mathbf{y} \in \mathfrak{E}$ y $\mathbf{z} \in \mathfrak{F}$ tales que $\mathbf{x} = \mathbf{y} + \mathbf{z}$. Luego por 2.30 $\mathfrak{E} + \mathbf{x} = \mathfrak{E} + \mathbf{y} + \mathbf{z} = \mathfrak{E} + \mathbf{z}$ y por lo tanto $\mathbf{z} \in \mathfrak{E} + \mathbf{x}$ o sea, $\mathbf{z} \in \mathfrak{E} \cap \mathfrak{F}$. Si \mathbf{z}' es otro vector en $\mathfrak{E} \cap \mathfrak{F}$ entonces, existe $\mathbf{a} \in \mathfrak{E}$ tal que $\mathbf{z}' = \mathbf{a} + \mathbf{x}$. De aquí $\mathbf{x} = -\mathbf{a} + \mathbf{z}'$ y por la unicidad de la descomposición de un vector en suma de vectores de espacios complementarios, $\mathbf{y} = -\mathbf{a}$ y $\mathbf{z} = \mathbf{z}'$. ■

2.34

Sea \mathfrak{F} un subespacio complementario a \mathfrak{E} . Entonces,
 $f: \mathfrak{F} \ni \mathbf{x} \mapsto (\mathfrak{E} + \mathbf{x}) \in \mathfrak{D}/\mathfrak{E}$
es un isomorfismo de espacios vectoriales.

Prueba. Por 2.33 la aplicación $f: \mathfrak{F} \ni \mathbf{x} \mapsto (\mathfrak{E} + \mathbf{x}) \in \mathfrak{D}/\mathfrak{E}$ tiene inversa (que a cada $\mathfrak{E} + \mathbf{x} \in \mathfrak{D}/\mathfrak{E}$ le hace corresponder el único vector en $(\mathfrak{E} + \mathbf{x}) \cap \mathfrak{F}$). Luego, f es biyectiva. Además,

$$\begin{aligned} f(\mathbf{x} + \mathbf{y}) &= \mathfrak{E} + (\mathbf{x} + \mathbf{y}) = (\mathfrak{E} + \mathbf{x}) + (\mathfrak{E} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \\ f(\lambda \mathbf{x}) &= \mathfrak{E} + (\lambda \mathbf{x}) = \lambda(\mathfrak{E} + \mathbf{x}) = \lambda f(\mathbf{x}) \end{aligned}$$

por lo que f es un isomorfismo. ■

Este isomorfismo es canónico. Luego, cualquier subespacio complementario a \mathfrak{E} es canómicamente isomorfo a $\mathfrak{D}/\mathfrak{E}$. Esta proposición también nos dice que la dimensión de $\mathfrak{D}/\mathfrak{E}$ es la misma que la de un complementario a \mathfrak{E} o sea, es $\dim \mathfrak{D} - \dim \mathfrak{E}$.

 Es posible (gracias al isomorfismo con los complementarios) desarrollar toda el álgebra lineal sin la introducción de los espacios cocientes. Si embargo, es más cómodo hacerlo con ellos. Además es importante que el lector se familiarize con el concepto, ya que, en el estudio de estructuras algebraicas más generales, no siempre existen estructuras “complementarias” (por ejemplo en la teoría de grupos). Por otro lado, las estructuras cocientes si se pueden construir. Por ejemplo, todo grupo tiene un cociente por cualquier subgrupo normal.

2.8 El espacio afín



Recordemos de la sección anterior que un **subespacio afín** del espacio vectorial \mathfrak{F} es el trasladado $\mathfrak{E} + \mathbf{x}$ de un subespacio vectorial \mathfrak{E} de \mathfrak{F} . La **dimensión** de $\mathfrak{E} + \mathbf{x}$ es por definición la dimensión de \mathfrak{E} .

Los subespacios afines de dimensión pequeña reciben nombres especiales según la tradición geométrica. Los de dimensión 0 se le llaman **puntos**. Los de dimensión 1 se le llaman **rectas** y los de dimensión 2 se le llaman **planos**.

Un punto es, en otras palabras, un subespacio afín $\mathfrak{E} + \mathbf{x}$ donde \mathfrak{E} es de dimensión cero. Pero el único subespacio vectorial de dimensión cero es $\{\mathbf{0}\}$. Luego, un punto es un conjunto de la forma $\{\mathbf{0}\} + \mathbf{x} = \{\mathbf{x}\}$. Esto nos da una biyección obvia $\{\mathbf{x}\} \mapsto \mathbf{x}$ entre los puntos del espacio afín y los vectores del espacio vectorial.

Al conjunto de todos los puntos del espacio vectorial \mathfrak{F} se le llama **el espacio afín** de \mathfrak{F} . Pero como, dirá el lector, si la diferencia entre $\{\mathbf{x}\}$ y \mathbf{x} es puramente formal entonces, ¿cuál es la diferencia entre el espacio afín y el espacio vectorial? La respuesta

es ambivalente pues es la misma pregunta que ¿que diferencia hay entre geometría y álgebra? Antes de Descartes eran dos cosas completamente distintas, después de él, son cosas muy parecidas.

Intuitivamente, un espacio afín es lo mismo que el espacio vectorial pero sin origen. El espacio afín de \mathbb{R}^2 es el plano de Euclides en el que estudiamos la geometría más clásica y elemental: congruencia de triángulos, teorema de Pitágoras, etc. Los resultados de esta geometría no dependen de cual es el origen de coordenadas. Dos triángulos son congruentes o no independientemente de en qué lugar del plano se encuentran.

A un nivel más alto, podemos pensar el espacio afín de un espacio vectorial como una estructura matemática adecuada para poder estudiar las propiedades de los espacios vectoriales que son invariantes por traslaciones, o sea, que no cambian al mover un objeto de un lugar a otro en el espacio.

La regla del paralelogramo

La conocida regla del paralelogramo para la suma de vectores en el plano \mathbb{R}^2 se generaliza a todas las dimensiones y sobre cualquier campo.

Postularemos que el conjunto vacío es un subespacio afín de dimensión -1 . Eso nos evitará muchos problemas en la formulación de nuestros resultados.

2.35 *La intersección de subespacios afines es un subespacio afín.*

Prueba. Sea x un punto en la intersección. Cada uno de los subespacios afines, es $E + x$ para cierto subespacio E del espacio vectorial. Si F es la intersección de todos estos subespacios entonces $F + x$ es la intersección de los subespacios afines. ■

Regla del paralelogramo

2.36 *Si a y b son vectores LI de un espacio vectorial entonces, $a + b$ es la intersección de los subespacios afines $\langle a \rangle + b$ y $\langle b \rangle + a$.*

Prueba. Obviamente $a + b \in (\langle a \rangle + b) \cap (\langle b \rangle + a)$. Por otro lado, ambos $\langle a \rangle + b$ y $\langle b \rangle + a$ tienen dimensión 1 y por lo tanto su intersección tiene dimensión 0 o 1. Si esta dimensión es cero entonces terminamos. Si esta dimensión es uno entonces $\langle a \rangle + b = \langle b \rangle + a$ y por lo tanto $a \in \langle a \rangle + b$. Por 2.30 tenemos que $\langle a \rangle + b = \langle a \rangle + a = \langle a \rangle$, Por lo tanto $b \in \langle a \rangle$ y esto contradice que $\{a, b\}$ es LI. ■

Cerradura afín

Para un conjunto A de puntos en el espacio afín, la **cerradura afín** de A es la intersección de todos los subespacios afines que contienen a A . A la cerradura afín de A la denotaremos por $[A]$. Esto la diferencia claramente de la cerradura lineal $\langle A \rangle$.

Para la cerradura afín, tenemos las mismas propiedades que para la cerradura lineal.

2.37

[A] es el subespacio afín más pequeño que contiene a **A**.

Prueba. **[A]** es un subespacio afín. Si **B** es un subespacio afín que contiene a **A** entonces $[A] \subseteq B$ pues para hallar **[A]** tuvimos que intersectar con **B**. ■

2.38

La cerradura afín cumple las siguientes propiedades:

1. $A \subseteq [A]$ (incremento)
2. $A \subseteq B \Rightarrow [A] \subseteq [B]$ (monotonía)
3. $[[A]] = [A]$ (idempotencia).

Prueba. Es exactamente la misma que para el caso de la cerradura lineal. ■

Generadores e independencia

Un conjunto de puntos **A** es **generador afín** si **[A]** es todo el espacio afín. Los conjuntos **afinmente independientes** los definiremos con el siguiente resultado.

Definición de conjuntos afinmente independientes

2.39

Sea **A** un conjunto de puntos. Las siguientes afirmaciones son equivalentes

1. Cualquier subconjunto propio de **A** genera un subespacio afín más pequeño que todo **A**.
2. Cualquier punto en **A** no está en la cerradura afín de los restantes.

Prueba. Es análoga a la prueba $(1 \Leftrightarrow 2)$ de la caracterización de conjuntos LI. ■

Los conjuntos de puntos que no son afinmente independientes los llamaremos **afinmente dependientes**.



Nuevamente para evitarnos largas oraciones, a los conjuntos afinmente independientes los llamaremos conjuntos AI y a los conjuntos afinmente dependientes los llamaremos conjuntos AD.

2.40

Todo sobreconjunto de un generador afín es un generador afín.

Todo subconjunto de un conjunto AI es AI.

Prueba. Sea **A** un generador afín y $B \supseteq A$. Tenemos $[A] \subseteq [B]$ y como **[A]** es todo el espacio afín entonces **[B]** también lo es.

Sea **A** que es AI y $B \subseteq A$. Si **B** fuera AD entonces existiría $\mathbf{b} \in B$ tal que $\mathbf{b} \in [B \setminus \mathbf{b}] \subseteq [A \setminus \mathbf{b}]$ y esto no puede ser pues **A** es AI. ■

Bases afines

Una **base afín** es un conjunto de puntos que es AI y generador afín. Ahora podríamos seguir por el mismo camino demostrando que las bases afines son los generadores afines más pequeños o los AI más grandes. Despues, el teorema de existencia de base, el lema del cambio para las bases afines, la prueba de que dos bases afines tienen el mismo cardinal, etc.

Este camino aunque se puede realizar, sin embargo, tiene dos desventajas. La primera es que aún no hemos definido combinaciones afines y estas se necesitan para probar todo esto. La segunda, y más obvia, es que todo esto es muy largo. Por suerte, hay una sencilla relación que enlaza las bases afines con las bases del espacio vectorial y esto nos ahorrará mucho trabajo.

2.4.1 *Sea $A = \{x_0, x_1, \dots, x_n\}$ un conjunto de puntos. Definimos $A_0 = \{x_1 - x_0, \dots, x_n - x_0\}$. Entonces, A es una base afín si y solo si A_0 es una base del espacio vectorial.*

Prueba. Veamos que $[A] = \langle A_0 \rangle + x_0$. Efectivamente, $\langle A_0 \rangle + x_0$ es un subespacio afín que contiene a A y por lo tanto $[A] \subseteq \langle A_0 \rangle + x_0$. Por otro lado $[A] - x_0$ es un subespacio por el origen que contiene a A_0 y por lo tanto $[A] - x_0 \supseteq \langle A_0 \rangle$.

De aquí inmediatamente obtenemos que A es generador afín si y solo si A_0 es un generador. Luego, podemos suponer por el resto de la prueba que tanto A como A_0 son generadores. Nos queda probar que A es AI si y solo si A_0 es LI.

Supongamos que A_0 es LD. Sea B_0 una base lineal tal que $B_0 \subset A_0$. Como al principio de la prueba $B = \{x_0\} \cup (B_0 + x_0)$ es un generador afín del espacio. Como $B_0 \neq A_0$ entonces, $B \neq A$ y por lo tanto A es AD.

Supongamos que A es AD. Entonces, hay un subconjunto propio B de A que genera al espacio. Sea $y \in B$. Como al principio de la prueba $B' = \{x_i - y \mid x_i \in B \setminus \{y\}\}$ es un generador lineal del espacio. Luego la dimensión del espacio es estrictamente menor que n y por lo tanto A_0 no puede ser LI. ■

Ahora podemos traspasar todos los resultados probados para las bases del espacio vectorial a las bases afines. En particular, es cierto que:

1. Las bases afines son los conjuntos generadores afines minimales por inclusión.
2. Las bases afines son los conjuntos AI maximales por inclusión.
3. Si A es un conjunto AI y B es un conjunto generador afín tal que $A \subseteq B$ entonces hay una base afín C tal que $A \subseteq C \subseteq B$.
4. Dos bases afines cualesquiera tienen el mismo número de puntos (uno más que la dimensión).

Las demostraciones de estos resultados son obvias dada la correspondencia entre las bases afines y las bases del espacio vectorial.

Ejercicio 56 Formule el lema del cambio para las bases afines.

El siguiente resultado es el análogo del lema Lema de Aumento de un Conjunto LI (2.11) para el caso afín y es consecuencia directa de la correspondencia entre las bases afines y las bases del espacio vectorial.

2.4.2

Si A es un conjunto AI entonces $A \cup b$ es AD si y solo si $b \in [A]$.



2.9 El caso de dimensión infinita

En la Sección 2.4 enunciamos el Teorema de Existencia de Bases (2.14) y la Equicardinalidad de las Bases (2.16) pero solo los probamos para el caso de que el espacio tiene dimensión finita. En esta sección demostramos estos resultados para los casos faltantes. Porque siempre aparece un curioso que quiere saber.

Como estas demostraciones dependen de resultados de Teoría de Conjuntos y una exposición de esta nos llevaría a escribir otro libro, lo haremos todo en forma minimalista: Antes de cada una de las dos pruebas daremos las definiciones y resultados exclusivamente que necesitamos. Los resultados complicados de Teoría de Conjuntos no los demostraremos.

El Lema de Zorn

Un **conjunto ordenado** es un conjunto con una **relación de orden**, o sea una relación reflexiva antisimétrica y transitiva (véase el glosario). Sea P un conjunto ordenado y A un subconjunto de P . Diremos que $x \in P$ es una **cota superior** de A si para cualquier $y \in A$ se cumple que $y \leq x$. Diremos que $x \in A$ es **elemento maximal** de A si para cualquier $y \in A$ se cumple que $x \not\leq y$. Se dice que A es una **cadena** si para cualesquiera $x, y \in A$ se cumple que $x \leq y$ o que $y \leq x$. Diremos que A está **inductivamente ordenado** si $A \neq \emptyset$ y cualquier cadena contenida en A tiene una cota superior que está en A . El siguiente resultado es clásico en teoría de conjuntos.

Lema de Zorn

2.4.3

Cualquier conjunto inductivamente ordenado tiene un elemento maximal.

Existencia de bases

Teorema de Existencia de Bases (caso general)

2.44 Sea N un conjunto generador y $L \subseteq N$ un conjunto LI. Entonces, existe una base M tal que $L \subseteq M \subseteq N$.

Prueba. Sean L y N como en las hipótesis. Denotemos

$$\mathcal{T} = \{M \mid M \text{ es linealmente independiente y } L \subseteq M \subseteq N\}.$$

El conjunto \mathcal{T} está naturalmente ordenado por inclusión. Sea M un elemento maximal de \mathcal{T} . Entonces $\forall x \in N \setminus M \quad M \cup x$ es dependiente y por el Lema de Aumento de un Conjunto LI (2.11) tenemos $N \subseteq \langle M \rangle$. Como N es generador, esto significa que M también lo es y por lo tanto M es una base.

Nuestra tarea es encontrar un elemento maximal de \mathcal{T} . Probaremos que el conjunto \mathcal{T} está inductivamente ordenado. Efectivamente, \mathcal{T} es no vacío ya que $L \in \mathcal{T}$. Sea $\{B_i\}_{i \in I}$ una cadena arbitraria en \mathcal{T} y denotemos $B = \bigcup_{i \in I} B_i$. El conjunto B es una cota superior de la cadena $\{B_i\}_{i \in I}$ y tenemos que convencernos que B está en \mathcal{T} . Como para cualquier i tenemos $B_i \subseteq N$ entonces, $B \subseteq N$. Supongamos que B es linealmente dependiente. Entonces, existe un subconjunto finito B' de B y una combinación lineal de B' igual a cero tal que todos sus coeficientes son no cero, o sea B' es linealmente dependiente. Como B' es finito, y $\{B_i\}_{i \in I}$ es una cadena entonces, tiene que existir i_0 tal que $B' \subseteq B_{i_0}$ y esto contradice que todo subconjunto de un conjunto linealmente independiente es linealmente independiente. Luego, \mathcal{T} está inductivamente ordenado y por el Lema de Zorn (2.43) el conjunto \mathcal{T} tiene un elemento maximal. ■

Cardinales

Dados dos conjuntos A y B denotaremos $|A| \leq |B|$ si existe una inyección de A en B . La relación $A \sim B$ definida como $|A| \leq |B|$ y $|B| \leq |A|$ es una relación de equivalencia entre todos los conjuntos. A la clase de equivalencia que contiene al conjunto A se le llama **cardinal** del conjunto A y se denota por $|A|$. Los cardinales están ordenados por la relación de orden que ya definimos. Los cardinales pueden ser finitos o infinitos. El cardinal infinito más pequeño es \aleph_0 que es el cardinal de los naturales (\aleph es la primera letra del alfabeto hebreo y se llama “álef”). La **suma** de cardinales se define como el cardinal de la unión de dos conjuntos disjuntos. El **producto** de cardinales se define como el cardinal del producto cartesiano de conjuntos.

Necesitaremos dos resultados acerca de los cardinales de conjuntos. El primero es muy sencillo y daremos una demostración para él.

2.45 Si $\forall i \quad |A_i| \leq t$ entonces, $\sum_{i \in I} |A_i| \leq t |I|$.

Prueba. Podemos pensar que los A_i son disjuntos. Sea T un conjunto de cardinal t . Por hipótesis existen inyecciones $f_i : A_i \hookrightarrow T$. Sea $\varphi : \bigcup_{i \in I} A_i \rightarrow T \times I$ definida por $\varphi(a) = (f_i(a), i)$ si $a \in A_i$. Por definición de φ si $\varphi(a) = \varphi(b)$ entonces, a y b están en el mismo conjunto A_i , $f_i(a) = f_i(b)$ y por lo tanto $a = b$. Luego, φ es inyectiva. ■

El siguiente resultado que necesitamos se demuestra (usando muchas cosas) en la Teoría de Conjuntos (véase por ejemplo: Kamke E., *Theory of sets*. Dover, New York, 1950. página 121). Nosotros omitiremos la prueba.

2.46 Si $|A|$ es infinito entonces, $\aleph_0 |A| = |A|$.

Equicardinalidad de las bases

Equicardinalidad de las Bases (caso infinito)

2.47 Dos bases cualesquiera tienen el mismo cardinal.

Prueba. Sean A y B dos bases. Ya probamos el caso en que una de las dos bases es finita. Luego, podemos suponer que A y B son infinitos. Como B es una base y debido a la finitud de las combinaciones lineales entonces $\forall a \in A$ el mínimo subconjunto $B_a \subseteq B$ tal que $a \in \langle B_a \rangle$ existe y es finito.

Construyamos la relación $R \subseteq A \times B$ de manera que $(a, b) \in R$ si y solo si $b \in B_a$. Tenemos $|B| \leq |R|$ ya que si hubiera un $b_0 \in B$ no relacionado con ningún elemento de A entonces, $A \subseteq \langle B \setminus b_0 \rangle$ y como A es base obtendríamos que $B \setminus b_0$ es generador lo que contradice que B es base.

Por otro lado, como $|A|$ es infinito y $|B_a|$ es finito entonces, usando 2.45 y 2.46 obtenemos

$$|B| \leq |R| = \sum_{a \in A} |B_a| \leq \aleph_0 |A| = |A|.$$

Luego $|B| \leq |A|$ y por simetría de nuestros razonamientos $|A| \leq |B|$. ■



Capítulo tercero

Transformaciones Lineales

as transformaciones lineales son uno de los objetos más estudiados y más importantes en las matemáticas. El objetivo de este capítulo es familiarizar al lector con el concepto de transformación lineal y sus propiedades básicas. Introduciremos las matrices y estudiaremos la relación de las mismas con las transformaciones lineales. Veremos los conceptos de nucleo e imagen de una transformación lineal y como estos se usan para reducir el estudio de las transformaciones lineales al estudio de las transformaciones lineales biyectivas.

3.1 Definición y ejemplos

Sean \mathfrak{E} y \mathfrak{F} dos espacios vectoriales sobre \mathbb{K} . Una función $f : \mathfrak{E} \rightarrow \mathfrak{F}$ se le llama **transformación lineal** de \mathfrak{E} en \mathfrak{F} si para cualesquier vectores \mathbf{a} y \mathbf{b} y cualquier escalar λ se cumplen las propiedades del recuadro a la derecha.

$$\begin{aligned} f(\mathbf{a} + \mathbf{b}) &= f(\mathbf{a}) + f(\mathbf{b}) \\ f(\lambda\mathbf{a}) &= \lambda f(\mathbf{a}) \end{aligned}$$

 Nuevamente, para no reescribir constantemente frases largas, en lugar de decir que f es una transformación lineal, diremos que f es una TL. En plural escribiremos TLs.

Imágenes de subespacios

 Toda TL transforma subespacios en subespacios.

Prueba. Sea $f : \mathfrak{E} \rightarrow \mathfrak{F}$ una TL y \mathfrak{E}' un subespacio de \mathfrak{E} . Denotemos $\mathfrak{F}' = f(\mathfrak{E}')$ la imagen de \mathfrak{E}' . Sean \mathbf{a} y \mathbf{b} vectores en \mathfrak{F}' . Existen vectores \mathbf{x}, \mathbf{y} tales que $f(\mathbf{x}) = \mathbf{a}$ y $f(\mathbf{y}) = \mathbf{b}$. Tenemos, $f(\mathbf{x} + \mathbf{y}) = \mathbf{a} + \mathbf{b}$ por lo que $\mathbf{a} + \mathbf{b} \in \mathfrak{F}'$. Sea ahora λ un escalar. Tenemos $f(\lambda\mathbf{x}) = \lambda\mathbf{a}$ por lo que $\lambda\mathbf{a} \in \mathfrak{F}'$. Luego, \mathfrak{F}' es un subespacio de \mathfrak{F} . ■



Las TLs NO necesariamente transforman conjuntos LI en conjuntos LI ni tampoco conjuntos generadores en conjuntos generadores.

Homotecias

Veamos el ejemplo más simple de TL. Si a cada vector \mathbf{x} de un espacio \mathbb{E} le hacemos corresponder el vector $2\mathbf{x}$ obtenemos la función $h_2 : \mathbb{E} \ni \mathbf{x} \mapsto 2\mathbf{x} \in \mathbb{E}$. Esta función es una TL ya que $h_2(\mathbf{a} + \mathbf{b}) = 2(\mathbf{a} + \mathbf{b}) = 2\mathbf{a} + 2\mathbf{b} = h_2(\mathbf{a}) + h_2(\mathbf{b})$

$$h_2(\lambda\mathbf{a}) = 2\lambda\mathbf{a} = \lambda 2\mathbf{a} = \lambda h_2(\mathbf{a})$$

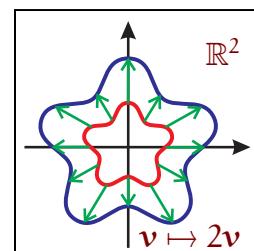
Observese que en el segundo renglón se usa la commutatividad del producto de escalares.

Lo mismo ocurre si en lugar del escalar 2 usamos cualquier otro escalar $\alpha \in \mathbb{K}$ obteniendo la TL $h_\alpha : \mathbb{E} \ni \mathbf{x} \mapsto \alpha\mathbf{x} \in \mathbb{E}$. A las funciones h_α se les llama **homotecias**.

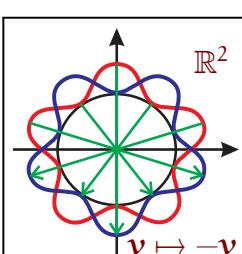
Hay dos casos particulares de homotecias especialmente importantes. Si $\alpha = 1$ entonces obtenemos la función **identidad** $\mathbf{x} \mapsto \mathbf{x}$. A esta función se la denotará por \mathbb{I} . Si $\alpha = 0$ entonces obtenemos la **función nula** $\mathbf{x} \mapsto \mathbf{0}$ que se denotará por \mathbb{O} .

En el caso del campo \mathbb{R} las homotecias tienen una interpretación geométrica muy clara. Si $\alpha > 0$ entonces cada punto $\mathbf{x} \in \mathbb{R}^n$ se transforma en el punto $\alpha\mathbf{x}$ que está en la misma recta por el origen que \mathbf{x} pero a una distancia del origen α veces mayor que \mathbf{x} . Esto quiere decir que h_α es la **dilatación** de razón α . Si $0 < \alpha < 1$ entonces h_α es una **contracción** de razón $1/\alpha$.

En la figura de la derecha observamos la dilatación $\mathbf{x} \mapsto 2\mathbf{x}$ en el plano cartesiano \mathbb{R}^2 . La curva interior (que es la gráfica de la función $5 + \sin 5\theta$ en coordenadas polares) se transforma en la misma curva pero del doble de tamaño ($10 + 2\sin 5\theta$).



Si $\alpha = -1$ entonces a la homotecia $h_{-1} : \mathbf{x} \mapsto -\mathbf{x}$ se le llama **función antipodal**. El nombre viene de la siguiente interpretación. Si trazamos una recta que pase por el centro de la Tierra intersecaremos la superficie en dos puntos. Estos puntos se dicen que son antípodas o sea, nuestros antípodas son las personas que viven "pies arriba" del "otro lado" de la Tierra. Si coordinatizamos la Tierra con unos ejes de coordenadas con origen en el centro de la misma, nuestros antípodas se obtienen multiplicando por -1 .



En la figura de la izquierda está representada la función antípodal en \mathbb{R}^2 . En ella hay dos curvas cerradas de cinco pétalos. La primera es la misma que la de la figura anterior ($10 + 2\sin 5\theta$). La segunda es la antípoda de la primera ($-10 - 2\sin 5\theta$). La figura es expresamente complicada para que el lector tenga la oportunidad de pensar un poco en que punto se va a transformar cada punto. Cualquier homotecia h_α con $\alpha < 0$ se puede representar como $h_{-1}(h_{-\alpha})$ por lo que h_α se puede interpretar como una dilatación seguida de la función antípodal. A pesar de su simplicidad las homotecias juegan un papel fundamental en la comprensión de las TL y este papel se debe a que ellas son todas las TL en dimensión 1.



Si $\dim \mathfrak{E} = 1$ entonces toda TL de \mathfrak{E} en \mathfrak{E} es una homotecia.

Prueba. Sea $\{\mathbf{a}\}$ una base de \mathfrak{E} y f una TL de \mathfrak{E} en \mathfrak{E} . Como $\{\mathbf{a}\}$ es una base tiene que existir un escalar $\lambda \in \mathbb{K}$ tal que $f(\mathbf{a}) = \lambda\mathbf{a}$. Sea $\mathbf{x} = \alpha\mathbf{a}$ un vector arbitrario en \mathfrak{E} . Como f es lineal tenemos $f(\mathbf{x}) = f(\alpha\mathbf{a}) = \alpha f(\mathbf{a}) = \alpha\lambda\mathbf{a} = \lambda\alpha\mathbf{a} = \lambda\mathbf{x}$ lo que demuestra que f es una homotecia. ■

Inmersiones

Hasta ahora las TL que hemos visto (las homotecias) están definidas de un espacio en si mismo. Las inmersiones son la TL más simples que están definidas de un espacio en otro distinto. Sea \mathfrak{E} un subespacio de \mathfrak{F} . A la función $i : \mathfrak{E} \ni \mathbf{x} \mapsto \mathbf{x} \in \mathfrak{F}$ se le llama **inmersión** de \mathfrak{E} en \mathfrak{F} . Las inmersiones son TLs inyectivas y son las restricciones a subespacios de la función identidad. No hay mucho que decir de las inmersiones excepto de que hay una para cada subespacio.

Proyecciones

Ahora supongamos al revés (de las inmersiones) que \mathfrak{F} es subespacio de \mathfrak{E} . ¿Habrá alguna TL natural de \mathfrak{E} en \mathfrak{F} ? Lo que podemos hacer es buscarnos un espacio \mathfrak{G} complementario de \mathfrak{F} (existe por 2.28) De esta manera $\mathfrak{E} = \mathfrak{F} \oplus \mathfrak{G}$ y por 2.29 tenemos que cada vector $\mathbf{x} \in \mathfrak{E}$ se descompone de manera única como $\mathbf{a} + \mathbf{b}$ donde \mathbf{a} es un vector en \mathfrak{F} y \mathbf{b} es un vector en \mathfrak{G} . Así para cada $\mathbf{x} \in \mathfrak{E}$ tenemos un único $\mathbf{a} \in \mathfrak{F}$ y esto nos da una función que denotaremos por $\pi_{\mathfrak{F}}$ y la llamaremos **proyección** de \mathfrak{E} en \mathfrak{F} a lo largo de \mathfrak{G} . Cualquier proyección $\pi_{\mathfrak{F}}$ restringida a \mathfrak{F} es la identidad por lo que necesariamente es sobreyectiva.



La notación $\pi_{\mathfrak{F}}$ sugiere erróneamente que esta función solo depende del subespacio \mathfrak{F} . Esto no es cierto, cualquier subespacio \mathfrak{F} tiene muchos subespacios complementarios diferentes y la proyección depende del subespacio complementario que escojamos.

Las proyecciones son transformaciones lineales.

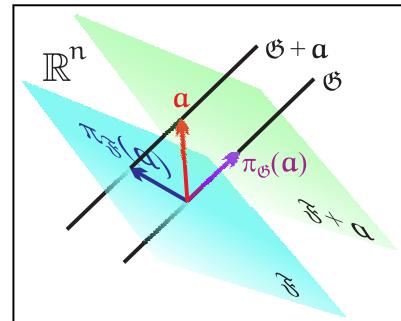
Prueba. Escojamos un subespacio \mathfrak{G} complementario de \mathfrak{F} . De esta manera $\pi_{\mathfrak{F}}$ es una función fija y bien definida.

Si \mathbf{x}, \mathbf{y} son dos vectores entonces hay descomposiciones únicas $\mathbf{x} = \pi_{\mathfrak{F}}(\mathbf{x}) + \pi_{\mathfrak{G}}(\mathbf{x})$ y $\mathbf{y} = \pi_{\mathfrak{F}}(\mathbf{y}) + \pi_{\mathfrak{G}}(\mathbf{y})$. Luego $\mathbf{x} + \mathbf{y} = (\pi_{\mathfrak{F}}(\mathbf{x}) + \pi_{\mathfrak{F}}(\mathbf{y})) + (\pi_{\mathfrak{G}}(\mathbf{x}) + \pi_{\mathfrak{G}}(\mathbf{y}))$. El primer sumando de la derecha está en \mathfrak{F} y el segundo en \mathfrak{G} . Por la unicidad de la descomposición necesariamente tenemos $\pi_{\mathfrak{F}}(\mathbf{x} + \mathbf{y}) = \pi_{\mathfrak{F}}(\mathbf{x}) + \pi_{\mathfrak{F}}(\mathbf{y})$.

Sea ahora $\lambda \in \mathbb{K}$. Tenemos $\lambda\mathbf{x} = \lambda\pi_{\mathfrak{F}}(\mathbf{x}) + \lambda\pi_{\mathfrak{G}}(\mathbf{x})$. Nuevamente, el primer sumando de la derecha está en \mathfrak{F} y el segundo en \mathfrak{G} y por la unicidad de la descomposición necesariamente tenemos $\pi_{\mathfrak{F}}(\lambda\mathbf{x}) = \lambda\pi_{\mathfrak{F}}(\mathbf{x})$. ■

¿Como son geométricamente las proyecciones? En esencia el problema es el siguiente. Dados dos espacios complementarios \mathfrak{F} , \mathfrak{G} y un vector \mathbf{a} como hallar un vector en \mathfrak{F} y otro en \mathfrak{G} que sumados den \mathbf{a} . Esto ya lo hicimos una vez cuando introducimos las coordenadas cartesianas en \mathbb{R}^2 . Dado un vector \mathbf{a} trazamos la recta paralela al eje y que pasa por \mathbf{a} y la intersección del eje x con esta recta es un vector $\pi_x(\mathbf{a})$. De manera análoga obtenemos $\pi_y(\mathbf{a})$ y sabemos que $\mathbf{a} = \pi_x(\mathbf{a}) + \pi_y(\mathbf{a})$.

En el caso general es exactamente igual. Esta construcción se ilustra en la figura de la derecha. Tenemos que \mathfrak{F} es un subespacio de dimensión k y uno de sus complementarios \mathfrak{G} es de dimensión $n - k$. Hay un solo subespacio afín paralelo a \mathfrak{F} que pasa por \mathbf{a} y este es $\mathfrak{F} + \mathbf{a}$. La intersección $(\mathfrak{F} + \mathbf{a}) \cap \mathfrak{G}$ es un solo punto y este punto es el vector $\pi_{\mathfrak{G}}(\mathbf{a})$. Análogamente se observa que $\pi_{\mathfrak{F}}(\mathbf{a}) = (\mathfrak{G} + \mathbf{a}) \cap \mathfrak{F}$. Como es lógico, no pudimos dibujar el espacio \mathbb{R}^n así que el lector deberá contentarse con el caso $n = 3$, $k = 2$ y con la prueba general.



3.4 Si \mathfrak{F} y \mathfrak{G} son complementarios entonces, para cualquier vector \mathbf{a} se cumple $\mathbf{a} = ((\mathfrak{G} + \mathbf{a}) \cap \mathfrak{F}) + ((\mathfrak{F} + \mathbf{a}) \cap \mathfrak{G})$.

Prueba. Como \mathfrak{F} y \mathfrak{G} son complementarios $(\mathfrak{G} + \mathbf{a}) \cap \mathfrak{F}$ es un solo punto que denominaremos \mathbf{x} . Sea $\mathbf{y} \in \mathfrak{G}$ tal que $\mathbf{a} = \mathbf{x} + \mathbf{y}$. Tenemos que $\mathfrak{F} + \mathbf{a} = \mathfrak{F} + \mathbf{x} + \mathbf{y} = \mathfrak{F} + \mathbf{y}$ por lo tanto $\mathbf{y} \in (\mathfrak{F} + \mathbf{a})$. Luego $\mathbf{y} = (\mathfrak{F} + \mathbf{a}) \cap \mathfrak{G}$. ■

3.2 Operaciones entre transformaciones lineales

Ahora, veremos que operaciones se definen naturalmente entre las TLs.

El espacio vectorial de las transformaciones lineales

Sea f una TL y λ un escalar. Denotaremos por λf a la función $\mathbf{a} \mapsto \lambda f(\mathbf{a})$. A esta operación se le llama **producto de un escalar por una TL**.

3.5 El producto de un escalar por una TL es una TL.

Prueba. Efectivamente, tenemos

$$\begin{aligned} (\lambda f)(\mathbf{a} + \mathbf{b}) &= \lambda f(\mathbf{a} + \mathbf{b}) = \lambda(f(\mathbf{a}) + f(\mathbf{b})) = (\lambda f)(\mathbf{a}) + (\lambda f)(\mathbf{b}) \\ (\lambda f)(\alpha \mathbf{a}) &= \lambda f(\alpha \mathbf{a}) = \lambda \alpha f(\mathbf{a}) = \alpha \lambda f(\mathbf{a}) = \alpha (\lambda f)(\mathbf{a}) \end{aligned}$$

lo que significa que λf es una TL. ■

Sean ahora f y g dos transformaciones lineales. Denotaremos por $f + g$ a la función

Sección 3.2 Operaciones entre transformaciones lineales

$\mathbf{a} \mapsto f(\mathbf{a}) + g(\mathbf{a})$. A esta operación se le llama **suma de TLs**.

3.6 *La suma de dos TLs es una TL.*

Prueba. Denotemos $\mathbf{h} = f + g$. Tenemos

$$\mathbf{h}(\alpha\mathbf{a}) = f(\alpha\mathbf{a}) + g(\alpha\mathbf{a}) = \alpha(f(\mathbf{a}) + g(\mathbf{a})) = \alpha\mathbf{h}(\mathbf{a})$$

$$\mathbf{h}(\mathbf{a} + \mathbf{b}) = f(\mathbf{a} + \mathbf{b}) + g(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + g(\mathbf{a}) + f(\mathbf{b}) + g(\mathbf{b}) = \mathbf{h}(\mathbf{a}) + \mathbf{h}(\mathbf{b})$$

lo que significa que \mathbf{h} es una TL. ■

Luego, podemos sumar transformaciones lineales y multiplicarlas por escalares. Los axiomas de espacio vectorial se comprueban de manera muy simple usando las definiciones. Por ejemplo, la prueba de la distributividad del producto por escalares con respecto a la suma es la siguiente: $(\lambda(f + g))(\mathbf{a}) = \lambda(f(\mathbf{a}) + g(\mathbf{a})) = \lambda f(\mathbf{a}) + \lambda g(\mathbf{a}) = (\lambda f + \lambda g)(\mathbf{a})$. Al espacio vectorial de todas las TLs del espacio \mathfrak{E} en el espacio \mathfrak{F} lo denotaremos por $\text{Mor}(\mathfrak{E}, \mathfrak{F})$. Esta notación es debido que a las transformaciones lineales también se les llama **morfismos de espacios vectoriales**. Debido a todo lo dicho es válido el siguiente resultado:

3.7 *Mor($\mathfrak{E}, \mathfrak{F}$) es un espacio vectorial.*

Composición de transformaciones lineales

Sean $f \in \text{Mor}(\mathfrak{E}, \mathfrak{F})$ y $g \in \text{Mor}(\mathfrak{F}, \mathfrak{G})$ dos TLs. La composición $\mathbf{h} = g \circ f$ se define como la función $\mathfrak{E} \ni \mathbf{a} \mapsto g(f(\mathbf{a})) \in \mathfrak{G}$. Demostremos que $\mathbf{h} = g \circ f \in \text{Mor}(\mathfrak{E}, \mathfrak{G})$ o sea, que \mathbf{h} es una TL.

3.8 *La composición de TLs es una TL.*

Prueba. Sea $\mathbf{h} = g \circ f$ la composición de dos TLs. Tenemos

$$\mathbf{h}(\mathbf{a} + \mathbf{b}) = g(f(\mathbf{a} + \mathbf{b})) = g(f(\mathbf{a}) + f(\mathbf{b})) = g(f(\mathbf{a})) + g(f(\mathbf{b})) = \mathbf{h}(\mathbf{a}) + \mathbf{h}(\mathbf{b})$$

$$\mathbf{h}(\alpha\mathbf{a}) = g(f(\alpha\mathbf{a})) = g(\alpha f(\mathbf{a})) = \alpha g(f(\mathbf{a})) = \alpha\mathbf{h}(\mathbf{a})$$

que es lo que se quería demostrar. ■

3.9 *La composición de TLs cumple las siguientes propiedades:*

1. $f \circ (g \circ h) = (f \circ g) \circ h$ *(asociatividad)*
2. $f \circ (g + h) = f \circ g + f \circ h$ *(distributividad a la izquierda)*
3. $(f + g) \circ h = f \circ h + g \circ h$ *(distributividad a la derecha)*
4. $f \circ \lambda g = \lambda f \circ g = \lambda(f \circ g)$ *(commuta con el producto por escalares)*

Prueba. Ya vimos en el Capítulo 1 que la composición es asociativa. Con

$$(f \circ (g + h))(a) = f((g + h)(a)) = f(g(a) + h(a)) =$$

$$= f(g(a)) + f(h(a)) = (f \circ g)(a) + (f \circ h)(a) = ((f \circ g) + (f \circ h))(a)$$

probamos la distributividad a la izquierda. Para la distributividad a la derecha usamos

$$((f + g) \circ h)(a) = (f + g)(h(a)) = f(h(a)) + g(h(a)) =$$

$$= (f \circ h)(a) + (g \circ h)(a) = ((f \circ h) + (g \circ h))(a)$$

Finalmente, probamos que la composición conmuta con el producto por escalares con

$$(f \circ \lambda g)(a) = f(\lambda g(a)) = \lambda f(g(a)) = (\lambda(f \circ g))(a) =$$

$$= \lambda f(g(a)) = (\lambda f)(g(a)) = (\lambda f \circ g)(a).$$

El lector debe ya saber encontrar por si mismo el porqué de la validez de cada una de las igualdades utilizadas en esta prueba. ■

El álgebra de operadores lineales

A una TL de un espacio vectorial en sí misma se le llama **operador lineal**. Nuevamente, usaremos la abreviatura OL para referirnos a los operadores lineales. Los OLs juegan (como veremos más adelante) un papel muy importante en el estudio de las TLs y por esto es que merecen un nombre especial. El conjunto de todos los OLs de \mathbb{E} se denotará por $\text{End } \mathbb{E}$. Lo hacemos así porque a los OLs también se les llama **endomorfismos** de un espacio vectorial. Por definición tenemos $\text{End } \mathbb{E} = \text{Mor}(\mathbb{E}, \mathbb{E})$. La principal diferencia entre las TLs y los OLs es que la operación de composición es una operación interna en espacio vectorial $\text{End } \mathbb{E}$. O sea, si componemos dos OLs, obtenemos otro OL.

Si un espacio vectorial cualquiera (el cual ya trae definidos la suma y el producto por escalares) tiene otra operación binaria $*$ que cumple los axiomas en el recuadro a la derecha entonces, se le llama **álgebra**. Observese que los primeros tres axiomas los podemos resumir en uno: la suma de vectores y $*$ definen un anillo en el espacio vectorial. El cuarto axioma lo que quiere decir es que para cualquier escalar λ y cualesquier vectores a y b se cumple que $\lambda a * b = a * \lambda b = \lambda(a * b)$.

Alg1) * es asociativa

Alg2) * es distributiva con respecto a +

Alg3) * tiene elemento neutro

Alg4) * conmuta con el producto por escalares

Ya vimos (3.9) que la operación de composición de OLs cumple los axiomas Alg1, Alg2 y Alg4. Para ver que $\text{End } \mathbb{E}$ es un álgebra solo nos queda comprobar que la composición tiene elemento neutro. Pero esto es obvio ya que la función identidad cumple que $f \circ \mathbb{I} = \mathbb{I} \circ f = f$. O sea, es el neutro para la composición. Hemos demostrado el siguiente resultado

 *El espacio vectorial $\text{End } \mathbb{E}$ en un álgebra con respecto a la composición.*

Hay otras dos álgebras importantes que deben ser muy conocidas por el lector. Primero, el conjunto de los polinomios con coeficientes reales $\mathbb{R}[x]$ es un espacio vectorial

sobre \mathbb{R} , pero además sabemos multiplicar polinomios. Este producto cumple todos los axiomas de Alg1-Alg4. Este ejemplo se generaliza a los polinomios sobre cualquier campo. El segundo ejemplo son los números complejos. Estos son un espacio vectorial de dimensión dos sobre los reales pero además sabemos multiplicar números complejos. La multiplicación de complejos también cumple todos los axiomas Alg1-Alg4.



Un álgebra se le llama **comutativa** si el producto de vectores es comutativo. El álgebra de los números complejos y el álgebra de polinomios sobre un campo son comutativas. Las álgebras **End** \mathfrak{E} casi nunca son comutativas (salvo en dimensión 1). Por ejemplo en el plano cartesiano \mathbb{R}^2 la rotación f en 45° y la reflección g en el eje y son (como veremos después) OLs. Sin embargo, $(g \circ f)(1, 0) = \frac{1}{\sqrt{2}}(-1, 1) \neq \frac{1}{\sqrt{2}}(-1, -1) = (f \circ g)(1, 0)$.



Un álgebra con división es un álgebra en la cual todo vector tiene inverso multiplicativo. El Teorema de Frobenius (demostrado en 1877) afirma que las álgebras con división de dimensión finita sobre los reales son \mathbb{R} , \mathbb{C} y \mathbb{H} (los cuaterniones), no hay más.

El grupo general lineal

Una función cualquiera es biyectiva si y solo si esta tiene inversa. En el capítulo anterior, cuando vimos los isomorfismos de espacios vectoriales, demostramos que si una TL tiene inversa entonces esta inversa también es una TL. En particular, la función inversa de un OL es un operador lineal. Un operador lineal se le llama **singular** si este no tiene inverso. En el caso contrario se le llama **no singular**. A los OLs no singulares también se les llama **automorfismos** del espacio vectorial. En otras palabras los automorfismos son los endomorfismos biyectivos.

Al conjunto de todos los OLs no singulares del espacio vectorial \mathfrak{E} se le denota por **GL** (\mathfrak{E}). La suma de OLs no singulares puede ser singular ya que, por ejemplo, la función nula cumple que $0 = f - f$. Sin embargo, la composición de OLs no singulares es siempre un OL no singular. Luego, la composición es una operación binaria en **GL** (\mathfrak{E}) que ya sabemos que es asociativa y tiene elemento neutro. Además, cada elemento tiene inverso ya que $f \circ f^{-1} = f^{-1} \circ f = I$. Luego, **GL** (\mathfrak{E}) es un grupo para la composición al cual se llama **grupo general lineal** del espacio vectorial \mathfrak{E} .

3.3 Extensiones lineales

Estudiaremos en esta sección una manera universal de construir TLs. Pero antes, veamos algunas consideraciones de tipo general acerca de funciones entre conjuntos arbitrarios.

Extensiones y restricciones

Sean A y B dos conjuntos y A' un subconjunto de A . Cada vez que se tiene una función $f : A \rightarrow B$ también se tiene una función $f' : A' \rightarrow B$ definida por $f'(a) = f(a)$ para cualquier $a \in A'$. A la función f' se le llama **restricción** de f a A' y la denotaremos por $f_{A'}$. Todas las diferentes restricciones de f se obtienen al tomar diferentes subconjuntos A' . Las inmersiones son las restricciones de la identidad.

Si h y g son dos funciones tales que h es una restricción de g entonces se dice que g es una **extensión** de h . Si está dada una función $g : A' \rightarrow B$ y A es un sobreconjunto de A' entonces, pueden haber muchas extensiones $h : A \rightarrow B$ de g , ya que podemos escoger arbitrariamente los valores de $h(x)$ para todos los $x \in A \setminus A'$.

Es un problema frecuente en matemáticas el encontrar extensiones que cumplan ciertas propiedades. En nuestro caso, debemos encontrar extensiones que sean TLs. Formulemos nuestro problema más precisamente. Sean \mathfrak{E} y \mathfrak{F} dos espacios vectoriales sobre \mathbb{K} y N un conjunto de vectores de \mathfrak{E} . Sea $h : N \rightarrow \mathfrak{F}$ una TL. Sabemos que $N \subset \mathfrak{E}$ y por lo tanto tenemos la restricción $h_N : N \rightarrow \mathfrak{F}$. ¿Será posible para cualquier función $g : N \rightarrow \mathfrak{F}$ encontrar una extensión $h : \mathfrak{E} \rightarrow \mathfrak{F}$ de g que sea TL? ¿Será única tal extensión? Veremos que ambas preguntas tienen respuesta positiva si N es una base.

Para demostrar la existencia de la extensión debemos construirla. Sea N una base de \mathfrak{E} y $g : N \rightarrow \mathfrak{F}$ una función arbitraria de N en \mathfrak{F} . Cualquier $x \in \mathfrak{E}$ se expresa de forma única como combinación lineal de N o sea $x = \sum_{i \in N} \alpha_i i$. A la función h del recuadro a la derecha se le llama **extensión lineal** de g . Observese que (como debe ser) la restricción de h a N es igual a g ya que si $x \in N$ entonces, la descomposición de x en la base N tiene coeficientes $\alpha_i = 0$ para $i \neq x$ y $\alpha_i = 1$ para $i = x$.

$$x \mapsto \sum_{i \in N} \alpha_i g(i)$$

3.11 Las extensiones lineales son transformaciones lineales.

Prueba. Tenemos

$$\begin{aligned} h(x+y) &= \sum_{i \in N} (\alpha_i + \beta_i) g(i) = \sum_{i \in N} \alpha_i g(i) + \sum_{i \in N} \beta_i g(i) = h(x) + h(y) \\ h(\lambda x) &= \sum_{i \in N} \lambda \alpha_i g(i) = \lambda \sum_{i \in N} \alpha_i g(i) = \lambda h(x) \end{aligned}$$

y esto prueba que h es una TL. ■

Para demostrar la unicidad de la extensión debemos convencernos de que dos TLs distintas no pueden coincidir en una base.

3.12 Las TLs están predeterminadas por sus valores en una base.

Prueba. Sean $f, g : \mathfrak{E} \rightarrow \mathfrak{F}$ dos TLs y N una base de \mathfrak{E} . Supongamos que $f(i) = g(i)$ para cualquier $i \in N$. Cualquier $x \in \mathfrak{E}$ se expresa de forma única como combinación

lineal de \mathbf{N} o sea $\mathbf{x} = \sum_{i \in \mathbf{N}} \alpha_i \mathbf{i}$. Luego

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}\left(\sum_{i \in \mathbf{N}} \alpha_i \mathbf{i}\right) = \sum_{i \in \mathbf{N}} \alpha_i \mathbf{f}(\mathbf{i}) = \sum_{i \in \mathbf{N}} \alpha_i \mathbf{g}(\mathbf{i}) = \mathbf{g}\left(\sum_{i \in \mathbf{N}} \alpha_i \mathbf{i}\right) = \mathbf{g}(\mathbf{x})$$

y por lo tanto las TLs \mathbf{f} y \mathbf{g} son iguales. ■

El isomorfismo entre $\mathfrak{F}^{\mathbf{N}}$ y $\text{Mor}(\mathfrak{E}, \mathfrak{F})$

Recordemos ahora de la Sección 2.2 que el conjunto de todas las funciones de \mathbf{N} en \mathfrak{F} es el conjunto de las \mathbf{N} -adas de vectores de \mathfrak{F} , que este es un espacio vectorial para la suma y el producto por escalares definidos por coordenadas y que se denota por $\mathfrak{F}^{\mathbf{N}}$.

Si \mathbf{N} es una base de \mathfrak{E} entonces, hemos construido una correspondencia biunívoca entre $\mathfrak{F}^{\mathbf{N}}$ y $\text{Mor}(\mathfrak{E}, \mathfrak{F})$. A cada \mathbf{N} -ada $\mathbf{h}_{\mathbf{N}} \in \mathfrak{F}^{\mathbf{N}}$ le corresponde su extensión lineal $\mathbf{h} \in \text{Mor}(\mathfrak{E}, \mathfrak{F})$ y a cada TL $\mathbf{h} \in \text{Mor}(\mathfrak{E}, \mathfrak{F})$ le corresponde $\mathbf{h}_{\mathbf{N}}$ su restricción a \mathbf{N} (que es una \mathbf{N} -ada). Veamos que esta correspondencia es un isomorfismo.

3.13 Si \mathbf{N} es una base de \mathfrak{E} entonces, la biyección
 $\tau : \text{Mor}(\mathfrak{E}, \mathfrak{F}) \ni \mathbf{h} \mapsto \mathbf{h}_{\mathbf{N}} \in \mathfrak{F}^{\mathbf{N}}$
 es un isomorfismo de espacios vectoriales.

Prueba. Solo nos queda probar que τ es una TL. Efectivamente, sean $\mathbf{h}, \mathbf{h}' \in \text{Mor}(\mathfrak{E}, \mathfrak{F})$ y λ un escalar. Tenemos $\tau(\lambda \mathbf{h}) = (\lambda \mathbf{h})_{\mathbf{N}} = \lambda \mathbf{h}_{\mathbf{N}} = \lambda \tau(\mathbf{h})$
 $\tau(\mathbf{h} + \mathbf{h}') = (\mathbf{h} + \mathbf{h}')_{\mathbf{N}} = \mathbf{h}_{\mathbf{N}} + \mathbf{h}'_{\mathbf{N}} = \tau(\mathbf{h}) + \tau(\mathbf{h}')$

que se cumplen por las definiciones de suma y producto por escalares de las \mathbf{N} -adas. ■

Un criterio de isomorfismo

Al establecer que los espacios $\text{Mor}(\mathfrak{E}, \mathfrak{F})$ y $\mathfrak{F}^{\mathbf{N}}$ son isomorfos es natural que esperemos que cualquier propiedad de las TL se traduzca de alguna u otra manera al lenguaje de las \mathbf{N} -adas de vectores. En este caso queremos hacer la traducción de la propiedad de una TL de ser o no un isomorfismo de espacios vectoriales. Ya vimos en el capítulo anterior que un isomorfismo transforma una base del dominio en una base del codominio. ¿Será esta propiedad suficiente para comprobar que una TL es un isomorfismo?. La respuesta es NO. Por ejemplo, la extensión lineal de la función definida en la base canónica de \mathbb{R}^3 como en el recuadro a la derecha transforma a esta base en la base canónica de \mathbb{R}^2 y sin embargo no es inyectiva. Nos falta la propiedad evidentemente necesaria de que la restricción de la TL debe ser inyectiva.

$(1, 0, 0) \mapsto (1, 0)$
$(0, 1, 0) \mapsto (0, 1)$
$(0, 0, 1) \mapsto (0, 1)$

3.14 Una TL es un isomorfismo si y solo si su restricción a una base es inyectiva y la imagen de esta restricción es una base.

Prueba. Ya hemos probado la necesidad. Para la suficiencia sea \mathbf{N} una base de \mathfrak{E} y $\mathbf{h}_N \in \mathfrak{F}^N$ una N -ada de vectores de \mathfrak{F} tal que sus coordenadas son todas diferentes (la inyectividad) y que el conjunto de sus coordenadas (la imagen) es una base $\mathbf{M} = \mathbf{h}(\mathbf{N})$ de \mathfrak{F} . Probemos que la extensión lineal $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{F}$ de \mathbf{h}_N es un isomorfismo. Efectivamente, si $\mathbf{x} = \sum_{i \in N} \alpha_i \mathbf{i}$ y $\mathbf{y} = \sum_{i \in N} \beta_i \mathbf{i}$ son dos vectores cualesquiera en \mathfrak{E} y $\mathbf{h}(\mathbf{x}) = \mathbf{h}(\mathbf{y})$ entonces,

$$\sum_{i \in N} \alpha_i \mathbf{h}(\mathbf{i}) = \sum_{i \in N} \beta_i \mathbf{h}(\mathbf{i})$$

y como todos los $\mathbf{h}(\mathbf{i}) = \mathbf{h}_i$ son diferentes, estas son dos combinaciones lineales iguales de la base \mathbf{M} . Luego, los coeficientes de estas combinaciones lineales tienen que coincidir $\alpha_i = \beta_i$ y por lo tanto $\mathbf{x} = \mathbf{y}$. Luego, \mathbf{h} es inyectiva.

Para ver que \mathbf{h} es sobreyectiva sea $\mathbf{z} \in \mathfrak{F}$. Como \mathbf{M} es una base de \mathfrak{F} existen $\gamma_i \in \mathbb{K}$ tales que $\mathbf{z} = \sum_{i \in N} \gamma_i \mathbf{h}(\mathbf{i})$ y por lo tanto $\mathbf{z} = \mathbf{h}(\mathbf{v})$ donde $\mathbf{v} = \sum_{i \in N} \gamma_i \mathbf{i}$. ■

3.4 Coordinatización de transformaciones lineales

Para darle coordenadas a una TL lo primero es darle coordenadas a los espacios entre los cuales está definida la TL. Sean \mathbf{N} y \mathbf{M} bases de \mathfrak{E} y \mathfrak{F} respectivamente. Tenemos los isomorfismos de coordinatización $\mathfrak{E} \leftrightarrow \mathbb{K}^{[N]}$ y $\mathfrak{F} \leftrightarrow \mathbb{K}^{[M]}$. Para cada $f \in \text{Mor}(\mathfrak{E}, \mathfrak{F})$ tenemos la composición $g : \mathbb{K}^{[N]} \rightarrow \mathfrak{E} \xrightarrow{f} \mathfrak{F} \rightarrow \mathbb{K}^{[M]}$ que es una TL en $\text{Mor}(\mathbb{K}^{[N]}, \mathbb{K}^{[M]})$.

Recíprocamente para cada $g \in \text{Mor}(\mathbb{K}^{[N]}, \mathbb{K}^{[M]})$ tenemos la composición $f : \mathfrak{E} \rightarrow \mathbb{K}^{[N]} \xrightarrow{g} \mathbb{K}^{[M]} \rightarrow \mathfrak{F}$ que es una TL en $\text{Mor}(\mathfrak{E}, \mathfrak{F})$. Es fácil ver y es intuitivamente claro que esta correspondencia biunívoca $\text{Mor}(\mathfrak{E}, \mathfrak{F}) \leftrightarrow \text{Mor}(\mathbb{K}^{[N]}, \mathbb{K}^{[M]})$ es un isomorfismo de espacios vectoriales.

$$\begin{array}{ccc} \mathfrak{E} & \xrightarrow{f} & \mathfrak{F} \\ \uparrow & & \downarrow \\ \mathbb{K}^{[N]} & \xrightarrow{g} & \mathbb{K}^{[M]} \end{array}$$

Podemos pensar a \mathbf{N} como la base canónica de $\mathbb{K}^{[N]}$. Luego, aplicando 3.13 obtenemos el isomorfismo $\text{Mor}(\mathbb{K}^{[N]}, \mathbb{K}^{[M]}) \leftrightarrow (\mathbb{K}^{[M]})^N = \mathbb{K}^{[M] \times N}$ que es el conjunto de las MN matrices tales que cada columna es una N -ada finita. Para el caso que más nos interesa en que \mathbf{N} y \mathbf{M} son bases finitas obtenemos $\text{Mor}(\mathfrak{E}, \mathfrak{F}) \leftrightarrow (\mathbb{K}^M)^N = \mathbb{K}^{MN}$. Sea $f \in \text{Mor}(\mathfrak{E}, \mathfrak{F})$. A la matriz α_{MN} que le corresponde a f mediante el isomorfismo construido se le llama **matriz de la TL f** en las bases \mathbf{M} y \mathbf{N} . Los resultados de la sección anterior nos dicen como construir α_{MN} dada f . Para cada $i \in N$ la columna α_{Mi} es el vector $f(i)$ coordinatizado en la base \mathbf{M} . O sea $f(i) = \sum_{a \in M} \alpha_{ai} a$.

Ejercicio 57 Sean $\mathfrak{E} \leftrightarrow \mathfrak{E}'$ y $\mathfrak{F} \leftrightarrow \mathfrak{F}'$ isomorfismos de espacios vectoriales. Construya un isomorfismo $\text{Mor}(\mathfrak{E}, \mathfrak{F}) \leftrightarrow \text{Mor}(\mathfrak{E}', \mathfrak{F}')$.

Sección 3.4 Coordinatización de transformaciones lineales

Ejercicio 58 Pruebe que la función $\mathbb{K}[x] \ni \sum_{i=0}^n a_i x^i \mapsto \sum_{i=0}^n a_i (x+1)^i \in \mathbb{K}[x]$ es un isomorfismo de espacios vectoriales. Construya algunas columnas de la matriz α_{MN} de esta TL en la base canónica del espacio de polinomios. Demuestre que las entradas de esta matriz están definidas por la ecuación recursiva $\alpha_{kn} = \alpha_{k,(n-1)} + \alpha_{(k-1),(n-1)}$ con las condiciones de frontera $\alpha_{kn} = 0$ si $k > n$ y $\alpha_{kn} = 1$ si $k = 0$ o $k = n$.

La fórmula $f(\mathbf{i}) = \sum_{\mathbf{a} \in M} \alpha_{\mathbf{ai}} \mathbf{a}$ nos dice también como construir f dada α_{MN} . Las imágenes de los $\mathbf{i} \in N$ las calculamos por la fórmula y a f la construimos por extensión lineal. O sea, si $\mathfrak{F} \ni x = \sum_{i \in N} \beta_i i$ entonces,

$$\mathfrak{F} \ni f(x) = \sum_{i \in N} \beta_i f(i) = \sum_{i \in N} \beta_i \sum_{\mathbf{a} \in M} \alpha_{\mathbf{ai}} \mathbf{a} = \sum_{\mathbf{a} \in M} \left(\sum_{i \in N} \alpha_{\mathbf{ai}} \beta_i \right) \mathbf{a}$$

La expresión $\sum_{i \in N} \alpha_{\mathbf{ai}} \beta_i$ es un escalar y hay uno para cada $\mathbf{a} \in M$ por lo que son las coordenadas de una M -ada. A esta M -ada se le llama **producto de la matriz α_{MN} por el vector β_N** y se denota por $\alpha_{MN} \beta_N$.

$$\alpha_{MN} \beta_N = \sum_{i \in N} \alpha_{Mi} \beta_i$$

Observese que este producto lo podemos escribir como en el recuadro. Esto quiere decir que este producto se obtiene multiplicando las columnas de α_{MN} por las correspondientes coordenadas de β_N y sumando los resultados. En otras palabras $\alpha_{MN} \beta_N$ es la combinación lineal de las columnas de α_{MN} cuyos coeficientes son las coordenadas de β_N .

En esta sección estudiaremos sistemáticamente el isomorfismo entre las TLs y las matrices repitiendo detalladamente todo lo dicho en esta introducción. Si el lector no entendió, le recomiendo seguir adelante y después releer esta introducción.

El producto escalar canónico

Sean α_N y β_N dos N -adas. El **producto escalar** de estas N -adas es el escalar del recuadro a la derecha. De esta manera, el producto escalar de dos vectores es un elemento del campo.

$$\alpha_N \beta_N = \sum_{i \in N} \alpha_i \beta_i$$

 No se debe confundir el “producto por un escalar” con el “producto escalar”. El primero es un producto de un escalar por un vector y el segundo es un producto de dos vectores. Más adelante veremos que hay otros productos definidos en cualquier espacio vectorial. Por esto a este producto lo llamaremos canónico y solamente está definido en el espacio vectorial de las N -adas para cierto conjunto finito de índices N .

3.15 *El producto escalar cumple las siguientes propiedades:*

1. $\mathbf{x}\mathbf{y} = \mathbf{y}\mathbf{x}$ (comutatividad)
2. $\mathbf{x}(\mathbf{y} + \mathbf{z}) = \mathbf{x}\mathbf{y} + \mathbf{x}\mathbf{z}$ (distributividad a la izquierda)
3. $(\mathbf{y} + \mathbf{z})\mathbf{x} = \mathbf{y}\mathbf{x} + \mathbf{z}\mathbf{x}$ (distributividad a la derecha)
4. $\mathbf{x}(\lambda\mathbf{y}) = (\lambda\mathbf{x})\mathbf{y} = \lambda(\mathbf{x}\mathbf{y})$ (commuta con el producto por escalares)

Ejercicio 59 Pruebe las principales propiedades del producto de **N**-adas (3.15).

Ejercicio 60 Busque tres vectores \mathbf{x} \mathbf{y} \mathbf{z} en \mathbb{R}^2 tales que $(\mathbf{x}\mathbf{y})\mathbf{z} \neq \mathbf{x}(\mathbf{y}\mathbf{z})$. [190]

Ejercicio 61 ¿Se puede definir el producto escalar canónico en $\mathbb{K}^{[N]}$?

Ejercicio 62 Pruebe que $\forall \alpha_N \in \mathbb{R}^N$ se cumple que $\alpha_N^2 = \alpha_N \alpha_N \geq 0$.

El producto de matrices

Sean α_{MN} y β_{NL} dos matrices. Observese que el conjunto de índices de las columnas de la primera, coincide con el conjunto de índices de los renglones de la segunda. Así, tanto un renglón α_{iN} como una columna β_{Nj} son vectores del espacio \mathbb{K}^N de **N**-adas y podemos formar su producto $\alpha_{iN}\beta_{Nj}$. Cuando hacemos esto, para todos los $i \in M$ y todos los $j \in L$ obtenemos una **ML**-matriz formada por todos estos productos. A esta matriz se le llama el **producto de las matrices** α_{MN} y β_{NL} y se denotará por $\alpha_{MN}\beta_{NL}$. Resumiendo, si $\gamma_{ML} = \alpha_{MN}\beta_{NL}$ entonces $\gamma_{ij} = \alpha_{iN}\beta_{Nj}$. Por ejemplo, si los conjuntos de índices son $M = \{1, 2\}$, $N = \{1, 2, 3\}$ y $L = \{1, 2\}$ entonces, en forma gráfica tenemos

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \end{pmatrix} = \begin{pmatrix} \alpha_{1N}\beta_{N1} & \alpha_{1N}\beta_{N2} \\ \alpha_{2N}\beta_{N1} & \alpha_{2N}\beta_{N2} \end{pmatrix}$$

y por definición de producto escalar de vectores tenemos

$$\begin{pmatrix} \alpha_{1N}\beta_{N1} & \alpha_{1N}\beta_{N2} \\ \alpha_{2N}\beta_{N1} & \alpha_{2N}\beta_{N2} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^3 \alpha_{1i}\beta_{i1} & \sum_{i=1}^3 \alpha_{1i}\beta_{i2} \\ \sum_{i=1}^3 \alpha_{2i}\beta_{i1} & \sum_{i=1}^3 \alpha_{2i}\beta_{i2} \end{pmatrix}.$$

Productos de matrices y vectores

Sean α_{MN} y β_{NL} dos matrices. Si el conjunto de indices L tiene un solo elemento entonces la matriz β_{NL} tiene una sola columna. En este caso podemos escribir $\beta_{NL} = \beta_{N1}$ y diremos que β_{N1} es un **vector columna** o una **N-ada columna**. Obviamente, podemos pensar que β_{N1} es una **N-ada** β_N . En este caso podemos hacer el producto de matrices $\alpha_{MN}\beta_{N1} = \alpha_{MN}\beta_N$ y este es el **producto de una matriz por un vector** definido al principio de esta sección. Análogamente se define el producto por el otro lado. Si el conjunto de indices M tiene un solo elemento entonces la matriz α_{MN} tiene un solo renglón. En este caso podemos escribir $\alpha_{MN} = \alpha_{1N}$ y diremos que α_{1N} es un **vector renglón** o **N-ada renglón**. Obviamente, podemos pensar que α_{1N} es una **N-ada** α_N . En este caso podemos hacer el producto de matrices $\alpha_{1N}\beta_{NL} = \alpha_N\beta_{NL}$ y esta es la definición del **producto de un vector por una matriz**.

En este libro, no haremos distinciones entre **N**-adas, **N**-adas columna y **N**-adas renglón o sea $\alpha_{1N} = \alpha_{N1} = \alpha_N$. Para esto, dimos las definiciones de producto de una matriz por un vector y al revés. Intuitivamente, el lector debe pensar que cuando aparezca un vector en un producto de matrices este se convierte en vector fila o columna según sea conveniente.

Sección 3.4 Coordinatización de transformaciones lineales



Claro, este abuso de la notación aunque es muy cómodo puede llevar (si nos ponemos pedantes) a contradicciones. Por ejemplo, podemos sumar dos \mathbf{N} -adas pero no podemos sumar una \mathbf{N} -ada columna con una \mathbf{N} -ada renglón.

Observese que, no solo el producto de matrices por vectores y al revés son casos particulares del producto de matrices, sino también el producto escalar canónico de dos \mathbf{N} -adas al constatar que $\alpha_{\mathbf{N}} \beta_{\mathbf{N}} = \alpha_{\mathbf{1N}} \beta_{\mathbf{N1}}$.

Ejercicio 63 Tome dos matrices con entradas enteras y multiplíquelas. Repita este ejercicio hasta que usted comprenda muy bien el concepto de producto de matrices.

La transformación lineal de una matriz

Sea α_{MN} una MN -matriz. Esta matriz define una función $\mathbb{K}^N \ni \beta_N \rightarrow \alpha_{MN}\beta_N \in \mathbb{K}^M$. Esta función es una TL como ya vimos al principio de esta sección utilizando el isomorfismo $\text{Mor}(\mathbb{K}^N, \mathbb{K}^M) \leftrightarrow \mathbb{K}^{MN}$. Sin embargo, la prueba directa de este hecho es muy sencilla.

3.16 *El multiplicar una matriz fija por \mathbf{N} -adas es una TL.*

Prueba. Sea α_{MN} una matriz cualquiera pero fija. Por las propiedades del producto escalar tenemos que para todo $i \in M$ se cumple que $\alpha_{iN}(\beta_N + \gamma_N) = \alpha_{iN}\beta_N + \alpha_{iN}\gamma_N$ y que $\alpha_{iN}(\lambda\beta_N) = \lambda(\alpha_{iN}\beta_N)$. Esto significa que son válidas las igualdades $\alpha_{MN}(\beta_N + \gamma_N) = \alpha_{MN}\beta_N + \alpha_{MN}\gamma_N$ y $\alpha_{MN}(\lambda\beta_N) = \lambda(\alpha_{MN}\beta_N)$. ■

La matriz de una transformación lineal

En la proposición anterior vimos que al multiplicar MN -matrices por \mathbf{N} -adas obtenemos ejemplos de TLs. Ahora queremos ver que estos son todos los ejemplos posibles, o sea, que cualquier TL de \mathbb{K}^N en \mathbb{K}^M se obtiene multiplicando por una MN -matriz. En realidad, esto ya lo probamos al principio de esta sección al construir el isomorfismo $\text{Mor}(\mathbb{K}^N, \mathbb{K}^M) \leftrightarrow \mathbb{K}^{MN}$. Sin embargo, aquí es más simple ya que tenemos bases canónicas de \mathbb{K}^N y \mathbb{K}^M . Sea $E = \{\mathbf{e}_i : i \in N\}$ la base canónica de \mathbb{K}^N . Recordemos que \mathbf{e}_i es la \mathbf{N} -ada con coordenadas $\delta_{ji} = 1$ si $i = j$ y $\delta_{ji} = 0$ si $i \neq j$. Sea $f : \mathbb{K}^N \rightarrow \mathbb{K}^M$ una TL. Denotemos $\alpha_{Mi} = f(\mathbf{e}_i) \in \mathbb{K}^M$. A la matriz α_{MN} cuyas columnas son las imágenes de la base canónica mediante la TL f la llamaremos **matriz de la TL f** .

3.17

Sea $f : \mathbb{K}^N \rightarrow \mathbb{K}^M$ una TL y α_{MN} su matriz. Entonces, para cualquier $\beta_N \in \mathbb{K}^N$ se cumple que $f(\beta_N) = \alpha_{MN}\beta_N$.

Prueba. Sea $f' : \beta_N \mapsto \alpha_{MN}\beta_N$. Sabemos que f y f' son TLs. Si $i \in N$ entonces, por definición de la base canónica y del producto de una matriz por un vector tenemos la igualdad del recuadro. Luego f y f' coinciden en la base canónica y por extensión lineal ambas son la misma función. ■

$$\alpha_{MN}e_i = \sum_{j \in N} \alpha_{Mj}\delta_{ji} = \alpha_{Mi}$$

Ejercicio 64 Halle la matriz de la rotación con ángulo α en \mathbb{R}^2 . [190]

Composición de TLs y producto de matrices

3.18

La matriz de la composición de dos TLs es igual al producto de las matrices de las TLs.

Prueba. Sean $f \in \text{Mor}(\mathbb{K}^N, \mathbb{K}^M)$ y $g \in \text{Mor}(\mathbb{K}^M, \mathbb{K}^L)$ dos TLs. Sean α_{MN} y β_{LM} las matrices de f y g respectivamente. Para cualquier $\gamma_N \in \mathbb{K}^N$ y cualquier $i \in L$ tenemos

$$\begin{aligned} \beta_{iM}(\alpha_{MN}\gamma_N) &= \sum_{j \in M} \beta_{ij}(\alpha_{jN}\gamma_N) = \sum_{j \in M} \beta_{ij} \sum_{k \in N} \alpha_{jk}\gamma_k = \\ &= \sum_{k \in N} \sum_{j \in M} \beta_{ij}\alpha_{jk}\gamma_k = \sum_{k \in N} (\beta_{iM}\alpha_{Mk})\gamma_k = (\beta_{iM}\alpha_{MN})\gamma_N \end{aligned}$$

y por lo tanto $\beta_{LM}(\alpha_{MN}\gamma_N) = (\beta_{LM}\alpha_{MN})\gamma_N$. Como $\gamma_N \mapsto \beta_{LM}(\alpha_{MN}\gamma_N)$ es la TL $g \circ f$ entonces, tenemos $(g \circ f)(\gamma_N) = (\beta_{LM}\alpha_{MN})\gamma_N$ que es lo que se quería probar. ■

3.19

El producto de matrices es asociativo, distribuye por ambos lados con la suma de matrices y commuta con el producto por un escalar.

Prueba. Sean f, g y h TLs cuyas matrices son α_{MN} , β_{LM} y γ_{KL} respectivamente. La matriz de $(h \circ g) \circ f$ es $(\gamma_{KL}\beta_{LM})\alpha_{MN}$. La matriz de $h \circ (g \circ f)$ es $\gamma_{KL}(\beta_{LM}\alpha_{MN})$. Como la composición de TLs es asociativa tenemos $(h \circ g) \circ f = h \circ (g \circ f)$ y por lo tanto $(\gamma_{KL}\beta_{LM})\alpha_{MN} = \gamma_{KL}(\beta_{LM}\alpha_{MN})$. Esto prueba la asociatividad.

Las demás propiedades se prueban exactamente igual o sea, se desprenden de las respectivas propiedades de las TLs y de la proposición 3.18. ■

Ejercicio 65 Pruebe la asociatividad del producto de matrices directamente de la definición de producto o sea, sin usar TLs. [190]

Sección 3.4 Coordinatización de transformaciones lineales

320

El espacio de todas las \mathbb{N}^N -matrices es un álgebra isomorfa a $\text{End}(\mathbb{K}^N)$.

Prueba. Ya sabemos que $\mathbb{K}^{\mathbb{N}^N}$ es un espacio vectorial. El resultado anterior hace la mayor parte del trabajo necesario para mostrar que $\mathbb{K}^{\mathbb{N}^N}$ es un álgebra. Solo falta el neutro para el producto que es la matriz de la identidad en \mathbb{K}^N . Esta matriz es $\mathbb{I}_{\mathbb{N}^N}$ que cumple que $\mathbb{I}_{ij} = \delta_{ij}$ (el delta de Kronecker) y que la llamaremos **matriz identidad**.

Además, ya sabemos que la aplicación que a un OL en \mathbb{K}^N le hace corresponder su matriz es un isomorfismo de espacios vectoriales. La proposición 3.18 completa la tarea de demostrar que esta aplicación es un isomorfismo de álgebras. ■

Matrices inversas

Sea $f \in \text{Mor}(\mathbb{K}^N, \mathbb{K}^M)$ y α_{MN} la matriz de f . La función f es biyectiva si y solo si, existe la TL f^{-1} tal que $f \circ f^{-1} = \mathbb{I}(\mathbb{K}^N)$ y $f^{-1} \circ f = \mathbb{I}(\mathbb{K}^M)$. A la matriz de la TL f^{-1} se le llama **matriz inversa** de α_{MN} y se denota por α_{MN}^{-1} . De la proposición 3.18 obtenemos que la matriz inversa cumple que $\alpha_{MN}^{-1} \alpha_{MN} = \mathbb{I}_{NN}$ y $\alpha_{MN} \alpha_{MN}^{-1} = \mathbb{I}_{MM}$. Observese que el conjunto de índices de las columnas de α_{MN}^{-1} es M y no N . Análogamente, el conjunto de índices de los renglones de α_{MN}^{-1} es N y no M .

De la definición es inmediato que una matriz tiene inversa si y solo si su TL es un isomorfismo de espacios vectoriales. En particular los conjuntos de índices N y M tienen que tener el mismo cardinal ya que estos cardinales son las dimensiones del dominio y el codominio de esta TL. O sea, la matriz debe ser **cuadrada**. Ahora nos preocuparemos en traducir nuestro criterio de isomorfismo 3.14 al lenguaje de matrices.

321

Una matriz cuadrada α_{MN} tiene inversa si y solo si sus columnas son todas diferentes y son una base de \mathbb{K}^M .

Prueba. Sea α_{MN} una matriz cuadrada y $f \in \text{Mor}(\mathbb{K}^N, \mathbb{K}^M)$ su TL. La restricción de f a la base canónica de \mathbb{K}^N es la N -ada de las columnas de la matriz α_{MN} . Por 3.14 la función f tiene inversa si y solo si esta restricción es inyectiva (las columnas diferentes) y su imagen (el conjunto de columnas) es una base. ■

Es posible probar un criterio análogo al anterior substituyendo las columnas por los renglones. Sin embargo, su prueba aquí se nos haría innecesariamente complicada. Mejor lo dejaremos para el próximo capítulo donde esto será una facil consecuencia de un resultado mucho más importante.

Ejercicio 66 Sean f y g las rotaciones del plano \mathbb{R}^2 en los ángulos α y β respectivamente. Use el ejercicio 64 para hallar las matrices en la base canónica de f , g y $f \circ g$. Use 3.18 para hallar fórmulas para el seno y el coseno de la suma de dos ángulos. [190]

Ejercicio 67 ¿Cuál es la matriz inversa a la matriz de una rotación?

Ejercicio 68 Sea f el OL en \mathbb{R}^2 que deja fijo a $(1, 0)$ y manda $(0, 1)$ en $(1, 1)$. ¿Cuál es su matriz? ¿Cuál es la matriz inversa?

3.5 Cambios de base

Es usual en las aplicaciones que sea conveniente realizar ciertos cálculos en un sistema de coordenadas y después realizar otros cálculos en otro sistema de coordenadas. En el álgebra lineal estos cambios de coordenadas son lineales o sea, la transformación que lleva unas coordenadas a otras es una TL.

Cambios de base en un espacio vectorial

Sean V y N dos bases del espacio \mathfrak{E} . Conocemos los isomorfismos de coordinatización $\mathbb{K}^V \leftrightarrow \mathfrak{E} \leftrightarrow \mathbb{K}^N$. Nuestro problema ahora es: dada una V -ada β_V que son las coordenadas del vector x en la base V , ¿cómo hallar las coordenadas γ_N de x en la base N ? En este caso las letras V y N tienen el sentido de que V es la base “vieja” y que N es la base “nueva”.

Sea α_{NV} la matriz cuyas columnas son los vectores de la base V expresados en las coordenadas de N . O sea, para cualquier $v \in V$ tenemos la fórmula en el recuadro a la derecha. A la matriz α_{NV} se le llama **matriz de cambio de base** (de V a N). Esta matriz no es otra cosa que la matriz del isomorfismo $\mathbb{K}^V \rightarrow \mathfrak{E} \rightarrow \mathbb{K}^N$.

$$v = \sum_{i \in N} \alpha_{iv} i$$

3.22

Si β_V es la V -ada de las coordenadas de un vector en la base V entonces, $\alpha_{NV}\beta_V$ es la N -ada de las coordenadas del mismo vector en la base N .

Prueba. Descompongamos $x \in \mathfrak{E}$ en las dos bases. Tenemos, $x = \sum_{v \in V} \beta_v v = \sum_{i \in N} \gamma_i i$ y por lo tanto

$$x = \sum_{v \in V} \beta_v \left(\sum_{i \in N} \alpha_{iv} i \right) = \sum_{i \in N} \left(\sum_{v \in V} \alpha_{iv} \beta_v \right) i = \sum_{i \in N} \gamma_i i$$

De la unicidad de las coordenadas de cualquier vector en la base N obtenemos la igualdad $\sum_{v \in V} \alpha_{iv} \beta_v = \gamma_i$ que es la que se necesitaba demostrar. ■

Ejemplo. Queremos calcular las coordenadas de un vector $u = (x, y) \in \mathbb{R}^2$ en la base $N = \{\mathbf{a}_1, \mathbf{a}_2\}$ donde $\mathbf{a}_1 = (2, 3)$ y $\mathbf{a}_2 = (1, 2)$. Las coordenadas (x, y) son las coordenadas de u en la base canónica. Luego, $V = \{\mathbf{e}_1, \mathbf{e}_2\}$ es la base canónica y para construir la matriz α_{NV} de cambio de base necesitamos las coordenadas de V en la base N . Estas coordenadas se pueden hallar resolviendo dos sistemas de ecuaciones lineales pero es más sencillo usar la siguiente argumentación. Como un cambio de base es un isomorfismo entonces la matriz α_{NV} tiene inversa que es la matriz de cambio de

la base \mathbf{N} a la base \mathbf{V} . Las columnas de esta matriz ya las tenemos, son \mathbf{a}_1 y \mathbf{a}_2 . Luego, denotando \mathbf{p} y \mathbf{q} las coordenadas que buscamos, o sea, $\mathbf{u} = \mathbf{p}\mathbf{a}_1 + \mathbf{q}\mathbf{a}_2$ tenemos:

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ 2y - 3x \end{pmatrix}.$$

y en estos cálculos lo único que no conoce el lector es como calcular la matriz inversa. Pero, esto lo pospondremos hasta el próximo capítulo.

Cambios de base en el espacio de transformaciones lineales

Veamos como cambia la matriz de una TL cuando cambian las bases. Sean \mathbf{V} , \mathbf{N} dos bases del espacio \mathfrak{E} y \mathbf{W} , \mathbf{M} dos bases del espacio \mathfrak{F} . Conocemos los isomorfismos de coordinatización $\mathbb{K}^{\mathbf{W}\mathbf{V}} \leftrightarrow \text{Mor}(\mathfrak{E}, \mathfrak{F}) \leftrightarrow \mathbb{K}^{\mathbf{M}\mathbf{N}}$. El problema ahora es: dada una $\mathbf{W}\mathbf{V}$ -matriz $\alpha_{\mathbf{W}\mathbf{V}}$ que es la matriz de la TL $f : \mathfrak{E} \rightarrow \mathfrak{F}$ en las bases \mathbf{V} y \mathbf{W} , ¿cómo hallar la matriz $\beta_{\mathbf{M}\mathbf{N}}$ de f en las bases \mathbf{N} y \mathbf{M} ? Nuevamente, \mathbf{V}, \mathbf{W} son las bases “viejas” y \mathbf{M}, \mathbf{N} son las bases nuevas.

Sea $\gamma_{\mathbf{N}\mathbf{V}}$ la matriz de cambio de base de \mathbf{V} a \mathbf{N} en \mathfrak{E} . Sea $\lambda_{\mathbf{M}\mathbf{W}}$ la matriz de cambio de base de \mathbf{W} a \mathbf{M} en \mathfrak{F} . O sea, para cualquier $\mathbf{v} \in \mathbf{V}$ y cualquier $\mathbf{w} \in \mathbf{W}$ tenemos las fórmulas en el recuadro a la derecha. Estas matrices no son otra cosa que las matrices de los isomorfismos de coordinatización $\mathbb{K}^{\mathbf{V}} \rightarrow \mathfrak{E} \rightarrow \mathbb{K}^{\mathbf{N}}$ y $\mathbb{K}^{\mathbf{W}} \rightarrow \mathfrak{F} \rightarrow \mathbb{K}^{\mathbf{M}}$.

$$\mathbf{v} = \sum_{i \in \mathbf{N}} \gamma_{iv} \mathbf{i}$$

$$\mathbf{w} = \sum_{j \in \mathbf{M}} \lambda_{jw} \mathbf{j}$$

Si $\alpha_{\mathbf{W}\mathbf{V}}$ **es la matriz de** f **en las bases** \mathbf{V} **y** \mathbf{W} **entonces,**
 $\lambda_{\mathbf{M}\mathbf{W}} \alpha_{\mathbf{W}\mathbf{V}} \gamma_{\mathbf{N}\mathbf{V}}^{-1}$ **es la matriz de** f **en las bases** \mathbf{N} **y** \mathbf{M} .

Prueba. Las columnas de $\alpha_{\mathbf{W}\mathbf{V}}$ son las imágenes por f de la base \mathbf{V} expresadas en la base \mathbf{W} o sea, $\forall \mathbf{v} \in \mathbf{V}$ se cumple la fórmula del recuadro a la derecha. Denotemos por $\beta_{\mathbf{M}\mathbf{N}}$ la matriz

$$f(\mathbf{v}) = \sum_{w \in \mathbf{W}} \alpha_{wv} \mathbf{w}$$

$f(\mathbf{i}) = \sum_{j \in \mathbf{M}} \beta_{ji} \mathbf{j}$ de f en las bases \mathbf{N}, \mathbf{M} . Las columnas de $\beta_{\mathbf{M}\mathbf{N}}$ son las imágenes por f de la base \mathbf{N} expresadas en la base \mathbf{M} . O sea, para cualquier $\mathbf{i} \in \mathbf{N}$ se cumple la fórmula del recuadro a la izquierda.

Substituyendo en la fórmula de la derecha las fórmulas que definen las matrices $\lambda_{\mathbf{M}\mathbf{W}}$ y $\gamma_{\mathbf{N}\mathbf{V}}$ obtenemos

$$\sum_{i \in \mathbf{N}} \gamma_{iv} f(\mathbf{i}) = \sum_{j \in \mathbf{M}} \left(\sum_{w \in \mathbf{W}} \lambda_{jw} \alpha_{wv} \right) \mathbf{j}$$

y en esta igualdad substituimos $f(\mathbf{i})$ por la fórmula de la izquierda para obtener

$$\sum_{j \in \mathbf{M}} \left(\sum_{i \in \mathbf{N}} \beta_{ji} \gamma_{iv} \right) \mathbf{j} = \sum_{j \in \mathbf{M}} \left(\sum_{w \in \mathbf{W}} \lambda_{jw} \alpha_{wv} \right) \mathbf{j}.$$

De la unicidad de las coordenadas de cualquier vector en la base \mathbf{M} obtenemos que para cualesquiera $\mathbf{j} \in \mathbf{M}$ y $\mathbf{v} \in \mathbf{V}$ se cumple que $\sum_{i \in \mathbf{N}} \beta_{ji} \gamma_{iv} = \sum_{w \in \mathbf{W}} \lambda_{jw} \alpha_{wv}$ y por lo tanto $\beta_{\mathbf{M}\mathbf{N}} \gamma_{\mathbf{N}\mathbf{V}} = \lambda_{\mathbf{M}\mathbf{W}} \alpha_{\mathbf{W}\mathbf{V}}$. Como $\gamma_{\mathbf{N}\mathbf{V}}$ es la matriz de un isomorfismo entonces, $\gamma_{\mathbf{N}\mathbf{V}}$ tiene inversa por lo que podemos despejar $\beta_{\mathbf{M}\mathbf{N}}$. ■

$$\begin{array}{ccc} \mathbb{K}^V & \xrightarrow{\alpha_{VV}} & \mathbb{K}^W \\ \gamma_{NV} \downarrow & & \downarrow \lambda_{MW} \\ \mathbb{K}^N & \xrightarrow{\beta_{MN}} & \mathbb{K}^M \end{array}$$

La proposición anterior la podemos interpretar gráficamente de la siguiente manera. Las matrices α_{VV} , β_{MN} , γ_{NV} , y λ_{MW} son las matrices de TLs entre espacios como se muestra en el diagrama a la izquierda. Se dice que **un diagrama de funciones es conmutativo** si cualesquiera dos caminos dirigidos entre dos cualesquiera conjuntos son funciones iguales.

En nuestro caso, el que el diagrama a la izquierda sea conmutativo lo quiere decir es que $\beta_{MN}\gamma_{NV} = \lambda_{MW}\alpha_{VV}$.

Cambios de base en el espacio de operadores lineales

Si $f \in \text{End}(\mathfrak{E}) = \text{Mor}(\mathfrak{E}, \mathfrak{E})$ es un OL entonces, no tiene sentido escoger bases diferentes para el dominio y el codominio ya que estos son iguales. Sean V, N dos bases de \mathfrak{E} . El problema ahora es: dada una WW -matriz α_{VV} que es la matriz del OL f en la base V , hallar la matriz β_{NN} de f en la base N . Sea γ_{NV} la matriz de cambio de base de V a N en \mathfrak{E} . En este caso, el diagrama es el de la derecha. Ya no hay que probar la conmutatividad de este ya que él, es un caso particular del anterior. Luego, $\beta_{NN}\gamma_{NV} = \gamma_{NV}\alpha_{VV}$ y despejando obtenemos que $\beta_{NN} = \gamma_{NV}\alpha_{VV}\gamma_{NV}^{-1}$.

$$\begin{array}{ccc} \mathbb{K}^V & \xrightarrow{\alpha_{VV}} & \mathbb{K}^V \\ \gamma_{NV} \downarrow & & \downarrow \gamma_{NV} \\ \mathbb{K}^N & \xrightarrow{\beta_{NN}} & \mathbb{K}^N \end{array}$$

Ejemplo. Sea f la TL del plano \mathbb{R}^2 que tiene la matriz del re-cuadro a la izquierda en la base $V = \{\mathbf{a}_1, \mathbf{a}_2\}$ donde $\mathbf{a}_1 = (2, 3)$ y $\mathbf{a}_2 = (1, 2)$. ¿Cuál será la matriz de f en la base canónica? La matriz de cambio de base a la base canónica es la que tiene como columnas a los vectores \mathbf{a}_1 y \mathbf{a}_2 . Luego, la matriz de f en la base canónica es

$$\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} \cos \alpha + 8 \sin \alpha & -5 \sin \alpha \\ 13 \sin \alpha & \cos \alpha - 8 \sin \alpha \end{pmatrix}$$

y esto es una advertencia de que un operador lineal puede tener en una base una matriz que es igual a la de la rotación en la base canónica y sin embargo no es una rotación.

3.6 El núcleo y la imagen de una TL

En esta sección queremos ver que para describir todas las transformaciones lineales nos es suficiente conocer las inmersiones, las proyecciones y los isomorfismos. Después, veremos interesantes consecuencias de este resultado.

Definiciones

Para esto comenzaremos con dos definiciones fundamentales. Sea $f : \mathfrak{E} \rightarrow \mathfrak{F}$ una TL. Al conjunto $\{y \in \mathfrak{F} \mid \exists x \in \mathfrak{E} \quad f(x) = y\}$ se le llama **imagen** de la TL. La imagen de f se denotará por $\text{Im } f$. Al conjunto $\{x \in \mathfrak{E} \mid f(x) = 0\}$ se le llama **núcleo** de la TL. El núcleo de f se denotará por $\ker f$. Esta notación es debido a que en inglés núcleo es

“kernel”. La imagen es el conjunto de los vectores en el codominio que tienen preimagen y el núcleo es el conjunto de los vectores en el dominio cuya imagen es el vector $\mathbf{0}$.

3.24

La imagen y el núcleo de una TL son subespacios.

Prueba. Sean \mathbf{x}, \mathbf{y} vectores en $\text{Im } f$ y $\lambda \in \mathbb{K}$. Por definición existen \mathbf{a}, \mathbf{b} tales que $f(\mathbf{a}) = \mathbf{x}$, $f(\mathbf{b}) = \mathbf{y}$. Como f es lineal tenemos $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}) = \mathbf{x} + \mathbf{y}$ y además $f(\lambda\mathbf{a}) = \lambda f(\mathbf{a}) = \lambda\mathbf{x}$. Esto quiere decir que $\text{Im } f$ es un subespacio.

Sean \mathbf{a}, \mathbf{b} vectores en $\ker f$ y $\lambda \in \mathbb{K}$. Tenemos $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$ y $f(\lambda\mathbf{a}) = \lambda f(\mathbf{a}) = \lambda\mathbf{0} = \mathbf{0}$. Esto quiere decir que $\ker f$ es un subespacio. ■



El núcleo y la imagen de una TL son subespacios de espacios diferentes. Si $f : \mathfrak{E} \rightarrow \mathfrak{F}$ entonces $\ker f \subset \mathfrak{E}$ e $\text{Im } f \subset \mathfrak{F}$. Solamente en el caso que la TL es un OL o sea, cuando $\mathfrak{E} = \mathfrak{F}$ el núcleo y la imagen son subespacios del mismo espacio. Sin embargo, como veremos más adelante, en este caso pasan cosas raras ya que, aunque estos subespacios tienen dimensiones complementarias ellos NO siempre son complementarios.

Transformaciones lineales con núcleo trivial

Observese que, como para cualquier TL se tiene que $f(\mathbf{0}) = \mathbf{0}$ entonces el vector cero siempre es un elemento del núcleo. Si el núcleo solo contiene al vector cero se dice que f tiene **núcleo trivial**. Cualquier TL lineal inyectiva tiene núcleo trivial ya que en este caso la preimagen de cualquier vector es única. Lo importante es que el recíproco también es cierto.

3.25

Una TL es inyectiva si y solo si su núcleo es trivial.

Prueba. Sea f una TL. Sean \mathbf{x}, \mathbf{y} dos vectores en el dominio de f . Tenemos

$$(f(\mathbf{x}) = f(\mathbf{y})) \Leftrightarrow (f(\mathbf{x}) - f(\mathbf{y}) = \mathbf{0}) \Leftrightarrow (f(\mathbf{x} - \mathbf{y}) = \mathbf{0}) \Leftrightarrow (\mathbf{x} - \mathbf{y} \in \ker f)$$

y por lo tanto el que existan dos vectores diferentes cuyas imágenes sean iguales es equivalente a la existencia de un vector no nulo en el núcleo de la TL. ■

Descomposición de transformaciones lineales

Ahora, demostraremos el resultado prometido al principio de la sección. Sea $f : \mathfrak{E} \rightarrow \mathfrak{F}$ una TL. Sea \mathfrak{K} un subespacio complementario cualquiera pero fijo del $\ker f$. Sea $f_{|\mathfrak{K}}$ la restricción de f al subespacio \mathfrak{K} . Denotemos por $i : \text{Im } f \hookrightarrow \mathfrak{F}$ a la inmersión del subespacio $\text{Im } f$ en \mathfrak{F} . Finalmente, denotemos por $\pi_{|\mathfrak{K}} : \mathfrak{E} \twoheadrightarrow \mathfrak{K}$ la proyección de \mathfrak{E} a \mathfrak{K} a lo largo del $\ker f$.

Teorema de Descomposición de una TL

3.26

$$f = i \circ f_{\mathfrak{K}} \circ \pi_{\mathfrak{K}} \text{ y } f_{\mathfrak{K}} \text{ es un isomorfismo.}$$

Prueba. Sea x un vector arbitrario en el dominio de f . Por 2.29 (página 53) existen unos únicos vectores $\mathbf{a} \in \mathfrak{K}$, $\mathbf{b} \in \ker f$ tales que $x = \mathbf{a} + \mathbf{b}$. De aquí obtenemos

$$(i \circ f_{\mathfrak{K}} \circ \pi_{\mathfrak{K}})(x) = i(f_{\mathfrak{K}}(\pi_{\mathfrak{K}}(x))) = i(f_{\mathfrak{K}}(\mathbf{a})) = f(\mathbf{a}) = f(\mathbf{a}) + f(\mathbf{b}) = f(x)$$

y con esto queda probado que $f = i \circ f_{\mathfrak{K}} \circ \pi_{\mathfrak{K}}$.

Para probar que $f_{\mathfrak{K}}$ es un isomorfismo sea $f(x) \in \operatorname{Im} f$. Por 2.29 existen unos únicos vectores $\mathbf{a} \in \mathfrak{K}$, $\mathbf{b} \in \ker f$ tales que $x = \mathbf{a} + \mathbf{b}$. Como $f_{\mathfrak{K}}(\mathbf{a}) = f(\mathbf{a}) = f(\mathbf{a}) + f(\mathbf{b}) = f(x)$ entonces $f_{\mathfrak{K}}$ es sobreyectiva. Si $f_{\mathfrak{K}}(\mathbf{a}) = f_{\mathfrak{K}}(\mathbf{b})$ entonces, $f_{\mathfrak{K}}(\mathbf{a} - \mathbf{b}) = \mathbf{0}$. Como $(\mathbf{a} - \mathbf{b}) \in \mathfrak{K} \cap \ker f$ entonces, $\mathbf{a} - \mathbf{b} = \mathbf{0}$ o sea $\mathbf{a} = \mathbf{b}$. Luego, $f_{\mathfrak{K}}$ es inyectiva. ■

Este teorema lo podemos visualizar más fácilmente si observamos el diagrama de la derecha. La primera afirmación del Teorema de Descomposición de una TL lo que dice es que este diagrama es comunitativo. La segunda afirmación nos dice que $f_{\mathfrak{K}}$ es un isomorfismo de espacios vectoriales.

$$\begin{array}{ccc} \mathfrak{E} & \xrightarrow{f} & \mathfrak{F} \\ \pi_{\mathfrak{K}} \downarrow & & \uparrow i \\ \mathfrak{K} & \xrightarrow{f_{\mathfrak{K}}} & \operatorname{Im} f \end{array}$$

3.27

Para cualquier transformación lineal $f : \mathfrak{E} \rightarrow \mathfrak{F}$,
 $\dim \mathfrak{E} = \dim \ker f + \dim \operatorname{Im} f$.

Prueba. Por el teorema anterior $\operatorname{Im} f$ es isomorfo a un complementario de $\ker f$. ■

Un criterio de isomorfismo

Recordemos un sencillo resultado de teoría de conjuntos: toda función inyectiva de un conjunto finito en otro con el mismo número de elementos es biyectiva. De hecho, esto lo usamos en el primer capítulo para probar que \mathbb{Z}_p es un campo para p primo. El objetivo ahora, es mostrar que “exactamente” el mismo resultado (simple pero muy útil) se cumple para espacios vectoriales de dimensión finita. Esto es una consecuencia del Teorema de Descomposición de una TL (3.26).

3.28

Sean \mathfrak{E} y \mathfrak{F} dos espacios de dimensiones finitas e iguales.

Una TL $f : \mathfrak{E} \rightarrow \mathfrak{F}$ es inyectiva si y solo si es sobreyectiva.

Prueba. Por 3.27 tenemos $\dim \ker f = \dim \mathfrak{E} - \dim \operatorname{Im} f$. Si f es sobreyectiva entonces, $\mathfrak{F} = \operatorname{Im} f$. Por hipótesis $\dim \mathfrak{F} = \dim \mathfrak{E}$. De aquí, debido a que todas las dimensiones son finitas, $\dim \ker f = 0$ y por lo tanto $\ker f = \{\mathbf{0}\}$. De 3.25 concluimos que f es inyectiva.

Si f es inyectiva entonces, la función $f : \mathfrak{E} \rightarrow \operatorname{Im} f$ es un isomorfismo y por lo tanto $\dim \mathfrak{E} = \dim \operatorname{Im} f$. Por hipótesis $\dim \mathfrak{F} = \dim \mathfrak{E}$. Como $\operatorname{Im} f$ es un subespacio de \mathfrak{F} entonces, aplicando 2.17 (página 42) obtenemos $\mathfrak{F} = \operatorname{Im} f$. ■

Como consecuencia de este resultado probaremos que para comprobar que dos operadores lineales f y g en un espacio de dimensión finita son el inverso uno del otro, solo es necesario comprobar una de las dos igualdades $g \circ f = \mathbb{I}$ o $f \circ g = \mathbb{I}$.

3.29

Sean $f, g : \mathfrak{E} \rightarrow \mathfrak{E}$ dos OLs de un espacio finito dimensional. Entonces, $f \circ g = \mathbb{I}$ si y solo si $g \circ f = \mathbb{I}$.

Prueba. La función identidad es sobreyectiva. Luego, si $f \circ g = \mathbb{I}$ entonces, f es sobreyectiva. Por 3.28 f es inyectiva y por lo tanto tiene inversa f^{-1} . Componiendo con f^{-1} obtenemos $f^{-1} \circ f \circ g = f^{-1} \circ \mathbb{I}$ y por lo tanto $g = f^{-1}$. ■

Descomposición canónica de transformaciones lineales

Para aplicar el Teorema de Descomposición de una TL necesitamos escoger (ya que hay muchos) un subespacio complementario \mathfrak{K} del núcleo de f . Ya vimos (véase 2.34) que cualquier tal subespacio es isomorfo al cociente $\mathfrak{E}/\ker f$. Queremos substituir \mathfrak{K} por $\mathfrak{E}/\ker f$ en el Teorema de Descomposición de una TL para así obtener otra versión del mismo que no dependa de escoger nada, o sea que sea canónico.

En el Teorema de Descomposición de una TL están involucradas tres funciones: la proyección $\pi_{\mathfrak{K}}$ (que es sobreyectiva), la restricción $f|_{\mathfrak{K}}$ (que es biyectiva) y la inmersión i (que es inyectiva). Esta última no depende de \mathfrak{K} y podemos quedarnos con ella. Así que todas nuestras incógnitas están representadas en el diagrama de la derecha.

$$\begin{array}{ccc} \mathfrak{E} & \xrightarrow{f} & \mathfrak{F} \\ ? \downarrow & & \uparrow i \\ \mathfrak{E}/\ker f & \longleftrightarrow & \text{Im } f \end{array}$$

¿Cuáles funciones podemos escoger para nuestras incógnitas? No tenemos muchas variantes. El espacio cociente $\mathfrak{E}/\ker f$ está formado por todos los subespacios afines $\ker f + \mathbf{x}$. El espacio $\text{Im } f$ está formado por todas las imágenes $f(\mathbf{x})$. Luego, la única posible respuesta a nuestra pregunta son las funciones definidas en el recuadro a la izquierda.

La primera de estas funciones tiene dominio \mathfrak{E} , codominio $\mathfrak{E}/\ker f$, se le llama función **natural** y se denota por “**nat**”. La función natural es una TL ya que

$$\text{nat}(\mathbf{x} + \mathbf{y}) = \ker f + \mathbf{x} + \mathbf{y} = \ker f + \mathbf{x} + \ker f + \mathbf{y} = \text{nat}(\mathbf{x}) + \text{nat}(\mathbf{y})$$

$$\text{nat}(\lambda \mathbf{x}) = \ker f + \lambda \mathbf{x} = \lambda \ker f + \lambda \mathbf{x} = \lambda (\ker f + \mathbf{x}) = \lambda \text{nat}(\mathbf{x})$$

y además es evidentemente sobreyectiva.

La segunda de estas funciones es nuestra preocupación fundamental ya que tenemos que probar que es un isomorfismo. Esta función tiene dominio $\text{Im } f$, codominio $\mathfrak{E}/\ker f$ y la denotaremos por g . La función g es una TL ya que

$$g(f(\mathbf{x}) + f(\mathbf{y})) = g(f(\mathbf{x} + \mathbf{y})) = \ker f + \mathbf{x} + \mathbf{y} = g(f(\mathbf{x})) + g(f(\mathbf{y}))$$

$$g(\lambda f(\mathbf{x})) = g(f(\lambda \mathbf{x})) = \ker f + \lambda \mathbf{x} = \lambda (\ker f + \mathbf{x}) = \lambda g(f(\mathbf{x}))$$

y es también evidentemente sobreyectiva.

¿Qué quiere decir que g es inyectiva? Lo que quiere decir es que si los subespacios

afines $\ker f + \mathbf{x}$ y $\ker f + \mathbf{y}$ son el mismo entonces, $f(\mathbf{x}) = f(\mathbf{y})$. Como para un subespacio afín E paralelo al $\ker f$ se cumple que $(E = \ker f + \mathbf{x}) \Leftrightarrow (\mathbf{x} \in E)$ entonces, la inyectividad de g es equivalente a que la función f sea constante en cualquier subespacio afín paralelo al $\ker f$.

330 Los subespacios afines paralelos a $\ker f$ son precisamente los conjuntos de vectores en que la función f es constante.

Prueba. Tenemos

$$(\mathbf{y} \in \ker f + \mathbf{x}) \Leftrightarrow (\exists \mathbf{a} \in \ker f \mid \mathbf{y} = \mathbf{a} + \mathbf{x}) \Leftrightarrow (f(\mathbf{y} - \mathbf{x}) = 0) \Leftrightarrow (f(\mathbf{y}) = f(\mathbf{x}))$$

lo que nos convence de la validéz del resultado. ■

Luego, g es un isomorfismo y por lo tanto tiene un isomorfismo inverso que denotaremos por f' . El isomorfismo f' es el que a un subespacio afín $\ker f + \mathbf{x}$ le hace corresponder $f(\mathbf{x})$. Así completamos nuestro diagrama como en el recuadro a la derecha. Solo nos falta demostrar que este es commutativo. Sin embargo esto es muy fácil porque la composición de funciones

$$\mathbf{x} \xrightarrow{\text{nat}} \ker f + \mathbf{x} \xrightarrow{f'} f(\mathbf{x}) \xrightarrow{i} f(\mathbf{x})$$

es evidentemente igual a la función f .

$$\begin{array}{ccc} \mathfrak{E} & \xrightarrow{f} & \mathfrak{F} \\ \text{nat} \downarrow & & \uparrow i \\ \mathfrak{E}/\ker f & \xrightarrow{f'} & \text{Im } f \end{array}$$

3.7 Trasformaciones semilineales y coalineaciones



Una función $f : \mathfrak{E} \rightarrow \mathfrak{F}$ de espacios vectoriales sobre \mathbb{K} se le llama *transformación semilineal* si existe un automorfismo $\lambda \mapsto \bar{\lambda}$ del campo \mathbb{K} tal que $\forall \mathbf{a}, \mathbf{b} \in \mathfrak{E}$ y $\forall \lambda \in \mathbb{K}$ se cumplen las propiedades del recuadro.

$$\begin{aligned} f(\mathbf{a} + \mathbf{b}) &= f(\mathbf{a}) + f(\mathbf{b}) \\ f(\lambda \mathbf{a}) &= \bar{\lambda} f(\mathbf{a}) \end{aligned}$$

Observese que las trasformaciones lineales son semilineales usando como automorfismo del campo la función identidad. Para construir otras trasformaciones semilineales necesitamos automorfismos del campo que no sean la identidad. El ejemplo más importante es la conjugación $\mathbf{a} + b\mathbf{i} \mapsto \mathbf{a} - b\mathbf{i}$ en el campo de los números complejos.

Ejercicio 69 Pruebe que para una transformación semilineal no nula f el correspondiente automorfismo del campo es único. [190]

Trasformaciones semilineales reales

En el caso del campo de los números reales no hay trasformaciones semilineales que no sean lineales debido al siguiente resultado:

3.31

El único automorfismo de \mathbb{R} es la identidad.

Prueba. Sea f un automorfismo de \mathbb{R} . Supongamos que $x > 0$ y sea $y = \sqrt{x}$ entonces, $f(x) = f(y^2) = f(y)^2 > 0$. En particular si $x = b - a > 0$ entonces $f(b - a) = f(b) - f(a) > 0$. En otras palabras, f es monótona $b > a \Rightarrow f(b) > f(a)$. Como f^{-1} también es un automorfismo entonces, tambien es monótona. Luego, f commuta con el supremo y el ínfimo (véase el ejercicio 70).

Como $f(1) = 1$ y 1 es generador del grupo aditivo de \mathbb{Z} obtenemos que f es la identidad en \mathbb{Z} . Como $f(a/b) = f(a)/f(b)$ obtenemos que f es la identidad en \mathbb{Q} . Sea x un irracional y denotemos por A el conjunto de todos los racionales menores que x . Sabemos que $x = \sup A$ y por lo tanto $f(x) = f(\sup A) = \sup f(A) = \sup A = x$. ■

Ejercicio 70 Sean R un conjunto ordenado, $f : R \rightarrow R$ una biyección tal que f y f^{-1} son monótonas y $A \subseteq R$ tal que $\sup A$ existe. Pruebe que $f(\sup A) = \sup f(A)$. [190]

Ejercicio 71 Sea $f \neq \text{Id}$ un automorfismo de \mathbb{C} . Pruebe que las siguientes afirmaciones son equivalentes: 1. f es la conjugación compleja. 2. f es continua. 3. f es la identidad en \mathbb{R} . [190]

Propiedades de las transformaciones semilineales

Al igual que las TL las transformaciones semilineales preservan subespacios.

3.32

Toda trasformación semilineal transforma subespacios en subespacios.

Prueba. Sea $f : E \rightarrow F$ una trasformación semilineal y \mathfrak{F} un subespacio de E . Sean $a, b \in \mathfrak{F}$ y $\alpha \in \mathbb{K}$. Tenemos $f(a + b) = f(a) + f(b)$ por lo que $f(\mathfrak{F})$ es cerrado para la suma. Sea $\lambda \in \mathbb{K}$ tal que $\bar{\lambda} = \alpha$. Tenemos $\alpha f(a) = \bar{\lambda} f(a) = f(\lambda a)$ o sea $f(E)$ es cerrado para el producto por escalares. ■

3.33

Toda trasformación semilineal transforma subespacios afines en subespacios afines.

Prueba. Sea $f : E \rightarrow F$ una trasformación semilineal y \mathfrak{F} un subespacio de E . Si $\mathfrak{F} + x$ es un subespacio afín entonces, $f(\mathfrak{F} + x) = f(\mathfrak{F}) + f(x)$ que es también un subespacio afín puesto que por el resultado anterior $f(\mathfrak{F})$ es un subespacio. ■

Automorfismos semilineales.

A diferencia de las TL las transformaciones semilineales no forman un espacio vectorial: la suma de transformaciones semilineales no necesariamente es semilineal. A las transformaciones semilineales $f : \mathfrak{E} \rightarrow \mathfrak{E}$ biyectivas las llamaremos **automorfismos semilineales**.

3.34

La composición de automorfismos semilineales es un automorfismo semilineal.

Prueba. Sean f y g dos automorfismos semilineales. De la misma manera que para los operadores lineales se prueba que $(f \circ g)(\mathbf{a} + \mathbf{b}) = (f \circ g)(\mathbf{a}) + (f \circ g)(\mathbf{b})$. Sean $\lambda \mapsto \tilde{\lambda}$ y $\lambda \mapsto \bar{\lambda}$ los automorfismos del campo correspondientes a f y g respectivamente. Tenemos que $(f \circ g)(\lambda \mathbf{a}) = f(\bar{\lambda} \mathbf{a}) = \tilde{\lambda} \mathbf{a}$ y la prueba termina al observar que $\lambda \mapsto \tilde{\lambda}$ siendo una composición de automorfismos del campo es automorfismo del campo. ■

Ejercicio 72 Sea $\lambda \mapsto \bar{\lambda}$ un automorfismo del campo \mathbb{K} . Pruebe que la transformación $\mathbb{K}^n \ni (x_1, \dots, x_n) \mapsto (\bar{x}_1, \dots, \bar{x}_n) \in \mathbb{K}^n$ es un automorfismo semilineal. A tales automorfismos semilineales los llamaremos *estandar*. Pruebe que toda transformación semilineal de \mathbb{K}^n en \mathbb{K}^n es la composición de un automorfismo semilineal estandar con una TL. [191]

3.35

La inversa de un automorfismo semilineal es un automorfismo semilineal..

Prueba. Sea $f : \mathfrak{E} \rightarrow \mathfrak{E}$ un automorfismo semilineal. Sean $\lambda \in \mathbb{K}$ y $\mathbf{x}, \mathbf{y} \in \mathfrak{E}$. Sean $\mathbf{a}, \mathbf{b} \in \mathfrak{E}$ tales que $f(\mathbf{a}) = \mathbf{x}$, $f(\mathbf{b}) = \mathbf{y}$. Tenemos

$$\begin{aligned} f^{-1}(\mathbf{x} + \mathbf{y}) &= f^{-1}(f(\mathbf{a}) + f(\mathbf{b})) = f^{-1}(f(\mathbf{a} + \mathbf{b})) = \mathbf{a} + \mathbf{b} = f^{-1}(\mathbf{x}) + f^{-1}(\mathbf{y}) \\ f^{-1}(\bar{\lambda} \mathbf{x}) &= f^{-1}(\bar{\lambda} f(\mathbf{a})) = f^{-1}(f(\bar{\lambda} \mathbf{a})) = \bar{\lambda} \mathbf{a} = \lambda f^{-1}(\mathbf{x}) \end{aligned}$$

solo queda observar que la función $\bar{\lambda} \mapsto \lambda$ es un automorfismo de \mathbb{K} . ■

Estos dos últimos resultados significan que el conjunto de todos los automorfismos semilineales forman un grupo respecto a la composición de funciones.

Coalineaciones

Una biyección $f : \mathfrak{E} \rightarrow \mathfrak{E}$ en un espacio vectorial se le llama **coalineación** si la imagen de cualquier subespacio afín es un subespacio afín y la preimagen de cualquier subespacio afín es un subespacio afín de \mathfrak{E} . En otras palabras, tanto f como f^{-1} transforman subespacios afines en subespacios afines. Obviamente, si f es una coalineación entonces f^{-1} también lo es. El siguiente resultado nos dice que la definición de

Sección 3.7 Trasformaciones semilineales y coalineaciones

coalineación es posible hacerla de diversas maneras.

3.36

Sea $f : \mathfrak{E} \rightarrow \mathfrak{F}$ una biyección entre dos espacios afines. Entonces, las siguientes afirmaciones son equivalentes:

1. f y f^{-1} transforman subespacios afines en subespacios afines.
2. f y f^{-1} transforman generadores afines en generadores afines.
3. f y f^{-1} transforman bases afines en bases afines.
4. f y f^{-1} transforman conjuntos AI en conjuntos AI.
5. f y f^{-1} comutan con la cerradura afín.

Prueba. (1 \Rightarrow 2) Sea A un conjunto generador afín de \mathfrak{E} y $B = f(A)$. Sea $C = f^{-1}(B)$ que es un subespacio afín por ser la preimagen de un subespacio. Tenemos, $B \subseteq [B]$ y por lo tanto $A = f^{-1}(B) \subseteq f^{-1}([B]) = C$. Luego $\mathfrak{E} = [A] \subseteq C$ y por lo tanto $C = \mathfrak{E}$. De aquí $f(\mathfrak{E}) = [B]$ y como f es sobreyectiva tenemos que B es generador. Por simetría, si B es generador entonces $f^{-1}(B)$ también lo es.

(2 \Rightarrow 3) Sea ahora A una base afín de \mathfrak{E} y $B = f(A)$. Sabemos que B es generador afín. Si B no fuera AI entonces existiría b tal que $B \setminus b$ es generador y por lo tanto, tendríamos que $f^{-1}(B \setminus b) = A \setminus f^{-1}(b)$ es generador afín. Esto contradeciría que A es una base afín. Por simetría, si B es una base afín entonces $f^{-1}(B)$ también lo es.

(3 \Rightarrow 4) Sea ahora A un conjunto AI. Sea A' una base afín que lo contiene. Sabemos que $f(A')$ es una base afín. Como $f(A) \subseteq f(A')$ tenemos que $f(A)$ es AI. Por simetría, si B es AI entonces $f^{-1}(B)$ también lo es.

(4 \Rightarrow 1) Sea A una base afín del subespacio afín $[A]$. Como A es AI entonces $B = f(A)$ también lo es. Si $b \in [A]$ entonces $A \cup b$ es AD y por lo tanto $B \cup f(b)$ también lo es. Luego, $f(b) \in [B]$ y por lo tanto $f[A] \subseteq [B]$. Si $b \in [B]$ entonces $B \cup b$ es AD y por lo tanto $A \cup f^{-1}(b)$ también lo es. Luego, $f^{-1}(b) \in [A]$ y por lo tanto $[B] \subseteq f[A]$. Esto significa que f transforma el subespacio $[A]$ en el subespacio $[B]$. Por simetría, f^{-1} también transforma subespacios en subespacios.

(5 \Rightarrow 1) Si $f[A] = [f(A)]$ entonces la imagen del subespacio afín $[A]$ es el subespacio afín $[f(A)]$. O sea, f trasforma subespacios en subespacios. Por simetría, f^{-1} también transforma subespacios en subespacios.

(1 \Rightarrow 5) Sea f una coalineación. Sea A un conjunto de puntos. Como f transforma subespacios afines en subespacios afines la restricción de f a $[A]$ es una coalineación del espacio afín $[A]$ en el espacio afín $f[A]$. Como esta restricción transforma generadores en generadores tenemos que $f(A)$ es generador de $f[A]$. Luego $f[A] = [f(A)]$. Para ver de que f^{-1} comuta con la cerradura afín usamos que f^{-1} es una coalineación. ■

Ejercicio 73 Sea A un conjunto con un operador de cerradura y $f : A \rightarrow A$ una biyección. Si f y f^{-1} commutan con el operador de cerradura entonces f es una isotonía del conjunto ordenado de cerrados.

Estructura de las coalineaciones

La composición de coalineaciones de un espacio afín es una coalineación. Lo mismo sucede con la inversa de una coalineación. Luego el conjunto de todas las coalineaciones de un espacio afín \mathfrak{F} es un grupo respecto a la composición de funciones.

Si $\dim \mathfrak{E} = 1$ entonces cualquier biyección de \mathfrak{E} en \mathfrak{E} es una coalineación ya que en este caso todos los subespacios afines son puntos y la condición de preservar puntos no es ninguna restricción.

Un importante subgrupo de colineaciones son las traslaciones o sea, las funciones de la forma $t_a : \mathfrak{F} \ni x \mapsto x + a \in \mathfrak{F}$ y de estas hay una para cada vector a en el espacio vectorial \mathfrak{F} . Como $t_a \circ t_b = t_{a+b}$ observamos que el subgrupo de traslaciones es isomorfo al grupo aditivo del espacio vectorial \mathfrak{F} .

Sea f una coalineación en \mathfrak{E} . Denotemos $a = f(\mathbf{0})$ y t_a la traslación $x \mapsto x + a$. Observemos que la coalineación $f' = t_{-a} \circ f$ es tal que $f'(\mathbf{0}) = \mathbf{0}$. Luego, cualquier coalineación es la composición $f = t_a \circ f'$ de una coalineación que preserva $\mathbf{0}$ seguida de una traslación.

Si f es un automorfismo semilineal entonces f transforma subespacios afines en subespacios afines y por lo tanto es una coalineación que preserva $\mathbf{0}$.

Ahora, comenzaremos a probar el resultado principal de esta sección: que en dimensión al menos 2 toda coalineación que preserva $\mathbf{0}$ es semilineal. En todo lo que sigue, \mathfrak{E} es un espacio vectorial de dimensión al menos 2 sobre el campo \mathbb{K} .

3.37 *Toda coalineacion en \mathfrak{E} preserva el paralelismo de rectas.*

Prueba. Sea f una coalineación y ℓ, ℓ' dos rectas paralelas diferentes. Como f es un automorfismo del conjunto ordenado de subespacios afines $\dim [\ell \cup \ell'] = \dim f[\ell \cup \ell'] = \dim [f(\ell) \cup f(\ell')]$ y por lo tanto $f(\ell)$ y $f(\ell')$ son coplanares.

También $\dim [\ell \cap \ell'] = \dim f[\ell \cap \ell'] = \dim [f(\ell) \cap f(\ell')]$ y por lo tanto $f(\ell)$ y $f(\ell')$ no se intersectan. ■

3.38 *Toda coalineacion en \mathfrak{E} que preserva $\mathbf{0}$ es un automorfismo del grupo aditivo de vectores.*

Prueba. Sea f una coalineación de \mathfrak{E} que preserva $\mathbf{0}$. Recordemos que para cualquier vector x el subespacio $\langle x \rangle$ es la recta por el origen que pasa por x . Y como f preserva $\mathbf{0}$ tenemos $f\langle x \rangle = \langle f(x) \rangle$ o sea, f también commuta con la cerradura lineal.

Sección 3.7 Trasformaciones semilineales y coalineaciones

Sean $\mathbf{a}, \mathbf{b} \in \mathfrak{E}$ dos vectores. Tenemos que probar que $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$. Esto es trivial si $\{\mathbf{a}, \mathbf{b}\}$ contiene a $\mathbf{0}$ por lo que podemos suponer que $\mathbf{a} \neq \mathbf{0}$ y $\mathbf{b} \neq \mathbf{0}$. Supongamos primero que $\{\mathbf{a}, \mathbf{b}\}$ es LI. Por la regla del paralelogramo sabemos que

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= (\langle \mathbf{a} \rangle + \mathbf{b}) \cap (\langle \mathbf{b} \rangle + \mathbf{a}), \\ f(\mathbf{a}) + f(\mathbf{b}) &= (\langle f(\mathbf{a}) \rangle + f(\mathbf{b})) \cap (\langle f(\mathbf{b}) \rangle + f(\mathbf{a})). \end{aligned}$$

Como f preserva el paralelismo $f(\langle \mathbf{a} \rangle + \mathbf{b})$ es una recta paralela a $f(\langle \mathbf{a} \rangle) = \langle f(\mathbf{a}) \rangle$ que pasa por $f(\mathbf{b})$, o sea $f(\langle \mathbf{a} \rangle + \mathbf{b}) = \langle f(\mathbf{a}) \rangle + f(\mathbf{b})$. Análogamente, $f(\langle \mathbf{b} \rangle + \mathbf{a}) = \langle f(\mathbf{b}) \rangle + f(\mathbf{a})$. Como f commuta con la intersección (el ínfimo) tenemos $f(\mathbf{a} + \mathbf{b}) = f(\langle \mathbf{a} \rangle + \mathbf{b}) \cap f(\langle \mathbf{b} \rangle + \mathbf{a}) = (\langle f(\mathbf{a}) \rangle + f(\mathbf{b})) \cap (\langle f(\mathbf{b}) \rangle + f(\mathbf{a})) = f(\mathbf{a}) + f(\mathbf{b})$ y esto concluye el caso de que $\{\mathbf{a}, \mathbf{b}\}$ es LI.

Es claro que si $\{\mathbf{a}, \mathbf{b}\}$ es LI entonces, también lo es $\{\mathbf{a} - \mathbf{b}, \mathbf{b}\}$ y por lo tanto

$$f(\mathbf{a}) = f(\mathbf{a} - \mathbf{b} + \mathbf{b}) = f(\mathbf{a} - \mathbf{b}) + f(\mathbf{b}).$$

Luego, si $\{\mathbf{a}, \mathbf{b}\}$ es LI entonces, $f(\mathbf{a} - \mathbf{b}) = f(\mathbf{a}) - f(\mathbf{b})$.

Supongamos que $\{\mathbf{a}, \mathbf{b}\}$ es LD. Entonces, $\langle \mathbf{a} \rangle = \langle \mathbf{b} \rangle$. Como $\dim \mathfrak{E} > 1$ existe $\mathbf{c} \notin \langle \mathbf{a} \rangle$ tal que los conjuntos $\{\mathbf{a}, \mathbf{c}\}, \{\mathbf{b}, \mathbf{c}\}$ son LI.

Si $\mathbf{b} \neq -\mathbf{a}$ entonces, $\{\mathbf{a} + \mathbf{c}, \mathbf{b} - \mathbf{c}\}$ es LI y por el caso anterior tenemos que

$$f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a} + \mathbf{c} + \mathbf{b} - \mathbf{c}) = f(\mathbf{a} + \mathbf{c}) + f(\mathbf{b} - \mathbf{c}) = f(\mathbf{a}) + f(\mathbf{b})$$

y en particular cuando $\mathbf{b} = \mathbf{a}$ obtenemos $f(2\mathbf{a}) = 2f(\mathbf{a})$.

Si $\mathbf{b} = -\mathbf{a}$ entonces, $\{\mathbf{a} + \mathbf{c}, \mathbf{b} + \mathbf{c}\}$ es LI y tenemos que

$$2f(\mathbf{c}) = f(2\mathbf{c}) = f(\mathbf{a} + \mathbf{c} + \mathbf{b} + \mathbf{c}) = f(\mathbf{a} + \mathbf{c}) + f(\mathbf{b} + \mathbf{c}) = f(\mathbf{a}) + f(\mathbf{b}) + 2f(\mathbf{c})$$

y por lo tanto $f(\mathbf{a}) + f(\mathbf{b}) = \mathbf{0} = f(\mathbf{0}) = f(\mathbf{a} + \mathbf{b})$. ■

Ejercicio 74 Complete la demostración del resultado anterior mostrando que si $\{\mathbf{a}, \mathbf{c}\}$ es LI, $\mathbf{b} = \rho \mathbf{a}$ y $\rho \neq -1$ entonces $\{\mathbf{a} + \mathbf{c}, \mathbf{b} - \mathbf{c}\}$ es un conjunto LI. [191]

 Si f es una coalineación en \mathfrak{E} que preserva $\mathbf{0}$ entonces, existe un automorfismo $\lambda \mapsto \bar{\lambda}$ del campo \mathbb{K} tal que $f(\lambda \mathbf{a}) = \bar{\lambda} f(\mathbf{a})$ para todo vector $\mathbf{a} \in \mathfrak{E}$.

Prueba. Sea $\mathbf{a} \in \mathfrak{E} \setminus \mathbf{0}$. Como f preserva $\mathbf{0}$ tenemos que $f(\langle \mathbf{a} \rangle) = \langle f(\mathbf{a}) \rangle$. Sabemos que $\alpha \mathbf{a} \in \langle \mathbf{a} \rangle$ y por lo tanto $f(\alpha \mathbf{a}) \in \langle f(\mathbf{a}) \rangle$. Como en $\langle f(\mathbf{a}) \rangle$ están precisamente los múltiplos de $f(\mathbf{a})$ existe un escalar que denotaremos α^a tal que $f(\alpha \mathbf{a}) = \alpha^a f(\mathbf{a})$. De esta manera está definida una función $\mathbb{K} \ni \alpha \mapsto \alpha^a \in \mathbb{K}$ que es biyectiva pues f es una biyección de $\langle \mathbf{a} \rangle$ en $\langle f(\mathbf{a}) \rangle$.

Queremos probar que α^a no depende de \mathbf{a} . O sea que para cualesquiera $\mathbf{a}, \mathbf{b} \in \mathfrak{E} \setminus \mathbf{0}$ tenemos que $\alpha^a = \alpha^b$. Supongamos que $\{\mathbf{a}, \mathbf{b}\}$ es LI. Usando 3.38 obtenemos que

$$\begin{aligned} \alpha^a f(\mathbf{a}) + \alpha^b f(\mathbf{b}) &= f(\alpha \mathbf{a}) + f(\alpha \mathbf{b}) = f(\alpha \mathbf{a} + \alpha \mathbf{b}) = \\ &= f(\alpha(\mathbf{a} + \mathbf{b})) = \alpha^{a+b} f(\mathbf{a} + \mathbf{b}) = \alpha^{a+b} f(\mathbf{a}) + \alpha^{a+b} f(\mathbf{b}) \end{aligned}$$

y como $\{f(\mathbf{a}), f(\mathbf{b})\}$ es LI obtenemos $\alpha^a = \alpha^{a+b} = \alpha^b$.

Supongamos que $\{\mathbf{a}, \mathbf{b}\}$ es LD entonces, como $\dim \mathfrak{E} > 1$ existe \mathbf{c} tal que $\{\mathbf{a}, \mathbf{c}\}$ y $\{\mathbf{c}, \mathbf{b}\}$ son dos conjuntos LI. Por el caso anterior $\alpha^{\mathbf{a}} = \alpha^{\mathbf{c}} = \alpha^{\mathbf{b}}$.

Como $\alpha^{\mathbf{a}}$ no depende de \mathbf{a} denotaremos $\bar{\alpha} = \alpha^{\mathbf{a}}$. Sabemos que para cualquier vector $\mathbf{a} \in \mathfrak{E}$ tenemos $f(\alpha\mathbf{a}) = \bar{\alpha}f(\mathbf{a})$. Solo falta ver que $\alpha \mapsto \bar{\alpha}$ es un automorfismo de \mathbb{K} .

Sea \mathbf{a} un vector no nulo. Usando 3.38 obtenemos que

$$\begin{aligned} (\overline{\alpha + \beta})f(\mathbf{a}) &= f((\alpha + \beta)\mathbf{a}) = f(\alpha\mathbf{a} + \beta\mathbf{a}) = f(\alpha\mathbf{a}) + f(\beta\mathbf{a}) = (\bar{\alpha} + \bar{\beta})f(\mathbf{a}) \\ (\overline{\alpha\beta})f(\mathbf{a}) &= f(\alpha\beta\mathbf{a}) = \bar{\alpha}f(\beta\mathbf{a}) = \bar{\alpha}\bar{\beta}f(\mathbf{a}) \end{aligned}$$

y como $f(\mathbf{a})$ es no nulo obtenemos $\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}$ y $\overline{\alpha\beta} = \bar{\alpha}\bar{\beta}$. ■

Resumiendo los dos últimos resultados obtenemos:

Caracterización de las coalineaciones

3.40

Si $\dim \mathfrak{E} \geq 2$ entonces, cualquier coalineación en \mathfrak{E} que preseva $\mathbf{0}$ es un automorfismo semilineal.

Combinando esto con 3.31 vemos que el caso de los reales es más sencillo.

3.41

Toda coalineación en un espacio vectorial real de dimensión dos o más es un automorfismo lineal seguido por una traslación.



Capítulo cuarto

Determinantes

 El determinante es cierta función que a cada matriz cuadrada le hace corresponder un elemento del campo. Todo lo que se digamos acerca de la importancia de los determinantes en las matemáticas es poco. Este es uno de los conceptos sin los cuales realmente no se puede entender nada en las matemáticas superiores. En este capítulo daremos la definición de los determinantes, estudiaremos sus propiedades, métodos de cálculo y principales aplicaciones.

4.1 Permutaciones

La definición de determinante de una matriz pasa inevitablemente por la definición del signo de una permutación. El lector debe entender bien esta sección para poder pasar al estudio de los determinantes.

El grupo simétrico

Sea N un conjunto *finito*. Una **permutación** de N es una biyección de N en N . Al conjunto de todas las permutaciones de N lo denotaremos por S_N . La composición de biyecciones es una biyección y toda biyección tiene inversa, por lo tanto, S_N es un grupo respecto a la composición. Al grupo (S_N, \circ) se le llama el **grupo simétrico** de N . Denotaremos por I_N a la función identidad que es el neutro de S_N . Es importante recordar nuestra notación de la composición $(\sigma \circ \omega)(a) = \sigma(\omega(a))$ ya que la composición no es commutativa.

 Si $|M| = |N|$ entonces los grupos S_M y S_N son isomorfos.

Prueba. Sean M y N son dos conjuntos del mismo cardinal. Si $\omega : M \rightarrow N$ es una biyección entonces fijándonos en el diagrama commutativo del recuadro a la derecha obtenemos una función $\Delta : S_M \ni \sigma \mapsto \omega\sigma\omega^{-1} \in S_N$. Observemos, que Δ tiene inversa $S_N \ni \rho \mapsto \omega^{-1}\rho\omega \in S_M$.

$$\begin{array}{ccc} M & \xleftarrow{\sigma} & M \\ \omega \downarrow & & \uparrow \omega^{-1} \\ N & \xrightarrow{\omega\sigma\omega^{-1}} & N \end{array}$$

Además, $\Delta(\sigma\theta) = \omega\sigma\theta\omega^{-1} = \omega\sigma\omega^{-1}\omega\theta\omega^{-1} = \Delta(\sigma)\Delta(\theta)$ y por lo tanto, Δ es un isomorfismo de los grupos \mathbb{S}_M y \mathbb{S}_N . ■

El lector debe interpretar este resultado como que, en el grupo \mathbb{S}_M podemos cambiarle el nombre a los elementos de M mediante la biyección $\delta : M \rightarrow N$ y obteniendo el grupo \mathbb{S}_N . En particular, el número de elementos de \mathbb{S}_N solo depende del número de elementos en N .

 *El número de permutaciones de un conjunto con n elementos es $n!$.*

Prueba. Para la prueba es más claro encontrar $\rho(n)$ el número de biyecciones $f : M \rightarrow N$ donde $|M| = |N| = n$. Si $n = 1$ entonces $\rho(n) = 1 = 1!$. Hagamos inducción en n . Si $i \in M$ entonces, el conjunto de biyecciones lo podemos partir en n partes disjuntas según cual sea $j = f(i)$. Cada parte tiene tantos elementos como biyecciones $f : M \setminus i \rightarrow N \setminus j$ y por hipótesis de inducción este número es $(n - 1)!$. Luego, $\rho(n) = n(n - 1)! = n!$. ■

Ejemplo. Supongamos que $N = \{1, 2, 3\}$. El resultado anterior nos dice que hay 6 permutaciones y estas son las siguientes:

$$\mathbb{I} = \begin{matrix} 1 \mapsto 1 \\ 2 \mapsto 2 \\ 3 \mapsto 3 \end{matrix}, \quad \alpha = \begin{matrix} 1 \mapsto 1 \\ 2 \mapsto 3 \\ 3 \mapsto 2 \end{matrix}, \quad \beta = \begin{matrix} 1 \mapsto 2 \\ 2 \mapsto 1 \\ 3 \mapsto 3 \end{matrix}, \quad \gamma = \begin{matrix} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 1 \end{matrix}, \quad \delta = \begin{matrix} 1 \mapsto 3 \\ 2 \mapsto 1 \\ 3 \mapsto 2 \end{matrix}, \quad \varepsilon = \begin{matrix} 1 \mapsto 3 \\ 2 \mapsto 2 \\ 3 \mapsto 1 \end{matrix}$$

\circ	α	β	γ	δ	ε
α	\mathbb{I}	γ	β	ε	δ
β	δ	\mathbb{I}	ε	α	γ
γ	ε	α	δ	\mathbb{I}	β
δ	β	ε	\mathbb{I}	γ	α
ε	γ	δ	α	β	\mathbb{I}

La permutación \mathbb{I} es la función identidad que es el neutro para la composición. Haciendo unos pocos cálculos obtenemos que la tabla de composición de estas permutaciones es la del recuadro. Observese que en cada renglón y columna todas las entradas son diferentes. Esta propiedad se cumple para la tabla de la operación de cualquier grupo ya que no es nada más que el reflejo de que $(a * b = a * c) \Rightarrow (b = c)$.

Ciclos y órbitas

Sea $M = \{x_0, \dots, x_{n-1}\} \subseteq N$. A la permutación σ mostrada en el recuadro a la derecha se le llama **ciclo de orden n** . A esta permutación se le denotará por (x_0, \dots, x_{n-1}) . Dos ciclos (x_0, \dots, x_{n-1}) y (y_0, \dots, y_{m-1}) se les llama **disjuntos** si ningún x_i es igual a algún y_j .

$$\sigma(y) = \begin{cases} x_{i+1 \bmod n} & \text{si } y = x_i \\ y & \text{si } y \notin M \end{cases}$$

Es importante notar que la composición de ciclos *disjuntos* es conmutativa (¡pruébelo!) pero la composición de ciclos en general no lo es. También debemos notar que debido a las propiedades de la función $\bmod n$ se tiene que (a, \dots, b, c) es la misma permutación que (c, a, \dots, b) , o sea, siempre podemos escoger el principio del ciclo.

A.3

Sea $\sigma \in S_N$ una permutación. La relación en N definida por $\exists n \in \mathbb{Z}$ tal que $a = \sigma^n(b)$ es de equivalencia.

Prueba. Tenemos $a = \sigma^0(a)$ y por lo tanto es reflexiva. Si $a = \sigma^n(b)$ y $b = \sigma^m(c)$ entonces $a = \sigma^{n+m}(c)$ y por lo tanto es transitiva. Si $a = \sigma^n(b)$ entonces, $b = \sigma^{-n}(a)$ por lo que la relación es simétrica. ■

A las clases de equivalencia de esta relación se le llaman **órbitas** de la permutación.

A.4

La restricción de una permutación a una órbita es un ciclo.

Prueba. Supongamos que M es una órbita. Sea $a \in M$. Tenemos $M = \{\sigma^n(a) \mid n \in \mathbb{Z}\}$. El conjunto M no puede ser infinito por lo que existe un natural p más pequeño tal que $\sigma^p(b)$ ya apareció antes en la sucesión $a, \sigma(a), \dots$. Observemos que $\sigma^p(a) = a$ porque si no, habría un número k mayor que cero y menor que p tal que $\sigma^p(a) = \sigma^k(a)$ o sea, $\sigma^k(a)$ tendría dos preimágenes diferentes $\sigma^{k-1}(a) \neq \sigma^{p-1}(a)$ y esto no puede de ser ya que σ es una biyección. Dividiendo con resto cualquier entero n entre p tenemos que $n = kp + r$ con $0 \leq r < p$. Luego, $\sigma^n(a) = \sigma^r(\sigma^{kp}(a)) = \sigma^r(a)$ y $M = \{\sigma^n(a) \mid n \in \mathbb{Z}_p\}$. Esto quiere decir que la restricción de σ a M es el ciclo $(a, \sigma(a), \dots, \sigma^{p-1}(a))$. ■

A.5

Toda permutación es composición de ciclos disjuntos.

Prueba. Solo tenemos que tomar los ciclos correspondientes a todas las órbitas y componerlos en cualquier orden. ■

Observese que la descomposición en ciclos disjuntos de una permutación es única salvo el orden de la composición.

Ejemplo. Sea $\sigma = (1, 2, 3, 4) \circ (4, 5) \circ (2, 6, 3)$. Estos ciclos no son disjuntos y la permutación es la siguiente

$$1 \mapsto 2, 2 \mapsto 6, 3 \mapsto 3, 4 \mapsto 5, 5 \mapsto 1, 6 \mapsto 4$$

Luego, la descomposición en ciclos disjuntos es $\sigma = (1, 2, 6, 4, 5) \circ (3)$.

Ejercicio 75 Tome una permutación y descompóngala en composición de ciclos disjuntos. Repita el ejercicio hasta que entienda bien los conceptos de órbita y de descomposición en ciclos disjuntos.

Ejercicio 76 ¿Cuantas órbitas tiene un ciclo? ¿Cuantas órbitas tiene la identidad? Si σ tiene k órbitas, ¿cuantas órbitas tiene σ^{-1} ?

El grupo alternante

Una permutación se le llama **par** si en su descomposición en ciclos disjuntos hay un número par de ciclos de orden par. En otro caso se le llama **ímpar**.

En particular, los ciclos pares son de orden impar. A los ciclos de orden 2 se les llama **transposiciones**. Las transposiciones son impares. La inversa de cualquier transposición es ella misma.

4.6 *Al componer una permutación con una transposición la paridad de la permutación cambia.*

Prueba. Sea σ una permutación y $\tau = (a, b)$ una transposición. Distingamos dos casos: que a y b están en una misma órbita de σ o que no.

Si están en la misma órbita M entonces escogiendo el principio del ciclo en M y reenumerando los elementos de N podemos lograr que $\tau = (1, k)$ con $k > 1$ y que la restricción de σ a M es $(1, 2, \dots, n)$ con $n \geq k$. Como $\tau(\sigma(k-1)) = 1$ y $\tau(\sigma(n)) = k$, obtenemos

$$(1, k) \circ (1, 2, \dots, n) = (1, \dots, k-1) \circ (k, \dots, n).$$

Si n es par entonces, $k-1$ y $n-k+1$ tienen la misma paridad por lo que la paridad de σ cambia. Si n es impar entonces $k-1$ y $n-k+1$ tienen diferente paridad por lo que la paridad de σ cambia.

Si están en diferentes órbitas M_1 y M_2 entonces escogiendo los principios de los ciclos en M_1 , M_2 y reenumerando los elementos de N podemos lograr que $\tau = (1, k)$ con $k > 1$ y que la restricción de σ a M_1 es $(1, \dots, k-1)$ y la restricción de σ a M_2 es (k, \dots, n) . Como $\tau(\sigma(k-1)) = k$ y $\tau(\sigma(n)) = 1$, obtenemos que

$$(1, k) \circ (1, \dots, k-1) \circ (k, \dots, n) = (1, 2, \dots, n)$$

y ya vimos que $(1, \dots, k-1) \circ (k, \dots, n)$ tiene paridad diferente que $(1, 2, \dots, n)$.

Con esto hemos demostrado que la paridad de $\tau \circ \sigma$ es diferente a la de σ . La demostración de que a paridad de $\sigma \circ \tau$ es diferente a la de σ es análoga. ■

4.7 *Toda permutación es composición de transposiciones.*

Prueba. Se comprueba fácilmente que $(x_0, \dots, x_{n-1}) = (x_{n-1}, x_0) \circ \dots \circ (x_2, x_0) \circ (x_1, x_0)$ y esto prueba que todo ciclo se descompone en composición de transposiciones. La prueba se completa porque toda permutación es composición de ciclos. ■

4.8 *Una permutación es par si y solo si es composición de un número par de transposiciones.*

Prueba. Las transposiciones son impares. Aplicando repetidamente 4.6 obtenemos que las composiciones de un número par de transposiciones son pares y las composiciones

de un número impar de transposiciones es impar. La prueba se completa con 4.7. ■



Hay muchas descomposiciones de una misma permutación en composición de transposiciones. El resultado anterior nos garantiza que en dos diferentes descomposiciones la paridad del número de transposiciones es la misma.

Al conjunto de todas las permutaciones pares de \mathbb{N} se le denotará por \mathbb{A}_N . La composición de dos permutaciones pares es par y la inversa de una permutación par es también par. Luego \mathbb{A}_N es un grupo para la composición y se le llama **grupo alterna**nte. Al conjunto de las permutaciones impares lo denotaremos por \mathbb{A}_N^- . Observese que \mathbb{A}_N y \mathbb{A}_N^- tienen la misma cantidad de permutaciones e igual a $n!/2$ ya que el componer con una transposición fija es una biyección entre \mathbb{A}_N y \mathbb{A}_N^- .

El signo de una permutación

En todo campo \mathbb{K} hay dos elementos notables, 1 y -1 . El primero es el neutro para el producto y el segundo es el opuesto para la suma del primero. Observese que estos dos elementos son diferentes si y solo si la característica del campo es diferente de 2. El **signo** de una permutación es la función $\text{sgn} : \mathbb{S}_N \rightarrow \mathbb{K}$ que es 1 si la permutación es par y es -1 si la permutación es impar.

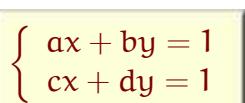
 ♦ $\text{sgn}(\pi \circ \rho) = \text{sgn } \pi \text{sgn } \rho$
♦ $\text{sgn } \pi^{-1} = \text{sgn } \pi$

Prueba. La composición de dos permutaciones de la misma paridad es una permutación par. La composición de dos permutaciones de diferente paridad es impar. Las órbitas de una permutación no cambian al tomar la inversa. ■

La función sgn jugará un papel vital en todo lo que sigue. En particular, en la definición de determinante de una matriz los signos de los sumandos están definidos precisamente mediante la función sgn . Por esto, el lector debe familiarizarse muy bien con la definición de esta función y sus propiedades.

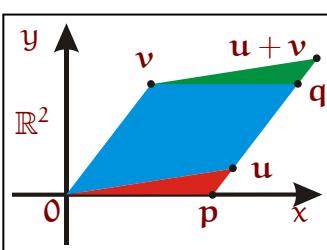
 El resultado 4.9 lo que quiere decir es que la función $\text{sgn} : \mathbb{S}_N \rightarrow \mathbb{K}$ es un morfismo del grupo \mathbb{S}_N al grupo multiplicativo del campo. Su imagen es $\{1, -1\}$ y su núcleo es \mathbb{A}_N si el campo es de característica diferente de dos.

4.2 Determinantes. Propiedades básicas

 Resolvamos el sistema de ecuaciones lineales de la izquierda. Denotemos $\Delta = ad - bc$. Despejando x en la primera ecuación y substituyendo en la segunda obtenemos y . Substituyendo este y en alguna de las ecuaciones obtendremos x . Esta solución es la del recuadro a la derecha.

$$x = \frac{d - b}{\Delta}$$

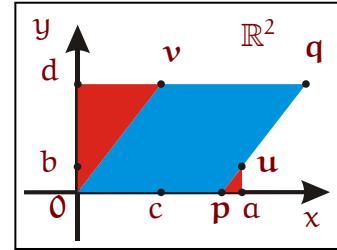
$$y = \frac{a - c}{\Delta}$$



Sean ahora $\mathbf{u} = (a, b)$ y $\mathbf{v} = (c, d)$ dos vectores en el plano \mathbb{R}^2 como se muestra en la figura a la izquierda. Estos dos vectores definen un paralelogramo cuyos vértices son $\mathbf{0}, \mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{v}$. Queremos calcular el área de este paralelogramo. Para esto, sea \mathbf{q} el punto de intersección de la recta $\mathbf{u}, \mathbf{u} + \mathbf{v}$ con la recta paralela al eje x que pasa por \mathbf{v} . Sea \mathbf{p} el punto de intersección de la recta $\mathbf{u}, \mathbf{u} + \mathbf{v}$ con el eje x . Es fácil ver que el triángulo $\mathbf{v}, \mathbf{q}, \mathbf{u} + \mathbf{v}$ es igual al triángulo $\mathbf{0}, \mathbf{p}, \mathbf{u}$. Luego, el paralelogramo $\mathbf{0}, \mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{v}$ tiene área igual a la del paralelogramo $\mathbf{0}, \mathbf{v}, \mathbf{q}, \mathbf{p}$. En la figura a la derecha los triángulos $\mathbf{p}, \mathbf{a}, \mathbf{u}$ y $\mathbf{0}, \mathbf{v}, \mathbf{d}$ tienen dos ángulos iguales o sea son congruentes. Por el Teorema de Tales tenemos que

$$(b : (a - p) = d : c) \Rightarrow (pd = ad - cb).$$

Sabemos que, pd es el área (base por altura) del paralelogramo $\mathbf{0}, \mathbf{v}, \mathbf{q}, \mathbf{p}$. Luego, hemos demostrado que el área del paralelogramo $\mathbf{0}, \mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{v}$ es igual a $\Delta = ad - bc$.



Estos dos ejemplos nos dan una idea de la importancia del número $\Delta = ad - bc$ para la solución de sistemas de dos ecuaciones lineales y para el cálculo de áreas en \mathbb{R}^2 . Los determinantes son la generalización de este número a dimensiones arbitrarias.

Definición de los determinantes

Sea \mathbf{N} un conjunto finito. Recordemos que $\mathbb{S}_{\mathbf{N}}$ es el conjunto de todas las permutaciones de \mathbf{N} . Si $\sigma \in \mathbb{S}_{\mathbf{N}}$ e $i \in \mathbf{N}$ entonces, denotaremos por σ_i la imagen de i por la permutación σ , o sea σ_i es una forma corta de escribir $\sigma(i)$.

El **determinante** de una matriz $\alpha_{\mathbf{NN}}$ es por definición el elemento del campo definido por la fórmula en el recuadro de la derecha. A esta fórmula se la conoce como **fórmula de Leibniz**.

$$\det \alpha_{\mathbf{NN}} = \sum_{\sigma \in \mathbb{S}_{\mathbf{N}}} \operatorname{sgn} \sigma \prod_{i \in \mathbf{N}} \alpha_{i\sigma_i}$$

En los capítulos anteriores ya acostumbramos al lector a las sumatorias en las cuales el conjunto de índices es un conjunto de vectores. A partir de ahora, el lector deberá acostumbrarse a usar sumatorias en las cuales el conjunto de índices es un conjunto de permutaciones. Es oportuno enfatizar que el orden en que se efectúe la suma no es relevante ya que la suma en cualquier campo es commutativa.



Recalquemos que el determinante de una matriz está definido solo cuando los conjuntos de índices de renglones y columnas coinciden y son finitos. Por este motivo *en este capítulo todos los espacios serán de dimensión finita y todos los conjuntos de índices serán finitos*.

Ejercicio 77 ¿Como se reescribiría la definición de determinante usando que $\mathbb{A}_{\mathbf{N}}$ es el conjunto de las permutaciones pares y $\mathbb{A}_{\mathbf{N}}^-$ el de las impares?

Determinantes de matrices pequeñas

Interpretemos esta definición para conjuntos N con pocos elementos. Si $|N| = 1$ entonces la matriz α_{NN} consta de una sola entrada y su determinante es esta entrada.

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} +$$

Si, digamos $N = \{1, 2\}$ entonces tenemos 2 permutaciones de N que son \mathbb{I} y $(1, 2)$. A la primera le corresponde el sumando $\alpha_{11}\alpha_{12}$ y a la segunda el sumando $-\alpha_{12}\alpha_{21}$. Gráficamente, cuando $N = \{1, 2\}$ el determinante es la suma de los dos productos que se muestran en el recuadro.

Pongamos ahora $N = \{1, 2, 3\}$. Hay 6 permutaciones de N y estas son \mathbb{I} , $(1, 2, 3)$, $(1, 3, 2)$, $(2, 3)$, $(1, 2)$ y $(1, 3)$. Las tres primeras son de signo positivo y se corresponden con los tres sumandos $\alpha_{11}\alpha_{22}\alpha_{33}$, $\alpha_{12}\alpha_{23}\alpha_{31}$ y $\alpha_{13}\alpha_{21}\alpha_{32}$. Gráficamente, estos tres sumandos se pueden representar por la diagonal principal de la matriz y los dos “triángulos” con lados paralelos a esta diagonal como se muestra en el recuadro de la derecha.

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$$

Las otras tres permutaciones tienen signo negativo y se corresponden con los sumandos $-\alpha_{11}\alpha_{23}\alpha_{32}$, $-\alpha_{12}\alpha_{21}\alpha_{33}$ y $-\alpha_{13}\alpha_{22}\alpha_{31}$. Gráficamente, estos tres sumandos se pueden representar por la diagonal alterna de la matriz y los dos “triángulos” con lados paralelos a esta diagonal como se muestra en el recuadro de la izquierda.

El número de sumandos en la definición del determinante es $|\mathbb{S}_N| = |N|!$. Este número crece rápidamente con $|N|$ como se ve en la siguiente tabla

n	1	2	3	4	5	6	7	8	9	10
$n!$	1	2	6	24	120	720	5,040	40,320	362,880	3'628,800

Por esto, calcular determinantes con directamente de la definición es muy inefficiente.

El determinante de la identidad

Ya vimos que el conjunto de matrices α_{NN} es un álgebra respecto al producto y suma de matrices y multiplicación por elementos del campo. El elemento neutro para el producto lo llamamos **matriz identidad**, denotamos \mathbb{I}_{NN} y es la matriz cuyas entradas son iguales al **delta de Kronecker** δ_{ij} definido en el recuadro a la derecha. Cuando el conjunto de

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

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

índices está ordenado, la matriz identidad se puede representar gráficamente como una que tiene unos en la diagonal y ceros en todas las demás entradas como se muestra en el recuadro a la izquierda para una matriz de dos renglones y columnas.

El determinante de la matriz identidad es 1.

$$\det \mathbb{I}_{NN} = 1$$

Prueba. Si la permutación $\sigma \in \mathbb{S}_N$ no es la identidad entonces hay un $i \in N$ tal que

$i \neq \sigma_i$ y por lo tanto $\prod_{i \in N} \delta_{i\sigma_i} = 0$. Luego,

$$\det \mathbb{I}_{NN} = \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \delta_{i\sigma_i} = \operatorname{sgn} (\mathbb{I}_N) \prod_{i \in N} \delta_{ii} = 1. \blacksquare$$

Matrices con filas nulas

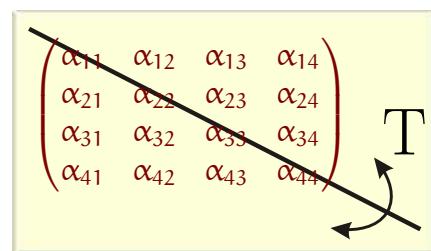
4.11

Si una matriz tiene una columna o un renglón nulo entonces, su determinante es cero.

Prueba. Cada sumando de los determinantes $\prod_{i \in N} \alpha_{i\sigma_i}$ contiene como factor una entrada de cada renglón. Si un renglón es nulo entonces, todos estos sumandos tienen un factor cero y por lo tanto la suma de todos ellos es cero. Lo mismo ocurre con las columnas. ■

El determinante de la transpuesta

Dada una matriz α_{MN} su **transpuesta** se define como la matriz $\beta_{NM} = \alpha_{MN}^T$ tal que $\beta_{ij} = \alpha_{ji}$. Gráficamente la operación de transponer una matriz es hacer una reflexión con respecto a la diagonal de la matriz. Observese que el conjunto de índices de los renglones de la matriz α_{NM}^T es M y no N como se podría pensar de los subíndices. La notación es así porque pensamos la transpuesta como la aplicación de la operación de transposición.



4.12 *El determinante no se altera al transponer una matriz.*

$$\det A = \det A^T$$

Prueba. Tenemos que demostrar que $\det \alpha_{NN} = \det \alpha_{NN}^T$. Efectivamente, por la definición

$$\begin{aligned} \det \alpha_{NN}^T &= \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \alpha_{\sigma_i i} = \left[\begin{array}{c} \text{cambio de variable} \\ \omega = \sigma^{-1} \end{array} \right] = \sum_{\omega \in S_N} \operatorname{sgn} \omega \prod_{i \in N} \alpha_{\omega^{-1}(i)i} = \\ &= \left[\begin{array}{c} \text{cambio de variable} \\ j = \omega^{-1}(i) \end{array} \right] = \sum_{\omega \in S_N} \operatorname{sgn} \omega \prod_{j \in N} \alpha_{j\omega_j} = \det \alpha_{NN}. \blacksquare \end{aligned}$$

El determinante del producto

La siguiente propiedad básica de los determinantes es probablemente la más importante para el álgebra lineal. Como la demostración de esta propiedad es laboriosa, le recomiendo al lector omitirla en una primera lectura. La complejidad de esta demostración es el precio que tenemos que pagar por dar una definición directa del determinante.

Teorema del determinante del producto

4.13

El determinante del producto de dos matrices es igual al producto de los determinantes de las dos matrices.

$$\det AB = \det A \det B$$

Prueba. Sean $A = \alpha_{NN}$ y $B = \beta_{NN}$. Para el objeto de esta demostración denotemos por \mathbb{F}_N el conjunto de todas las funciones de N en N . Claramente $S_N \subset \mathbb{F}_N$. Sea, además, $T_N \stackrel{\text{def}}{=} \mathbb{F}_N \setminus S_N$ el conjunto de todas las funciones no biyectivas de N en N .

Por la definición de determinante y de producto de matrices tenemos

$$\det AB = \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \sum_{j \in N} \alpha_{ij} \beta_{j\sigma_i}$$

y usando la fórmula $\prod_{i \in N} \sum_{j \in N} \gamma_{ij} = \sum_{\omega \in \mathbb{F}_N} \prod_{i \in N} \gamma_{i\omega_i}$ (Sec. 1.6, pág. 22) obtenemos

$$\det AB = \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \sum_{\omega \in T_N} \prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i} + \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \sum_{\omega \in S_N} \prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i}$$

Denotando por ∇ el primer sumando nuestro determinante se convierte en

$$\begin{aligned} &= \nabla + \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \sum_{\omega \in S_N} \prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i} = \left[\begin{array}{c} \text{cambio de variable} \\ \sigma = \rho \circ \omega \end{array} \right] = \\ &= \nabla + \sum_{\rho \in S_N} \sum_{\omega \in S_N} \operatorname{sgn} (\rho \circ \omega) \prod_{i \in N} \alpha_{i\omega_i} \prod_{j \in N} \beta_{\omega_j \rho(\omega_j)} = \left[\begin{array}{c} \text{cambio de var} \\ k = \omega_j \end{array} \right] = \\ &= \nabla + \left(\sum_{\omega \in S_N} \operatorname{sgn} \omega \prod_{i \in N} \alpha_{i\omega_i} \right) \left(\sum_{\rho \in S_N} \operatorname{sgn} \rho \prod_{k \in N} \beta_{k\rho_k} \right) = \nabla + \det A \det B \end{aligned}$$

O sea, para completar la demostración tenemos que probar que $\nabla = 0$. Para esto recordemos que A_N es el subgrupo de las permutaciones pares e A_N^- es el conjunto de todas las permutaciones impares. Si observamos detenidamente la definición de ∇

$$\nabla = \sum_{\omega \in T_N} \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i}$$

vemos que para probar $\nabla = 0$ es suficiente construir para cada $\omega \in T_N$ una biyección $f : A_N \rightarrow A_N^-$ tal que si $\theta = f(\sigma)$ entonces,

$$\prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i} = \prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \theta_i}$$

Esto nos garantizará que cada sumando positivo se cancele con otro negativo.

Sea $\omega \in T_N$ arbitraria pero fija. Como ω no es una biyección existen $j, k \in N$ tales que $\omega(j) = \omega(k)$. Sea $t \in A_N^-$ la transposición que intercambia j y k . Sea $f : A_N \ni \sigma \mapsto \sigma \circ t \in A_N^-$. La función f es biyectiva ya que su inversa es $\theta \mapsto \theta \circ t$. Además, tenemos

$$\prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i (\sigma \circ t)_i} = \alpha_{j\omega_j} \beta_{\omega_j \sigma_k} \alpha_{k\omega_k} \beta_{\omega_k \sigma_j} \prod_{i \in N \setminus \{j, k\}} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i} = \prod_{i \in N} \alpha_{i\omega_i} \beta_{\omega_i \sigma_i}$$

donde la última igualdad es válida porque $\omega_j = \omega_k$. ■

Matrices con filas iguales

A.14 Si una matriz tiene dos columnas o dos renglones iguales entonces su determinante es cero.

Prueba. Supongamos que para α_{NN} se tiene que $\alpha_{jM} = \alpha_{kM}$ o sea, los renglones j y k son iguales. Todo sumando del determinante depende exactamente de una entrada en el renglón j y de otra en el renglón j . Luego, podemos escribir

$$\det \alpha_{NN} = \sum_{\sigma \in S_N} \alpha_{j\sigma_j} \alpha_{k\sigma_k} \operatorname{sgn} \sigma \prod_{i \in N \setminus \{j, k\}} \alpha_{i\sigma_i}.$$

Denotemos ρ la transposición (j, k) . La función $\Phi : S_N \ni \sigma \mapsto \sigma \circ \rho \in S_N$ es una biyección y el sumando correspondiente a $\sigma \circ \rho$ es igual a

$$\alpha_{j\sigma_k} \alpha_{k\sigma_j} \operatorname{sgn}(\sigma \circ \rho) \prod_{i \in N \setminus \{j, k\}} \alpha_{i\sigma_i}$$

pero como $\alpha_{j\sigma_k} = \alpha_{k\sigma_k}$ y $\alpha_{k\sigma_j} = \alpha_{j\sigma_j}$ entonces, $\alpha_{j\sigma_k} \alpha_{k\sigma_j} \operatorname{sgn}(\sigma \circ \rho) = -\alpha_{j\sigma_j} \alpha_{k\sigma_k} \operatorname{sgn} \sigma$. Esto significa que Φ es una biyección que transforma a un sumando en su negativo y por lo tanto $\det \alpha_{NN} = 0$. La prueba para las columnas se obtiene transponiendo la matriz. ■

Matrices de permutaciones

Sean M y N dos conjuntos finitos y $\phi : M \rightarrow N$ una biyección. Denotaremos por Φ_{MN} a la matriz cuyas entradas están definidas como en el recuadro. A esta matriz la llamaremos **matriz de la biyección** ϕ . Como ϕ es una biyección entonces la matriz Φ_{MN} tiene la misma cantidad de renglones y de columnas. Además, en cada columna y cada renglón de Φ_{MN} hay exactamente una entrada igual a 1 y las demás son cero.

$$\Phi_{ij} = \begin{cases} 1 & \text{si } j = \phi(i) \\ 0 & \text{en otro caso} \end{cases}$$

Haremos un buen uso de las matrices de biyecciones en la próxima sección. Aquí estaremos interesados solo en el caso particular cuando $N = M$, o sea, que ϕ es una permutación. En este caso a Φ_{NN} se le llama **matriz de la permutación** ϕ .

A.15 Si ϕ y σ son dos permutaciones de N entonces, el producto $\Phi_{NN} \sigma_{NN}$ de las matrices de permutaciones es la matriz de la permutación $\sigma \circ \phi$.

Prueba. Denotemos $\gamma_{NN} \stackrel{\text{def}}{=} \Phi_{NN} \sigma_{NN}$. Tenemos que

$$\gamma_{ij} = \sum_{k \in N} \Phi_{ik} \sigma_{kj} = \sigma_{\phi(i)j} = \begin{cases} 1 & \text{si } j = \sigma(\phi(i)) \\ 0 & \text{en otro caso} \end{cases}$$

y esta es la definición de la matriz de la permutación $\sigma \circ \phi$. ■

El resultado anterior se puede escribir de forma corta como $\Phi_{NN} \sigma_{NN} = (\sigma \circ \phi)_{NN}$. ¿Qué querrá decir Φ_{NN}^{-1} ? Hay dos formas de interpretarlo y en el siguiente resultado se

demuestra que ambas producen la misma matriz.

4.16

La matriz de la inversa de una permutación es igual a la inversa de la matriz de la permutación.

$$(\phi^{-1})_{NN} = (\phi_{NN})^{-1}$$

Prueba. Sea ϕ una permutación ϕ_{NN} su matriz y γ_{NN} la matriz de ϕ^{-1} . Por 4.15 tenemos que $\phi_{NN}\gamma_{NN} = (\phi^{-1} \circ \phi)_{NN} = I_{NN}$. O sea, $\gamma_{NN} = (\phi_{NN})^{-1}$. ■

Ejercicio 78 ¿Cuál es la TL de la matriz de una permutación? ¿Qué tienen que ver 4.15 y 4.16 con que la correspondencia entre matrices y TL es un morfismo de álgebras?

Ejercicio 79 Demuestre que la transpuesta de la matriz de la permutación ϕ es la matriz de la permutación ϕ^{-1} .

Ejercicio 80 Pruebe que $(AB)^T = B^T A^T$ y $(A^T)^{-1} = (A^{-1})^T$.

4.17

El determinante de la matriz de una permutación es igual al signo de la permutación.

$$\det \phi_{NN} = \operatorname{sgn} \phi$$

Prueba. Sea ϕ una permutación y ϕ_{NN} su matriz. Por definición tenemos que

$$\det \phi_{NN} = \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \phi_{i\sigma_i}.$$

Por otro lado, por definición de ϕ_{NN} tenemos que

$$\prod_{i \in N} \phi_{i\sigma_i} = \begin{cases} 1 & \text{si } \sigma = \phi \\ 0 & \text{en otro caso} \end{cases}.$$

Usando estas dos igualdades obtenemos la prueba. ■

Permutaciones de columnas y renglones

Bueno, ¿y qué pasa si multiplicamos una matriz arbitraria por una matriz de permutaciones?

4.18

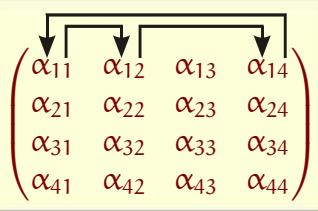
Sea α_{MN} cualquier matriz y ϕ_{NN} la matriz de la permutación ϕ . El resultado del producto $\alpha_{MN}\phi_{NN}$ es la matriz α_{MN} a la que se le han permutado las columnas usando la permutación ϕ .

Prueba. Sea $\gamma_{MN} = \alpha_{MN}\phi_{NN}$. Tenemos

$$\gamma_{ij} = \sum_{k \in N} \alpha_{ik} \phi_{kj} = \alpha_{i\phi^{-1}(j)}$$

y por lo tanto $\gamma_{Mj} = \alpha_{M\phi^{-1}(j)}$ o lo que es lo mismo $\gamma_{M\phi(j)} = \alpha_{Mj}$. Más descriptivamente,

la columna j de la matriz α_{MN} tiene índice $\phi(j)$ en la matriz γ_{MN} . ■



En este recuadro se ilustra graficamente que es lo que pasa cuando se permuta una matriz genérica de 4 renglones y columnas usando la permutación $1 \mapsto 2 \mapsto 4 \mapsto 1$. Los principios de las flechas marcan las columnas que se moverán y los finales de las flechas marcan el lugar donde quedarán las columnas.

4.19

Al permutar las columnas de una matriz con la permutación ϕ , el determinante se multiplica por un factor igual a $\operatorname{sgn} \phi$.

Prueba. Al permutar la columnas lo que estamos haciendo es multiplicar la matriz por la matriz de ϕ . Por 4.17 el determinante de la matriz de ϕ es igual al signo de ϕ . Usando el Teorema del determinante del producto concluimos la demostración. ■

¿Qué pasa con los renglones? Pues lo mismo, solo hay que multiplicar por el otro lado. El lector puede modificar los razonamientos anteriores para el caso de los renglones. También puede tomar la vía rápida: permutar los renglones de una matriz es lo mismo que permutar las columnas de la transpuesta y entonces $(\alpha_{MN}^T \phi_{MM})^T = \phi_{MM}^T \alpha_{MN} = \phi_{MM}^{-1} \alpha_{MN}$ (véanse los ejercicios 79 y 80). Esto quiere decir que si multiplicamos por la izquierda con la matriz de permutaciones ϕ_{MM}^{-1} entonces los renglones se permutan mediante la permutación ϕ y el determinante cambia igual (ya que $\operatorname{sgn} \phi = \operatorname{sgn} \phi^{-1}$).

4.3 Expansión de Laplace

En esta sección encontraremos una descomposición de los determinantes como una combinación lineal de determinantes de matrices más pequeñas. Despues veremos importantes consecuencias de esta expansión, en particular que una matriz tiene inversa si y solo si esta tiene determinante diferente de cero.

Cambios de índices

Sea α_{MN} una matriz y $\phi : N \rightarrow L$ una biyección. Sea ϕ_{NL} la matriz de la biyección ϕ (véase la página 102). Podemos construir una matriz β_{ML} por la fórmula $\beta_{ML} = \alpha_{MN} \phi_{NL}$. A esta operación la llamaremos **cambio de índices de las columnas** de la matriz α_{MN} mediante la biyección ϕ . De la misma manera se definen los **cambios de índices de los renglones**. Observese que las permutaciones de las columnas y los renglones son un caso particular de los cambios de índices cuando $N = L$. Si $N \neq L$ entonces, podemos pensar que hacer un cambio de índices de las columnas es darle nuevos nombres a estas.

Una **matriz cuadrada** es una matriz cuyas cantidades de renglones y columnas coinciden. A este número común se le llama **orden de la matriz**. Necesitaremos los

cambios de índices para definir los determinantes de las matrices cuadradas. Así, si $\phi : \mathbf{N} \rightarrow \mathbf{M}$ es una biyección entonces, podríamos definir $\det \alpha_{MN} = \det \alpha_{MN}\phi_{NM}$. El único “pero” es que, en principio, esta definición no solo depende de la matriz α_{NM} sino también de la biyección ϕ . El cuanto depende esta definición de ϕ lo responde la siguiente proposición.

4.20

Si ϕ y φ son dos biyecciones de \mathbf{N} en \mathbf{M} entonces, $\det \alpha_{MN}\phi_{NM} = \text{sgn}(\phi \circ \varphi^{-1}) \det \alpha_{MN}\varphi_{NM}$.

Prueba. Usando Teorema del determinante del producto (4.13) obtenemos que

$$\det \alpha_{MN}\phi_{NM} = \det \alpha_{MN}\varphi_{NM}\varphi_{NM}^{-1}\phi_{NM} = \det \alpha_{MN}\varphi_{NM} \det \varphi_{NM}^{-1}\phi_{NM}$$

Ahora, observese que $\phi \circ \varphi^{-1}$ es una permutación de \mathbf{M} y que la matriz de esta permutación es $\varphi_{NM}^{-1}\phi_{NM}$ cuyo determinante es igual (por 4.17) a $\text{sgn}(\phi \circ \varphi^{-1})$. ■



No podemos poner la conclusión de este resultado como $\text{sgn } \phi \det \alpha_{MN}\phi_{NM} = \text{sgn } \varphi \det \alpha_{MN}\varphi_{NM}$ ya que como ϕ y φ son biyecciones de \mathbf{N} en \mathbf{M} ninguna de las dos tiene signo.

Como el signo de cualquier permutación es o **1** o **-1** esto quiere decir que el determinante $\det \alpha_{NM}$ está definido “salvo signo” o sea, que hay un elemento $a \in \mathbb{K}$ tal que el determinante es a o es $-a$. En un campo de característica dos esto es irrelevante ya que en este caso **1 = -1**. Sin embargo en los casos más importantes (\mathbb{R} y \mathbb{C}) de que el campo sea de característica diferente de dos tenemos una ambigüedad al definir el determinante $\det \alpha_{NM}$.

Esta ambigüedad se resuelve en diferentes casos de varias maneras. En muchos casos no nos interesa el valor concreto $\det \alpha_{MN}$ sino solamente saber si es cierto o no que $\det \alpha_{MN} = 0$. Si este es el caso entonces, en realidad no nos importa que biyección se escogió para calcular el determinante. Por ejemplo, una matriz cuadrada α_{MN} se le llama **singular** si $\det \alpha_{MN} = 0$, en el caso contrario se le llama **no singular**. Es claro, que el ser singular o no singular NO depende de la biyección escogida para definir $\det \alpha_{MN}$. Otro ejemplo, es que si una matriz tiene una columna o renglón nulo entonces su determinante es cero.

En otros casos lo importante no es el valor de algún determinante sino una igualdad entre estos. Al cambiar los índices en ambos lados de la igualdad los determinantes cambian en el mismo signo y la igualdad es cierta independientemente de los cambios de índices escogidos. Por ejemplo, las igualdades $\det \alpha_{MN} = \det \alpha_{MN}^T$ y $\det (\alpha_{LM}\beta_{MN}) = \det \alpha_{LM} \det \beta_{MN}$ son válidas independientemente de los cambios de índices usados para definir los determinantes.

En otros casos hay una biyección natural que nos dice cual debe ser el valor de $\det \alpha_{MN}$. Esto sucede por ejemplo si los conjuntos \mathbf{N} y \mathbf{M} son conjuntos de naturales. En este caso podemos siempre escoger la única biyección que conserva el orden. Por

ejemplo si $N = \{2, 3, 7\}$ y $M = \{1, 4, 5\}$ entonces la biyección es $2 \rightarrow 1, 3 \rightarrow 4, 7 \rightarrow 5$.



¿Y por que no escoger de los dos posibles valores el que sea mayor que cero? Primero, a veces se necesita el valor del signo del determinante y segundo ¿qué quiere decir que $0 < x \in \mathbb{Z}_5$? La desigualdad $0 < x$ solo tiene sentido si el campo está ordenado. Este es el caso de \mathbb{R} pero no el de \mathbb{Z}_p ni el de \mathbb{C} .



En muchos de los textos de álgebra lineal esta ambigüedad se resuelve postulando que los conjuntos de índices siempre tienen que ser conjuntos de naturales. Esto no solamente es innecesario, sino que también hace la exposición más compleja y conceptualmente menos clara. Por ejemplo, las matrices de cambio de base están naturalmente indexadas por conjuntos de vectores que no poseen un orden natural.

Complementos algebraicos

Estos razonamientos nos interesan ahora por el siguiente caso particular en el cual todo es más fácil. Sea α_{NN} una matriz y sean $i, j \in N$. La matriz $\alpha_{N \setminus i \ N \setminus j}$ se obtiene de la matriz α_{NN} eliminando el renglón i y la columna j . ¿Habrá una manera natural de definir el determinante de $\alpha_{N \setminus i \ N \setminus j}$? Veremos que sí.

Sea $\varphi : N \setminus j \rightarrow N \setminus i$ una biyección cualquiera. Podemos definir $\varphi(j) = i$ y de esta manera φ se convierte en una permutación de N . Definamos el determinante $\det \alpha_{N \setminus i \ N \setminus j}$ con la expresión $\boxed{\operatorname{sgn} \varphi \det \alpha_{N \setminus i \ N \setminus j} \varphi_{N \setminus j N \setminus i}}$ del recuadro a la derecha. Parecería que no hemos hecho nada ya que en principio esta definición depende de φ .



La expresión $\operatorname{sgn} \varphi \det \alpha_{N \setminus i \ N \setminus j} \varphi_{N \setminus j N \setminus i}$ no depende de φ .

Prueba. Sea ω otra permutación de N tal que $\omega(j) = i$. Aquí hay que tener cuidado. Por un lado ω es una biyección de $N \setminus j$ en $N \setminus i$ y por otro es una permutación de N . Otro tanto ocurre con φ . Para evitar confusiones, a las permutaciones de N las denotaremos en esta demostración por $\hat{\omega}$ y $\hat{\varphi}$ respectivamente. Por 4.20 tenemos que $\det \alpha_{N \setminus i \ N \setminus j} \varphi_{N \setminus j N \setminus i} = \operatorname{sgn}(\varphi \circ \omega^{-1}) \det \alpha_{N \setminus i \ N \setminus j} \omega_{N \setminus j N \setminus i}$. Observese que $\varphi \circ \omega^{-1}$ es una permutación de $N \setminus i$. ¿Que pasa con $\hat{\omega} \circ \hat{\varphi}^{-1}$? Pues, en todo elemento de $N \setminus i$ coincide con $\omega \circ \varphi^{-1}$ y además $\hat{\varphi} \circ \hat{\omega}^{-1}(i) = i$. Luego $\operatorname{sgn}(\varphi \circ \omega^{-1}) = \operatorname{sgn}(\hat{\varphi} \circ \hat{\omega}^{-1})$ y por las propiedades de la función signo se tiene que $\operatorname{sgn}(\hat{\varphi} \circ \hat{\omega}^{-1}) = \operatorname{sgn} \hat{\varphi} \operatorname{sgn} \hat{\omega}$. Luego $\det \alpha_{N \setminus i \ N \setminus j} \varphi_{N \setminus j N \setminus i} = \operatorname{sgn} \hat{\varphi} \operatorname{sgn} \hat{\omega} \det \alpha_{N \setminus i \ N \setminus j} \omega_{N \setminus j N \setminus i}$ y pasando $\operatorname{sgn} \hat{\varphi}$ al otro lado de la igualdad terminamos la prueba. ■

Ahora, podemos dar la siguiente definición. Si α_{ij} es una entrada de la matriz $A = \alpha_{NN}$ entonces el **complemento algebraico** de α_{ij} es $\alpha_{ij}^* = \operatorname{sgn} \varphi \det \alpha_{N \setminus i \ N \setminus j} \varphi_{N \setminus j N \setminus i}$ donde φ es cualquier permutación de N tal que $\varphi(j) = i$. A los complementos algebraicos también se les llama **cofactores**. La matriz α_{NN}^* cuyas entradas son los complementos algebraicos α_{ij}^* es la **matriz de los complementos algebraicos** de α_{NN} o **matriz de cofactores** de α_{NN} .

La expansión de un determinante por sus renglones

Teorema de Expansión de Laplace

4.22

Sea α_{iN} un renglón arbitrario de la matriz α_{NN} . El determinante de α_{NN} es igual a la suma de las entradas del renglón por sus cofactores.

$$\det \alpha_{NN} = \sum_{j \in N} \alpha_{ij} \alpha_{ij}^*$$

Prueba. Si $i \in N$ entonces tenemos la partición del grupo simétrico en el recuadro a la derecha donde $S_N^{i \rightarrow j} = \{\sigma \in S_N \mid \sigma(i) = j\}$. Luego, aplicando la definición de determinante y sacando factor común obtenemos

$$\det \alpha_{NN} = \sum_{j \in N} \alpha_{ij} \sum_{\sigma \in S_N^{i \rightarrow j}} \operatorname{sgn} \sigma \prod_{n \in N \setminus i} \alpha_{n\sigma_n}$$

$$\bigcup_{j \in N} S_N^{i \rightarrow j}$$

Sea ω una permutación de N tal que $\omega(j) = i$. Hagamos el cambio de variable $\rho = \omega \circ \sigma$ en la sumatoria interior. Tenemos $\rho(i) = i$ o sea, $\rho \in S_N^{i \rightarrow i} = S_{N \setminus i}$. Luego,

$$\begin{aligned} \det \alpha_{NN} &= \sum_{j \in N} \alpha_{ij} \operatorname{sgn} \omega \sum_{\rho \in S_{N \setminus i}} \operatorname{sgn} \rho \prod_{n \in N \setminus i} \alpha_{n\omega^{-1}(\rho_n)} \\ &= \sum_{j \in N} \alpha_{ij} \operatorname{sgn} \omega \det \alpha_{N \setminus i \ N \setminus j} \omega_{N \setminus j \ N \setminus i} = \sum_{j \in N} \alpha_{ij} \alpha_{ij}^* \end{aligned}$$

donde la última igualdad se obtiene por la definición de α_{ij}^* . ■

Como es natural hay un teorema exactamente igual a este correspondiente a las columnas. El lector debe observar que este teorema se expresa en forma más compacta usando el producto escalar canónico de N -adas: $\det \alpha_{NN} = \alpha_{iN} \alpha_{iN}^*$.

La expansión de Laplace en forma gráfica

El Teorema de expansión de Laplace es la primera herramienta (aparte de la definición) de que disponemos para calcular determinantes. Este reduce el problema a calcular varios determinantes de matrices más pequeñas. Es especialmente útil cuando hay renglones (o columnas) con muchos ceros.

Sin embargo, todavía debemos precisar el como calcular los cofactores en forma gráfica. Si bien, la definición de estos es sencilla necesitamos una biyección simple de $N \setminus i$ en $N \setminus j$ que facilite los cálculos cuando N está ordenado o sea $N = \{1, \dots, n\}$.

1	3	4	5	6	7	8	...	n
↓	↓	↓	↓	↓	↓	↓	↓	
1	2	3	4	5	6	8	...	n

Sea por ejemplo $i = 2$ y $j = 7$. En este caso, la biyección natural de $N \setminus 2$ en $N \setminus 7$ es la que se muestra en el recuadro y la llamaremos φ . Tenemos que definir $\varphi(2) = 7$ para completar una permutación de N .

Sabemos que φ transforma $7 \mapsto 6 \mapsto 5 \mapsto 4 \mapsto 3 \mapsto 2 \mapsto 7$ y que deja fijos a todos los demás elementos de N . Luego, φ es un ciclo de longitud $j - i + 1$ (si $i > j$ entonces es un ciclo de longitud $i - j + 1$).

Como todo ciclo de orden par tiene signo negativo y todo de orden impar tiene signo positivo obtenemos $\operatorname{sgn} \varphi = (-1)^{i+j}$ lo que es muy sencillo ya que es la “**regla del tablero de ajedrez**”. A la derecha se muestra el $\operatorname{sgn} \varphi$ para una matriz con 4 renglones y columnas. Observese el parecido con el tablero.

+	-	+	-
-	+	-	+
+	-	+	-
-	+	-	+

Con estas permutaciones, los complementos algebraicos tienen una interpretación gráfica muy sencilla. Si se quiere sacar el complemento algebraico de α_{ij} entonces

táchese el renglón i , táchese la columna j , sáquese el determinante de la matriz que queda y finalmente multiplíquese por el signo de la regla del tablero de ajedrez. Así por ejemplo, en el recuadro se muestra una matriz de orden cuatro en la cual estamos calculando el complemento algebraico de α_{23} .

$$-\det \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \cancel{\alpha_{23}} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \cancel{\alpha_{43}} & \alpha_{44} \end{pmatrix}$$



Al matemático y físico Pierre-Simon Laplace (Francia 1749-1827) se le conoce fundamentalmente por sus trabajos en mecánica celeste, ecuaciones diferenciales y teoría de las probabilidades. El publicó la demostración del Teorema de Expansión de Laplace (4.22) en 1772 en un artículo donde estudiaba las órbitas de los planetas interiores del sistema solar. En este artículo, Laplace analizó el problema de solución de sistemas de ecuaciones lineales mediante el uso de determinantes.

Ejercicio 81 Tóme una matriz y calcule su determinante usando el teorema de descomposición de Laplace. Repita hasta que usted esté satisfecho.

Ejercicio 82 Demuestre que el determinante de la matriz $\alpha_{ij} = x_j^{i-1}$ es igual a la expresión en el recuadro a la derecha. A este determinante se le conoce como **determinante de Vandermonde** y la matriz tiene el mismo nombre. [191]

$$\prod_{1 \leq i < j \leq n} (x_j - x_i)$$



La expansión de Laplace reduce el cálculo del determinante de una matriz al cálculo de determinantes de matrices más pequeñas. Por esto es también usada para dar la definición de determinante mediante inducción. En este caso, la fórmula de Leibniz es una consecuencia de esta definición.

Multinearidad de los determinantes

El determinante se puede ver como una función del espacio de matrices en el campo. ¿Será el determinante una TL? ¿Se cumplirá que $\det(A + B) = \det A + \det B$? La respuesta es NO. Para probar que no, basta ver un ejemplo. Para las siguientes matrices

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

tenemos $\det A + \det B = 0 \neq 1 = \det(A + B)$. Sin embargo, los determinantes cumplen una propiedad bastante parecida.

Sea \mathfrak{E} un espacio vectorial sobre el campo \mathbb{K} . A las TL de \mathfrak{E} en \mathbb{K} se le llama **funcionales lineales** del espacio \mathfrak{E} . En esta terminología vemos que la propiedad $\det(A + B) \neq \det A + \det B$ lo que quiere decir es que el determinante NO es un funcional lineal del espacio vectorial de las matrices. Pasemos a considerar una función f con valor en \mathbb{K} de dos variables x, y que toman sus valores en \mathfrak{E} o sea, $f : \mathfrak{E}^2 \ni (x, y) \mapsto f(x, y) \in \mathbb{K}$. Como \mathfrak{E}^2 es un espacio vectorial sobre \mathbb{K} entonces podría suceder que f sea un funcional lineal. Sin embargo, hay una propiedad un poco más sutil que podría cumplir la función f y es que sea **lineal en la primera variable**. En otras palabras, que $\forall y \in \mathfrak{E}$ se cumple que $f(a + b, y) = f(a, y) + f(b, y)$ y $f(\lambda a, y) = \lambda f(a, y)$. Análogamente, se define la propiedad de que f sea **lineal en la segunda variable**. Si f es lineal en las dos variables entonces, se dice que f es un **funcional bilineal** (el “bi” es porque son dos variables).

Evidentemente todo esto se puede hacer cuando tenemos muchas variables en cuyo caso nos referiremos a los **funcionales multilineales**. Más rigurosamente, sea $f : \mathfrak{E}^N \rightarrow \mathbb{K}$ una función donde N es un conjunto de variables. Sea $i \in N$ una variable. Podemos pensar a f como una función de dos variables $f : \mathfrak{E}^{N \setminus i} \oplus \mathfrak{E} \rightarrow \mathbb{K}$. Si $\forall y \in \mathfrak{E}^{N \setminus i}$ la función $\mathfrak{E} \ni x \mapsto f(y, x) \in \mathbb{K}$ es un funcional lineal entonces, diremos que f es **lineal en la variable i** . Diremos que f es un **funcional multilínea** si f es lineal en todas sus variables. Por ejemplo, la función $f : \mathbb{R}^3 \ni (x, y, z) \mapsto x + y + z \in \mathbb{R}$ es un funcional lineal. La función $g : \mathbb{R}^3 \ni (x, y, z) \mapsto xyz \in \mathbb{R}$ es un funcional multilínea porque $(x + x')yz = xyz + x'yz$ y también para las otras variables. Observese que f no es multilínea y que g no es lineal.

Ahora queremos ver que los determinantes son funcionales multilíneales. El espacio de matrices \mathbb{K}^{NN} es isomorfo a $(\mathbb{K}^N)^N$. Un isomorfismo de estos es el que a cada matriz le hace corresponder la N -ada de sus renglones. Luego, podemos pensar el determinante como un funcional cuyas variables son los renglones y que toma valores en el campo.

4.23 *El determinante es un funcional multilínea de los renglones.*

Prueba. Para probar que un funcional es multilínea hay que probar que es lineal en cada variable. Sea $i \in N$ arbitrario pero fijo en toda esta prueba. Sea $A = \alpha_{NN}$ una matriz tal que su i -ésimo renglón es la suma de dos N -adas x_N y y_N . Sea B la misma matriz que A excepto que su i -ésimo renglón es x_N . Sea C la misma matriz que A excepto que su i -ésimo renglón es y_N . Tenemos que probar que $\det A = \det B + \det C$. (Si el lector no entiende porqué entonces, debe regresar al principio de esta sección y volver a pensar en la definición de funcional multilínea.) Usando la descomposición de Laplace por el renglón i obtenemos

$$\det \alpha_{NN} = \alpha_{iN} \alpha_{iN}^* = (x_N + y_N) \alpha_{iN}^* = x_N \alpha_{iN}^* + y_N \alpha_{iN}^* = \det B + \det C$$

donde la última igualdad se cumple porque los cofactores de las entradas del

renglón i son los mismos en las matrices A , B y C (recuerdese que hay que “tachar” el renglón i).

Sea ahora $A = \alpha_{NN}$ una matriz tal que su i -ésimo renglón es λx_N . Sea B la misma matriz que A excepto que su i -ésimo renglón es x_N . Usando la descomposición de Laplace por el renglón i obtenemos $\det \alpha_{NN} = \alpha_{iN} \alpha_{iN}^* = \lambda x_N \alpha_{iN}^* = \lambda \det B$. ■

Por transposición el determinante es también multilineal en las columnas.



Los determinantes son los únicos funcionales multilineales de los renglones que son iguales a 1 en la matriz identidad y que cambian por un factor de $\text{sgn } \theta$ cuando se permutan los renglones con la permutación θ . Esto permite dar una definición alternativa de los determinantes. El primer problema aquí es demostrar la existencia del determinante.

La inversa de una matriz

Recordemos que, dada una matriz α_{MN} decimos que α_{MN}^{-1} es la **matriz inversa** de α_{MN} si se cumple que $\alpha_{MN} \alpha_{MN}^{-1} = I_{MM}$ y $\alpha_{MN}^{-1} \alpha_{MN} = I_{NN}$. Mediante el isomorfismo de las matrices con las TLs concluimos que α_{MN} y α_{MN}^{-1} son las matrices de dos TLs una la inversa de otra y por lo tanto estas TLs son isomorfismos. Luego, si α_{MN} tiene inversa entonces, ella es cuadrada.

4.24 Si las matrices β_{NL} y α_{MN} tienen inversa entonces, $(\alpha_{MN} \beta_{NL})^{-1} = \beta_{NL}^{-1} \alpha_{MN}^{-1}$.

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

Prueba. Usando el isomorfismo de las matrices con las TL la prueba se reduce a la ya conocida por nosotros igualdad $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$. ■

Sea $\varphi : N \rightarrow M$ cualquier biyección y φ_{NM} la matriz de esa biyección. La matriz φ_{NM} siempre tiene inversa e igual a la matriz de la biyección φ^{-1} . Usando el resultado anterior obtenemos que $\varphi_{NM} (\alpha_{MN} \varphi_{NM})^{-1} = \alpha_{MN}^{-1}$. O sea, cualquier cambio de índices apropiado reduce el cálculo de matrices inversas al caso en que el conjunto de columnas coincide con el conjunto de renglones. Por esto, nos olvidaremos un poco de los índices para poder enunciar más fácilmente nuestros resultados.

La siguiente proposición nos dice que en ciertos casos para probar que una matriz es inversa de otra es suficiente comprobar una de las dos igualdades involucradas en la definición. Aquí, la clave es que las matrices son finitas y cuadradas lo que significa que sus TLs son entre espacios de la misma dimensión finita.

4.25 Si $AB = I$ entonces, $BA = I$.

Prueba. Sean $f, g : K^M \rightarrow K^M$ las TLs de las matrices A y B respectivamente. Mediante el isomorfismo de las matrices con las TLs concluimos que $f \circ g = I$. De 3.29 (página 85) obtenemos que $g \circ f = I$ y por lo tanto $BA = I$. ■

4.26

Si una matriz tiene inversa entonces ella es no singular. Además, $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$.

Prueba. Por definición de matriz inversa y el Teorema del determinante del producto tenemos que $1 = \det \mathbb{I} = \det(\mathbf{A}\mathbf{A}^{-1}) = \det \mathbf{A} \det \mathbf{A}^{-1}$. De esta igualdad obtenemos que $\det \mathbf{A} \neq 0$ y $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$. ■

Para probar que el recíproco de esta afirmación también es cierto, lo que haremos es construir la matriz inversa en el caso de que el determinante sea diferente de cero.

4.27

Si \mathbf{A} es una matriz no singular entonces, su inversa es la transpuesta de la matriz de sus cofactores dividida por el determinante de \mathbf{A} .

$$\mathbf{A}^{-1} = \frac{\mathbf{A}^{*T}}{\det \mathbf{A}}$$

Prueba. Para hacer la demostración lo que hay que hacer es multiplicar. Sea $\alpha_{MM} = \mathbf{A}$ y $\mathbf{B} = \beta_{MM} = (\det \mathbf{A})^{-1} \mathbf{A}^{*T}$. Tenemos $\beta_{Mj} = (\det \mathbf{A})^{-1} \alpha_{jM}^*$ y de la definición de producto de matrices obtenemos que la entrada ij de la matriz \mathbf{AB} es igual a $\alpha_{iM} \beta_{Mj} = (\det \mathbf{A})^{-1} \alpha_{iM} \alpha_{jM}^*$. Si $i = j$ entonces $\alpha_{iM} \alpha_{jM}^*$ es la expansión de Laplace del $\det \mathbf{A}$ por el renglón i de lo que obtenemos que $\alpha_{iM} \beta_{Mj} = 1$.

Solo queda probar que $\alpha_{iM} \alpha_{jM}^* = 0$ si $i \neq j$. Si nos fijamos atentamente, vemos que esta es la expansión de Laplace por el renglón j del determinante de la matriz \mathbf{C} obtenida de la matriz \mathbf{A} substituyendo el renglón j por el renglón i . Como la matriz \mathbf{C} tiene dos renglones iguales entonces su determinante es cero y necesariamente $\alpha_{iM} \alpha_{jM}^* = 0$. ■

El determinante de un operador lineal

Sea $f \in \text{End } \mathfrak{E}$ un OL. Si escogemos una base de \mathfrak{E} entonces en esta base el OL f tiene por coordenadas una matriz \mathbf{A} . Podríamos definir $\det f = \det \mathbf{A}$. Sin embargo, no nos queda claro si esta definición depende o no de como hallamos escogido la base.

4.28

Si \mathbf{A} y \mathbf{B} son las matrices de $f \in \text{End } \mathfrak{E}$ en dos bases entonces, $\det \mathbf{A} = \det \mathbf{B}$.

Prueba. Sean $\mathbf{A} = \alpha_{MM}$ y $\mathbf{B} = \beta_{NN}$ las matrices de f en las bases M y N respectivamente. Por la fórmula del cambio de bases para matrices (3.23) tenemos $\mathbf{B} = \gamma_{NM} \mathbf{A} \gamma_{NM}^{-1}$ donde γ_{NM} es la matriz de cambio de la base M a la base N . Tenemos

$$\det \beta_{NN} = \det(\gamma_{NM} \alpha_{MM} \gamma_{NM}^{-1}) = \det \gamma_{NM} \det \alpha_{MM} \det \gamma_{NM}^{-1} = \det \alpha_{MM}$$

ya que por 4.26 la igualdad $\det \gamma_{NM} (\det \gamma_{NM})^{-1} = 1$ es cierta e independiente del cambio de índices que se utilice para definir el determinante de γ_{NM} . ■



El determinante de un OL f en un espacio de dimensión finita! es por definición el determinante de su matriz en alguna base y este escalar no depende de la base escogida.

Esta definición nos permite traducir propiedades de las matrices a los OLs. Por ejemplo, un OL es biyectivo si y solo si su determinante es diferente de cero.



El determinante de un OL en \mathbb{R}^n tiene una importante interpretación geométrica: Si A es un subconjunto de \mathbb{R}^n que tiene “volumen n -dimensional” (su medida de Lebesgue) igual a $\text{vol } A \in \mathbb{R}^+$ y f es un OL entonces $\text{vol } f(A) = |\det f| \text{vol } A$. En particular, los OLs de determinante 1 o -1 son los que “preservan el volumen”. Si el lector se intimida con lo de “volumen n -dimensional”, entonces es mejor que piense en el área en el plano \mathbb{R}^2 .



El signo del determinante de un OL en \mathbb{R}^n tiene otra importante interpretación geométrica. El espacio \mathbb{R}^n es un espacio “orientado”. Esto, intuitivamente, lo que quiere decir es que por mucho que alguien trate de convertir su mano derecha en su mano izquierda no lo logrará, a menos que use un espejo. Matemáticamente, las reflexiones invierten la “orientación” del espacio. Todo lo que es derecho se convierte en izquierdo y recíprocamente. Por eso las reflexiones tienen en \mathbb{R}^3 determinante negativo. Los OLs en \mathbb{R}^n que tienen determinante positivo son los que preservan la “orientación” y los que tienen determinante negativo son los que invierten la “orientación”

4.4 La expansión generalizada de Laplace



Ahora generalizaremos dramáticamente el teorema de expansión de Laplace. Sea α_{NN} una matriz y I, J subconjuntos de N de la misma cardinalidad. Para ahorrarnos mucha escritura, denotemos $I' = N \setminus I$ y $J' = N \setminus J$. Además, denotemos por $\nabla_{IJ} = \det \alpha_{IJ}$ el determinante de la matriz cuyas columnas son las de J y cuyos renglones son los de I . Claro, este determinante está definido solo salvo signo. De los signos no nos preocuparemos ahora sino solo más adelante. Por otro lado, denotemos por $\Delta_{IJ} = \det \alpha_{I'J'}$ el determinante de la matriz cuyas columnas son las que no están en J y cuyos renglones son los que no están en I (no nos preocuparemos por los signos). En estas notaciones, un sumando del Teorema de expansión de Laplace es de la forma $\nabla_{IJ}\Delta_{IJ}$ donde I y J son subconjuntos de cardinal 1. Nos gustaría tener un teorema de expansión donde los subconjuntos sean de un cardinal fijo pero arbitrario.

Para esto, ahora sí nos tenemos que preocupar por los signos. Sin embargo, la siguiente proposición nos dice que esto realmente no es un problema. Sea ϕ cualquier permutación de N tal que $\phi(J) = I$ y por lo tanto $\phi(J') = I'$. La restricción de ϕ a J y la restricción de ϕ a J' son biyecciones y sus matrices se denotarán por ϕ_{JI} y $\phi_{J'I'}$ respectivamente.

4.29
La expresión $\text{sgn } \phi \det \alpha_{IJ} \phi_{JI} \det \alpha_{I'J'} \phi_{J'I'}$ no depende de la permutación ϕ .

Prueba. Sea φ otra permutación tal que $\varphi(J) = I$. Observese que $\rho = \phi \circ \varphi^{-1}$ es

una permutación de \mathbf{N} tal que $\rho(I) = I$ y $\rho(I') = I'$. Sea x el signo de la restricción de ρ a I y y el signo de la restricción de ρ a I' . Como los conjuntos I y I' son disjuntos tenemos que $\text{sgn}(\phi \circ \varphi^{-1}) = xy$.

Por 4.20 tenemos

$\det \alpha_{IJ}\phi_{JI} = x \det \alpha_{IJ}\varphi_{JI}$ y $\det \alpha_{I'J'}\phi_{J'I'} = y \det \alpha_{I'J'}\varphi_{J'I'}$. Multiplicando estas dos igualdades obtenemos

$$\det \alpha_{IJ}\phi_{JI} \det \alpha_{I'J'}\phi_{J'I'} = \text{sgn}(\phi \circ \varphi^{-1}) \det \alpha_{IJ}\varphi_{JI} \det \alpha_{I'J'}\varphi_{J'I'}$$

y usando las propiedades de la función sgn obtenemos la prueba. ■

Ahora, para I, J subconjuntos de \mathbf{N} definamos ∇_{IJ} y Δ_{IJ} con las fórmulas de la derecha donde ϕ denota una permutación (arbitraria pero la misma para las dos definiciones) tal que $\phi(J) = I$ (y en consecuencia $\phi(J') = I'$). Esta definición no es correcta en el sentido de que ambos ∇_{IJ} y Δ_{IJ} dependen de ϕ . Sin embargo $\nabla_{IJ}\Delta_{IJ}$ no depende de ϕ y esto es lo importante.

$$\begin{aligned}\nabla_{IJ} &= \det \alpha_{IJ}\phi_{JI} \\ \Delta_{IJ} &= \text{sgn } \phi \det \alpha_{I'J'}\phi_{J'I'}\end{aligned}$$

Expansión Generalizada de Laplace

4.30

Si I un conjunto de p renglones de la matriz α_{NN} entonces, $\det \alpha_{NN} = \sum \nabla_{IJ}\Delta_{IJ}$ donde la suma recorre todos los subconjuntos de columnas J de cardinalidad p .

$$\det \alpha_{NN} = \sum_{|J|=p} \nabla_{IJ}\Delta_{IJ}$$

Prueba. Si $I \subseteq \mathbf{N}$ y $|I| = p$ entonces, tenemos la partición del grupo simétrico a la derecha donde $\mathbb{S}_N^{I \rightarrow J} = \{\sigma \in \mathbb{S}_N \mid \sigma(I) = J\}$. Aplicando la definición de determinante obtenemos

$$\det \alpha_{NN} = \sum_{|J|=p} \sum_{\sigma \in \mathbb{S}_N^{I \rightarrow J}} \text{sgn } \sigma \prod_{n \in \mathbf{N}} \alpha_{n\sigma_n}$$

$$\bigcup_{|J|=p} \mathbb{S}_N^{I \rightarrow J}$$

Sea ϕ una permutación de \mathbf{N} tal que $\phi(J) = I$. Hagamos el cambio de variable $\rho = \phi \circ \sigma$. Entonces,

$$\det \alpha_{NN} = \sum_{|J|=p} \text{sgn } \phi \sum_{\rho \in \mathbb{S}_N^{I \rightarrow I}} \text{sgn } \rho \prod_{n \in \mathbf{N}} \alpha_{i\rho^{-1}(\rho_n)}$$

Como $\rho(I) = I$ entonces, $\rho(I') = I'$. Luego ρ se puede descomponer en la composición de dos permutaciones $\theta \circ \omega$ donde $\theta \in \mathbb{S}_I$ y $\omega \in \mathbb{S}_{I'}$. Substituyendo ρ por $\theta \circ \omega$ la suma se convierte en suma doble y si sacamos factores comunes obtendremos:

$$\det \alpha_{NN} = \sum_{|J|=p} \text{sgn } \phi \left(\sum_{\theta \in \mathbb{S}_I} \text{sgn } \theta \prod_{n \in I} \alpha_{i\theta^{-1}(\theta_n)} \right) \left(\sum_{\omega \in \mathbb{S}_{I'}} \text{sgn } \omega \prod_{n \in I'} \alpha_{i\omega^{-1}(\omega_n)} \right)$$

Lo contenido en el primer paréntesis es $\det \alpha_{I\phi(J)} = \nabla_{IJ}$. Lo contenido en el segundo paréntesis es $\det \alpha_{I'\phi(J')}$ y este determinante multiplicado por $\text{sgn } \phi$ es Δ_{IJ} . ■

Como ya es costumbre tediosa, hagamos la observación de que también es válida la Expansión Generalizada de Laplace cuando la hacemos por un conjunto de columnas.

Matrices diagonales y triangulares por bloques

Sea α_{MM} una matriz. Supongamos que existe una partición $M_1 \cup \dots \cup M_t = M$ y que $\alpha_{ij} = 0$ si i y j pertenecen a bloques diferentes. Entonces decimos que α_{MM} es **diagonal por bloques**. Si cada M_i contiene un solo índice entonces decimos que α_{MM} es **diagonal**.

Supongamos ahora que, para la partición $M_1 \cup \dots \cup M_t = M$ se cumple que si $i \in M_p$, $j \in M_q$ y $p < q$ entonces $\alpha_{ij} = 0$. En este caso, decimos que α_{MM} es **triangular por bloques**. Si cada M_i tiene un solo índice entonces decimos que α_{MM} es **triangular**. Es claro de las definiciones que toda matriz diagonal por bloques es triangular por bloques. En particular, toda matriz diagonal es triangular.

4.3 Si α_{MM} es una matriz triangular por bloques entonces,

$$\det \alpha_{MM} = \prod_{i=1}^t \det \alpha_{M_i M_i}.$$

Prueba. Haremos la prueba por inducción en el número de bloques t . Para $t = 1$ la proposición es trivial. Denotemos $M' = M \setminus M_1$. Aplicando la expansión generalizada de Laplace al conjunto de renglones $I = M_1$ obtenemos $\det \alpha_{MM} = \sum_{|J|=|I|} \nabla_{IJ} \Delta_{IJ}$.

Si $J \neq I$ entonces, en α_{IJ} hay una columna $j \notin M_1$ y por lo tanto $j \in M_q$ con $1 < q$. Luego, por definición de matriz triangular, $\forall i \in I = M_1$ se tiene que $\alpha_{ij} = 0$. O sea, la columna j es cero en α_{IJ} e independientemente de la biyección que se escoja para calcular el determinante tenemos $\nabla_{IJ} = 0$. Luego, $\det \alpha_{MM} = \nabla_{II} \Delta_{II} = \det \alpha_{M_1 M_1} \det \alpha_{M' M'}$. La matriz M' es triangular por bloques con un bloque menos y por hipótesis de inducción $\det \alpha_{M' M'} = \prod_{i=2}^t \det \alpha_{M_i M_i}$. ■

Ejercicio 83 ¿Cuáles de las siguientes matrices son triangulares?

$$A = \begin{pmatrix} a & 0 & 0 \\ 1 & b & 0 \\ 1 & 1 & c \end{pmatrix} B = \begin{pmatrix} a & 1 & 1 \\ 0 & b & 1 \\ 0 & 0 & c \end{pmatrix} C = \begin{pmatrix} a & 1 & 1 \\ 0 & b & 0 \\ 0 & 1 & c \end{pmatrix}$$

[191]

Ejercicio 84 Sean $M = \{1, \dots, 5\}$ y α_{MM} una matriz triangular por los bloques $M_1 = \{1, 2\}$, $M_2 = \{3\}$ y $M_3 = \{4, 5\}$. ¿Cuál es el aspecto de α_{MM} en forma gráfica?

[192]

La expansión generalizada de Laplace en forma gráfica

Ahora, precisaremos los signos de las biyecciones en la expansión generalizada de Laplace cuando la matriz está dada en forma gráfica. Como esto es útil solo para calcular el determinante de matrices concretas y hacerlo así es por lo general extremadamente ineficiente y complicado, recomiendo omitir en una primera lectura lo que resta de esta sección.

Si $N = \{1, \dots, n\}$ entonces, entre dos cualesquiera subconjuntos $I, J \subseteq N$ del mismo

cardinal hay una biyección natural que conserva el orden. Para esto, introducimos la notación $\{m_1, m_2, \dots, m_t\}_<$ para señalar que $\forall p \in \{2, \dots, t\}$ se tiene que $i_{p-1} < i_p$. Ahora, si $I = \{i_1, \dots, i_t\}_<$ y $J = \{j_1, \dots, j_t\}_<$ entonces, la biyección es $\phi_1 : I \ni i_p \mapsto j_p \in J$. También, tenemos una biyección natural entre los complementos de I y J . Para esto sean $K = N \setminus I = \{k_1, \dots, k_s\}_<$ y $L = N \setminus J = \{\ell_1, \dots, \ell_s\}_<$ y definamos $\phi_2 : K \ni k_p \mapsto \ell_p \in L$. Luego, para cualesquiera subconjuntos $I, J \subseteq N$ del mismo cardinal la permutación $\phi_I^J = \phi_1 \cup \phi_2$ cumple lo necesario para calcular el signo de un sumando de la expansión generalizada de Laplace, o sea $\phi_I^J(I) = J$ y $\phi_I^J(N \setminus I) = N \setminus J$.

Observese, que la biyección ϕ_1 es la natural de renglones a columnas que se obtiene si en una matriz α_{NN} tachamos todos los renglones con índices en K y todas las columnas con índices en L . De la misma manera ϕ_2 es la biyección de renglones a columnas si tachamos todos los renglones con índices en I y todas las columnas con índices en J .

Calculemos el signo de ϕ_I^J . Al estudiar la expansión (normal) de Laplace en forma gráfica vimos que si $I = \{i\}$ y $J = \{j\}$ entonces, $\phi_I^J = \phi_i^j$ es el ciclo que se muestra a la derecha y por lo tanto $\operatorname{sgn} \phi_i^j = (-1)^{i+j}$ (la regla del “tablero de ajedrez”).

$$\begin{cases} (j, j-1, \dots, i+1, i) & \text{si } j > i \\ (j, j+1, \dots, i-1, i) & \text{si } j < i \end{cases}$$

4.32 $\operatorname{sgn} \phi_I^J = \operatorname{sgn} \phi_{i_1}^{j_1} \operatorname{sgn} \phi_{I \setminus i_1}^{J \setminus j_1}.$

Prueba. Sea $\theta = (\phi_I^J)^{-1} \circ \phi_{I \setminus i_1}^{J \setminus j_1}$. Tenemos que demostrar que $\operatorname{sgn} \theta = (-1)^{i_1 + j_1}$. Para esto, recordemos que $K = N \setminus I = \{k_1, \dots, k_s\}_<$ y $L = N \setminus J = \{\ell_1, \dots, \ell_s\}_<$. Sea p el menor índice tal que $i_1 < k_p$ y sea q el menor índice tal que $j_1 < \ell_q$ (es admisible que

$(k_p, k_{p+1}, \dots, k_{q-1}, i_1)$	$\text{si } p < q$
$(k_{p-1}, k_{p-2}, \dots, k_q, i_1)$	$\text{si } p > q$
\vdots	$\text{si } p = q$

$p = s+1$ y/o $q = s+1$). Usando la definición de θ calculamos que esta permutación es igual a la del recuadro a la izquierda. Luego, θ es siempre un ciclo de longitud $|p - q| + 1$ y por lo tanto $\operatorname{sgn} \theta = (-1)^{p+q}$.

Como i_1 y j_1 son los elementos más pequeños de I y J respectivamente entonces, tenemos que $\{k_1, \dots, k_{p-1}, i_1\}_< = \{1, \dots, p-1, p\}$ y $\{\ell_1, \dots, \ell_{q-1}, j_1\}_< = \{1, \dots, q-1, q\}$ y de aquí finalmente, $p = i_1$ y $q = j_1$. ■

Iterando este resultado obtenemos $\operatorname{sgn} \phi_I^J = \operatorname{sgn} \phi_{i_1}^{j_1} \operatorname{sgn} \phi_{i_2}^{j_2} \dots \operatorname{sgn} \phi_{i_t}^{j_t}$. Observese que $\alpha_{i_r j_r}$ son las entradas en la diagonal de la submatriz α_{IJ} de α_{NM} . Luego, para hallar $\operatorname{sgn} \phi_I^J$ lo que hay que hacer es multiplicar los signos de la regla del tablero de ajedrez de los elementos en la diagonal de α_{IJ} . También se puede hacer, calculando la paridad de $\sum_{i \in I} i + \sum_{j \in J} j$.

Ejemplo. Calculemos el determinante de una matriz α_{NN} del recuadro a la izquierda haciendo la expansión generalizada de Laplace por el conjunto I del segundo y cuarto renglones. Tenemos que recorrer todos los subconjuntos J de dos columnas.

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

Observese, que si la columna 4 no está en J entonces la matriz α_{IJ} tiene dos renglones iguales por lo que todos los sumandos correspondientes en la expansión serán cero. Además, para $J = \{3, 4\}$ la matriz $\alpha_{N \setminus I N \setminus J}$ tiene dos renglones iguales por lo que también el correspondiente sumando es cero.

Solo nos quedan dos sumandos cuando $J = \{1, 4\}$ y cuando $J = \{2, 4\}$. Para ellos podemos extraer del tablero de ajedrez las submatrices del recuadro a la derecha y multiplicando los signos en las diagonales obtenemos $\operatorname{sgn} \phi_I^{\{1,4\}} = -1$ y $\operatorname{sgn} \phi_I^{\{2,4\}} = 1$. Luego, por la expansión generalizada de Laplace el determinante de nuestra matriz es igual a

$$\left| \begin{array}{cc|cc} 2 & 2 & 4 & 1 \\ 2 & 4 & 4 & 2 \end{array} \right| - \left| \begin{array}{cc|cc} 1 & 2 & 5 & 1 \\ 1 & 4 & 5 & 2 \end{array} \right| = 4 \times 4 - 2 \times 5 = 6$$

$J = \{1, 4\}$	$J = \{2, 4\}$
$(\begin{array}{cc} - & + \\ - & + \end{array})$	$(\begin{array}{cc} + & + \\ + & + \end{array})$

donde usamos las barras verticales para ahorrar espacio al denotar los determinantes.

4.5 El rango de una matriz

En esta sección definiremos las bases de una matriz como las submatrices más grandes con determinante diferente de cero. Probaremos que las bases de una matriz definen y están definidas por las bases de los espacios de columnas y renglones. Esto es el fundamento de los diversos métodos de solución de los sistemas de ecuaciones lineales.

Matrices no singulares

Agrupemos en un solo resultado lo que hemos probado para las matrices no singulares.

Caracterización de Matrices No Singulares

Para cualquier matriz cuadrada A las siguientes afirmaciones son equivalentes:

1. A tiene inversa,
3. los renglones de A son LI,
2. A es no singular,
4. las columnas de A son LI.

Prueba. La implicación $1 \Rightarrow 2$ es el resultado 4.26. La implicación $2 \Rightarrow 1$ es el resultado 4.27. La equivalencia $1 \Leftrightarrow 4$ es el resultado 3.21 (página 79). De aquí, $2 \Leftrightarrow 4$ y como los determinantes no cambian por transposición obtenemos $2 \Leftrightarrow 3$. ■

Espacios de columnas y renglones

Al subespacio de \mathbb{K}^N generado por los renglones de la matriz α_{MN} se le llama **espacio de renglones de la matriz**. Aquí tenemos un pequeño problema de lenguaje:

es posible que haya renglones iguales y los conjuntos no distinguen elementos iguales. Por esto, diremos que un **conjunto de renglones es LI** si ellos son distintos dos a dos y es LI en el espacio de renglones de la matriz. Un conjunto de renglones distintos dos a dos que es una base del espacio de renglones lo llamaremos **base de los renglones de α_{MN}** . Análogamente se define el **espacio de columnas de una matriz**, los **conjuntos de columnas LI** y las **bases de las columnas**.

4.34 *El espacio de columnas de α_{MN} es la imagen de la TL de α_{MN} .*

Prueba. Sea $f: \mathbb{K}^N \ni \beta_N \mapsto \alpha_{MN}\beta_N \in \mathbb{K}^M$ la TL de la matriz α_{MN} . La imagen de f es igual a $\langle f(B) \rangle$ donde B es cualquier base de \mathbb{K}^N . En particular si tomamos B igual a la base canónica obtenemos que $f(B)$ es el conjunto de columnas de α_{MN} . ■

Debido a este resultado, frecuentemente al espacio de columnas de una matriz se le llama **imagen de la matriz**. Obviamente, el espacio de renglones de una matriz es la imagen de la TL correspondiente a la transpuesta de la matriz.

Lema de aumento de matrices no singulares

El siguiente lema es parte importante de las demostraciones de los que siguen. Su demostración es clásica e ingeniosa. Este lema es en cierta manera análogo al lema de aumento de conjuntos LI que vimos en el Capítulo 2.

Lema de Aumento de Submatrices No Singulares

4.35 *Sea α_{IJ} una submatriz cuadrada no singular de α_{MN} . Sea $m \in M \setminus I$ tal que el conjunto de renglones indexado por $I \cup m$ es LI. Entonces, existe $n \in N$ tal que la matriz $\alpha_{I \cup m \cup n}$ es no singular.*

Prueba. Denotemos $M' = I \cup m$ y $B_n = \alpha_{M' \cup n}$. Al absurdo, supongamos que $\forall n \in N \setminus J$ se cumple que $\det B_n = 0$. Si $n \in J$ entonces, denotemos por B_n la matriz $\alpha_{M' \cup J}$ a la que artificialmente le agregamos otra columna n . En este caso B_n tiene dos columnas iguales y por lo tanto también $\det B_n = 0$. Para cualquier n , podemos pensar la matriz B_n gráficamente como se muestra en el recuadro. Sean B_{in} los cofactores de las entradas de la última columna en la matriz B_n . Observemos que en la definición de los B_{in} no están involucradas las entradas de la última columna. Luego, B_{in} no depende de n y podemos denotar $B_{in} = \beta_i \in \mathbb{K}$. De la expansión de Laplace por la última columna en la matriz B_n obtenemos:

$$0 = \det B_n = \sum_{i \in M'} \alpha_{in} B_{in} = \sum_{i \in M'} \alpha_{in} \beta_i.$$

$$B_n = \begin{pmatrix} \alpha_{IJ} & \alpha_{In} \\ \alpha_{mJ} & \alpha_{mn} \end{pmatrix}$$

Como esto es válido $\forall n \in \mathbb{N}$ concluimos que $\beta_M, \alpha_{M'N} = 0_N$. Como $\beta_m = \det \alpha_{IJ} \neq 0$ esto significa que los renglones de $\alpha_{M'N}$ están en una combinación lineal nula con coeficientes no todos nulos. Esto contradice la hipótesis de que sus renglones son LI. ■

Bases de una matriz

Sea α_{MN} una matriz. Entre las submatrices cuadradas no singulares de α_{MN} tenemos la relación de contención

$$(\alpha_{I'J'} \subseteq \alpha_{IJ}) \Leftrightarrow (I' \subseteq I \text{ y } J' \subseteq J)$$

Una **base** de α_{MN} es una submatriz no singular maximal por contención.

4.36

Si α_{IJ} es una base de una matriz α_{MN} entonces el conjunto de renglones indexado por I es una base del espacio de renglones de α_{MN} .

Prueba. Sea α_{IJ} una submatriz de α_{MN} . Es conveniente denotar $X = M \setminus I$ y $Y = N \setminus J$ y pensar que α_{MN} es de la forma en el recuadro a la derecha. Supongamos que α_{IJ} es cuadrada y que es una base de α_{MN} . Entonces, por la Caracterización de Matrices No Singulares (4.33) los renglones de α_{IJ} son LI y con más razón los renglones de α_{IN} son LI.

$$\alpha_{MN} = \begin{pmatrix} \alpha_{IJ} & \alpha_{IY} \\ \alpha_{XJ} & \alpha_{XY} \end{pmatrix}$$

Si los renglones de α_{IN} no son una base de los renglones de α_{MN} entonces, existe otro renglón indexado por $m \in X$ tal que el conjunto de renglones indexado por $I \cup m$ es LI y por el Lema de Aumento de Submatrices No Singulares (4.35) existe $n \in Y$ tal que $\alpha_{I \cup m, J \cup n}$ es no singular y esto contradice la suposición de que α_{IJ} es una base. Luego, los renglones de α_{IN} son una base de los renglones. ■

Una consecuencia importante de 4.36 es la siguiente.

Teorema del rango

4.37

Para cualquier matriz los siguientes tres números coinciden:

1. La dimensión de su espacio de renglones.
2. La dimensión de su espacio de columnas.
3. El número de renglones y columnas en cualquier base de la matriz.

Prueba. La igualdad (1) = (3) es consecuencia inmediata de 4.36. La igualdad (2) = (3) también la obtenemos de 4.36 teniendo en cuenta que los determinantes no cambian al transponer una matriz. ■

El **rango** de una matriz es por definición el número común del resultado anterior.

Ejercicio 85 Sean $\mathfrak{E} = \mathfrak{F} \oplus \mathfrak{G}$ espacios vectoriales y $x_i = (y_i, z_i)$ vectores donde $x_i \in \mathfrak{E}$, $y_i \in \mathfrak{F}$ y $z_i \in \mathfrak{G}$. Pruebe que si los y_i son LI entonces los x_i son LI. Esto es lo que se usa en la prueba de 4.36 cuando se usa la frase “con más razón”.

4.6 Sistemas de ecuaciones lineales

Sea $A = \alpha_{NM}$ una matriz y $x_M = x_{M1}$ una columna. El producto de estas matrices es nuevamente una columna $\alpha_{NM}x_{M1} = b_{N1} = b_N$. Si desarrollamos este producto por la definición de producto de matrices entonces obtenemos para cada $i \in N$ la igualdad en el recuadro a la derecha.

$$\sum_{j \in M} \alpha_{ij}x_j = b_i$$

Como ya es usual podemos interpretar la columna x_M como un vector x en el espacio vectorial \mathbb{K}^M y a b_N como un vector b en el espacio \mathbb{K}^N y en estas notaciones la igualdad se escribe en una forma más simple $Ax = b$. Supongamos que A es una matriz fija. Si hacemos variar x en \mathbb{K}^M entonces esta igualdad la podemos pensar como una TL de \mathbb{K}^M en \mathbb{K}^N . Como es lógico, si conocemos x entonces como sabemos multiplicar matrices hallamos $b = Ax$. Más difícil es el problema inverso, si se conoce b entonces, ¿cómo hallar x tal que $Ax = b$?

A una igualdad de la forma $Ax = b$ donde A y b son conocidos y x es incógnita se le llama **sistema de ecuaciones lineales**. ¿Porqué sistema? ¿Porqué ecuaciones en plural si nada más tenemos UNA?. La respuesta a estas preguntas no es gran misterio, es un problema histórico. Cuando sobre la faz de la Tierra aún no vivían las matrices y los vectores, los humanos necesitaban encontrar unos numeritos x_j tales que para cualquier $i \in N$ se cumpla que $\sum_{j \in M} \alpha_{ij}x_j = b_i$. Obsérvese que se necesitan encontrar $|M|$ numeritos. En el caso de que $|N| = 1$ se decía que tenemos que resolver una **ecuación lineal**. Para el caso $|N| > 1$ los numeritos x_j deberían de cumplir todas y cada una de las ecuaciones (para cada $i \in N$) y entonces, se decía que se necesitaba resolver un sistema (conjunto, colección, cualquier cosa que signifique que hay muchas) de ecuaciones lineales. De hecho, para acercarnos más a la verdad, debemos substituir en todo lo dicho los “se decía” por “se dice” en presente. Luego, hay dos formas de ver los sistemas de ecuaciones lineales:

- ◆ Tenemos que encontrar $|M|$ elementos del campo x_j tales que para cualquier i , se cumple la igualdad $\sum_{j \in M} \alpha_{ij}x_j = b_i$,
- ◆ Tenemos que encontrar un vector $x \in \mathbb{K}^M$ tal que $Ax = b$.

Ambas formas son en esencia la misma ya que encontrar un vector x es lo mismo que encontrar todas sus coordenadas x_j . Sin embargo, la elegancia de la segunda forma hace que nuestra manera de pensarlo sea mucho más simple y sistemática (a costo de aprender sobre matrices y vectores). Por ejemplo, si ocurre que la

matriz \mathbf{A} es no singular entonces multiplicando por \mathbf{A}^{-1} a la izquierda obtenemos $\mathbf{Ax} = \mathbf{b} \Rightarrow \mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b} \Rightarrow \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ y como ya sabemos encontrar la matriz inversa y sabemos multiplicar matrices la solución del sistema de ecuaciones lineales está a la mano. Esto no significa que ya sepamos resolver todos los sistemas de ecuaciones lineales ya que para que una matriz sea no singular se necesita primero que sea cuadrada y que además el determinante sea diferente de cero.

Regla de Cramer

Sea $\mathbf{Ax} = \mathbf{b}$ un sistema de ecuaciones lineales. A la matriz \mathbf{A} se le llama **matriz del sistema** y al vector \mathbf{b} lo llamaremos **vector de coeficientes libres**. Denotaremos por $\mathbf{A}(j)$ la matriz obtenida de la matriz del sistema substituyendo la j -ésima columna por el vector de coeficientes libres. Un ejemplo de esta sustitución es el siguiente

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}, \quad \mathbf{A}(2) = \begin{pmatrix} 1 & 7 & 3 \\ 4 & 8 & 6 \end{pmatrix}.$$

Regla de Cramer

4.38

Si la matriz \mathbf{A} es no singular entonces, la j -ésima coordenada x_j del vector \mathbf{x} es igual al determinante de $\mathbf{A}(j)$ dividido entre el determinante de \mathbf{A} .

$$x_j = \frac{\det \mathbf{A}(j)}{\det \mathbf{A}}$$

Prueba. Sea α_{NM} es una matriz no singular. Ya observamos que en este caso necesariamente $\mathbf{x} = \alpha_{NM}^{-1} \mathbf{b}$. Denotemos por β_{ij} las entradas de α_{NM}^{-1} entonces, tenemos $x_j = \beta_{jN} b_N$. Por 4.27 tenemos que $\beta_{jN} = \alpha_{Nj}^* / \det \alpha_{NM}$ y de aquí $x_j \det \alpha_{NM} = b_N \alpha_{Nj}^*$. Para terminar la demostración basta observar que la expresión a la derecha de la igualdad es la expansión de Laplace de $\mathbf{A}(j)$ por la columna j . ■



Gabriel Cramer (Suiza 1704-1752) publicó su famosa regla en el artículo “Introducción al análisis de las curvas algebraicas” (1750). Esta surgió del deseo de Cramer de encontrar la ecuación de una curva plana que pasa por un conjunto de puntos dado. El escribe su regla en un apéndice del artículo y no da prueba alguna para ella. Este es uno de los orígenes históricos del concepto de determinante. Es curioso (y educativo) ver con qué palabras Cramer formula su regla:

“Uno encuentra el valor de cada indeterminada formando n fracciones el común denominador de las cuales tiene tantos términos como las permutaciones de n cosas.”

Cramer continúa explicando como se calculan estos términos como productos de ciertos coeficientes de las ecuaciones y como se determina el signo. El también señala

como los numeradores de las fracciones se pueden encontrar substituyendo ciertos coeficientes de los denominadores por los coeficientes libres del sistema.

Para nosotros, con más de dos siglos de ventaja, es mucho más fácil. Las “ n fracciones” son $\det \mathbf{A}(j) / \det \mathbf{A}$ y el “común denominador” es $\det \mathbf{A}$ (que tiene tantos sumandos como las permutaciones de n cosas).

La Regla de Cramer (4.38) tiene fundamentalmente un interés histórico por ser uno de los orígenes del concepto de determinante. Es necesario recalcar que esta regla solo funciona para el caso de que la matriz es no singular. O sea, cuando en el sistema se tienen tantas incógnitas como ecuaciones y el determinante de la matriz del sistema no es nulo.

Existencia de soluciones

Cuando se tiene un sistema de ecuaciones se deben contestar tres preguntas:

- ◆ ¿Existe una solución?
- ◆ ¿Cómo encontrar una solución en el caso de que exista?
- ◆ ¿Cómo encontrar TODAS las soluciones?

La Regla de Cramer da la respuesta a las tres preguntas para cuando \mathbf{A} es una matriz no singular: la solución existe, es única y se halla por la Regla de Cramer. Ahora responderemos en general la primera pregunta.

Primero, una definición. Si $\mathbf{Ax} = \mathbf{b}$ es un sistema de ecuaciones lineales entonces la **matriz ampliada** del sistema denotada por $(\mathbf{A}|\mathbf{b})$ es la matriz que se obtiene de la matriz del sistema añadiendo el vector columna de coeficientes libres \mathbf{b} .

Teorema de Existencia de Soluciones

4.39

El sistema $\mathbf{Ax} = \mathbf{b}$ tiene una solución si y solo si el rango de la matriz del sistema coincide con el rango de la matriz ampliada.

Prueba. Por definición de rango de una matriz siempre se tiene que $\text{rank } \mathbf{A} \leq \text{rank } (\mathbf{A}|\mathbf{b})$. Que coincidan es equivalente por el Teorema del rango (4.37) a que \mathbf{b} sea combinación lineal de las columnas de \mathbf{A} . El que \mathbf{b} sea combinación lineal de las columnas de \mathbf{A} es equivalente a la existencia de escalares x_j tales que $\alpha_{iN}x_N = b_i$ o sea a que $\mathbf{Ax} = \mathbf{b}$ tenga al menos una solución. ■

Eliminación de ecuaciones dependientes

Lema de Eliminación de Ecuaciones Dependientes

4.40 Sea I el conjunto de índices de una base de renglones de la matriz ampliada $(\alpha_{MN}|b_M)$. Entonces, el conjunto de soluciones de $\alpha_{MN}x_N = b_M$ es exactamente el mismo que el de $\alpha_{IN}x_N = b_I$.

Prueba. Como $\alpha_{MN}x_N = b_M$ tiene más ecuaciones que $\alpha_{IN}x_N = b_I$ entonces cada solución de $\alpha_{MN}x_N = b_M$ es solución de $\alpha_{IN}x_N = b_I$.

Recíprocamente, sea x_N tal que $\alpha_{IN}x_N = b_I$ y $m \in M \setminus I$. El renglón $(\alpha_{mN}|b_m)$ es combinación lineal de los indexados por I . Luego existen escalares λ_i tales que se cumplen las igualdades a la derecha. Multiplicando la primera igualdad por x_N y usando la definición de x_N obtenemos $\alpha_{mN}x_N = \lambda_I \alpha_{IN}x_N = \lambda_I b_I = b_m$ y por lo tanto x_N satisface todas las ecuaciones del sistema $\alpha_{MN}x_N$. ■

$$\begin{aligned}\alpha_{mN} &= \lambda_I \alpha_{IN} \\ b_m &= \lambda_I b_I\end{aligned}$$

Otra manera, más algorítmica, de ver este lema es que si tenemos un sistema de ecuaciones $\alpha_{MN}x_N = b_M$ y un renglón de este sistema es combinación lineal de los demás entonces, debemos eliminar la ecuación correspondiente a este renglón. El Lema de Eliminación de Ecuaciones Dependientes (4.40) nos garantiza que esta operación no altera el conjunto de soluciones. Repitiendo esta operación una cierta cantidad de veces, obtenemos una matriz ampliada con renglones LI.

El núcleo y la imagen de una matriz

Sea α_{MN} una MN -matriz. Ella es la matriz de la TL $f : \mathbb{K}^N \ni x_N \mapsto \alpha_{MN}x_N \in \mathbb{K}^M$. Luego, podemos definir la **imagen de la matriz** α_{MN} como $\text{Im } \alpha_{MN} = \text{Im } f$ y el **núcleo de la matriz** α_{MN} como $\ker \alpha_{MN} = \ker f$. Pasemos ahora a analizar la imagen. Por definición de imagen tenemos $\text{Im } \alpha_{MN} = \{\beta_M \mid \exists x_N \quad \alpha_{MN}x_N = \beta_M\}$. O sea, $\text{Im } \alpha_{MN}$ es el conjunto de vectores β_M tales que el sistema de ecuaciones lineales $\alpha_{MN}x_N = \beta_M$ tiene al menos una solución. Tenemos $\alpha_{MN}x_N = \sum_{i \in N} \alpha_{Mi}x_i$ que es una combinación lineal de las columnas. Luego, $\text{Im } \alpha_{MN}$ es el subespacio generado por las columnas de la matriz α_{MN} . Además, si γ_N es una solución de $\alpha_{MN}x_N = \beta_M$ entonces, 3.30 (página 86) nos dice que el conjunto de todas sus soluciones es el subespacio afín $\gamma_N + \ker \alpha_{MN}$. Resumamos todo lo dicho en 4.41 para referencias futuras.

- 4.41**
- ◆ El subespacio $\ker \alpha_{MN}$ es el conjunto de soluciones de $\alpha_{MN}x_N = 0_M$.
 - ◆ El subespacio $\text{Im } \alpha_{MN}$ es el generado por las columnas de α_{MN} .
 - ◆ Si γ_N es una solución de $\alpha_{MN}x_N = \beta_M$ entonces, el conjunto de todas sus soluciones es el subespacio afín $\gamma_N + \ker \alpha_{MN}$.

Bases del subespacio afín de soluciones

Ahora daremos la solución general de los sistemas de ecuaciones lineales. Sea $\alpha_{MN}x_N + b_M$ un sistema de ecuaciones. Primero, podemos suponer que la matriz del sistema α_{MN} tiene el mismo rango que la matriz ampliada $[\alpha_{MN}|b_M]$ porque si no entonces, el conjunto de soluciones es vacío. Segundo, podemos suponer que la matriz ampliada $[\alpha_{MN}|b_M]$ tiene renglones linealmente independientes porque si no, entonces podemos descartar aquellos que son combinación lineal de otros.

Luego existe un conjunto de columnas $J \subseteq M$ tal que α_{MJ} es una base de la matriz ampliada $[\alpha_{MN}|b_M]$. Denotando por Y al conjunto de columnas que no están en la base

$\alpha_{MJ}x_J + \alpha_{MY}x_Y = b_M$ podemos representar nuestro sistema de ecuaciones lineales como se muestra en el recuadro a la izquierda.

Aquí las incógnitas x_J y x_Y son aquellas que se corresponden con las columnas en J y Y respectivamente. Por ejemplo, en el sistema de ecuaciones

$$\begin{aligned} 2x_1 + 1x_2 + 1x_3 + 0x_4 &= 5 \\ 3x_1 + 2x_2 + 0x_3 + 1x_4 &= 5 \end{aligned}$$

podemos tomar J al conjunto de las dos primeras columnas y Y al conjunto de la tercera y cuarta columnas. Luego a este sistema lo podemos escribir como

$$\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}.$$

Ahora, como la matriz α_{MJ} tiene inversa, lo que hacemos es simplemente despejar x_J y obtenemos $x_J = \alpha_{MJ}^{-1}b_M - \alpha_{MJ}^{-1}\alpha_{MY}x_Y$. En nuestro ejemplo obtenemos

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 5 \end{pmatrix} - \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2x_3 + x_4 + 5 \\ 3x_3 - 2x_4 - 5 \end{pmatrix}$$

Esto describe en forma parámetrica el conjunto de todas las soluciones, para cualquier vector x_Y hay una solución $(\alpha_{MJ}^{-1}b_M - \alpha_{MJ}^{-1}\alpha_{MY}x_Y, x_Y)$ del sistema. Otra manera de describir el conjunto solución es ponerlo en la forma descrita en 4.41. Para encontrar una solución ponemos $x_Y = \mathbf{0}$ y en nuestro ejemplo el vector solución es $(5, -5, 0, 0)$. Para encontrar el núcleo de la matriz del sistema ponemos $b_M = \mathbf{0}$. Para encontrar una base del núcleo recorremos con x_Y a la base canónica de \mathbb{K}^Y . En nuestro ejemplo

$$(x_3, x_4) = (1, 0) \Rightarrow (x_1, x_2) = (3, -2)$$

$$(x_3, x_4) = (0, 1) \Rightarrow (x_1, x_2) = (6, -7)$$

y finalmente obtenemos que el conjunto de soluciones (x_1, x_2, x_3, x_4) es el subespacio afín

$$(5, -5, 0, 0) + \rho(1, 0, 3, -2) + \tau(0, 1, 6, -7)$$

donde ρ y τ son escalares cualesquiera.

4.7 Método de eliminación de Gauss

La idea del método de eliminación de Gauss es realizar ciertas transformaciones de las matrices que no cambian lo que se pretende calcular y que convierte a las matrices en otras con muchas entradas iguales a cero. Este es el método más universal y eficiente (al menos manualmente) para calcular los determinantes, el rango, el núcleo, las matrices inversas, etc. Aquí, estudiaremos brevemente este método.

Transformaciones elementales

Sea α_{MN} una matriz. Sean α_{iN} , α_{jN} dos renglones de α_{MN} y $\lambda \in \mathbb{K}$ un escalar. Denotemos $\beta_{jN} = \lambda\alpha_{iN} + \alpha_{jN}$ y sea β_{MN} la matriz obtenida de α_{MN} reemplazando el renglón α_{jN} por la N-ada β_{jN} . A la operación que dados α_{MN} , i , j y λ obtenemos β_{MN} se le llama **transformación elemental de los renglones**. Otra manera útil de pensar las transformaciones elementales de los renglones es que en la matriz α_{MN} al renglón α_{jN} le sumamos el renglón α_{iN} multiplicado por λ . De forma análoga, se definen las **transformaciones elementales de las columnas**. Una **transformación elemental** es de renglones o de columnas.

A.4.2 *Las transformaciones elementales no cambian los determinantes.*

Prueba. Sean α_{MM} , β_{MM} y γ_{MM} matrices que se diferencian solo en el renglón indexado por j para el cual se tiene que $\beta_{jM} = \lambda\alpha_{iM} + \alpha_{jM}$ y $\gamma_{jM} = \alpha_{iM}$. Como el determinante es un funcional lineal de los renglones tenemos $\det \beta_{MM} = \lambda \det \gamma_{MM} + \det \alpha_{MM}$ y como la matriz γ_{MM} tiene dos renglones iguales entonces, $\det \beta_{MM} = \det \alpha_{MM}$. La prueba termina al recordar que el determinante no cambia por transposición de la matriz. ■

Las bases y el rango de una matriz dependen exclusivamente de los determinantes de sus submatrices y por lo tanto no son afectados por las transformaciones elementales. Sin embargo, las trasformaciones elementales de las columnas cambian los núcleos y el subespacio afín solución de un sistema de ecuaciones lineales. Para las transformaciones elementales de los renglones la situación es diferente.

A.4.3 *Las transformaciones elementales de los renglones de la matriz ampliada no cambian el subespacio afín solución de un sistema de ecuaciones lineales.*

Prueba. Sea $\alpha_{MN}x_N = b_M$ un sistema de ecuaciones lineales. Sea $\beta_{MN}x_N = c_M$ otro sistema que se diferencia solo en la ecuación j para la cual se tiene $\beta_{jN} = \lambda\alpha_{iN} + \alpha_{jN}$ y $c_j = \lambda b_i + b_j$. Si γ_N es una solución del primer sistema entonces, $\beta_{jN}\gamma_N = \lambda\alpha_{iN}\gamma_N + \alpha_{jN}\gamma_N = \lambda b_i + b_j = c_j$ por lo que γ_N es una solución del segundo sistema. Si γ_N es una solución del segundo sistema entonces, $\alpha_{jN}\gamma_N = \beta_{jN}\gamma_N - \lambda\alpha_{iN}\gamma_N = c_j - \lambda b_i = b_j$ por

lo que γ_N es una solución del primer sistema. ■

Ejemplo

El método de Gauss se puede precisar en todo detalle para convertirlo en un algoritmo programable en una computadora. Pero como tratamos de enseñar a humanos y no a computadoras la mejor manera no es dar todos los detalles sino dejar a la creatividad individual y a la imitación de ejemplos el como aplicar las transformaciones elementales.

La matriz ampliada siguiente arriba a la izquierda representa un sistema de ecuaciones lineales. Despues de las 4 transformaciones elementales de los renglones que se muestran obtenemos la matriz de abajo a la derecha. La solución del sistema de esta última es obvia.

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 2 & 3 & 4 & 2 \\ 3 & 3 & 1 & 1 \end{array} \right) \xrightarrow{r_2 := r_2 - 2r_1} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 4 & -4 \\ 3 & 3 & 1 & 1 \end{array} \right) \xrightarrow{r_3 := r_3 - 3r_1} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 4 & -4 \\ 0 & 0 & 1 & -8 \end{array} \right)$$

$$\xrightarrow{r_2 := r_2 - 4r_3} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 28 \\ 0 & 0 & 1 & -8 \end{array} \right) \xrightarrow{r_1 := r_1 - r_2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -25 \\ 0 & 1 & 0 & 28 \\ 0 & 0 & 1 & -8 \end{array} \right)$$

Aquí las transformaciones elementales realizadas están marcadas en las flechas. Por ejemplo, la primera es $r_2 := r_2 - 2r_1$ lo que significa que al segundo renglón le sumamos el primero multiplicado por -2 . Obsérvese que después de dos transformaciones ya vemos que el determinante de la matriz del sistema es 1 porque en una matriz triangular el determinante es el producto de las entradas diagonales.

El caso general

Para el cálculo de los determinantes no hay mucho más que decir. Podemos utilizar transformaciones elementales de columnas y renglones para convertir nuestra matriz cuadrada en matriz triangular. Si en el proceso nos aparece un renglón o columna cero entonces el determinante es cero.

Para resolver los sistemas de ecuaciones lineales trabajamos con la matriz ampliada y solo podemos realizar transformaciones elementales de los renglones. Podemos además multiplicar un renglón por un escalar porque esto no afecta la ecuación correspondiente al renglón. Si nos aparece un renglón cero podemos descartarlo como combinación lineal de los otros. Si nos aparece un renglón que es cero en la matriz del sistema pero su coeficiente libre no es cero entonces el sistema no tiene solución porque el rango de la matriz ampliada es mayor que el rango de la matriz del sistema. Finalmente, si en

algún momento nos aparece una configuración del tipo

$$\left(\begin{array}{cccccc|c} a & * & \cdots & * & * & * & * \\ 0 & b & \cdots & * & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & c & * & * & * \end{array} \right)$$

donde a, b, \dots, c son diferentes de cero entonces las primeras columnas forman una base de la matriz y tenemos más incógnitas que ecuaciones. En este caso debemos seguir diagonalizando la parte correspondiente a las primeras columnas hasta obtener

$$\left(\begin{array}{cccccc|c} 1 & 0 & \cdots & 0 & * & * & * \\ 0 & 1 & \cdots & 0 & * & * & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & * & * & * \end{array} \right).$$

Para este sistema procedemos como en la sección anterior y pasamos restando las incógnitas que sobran como parámetros hacia la parte derecha del sistema de ecuaciones. Tomemos el mismo ejemplo de la sección anterior

$$\left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 5 \\ 3 & 2 & 0 & 1 & 5 \end{array} \right) \xrightarrow{r_2 := 3r_2 - 2r_1} \left(\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 5 \\ 0 & 1 & -3 & 2 & -5 \end{array} \right) \xrightarrow{r_1 := \frac{r_1 - r_2}{2}} \left(\begin{array}{cccc|c} 1 & 0 & 2 & -1 & 5 \\ 0 & 1 & -3 & 2 & -5 \end{array} \right)$$

y por lo tanto la solución de este sistema es $x_1 = 5 - 2x_3 + x_4$ y $x_2 = -5 + 3x_3 - 2x_4$ donde x_3 y x_4 pueden tomar cualquier valor.

Aquí es cuando el lector se preguntará ¿Para qué diagonalizar la primera y segunda columna si ya la tercera y cuarta están en forma diagonal? Pues tiene toda la razón, simplemente está escogiendo otra base de la matriz. La solución del sistema se obtiene directamente de la matriz ampliada original en la forma $x_3 = 5 - 2x_1 - x_2$ y $x_4 = 5 - 3x_1 - 2x_2$ donde x_1 y x_2 pueden tomar cualquier valor. Que es mejor es cuestión de gustos u otras necesidades. Por ejemplo, si se necesita substituir x_1 y x_2 en otras fórmulas entonces, la mejor solución es la primera. Si por el contrario se necesita substituir x_3 y x_4 en otras fórmulas entonces, la mejor solución es la segunda.

Solución de ecuaciones matriciales, matriz inversa

Si tenemos varios sistemas de ecuaciones lineales $\alpha_{MN}x_{N1} = b_{M1}, \dots, \alpha_{MN}x_{N\ell} = b_{M\ell}$ todos con la misma matriz del sistema α_{MN} entonces, podemos denotar $L = \{1, \dots, \ell\}$ y escribirlos todos en la forma $\alpha_{MN}x_{NL} = b_{ML}$. Esta ecuación matricial tiene la matriz ampliada $[\alpha_{MN}|b_{ML}]$ y podemos aplicarle a esta matriz la eliminación de Gauss para resolver todos nuestros sistemas al mismo tiempo. Esto lo podremos hacer porque en la eliminación de Gauss no se reordenan columnas y solo se aplican transformaciones elementales a los renglones, así que no nos importa cuantas columnas halla después de la barra vertical.

En particular, si $M = N$ entonces, la única solución posible (si existe) a la ecuación matricial $\alpha_{MM}x_{MM} = I_{MM}$ sería $x_{MM} = \alpha_{MM}^{-1}$. Esta es la forma en que con el método

de eliminación de Gauss se calculan las matrices inversas. Por ejemplo:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{\substack{r_2 := r_2 - 2r_1 \\ r_3 := r_3 - 3r_1}} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 4 & -2 & 1 & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{array} \right) \xrightarrow{r_2 := r_2 - 4r_3}$$

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 10 & 1 & -4 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{array} \right) \xrightarrow{r_1 := r_1 - r_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -9 & -1 & 4 \\ 0 & 1 & 0 & 10 & 1 & -4 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{array} \right)$$

y por lo tanto

$$\left(\begin{array}{ccc} 1 & 1 & 0 \\ 2 & 3 & 4 \\ 3 & 3 & 1 \end{array} \right) \left(\begin{array}{ccc} -9 & -1 & 4 \\ 10 & 1 & -4 \\ -3 & 0 & 1 \end{array} \right) = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right).$$

Ejercicio 86 Tome un sistema de ecuaciones lineales y resuelvalo por el método de eliminación de Gauss. Repita cuantas veces le sea necesario.

No está de más recalcar aquí que el método de eliminación de Gauss funciona en espacios vectoriales sobre cualquier campo sea este \mathbb{Q} , \mathbb{R} , \mathbb{C} , \mathbb{Z}_p u otro campo cualquiera. Solo hay que realizar las operaciones de suma, producto y división como están definidas en el campo en particular.



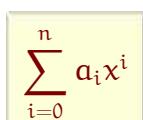
Capítulo quinto

Polinomios

 A pesar de que técnicamente, el estudio de los polinomios no es parte del Álgebra Lineal, ellos son una herramienta ineludible para la clasificación de los operadores lineales. Es por esto que necesitamos aprender algunas de las cosas más básicas sobre ellos.

5.1 Polinomios sobre campos

Hay muchas maneras de ver los polinomios. La manera más sencilla de verlos es que un **polinomio de grado n** en la literal x es una expresión formal del tipo en el recuadro a la derecha donde los **coeficientes** a_0, \dots, a_n son elementos de cierto campo \mathbb{K} y $a_n \neq 0$. Al coeficiente a_n se le llama **coeficiente principal** del polinomio. Todos los elementos del campo son polinomios de grado cero. Dos polinomios se consideran iguales solo cuando todos sus coeficientes son iguales.



En la definición anterior no encaja el **polinomio cero** cuyos coeficientes son todos cero. Es el único polinomio con coeficiente principal cero y su grado no está bien definido. Algunos de los resultados formulados en esta sección, obviamente no son válidos para el polinomio cero. El lector interesado en cuales sí y cuales no, deberá pensarlo por si mismo en cada caso.

Suma y producto de polinomios

Aunque el lector seguramente conoce las definiciones de suma y producto de polinomios, nos parece apropiado recordar el porqué de las mismas. Si interpretamos a x como un elemento del campo \mathbb{K} entonces, $\sum a_i x^i$ también es un elemento del campo \mathbb{K} . Esto quiere decir que un polinomio define una función $\mathbb{K} \ni x \mapsto \sum a_i x^i \in \mathbb{K}$ y en esta interpretación x no es una literal sino una variable. Siendo esta interpretación de los polinomios fundamental, necesitamos que la suma y producto de polinomios concuerde con la definición de suma y producto de funciones $(f + g)(x) = f(x) + g(x)$,

$(fg)(x) = f(x)g(x)$. Por esto, tenemos

$$\sum_{i=0}^n a_i x^i + \sum_{i=0}^n b_i x^i = \sum_{i=0}^n (a_i x^i + b_i x^i) = \sum_{i=0}^n (a_i + b_i) x^i$$

donde la primera igualdad se da por asociatividad y commutatividad y la segunda por distributividad. O sea, interpretamos al polinomio como la imagen por la función de evaluación de la variable x que toma sus valores en el campo. Para el producto, usando la forma general de la ley distributiva tenemos

$$\left(\sum_{i=0}^n a_i x^i \right) \left(\sum_{j=0}^m b_j x^j \right) = \sum_{i=0}^n \sum_{j=0}^m a_i b_j x^{i+j} = \sum_{k=0}^{n+m} \left(\sum_{i=\max(0, k-m)}^{\min(n, k)} a_i b_{k-i} \right) x^k$$

donde la segunda igualdad se obtiene haciendo el cambio de variable $k = i + j$ y usando asociatividad, commutatividad y distributividad. Se puede saber sumar y multiplicar polinomios sin saberse estas fórmulas. Lo importante es saber aplicar sistemáticamente asociatividad, commutatividad y distributividad para obtener el resultado deseado. El conjunto de todos los polinomios sobre un campo arbitrario \mathbb{K} forma un anillo commutativo (para ser campo solo le falta la existencia de inversos multiplicativos). A este anillo se lo denota por $\mathbb{K}[x]$.

El lector debe observar en la fórmula del producto que el coeficiente principal del producto de dos polinomios es $a_n b_m x^{n+m}$. Como en un campo no hay divisores de cero entonces $a_n b_m \neq 0$ y por lo tanto *el grado del producto de dos polinomios es igual a la suma de sus grados*.

La función de evaluación

Sea $p(x) = \sum_{i=0}^n a_i x^i$ un polinomio en $\mathbb{K}[x]$. Sea b un elemento arbitrario del campo. Obviamente $\sum_{i=0}^n a_i b^i$ es también un elemento del campo ya que la suma, la multiplicación y las potencias están bien definidas en un campo arbitrario. Es natural denotar $p(b) = \sum_{i=0}^n a_i b^i$. Esto nos da una función $\mathbb{K} \ni b \mapsto p(b) \in \mathbb{K}$ llamada la **función de evaluación** del polinomio p .

Recíprocamente, diremos que una función $f : \mathbb{K} \rightarrow \mathbb{K}$ es **polinomial** si existe un polinomio $p(x) \in \mathbb{K}[x]$ tal que f es la función de evaluación del polinomio p , o sea que $\forall b \in \mathbb{K}$ se tiene que $f(b) = p(b)$.

Identificar los polinomios con las funciones polinomiales es un craso error. Así por ejemplo, hay solo 4 funciones de \mathbb{Z}_2 en \mathbb{Z}_2 pero hay una cantidad infinita de polinomios en $\mathbb{Z}_2[x]$. Sin embargo, si el campo es infinito esto no sucede.



Un polinomio de grado n está predeterminado por su evaluación en $n+1$ diferentes elementos del campo.

Prueba. Sea $p(x) = \sum_{i=0}^n a_i x^i$ y b_0, \dots, b_n diferentes elementos del campo entonces,

en forma matricial tenemos

$$\begin{pmatrix} b_0^n & b_0^{n-1} & \cdots & b_0 & 1 \\ b_1 & b_1 & \cdots & b_1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_n & b_n & \cdots & b_n & 1 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{pmatrix} = \begin{pmatrix} p(b_0) \\ p(b_1) \\ \vdots \\ p(b_n) \end{pmatrix}.$$

Si conocemos las evaluaciones $p(b_0), \dots, p(b_n)$ entonces, para encontrar los coeficientes a_0, \dots, a_n tenemos que resolver este sistema de ecuaciones lineales. La matriz de este sistema es una matriz de Vandermonde cuyo determinante es diferente de cero si y solo si todos los b_i son diferentes (véase el ejercicio 82). Esto quiere decir que el sistema tiene una solución única, o sea los a_i están predeterminados por los $p(b_i)$. ■

De este resultado se deduce fácilmente, que en el caso de campos infinitos, la correspondencia entre funciones polinomiales y polinomios es biyectiva y más aún, que esta correspondencia es un isomorfismo de anillos. Por esto, es frecuente que se identifiquen los polinomios reales con las funciones polinomiales reales.



Un problema que frecuentemente aparece en aplicaciones de las matemáticas es que se tiene una función real de la cual no se sabe mucho salvo su valor en n puntos. Entonces esta función se “aproxima” con el único polinomio de grado $n - 1$ que pasa por estos puntos. Hay diferentes fórmulas para esto pero todas son equivalentes a sacar la matriz inversa de la matriz de Vandermonde. Estas fórmulas se diferencian en que tan “buenas” y/o “estables” son computacionalmente.

División de polinomios

División con Resto de Polinomios

5.2

Sea q un polinomio de grado al menos 1 y sea p otro polinomio.

Existen polinomios c y r tales que:

1. $p = cq + r$
2. El grado de r es estrictamente menor que el grado de q .

Prueba. Sea $p = \sum a_i x^i$ un polinomio de grado n y $q = \sum b_i x^i$ un polinomio de grado $m \geq 1$. Si $n < m$ entonces poniendo $c = 0$ y $r = p$ terminamos.

Supongamos $m \leq n$. Sea c_1 como en el recuadro a la izquierda. Sacando cuentas nos podemos convencer de que el grado de $r_1 = p - c_1 q$ es estrictamente menor que el grado de p . Si el grado de r_1 es menor que el grado de q entonces ya terminamos, si no entonces, haciendo cálculos similares podemos escribir $r_1 = c_2 q + r_2$ y vamos disminuyendo el grado de r_i hasta que este sea menor que el grado de q .

Luego, existe un i tal que $p = (c_1 + \dots + c_i)q + r_i$ y el grado de r_i es estrictamente menor que el grado de q . ■

Al polinomio c se le llama **cociente** de la división de p entre q . Al polinomio r se le llama **resto** de la división de p entre q .

$$c_1 = \frac{x^{n-m} a_n}{b_m}$$

Divisibilidad

Sean p y q dos polinomios. Si existe un polinomio c tal que $p = cq$ entonces se dice que q divide a p , o que q es un divisor de p , o que p es un múltiplo de q . Obviamente, cualquier polinomio no cero de grado cero divide a cualquier otro polinomio (en un campo hay inversos).



Para denotar la relación de divisibilidad entre polinomios usaremos los símbolos “ \dashv ” y “ \vdash ”. O sea, $p \dashv q$ significa que p divide a q . Por otro lado $p \vdash q$ se lee como p es múltiplo de q .



La tradición exige que la relación de divisibilidad se denote por el símbolo “ $|$ ”. Nosotros no seguiremos la tradición por dos razones. Primero en este libro ese símbolo se utiliza sistemáticamente para denotar “tal que”. Segundo, ese símbolo sugiere que la relación es simétrica y no lo es para nada, todo lo contrario. Sin embargo, es importante que el lector conozca esto para poder leer satisfactoriamente otros libros.

5.3 *Si $p \dashv q$ y $q \dashv p$ entonces existe un elemento del campo α tal que $p = \alpha q$.*

Prueba. Tenemos que existen polinomios a y b tales que $p = aq$ y $q = bp$ y por lo tanto $p = abp$. El grado de la parte derecha de esta igualdad es la suma de los grados de a , b y p . Así vemos que necesariamente los polinomios a y b tienen que tener grado cero, o sea, son elementos del campo. ■

La relación de divisibilidad entre polinomios es obviamente reflexiva y transitiva. pero no es ni simétrica ni antisimétrica.

Debido a que en un campo hay siempre inversos multiplicativos, todo polinomio p se expresa de forma única como $\alpha p'$ donde α es el coeficiente principal y p' es un polinomio mónico o sea, un polinomio cuyo coeficiente principal es igual a 1. Del resultado anterior se puede deducir fácilmente que la relación de divisibilidad entre polinomios monicos es antisimétrica y por lo tanto es una relación de orden (parcial).

Este simple hecho será fundamental para nosotros ya que frecuentemente usaremos la antisimetría para probar que dos polinomios monicos son iguales.

Además, esto significa que el conjunto de polinomios con la relación de divisibilidad es un conjunto parcialmente ordenado y por lo tanto a él se aplican los conceptos tales como máximo y mínimo; cotas superiores e inferiores; elementos maximales y minimales; supremos e ínfimos (véase el glosario).

Especial interés tienen el supremo y el ínfimo los cuales para este caso se traducen como el mínimo común múltiplo y el máximo común divisor.

Ejercicio 87 Encuentre todos los divisores mónicos del polinomio $(x + 1)^2(x + 2)$. Organícelos de tal manera que se vea claramente la relación de orden entre ellos.

Factores y raíces

Diremos que q es un **factor** de p si $q \mid p$, el grado de q es al menos uno y q es mónico. Los ceros de la función de evaluación son las **raíces** del polinomio p , o sea son los elementos del campo b tales que $p(b) = 0$. Al conjunto de todas las raíces de p lo llamaremos **núcleo** de p y lo denotaremos por $\ker p$ (en inglés “núcleo” es “kernel”). Las raíces y los factores de un polinomio están enlazados por el siguiente resultado:

5.4 *Para que b sea una raíz de p es necesario y suficiente que $(x - b)$ sea un factor de p .*

Prueba. Dividamos con resto p entre $(x - b)$. Sean c y r tales que $p(x) = c(x)(x - b) + r$. Como $(x - b)$ es de grado uno r tiene que ser de grado cero. Evaluando en b obtenemos $p(b) = c(b)(b - b) + r = r$. Luego, si b es raíz entonces, $r = 0$ y recíprocamente. ■

Sea b una raíz del polinomio p . Si $n \geq 1$ es el mayor natural tal que $(x - b)^n$ es factor de p entonces a n se le llama **multiplicidad** de la raíz b . Es muy incómodo trabajar con el concepto de multiplicidad. Por ejemplo, tomemos la afirmación: “Si b_1 y b_2 son dos raíces del polinomio p entonces $(x - b_1)(x - b_2)$ es un factor de p ”. Esta afirmación es cierta no solo para cuando $b_1 \neq b_2$ sino también cuando son iguales pero la raíz tiene multiplicidad mayor que 2.



Es mucho más cómodo pensar que si una raíz tiene multiplicidad n entonces hay n “diferentes” raíces todas del mismo “valor”. Este abuso del lenguaje será común en este libro y a partir de ahora no tendremos necesidad de usar continuamente el concepto de multiplicidad. Le dejamos al lector interesado la desagradable tarea, de ajustar los hechos expuestos a un lenguaje más riguroso.

Ahora el núcleo de un polinomio no es exactamente un conjunto sino una “colección” de elementos del campo en la cual puede haber elementos repetidos. Así por ejemplo tenemos $\text{Ker}(x^3 - 6x^2 + 9x - 4) = \{1, 1, 4\}$. Ajustada nuestra terminología, podemos establecer una importante consecuencia del resultado anterior.

5.5 *Un polinomio de grado $n \geq 1$ tiene a lo más n raíces.*

Prueba. Si el polinomio p tiene como raíces a b_1, \dots, b_{n+1} entonces p tiene, por 5.4, como factor a $\prod_{i=1}^{n+1} (x - b_i)$ que es un polinomio de grado $n + 1$. Esto contradice que p tiene grado n . ■

Ejercicio 88 Sea G un subgrupo finito del grupo multiplicativo de un campo. Calcule el producto de todos los elementos de G . [192]

Ejercicio 89 Demuestre que todo subgrupo finito con q elementos del grupo multiplicativo de un campo es isomorfo a $(\mathbb{Z}_q, +)$. [192]

Ideales de polinomios

Un conjunto I de polinomios se llama **ideal** si se cumplen las dos siguientes propiedades:

- ◆ Si $p, q \in I$ entonces $p + q \in I$.
- ◆ Si $p \in I$ y $r \in \mathbb{K}[x]$ entonces $rp \in I$.

En otras palabras, la suma es una operación binaria dentro del ideal y cualquier múltiplo de un elemento del ideal también está en el ideal. Los ideales más sencillos son los que se forman tomando todos los múltiplos de un polinomio fijo p . Si qp y $q'p$ son dos tales múltiplos entonces $qp + q'p = (q + q')p$ también es un múltiplo de p . Esto demuestra la propiedad 1. La propiedad 2 se cumple obviamente. Estos ideales se les llama **ideales principales**. Lo asombroso es que todo ideal es así.

5.6 *Todo ideal de polinomios es principal.*

Prueba. Sea I un ideal. Sea ahora m un polinomio no nulo de grado mínimo tal que $m \in I$ y denotemos por el conjunto de todos los múltiplos de m o sea, $I_m = \{\alpha m \mid \alpha \in \mathbb{K}[x]\}$. Por definición de ideal se tiene que $I_m \subseteq I$. Probemos que $I_m \supseteq I$. Efectivamente, si $g \in I$ entonces dividiendo g entre m obtenemos polinomios c, r tales que $g = cm + r$ donde el grado de r es menor que el grado de m . Tenemos $r = g - cm$ y por definición de ideal $r \in I$. De la minimalidad del grado de m obtenemos $r = 0$. y por lo tanto $g = cm \in I_m$. Esto prueba que $I_m = I$ o sea, que I es principal. ■

Observese que fácilmente podemos definir los ideales en cualquier anillo commutativo. Sin embargo, no siempre cualquier ideal es principal. Este es el caso por ejemplo, para el anillo de polinomios de dos variables $\mathbb{K}[x, y]$.

Una primera consecuencia de 5.6 es el Teorema de Bezout, el cual tendremos muchísimas oportunidades para utilizarlo.

Teorema de Bezout

5.7 *Sean p y q dos polinomios sin factores comunes. Existen polinomios α y β tales que $\alpha p + \beta q = 1$.*

Prueba. Denotemos por $I_{pq} = \{\alpha p + \beta q \mid \alpha, \beta \in \mathbb{K}[x]\}$. Probemos que I_{pq} es un ideal.

Veamos primero que la suma de elementos de I_{pq} está en I_{pq} . Efectivamente,

$$(\alpha p + \beta q) + (\alpha' p + \beta' q) = (\alpha + \alpha') p + (\beta + \beta') q = \alpha'' p + \beta'' q$$

donde $\alpha'' = (\alpha + \alpha')$ y $\beta'' = (\beta + \beta')$. Ahora, comprobemos que los múltiplos de los elementos de I_{pq} están en I_{pq} . Tenemos, $\gamma (\alpha p + \beta q) = \gamma \alpha p + \gamma \beta q = \alpha' p + \beta' q$ donde $\alpha' = \gamma \alpha$ y $\beta' = \gamma \beta$ y con esto concluimos que I_{pq} es un ideal.

Como todo ideal de polinomios es principal, existe m tal que $I_{pq} = \{ \alpha m \mid \alpha \in \mathbb{K}[x] \}$. Como $p, q \in I_{pq}$, existen polinomios α y β tales que $p = \alpha m$ y $q = \beta m$. Como p y q no tienen factores comunes, esto significa que m es de grado cero y por lo tanto $I_{pq} = \mathbb{K}[x]$. En particular, $1 \in I_{pq}$. ■



Etienne Bezout (Francia, 1730-1783). Famoso en su época sobre todo por los seis volúmenes de su libro de texto “*Cours complet de mathématiques à l'usage de marine et de l'artillerie*” que por muchos años fueron los libros que estudiaban los que aspiraban a ingresar a la “École Polytechnique”. Su investigación matemática la dedicó al estudio de los determinantes y de las soluciones de ecuaciones polinomiales.

Ejercicio 90 Demuestre el teorema de Bezout para \mathbb{Z} : Sean p y q dos enteros sin factores comunes. Entonces, existen dos enteros α y β tales que $\alpha p + \beta q = 1$. [193]

Ejercicio 91 Demuestre que si p y q son dos polinomios tales que el máximo común divisor de p y q es r entonces, existen polinomios α y β tales que $\alpha p + \beta q = r$. [193]

Unicidad de la factorización en irreducibles.

Diremos que un factor q es **factor propio** del polinomio mónico p , si $p \neq q$.

Un polinomio se le llama **irreducible** si este es mónico, tiene grado al menos 1 y no tiene factores propios. En otras palabras, cuando no se puede descomponer no trivialmente en producto de dos factores. Cualquier polinomio p se puede descomponer como $\alpha p_1 p_2 \dots p_n$ donde α es su coeficiente principal y $p_1 \dots p_n$ son polinomios irreducibles. *La prueba de esto es obvia.* Si un polinomio no es irreducible puedo descomponerlo en producto de dos factores. Si estos factores no son irreducibles puedo descomponer en factores cada uno de ellos y así sucesivamente llegamos a factores irreducibles.

Si p y q son dos polinomios y r es un factor irreducible de pq entonces r es un factor de p o de q .

Prueba. Supongamos que r no es un factor de p y probemos que es un factor de q . Como r es irreducible p y r no tienen factores comunes. Por el teorema de Bezout existen polinomios α y β tales que $\alpha r + \beta p = 1$. Multiplicando por q obtenemos que $\alpha rq + \beta pq = q$. Como la parte izquierda de esta igualdad se divide entre r entonces r

es un factor de q . ■

5.9

Sea $p = \alpha p_1 \cdots p_n$ una descomposición en factores irreducibles de p . Si q es cualquier factor irreducible de p entonces, q es igual a alguno de los p_i .

Prueba. Por inducción en n . Si $n = 1$ entonces q divide a p_1 y como p_1 y q son irreducibles obtenemos que $p_1 = q$. Sea $n > 1$. Por 5.8 o q es un factor de p_n o q es un factor de $p_1 \cdots p_{n-1}$. En el primer caso $q = p_n$ y en el segundo el resultado se sigue por hipótesis de inducción. ■

Teorema de Factorización de Polinomios

5.10

Cualquier polinomio p se descompone como $\alpha p_1 p_2 \cdots p_n$ donde α es su coeficiente principal y $p_1 \cdots p_n$ son polinomios irreducibles. Esta descomposición es única salvo el orden de los factores.

Prueba. Solamente debemos probar la unicidad. Además, sacando como factor el coeficiente principal podemos suponer que p es mónico. Si p es de grado 1 entonces no hay nada que probar ya que entonces p es irreducible y la única descomposición de p es el mismo.

Supongamos que el teorema es cierto para cualquier polinomio de grado estrictamente menor que k . Sea p mónico de grado k que tiene dos descomposiciones en irreducibles $p_1 \cdots p_n = p'_1 \cdots p'_m$. Como p_n es irreducible entonces de 5.9 obtenemos que tiene que existir un j tal que p_n es un factor de p'_j . Como p'_j es irreducible $p'_j = p_n$. Luego, para $p/p_n = p/p'_j$ tenemos dos descomposiciones que por hipótesis de inducción son iguales salvo orden. ■

El conjunto ordenado de polinomios monómicos

5.11

Sea $p = p_1^{j_1} p_2^{j_2} \cdots p_m^{j_m}$ la descomposición factores irreducibles del polinomio mónico p . Entonces, todos los divisores monómicos de p son de la forma $p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ con $0 \leq k_i \leq j_i$.

Prueba. Sea q un divisor mónico de $p = p_1^{j_1} p_2^{j_2} \cdots p_m^{j_m}$. Por el resultado 5.9 y la transitividad de la relación de divisibilidad cualquier factor irreducible de q es factor irreducible de p y por lo tanto $q = p_1^{k_1} p_2^{k_2} \cdots p_m^{k_m}$ para ciertos $k_i \geq 0$. Supongamos que $k_1 > j_1$. Entonces como p_1 no divide a $p_2^{j_2} \cdots p_m^{j_m}$ obtenemos que $p_1^{k_1}$ no divide a p . Esto contradice que q divide a p . Este mismo argumento funciona si suponemos que para cierto i se tiene que $k_i > j_i$. ■

Sean p y q dos polinomios monómicos. Sea $\{p_1, \dots, p_m\}$ el conjunto de los polinomios

irreducibles que son factores de p ó de q . Entonces por el Teorema de Factorización de Polinomios encontramos descomposiciones únicas $p = p_1^{j_1} \dots p_m^{j_m}$ y $q = p_1^{k_1} \dots p_m^{k_m}$ donde los exponentes son naturales algunos posiblemente iguales a cero.

Si $p = p_1^{j_1} \dots p_m^{j_m}$ y $q = p_1^{k_1} \dots p_m^{k_m}$ son dos polinomios mónicos descompuestos de la manera anterior, entonces de 5.11 concluimos que $p_1^{\max(j_1, k_1)} \dots p_m^{\max(j_m, k_m)}$ es el mínimo común múltiplo de p y q ; y que $p_1^{\min(j_1, k_1)} \dots p_m^{\min(j_m, k_m)}$ es el máximo común divisor de p y q . Esto prueba que *cualquier dos polinomios tienen máximo común divisor y mínimo común múltiplo*.

Este mismo argumento se puede usar para convencernos de que cualquier conjunto *finito* de polinomios tiene mínimo común múltiplo y máximo común divisor. Este asunto es un poco más complejo cuando el conjunto de polinomios es *infinito*. Así por ejemplo, el conjunto de polinomios $\{x, x^2, x^3, \dots\}$ no tiene mínimo común múltiplo. Sin embargo, *cualquier* conjunto de polinomios (finito o infinito) si tiene máximo común divisor. La prueba de este hecho es en esencia la misma que la anterior, solo hay que observar que *cualquier* conjunto de naturales (finito o infinito) tiene mínimo.

Ejercicio 92 Pruebe que cualquier conjunto de polinomios mónicos tiene máximo común divisor.

Ejercicio 93 Sea $p = p_1^{j_1} p_2^{j_2} \dots p_m^{j_m}$ la descomposición factores irreducibles del polinomio mónico p . ¿Cuántos divisores mónicos tiene p ?

Desarrollo de Taylor



Brook Taylor (Inglaterra 1685-1731). Entre las contribuciones de este matemático, se destacan: La invención de la rama de las matemáticas que hoy en día se conoce como Cálculo en Diferencias, la invención de la integración por partes, y la fórmula llamada por su nombre. En 1772 Lagrange proclamó esta fórmula como “el principio básico del cálculo diferencial”. A esta fórmula, enunciada en la siguiente proposición, se le llama **Desarrollo de Taylor alrededor** del punto x_0 . Como el lector sabe de los cursos de cálculo, este desarrollo es mucho más general y se cumple en cierta clase de funciones. Sin embargo para polinomios, su demostración es independiente del campo y puramente algebraica (no requiere del concepto de continuidad).

Desarrollo de Taylor

Para cualquier polinomio $p(x)$ y cualquier elemento del campo x_0 existen unos únicos coeficientes $\alpha_0, \alpha_1, \dots, \alpha_n$ tales que $p(x) = \alpha_0 + \alpha_1(x - x_0) + \dots + \alpha_n(x - x_0)^n$.

Prueba. Sea $p(x) = \sum a_k x^k$ un polinomio de grado n . Si $x_0 = 0$ entonces, $\alpha_k = a_k$. Supongamos que $x_0 \neq 0$. Por el binomio de Newton tenemos

$$\sum_{j=0}^n \alpha_j (x - x_0)^j = \sum_{j=0}^n \alpha_j \sum_{k=0}^j \binom{j}{k} x^k (-x_0)^{j-k} = \sum_{k=0}^n \left(\sum_{j=k}^n \binom{j}{k} (-x_0)^{j-k} \alpha_j \right) x^k$$

y para encontrar los coeficientes α_j tenemos el sistema de ecuaciones lineales $a_k = \sum_{j=k}^n b_{kj} \alpha_j$ donde $b_{kj} = \binom{j}{k} (-x_0)^{j-k}$. Observese que $b_{kk} = 1$ por lo que la matriz de nuestro sistema de ecuaciones es triangular con unos en la diagonal y por lo tanto su determinante es distinto de cero. Luego, el sistema tiene solución única. ■

En otras palabras el desarrollo de Taylor para polinomios lo que afirma es que $\{1, (x - x_0), (x - x_0)^2, \dots\}$ es una base del espacio vectorial de polinomios. Esto no es algo muy notable. Lo que sí es notable es la forma que toman los coeficientes α_i en términos de las derivadas del polinomio (véanse los ejercicios que siguen).

Ejercicio 94 Pruebe que si $P = \{p_0, p_1, p_2, \dots\}$ es un conjunto de polinomios tales que $\forall i \in \mathbb{N}$ el grado de p_i es i entonces, P es una base del espacio vectorial de polinomios.

Ejercicio 95 Demuestre que la expresión $\sum_{j=k}^i (-1)^{j-k} \binom{j}{k} \binom{i}{j}$ es igual a la función delta de Kronecker δ_{ik} o sea, es cero si $i \neq k$ y es uno si $i = k$. [193]

Ejercicio 96 Demuestre que los coeficientes α_j del Desarrollo de Taylor (5.12) del polinomio $\sum a_k x^k$ son iguales a $\beta_j = \sum_{i=j}^n \binom{i}{j} a_i x_0^{i-j}$. [194]

Ejercicio 97 Pruebe que los coeficientes α_j del Desarrollo de Taylor (5.12) del polinomio $p(x)$ son iguales a $p^{(j)}(x_0)/j!$ donde $p^{(j)}(x)$ denota el polinomio derivado j veces. [194]

5.2 Polinomios complejos. Teorema de Gauss



Johann Carl Friedrich Gauss (Alemania 1777-1855) fué el más grande matemático de su época. Este teorema que lleva su nombre fué demostrado por primera vez en su tesis doctoral (1799). Este teorema es conocido como el “teorema fundamental del álgebra”. En este libro hemos mencionado otros resultados de Gauss y cualquiera que se dedique a estudiar matemáticas, estadísticas, física o astronomía oirá de este científico en más de una ocasión.

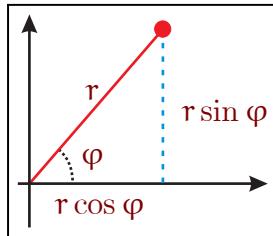
En esta sección demostraremos que el campo de los números complejos es algebraicamente cerrado. Como esta demostración no es algebraica sino analítica necesitaremos introducir algunos conceptos básicos de análisis complejo. Para esto, presupondremos que el lector conoce los correspondientes conceptos de análisis real.

Si bien, el contenido de esta sección no es básico para la comprensión del álgebra

lineal, por otro lado, si es fundamental que el lector conosca a la perfección el enunciado del teorema de Gauss: *todo polinomio complejo de grado al menos 1 tiene una raíz.*

Forma polar. Igualdad de Moivre

Todo número complejo z se puede representar en la **forma polar** $z = r(\cos \varphi + i \sin \varphi)$ donde r es la longitud del vector \overrightarrow{Oz} en el plano complejo y φ es el ángulo que forma dicho vector con el eje real de este plano (vease la figura). Al número r se le llama **módulo** del número complejo y es común que se denote por $\|z\|$. Al ángulo φ se le llama **argumento** del número complejo. La forma polar hace mucho más fácil calcular el producto y las potencias de números complejos.



5.13

Para hallar el producto de dos números complejos hay que multiplicar sus módulos y sumar sus argumentos.

Prueba. Sean $r(\cos \varphi + i \sin \varphi)$ y $\rho(\cos \psi + i \sin \psi)$ dos complejos. Tenemos

$$r(\cos \varphi + i \sin \varphi) \rho(\cos \psi + i \sin \psi) =$$

$$\rho((\cos \psi \cos \varphi - \sin \psi \sin \varphi) + (\cos \psi \sin \varphi + \cos \varphi \sin \psi)i) =$$

$$\rho(\cos(\varphi + \psi) + i \sin(\varphi + \psi))$$

que es lo que se necesitaba probar. ■

$$z^n = r^n (\cos n\varphi + i \sin n\varphi)$$

Aplicando este resultado al caso de la potencia de números complejos obtenemos la igualdad mostrada en el recuadro a la izquierda. Esta igualdad se conoce como la **Igualdad de Moivre**. Una de sus consecuencias más importantes es el siguiente resultado.

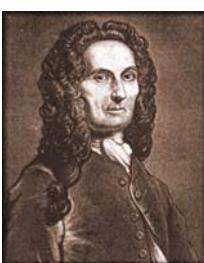
5.14

Los polinomios $z^n - a$ tienen exactamente n raíces complejas.

Prueba. Sea $a = r(\cos \varphi + i \sin \varphi)$. Para $k \in \{0, 1, \dots, n-1\}$ denotemos x_k el número complejo con módulo $\sqrt[n]{r}$ y argumento $(\varphi + 2k\pi)/n$. Por la Igualdad de Moivre tenemos

$$x_k^n = (\sqrt[n]{r})^n \left(\cos n \frac{\varphi + 2k\pi}{n} + i \sin n \frac{\varphi + 2k\pi}{n} \right) = r(\cos \varphi + i \sin \varphi) = a$$

que es lo que se necesitaba probar. ■



Abraham de Moivre (Francia 1667-1754). Uno de los fundadores de la Geometría Analítica y de la Teoría de las Probabilidades. A pesar de su exelencia científica, nunca tuvo la oportunidad de tener un puesto en alguna universidad. Sus ingresos provenían de dar clases privadas de Matematicas y murió en la pobreza. Moivre también es famoso por haber predicho el día de su muerte. Descubrió que dormía 15 minutos más cada noche y sumando la progresión aritmética calculó que moriría en el día que dormiría 24 horas. ¡Tuvo razón!

Ejercicio 98 Pruebe que el conjunto de raíces complejas del polinomio $z^n - 1$ es un grupo para el producto. ¿Qué grupo es este? [194]

Continuidad

Una función $f : \mathbb{C} \rightarrow \mathbb{C}$ es **continua en el punto z_0** si para todo real positivo ε existe otro real positivo δ tal que se cumple que $(\|z - z_0\| < \delta) \Rightarrow (\|f(z) - f(z_0)\| < \varepsilon)$. Una función continua en todo punto del plano complejo se le llama **continua**.

5.15

La función módulo $z \mapsto \|z\|$ es continua.

Prueba. Veamos la desigualdad $\|z - z_0\| \geq \||z| - \|z_0|\|$. Esta desigualdad es equivalente a la siguiente: “En un triángulo el valor absoluto de la diferencia de dos de sus lados siempre es menor o igual que el tercero”. La prueba de esta la dejaremos en calidad de ejercicio. Por esta desigualdad, tenemos que $\forall \varepsilon > 0 \exists \delta = \varepsilon$ tal que si $\|z - z_0\| < \delta$ entonces, $\||z| - \|z_0|\| \leq \|z - z_0\| < \varepsilon$ y esto prueba nuestra tesis. ■

Ejercicio 99 Pruebe que en un triángulo el valor absoluto de la diferencia las longitudes de dos de sus lados siempre es menor o igual que la longitud del tercero. [194]

5.16

La suma y el producto de funciones continuas en un punto z_0 son continuas en el punto z_0 .

Prueba. Sean f y g funciones continuas en z_0 . Con el objetivo de ahorrar espacio denotemos $f_z = f(z)$, $f_0 = f(z_0)$, $g_z = g(z)$ y $g_0 = g(z_0)$. Para todo $\varepsilon > 0$ existen δ_1 y δ_2 tales que

$$\begin{aligned} (\|z - z_0\| < \delta_1) &\Rightarrow (\|f_z - f_0\| < \varepsilon) \\ (\|z - z_0\| < \delta_2) &\Rightarrow (\|g_z - g_0\| < \varepsilon) \end{aligned}$$

y en particular para el mínimo (que llamaremos δ) de δ_1 y δ_2 se cumplen las dos desigualdades a la derecha. Sumando y aplicando la desigualdad del triángulo obtenemos:

$$\theta = 2\epsilon > \|f_z - f_0\| + \|g_z - g_0\| \geq \|f_z - f_0 + g_z - g_0\| = \|(f+g)(z) - (f+g)(z_0)\|$$

Luego, para todo $\theta > 0$ existe δ tal que $(|z - z_0| < \delta) \Rightarrow |(f+g)(z) - (f+g)(z_0)| < \theta$ con lo que se prueba que la suma de continuas es continua.

Por otro lado, por la desigualdad del triángulo y 5.13 tenemos

$$\begin{aligned} \|(fg)(z) - (fg)(z_0)\| &= \|f_z g_z - f_0 g_0\| = \\ \|(f_z - f_0)(g_z - g_0) + (f_z - f_0)g_0 + (g_z - g_0)f_0\| &\leq \\ \|f_z - f_0\| \|g_z - g_0\| + \|f_z - f_0\| \|g_0\| + \|g_z - g_0\| \|f_0\| &< \\ < \epsilon^2 + \epsilon |g_0| + \epsilon |f_0| &< (1 + |g_0| + |f_0|) \epsilon = c\epsilon = \theta \end{aligned}$$

donde la última desigualdad se da para $\epsilon < 1$. Como c es una constante que no depende de z obtenemos que para todo $\theta > 0$ existe δ tal que $(|z - z_0| < \delta) \Rightarrow |(fg)(z) - (fg)(z_0)| < \theta$ lo que prueba que el producto de continuas es continua. ■

5.17

Para cualquier polinomio complejo p la función de evaluación $\mathbb{C} \ni z \mapsto p(z) \in \mathbb{C}$ es continua.

Prueba. La función de evaluación de un polinomio se obtiene usando sumas y productos de funciones constantes y la función identidad $f(z) = z$. Estas funciones son continuas y de 5.16 obtenemos la prueba. ■

5.18

Sea g una función continua en el punto z_0 y f una función continua en el punto $g(z_0)$ entonces, la composición $f \circ g$ es continua en el punto z_0 .

5.19

El módulo de un polinomio es una función continua.

Prueba. Por 5.18 y porque los polinomios y el módulo son funciones continuas. ■

Límite de sucesiones complejas

Una sucesión de números complejos $\{z_j\} = \{a_j + b_j i\}$ tiene **límite** $z = a + bi$ si las sucesiones reales $\{a_j\}$ y $\{b_j\}$ tienen límites a a y b respectivamente. Esto es equivalente a que los módulos y los argumentos de la sucesión converjan al módulo y al argumento

del límite. También, esto es equivalente a que $\forall \varepsilon > 0 \exists N$ tal que $\forall k > N \|\mathbf{z}_k - \mathbf{z}\| < \varepsilon$. Una sucesión es **convergente** si esta tiene límite. Por una propiedad análoga para las sucesiones reales se tiene que toda subsucesión de una sucesión convergente es convergente y converge al mismo límite. Una sucesión $\{\mathbf{z}_j\} = \{\mathbf{a}_j + \mathbf{b}_j i\}$ es **acotada** si ambas sucesiones $\{\mathbf{a}_j\}$ y $\{\mathbf{b}_j\}$ son acotadas. Esto es equivalente a que la sucesión real $\{\|\mathbf{z}_j\|\}$ sea acotada. Es claro que una sucesión no acotada no puede tener límite. También, que toda sucesión acotada tiene una subsucesión convergente pues esta misma propiedad se cumple para sucesiones reales. Expongamos otras dos propiedades que son un poco más difíciles de probar.

5.20

Sea f una función continua en z_0 y $\{\mathbf{z}_k\}$ una sucesión de números complejos que converge a z_0 . Entonces, $\lim f(\mathbf{z}_k) = f(\lim \mathbf{z}_k)$.

Prueba. Como f es continua en z_0 y $\lim \mathbf{z}_k = z_0$ entonces tenemos que

$$\begin{aligned} \forall \varepsilon > 0 \quad &\exists \delta \quad (\|\mathbf{z} - z_0\| < \delta) \Rightarrow (\|f(\mathbf{z}) - f(z_0)\| < \varepsilon) \\ \forall \delta > 0 \quad &\exists N \quad (k > N) \Rightarrow (\|\mathbf{z}_k - z_0\| < \delta) \end{aligned} .$$

Por la transitividad de la implicación obtenemos que

$$\forall \varepsilon > 0 \quad \exists N \quad (k > N) \Rightarrow \|f(\mathbf{z}_k) - f(z_0)\| < \varepsilon$$

y esto quiere decir que $\lim f(\mathbf{z}_k) = f(z_0)$. ■

5.21

Si $\{\mathbf{z}_k\}$ es no acotada y p es un polinomio de grado al menos 1 entonces, la sucesión $\{\|p(\mathbf{z}_k)\|\}$ es no acotada.

Prueba. Sea $p(z)$ un polinomio de grado $n > 1$. Por la desigualdad triangular, tenemos

$$\|p(z)\| \geq \sum_{i=0}^n \|a_i z^i\| = \sum_{i=0}^n \|a_i\| \|z\|^i = \|a_n\| \|z\|^n + \sum_{i=0}^{n-1} \|a_i\| \|z\|^i \geq \|a_n\| \|z\|^n$$

Como $\{\mathbf{z}_k\}$ no es acotada, tampoco lo son $\{\|a_n\| \|\mathbf{z}_k\|^n\}$ y $\{\|p(\mathbf{z}_k)\|\}$. ■

Teorema de Gauss

Ya estamos listos para demostrar el Teorema de Gauss pero antes, veamos un resultado preliminar que hace la mayoría del trabajo.

5.22

Sea p un polinomio complejo de grado al menos 1. Si z_0 no es una raíz de p entonces, existe $z \in \mathbb{C}$ tal que $\|p(z)\| < \|p(z_0)\|$.

Prueba. Hagamos el desarrollo de Taylor de $p(z)$ alrededor del punto z_0 . Tenemos

$$p(z) = \sum_{j=0}^n \alpha_j (z - z_0)^j = p(z_0) + \sum_{j=1}^n \alpha_j (z - z_0)^j$$

ya que $\alpha_0 = p(z_0)$. Sea α_k el primero de los $\{\alpha_j \mid j > 0\}$ diferente de cero y escojamos $z = z_0 + t\theta$ donde θ es una (véase 5.14) de las raíces de la ecuación $\alpha_k x^k + p(z_0) = 0$ y t es un real tal que $0 < t < 1$ que definiremos después. Por la definición de θ , z y k tenemos

$$p(z) = p(z_0) + \alpha_k \theta^k t^k + \sum_{j=k+1}^n \alpha_j \theta^j t^j = p(z_0) (1 - t^k) + \sum_{j=k+1}^n \alpha_j \theta^j t^j$$

y por lo tanto, de la desigualdad del triángulo obtenemos:

$$\|p(z)\| \leq (1 - t^k) \|p(z_0)\| + \sum_{j=k+1}^n \|\alpha_j \theta^j\| t^j = \|p(z_0)\| + t^k q(t)$$

donde $q(t)$ denota el polinomio (con coeficientes reales) del recuadro a la derecha. Observemos que se cumple la desigualdad $q(0) = -\|p(z_0)\| < 0$.

$$-\|p(z_0)\| + \sum_{j=k+1}^n \|\alpha_j \theta^j\| t^{j-k}$$

Por continuidad (de los polinomios reales) existe un $t_0 > 0$ suficientemente pequeño tal que $q(t_0) < 0$. Luego, $\|p(z_0 + t_0 \theta)\| \leq \|p(z_0)\| + t_0^k q(t_0) < \|p(z_0)\|$. ■

Ejercicio 100 ¿Dónde se usa en la demostración anterior que $t < 1$? [194]

Teorema de Gauss

5.23

Todo polinomio de grado mayor que cero tiene al menos una raíz compleja.

Prueba. Sea p un polinomio. Denotemos $A = \{\|p(z)\| : z \in \mathbb{C}\}$. El conjunto A es un conjunto de reales acotado inferiormente pues $\|p(z)\| \geq 0$. Luego (por un teorema clásico de análisis matemático), A tiene un ínfimo que denotaremos por μ .

Demostremos que μ es el mínimo o sea, que $\mu \in A$. Como μ es ínfimo hay una sucesión $\{a_j\}$ de elementos de A que converge a μ y por lo tanto hay una sucesión de complejos $\{z_j\}$ tales que $\lim \|p(z_j)\| = \mu$. Si la sucesión $\{z_j\}$ no estuviera acotada entonces, por 5.21 la sucesión $\{\|p(z_j)\|\}$ tampoco lo sería lo que contradice que esta sucesión converge a μ . Luego, $\{z_j\}$ es acotada y podemos escoger una subsucesión convergente que podemos suponer la misma. Denotemos $y = \lim z_j$. Como el módulo de un polinomio es una función continua entonces, por 5.20 tenemos $\mu = \lim \|p(z_j)\| = \|p(\lim z_j)\| = \|p(y)\|$. Luego, $\mu \in A$.

Si $\mu \neq 0$ entonces, por 5.22 existiría un y' tal que $\|p(y')\| < \|p(y)\| = \mu$ lo que contradice que μ es el mínimo de A . Luego, $\|p(y)\| = 0$ y por lo tanto p tiene una

raíz. ■

5.3 Factorización de polinomios complejos y reales

En esta sección utilizaremos el teorema de Gauss para averiguar cuales son todos los polinomios irreducibles con coeficientes complejos y los polinomios irreducibles con coeficientes reales. Esto nos dará la posibilidad de encontrar la descomposición en factores (única salvo orden de los factores) de los polinomios complejos y reales.

Caso Complejo

Clasificación de los polinomios complejos irreducibles

5.24

Los polinomios complejos irreducibles son exactamente los mónicos de grado uno.

Prueba. Al absurdo supongamos que un polinomio irreducible tiene grado mayor que uno. Por el Teorema de Gauss (5.23) este polinomio tiene una raíz α . Por 5.4 el polinomio se divide entre $(x - \alpha)$ y esto contradice que el polinomio es irreducible. ■

Este resultado nos da la posibilidad de factorizar completamente los polinomios complejos. Por el Teorema de Factorización de Polinomios (5.10) cada polinomio complejo $p(x)$ se tiene que descomponer como en el recuadro a la izquierda. En esta fórmula, a_n es el coeficiente principal del polinomio. Los complejos α_j son las diferentes raíces de $p(x)$. El natural n_j es la multiplicidad de la raíz α_j y $n = \sum n_i$ es el grado de $p(x)$. Nuevamente, por el Teorema de Factorización de Polinomios (5.10) esta descomposición es única salvo orden de los factores.

Caso real

Recordemos que si $a+bi$ es un número complejo entonces su **complejo conjugado** es $a-bi$. Denotaremos por \bar{z} el complejo conjugado del número complejo z . Es fácil comprobar que la operación de conjugación cumple las propiedades del recuadro a la derecha.

1. $z \in \mathbb{R} \Rightarrow \bar{z} = z$
2. $\bar{z+u} = \bar{z} + \bar{u}$
3. $\bar{zu} = \bar{z}\bar{u}$

Las propiedad 1 es trivial de la definición. Las propiedades 2 y 3 significan que la conjugación compleja es un automorfismo del campo de los números complejos.

5.25

Sea p un polinomio con coeficientes reales y α una raíz compleja de p . Entonces $\bar{\alpha}$ es también una raíz de p .

Prueba. Como α es raíz de $p = \sum a_i x^i$ y por las propiedades de la conjugación tenemos

$$0 = \bar{0} = \overline{\sum a_i \alpha^i} = \sum \overline{a_i} \bar{\alpha}^i = \sum a_i \bar{\alpha}^i$$

que es lo que se quería probar. ■

Clasificación de los polinomios reales irreducibles

5.26

Si p es un polinomio real irreducible entonces p es de la forma $x - \alpha$ o es de la forma $(x - a)^2 + b^2$ con $b \neq 0$.

Prueba. Si p es de grado 1 entonces necesariamente es igual a $x - \alpha$ para cierto número real α . Si p es de grado 2 entonces por el teorema de Gauss este tiene un factor $x - \alpha$ para cierto α complejo. Si α fuera real esto contradeciría que p es irreducible. Luego, $\alpha = a + bi$ con $b \neq 0$. Por la proposición anterior $a - bi$ también es una raíz de p por lo que

$$p(x) = (x - (a + bi))(x - (a - bi)) = (x - a)^2 + b^2$$

Si p es de grado al menos 3 entonces, por el teorema de Gauss este tiene una raíz compleja α . Si α fuera real entonces $x - \alpha$ sería un factor de p . Si $\alpha = a + bi$ con $b \neq 0$ entonces $(x - a)^2 + b^2$ sería un factor de p . En ambos casos se contradice la suposición de que p es irreducible. ■

Ahora, ya podemos descomponer completamente en factores los polinomios con coeficientes reales. Cada polinomio real $p(x)$ se tiene que expresar como:

$$p(x) = a_n \prod_{j=1}^k (x - \alpha_j)^{n_j} \prod_{\ell=1}^l \left((x - a_\ell)^2 + b_\ell^2 \right)^{m_\ell}$$

En esta fórmula, a_n es el coeficiente principal del polinomio. Los números reales α_j son las diferentes raíces reales de $p(x)$. El número natural n_j es la multiplicidad de la raíz α_j . Los números complejos $(a_\ell + b_\ell i)$ y $(a_\ell - b_\ell i)$ son las diferentes raíces complejas de $p(x)$. El número natural m_ℓ es la multiplicidad de la raíz compleja $(a_\ell + b_\ell i)$. Obviamente, $n = \sum n_i + 2 \sum m_\ell$ es el grado de $p(x)$. Nuevamente, por el Teorema de Factorización de Polinomios (5.10) esta descomposición es única salvo orden de los factores.

La diferencia fundamental entre los números reales y los números complejos se expresa de manera muy evidente en los resultados de esta sección. Esta diferencia hará que posteriormente nos sea mucho más fácil clasificar los operadores lineales en un espacio vectorial complejo que en un espacio vectorial real. En general, todo es más fácil para los campos que cumplen el Teorema de Gauss o sea, los campos algebraicamente

cerrados.

5.4 Campos de fracciones. Funciones racionales



Sea $(A, +, \cdot)$ un anillo conmutativo y denotemos por 0 el neutro aditivo y por 1 el neutro multiplicativo respectivamente. Queremos construir un campo que contenga a A . Esta es una situación análoga a cuando construimos el campo de los racionales \mathbb{Q} para que contenga al anillo conmutativo \mathbb{Z} . ¿Funcionará esta misma construcción en el caso general? Investiguemos para lograr una respuesta.

Campos de fracciones

Lo primero es definir las fracciones. Consideraremos conjunto de todas las **fracciones** a/b donde $a \in A$ y $b \in A \setminus \{0\}$. Dos fracciones las consideraremos iguales si se cumple la igualdad en el recuadro a la derecha.

$$\left(\frac{a}{b} = \frac{c}{d} \right) \Leftrightarrow (ad = bc)$$

Ejercicio 102 Pruebe que la igualdad de fracciones es una relación de equivalencia.
[194]

$$\frac{a c}{b d} = \frac{a c}{b d}$$

Ahora, necesitamos definir las operaciones entre fracciones. Primero el producto porque es el más fácil. Definamos el producto por la fórmula en el recuadro a la izquierda. Nos tenemos que convencer primero de que esta definición es correcta. O sea que si las fracciones son iguales sus productos son iguales. Más precisamente, necesitamos ver que se cumple lo siguiente:

$$\left(\frac{a}{b} = \frac{a'}{b'} \text{ y } \frac{c}{d} = \frac{c'}{d'} \right) \Rightarrow \left(\frac{ac}{bd} = \frac{a'c'}{b'd'} \right)$$

y efectivamente de las hipótesis de esta implicación tenemos $ab' = a'b$, $cd' = c'd$ y multiplicando estas dos igualdades obtenemos $acb'd' = a'c'bd$, lo que es equivalente a la tesis de la implicación.

Este producto es conmutativo y asociativo ya que ambas propiedades se cumplen en los “numeradores” y en los “denominadores”. La fracción $1/1$ es neutro para el producto y cualquier fracción a/b donde $a \neq 0$ tiene inverso multiplicativo b/a ya que $ab/ba = 1/1$ por la conmutatividad del producto en A . Todo esto significa que el conjunto de fracciones con numerador distinto de cero forman un grupo conmutativo.

Definamos ahora la suma de fracciones por la fórmula en el recuadro a la derecha. Para convencernos que la definición es correcta tenemos que probar que las sumas de fracciones iguales son iguales. O sea que:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd}$$

$$\left(\frac{a}{b} = \frac{a'}{b'} \text{ y } \frac{c}{d} = \frac{c'}{d'} \right) \Rightarrow \left(\frac{ad + cb}{bd} = \frac{a'd' + c'b'}{b'd'} \right)$$

y efectivamente haciendo algunos cálculos inteligentes obtenemos

$$\left(\begin{array}{l} ab' = a'b \\ cd' = c'd \end{array} \right) \Rightarrow \left(\begin{array}{l} 0 = (ab' - a'b) dd' = \\ = (c'd - cd') bb' = 0 \end{array} \right) \Rightarrow \left(\begin{array}{l} (ad + cb) b'd' = \\ = (a'd' + c'b') bd \end{array} \right)$$

que era lo que se necesitaba probar.

La comprobación de que esta suma es commutativa es obvia. La asociatividad se comprueba calculando que $\frac{a}{b} + \frac{c}{d} + \frac{u}{v} = \frac{adv + cbv + ubd}{bdv}$

independientemente del orden en que se haga la suma. La fracción $0/1$ es neutro para la suma y cualquier fracción a/b tiene opuesto $-a/b$ (para esto último es necesario observar que $0/1 = 0/v$ para todo $v \neq 0$). Luego, el conjunto de todas las fracciones es un grupo abeliano respecto a la suma.

Solo nos falta la distributividad para comprobar que las fracciones con las operaciones definidas forman un campo y esto se comprueba con los siguientes cálculos:

$$\frac{u}{v} \left(\frac{a}{b} + \frac{c}{d} \right) = \frac{uad + ucb}{vbd} = \frac{uad}{vbd} + \frac{ucb}{vbd} = \frac{ua}{vb} + \frac{uc}{vd} = \frac{u}{v} \frac{a}{b} + \frac{u}{v} \frac{c}{d}.$$

Por último observemos que si identificamos al elemento $a \in A$ con la fracción $a/1$ podemos pensar que el anillo A es un subconjunto de las fracciones ya que la suma y el producto de estas fracciones coinciden con la suma y el producto dentro de A .

Hemos hecho todas estas pruebas detalladamente para no equivocarnos al afirmar que esto que hemos demostrado sirve para cualquier anillo commutativo. El lector debería analizar cuidadosamente cada igualdad en los razonamientos anteriores para convencerse de que todas ellas se desprenden de los axiomas de anillo commutativo y de las definiciones de las operaciones entre fracciones. Al conjunto de todas las fracciones con las operaciones así definidas se le llama el **campo de fracciones** del anillo commutativo A .



MENTIRA, no hay tal campo de fracciones para cualquier anillo commutativo. ¿Puede usted encontrar el error? Si cree que puede regrese arriba y búskuelo, si no, siga leyendo.

El problema es el siguiente. Al definir el producto $(a/b)(c/d) = (ac/bd)$ con b y d distintos de cero supusimos que necesariamente bd es DISTINTO DE CERO. Esto no es cierto en cualquier anillo commutativo, por ejemplo en \mathbb{Z}_6 tenemos $2 \times 3 = 0$. Si en el anillo hay tales elementos no podemos definir adecuadamente el producto de fracciones (tampoco la suma). Ya vimos, que si un anillo commutativo es tal que para cualesquiera b y d distintos de cero se tiene que bd es distinto de cero entonces, se dice que este anillo es un dominio de integridad. Ahora si, *todo dominio de integridad tiene su campo de fracciones*. El ejemplo evidente de dominio de integridad es \mathbb{Z} . Su campo de fracciones es \mathbb{Q} .

Funciones racionales

El ejemplo por el cual hemos escrito esta sección es el siguiente:

Todo anillo de polinomios con coeficientes en un campo es un dominio de integridad.

Prueba. Tenemos que probar que el producto de dos polinomios diferentes de cero es diferente de cero (recuérdese que un polinomio es cero cuando todos sus coeficientes son cero).

Sean $p(x) = \sum_{i=0}^n a_i x^i$ y $q(x) = \sum_{i=0}^m b_i x^i$ dos polinomios cualesquiera de grados n y m respectivamente. Denotemos $p(x) q(x) = \sum_{i=0}^{n+m} c_i x^i$. Por la fórmula del producto de polinomios, tenemos $c_{n+m} = a_n b_m$. Como $a_n \neq 0$, $b_m \neq 0$ y todo campo es dominio de integridad obtenemos $c_{n+m} \neq 0$ y por lo tanto $p(x) q(x) \neq 0$. ■



Observese que en la demostración no se usó el hecho de que en el campo \mathbb{K} hay inversos multiplicativos. Eso quiere decir que de hecho, hemos demostrado algo mucho más fuerte:

Los anillos de polinomios con coeficientes en un dominio de integridad son dominios de integridad.

Como el anillo de polinomios $\mathbb{K}[x]$ es un dominio de integridad este tiene su campo de fracciones que se denota por $\mathbb{K}(x)$. Nótese la diferencia entre $\mathbb{K}[x]$ y $\mathbb{K}(x)$. A los elementos de $\mathbb{K}(x)$ se les llama **funciones racionales** (en la variable x). Las funciones racionales son fracciones de polinomios $p(x)/q(x)$ que se suman y multiplican mediante las reglas a las que todos estamos acostumbrados.

Ejercicio 103 ¿Conoce usted un campo infinito de característica 2? [195]

Ejercicio 104 ¿Qué pasa si construimos el campo de fracciones de un campo? [195]



Capítulo sexto

Descomposición de Operadores Lineales

 Hay dos teoremas que son muy conocidos por el lector y que son muy parecidos. El primero es que todo número natural se descompone de forma única salvo orden de los factores en producto de números primos. El segundo es que todo polinomio se descompone de forma única salvo orden de los factores en producto de polinomios irreducibles. Para los operadores lineales hay un teorema parecido *todo operador lineal se descompone en suma directa de OLs irreducibles, el “tipo” de tal descomposición es único*. En este capítulo daremos la demostración de este teorema y lo que es más importante, encontraremos cuales son todos los OLs irreducibles de un espacio vectorial de dimensión finita.

6.1 Suma directa de operadores lineales

Recordemos que el acrónimo OL significa para nosotros un operador lineal, o sea una transformación lineal de un espacio en si mismo. Con otras palabras, un endomorfismo del espacio. Usaremos sistemáticamente este acrónimo. Sin embargo, al conjunto de todos los OLs de un espacio vectorial \mathfrak{E} lo denotaremos por $\text{End}(\mathfrak{E})$. Ya vimos que el conjunto $\text{End}(\mathfrak{E})$ es un álgebra para la suma de funciones, la multiplicación por escalares y la composición de OLs. El conjunto $\text{End}(\mathfrak{E})$ contiene el neutro para la suma que es el operador lineal \mathbb{O} que a cualquier vector le hace corresponder el vector $\mathbf{0}$. También contiene al neutro para la composición de funciones \mathbb{I} que es la identidad $\mathbb{I}(\mathbf{a}) = \mathbf{a}$.

En este capítulo todos los espacios vectoriales serán de dimensión finita. Si $f : \mathfrak{E} \rightarrow \mathfrak{E}$ y $g : \mathfrak{F} \rightarrow \mathfrak{F}$ son dos OLs entonces, a la función

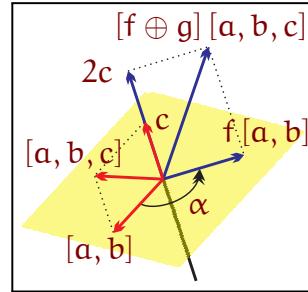
$$f \oplus g : \mathfrak{E} \oplus \mathfrak{F} \ni (\mathbf{a}, \mathbf{b}) \mapsto (f(\mathbf{a}), g(\mathbf{b})) \in \mathfrak{E} \oplus \mathfrak{F}$$

se le llama **suma directa** de los OLs f y g . Es fácil comprobar que la suma directa de dos OLs es un OL.

Ejercicio 105 Pruebe que la suma directa de OLs es siempre un OL. [195]

El lector no debe confundir la suma directa de OLs con la habitual suma de funciones. Si $f : \mathbb{E} \rightarrow \mathbb{E}$ y $g : \mathbb{E} \rightarrow \mathbb{E}$ son dos OLs entonces su suma se define como $(f + g)(\mathbf{a}) = f(\mathbf{a}) + g(\mathbf{a})$.

Veamos ahora un ejemplo. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la rotación en el ángulo α en contra de las manecillas del reloj o en otras palabras $(1, 0) \mapsto (\cos \alpha, \sin \alpha)$ y $(0, 1) \mapsto (-\sin \alpha, \cos \alpha)$. Sea $g : \mathbb{R} \rightarrow \mathbb{R}$ la dilatación de factor 2 o sea $z \mapsto 2z$. El espacio $\mathbb{R}^2 \oplus \mathbb{R}$ lo podemos pensar como \mathbb{R}^3 en el cual el primer sumando es el plano x, y y el segundo es el eje z . Sabemos como transforma f el plano x, y (rotando) y como transforma g el eje z (dilatando). Vease la figura a la derecha.



Ahora si (a, b, c) es un vector arbitrario de \mathbb{R}^3 entonces podemos rotar a (a, b) en el plano xy obteniendo $(a \cos \alpha - b \sin \alpha, b \cos \alpha + a \sin \alpha)$ y dilatar c en el eje z obteniendo $2c$. De esta manera, obtenemos el OL de todo \mathbb{R}^3 que a (a, b, c) le hace corresponder $(a \cos \alpha - b \sin \alpha, b \cos \alpha + a \sin \alpha, 2c)$. Este OL es precisamente $f \oplus g$.

Ejercicio 106 Dado $f \oplus g \in \text{End}(\mathbb{E} \oplus \mathfrak{F})$ podemos definir a $f' = f \oplus \mathbb{I}$ y $g' = \mathbb{I} \oplus g$. Pruebe que $f \oplus g = f' \circ g' = g' \circ f'$. Esto significa que podemos pensar la suma directa como la composición de dos OLs que comutan. [195]**Subespacios invariantes, componentes irreducibles**

A partir de ahora y *en todo este capítulo* la función $h : \mathbb{E} \rightarrow \mathbb{E}$ es un OL del espacio vectorial *finito dimensional* \mathbb{E} . La **dimensión** de h es por definición la dimensión de \mathbb{E} .

La simplicidad de la operación de suma directa de OLs nos lleva a querer descomponer h en suma directa de otros OLs. La pregunta es: ¿Dado h será posible encontrar OLs f y g tales que $h = f \oplus g$? Detallemos un poco más el problema. Si \mathfrak{F} y \mathfrak{G} son subespacios complementarios de \mathbb{E} entonces, por el isomorfismo canónico entre la suma de subespacios complementarios y la suma directa de subespacios tenemos $\mathbb{E} = \mathfrak{F} + \mathfrak{G} = \mathfrak{F} \oplus \mathfrak{G}$. La pregunta es ¿cuando existen OLs $f : \mathfrak{F} \rightarrow \mathfrak{F}$ y $g : \mathfrak{G} \rightarrow \mathfrak{G}$ tales que $h = f \oplus g$?

Supongamos que efectivamente $h = f \oplus g$ y sea \mathbf{a} un vector en \mathfrak{F} . Por definición de $f \oplus g$ para calcular $h(\mathbf{a})$ tenemos que expresar \mathbf{a} como suma de un vector en \mathfrak{F} y otro en \mathfrak{G} . Esta descomposición es única y en este caso es $\mathbf{a} = \mathbf{a} + \mathbf{0}$ por lo que $h(\mathbf{a}) = f(\mathbf{a}) + g(\mathbf{0}) = f(\mathbf{a}) + \mathbf{0} = f(\mathbf{a})$. De aquí deducimos que $h(\mathbf{a}) \in \mathfrak{F}$ y como esto se hizo para un vector arbitrario en \mathfrak{F} obtenemos que $h(\mathfrak{F}) = \{h(\mathbf{a}) \mid \mathbf{a} \in \mathfrak{F}\} \subseteq \mathfrak{F}$.

De la misma manera se demuestra que $\mathbf{h}(\mathfrak{G}) \subseteq \mathfrak{G}$.

Esto nos lleva a la siguiente definición. Diremos que $\mathfrak{F} \subseteq \mathfrak{E}$ es un **subespacio invariante** de \mathbf{h} (o que \mathfrak{F} es **\mathbf{h} -invariante**) si se cumple que $\mathbf{h}(\mathfrak{F}) \subseteq \mathfrak{F}$.

6.1

Un OL se descompone como suma directa de dos OLs si y solo si él tiene dos subespacios invariantes complementarios.

Prueba. Ya demostramos la necesidad. Demostremos la suficiencia. Para esto supongamos que $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ tiene dos subespacios invariantes complementarios \mathfrak{F} y \mathfrak{G} . Tenemos $\mathfrak{E} = \mathfrak{F} \oplus \mathfrak{G}$, $\mathbf{h}(\mathfrak{F}) \subseteq \mathfrak{F}$ y $\mathbf{h}(\mathfrak{G}) \subseteq \mathfrak{G}$. Sean $f : \mathfrak{F} \rightarrow \mathfrak{F}$ y $g : \mathfrak{G} \rightarrow \mathfrak{G}$ las restricciones de la función \mathbf{h} a los subespacios \mathfrak{F} y \mathfrak{G} respectivamente. Obviamente f y g son OLs. Sea x un vector arbitrario en \mathfrak{E} . Existen unos únicos $\mathbf{a} \in \mathfrak{F}$, $\mathbf{b} \in \mathfrak{G}$ tales que $x = \mathbf{a} + \mathbf{b}$ y por linearidad tenemos $\mathbf{h}(x) = \mathbf{h}(\mathbf{a} + \mathbf{b}) = \mathbf{h}(\mathbf{a}) + \mathbf{h}(\mathbf{b}) = f(\mathbf{a}) + g(\mathbf{b})$ por lo que $\mathbf{h} = f \oplus g$. ■

Acabamos de traducir nuestro problema original al de la existencia de subespacios invariantes complementarios pero hay un caso degenerado que no nos facilita en nada las cosas. Todo el espacio \mathfrak{E} es invariante pues obviamente $\mathbf{h}(\mathfrak{E}) \subseteq \mathfrak{E}$. Igualmente el subespacio $\{\mathbf{0}\}$ formado solo por el origen es invariante ya que $\mathbf{h}(\mathbf{0}) = \mathbf{0}$. Además, los subespacios $\{\mathbf{0}\}$ y \mathfrak{E} son complementarios y en este caso \mathbf{h} se descompone como la suma directa de $f : \mathbf{0} \mapsto \mathbf{0}$ y $g = \mathbf{h}$. Tal descomposición siempre existe, pero no nos da nada ya que $g = h$.

Un subespacio se le llama **no trivial** si él no es ni todo el espacio y ni el origen. Un OL se le llama **reducible** si él tiene dos subespacios invariantes complementarios no triviales y en otro caso se le llama **irreducible**. A la restriccion de \mathbf{h} a un subespacio invariante no trivial que tiene un complementario invariante se le llama **componente** de \mathbf{h} . En otras palabras, f es una componente de \mathbf{h} si existe g tal que $\mathbf{h} = f \oplus g$ y esta descomposición es no trivial. Los OLs irreducibles son exactamente aquellos cuya única descomposición como suma directa de dos es la trivial o sea, aquellos que no tienen componentes. Si un OL es reducible entonces, este se descompone como suma directa de dos componentes. Si alguna de las componentes es reducible entonces podemos descomponerla. Así, vemos que cualquier OL es suma directa de componentes irreducibles.

Ejercicio 107 Sea α un escalar no nulo y f un OL. Pruebe que si f es irreducible entonces αf es irreducible. [195]

Ejercicio 108 Se dice que dos operadores f y g son conjugados si existe un OL biyectivo ρ tal que $f = \rho \circ g \circ \rho^{-1}$. Pruebe que la relación de conjugación es de equivalencia. [195]

Ejercicio 109 Pruebe que f y g son conjugados si y solo si existen bases A y B del espacio tales que la matriz de f en la base A es igual a la matriz de g en la base B . [195]

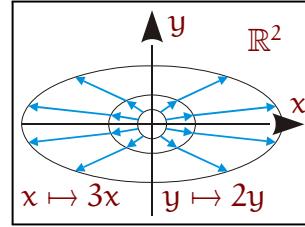
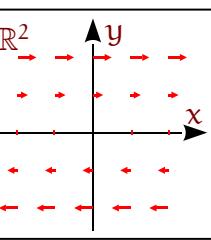
Ejercicio 110 Pruebe que si f es irreducible y g es un conjugado de f entonces, g también es irreducible. [195]

Ejemplos en dimensión 2

Todo OL de dimensión 1 es irreducible pues los únicos subespacios posibles son los triviales. Veamos que sucede en dimensión 2. Supongamos que \mathcal{E} es de dimensión 2. Si h es reducible entonces, $\mathcal{E} = \mathfrak{F} \oplus \mathfrak{G}$ y necesariamente \mathfrak{F} y \mathfrak{G} son subespacios h -invariantes de dimensión 1. Por la proposición 3.2 las restricciones f y g de h a \mathfrak{F} y \mathfrak{G} respectivamente son homotecias, o sea, existen escalares α y β tales que f es la multiplicación por α y g es la multiplicación por β .

Si $\{\mathbf{x}\}$ es una base de \mathfrak{F} y $\{\mathbf{y}\}$ es una base de \mathfrak{G} entonces la matriz de h en la base $\{\mathbf{x}, \mathbf{y}\}$ es la del recuadro a la izquierda, o sea es diagonal. Hemos demostrado que cualquier operador lineal reducible de dimensión 2 cumple que existe una base en la cual su matriz es diagonal. En la figura de la derecha está representada el OL de este tipo cuando $\mathcal{E} = \mathbb{R}^2$, $\{\mathbf{x}, \mathbf{y}\}$ es la base canónica, $\alpha = 3$ y $\beta = 2$.

Una rotación de \mathbb{R}^2 en un ángulo α en contra de las manecillas del reloj es obviamente irreducible para $\alpha \notin \{0^\circ, 180^\circ\}$ ya que en este caso, ninguna recta por el origen se queda en su lugar. O sea, las rotaciones no tienen subespacios invariantes no triviales.



Otro ejemplo es el que surge si a un cuadrado le aplicamos dos fuerzas en sentido contrario a dos de sus lados opuestos. Más precisamente, al OL que tiene como matriz en la base canónica de \mathbb{R}^2 la del recuadro a la derecha la llamaremos **λ -deslizamiento**. La figura de la izquierda muestra intuitivamente como se mueven los puntos de \mathbb{R}^2 al aplicar un λ -deslizamiento. De esta figura, se ve que la única recta por el origen que es invariante es el eje x . Luego un λ -deslizamiento es irreducible ($\lambda \neq 0$) ya que el eje x no tiene complementario invariante.

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

Ejercicio 111 ¿Será cierto que si un operador es irreducible entonces es biyectivo? [195]

Las matrices y los subespacios invariantes

Sea \mathfrak{F} es un subespacio invariante del OL $h : \mathcal{E} \rightarrow \mathcal{E}$. Sea \mathfrak{G} un subespacio complementario a \mathfrak{F} . Escojamos bases A y B de \mathfrak{F} y \mathfrak{G} respectivamente. Sabemos que $A \cup B$

es una base de todo el espacio. Sea \mathbf{x} un vector en la base \mathbf{A} . Podemos hallar las coordenadas de $\mathbf{h}(\mathbf{x})$ en la base $\mathbf{A} \cup \mathbf{B}$

$$\mathbf{h}(\mathbf{x}) = \sum_{\mathbf{a} \in \mathbf{A}} \alpha_{\mathbf{a}} \mathbf{a} + \sum_{\mathbf{b} \in \mathbf{B}} \beta_{\mathbf{b}} \mathbf{b}$$

y como $\mathbf{h}(\mathbf{x}) \in \langle \mathbf{A} \rangle$ entonces, todas las coordenadas $\beta_{\mathbf{b}}$ son iguales a cero.

(M	*
0	*	
)		

Estas coordenadas forman una columna de la matriz de \mathbf{h} en la base $\mathbf{A} \cup \mathbf{B}$ y por lo tanto despues de ordenar las bases, esta matriz tiene que verse como en el recuadro a la izquierda. La matriz \mathbf{M} es la de la restriccion de \mathbf{h} a \mathfrak{F} . El $\mathbf{0}$ representa una submatriz con entradas cero con $|\mathbf{A}|$ columnas y $|\mathbf{B}|$ renglones. Finalmente, los “*” representan submatrices de las dimensiones apropiadas. Resumiendo para futuras referencias:

6.2

Si \mathfrak{F} es un subespacio invariante de \mathbf{h} y \mathbf{A} es una base de todo el espacio que contiene a una base \mathbf{B} de \mathfrak{F} entonces, la matriz de \mathbf{h} en la base \mathbf{A} es triangular por bloques y el bloque superior izquierdo es la matriz de la restriccion de \mathbf{h} a \mathfrak{F} en la base \mathbf{B} .

Si además, \mathfrak{G} es \mathbf{h} invariante entonces, el mismo razonamiento nos hace ver que el “*” superior derecho es una submatriz con entradas cero con $|\mathbf{B}|$ columnas y $|\mathbf{A}|$ renglones. En este caso la matriz tiene que verse como en el recuadro a la derecha. La matriz \mathbf{M}' es la de la restriccion de \mathbf{h} a \mathfrak{G} . Resumiendo para futuras referencias:

(M	0
0		M'

6.3

Si \mathfrak{F} y \mathfrak{G} son subespacios invariantes complementarios de \mathbf{h} y los conjuntos \mathbf{A} y \mathbf{B} son bases de \mathfrak{F} y \mathfrak{G} respectivamente entonces, la matriz de \mathbf{h} en la base $\mathbf{A} \cup \mathbf{B}$ es diagonal por bloques y los bloques diagonales son las matrices de las restricciones de \mathbf{h} a \mathfrak{F} en la base \mathbf{A} y de \mathbf{h} a \mathfrak{G} en la base \mathbf{B} respectivamente.

Ejercicio 112 Demuestre que $(\mathbf{f} \oplus \mathbf{g})^2 = \mathbf{f}^2 \oplus \mathbf{g}^2$. Traduzca esta igualdad al lenguaje de matrices.

6.2 Polinomios de operadores lineales

En esta sección introduciremos una herramienta para el cálculo de los subespacios invariantes de un OL a saber, los polinomios de OLs. Se recomienda que el lector lea (o vuelva a leer) la sección 5.1 en la cual se introducen los ideales de polinomios y se

demuestra que todo ideal de $\mathbb{K}[x]$ es principal.

El morfismo de $\mathbb{K}[x]$ en $\text{End}(\mathfrak{E})$

Para cualquier número natural n el operador h^n se define como en el recuadro a la derecha. De aquí, como sabemos sumar OLs y multiplicar por escalares a los OLs, vemos que una expresión como por ejemplo $h^2 - 2h^1 + 5\mathbb{I}$ es un OL que conocemos si tan solo conocemos a h .

$$h^n = \begin{cases} \overbrace{h \circ \dots \circ h}^{n \text{ veces}} & \text{si } n > 0 \\ \mathbb{I} & \text{si } n = 0 \end{cases}$$

En general si $p(x) = \sum_{i=0}^n \alpha_i x^i \in \mathbb{K}[x]$ es un polinomio arbitrario con coeficientes en el campo del espacio vectorial \mathfrak{E} , entonces $p_h = \sum_{i=0}^n \alpha_i h^i$ es un OL bien definido. Al proceso de substituir la variable x por el OL h y de esta manera convertir el polinomio $p(x)$ en el OL p_h se le llama **evaluación** de p en h .



El lector debe prestar atención a la notación. Usaremos p_h y no $p(h)$ aunque al parecer esta última es más natural. El problema es que tendremos que evaluar polinomios en operadores lineales y a su vez estos en vectores. Si usaramos la notación $p(h)$ tendríamos que escribir $p(h)(a)$ y estos son demasiados paréntesis en comparación con $p_h(a)$.

Recordemos de la sección 3.2 que un álgebra es un espacio vectorial con un producto de vectores asociativo, distributivo, con elemento neutro y que conmuta con el producto por escalares. Ya vimos, que el espacio vectorial $\text{End}(\mathfrak{E})$ es un álgebra para la composición de OL. También vimos que el espacio vectorial de todos los polinomios $\mathbb{K}[x]$ es un álgebra para el producto de polinomios.

Una **subálgebra** es un subconjunto de una álgebra que es álgebra para las operaciones inducidas en el subconjunto. Un subconjunto de una álgebra es subálgebra cuando es subespacio del espacio vectorial, es cerrado para el producto y contiene el 1 .

Una transformación lineal entre dos álgebras es un **morfismo de álgebras** si esta conmuta con el producto y preserva el 1 .

6.4

La función de evaluación de polinomios en un operador lineal es un morfismo de álgebras.

Prueba. La función de evaluación $\mathbb{K}[x] \ni p(x) \mapsto p_h \in \text{End}(\mathfrak{E})$ es una función cuyo dominio y codominio son álgebras. Sean $\alpha_0, \dots, \alpha_n$ y β_0, \dots, β_n los coeficientes de dos polinomios p y q respectivamente. Tenemos la misma cantidad de coeficientes en ambos polinomios ya que siempre podemos agregar suficientes ceros. Sea λ un escalar. Tenemos

$$(\lambda p)_h = \sum_{i=0}^n \lambda \alpha_i h^i = \lambda \sum_{i=0}^n \alpha_i h^i = \lambda p_h$$

$$(p + q)_h = \sum_{i=0}^n (\alpha_i + \beta_i) h^i = \sum_{i=0}^n \alpha_i h^i + \sum_{i=0}^n \beta_i h^i = p_h + q_h$$

y esto muestra que la evaluación en \mathbf{h} es una TL.

Por definición de producto de polinomios, tenemos

$$\begin{aligned} (\mathbf{p}\mathbf{q})_{\mathbf{h}} &= \left(\sum_{i=0}^n \sum_{j=0}^n \alpha_i \beta_j x^{i+j} \right)_{\mathbf{h}} = \sum_{i=0}^n \sum_{j=0}^n \alpha_i \beta_j \mathbf{h}^{i+j} = \sum_{i=0}^n \sum_{j=0}^n \alpha_i \beta_j (\mathbf{h}^i \circ \mathbf{h}^j) = \\ &= \sum_{i=0}^n \sum_{j=0}^n (\alpha_i \mathbf{h}^i \circ \beta_j \mathbf{h}^j) = \sum_{i=0}^n \alpha_i \mathbf{h}^i \circ \sum_{j=0}^n \beta_j \mathbf{h}^j = \mathbf{p}_{\mathbf{h}} \circ \mathbf{q}_{\mathbf{h}} \end{aligned}$$

y finalmente, con $\mathbf{1}_{\mathbf{h}} = (x^0)_{\mathbf{h}} = \mathbf{h}^0 = \mathbb{I}$ terminamos la prueba. ■

La subálgebra $\mathbb{K}[\mathbf{h}]$

El morfismo de evaluación en \mathbf{h} tiene como imagen el conjunto de todos los OL que son la evaluación en \mathbf{h} de algún polinomio de $\mathbb{K}[x]$. Este conjunto de OLs se denotará por $\mathbb{K}[\mathbf{h}]$. Esto refleja que en la evaluación lo que hacemos es substituir la variable x por el OL \mathbf{h} .

6.5

$\mathbb{K}[\mathbf{h}]$ es una subálgebra conmutativa del álgebra de operadores lineales.

Prueba. Como la imagen de una transformación lineal es un subespacio, obtenemos que $\mathbb{K}[\mathbf{h}]$ es un subespacio de $\text{End}(\mathfrak{E})$. Como $\mathbf{1}(\mathbf{h}) = \mathbb{I}$, obtenemos que $\mathbb{I} \in \mathbb{K}[\mathbf{h}]$. Como $\mathbf{p}_{\mathbf{h}} \circ \mathbf{q}_{\mathbf{h}} = (\mathbf{p}\mathbf{q})_{\mathbf{h}}$ obtenemos que $\mathbb{K}[\mathbf{h}]$ es cerrado para la composición. La conmutatividad se sigue de la conmutatividad del producto de polinomios. Efectivamente, $\mathbf{p}_{\mathbf{h}} \circ \mathbf{q}_{\mathbf{h}} = (\mathbf{p}\mathbf{q})_{\mathbf{h}} = (\mathbf{q}\mathbf{p})_{\mathbf{h}} = \mathbf{q}_{\mathbf{h}} \circ \mathbf{p}_{\mathbf{h}}$ y por lo tanto, los OLs en $\mathbb{K}[\mathbf{h}]$ comutan entre ellos. ■

La conmutatividad de $\mathbb{K}[\mathbf{h}]$ es un hecho trivial pero muy notable. Los operadores lineales en general, no son conmutativos para la composición. Sin embargo, los que están en $\mathbb{K}[\mathbf{h}]$ sí comutan entre ellos. Esto jugará un papel importante en lo que sigue.



Cada vez que se tiene un morfismo de álgebras, la imagen de este morfismo es una subálgebra del codominio del morfismo. Si el dominio del morfismo es un álgebra conmutativa entonces la imagen es una subálgebra conmutativa.

El polinomio mínimo

El morfismo de evaluación en \mathbf{h} tiene como núcleo el conjunto de todos los polinomios \mathbf{p} tales que $\mathbf{p}_{\mathbf{h}} = \mathbb{O}$.

6.6

El núcleo del morfismo de evaluación en \mathbf{h} es un ideal de $\mathbb{K}[x]$.

Prueba. El núcleo de una transformación lineal es un subespacio y por lo tanto si $\mathbf{p}_{\mathbf{h}} = \mathbf{q}_{\mathbf{h}} = \mathbb{O}$ entonces $(\mathbf{p} + \mathbf{q})_{\mathbf{h}} = \mathbb{O}$. Sea $\mathbf{r}(x)$ cualquier polinomio. Necesitamos mostrar que $(\mathbf{r}\mathbf{q})_{\mathbf{h}} = \mathbb{O}$. Para cualquier $\mathbf{a} \in \mathfrak{E}$ tenemos

$$(\mathbf{r}\mathbf{q})_{\mathbf{h}}(\mathbf{a}) = (\mathbf{r}_{\mathbf{h}} \circ \mathbf{q}_{\mathbf{h}})(\mathbf{a}) = \mathbf{r}_{\mathbf{h}}(\mathbf{q}_{\mathbf{h}}(\mathbf{a})) = \mathbf{r}_{\mathbf{h}}(\mathbf{0}) = \mathbf{0}$$

y esto es todo lo que queríamos probar. ■

Por 5.6 todo ideal de $\mathbb{K}[x]$ es principal y por lo tanto existe un único polinomio h (nótese la letra gótica) de coeficiente principal 1 (o sea, mónico) tal que si $p_h = \mathbb{O}$ entonces, p es un múltiplo de h . En símbolos matemáticos $\{p(x) \in \mathbb{K}[x] \mid p_h = \mathbb{O}\} = \{qh \mid q \in \mathbb{K}[x]\}$. Al polinomio h se le llama **polinomio mínimo** de h . El polinomio mínimo de h es el polinomio mónico de grado más pequeño que al evaluarlo en h se obtiene el OL nulo \mathbb{O} . Por ahora, no sabemos mucho del polinomio mínimo de h , solo sabemos que existe y que es único.

Otra manera más descriptiva de ver el polinomio mínimo es la siguiente. Consideraremos la sucesión infinita de operadores $\mathbb{I}, h, h^2, \dots$. Todos los elementos de esta sucesión no pueden ser LI en el espacio vectorial $\text{End}(\mathfrak{E})$ porque este espacio es de dimensión finita e igual a $(\dim \mathfrak{E})^2$. Esto quiere decir que hay un primer natural n y unos coeficientes escalares α_i tales que $h^n = \alpha_0 h^0 + \alpha_1 h^1 + \dots + \alpha_{n-1} h^{n-1}$. Denotando $p(x) = x^n - \alpha_{n-1} x^{n-1} - \dots - \alpha_1 x^1 - \alpha_0 x^0$ vemos que $p_h = \mathbb{O}$ y que este es el único polinomio mónico de grado más pequeño que cumple esto. Luego, p es el polinomio mínimo de h .

El período de un vector

Si p es un polinomio entonces p_h es un OL. Si a es un vector entonces a la imagen por p_h de a se le denotará por $p_h(a)$. Luego, $p_h(a)$ se obtiene en dos pasos: primero tomamos el polinomio p lo evaluamos en h y así obtenemos p_h ; segundo la función p_h la evaluamos en a y así obtenemos $p_h(a)$.

6.7

Para cualquier vector a , el conjunto de todos los polinomios p tales que $p_h(a) = \mathbf{0}$, es un ideal.

Prueba. Si $p_h(a) = q_h(a) = \mathbf{0}$ y r es cualquier polinomio entonces
 $(p+q)_h(a) = (p_h + q_h)(a) = p_h(a) + q_h(a) = \mathbf{0}$
 $(rp)_h(a) = r_h(p_h(a)) = r_h(\mathbf{0}) = \mathbf{0}$

y esto es todo lo que se requería probar. ■

Nuevamente, como todo ideal de polinomios es principal entonces, existe un único polinomio q tal que el conjunto $\{p \in \mathbb{K}[x] \mid p_h(a) = \mathbf{0}\}$ es exactamente el conjunto de todos polinomios que son múltiplos de q . Al polinomio q se le llama el **h -período** del vector a o sencillamente el **período** de a si está implícito cual es el operador h . En otras palabras, el período de a es el polinomio q mónico de grado más pequeño tal que $q_h(a) = \mathbf{0}$.

Más descriptivamente. En la sucesión infinita de vectores $a, h(a), h^2(a), \dots$ todos los elementos no pueden ser LI porque el espacio \mathfrak{E} es de dimensión finita. Esto quiere decir que hay un primer natural n y unos coeficientes escalares α_i tales que $h^n = \alpha_0 h^0(a) + \dots + \alpha_{n-1} h^{n-1}(a)$. Denotando $p(x) = x^n - \alpha_{n-1} x^{n-1} - \dots - \alpha_1 x^1 - \alpha_0 x^0$

vemos que $p_h(\mathbf{a}) = \mathbf{0}$ y que este es el único polinomio mónico de grado más pequeño que cumple esto. Luego, $p(x)$ es el período de \mathbf{a} .

Ejercicio 113 Pruebe que $\mathbf{0}$ es el único vector cuyo período es de grado cero. [195]

Ejercicio 114 Pruebe que los vectores no nulos cuyo período es el polinomio x son exactamente aquellos que están en el núcleo de h . [196]

Anuladores

Sea ahora $A \subseteq \mathfrak{E}$ un conjunto arbitrario de vectores. El **h -anulador** de A es el conjunto de polinomios $\{p \in \mathbb{K}[x] \mid \forall \mathbf{a} \in A \ p_h(\mathbf{a}) = \mathbf{0}\}$. Previamente ya habíamos considerado dos anuladores. En el caso de que A es todo el espacio entonces el anulador es el ideal usado para definir el polinomio mínimo. En el caso de que A es un solo vector entonces el anulador es el ideal usado para definir el período del vector. Análogamente a las pruebas de 6.6 y 6.7 podemos probar que cualquier anulador es un ideal. Esto nos permite definir el **h -período** de un conjunto de vectores como el polinomio generador de su anulador. El anulador de un conjunto de vectores es el conjunto de polinomios que son múltiplos de su período.



De aquí en lo adelante denotaremos por $\text{per}_h(A)$ al h -período de un conjunto de vectores A . Así, $\text{per}_h(\mathbf{a})$ es el h -período del vector \mathbf{a} y $\text{per}_h(\mathfrak{E})$ es el h -período de todo el espacio o sea, el polinomio mínimo de h .

Propiedades del período.

6.8

El h -anulador de A es el conjunto de los múltiplos comunes a los períodos de los vectores en A .

Prueba. Sea $A^0 = \{p \in \mathbb{K}[x] \mid \forall \mathbf{a} \in A \ p_h(\mathbf{a}) = \mathbf{0}\}$ el h -anulador de A . Sea $A^1 = \{p \in \mathbb{K}[x] \mid \forall \mathbf{a} \in A \ p \vdash \text{per}_h(\mathbf{a})\}$ el conjunto de los múltiplos comunes a los períodos de los vectores en A .

Si $p \in A^1$ y $\mathbf{a} \in A$, entonces existe q tal que $p = q \text{per}_h(\mathbf{a})$ y por lo tanto $p_h(\mathbf{a}) = q_h(\text{per}_h(\mathbf{a})) = q_h(\mathbf{0}) = \mathbf{0}$. Luego, $A^1 \subseteq A^0$.

Recíprocamente, sean $p \in A^0$ y $\mathbf{a} \in A$ entonces $p_h(\mathbf{a}) = \mathbf{0}$ y por definición de período $\text{per}_h(\mathbf{a}) \vdash p$. Luego, $A^0 \subseteq A^1$. ■

Una consecuencia directa de esto son los siguientes dos resultados:

6.9

$\text{per}_h(A)$ es el mínimo común múltiplo de los períodos de los vectores en A .

6.10

El polinomio mínimo de \mathbf{h} es el mínimo común múltiplo de los períodos de todos los vectores.

6.11

Si $\mathbf{h} = \mathbf{f} \oplus \mathbf{g}$ entonces el polinomio mínimo de \mathbf{h} es igual al mínimo común múltiplo de los polinomios mínimos de \mathbf{f} y \mathbf{g} .

Prueba. Demostraremos que el \mathbf{h} -anulador de todo el espacio es igual al conjunto de los múltiplos comunes de los polinomios mínimos de \mathbf{f} y \mathbf{g} . Lo que se quiere demostrar es una consecuencia directa de esto.

Sea $\mathfrak{E} = \mathfrak{E}_1 \oplus \mathfrak{E}_2$ la descomposición en subespacios invariantes de tal manera que \mathbf{f} y \mathbf{g} son las restricciones de \mathbf{h} a \mathfrak{E}_1 y \mathfrak{E}_2 respectivamente. Sean $\mathbf{a} \in \mathfrak{E}_1$ y $\mathbf{b} \in \mathfrak{E}_2$. Cualquier vector en \mathfrak{E} es de la forma $\mathbf{a} + \mathbf{b}$.

Si \mathbf{p} es un común múltiplo de los polinomios mínimos de \mathbf{f} y \mathbf{g} entonces $p_{\mathbf{h}}(\mathbf{a} + \mathbf{b}) = p_{\mathbf{f}}(\mathbf{a}) + p_{\mathbf{g}}(\mathbf{b}) = \mathbf{0}$ y por lo tanto \mathbf{p} está en el \mathbf{h} -anulador de todo el espacio.

Recíprocamente si \mathbf{p} está en el \mathbf{h} -anulador de todo el espacio, entonces tiene que anular a todos los vectores en \mathfrak{E}_1 y por lo tanto es un múltiplo del polinomio mínimo de \mathbf{f} . Por la misma razón es un múltiplo del polinomio mínimo de \mathbf{g} . ■

monotonía del período

6.12

Si $A \subseteq B$ entonces, $\text{per}_{\mathbf{h}}(A) \dashv \text{per}_{\mathbf{h}}(B)$.

Prueba. Sea A^0 el \mathbf{h} -anulador de A . Si \mathbf{p} es el período de B entonces $p_{\mathbf{h}}(\mathbf{b}) = \mathbf{0}$ para cualquier $\mathbf{b} \in B$ y por lo tanto $p_{\mathbf{h}}(\mathbf{a}) = \mathbf{0}$ para cualquier $\mathbf{a} \in A$. Luego $\mathbf{p} \in A^0$ y por lo tanto es un múltiplo del período de A . ■

6.3 Subespacios radicales

Núcleos de polinomios de operadores lineales

Los núcleos de los polinomios evaluados en un OL son un objeto importante para la descomposición de ese OL en componentes irreducibles. Su importancia está dada por el siguiente resultado.

invariancia de los núcleos

6.13

El núcleo de cualquier operador en $\mathbb{K}[\mathbf{h}]$ es un subespacio \mathbf{h} -invariante.

Prueba. Sabemos que el núcleo de cualquier OL es un subespacio. Demostremos la invariancia. Sea p un polinomio y $a \in \ker p_h$. Entonces $p_h(a) = 0$. Por la commutatividad de los OL en $\mathbb{K}[h]$ tenemos $p_h(h(a)) = h(p_h(a)) = h(0) = 0$. O sea, $h(a) \in \ker p_h$. ■

monotonía de los núcleos

6.14

Si $p \dashv q$ entonces $\ker p_h \subseteq \ker q_h$.

Prueba. Sea $q = p'p$. Si $x \in \ker p_h$ entonces $q_h(x) = p'_h(p_h(x)) = p'_h(0) = 0$ y por lo tanto $x \in \ker q_h$. ■

En lo que sigue muy frecuentemente nos encontraremos parejas de polinomios sin factores comunes. Necesitamos hacer un aparte para hablar de estas parejas. Primero, les daremos un nombre más corto. Dos polinomios p y q se les llama **coprimos** si cualquier divisor común a ambos es de grado 0. O sea, no tienen factores comunes no triviales. Dos polinomios p y q son coprimos si y solo si los factores irreducibles de p son diferentes a los factores irreducibles de q .

Lema de Descomposición de Núcleos

6.15

Si p y q son polinomios coprimos entonces, $\ker(pq)_h = \ker p_h \oplus \ker q_h$.

Prueba. Todos los núcleos involucrados son subespacios. Lo que hay que probar es que $\ker p_h \cap \ker q_h = \{0\}$ y que $\ker(pq)_h = \ker p_h + \ker q_h$. O sea, es una suma directa.

Por el Teorema de Bezout existen polinomios r, s tales que $rp + sq = 1$.

Sea $x \in \ker p_h \cap \ker q_h$. Tenemos que

$$x = 1_h(x) = (rp + sq)_h(x) = r_h(p_h(x)) + s_h(q_h(x)) = r_h(0) + s_h(0) = 0$$

lo que prueba que $\ker p_h \cap \ker q_h = \{0\}$.

Sea $x \in \ker(pq)_h$ y denotemos $y = (sq)_h(x)$, $z = (rp)_h(x)$. Como $rp + sq = 1$ tenemos $z + y = x$. Además, por la commutatividad tenemos que

$$p_h(y) = (psq)_h(x) = s_h((pq)_h(x)) = s_h(0) = 0$$

$$q_h(z) = (qrp)_h(x) = r_h((pq)_h(x)) = r_h(0) = 0$$

o sea, $y \in \ker p_h$ y $z \in \ker q_h$. Luego $\ker(pq)_h \subseteq \ker p_h + \ker q_h$.

Para probar la otra inclusión sean $a \in \ker p(h)$ y $b \in \ker q(h)$. Tenemos que

$$\begin{aligned} (pq)_h(a + b) &= p_h(q_h(a + b)) = p_h(q_h(a) + q_h(b)) = \\ &= p_h(q_h(a)) = q_h(p_h(a)) = q_h(0) = 0 \end{aligned}$$

y por lo tanto $\ker p_h + \ker q_h \subseteq \ker(pq)_h$. ■

Operadores lineales radicales

Por la invariancia de los núcleos (6.13) y el Lema de Descomposición de Núcleos (6.15) si $(pq)_h = \mathbb{O}$ y p, q son coprimos entonces, h tiene dos subespacios invariantes

complementarios $\ker p_h$ y $\ker q_h$ y podríamos tener (si la descomposición es no trivial) que h es reducible. El candidato que tenemos para el polinomio pq es el polinomio mínimo de h .

Un polinomio p se descompone en producto de dos factores coprimos si y solo si p tiene al menos dos factores irreducibles distintos. Esto nos lleva a la siguiente definición. Diremos que h es **radical de tipo p** si el polinomio mínimo de h tiene un solo factor irreducible e igual a p . Equivalentemente, h es radical si el período de cualquier vector es de la forma p^m donde p es un polinomio mónico sin factores no triviales.



En la teoría de anillos el radical de un ideal I es el conjunto de todos los elementos x tales que para cierto natural m se tiene que $x^m \in I$. El radical de cualquier ideal es un ideal. En este lenguaje, un operador es radical de tipo p (irreducible) si el radical del anulador del espacio es el ideal generado por p .

6.16

Si h es irreducible entonces, es radical.

Prueba. Si h no es radical entonces el polinomio mínimo h de h se descompone no trivialmente como un producto $h = pq$ donde p y q son coprimos. Por el Lema de Descomposición de Núcleos (6.15) tenemos $\mathcal{E} = \ker(pq)_h = \ker p_h \oplus \ker q_h$ siendo $\ker p_h$ y $\ker q_h$ subespacios invariantes de h (por la invariancia de los núcleos (6.13)). Como p es un factor propio de h que es el mínimo común múltiplo de los períodos de todos los vectores entonces, $\ker p_h \neq \mathcal{E}$ y por la misma razón $\ker q_h \neq \mathcal{E}$. Esto quiere decir que la descomposición $\ker p_h \oplus \ker q_h$ es no trivial y por lo tanto h es reducible. ■

Este resultado no alcanza para caracterizar a los operadores lineales irreducibles. Por ejemplo la identidad en \mathbb{R}^2 tiene polinomio mínimo $x - 1$ o sea, es radical. Sin embargo, es evidentemente la suma directa de la identidad en el eje x y la identidad en el eje y .

Componentes radicales

Un vector se le llama **radical de tipo p** si su período es igual a p^m y p es irreducible. Un conjunto de vectores se le llama **radical de tipo p** si todos sus vectores son radicales de tipo p .

6.17

Si p es factor irreducible de multiplicidad m del polinomio mínimo de h entonces, $\ker p_h^m$ es el conjunto de todos los vectores radicales de tipo p .

Prueba. Sea a un vector de período p^k . Si $k > m$ entonces $p^k = \text{per}_h(a)$ no divide al polinomio mínimo y esto no puede ser. Luego $k \leq m$. De la monotonía de los núcleos (6.14) obtenemos $\ker p_h^k \subseteq \ker p_h^m$ y por lo tanto $a \in \ker p_h^m$. Recíprocamente, si $a \in \ker p_h^m$ entonces $p_h^m(a) = 0$ y por lo tanto $\text{per}_h(a)$ es un divisor de p^m . Como

p es irreducible, necesariamente $\text{per}_h(a)$ es igual a p^k para cierto $k \leq m$. ■

Por el teorema de descomposición de un polinomio en factores irreducibles el polinomio mínimo de h es igual a un producto $\prod_{p \in P} p^{m_p}$ donde P es el conjunto de sus factores irreducibles y m_p es la multiplicidad del polinomio irreducible p . Por el resultado anterior, los espacios $\ker p_h^{m_p}$ son los **subespacios radicales maximales** de h . De la invariancia de los núcleos (6.13) sabemos que los subespacios radicales maximales son invariantes. A la restricción de h a un subespacio radical maximal se le llama **componente radical** de h .

Teorema de Descomposición en Componentes Radicales

6.18 *Todo operador lineal es la suma directa de sus componentes radicales.*

Prueba. Sea $h = \prod_{p \in P} p^{m_p}$ el polinomio mínimo de h . Si en P hay un solo polinomio irreducible entonces h es radical y no hay nada que probar. Hagamos inducción en el número de polinomios irreducibles en P . Sea r un polinomio irreducible fijo pero arbitrario en P . Denotemos $q = r^{m_r}$ y $q' = \prod_{p \in P \setminus r} p^{m_p}$. Tenemos que $h = qq'$ y que q, q' son coprimos. Del Lema de Descomposición de Núcleos (6.15) obtenemos la descomposición en subespacios invariantes complementarios $\mathfrak{E} = \mathfrak{F} \oplus \mathfrak{G}$ donde $\mathfrak{F} = \ker q_h$ y $\mathfrak{G} = \ker q'_h$.

Sean f y g las restricciones de h a \mathfrak{F} y \mathfrak{G} respectivamente. Tenemos que $f = f \oplus g$ y que f es una componente radical. De 6.11 el polinomio mínimo de g es q . Como q tiene menos factores irreducibles que h , podemos aplicar hipótesis de inducción. ■

Como la descomposición de un polinomio en polinomios irreducibles es única salvo orden de los factores entonces, la descomposición de un operador lineal en componentes radicales es única salvo orden de los sumandos.

Existencia de un vector de período máximo

La descomposición en componentes radicales nos permite limitarnos a considerar el caso en que el operador lineal h es radical. Esto simplifica mucho las cosas por la simplicidad de la relación de divisibilidad entre los divisores de los polinomios tipo p^n cuando p es irreducible. Esta relación orden es total. Por ejemplo, si A es un conjunto de divisores de p^n entonces se cumple que el mínimo común múltiplo de A es un polinomio en A .

6.19 *Para cualquier operador lineal h existe un vector tal que su período es igual al polinomio mínimo de h .*

Prueba. Sea h el polinomio mínimo de $h : \mathfrak{E} \rightarrow \mathfrak{E}$. Primero para el caso cuando $h = p^n$ con p irreducible. El polinomio h es el mínimo común múltiplo de los períodos de todos los vectores y todos estos son divisores de p^n . Por la observación que precede

a este resultado tiene que ocurrir que uno de esos períodos es p^n .

El caso general por inducción en el número de componentes radicales. Descompongamos $h = f \oplus g$ donde f es una componente radical. Esta descomposición se corresponde con la descomposición $E = F \oplus G$ en subespacios invariantes. Los operadores f y g son las restricciones de h a F y G respectivamente. Los polinomios mínimos f y g de f y g respectivamente son coprimos y cumplen que $h = fg$. Por hipótesis de inducción existen vectores a y b en F y G respectivamente tales que $f = \text{per}_h(a)$ y $g = \text{per}_h(b)$.

Denotemos $p = \text{per}_h(a - b)$. Tenemos $p \vdash h = fg$. Además, $(p_h(a - b) = 0) \Rightarrow (p_h(a) = p_h(b))$. Como F y G son invariantes $F \ni p_h(a) = p_h(b) \in G$. Como F y G son complementarios $p_h(a) = p_h(b) = 0$. Luego, $p \vdash \text{per}_h(a) = f$ y $p \vdash \text{per}_h(b) = g$. Como f y g son coprimos entonces, $p \vdash fg = h$. ■

6.4 Subespacios cílicos

h-combinaciones

Sea $h : E \rightarrow E$ un operador lineal. Sea $V = \{v_1, \dots, v_n\} \subseteq E$ un conjunto de vectores. Una **h-combinación** de V es un vector de la forma

$$p_h(v_1) + q_h(v_2) + \cdots + r_h(v_n)$$

donde los **coeficientes** p, q, \dots, r son polinomios arbitrarios en $\mathbb{K}[x]$.

Le dejamos al lector dar la definición para el caso de que V es infinito. En este caso, hay que exigir soporte finito, o sea que el conjunto de coeficientes no nulos sea finito.

Las **h**-combinaciones son combinaciones lineales en el caso de que los coeficientes sean polinomios de grado cero. Recordemos que $\langle V \rangle$ denota el conjunto de todas las combinaciones lineales de V . Denotemos por $\langle V \rangle_h$ el conjunto de todas las **h**-combinaciones de V . La observación anterior significa que $\langle V \rangle \subseteq \langle V \rangle_h$.

Conjuntos **h**-generadores

6.20

El conjunto de todas las **h**-combinaciones de V es un subespacio **invariante**.

Prueba. Sea λ un escalar. Sean $a = p_h(v_1) + \cdots + q_h(v_n)$ y $b = r_h(v_1) + \cdots + s_h(v_n)$ dos **h**-combinaciones de V . Entonces,

$$a + b = p'_h(v_1) + \cdots + q'_h(v_n) \quad \text{donde } p' = p + r, \dots, q' = q + s,$$

$$\lambda a = p'_h(v_1) + \cdots + q'_h(v_n) \quad \text{donde } p' = \lambda p, \dots, q' = \lambda q,$$

$$h(a) = p'_h(v_1) + \cdots + q'_h(v_n) \quad \text{donde } p'(x) = xp(x), \dots, q'(x) = xq(x).$$

Las dos primeras igualdades prueban que es un subespacio. La tercera muestra que es invariante. ■

Ya es la tercera vez que nos tropezamos con una situación similar. Las anteriores fueron la cerradura lineal y la cerradura afín. Los ejercicios siguientes tres resultados muestran que la función $V \mapsto \langle V \rangle_h$ es un operador de cerradura que llamaremos **h -cerradura**. Las pruebas son completamente análogas a las que dimos en el Capítulo 2 para la cerradura lineal.

Ejercicio 115 Pruebe que la intersección de subespacios invariantes es invariante.

Ejercicio 116 Pruebe que $\langle V \rangle_h$ es la intersección de todos los subespacios invariantes que contienen a V .

Ejercicio 117 Pruebe que la h -cerradura cumple las siguientes propiedades:

- ◆ $V \subseteq \langle V \rangle_h$ (incremento),
- ◆ $V' \subseteq V \Rightarrow \langle V' \rangle_h \subseteq \langle V \rangle_h$ (monotonía),
- ◆ $\langle \langle V \rangle_h \rangle_h = \langle V \rangle_h$ (idempotencia).

A $\langle V \rangle_h$ le llamaremos el subespacio invariante **h -generado** por V . Si $\langle V \rangle_h$ es todo el espacio diremos que V es un **h -generador**. Obviamente los sobreconjuntos de h -generadores y los conjuntos generadores (sin la h) son h -generadores.

$$\text{per}_h \langle V \rangle_h = \text{per}_h V.$$

Prueba. Tenemos $V \subseteq \langle V \rangle_h$ y por la monotonía del período (6.12) sabemos que $\text{per}_h V \dashv \text{per}_h \langle V \rangle_h$. Denotemos $q = \text{per}_h V$. Si $x \in \langle V \rangle_h$ entonces, x es una h -combinación $x = p_h(v_1) + \dots + r_h(v_n)$, donde $\{v_1, \dots, v_n\} \subseteq V$. De la linearidad y commutatividad obtenemos que

$$q_h(x) = q_h(p_h(v_1)) + \dots + q_h(r_h(v_n)) = p_h(q_h(v_1)) + \dots + r_h(q_h(v_n)) = 0.$$

Luego, $\text{per}_h V$ está en el anulador de $\langle V \rangle_h$ y por lo tanto $\text{per}_h V \vdash \text{per}_h \langle V \rangle_h$. ■

Subespacios cíclicos

Un subespacio invariante se le llama **h -cíclico** si este está h -generado por un solo vector. Los subespacios cíclicos son los h -análogos de las rectas por el origen que están generadas (sin la h) por un solo vector. Al operador h se le llama **cíclico** si todo el espacio es h -cíclico.

El siguiente resultado es “análogo” a que si una recta por el origen está generada por a entonces también está generada por los múltiplos de a .

6.22 Si q es un polinomio coprimo con $\text{per}_h a$ entonces, $\langle a \rangle_h = \langle q_h(a) \rangle_h$ y $\text{per}_h a = \text{per}_h q_h(a)$.

Prueba. Denotemos $p = \text{per}_h a$. Sea q un polinomio coprimo con p y denotemos

$\mathbf{a}' = \mathbf{q}_h(\mathbf{a})$. Demostremos que existe un polinomio \mathbf{r} tal que $\mathbf{a} = \mathbf{r}_h(\mathbf{a}')$. Efectivamente por el Teorema de Bezout existen polinomios \mathbf{s} y \mathbf{r} tales que $\mathbf{s}\mathbf{p} + \mathbf{r}\mathbf{q} = \mathbf{1}$ y por lo tanto

$$\mathbf{r}_h(\mathbf{a}') = (\mathbf{r}\mathbf{q})_h(\mathbf{a}) = (\mathbf{s}\mathbf{p})_h(\mathbf{a}) + (\mathbf{r}\mathbf{q})_h(\mathbf{a}) = \mathbb{I}(\mathbf{a}) = \mathbf{a}$$

y así el polinomio \mathbf{r} cumple lo que queremos.

Luego $\mathbf{a}' \in \langle \mathbf{a} \rangle_h$ y $\mathbf{a} \in \langle \mathbf{a}' \rangle_h$. Usando la monotonía y la idempotencia de la h -cerradura obtenemos que $\langle \mathbf{a}' \rangle_h \subseteq \langle \mathbf{a} \rangle_h$ y $\langle \mathbf{a} \rangle_h \in \langle \mathbf{a}' \rangle_h$.

Por esto, usando 6.21 obtenemos que

$$\text{per}_h \mathbf{q}_h(\mathbf{a}) = \text{per}_h \langle \mathbf{q}_h(\mathbf{a}) \rangle_h = \text{per}_h \langle \mathbf{a} \rangle_h = \text{per}_h \mathbf{a}.$$

Ahora veremos que los subespacios cílicos pueden tener dimensión grande. Esto significa que la analogía con las rectas por el origen no hay que llevarla demasiado lejos.

6.23

Si el período de \mathbf{a} es de grado n , entonces el conjunto de vectores $\{\mathbf{a}, h(\mathbf{a}), \dots, h^{n-1}(\mathbf{a})\}$ es una base de $\langle \mathbf{a} \rangle_h$.

Prueba. Denotemos $\mathbf{p} = \text{per}_h \mathbf{a}$ y $\mathbf{B} = \{\mathbf{a}, h(\mathbf{a}), \dots, h^{n-1}(\mathbf{a})\}$. Tenevamos que convencernos que \mathbf{B} es una base del subespacio $\langle \mathbf{a} \rangle_h$. Si hubiera una combinación lineal

$$\beta_0 \mathbf{a} + \beta_1 h(\mathbf{a}) + \dots + \beta_{n-1} h^{n-1}(\mathbf{a}) = \mathbf{0}$$

con no todos sus coeficientes nulos entonces, el polinomio no nulo

$$\mathbf{q}(x) = \beta_0 + \beta_1 x + \dots + \beta_{n-1} x^{n-1}$$

sería tal que $\mathbf{q}_h(\mathbf{a}) = \mathbf{0}$ y esto contradice (\mathbf{q} tiene grado menor que n) que los polinomios en el anulador de \mathbf{a} son los múltiplos de \mathbf{p} . Luego, \mathbf{B} es LI.

Por otro lado, para cualquier vector $\mathbf{x} \in \langle \mathbf{a} \rangle_h$ existe un polinomio \mathbf{q} tal que $\mathbf{x} = \mathbf{q}_h(\mathbf{a})$. Efectuando la división con resto obtenemos $\mathbf{q} = \mathbf{cp} + \mathbf{r}$ donde el grado de \mathbf{r} es estrictamente menor que n y por lo tanto $\mathbf{r}_h(\mathbf{a}) \in \langle \mathbf{B} \rangle$.

Tenemos que

$$\mathbf{r}_h(\mathbf{a}) = (\mathbf{q} - \mathbf{cp})_h(\mathbf{a}) = \mathbf{q}_h(\mathbf{a}) - \mathbf{c}_h(p_h(\mathbf{a})) = \mathbf{q}_h(\mathbf{a}) = \mathbf{x}$$

Esto significa que $\langle \mathbf{B} \rangle = \langle \mathbf{a} \rangle_h$ y por lo tanto es una base de $\langle \mathbf{a} \rangle_h$. ■

Una consecuencia obvia de este resultado es la siguiente.

6.24

dim $\langle \mathbf{a} \rangle_h$ es igual al grado del período de \mathbf{a} .

Conjuntos h -independientes

Un conjunto de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ no nulos se le llama **h -independiente** si

$$(p_h(\mathbf{v}_1) + \dots + q_h(\mathbf{v}_n)) = \mathbf{0} \Rightarrow (p_h(\mathbf{v}_1) = \dots = q_h(\mathbf{v}_n) = \mathbf{0})$$

Esto es equivalente a que si una h -combinación de ellos es cero entonces, para todo i el coeficiente de \mathbf{v}_i es un múltiplo del período de \mathbf{v}_i . Al considerar coeficientes de grado cero vemos que los conjuntos h -independientes siempre son linealmente independien-

tes. En particular, el número de elementos en un conjunto \mathbf{h} -independiente no puede sobrepasar la dimensión del espacio.



Es posible imaginar tres definiciones distintas de \mathbf{h} -independencia:

$$(p_h(v_1) + \dots + q_h(v_n) = 0) \Rightarrow (p = \dots = q = 0)$$

$$(p_h(v_1) + \dots + q_h(v_n) = 0) \Rightarrow (p_h = \dots = q_h = 0)$$

$$(p_h(v_1) + \dots + q_h(v_n) = 0) \Rightarrow (p_h(v_1) = \dots = q_h(v_n) = 0)$$

En la primera se pide que los polinomios sean cero. En la segunda que los operadores lineales sean cero. En la tercera que los vectores sean cero. *Las dos primeras son erróneas.*

6.25

Si $V = \{v_1, \dots, v_n\}$ es \mathbf{h} -independiente entonces $\langle V \rangle_h = \langle v_1 \rangle_h \oplus \dots \oplus \langle v_n \rangle_h$.

Prueba. Por inducción en el natural n . Si $n = 1$ el resultado es obvio. Denotemos $V' = \{v_2, \dots, v_n\}$. Tenemos que probar que $\langle V \rangle_h = \langle v_1 \rangle_h + \langle V' \rangle_h$ y que $\langle v_1 \rangle_h \cap \langle V' \rangle_h = 0$.

Si $\alpha \in \langle V \rangle_h$ entonces existen coeficientes polinomiales tales que $\alpha = p_h(v_1) + q_h(v_2) + \dots + r_h(v_n)$. Obviamente, $\alpha' = p_h(v_1) \in \langle v_1 \rangle_h$, $\alpha'' = q_h(v_2) + \dots + r_h(v_n) \in \langle V' \rangle_h$ y $\alpha = \alpha' + \alpha''$. Luego, $\langle V \rangle_h \subseteq \langle v_1 \rangle_h + \langle V' \rangle_h$. La otra inclusión se demuestra igual.

Supongamos que $\alpha \in \langle v_1 \rangle_h \cap \langle V' \rangle_h$. Entonces, existen polinomios tales que

$$p_h(v_1) = \alpha = q_h(v_2) + \dots + r_h(v_n)$$

por lo tanto

$$p_h(v_1) - q_h(v_2) - \dots - r_h(v_n) = 0.$$

Como V es \mathbf{h} -independiente entonces, $\alpha = p_h(v_1) = 0$. Luego, $\langle V \rangle_h = \langle v_1 \rangle_h \oplus \langle V' \rangle_h$ y usando la hipótesis de inducción terminamos la prueba ■

\mathbf{h} -bases

El resultado anterior es bueno para nosotros. Si solo V además de \mathbf{h} -independiente fuera \mathbf{h} -generador, obtendríamos una descomposición de todo el espacio en suma directa de subespacios invariantes \mathbf{h} -generados por un solo vector o sea cíclicos.

Una \mathbf{h} -base es un conjunto de vectores que es \mathbf{h} -independiente y \mathbf{h} -generador. El único problema es que no sabemos si existen o no las \mathbf{h} -bases. Veremos en la siguiente sección que sí existen, sin embargo, el demostrarlo no es tan sencillo como en el caso de las bases ordinarias.

Por lo pronto, nos conformaremos con dos consecuencias obvias de lo ya demostrado:

6.26

Si A es una \mathbf{h} -base entonces, la dimensión del espacio es igual a la suma de los grados de los \mathbf{h} -períodos de los vectores en A .

Prueba. Es consecuencia de que la dimensión de la suma directa es la suma de las dimensiones y de que la dimensión de un subespacio cíclico es igual al grado del período

de su \mathbf{h} -generador (6.23). ■

6.27

Si \mathbf{A} es una \mathbf{h} -base entonces, el polinomio mínimo de \mathbf{h} es igual al mínimo común múltiplo de los \mathbf{h} -períodos de los vectores en \mathbf{A} .

Prueba. Es consecuencia de que el período de la suma directa es el mínimo común múltiplo de los períodos de los sumandos (6.11) y de 6.21. ■

6.5 Descomposición en subespacios cíclicos radicales.

Entre todos los subespacios \mathbf{h} -cíclicos los más grandes son aquellos que están \mathbf{h} -generados por vectores cuyo período tiene grado lo más grande posible (6.23), o sea, aquellos cuyo período es igual al polinomio mínimo. A estos subespacios les llamaremos \mathbf{h} -cíclicos **maximales**. Por 6.19 siempre existe algún subespacio \mathbf{h} -cíclico maximal.

La estrategia para deshacernos del problema de que no cualquier subespacio invariante tiene complementario invariante es limitarnos a los operadores radicales, fijarnos solamente en los subespacios invariantes maximales y construir un complementario que sí es invariante. Para eso necesitaremos usar espacios cocientes.

El espacio cociente por un subespacio invariante

Sea $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ un OL y \mathfrak{F} un subespacio \mathbf{h} -invariante o sea $\mathbf{h}(\mathfrak{F}) \subseteq \mathfrak{F}$. El espacio cociente $\mathfrak{E}/\mathfrak{F}$ es el conjunto de todos los subespacios afines paralelos a \mathfrak{F} . Lo que queremos ver es que si los vectores \mathbf{v} y \mathbf{u} están en un mismo subespacio afín paralelo a \mathfrak{F} entonces $\mathbf{h}(\mathbf{v})$ y $\mathbf{h}(\mathbf{u})$ cumplen exactamente lo mismo.

Efectivamente, si \mathbf{v} y \mathbf{u} están en un mismo subespacio afín paralelo a \mathfrak{F} entonces $\mathbf{v} - \mathbf{u} \in \mathfrak{F}$. Como \mathfrak{F} es \mathbf{h} -invariante entonces $\mathbf{h}(\mathbf{v}) - \mathbf{h}(\mathbf{u}) = \mathbf{h}(\mathbf{v} - \mathbf{u}) \in \mathfrak{F}$ y por lo tanto $\mathbf{h}(\mathbf{v})$ y $\mathbf{h}(\mathbf{u})$ están en un mismo subespacio afín paralelo a \mathfrak{F} . Nótese que aquí el argumento crucial es que \mathfrak{F} es \mathbf{h} -invariante.

Definiremos la función $\tilde{\mathbf{h}}$ en el espacio cociente con la igualdad $\tilde{\mathbf{h}}(\mathbf{v} + \mathfrak{F}) = \mathbf{h}(\mathbf{v}) + \mathfrak{F}$. Por la observación precedente, $\tilde{\mathbf{h}}$ está bien definida, o sea, si $\mathbf{v} + \mathfrak{F} = \mathbf{u} + \mathfrak{F}$ entonces, $\mathbf{h}(\mathbf{v}) + \mathfrak{F} = \mathbf{h}(\mathbf{u}) + \mathfrak{F}$. Observese que $\tilde{\mathbf{h}}$ es un OL en $\mathfrak{E}/\mathfrak{F}$ ya que

$$\begin{aligned}\tilde{\mathbf{h}}(\mathbf{v} + \mathfrak{F} + \mathbf{u} + \mathfrak{F}) &= \tilde{\mathbf{h}}(\mathbf{v} + \mathbf{u} + \mathfrak{F}) = \mathbf{h}(\mathbf{v} + \mathbf{u}) + \mathfrak{F} = \\ &= \mathbf{h}(\mathbf{v}) + \mathfrak{F} + \mathbf{h}(\mathbf{u}) + \mathfrak{F} = \tilde{\mathbf{h}}(\mathbf{v} + \mathfrak{F}) + \tilde{\mathbf{h}}(\mathbf{u} + \mathfrak{F}); \\ \tilde{\mathbf{h}}(\lambda(\mathbf{v} + \mathfrak{F})) &= \tilde{\mathbf{h}}(\lambda\mathbf{v} + \mathfrak{F}) = \mathbf{h}(\lambda\mathbf{v}) + \mathfrak{F} = \lambda\mathbf{h}(\mathbf{v}) + \mathfrak{F} = \lambda(\mathbf{h}(\mathbf{v}) + \mathfrak{F}) = \lambda\tilde{\mathbf{h}}(\mathbf{v} + \mathfrak{F}).\end{aligned}$$

Necesitaremos conocer mejor los OLs del espacio cociente $\text{End}(\mathfrak{E}/\mathfrak{F})$.

6.28

El conjunto $\text{End}^{\tilde{\mathfrak{F}}}(\mathfrak{E})$ de todos los operadores lineales que dejan invariante $\tilde{\mathfrak{F}}$ es una subálgebra de $\text{End}(\mathfrak{E})$.

Prueba. Sean f y g dos operadores en $\text{End}^{\tilde{\mathfrak{F}}}(\mathfrak{E})$. Sea λ un escalar. Para cualquier vector $v \in \mathfrak{F}$ tenemos $f(v) \in \tilde{\mathfrak{F}}$ y $g(v) \in \tilde{\mathfrak{F}}$ y por lo tanto

$$(g + f)(v) = g(v) + f(v) \in \tilde{\mathfrak{F}}; \quad \mathbb{I}(v) \in \tilde{\mathfrak{F}};$$

$$(g \circ f)(v) = g(f(v)) \in \tilde{\mathfrak{F}}; \quad (\lambda f)(v) = \lambda f(v) \in \tilde{\mathfrak{F}}.$$

Luego, $\text{End}^{\tilde{\mathfrak{F}}}(\mathfrak{E})$ es una subálgebra. ■

6.29

La función $\text{End}^{\tilde{\mathfrak{F}}}(\mathfrak{E}) \ni h \mapsto \tilde{h} \in \text{End}(\mathfrak{E}/\tilde{\mathfrak{F}})$ es un morfismo de álgebras.

Prueba. Probaremos que $h \mapsto \tilde{h}$ es morfismo para la composición

$$\begin{aligned}\widetilde{f \circ g}(v + \tilde{\mathfrak{F}}) &= (f \circ g)(v) + \tilde{\mathfrak{F}} = f(g(v)) + \tilde{\mathfrak{F}} = \tilde{f}(g(v) + \tilde{\mathfrak{F}}) = \\ &= \tilde{f}(g(v) + \tilde{\mathfrak{F}}) = \tilde{f}(\tilde{g}(v + \tilde{\mathfrak{F}})) = (\tilde{f} \circ \tilde{g})(v + \tilde{\mathfrak{F}});\end{aligned}$$

y para la suma

$$\begin{aligned}\widetilde{f + g}(v + \tilde{\mathfrak{F}}) &= (f + g)(v) + \tilde{\mathfrak{F}} = f(v) + \tilde{\mathfrak{F}} + g(v) + \tilde{\mathfrak{F}} = \\ &= \tilde{f}(v + \tilde{\mathfrak{F}}) + \tilde{g}(v + \tilde{\mathfrak{F}}) = (\tilde{f} + \tilde{g})(v + \tilde{\mathfrak{F}}); \end{aligned}$$

y para el producto por escalares

$$\begin{aligned}\widetilde{\lambda f}(v + \tilde{\mathfrak{F}}) &= (\lambda f)(v) + \tilde{\mathfrak{F}} = f(\lambda v) + \tilde{\mathfrak{F}} = \tilde{f}(\lambda v + \tilde{\mathfrak{F}}) = \\ &= \tilde{f}(\lambda(v + \tilde{\mathfrak{F}})) = (\lambda \tilde{f})(v + \tilde{\mathfrak{F}}); \end{aligned}$$

y que preserva la identidad $\widetilde{\mathbb{I}}(v + \tilde{\mathfrak{F}}) = \mathbb{I}(v) + \tilde{\mathfrak{F}} = v + \tilde{\mathfrak{F}} = \mathbb{I}(v + \tilde{\mathfrak{F}})$. ■

Ejercicio 118 Muestre que el morfismo $h \mapsto \tilde{h}$ es sobreyectivo y su núcleo está formado por aquellos operadores lineales cuya imagen es $\tilde{\mathfrak{F}}$. [196]

Polinomios y el espacio cociente

Sea $\tilde{\mathfrak{F}} \subseteq \mathfrak{E}$ un subespacio h -invariante. Ya vimos como se define \tilde{h} en $\mathfrak{E}/\tilde{\mathfrak{F}}$ a saber $\tilde{h}(v + \tilde{\mathfrak{F}}) = h(v) + \tilde{\mathfrak{F}}$ y comprobamos que esta definición es correcta debido a la h -invariancia de $\tilde{\mathfrak{F}}$. Para poder definir $\widetilde{p_h}(v + \tilde{\mathfrak{F}}) = p_h(v) + \tilde{\mathfrak{F}}$ necesitamos que $\tilde{\mathfrak{F}}$ sea p_h -invariante. Por suerte, esto se cumple automáticamente.

6.30

Sea p un polinomio. Si $\tilde{\mathfrak{F}}$ es h -invariante entonces también es p_h -invariante.

Prueba. Como el conjunto de los operadores que preservan $\tilde{\mathfrak{F}}$ es una subálgebra del álgebra de todos los operadores entonces todos los h^n están ahí, también todas las

combinaciones lineales de estos o sea, todos los polinomios evaluados en \mathbf{h} . ■

Luego, $\widetilde{p_h}$ es un OL bien definido en el espacio cociente. Por otro lado, si evaluamos el polinomio p en el operador $\tilde{\mathbf{h}}$ obtendremos $p_{\tilde{\mathbf{h}}}$ que es otro OL bien definido en el espacio cociente.

6.31

$$\widetilde{p_h} = p_{\tilde{\mathbf{h}}}$$

Prueba. Sea $p(x) = \sum \alpha_i x^i$. Como la función $\mathbf{h} \mapsto \tilde{\mathbf{h}}$ es un morfismo de álgebras, tenemos

$$\widetilde{p_h} = \sum_{i=0}^n \widetilde{\alpha_i h^i} = \sum_{i=0}^n \widetilde{\alpha_i} \widetilde{h^i} = \sum_{i=0}^n \alpha_i \widetilde{h^i} = \sum_{i=0}^n \alpha_i (\tilde{\mathbf{h}})^i = p_{\tilde{\mathbf{h}}}$$

y esto es lo que se necesitaba probar. ■

Esto es muy cómodo, ya que en el cociente tenemos la fórmula $\widetilde{p_h}(\mathbf{v} + \mathfrak{F}) = \widetilde{p_h}(\mathbf{v} + \mathfrak{F}) = p_h(\mathbf{v}) + \mathfrak{F}$.

El período en el espacio cociente

6.32

Si $\mathbf{b} \in \mathbf{v} + \mathfrak{F}$ entonces $\text{per}_h(\mathbf{b}) \vdash \text{per}_{\tilde{h}}(\mathbf{v} + \mathfrak{F})$.

Prueba. Sea $\mathbf{b} \in \mathbf{v} + \mathfrak{F}$. Si $p = \text{per}_h(\mathbf{b})$ entonces $\widetilde{p_h}(\mathbf{b} + \mathfrak{F}) = p_h(\mathbf{b}) + \mathfrak{F} = \mathbf{0} + \mathfrak{F} = \mathfrak{F}$ y por lo tanto $p \vdash \text{per}_{\tilde{h}}(\mathbf{b} + \mathfrak{F})$. La prueba concluye observando que $\mathbf{v} + \mathfrak{F} = \mathbf{b} + \mathfrak{F}$. ■

Ejercicio 119 Pruebe que $\text{per}_h(\mathbf{b}) \dashv \text{per}_{\tilde{h}}(\mathbf{b} + \mathfrak{F})$ si y solo si $\langle \mathbf{b} \rangle_h \cap \mathfrak{F} = \{0\}$. [196]

El resultado que sigue no es trivial y consiste en que bajo ciertas condiciones es posible encontrar un $\mathbf{b} \in \mathbf{v} + \mathfrak{F}$ tal que $\text{per}_h(\mathbf{b}) \dashv \text{per}_{\tilde{h}}(\mathbf{v} + \mathfrak{F})$ y por lo tanto $\text{per}_h(\mathbf{b}) = \text{per}_{\tilde{h}}(\mathbf{v} + \mathfrak{F})$.

Lema del Período

6.33

Sea $\mathbf{h} \in \text{End}(\mathfrak{E})$ radical y $\mathfrak{F} = \langle \mathbf{a} \rangle_h$ un subespacio \mathbf{h} -cíclico maximal. Para cualquier $\mathbf{v} + \mathfrak{F} \in \mathfrak{E}/\mathfrak{F}$ existe $\mathbf{b} \in \mathbf{v} + \mathfrak{F}$ tal que $\text{per}_h(\mathbf{b}) \dashv \text{per}_{\tilde{h}}(\mathbf{v} + \mathfrak{F})$.

Prueba. Sea p^n (con p irreducible) el polinomio mínimo de \mathbf{h} . Como $\mathfrak{F} = \langle \mathbf{a} \rangle_h$ es maximal, el \mathbf{h} -período de \mathbf{a} es p^n . Los \mathbf{h} -períodos de todos los vectores son potencias de p . Los $\tilde{\mathbf{h}}$ -períodos de todos $\mathbf{v} + \mathfrak{F}$ son potencias de p ya que por 6.32 estos son divisores de los períodos de los vectores.

Sea $\mathbf{v} + \mathfrak{F} \in \mathfrak{E}/\mathfrak{F}$ de $\tilde{\mathbf{h}}$ -período \mathbf{p}^m . Sabemos que $m \leq n$. Veamos que existe un natural k y un vector \mathbf{a}' tales que se cumplen las propiedades del recuadro a la derecha. Efectivamente, sabemos que $p_h^m(\mathbf{v}) + \mathfrak{F} = p_{\tilde{\mathbf{h}}}^m(\mathbf{v} + \mathfrak{F}) = \mathfrak{F}$ o lo que es lo mismo $p_h^m(\mathbf{v}) \in \mathfrak{F} = \langle \mathbf{a} \rangle_h$. Luego, existe un polinomio r tal que $p_h^m(\mathbf{v}) = r_h(\mathbf{a})$. Usando el teorema de descomposición de un polinomio en producto de irreducibles, existen un polinomio q coprimo con p y un natural $k \geq 0$ tales que $r = p^k q$. Por 6.22 el vector $\mathbf{a}' \stackrel{\text{def}}{=} q_h(\mathbf{a})$ es también un \mathbf{h} -generador de \mathfrak{F} y \mathbf{a}' tiene el mismo \mathbf{h} -período que \mathbf{a} o sea, \mathbf{p}^n . Además $p_h^m(\mathbf{v}) = r_h(\mathbf{a}) = p_h^k(q_h(\mathbf{a})) = p_h^k(\mathbf{a}')$ y esto demuestra la igualdad (*).

$$\begin{aligned}\mathfrak{F} &= \langle \mathbf{a}' \rangle_h \\ \text{per}_h(\mathbf{a}') &= \mathbf{p}^n \\ p_h^m(\mathbf{v}) &= p_h^k(\mathbf{a}') \quad (*)\end{aligned}$$

Si ocurriera que $m > k$ entonces, aplicando p_h^{n-m} a la igualdad (*) y observando que \mathbf{p}^n es el polinomio mínimo, obtendríamos que $0 = p_h^n(\mathbf{v}) = p_h^{n-m+k}(\mathbf{a}')$. Como $n - m + k < n$ esto contradeciría que el \mathbf{h} -período de \mathbf{a}' es \mathbf{p}^n .

Luego, $m \leq k$ y esto nos sirve para definir el siguiente vector

$$\mathbf{b} \stackrel{\text{def}}{=} \mathbf{v} - p_h^{k-m}(\mathbf{a}').$$

Observese que $\mathbf{b} \in \mathbf{v} + \mathfrak{F}$ ya que $p_h^{k-m}(\mathbf{a}') \in \langle \mathbf{a}' \rangle_h = \mathfrak{F}$. Aplicando p_h^m a la definición de \mathbf{b} , obtenemos que

$$p_h^m(\mathbf{b}) = p_h^m(\mathbf{v}) - p_h^k(\mathbf{a}') \stackrel{(*)}{=} 0$$

por lo que $\text{per}_h(\mathbf{b}) \dashv \mathbf{p}^m$. ■

Existencia de \mathbf{h} -bases

Lema del Cociente

 Sea $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ un operador radical y $\mathfrak{F} = \langle \mathbf{a} \rangle_h$ un subespacio \mathbf{h} -cíclico maximal. Si $\{\mathbf{v}_1 + \mathfrak{F}, \dots, \mathbf{v}_m + \mathfrak{F}\}$ es una $\tilde{\mathbf{h}}$ -base del espacio cociente $\mathfrak{E}/\mathfrak{F}$ entonces, existen vectores $\mathbf{b}_1 \in \mathbf{v}_1 + \mathfrak{F}, \dots, \mathbf{b}_m \in \mathbf{v}_m + \mathfrak{F}$ tales que $\{\mathbf{a}, \mathbf{b}_1, \dots, \mathbf{b}_m\}$ es una \mathbf{h} -base de \mathfrak{E} .

Prueba. Sea \mathbf{a} un \mathbf{h} -generador de \mathfrak{F} . Aplicando el Lema del Período (6.33) a $\mathbf{v}_1 + \mathfrak{F}, \dots, \mathbf{v}_m + \mathfrak{F}$, obtenemos $\mathbf{b}_1 \in \mathbf{v}_1 + \mathfrak{F}, \dots, \mathbf{b}_m \in \mathbf{v}_m + \mathfrak{F}$ tales que el \mathbf{h} -período de \mathbf{b}_i es igual al $\tilde{\mathbf{h}}$ -período de $\mathbf{v}_i + \mathfrak{F}$. Denotemos $B \stackrel{\text{def}}{=} \{\mathbf{a}, \mathbf{b}_1, \dots, \mathbf{b}_m\}$. Observese que

$$\overline{B \setminus \mathbf{a}} \stackrel{\text{def}}{=} \{\mathbf{b}_1 + \mathfrak{F}, \dots, \mathbf{b}_m + \mathfrak{F}\} = \{\mathbf{v}_1 + \mathfrak{F}, \dots, \mathbf{v}_m + \mathfrak{F}\}$$

es una $\tilde{\mathbf{h}}$ -base de $\mathfrak{E}/\mathfrak{F}$. Probemos que B es \mathbf{h} -generador de \mathfrak{E} . Si $\mathbf{x} \in \mathfrak{E}$ entonces $\mathbf{x} + \mathfrak{F} \in \langle \overline{B \setminus \mathbf{a}} \rangle_{\tilde{\mathbf{h}}}$, o sea, existen coeficientes polinomiales tales que

$$\mathbf{x} + \mathfrak{F} = q_{\tilde{\mathbf{h}}}(\mathbf{b}_1 + \mathfrak{F}) + \dots + r_{\tilde{\mathbf{h}}}(\mathbf{b}_m + \mathfrak{F}) = q_h(\mathbf{b}_1) + \dots + r_h(\mathbf{b}_m) + \mathfrak{F}$$

o lo que es lo mismo $\mathbf{x} - q_h(\mathbf{b}_1) - \dots - r_h(\mathbf{b}_m) \in \mathfrak{F} = \langle \mathbf{a} \rangle_h$. Luego, existe otro polinomio s tal que $\mathbf{x} - q_h(\mathbf{b}_1) - \dots - r_h(\mathbf{b}_m) = s_h(\mathbf{a})$ y por lo tanto

$$\mathbf{x} = s_h(\mathbf{a}) + q_h(\mathbf{b}_1) + \dots + r_h(\mathbf{b}_m) \in \langle B \rangle_h.$$

Para probar que \mathbf{B} es \mathbf{h} -independiente supongamos que existen polinomios tales que

$$\mathbf{0} = s_h(\mathbf{a}) + q_h(\mathbf{b}_1) + \cdots + r_h(\mathbf{b}_m). \quad (*)$$

Pasando al cociente (sumando \mathfrak{F}) obtenemos

$$\begin{aligned}\mathbf{0} + \mathfrak{F} &= s_h(\mathbf{a}) + \mathfrak{F} + q_h(\mathbf{b}_1) + \mathfrak{F} + \cdots + r_h(\mathbf{b}_m) + \mathfrak{F} = \\ &= q_{\tilde{h}}(\mathbf{b}_1 + \mathfrak{F}) + \cdots + r_{\tilde{h}}(\mathbf{b}_m + \mathfrak{F})\end{aligned}$$

y como $\overline{\mathbf{B} \setminus \mathbf{a}}$ es independiente obtenemos que

$$q_{\tilde{h}}(\mathbf{b}_1 + \mathfrak{F}) = \cdots = r_{\tilde{h}}(\mathbf{b}_m + \mathfrak{F}) = \mathbf{0} + \mathfrak{F}.$$

Como el \mathbf{h} -período de \mathbf{b}_i es igual al $\tilde{\mathbf{h}}$ -período de $\mathbf{b}_i + \mathfrak{F}$ concluimos que

$$q_h(\mathbf{b}_1) = \cdots = r_h(\mathbf{b}_m) = \mathbf{0}.$$

Substituyendo esto en la igualdad $(*)$ obtenemos $s_h(\mathbf{a}) = \mathbf{0}$. ■

Teorema de Existencia de \mathbf{h} -bases

6.35

Cualquier operador lineal radical \mathbf{h} tiene una \mathbf{h} -base.

Prueba. Sea $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ un operador radical y $\mathfrak{F} = \langle \mathbf{a} \rangle_{\mathbf{h}}$ subespacio \mathbf{h} -cíclico maximal. Hagamos inducción en $\dim \mathbf{h}$. Si $\dim \mathbf{h} = 1$ entonces, $\mathfrak{F} = \mathfrak{E}$ y por lo tanto $\{\mathbf{a}\}$ es una \mathbf{h} -base.

Supongamos $\dim \mathbf{h} > 1$. Si $\mathfrak{F} = \mathfrak{E}$ entonces, otra vez $\{\mathbf{a}\}$ es una \mathbf{h} -base. Si no, entonces $\mathfrak{E}/\mathfrak{F}$ es no trivial y $\dim \mathfrak{E}/\mathfrak{F} < \dim \mathbf{h}$. Por hipótesis de inducción $\mathfrak{E}/\mathfrak{F}$ tiene una $\tilde{\mathbf{h}}$ -base y por el Lema del Cociente (6.34) existe una \mathbf{h} -base de \mathfrak{E} . ■

Ejercicio 120 Demuestre la siguiente afirmación. Sean $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ un OL y $\mathfrak{E} = \mathfrak{F} \oplus \mathfrak{G}$ una descomposición de \mathfrak{E} en subespacios \mathbf{h} -invariantes. Si \mathbf{A} es una \mathbf{h} -base de \mathfrak{F} y \mathbf{B} es una \mathbf{h} -base de \mathfrak{G} entonces $\mathbf{A} \cup \mathbf{B}$ es una \mathbf{h} -base de \mathfrak{E} . [196]

Ejercicio 121 Use el ejercicio anterior, el Teorema de Descomposición en Componentes Radicales (6.18) y el Teorema de Existencia de \mathbf{h} -bases (6.35) para probar todo operador lineal \mathbf{h} tiene una \mathbf{h} -base.

Teorema de Descomposición en Componentes Radicales Cílicas

6.36

Todo operador lineal es suma directa de componentes radicales cílicas.

Prueba. Sea $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ un OL. Si \mathbf{h} es radical entonces \mathfrak{E} tiene una \mathbf{h} -base $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$. Por 6.25 $\mathfrak{E} = \langle \mathbf{a}_1 \rangle_{\mathbf{h}} \oplus \cdots \oplus \langle \mathbf{a}_n \rangle_{\mathbf{h}}$. Por 6.20 todos los $\langle \mathbf{a}_i \rangle_{\mathbf{h}}$ son subespacios invariantes. Denotando por f_i la restricción de \mathbf{h} a $\langle \mathbf{a}_i \rangle_{\mathbf{h}}$ obtenemos $\mathbf{h} = f_1 \oplus \cdots \oplus f_n$. Si \mathbf{h} no es radical entonces usando el Teorema de Descomposición en Componentes Radicales (6.18) lo descomponemos en componentes radicales y posteriormente cada componente radical la descomponemos en componentes cílicas. ■

Unicidad de la descomposición

A diferencia de la descomposición en componentes radicales una descomposición en componentes radicales cíclicas no es única. Esto sucede porque \mathbf{h} -bases pueden haber muchas. Por ejemplo para la identidad en \mathbb{R}^2 cualesquiera dos rectas diferentes por el origen forman una descomposición en subespacios invariantes radicales cíclicos. Sin embargo, lo que sí es único es el “tipo” de la descomposición.

Si $\mathbf{h} = \mathbf{f}_1 \oplus \cdots \oplus \mathbf{f}_n$ es una descomposición entonces, a la sucesión $\mathbf{p}, \mathbf{q}, \dots, \mathbf{r}$ de los polinomios mínimos de las componentes se le llama el **tipo de la descomposición**. Dos tipos se consideran iguales si uno se puede obtener del otro reordenando la sucesión.

Teorema de Unicidad del Tipo

6.37

Dos descomposiciones cualesquiera en componentes radicales cíclicas tienen el mismo tipo.

Antes de demostrar nuestro teorema de unicidad necesitamos demostrar tres resultados auxiliares sencillos. El lector ávido puede saltárselos y regresar a ellos en la medida de sus necesidades. En los siguientes tres resultados $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ es un operador radical con polinomio mínimo \mathbf{p}^n (no se necesita que sea cíclico).

6.38

$$\text{per}_{\mathbf{h}} \mathbf{x} = \mathbf{p} \times \text{per}_{\mathbf{h}} \mathbf{p}_{\mathbf{h}}(\mathbf{x}).$$

Prueba. Los períodos de los vectores son divisores del polinomio mínimo, en particular, el período de \mathbf{x} es \mathbf{p}^k para cierto $k \leq n$. Tenemos $\mathbf{p}_{\mathbf{h}}^{k-1}(\mathbf{p}_{\mathbf{h}}(\mathbf{x})) = \mathbf{p}_{\mathbf{h}}^k(\mathbf{x}) = \mathbf{0}$. Por otro lado $\mathbf{p}_{\mathbf{h}}^{k-2}(\mathbf{p}_{\mathbf{h}}(\mathbf{x})) = \mathbf{p}_{\mathbf{h}}^{k-1}(\mathbf{x}) \neq \mathbf{0}$. Luego, \mathbf{p}^{k-1} es el polinomio mónico de grado más pequeño que anula a $\mathbf{p}_{\mathbf{h}}(\mathbf{x})$. Un lector cuidadoso debería analizar los casos $k \in \{0, 1\}$ para los cuales esta prueba es formalmente incorrecta. ■

6.39

$\mathbf{p}_{\mathbf{h}}(\mathfrak{E})$ es un subespacio invariante y $\dim \mathbf{p}_{\mathbf{h}}(\mathfrak{E}) < \dim \mathfrak{E}$.

Prueba. El subespacio $\mathfrak{F} = \mathbf{p}_{\mathbf{h}}(\mathfrak{E}) = \text{Im } \mathbf{p}_{\mathbf{h}}$ es invariante ya que $\mathbf{h}(\mathbf{p}_{\mathbf{h}}(\mathbf{x})) = \mathbf{p}_{\mathbf{h}}(\mathbf{h}(\mathbf{x}))$. Sea \mathbf{x} un vector no nulo de período \mathbf{p}^k (claramente estamos asumiendo que $\mathfrak{E} \neq \{\mathbf{0}\}$). Si $k = 1$ entonces, $\mathbf{x} \in \ker \mathbf{p}_{\mathbf{h}}$. Si $k > 1$ entonces, el vector no nulo $\mathbf{y} = \mathbf{p}_{\mathbf{h}}^{k-1}(\mathbf{x})$ es tal que $\mathbf{p}_{\mathbf{h}}(\mathbf{y}) = \mathbf{0}$. De aquí, el OL $\mathbf{p}_{\mathbf{h}}$ tiene núcleo no trivial y por lo tanto $\mathfrak{F} \neq \mathfrak{E}$. Luego, $\dim \mathfrak{F} < \dim \mathfrak{E}$. ■

6.40

Si X es una \mathbf{h} -base de \mathfrak{E} entonces $\mathbf{p}_{\mathbf{h}}(X) \setminus \{\mathbf{0}\}$ es una \mathbf{h} -base de $\mathbf{p}_{\mathbf{h}}(\mathfrak{E})$.

Prueba. Para todo $\mathbf{x} \in \mathfrak{E}$ denotemos $\bar{\mathbf{x}} \stackrel{\text{def}}{=} \mathbf{p}_{\mathbf{h}}(\mathbf{x})$. Sea $X = \{\mathbf{u}, \dots, \mathbf{v}\}$ una \mathbf{h} -base de

E. Comprobaremos que $\bar{X} \stackrel{\text{def}}{=} \{\bar{x} \mid x \in X \text{ y } \bar{x} \neq 0\}$ es una \mathbf{h} -base de \mathfrak{F} . Efectivamente, si $\bar{y} \in \mathfrak{F} \stackrel{\text{def}}{=} p_h(\mathfrak{E})$ entonces, como X es \mathbf{h} -generador, existen polinomios tales que:

$$\bar{y} = \overline{q_h(u) + \cdots + r_h(v)} = q_h(\bar{u}) + \cdots + r_h(\bar{v}).$$

Es claro que en la suma de la derecha podemos descartar aquellos sumandos para los cuales $\bar{x} = 0$. Luego, $\mathfrak{F} \subseteq \langle \bar{X} \rangle_{\mathbf{h}}$ o sea, \bar{X} es un \mathbf{h} -generador de \mathfrak{F} . Supongamos que para ciertos polinomios q, \dots, r se tiene que

$$0 = q_h(\bar{u}) + \cdots + r_h(\bar{v}) = (qp)_h(u) + \cdots + (rp)_h(v).$$

Como X es \mathbf{h} -independiente, entonces

$$0 = (qp)_h(u) = q_h(\bar{u}) = \cdots = (rp)_h(v) = r_h(\bar{v})$$

y por lo tanto \bar{X} es \mathbf{h} -independiente. Luego, \bar{X} es una \mathbf{h} -base de \mathfrak{F} . ■

Ya estamos listos para probar el teorema de unicidad.

Prueba. (Del Teorema de Unicidad del Tipo) De la definición del tipo y de la unicidad de la descomposición en componentes radicales queda claro que es suficiente demostrar el teorema para OLs radicales $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ cuyo polinomio mínimo es una potencia de un polinomio irreducible que denotaremos por p .

Para cualquier vector x denotaremos $\bar{x} \stackrel{\text{def}}{=} p_h(x)$. Para cualquier conjunto de vectores X denotaremos $\bar{X} \stackrel{\text{def}}{=} \{\bar{x} \mid x \in X \text{ y } \bar{x} \neq 0\}$.

Sean $\mathfrak{E} = \langle \mathbf{a}_1 \rangle_{\mathbf{h}} \oplus \cdots \oplus \langle \mathbf{a}_n \rangle_{\mathbf{h}} = \langle \mathbf{b}_1 \rangle_{\mathbf{h}} \oplus \cdots \oplus \langle \mathbf{b}_m \rangle_{\mathbf{h}}$ dos descomposiciones en subespacios cíclicos. Como \mathbf{h} es radical los períodos de todos sus vectores son potencias del polinomio irreducible p . Sean p^{n_i} los períodos de los \mathbf{a}_i y p^{m_i} los períodos de los \mathbf{b}_i . Los tipos de estas dos descomposiciones son p^{n_1}, \dots, p^{n_n} y p^{m_1}, \dots, p^{m_m} respectivamente. Reordenando los sumandos podemos suponer que $n_1 \geq \cdots \geq n_n$ y que $m_1 \geq \cdots \geq m_m$. También podemos suponer que $n \geq m$. Tenemos que probar que $n = m$ y que los n_i son iguales a los m_i .

Hagamos la prueba del teorema por inducción en $\dim \mathbf{h}$. Si $\dim \mathbf{h} = 1$ entonces, necesariamente $n = m = 1$ y $n_1 = m_1$.

Los conjuntos $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ y $B = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ son dos \mathbf{h} -bases de \mathfrak{E} . Por 6.40 tenemos que \bar{A} y \bar{B} son bases de $\mathfrak{F} = p_h(\mathfrak{E})$. Sea $k \in \{0, \dots, n\}$ el mayor índice tal que $n_k > 1$. Sea $\ell \in \{0, \dots, m\}$ el mayor índice tal que $m_\ell > 1$.

Si $i > k$ entonces, $\text{per}_{\mathbf{h}}(\mathbf{a}_i) = p$ y por lo tanto $\bar{a}_i = 0$. Luego, $\bar{A} = \{\bar{a}_1, \dots, \bar{a}_k\}$. Un argumento análogo nos dice que $\bar{B} = \{\bar{b}_1, \dots, \bar{b}_\ell\}$.

Luego, $\mathfrak{F} = \langle \bar{a}_1 \rangle_{\mathbf{h}} \oplus \cdots \oplus \langle \bar{a}_k \rangle_{\mathbf{h}} = \langle \bar{b}_1 \rangle_{\mathbf{h}} \oplus \cdots \oplus \langle \bar{b}_\ell \rangle_{\mathbf{h}}$ son dos descomposiciones de \mathfrak{F} en subespacios cílico-radicales. Por 6.38 los tipos de estas descomposiciones son $p^{n_1-1}, \dots, p^{n_k-1}$ y $p^{m_1-1}, \dots, p^{m_\ell-1}$. Como por 6.39 $\dim \mathfrak{F} < \dim \mathfrak{E}$, podemos aplicar hipótesis de inducción y así obtenemos que

$$k = \ell, n_1 - 1 = m_1 - 1, \dots, n_k - 1 = m_k - 1.$$

O sea, todos los n_i son iguales a los m_i desde $i = 1$ hasta $k = \ell$.

Sea ∇ el grado de p . De 6.26 obtenemos que

$$(n_1 + \cdots + n_n) \nabla = \dim \mathfrak{E} = (m_1 + \cdots + m_m) \nabla$$

y por lo tanto

$$n_{k+1} + \cdots + n_n = m_{k+1} + \cdots + m_m.$$

Todos los n_i y los m_i para $i > k$ son iguales a 1 y por lo tanto $n = m$. Con esto, para todo i el natural n_i es igual a m_i . ■

Ejercicio 122 Sean $p = \text{per}_h(\mathbf{a})$ y $q = \text{per}_h(\mathbf{b})$ los períodos de dos vectores. Demuestre que si p y q son coprimos entonces, $\langle \mathbf{a} + \mathbf{b} \rangle_h = \langle \mathbf{a} \rangle_h \oplus \langle \mathbf{b} \rangle_h$. [196]

Ejercicio 123 Use el ejercicio anterior y el Teorema de Descomposición en Componentes Radicales Cílicas (6.36) para probar que cualquier OL tiene una descomposición en OL cílicos $f_1 \oplus \dots \oplus f_t$ en la cual el polinomio mínimo de cada f_i divide al polinomio mínimo del siguiente. [197]

Ejercicio 124 Usando el Teorema de Unicidad del Tipo (6.37) demuestre que el tipo de las descomposiciones introducidas en el ejercicio anterior es único. [197]

Estructura de los operadores cílico-radicales

Ahora queremos conocer *todos* los subespacios invariantes de los operadores cílico radicales para poder demostrar que estos son irreducibles. En los tres siguientes resultados $h : \mathfrak{E} \rightarrow \mathfrak{E}$ es un operador cílico-radical con polinomio mínimo p^n .

6.41

Si x es un vector de período p^k con $k < n$ entonces, existe un vector y tal que $x = p_h(y)$ y $\text{per}_h(y) = p^{k+1}$.

Prueba. Sea \mathbf{a} un vector h -generador. Existe un polinomio q tal que $x = q_h(\mathbf{a})$. Si q fuera coprimo con p entonces el período de x fuera el mismo que el de \mathbf{a} y eso no es cierto por hipótesis. Luego $q = pr$ y si ponemos $y = r_h(\mathbf{a})$ obtenemos que $x = p_h(y)$. Por 6.38 $\text{per}_h(y) = p^{k+1}$. ■

6.42

Si $\text{per}_h(x) = p^k$ entonces $\langle x \rangle_h = \ker p_h^k$.

Prueba. Sea x de período p^k . Aplicando repetidamente 6.41 existe un vector \mathbf{a} tal que $x = p_h^{n-k}(\mathbf{a})$ y $\text{per}_h(\mathbf{a}) = p^n$ y por lo tanto \mathbf{a} es un h -generador de todo el espacio.

Como $x \in \ker p_h^k$ por monotonía de la h -cerradura tenemos que $\langle x \rangle_h \subseteq \ker p_h^k$. Recíprocamente, sea $\mathbf{b} \in \ker p_h^k$ entonces, $\text{per}_h \mathbf{b} \dashv p^k$ y por 6.41 existe \mathbf{y} tal que $\mathbf{b} = p_h^{n-k}(\mathbf{y})$. Sea q tal que $\mathbf{y} = q_h(\mathbf{a})$. Entonces $\mathbf{b} = p_h^{n-k}(q_h(\mathbf{a})) = q_h(x)$. Luego, $\ker p_h^k \subseteq \langle x \rangle_h$. ■

6.43

Los subespacios $\ker p_h^k$, $k \in \{0, 1, \dots, n\}$ son los únicos subespacios invariantes de h .

Prueba. Sea \mathfrak{F} cualquier subespacio \mathbf{h} -invariante de \mathfrak{E} . Sea $k \leq n$ tal que $\text{per}_{\mathbf{h}} \mathfrak{F} = p^k$. De aquí, $\mathfrak{F} \subseteq \ker \mathbf{p}_{\mathbf{h}}^k$. Por 6.19 existe un vector $x \in \mathfrak{F}$ de período p^k . Por monotonía de la \mathbf{h} -cerradura $\langle x \rangle_{\mathbf{h}} \subseteq \mathfrak{F}$. Por 6.42 $\langle x \rangle_{\mathbf{h}} = \ker \mathbf{p}_{\mathbf{h}}^k$. ■

Teorema de Caracterización de los OLs irreducibles

6.44

Un operador lineal es irreducible si y solo si es cíclico y radical.

Prueba. Sea \mathbf{h} un OL. Si \mathbf{h} no es cíclico radical entonces por el Teorema de Descomposición en Componentes Radicales Cíclicas (6.36) este se descompone no trivialmente. Si \mathbf{h} es cíclico radical entonces, por 6.43 sus subespacios invariantes son del tipo $\ker \mathbf{p}_{\mathbf{h}}^k$. Por monotonía de los núcleos (6.14), dados dos de estos, siempre hay uno incluido dentro del otro. Luego, \mathbf{h} no tiene una pareja de subespacios invariantes complementarios no triviales. ■

6.6 Polinomio característico

Rectas invariantes

Los subespacios invariantes no triviales más sencillos que nos podemos imaginar son los de dimensión 1 o sea, las rectas por el origen. Si $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ es un OL y \mathfrak{F} es invariante de dimensión 1 entonces, la restricción de \mathbf{h} a \mathfrak{F} es una homotecia. Sea $\{\mathbf{a}\}$ una base de \mathfrak{F} . Entonces, existe un escalar λ (la razón de la homotecia) tal que $\mathbf{h}(\mathbf{a}) = \lambda \mathbf{a}$.

Recíprocamente, supongamos que existen un escalar λ y un vector *no nulo* \mathbf{a} tales que $\mathbf{h}(\mathbf{a}) = \lambda \mathbf{a}$ entonces, al vector \mathbf{a} se le llama **vector propio** del operador \mathbf{h} y al escalar λ se le llama **valor propio** de \mathbf{h} . También decimos que λ es el valor propio **correspondiente** al vector propio \mathbf{a} . Observese que el valor propio correspondiente a un vector es único pues de $\lambda \mathbf{a} = \lambda' \mathbf{a}$ obtenemos $(\lambda - \lambda') \mathbf{a} = 0$ y como $\mathbf{a} \neq 0$ esto implica que $\lambda = \lambda'$.

6.45

La recta $\langle \mathbf{a} \rangle$ es \mathbf{h} -invariante si y solo si \mathbf{a} es un vector propio de \mathbf{h} .

Prueba. Ya vimos la parte “solo si”. Supongamos que \mathbf{a} es un vector propio. Sea λ su correspondiente valor propio. Por linearidad de \mathbf{h} , para cualquier escalar α se cumple que $\mathbf{h}(\alpha \mathbf{a}) = \alpha \mathbf{h}(\mathbf{a}) = \alpha(\lambda \mathbf{a}) = (\alpha \lambda) \mathbf{a}$ y esto significa que $\langle \mathbf{a} \rangle$ es \mathbf{h} -invariante. ■

El polinomio característico de un operador lineal

Recordemos que el determinante de un OL está bien definido como el determinante de su matriz en una base arbitraria.

El hecho de que $\det(\mathbf{h}(\mathbf{a})) = \lambda\mathbf{a}$ lo podemos expresar como que el operador $\lambda\mathbb{I} - \mathbf{h}$ evaluado en el vector \mathbf{a} es cero. O sea el vector está en el núcleo de $\lambda\mathbb{I} - \mathbf{h}$. Esto quiere decir que $\ker(\lambda\mathbb{I} - \mathbf{h})$ es el conjunto de todos los vectores propios correspondientes a λ . Sabemos que $\ker(\lambda\mathbb{I} - \mathbf{h}) \neq \{\mathbf{0}\}$ si y solo si $\det(\lambda\mathbb{I} - \mathbf{h}) = 0$ (véase 4.26, 4.27, 3.25 y 3.28). Luego, λ es un valor propio de \mathbf{h} si y solo si $\det(\lambda\mathbb{I} - \mathbf{h}) = 0$.

Así, para calcular los valores propios de \mathbf{h} podríamos probar todos los elementos λ del campo y comprobar si $\det(\lambda\mathbb{I} - \mathbf{h}) = 0$. Esta estrategia es imposible de realizar si el campo \mathbb{K} es infinito. Por esto es mejor considerar la función $\mathbb{K} \ni x \mapsto \det(x\mathbb{I} - \mathbf{h}) \in \mathbb{K}$ y encontrar sus raíces.

6.46 $\det(x\mathbb{I} - \mathbf{h})$ es un polinomio mónico de grado $\dim \mathbf{h}$ en la variable x .

Prueba. Sabemos que el determinante de un OL no depende de la base en el cual se calcule. Tomemos una base cualquiera \mathbf{N} del espacio vectorial. Tenemos $|\mathbf{N}| = \dim \mathbf{h}$. Sea $\alpha_{\mathbf{NN}}$ la matriz de \mathbf{h} en la base \mathbf{N} . La matriz de $x\mathbb{I} - \mathbf{h}$ en la base \mathbf{N} es $\beta_{\mathbf{NN}} = x\delta_{\mathbf{NN}} - \alpha_{\mathbf{NN}}$ donde $\delta_{\mathbf{NN}}$ es el delta de Kronecker (la matriz identidad). Las entradas de $\beta_{\mathbf{NN}}$ son $\beta_{ii} = x - \alpha_{ii}$ y $\beta_{ij} = -\alpha_{ij}$ para cualquier $j \neq i$.

El determinante lo calculamos por la definición

$$\det \beta_{\mathbf{NN}} = \sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \beta_{i\sigma_i}.$$

Como los β_{ij} son polinomios, el producto y la suma de polinomios son polinomios, tenemos que $\det \beta_{\mathbf{NN}}$ es un polinomio. Observemos además que $\beta_{\mathbf{NN}}$ solo contiene la variable x en la diagonal. Por esto, la potencia más grande de x en $\det \beta_{\mathbf{NN}}$ se obtiene cuando la permutación σ es la identidad, o sea en el producto

$$\prod_{i \in N} \beta_{ii} = \prod_{i \in N} (x - \alpha_{ii}).$$

Por la ley distributiva vemos que es de grado $|\mathbf{N}|$ y tiene coeficiente principal 1. ■

Al polinomio $\det(x\mathbb{I} - \mathbf{h})$ se le llama **polinomio característico** de \mathbf{h} . De esta manera, los valores propios de \mathbf{h} son las raíces del polinomio característico. Si $\alpha_{\mathbf{NN}}$ es la matriz de \mathbf{h} en la base \mathbf{N} entonces, los vectores propios correspondientes a un valor propio λ se pueden hallar resolviendo el sistema de ecuaciones lineales $(\lambda\delta_{\mathbf{NN}} - \alpha_{\mathbf{NN}})\mathbf{a}_{\mathbf{N}} = \mathbf{0}_{\mathbf{N}}$. Este es un sistema de ecuaciones lineales y su conjunto de soluciones es $\ker(\lambda\mathbb{I} - \mathbf{h})$. El conjunto $\ker(\lambda\mathbb{I} - \mathbf{h})$ es un subespacio invariante (porque es el nucleo de un polinomio evaluado en \mathbf{h}) que se conoce como el **subespacio propio** correspondiente al valor propio λ . La restricción de \mathbf{h} a un subespacio propio es una homotecia. De hecho, los subespacios propios son los subespacios invariantes más grandes en los que \mathbf{h} es una homotecia.

Ejercicio 125 A la suma de las entradas en la diagonal de una matriz se le llama **traza** de la matriz. Demuestre que la traza es una propiedad de los operadores lineales, o sea que la traza no cambia al hacer un cambio de base. [197]

El polinomio característico y el polinomio mínimo

Para ver como se relacionan el polinomio característico y el polinomio mínimo daremos la siguiente definición. Sea $p = x^n + \alpha_{n-1}x^{n-1} + \dots + \alpha_1x + \alpha_0$ un polinomio mónico. A la matriz cuadrada de orden n que se muestra en el recuadro a la derecha se le llama **matriz acompañante del polinomio p** .

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & -\alpha_{n-2} \\ 0 & 0 & \dots & 1 & -\alpha_{n-1} \end{pmatrix}$$

6.47

El polinomio característico de la matriz acompañante del polinomio p es igual a p .

Prueba. Por definición el polinomio característico de la matriz acompañante del polinomio $p = x^n + \alpha_{n-1}x^{n-1} + \dots + \alpha_1x + \alpha_0$ es el determinante de la matriz que se muestra a la derecha. Calculémoslo. Triangularizando con el método de eliminación de Gauss obtenemos que este determinante es igual al de la siguiente matriz:

$$\begin{pmatrix} x & 0 & \dots & 0 & \alpha_0 \\ 0 & x & \dots & 0 & \alpha_1 + \alpha_0x^{-1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x & \alpha_{n-2} + \alpha_{n-3}x^{-1} + \dots + \alpha_1x^{-n+3} + \alpha_0x^{-n+2} \\ 0 & 0 & \dots & 0 & x + \alpha_{n-1} + \alpha_{n-2}x^{-1} + \dots + \alpha_1x^{-n+2} + \alpha_0x^{-n+1} \end{pmatrix}.$$

$$\begin{pmatrix} x & 0 & \dots & 0 & \alpha_0 \\ -1 & x & \dots & 0 & \alpha_1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & x & \alpha_{n-2} \\ 0 & 0 & \dots & -1 & x + \alpha_{n-1} \end{pmatrix}$$

Multiplicando las entradas en la diagonal obtenemos p . El lector debe observar el uso de potencias negativas de la variable. Esto quiere decir que el cálculo lo hicimos en el campo de funciones racionales $\mathbb{K}(x)$. ■

6.48

Sea h un OL cíclico con polinomio mínimo p de grado n . La matriz de h en la base $\{\mathbf{a}, h(\mathbf{a}), \dots, h^{n-1}(\mathbf{a})\}$ es la matriz acompañante de p .

Prueba. Denotemos $B = \{\mathbf{v}_0, \dots, \mathbf{v}_{n-1}\} \stackrel{\text{def}}{=} \{\mathbf{a}, h(\mathbf{a}), \dots, h^{n-1}(\mathbf{a})\}$. Sabemos por 6.23 que B es una base. Veamos como transforma h a los vectores de la base B . Tenemos que para $k \in \{0, \dots, n-2\}$ se cumple que $h(\mathbf{v}_k) = \mathbf{v}_{k+1}$ lo que justifica las $n-1$ primeras columnas de la matriz acompañante. Por otro lado si $p = x^n + \alpha_{n-1}x^{n-1} +$

$\cdots + \alpha_1x + \alpha_0$ es el polinomio mínimo de \mathbf{h} . entonces

$$\mathbf{h}(\mathbf{v}_{n-1}) = \mathbf{h}^n(\mathbf{a}) = p_{\mathbf{h}}(\mathbf{a}) - \sum_{i=0}^{n-1} \alpha_i h^i(\mathbf{a}) = 0 - \sum_{i=0}^{n-1} \alpha_i v_i$$

y esto justifica la última columna de la matriz acompañante. ■

Una consecuencia inmediata de 6.48 y 6.47 es el siguiente:

6.49

Si \mathbf{h} es cíclico entonces, los polinomios mínimo y característico de \mathbf{h} coinciden.

6.50

Si $\mathbf{h} = \mathbf{f} \oplus \mathbf{g}$ entonces, el polinomio característico de \mathbf{h} es igual al producto de los polinomios característicos de \mathbf{f} y \mathbf{g} .

$$\left(\begin{array}{|c|c|} \hline M & 0 \\ \hline 0 & M' \\ \hline \end{array} \right)$$

Prueba. Sean \mathfrak{F} y \mathfrak{G} los subespacios invariantes en los cuales estan definidas \mathbf{f} y \mathbf{g} . Por 6.3, si \mathbf{A} es una base de \mathfrak{F} y \mathbf{B} es una base de \mathfrak{G} entonces, en la base de todo el espacio $\mathbf{A} \cup \mathbf{B}$, la matriz de $x\mathbb{I} - \mathbf{h}$ es diagonal por bloques. El determinante de cualquier matriz diagonal por bloques es igual al producto de los determinantes de los bloques diagonales. El determinante del bloque superior izquierdo es el de $x\mathbb{I} - \mathbf{f}$ o sea, el polinomio característico de \mathbf{f} . El determinante del bloque inferior derecho es el de $x\mathbb{I} - \mathbf{g}$ o sea, el polinomio característico de \mathbf{g} . ■

Teorema de Hamilton-Caley-Frobenius

6.51

El polinomio característico es un múltiplo del polinomio mínimo. Estos dos polinomios tienen los mismos factores irreducibles.

Prueba. Por el Teorema de Descomposición en Componentes Radicales Cíclicas (6.36) todo OL tiene una descomposición en componentes cíclicas. Por 6.49 los polinomios mínimos y característicos coinciden en cada una de las componentes. Por 6.50 el polinomio característico es el producto de los de las componentes. Por 6.11 el polinomio mínimo es el mínimo común múltiplo de los de las componentes. El mínimo común múltiplo es siempre un divisor del producto. El mínimo común múltiplo siempre tiene los mismos factores irreducibles que el producto. ■

La primera afirmación se conoce en la literatura como Teorema de Hamilton-Caley y es equivalente por definición de polinomio mínimo a que el polinomio característico de \mathbf{h} , evaluado en \mathbf{h} es el operador nulo. La segunda se conoce como Teorema de Frobenius. Una curiosidad histórica es que fué Frobenius el que dió la primera prueba completa del Teorema de Hamilton-Caley.

Es posible dar muchas diferentes demostraciones del Teorema de Hamilton-Caley que no pasan por la descomposición de un operador en componentes cíclicas (que es

el resultado “duro” de esta teoría). La idea de una de ellas es tomarse un vector \mathbf{a} , observar que en el subespacio invariante $\langle \mathbf{a} \rangle_h$ los polinomios característico y mínimo coinciden y por lo tanto el polinomio característico está en el h -anulador de \mathbf{a} . Como el vector \mathbf{a} es arbitrario entonces el polinomio característico está en el h -anulador de todo el espacio y por lo tanto es un múltiplo del polinomio mínimo.

El Teorema de Frobenius se prueba fácilmente sobre los complejos ya que lo esencial aquí, es saber que si p es un factor irreducible del polinomio característico entonces, $\ker(p_h) \neq \emptyset$ o sea, existen vectores de período p . Esto es inmediato en el caso complejo porque todos los irreducibles son de grado 1 y todo valor propio tiene al menos un vector propio correspondiente. En el caso general, para que esto funcione, es necesaria la introducción del campo de descomposición de un polinomio y esto queda fuera de los objetivos de este libro.

Por otro lado, el saber a priori la veracidad de estos dos teoremas no ayuda en mucho para la prueba de la existencia de h -bases, o lo que es lo mismo, la existencia de una descomposición en componentes cíclicas.

Ejercicio 126 Un OL es **diagonalizable** si existe una base en la cual su matriz es diagonal. Pruebe que un OL es diagonalizable *si y solo si* su polinomio mínimo es un producto de **diferentes** polinomios de grado 1. [198]

Ejercicio 127 Pruebe que si un operador lineal en dimensión n tiene n **diferentes** valores propios entonces es diagonalizable. [198]

Ejercicio 128 Un OL es **triangulable** si existe una base ordenada en la cual su matriz es triangular. Pruebe que un OL es triangulable *si y solo si* existe una cadena de subespacios invariantes $\{\mathbf{0}\} \subseteq \mathfrak{E}_1 \subseteq \dots \subseteq \mathfrak{E}_n = \mathfrak{E}$ tales que $\dim \mathfrak{E}_k = k$. [198]

Ejercicio 129 Pruebe que $f \oplus g$ es triangulable *si y solo si* f y g lo son. [198]

Ejercicio 130 Pruebe que un OL es triangulable *si y solo si* los factores irreducibles de su polinomio característico (que son los mismos del mínimo) son polinomios de grado 1. En particular, todo OL complejo es triangulable. [199]

 Los OL sobre \mathbb{C} que no son diagonalizables tienen medida de Lebesgue cero. Estos están contenidos en una hipersuperficie (definida por el discriminante del polinomio característico) y por lo tanto, en cualquier vecindad de cualquier operador casi todos son diagonalizables. Este hecho tiene importantes consecuencias para las aplicaciones en las ciencias naturales.

6.7 Formas normales

Un OL h siempre se descompone en suma directa de sus componentes radicales cíclicas. Por lo tanto, existen bases del espacio vectorial en las cuales la matriz de h es diagonal por bloques. Los bloques son las matrices de las componentes cílico-radicales de h . Nuestra tarea ahora es encontrar bases en las cuales la matriz de un operador

cíclico-radical es lo más “simple” posible. Estas matrices dependen mucho de cual es el polinomio mínimo (que para operadores cílicos son iguales al característico) del operador lineal. Por esto, empezaremos con casos particulares. Esto significa que en lo que sigue usaremos los mismos argumentos varias veces.

Forma normal de Jordán



Camille Jordan (Lyon 1838 - París 1922) Matemático e ingeniero. Conocido tanto por su trabajo en la teoría de grupos como por su influyente libro “Cours d’analyse”. Su libro “Traité des substitutions et des équations algébriques” publicado en 1870 contiene su descubrimiento de las ahora conocidas como formas normales de Jordán. Fué el primero que realmente entendió a Galois después de la publicación póstuma de los trabajos de Galois en 1846 y contribuyó mucho a que la Teoría de Grupos y la Teoría de Galois fueran parte de la corriente principal de las matemáticas.

Una matriz del tipo que se muestra en el recuadro de la derecha se le llama **celda de Jordán**. Observese que el polinomio característico de una celda de Jordan es igual a $(x - \lambda)^n$ y su polinomio mínimo es el mismo. Esto último se puede probar directamente de Hamilton-Caley y haciendo algunos cálculos rutinarios.

$$\begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & \lambda \end{pmatrix}$$

Como no necesitaremos esto, no haremos los cálculos. Sin embargo, el lector puede hacerse una idea de la prueba general con lo siguiente:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, A^4 = 0.$$

6.52

Si \mathbf{h} es un operador cíclico-radical con polinomio mínimo $(x - \lambda)^n$ entonces, en cierta base, la matriz de \mathbf{h} es una celda de Jordan.

Prueba. Denotemos por \mathbf{p} al polinomio $(x - \lambda)$. Sea $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ cíclico-radical con polinomio mínimo \mathbf{p}^n . Sabemos que $\dim \mathfrak{E} = n$. Sea \mathbf{a} tal que $\langle \mathbf{a} \rangle_{\mathbf{h}} = \mathfrak{E}_n$. Tal vector existe ya que \mathbf{h} es cíclico. Sabemos que $\text{per}_{\mathbf{h}}(\mathbf{a}) = \mathbf{p}^n$.

Para $k \in \{1, \dots, n\}$ denotemos $\mathbf{a}_k = \mathbf{p}_h^{n-k}(\mathbf{a})$. Sea $A \stackrel{\text{def}}{=} \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ y demostremos que A es linealmente independiente y por lo tanto una base del espacio.

Supongamos que para ciertos escalares $\sum \beta_k \mathbf{a}_k = \mathbf{0}$. Entonces

$$\mathbf{0} = \sum_{k=1}^n \beta_k \mathbf{a}_k = \sum_{k=1}^n \beta_k \mathbf{p}_h^{n-k}(\mathbf{a}) = \left(\sum_{j=0}^{n-1} \beta_{n-j} \mathbf{p}^j \right)_h (\mathbf{a}) \stackrel{\text{def}}{=} \mathbf{q}_h(\mathbf{a})$$

El polinomio \mathbf{q} es de grado estrictamente menor que el grado del período de \mathbf{a}_n y por

lo tanto $q = 0$. Además, en la familia de polinomios p^j hay exactamente un polinomio de grado t para cualquier $t \in \{0, \dots, n-1\}$. Luego, estos polinomios son linealmente independientes y de $q = 0$ concluimos que todos los β_{kj} son cero.

Veamos como es la matriz de h en esta base. Para $k \in \{1, \dots, n-1\}$ tenemos $a_k = p_h(a_{k+1}) = h(a_{k+1}) - \lambda a_{k+1}$ y despejando obtenemos $h(a_{k+1}) = a_k + \lambda a_{k+1}$. En otras palabras, en la base A el vector $h(a_{k+1})$ tiene coordenadas $(0, \dots, 1, \lambda, \dots, 0)$ donde el 1 aparece en el índice k y λ aparece en el índice $k+1$. Estas son las columnas $2, \dots, n$ de una celda de Jordán.

Para ver como es la primera columna, observemos que $a_1 \in \mathfrak{E}_1 = \ker p_h$ y por lo tanto $h(a_1) - \lambda a_1 = 0$. En otras palabras, en la base A el vector $h(a_1)$ tiene coordenadas $(\lambda, \dots, 0)$ y esa es la primera columna de una celda de Jordán. ■

El resultado anterior se aplica para operadores cíclico-radicales cuyo polinomio mínimo es potencia de un polinomio de grado 1 . Como por el Teorema de Gauss, los polinomios irreducibles en $\mathbb{C}[x]$ siempre son de grado 1 , entonces este siempre es el caso para los OL cíclico-radicales definidos en un espacio vectorial sobre los complejos.

Conjugando esto con la descomposición de un operador lineal en sus componentes cíclico-radicales obtenemos el siguiente:

Forma normal de Jordán

6.5.9 Si h es un operador lineal en un espacio vectorial finito dimensional sobre el campo de los números complejos entonces, existe una base del espacio vectorial en la cual la matriz del operador es diagonal por bloques y los bloques son celdas de Jordán.

Forma normal real

En el caso de los operadores lineales en un espacio vectorial sobre el campo \mathbb{R} de los números reales el problema es un poco más complicado ya que tenemos más polinomios de grado 2 que son irreducibles. Más precisamente, un polinomios del tipo $(x - \alpha)^2 + \beta^2$ es irreducible si $\beta \neq 0$. Denotemos

$$\Lambda = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

La matriz Λ tiene polinomio característico $(x - \alpha)^2 + \beta^2$ y como este polinomio es irreducible entonces por Hamilton-Caley su polinomio mínimo es el mismo. La matriz Λ es muy parecida a la de una rotación en el plano \mathbb{R}^2 . De hecho, si $\alpha^2 + \beta^2 = 1$ entonces, es la matriz de una rotación de un ángulo igual a $\arccos \alpha$.

Una matriz del tipo que se muestra en el recuadro de la derecha se le llama **celda cuadrática**. Observese que el polinomio característico de una celda cuadrática es igual a $((x - \alpha)^2 + \beta^2)^n$. El lector debe notar el parecido con las celdas de Jordán. El parecido está dado por la correspondencia de símbolos $\lambda \leftrightarrow \Lambda$, $1 \leftrightarrow I$ y $0 \leftrightarrow O$. La diferencia fundamental está en que en el caso de las celdas de Jordán las entradas son elementos del campo y aquí son matrices de orden 2.

$$\begin{pmatrix} \Lambda & I & \cdots & O \\ O & \Lambda & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & I \\ O & O & \cdots & \Lambda \end{pmatrix}$$

6.54

Si $((x - \alpha)^2 + \beta^2)^n$ con $\beta \neq 0$ es el polinomio mínimo de un operador cíclico-radical h entonces, en cierta base, la matriz de h es una celda cuadrática.

Prueba. Denotemos por p al polinomio $((x - \alpha)^2 + \beta^2)$. Sea $h : E \rightarrow E$ cíclico radical con polinomio mínimo p^n . Sabemos que $\dim E = n$. Sea a tal que $\langle a \rangle_h = E_n$. Tal vector existe ya que h es cíclico. Sabemos que $\text{per}_h(a) = p^n$.

Denotemos por r al polinomio $(x - \alpha)/\beta$. El lector puede comprobar fácilmente que se cumple la siguiente igualdad polinomial

$$xr = \frac{p}{\beta} - \beta + \alpha r \quad \dagger$$

Para $k \in \{1, \dots, n\}$ denotemos $a_k = (p/\beta)_h^{n-k}(a)$ y $b_k = r_h(a_k)$. Sea $A \stackrel{\text{def}}{=} \{a_1, b_1, \dots, a_n, b_n\}$ y demostremos que A es linealmente independiente y por lo tanto una base del espacio.

Supongamos que para ciertos escalares $\sum \omega_k a_k + \sum \rho_k b_k = 0$. Entonces

$$0 = \sum_{k=1}^n (\omega_k a_k + \rho_k b_k) = \left(\sum_{j=0}^{n-1} \left(\frac{\omega_{n-j} + \rho_{n-j} r}{\beta} \right) p^j \right)_h (a) \stackrel{\text{def}}{=} q_h(a).$$

El polinomio q es de grado estrictamente menor que el grado del período de a y por lo tanto $q = 0$. Además, en la familia de polinomios p^j, rp^j hay exactamente un polinomio de grado t para cualquier $t \in \{0, \dots, 2n-1\}$. Luego, estos polinomios son linealmente independientes y de $q = 0$ concluimos que todos los coeficientes son cero.

Veamos como es la matriz de h en la base ordenada A . Tenemos $b_k = r_h(a_k) = \frac{1}{\beta}(h(a_k) - \alpha a_k)$ y despejando $h(a_k) = \alpha a_k + \beta b_k$. Esto justifica todas las columnas con índice impar de la celda cuadrática. Además, usando la igualdad polinomial \dagger tenemos

$$h(b_k) = (xr)_h(a_k) = \left(\frac{p}{\beta} \right)_h(a_k) - \beta a_k + \alpha r(a_k) = a_{k-1} - \beta a_k + \alpha b_k$$

Esta igualdad es válida incluso para $k = 1$ si definimos $a_0 = 0$. Esto justifica todas las columnas con índice par de la celda cuadrática. ■

El resultado anterior es válido sobre cualquier campo (por ejemplo \mathbb{Q}) en el cual $(x - \alpha)^2 + \beta^2$ sea un polinomio irreducible. Para \mathbb{R} esto termina el análisis ya que todo

polinomio real irreducible es de esta forma o es de grado 1.

Conjugando esto con la descomposición de un operador lineal en sus componentes cílico-radicales obtenemos el siguiente:

Forma normal real

6.55

Si \mathbf{h} es un operador lineal en un espacio vectorial finito dimensional sobre el campo de los números reales entonces, existe una base del espacio vectorial en la cual la matriz del operador es diagonal por bloques y los bloques son celdas de Jordán o celdas cuadráticas.

Forma normal canónica

Ahora analizaremos el caso general en el cual el polinomio irreducible es arbitrario. Denotaremos por \mathbf{p} al polinomio $\alpha_0 + \alpha_1x + \dots + \alpha_{m-1}x^{m-1} + x^m$. Así, \mathbf{p} es mónico de grado m y supondremos que \mathbf{p} es irreducible en nuestro campo. En este caso, nuestros bloques para construir las celdas son las matrices de orden m siguientes

$$\Lambda = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -\alpha_{m-2} \\ 0 & 0 & \cdots & 1 & -\alpha_{m-1} \end{pmatrix}, \quad I = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \quad O = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

El lector debe reconocer que Λ es la matriz acompañante del polinomio \mathbf{p} .

Una matriz del tipo que se muestra en el recuadro de la derecha donde los bloques son los definidos previamente se le llama **celda canónica**. Usando 6.47 obtenemos que el polinomio característico de una celda canónica es igual a \mathbf{p}^n . La diferencia con las celdas cuadráticas es que los bloques son diferentes aunque estén denotados con las mismas letras.

$$\begin{pmatrix} \Lambda & I & \cdots & O \\ O & \Lambda & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & I \\ O & O & \cdots & \Lambda \end{pmatrix}$$

Observese que las celdas canónicas para $\mathbf{p} = (x - \lambda)$ son celdas de Jordán. Sin embargo, las celdas canónicas y las celdas cuadráticas son diferentes. Efectivamente, supongamos $m = 2$ y $\mathbf{p} = (x - \alpha)^2 + \beta^2 = x^2 - 2\alpha x + \gamma$ donde $\gamma = \alpha^2 + \beta^2$ entonces, en la desigualdad

$$\begin{pmatrix} 0 & -\gamma & 0 & 1 \\ 1 & 2\alpha & 0 & 0 \\ 0 & 0 & 0 & -\gamma \\ 0 & 0 & 1 & 2\alpha \end{pmatrix} \neq \begin{pmatrix} \alpha & -\beta & 0 & 1 \\ \beta & \alpha & 0 & 0 \\ 0 & 0 & \alpha & -\beta \\ 0 & 0 & \beta & \alpha \end{pmatrix}$$

la celda canónica es la de la izquierda y la celda cuadrática es la de la derecha. La ventaja de las celdas cuadráticas es que son más fáciles de intuir geométricamente mediante rotaciones. La ventaja de las celdas canónicas es que siempre funcionan.

6.56

Si \mathbf{h} es un operador cíclico-radical con polinomio mínimo \mathbf{p}^n entonces, en cierta base, la matriz de \mathbf{h} es una celda canónica.

Prueba. Sea \mathbf{p} mónico de grado \mathbf{m} . Sea $\mathbf{h} : \mathfrak{E} \rightarrow \mathfrak{E}$ cíclico radical con polinomio mínimo \mathbf{p}^n . Tenemos que $\dim \mathfrak{E}_k = \mathbf{mk}$. Sea \mathbf{a} tal que $\langle \mathbf{a} \rangle_{\mathbf{h}} = \mathfrak{E}_{\mathbf{n}}$. Tal vector existe ya que \mathbf{h} es cíclico. Sabemos que $\text{per}_{\mathbf{h}}(\mathbf{a}) = \mathbf{p}^n$. Para $k \in \{1, \dots, n\}$ y $j \in \{0, \dots, \mathbf{m}-1\}$, denotemos \mathbf{a}_k^j como en el recuadro a la derecha.

Tenemos \mathbf{nm} vectores \mathbf{a}_k^j y necesitamos convencernos que estos forman una base del espacio, o lo que es lo mismo, que son linealmente independientes. Sean β_{kj} escalares tales que $\sum \beta_{kj} \mathbf{a}_k^j = \mathbf{0}$. Tenemos

$$\mathbf{0} = \sum_{k,j} \beta_{kj} \mathbf{a}_k^j = \sum_{k,j} \beta_{kj} (x^j \mathbf{p}^{n-k})_{\mathbf{h}} (\mathbf{a}) = \left(\sum_{k,j} \beta_{kj} x^j \mathbf{p}^{n-k} \right)_{\mathbf{h}} (\mathbf{a}) \stackrel{\text{def}}{=} \mathbf{q}_{\mathbf{h}} (\mathbf{a}).$$

El polinomio \mathbf{q} es de grado estrictamente menor que el grado del período de \mathbf{a}_n y por lo tanto $\mathbf{q} = \mathbf{0}$. Además, en la familia de polinomios $x^j \mathbf{p}^k$ hay exactamente un polinomio de grado \mathbf{t} para cualquier $\mathbf{t} \in \{0, \dots, \mathbf{nm}-1\}$. Luego, estos polinomios son linealmente independientes y de $\mathbf{q} = \mathbf{0}$ concluimos que todos los β_{kj} son cero.

Ordenemos nuestra base de la siguiente manera

$$\mathbf{a}_1^0, \mathbf{a}_1^1, \dots, \mathbf{a}_1^{\mathbf{m}-1}, \mathbf{a}_2^0, \mathbf{a}_2^1, \dots, \mathbf{a}_2^{\mathbf{m}-1}, \dots, \mathbf{a}_n^0, \mathbf{a}_n^1, \dots, \mathbf{a}_n^{\mathbf{m}-1}$$

y veamos como es la matriz de \mathbf{h} en esta base ordenada.

Para $j \in \{0, \dots, \mathbf{m}-2\}$ tenemos

$$\mathbf{h}(\mathbf{a}_k^j) = \mathbf{h}((x^j \mathbf{p}^{n-k})_{\mathbf{h}} (\mathbf{a}_n)) = \mathbf{h}(\mathbf{h}^j(\mathbf{p}_{\mathbf{h}}^{n-k}(\mathbf{a}_n))) = \mathbf{a}_k^{j+1}$$

y esto justifica las columnas de la celda canónica cuyos índices no son divisibles entre \mathbf{m} que son del tipo $(0, \dots, 0, 1, 0, \dots, 0)$.

Denotemos los coeficientes de \mathbf{p} por α_i , o sea, $\mathbf{p} = \alpha_0 + \alpha_1 x + \dots + \alpha_{m-1} x^{m-1} + x^m$ y definamos $\mathbf{r} \stackrel{\text{def}}{=} \mathbf{p} - x^m$. Tenemos

$$\begin{aligned} \mathbf{h}(\mathbf{a}_k^{\mathbf{m}-1}) &= (x^m \mathbf{p}^{n-k})_{\mathbf{h}} (\mathbf{a}_n) = ((\mathbf{p} - \mathbf{r}) \mathbf{p}^{n-k})_{\mathbf{h}} (\mathbf{a}_n) = \\ &= \mathbf{p}_{\mathbf{h}}^{n-(k-1)} (\mathbf{a}_n) - (\mathbf{r} \mathbf{p}^{n-k})_{\mathbf{h}} (\mathbf{a}_n) = \\ &= \mathbf{a}_{k-1}^0 - \alpha_0 \mathbf{a}_k^0 - \alpha_1 \mathbf{a}_k^1 - \dots - \alpha_{m-1} \mathbf{a}_k^{m-1} \end{aligned}$$

la cual es válida incluso para $k = 1$ si definimos $\mathbf{a}_0^0 = \mathbf{0}$. Esto justifica las columnas de la celda canónica cuyos índices son divisibles entre \mathbf{m} que son del tipo $(0, \dots, 0, 1, 0, \dots, 0, -\alpha_0, \dots, \alpha_{m-1}, 0, \dots, 0)$. ■

Conjugando esto con la descomposición de un operador lineal en sus componentes cíclico-radicales obtenemos el siguiente:

Forma normal canónica

6.57

Si \mathbf{h} es un operador lineal en un espacio vectorial finito dimensional sobre cualquier campo entonces, existe una base en la cual la matriz del operador es diagonal por bloques y los bloques son celdas canónicas.



Definamos el **tipo de un OL** como el tipo de una de sus descomposiciones en componentes radicales cíclicas. Hay muchísima libertad al construir formas normales. Por esto, no hay que sobredimensionar la importancia de las diferencias entre unas y otras. Lo importante y común a todas las formas normales es que ellas *solo dependen del tipo del operador* y no del operador en si mismo.

Ejercicio 131 Se dice que dos OL \mathbf{f} y \mathbf{g} son conjugados si existe un OL invertible \mathbf{h} tal que $\mathbf{f} = \mathbf{h} \circ \mathbf{g} \circ \mathbf{h}^{-1}$. Demuestre que dos OL tienen el mismo tipo *si y solo si* son conjugados. [199]



Soluciones de ejercicios selectos

Ejercicio 2 (Sección 1.1 página 2) Una operación unaria en el conjunto A es una función de A en A . Por ejemplo para cada entero x hay otro entero $-x$. De esta manera la operación $x \mapsto -x$ es una operación unaria. Otros ejemplos comunes son el hallar el complemento de un conjunto o el hallar el complejo conjugado de un número complejo.

Ejercicio 3 (Sección 1.1 página 2) El área del triángulo con lados a, b y c no es una operación ternaria en \mathbb{R} ya que no para cualesquiera a, b y c existe un triángulo con lados de esas longitudes.

Ejercicio 4 (Sección 1.1 página 2) La fórmula $(a + b)(x + y)$ se expresa en notación sufija como $ab + xy + x$.

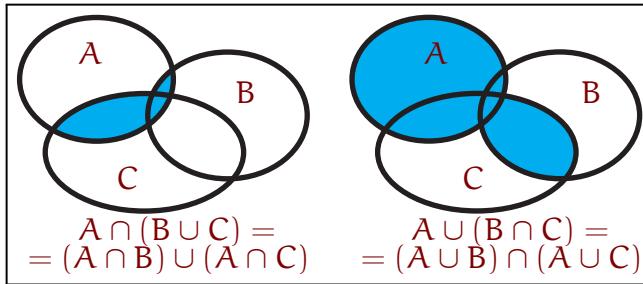
Ejercicio 5 (Sección 1.1 página 3) La suma, el producto, el máximo común divisor el mínimo común múltiplo y el máximo son operaciones conmutativas. Las demás no lo son.

Ejercicio 6 (Sección 1.1 página 3) Ya vimos en el texto que la exponentiación no es asociativa. Observemos que $4 = 4 - (2 - 2) \neq (4 - 2) - 2 = 0$ por lo que la resta no es asociativa. Tenemos $4 = 4 / (2/2) \neq (4/2)/2 = 1$ y por lo tanto la división no es asociativa. Tampoco es asociativo el logaritmo. El resto de las operaciones mencionadas sí son asociativas.

Ejercicio 7 (Sección 1.1 página 4) La resta no tiene elemento neutro ya que de $a - e = a$ se sigue que $e = 0$ pero por otro lado $0 - a = -a \neq a$. Lo mismo ocurre con la división. La operación de exponentiación no tiene elemento neutro ya que $e^1 = 1 \Rightarrow e = 1$ pero $1^2 = 1 \neq 2$. Lo mismo ocurre con el logaritmo. El neutro de la suma es el cero, el neutro del producto es el uno. El 1 también es el neutro para el mínimo común múltiplo. La operación de máximo común divisor no tiene neutro.

Ejercicio 8 (Sección 1.1 página 5) Es importante señalar que el inverso tiene sentido solo si hay neutro. El inverso de a para la suma es $-a$, para el producto es $\frac{1}{a}$. Aunque el mínimo común múltiplo tiene neutro 1, esta operación no tiene inversos.

Ejercicio 9 (Sección 1.1 página 5) Fijémonos en un conjunto U y en el conjunto de todos los subconjuntos de U que denotaremos por 2^U . En 2^U están bien definidas las operaciones de unión e intersección de conjuntos. Estas operaciones cumplen las propiedades requeridas como se puede observar en los dos siguientes diagramas de Venn.



Ejercicio 11 (Sección 1.2 página 8) No porque $(1, 0) \times (0, 1) = (0, 0)$ lo que muestra que \mathbb{K}^2 no es dominio de integridad y por lo tanto no es un campo.

Ejercicio 12 (Sección 1.2 página 9) Supongamos que \sqrt{n} es un racional. Descomponiendo su numerador y denominador en factores primos obtenemos que $\sqrt{n} = p_1^{n_1} p_2^{n_2} \dots p_t^{n_t}$ para ciertos naturales primos $p_1 p_2 \dots p_t$ y ciertos números enteros n_1, n_2, \dots, n_t . Pero entonces, $n = p_1^{2n_1} p_2^{2n_2} \dots p_t^{2n_t}$ y por lo tanto $\forall i \in \{1, \dots, t\} \quad 2n_i \geq 0$. Luego, todos los n_i son naturales lo que significa que \sqrt{n} es natural.

Ejercicio 13 (Sección 1.2 página 9) Tenemos $1 < \sqrt{2} < 2$ y por lo tanto $\sqrt{2}$ no es natural. Por el ejercicio anterior, no es racional.

Ejercicio 14 (Sección 1.2 página 9) Siempre podemos medir un segmento de recta con cada vez más precisión (al menos teóricamente). Cada medición nos da un número racional. Luego, la longitud del segmento es un límite de racionales por lo que es un real.

Ejercicio 18 (Sección 1.3 página 12) Por 1.1.4 sabemos que $f(0) = 0$ y $f(1) = 1$. Si $f(a) = 0$ y a no fuera el cero de A entonces tendríamos $1 = f(1) = f(aa^{-1}) = f(a)f(a^{-1}) = 0f(a^{-1}) = 0$ lo que no puede ser, o sea si $f(a) = 0$ entonces $a = 0$. Luego, si $f(x) = f(y)$ entonces, $f(x - y) = 0$ y por lo tanto $x = y$.

Ejercicio 20 (Sección 1.3 página 14) Si f no es inyectiva entonces existen $x \neq y$ tales que $f(x) = f(y)$ y por lo tanto $g(f(x)) = g(f(y))$. Luego, $g \circ f$ no puede ser la identidad.

Si g no es sobreyectiva entonces existe x tal que $\forall y$ se tiene que $g(y) \neq x$ en particular $\forall f(z)$ se tiene que $g(f(z)) \neq x$ y por lo tanto $g \circ f$ no es sobreyectiva. Luego, $g \circ f$ no puede ser la identidad.

Ejercicio 22 (Sección 1.4 página 17)

La tabla de multiplicar en \mathbb{Z}_5 es la derecha. Luego el elemento inverso de 1 es 1 , el elemento inverso de 2 es 3 , el elemento inverso de 3 es 2 y el elemento inverso de 4 es 4 . Observese que en la tabla de multiplicar de $\mathbb{Z}_5 \setminus \{0\}$ en cada columna y en cada renglón nos encontramos con todos los elementos posibles. Esto es una propiedad de las operaciones binarias de los grupos. De hecho, un conjunto con una operación binaria asociativa y con elemento neutro es un grupo si y solo si su tabla de multiplicar cumple esta propiedad.

•	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Ejercicio 23 (Sección 1.4 página 17) La prueba de que es un anillo conmutativo es directa de las definiciones de suma y producto. Para ver que es un campo observamos que el producto $(a, b)(x, y) = (ax + 7by, ay + bx)$ tiene neutro $(1, 0)$. Para hallar los inversos multiplicativos resolvemos el sistema de ecuaciones lineales

$$\begin{cases} ax + 7by = 1 \\ ay + bx = 0 \end{cases}$$

que tiene como solución única $x = a/\Delta$, $y = -b/\Delta$ donde $\Delta = a^2 - 7b^2$. Nos queda comprobar que si $\Delta = 0$ entonces $a = b = 0$. Efectivamente, observando la siguiente tabla calculada en \mathbb{Z}_{11}

x	1	2	3	4	5	6	7	8	9	10
x^2	1	4	9	5	3	3	5	9	4	1
$7x^2$	7	6	8	2	10	10	2	8	6	7

vemos que ningún a^2 puede ser igual a un $7b^2$.

Ejercicio 24 (Sección 1.5 página 19) No, porque t no es un elemento del campo al contrario, t es un número natural. Lo que quiere decir tx es $x + \dots + x$, t veces.

Ejercicio 26 (Sección 1.5 página 19) De que $a = -a$ obtenemos que

$$0 = a + a = 1a + 1a = (1 + 1)a$$

Como todo campo es dominio de integridad obtenemos que $a = 0$ o $1 + 1 = 0$. Si la característica del campo es diferente de 2 entonces, $1 + 1 \neq 0$ por lo que la afirmación del ejercicio es cierta. Si la característica del campo es 2 entonces, $1 + 1 = 0$ y por lo tanto $a = -a$ para cualquier elemento del campo.

Ejercicio 27 (Sección 1.6 página 21) En cualquier campo $(xy)^p = x^p y^p$. Por el binomio de Newton $(x+y)^p = \sum_{k=0}^p \frac{p!}{k!(p-k)!} x^k y^{p-k}$. El coeficiente binomial $p!/k!(p-k)!$ se divide entre p si $k \in [1, \dots, p-1]$. Esto quiere decir (ver 1.15) que en un campo de característica p todos los sumandos del binomio de Newton excepto el primero y el último son cero. Luego $(x+y)^p = x^p + y^p$. Esto muestra que $x \mapsto x^p$ es un morfismo. Como todo morfismo de campos es inyectivo (ver ejercicio 18) y toda función inyectiva de un conjunto finito en si mismo es sobreyectiva obtenemos que este morfismo es automorfismo para campos finitos.

Ejercicio 28 (Sección 1.7 página 23) Usando la distributividad por la derecha e izquierda obtenemos que:

$$(a+1)(b+1) = (a+1)b + a + 1 = ab + b + a + 1$$

$$(a+1)(b+1) = a(b+1) + b + 1 = ab + a + b + 1$$

Igualando y cancelando ab por la izquierda y 1 por la derecha obtenemos que $b+a = a+b$.

Ejercicio 29 (Sección 1.7 página 24) Las seis igualdades necesarias se pueden obtener de la siguiente manera

$$\begin{aligned} (ijk = -1) &\Rightarrow (ijkk = -k) & \Rightarrow (ij = k) &\Rightarrow (ijj = kj) &\Rightarrow (kj = -i) \\ (ijk = -1) &\Rightarrow (iijk = -i) & \Rightarrow (jk = i) &\Rightarrow (jkk = ik) &\Rightarrow (ik = -j) \\ (ijk = -1) &\Rightarrow (kijkkj = -kkj) & \Rightarrow (ki = j) &\Rightarrow (kii = ji) &\Rightarrow (ji = -k) \end{aligned}$$

Ejercicio 30 (Sección 1.7 página 24) Por definición de producto de quaternios

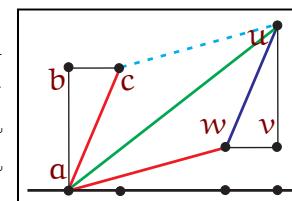
$$\left((a + bi + cj + dk)^2 = -1 \right) \Leftrightarrow \left(\begin{array}{l} a^2 - (b^2 + c^2 + d^2) = -1 \\ ab = ac = ad = 0 \end{array} \right).$$

Como a es un real $b^2 + c^2 + d^2 \neq 0$ por lo que cualquier solución de estas ecuaciones requiere que $a = 0$.

Ejercicio 31 (Sección 1.7 página 24) En los quaternios $(x-i)(x+i) = x^2 - ix + xi + 1 \neq x^2 + 1$ lo que quiere decir que no es posible definir correctamente los polinomios con coeficientes quaterniónicos.

Ejercicio 32 (Sección 2.2 página 29)

El triángulo abc es semejante al triángulo uvw de hecho son congruentes porque $\overline{ac} = \overline{uw}$ (lados opuestos de un paralelogramo tienen la misma longitud). Luego $\overline{bc} = \overline{vw}$ y por lo tanto la coordenada en x del vector \overrightarrow{au} es igual a la de \overrightarrow{aw} más \overrightarrow{bc} . La prueba para la otra coordenada es igual.



Ejercicio 34 (Sección 2.2 página 29) En la expresión $\alpha(\beta\mathbf{a})$ se utiliza un solo tipo de operación (la de un escalar por un vector). En la expresión $(\alpha\beta)\mathbf{a}$ se utilizan dos operaciones diferentes primero la del producto de escalares y después la de un escalar por un vector.

Ejercicio 35 (Sección 2.2 página 29) Si $\alpha \neq 0$ entonces tiene inverso y por lo tanto

$$\alpha\mathbf{a} = \mathbf{0} \Rightarrow \mathbf{a} = \alpha^{-1}\alpha\mathbf{a} = \alpha^{-1}\mathbf{0} = \mathbf{0}$$

Ejercicio 36 (Sección 2.2 página 29) El mínimo número de elementos en un espacio vectorial es 1 ya que al menos necesita tener el neutro para la suma de vectores. Y efectivamente un conjunto formado por un solo elemento $\mathbf{0}$ es un espacio vectorial sobre cualquier campo definiendo $\alpha\mathbf{0} = \mathbf{0}$ para cualquier α en el campo.

Ejercicio 38 (Sección 2.3 página 37) Supongamos que $x \in \langle N \cup y \rangle \setminus \langle N \rangle$. Entonces, existe una combinación lineal $x = \alpha_y y + \sum_{i \in N} \alpha_i i$. Tenemos que $\alpha_y \neq 0$ (ya que $x \notin \langle N \rangle$). Luego, despejando y obtenemos que $y \in \langle N \cup x \rangle$.

Ejercicio 39 (Sección 2.4 página 39) Solo en la prueba de que $2 \Rightarrow 4$ al despejar a .

Ejercicio 40 (Sección 2.4 página 41) 1.- El conjunto vacío es LI y todo el espacio \mathfrak{E} es generador. Luego existe N tal que $\emptyset \subseteq N \subseteq \mathfrak{E}$ que es base. 2.- Si M es LI entonces existe una base N tal que $M \subseteq N \subseteq \mathfrak{E}$. 3.- Si L es generador entonces existe una base N tal que $\emptyset \subseteq N \subseteq L$.

Ejercicio 41 (Sección 2.4 página 42) Sea A una base de \mathfrak{E} . Como \mathfrak{F} es generador y $A \subseteq \mathfrak{F}$ es LI entonces por el Teorema de Existencia de Bases (2.14) existe una base B de \mathfrak{F} que contiene a A . De aquí tenemos $\dim \mathfrak{E} = |A| \leq |B| = \dim \mathfrak{F}$. Esta prueba es válida independientemente de si los cardinales de A y B son finitos o no.

Ejercicio 42 (Sección 2.4 página 42) El conjunto de todos los polinomios $\sum a_i x^i$ tales que $a_0 = 0$ es un subespacio de $\mathbb{K}[x]$ de la misma dimensión que $\mathbb{K}[x]$.

Ejercicio 43 (Sección 2.4 página 44) La diferencia entre el espacio de las $(N \times M)$ -adas y las NM -matrices es solo en las notaciones. Luego, el espacio de las NM -matrices tiene dimensión $|N \times M| = |N| |M|$. Para cada pareja i, j con $i \in N$ y $j \in M$ hay una matriz e_{ij} en la base canónica la cual tiene su entrada ij igual a 1 y todas las demás igual a cero.

Ejercicio 44 (Sección 2.5 página 44) Sean $\mathfrak{E} \xrightarrow{f} \mathfrak{F} \xrightarrow{g} \mathfrak{G}$ transformaciones lineales. Tenemos que $f(g(x+y)) = f(g(x)+g(y)) = f(g(x))+f(g(y))$ y $f(g(\lambda x)) = f(\lambda g(x)) = \lambda f(g(x))$ y esto era todo lo que se quería probar.

Ejercicio 46 (Sección 2.5 página 48) El conjunto \mathfrak{E} es contable y es un espacio vectorial sobre \mathbb{Q} . Si existe un racional α tal que $y = \alpha x$ entonces la dimensión de \mathfrak{E} sobre \mathbb{Q} es 1. Si por el contrario α no existe entonces la dimensión de \mathfrak{E} sobre \mathbb{Q} es 2. El conjunto \mathfrak{E} no es un espacio vectorial sobre \mathbb{R} .

Ejercicio 48 (Sección 2.6 página 52) 1.- Es fácil probar que la aplicación $\mathfrak{E} \oplus \mathfrak{F} \ni (a, b) \mapsto (b, a) \in \mathfrak{F} \oplus \mathfrak{E}$ es un isomorfismo canónico. 2.- Comutatividad. 3.- El isomorfismo canónico es $((a, b), c) \mapsto (a, (b, c)) \mapsto (a, b, c)$. 4.- Si. La aplicación $\mathfrak{E} \oplus \{0\} \ni (a, 0) \mapsto a \in \mathfrak{E}$ es un isomorfismo canónico. Por esto podemos pensar que $\mathfrak{E} \oplus \{0\} = \mathfrak{E}$.

Ejercicio 49 (Sección 2.6 página 53) Denotemos por \mathfrak{S} a todo el espacio. Como cada vector se expresa como $a + b$ con $a \in \mathfrak{E}$ y $b \in \mathfrak{F}$ tenemos $\mathfrak{S} = \mathfrak{E} + \mathfrak{F}$. Supongamos que $a \neq 0 \in \mathfrak{E} \cap \mathfrak{F}$ entonces $0 = 0 + 0 = a - a$ lo que contradice que la descomposición

de $\mathbf{0}$ es única. Luego $\mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$ y por lo tanto $\mathfrak{S} = \mathfrak{E} \oplus \mathfrak{F}$.

Ejercicio 55 (Sección 2.7 página 56) Si $E = \mathfrak{E} + \mathbf{x}$ y $F = \mathfrak{F} + \mathbf{y}$ son dos subespacios paralelos o no entonces $E + F$ es igual a $(\mathfrak{E} + \mathfrak{F}) + (\mathbf{x} + \mathbf{y})$ o sea, es un espacio afín paralelo a $\mathfrak{E} + \mathfrak{F}$. Si E es paralelo a F entonces, $\mathfrak{E} = \mathfrak{F}$ y por lo tanto $\mathfrak{E} + \mathfrak{F} = \mathfrak{E}$.

Ejercicio 60 (Sección 3.4 página 76) Tómese $\mathbf{x} = (1, 0)$, $\mathbf{y} = (0, 1)$ y $\mathbf{z} = (1, 1)$. Tenemos $(\mathbf{x}\mathbf{y})\mathbf{z} = 0(1, 1) = (0, 0)$ y $\mathbf{x}(\mathbf{y}\mathbf{z}) = (1, 0)1 = (1, 0)$ por lo que $(\mathbf{x}\mathbf{y})\mathbf{z} \neq \mathbf{x}(\mathbf{y}\mathbf{z})$.

Ejercicio 64 (Sección 3.4 página 78) $\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$

Ejercicio 65 (Sección 3.4 página 78) Para cualesquiera $n \in \mathbb{N}$ y $k \in \mathbb{K}$ tenemos

$$(\alpha_{nM} \beta_{ML}) \gamma_{Lk} = \sum_{l \in L} (\alpha_{nM} \cdot \beta_{ML}) \gamma_{lk} = \sum_{l \in L} \sum_{m \in M} \alpha_{nm} \beta_{ml} \gamma_{lk} = \\ \sum_{m \in M} \alpha_{nm} \sum_{l \in L} \beta_{ml} \gamma_{lk} = \sum_{m \in M} \alpha_{nm} (\beta_{ML} \cdot \gamma_{Lk}) = \alpha_{nM} (\beta_{ML} \gamma_{Lk})$$

Ejercicio 66 (Sección 3.4 página 79) Las matrices de f , g y $f \circ g$ tienen que cumplir:

$$\begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} = \\ = \begin{pmatrix} \cos \alpha \cos \beta - \sin \alpha \sin \beta & -\cos \alpha \sin \beta - \sin \alpha \cos \beta \\ \cos \alpha \sin \beta + \sin \alpha \cos \beta & \cos \alpha \cos \beta - \sin \alpha \sin \beta \end{pmatrix}$$

y por lo tanto

$$\begin{aligned} \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin(\alpha + \beta) &= \cos \alpha \sin \beta + \sin \alpha \cos \beta \end{aligned}$$

Ejercicio 69 (Sección 3.7 página 86) Sea f semilineal y σ y ρ dos automorfismos del campo tales que para cualquier escalar λ y cualquier vector \mathbf{a} se cumple que $f(\lambda \mathbf{a}) = \sigma(\lambda) f(\mathbf{a}) = \rho(\lambda) f(\mathbf{a})$. Podemos escoger \mathbf{a} tal que $f(\mathbf{a}) \neq \mathbf{0}$ y por lo tanto $((\sigma(\lambda) - \rho(\lambda)) f(\mathbf{a}) = \mathbf{0}) \Rightarrow (\sigma(\lambda) = \rho(\lambda))$.

Ejercicio 70 (Sección 3.7 página 87) Supongamos que $\sup A$ existe. Entonces, $\forall b \in \mathbb{R}$ tenemos $(\forall a \in A \ b \geq a) \Rightarrow (b \geq \sup A)$. Usando que f y f^{-1} son monótonas vemos que $\forall f(b) \in f(\mathbb{R}) = \mathbb{R}$ tenemos $(\forall f(a) \in f(A) \ f(b) \geq f(a)) \Leftrightarrow (\forall a \in A \ b \geq a) \Rightarrow (b \geq \sup A) \Rightarrow (f(b) \geq f(\sup A))$ y esto prueba que $f(\sup A) = \sup f(A)$.

Ejercicio 71 (Sección 3.7 página 87) (1 \Rightarrow 2) La conjugación compleja es continua.
(2 \Rightarrow 3) Tenemos (véase la prueba de 3.31) que f es la identidad en \mathbb{Q} . De que los reales son los límites de los racionales y f es continua obtenemos que f es la identidad en \mathbb{R} .
(3 \Rightarrow 1) Si f es la identidad en \mathbb{R} entonces, $f(a + bi) = a + bf(i)$. Como $f(i)^2 = f(i^2) = f(-1) = -1$ y $f(i) \neq i$ entonces, $f(i) = -i$.

Para ver que hay muchos automorfismos de \mathbb{C} que no son la conjugación compleja véase por ejemplo: Yale, Paul B., *Automorphisms of the complex numbers*, Mathematics

Magazine, May-June 1966, 135–141.

Ejercicio 72 (Sección 3.7 página 88) La función $(x_1, \dots, x_n) \mapsto (\bar{x}_1, \dots, \bar{x}_n)$ es biyectiva ya que $\lambda \mapsto \bar{\lambda}$ es biyectiva. Además, tenemos

$$\begin{aligned}(\bar{x}_1 + \bar{y}_1, \dots, \bar{x}_n + \bar{y}_n) &= (\bar{x}_1 + \bar{y}_1, \dots, \bar{x}_n + \bar{y}_n) = (\bar{x}_1, \dots, \bar{x}_n) + (\bar{y}_1, \dots, \bar{y}_n) \\(\bar{\lambda}x_1, \dots, \bar{\lambda}x_n) &= (\bar{\lambda}\bar{x}_1, \dots, \bar{\lambda}\bar{x}_n) = \bar{\lambda}(\bar{x}_1, \dots, \bar{x}_n)\end{aligned}$$

y esto muestra que la función es un automorfismo semilineal.

Si f es una transformación semilineal arbitraria cuyo automorfismo del campo es σ entonces, su composición con el automorfismo semilineal estandar correspondiente a σ^{-1} es una transformación lineal.

Ejercicio 74 (Sección 3.7 página 91) Supongamos que $0 = \alpha a + \alpha c + \beta b - \beta c = (\alpha - \beta)c + (\alpha + \beta\rho)a$. Como $\{a, c\}$ es LI entonces, $\alpha = \beta$ y $\alpha(1 + \rho) = 0$. Como $\rho \neq -1$ entonces, $0 = \alpha = \beta$.

Ejercicio 82 (Sección 4.3 página 108) Es saludable poner la matriz de Vandermonde en

forma gráfica como se muestra en el recuadro a la derecha. Denotemos por $v(x_1, x_2, \dots, x_n)$ al determinante de esta matriz. Denotemos

$$f(x_1, x_2, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

Veamos que $v(x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n)$.

$$\left(\begin{array}{cccc} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \cdots & \cdots & \cdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{array} \right)$$

Por inducción en n . Para $n = 1$ tenemos $f(x_1) = v(x_1) = 1$. Supongamos que está probado hasta $n - 1$. Pongamos $y = x_n$. Tenemos que probar que $v(x_1, \dots, x_{n-1}, y) = f(x_1, \dots, x_{n-1}, y)$. Para esto veamos ambos lados de la igualdad como polinomios en la variable y . Ambos polinomios tienen grado $n - 1$.

Haciendo la expansión de Laplace por la última columna vemos que el coeficiente principal de $v(x_1, \dots, x_{n-1}, y)$ es igual a $v(x_1, \dots, x_{n-1})$. Por otro lado

$$f(x_1, \dots, x_{n-1}, y) = \prod_{1 \leq i < j \leq n-1} (x_j - x_i) \prod_{1 \leq i \leq n-1} (y - x_i)$$

que tiene coeficiente principal $f(x_1, \dots, x_{n-1})$. Por hipótesis de inducción $v(x_1, \dots, x_{n-1}) = f(x_1, \dots, x_{n-1})$ y por lo tanto nuestros dos polinomios tienen el mismo coeficiente principal.

Además, las raíces de $v(x_1, \dots, x_{n-1}, y)$ son x_1, \dots, x_{n-1} porque evaluando y en estos valores obtenemos una matriz con dos columnas iguales. Es obvio que x_1, \dots, x_{n-1} son también las raíces de $f(x_1, \dots, x_{n-1}, y)$.

Luego, $v(x_1, \dots, x_{n-1}, y)$ y $f(x_1, \dots, x_{n-1}, y)$ son dos polinomios del mismo grado, con los mismos coeficientes principales y las mismas raíces y por lo tanto son polinomios iguales.

Ejercicio 83 (Sección 4.4 página 114) Las tres. Para la matriz A los bloques son

$M_1 = \{1\}$, $M_2 = \{2\}$ y $M_3 = \{3\}$. Para la matriz B los bloques son $M_1 = \{3\}$, $M_2 = \{2\}$ y $M_3 = \{3\}$. Para la matriz C los bloques son $M_1 = \{2\}$, $M_2 = \{3\}$ y $M_3 = \{1\}$ ya que $\alpha_{23} = \alpha_{21} = \alpha_{31} = 0$. Los determinantes de las tres matrices son iguales a abc .

Ejercicio 84 (Sección 4.4 página 114)

El aspecto es el del recuadro a la derecha. Aquí hemos denotado por * las entradas que pueden ser diferentes de cero. Las entradas con 0 son las que obligatoriamente tienen que ser cero. Además, hemos dividido con rectas los tres bloques de la matriz. El determinante de una matriz de este tipo es igual al producto de los determinantes de sus tres bloques diagonales.

*	*	0	0	0
*	*	0	0	0
*	*	*	0	0
*	*	*	*	*
*	*	*	*	*

Ejercicio 88 (Sección 5.1 página 134) Como G es un subgrupo el elemento neutro $1 \in G$ y los inversos de los elementos en G están en G . Como los inversos son únicos y $(a^{-1})^{-1} = a$ entonces la función $f: G \ni a \mapsto a^{-1} \in G$ es una biyección tal que f^2 es la identidad. Esto nos permite partir G en tres partes disjuntas $G = A \cup B \cup C$ tales que $f(A) = B$ y $C = \{a \in G \mid a^{-1} = a\}$. Por esto

$$\prod_{a \in A} a = \prod_{a \in A} (a^{-1})^{-1} = \left(\prod_{a \in B} a^{-1} \right)^{-1} = \left(\prod_{a \in B} a \right)^{-1}$$

y por lo tanto

$$\rho \stackrel{\text{def}}{=} \prod_{a \in G} a = \prod_{a \in A} a \prod_{a \in B} a \prod_{a \in C} a = \prod_{a \in C} a$$

Por otro lado si $a \in C$ entonces tiene que ser raíz del polinomio $x^2 - 1$ y en cualquier campo este polinomio tiene a lo más las raíces 1 y -1 (véase 5.5). Luego si $-1 \in G$ entonces, $\rho = -1$ y si $-1 \notin G$ entonces $\rho = 1$.

Ejercicio 89 (Sección 5.1 página 134) Sea \mathbb{K} un campo y G un subgrupo finito de $(\mathbb{K} \setminus 0, \bullet)$. Si $x \in G$ entonces al natural n más pequeño tal que $x^n = 1$ lo llamaremos orden de x . En este caso todos los elementos de $\langle x \rangle = \{1, x, x^2, \dots, x^{n-1}\}$ son diferentes y la función $f: (\langle x \rangle, \bullet) \ni x^i \mapsto i \in (\mathbb{Z}_n, +)$ es un isomorfismo de grupos. Luego, lo que hay que probar es que en G existe un elemento de orden $q = |G|$.

1. Sea F un subgrupo de G . Para $x \in G$ el conjunto $xF = \{xf \mid f \in F\}$ se le llama clase lateral de x . Tenemos que si $f \in F$ entonces $xF = F$ ya que F es un subgrupo. De

$$(y \in xF) \Rightarrow (y = xf) \Rightarrow (x = yf^{-1}) \Rightarrow (xF = yf^{-1}F) \Rightarrow (xF = yF)$$

$$(y \in xF \cap zF) \Rightarrow (xF = yF = zF)$$

obtenemos que las clases laterales forman una partición de G . La función $F \ni f \mapsto xf \in xF$ es la inversa de $xF \ni xf \mapsto f \in F$ y por lo tanto cualquier clase lateral tiene la misma cantidad de elementos que F . Luego $|F|$ es un divisor de $|G|$ y el cociente $|G| \div |F|$ es el número de clases laterales. Este hecho se conoce como **Teorema de Lagrange**. En particular, el orden de cada $x \in G$ es un divisor de

$$q = |G|.$$

2. Sea x de orden n y denotemos $\varphi(n)$ el número de elementos de orden n en $\langle x \rangle$. Como x tiene orden n entonces, $\varphi(n) \geq 1$. La función $\varphi(n)$ no depende de x pues cualesquiera dos x, y de orden n son tales que los grupos $\langle x \rangle$ y $\langle y \rangle$ son isomorfos. Luego, $\varphi(n)$ solo depende del natural n .
- Como cualquier $z \in \langle x \rangle$ tiene cierto orden que por el Teorema de Lagrange es un divisor de n y el número de elementos en $\langle x \rangle$ es n entonces, $\sum_{d|n} \varphi(d) = n$ donde la suma recorre todos los divisores de n . A φ se la conoce como **Función de Euler**.
3. Denotemos por $\rho(n)$ al número de elementos de orden n en nuestro grupo G . Como cualquier $z \in G$ tiene cierto orden que por el Teorema de Lagrange es un divisor de $|G| = q$ entonces, $\sum_{d|q} \rho(d) = q$ donde la suma recorre los divisores de q .
4. Si en G no hay un elemento de orden n entonces $\rho(n) = 0$. Si por el contrario $x \in G$ tiene orden n entonces, $\forall y = x^k \in \langle x \rangle$ tenemos que $y^n = (x^k)^n = (x^n)^k = 1$. Luego, todos los n elementos de $\langle x \rangle$ son raíces del polinomio $z^n - 1$. Como el polinomio $z^n - 1$ no puede tener más de n raíces en el campo \mathbb{K} (véase 5.5) entonces, todos los elementos de G de orden n están en $\langle x \rangle$ y por lo tanto $\rho(n) = \varphi(n)$. De aquí, para cualquier n se tiene que $\varphi(n) - \rho(n) \geq 0$.
5. De (2) y (3) tenemos que

$$0 = \sum_{d|q} \varphi(d) - \sum_{d|q} \rho(d) = \sum_{d|q} (\varphi(d) - \rho(d))$$

y como por (4) $\varphi(d) - \rho(d) \geq 0$ entonces, para cualquier d divisor de $q = |G|$ tenemos $\varphi(d) = \rho(d)$. En particular, $\rho(q) = \varphi(q) \geq 1$. Esto significa que en G hay un elemento de orden $q = |G|$.

Ejercicio 90 (Sección 5.1 página 135) Idea: todo lo demostrado para polinomios se traduce tal cual para los enteros. La única diferencia es que en lugar de usar el concepto de grado de un polinomio hay que usar el concepto de valor absoluto de un entero.

Ejercicio 91 (Sección 5.1 página 135) Si r es el máximo común divisor de p y q entonces los polinomios $p' = p/r$ y $q' = q/r$ no tienen factores comunes. Por el Teorema de Bezout existen polinomios α y β tales que $\alpha p' + \beta q' = 1$. La prueba concluye multiplicando esta igualdad por r .

Ejercicio 95 (Sección 5.1 página 138) Si $i < k$ entonces la suma es vacía y por lo tanto es cero. Si $i = k$ entonces, hay un solo sumando en el cual $i = j = k$ y este sumando es uno. Supongamos $i > k$. Haciendo los cambios de variables $t = j - k$ y

$n = i - k$ obtenemos

$$\sum_{j=k}^i (-1)^{j-k} \binom{j}{k} \binom{i}{j} = \sum_{t=0}^n (-1)^t \frac{(n+k)!}{k! t! (t-n)!} = \binom{n+k}{k} \sum_{t=0}^n (-1)^t \binom{n}{t}$$

y esta última expresión es cero ya que $\sum_{t=0}^n (-1)^t \binom{n}{t}$ es el binomio de Newton de $(x-1)^n$ evaluado en $x=1$.

Ejercicio 96 (Sección 5.1 página 138) De la prueba del Desarrollo de Taylor (5.12) sabemos que si $k \in \{0, \dots, n\}$ entonces, $a_k = \sum_{j=k}^n b_{kj} \alpha_j$ donde $b_{kj} = \binom{j}{k} (-x_0)^{j-k}$. Por el ejercicio anterior tenemos

$$\sum_{j=k}^n b_{kj} \beta_j = \sum_{i=k}^n \sum_{j=k}^i (-1)^{j-k} \binom{j}{k} \binom{i}{j} a_i x_0^{i-k} = \sum_{i=k}^n \delta_{ik} a_i x_0^{i-k} = a_k$$

Luego, $\sum_{j=k}^n b_{kj} \beta_j = a_j$ y como el sistema de ecuaciones lineales $a_k = \sum_{j=k}^n b_{kj} \alpha_j$ tiene solución única obtenemos que $\alpha_j = \beta_j$.

Ejercicio 97 (Sección 5.1 página 138) Formalmente (¡y en cualquier campo!) la derivada de $p(x) = \sum_{i=0}^n a_i x^i$ es $p^{(1)}(x) = \sum_{i=1}^n i a_i x^{i-1}$ y por lo tanto $p^{(j)}(x) = \sum_{i=j}^n \frac{i!}{(i-j)!} a_i x^{i-j} = j! \sum_{i=j}^n \binom{i}{j} a_i x_0^{i-j}$. Del ejercicio anterior obtenemos la prueba.

Ejercicio 98 (Sección 5.2 página 140) Las raíces del polinomio $z^n - 1$ son $x_k = \cos k \frac{2\pi}{n} + i \sin k \frac{2\pi}{n}$ para $k \in \{0, \dots, n-1\}$. Tenemos

$$x_k x_m = \cos \left((k+m) \frac{2\pi}{n} \right) + i \sin \left((k+m) \frac{2\pi}{n} \right) = \cos \ell \frac{2\pi}{n} + i \sin \ell \frac{2\pi}{n} = x_\ell$$

donde $\ell = (k+m) \bmod n$. Para el producto, estas raíces forman un grupo isomorfo a \mathbb{Z}_n .

Ejercicio 99 (Sección 5.2 página 140) Sean a, b y c tres lados de un triángulo. Por simetría solo tenemos que probar que $|a - b| \leq c$. Podemos suponer que $a \geq b$. Por la desigualdad del triángulo $b + c \geq a$ y de aquí $c \geq a - b$.

Ejercicio 100 (Sección 5.2 página 143) Para usar $\|(1-t^k)p(z_0)\| = (1-t^k)\|p(z_0)\|$

Ejercicio 101 (Sección 5.3 página 144) Para la propiedad 2 vemos que

$$\begin{aligned} \overline{(a+bi)+(c+di)} &= \overline{(a+c)+(b+d)i} = \\ &= (a+c)-(b+d)i = (a-bi)+(c-di) = \overline{(a+bi)}+\overline{(c+di)} \end{aligned}$$

Finalmente, para probar la propiedad 3 vemos que

$$\begin{aligned} \overline{(a+bi)(c+di)} &= \overline{(ac-bd)+(ad+bc)i} = \\ &= (ac-bd)-(ad+bc)i = (a-bi)(c-di) = \overline{(a+bi)}\overline{(c+di)} \end{aligned}$$

Ejercicio 102 (Sección 5.4 página 146) Tenemos $ab = ba$ por lo que $(a, b) \sim (a, b)$ lo que significa que la relación es reflexiva. De $ad = bc$ obtenemos $cb = da$. Esto significa que si $(a, b) \sim (c, d)$ entonces $(c, d) \sim (a, b)$ por lo que la relación es simétrica. Si $ad = bc$ y $cf = de$ entonces $adcf = bcde$ por lo que $af = be$. Esto significa que

si $(a, b) \sim (c, d)$ y $(c, d) \sim (e, f)$ entonces $(a, b) \sim (e, f)$ por lo que la relación es transitiva.

Ejercicio 103 (Sección 5.4 página 148) El campo de funciones racionales $\mathbb{Z}_2(x)$ contiene a \mathbb{Z}_2 y por lo tanto es de característica 2. Este campo tiene evidentemente un número infinito de elementos.

Ejercicio 104 (Sección 5.4 página 148) El campo de fracciones de un campo es él mismo.

Ejercicio 105 (Sección 6.1 página 150) Sean $(a, b), (a', b') \in \mathfrak{E} \oplus \mathfrak{F}$. Tenemos

$$\begin{aligned}(f \oplus g)((a, b) + (a', b')) &= (f \oplus g)(a + a', b + b') = (f(a + a'), g(b + b')) = \\ &= (f(a) + f(a'), g(b) + g(b')) = (f(a), g(b)) + (f(a'), g(b')) = \\ &= (f \oplus g)(a, b) + (f \oplus g)(a', b')\end{aligned}$$

con lo que tenemos la primera propiedad de linearidad. Para la segunda vemos que

$$(f \oplus g)(\lambda(a, b)) = (f \oplus g)(\lambda a, \lambda b) = \lambda(f(a), g(b)) = \lambda((f \oplus g)(a, b)).$$

Ejercicio 106 (Sección 6.1 página 150)

$$(f' \circ g')(a, b) = f'(g'(a, b)) = f'(a, g(b)) = (f(a), g(b)) = (f \oplus g)(a, b).$$

Ejercicio 107 (Sección 6.1 página 151) Si \mathfrak{E} es un subespacio y $f(\mathfrak{E}) \subseteq \mathfrak{E}$ entonces $\alpha f(\mathfrak{E}) \subseteq \alpha \mathfrak{E} = \mathfrak{E}$. Recíprocamente, si $\alpha f(\mathfrak{E}) \subseteq \mathfrak{E}$ entonces $f(\mathfrak{E}) \subseteq \alpha^{-1}\mathfrak{E} = \mathfrak{E}$. Luego, f y αf tienen los mismos subespacios invariantes.

Ejercicio 108 (Sección 6.1 página 151) Reflexividad: $f = \text{Id} \circ f \circ \text{Id}^{-1}$. Simetría: Si $f = \rho \circ g \circ \rho^{-1}$ entonces $g = \rho^{-1} \circ f \circ (\rho^{-1})^{-1}$. Transitividad: Si $f = \rho \circ g \circ \rho^{-1}$ y $g = \omega \circ h \circ \omega^{-1}$ entonces $f = \rho \circ \omega \circ h \circ \omega^{-1} \circ \rho^{-1} = (\rho \circ \omega) \circ h \circ (\rho \circ \omega)^{-1}$.

Ejercicio 109 (Sección 6.1 página 151) Sea L la matriz de f en la base A y M la matriz de cambio de base de A a B . En la base B la matriz del operador f es MLM^{-1} (véase la página 82). Si ρ es el operador de la cuya matriz en la base B es M entonces, $f = \rho \circ g \circ \rho^{-1}$.

Ejercicio 110 (Sección 6.1 página 152) Si g es reducible entonces tiene un par de subespacios g -invariantes no triviales complementarios \mathfrak{F} y \mathfrak{G} . Si ρ es un OL biyectivo entonces $\rho(\mathfrak{F})$ y $\rho(\mathfrak{G})$ son no triviales y complementarios. Además, si $f = \rho \circ g \circ \rho^{-1}$ entonces

$$f(\rho(\mathfrak{F})) = (\rho \circ g \circ \rho^{-1})(\rho(\mathfrak{F})) = (\rho \circ g)(\mathfrak{F}) \subseteq \rho(\mathfrak{F})$$

O sea, $\rho(\mathfrak{F})$ es f -invariante. Lo mismo ocurre con $\rho(\mathfrak{G})$ y concluimos que f es reducible.

Ejercicio 111 (Sección 6.1 página 152) Falso, un 0-deslizamiento en \mathbb{R}^2 es irreducible pero no es biyectivo.

Ejercicio 113 (Sección 6.2 página 157) Sea p el período de a . Si p es de grado cero entonces, $p = 1$ y por lo tanto $0 = p_h(a) = h^0(a) = a$.

Ejercicio 114 (Sección 6.2 página 157) Si $p = x$ entonces $p_h(a) = h(a)$ por lo que $p_h(a) = 0$ si y solo si $a \in \ker h$.

Ejercicio 118 (Sección 6.5 página 167) Sea f un operador lineal en $\text{End}(\mathfrak{E}/\mathfrak{F})$ y \mathfrak{G} un subespacio complementario a \mathfrak{F} . Definimos h como la función nula en \mathfrak{F} , como $h(a) = \mathfrak{G} \cap f(a + \mathfrak{F})$ en \mathfrak{G} y en todo el espacio por extensión lineal. Para ver que h es lineal solo hay que comprobar que es lineal en \mathfrak{G} . Esto es trivial porque en \mathfrak{G} la función h cumple $h = \text{nat}^{-1} \circ f \circ \text{nat}$ donde $a \xrightarrow{\text{nat}} a + \mathfrak{F}$ es el isomorfismo canónico entre el cociente y un complementario. Para ver que $f = h$ basta comprobarlo para los vectores a que estan en \mathfrak{G} . Pero $f(a + \mathfrak{F}) = h(a) + \mathfrak{F} = \mathfrak{F}$ ya que $f = \text{nat} \circ h \circ \text{nat}^{-1}$. Para calcular el núcleo comprobamos que para cualquier vector a se cumple que

$$(h(a) = \mathbb{O}(a)) \Leftrightarrow (h(a) + \mathfrak{F} = \mathfrak{F}) \Leftrightarrow (h(a) \in \mathfrak{F}).$$

Ejercicio 119 (Sección 6.5 página 168) Supongamos que $\langle b \rangle_h \cap \mathfrak{F} = \{0\}$. Sea $p = \text{per}_h(b + \mathfrak{F})$. Tenemos que $p_h(b + \mathfrak{F}) = p_h(b) + \mathfrak{F} = \mathfrak{F}$ y por lo tanto $p_h(b) \in \mathfrak{F}$. Por hipótesis, $p_h(b) = 0$. Luego, $p \vdash \text{per}_h(b)$.

Supongamos $\text{per}_h(b) \dashv \text{per}_h(b + \mathfrak{F})$ y sea $q_h(b) \in \mathfrak{F}$. Tenemos, $q_h(b + \mathfrak{F}) = q_h(b) + \mathfrak{F} = \mathfrak{F}$ y por lo tanto $q \vdash \text{per}_h(b + \mathfrak{F}) \vdash \text{per}_h(b)$. Luego, $q_h(b) = 0$.

Ejercicio 120 (Sección 6.5 página 170) Es fácil ver que $\langle A \cup B \rangle_h \supseteq \langle A \rangle_h + \langle B \rangle_h = \mathfrak{F} + \mathfrak{G} = \mathfrak{E}$ y por lo tanto $A \cup B$ es h -generador. Comprobemos que $A \cup B$ es h -independiente. Observese que $A \cap B = \emptyset$ ya que 0 no está en ningún conjunto h -independiente y $A \cap B \subseteq \mathfrak{F} \cap \mathfrak{G} = \{0\}$. Luego una h -combinación de $A \cup B$ es de la forma

$$p_h(a_1) + \cdots + q_h(a_n) + r_h(b_1) + \cdots + s_h(b_m)$$

donde $A = \{a_1, \dots, a_n\}$ y $B = \{b_1, \dots, b_m\}$. Si esta h -combinación es igual cero entonces

$$\mathfrak{F} = \langle A \rangle_h \ni p_h(a_1) + \cdots + q_h(a_n) = -r_h(b_1) - \cdots - s_h(b_m) \in \langle B \rangle_h = \mathfrak{G}$$

y por lo tanto,

$$p_h(a_1) + \cdots + q_h(a_n) = r_h(b_1) + \cdots + s_h(b_m) = 0$$

y como A y B son h -independientes deducimos que

$$p_h(a_1) = \cdots = q_h(a_n) = r_h(b_1) = \cdots = s_h(b_m) = 0.$$

Ejercicio 122 (Sección 6.5 página 173) Demostremos primero que $\langle a \rangle_h \cap \langle b \rangle_h = \{0\}$. Tenemos para cualquier polinomio r que

$$(p_h r_h(a) = r_h(p_h(a))) = 0 \Rightarrow (r_h(a) \in \ker p_h)$$

y por lo tanto $\langle a \rangle_h \subseteq \ker p_h$. Análogamente $\langle b \rangle_h \subseteq \ker q_h$. Por el Lema de Descomposición de Núcleos (6.15) sabemos que $\ker p_h \cap \ker q_h = \{0\}$ y por lo tanto $\langle a \rangle_h \cap \langle b \rangle_h = \{0\}$.

De $r_h(a + b) = r_h(a) + r_h(b)$ concluimos que $\langle a + b \rangle_h \subseteq \langle a \rangle_h + \langle b \rangle_h$. Demostraremos que ambos subespacios tienen la misma dimensión (finita) y por lo tanto son iguales.

Como $\langle a \rangle_h \cap \langle b \rangle_h = \{0\}$ el conjunto $\{a, b\}$ es una h -base de $\langle a \rangle_h + \langle b \rangle_h$ y por 6.26 $\dim(\langle a \rangle_h + \langle b \rangle_h)$ es la suma de los grados de p y q . De la misma manera que en la

prueba de 6.19 podemos demostrar que $\text{per}_h(\mathbf{a} + \mathbf{b}) = \mathbf{p}\mathbf{q}$. Por 6.26 $\dim \langle \mathbf{a} + \mathbf{b} \rangle_h$ es igual al grado de $\mathbf{p}\mathbf{q}$ o sea, igual a la suma de los grados de \mathbf{p} y \mathbf{q} .

Ejercicio 123 (Sección 6.5 página 173)

Sean $\mathbf{p}, \mathbf{q}, \dots, \mathbf{r}$ los factores irreducibles del polinomio mínimo del OL \mathbf{h} . Organizaremos los polinomios en el tipo de una descomposición de \mathbf{h} en componentes radicales cíclicas en una tabla donde $\ell_i \geq \ell_{i+1}$, $m_i \geq m_{i+1}$ y $n_i \geq n_{i+1}$. Para hacer nuestra tabla cuadrada puede ser que necesitemos hacer algunos de los exponentes iguales a cero. O sea, hacia la derecha de la tabla pueden haber polinomios iguales a 1.

\mathbf{p}^{ℓ_1}	\mathbf{p}^{ℓ_2}	\dots	\mathbf{p}^{ℓ_t}
\mathbf{q}^{m_1}	\mathbf{q}^{m_2}	\dots	\mathbf{q}^{m_t}
\vdots	\vdots	\dots	\vdots
\mathbf{r}^{n_1}	\mathbf{r}^{n_2}	\dots	\mathbf{r}^{n_t}

Cada entrada diferente de 1 de esta tabla se corresponde con una componente radical cíclica de \mathbf{h} . Los renglones de la tabla corresponden a las componentes radicales de \mathbf{h} . Para cada columna $i \in \{1, \dots, t\}$ denotemos por \mathbf{g}_i a la suma directa de las componentes irreducibles de \mathbf{h} correspondientes a las entradas de la columna. Por el 6.11 el polinomio mínimo de \mathbf{g}_i es igual a $\mathbf{p}^{\ell_i} \mathbf{q}^{m_i} \dots \mathbf{r}^{n_i}$ y por lo tanto el polinomio mínimo de cada \mathbf{g}_i es un múltiplo del polinomio mínimo de \mathbf{g}_{i+1} . Por el ejercicio anterior los \mathbf{g}_i son cíclicos. Denotando $f_i = g_{t-i}$ obtenemos la descomposición deseada.

Ejercicio 124 (Sección 6.5 página 173) El tipo de tales descomposiciones está formado por los productos de los polinomios en las columnas de la tabla del ejercicio anterior. De la unicidad del tipo de las descomposiciones en componentes irreducibles se desprende la unicidad de la tabla salvo reordenamiento de los renglones.

Ejercicio 125 (Sección 6.6 página 175) Para probar esto, demostraremos que la traza es uno de los coeficientes del polinomio característico. Como el polinomio característico es invariante bajo cambios de base, entonces lo mismo es válido para sus coeficientes.

Sea $\mathbf{A} = \alpha_{NN}$ una matriz. Su polinomio característico es

$$\det(x\mathbb{I} - \mathbf{A}) = \sum_{\sigma \in S_N} \text{sgn } \sigma \prod_i (x - \alpha_{ii}) \prod_k (-\alpha_{k\sigma_k})$$

donde el primer producto recorre todos los índices $i \in \mathbb{N}$ tales que $\sigma_i = i$ y el segundo los índices $k \in \mathbb{N}$ tales que $\sigma_k \neq k$.

Como cualquier permutación σ que no sea la identidad tiene al menos dos k tales que $\sigma_k \neq k$, entonces todos los sumandos correspondientes a permutaciones que no sean la identidad son polinomios de grado menor o igual que $|N| - 2$. Luego, los coeficientes en x^n y x^{n-1} del polinomio característico coinciden con los correspondientes coeficientes del polinomio

$$\prod_{i \in N} (x - \alpha_{ii}) = x^n - \left(\sum_{i \in N} \alpha_{ii} \right) x^{n-1} + \dots$$

y así vemos, que el coeficiente en x^{n-1} es igual a la traza multiplicada por -1 .

Ejercicio 126 (Sección 6.6 página 178) Si el polinomio mínimo es un producto de diferentes polinomios de grado 1 entonces las componentes radicales son homotecias ya que para cualquier vector \mathbf{a} en el subespacio radical de tipo $x - \lambda$ se cumple que $\mathbf{h}(\mathbf{a}) - \lambda\mathbf{a} = \mathbf{0}$. Las homotecias tienen matrices diagonales en cualquier base y la suma directa de operadores diagonalizables es diagonalizable.

Si en cierta base \mathbf{N} el OL \mathbf{h} tiene matriz diagonal $\alpha_{\mathbf{NN}}$ entonces, podemos denotar $\mathbf{N}_\lambda = \{\mathbf{i} \in \mathbf{N} \mid \alpha_{ii} = \lambda\}$. También denotemos \mathbf{h}_λ a la restricción de \mathbf{h} al subespacio generado por \mathbf{N}_λ . Los operadores \mathbf{h}_λ son homotecias porque en la base \mathbf{N}_λ su matriz es $\lambda\mathbb{I}$ y por lo tanto el polinomio mínimo de \mathbf{h}_λ es $x - \lambda$. Tenemos que \mathbf{h} es la suma directa de los \mathbf{h}_λ y que su polinomio mínimo es el producto de los $x - \lambda$ ya que todos estos son diferentes.

Ejercicio 127 (Sección 6.6 página 178) Si $\lambda_1, \dots, \lambda_n$ son n diferentes valores propios de un operador lineal en dimensión n entonces su polinomio característico es $(x - \lambda_1) \cdots (x - \lambda_k)$ y por el Teorema de Hamilton-Caley-Frobenius este coincide con su polinomio mínimo. Por el ejercicio anterior el operador es diagonalizable.

Ejercicio 128 (Sección 6.6 página 178) Supongamos que \mathbf{h} es triangulable entonces existe una base $\mathbf{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ en la cual la matriz de \mathbf{h} es triangular. Denotemos $\mathfrak{E}_k = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$, evidentemente $\{\mathbf{0}\} \subseteq \mathfrak{E}_1 \subseteq \cdots \subseteq \mathfrak{E}_n = \mathfrak{E}$ y $\dim \mathfrak{E}_k = k$. Además $\{\mathbf{h}(\mathbf{v}_1), \dots, \mathbf{h}(\mathbf{v}_k)\} \subseteq \mathfrak{E}_k$ debido a la forma triangular de la matriz de \mathbf{h} . Luego, los \mathfrak{E}_k son invariantes.

Recíprocamente supongamos que una cadena de subespacios invariantes $\{\mathbf{0}\} \subseteq \mathfrak{E}_1 \subseteq \cdots \subseteq \mathfrak{E}_n = \mathfrak{E}$ cumple que $\dim \mathfrak{E}_k = k$. Escojamos $\mathbf{v}_1 \in \mathfrak{E}_1$ y para $k \in \{2, \dots, n\}$ escojamos $\mathbf{v}_k \in \mathfrak{E}_k \setminus \mathfrak{E}_{k+1}$. El conjunto $\mathbf{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es LI ya que ningún \mathbf{v}_k es combinación lineal de los anteriores. Luego $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ es una base de \mathfrak{E}_k .

Como \mathfrak{E}_k es invariante entonces por 6.2, en la base \mathbf{B} la matriz de \mathbf{h} tiene las entradas cuyas columnas están indexadas por $\{1, \dots, k\}$ y cuyos renglones están indexados por $\{k+1, \dots, n\}$; iguales a cero. En particular, $\forall k$ en la columna k solo pueden ser distintos de cero las entradas correspondientes a los renglones $\{1, \dots, k\}$.

Ejercicio 129 (Sección 6.6 página 178) Si \mathbf{f} y \mathbf{g} son triangulables entonces en cierta base la matriz de \mathbf{f} es diagonal con dos bloques y cada uno de ellos triangular. Luego, \mathbf{f} es triangulable.

Recíprocamente, sea $\{\mathbf{0}\} \subseteq \mathfrak{E}_1 \subseteq \cdots \subseteq \mathfrak{E}_n = \mathfrak{E}$ la cadena de subespacios invariantes que hace que $\mathbf{h} \stackrel{\text{def}}{=} \mathbf{f} \oplus \mathbf{g}$ sea triangulable (véase el ejercicio anterior).

Sea $\mathfrak{E} = \mathfrak{F} \oplus \mathfrak{G}$ la descomposición en subespacios invariantes de tal manera que \mathbf{f} es la restricción de \mathbf{h} al subespacio \mathfrak{F} . Sea π la proyección a \mathfrak{F} a lo largo de \mathfrak{G} , o sea, si $\mathbf{x} = \mathbf{y} + \mathbf{z}$ con $\mathbf{y} \in \mathfrak{F}$ y $\mathbf{z} \in \mathfrak{G}$ entonces $\pi(\mathbf{x}) = \mathbf{y}$.

Para $k \in \{1, \dots, n\}$ denotemos $\mathfrak{F}_k = \pi(\mathfrak{E}_k)$. Los subespacios \mathfrak{F}_k cumplen que

$$\{\mathbf{0}\} \subseteq \mathfrak{F}_1 \subseteq \cdots \subseteq \mathfrak{F}_n = \mathfrak{F}.$$

Probemos que los espacios \mathfrak{F}_k son invariantes. Para esto sea $x = y + z$ con $y \in \mathfrak{F}$ y $z \in \mathfrak{G}$. Por linearidad tenemos $h(x) = h(y) + h(z)$. Como \mathfrak{F} y \mathfrak{G} son invariantes entonces $h(y) \in \mathfrak{F}$ y $h(z) \in \mathfrak{G}$ y por definición de π tenemos que

$$\pi(h(x)) = h(y) = h(\pi(x))$$

esto quiere decir que π commuta con h . Luego

$$h(\mathfrak{F}_k) = h(\pi(\mathfrak{E}_k)) = \pi(h(\mathfrak{E}_k)) \subseteq \pi(\mathfrak{E}_k) = \mathfrak{F}_k$$

y por lo tanto los \mathfrak{F}_k son invariantes.

Por otro lado, como $\dim \mathfrak{E}_k = 1 + \dim \mathfrak{E}_{k-1}$ tenemos dos alternativas posibles

$$\mathfrak{F}_k = \mathfrak{F}_{k-1} \quad \text{o} \quad \dim \mathfrak{F}_k = 1 + \dim \mathfrak{F}_{k-1}$$

y si en la cadena de subespacios invariantes

$$\{0\} \subseteq \mathfrak{F}_1 \subseteq \cdots \subseteq \mathfrak{F}_n = \mathfrak{F}$$

eliminamos las repeticiones obtendremos otra cadena de subespacios invariantes

$$\{0\} \subseteq \mathfrak{F}'_1 \subseteq \cdots \subseteq \mathfrak{F}'_m = \mathfrak{F}$$

en la cual se cumple que $\dim \mathfrak{F}'_k = k$. Por el ejercicio anterior esto quiere decir que f es triangulable.

Ejercicio 130 (Sección 6.6 página 178) Por Teorema de Descomposición en Componentes Radicales Cíclicas (6.36) y el ejercicio anterior solo hay que probarlo para operadores cílicos radicales. Sea h cílico radical con polinomio mínimo p^n . La única cadena de subespacios invariantes posible es por 6.43 la siguiente

$$\{0\} \subseteq \ker p_h \subseteq \cdots \subseteq \ker p_h^k \subseteq \cdots \ker p_h^n = \mathfrak{E}.$$

Por 6.42 y 6.23 tenemos que $\dim \ker p_h^k$ es igual al grado de p^k . Luego, esta cadena cumple los requisitos del ejercicio 128 si y solo si el grado de p es 1.

En los complejos, todo polinomio irreducible es de grado 1.

Ejercicio 131 (Sección 6.7 página 184) Sean f y g del mismo tipo. Entonces, usando la Forma normal canónica (6.57) vemos que existen bases del espacio en las cuales f y g tienen la misma matriz. Esto quiere decir que en la misma base f y g tienen matrices A y CAC^{-1} donde C es la matriz adecuada de cambio de bases. La matriz C es la matriz de un OL invertible h por lo que $f = h \circ g \circ h^{-1}$.

Recíprocamente, supongamos que $f = h \circ g \circ h^{-1}$. Si A , B y C son las matrices de f , g y h respectivamente en cierta base entonces $A = CBC^{-1}$. Luego, en otra base (definida por el cambio de base C), la matriz de g es A . Por lo tanto, existen bases en las cuales las matrices de g y f son las mismas. Como el tipo de un operador se puede calcular por su matriz sin usar la base en la que está definida entonces, los tipos de f y g son el mismo.

Glosario

de conceptos

Álgebra.	Espacio vectorial con un producto de vectores asociativo, distributivo, con elemento neutro y que commuta con el producto por escalares	70, 79, 99
Álgebra comutativa.	Álgebra en la cual el producto de vectores es comutativo	71, 155
Anillo.	Conjunto con dos operaciones binarias denotadas por $+$ y \bullet que es grupo abeliano para la suma, que el producto es asociativo con elemento neutro y distributivo respecto a la suma	7, 12, 70
Anillo comutativo.	Anillo en el cual el producto es comunitativo	7, 15, 22, 130, 146
Anulador de un conjunto de vectores.	La intersección de los anuladores de todos los vectores en el conjunto	157
Anulador de un OL.	El anulador de todo el espacio	157
Anulador de un vector.	El conjunto de polinomios p tales que $p_h(a) = 0$	157
Argumento de un complejo.	El ángulo que forma el vector \vec{Oz} con el eje real del plano complejo	139
Asociatividad.	Propiedad de algunas operaciones binarias que consiste en que $a \circ (b \circ c) = (a \circ b) \circ c$ para cualesquiera elementos a , b y c del conjunto en el cual está definida la operación	3, 14, 20, 69, 78, 130
Asociatividad del producto por escalares.	Axioma de espacio vectorial: $(\beta a) = (\alpha\beta) a$	28
Automorfismo.	Endomorfismo biyectivo	13
Automorfismo de Frobenius.	La función $\mathbb{K} \ni x \mapsto x^p \in \mathbb{K}$ donde \mathbb{K} es un campo finito de característica $p > 0$	21
Automorfismo semilineal.	Transformación semilineal biyectiva de un espacio en si mismo	88
Base.	Conjunto de vectores que es generador y LI. Conjunto generador minimal. Conjunto LI maximal	40, 45, 49, 62, 72, 80, 117
Base canónica.	El conjunto de \mathbb{N} -adas $\{e_i \mid i \in \mathbb{N}\}$ donde la j -ésima coordenada de e_i es el delta de Kronecker δ_{ij}	43, 77, 82
Base de las columnas.	Conjunto de columnas diferentes que son una base del espacio de columnas	116
Base de los renglones.	Conjunto de renglones diferentes que son una base del espacio de renglones	116
Base de una matriz.	Submatriz no singular maximal por contención	118, 124
Binomio de Newton.	Fórmula para expandir $(x + y)^n$	21, 138
Cadena.	Subconjunto de un conjunto ordenado que está totalmente ordenado	*, 61
Cambio de índices.	Biyección mediante la cual los índices de una \mathbb{N} -ada (o columnas, o renglones) obtienen nuevos nombres. Es la operación en que una biyección $\omega : \mathbb{N} \rightarrow L$ se le aplica a las columnas de una matriz α_{MN} pa-	

ra obtener la matriz $\beta_{ML} = \alpha_{M\omega(N)}$ que cumple que $\beta_{ij} = \alpha_{i\omega^{-1}(j)}$. Análogamente se definen los cambios de índices de los renglones 104

Campo. Anillo commutativo en el cual todo elemento diferente de cero tiene inverso 8, 16, 17

Campo algebraicamente cerrado.

Campo en el cual todo polinomio de grado al menos uno tiene una raíz. Campo el cual los polinomios irreducibles son todos de grado uno 10, 47, 138, 145

Campo de fracciones.

Entre las fracciones de un dominio de integridad se define la suma y el producto exactamente de la misma manera que se definen estas operaciones en \mathbb{Q} . Para estas operaciones el conjunto de fracciones es un campo 147

Campo ordenado.

Campo con una relación de orden en el cual todo elemento es comparable con el cero. Los elementos mayores que cero se les llama positivos y los menores que cero negativos. Los axiomas que debe cumplir la relación de orden son:

1. El opuesto de un positivo es negativo,
2. El opuesto de un negativo es positivo,
3. La suma de positivos es positiva,
4. El producto de positivos es positivo.

Los campos \mathbb{R} y \mathbb{Q} están ordenados. Cualquier campo ordenado tiene que ser de característica cero. Además, \mathbb{C} no se puede ordenar *, 9, 106

Campo primo. Campo cuyo único subcampo es el mismo 17

Característica de un campo.

Es cero si el campo contiene como subcampo a \mathbb{Q} . Es igual al número primo p si el campo contiene como subcampo a \mathbb{Z}_p 19, 97, 105, 148

Cerradura lineal. El conjunto de todas las combinaciones lineales de un conjunto de vectores 37, 45, 49

Ciclo.

Permutación σ del grupo simétrico de N que cambia un conjunto $\{x_0, \dots, x_{n-1}\} \subseteq N$ según la regla $\sigma(x_i) = x_{i+1 \text{ mod } n}$ y deja todos los demás elementos de N fijos 94, 107, 115

Ciclos disjuntos.

Ciclos tales que los conjuntos de elementos que ellos no dejan fijos no se intersectan 94

Clases de equivalencia. Si \approx es una relación de equivalencia en A entonces, las clases de equivalencia son los subconjuntos de A para los cuales $a \approx b$ si y solo si a y b están en la misma clase * , 55, 62

Coalineación. Función biyectiva de un espacio vectorial en si mismo que tal que f y f^{-1} preservan subespacios afines. 89

Cociente.

Al efectuar la división con resto de un elemento p de un anillo commutativo (\mathbb{Z} , $\mathbb{K}[x]$) entre otro q obtenemos la igualdad $p = cq + r$. Al elemento c se le llama cociente de la división de p entre q 131

Codominio de una función.

Sea $f : A \rightarrow B$ una función. Al conjunto B se le llama codominio de la función f 11, 83

Coeficiente principal. El coeficiente diferente de cero de un polinomio que multiplica a la potencia más grande de la variable 129

Coeficientes de un polinomio.

(véase polinomio) 129

Cofactor. De la entrada α_{ij} es el escalar $\operatorname{sgn} \omega \det \alpha_{N \setminus i \omega(N \setminus j)}$ donde ω es cualquier permutación de N tal que $\omega(j) = i$ 106

Columna de una matriz.

Dada una matriz α_{NM} y $j \in M$ a la N -ada α_{Nj} (j está fijo, los elementos de N varían) se le llama j -ésima columna de la matriz α_{NM} 33, 76

Combinación lineal.

Los vectores de la forma $\sum_{i \in N} \alpha_i i$ (donde solo un número finito de los α_i es diferente de

cero)	35, 49	3
Complemento algebraico.		
Cofactor	106	66
Componente radical.		
Restricción del OL a un subespacio radical maximal	161	29
Composición de funciones.		
Operación entre dos funciones que consiste en aplicar primero una función y después la otra	14, 44, 69, 78, 93	143
Conjunto acotado inferiormente.		
Conjunto que tiene una cota inferior	* ,	
Conjunto acotado superiormente.		
Conjunto que tiene una cota superior ..*		
Conjunto generador.	Conjunto de vectores cuya cerradura lineal es todo el espacio	38, 45
Conjunto h-independiente.	Conjunto de vectores no nulos tales que los sumandos de cualquier h-combinación nula son nulos	164
Conjunto inductivamente ordenado.		
Conjunto ordenado no vacío dentro del cual cada cadena tiene cota superior ..	61	
Conjunto LD.	Acrónimo de conjunto linealmente dependiente	39
Conjunto LI.	Acrónimo de conjunto linealmente independiente	39
Conjunto linealmente dependiente.		
Conjunto de vectores que no es LI ..	39	
Conjunto linealmente independiente.		
Conjunto de vectores A tal que $\forall a \in A$ se cumple que $\langle A \setminus a \rangle \neq \langle A \rangle$	39, 116	
Conjunto ordenado.	Conjunto en el cual está definida una relación de orden	*
		61
Conjunto totalmente ordenado.		
Conjunto ordenado en el cual dos elementos cualesquiera son comparables ..*		
Commutatividad.	Propiedad de algunas operaciones binarias que consiste en que $a \circ b = b \circ a$	
✓ elementos a y b del conjunto en el cual está definida la operación	3	
Contracción.	Homotecia cuyo escalar es un real mayor que 0 y menor que 1 ..	
Coordenada.		
De la n -ada (a_1, a_2, \dots, a_n) es alguno de los a_i . De la N -ada a_N es alguno de los a_i para $i \in N$	29	
Coordinatización.		
Dada una base N de \mathfrak{F} , es el isomorfismo $\mathfrak{F} \ni \sum_{i \in N} \alpha_i i \mapsto \alpha_N \in \mathbb{K}^{\{N\}}$	46, 80	
Cota inferior.	Una cota inferior de un subconjunto A de un conjunto ordenado B es un elemento b de B tal que cualquier elemento de A es mayor o igual que b	*
		143
Cota superior.	Una cota superior de un subconjunto A de un conjunto ordenado B es un elemento b de B tal que cualquier elemento de A es menor o igual que b	*
		61
Defina la multiplicidad de una raíz.		
		133
Delta de Kronecker.		
Función δ_{ij} de dos variables que toma valor 1 si $i = j$ y toma valor 0 si $i \neq j$	138, 79, 99	
Desarrollo de Taylor.		
Expresión de una función como serie de potencias $f(x) = \sum a_i (x - x_0)^i$. El coeficiente de la serie a_i es la i -ésima derivada de f evaluada en el punto x_0	137, 142	
Determinante.	El determinante de α_{NN} es	
	$\sum_{\sigma \in S_N} \operatorname{sgn} \sigma \prod_{i \in N} \alpha_{i\sigma_i}$	98, 21, 124
Determinante de un OL.		
Determinante de su matriz en una base. No depende de la base escogida	111	
Diagrama.	Reperesentación gráfica de una familia de conjuntos y de funciones entre ellos mediante la utilización de flechas	
		82
Diagrama conmutativo.		

Diagrama en el cual cualesquiera dos caminos dirigidos (que siguen el orden de las flechas) de un conjunto a otro representan dos descomposiciones de la misma función como composición de las funciones de las flechas de los caminos	82	incógnito	119
Dilatación. Homotecia cuyo escalar es un real mayor que 1	66	Ecuación matricial. Ecuación $\mathbf{AX} = \mathbf{B}$ donde las matrices \mathbf{A} y \mathbf{B} son conocidas y la matriz \mathbf{X} es incógnita	126
Dimensión.		Elemento maximal. Elemento tal que cualquier otro no es mayor que el	*, 40, 61, 118
El cardinal de cualquier base.....	42, 50, 55, 84	Elemento minimal. Elemento tal que cualquier otro no es menor que el *, 40	
Dimensión de un OL. Dimensión del espacio donde está definido el operador	150	Elementos comparables. Dos elementos a y b de un conjunto ordenado para los cuales $a \preceq b$ o $b \preceq a$ o los dos	*, 40
Dimensión de un subespacio afín. Si E es una traslación del subespacio \mathfrak{E} entonces la dimensión de E es la dimensión de \mathfrak{E}	55	Endomorfismo. Morfismo cuyo dominio y codominio coinciden	13
Distributividad.		Entensión lineal de una función.	
Relación de una operación binaria \diamond con respecto a otra \circ que consiste en que las igualdades		Extensión que es transformación lineal. Si la función tiene como dominio una base, entonces la extensión lineal existe y es única	72
$a \diamond (b \circ c) = (a \diamond b) \diamond (a \diamond c)$			
$(b \circ c) \diamond a = (b \diamond a) \circ (c \diamond a)$			
se cumplen para cualesquiera elementos a, b y c del conjunto en el cual está definida la operación	5, 20, 69, 78	Entrada de una matriz.	
Distributividad del producto por escalares.		Coordenada de una NM-ada	33
Axioma de espacio vectorial:		Epimorfismo. Morfismo sobreyectivo	13
$\alpha(a + b) = \alpha a + \alpha b$ y $(\alpha + \beta)a = \alpha a + \beta a$		Escalar. Elemento del campo (¡la escala!) sobre el cual está definido el espacio vectorial	28, 28
Divisor. Para dos polinomios p y q se dice que q es un divisor de p si $\exists c$ tal que $q = cp$	132	Espacio cociente. Si \mathfrak{E} es un subespacio de \mathfrak{F} entonces, el espacio cociente $\mathfrak{F}/\mathfrak{E}$ es el conjunto de todos los subespacios afines paralelos a \mathfrak{E} dotado con la suma de subespacios afines y el producto de subespacios afines por escalares	56, 85
Dominio de integridad. Anillo commutativo en el cual el producto de elementos diferentes de cero es diferente de cero	15, 19, 147	Espacio de columnas. El espacio generado por las columnas de una matriz	116
Dominio de una función. Sea $f: A \rightarrow B$ una función. Al conjunto A se le llama dominio de la función f . 11, 83		Espacio de renglones. El espacio generado por los renglones de una matriz	116
Ecuación lineal.		Espacio vectorial. Grupo abeliano con un producto por escalares distributivo, asociativo y con neutro	28, 46, 54, 56, 69
Ecuación $\alpha_N x_N = \beta_N$ donde las N -adas α_N y β_N son conocidos y el vector x_N es		Extensión de una función. Si f es una restricción de g entonces g es	

una extensión de f	72	
Factor de un polinomio.	Divisor	
mónico de grado al menos 1	133	
Factor propio.		
Factor mónico diferente al polinomio	135	
Forma polar de un complejo.		
Es la expresión $r(\cos \varphi + i \sin \varphi)$ donde r es el módulo de z y φ es el argumento de z	139	
Fórmula multinomial.	Fórmula	
para expandir la n -sima potencia de una suma	21	
Fracción.	Pareja de elementos	
de un dominio de integridad (por ejemplo $\mathbb{K}(x)$ o \mathbb{Z}) que se denota por a/b .		
Dos fracciones a/b y c/d se consideran iguales si $ad = bc$	146, 8	
Función.	Informalmente	
es una regla o procedimiento mediante el cual para cada elemento de un conjunto a podemos obtener (calcular) otro único elemento $f(a)$ de otro conjunto y que llamamos la imagen de a mediante la función f . Formalmente, una función $f : A \rightarrow B$ es un subconjunto del producto cartesiano $f \subseteq A \times B$ que cumple que para cualquier $a \in A$ existe una única pareja $(a, b) \in f$*		
Función antipodal.	La función	
$x \mapsto -x$	66	
Función biyectiva.	Función inyectiva y sobreyectiva	* , 93
Función continua.	Función continua en todos los puntos de su dominio	140
Función continua en un punto.	La función f es continua en z_0 si $\forall \varepsilon > 0 \exists \delta$ tal que $\ z - z_0\ < \delta \Rightarrow \ f(z) - f(z_0)\ < \varepsilon$	140
Función de evaluación.		
De un polinomio $p(x)$ es la función:		
$p : \mathbb{K} \ni b \mapsto p(b) \in \mathbb{K}$	131	
Función identidad.	La que a todo elemento le corresponde el mismo	66
Función inversa.		
La inversa de una función $f : A \rightarrow B$ es una función g tal que $f \circ g = \mathbb{I}_B$ y $g \circ f = \mathbb{I}_A$. Una función tiene inversa si y solo si esta es biyectiva	* , 44	
Función inyectiva.	Cada elemento de la imagen tiene una única preimagen	* , 83
Función nula.	La que a todo elemento le corresponde el cero	66
Función racional.		
Fracción de dos polinomios	148	
Función sobreyectiva.		
Función tal que su imagen coincide con su codominio	* , 84	
Funcional.	Función de un espacio vectorial en su campo	109
Funcional bilineal.		
Funcional lineal en sus dos variables	109	
Funcional lineal.		
Funcional que es una TL	109	
Funcional lineal en una variable.		
Función de muchas variables con imágenes en un campo que para cualesquiera valores fijos de las demás variables determina un funcional lineal de la variable que se trata	109	
Funcional multilineal.		
Funcional lineal en todas sus variables	109	
Grado de un polinomio.	La potencia mayor de la variable con coeficiente diferente de cero. Si el grado es cero el polinomio es un elemento del campo. No está definido el grado del polinomio cero	129
Grupo.	Conjunto con una operación binaria asociativa que tiene elemento neutro y en el cual todo elemento tiene inverso	7, 11, 14, 57, 71, 93, 134
Grupo abeliano.	Grupo en el cual la operación es commutativa	7
Grupo alternante.		
El subgrupo (del grupo simétrico) de to-		

das las permutaciones pares	97	
Grupo general lineal.	El grupo de automorfismos de un espacio vectorial. El grupo de operadores lineales no singulares	71
Grupo simétrico.	El grupo de todas las permutaciones con la operación de composición	93
h-base.	Conjunto h -independiente y h -generador. Todo OL h tiene h -bases	165
h-cerradura.	Conjunto de todas las h -combinaciones	163
h-combinación.	Vector de la forma $p_h(v_1) + q_h(v_2) + \dots + r_h(v_n)$ donde p, q, ..., r son polinomios	162
h-generador.	Conjunto de vectores que h -genera a todo el espacio	163
Homotecia.	Transformación lineal que consiste en la multiplicación por un escalar	66
Ideal.	Subconjunto I de un anillo A tal que	
1. $\forall p, q \in I \quad p + q \in I,$		
2. $\forall p \in I \quad \forall r \in A \quad rp \in I$	134	
Ideal principal.	Ideal formado por todos los múltiplos de un polinomio	134
Idempotencia.	Propiedad de una función que consiste en que $f \circ f = f^2 = f$	37
Igualdad de Moivre.	Fórmula para calcular la n -ésima potencia de un número complejo en forma polar	139
Imagen de un elemento.	(véase: Función)	*
Imagen de una función.	El conjunto de los elementos del codominio que tienen alguna preimagen. Se denota por Im f	*, 82
Imagen de una matriz.	La imagen de su transformación lineal. Su espacio de columnas	122
Ínfimo.	El máximo de las cotas superiores	*, 143
Ínfimo de un conjunto de reales.		
El número máximo x tal que $x \leq a$ para cualquier a ∈ A . Para el ínfimo x de A siempre existe una sucesión $\{a_k\}$ de elementos de A cuyo límite es x		143
Inmersión.	Restricción de la función identidad a un subconjunto	67, 83
Inverso.	De un elemento a para la operación binaria o es un elemento b que cumple que $a \circ b = b \circ a = e$ para cualquier elemento a del conjunto en el cual está definida la operación. Si un elemento tiene inverso entonces este es único. En los anillos y campos es el inverso para el producto	5
Isomorfismo.	Morfismo biyectivo. La inversa de cualquier isomorfismo es también un isomorfismo	12, 51
Isomorfismo canónico.	Isomorfismo cuya construcción no depende de escoger algo (una base, un complementario, etc.)	51, 57
Isomorfismo de espacios vectoriales.		
Transformación lineal biyectiva		44, 73
Isomorfismo lineal.	Isomorfismo de espacios vectoriales	44
Límite de una sucesión.		
La sucesión $\{z_k\}$ tiene límite z si $\forall \varepsilon > 0 \exists N$ tal que $\forall k > N \quad \ z_k - z\ < \varepsilon$		142
Matriz.	Es una NM -ada o sea un conjunto indexado por dos conjuntos	33
Matriz acompañante de un polinomio.		
Matriz cuadrada cuya última columna es el vector de coeficientes del polinomio multiplicado por -1, la paralela inferior a la diagonal contiene unos y todas las demás entradas son cero.		176
Matriz ampliada.	Del sistema $Ax = b$ es la matriz $(A b)$	121
Matriz cuadrada.		
Matriz con la misma cantidad de columnas y renglones		105, 79
Matriz de cambio de base.		

Escogidas dos bases V y N de un espacio vectorial la matriz α_{NV} cuyas columnas son los coeficientes de las expresiones de la base V como combinación lineal de la base N . O sea, para cualquier $v \in V$ se cumple que $v = \sum_{i \in N} \alpha_{iv} i$. Si β_V son las coordenadas de un vector en la base V entonces $\alpha_{NV}\beta_V$ son las coordenadas del mismo vector en la base N 80	
Matriz de permutación. La matriz que se obtiene al permutar las columnas o los renglones de la matriz identidad. Matriz que tiene exactamente un 1 en cada renglón y columna y todas las demás entradas son cero 102	
Matriz de una transformación lineal.	
Escogidas una base N en el dominio y otra M en el codominio de una TL f la matriz α_{MN} cuyas columnas son los coeficientes de las expresiones de las imágenes de la base N como combinación lineal de la base M . O sea, para cualquier $j \in N$ se cumple que $f(j) = \sum_{i \in M} \alpha_{ij} i$ 77, 74, 81	
Matriz del sistema. La matriz A del sistema de ecuaciones lineales $Ax = b$ 120	
Matriz diagonal. Matriz α_{MM} en la cual si $i \neq j$ entonces, $\alpha_{ij} = 0$ 114	
Matriz diagonal por bloques. Matriz α_{MM} en la cual hay una partición del conjunto de índices $M_1 \cup \dots \cup M_t = M$ y que $\alpha_{ij} = 0$ si i y j pertenecen a bloques diferentes 114	
Matriz identidad.	
La matriz I_{NN} cuyas entradas son el delta de Kronecker: unos en la diagonal y ceros en el resto de las entradas. La matriz de la transformación lineal identidad 79, 99	
Matriz inversa. La matriz cuadrada α_{MN} es la inversa de β_{NM} si $\beta_{NM}\alpha_{MN} = I_{NN}$ y $\alpha_{MN}\beta_{NM} = I_{MM}$.	
Si ambos N y M son finitos entonces, basta comprobar una de las dos igualdades. Una matriz cuadrada tiene inversa si y solo si su determinante es diferente de cero 79, 110, 116, 126	
Matriz singular.	
Matriz cuadrada con determinante igual a cero 105, 116, 120	
Matriz triangular.	
Matriz triangular por bloques en la cual cada bloque es de cardinalidad 1 114	
Matriz triangular inferior.	
Matriz α_{MM} en la cual $M = \{1, \dots, m\}$ y tal que en su representación gráfica todas las entradas por encima de la diagonal son cero 114	
Matriz triangular por bloques. Matriz α_{MM} en la cual hay una partición del conjunto de índices $M_1 \cup \dots \cup M_t = M$ y que $\alpha_{ij} = 0$ si $i \in M_p$, $j \in M_q$ y $p < q$ 114	
Matriz triangular superior.	
Matriz α_{MM} en la cual $M = \{1, \dots, m\}$ y tal que en su representación gráfica todas las entradas por debajo de la diagonal son cero 114	
Máximo. El máximo de un subconjunto A de un conjunto ordenado B es una cota superior que pertenece a A . Si existe entonces, el máximo es único *	
Mínimo. El mínimo de un subconjunto A de un conjunto ordenado B es una cota inferior que pertenece a A . Si existe entonces, el mínimo es único *, 143	
Módulo de un complejo. La longitud del vector \vec{Oz} en el plano complejo 139	
Monomorfismo. Morfismo inyectivo. 13	
Monotonía.	
Propiedad de una función de un conjunto ordenado a otro que consiste en que $x \leq y \Rightarrow f(x) \leq f(y)$ 37	
Morfismo. Es una función $f : A \rightarrow B$ que cumple que $f(x \circ y) = f(x) \bullet f(y)$ donde	

○ es una operación definida en A y ● es una operación definida en B 10, 44	
Morfismo de álgebras. TL que commuta con el producto y preserva el 1 154	
Morfismo de anillos. Función entre dos anillos que es morfismo para la suma y para el producto 12	
Morfismo de campos. Función entre dos campos que es morfismo para la suma y para el producto 12	
Morfismo de espacios vectoriales. Función f de un espacio vectorial a otro sobre el mismo campo que es un morfismo de los grupos abelianos de vectores y que cumple $f(\alpha a) = \alpha f(a)$ para cualquier vector a y cualquier escalar α . 44	
Morfismo de grupos. Morfismo cuyo dominio y codominios son grupos 11, 93	
Multiplicidad de una raíz. El elemento del campo α es una raíz de multiplicidad n del polinomio p si n es el mayor natural tal que $p \vdash (x - \alpha)^n$. 133	
Múltiplo. Para dos polinomios p y q se dice que p es un múltiplo de q si $\exists c$ tal que $p = cq$ 132	
n-ada. Elemento (a_1, a_2, \dots, a_n) del producto cartesiano Aⁿ 29	
N-ada. Función en una notación diferente, $\alpha_N : N \ni i \rightarrow \alpha_i \in A$. A N se le llama el conjunto de índices y a las α_i se le llaman coordenadas 31, 76	
Neutro. De una operación binaria \circ es un elemento e que cumple que $a \circ e = e \circ a = a$ para cualquier a del conjunto en el cual está definida la operación. Si una operación tiene neutro entonces este es único 4	
Núcleo de un morfismo. El el caso de grupos la preimagen del neutro. Para anillos y espacios vectoriales la preimagen del cero 83	
Núcleo de un polinomio.	
La colección de sus raíces. Cada raíz aparece tantas veces como su multiplicidad 133	
Núcleo de una matriz. El núcleo de su transformación lineal. El conjunto solución de un sistema de ecuaciones con vector de coeficientes libres igual a cero 122	
Núcleo de una TL. La preimagen del cero 83	
Núcleo trivial. Núcleo igual a {0} 83	
OL. Acrónimo de operador lineal 70	
Operación binaria. Función mediante la cual para cualesquiera dos elementos a y b de un conjunto A podemos encontrar otro elemento de A que es el resultado de la operación 2	
Operador cíclico. OL h tal que todo el espacio es h -cíclico 163	
Operador diagonalizable. OL tal que existe una base en la cual su matriz es diagonal 178	
Operador irreducible. OL que no se puede decomponer como suma directa 151	
Operador lineal. Transformación lineal de un espacio en si mismo 70	
Operador radical. OL cuyo polinomio mínimo (o característico) es potencia de un polinomio irreducible 160	
Operador singular. OL no biyectivo 71	
Operador triangulable. OL tal que existe una base ordenada en la cual su matriz es triangular 178	
Opuesto. El inverso de un elemento de un anillo o campo respecto a la suma 7	
Órbita de una permutación. Clase de equivalencia de la relación $(a \sim b) \Leftrightarrow (a = \sigma^n(b))$ 95	
Orden de un ciclo. El número de elementos que un ciclo no deja fijos 94	
Orden de una matriz. El número de renglones y columnas de	

una matriz cuadrada	105	
Período de un conjunto de vectores.		
El generador del anulador del conjunto de vectores	157	
Período de un vector.		
El polinomio mónico p más pequeño tal que $p_h(a) = 0$. El generador del anulador del vector	156	
Permutación. Biyección de un conjunto finito en si mismo	93	
Permutación impar. Permutación que se descompone en la composición de un número impar de transposiciones	97	
Permutación par. Permutación que se descompone en la composición de un número par de transposiciones ...	97	
Plano.		
Subespacio afín de dimensión dos	55, 35	
Polinomio.	Expresión formal	
$\sum_{i=0}^n a_i x^i$ donde los coeficientes a_i son elementos de un campo	129	
Polinomio característico.	El polinomio $\det(xI - h)$. Es un múltiplo del polinomio mínimo y tiene los mismos factores irreducibles que el polinomio mínimo	175
Polinomio irreducible.	Polinomio mónico de grado al menos 1 sin factores propios	135
Polinomio mínimo.	El polinomio mónico más pequeño que anula un OL. El generador del anulador del operador	156
Polinomio mónico.	Polinomio cuyo coeficiente principal es 1	135
Preimagen de un elemento.	Sea b un elemento del codominio de f . La preimagen de b es el conjunto de todos x en dominio tales que $f(x) = b$	*, 83
Producto de matrices.		
Si α_{MN} y β_{NL} son dos matrices entonces $\gamma_{ML} = \alpha_{MN}\beta_{N,L}$ es la matriz cuyas entradas son los productos escala-		
lares canónicos $\gamma_{ij} = \alpha_{iN}\beta_{Nj}$ de los renglones de α_{MN} por las columnas de β_{NL}	76, 101	
Producto de un escalar por una función		
Si en el codominio de f está definido un producto por escalares entonces λf es la función tal que $(\lambda f)(a) = f(\lambda a)$	68	
Producto de un vector por una matriz		
Es la combinación lineal de los renglones de la matriz cuyos coeficientes son las coordenadas del vector	76	
Producto de una matriz por un vector		
Es la combinación lineal de las columnas de la matriz cuyos coeficientes son las coordenadas del vector	76	
Producto escalar.	Producto bilineal de dos vectores cuyo resultado es un escalar	75
Producto escalar canónico.	Producto escalar definido en $\mathbb{K}^{[N]}$ como:	
	$\alpha_N \beta_N = \sum_{i \in N} \alpha_i \beta_i$	75
Proyección.	Si $C = A \times B$ entonces, la función $f : c = (a, b) \mapsto a$ se le llama proyección de C a A a lo largo de B . En particular, nosotros la usamos en el caso de que $E = F \oplus G$ (la suma directa es un producto cartesiano!). En este caso las proyecciones son TLs	67, 83
Punto.	Subespacio afín de dimensión cero	55
Raíz de un polinomio.	Elemento del campo en el cual la función de evaluación del polinomio toma valor cero	133
Rango de una matriz.	El orden de sus bases.	
	La dimensión de su espacio de columnas.	
	La dimensión de su espacio de renglones.	119
Recta.		
Subespacio afín de dimensión uno	55, 35	
Regla del tablero de ajedrez.		

Regla mediante la cual se hallan los signos de los cofactores de una matriz en forma gráfica. El signo del cofactor de α_{ij} es $(-1)^{i+j}$	108	el vector \mathbf{b} son conocidos y el vector \mathbf{x} es incógnito	119	
Relación antisimétrica.		Subálgebra.	Subespacio vectorial que contiene al 1 y en el cual el producto es interno	
Si $\mathbf{a} \approx \mathbf{b}$ y $\mathbf{b} \approx \mathbf{a}$ entonces, $\mathbf{a} = \mathbf{b}$	*		154	
Relación de equivalencia.	Relación simétrica, reflexiva y transitiva *	55	
Relación de orden.	Relación antisimétrica, reflexiva y transitiva *	61, 118	
Relación en un conjunto.	Subconjunto del producto cartesiano $A \times A = A^2$ *		
Relación reflexiva.	Para todo a se tiene $a \approx a$ *		
Relación simétrica.	$(a \approx b) \Rightarrow (b \approx a)$ *		
Relación transitiva.	Si $a \approx b$ y $b \approx c$ entonces, $b \approx a$ *		
Renglón de una matriz.	Dada una matriz α_{NM} e $i \in N$ a la M -ada α_{iM} (i está fijo, los elementos de M varían) se le llama i -ésimo renglón de la matriz α_{NM} 33, 76	Subespacio afín.	
Resto.	Al efectuar la división con resto de un elemento p de un anillo comunitativo ($\mathbb{Z}, \mathbb{K}[x]$) entre otro q obtenemos la igualdad $p = cq + r$. Al elemento r se le llama resto de la división de p entre q 131, 14	Traslación de un subespacio	
Restricción de una función.			54, 68, 86	
Una función $f: A' \rightarrow B$ se le llama restricción de $g: A \rightarrow B$ si $A' \subseteq A$ y para cualquier $a \in A'$ se cumple $f(a) = g(a)$ 72	Subespacio generado.	Cerradura lineal	
Serie.	Expresión formal del tipo $\sum_{i=0}^{\infty} a_i x^i$. A los elementos del campo a_i se le llaman coeficientes de la serie 30 37	
Signo de una permutación.	Función que a cada permutación le hace corresponder 1 si esta es par y -1 si esta es impar 97	Subespacio h-cíclico.	Subespacio h -generado por un solo vector
Sistema de ecuaciones lineales.	Ecuación $A\mathbf{x} = \mathbf{b}$ donde la matriz A y	 163	
		Subespacio h-generado.	La h -cerradura de un conjunto de vectores	
		Subespacio invariante.	Subespacio donde está bien definida la restricción de un OL	
		Subespacio radical.	Subespacio invariante en el cual el operador es radical	
		Subespacios afines paralelos.	Dos subespacios afines son paralelos si uno es traslación del otro	
		 55	
		Subespacios complementarios.	Dos subespacios cuya suma es todo el espacio y cuya intersección es el origen	
		 52, 54, 56, 67, 83	
		Subgrupo.	Subconjunto de un grupo que es un grupo para la misma operación del grupo más grande	
		 17	
		Submatriz.	Dada una matriz α_{NM} a cualquier matriz $\alpha_{N'M'}$ tal que $N' \subseteq N$ y $M' \subseteq M$ se le llama submatriz de α_{NM}	
		 33, 118	
		Sucesión.	Elemento $(a_0, a_1, \dots, a_n, \dots)$ del	

conjunto $A^{\mathbb{N}}$ de funciones $f : \mathbb{N} \rightarrow A$	30
Sucesión acotada.	
La sucesión $\{z_k\}$ es acotada si existe un real M tal que $\forall k \quad \ z_k\ \leq M$	142
Sucesión convergente.	
Una sucesión que tiene límite	142
Suma de conjuntos.	
Si A y B son conjuntos y entre sus elementos hay una operación de suma entonces:	
$A + B = \{a + b \mid a \in A \text{ y } b \in B\}$	49
Suma de funciones.	
Si en el codominio mutuo de f y g está definida una suma entonces $f + g$ es la función tal que $(\lambda + g)(a) = f(a) + g(a)$	69
Suma directa de espacios.	
El producto cartesiano de los subespacios con la suma definida por coordenadas y también el producto por escalares. Si la intersección de dos subespacios tiene dimensión cero entonces, la suma directa de ellos es canónicamente isomorfa a la suma	50
Suma directa de OL.	
OL definido por coordenadas en la suma directa de espacios	149
Supremo.	
El mínimo de las cotas inferiores	*
Tensor de exponente n.	
Un conjunto indexado por n conjuntos de índices	34
Tipo de una descomposición.	
Si $h = f_1 \oplus \dots \oplus f_n$ es una descomposición de h entonces, su tipo es la sucesión de los polinomios mínimos de f_1, \dots, f_n . Dos tipos son iguales si uno se puede obtener del otro reordenando la sucesión. En el tipo pueden haber polinomios iguales	171
TL.	
Acrónimo de transformación lineal	65
Transformación elemental.	
Una transformación elemental de los renglones de una matriz consiste en sumarle a un renglón otro multiplicado por un escalar. Análogamente se definen las transformaciones elementales de las columnas	124
Transformación lineal.	
Morfismo de espacios vectoriales	44, 65, 72
Transformación lineal de una matriz.	
Dada la matriz α_{MN} es la función:	
$\mathbb{K}^N \ni \beta_N \rightarrow \alpha_{MN}\beta_N \in \mathbb{K}^M$	77
Transformación semilineal.	
Función de un espacio vectorial a otro que es morfismo para la suma y que cumple que $f(\lambda x) = \bar{\lambda}f(x)$ para cierto automorfismo $\lambda \mapsto \bar{\lambda}$ del campo de escalares	86
Transposición.	
Ciclo de orden dos	96
Transpuesta.	
Matriz que se obtiene de otra intercambiando los subíndices de sus entradas, o sea si $A = \alpha_{NM}$ y $A^T = B = \beta_{MN}$ entonces $\alpha_{ij} = \beta_{ji}$	100
Traslación de un conjunto de vectores.	
Si A es un conjunto de vectores y x otro vector entonces $A + x$ es la traslación de A con el vector x	54
Valor propio.	
Escalar tal que para cierto vector propio a se tiene que $h(a) = \lambda a$. Raíz del polinomio característico	174
Vector.	
Elemento de un espacio vectorial. En \mathbb{R}^n son los segmentos dirigidos del origen a un punto del espacio	28, 28
Vector columna.	
Matriz con una sola columna	76
Vector de coeficientes libres.	
El vector b del sistema de ecuaciones lineales $Ax = b$	120
Vector propio.	
Vector a tal que la recta $\langle a \rangle$ es invariante	174
Vector renglón.	
Matriz con un solo renglón	76

Notaciones

\forall	Para todo.
\exists	Existe.
$\exists!$	Existe y es único.
$P \Rightarrow Q$	Si P entonces Q . P implica Q .
	P es suficiente para Q .
$P \Leftarrow Q$	P solo si Q . P es necesario para Q .
$P \Leftrightarrow Q$	P si y solo si Q . P y Q son equivalentes. P es necesario y suficiente para Q .

$b \in B$	b pertenece a B .
$B \ni b$	B contiene a b .
$A \subseteq B$	A es subconjunto de B .
$A \supseteq B$	A es sobreconjunto de B .
$A \subsetneq B$	A es subconjunto propio de B . O sea, $A \subseteq B$ y $A \neq B$.
$A \supsetneq B$	A es sobreconjunto propio de B . O sea, $A \supseteq B$ y $A \neq B$.
$A \cup B$	La unión de A y B .
$A \cap B$	La intersección de A y B .
$A \setminus B$	La resta de conjuntos. El conjunto de los elementos de A que no pertenecen a B .
$A \times B$	El producto cartesiano de dos conjuntos. El conjunto de todas las parejas (a, b) con $a \in A$ y $b \in B$.
2^A	El conjunto de todos los subconjuntos del conjunto A . El conjunto de todas las funciones $f : A \rightarrow \{0, 1\}$.
A^n	El producto cartesiano de A consigo mismo n veces. El conjunto

de todos las n -adas (a_1, \dots, a_n) donde todos los a_i están en A .
 A^N El conjunto de todas las funciones $f : N \rightarrow A$. El conjunto de todos las N -adas a_N donde $\forall i \in N$ el elemento a_i pertenece a A . Si $N = \{1, \dots, n\}$ entonces $A^N = A^n$.

0	El vector cero. El origen de coordenadas.
0	Cero es elemento neutro para la suma de un anillo o campo.
1	Uno es el elemento neutro para el producto de un anillo o campo.
-1	Menos uno es el opuesto de 1 en un anillo o campo.
I_{NN}	Matriz identidad.
I	La función identidad. La permutación identidad
\emptyset	La función nula. La matriz nula
\aleph_0	El cardinal de N .

K	Un campo arbitrario.
K^n	El espacio de las n -adas.
K^N	El espacio de las N -adas.
$K[x]$	El álgebra de polinomios en la variable x con coeficientes en K .
$K[[x]]$	El álgebra de series en la variable x con coeficientes en K .
$K(x)$	El campo de funciones racionales en x con coeficientes en K .
N	El conjunto de los naturales.
Z	El anillo de los enteros.
Z_n	El anillo de restos módulo n . Para n primo Z_n es un campo.

\mathbb{Q}	El campo de los racionales.
\mathbb{R}	El campo de los reales.
\mathbb{C}	El campo de los complejos.
\mathbb{S}_N	El grupo simétrico de \mathbb{N} .
\mathbb{A}_N	El grupo alternante de \mathbb{N} .
\mathbb{A}_N^-	Las permutaciones impares de \mathbb{N} .

$f : A \rightarrow B$	Una función que tiene dominio A y codominio B .
$f : A \ni a \mapsto f(a) \in B$	Usada para denotar funciones concretas por ejemplo,
$f : \mathbb{R} \ni a \mapsto a^2 \in \mathbb{R}^+$	es la función con dominio \mathbb{R} y codominio \mathbb{R}^+ que a cada real le hace corresponder su cuadrado.

$p \bmod q$	El resto de la división de p entre q . Está bien definido en cualquier anillo con división. Por ejemplo, en los números enteros y en los polinomios con coeficientes en un campo.
$n!$	El factorial del natural n . Se define como $n! = 1 \times 2 \times \dots \times n$.
$ A $	El cardinal del conjunto A . Puede de ser finito o infinito.
$\ z\ $	El módulo del número complejo z .

$\lim z_k$	El límite de la sucesión $\{z_k\}$.
$\dim E$	La dimensión de E .
$\operatorname{sgn} \pi$	El signo de la permutación π .
$\det A$	El determinante de la matriz A .
$\operatorname{Im} f$	La imagen de la función f .
$\ker f$	El nucleo del morfismo f .

$A+B$	La suma de dos conjuntos de vectores.
λA	El producto de un escalar por un conjunto de vectores.
$A+x$	El conjunto de vectores A trasladado mediante el vector x . Si A es un subespacio entonces, es un subespacio afín.
$E+F$	La suma directa de dos espacios. Si E y F son subespacios que se intersectan solo en el origen entonces es canónicamente isomorfa a la suma de E y F .
α_{MN}	Una matriz cuyos renglones están indexados por M y cuyas columnas están indexadas por N .
α_{ij}	La entrada α_{ij} de una matriz α_{MN} .
α_{iN}	El i -ésimo renglón de la matriz α_{NM} .
α_{Mj}	La j -ésima columna de la matriz α_{NM} .
$\alpha_{M\omega(N)}$	Si ω es una permutación de N , es la matriz α_{MN} cuyas columnas están permutadas mediante ω . Si $\omega : N \rightarrow L$ es una biyección, es la matriz cuyas columnas están indexadas por L mediante el cambio de índices ω .
$\alpha_{\omega(M)N}$	Lo mismo que el anterior pero para los renglones
α_{ij}^*	El cofactor de la entrada α_{ij} de una matriz cuadrada.
A^*	La matriz de cofactores de A .
A^T	La matriz transpuesta de A .
$(A \mathbf{b})$	La matriz ampliada del sistema $Ax = \mathbf{b}$.



Alfabeto gótico

A	B	C	D	E	F	G	H	I	J	K	L	M
ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ

a	b	c	d	e	f	g	h	i	j	k	l	m
ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ

n	o	p	q	r	s	t	u	v	w	x	y	z
ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ	ȝ

Alfabeto caligráfico

A	B	C	D	E	F	G	H	I	J	K	L	M
ѧ	ܶ	ܳ	ܴ	ܵ	ܶ	ܷ	ܸ	ܹ	ܺ	ܻ	ܼ	ܾ

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
ܿ	ܰ	ܫ	ܭ	ܮ	ܯ	ܱ	ܲ	ܳ	ܵ	ܶ	ܷ	ܸ

Alfabeto griego

Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι
α	β	γ	δ	ε ε	ζ	η	θ θ	ι
alfa	beta	gamma	delta	épsilon	zeta	eta	zita	iota

Κ	Λ	Μ	Ν	Ξ	Ο	Π	Ρ	Σ
κ κ	λ	μ	ν	ξ	ο	π	ρ	σ
kappa	lamda	mu	nu	xi	ómicron	pi	ro	sigma

Τ	Υ	Φ	Χ	Ψ	Ω
τ	υ	φ φ	χ	ψ	ω
tau	úpsilon	fi	chi	psi	omega

Guía de estudio

Capítulo 1

1. Defina los siguientes conceptos:
 - Operación binaria.
 - Propiedad conmutativa.
 - Propiedad asociativa.
 - Propiedad distributiva.
 - Elemento neutro.
 - Elemento inverso.
2. En cualquier campo $0x = 0$.
3. Defina los siguientes conceptos:
 - Grupo y grupo abeliano.
 - Anillo y anillo conmutativo.
 - Campo.
4. Defina los morfismos.
5. Los morfismos preservan las operaciones binarias.
6. Los morfismos preservan la conmutatividad.
7. Los morfismos preservan la asociatividad.
8. Los morfismos preservan el neutro.
9. Los morfismos preservan el inverso.
10. Los morfismos preservan la distributividad.
11. La inversa de un isomorfismo es un isomorfismo.
12. Defina la suma y el producto en \mathbb{Z}_n .
13. $(\mathbb{Z}_n, +, \bullet)$ es un anillo conmutativo.
14. Todo campo es un dominio de integridad.
15. \mathbb{Z}_n es un dominio de integridad si y solo si n es primo.
16. Todo dominio de integridad finito es un campo.
17. Definición de subcampo.
18. Definición de campo primo

19. Todo campo \mathbb{K} contiene un único subcampo primo que está contenido en cualquier subcampo de \mathbb{K} .
20. \mathbb{Q} es un campo primo.
21. Los campos \mathbb{Z}_p son primos.
22. Teorema de Clasificación de Campos Primos.
23. Defina la característica de un campo.
24. En un campo de característica t para cualquier elemento x se cumple que $tx = 0$.
25. Demuestre la fórmula del binomio de Newton en base a la fórmula multinomial.

Capítulo 2

1. Definición de espacio vectorial.
2. ¿Qué es una n -ada? ¿Qué es una N -ada? ¿Cuál es su conjunto de índices? ¿Qué son sus coordenadas?
3. ¿Qué es una NM -matriz? ¿Qué es una entrada, renglón, columna?
4. Definición de subespacio.
5. Los subespacios son los conjuntos cerrados para la suma y el producto por escalares.
6. La unión de subespacios no es un subespacio.
7. La intersección de subespacios es un subespacio.
8. Definición de combinación lineal y sus coeficientes.
9. El conjunto de todas las combinaciones lineales de un conjunto de vectores es un subespacio.
10. Los siguientes tres conjuntos de vectores coinciden:

- El conjunto de todas las combinaciones lineales de \mathbb{N} .
 - La intersección de todos los subespacios que contienen a \mathbb{N} .
 - El subespacio más pequeño que contiene a \mathbb{N} .
11. Propiedades básicas de la cerradura lineal:
- incremento,
 - monotonía,
 - idempotencia.
12. Todo sobreconjunto de un conjunto generador es generador.
13. Teorema de caracterización de conjuntos LI.
14. Todo subconjunto de un conjunto LI es LI.
15. Lema de aumento de un conjunto LI.
16. Teorema de caracterización de bases.
17. Teorema de existencia de bases.
18. Propiedad del cambio de las bases.
19. Dos bases cualesquiera de un espacio vectorial tienen el mismo cardinal.
20. Si \mathfrak{E} es un subespacio de \mathfrak{F} y $\dim \mathfrak{E} = \dim \mathfrak{F} < \infty$ entonces, $\mathfrak{E} = \mathfrak{F}$.
21. Describa las bases canónicas de los espacios vectoriales $\mathbb{K}^{[N]}$ y $\mathbb{K}[x]$.
22. La inversa de un isomorfismo lineal es un isomorfismo lineal.
23. Un isomorfismo transforma conjuntos LI en conjuntos LI, conjuntos generadores en conjuntos generadores y bases en bases.
24. Describa el isomorfismo de coordinatización en una base.
25. Dos espacios vectoriales sobre un mismo campo son isomorfos si y solo si tienen la misma dimensión.
26. Todo espacio vectorial finito dimensional es isomorfo a \mathbb{K}^n .
27. El número de elementos en un campo finito es potencia de un número primo.
28. $\langle \mathfrak{E} \cup \mathfrak{F} \rangle = \{ \mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathfrak{E}, \mathbf{b} \in \mathfrak{F} \}$
29. La igualdad modular (incluye la demostración del lema).
30. Si \mathfrak{E} y \mathfrak{F} son subespacios tales que $\mathfrak{E} \cap \mathfrak{F} = \{\mathbf{0}\}$ entonces la función $\mathfrak{E} \oplus \mathfrak{F} \ni (\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} + \mathbf{b} \in \mathfrak{E} + \mathfrak{F}$ es un isomorfismo de espacios vectoriales.
31. Que es un isomorfismo canónico. Demuestre que $\mathfrak{E} \oplus \mathfrak{F}$ y $\mathfrak{F} \oplus \mathfrak{E}$ son canómicamente isomorfos.
32. Todo subespacio tiene complementario.
33. Si \mathfrak{E} y \mathfrak{F} son dos subespacios complementarios entonces cada vector \mathbf{x} se expresa de forma única como $\mathbf{x} = \mathbf{a} + \mathbf{b}$ donde $\mathbf{a} \in \mathfrak{E}$ y $\mathbf{b} \in \mathfrak{F}$.
34. Defina los subespacios afines.
35. ¿Que es el paralelismo de subespacios afines?
36. Todo subespacio afín es paralelo a un solo subespacio vectorial.
37. Dos diferentes subespacios afines paralelos no se intersectan
38. ¿Que es el espacio cociente? ¿Cuales son sus operaciones?
39. Cualquier subespacio complementario a \mathfrak{E} intersecta al subespacio afín $(\mathfrak{E} + \mathbf{x})$ en un solo punto.
40. $\mathfrak{D}/\mathfrak{E}$ es canómicamente isomorfo a cualquier complementario de \mathfrak{E} .

Capítulo 3

1. Definición de transformación lineal.
2. Toda TL transforma subespacios en subespacios.
3. Toda TL de un espacio de dimensión 1 es una homotecia.
4. Definición de proyección. Toda proyección es una TL.
5. El producto de un escalar por una TL es una TL.
6. La suma de dos TLs es una TL.
7. La composición de TLs es una TL.
8. Propiedades de la composición de TLs: asociatividad, distributividad y commutatividad con el producto por escalares.

9. Definición de Álgebra. De tres ejemplos de álgebras. ¿Qué es el grupo general lineal?
10. Las extensiones lineales son TLs.
11. Las TLs están predeterminadas por sus valores en una base.
12. Si \mathbf{N} es una base de \mathfrak{E} entonces, la función que a cada TL $\mathbf{h} \in \text{Hom}(\mathfrak{E}, \mathfrak{F})$ le hace corresponder su restricción $\mathbf{h}_{\mathbf{N}} \in \mathfrak{F}$ es un isomorfismo de espacios vectoriales.
13. Una TL es un isomorfismo si y solo si su restricción a una base es inyectiva y la imagen de esta restricción es una base.
14. Propiedades del producto escalar canónico de \mathbf{N} -adas: comunitatividad, distributividad y comunitatividad con el producto por escalares.
15. Definición del producto de matrices.
16. ¿Cuál es la TL definida por una matriz?
¿Cuál es la matriz de una TL?
17. La matriz de la composición de dos TLs es igual al producto de las matrices de las TLs.
18. Defina las matrices de cambio de base.
19. La \mathbf{N} -ada de las coordenadas de un vector en la base nueva se obtiene multiplicando la matriz de cambio de base por la \mathbf{V} -ada de las coordenadas del vector en la base vieja.
20. Sea $\mathbf{f} : \mathfrak{E} \rightarrow \mathfrak{F}$ una TL. Sean \mathbf{B} y \mathbf{C} matrices de cambio de base en \mathfrak{E} y \mathfrak{F} respectivamente. Si \mathbf{A} es la matriz de \mathbf{f} en las bases viejas entonces \mathbf{CAB}^{-1} es la matriz de \mathbf{f} en las bases nuevas.
21. Definición de núcleo e imagen de una TL.
22. La imagen y el núcleo son subespacios.
23. Una TL es inyectiva si y solo si su núcleo es trivial.
24. Si \mathfrak{K} es un subespacio complementario a $\ker \mathbf{f}$ entonces, la restricción de \mathbf{f} a \mathfrak{K} es inyectiva.
25. Sean \mathfrak{E} y \mathfrak{F} dos espacios tales que $\dim \mathfrak{E} = \dim \mathfrak{F} < \infty$. Una TL de \mathfrak{E} a

- \mathfrak{F} es inyectiva si y solo si es sobreyectiva.
26. Los subespacios afines paralelos a $\ker \mathbf{f}$ son precisamente los conjuntos de vectores en que la TL \mathbf{f} es constante.

Capítulo 4

1. Si $|\mathbf{M}| = |\mathbf{N}|$ entonces, los grupos simétricos $\mathbb{S}_{\mathbf{M}}$ y $\mathbb{S}_{\mathbf{N}}$ son isomorfos.
 2. El número de permutaciones de un conjunto con n elementos es $n!$.
 3. Que son las órbitas de una permutación.
 4. La restricción de una permutación a una órbita es un ciclo.
 5. Definición de permutaciones pares e impares. Definición del signo.
 6. La composición de una permutación con una transposición cambia la paridad de la permutación.
 7. Toda permutación es composición de transposiciones.
 8. Pruebe que $\text{sgn}(\pi \circ \rho) = \text{sgn } \pi \text{sgn } \rho$ y que $\text{sgn } \pi^{-1} = \text{sgn } \pi$.
 9. Definición de determinante de una matriz. Regla del triángulo para los determinantes de orden 3.
 10. El determinante de la matriz identidad es 1.
 11. Si una matriz tiene una columna o un renglón nulo entonces, su determinante es cero.
 12. El determinante no se altera al transponer una matriz.
 13. Si una matriz tiene dos renglones iguales entonces, su determinante es cero.
 14. Si ϕ y φ son dos cambios de índices de \mathbf{N} en \mathbf{M} entonces, se cumple la igualdad:
- $$\det \alpha_{\mathbf{M}\phi(\mathbf{N})} = \text{sgn}(\phi \circ \varphi^{-1}) \det \alpha_{\mathbf{M}\varphi(\mathbf{N})}.$$
15. Definición de matriz no singular.
 16. Definición de cofactores de una entrada de una matriz. Pruebe que los cofactores no dependen del cambio de índices usado para calcular el determinante.
 17. Teorema de expansión de Laplace.

18. Definición de funcional multilineal
19. El determinante es un funcional multilineal de los renglones.
20. $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$.
21. $\mathbf{A}^{-1} = \mathbf{A}^{*\top} / \det \mathbf{A}$.
22. El determinante de un OL no depende de la base que se use para calcularlo.
23. Enuncie la expansión generalizada de Laplace
24. El determinante de una matriz triangular por bloques es igual al producto de los determinantes de sus bloques.
25. Enuncie la caracterización de matrices no singulares.
26. Lema de aumento de submatrices no singulares.
27. Si α_{IJ} es una base de una matriz α_{MN} entonces el conjunto de renglones indexado por I es una base del espacio de renglones de α_{MN} .
28. Enuncie el Teorema del rango.
29. Regla de Cramer.
30. Teorema de existencia de soluciones.
31. Lema de eliminación de ecuaciones dependientes.
32. Describa el procedimiento general de solución de los sistemas de ecuaciones lineales.
33. Las transformaciones elementales no cambian los determinantes.
34. Las transformaciones elementales de los renglones de la matriz ampliada no cambian el subespacio afín solución de un sistema de ecuaciones lineales.
35. Resuelva un sistema de ecuaciones lineales por el método de Gauss.
36. Encuentre la inversa de una matriz por el método de eliminación de Gauss

Capítulo 5

1. Defina los polinomios sobre un campo, sus coeficientes, el grado y el coeficiente principal.
2. Defina la suma y el producto de polino-

- mios.
3. Defina la función de evaluación de un polinomio.
 4. Un polinomio de grado n está predeterminado por su evaluación en $n+1$ diferentes elementos del campo.
 5. División con resto de polinomios.
 6. Defina los divisores, los múltiplos y los factores de un polinomio.
 7. Si $p \dashv q$ y $q \dashv p$ entonces existe un elemento del campo α tal que $p = \alpha q$.
 8. Defina los divisores, los múltiplos y los factores de un polinomio.
 9. Defina las raíces de un polinomio.
 10. Para que b sea una raíz de p es necesario y suficiente que $(x - b)$ sea un factor de p .
 11. Un polinomio de grado n tiene a lo más n raíces.
 12. Defina los ideales de un anillo y los ideales principales.
 13. Todo ideal de polinomios es principal.
 14. Teorema de Bezout.
 15. Defina factores propios, polinomios mónicos, factores irreducibles.
 16. Si p y q son dos polinomios y r es un factor irreducible de pq entonces r es un factor de p o de q .
 17. Sea $p = \alpha p_1 \cdots p_n$ una descomposición en factores irreducibles de p . Si q es cualquier factor irreducible de p entonces, q es igual a alguno de los p_i .
 18. Teorema de Factorización de Polinomios.
 19. Desarrollo de Taylor.
 20. Enuncie el Teorema de Gauss.
 21. Usando el Teorema de Gauss demuestre la clasificación de los polinomios complejos irreducibles.
 22. Demuestre la clasificación de los polinomios reales irreducibles.

Capítulo 6

1. Defina la suma directa de operadores lineales.

2. Defina los subespacios invariantes de un operador lineal.
3. Un OL se descompone como suma directa de dos OLs si y solo si él tiene dos subespacios invariantes complementarios.
4. Defina los operadores lineales irreducibles.
5. Demuestre que todo operador lineal se descompone en suma de irreducibles.
6. Demuestre que si un operador lineal se descompone en suma directa de otros dos entonces, en cierta base su matriz es diagonal por bloques.
7. Defina la función de evaluación de un polinomio en un operador lineal.
8. La función de evaluación de polinomios en un operador lineal es un morfismo de álgebras.
9. La imagen de la función de evaluación de polinomios es una subálgebra conmutativa del álgebra de operadores lineales.
10. El núcleo del morfismo de evaluación en \mathbf{h} es un ideal de $\mathbb{K}[x]$.
11. Defina el polinomio mínimo de un operador lineal.
12. Defina el anulador de un conjunto de vectores.
13. Defina el \mathbf{h} -período de un conjunto de vectores.
14. El \mathbf{h} -anulador de \mathbf{A} es el conjunto de los múltiplos comunes a los períodos de los vectores en \mathbf{A} .
15. Si $\mathbf{h} = \mathbf{f} \oplus \mathbf{g}$ entonces el polinomio mínimo de \mathbf{h} es igual al mínimo común múltiplo de los polinomios mínimos de \mathbf{f} y \mathbf{g} .
16. Monotonía del período.
17. Invariancia de los núcleos.
18. Monotonía de los núcleos.
19. Lema de Descomposición de Núcleos.
20. Si \mathbf{h} es irreducible entonces, es radical.
21. Enuncie el Teorema de Descomposición en Componentes Radicales (sin demostración).
22. Para cualquier operador lineal \mathbf{h} existe un vector tal que su período es igual al polinomio mínimo de \mathbf{h} .
23. Definición de \mathbf{h} -combinación y sus coeficientes.
24. El conjunto de todas las \mathbf{h} -combinaciones de \mathbf{V} es un subespacio invariante.
25. Definición de subespacio \mathbf{h} -generado y conjunto de vectores \mathbf{h} -generador.
26. $\text{per}_{\mathbf{h}} \langle \mathbf{V} \rangle_{\mathbf{h}} = \text{per}_{\mathbf{h}} \mathbf{V}$.
27. Definición de subespacio \mathbf{h} -cíclico y de operador cíclico.
28. Si \mathbf{q} es un polinomio coprimo con $\text{per}_{\mathbf{h}} \mathbf{a}$ entonces $\langle \mathbf{a} \rangle_{\mathbf{h}} = \langle \mathbf{q}_{\mathbf{h}}(\mathbf{a}) \rangle_{\mathbf{h}}$.
29. Si el período de \mathbf{a} es de grado n , entonces el conjunto de vectores $\{\mathbf{a}, \mathbf{h}(\mathbf{a}), \dots, \mathbf{h}^{n-1}(\mathbf{a})\}$ es una base de $\langle \mathbf{a} \rangle_{\mathbf{h}}$.
30. Definición de conjunto \mathbf{h} -independiente.
31. Si $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es \mathbf{h} -independiente entonces $\langle \mathbf{V} \rangle_{\mathbf{h}} = \langle \mathbf{v}_1 \rangle_{\mathbf{h}} \oplus \dots \oplus \langle \mathbf{v}_n \rangle_{\mathbf{h}}$.
32. Definición de \mathbf{h} -base.
33. El conjunto $\text{End}^{\mathfrak{F}}(\mathfrak{E})$ de todos los operadores lineales que dejan invariante \mathfrak{F} es una subálgebra de $\text{End}(\mathfrak{E})$.
34. La función $\text{End}^{\mathfrak{F}}(\mathfrak{E}) \ni \mathbf{h} \mapsto \tilde{\mathbf{h}} \in \text{End}(\mathfrak{E}/\mathfrak{F})$ es un morfismo de álgebras.
35. Pruebe que $\widetilde{\mathbf{p}_{\mathbf{h}}} = \mathbf{p}_{\tilde{\mathbf{h}}}$.
36. Si $\mathbf{b} \in \mathbf{v} + \mathfrak{F}$ entonces $\text{per}_{\mathbf{h}}(\mathbf{b}) \vdash \text{per}_{\tilde{\mathbf{h}}}(\mathbf{v} + \mathfrak{F})$.
37. Lema del período (sin demostración)
38. Lema del cociente.
39. Teorema de Existencia de \mathbf{h} -bases.
40. Teorema de Descomposición en Componentes Radicales Cíclicas
41. Defina el tipo de una descomposición.
42. Teorema de Unicidad del Tipo (sin demostración).
43. Teorema de Caracterización de los OLs irreducibles.
44. Definición de vector propio y de valor propio.
45. La recta $\langle \mathbf{a} \rangle$ es \mathbf{h} -invariante si y solo si \mathbf{a} es un vector propio de \mathbf{h} .

46. Definición de polinomio característico.
Cual es el grado del polinomio característico (sin demostración).
47. Demuestre que las raíces del polinomio característico son los valores propios.
48. Definición de matriz acompañante de un polinomio.
49. Sea \mathbf{h} un OL cíclico con polinomio mínimo p de grado n . La matriz de \mathbf{h} en la base $\{\mathbf{a}, \mathbf{h}(\mathbf{a}), \dots, \mathbf{h}^{n-1}(\mathbf{a})\}$ es la matriz acompañante de p .

50. Si $\mathbf{h} = \mathbf{f} \oplus \mathbf{g}$ entonces, el polinomio característico de \mathbf{h} es igual al producto de los polinomios característicos de \mathbf{f} y \mathbf{g} .
51. Teorema de Hamilton-Caley-Frobenius.
52. Definición de celda de Jordán.
53. Si \mathbf{h} es un operador cíclico-radical con polinomio mínimo $(x - \lambda)^n$ entonces, en cierta base, la matriz de \mathbf{h} es una celda de Jordan.
54. Forma normal de Jordán

